GRAND VALLEY STATE UNIVERSITY
SCHOOL OF PUBLIC, NONPROFIT, AND HEALTH ADMINISTRATION

Lean Six Sigma Black Belt Training

Featuring Examples from Minitab 18

1.0 Define Phase

1.1 Overview of Six Sigma

Black Belt Training: Define Phase

1.1 Six Sigma Overview
- 1.1.1 What is Six Sigma
- 1.1.2 Six Sigma History
- 1.1.3 Six Sigma Approach Y = f(x)
- 1.1.4 Six Sigma Methodology
- 1.1.5 Roles and Responsibilities

1.2 Six Sigma Fundamentals
- 1.2.1 Defining a Process
- 1.2.2 VOC and CTQs
- 1.2.3 QFD
- 1.2.4 Cost of Poor Quality (COPQ)
- 1.2.5 Pareto Analysis (80 : 20 rule)

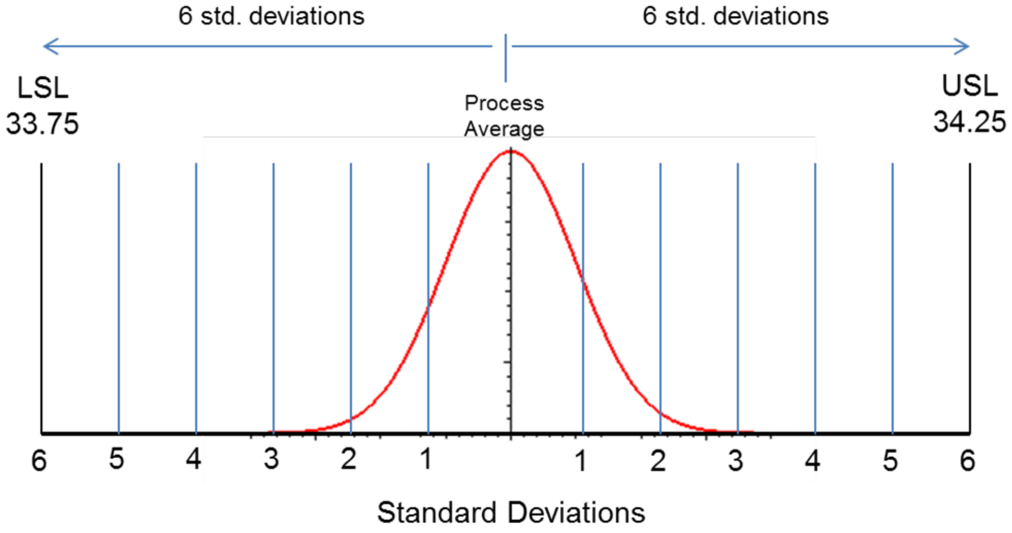
1.3 Lean Six Sigma Projects
- 1.3.1 Six Sigma Metrics
- 1.3.2 Business Case and Charter
- 1.3.3 Project Team Selection
- 1.3.4 Project Risk Management
- 1.3.5 Project Planning

1.4 Lean Fundamentals
- 1.4.1 Lean and Six Sigma
- 1.4.2 History of Lean
- 1.4.3 The Seven Deadly Muda
- 1.4.4 Five-S (5S)

1.1.1 What is Six Sigma

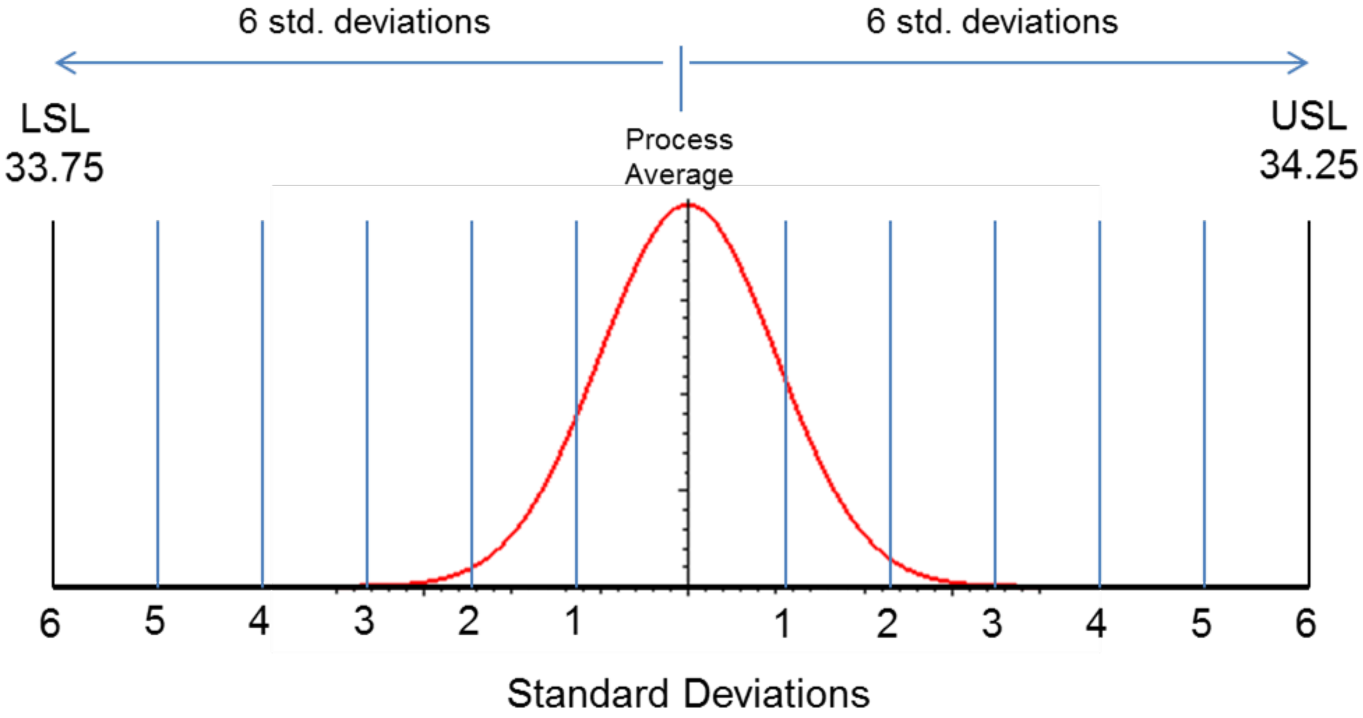
What is Six Sigma?

- What is “sigma”?
  - In statistics, **sigma** (σ) refers to “standard deviation,” which is a measure of variation.
  - You will come to learn that variation is the enemy of any quality process. We need to understand, manage, and minimize process variation.
- What is “Six Sigma”?
  - **Six Sigma** is an aspiration or goal of process performance.
  - A Six Sigma “goal” is for a process average to operate approximately 6σ away from customer’s high and low specification limits.



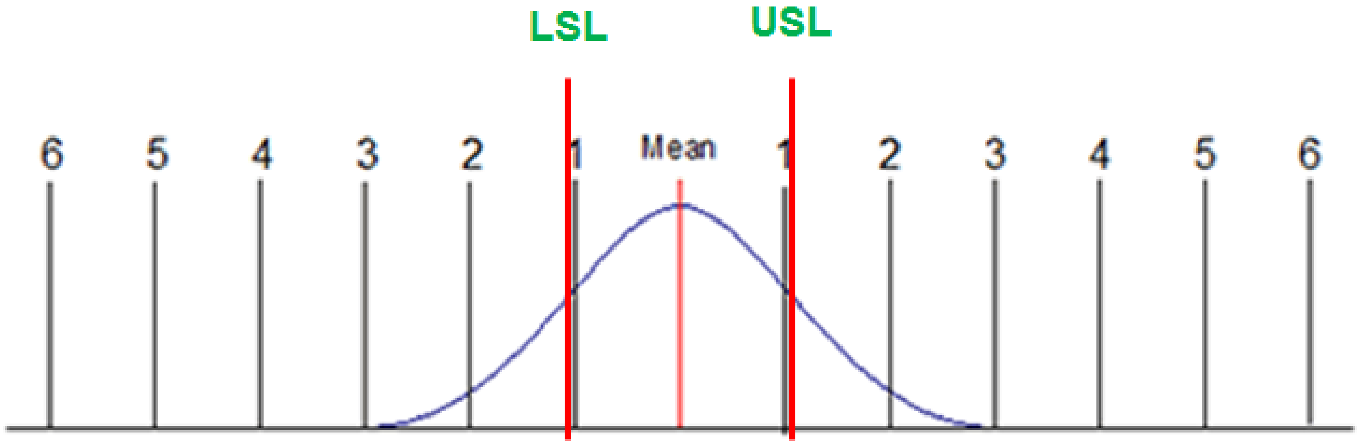
What is Six Sigma?

- A process whose average is about 6σ away from the customer's high and low specification limits has abundant room to "float" before approaching the customer's specification limits.
- A Six Sigma process only yields 3.4 defects for every million opportunities! In other words, 99.9997% of the products are defect-free!



What is Six Sigma: Sigma Level

- **Sigma level** measures how many "sigma" there are between your process average and the nearest customer specification.
- Let us assume that your customers upper and lower specifications limits (USL & LSL) were narrower than the width of your process spread.
- The USL & LSL below stay about 1 standard deviation away from the process average. Therefore, this process operates at **1 sigma**.



What is Six Sigma: Sigma Level

- A process operating at 1 sigma has a defect rate of approximately 70%.

  - This means that the process will generate defect-free products only 30% of the time.

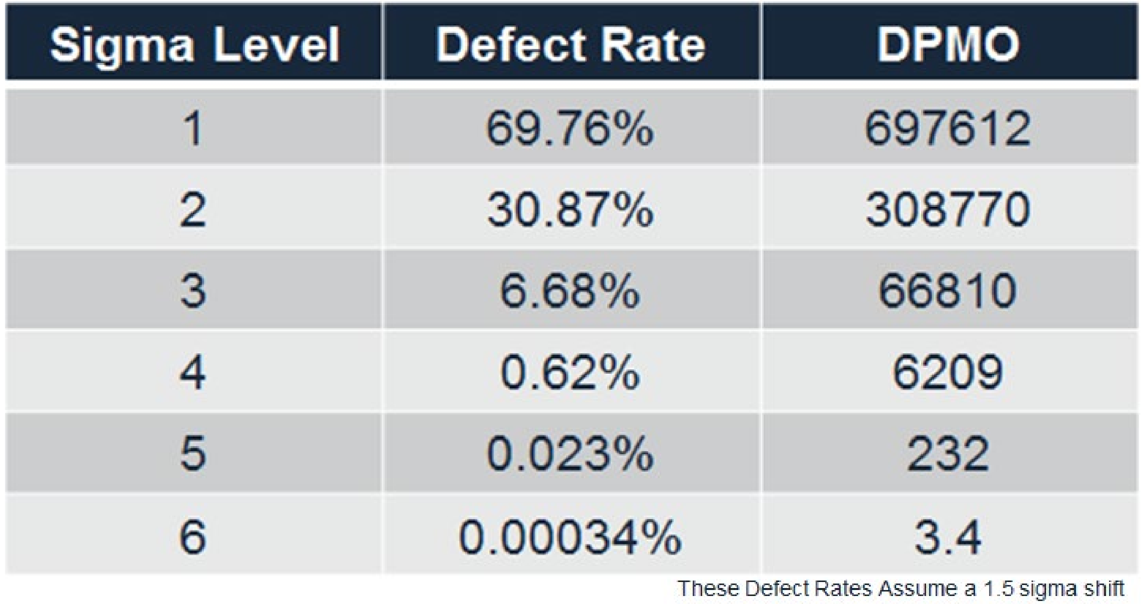
- What about processes with more than 1 sigma level?

- A higher sigma level means a lower defect rate.

- Let us take a look at the defect rates of processes at different sigma levels.

What is Six Sigma: Sigma Level

- This table shows each sigma level’s corresponding defect rate and DPMO (defects per million opportunities).
- The higher the sigma level, the lower the defective rate and DPMO.



Sigma Level	Defect Rate	DPMO
1	69.76%	697612
2	30.87%	308770
3	6.68%	66810
4	0.62%	6209
5	0.023%	232
6	0.00034%	3.4

These Defect Rates Assume a 1.5 sigma shift

- How does this translate into things you might easily relate to?

What is Six Sigma: Sigma Level

- Let us take a look at processes operating at 3 sigma.
- 3 sigma processes have a defect rate of approximately 7%. What would happen if processes operated at 3 sigma?
  - Virtually no modern computer would function*.
  - 10,800,000 health care claims would be mishandled each year.
  - 18,900 US savings bonds would be lost every month.
  - 54,000 checks would be lost each night by a single large bank.
  - 4,050 invoices would be sent out incorrectly each month by a modest-sized telecommunications company.
  - 540,000 erroneous call details would be recorded each day from a regional telecommunications company.
  - 270 million erroneous credit card transactions would be recorded each year in the United States.

*(*http://www.qualityamerica.com)*

What is Six Sigma: Sigma Level

- What if processes operated with 1% defect rate?
    - 20,000 lost articles of mail per hour*.
    - Unsafe drinking water almost 15 minutes per day.
    - 5,000 incorrect surgical operations per week.
    - Short or long landings at most major airports each day.
    - 200,000 wrong drug prescriptions each year.
    - No electricity for almost 7 hours per month.

- Even at 1% defect rate, some processes would be unacceptable to you and many others.

- **So what is Six Sigma?**
    - Sigma level is the measure!
    - Six is the goal!

(Implementing Six Sigma – Forest W. Breyfogle III)*

What is Six Sigma: The Methodology

- Six Sigma itself is the **goal**, not the method.

- In order to achieve Six Sigma, you need to improve your process performance by:
  - Minimizing the process variation so that your process has enough room to fluctuate within customer's spec limits
  - Shifting your process average so that it is centered between your customer's spec limits.

- Accomplishing these two process improvements *(along with stabilization and control),* you can achieve Six Sigma.

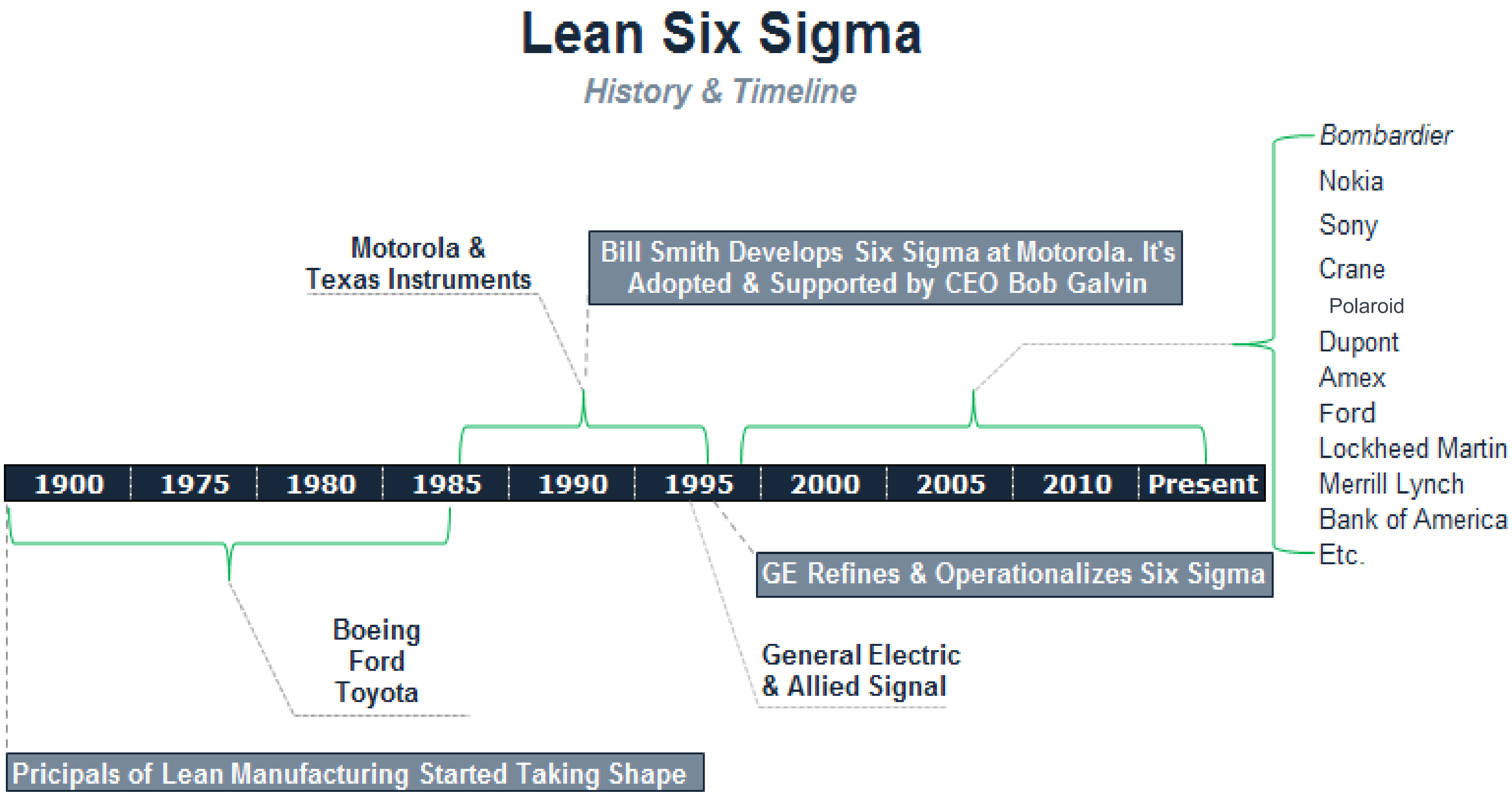
- DMAIC is the systematic methodology prescribed to achieve Six Sigma.

What is Six Sigma: The Methodology

- DMAIC is a systematic and rigorous methodology that can be applied to any process in order to achieve Six Sigma.
- It consists of 5 phases of a project:
  - **D**efine
  - **M**easure
  - **A**nalyze
  - **I**mprove
  - **C**ontrol.
- You will be heavily exposed to many concepts, tools, and examples of the DMAIC methodology through this training.
- You will be capable of applying the DMAIC methodology to improve the performance of any process at the completion of the curriculum.

1.1.2 Six Sigma History

Six Sigma History



Lean Six Sigma

History & Timeline

1900 | 1975 | 1980 | 1985 | 1990 | 1995 | 2000 | 2005 | 2010 | Present

Motorola &
Texas Instruments

Bill Smith Develops Six Sigma at Motorola. It's Adopted & Supported by CEO Bob Galvin

Bombardier
Nokia
Sony
Crane
Polaroid
Dupont
Amex
Ford
Lockheed Martin
Merrill Lynch
Bank of America
Etc.

GE Refines & Operationalizes Six Sigma

Boeing
Ford
Toyota

General Electric
& Allied Signal

Pricipals of Lean Manufacturing Started Taking Shape

Six Sigma History

- The "Six Sigma" terminology was originally adopted by Bill Smith at Motorola in the late 1980s as a quality management methodology.
- As the "Father of Six Sigma," Bill forged the path for Six Sigma through Motorola's CEO Bob Galvin who strongly supported Bill's passion and efforts.
- Starting from the late 1980s, Motorola extensively applied Six Sigma as a process management discipline throughout the company, leveraging Motorola University.
- In 1988, Motorola was recognized with the prestigious Malcolm Baldrige National Quality Award for its achievements in quality improvement.

Six Sigma History

- Six Sigma has been widely adopted by companies as an efficient way of improving the business performance since General Electric implemented the methodology under the leadership of Jack Welch in the 1990s.

- As GE connected Six Sigma results to its executive compensation and published the financial benefits of Six Sigma implementation in their annual report, Six Sigma became a highly sought-after discipline of quality.

Six Sigma History

- Most Six Sigma programs cover the aspects, tools, and topics of Lean or Lean Manufacturing.

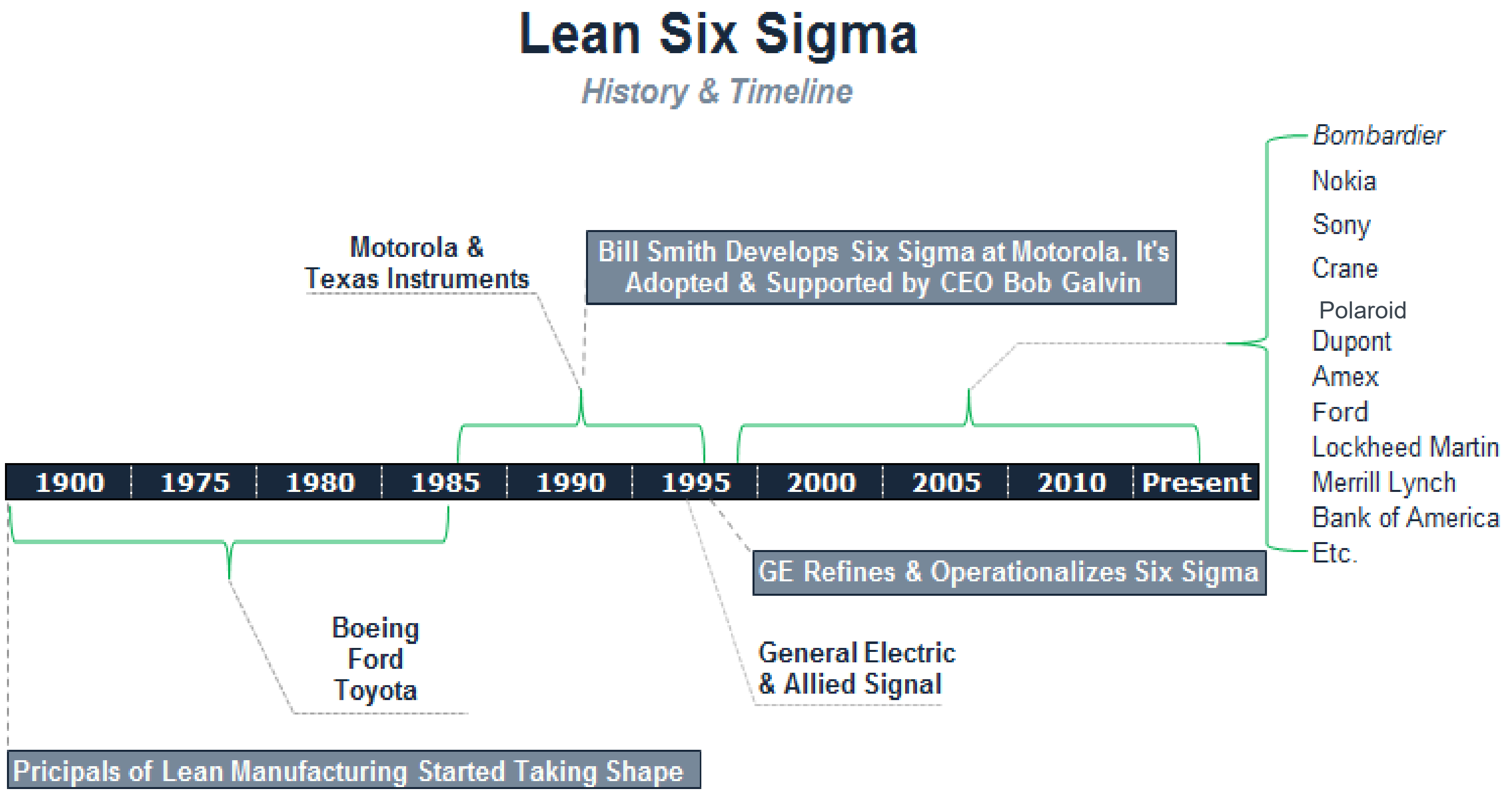
- The two work hand in hand, benefitting each other.
  - Six Sigma focuses on minimizing process variability, shifting the process average, and delivering within customer's specification limits.
  - Lean focuses on eliminating waste and increasing efficiency.

- Lean and its popularity began to form and gain significant traction in the mid 1960s with the Toyota initiative "TPS" or Toyota Production System.

- The concepts and methodology of Lean, however, were fundamentally applied much earlier by both Ford and Boeing in the early 1900s.

Six Sigma History

- Despite the criticism and immaturity of Six Sigma in many aspects, its history continues to be written with every company and organization striving to improve its business performance.



Lean Six Sigma

History & Timeline

1900 | 1975 | 1980 | 1985 | 1990 | 1995 | 2000 | 2005 | 2010 | Present

Pricipals of Lean Manufacturing Started Taking Shape

Boeing
Ford
Toyota

Motorola &
Texas Instruments

Bill Smith Develops Six Sigma at Motorola. It's Adopted & Supported by CEO Bob Galvin

General Electric
& Allied Signal

GE Refines & Operationalizes Six Sigma

Bombardier
Nokia
Sony
Crane
Polaroid
Dupont
Amex
Ford
Lockheed Martin
Merrill Lynch
Bank of America
Etc.

1.1.3 Six Sigma Approach

Six Sigma Approach: Y = f(x)

- The Six Sigma approach to problem solving uses a transfer function.
- A **transfer function** is a mathematical expression of the relationship between the inputs and outputs of a system.
- **Y = f(x)** is the relational transfer function that is used by all Six Sigma practitioners.
- It is absolutely critical that you understand and embrace this concept.

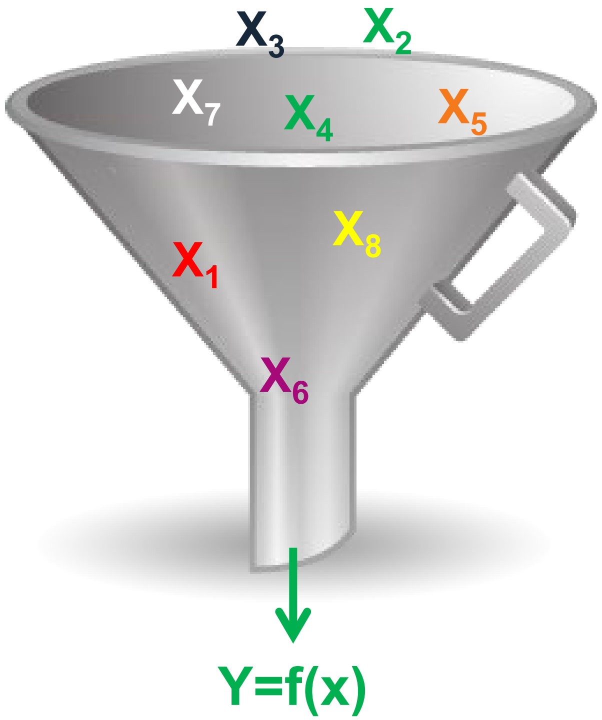
Six Sigma Approach: Y = f(x)

- "Y" refers to the measure or output of a process.
  - Y is usually your primary metric
  - Y is the measure of process performance that you are trying to improve.
- f(x) means "function of x."
  - x's are factors or inputs that affect the Y
- Combined, the Y = f(x) statement reads "Y is a function of x."
- In simple terms: "My process performance is dependent on certain x's."
- The objective in a Six Sigma project is to identify the critical x's that have the most influence on the output (Y) and adjust them so that the Y improves.

Six Sigma Approach: Y = f(x)

- Let us look at a simple example of a pizza delivery company that desires to meet customer expectations of on-time delivery.
  - Measure = on-time pizza deliveries
    - Y = percent of on-time deliveries
  - f(x) would be the x’s or factors that heavily influence timely deliveries
    - x1: might be traffic
    - x2: might be the number of deliveries per driver dispatch
    - x3: might be the accuracy of directions provided to the driver
    - x4: might be the reliability of the delivery vehicle
    - etc.
- The statement Y = f(x) in this example will refer to the proven x’s determined through the steps of a Six Sigma project.

Six Sigma Approach: Y = f(x)



- With this approach, all potential x’s are evaluated throughout the DMAIC methodology.
- The x’s should be narrowed down until the vital few x’s that significantly influence “on-time pizza deliveries” are identified!

Six Sigma Approach: Y = f(x)

- This approach to problem solving will take you through the process of determining all potential x's that **might** influence on-time deliveries and then determining through measurements and analysis which x's **do** influence on-time deliveries.
- Those significant x's become the ones used in the Y = f(x) equation.

- The Y = f(x) equation is a very powerful concept and requires the ability to measure your output and quantify your inputs.
- Measuring process inputs and outputs is crucial to effectively determining the significant influences to any process.

1.1.4 Six Sigma Methodology

Six Sigma Methodology

- Six Sigma follows a methodology that is conceptually rooted in the principles of a five-phase project.
- Each phase has a specific purpose and specific tools and techniques that aid in achieving the phase objectives.

- The 5 phases of DMAIC:
  1. **D**efine
  2. **M**easure
  3. **A**nalyze
  4. **I**mprove
  5. **C**ontrol

Six Sigma Methodology

Define
- Define and quantify the problem and objective

Measure
- Baseline process, validate measures, and identify all possible x's and dependent Y's

Analyze
- Analyze and validate causes identifying critical factors

Improve
- Develop solutions

Control
- Implement and sustain solutions

Six Sigma Methodology: Define Phase

- The goal of the **Define** phase is to establish a solid foundation and business case for a Six Sigma project.

- Define is arguably the most important aspect of any Six Sigma project.

- All successful projects start with a current state challenge or problem that can be articulated in a quantifiable manner.
  - It is not enough to just know the problem, you must quantify it and also determine the goal.

- Once problems and goals are identified and quantified, the rest of the define phase will be about valuation, team, scope, project planning, timeline, stakeholders, Voice Of the Customer (VOC), and Voice Of the Business (VOB).

Six Sigma Methodology: Define Phase

- **Define Phase Tools and Deliverables**
  - Project Charter – Establish the:
    - Business Case
    - Problem Statement
    - Project Objective
    - Project Scope
    - Project Timeline
    - Project Team.

  - Stakeholder Assessment
  - High-Level Pareto Chart Analysis
  - High-Level Process Map
  - VOC/VOB and CTQs Identified and Defined
  - Financial Assessment

Six Sigma Methodology: Measure Phase

- The goal of the **Measure** phase is to gather baseline information about the process (process performance, inputs, measurements, customer expectations etc.).
- Throughout the Measure phase you will seek to achieve a few important objectives:
  - Gather All Possible x's
  - Assess Measurement System and Data Collection Requirements
  - Validate Assumptions
  - Validate Improvement Goals
  - Determine COPQ (Cost of Poor Quality)
  - Refine Process Understanding
  - Determine Process Stability
  - Determine Process Capability.

Six Sigma Methodology: Measure Phase

- **Measure Phase Tools and Deliverables**
  - Process Maps, SIPOC, Value Stream Maps
  - Failure Modes and Effects Analysis (FMEA)
  - Cause-and-Effect Diagram
  - XY Matrix
  - Six Sigma Statistics
    - Basic Statistics
    - Descriptive Statistics
  - Measurement Systems Analysis
    - Variable and/or Attribute Gage R&R
    - Gage Linearity and Accuracy or Stability
  - Basic Control Charts
  - Process Capability (Cpk, Ppk) and Sigma Levels
  - Data Collection Plan

Six Sigma Methodology: Analyze Phase

- The **Analyze** phase is all about establishing verified drivers.

- In the DMAIC methodology, the Analyze phase uses statistics and higher-order analytics to discover relationships between process performance and process inputs (in other words, what are the root causes or drivers of the improvement effort).

- Ultimately, the Analyze phase establishes a reliable hypothesis for improvement solutions.
  - Establish the Transfer Function Y = f(x)
  - Validate the List of Critical x's and Impacts
  - Create a Beta Improvement Plan (e.g., pilot plan).

Six Sigma Methodology: Analyze Phase

- **Analyze Phase Tools and Deliverables**
  - The Analyze phase is about proving and validating critical x's using the appropriate and necessary analysis techniques. Examples include:
    - Hypothesis Testing
      - Parametric and Non-Parametric
    - Regression
      - Simple Linear Regression
      - Multiple Linear Regression
  - The Analyze phase is also about establishing a set of solution hypotheses to be tested and further validated in the Improve phase.

Six Sigma Methodology: Improve Phase

- The goal of the **Improve** phase is. . .you guessed it! "make the improvement." Improve is about designing, testing, and implementing your solution.
- To this point you have defined the problem and objective of the project, brainstormed possible x's, analyzed and verified critical x's. Now it's time to make it real!
  - Statistically Proven Results from Active Study/Pilot
  - Improvement/Implementation Plan
  - Updated Stakeholder Assessment
  - Revised Business Case with Return on Investment (ROI)
  - Risk Assessment/Updated FMEA
  - New Process Capability and Sigma.

Six Sigma Methodology: Improve Phase

- **Improve Phase Tools and Deliverables**
  - Any Appropriate Tool from Previous Phases

  - Design of Experiment (DOE)
    - Full Factorial
    - Fractional Factorial

  - Pilot or Planned Study Using:
    - Hypothesis Testing
    - Valid Measurement Systems

  - Implementation Plan

Six Sigma Methodology: Control Phase

- The last of the 5 core phases of the DMAIC methodology is the **Control** phase.

- The goal of the Control phase is to establish automated and managed mechanisms to maintain and sustain your improvement.

- A successful control plan also establishes a reaction and mitigation plan as well as an accountability structure.

Six Sigma Methodology: Control Phase

- **Control Phase Tools and Deliverables**
  - Statistical Process Control (SPC/Control Charts)
    - IMR, XbarS, XbarR, P, NP, U, C etc.

  - Control Plan Documents
    - Control Plan
    - Training Plan
    - Communication Plan
    - Audit Checklist

  - Lean Control Methods
    - Poka-Yoke
    - Five-S
    - Kanban

Six Sigma Methodology

Six Sigma DMAIC Roadmap

Define
- **Goal**: Problem Statement, Objective, Business Case, Project Scope, Team
- **Main Tools**: Project Charter, Pareto, Process Maps

Measure
- **Goal**: Brainstorm/Prioritize Possible x's, Validate measurement, Capability
- **Tools**: Basic Statistics, C & E, XY Matrix, Capability Analysis, MSA, Process Maps, Control Charts

Analyze
- **Goal**: Identify critical x's
- **Tools**: Hypotheses Tests (Normal/Non Normal), Regression and Correlation

Improve
- **Goal**: Design, Test, and Implement Improvement
- **Tools**: DOE, Implementation/Change/Communication Plan

Control
- **Goal**: Lock-in the Improvement
- **Tools**: Control Plan, Poka-Yoke, SPC, SOPs, Training Plans etc.

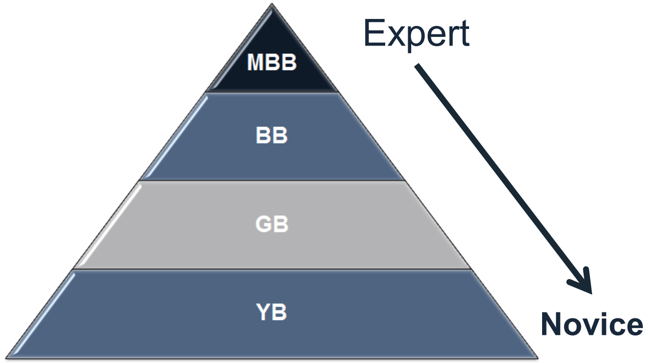
1.1.5 Roles and Responsibilities

Roles and Responsibilities

- The various roles in a Six Sigma program are commonly referred to as “Belts.”
- In addition to Belts, there are also other key roles with specific responsibilities.
- Let us explore the different roles and their corresponding responsibilities in a Six Sigma program.

Roles and Responsibilities

- Each of the four Six Sigma belts represents a different level of expertise in the field of Six Sigma.
  - Six Sigma Master Black Belt (MBB)
  - Six Sigma Black Belt (BB)
  - Six Sigma Green Belt (GB)
  - Six Sigma Yellow Belt (YB)



- In addition to Belts, there are other critical and complementary roles:
  - Champions
  - Sponsors
  - Stakeholders
  - Subject Matter Experts (SMEs).

Roles and Responsibilities: MBB

- The **Master Black Belt** (MBB) is the most experienced, educated, and capable Six Sigma expert.
- A typical MBB has managed dozens of Black Belt level projects.
- The MBB can simultaneously lead multiple Six Sigma Belt projects while mentoring and certifying Black Belt and Green Belt candidates.
- The MBB typically works with high-level operations directors, senior executives, and business managers to help with assessing and planning business strategies and tactics.

Roles and Responsibilities: MBB

- MBB commonly advises management team on the cost of poor quality of an operation and consults on methods to improve business performance.

- **Typical Responsibilities of a MBB**
  - Identifies and defines the portfolio of projects required to support a business strategy
  - Establishes scope, goals, timelines, and milestones
  - Assigns and marshals resources
  - Trains and mentors Green Belts and Black Belts
  - Facilitates tollgates or checkpoints for Belt candidates
  - Reports-out/updates stakeholders and executives
  - Establishes organization’s Six Sigma strategy/roadmap
  - Leads the implementation of Six Sigma.

Roles and Responsibilities: BB

- The **Black Belt** (BB) is the most active and valuable experienced Six Sigma professional among all the Six Sigma Belts.

- A typical BB has
  - led multiple projects
  - trained and mentored various Green Belts candidates
  - understood how to define a problem and drive effective solution.

- The BB is well rounded in terms of project management, statistical analysis, financial analysis, meeting facilitation, prioritization, and a range of other value-added capabilities, which makes a BB highly valuable asset in the business world.

Roles and Responsibilities: BB

- BBs commonly serve as the dedicated resource continuing their line management role while simultaneously achieving a BB certification.

- **Typical Responsibilities of a BB**
  - Project Management
    - Defines projects, scope, teams etc.
    - Marshals resources
    - Establishes goals, timelines, and milestones
    - Provides reports and/or updates to stakeholders and executives.

Roles and Responsibilities: BB

- **Typical Responsibilities of a BB** *(continued)*
  - Task Management
    - Establishes the team’s Lean Sigma roadmap
    - Plans and implements the use of Lean Sigma tools
    - Facilitates project meetings
    - Does project management of the team’s work
    - Manages progress toward objectives.

  - Team Management
    - Chooses or recommend team members
    - Defines ground rules for the project team
    - Coaches, mentors, and directs project team
    - Coaches other Six Sigma Belts
    - Manages the team’s organizational interfaces.

Roles and Responsibilities: GB

- The **Green Belt** (GB) is considered as a less intense version of Six Sigma professional than the Black Belt (BB).

- A GB is exposed to all the comprehensive aspects of Six Sigma with less focus on the statistical theories and some other advanced analytical methodologies such as Design of Experiment (DOE).

- When it comes to project management, a GB has almost the same responsibilities as a BB.

- In general, the GB works on less complicated and challenging business problems than a BB.

Roles and Responsibilities: GB

- **Typical Responsibilities of a Green Belt**
  - Project Management
    - Defines the project, scope, team etc.
    - Marshals resources
    - Sets goals, timelines, and milestones
    - Reports-out/updates stakeholders and executives.
  - Task Management
    - Establishes the team’s Lean Sigma Roadmap
    - Plans and implements the use of Lean Sigma tools
    - Facilitates project meetings
    - Does Project Management of the team’s work
    - Manages progress toward objectives.
  - Team Management
    - Chooses or recommends team members
    - Defines ground rules for the project team
    - Coaches, mentors, and directs project team
    - Coaches other Six Sigma Belts
    - Manages the team’s organizational interfaces.

Roles and Responsibilities: YB

- The **Yellow Belt** (YB) understands the basic objectives and methods of a Six Sigma project.
- YB has an elementary understanding about what other Six Sigma Belts (GB, BB, MBB) are doing to help them succeed.
- In a Six Sigma project, YB usually serves as a subject matter expert regarding some aspects of the process or project.
- Supervisors, managers, directors, and sometimes executives are usually trained at the YB level.

Roles and Responsibilities: YB

- **Typical Responsibilities of a Yellow Belt**
  - Helps define process scope and parameters
  - Contributes to team selection process
  - Assists in information and data collection
  - Participates in experiential analysis sessions (FMEA, Process Mapping, Cause and Effect etc.)
  - Assists in assessing and developing solutions
  - Delivers solution implementations.

Roles and Responsibilities: Champions & Sponsors

- **Champions and sponsors** are those individuals (directors, executives, managers etc.) chartering, funding, or driving the Six Sigma projects that BBs and GBs are conducting.

- Champions and sponsors need to have a basic understanding of the concepts, tools, and techniques involved in the DMAIC methodology so that they can provide proper support and direction.

Roles and Responsibilities: Champions & Sponsors

- Champions and sponsors play critical roles in the successful deployment of Six Sigma.
- Strong endorsement of Six Sigma from the leadership team is critical for success.

- **Typical Responsibilities of a Champion/Sponsor**
  - Maintains a strategic oversight
  - Establishes strategy and direction for a portfolio of projects
  - Clearly defines success
  - Provides resolution for issues such as resources or politics
  - Establishes routine tollgates or project reviews
  - Clears the path for solution implementation
  - Assists in project team formation.

Roles and Responsibilities: Stakeholders

- **Stakeholders** are usually the recipients or beneficiaries of the success of a Six Sigma project.

- Stakeholders are individuals owning the process, function, or production/service line that a Six Sigma Belt focuses on improving the performance of.

- BBs and GBs need to keep strong working relationships with stakeholders because without their support, it would be extremely difficult to make the Six Sigma project a success.

Roles and Responsibilities: SMEs

- **Subject Matter Experts** (SMEs) are commonly known as the experts of the process or subject matter.

- Six Sigma Belts should proactively look to key SMEs to round out their working project team.

- SMEs play critical roles to the success of a project.
  - Based on SMEs' extensive knowledge about the process, they have the experience to identify which solutions can work and which cannot work.
  - SMEs who simply do not speak up can hurt the chances of the process' success.
  - SMEs are also the same people who prefer to keep the status quo. Six Sigma Belts may find many of them unwilling to help implement the changes.

Roles and Responsibilities

- Throughout this module we have reviewed the various common roles and corresponding responsibilities in any Six Sigma program:
  - Six Sigma Master Black Belt
  - Six Sigma Black Belt
  - Six Sigma Green Belt
  - Six Sigma Yellow Belt
  - Champion and Sponsors
  - Stakeholders
  - Subject Matter Experts (SMEs)

- These Six Sigma belts and other roles are designed to deliver value to the business effectively and successfully.

1.2. Six Sigma Fundamentals

Black Belt Training: Define Phase

1.1 Six Sigma Overview
- 1.1.1 What is Six Sigma
- 1.1.2 Six Sigma History
- 1.1.3 Six Sigma Approach Y = f(x)
- 1.1.4 Six Sigma Methodology
- 1.1.5 Roles and Responsibilities

1.2 Six Sigma Fundamentals
- 1.2.1 Defining a Process
- 1.2.2 VOC & CTQs
- 1.2.3 QFD
- 1.2.4 Cost of Poor Quality (COPQ)
- 1.2.5 Pareto Analysis (80:20 rule)

1.3 Lean Six Sigma Projects
- 1.3.1 Six Sigma Metrics
- 1.3.2 Business Case and Charter
- 1.3.3 Project Team Selection
- 1.3.4 Project Risk Management
- 1.3.5 Project Planning

1.4 Lean Fundamentals
- 1.4.1 Lean and Six Sigma
- 1.4.2 History of Lean
- 1.4.3 The Seven Deadly Muda
- 1.4.4 Five-S (5S)

1.2.1 Defining a Process

Defining a Process

- The basic method of defining and understanding a process is the **process map**.
- Process maps help determine where and how a process begins as well as all the steps and decisions in between.
- By learning the various types and methods of process maps, you can become adept at setting project scopes, identifying value-added and non-value-added steps, identifying problems in a process, etc.
- This module covers:
  - High-level process maps
  - Detailed process maps
  - Functional maps.
- In the Measure section we will touch on several other types and methods of process mapping.

What is a Process Map?

- A **process map** is a graphical representation of a process flow.
- It illustrates how the business process is accomplished step by step.
- It describes how the materials or information sequentially flow from one business entity to the next.
- It illustrates who is responsible for what between the process boundaries.
- It depicts the inputs and outputs of each individual process step.

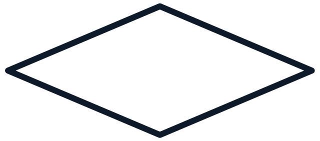
- Always encourage your project team to map the current state of the process instead of the ideal state. Be honest with each other!

Process Map Basic Symbols

- The following four symbols are the most commonly used symbols in a process map.

Terminator (Oval):
Shows the start and end points in the process.

Process (Rectangle):
Indicates a single process step.



Decision (Diamond):
Indicates a question with two choices (e.g. Yes/No)

Flow Line (Arrow):
Shows the direction of the process flow.

Additional Process Symbols

- Additional Process Symbols:

Alternative Process:
Indicates a process step as an alternate of a normal step.

Predefined Process:
Indicates a formally-defined process step. Other documentation or instruction is needed to support further details of the step.

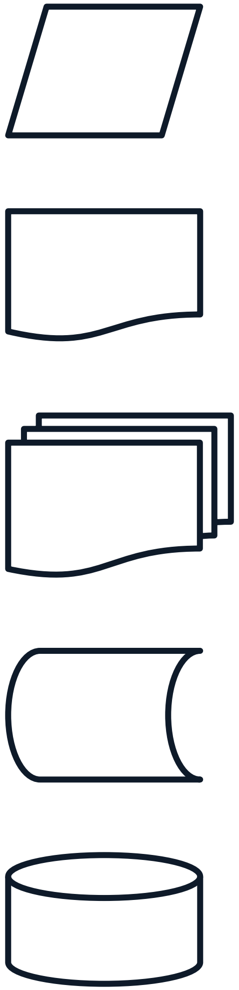
Manual Operation:
Indicates a process step conducted manually.

Preparation:
Indicates a preparation step.

Delay:
Indicates a waiting period in the process.

Additional Process Symbols

- Additional File and Information Related Symbols:



Data (I/O):
Shows the inputs and outputs of a process.

Document:
Indicates a process step that results in a document.

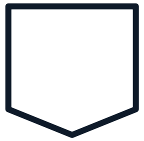
Multi-Document:
Indicates a process step that results in multiple documents.

Stored Data:
Indicates a process step that stores data.

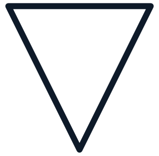
Magnetic Disk:
Indicates a database.

Additional Process Symbols

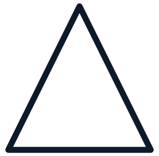
- Additional Control of Flow Symbols:



Off-Page Connector:
Indicates the process flow continues onto another page.



Merge:
Indicates multiple processes merge into one.



Extract:
Indicates a process splits into multiple parallel processes.

Or:
Indicates a single data processing flow diverges to multiple branches with different criteria requirements.

Summing Junction:
Indicates multiple data processing flows converge into one.

How to Plot a Process Map

- Step 1: Define the boundaries of the process you want to map.
  - A process map can depict the flow of an entire process or a segment of it.
  - You need to identify and define the beginning and ending points of the process before starting to plot.
  - Use operational definitions where possible.

How to Plot a Process Map

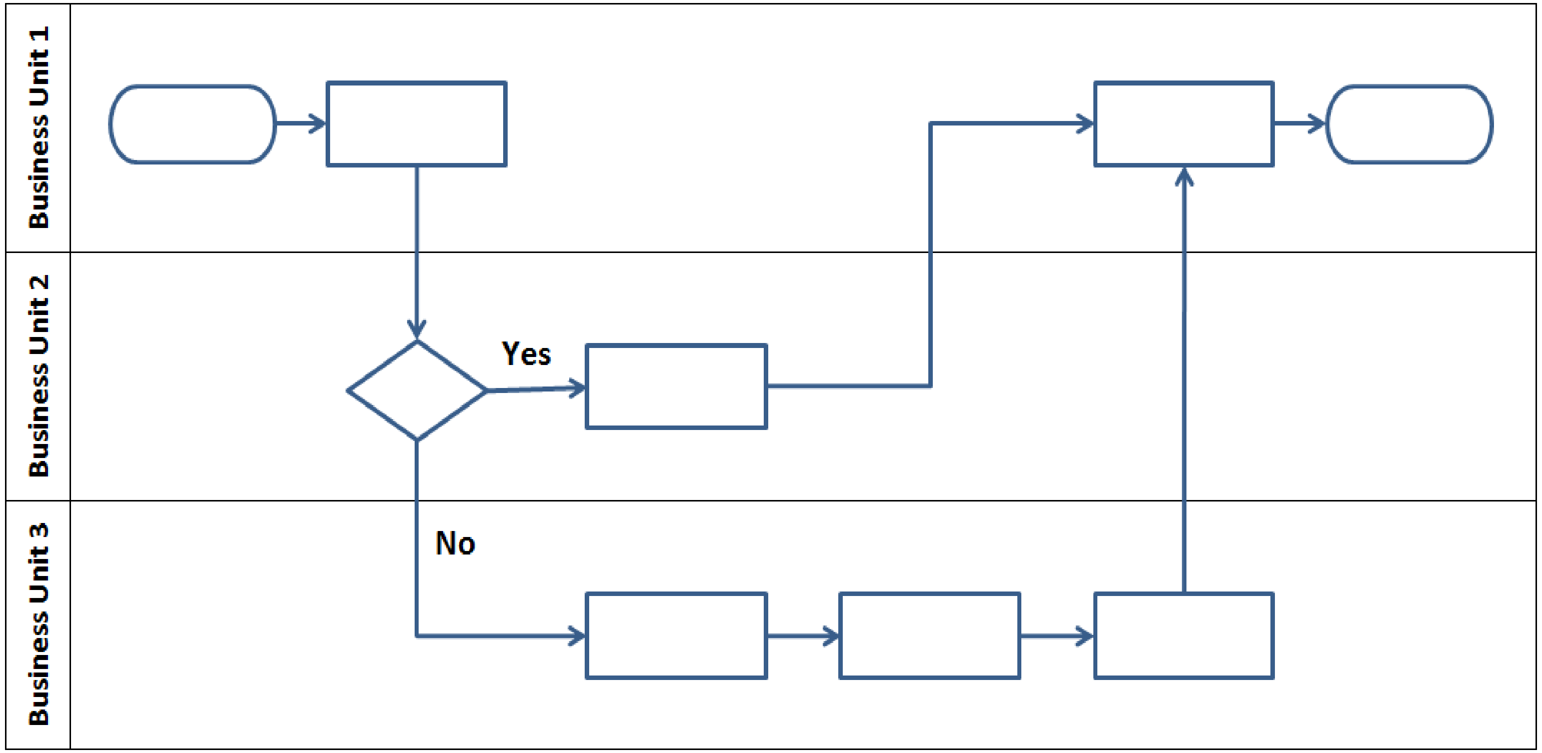
- Step 2: Define and sort the process steps with the flow.
  - Consult with process owners and SMEs or observe the process in action to understand how the process is actually performed.
  - Record the process steps and sort them according to the order of their occurrence.

How to Plot a Process Map

- Step 3: Fill the step information into the appropriate process symbols and plot the diagram.
  - In the team meeting of process mapping, place the sticky notes with different colors on a white board so you can move them around while the map is under-construction.
  - The flow lines can be plotted directly on the white board.
  - Decision steps. Rotate the sticky note 45 degrees.
  - When the map is completed on the white board, record the map using Excel, PowerPoint, Visio, Quality Companion, or other preferred software.

How to Plot a Process Map

- Step 3:
  - To illustrate the responsibility of different organizations involved in the process, use a Swim Lane Process Map.



How to Plot a Process Map

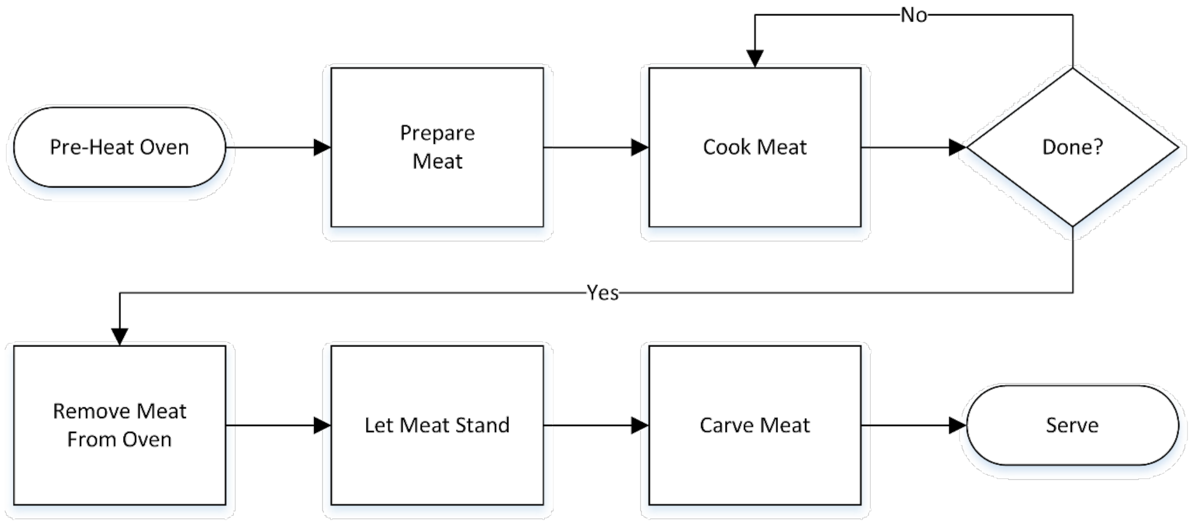
- Step 4: Identify and record inputs/outputs and their corresponding specifications for each process step.
  - The process map helps in understanding and documenting Y = f(x) of a process, where Y represents the outputs and x represents the inputs.
  - The inputs of each process step can be controllable or non-controllable, standardized operational procedures, or noise.
  - Inputs are the source of variation in the process and need to be analyzed qualitatively and quantitatively.
  - The outputs of each process step can be products, information, services, etc. They are the little Y’s within the process.

How to Plot a Process Map

- Step 5: Evaluate the process map and adjust if needed.
  - If the process is too complicated to be covered in one single process map, you may create additional detailed sub-process maps for further information.
  - Number the process steps in the order of their occurrence for clarity.

High Level Process Map

- Most high-level process maps are also referred to as **flow charts**.
- The key to a high-level process map is to over-simplify the process being depicted so that it can be understood in its most generic form.
- As a general rule, high-level process maps should be no more than 4–6 steps.
- Below is an oversimplified version of a high-level process map for cooking a 10lb prime rib for a dozen holiday guests.



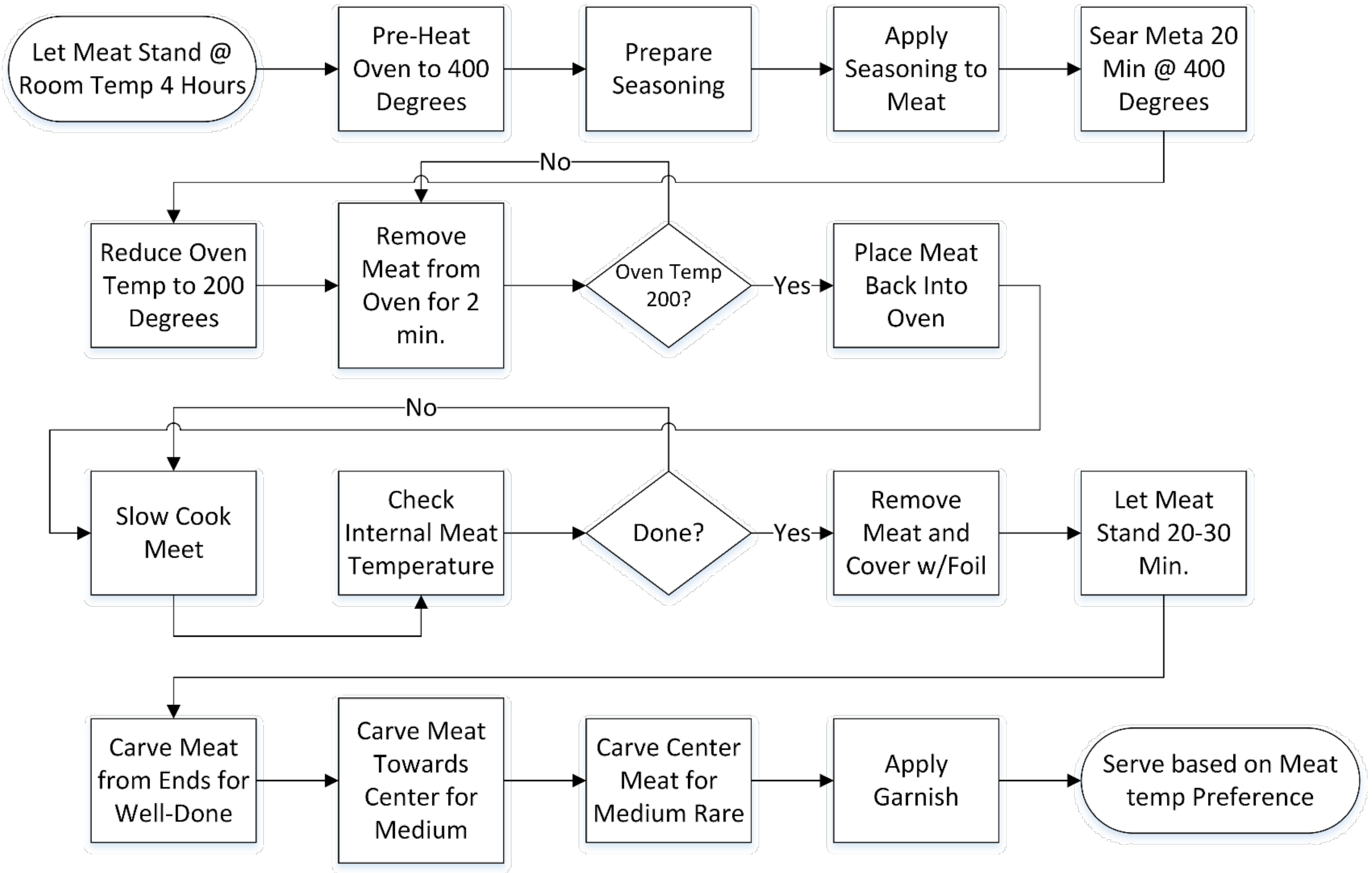
Detailed Process Map

- Detailed process maps or multi-level maps take the high-level map much further.
- Detailed maps can be 2–4 levels deeper than your high-level process map.
- A good guideline used to help create the second level is to take each step in the high level map and break it down into 2–4 steps (no more).

- Repeat this process (level 3, level 4 etc.) until reaching the desired level of detail.
- Some detailed maps are 2 or 3 levels deep, others can be 5–6 levels deep. Obviously, the deeper the levels, the more complex and the more burdensome.

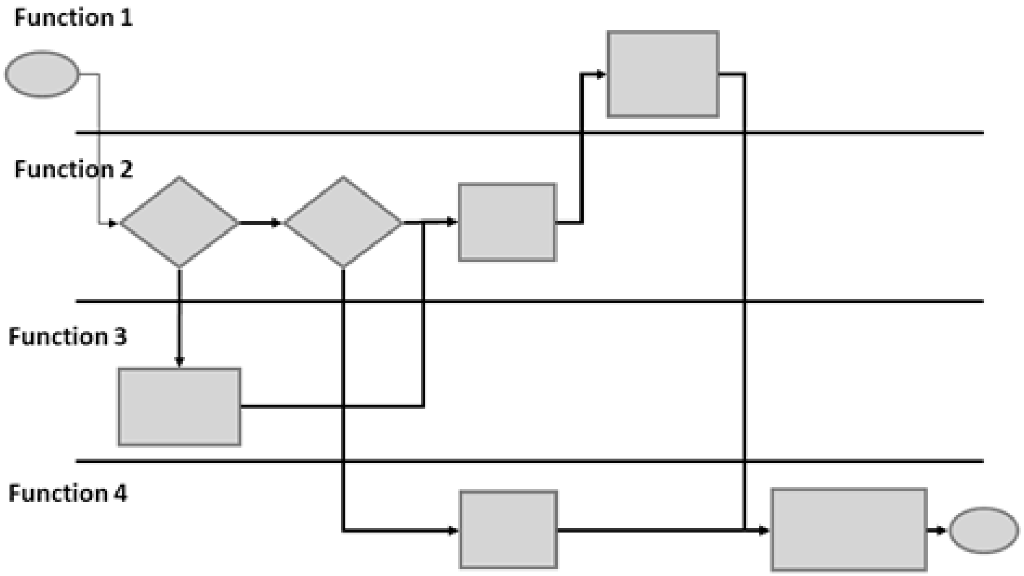
Detailed Process Map

- At right is our prime rib cooking example at **level 2 detail**.
- This process map has a few more decision points and process steps.
- You can see that going only one more level deep adds a fair amount of information to the process map.



Functional Process Map

- The functional map adds dimension to the high-level or detailed map.
- The dimension added is identifying which function or job performs the step or makes the decision.
- Below is a generic example of a functional map. Note that functions are identified in horizontal lanes and each process step is placed in the appropriate lane based on which function performs the step.



1.2.2 VOC and CTQs

Voice of the Customer

- **VOC** stands for “Voice of the Customer.”
- Voice of the customer is a term used for a data-driven plan to discover customer wants and needs.
- VOC is an important component to a successful Six Sigma project.
- There are also other “Voices” that need to be heard when conducting projects. The 3 primary forms are:
  - VOC: Voice of the Customer
  - VOB: Voice of the Business
  - VOA: Voice of the Associate.

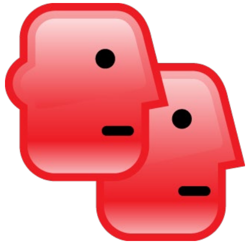
Gathering VOC

- Gathering VOC should be performed methodically.
- The two most popular methods of collecting VOC are
  1. Indirect
  2. Direct.
- 1. Indirect data collection for VOC involves passive information exchange:
  - Warranty claims
  - Customer complaints/compliments
  - Service calls
  - Sales reports.

- 2. Indirect methods are less effective, sometimes dated, require heavy interpretation, and are also more difficult to confirm.

Gathering VOC

- Direct data collection methods for VOC are active and planned customer engagements:
  - Conducting interviews
  - Conducting customer surveys
  - Conducting market research
  - Hosting focus groups.



- Direct methods are typically more effective for several reasons:
  - Less need to interpret meaning
  - Researchers can go a little deeper when interacting with customers
  - Customers are aware of their participation and will respond better upon follow-up
  - Researchers can properly plan engagements (questions, sample size, information collection techniques etc.).

Gathering VOC

- Gathering VOC requires consideration of many factors such as product or services types, customer segments, manufacturing methods or facilities etc.
- All this information will influence the sampling strategy.

- Consider which factors are important and build a sample size plan around them.
- Also, consider response rates and adjust the initial sample strategy to ensure adequate input is received.

- Once a sampling plan is in place, collect data via the direct and indirect methods discussed earlier.
- After gathering VOC it will be necessary to translate it into something meaningful: CTQs.

Critical to Quality: CTQ

- **CTQ** stands for Critical to Quality.
- CTQs are translated from VOC or "voice of the customer" feedback.
- VOC is often vague, emotional, or simply a generalization about products or services.
- CTQs are the quantifiable, measureable, and meaningful translations of VOC.
- Organizing VOC helps to identify CTQs.

- One effective way to organize VOC is to group or bucket it using an **affinity diagram**.
- Affinity diagrams are ideal for large amounts of soft data resulting from brainstorming sessions or surveys.

Affinity Diagram: Building a CTQ Tree

- Steps for conducting an Affinity Diagram exercise:
    - Step 1: Clearly define the question or focus of the exercise (“Why are associates late for work?”).
    - Step 2: Record all participant responses on note cards or sticky notes (this is the sloppy part, record everything!).
    - Step 3: Lay out all note cards or post the sticky notes onto a wall.
    - Step 4: Look for and identify common themes.
    - Step 5: Begin moving the note cards or sticky notes into the themes until all responses are allocated.
    - Step 6: Re-evaluate and make adjustments.

Affinity Diagram: Building a CTQ Tree

- Define the question or focus
- Record responses on note cards or sticky notes
- Display all note cards or sticky notes on a wall if necessary.

Why are Associates late to work?

Slept Late

Had to take kids to school

Traffic Accident

Long line at Starbucks

Traffic

Stop light was out

Couldn't find a parking place

Kids were sick

Power went out

Long wait for the elevator

Dr. Appointment

Dog threw up

Forgot to set alarm

Affinity Diagram: Building a CTQ Tree

- Look for and identify common themes within the responses.

Why are Associates late to work?

Delays at Home

Delays on the Commute

Kid/Pet Related Delays

Delays in the building

Slept Late

Had to take kids to school

Traffic Accident

Long line at Starbucks

Traffic

Stop light was out

Couldn’t find a parking place

Kids were sick

Power went out

Long wait for the elevator

Dr. Appointment

Dog threw up

Forgot to set alarm

Affinity Diagram: Building a CTQ Tree

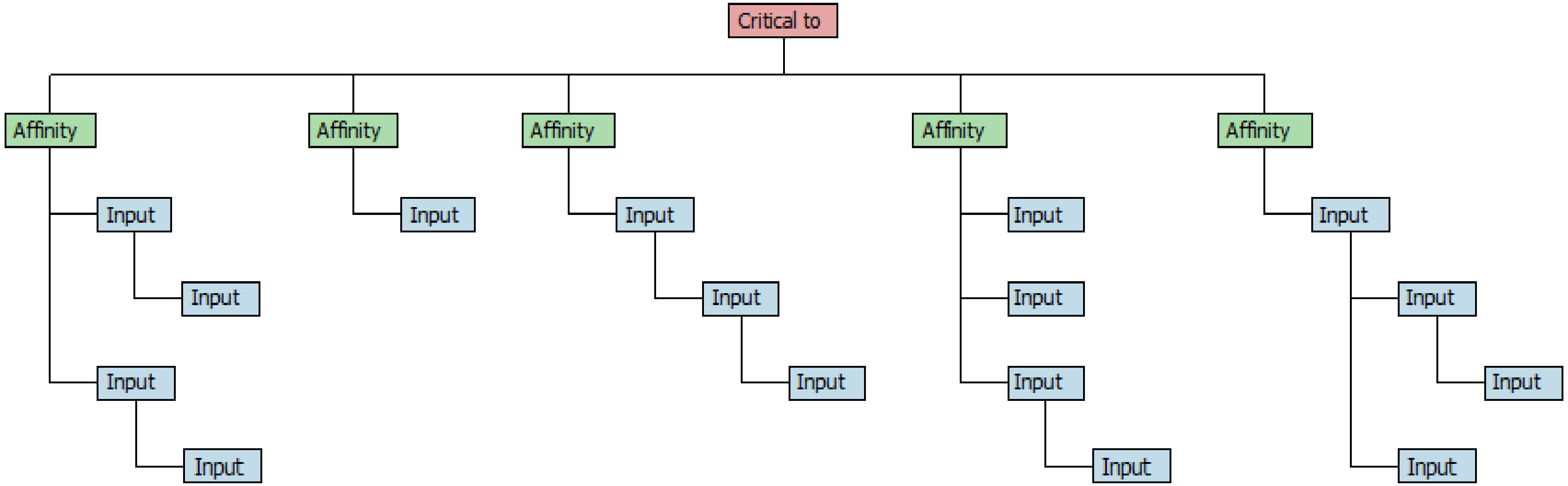
- Group note cards or sticky notes into themes until all responses are allocated.
- Re-evaluate and make final adjustments.

Why are Associates late to work?

Delays at Home	Delays on the Commute	Kid/Pet Related Delays	Delays in the building
Slept Late	Traffic Accident	Had to take kids to school	Long line at Starbucks
Power went out	Stop light was out	Kids were sick	Couldn't find a parking place
Dog threw up	Traffic	Dr. Appointment	Long wait for the elevator
Forgot to set alarm			

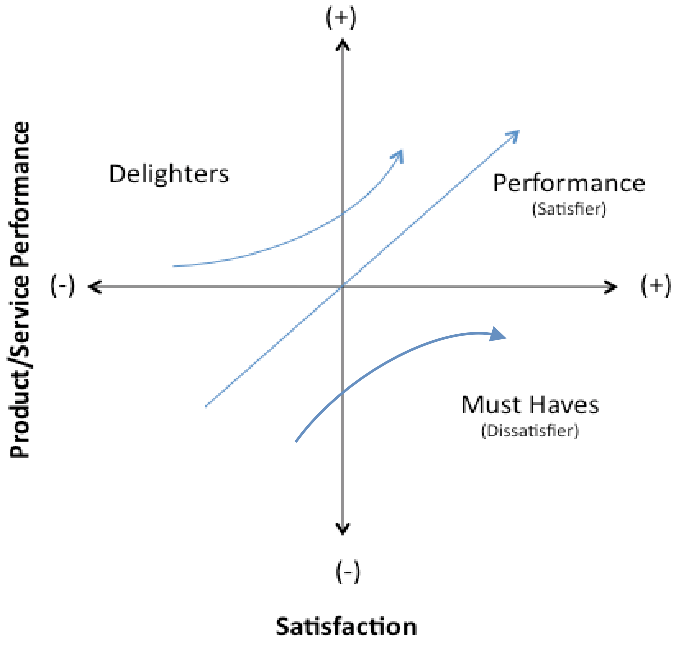
CTQ Tree

- Example of a generic CTQ tree transposed from a white board to a software package.



Kano

- Another VOC categorization technique is the Kano.
- The Kano model was developed by Noriaki Kano in the 1980s.
- The Kano model is a graphic tool that further categorizes VOC and CTQs into 3
  distinct groups:
  - Must Haves
  - Performance Attributes
  - Delighters.



- The Kano helps to identify CTQs that add incremental value vs. those that are simply requirements and having more is not necessarily better.

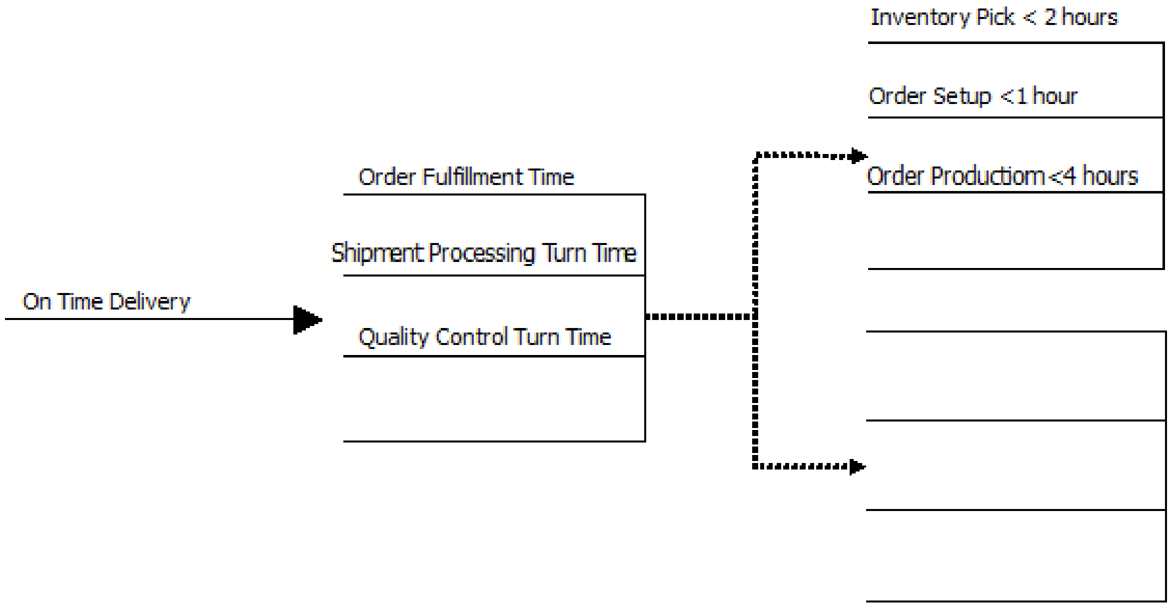
Validating VOC and CTQs

- After determining all CTQs, confirm them with the customer.
- Confirming can be accomplished by conducting surveys through one or more of the following methods:
  - Group sessions
  - One-on-one meetings
  - Phone interviews
  - Electronic means (chat, email, social media etc.)
  - Physical mail.
- Consider your confirming audience and try to avoid factors that may influence or bias responses such as inconvenience or overly burdensome time commitments.

Translating CTQs to Requirements

- Lastly, CTQs must be transformed into **specifics** that can be built upon in a process.

- A **requirements tree** translates CTQs to meaningful and measureable requirements for production processes and products.



1.2.3 Quality Function Deployment

History of QFD

- Developed by Shigeru Mizuno (1910–1989) and Yoji Akao (b. 1928) in Japan. Quality Function Deployment (QFD) aims to design products that assure customer satisfaction and value – the first time and every time.

- The QFD framework can be used for translating actual customer statements and needs ("The voice of the customer") into actions and designs to build and deliver a quality product.

What is QFD?

- **Quality Function Deployment (QFD)** is a construction methodology and quantification tool used to identify and measure customer’s requirements and transform them into meaningful and measureable parameters.

- QFD helps to prioritize actions to advance process or product to meet customer’s anticipations.

- QFD is an excellent tool for contact between cross-functional groups.

Purpose of QFD

The quality function deployment has many purposes. Among the most important are:

- Market analysis to establish needs and expectations
- Examination of competitors’ abilities
- Identification of key factors for success
- Translation of key factors into product and process characteristics.

Phases of QFD

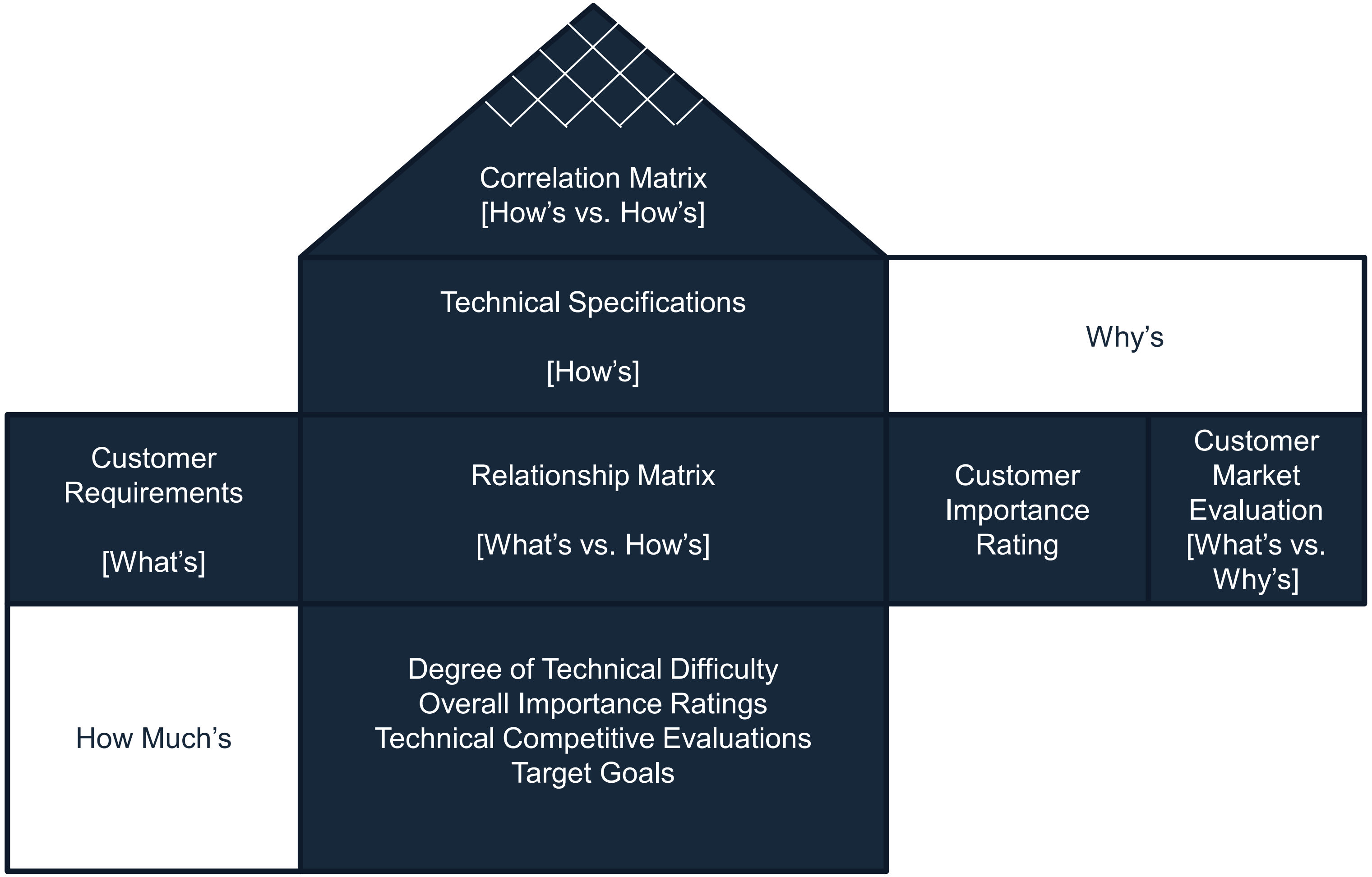
Four Key Phases of QFD

- **Phase I:** Product Planning Including the “House of Quality” (Requirements Engineering Life Cycle)
- **Phase II:** Product Design (Design Life Cycles)
- **Phase III:** Process Planning (Implementation Life Cycle)
- **Phase IV:** Process Control (Testing Life Cycle)

How to build a House of Quality

- Determine Customer Requirements (*What's* from VOC/CTQ)
- Technical Specifications/Design Requirements (*How's*)
- Develop Relationship Matrix (*What's* and *How's*)
- Prioritize Customer Requirements
- Conduct Competitive Assessments
- Develop Interrelationship (*How's*)
- Prioritize Design Requirements

House of Quality



Correlation Matrix
[How's vs. How's]

Technical Specifications

[How's]

Why's

Customer Requirements

[What's]

Relationship Matrix

[What's vs. How's]

Customer Importance Rating

Customer Market Evaluation [What's vs. Why's]

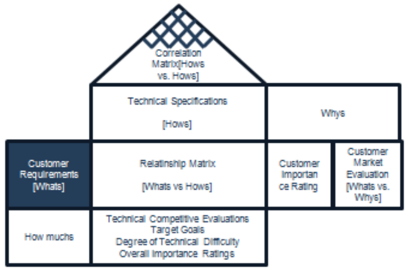
How Much's

Degree of Technical Difficulty
Overall Importance Ratings
Technical Competitive Evaluations
Target Goals

Step 1: Determine Customer Requirements

- Identify the important customer requirements. These are the “What’s” and are typically determined through the VOC/CTQ process.

- Use the results from your requirements tree diagram as inputs for the customer requirements in your HOQ.

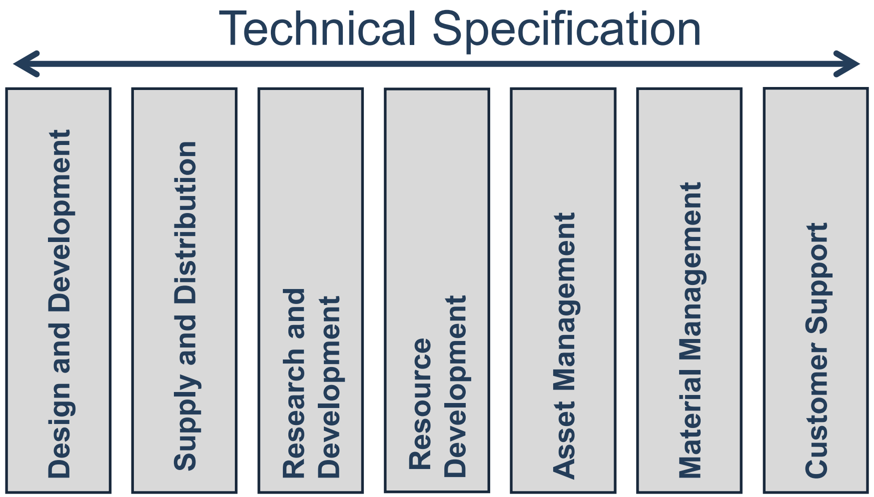



Customer Requirements

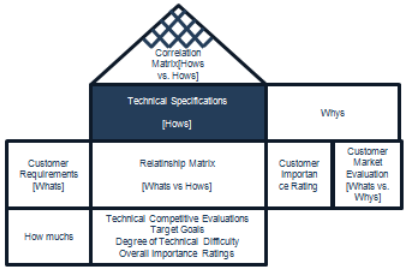
- On-Time Delivery
- Increase Product Quality
- Alternative Solutions
- Reduce Turn-Around Time
- Optimize Resource
- Reduce COPQ
- High Productivity

Step 2: Technical Specification

- Potential choices for product features
- Voice of Designers or Engineers
- Each “What” item must be refined to “How’s”

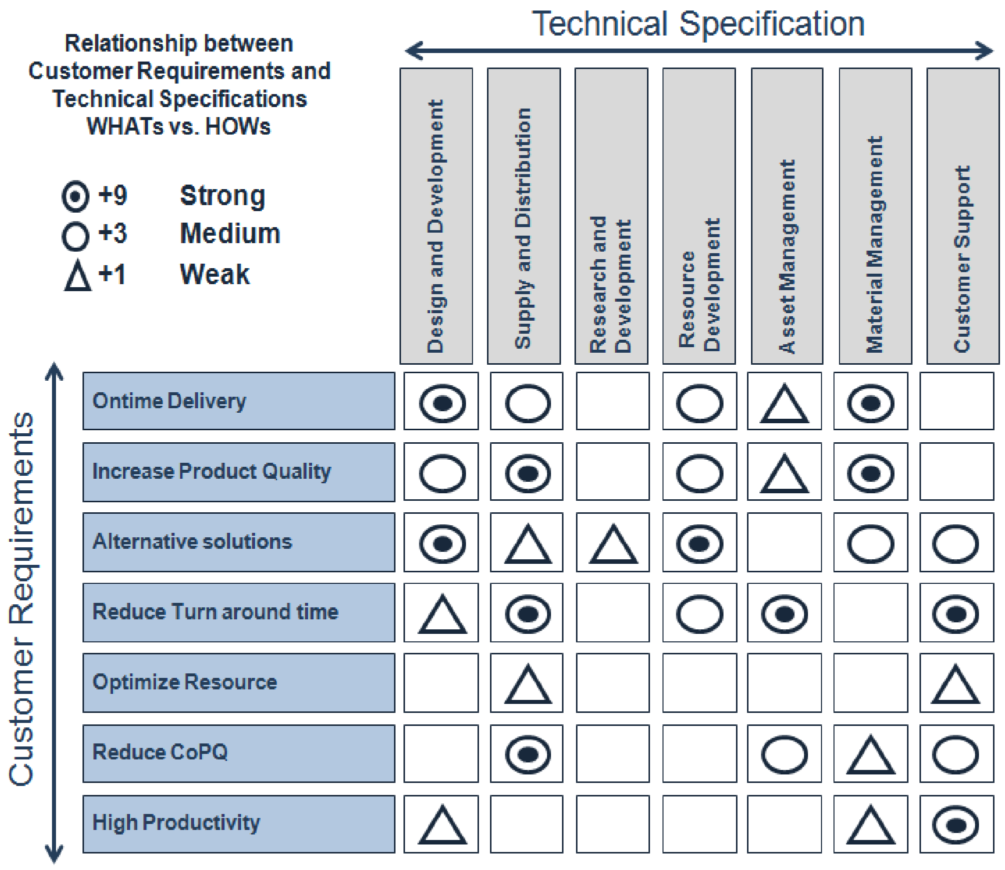

Technical Specification

- Design and Development
- Supply and Distribution
- Research and Development
- Resource Development
- Asset Management
- Material Management
- Customer Support



Step 3: Develop Relationship Matrix (What’s & How’s)

- This is the center portion of the house. Each cell represents how each technical specification relates to each customer requirement.



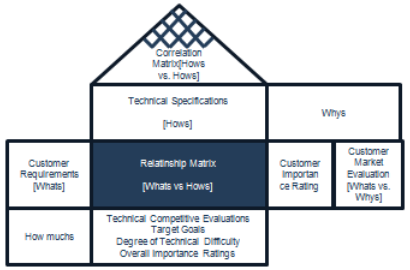
Relationship between Customer Requirements and Technical Specifications WHATs vs. HOWs

⦿ +9 **Strong**
○ +3 **Medium**
△ +1 **Weak**

Technical Specification

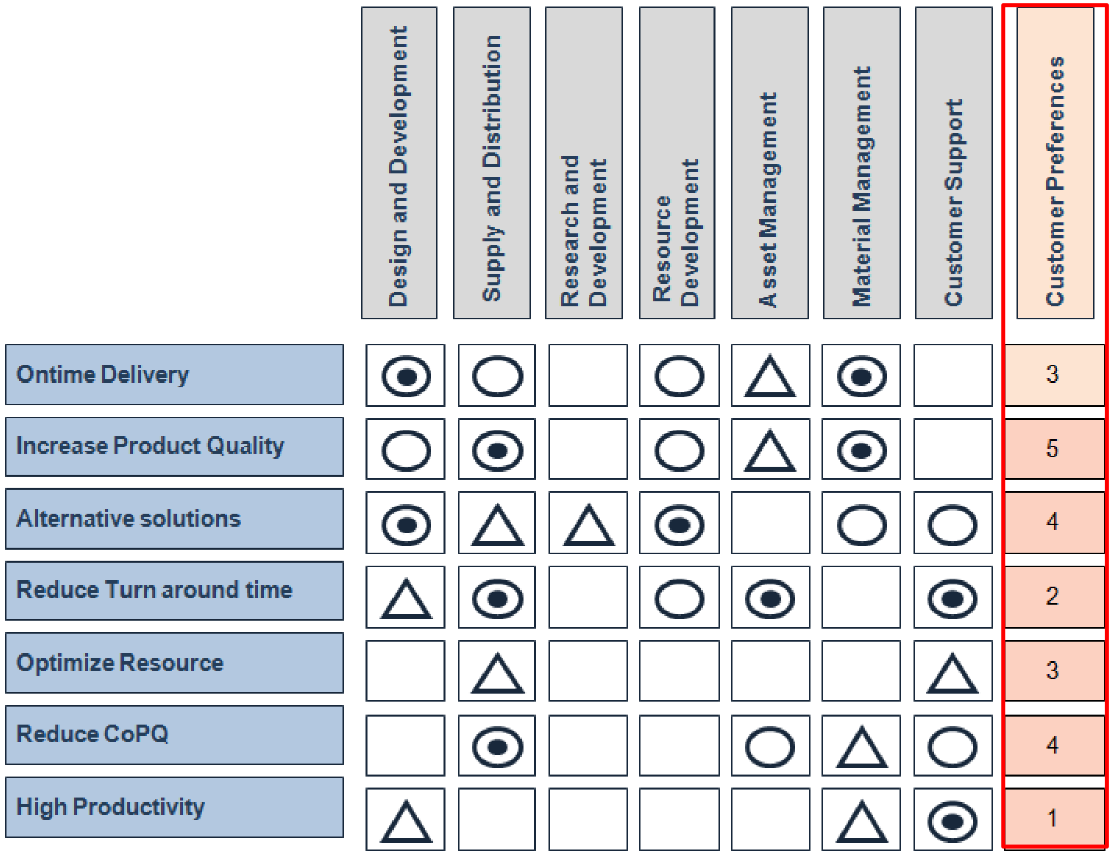
Customer Requirements

	Design and Development	Supply and Distribution	Research and Development	Resource Development	Asset Management	Material Management	Customer Support
Ontime Delivery	⦿	○		○	△	⦿	
Increase Product Quality	○	⦿		○	△	⦿	
Alternative solutions	⦿	△	△	⦿		○	○
Reduce Turn around time	△	⦿		○	⦿		⦿
Optimize Resource		△					△
Reduce CoPQ		⦿			○	△	○
High Productivity	△					△	⦿

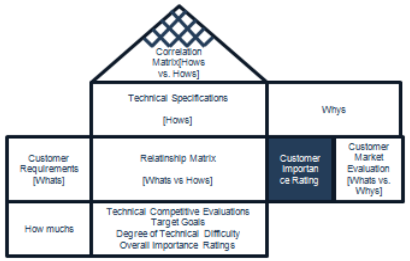


Step 4: Prioritize Customer Requirements

- This is the right portion of the house. Each cell represents customer requirements based on relative importance to customers and perceptions of competitive performance.

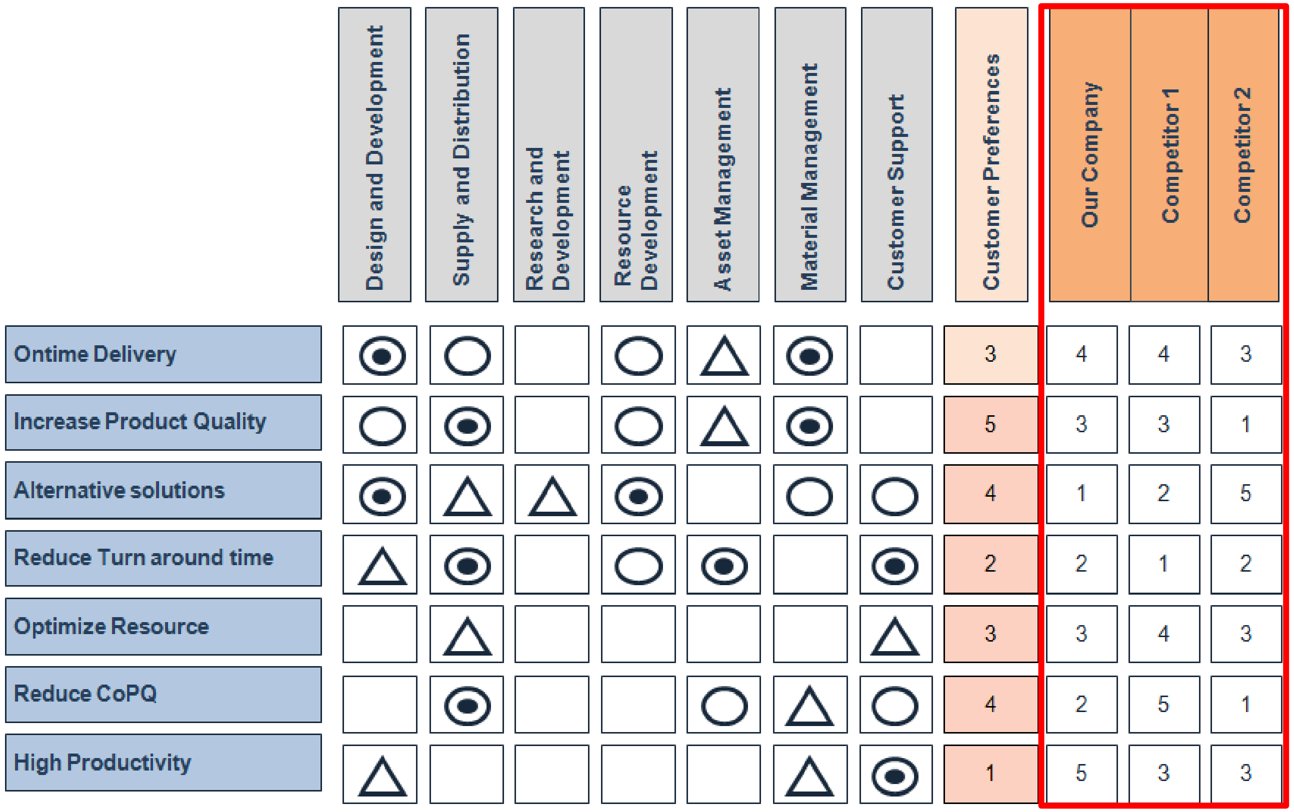


	Design and Development	Supply and Distribution	Research and Development	Resource Development	Asset Management	Material Management	Customer Support	Customer Preferences
Ontime Delivery	⊙	○		○	△	⊙		3
Increase Product Quality	○	⊙		○	△	⊙		5
Alternative solutions	⊙	△	△	⊙		○	○	4
Reduce Turn around time	△	⊙		○	⊙		⊙	2
Optimize Resource		△					△	3
Reduce CoPQ		⊙			○	△	○	4
High Productivity	△					△	⊙	1

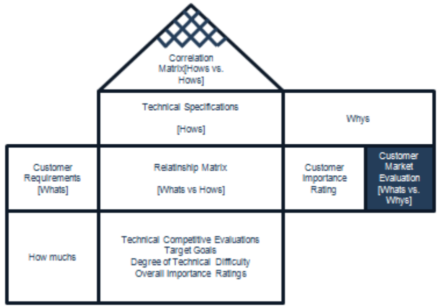


Step 5: Competitive Assessments

- This is the extreme right portion of the house. Comparison of the organization's product to competitors' products.



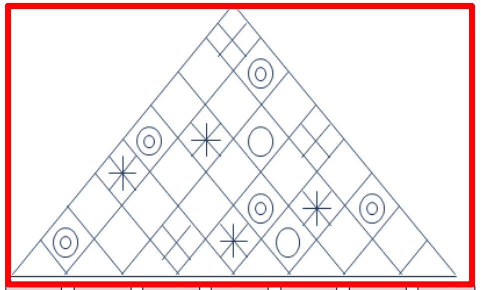
	Design and Development	Supply and Distribution	Research and Development	Resource Development	Asset Management	Material Management	Customer Support	Customer Preferences	Our Company	Competitor 1	Competitor 2
Ontime Delivery	◉	○		○	△	◉		3	4	4	3
Increase Product Quality	○	◉		○	△	◉		5	3	3	1
Alternative solutions	◉	△	△	◉		○	○	4	1	2	5
Reduce Turn around time	△	◉		○	◉		◉	2	2	1	2
Optimize Resource		△					△	3	3	4	3
Reduce CoPQ		◉			○	△	○	4	2	5	1
High Productivity	△					△	◉	1	5	3	3



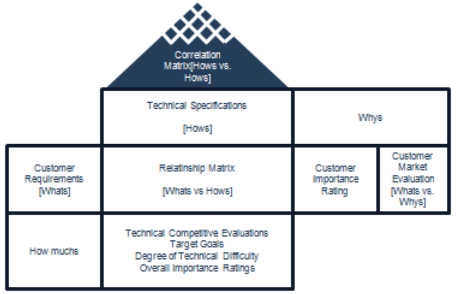
Step 6: Correlation Matrix

- This is the top portion of the house. It identifies the way "how" items either support (positive) or conflict (negative) with one another.

+9 ◎ Strong Positive
+3 ○ Positive
-3 × Negative
-9 ✳ Strong Negative

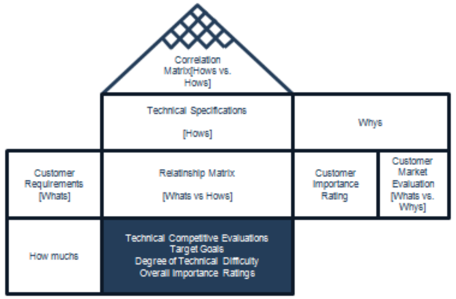


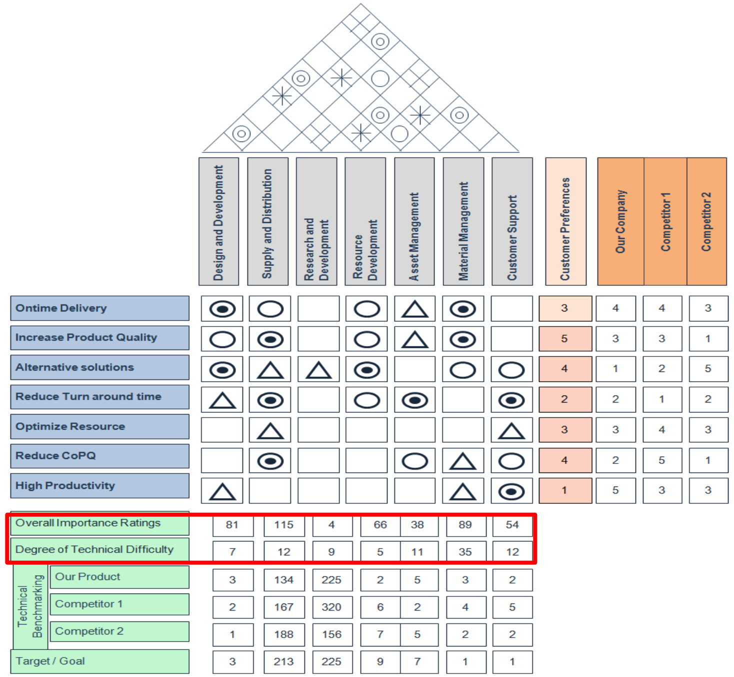
	Design and Development	Supply and Distribution	Research and Development	Resource Development	Asset Management	Material Management	Customer Support	Customer Preferences	Our Company	Competitor 1	Competitor 2
Ontime Delivery	◉	○		○	△	◉		3	4	4	3
Increase Product Quality	○	◉		○	△	◉		5	3	3	1
Alternative solutions	◉	△	△	◉		○	○	4	1	2	5
Reduce Turn around time	△	◉		○	◉		◉	2	2	1	2
Optimize Resource		△					△	3	3	4	3
Reduce CoPQ		◉			○	△	○	4	2	5	1
High Productivity	△					△	◉	1	5	3	3



Step 7: Prioritize Design Requirements

- Overall Importance Ratings
- Function of relationship ratings and customer prioritization ratings
- Technical Difficulty Assessment
- Similar to customer market competitive evaluations but conducted by the technical team

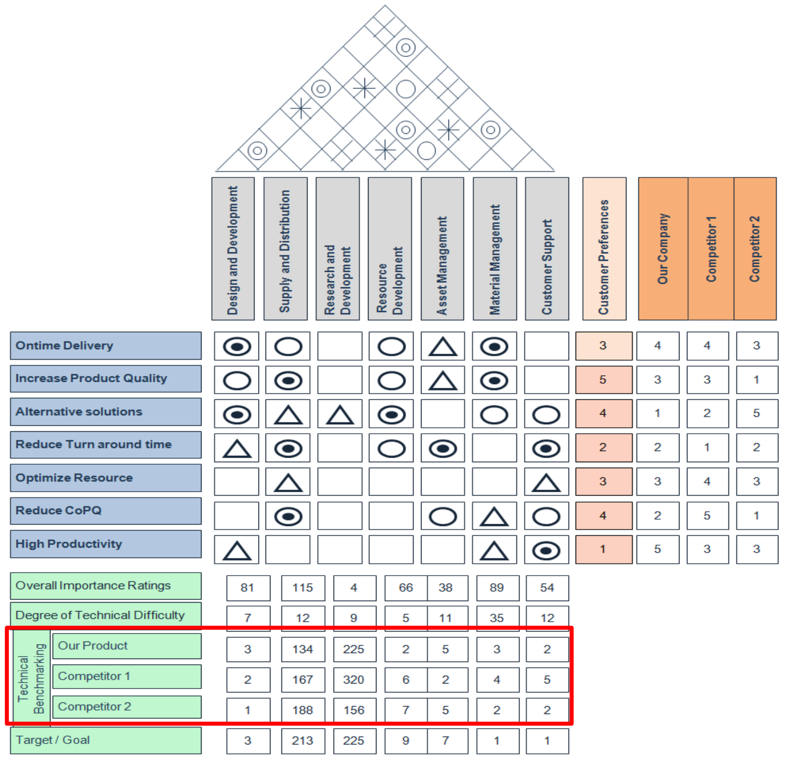




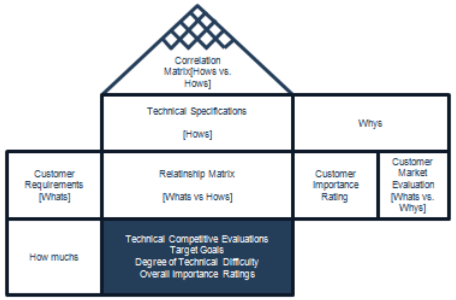
		Design and Development	Supply and Distribution	Research and Development	Resource Development	Asset Management	Material Management	Customer Support	Customer Preferences	Our Company	Competitor 1	Competitor 2
Ontime Delivery		◉	○		○	△	◉		3	4	4	3
Increase Product Quality		○	◉		○	△	◉		5	3	3	1
Alternative solutions		◉	△	△	◉		○	○	4	1	2	5
Reduce Turn around time		△	◉		○	◉		◉	2	2	1	2
Optimize Resource			△					△	3	3	4	3
Reduce CoPQ			◉			○	△	○	4	2	5	1
High Productivity		△					△	◉	1	5	3	3
Overall Importance Ratings		81	115	4	66	38	89	54				
Degree of Technical Difficulty		7	12	9	5	11	35	12				
Technical Benchmarking	Our Product	3	134	225	2	5	3	2				
	Competitor 1	2	167	320	6	2	4	5				
	Competitor 2	1	188	156	7	5	2	2				
Target / Goal		3	213	225	9	7	1	1				

Step 7: Prioritize Design Requirements

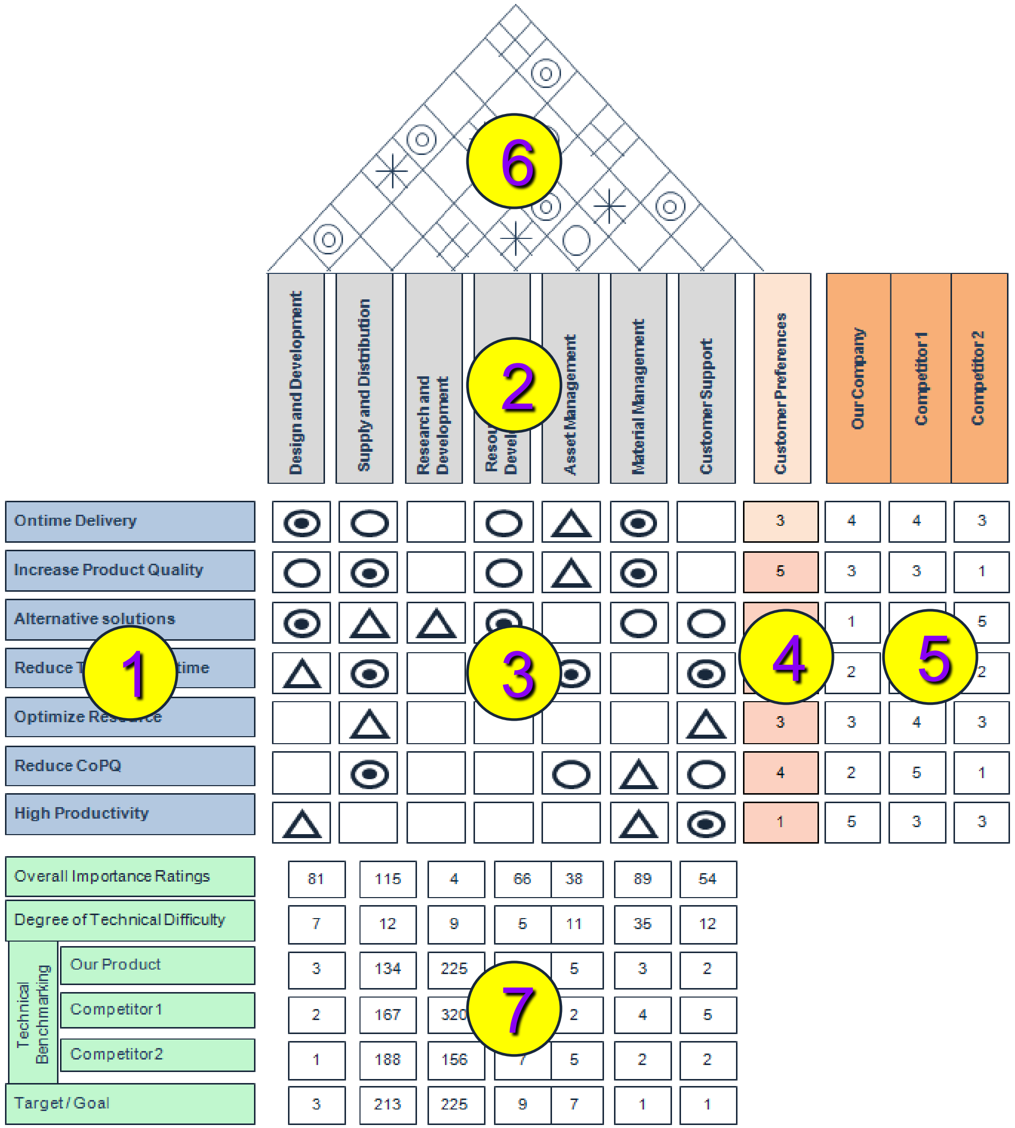
- Technical Specification Competitive Evaluation
- Helps to establish the feasibility and realization of each “how” item
- Target Goals
- How much is good enough to satisfy the customer



	Design and Development	Supply and Distribution	Research and Development	Resource Development	Asset Management	Material Management	Customer Support	Customer Preferences	Our Company	Competitor 1	Competitor 2
Ontime Delivery	◉	○		○	△	◉		3	4	4	3
Increase Product Quality	○	◉		○	△	◉		5	3	3	1
Alternative solutions	◉	△	△	◉		○	○	4	1	2	5
Reduce Turn around time	△	◉		○	◉		◉	2	2	1	2
Optimize Resource		△					△	3	3	4	3
Reduce CoPQ		◉			○	△	○	4	2	5	1
High Productivity	△					△	◉	1	5	3	3
Overall Importance Ratings	81	115	4	66	38	89	54				
Degree of Technical Difficulty	7	12	9	5	11	35	12				
Technical Benchmarking: Our Product	3	134	225	2	5	3	2				
Technical Benchmarking: Competitor 1	2	167	320	6	2	4	5				
Technical Benchmarking: Competitor 2	1	188	156	7	5	2	2				
Target / Goal	3	213	225	9	7	1	1				



House of Quality



	Design and Development	Supply and Distribution	Research and Development	Resou Devel	Asset Management	Material Management	Customer Support	Customer Preferences	Our Company	Competitor 1	Competitor 2
Ontime Delivery	◉	○		○	△	◉		3	4	4	3
Increase Product Quality	○	◉		○	△	◉		5	3	3	1
Alternative solutions	◉	△	△	◉		○	○	[illegible]	1	[illegible]	5
Reduce T[illegible]time	△	◉		[illegible]	◉		◉	[illegible]	2	[illegible]	2
Optimize Re[illegible]		△					△	3	3	4	3
Reduce CoPQ		◉			○	△	○	4	2	5	1
High Productivity	△					△	◉	1	5	3	3

Overall Importance Ratings		81	115	4	66	38	89	54
Degree of Technical Difficulty		7	12	9	5	11	35	12
Technical Benchmarking	Our Product	3	134	225	[illegible]	5	3	2
	Competitor 1	2	167	320	[illegible]	2	4	5
	Competitor2	1	188	156	7	5	2	2
Target / Goal		3	213	225	9	7	1	1

Pros of QFD

- Focuses the design of the product or process on satisfying customer’s needs and wants.

- Improves the contact channels between customers, advertising, research and improvement, quality and production departments, which sustains better decision making.

- Reduces the new product development project period and cost.

Cons of QFD

- The relationship matrix can be too obscure with many process inputs and/or many customer constraints.
- It can be very complicated and difficult to implement without experience.
- If throughout the process new ideas, specifications, or requirements are not discovered, you run the risk of losing team members’ trust in the process.

QFD Summary

When used properly, the quality function deployment is an extremely valuable approach to product/process design.

There are many benefits of QFD that can only be realized when each step of the process is completed thoroughly:

- Logical way of obtaining information and presenting it
- Smallest product development cycle
- Considerably condensed start-up costs
- Fewer engineering alterations
- Reduced chance of supervision during design process
- Collaborating environment
- Preserving everything in characters.

1.2.4 Cost of Poor Quality

Cost of Poor Quality

- **Cost of Poor Quality** (COPQ) is the expense incurred due to waste, inefficiencies, and defects.

- The COPQ has been proven to range from 5% to 30% of gross sales for most companies.
- The COPQ can be staggering when considering process inefficiencies, hidden factories, defective products, rework, scrap, etc.
- Understanding COPQ and where to look for it will help uncover process inefficiencies, defects, and hidden factories within your business.

Cost of Poor Quality

- There are 7 common forms of waste that are often referred to as the "**7 deadly muda.**"
- Technically, there are more than 7 forms of waste but if you can remember these you will capture over 90% of your waste.

1. Defects
2. Overproduction
3. Over-Processing
4. Inventory
5. Motion
6. Transportation
7. Waiting

Cost of Poor Quality

- The “7 deadly muda” are very important to understand. They are the best way to identify the COPQ.
- The presence of any muda causes many other forms of inefficiencies and hidden factories to manifest themselves.
- There are four key categories of costs related to muda:
  1. Costs Related to Production
  2. Costs Related to Prevention
  3. Costs Related to Detection
  4. Costs Related to Obligation

COPQ: Costs Related to Production

- Costs related to **production** are the direct costs of the presence of muda. These forms of COPQ are usually understood and easily observable. They are in fact the “7 deadly muda” themselves.
    1. Defects
    2. Overproduction
    3. Over-Processing
    4. Inventory
    5. Motion
    6. Transportation
    7. Waiting

COPQ: Costs Related to Prevention

- Costs related to the **prevention** of muda are those associated with trying to reduce or eliminate any of the “7 deadly muda.”
  - Costs for error proofing methods or devices
  - Costs for process improvement and quality programs
  - Costs for training and certifications
  - etc.

- *Any costs* directly associated with the prevention of waste and defects should be included in the COPQ calculation.

COPQ: Costs Related to Detection

- Costs related to the **detection** of muda are those associated with trying to find or observe any of the “7 deadly muda.”
  - Costs for sampling
  - Costs for quality control check points
  - Costs for inspection costs
  - Costs for cycle counts or inventory accuracy inspections
  - etc.
- Any costs directly associated with the detection of waste and defects should be included in the COPQ calculation.

COPQ: Costs Related to Obligation

- Costs related to **obligation** are those associated with addressing the muda that reaches a customer.
  - Repair costs
  - Warranty costs
  - Replacement costs
  - Customer returns and customer service overhead
  - etc.

- Any costs directly associated with customer obligations should be included in the COPQ calculation.

COPQ: Types of Cost

- There are two types of costs to be considered when determining COPQ
  1. **Hard Costs**
     - Tangible costs that can be traced to the income statement
  2. **Soft Costs**
     - Intangible costs: avoidance, opportunity costs, lost revenue etc.

- Calculating the COPQ
  1. Determine the types of waste that are present in your process
  2. Estimate the frequency of waste that occurs
  3. Estimate the cost per event, item, or time frame
  4. Do the math.

1.2.5 Pareto Charts and Analysis

Pareto Principle

- The **Pareto principle** is commonly known as the “law of the vital few” or “80:20 rule.”
- It means that the majority (approximately 80%) of effects come from a few (approximately 20%) of the causes.
- This principle was first introduced in early 1900s and has been applied as a rule of thumb in various areas.
- Example of applying the Pareto principle:
  - 80% of the defects of a process come from 20% of the causes.
  - 80% of sales come from 20% of customers.

Pareto Principle

- The Pareto principle helps us to focus on the vital few items that have the most significant impact.
- In concept, it also helps us to prioritize potential improvement efforts.
- Since this 80:20 rule was originally based upon the works of Wilfried Fritz Pareto (or Vilfredo Pareto), the Pareto principle and references to it should be capitalized because Pareto refers to a person (proper noun).
  - Mr. Pareto is also credited for many works associated with the 80:20, some more loosely than others:
    - Pareto's Law
    - Pareto efficiency
    - Pareto distribution etc.

Pareto Charts

- A **Pareto chart** is a chart of descending bars with an ascending cumulative line on the top.
  - **Sum or Count**:

  The descending bars on a Pareto chart may be set on a scale that represents the total of all bars or relative to the biggest bucket, depending on the software you are using.

  - **Percent to Total**: A Pareto chart shows the percentage to the total for individual bars.

  - **Cumulative Percentage**: A Pareto chart also shows the cumulative percentage of each additional bar. The data points of all cumulative percentages are connected into an ascending line on the top of all bars.

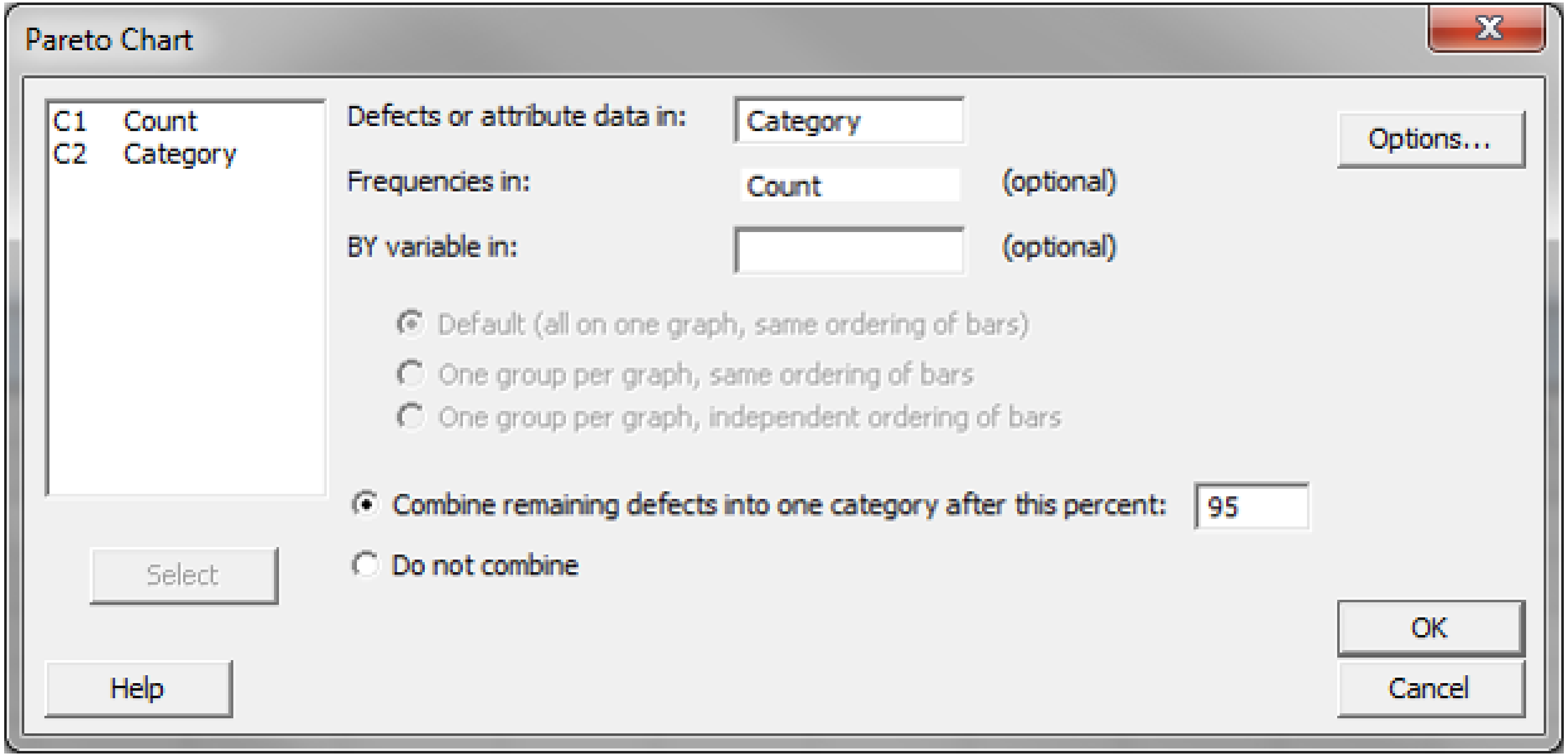
Pareto Charts

- Case study time!
  - Next we will use Minitab to run Pareto charts on exactly the same data set.
  - Open the Minitab data file labeled Pareto and follow the instructions over the next few pages to run Pareto charts in Minitab.

Create a Pareto Chart in Minitab

- Steps to generate a Pareto chart using Minitab:
  1. Open the spreadsheet with the count data for individual categories.
  2. Click on Stat → Quality Tools → Pareto Chart.
  3. A new window with the title "Pareto Chart" pops up.
  4. Select "Category" into the box "Defects or attribute data in" and "Count" into the box "Frequency in."
  5. Click "OK."
  6. The Pareto chart is created in a newly-generated window.

Create a Pareto Chart in Minitab


Pareto Chart

C1 Count
C2 Category

Defects or attribute data in: Category

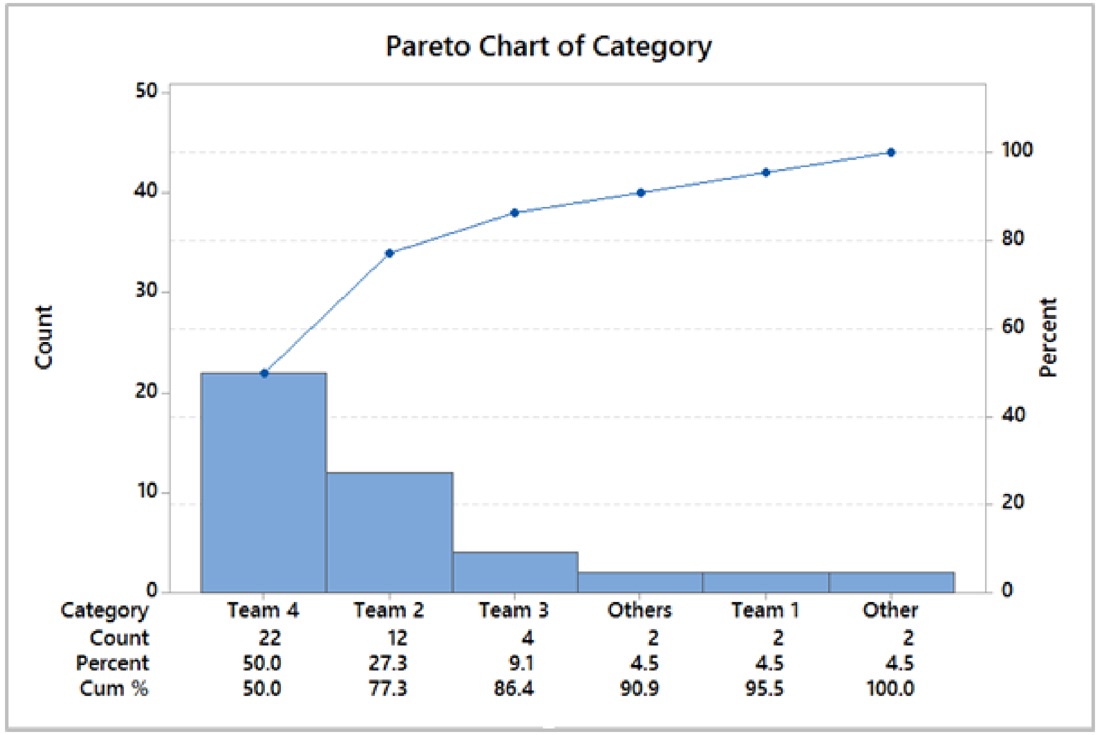
Frequencies in: Count (optional)

BY variable in: (optional)

● Default (all on one graph, same ordering of bars)
○ One group per graph, same ordering of bars
○ One group per graph, independent ordering of bars

● Combine remaining defects into one category after this percent: 95
○ Do not combine

Create a Pareto Chart in Minitab



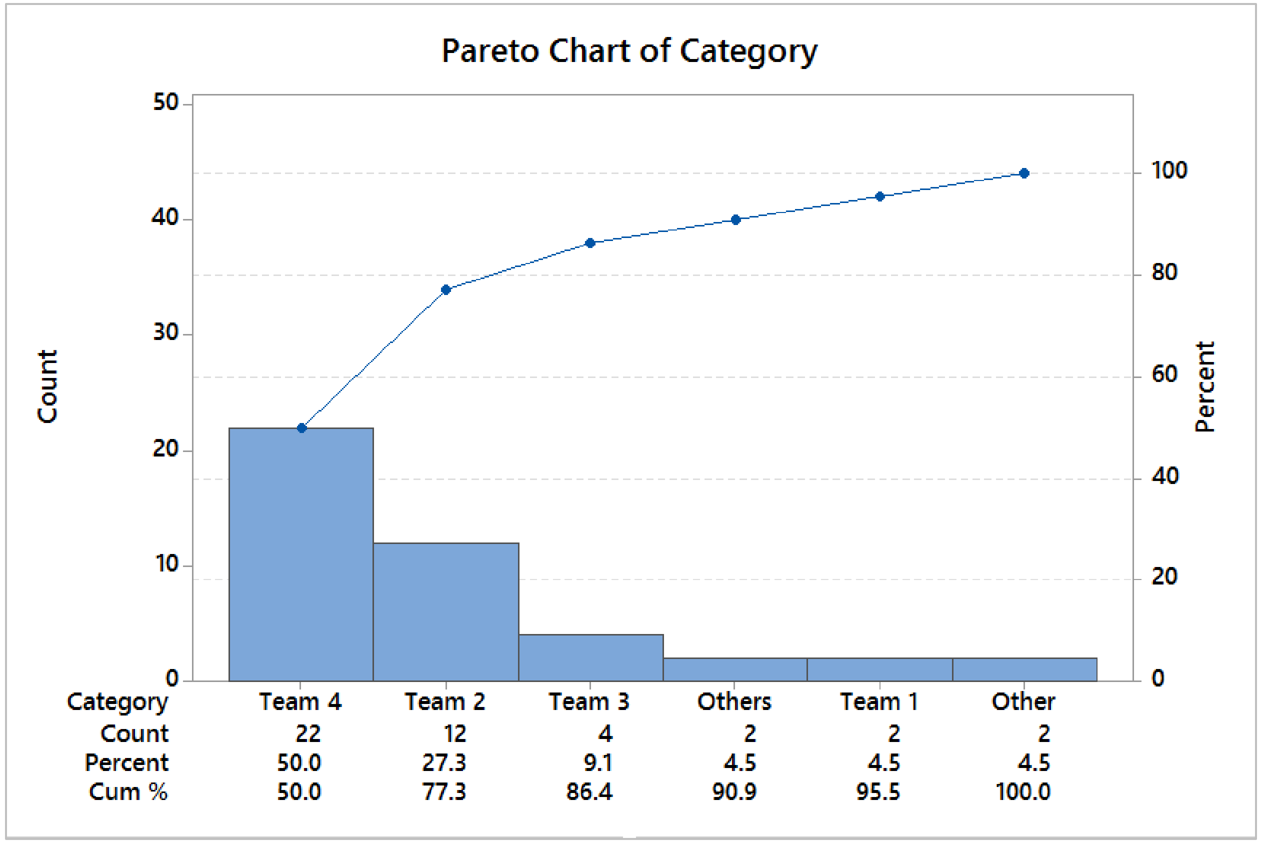
Category	Team 4	Team 2	Team 3	Others	Team 1	Other
Count	22	12	4	2	2	2
Percent	50.0	27.3	9.1	4.5	4.5	4.5
Cum %	50.0	77.3	86.4	90.9	95.5	100.0

- The above Pareto chart generated in Minitab presents the count of defective products by team.
- The bars are descending on a scale with the peak at 50, which is approximately the total count of all defective products for all teams.
- The table below the chart contains counts, individual percentages, and cumulative percentages.
- The cumulative percentages are the red data points driving the red line that spans across the graphic.

Pareto Analysis

- The **Pareto analysis** is used to identify the root causes by using multiple Pareto charts.
- In Pareto analysis, we drill down into the bigger buckets of defects and identify the root causes of defects that contribute heavily to total defects.
- This "drill down" approach effectively solves a significant portion of the problem.
- Next you will see an example of three-level Pareto analysis.
  - The second-level Pareto is a Pareto chart that is a subset of the tallest bar on the first Pareto.
  - The third-level Pareto is a subset of the tallest bar of the second-level Pareto.

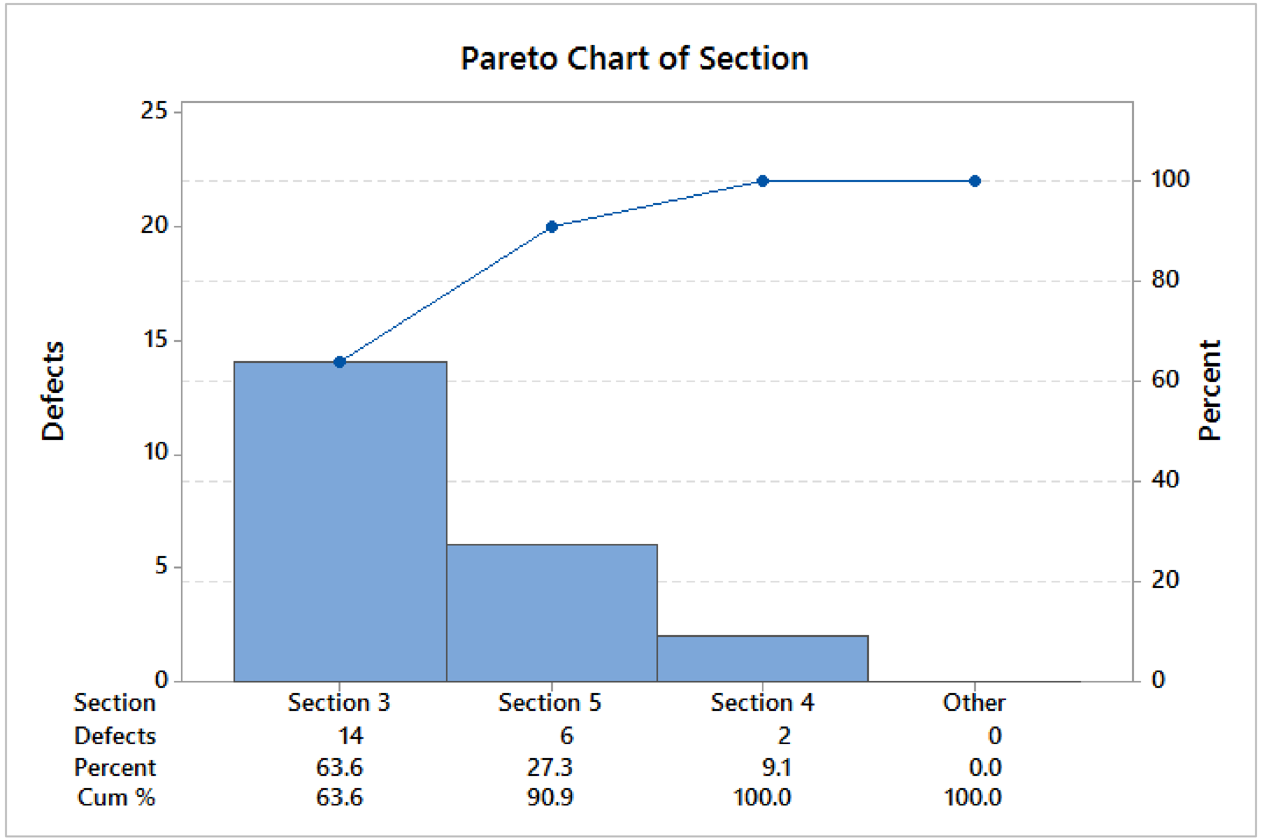
Pareto Analysis: First Level



Category	Team 4	Team 2	Team 3	Others	Team 1	Other
Count	22	12	4	2	2	2
Percent	50.0	27.3	9.1	4.5	4.5	4.5
Cum %	50.0	77.3	86.4	90.9	95.5	100.0

- First-level Pareto
- Shows the count of defective items by team
- Next level will only show the defective items of team 4

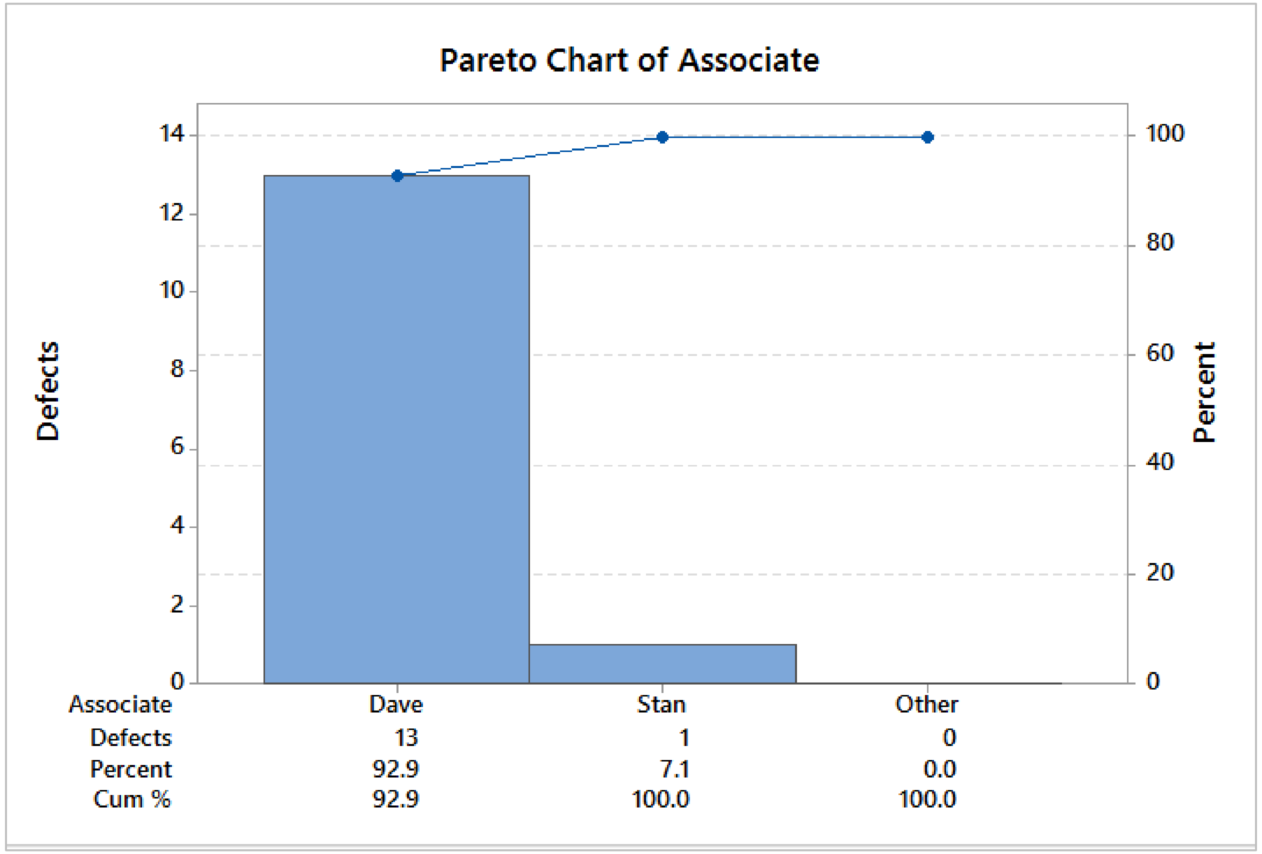
Pareto Analysis: Second Level



Section	Section 3	Section 5	Section 4	Other
Defects	14	6	2	0
Percent	63.6	27.3	9.1	0.0
Cum %	63.6	90.9	100.0	100.0

- Second-level Pareto
- Shows the count of the defective items by section for only team 4
- Next level will only show the defective items of section 3

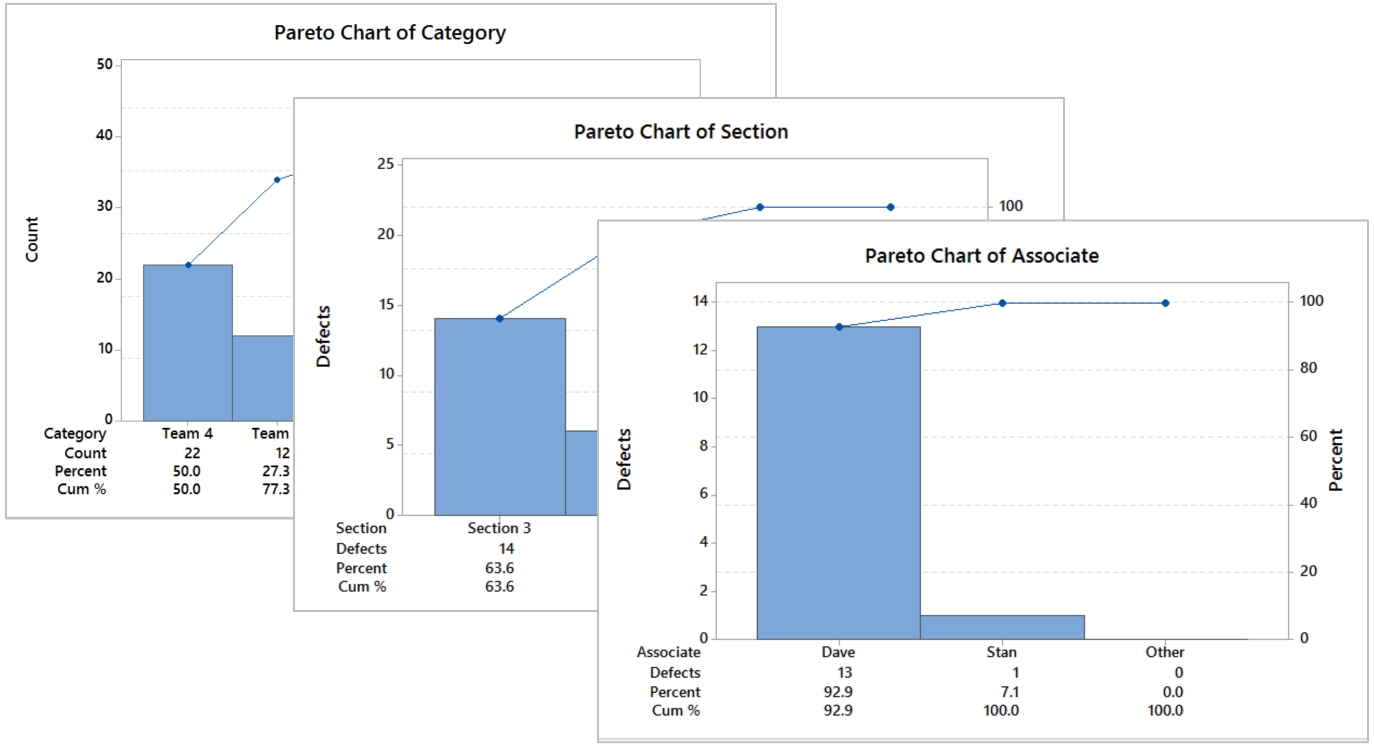
Pareto Analysis: Third Level



- Third-level Pareto
- Shows the count of defective items by associate for only section 3 of team 4
- Next level will only show the defective items of Dave

Pareto Analysis: Conclusion

- After drilling down three levels we find that most of the defective products are from Dave who is in Section 3 of Team 4.
- Determining what Dave might be doing differently and solving that problem can potentially fix about 30% of the entire defective products (13/44).



1.3 Six Sigma Projects

Black Belt Training: Define Phase

1.1 Six Sigma Overview

- 1.1.1 What is Six Sigma
- 1.1.2 Six Sigma History
- 1.1.3 Six Sigma Approach Y = f(x)
- 1.1.4 Six Sigma Methodology
- 1.1.5 Roles and Responsibilities

1.2 Six Sigma Fundamentals

- 1.2.1 Defining a Process
- 1.2.2 VOC and CTQs
- 1.2.3 QFD
- 1.2.4 Cost of Poor Quality (COPQ)
- 1.2.5 Pareto Analysis (80:20 rule)

1.3 Lean Six Sigma Projects

- 1.3.1 Six Sigma Metrics
- 1.3.2 Business Case and Charter
- 1.3.3 Project Team Selection
- 1.3.4 Project Risk Management
- 1.3.5 Project Planning

1.4 Lean Fundamentals

- 1.4.1 Lean and Six Sigma
- 1.4.2 History of Lean
- 1.4.3 The Seven Deadly Muda
- 1.4.4 Five-S (5S)

1.3.1 Six Sigma Metrics

Six Sigma Metrics

- There are many Six Sigma metrics and/or measures of performance used by Six Sigma practitioners.

- In addition to the ones we will cover here, several others (Sigma level, Cp, Cpk, Pp, Ppk, takt time, cycle time, utilization etc.) will be covered in other modules throughout this training.

- The Six Sigma metrics of interest here in the **define phase** are:
  - Defects per Unit (DPU)
  - Defects per Million Opportunities (DPMO)
  - Yield (Y)
  - Rolled Throughput Yield (RTY).

Defects per Unit: DPU

- **DPU** stands for “Defects per Unit“
- DPU is the basis for calculating DPMO and RTY, which we will cover in the next few pages.
- DPU is found by dividing total defects by total units.
  - ***DPU = D/U***
- For example, if you have a process step that produces an average of 65 defects for every 598 units, then your DPU = 65/598 = 0.109.

DPMO: Defects per Million Opportunities

- **DPMO** is one of the few important Six Sigma metrics that you should get comfortable with if you are associated with Six Sigma.

- In order to understand DPMO it is best if you first understand both the nomenclature and the nuances such as the difference between defect and defective.

- **Nomenclature**
    - **Defects = D**
    - **Unit = U**
    - **Opportunity to have a defect = O**

DPMO: Defects per Million Opportunities

- In order to properly discuss DPMO, we must first explore the differences between "defects" and "defective."
  - **Defective**
    - Defective suggests that the value or function of the entire unit or product has been compromised.
    - Defective items will always have at least one defect. Typically, however, it takes multiple defects and/or critical defects to cause an item to be defective.
  - **Defect**
    - A defect is an error, mistake, flaw, fault, or some type of imperfection that reduces the value of a product or unit.
    - A single defect may or may not render the product or unit "defective" depending on the specifications of the customer.
  - **Summary**
    - Defect means that part of a unit is bad.
    - Defective means that the whole unit is bad.

DPMO: Defects per Million Opportunities

- Now let us turn our attention to defining "opportunities" so that we can fully understand Defects per Million Opportunities (DPMO).
  - **Opportunities**
    - Opportunities are the total number of possible defects.
    - Therefore, if a unit has 6 possible defects, then each unit produced is equal to 6 defect opportunities.
    - If we produce 100 units, then there are 600 defect opportunities.

DPMO: Defects per Million Opportunities

- Calculating Defects per Million Opportunities
- **The equation is DPMO = (D/(U × O)) × 1,000,000**
- Example: Let us assume:
    - There are 6 defect opportunities per unit
    - There are an average of 4 defects every 100 units.

- Opportunities = 6 × 100 = 600
- Defect rate = 4/600
- DPMO = 4/600 × 1,000,000 = 6,667

DPMO: Defects per Million Opportunities

- **What is the reason or significance of 1,000,000?**
- Converting defect rates to a per million value becomes necessary when the performance of your process approaches Six Sigma.

- When this happens, the number of defects shrinks to virtually nothing. In fact, if you recall from the “What is Six Sigma” module, sigma is 3.4 defects per million opportunities.

- By using 1,000,000 opportunities as the barometer we have the resolution in the measurement to count defects all the way up to Six Sigma.

RTY: Rolled Throughput Yield

- **Rolled Throughput Yield (RTY)** is a process performance measure that provides insight into the cumulative effects of an entire process.

- RTY measures the yield for each of several process steps and provides the probability that a unit will come through that process defect-free.

- RTY allows us to expose the "hidden factory" by providing visibility into the yield of each process step.

- This helps us identify the poorest performing process steps and gives us clues into where to look to find the most impactful process improvement opportunities.

RTY: Rolled Throughput Yield

- **Calculating RTY:**
- RTY is found by multiplying the yields of each process step.

- Let us take the 5-step process below and calculate the RTY using the multiplication method mentioned above.

$Y_1 = 0.90$ → $Y_2 = 0.91$ → $Y_3 = 0.99$ → $Y_4 = 0.98$ → $Y_5 = 0.97$

- The calculation is: RTY = 0.90 × 0.91 × 0.99 × 0.98 × 0.97 = 0.77

- Therefore, RTY = 77%.

RTY: Rolled Throughput Yield

- You may have noticed that in order to calculate RTY we must determine the yield for each process step.

- Before we get into calculating yield, there are a few abbreviations that need to be declared.

  - Abbreviations
    - Defects = **D**
    - Unit = **U**
    - Defects per Unit = **DPU**
    - Yield = **Y**
    - e = 2.71828 *(mathematical constant)*

RTY: Rolled Throughput Yield

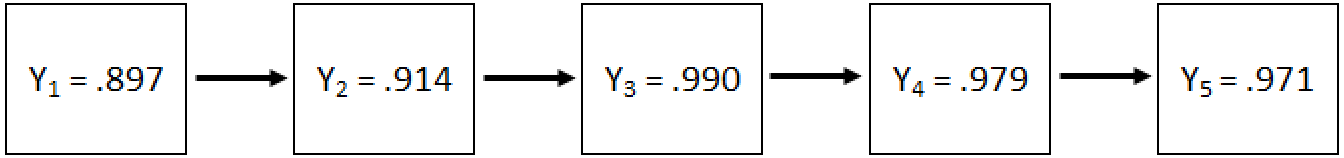
- **Calculating Yield**
- The **Yield** of a process step is the success rate of that step or the probability that the process step produces no defects.

- In order to calculate the yield, we need to know the DPU and then we can apply it to the yield equation below.

$$Yield = e^{-dpu}$$

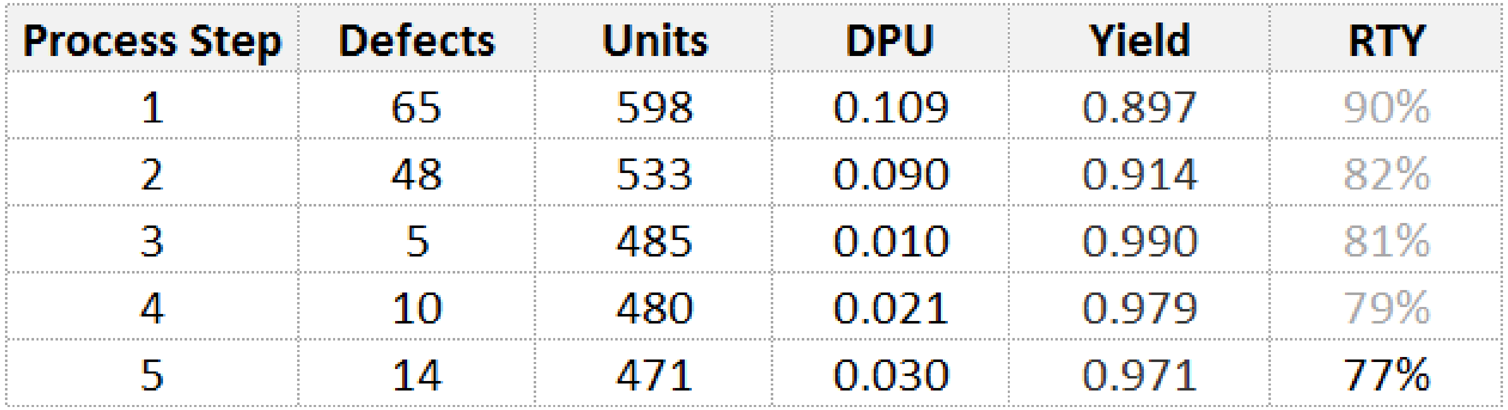
- Example
  - Let us assume a process step has a DPU of 0.109 (65/598)
  - Yield = 2.718 ^ -0.109 = 0.8967. Rounded, Yield = 90%.

RTY: Rolled Throughput Yield



$Y_1 = .897$ → $Y_2 = .914$ → $Y_3 = .990$ → $Y_4 = .979$ → $Y_5 = .971$

- Below is a table using the above process yield data that we used in the earlier RTY calculation.
- This table allows us to see the DPU and yield of each step as well as the RTY for the whole process.



Process Step	Defects	Units	DPU	Yield	RTY
1	65	598	0.109	0.897	90%
2	48	533	0.090	0.914	82%
3	5	485	0.010	0.990	81%
4	10	480	0.021	0.979	79%
5	14	471	0.030	0.971	77%

RTY: Using an Estimate of Yield

- Calculating RTY using **Yield estimation**
  - It is possible to “estimate” yield by taking the inverse of DPU or simply subtracting DPU from 1.
  - Yield Estimation = 1 − DPU
    - Example Yield Estimate for process step 1: 1 − 0.10870 = 0.90
  - RTY using the Yield Estimation Method
    - RTY = 0.891 × 0.91 × 0.99 × 0.979 × 0.97 = 0.76 = 76%

Process Step	Defects	Units	DPU	Yield$_{estimate}$	RTY$_{estimate}$
1	65	598	0.109	0.891	89%
2	48	533	0.090	0.910	81%
3	5	485	0.010	0.990	80%
4	10	480	0.021	0.979	79%
5	14	471	0.030	0.970	76%

NOTE: This method assumes max DPU is 1. Although this is not technically true, in practice its extremely rare to find processes with DPU greater than 1. If it is the case this method will not work.

RTY: Using First Time Yield (FTY)

- In the previous discussion we calculated RTY using Yields calculated based on “**Defects**” per unit. What if we’re counting “**Defectives**”?
  - **First Time Yield** (FTY) is the probability of producing a product or service defect free before rework
    - $\mathrm{FTY} = \frac{\#\ Non\text{-}Defective\ Units\ Produced\ Before\ Rework}{\#\ Units\ Inspected}$
- FTY is also a good measure of process performance. It accounts for the hidden factory or rework in the process. Its simply based on defective units instead of defects per unit.
    - You can still calculate RTY using First Time Yield:
    - $\mathrm{RTY} = \mathrm{FTY}_1 * \mathrm{FTY}_2 * \mathrm{FTY}_3 * \mathrm{FTY}_4 * \mathrm{FTY}_5$

1.3.2 Business Case and Charter

Business Case and Project Charter

- Earlier we stated that DMAIC is a structured and rigorous methodology designed to be repeatedly applied to *any* process in order to achieve Six Sigma.

- We also stated that DMAIC was a methodology that refers to 5 phases of a project.
  - Define, Measure, Analyze, Improve, and Control

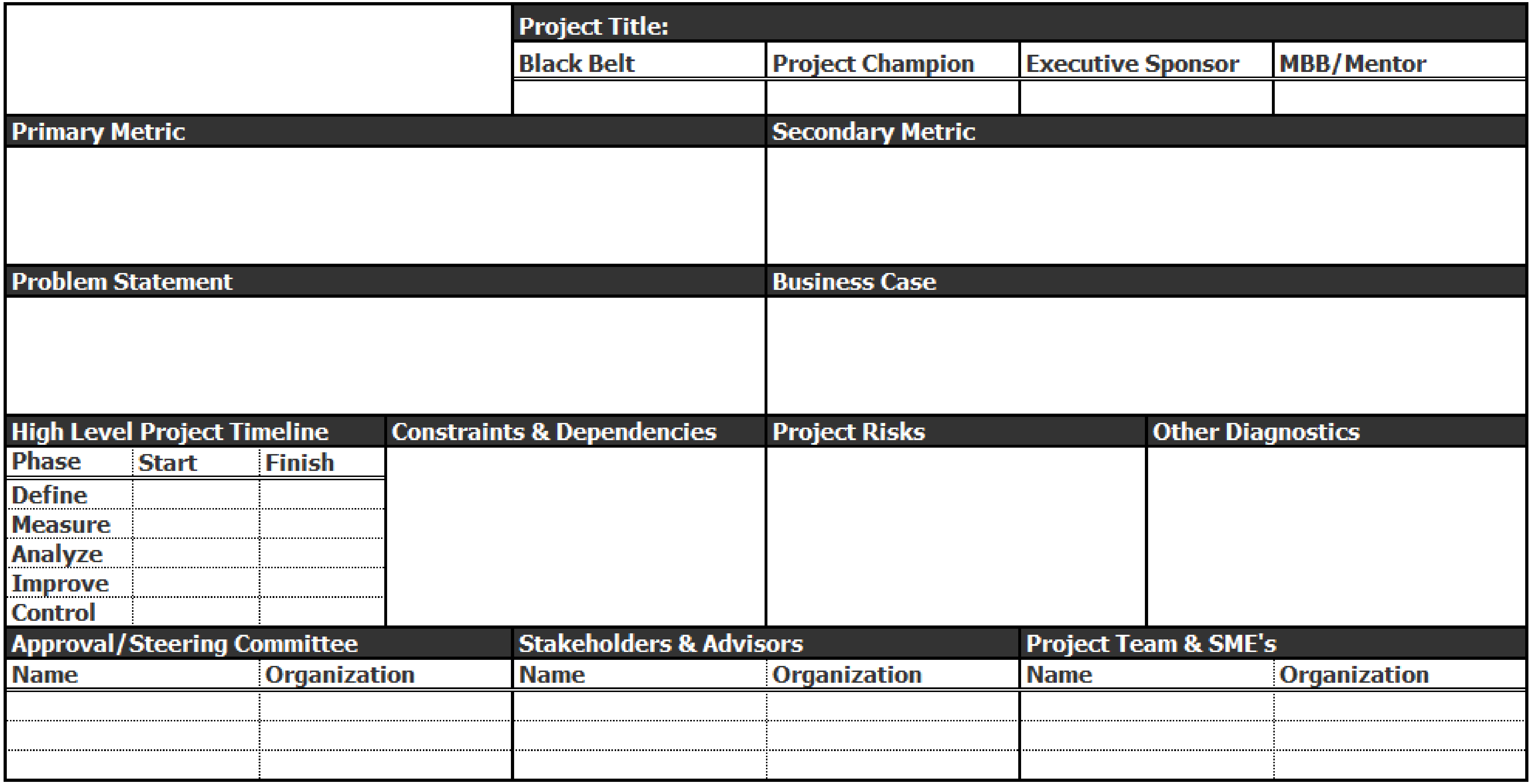
- Given that the premise of the DMAIC methodology is project-based, we must take the necessary steps to define and initiate a project, hence the need for. . .
  - **Project Charters**

Project Charter

- The purpose of a **project charter** is to provide vital information about a project in a quick and easy-to-comprehend manner.

- Project charters are used to get approval and buy-in for projects and initiatives as well as declaring:
  - The scope of work
  - Project teams
  - Decision authorities
  - Project lead
  - Success measures

Project Charter Template



Project Title:

Black Belt	Project Champion	Executive Sponsor	MBB/Mentor

Primary Metric	Secondary Metric

Problem Statement	Business Case

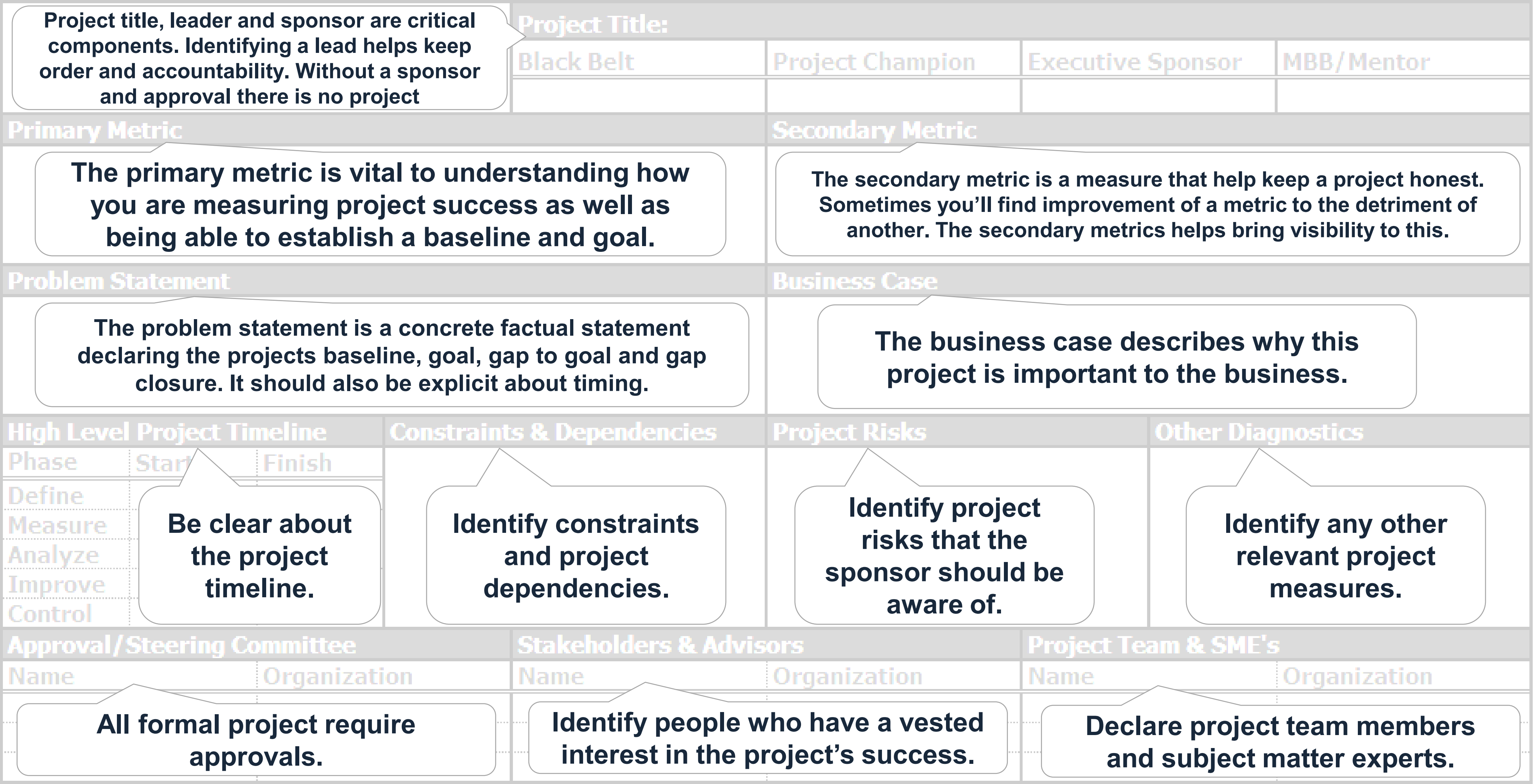
High Level Project Timeline

Phase	Start	Finish
Define		
Measure		
Analyze		
Improve		
Control		

Constraints & Dependencies	Project Risks	Other Diagnostics

Approval/Steering Committee		Stakeholders & Advisors		Project Team & SME's	
Name	Organization	Name	Organization	Name	Organization

Project Charter Key Elements



Project Title:

Black Belt | Project Champion | Executive Sponsor | MBB/Mentor

Project title, leader and sponsor are critical components. Identifying a lead helps keep order and accountability. Without a sponsor and approval there is no project

Primary Metric

The primary metric is vital to understanding how you are measuring project success as well as being able to establish a baseline and goal.

Secondary Metric

The secondary metric is a measure that help keep a project honest. Sometimes you'll find improvement of a metric to the detriment of another. The secondary metrics helps bring visibility to this.

Problem Statement

The problem statement is a concrete factual statement declaring the projects baseline, goal, gap to goal and gap closure. It should also be explicit about timing.

Business Case

The business case describes why this project is important to the business.

High Level Project Timeline

Phase	Start	Finish
Define		
Measure		
Analyze		
Improve		
Control		

Be clear about the project timeline.

Constraints & Dependencies

Identify constraints and project dependencies.

Project Risks

Identify project risks that the sponsor should be aware of.

Other Diagnostics

Identify any other relevant project measures.

Approval/Steering Committee

Name | Organization

All formal project require approvals.

Stakeholders & Advisors

Name | Organization

Identify people who have a vested interest in the project's success.

Project Team & SME's

Name | Organization

Declare project team members and subject matter experts.

Project Charter: Key Elements

- Title
  - Projects should have a name, title, or some reference that identifies them.
  - Branding can be an important ingredient in the success of a project so be sure your project has a reference name or title.

- Leader
  - Any projects needs a declared leader or someone who is responsible for project's execution and success.
  - You may hear references to RACI throughout in your Six Sigma journey.
  - **RACI** stands for **R**esponsible, **A**ccountable, **C**onsulted, **I**nformed and identifies the people that play those roles.
  - Every project must have declared leaders indicating who is responsible and who is accountable.

Project Charter: Key Elements

- Business Case
  - A **business case** is the quantifiable reason why the project is important.
  - Business cases help shed light on problems. They explain why a business should care.
  - Business cases must be quantified and stated succinctly.
  - COPQ is a key method of quantification for any business case.

Project Charter: Key Elements

- Problem Statement and Objective
  - A properly written problem statement has an objective statement woven into it.
  - There should be no question as to the current state or the goal.
  - A gap should be declared, the gap being the difference between the present state and the goal state.
  - The project objective should be to close the gap or reduce the gap by some reasonable amount.
  - Valuation or COPQ is the monetary value assigned to the gap.
  - Lastly, a well-written problem statement refers to a timeline expected to be met.

Project Charter: Problem Statement Examples

- Currently, process defect rates are 17% with a goal of 2%. This represents a gap of 15%, costing the business $7.4 million dollars. The goal of this project is to reduce this gap by 50% before Nov 2019 putting process defect rates at 9.5% and saving $3.7MM.

- Process cycle time has averaged 64 minutes since Q1 2019. However, production requirements put the cycle time goals at 48 min. This 16-min gap is estimated to cost the business $296,000. The goal of this project is reduce cycle time by 16 min. by Q4 2019 and capture all $296,000 cost savings.

Project Charter: Key Elements

- Metrics
    - A measure of success is an absolute for any project.
    - Metrics give clarity to the purpose of the work.
    - Metrics establish how the initiative will be judged.
    - Metrics establish a baseline or “starting point.”
    - For Six Sigma projects…**metrics are mandatory!**

Project Charter: Key Elements

- Primary Metric
  - The **primary metric** is a generic term for a Six Sigma project's most important measure of success. The primary metric is defined by the Black Belt, GB, MBB, or Champion.
  - A primary metric is an absolute MUST for any project and it should not be taken lightly. Here are a few characteristics of good primary metrics.
  - Primary metrics should be:
    - tied to the problem statement
    - measureable
    - expressed with an equation
    - aligned to business objectives
    - tracked at the proper frequency (hourly, daily, weekly, monthly etc.)
    - expressed pictorially over time with a run chart, time series, or control chart
    - validated with an MSA.

Project Charter: Key Elements

- The primary metric is the reason for your work.

- It is the success indicator.

- It is your beacon.

- The primary metric is of utmost importance and should be improved, **but** not at the expense of your secondary metric.

Project Charter: Key Elements

- Secondary Metric
  - The **secondary metric** is the thing you do not want sacrificed on behalf of a primary improvement.
  - A secondary metric is one that makes sure problems are not just "changing forms" or "moving around."
  - The secondary metric keeps us honest and ensures we are not sacrificing too much for our primary metric.
  - If your primary metric is a cost or speed metric, then your **secondary metric** should probably be some quality measure.
    - Example: If you were accountable for saving energy in an office building and your primary metric was energy consumption then you *could* shut off all the lights and the HVAC system and save tons of energy. . .except that your secondary metrics are probably comfort and functionality of the work environment.

Project Charter: Key Elements

- Elements of a Good Project Charters *(continued)*
  - Scope Statement – defined by high-level process map
  - Stakeholders Identified – who is affected by the project
  - Approval Authorities Identified – who makes the final call
  - Review Committees Defined – who is on the review team
  - Risks and Dependencies Highlighted – identify risks and critical path items
  - Project Team Declared – declare team members
  - Project Timeline Estimated – set high-level timeline expectations.

1.3.3 Project Team Selection

Project Team Selection

- Six Sigma project team selection is the cornerstone of a successful Six Sigma project.

- Teams and Team Success
  - A **team** is a group of people who share complementary skills and experience.
  - A team will be dedicated to consistent objectives.
  - Winning teams share similar and coordinated goals.
  - Teams often execute common methods or approaches.
  - Team members hold each other accountable for achieving shared goals.

Project Team Selection

- What makes a team successful?
  - Shared goals
  - Commitment
  - Leadership
  - Respect
  - Effective communication
  - Autonomy
  - Diversity (capabilities, knowledge, skills, experience etc.)
  - Adequate resources.

Project Team Selection

- Keys to Team Success
  - Agreed focus on the goal or the problem at hand
    - Focus on problems that have meaning to the business
    - Focus on solvable problems within the scope of influence; a successful team does not seek unattainable solutions.
  - Team Selection
    - Selected teammates have proper skills and knowledge
    - Adequately engaged management
    - Appropriate support and guidance from their direct leader
  - Successful teams use reliable methods
    - Follow the prescribed DMAIC methodology
    - Manage data, information, and statistical evidence
  - Successful teams always have exceeds players
    - Winning teams typically
      - Have unusually high standards.
      - Have greater expectations of themselves and each other.
      - Do not settle for average or even above average results.

Project Team Selection

- Principles of Team Selection:
  - Select team members based on
    - Skills required to achieve the objective
    - Experience (Subject Matter Expertise)
    - Availability and willingness to participate
    - Team size (usually 4–8 members)
      - Don’t go at it alone!
      - Don’t get too many cooks in the kitchen!
    - Members’ ability to navigate
      - The process
      - The company
      - The political landscape
  - Be sure to consider the inputs of others
    - Heed advice
    - Seek guidance

Project Team Development

- All teams experience the following four stages of development. It is helpful to understand these phases so that you can anticipate what your team is going to experience.

- The four stages of team development process:
  - Forming
  - Storming
  - Norming
  - Performing

- Teammates seek something different at each stage:
  - In the forming stage they seek inclusion
  - In the storming stage they seek direction and guidance
  - In the norming stage they seek agreement
  - In the performing stage they seek results.

Project Team Development

- Patterns of a team in the **Forming** stage:
  - Roles and responsibilities are unclear
  - Process and procedures are ignored
  - Scope and parameter setting is loosely attempted
  - Discussions are vague and frustrating
  - There is a high dependence on leadership for guidance
- Patterns of a team in the **Storming** stage:
  - Attempts to skip the research and jump to solutions
  - Impatience for some team members regarding lack of progress
  - Arguments about decisions and actions of the team
  - Team members establish their position
  - Subgroups or small teams form
  - Power struggles exist and resistance is present

Project Team Development

- Patterns of a team in the **Norming** stage:
  - Agreement and consensus start to form
  - Roles and responsibilities are accepted
  - Team members’ engagement increases
  - Social relationships begin to form
  - The leader becomes more enabling and shares authority

- Patterns of a team in the **Performing** stage:
  - Team is directionally aware and agrees on objectives
  - Team is autonomous
  - Disagreements are resolved within the team
  - Team forms above average expectations of performance

Project Team Development

- Well-structured and energized project teams are the essential components of any successful Six Sigma project.

- To have better chances of executing the project successfully, you will need to understand and effectively manage the team development process.

1.3.4 Project Risk Management

Risk

- Risk is defined as a future event that can impact the task/project if it occurs.

What is Project Risk Management?

- The main purpose of **risk management** is to foresee potential risks that may inhibit the project deliverables from being delivered on time, within budget, and at the appropriate level of quality, and then to mitigate these risks by creating, implementing, and monitoring *contingency plans*.

- Risk management is concerned with identifying, assessing, and monitoring project risks before they develop into issues and impact the project.

- **Risk analysis** helps to identify and manage potential problems that could impact key business initiatives or project goals.

Three Basic Parameters of Risk Analysis

- Risk Assessment:
  - The process of identifying and evaluating risks, whether in absolute or relative terms.
- Risk Management:
  - Project risk management is the effort of responding to risks throughout the life of a project and in the interest of meeting project goals and objectives.
- Risk Communication:
  - Communication plays a vital role in the risk analysis process because it leads to a good understanding of risk assessment and management decisions.

Why is Risk Analysis Necessary?

What can happen if you omit the risk analysis?

- Vulnerabilities cannot be detected
- Mitigation plans are introduced without proper justification
- Customer dissatisfaction
- Not meeting project goals
- Remake the whole system
- Huge cost and time loss

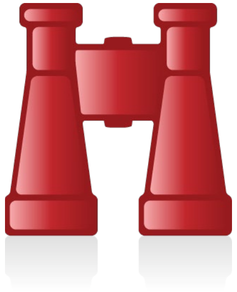
Project Risk Analysis Steps

The project risk analysis process consists of the following steps that evolve through the life cycle of a project.

- Risk Identification:
  - Identify risks and risk categories, group risks, and define ownership.
- Risk Assessment:
  - Evaluate and estimate the possible impacts and interactions of risks.
- Response Planning:
  - Define mitigation and reaction plans.
- Mitigation Actions:
  - Implement action plans and integrate them into the project.
- Tracking and Reporting:
  - Provide visibility to all risks.
- Closing:
  - Close the identified risk.

Risk Identification

The first action of risk management is the identification of individual events that the project may encounter during its lifecycle.



The identification step comprises:

- Identify the risks
- Categorize the risks
- Match the identified risks to categories
- Define ownership for managing the risks.

Risk Identification

- Source of Risk:
- Identification of risk sources provides a basis for systematically examining changing situations over time to uncover circumstances that impact the ability of the project to meet its objectives.

Risk Identification

Source of Risk	Description
Human Resources	The risks originated from human resources (e.g., availability, skill etc.)
Physical Resources	The risks originated from physical resources (e.g., hardware or software, availability of the required number at the right time etc.)
Technology	The risks originated from technology (e.g., development environment, new or complex technologies, performance requirements, tools etc.)
Suppliers	The risks are associated with a supplier (e.g., delays in supplies, capability of suppliers etc.)
Customer	The risks derived from the customer (e.g., unclear requirements, requirement volatility, change in project scope, delays in response etc.)
Security	The risks are associated with information security, security of personnel, security of assets, and security of intellectual property
Legal	The risks are associated with legal issues that may impact the project
Project management	The risks are associated with project management processes, organizational maturity, and ability

Risk Identification

Risk Parameters:

Parameters for evaluating, categorizing, and prioritizing risks include the following:

- Risk likelihood (i.e., probability of risk occurrence)
- Risk consequence (i.e., impact and severity of risk occurrence)
- Thresholds to trigger management activities.

Risk Assessment

The risk assessment consists of evaluating the range of possible impacts should the risk occur.

Follow these steps when assessing risks:

1. Define the various impacts of each risk
2. Rate each impact based on a logical severity level
3. Sort and evaluate risks by severity level
4. Determine if any controls already exist
5. Define potential mitigation actions.

Risk Mitigation Planning

The risk owners are responsible for planning and implementing mitigation actions with support from the project team.

- All team members, inclusive of partners and suppliers, may be requested to identify and develop mitigation measures for identified risks.
- The project core team members are responsible for identifying an appropriate action owner for each identified risk.
- After mitigation actions are defined, the project core team will review the actions.
- The risk owner must track all mitigation actions and expected completion dates.
- The risk owner and the project core team members must hold all action owners accountable for the risk mitigation planning.

Risk Mitigation Action Implementation

- The **action implementation** is the responsibility of the risk owner.

- The action owners are responsible for the execution of the tasks or activities necessary to complete the mitigation action and eliminate or minimize the risk.

- The risk owner or the project manager will monitor completion dates of the mitigation action implementation.

Risk Occurrence and Contingency Plans

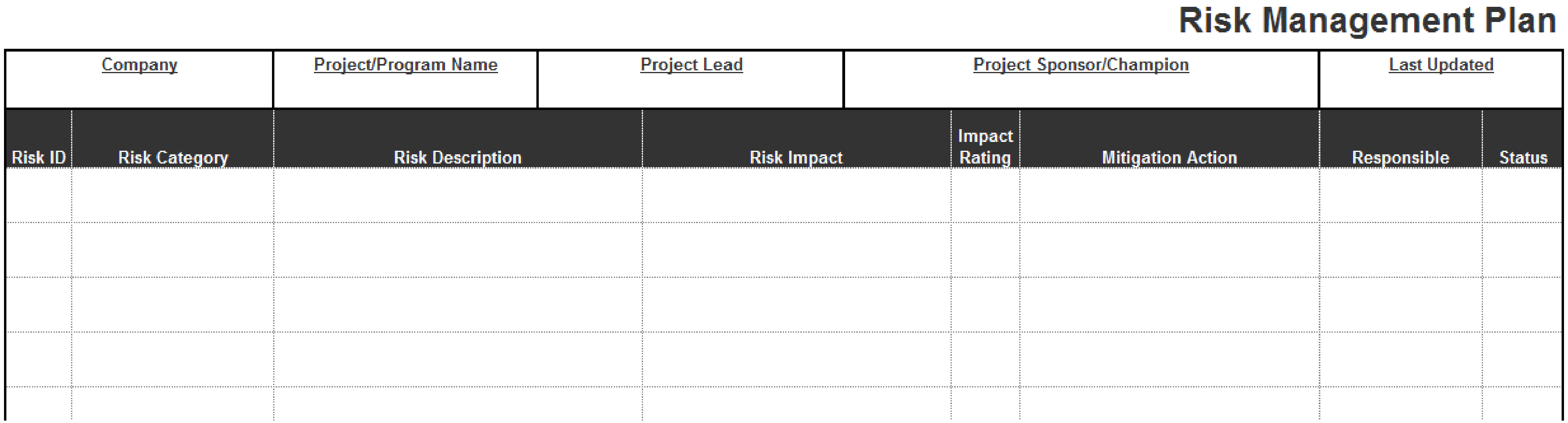
- Whenever any risk occurs, the project team should implement **contingency plans** to ensure that project deliverables can be met.

- The details of each occurrence should be recorded in the risk register or other tracking tool.

- The **risk register** or **risk management plan** *(see next slide)* will be maintained by the project manager and reviewed on a regular basis.

Risk Tracking and Reporting

- **Risk tracking and reporting** provides critical visibility to all risks.
- Risk owners must report on the status of their mitigation actions.
- Depending on the risk severity, project managers need to report the risk status of each category of risk to senior management.



Risk Management Plan

Company	Project/Program Name	Project Lead	Project Sponsor/Champion	Last Updated

Risk ID	Risk Category	Risk Description	Risk Impact	Impact Rating	Mitigation Action	Responsible	Status

Risk Closure

- The risk owners are responsible for recommending the risk closure to the project manager.

- A risk is *closed* only when the item is not considered a risk to the project anymore.

- When a risk is closed, the project manager needs to update the risk status in the lessons learned document.

Risk Analysis Features

The risk analysis should be:

- Systematic
- Comprehensive
- Data driven
- Adherent to evidence
- Logically sound
- Practically acceptable
- Open to critique
- Easy to understand.

Project Risk Analysis Advantages

- Helps strategic and business planning
- Meets customer requirements
- Reduces schedule slips and cost overruns
- Promotes an effective usage of resources
- Promotes continuous improvement
- Helps to achieve project goals
- Minimizes surprises from customers and stakeholders
- Allows a quick grasp of new opportunities
- Enhances communication
- Reassures stakeholders that the project stays on track.

1.3.5 Project Planning

What is Project Management?

- **Project management** is the process of defining, planning, organizing, managing, leading, securing, and optimizing the resources to achieve a set of planned goals and objectives.
- It is the application of knowledge, skills, tools, and techniques to project activities in order to meet project requirements.

What is a Project Plan?

- A **project plan** is a crucial step in project management for achieving a project's goals.
- A project plan is a formal approved document used to guide and execute project tasks.
- It provides an overall framework for managing project tasks, schedules, and costs.
- A project plan is a coordinating tool and communication device that helps teams, contractors, customers, and organizations define the crucial aspects of a project or program.

Project Planning Stages

1. **Determine project scope and objectives**: Explore opportunities, identify and prioritize needs, consider project solutions.
2. **Plan the project**: Identify input and resources requirements such as human resources, materials, software, hardware, and budgets.
3. **Prepare the project proposal:** Based on stakeholder feedback, plan the necessary resources, timeline, budget etc.
4. **Implement the project:** Implement the project by engaging responsible resources and parties. Ensure execution and compliance of the defined plans.
5. **Evaluate the project:** Regularly review progress and results. Measure the project's effectiveness against *quantifiable* requirements.

Planning and Scheduling Objectives

- To optimize the use of resources (both human and other resources).
- To increase productivity.
- To achieve desired schedules and deliverables.
- To establish an approach to minimize long-term maintenance costs.
- To minimize the chaos and productivity losses resulting from planned production schedules, priority changes, and non-availability of resources.
- To assess current needs and future challenges.

Project Planning Activities

- Statement of work (SOW)
- Work breakdown structure (WBS)
- Resource estimation plan
- Project schedule
- Budget or financial plan
- Communication plan
- Risk management plan

Project Planning Activities: SOW

Statement of Work (SOW) is a formal document often accompanying a contract that outlines specific expectations, limitations, resources and work guidelines intended to define the work or project.

SOW's:

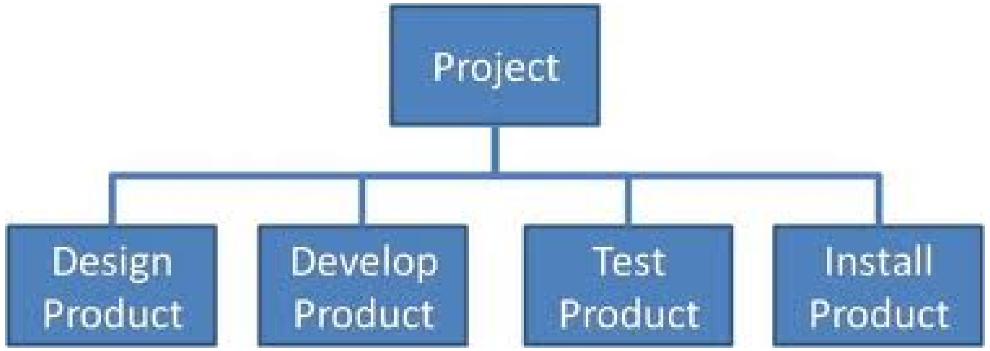
- Define the scope of the project.
- Establish expectations and parameters of the project.
- Identify technical requirements for the project
- Provide guidance on materials to be used.
- Establish timeline expectations.
- Etc.

Project Planning Activities: WBS

Work Breakdown Structure (WBS) is a decomposition of project components into small and logical bodies of work or tasks.

Work Breakdown Structures:

- Identify all required components of a project.
- Cascades components into sub-components and tasks.
- WBS's are not by themselves project plans or schedules but they are a necessary step to help establish project plans and timelines.
- WBS's also enable logical reporting and summarizations of project progress


Project

Design Product

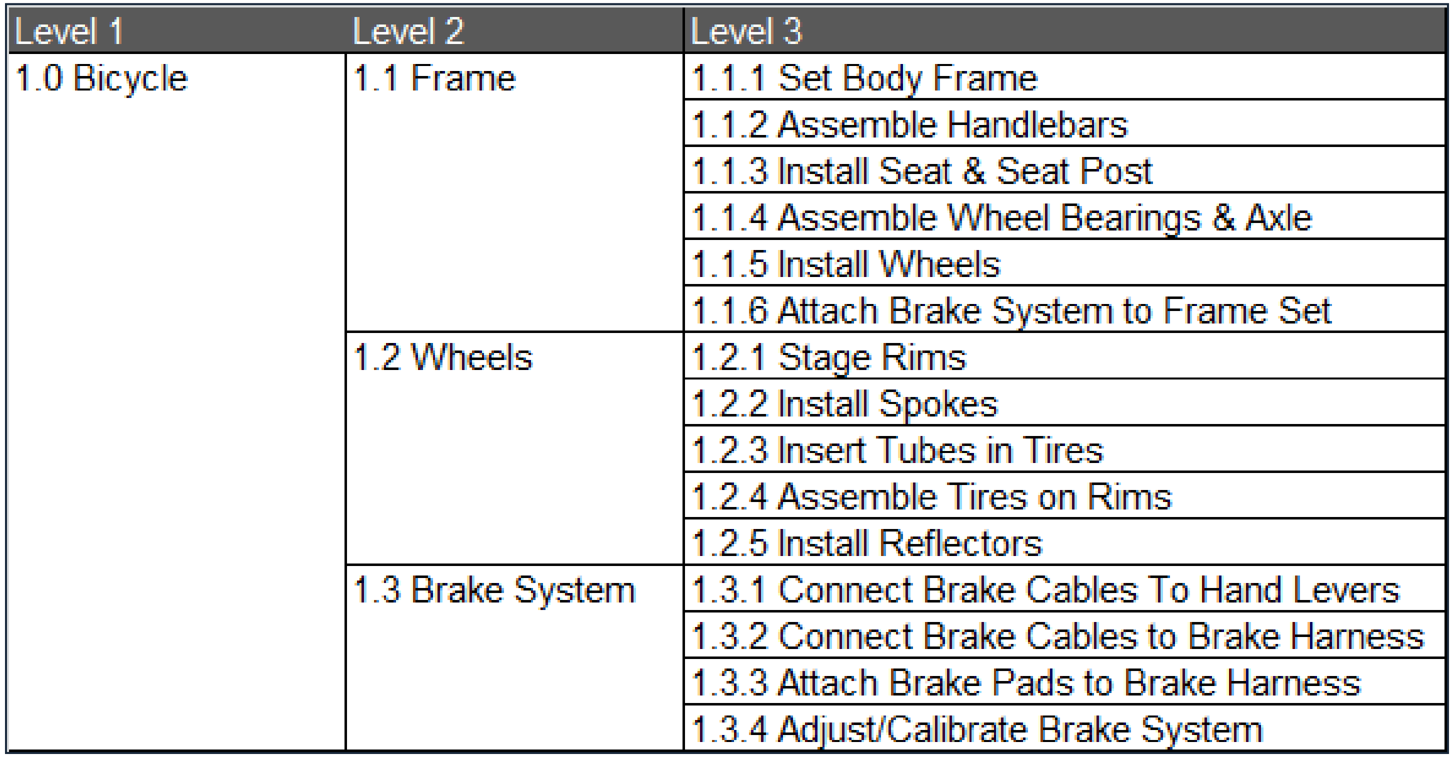
Develop Product

Test Product

Install Product

Project Planning Activities: WBS

Work Breakdown Structure Example: below is an example of a WBS for the assembly of a bicycle.

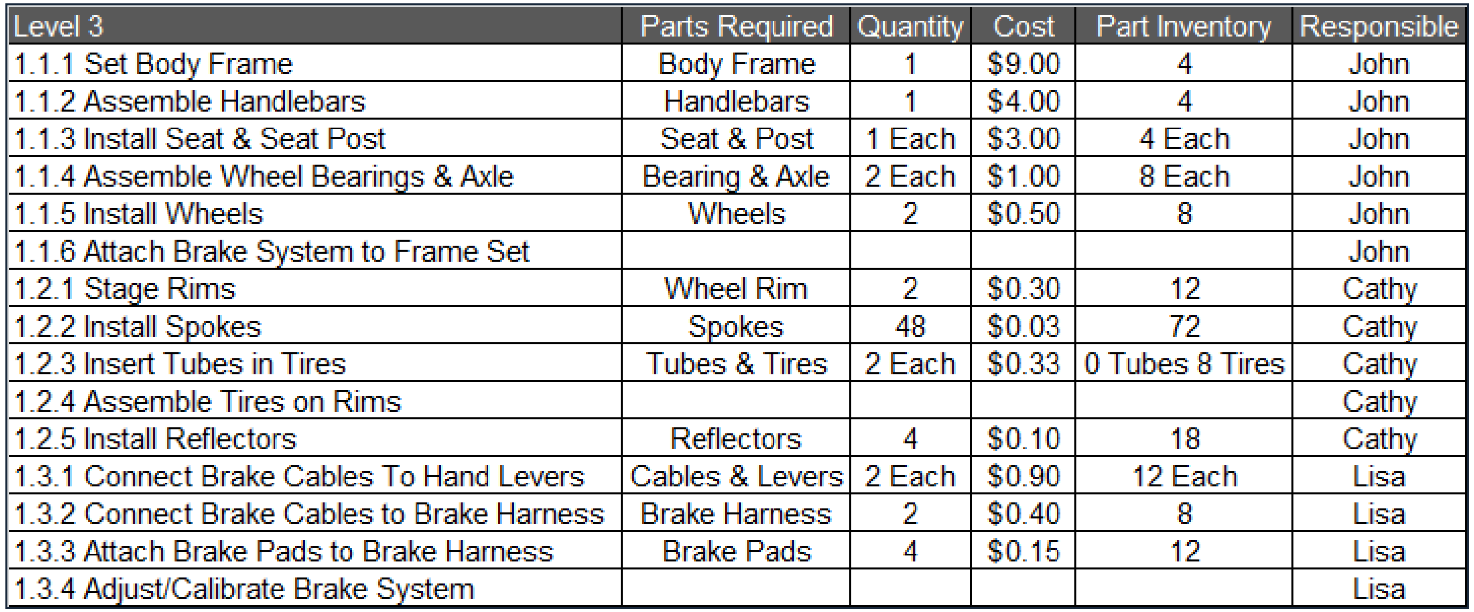


Level 1	Level 2	Level 3
1.0 Bicycle	1.1 Frame	1.1.1 Set Body Frame
		1.1.2 Assemble Handlebars
		1.1.3 Install Seat & Seat Post
		1.1.4 Assemble Wheel Bearings & Axle
		1.1.5 Install Wheels
		1.1.6 Attach Brake System to Frame Set
	1.2 Wheels	1.2.1 Stage Rims
		1.2.2 Install Spokes
		1.2.3 Insert Tubes in Tires
		1.2.4 Assemble Tires on Rims
		1.2.5 Install Reflectors
	1.3 Brake System	1.3.1 Connect Brake Cables To Hand Levers
		1.3.2 Connect Brake Cables to Brake Harness
		1.3.3 Attach Brake Pads to Brake Harness
		1.3.4 Adjust/Calibrate Brake System

Project Planning Activities: Resource Planning

Resource Estimation Plan

- Estimate Resource Requirements (Use Your WBS)
  - Parts, Hardware, Software, Human Resources etc.
- Plan resources
  - Establish who is responsible for what and when.
  - Determine quantity requirements and delivery dates etc.

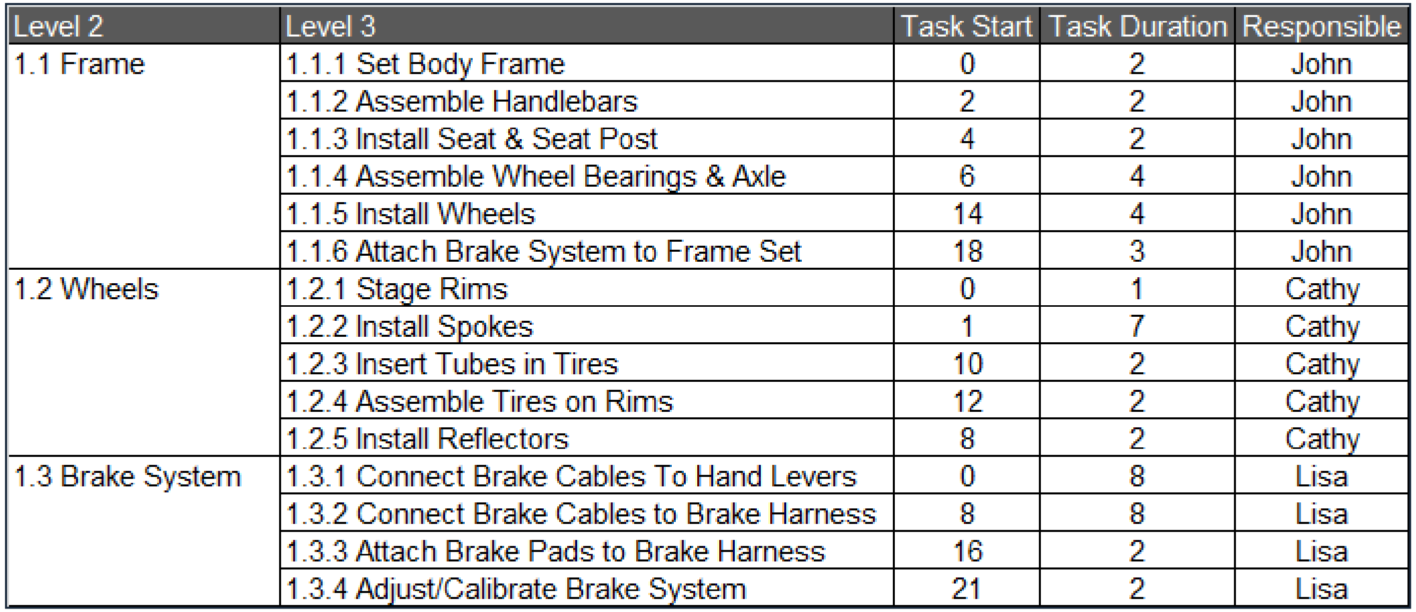


Level 3	Parts Required	Quantity	Cost	Part Inventory	Responsible
1.1.1 Set Body Frame	Body Frame	1	$9.00	4	John
1.1.2 Assemble Handlebars	Handlebars	1	$4.00	4	John
1.1.3 Install Seat & Seat Post	Seat & Post	1 Each	$3.00	4 Each	John
1.1.4 Assemble Wheel Bearings & Axle	Bearing & Axle	2 Each	$1.00	8 Each	John
1.1.5 Install Wheels	Wheels	2	$0.50	8	John
1.1.6 Attach Brake System to Frame Set					John
1.2.1 Stage Rims	Wheel Rim	2	$0.30	12	Cathy
1.2.2 Install Spokes	Spokes	48	$0.03	72	Cathy
1.2.3 Insert Tubes in Tires	Tubes & Tires	2 Each	$0.33	0 Tubes 8 Tires	Cathy
1.2.4 Assemble Tires on Rims					Cathy
1.2.5 Install Reflectors	Reflectors	4	$0.10	18	Cathy
1.3.1 Connect Brake Cables To Hand Levers	Cables & Levers	2 Each	$0.90	12 Each	Lisa
1.3.2 Connect Brake Cables to Brake Harness	Brake Harness	2	$0.40	8	Lisa
1.3.3 Attach Brake Pads to Brake Harness	Brake Pads	4	$0.15	12	Lisa
1.3.4 Adjust/Calibrate Brake System					Lisa

Project Planning Activities: Project Scheduling

Project Scheduling:

- Assign beginning times to each activity in the WBS
  (days are used for start and durations times in example below).
- Assign duration times to each activity in the WBS.
- Identify People Responsible and set completion dates.
- Represent schedules as Gantt charts or network diagrams.
- Identify critical dependencies between tasks.

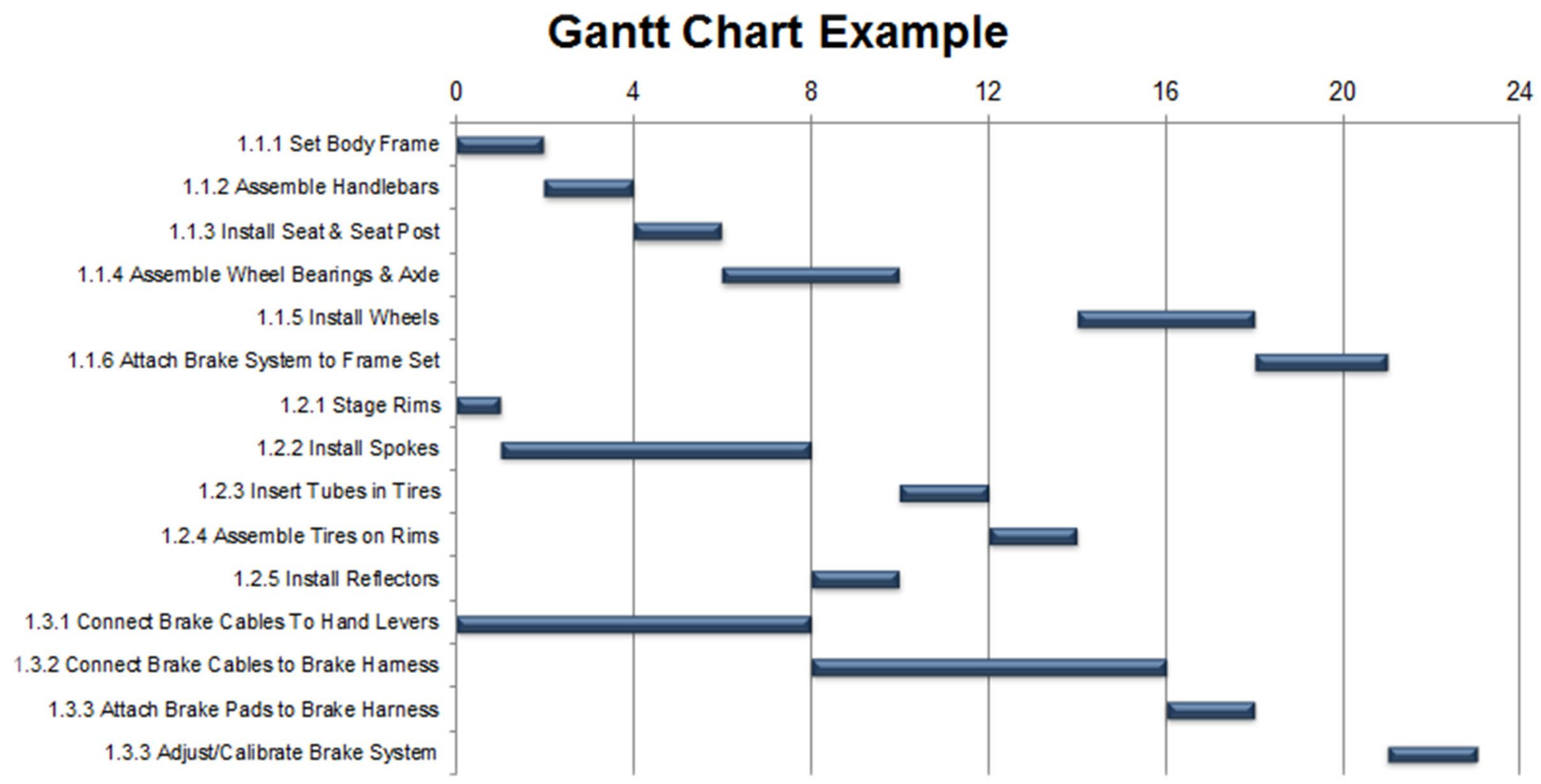


Level 2	Level 3	Task Start	Task Duration	Responsible
1.1 Frame	1.1.1 Set Body Frame	0	2	John
	1.1.2 Assemble Handlebars	2	2	John
	1.1.3 Install Seat & Seat Post	4	2	John
	1.1.4 Assemble Wheel Bearings & Axle	6	4	John
	1.1.5 Install Wheels	14	4	John
	1.1.6 Attach Brake System to Frame Set	18	3	John
1.2 Wheels	1.2.1 Stage Rims	0	1	Cathy
	1.2.2 Install Spokes	1	7	Cathy
	1.2.3 Insert Tubes in Tires	10	2	Cathy
	1.2.4 Assemble Tires on Rims	12	2	Cathy
	1.2.5 Install Reflectors	8	2	Cathy
1.3 Brake System	1.3.1 Connect Brake Cables To Hand Levers	0	8	Lisa
	1.3.2 Connect Brake Cables to Brake Harness	8	8	Lisa
	1.3.3 Attach Brake Pads to Brake Harness	16	2	Lisa
	1.3.4 Adjust/Calibrate Brake System	21	2	Lisa

Project Planning Activities: Project Scheduling

Project Schedule – Gantt Chart:

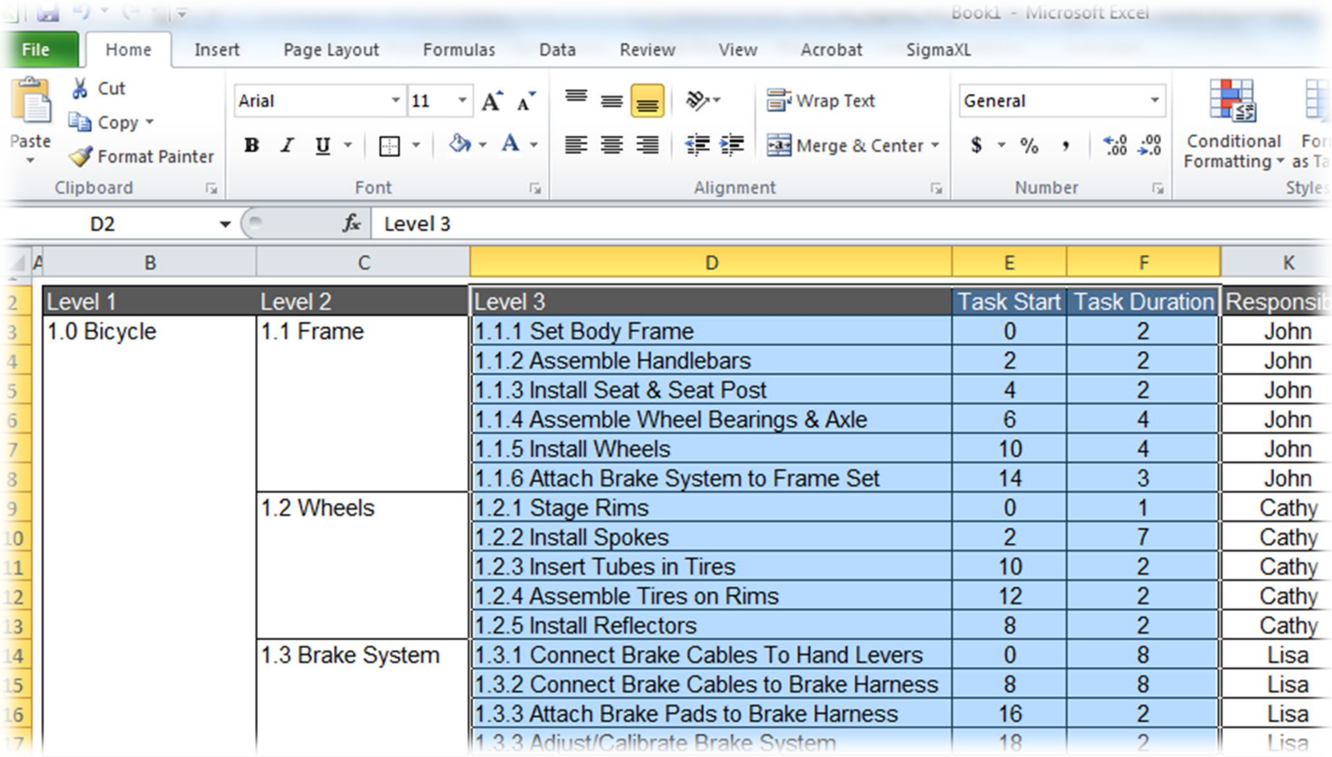
- The advantage of a Gantt chart is its ability to display the status of each task/activity at a glance.
- Because it is a graphic representation, it is easy to demonstrate the schedule of tasks and timelines to all the stakeholders.



Project Scheduling: Preparing a Gantt Chart in Excel

There are many tools available to create Gantt charts. Some of the most common and compatible for businesses are: Microsoft Project, Visio and Excel. In Microsoft Excel we use the bicycle example:

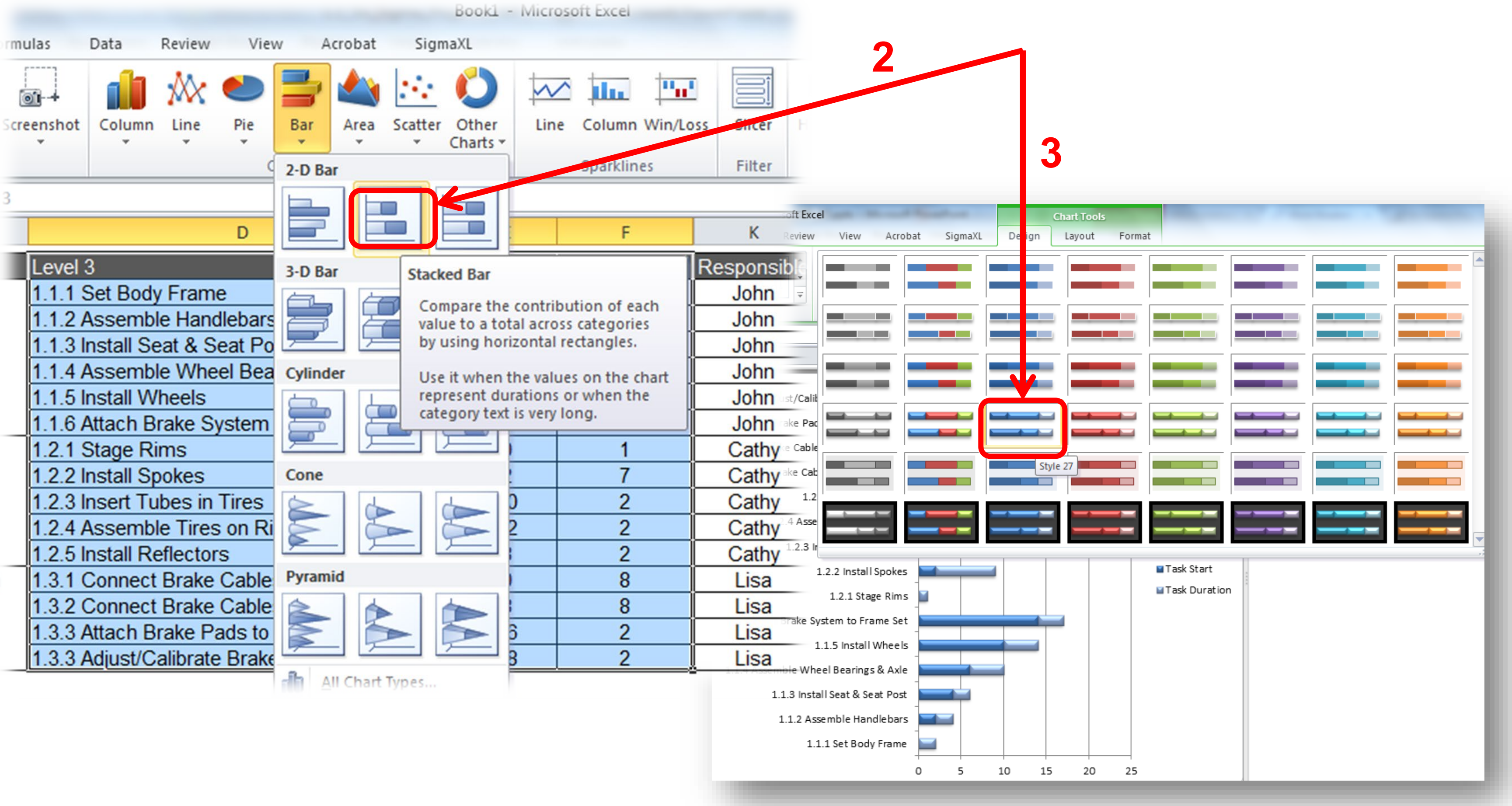
1. Select rows & columns of your tasks, start times and durations.



Level 1	Level 2	Level 3	Task Start	Task Duration	Responsi
1.0 Bicycle	1.1 Frame	1.1.1 Set Body Frame	0	2	John
		1.1.2 Assemble Handlebars	2	2	John
		1.1.3 Install Seat & Seat Post	4	2	John
		1.1.4 Assemble Wheel Bearings & Axle	6	4	John
		1.1.5 Install Wheels	10	4	John
		1.1.6 Attach Brake System to Frame Set	14	3	John
	1.2 Wheels	1.2.1 Stage Rims	0	1	Cathy
		1.2.2 Install Spokes	2	7	Cathy
		1.2.3 Insert Tubes in Tires	10	2	Cathy
		1.2.4 Assemble Tires on Rims	12	2	Cathy
		1.2.5 Install Reflectors	8	2	Cathy
	1.3 Brake System	1.3.1 Connect Brake Cables To Hand Levers	0	8	Lisa
		1.3.2 Connect Brake Cables to Brake Harness	8	8	Lisa
		1.3.3 Attach Brake Pads to Brake Harness	16	2	Lisa
		1.3.3 Adjust/Calibrate Brake System	18	2	Lisa

Project Scheduling: Preparing a Gantt Chart in Excel

2. On the Insert tab select Bar > Stacked Bar.
3. On the Design tab choose a design style from the style dropdown.



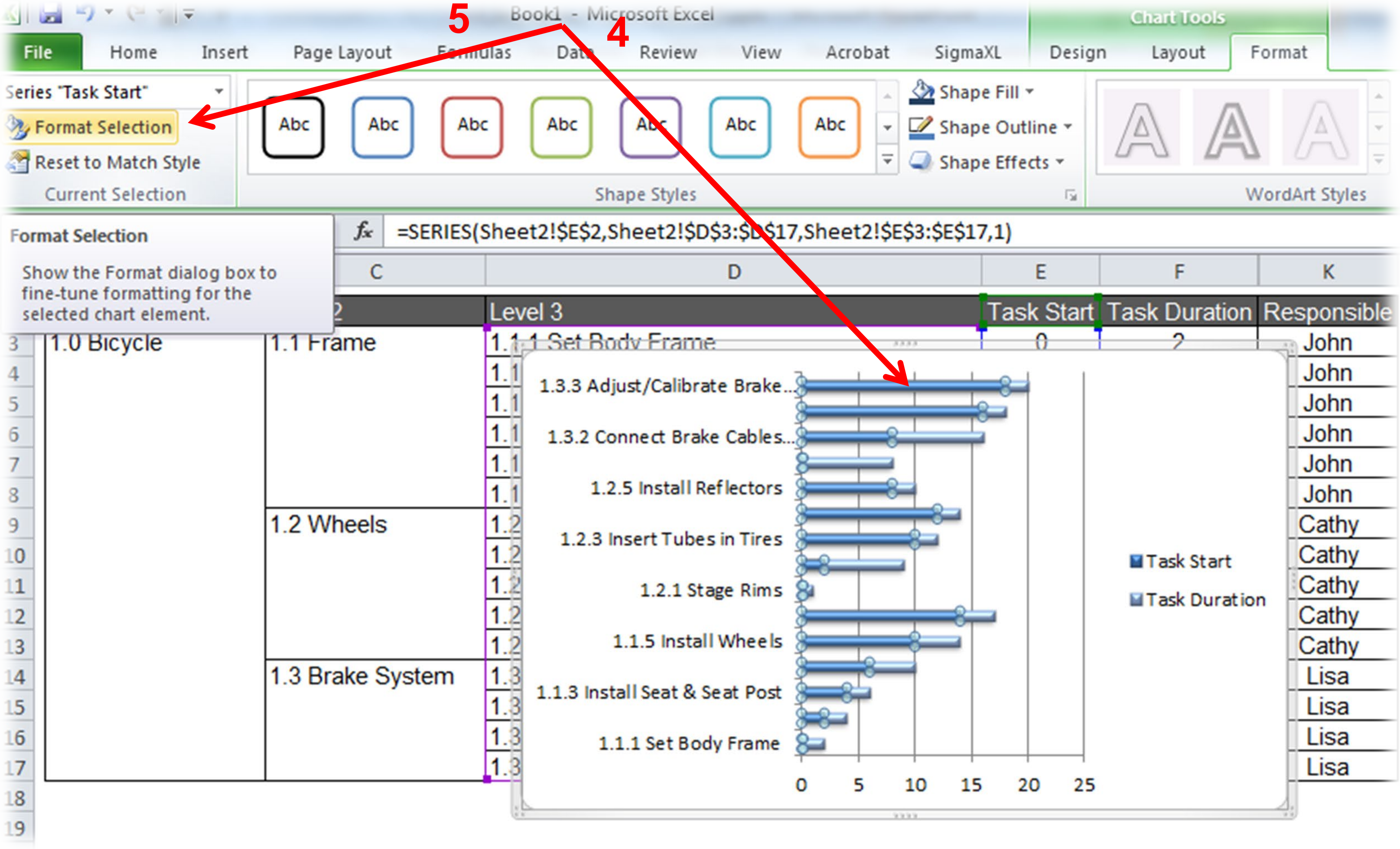
Stacked Bar

Compare the contribution of each value to a total across categories by using horizontal rectangles.

Use it when the values on the chart represent durations or when the category text is very long.

Project Scheduling: Preparing a Gantt Chart in Excel

4. Activate the “start tasks” on your graph.
5. From the format tab select “format selection”.



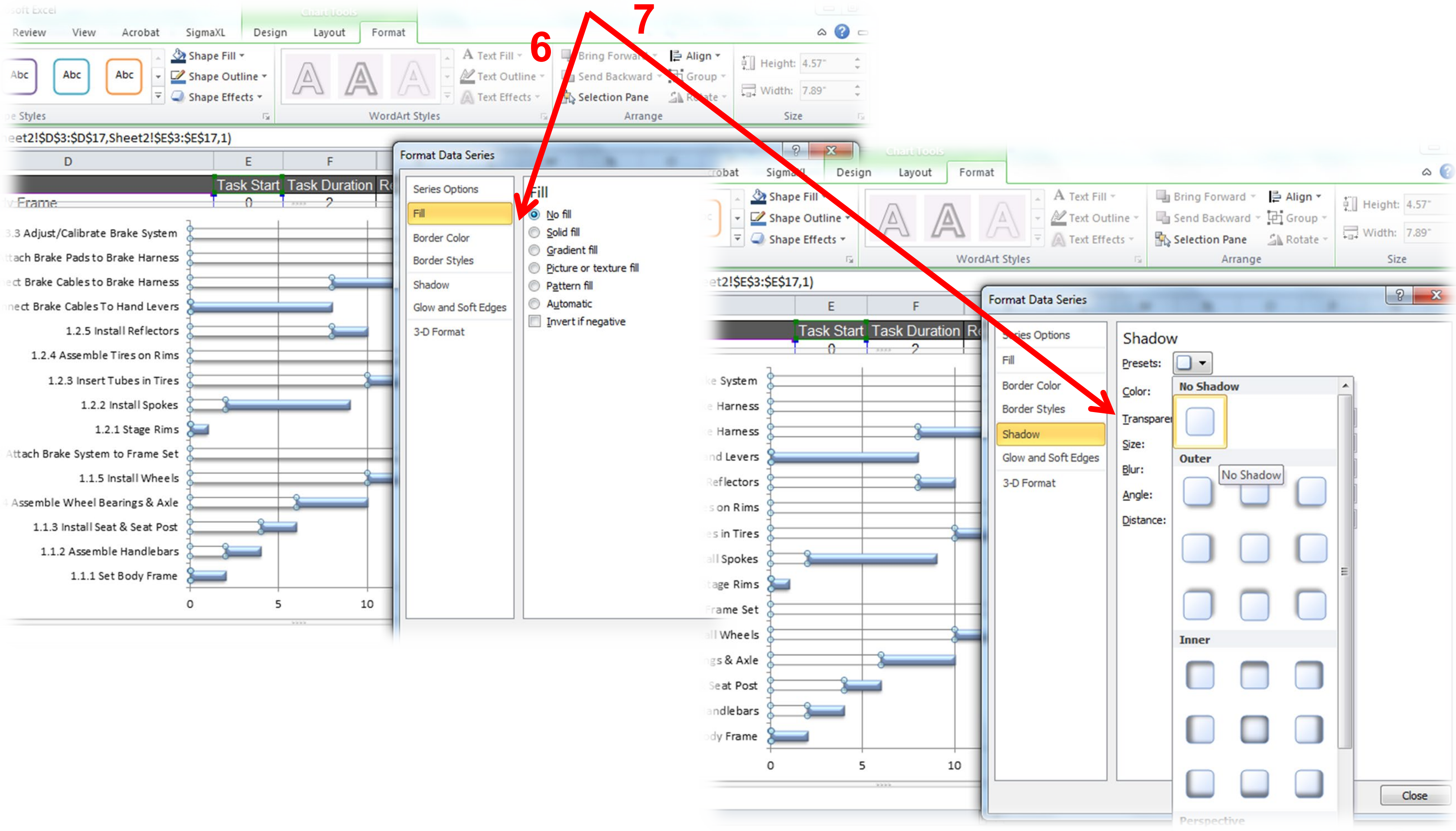
=SERIES(Sheet2!$E$2,Sheet2!$D$3:$D$17,Sheet2!$E$3:$E$17,1)

Format Selection

Show the Format dialog box to fine-tune formatting for the selected chart element.

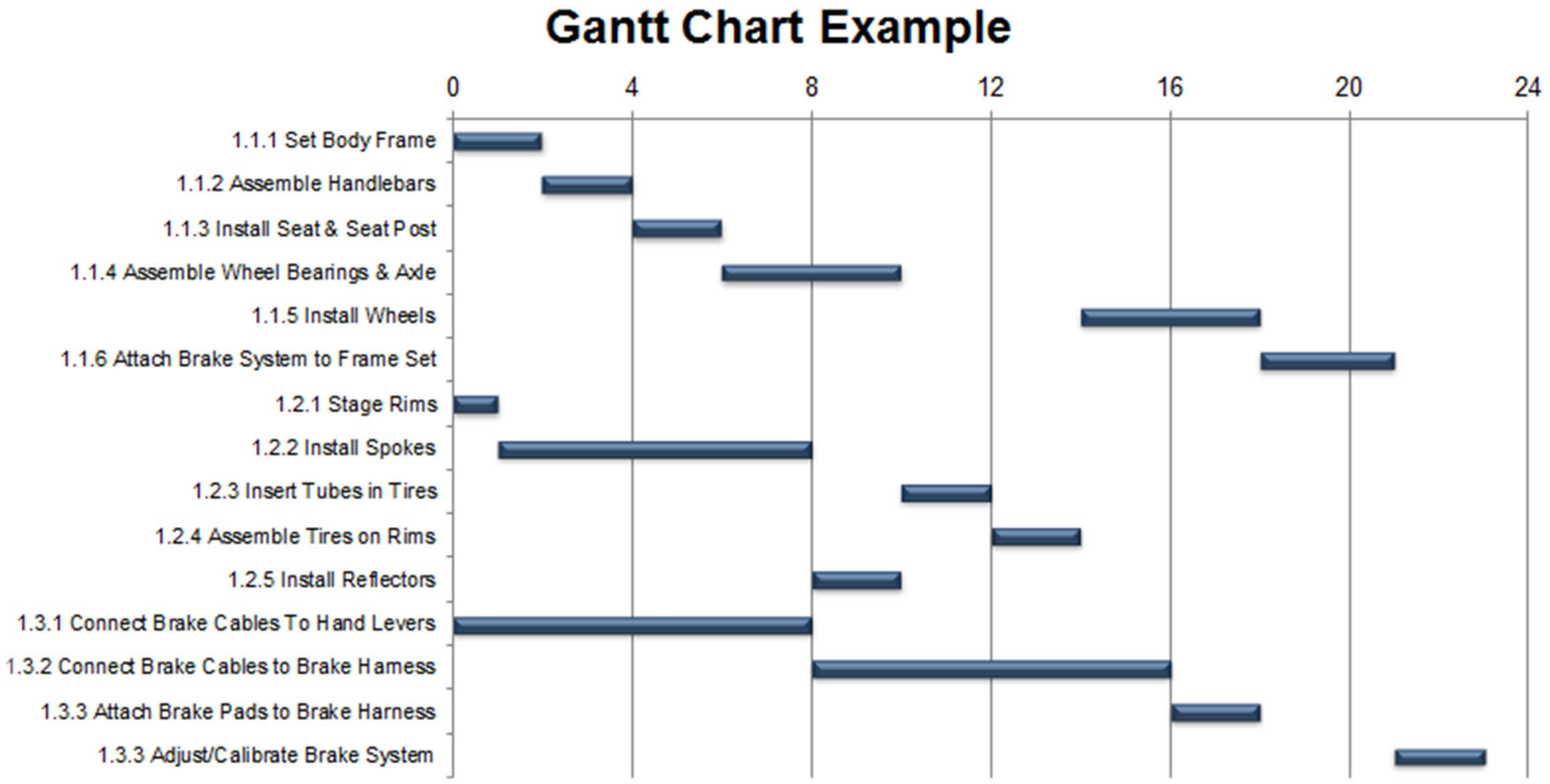
Project Scheduling: Preparing a Gantt Chart in Excel

6. Select Fill > No Fill.
7. Select Shadow > Presets > No Shadow > Close.



Project Scheduling: Preparing a Gantt Chart in Excel

8. Format your new Gantt chart to your own preferences.



Gantt Chart Example

Project Scheduling: Critical Path Method

Critical Path Method (CPM) is a project modeling technique used to identify the set of activities that are most influential to a projects completion timeline.

- Critical Path tasks are those that others are dependent upon.
- Project timelines cannot be shortened without shortening the tasks or activities that are identified as critical path items.

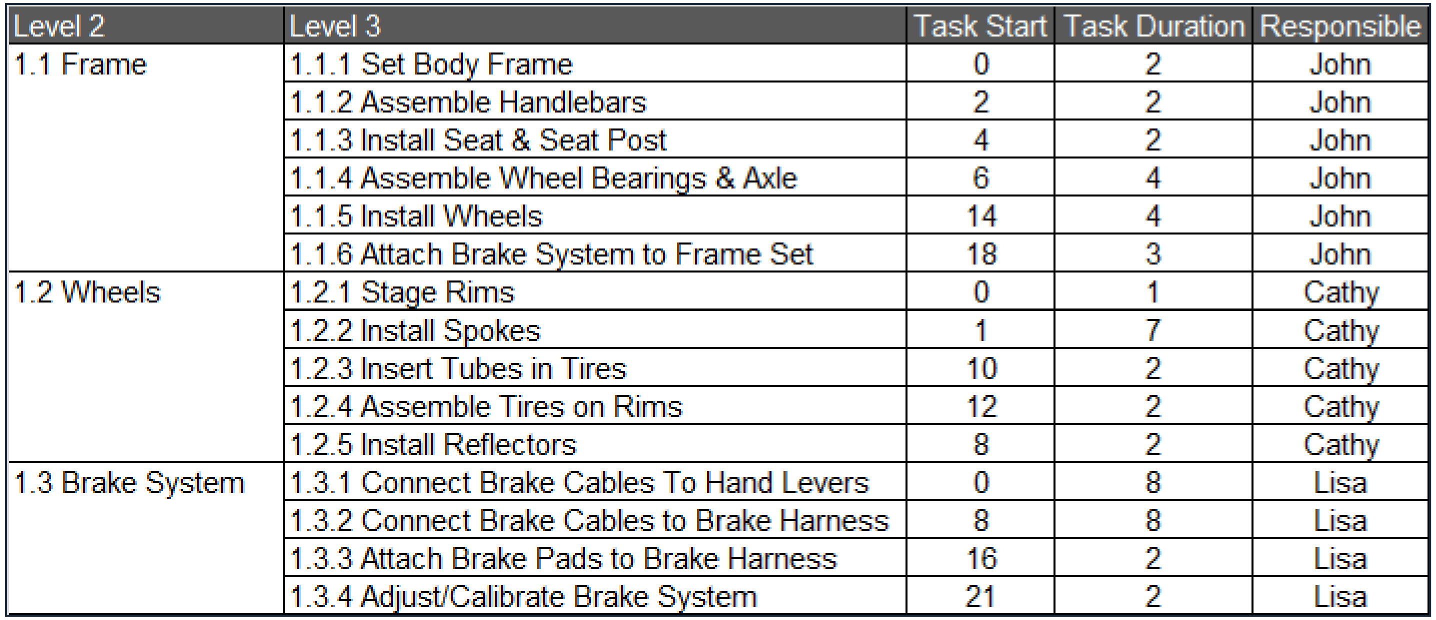
Steps To Using The Critical Path Method

1. List all activities necessary to complete the project.
2. Determine the time or duration of each activity.
3. Identify the dependencies between the activities.

Use Your Work Breakdown Structure (WBS).

Project Scheduling: Critical Path Method

Critical Path Method Example: Let’s continue with our bicycle example. Note that the tasks in the WBS are already set to start based on their dependencies.



Level 2	Level 3	Task Start	Task Duration	Responsible
1.1 Frame	1.1.1 Set Body Frame	0	2	John
	1.1.2 Assemble Handlebars	2	2	John
	1.1.3 Install Seat & Seat Post	4	2	John
	1.1.4 Assemble Wheel Bearings & Axle	6	4	John
	1.1.5 Install Wheels	14	4	John
	1.1.6 Attach Brake System to Frame Set	18	3	John
1.2 Wheels	1.2.1 Stage Rims	0	1	Cathy
	1.2.2 Install Spokes	1	7	Cathy
	1.2.3 Insert Tubes in Tires	10	2	Cathy
	1.2.4 Assemble Tires on Rims	12	2	Cathy
	1.2.5 Install Reflectors	8	2	Cathy
1.3 Brake System	1.3.1 Connect Brake Cables To Hand Levers	0	8	Lisa
	1.3.2 Connect Brake Cables to Brake Harness	8	8	Lisa
	1.3.3 Attach Brake Pads to Brake Harness	16	2	Lisa
	1.3.4 Adjust/Calibrate Brake System	21	2	Lisa

- For example, Installing the wheels in step 1.1.5 requires the whole set of steps in 1.2 to be completed so 1.1.5 starts on day 14 when the wheels are ready.

Project Scheduling: Critical Path Method

Continuing this example; the last task, 1.3.4 Adjusting the Brake System, completes on day 23 and can't begin until day 21 because the brake system must be attached to the bike frame before being calibrated.

1.1.6 Completes on day 21

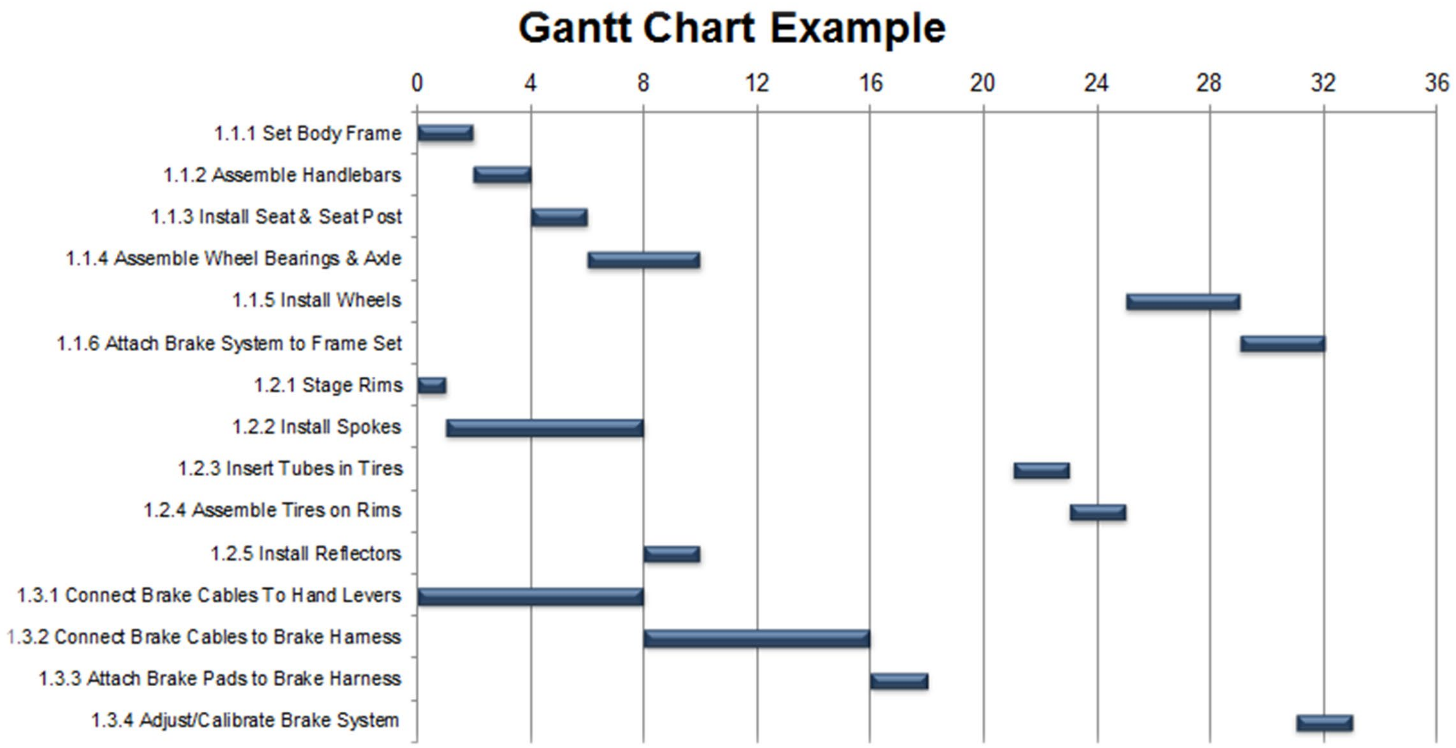
1.3.4 Starts on day 21

	Level 3	Task Start	Task Duration	Responsible
	1.1.1 Set Body Frame	0	2	John
	1.1.2 Assemble Handlebars	2	2	John
	1.1.3 Install Seat & Seat Post	4	2	John
	1.1.4 Assemble Wheel Bearings & Axle	6	4	John
	1.1.5 Install Wheels	14	4	John
	1.1.6 Attach Brake System to Frame Set	18	3	John
1.2 Wheels	1.2.1 Stage Rims	0	1	Cathy
	1.2.2 Install Spokes	1	7	Cathy
	1.2.3 Insert Tubes in Tires	10	2	Cathy
	1.2.4 Assemble Tires on Rims	12	2	Cathy
	1.2.5 Install Reflectors	8	2	Cathy
	1.3.1 Connect Brake Cables To Hand Levers	0	8	Lisa
	1.3.2 Connect Brake Cables to Brake Harness	8	8	Lisa
	1.3.3 Attach Brake Pads to Brake Harness	16	2	Lisa
	1.3.4 Adjust/Calibrate Brake System	21	2	Lisa

Project Scheduling: Critical Path Method

Earlier if you noticed when we conducted our resource plan that we were out of tire tubes then you might have managed your resources differently.

Let’s assume the tubes take 21 days to receive from the supplier. Below is the adjustment to our Gantt chart. Notice that the project timeline is pushed out by 8 days, finishing on day 31 instead of day 23.



Project Scheduling: PERT

Program Evaluation & Review Technique (PERT): is a method of evaluation that can be applied to time or cost.

- PERT provides a weighted assessment of time or cost
- PERT uses 3 parameters for estimation:
  - Optimistic or Best Case Scenario represented by ‘O’.
  - Pessimistic or Worst Case Scenario represented by ‘P’.
  - Most Likely Scenario represented by ‘ML’.
- PERT Equation:

$$(O + (4*ML) + P) / 6$$

  - The PERT equation provides for heavier weighting of the most likely scenario but also considers the best and worst cases. In the event that a best or worst case scenario is an extreme situation, the PERT will account for it in a weighted manner.

Project Scheduling: PERT

Let’s apply the PERT formula to our estimates of receiving our needed tire tubes from the supplier and then change our schedule based on the result.

- Your experience with this supplier tells you that they typically over estimate the time required to deliver. Therefore, their 21 day lead time on the tubes should be considered a “worst case scenario”.
- Procurement has indicated that tubes have arrived in as few as 6 days from this supplier. So, your best case scenario is 6 days.
- The most likely scenario you decide will be the median of this suppliers delivery time which is 10 days.
- Therefore:

(O + (4*ML) + P) / 6

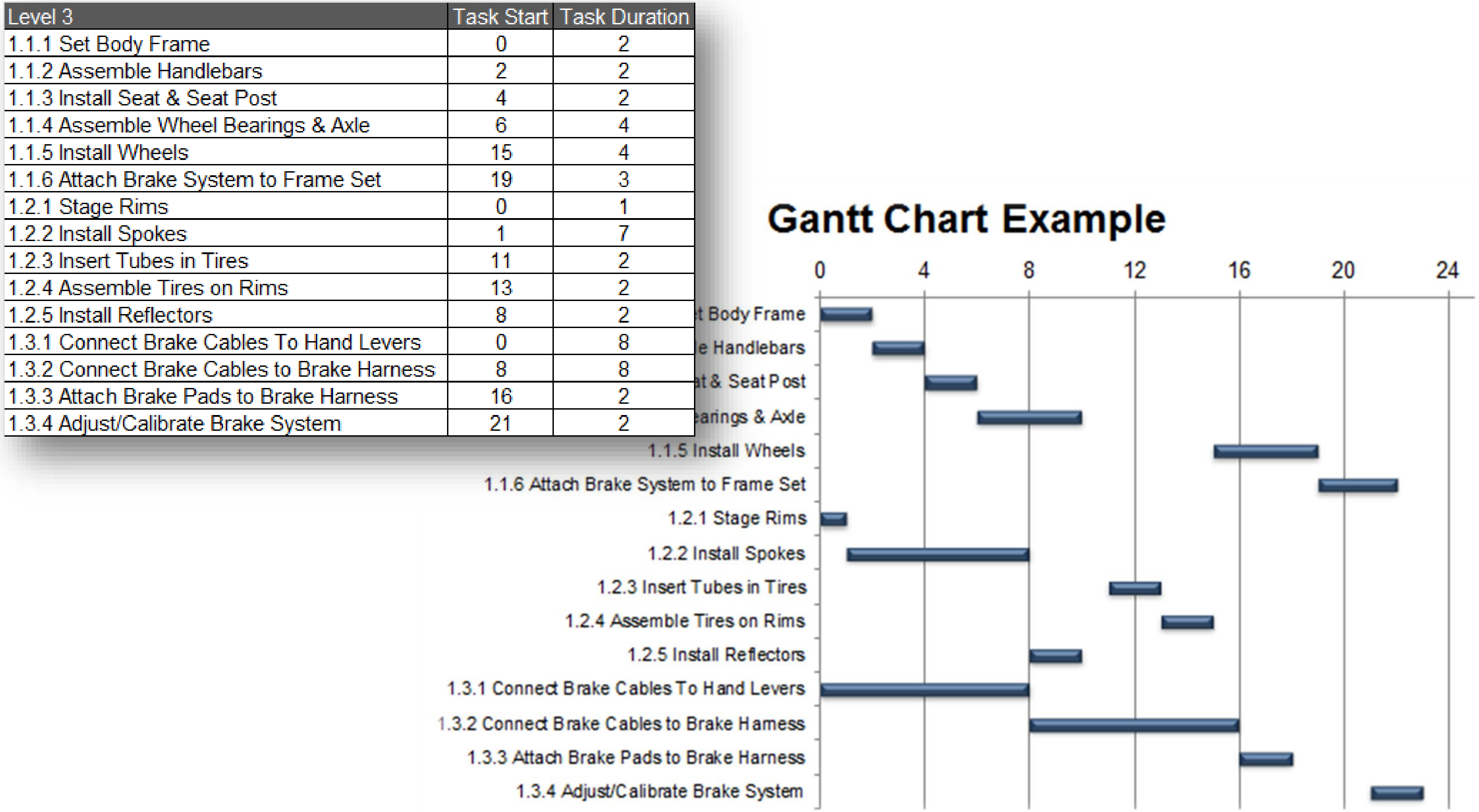
(6 + (4*10) + 21) / 6

(6 + 40 + 21) / 6

67 / 6 = 11

Project Scheduling: PERT

Using 11 days in your project plan instead of 21 gives you a more confident estimate of time and actually removes tire tubes as the primary detracting critical path item.

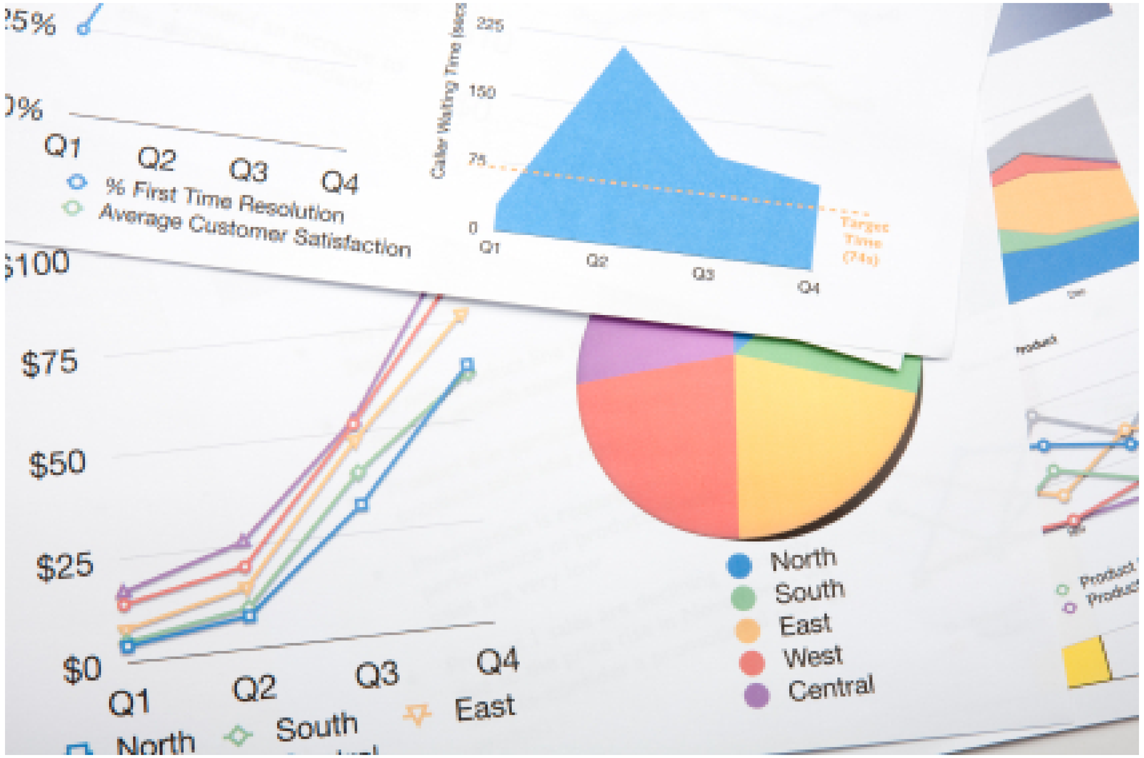


Level 3	Task Start	Task Duration
1.1.1 Set Body Frame	0	2
1.1.2 Assemble Handlebars	2	2
1.1.3 Install Seat & Seat Post	4	2
1.1.4 Assemble Wheel Bearings & Axle	6	4
1.1.5 Install Wheels	15	4
1.1.6 Attach Brake System to Frame Set	19	3
1.2.1 Stage Rims	0	1
1.2.2 Install Spokes	1	7
1.2.3 Insert Tubes in Tires	11	2
1.2.4 Assemble Tires on Rims	13	2
1.2.5 Install Reflectors	8	2
1.3.1 Connect Brake Cables To Hand Levers	0	8
1.3.2 Connect Brake Cables to Brake Harness	8	8
1.3.3 Attach Brake Pads to Brake Harness	16	2
1.3.4 Adjust/Calibrate Brake System	21	2

Project Planning Activities

Budget or Financial Plan

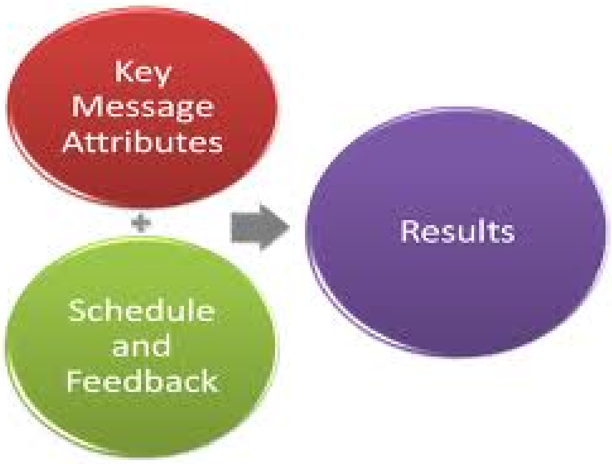
- Planned expenses
- Planned revenues
- Budget forecast



Project Planning Activities

Communication Plan

- Establish communication procedures among management, team members, and relevant stakeholders.
- Determine the communication schedule.
- Define the acceptable modes of communication.



Project Planning Activities

Risk Management Plan

- Identify the sources of project risks and estimate the effects of those risks.
- Risks might arise from new technology, availability of resources, lack of inputs from customers, business risks etc.
- Assess the impact of risk to the customers/stakeholders.
- Calculate the probability of risk occurrence based on previous similar projects or industry benchmarks
- Initiate mitigation and contingency plans
- Review risks on a periodic basis

Project Planning Tools Advantages

- Project planning tools are very useful to organize and communicate project plans, status, and projections.
- They help link tasks and sub-tasks or other work elements to get a whole view of what needs to be accomplished.
- They allow a more objective comparison of alternative solutions and provide consistent coverage of responsibilities.
- They allow for effective scope control and change management.
- They facilitate effective communication with all project participants and stakeholders.
- They help define management reviews.
- They act as an effective monitoring mechanism for the project.
- They establish project baselines for progress reviews and control points.

1.4 Lean Fundamentals

Black Belt Training: Define Phase

1.1 Six Sigma Overview
- 1.1.1 What is Six Sigma
- 1.1.2 Six Sigma History
- 1.1.3 Six Sigma Approach Y = f(x)
- 1.1.4 Six Sigma Methodology
- 1.1.5 Roles and Responsibilities

1.2 Six Sigma Fundamentals
- 1.2.1 Defining a Process
- 1.2.2 VOC and CTQs
- 1.2.3 QFD
- 1.2.4 Cost of Poor Quality (COPQ)
- 1.2.5 Pareto Analysis (80:20 rule)

1.3 Lean Six Sigma Projects
- 1.3.1 Six Sigma Metrics
- 1.3.2 Business Case and Charter
- 1.3.3 Project Team Selection
- 1.3.4 Project Risk Management
- 1.3.5 Project Planning

1.4 Lean Fundamentals
- 1.4.1 Lean and Six Sigma
- 1.4.2 History of Lean
- 1.4.3 The Seven Deadly Muda
- 1.4.4 Five-S (5S)

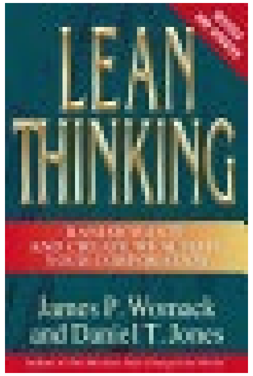
1.4.1 Lean and Six Sigma

What is Lean?

- A **lean enterprise** intends to eliminate waste and allow only value to be pulled through its system.
- **Lean manufacturing** is characterized by:
  - Identifying and driving value
  - Establishing flow and pull systems
  - Creating production availability and flexibility
  - Zero waste
- Waste Elimination
  - Waste identification and elimination is critical to any successful lean enterprise.
  - Elimination of waste enables flow, drives value, cuts cost, and provides flexible and available production.

The 5 Lean Principles

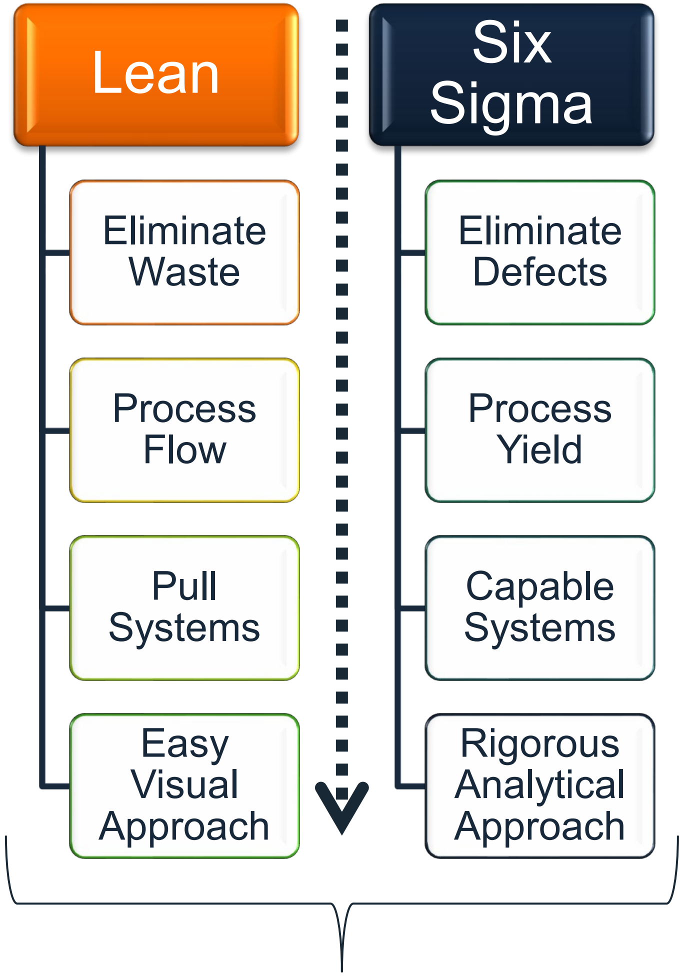
- The following 5 principles of lean are taken from the book *Lean Thinking* (1996) by James P. Womack and Daniel T. Jones.



1. Specify value desired by customers.
2. Identify the value stream.
3. Make the product flow continuous.
4. Introduce pull systems where continuous flow is possible.
5. Manage toward perfection so that the number of steps and the amount of time and information needed to serve the customer continually falls.

Lean & Six Sigma

- Lean and Six Sigma both have the objectives of producing high value (**quality**) at lower costs (**efficiency**).
- They approach these objectives in somewhat different manners but in the end, both Lean and Six Sigma drive out waste, reduce defects, improve processes, and stabilize the production environment.
- Lean and Six Sigma are a perfect combination of tools for improving quality and efficiency.

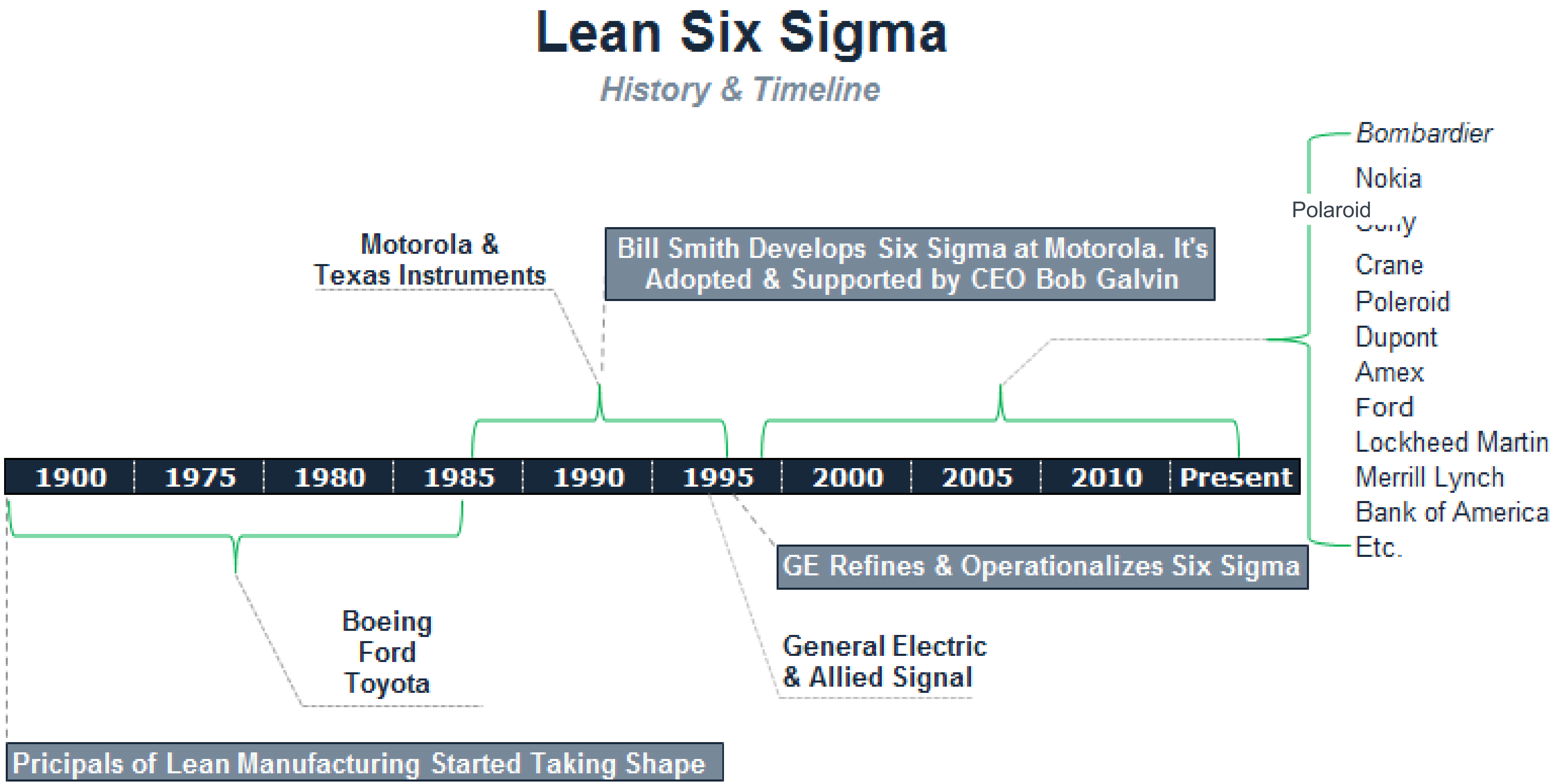


Lean	Six Sigma
Eliminate Waste	Eliminate Defects
Process Flow	Process Yield
Pull Systems	Capable Systems
Easy Visual Approach	Rigorous Analytical Approach

Quality & Value for the Customer
Efficiency for the Business

1.4.2 History of Lean

History of Lean



Lean Six Sigma

History & Timeline

1900 | 1975 | 1980 | 1985 | 1990 | 1995 | 2000 | 2005 | 2010 | Present

Pricipals of Lean Manufacturing Started Taking Shape

Boeing
Ford
Toyota

Motorola &
Texas Instruments

Bill Smith Develops Six Sigma at Motorola. It's Adopted & Supported by CEO Bob Galvin

GE Refines & Operationalizes Six Sigma

General Electric
& Allied Signal

Bombardier
Nokia
Polaroid
Crane
Poleroid
Dupont
Amex
Ford
Lockheed Martin
Merrill Lynch
Bank of America
Etc.

History of Lean

- Lean thinking originated, as far as is known, the 1400s.
- Henry Ford established the first mass production system in 1913 by combining standard parts, conveyors, and work flow.
- Decades later, Kiichiro Toyoda and Taiichi Ohno at Toyota improved and implemented various new concepts and tools (e.g., value stream, takt time, kanban etc.) based on Ford's effort.
- Toyota developed what is known today as the Toyota Production System (TPS) based on lean principles.

History of Lean

- Starting in the mid 1990s, Lean became extensively recognized and implemented when more and more Fortune 100 companies began to adopt Lean and Six Sigma.
- The term “Lean manufacturing” was introduced by James Womack in the 1990s.
- Lean and Six Sigma share similar objectives, work hand in hand, and have benefited from one another in the past 30 years.

1.4.3 Seven Deadly Muda

The 7 Deadly Muda

- The Japanese word for waste is "muda."
- There are 7 commonly recognized forms of waste, often referred to as the "7 deadly muda."

1. Defects
2. Overproduction
3. Over-Processing
4. Inventory
5. Motion
6. Transportation
7. Waiting

The 7 Deadly Muda: Defects

- Defects or defectives are an obvious waste for any working environment or production system.
- Defects require rework during production and/or after the product is returned from an unhappy customer.
- Some defects are difficult to solve and they create “workarounds” and hidden factories.
- Eliminating defects is a sure way to improve product quality, customer satisfaction, and production costs.

The 7 Deadly Muda: Overproduction

- Overproduction is wasteful because your system expends energy and resources to produce more materials than the customer or next function requires.

- Overproduction is one of the most detrimental of the seven deadly muda because it leads to many others:
- Inventory
- Transportation
- Waiting etc.

The 7 Deadly Muda: Over-processing

- **Over-processing** occurs any time more work is done than is required by the next process step, operation, or consumer.

- Over-processing also includes being over capacity (scheduling more workers than required or having more machines than necessary).

- Another form of over processing can be buying tools or software that are overkill (more precise, complex, or expensive than required).

The 7 Deadly Muda: Inventory

- **Inventory** is an often overlooked waste. Look at the picture above and imagine all the time, materials, and logistics that went into establishing such an abundance of inventory.

- If this were your personal business, and inventory velocity was not matched with production, how upset would you be?

The 7 Deadly Muda: Motion

- **Motion** is another form of waste often occurring as a result of poor setup, configuration, or operating procedures.
- Wasted motion can be experienced by machines or humans.
- Wasted motion is very common with workers who are unaware of the impact of small unnecessary movements in repetitive tasks.
- Wasted motion is exaggerated by repetition or recurring tasks.

The 7 Deadly Muda: Transportation

- **Transportation** is considered wasteful because it does *nothing* to add value or transform the product.
- Imagine for a moment driving to and from work twice before getting out of your car to go into work. . .

- That is waste in the form of transportation.
- The less driving you have to do, the better.
- In a similar way, the less transportation a product has to endure, the better. There would be fewer opportunities for delay, destruction, loss, damage etc.

The 7 Deadly Muda: Waiting

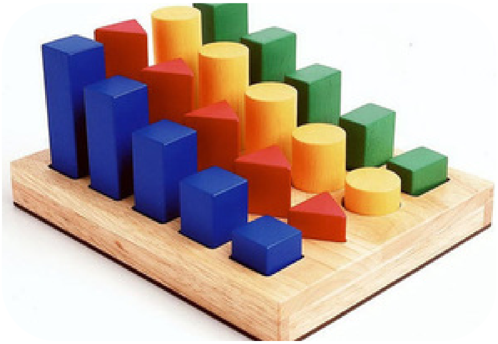
- **Waiting** is an obvious form of waste and is typically a symptom of an upstream problem.
- Waiting is usually caused by inefficiency, bottlenecks, or poorly-designed work flows within the value stream.
- Waiting can also be caused by inefficient administration.
- Reduction in waiting time will require thoughtful applications of lean and process improvement.

1.4.4 Five-S (5S)

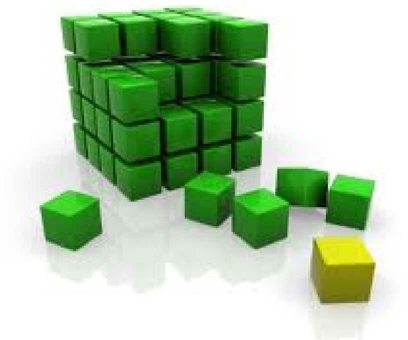
What is 5S?

- **5S** is systematic method to organize, order, clean, and standardize a workplace…and to keep it that way!
  - 5S is a methodology of organizing and improving the work environment.
- 5S is summarized in five Japanese words all starting with the letter S:
  - **Seiri** (sorting)
  - **Seiton** (straightening)
  - **Seiso** (shining)
  - **Seiketsu** (standardizing)
  - **Shisuke** (sustaining)
- 5S was originally developed in Japan and is widely used to optimize the workplace to increase productivity and efficiency.

Five-S (5S)



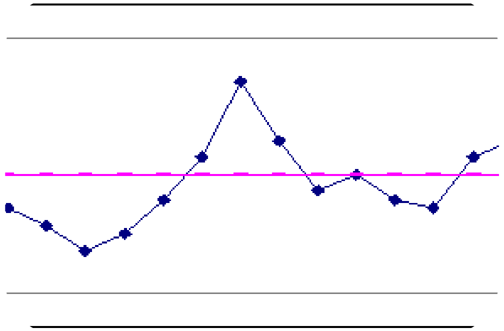
Sort



Set in Order

Shine

Standardize



Sustain

Goals of 5S

- Reduced waste
- Reduced cost
- Establish a work environment that is:
  - self-explaining
  - self-ordering
  - self-regulating
  - self-improving.
  - Where there is/are **no more:**
    - Wandering and/or searching
    - Waiting or delaying
    - Secret hiding spots for tools
    - Obstacles or detours
    - Extra pieces, parts, materials etc.
    - Injuries
    - Waste.

Benefits of 5S Systems

- Reduced changeovers
- Reduced defects
- Reduced waste
- Reduced delays
- Reduced injuries
- Reduced breakdowns
- Reduced complaints
- Reduced red ink
- Higher quality
- Lower costs
- Safer work environment
- Greater associate and equipment capacity.

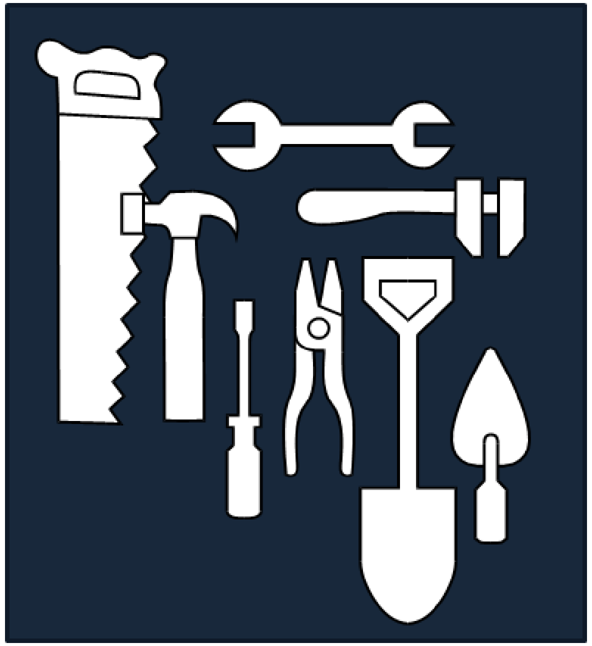
Reported Results of 5S Systems

- Cut in floor space: 60%
- Cut in flow distance: 80%
- Cut in accidents: 70%
- Cut in rack storage: 68%
- Cut in number of forklifts: 45%
- Cut in machine changeover time: 62%
- Cut in annual physical inventory time: 50%
- Cut in classroom training requirements: 55%
- Cut in nonconformance in assembly: 96%
- Increase in test yields: 50%
- Late deliveries: 0%
- Increase in throughput: 15%

Sorting (Seiri)

- Go through all the tools, parts, equipment, supply, and material in the workplace.
- Categorize them into two major groups: needed and unneeded.
- Eliminate the unneeded items from the workplace. Dispose of or recycle those items.
- Keep the needed items and sort them in the order of priority. **When in doubt…throw it out!**

Straightening (Seiton)



- **Straightening** in 5S is also called **setting in order.**
- Label each needed item.
- Store items at their best locations so that the workers can find them easily whenever they needed any item.
- Reduce the motion and time required to locate and obtain any item whenever it is needed.
- Promote an efficient work flow path.
- Use visual aids like the tool board image on this page.

Shining (Seiso)

- **Shining** in 5S is also called **sweeping**.
- Clean the workplace thoroughly.
- Maintain the tidiness of the workplace.
- Make sure every item is located at the specific location where it should be.
- Create the ownership in the team to keep the work area clean and organized.

Standardizing (Seiketsu)

- **Standardize** the workstation and the layout of tools, equipment and parts.

- Create identical workstations with a consistent way of storing the items at their specific locations so that workers can be moved around to any workstation any time and perform the same task.

Sustaining (Shisuke)

- **Sustaining** in 5S is also called **self-discipline**.
- Create the culture in the team to follow the first four S’s consistently.
- Avoid falling back to the old ways of cluttered and unorganized work environment.
- Keep the momentum of optimizing the workplace.
- Promote innovations of workplace improvement.
- Sustain the first fours S’s using:
  - 5S Maps
  - 5S Schedules
  - 5S Job cycle charts
  - Integration of regular work duties
  - 5S Blitz schedules
  - Daily workplace scans.

Simplified Summary of 5S

1. **Sort** – "when in doubt, move it out."
2. **Set in Order** – Organize all necessary tools, parts, and components of production. Use visual ordering techniques wherever possible.
3. **Shine** – Clean machines and/or work areas. Set regular cleaning schedules and responsibilities.
4. **Standardize** – Solidify previous three steps, make 5S a regular part of the work environment and everyday life.
5. **Sustain** – Audit, manage, and comply with established 5S guidelines for your business or facility.

Five-S (5S)

- A few words about 5S and the Lean Enterprise
  - As a method, 5S generates immediate improvements.
  - 5S is one of many effective lean methods that create observable results.
  - It is tempting to implement 5S alone without considering the entire value stream.
  - However, it is advisable to consider a well-planned lean manufacturing approach to the entire production system.

2.0 Measure Phase

Black Belt Training: Measure Phase

2.1 Process Definition
- 2.1.1 Cause and Effect Diagrams
- 2.1.2 Cause and Effects Matrix
- 2.1.3 Process Mapping
- 2.1.4 FMEA: Failure Modes & Effects Analysis
- 2.1.5 Theory of Constraints

2.2 Six Sigma Statistics
- 2.2.1 Basic Statistics
- 2.2.2 Descriptive Statistics
- 2.2.3 Distributions and Normality
- 2.2.4 Graphical Analysis

2.3 Measurement System Analysis
- 2.3.1 Precision and Accuracy
- 2.3.2 Bias, Linearity, and Stability
- 2.3.3 Gage R&R
- 2.3.4 Variable and Attribute MSA

2.4 Process Capability
- 2.4.1 Capability Analysis
- 2.4.2 Concept of Stability
- 2.4.3 Attribute and Discrete Capability
- 2.4.4 Monitoring Techniques

2.1 Process Definition

Black Belt Training: Measure Phase

2.1 Process Definition
- 2.1.1 Cause and Effect Diagrams
- 2.1.2 Cause and Effects Matrix
- 2.1.3 Process Mapping
- 2.1.4 FMEA: Failure Modes & Effects Analysis
- 2.1.5 Theory of Constraints

2.2 Six Sigma Statistics
- 2.2.1 Basic Statistics
- 2.2.2 Descriptive Statistics
- 2.2.3 Distributions and Normality
- 2.2.4 Graphical Analysis

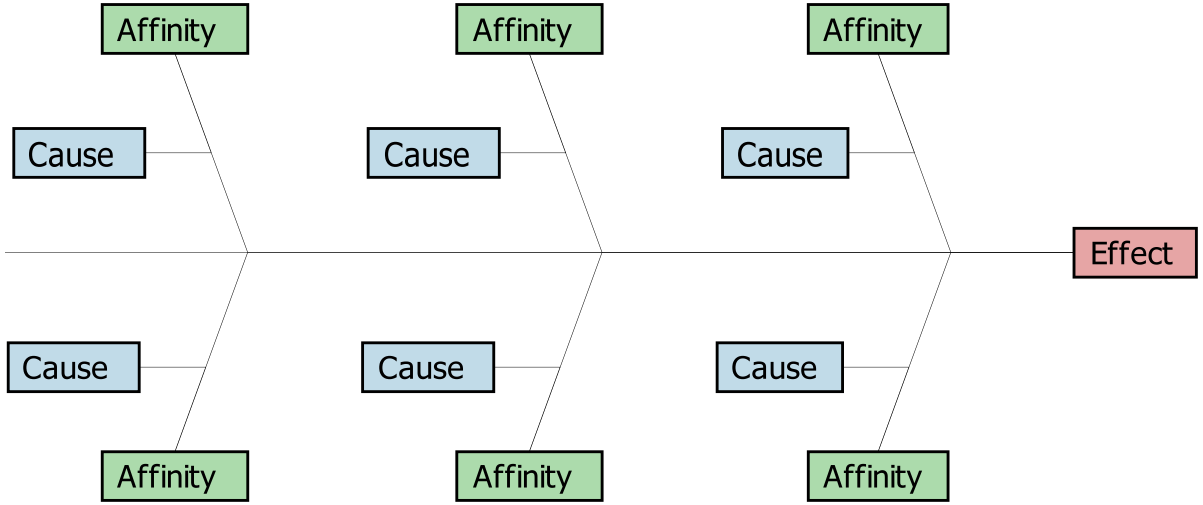
2.3 Measurement System Analysis
- 2.3.1 Precision and Accuracy
- 2.3.2 Bias, Linearity, and Stability
- 2.3.3 Gage R&R
- 2.3.4 Variable and Attribute MSA

2.4 Process Capability
- 2.4.1 Capability Analysis
- 2.4.2 Concept of Stability
- 2.4.3 Attribute and Discrete Capability
- 2.4.4 Monitoring Techniques

2.1.1 Cause and Effect Diagram

What is a Cause and Effect Diagram?

- A **cause and effect diagram** is also called a *Fishbone Diagram* or *Ishikawa Diagram*. It was created by Kaoru Ishikawa and is used to identify, organize, and display the potential causes of a specific effect or event in a graphical way similar to a fishbone.
- It illustrates the relationship between one specified event (output) and its categorized potential causes (inputs) in a visual and systematic way.



Major Categories of Potential Causes

- P4ME
  - **People**: People who are involved in the process
  - **Methods**: How the process is completed (e.g., procedures, policies, regulations, laws)
  - **Machines**: Equipment or tools needed to perform the process
  - **Materials**: Raw materials or information needed to do the job
  - **Measurements**: Data collected from the process for inspection or evaluation
  - **Environment**: Surroundings of the process (e.g., location, time, culture).

How to Plot a Cause and Effect Diagram

- Step 1: Identify and define the effect/event being analyzed.
  - Clearly state the operational definition of the effect/event of interest.
  - The event can be the positive outcome desired or negative problem targeted to solve.
  - Enter the effect/event in the end box of the Fishbone diagram and draw a spine pointed to it.

How to Plot a Cause and Effect Diagram

- Step 1

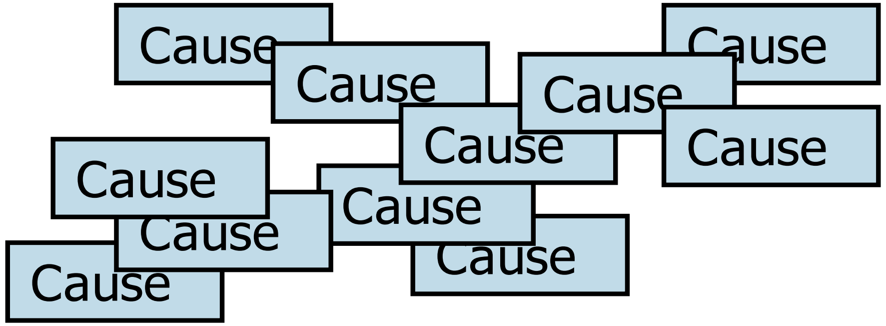
Effect or Event Being Analyzed

How to Plot a Cause and Effect Diagram

- Step 2: Brainstorm the potential causes or factors of the effect/event occurring.
  - Identify any factors with a potential impact on the effect/event and include them in this step.
  - Put all the identified potential causes aside for use later.

How to Plot a Cause and Effect Diagram

- Step 2

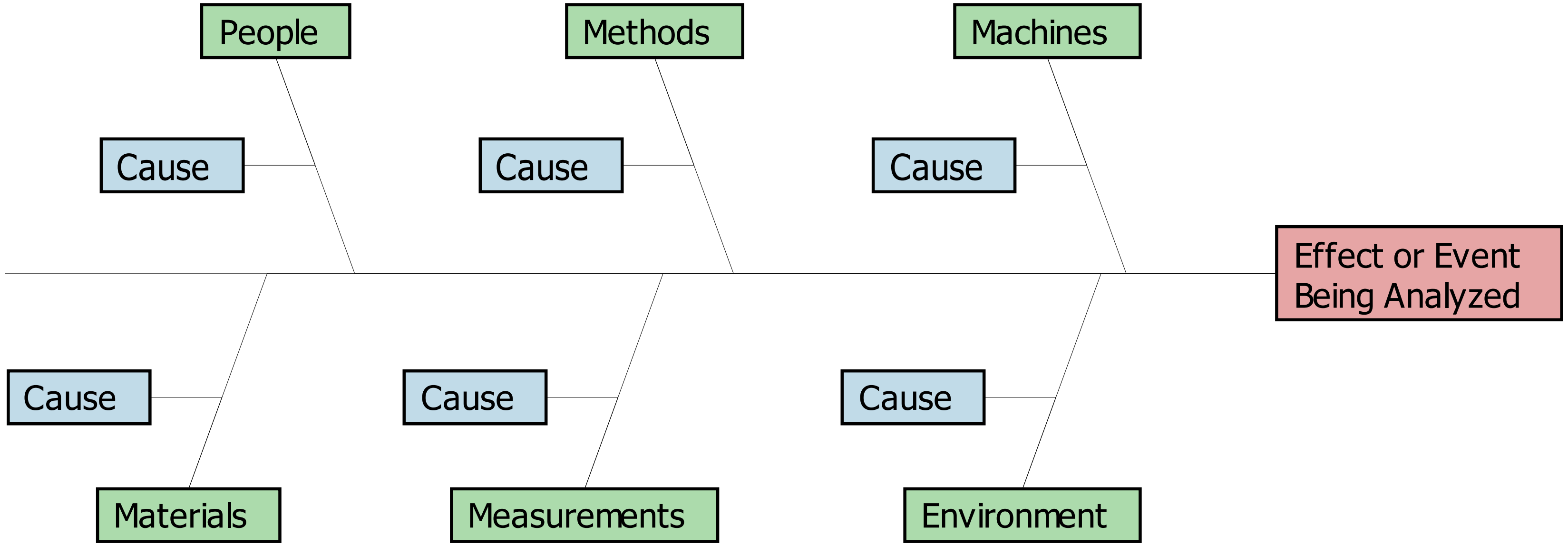


How to Plot a Cause and Effect Diagram

- Step 3: Identify the main categories of causes and group the potential causes accordingly.
  - Besides P4ME (i.e., people, methods, machines, materials, measurements, and environment), you can group potential causes into other customized categories.
  - Below each major category, you can define sub-categories and then classify them to help you visualize the potential causes.
  - Enter each cause category in a box and connect the box to the spine. Link each potential cause to its corresponding cause category.

How to Plot a Cause and Effect Diagram

- Step 3



How to Plot a Cause and Effect Diagram

- Step 4: Analyze the cause and effect diagram.
  - A cause and effect diagram includes all the possible factors of the effect/event being analyzed.
  - Use a Pareto chart to filter causes the project team needs to focus on.
  - Identify causes with high impact that the team can take action upon.
  - Determine how to measure causes and effects quantitatively. Prepare for further statistical analysis.

Benefits to Using Cause and Effect Diagram

- Helps to quickly identify and sort the potential causes of an effect.
- Provides a systematic way to brainstorm potential causes effectively and efficiently.
- Identifies areas requiring data collection for further quantitative analysis.
- Locates “low-hanging fruit.”

Limitation of Cause and Effect Diagrams

- A cause and effect diagram only provides qualitative analysis of correlation between each cause and the effect.
- One cause and effect diagram can only focus on *one* effect or event at a time.
- Further statistical analysis is required to quantify the relationship between various factors and the effect and identify the root causes.

Cause and Effect Diagram Example

- *Case study*:
  - A real estate company is interested to find the root causes of high energy costs of its properties.
  - The cause and effect diagram is used to identify, organize, and analyze the potential root causes.

Cause and Effect Diagram Example

- Step 1: Identify and define the effect/event being analyzed: high energy costs of buildings.

High Energy Cost of Buildings

Cause and Effect Diagram Example

- Step 2: Brainstorm the potential causes or factors of the high energy costs.

Bad energy consuming habits

Aged HVAC units

Poor maintenance of HVAC units

Most lights are still on during non-business hours

Inaccurate metering

Building materials

Building air leakage

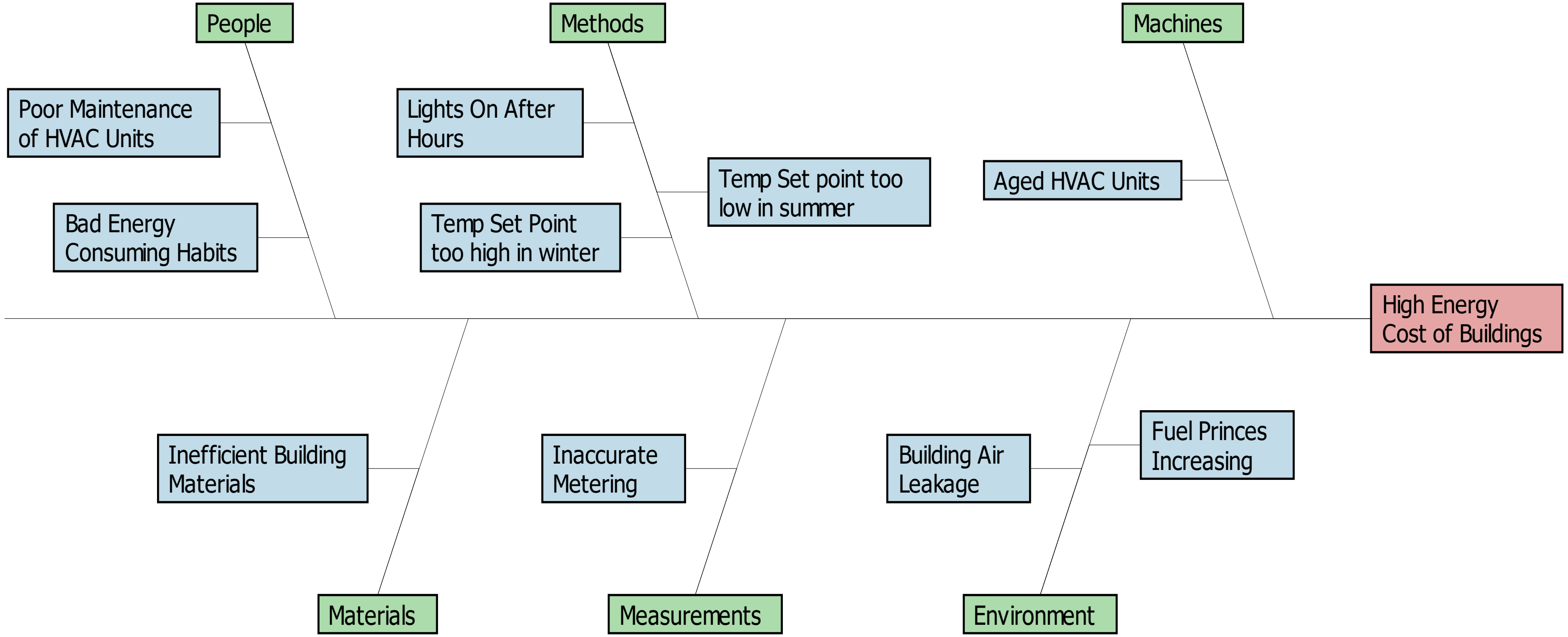
Temperature set point is too high in winter

Temperature set point is too low in summer

Fuel prices increasing

Cause and Effect Diagram Example

- Step 3: Identify the main categories of causes and group the potential causes accordingly.



Cause and Effect Diagram Example

- Step 4: Analyze the cause and effect (C&E) diagram.
  - After completing the C&E diagram, the real estate company conducts further research on each potential root cause.
  - It is discovered that:
    - The utility metering is accurate
    - The building materials are fine and there is not significant amount of air leakage from the building
    - The fuel prices increased recently but were negligible
    - Most lights are off during the non-business hours except that some lights have to be on for security purposes
    - The temperature set points in the summer and winter are both adequate and reasonable
    - The high energy costs are probably caused by the poor HVAC maintenance on aged units and the wasteful energy consuming habits.
- Next, the real estate company needs to collect and analyze the data to check whether root causes identified in the C&E diagram are statistically the causes of the high energy costs.

2.1.2 Cause and Effects Matrix

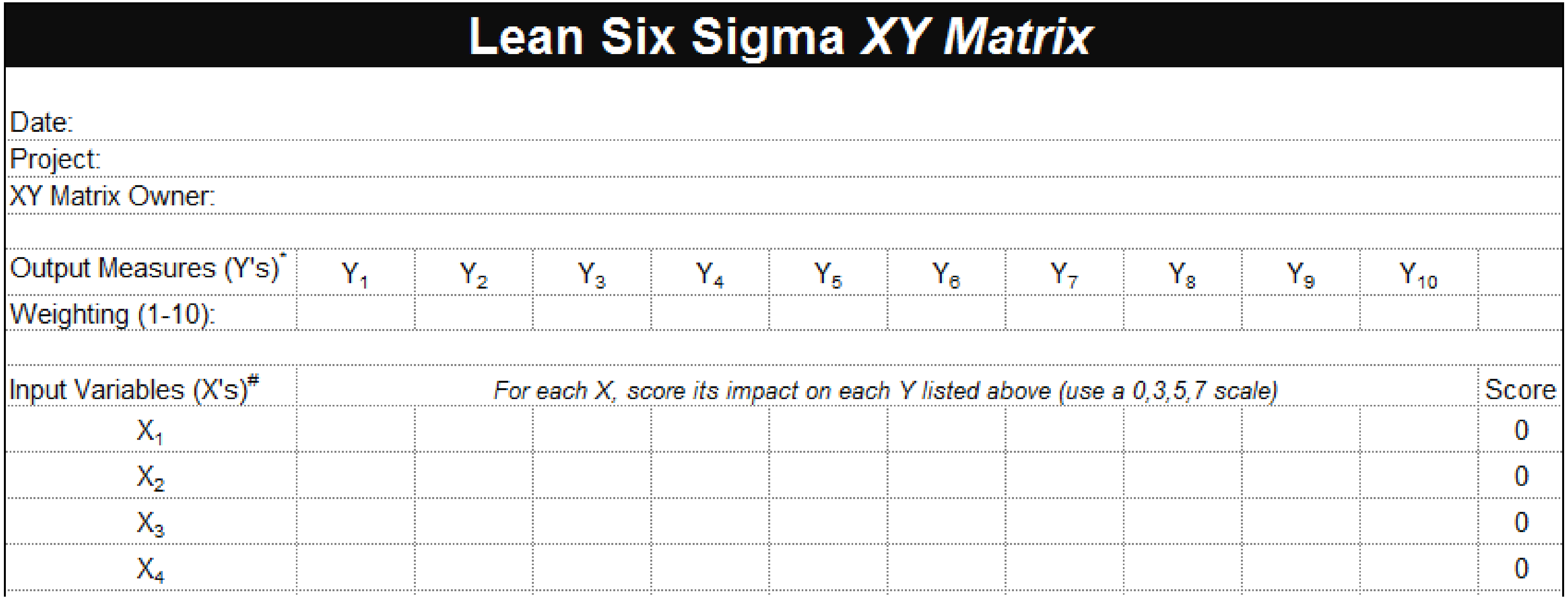
What is a Cause and Effect Matrix?

- The **cause and effect matrix** (XY Matrix) is a tool to help subjectively quantify the relationship of several X's to several Y's.

- Among the Y's under consideration, two important ones should be the *primary* and *secondary metrics* of your Six Sigma project.

- The X's should be derived from your cause and effect diagram. Let us take a peek as what it looks like on the next page.

Cause and Effects Matrix



Lean Six Sigma *XY Matrix*

Date:
Project:
XY Matrix Owner:

Output Measures (Y's)*	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	
Weighting (1-10):											
Input Variables (X's)#	*For each X, score its impact on each Y listed above (use a 0,3,5,7 scale)*										Score
X_1											0
X_2											0
X_3											0
X_4											0

How to Use a Cause and Effect Matrix

1. Across the top enter your output measures. These are the Y’s that are important to your project.

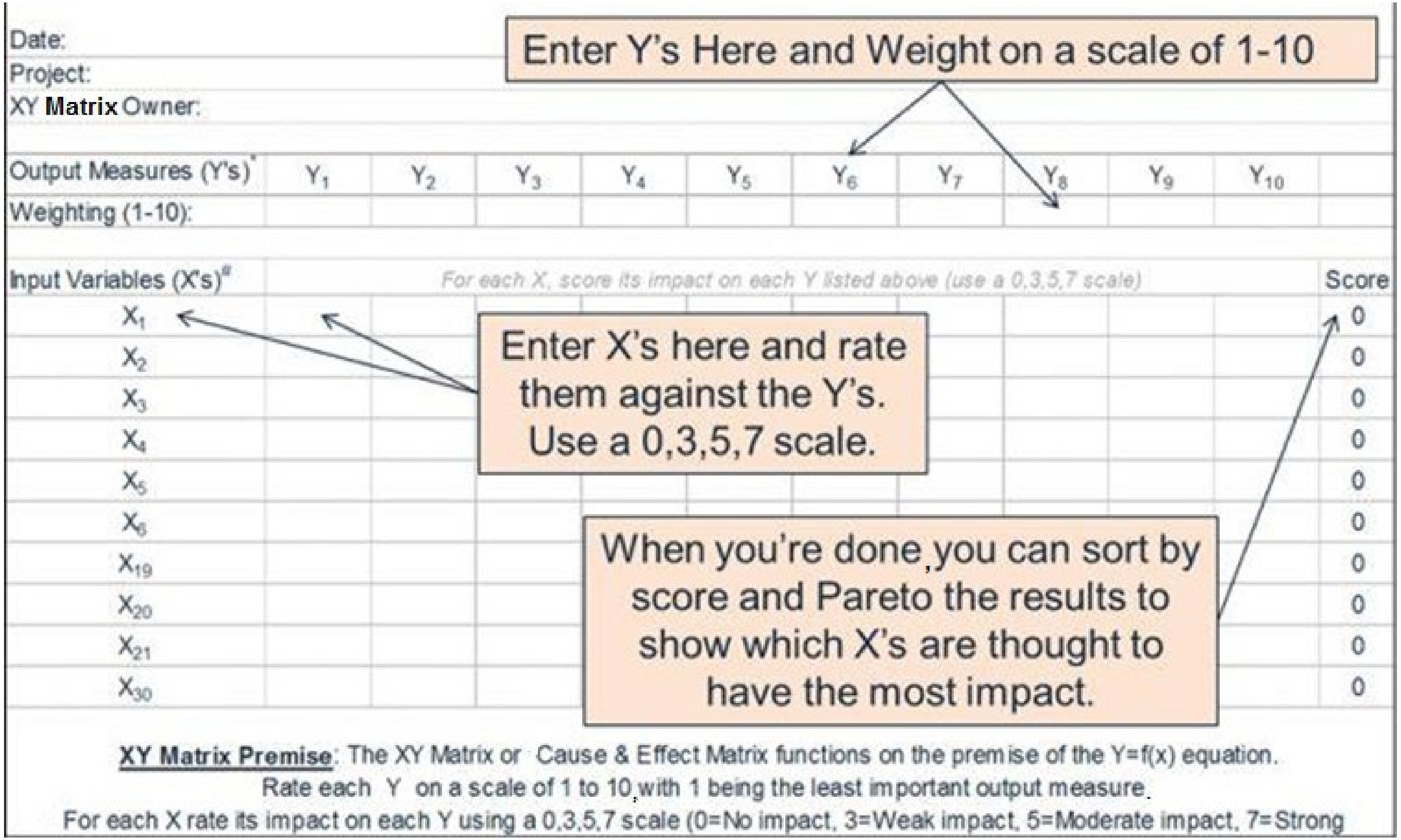
2. Next, give each Y a weight. Use a 1–10 scale, 1 being least important and 10 most important.

3. Below, in the leftmost column, enter all the variables you identified with your cause and effect diagram.

4. Within the matrix itself, rate the strength of the relationship between the X in the row and the corresponding Y in that column. Use a scale of 0, 3, 5, and 7.

5. Lastly, sort the “Score” column to order the most important X’s first.

Cause and Effect Matrix Notes



Enter Y's Here and Weight on a scale of 1-10

Date:											
Project:											
XY **Matrix** Owner:											
Output Measures (Y's)*	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	
Weighting (1-10):											
Input Variables (X's)#	*For each X, score its impact on each Y listed above (use a 0,3,5,7 scale)*										Score
X_1											0
X_2											0
X_3											0
X_4											0
X_5											0
X_6											0
X_{19}											0
X_{20}											0
X_{21}											0
X_{30}											0

Enter X's here and rate them against the Y's. Use a 0,3,5,7 scale.

When you're done, you can sort by score and Pareto the results to show which X's are thought to have the most impact.

XY Matrix Premise: The XY Matrix or Cause & Effect Matrix functions on the premise of the Y=f(x) equation.
Rate each Y on a scale of 1 to 10, with 1 being the least important output measure.
For each X rate its impact on each Y using a 0,3,5,7 scale (0=No impact, 3=Weak impact, 5=Moderate impact, 7=Strong

After You Have Completed the C&E Matrix

After you have completed your cause and effects matrix, build a strategy for validating and/or eliminating the x's as significant variables to the Y=f(x) equation.

- Build a data collection plan
- Prepare and execute planned studies
- Perform analytics
- Review results with SMEs
- etc.

2.1.3 Process Mapping

What is a Process Map?

- **Process mapping** is an important aspect to begin understanding a process. It provides a visual display of the steps required to produce a good or service.
- There are various forms of process mapping techniques that are used in the DMAIC methodology. We have already covered three common and useful maps in our “Process Definition” section of the Define phase:
  - High Level Process Map
  - Detailed Process Map
  - Functional Process Map
- Now we will explore a few other value process mapping techniques:
  - SIPOC
  - Value Stream
  - Spaghetti Diagram
  - Thought Process Map

What is SIPOC?

- A SIPOC (Suppliers-Input-Process-Output-Customers) is a high-level visualization tool to help identify and link the different components in a process.

- It is usually applied in the Measure phase in order to better understand the current state of the process and define the scope of the project.

Key Components of a SIPOC

- **S**uppliers: vendors who provide the raw material, services, and information. Customers can also be suppliers sometimes.
- **I**nput: the raw materials, information, equipment, services, people, environment involved in the process.
- **P**rocess: the high-level sequence of actions and decisions that results in the services or products delivered to the customers.
- **O**utput: the services or products delivered to the customers and any other outcomes of the process.
- **C**ustomers: the end users or recipients of the services or products.

How to Plot a SIPOC Diagram

- The first method:
  - Step 1: Create a template that can contain the information of the five key components in a clear way.
  - Step 2: Plot a high-level process map that covers five steps at maximum.
  - Step 3: Identify the outputs of the process.
  - Step 4: Identify the receipt of the process.
  - Step 5: Brainstorm the inputs required to run each process step.
  - Step 6: Identify the suppliers who provide the inputs.

How to Plot a SIPOC Diagram

- The second method:
  - Step 1: Create a template that can contain the information of the five key components in a clear way.
  - Step 2: Identify the receipt of the process.
  - Step 3: Identify the outputs of the process.
  - Step 4: Plot a high-level process map that covers five steps at maximum.
  - Step 5: Brainstorm the inputs required to run each process step.
  - Step 6: Identify the suppliers who provide the inputs.

Benefits of SIPOC Diagrams

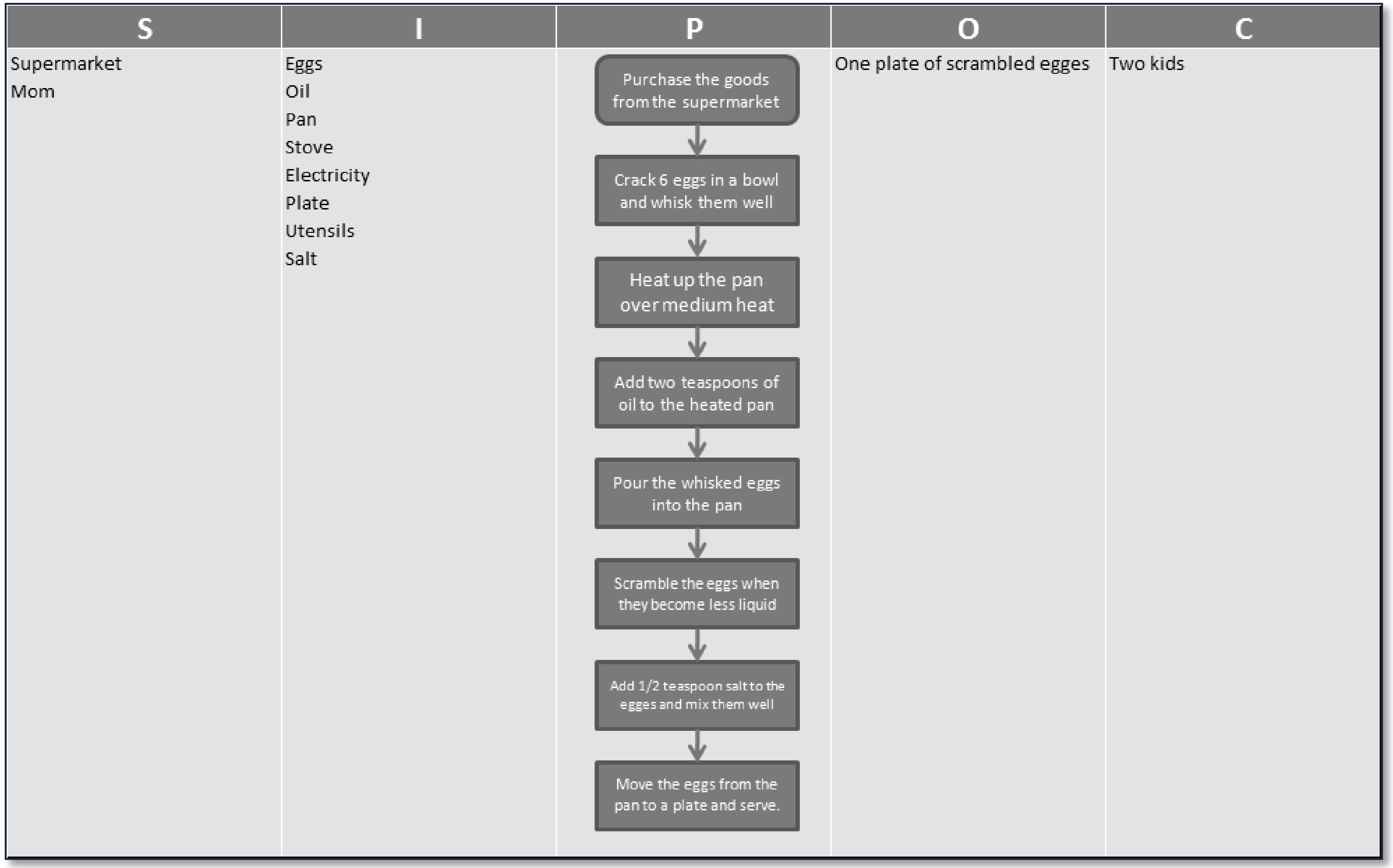
- A SIPOC diagram provides more detailed information than process maps and it demonstrates how each component gets involved in the process.

- It helps visualize and narrow the project scope.

- It serves as a great communication tool to help different process owners understand the entire process, their specific roles and responsibilities.

SIPOC Diagram Example

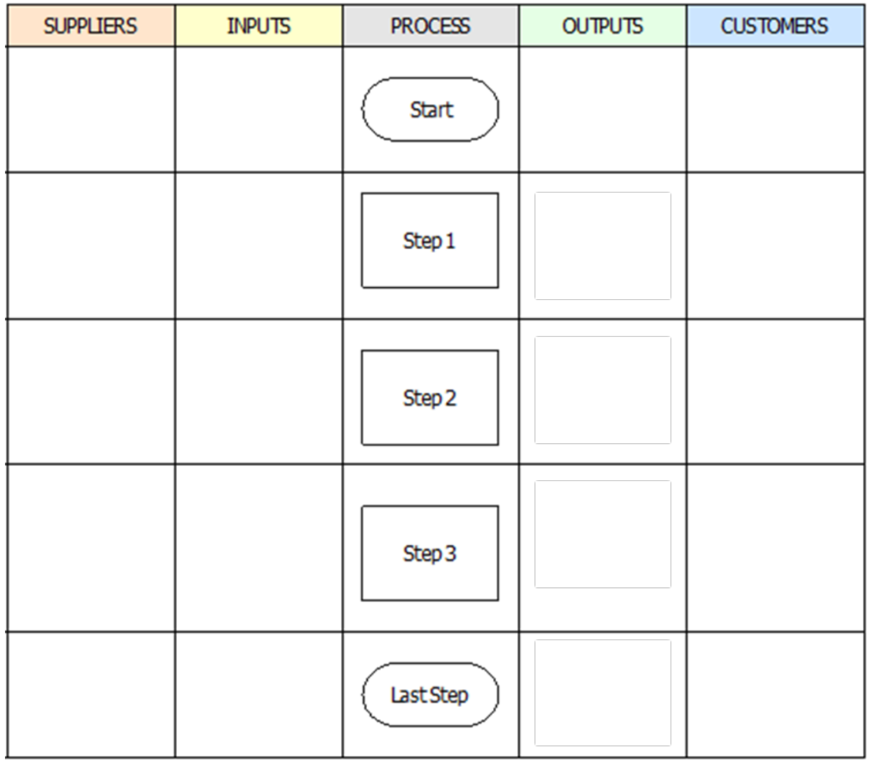
- Example of plotting a SIPOC diagram for Mom cooking scrambled eggs for two kids



S	I	P	O	C
Supermarket Mom	Eggs Oil Pan Stove Electricity Plate Utensils Salt	Purchase the goods from the supermarket ↓ Crack 6 eggs in a bowl and whisk them well ↓ Heat up the pan over medium heat ↓ Add two teaspoons of oil to the heated pan ↓ Pour the whisked eggs into the pan ↓ Scramble the eggs when they become less liquid ↓ Add 1/2 teaspoon salt to the egges and mix them well ↓ Move the eggs from the pan to a plate and serve.	One plate of scrambled egges	Two kids

Creating a SIPOC

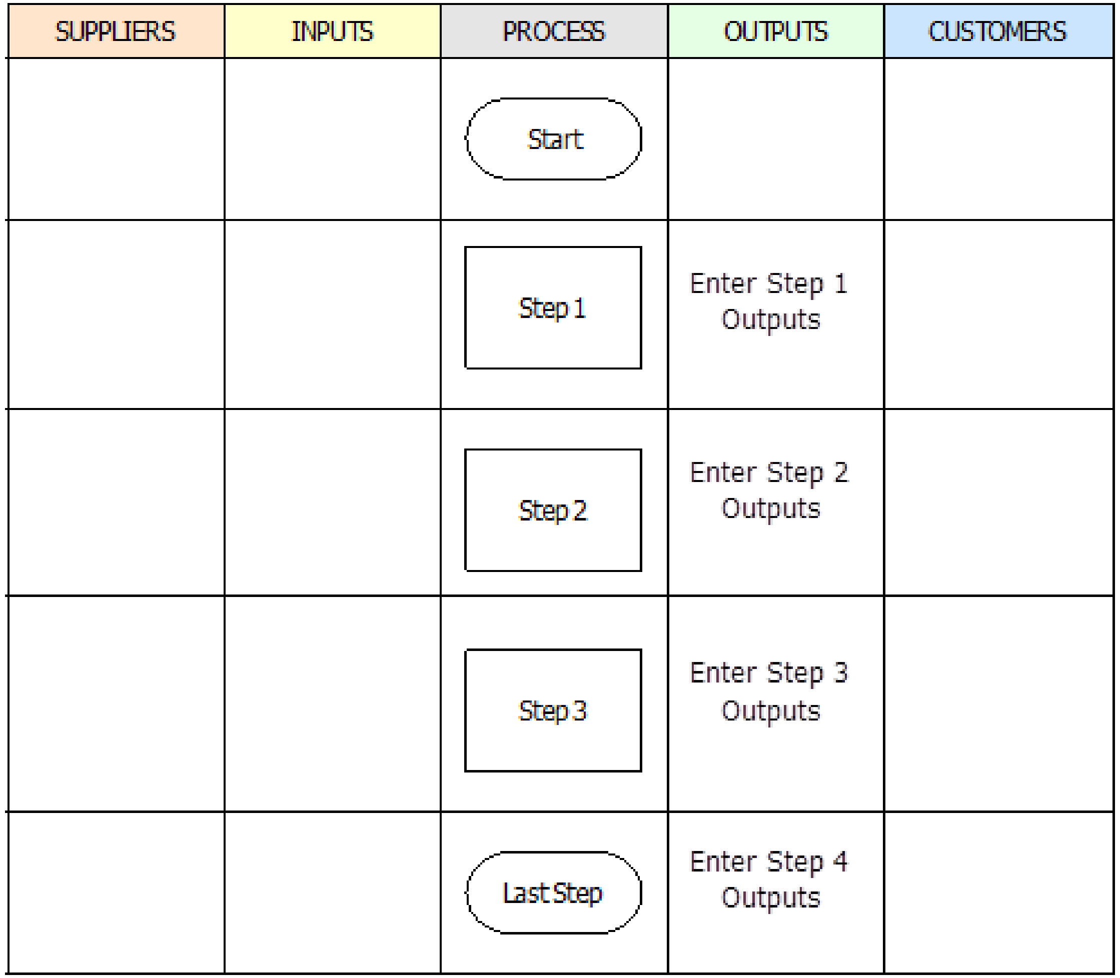
- Step 1: Vertically List High-Level Process
  - If you followed the general rules for a high-level process map, then you should have no more than 4–6 steps for your process.
  - List those steps in a vertical manner as depicted below.



SUPPLIERS	INPUTS	PROCESS	OUTPUTS	CUSTOMERS
		Start		
		Step 1		
		Step 2		
		Step 3		
		Last Step		

Creating a SIPOC

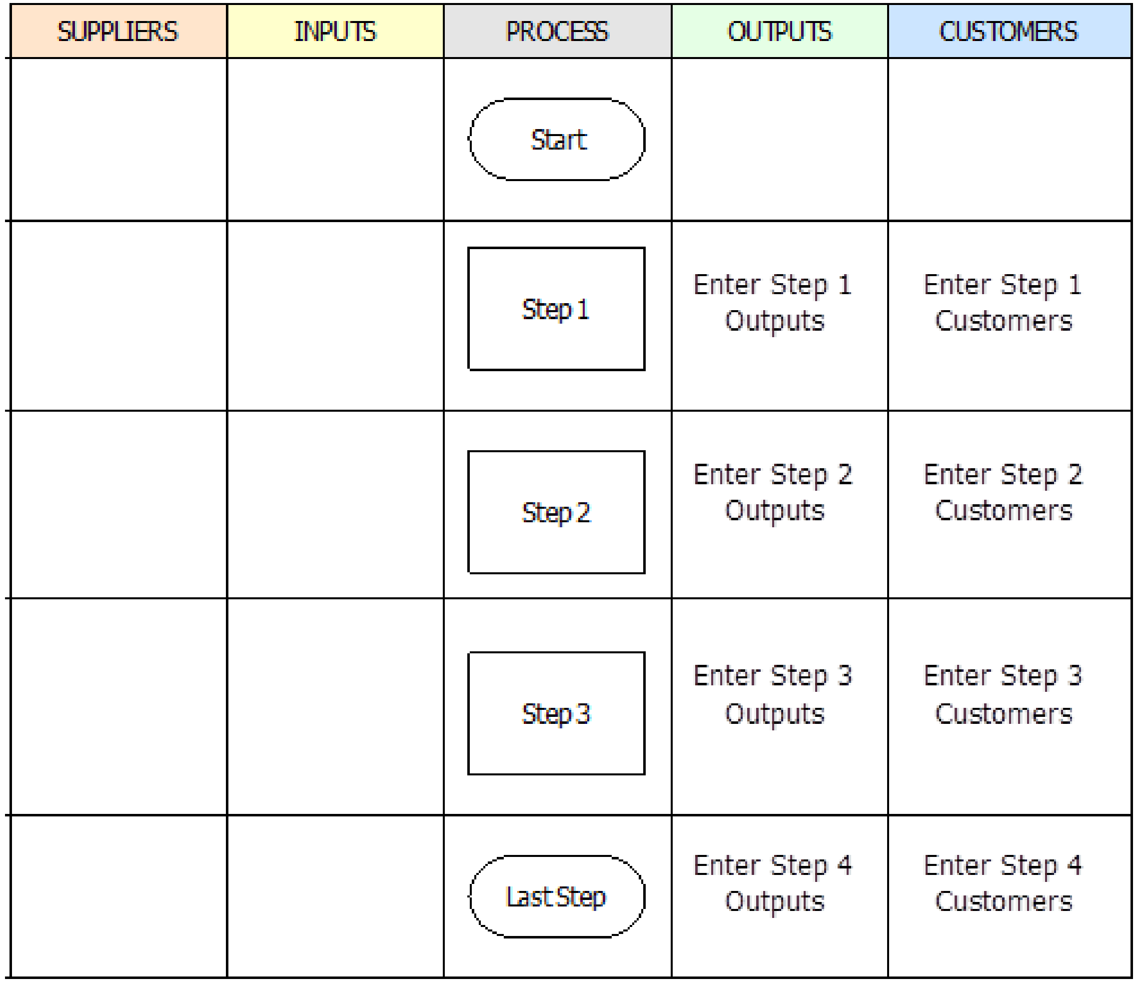
- Step 2: List Process Outputs



SUPPLIERS	INPUTS	PROCESS	OUTPUTS	CUSTOMERS
		Start		
		Step 1	Enter Step 1 Outputs	
		Step 2	Enter Step 2 Outputs	
		Step 3	Enter Step 3 Outputs	
		Last Step	Enter Step 4 Outputs	

Creating a SIPOC

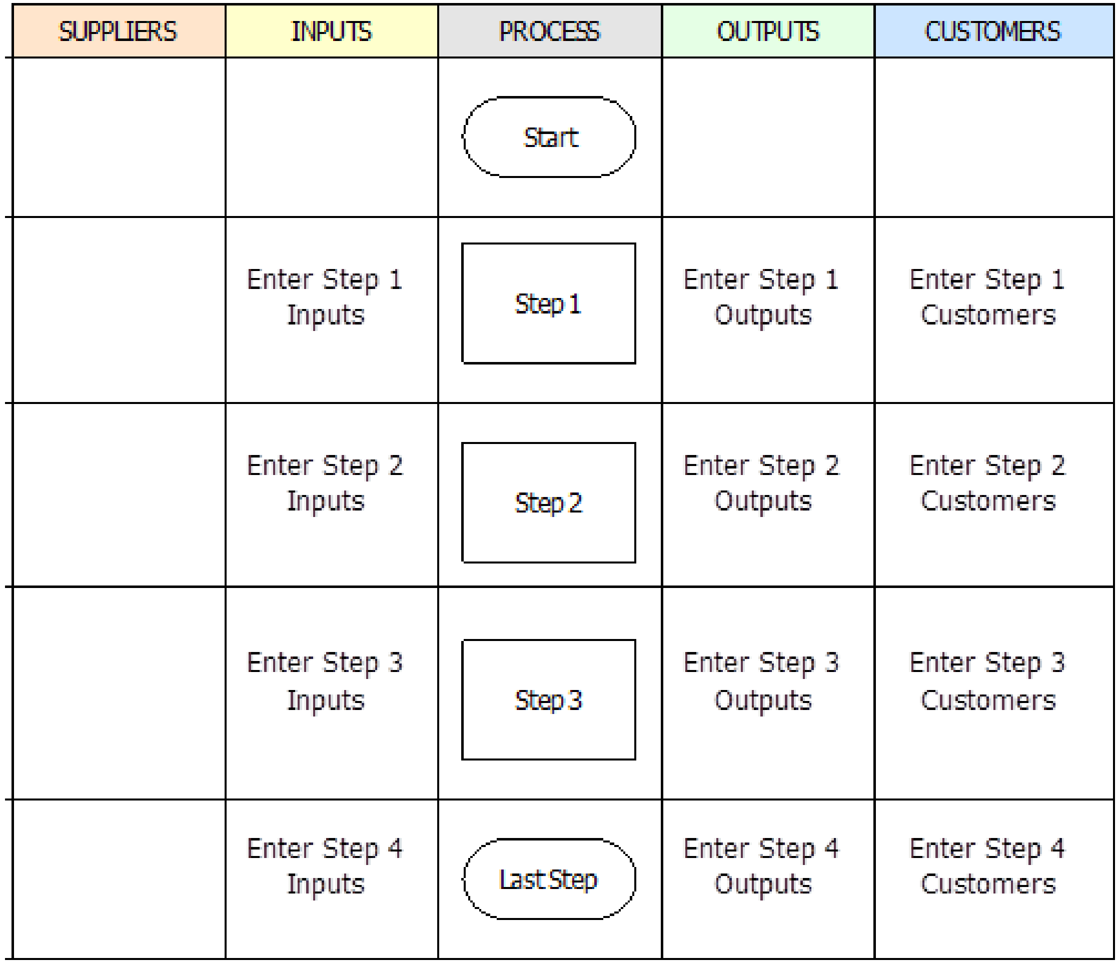
- Step 3: List Output Customers



SUPPLIERS	INPUTS	PROCESS	OUTPUTS	CUSTOMERS
		Start		
		Step 1	Enter Step 1 Outputs	Enter Step 1 Customers
		Step 2	Enter Step 2 Outputs	Enter Step 2 Customers
		Step 3	Enter Step 3 Outputs	Enter Step 3 Customers
		Last Step	Enter Step 4 Outputs	Enter Step 4 Customers

Creating a SIPOC

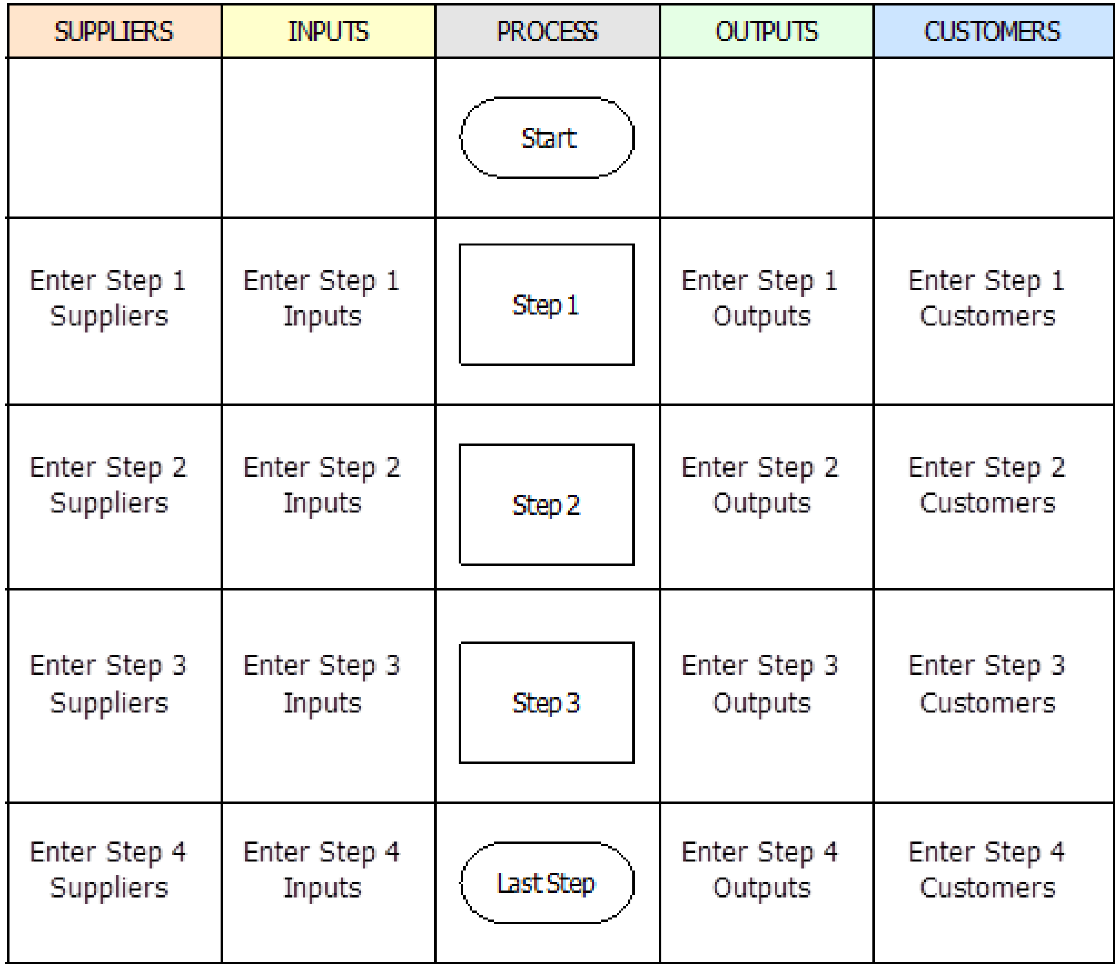
- Step 4: List Process Inputs



SUPPLIERS	INPUTS	PROCESS	OUTPUTS	CUSTOMERS
		Start		
	Enter Step 1 Inputs	Step 1	Enter Step 1 Outputs	Enter Step 1 Customers
	Enter Step 2 Inputs	Step 2	Enter Step 2 Outputs	Enter Step 2 Customers
	Enter Step 3 Inputs	Step 3	Enter Step 3 Outputs	Enter Step 3 Customers
	Enter Step 4 Inputs	Last Step	Enter Step 4 Outputs	Enter Step 4 Customers

Creating a SIPOC

- Step 5: List Suppliers of Inputs



SUPPLIERS	INPUTS	PROCESS	OUTPUTS	CUSTOMERS
		Start		
Enter Step 1 Suppliers	Enter Step 1 Inputs	Step 1	Enter Step 1 Outputs	Enter Step 1 Customers
Enter Step 2 Suppliers	Enter Step 2 Inputs	Step 2	Enter Step 2 Outputs	Enter Step 2 Customers
Enter Step 3 Suppliers	Enter Step 3 Inputs	Step 3	Enter Step 3 Outputs	Enter Step 3 Customers
Enter Step 4 Suppliers	Enter Step 4 Inputs	Last Step	Enter Step 4 Outputs	Enter Step 4 Customers

SIPOC Benefits

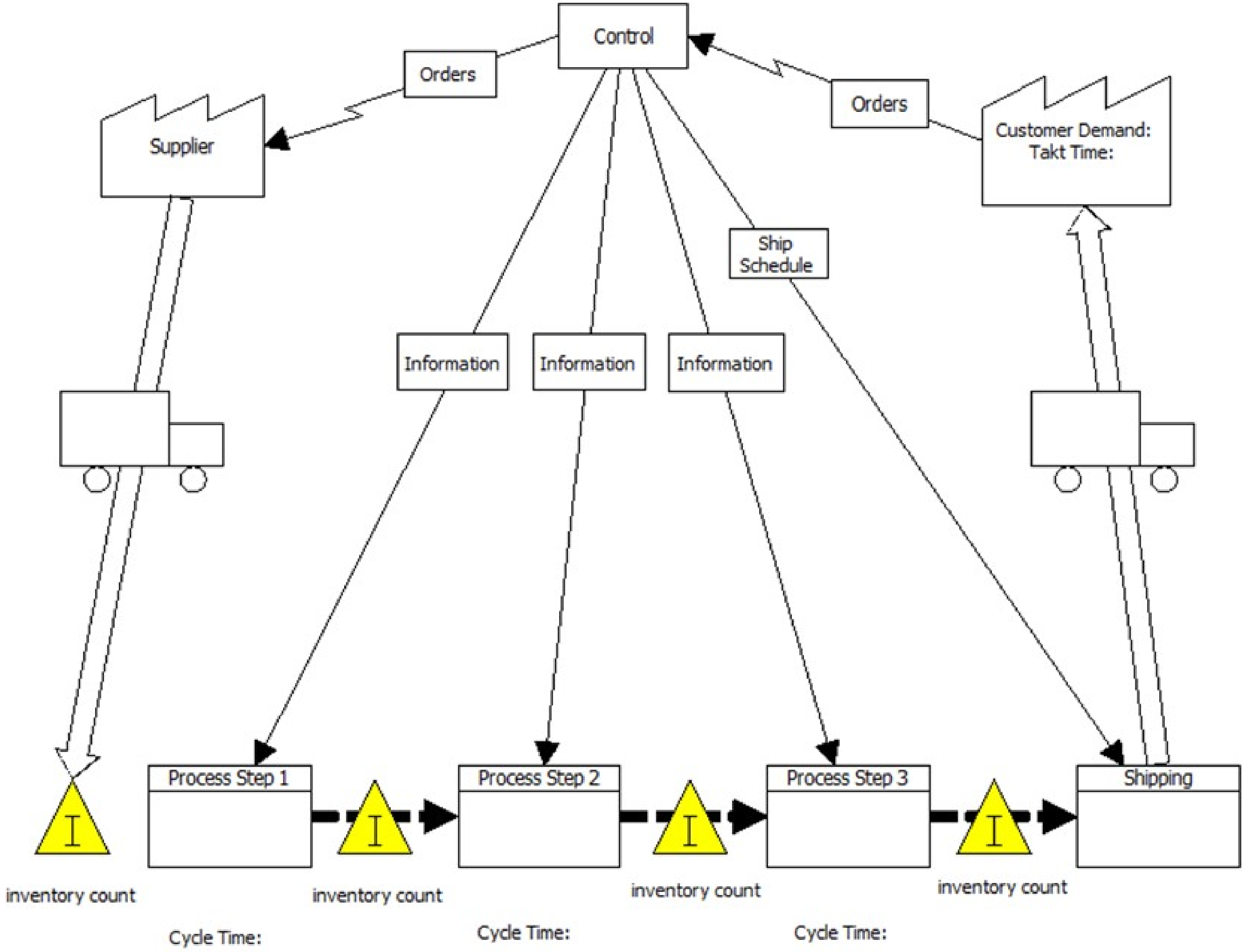
- Visually communicate project scope
- Identify key inputs and outputs of a process
- Identify key suppliers and customers of a process
- Verify:
  - Inputs match outputs for upstream processes
  - Outputs match inputs for downstream processes.
- This type of mapping is effective for identifying opportunities for improvement of your process.

- If you have completed your high-level process map, follow the outlined steps to create a process map of **S**uppliers, **I**nputs, **P**rocess, **O**utputs, and **C**ustomer.

What is Value Stream Mapping?

- **Value stream mapping** is a method to visualize and analyze the path of how information and raw materials are transformed into products or services customers receive.

- It is used to identify, measure, and decrease the non-value-adding steps in the current process.



Non-Value-Added Activities

- **Non-value-adding activities** are activities in a process that do not add any other value to the products or services customers demand.

- Example of non-value-adding activities:
  - Rework
  - Overproduction
  - Excess transportation
  - Excess stock
  - Waiting
  - Unnecessary motion.

- Not all non-value-adding activities are unnecessary.

How to Plot a Value Stream Map

- Plot the entire high-level process flow from when the customer places the order to when the customer receives the products or services in the end.
- A value stream map requires more detailed information for each step than the standard process map.
  - Cycle time
  - Preparation time
  - Actual working time
  - Available time
  - Scrap rate
  - Rework rate
  - Number of operators
- Assess the value stream map of current process, identity and eliminate the waste.

Spaghetti Chart

- A **spaghetti chart** is a graphical tool to map out the physical flow of materials, information, and people involved in a process. It can also reflect the distances between multiple workstations the physical flow has been through.

- A process that has not been streamlined has messy and wasteful movements of materials, information, and people, resembling a bowl of cooked spaghetti.

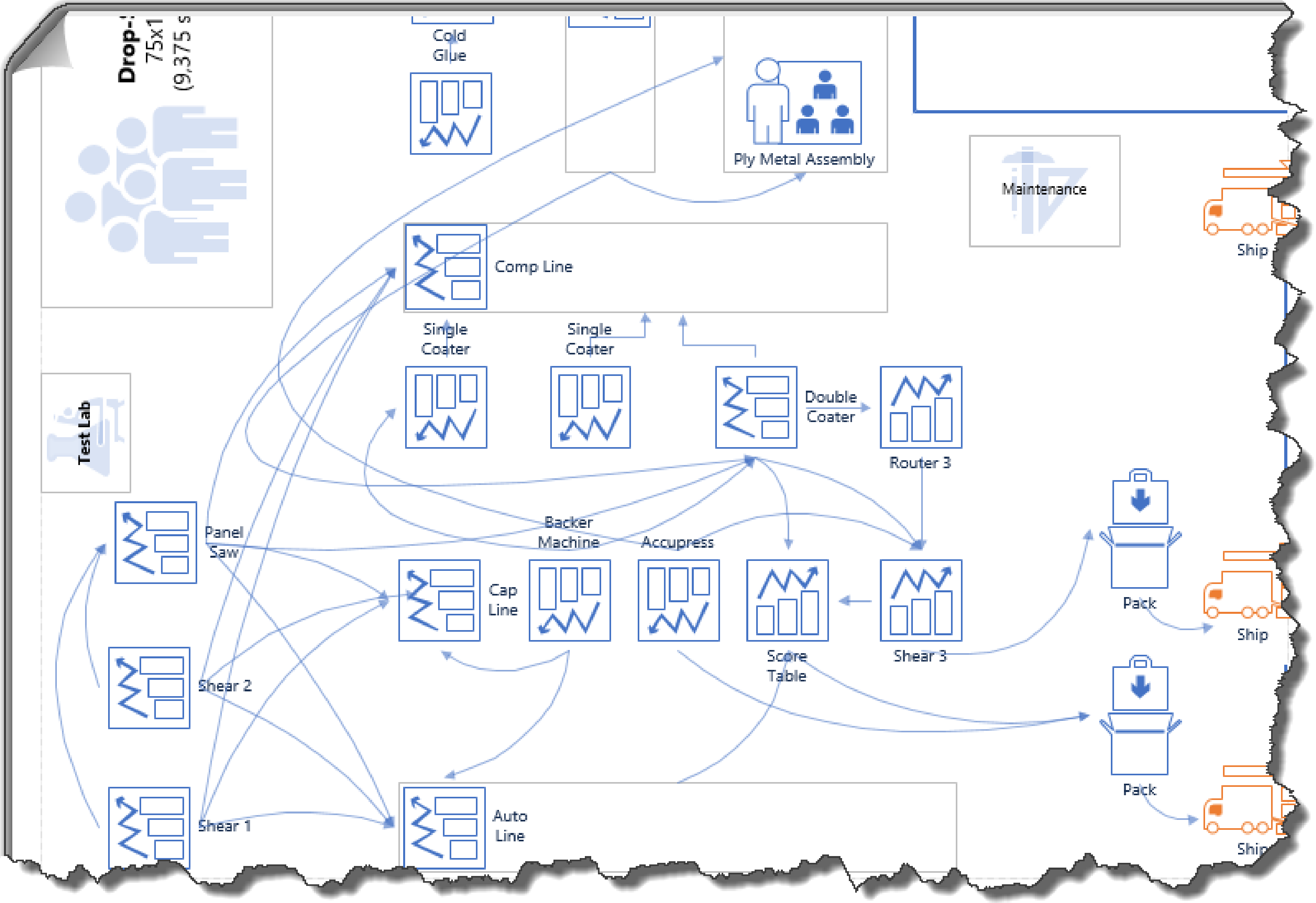
How to Plot a Spaghetti Chart

- Step 1: Create a map of the work area layout.

- Step 2: Observe the current work flow and draw the actual work path from the very beginning of work to the end when products exit the work area.

- Step 3: Analyze the spaghetti chart and identify improvement opportunities.

Spaghetti Chart Example



Thought Process Mapping

- A **thought process map** is a graphical tool to help brainstorm, organize, and visualize the information, ideas, questions, or thoughts regarding reaching the project goal.

- It's a popular tool often used at the beginning of a project in order to:
    - identify knowns and unknowns
    - communicate assumptions and risks
    - discover potential problems and solutions
    - identify resources, information, and actions required to meet the goal
    - present relationship of thoughts.

How to Plot a Thought Process Map

- Step 1: Define the project goal.

- Step 2: Brainstorm knowns and unknowns about the project.

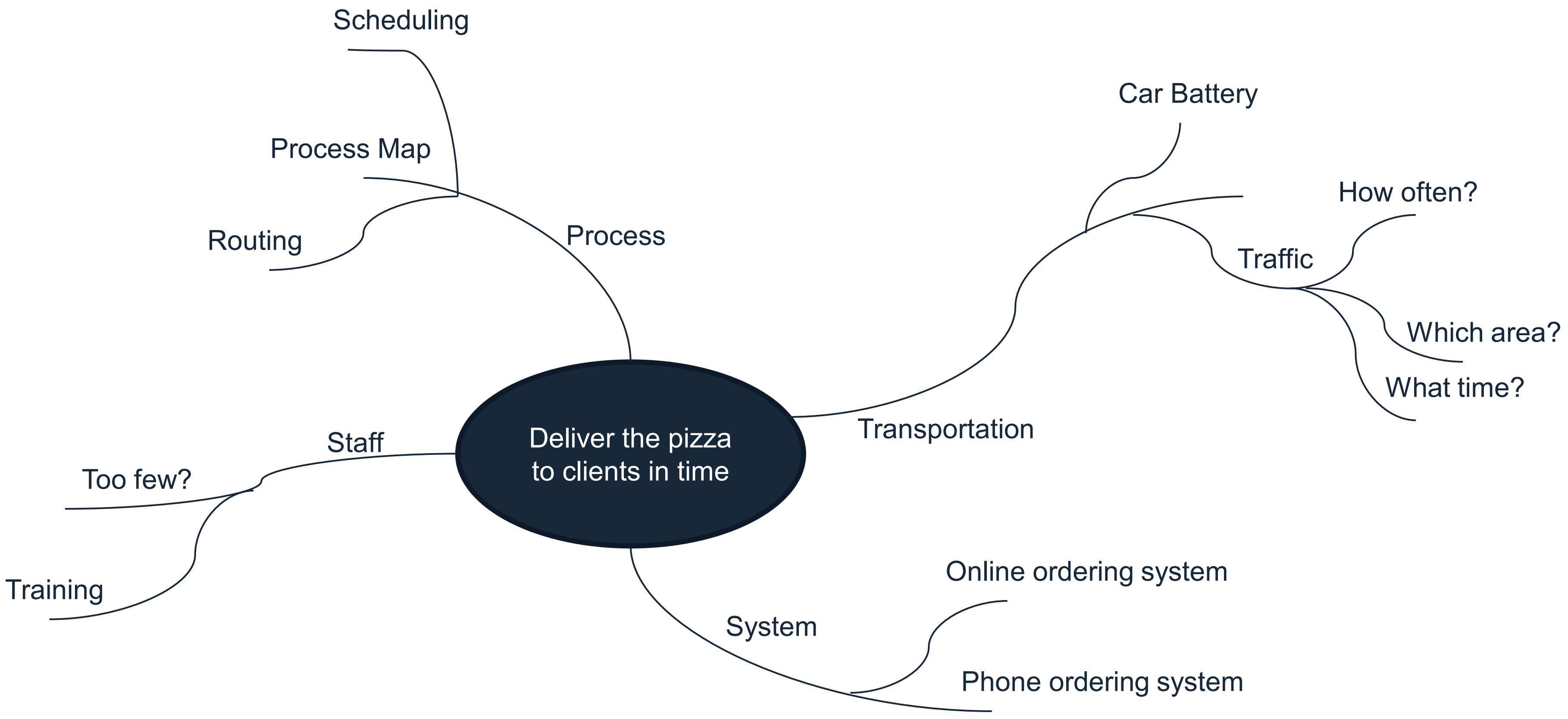
- Step 3: Brainstorm questions and group the unknowns and questions into five phases (**D**efine, **M**easure, **A**nalyze, **I**mprove, and **C**ontrol).

- Step 4: Sequence the questions below the project goal and link the related questions.

- Step 5: Identify tools or methods that would be used to answer the questions.

- Step 6: Repeat steps 3 to 5 as the project continues.

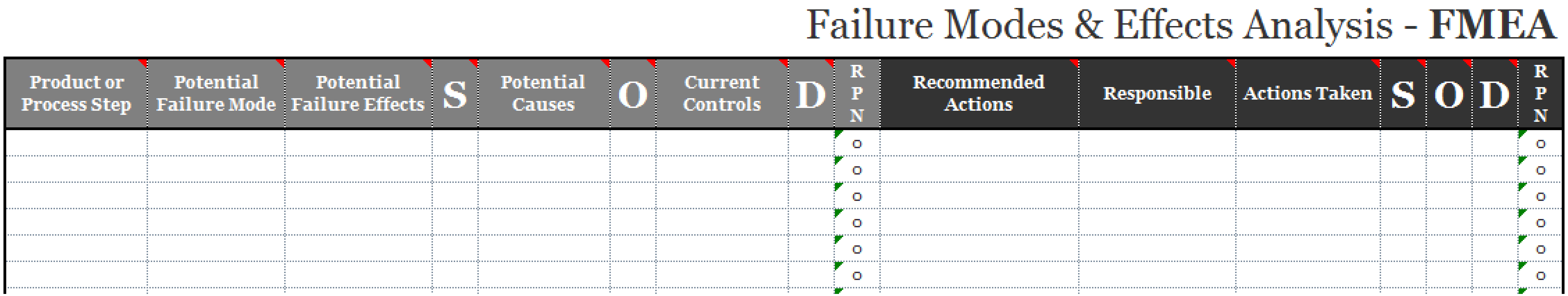
Thought Process Map Example



2.1.4 FMEA

What is FMEA?

- The FMEA (Failure Modes and Effects Analysis) is an analysis technique to identify, evaluate, and prioritize a potential deficiency in a process so that the project team can design action plans to reduce the probability of the failure/deficiency occurring.



Failure Modes & Effects Analysis - FMEA

Product or Process Step	Potential Failure Mode	Potential Failure Effects	S	Potential Causes	O	Current Controls	D	RPN	Recommended Actions	Responsible	Actions Taken	S	O	D	RPN
								0							0
								0							0
								0							0
								0							0
								0							0
								0							0

- FMEA is completed in cross-functional brainstorming sessions in which attendees have a good understanding of the entire process or of a segment of it.

Basic FMEA Terms

- **Process Functions**
  - Process steps depicted in the process map. FMEA is based on a process map and one step/function is analyzed at a time.
- **Failure Modes**
  - Potential and actual failure in the process function/step. It usually describes the way in which failure occurs. There might be more than one failure mode for one process function.
- **Failure Effects**
  - Impact of failure modes on the process or product. One failure mode might trigger multiple failure effects.
- **Failure Causes**
  - Potential defect of the design that might result in the failure modes occurring. One failure mode might have multiple potential failure causes.
- **Current Controls**
  - Procedures currently conducted to prevent failure modes from happening or to detect the failure mode occurring.

Basic FMEA Terms

- **Severity Score**
  - The seriousness of the consequences of a failure mode occurring.
  - Ranges from 1 to 10, with 10 indicating the most severe consequence.
- **Occurrence Score**
  - The frequency of the failure mode occurring.
  - Ranges from 1 to 10, with 10 indicating the highest frequency.
- **Detection Score**
  - How easily failure modes can be detected.
  - Ranges from 1 to 10, with 10 indicating the most difficult detection.

Basic FMEA Terms

- **RPN (Risk Prioritization Number)**
  - The product of the severity, occurrence, and detection scores.
  - Ranges from 1 to 1000.
  - The higher RPN is, the more focus the particular step/function needs.

- **Recommended Actions**
  - The action plan recommended to reduce the probability of failure modes occurring.

How to Conduct an FMEA

- Step 1: List the critical functions of the process based on the process map created.
- Step 2: List all potential failure modes that might occur in each function. One function may have multiple potential failures.
- Step 3: List all potential failure effects that might affect the process or product.
- Step 4: List all possible causes that may lead to the failure mode happening.
- Step 5: List the current control procedures for each failure mode.

How to Conduct an FMEA

- Step 6: Determine the severity rating for each potential failure mode.
- Step 7: Determine the occurrence rating for each potential failure cause.
- Step 8: Determine the detection rating for each current control procedure.
- Step 9: Calculate RPN (Risk Prioritization Number).
- Step 10: Rank the failures using RPN and determine the precedence of problems or critical inputs of the process. A Pareto chart might help to focus on the failure modes with high RPNs. The higher the RPN, the higher the priority the correction action plan.

How to Conduct an FMEA

- Step 11: Brainstorm and create recommended action plans for each failure mode.
- Step 12: Determine and assign the task owner and projected completion date to take actions.
- Step 13: Determine the new severity rating if the actions are taken.
- Step 14: Determine the new occurrence rating if the actions are taken.
- Step 15: Determine the new detection rating if the actions are taken.
- Step 16: Update the RPN based on new severity, occurrence, and detection ratings.

FMEA Example

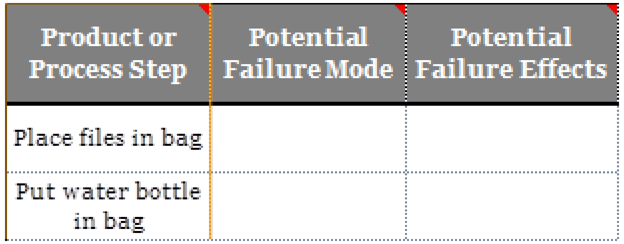
Case study:

- Joe is trying to identify, analyze, and eliminate the failure modes he experienced in the past when preparing his work bag before heading to the office every morning. He decides to run an FMEA for his process of work bag preparation.

- There are only two steps involved in the process.
  - Putting the work files in the bag
  - Putting a water bottle in the bag.

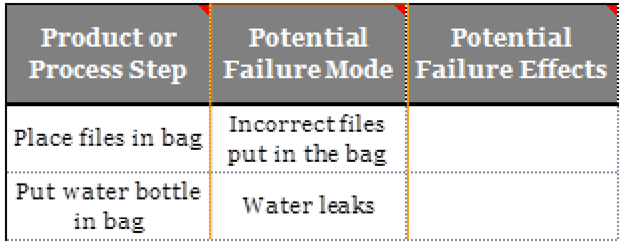
FMEA Example

- Step 1: List the critical functions of the process based on the process map created.



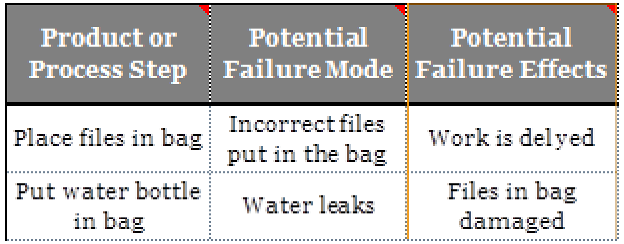
Product or Process Step	Potential Failure Mode	Potential Failure Effects
Place files in bag		
Put water bottle in bag		

- Step 2: List all the potential failure modes for each function.



Product or Process Step	Potential Failure Mode	Potential Failure Effects
Place files in bag	Incorrect files put in the bag	
Put water bottle in bag	Water leaks	

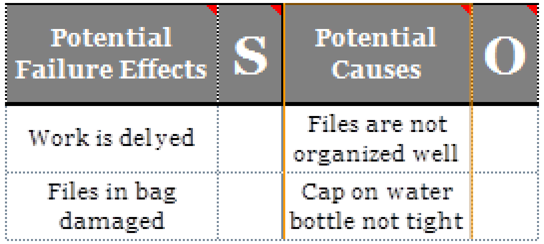
- Step 3: List potential failure effects that might affect the process.



Product or Process Step	Potential Failure Mode	Potential Failure Effects
Place files in bag	Incorrect files put in the bag	Work is delyed
Put water bottle in bag	Water leaks	Files in bag damaged

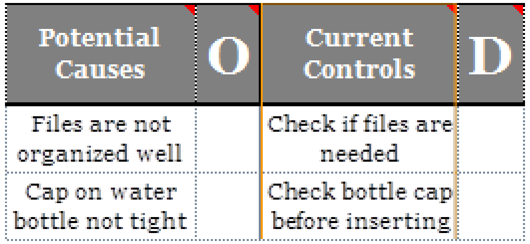
FMEA Example

- Step 4: List all possible causes to the failure mode.



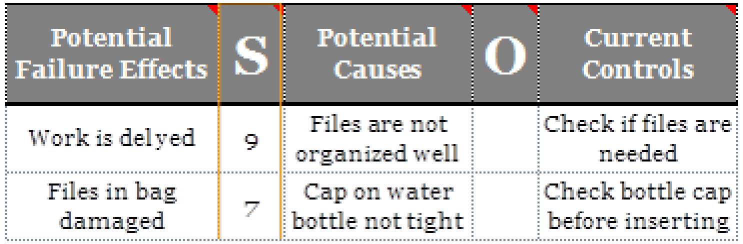
Potential Failure Effects	S	Potential Causes	O
Work is delyed		Files are not organized well	
Files in bag damaged		Cap on water bottle not tight	

- Step 5: List any control procedures for each failure mode.



Potential Causes	O	Current Controls	D
Files are not organized well		Check if files are needed	
Cap on water bottle not tight		Check bottle cap before inserting	

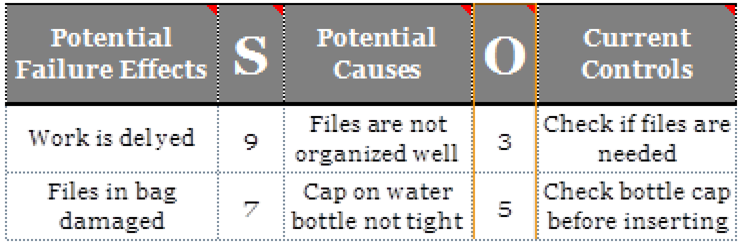
- Step 6: Determine the severity rating for each failure mode.



Potential Failure Effects	S	Potential Causes	O	Current Controls
Work is delyed	9	Files are not organized well		Check if files are needed
Files in bag damaged	7	Cap on water bottle not tight		Check bottle cap before inserting

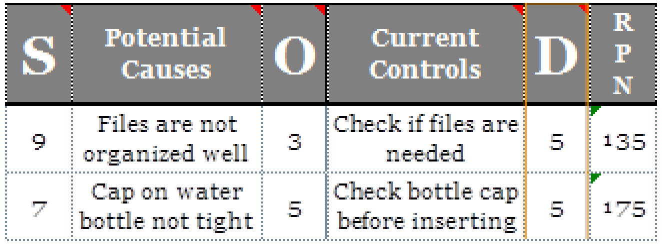
FMEA Example

- Step 7: Determine the occurrence rating for each failure cause.



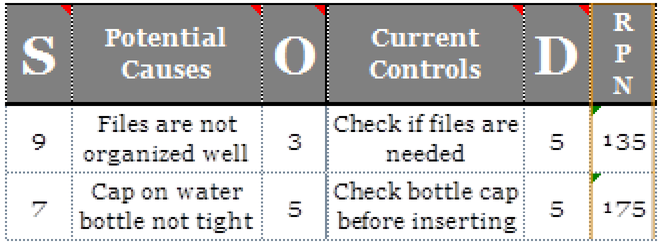
Potential Failure Effects	S	Potential Causes	O	Current Controls
Work is delyed	9	Files are not organized well	3	Check if files are needed
Files in bag damaged	7	Cap on water bottle not tight	5	Check bottle cap before inserting

- Step 8: Determine the detection rating for each control.



S	Potential Causes	O	Current Controls	D	RPN
9	Files are not organized well	3	Check if files are needed	5	135
7	Cap on water bottle not tight	5	Check bottle cap before inserting	5	175

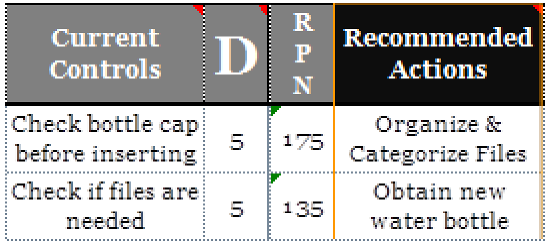
- Step 9: Calculate the RPN (Risk Prioritization Number).



S	Potential Causes	O	Current Controls	D	RPN
9	Files are not organized well	3	Check if files are needed	5	135
7	Cap on water bottle not tight	5	Check bottle cap before inserting	5	175

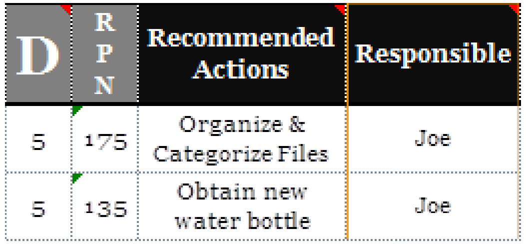
FMEA Example

- Step 10: Rank the failures using the RPN and determine the precedence of problems or critical inputs of process.



Current Controls	D	RPN	Recommended Actions
Check bottle cap before inserting	5	175	Organize & Categorize Files
Check if files are needed	5	135	Obtain new water bottle

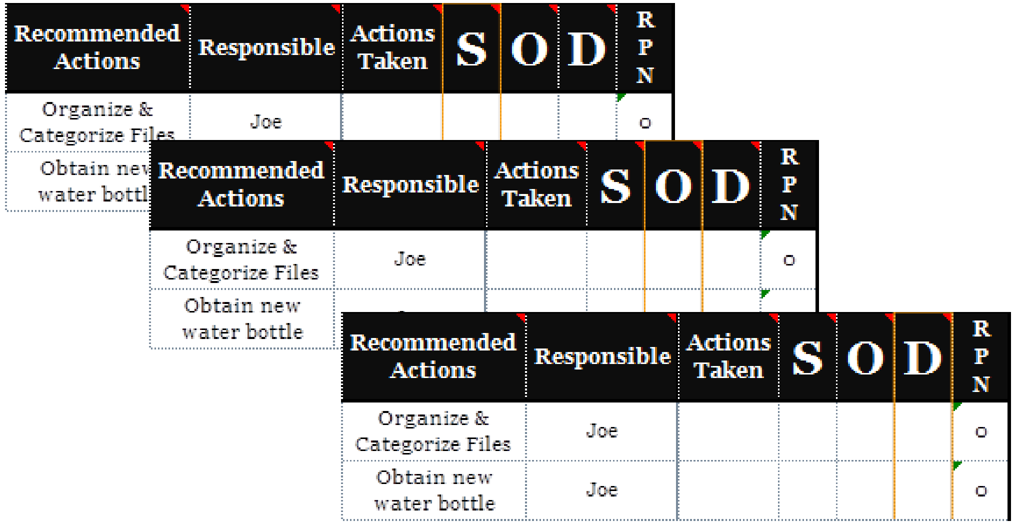
- Step 11: Brainstorm and create recommended action plans.
- Step 12: Determine and assign owners with completion dates.



D	RPN	Recommended Actions	Responsible
5	175	Organize & Categorize Files	Joe
5	135	Obtain new water bottle	Joe

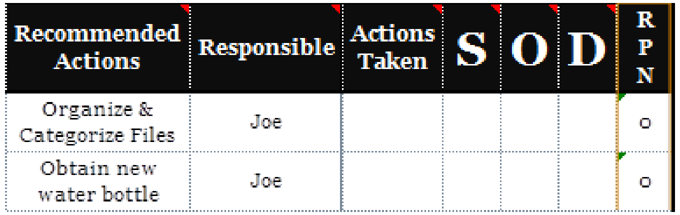
FMEA Example

- Steps 13-15: Determine new severity, occurrence and detection ratings if actions are taken.



Recommended Actions	Responsible	Actions Taken	S	O	D	RPN
Organize & Categorize Files	Joe					0
Obtain new water bottle	Joe					0

- Step 16: Update RPN based on new ratings.



Recommended Actions	Responsible	Actions Taken	S	O	D	RPN
Organize & Categorize Files	Joe					0
Obtain new water bottle	Joe					0

2.1.5 Theory of Constraints

What is the Theory of Constraints?

- Processes, systems, and organizations are vulnerable to their weakest part.
- Any manageable system is limited by constraints in its ability to produce more (and there is always at least one constraint).

Performance Measures

Making sound financial decisions based on these three measures is a critical requirement.

- **Throughput** – rate at which a system generates money through sales.
- **Operational Expense** – money spent by the system to turn inventory into throughput.
- **Inventory** – money the system has invested in purchasing things it intends to sell.

Five Focusing Steps

Objective: To ensure ongoing improvement efforts are focused on the constraints of a system.

1. Identify the system's constraint(s).
2. Decide how to exploit the constraint(s).
3. Subordinate everything else to the decision in step 2.
4. Elevate the constraint(s).
5. If in previous steps a constraint has been broken, return to step 1, but do not allow inertia to cause a system's constraint.

Logical Thinking Processes

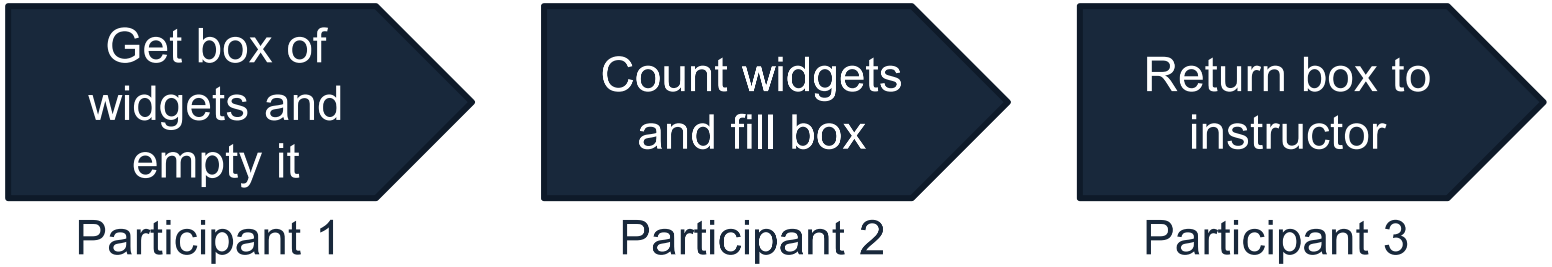
	Focusing Step	Thinking Process	Tools
1	Identify the system’s constraint(s)	• Identify the problems • Find the root causes	Cause and effect diagram
2	Decide how to exploit the constraint(s)	• Develop a solution	Future reality tree
3	Subordinate everything else to the decision in step 2	• Identify the conflict preventing the solution • Remove the conflict	Evaporating cloud
4	Elevate the constraint	• Construct and execute an implementation plan	Prerequisite tree Transition tree
5	If in previous steps a constraint has been broken, return to step 1, but do not allow inertia to cause a system’s constraint		

Simulation Exercise

Resources needed:

- 3 “production line” participants
- 1 timer per each “production line” participant
- 5 small boxes of 15 widgets each (paperclips, pens/pencils, candy, etc.)

Widget Value Chain:



Get box of widgets and empty it

Participant 1

Count widgets and fill box

Participant 2

Return box to instructor

Participant 3

Simulation Exercise

Widget Value Chain:

Get box of widgets and empty it

Participant 1

Count widgets and fill box

Participant 2

Return box to instructor

Participant 3

Five focusing steps:

1. Identify the system's constraint(s).
2. Decide how to exploit the constraint(s). Quick improvements. Use what you have.
3. Subordinate everything else to the decision in step 2.
4. Elevate the constraint(s). If more improvement is needed, what actions are required?
5. If in previous steps a constraint has been broken, return to step 1, but do not allow inertia to cause a system's constraint.

2.2 Six Sigma Statistics

Black Belt Training: Measure Phase

2.1 Process Definition
- 2.1.1 Cause and Effect Diagrams
- 2.1.2 Cause and Effects Matrix
- 2.1.3 Process Mapping
- 2.1.4 FMEA: Failure Modes & Effects Analysis
- 2.1.5 Theory of Constraints

2.2 Six Sigma Statistics
- 2.2.1 Basic Statistics
- 2.2.2 Descriptive Statistics
- 2.2.3 Distributions and Normality
- 2.2.4 Graphical Analysis

2.3 Measurement System Analysis
- 2.3.1 Precision and Accuracy
- 2.3.2 Bias, Linearity, and Stability
- 2.3.3 Gage R&R
- 2.3.4 Variable and Attribute MSA

2.4 Process Capability
- 2.4.1 Capability Analysis
- 2.4.2 Concept of Stability
- 2.4.3 Attribute and Discrete Capability
- 2.4.4 Monitoring Techniques

2.2.1 Basic Statistics

What is Statistics?

- **Statistics** is the science of collection, analysis, interpretation, and presentation of data.

- In Six Sigma, we apply statistical methods and principles to quantitatively measure and analyze the process performance to reach statistical conclusions and help solve business problems.

Types of Statistics

- Descriptive Statistics
  - Describing what was going on
- Inferential Statistics
  - Making inferences from the data at hand to more general conditions

Descriptive Statistics

- **Descriptive statistics** is applied to describe the main characteristics of a collection of data.
- Descriptive statistics summarizes the features of the data quantitatively.
- Descriptive statistics is descriptive only and it does not make any generalizations beyond the data at hand.
- The data used for descriptive statistics are for the purpose of representing or reporting.

Inferential Statistics

- **Inferential statistics** is applied to infer the characteristics or relationships of the populations from which the data are collected.

- Inferential statistics draws statistical conclusions about the population by analyzing the sample data subject to random variation.

- A complete data analysis includes both descriptive statistics and inferential statistics.

Statistics vs. Parameters

- The word *statistic* refers to a numeric measurement calculated using a sample data set, for example, sample mean or sample standard deviation. Its plural is *statistics* (the same spelling as “statistics” which refers to the scientific discipline).

- The *parameter* refers to a numeric metric describing the population, for example, population mean and population standard deviation. Unless you have the full data set of the population, you will not be able to know the population parameters.

Continuous Variable vs. Discrete Variable

- Continuous Variable
  - Measured
  - There is an infinite number of values possible
  - Examples: temperature, height, weight, money, time

- Discrete Variable
  - Counted
  - There is a finite number of values available
  - Examples: count of people, count of countries, count of defects, count of defectives

Types of Data

- Nominal
  - Categorical data
  - Examples: a set of colors, the social security number
- Ordinal
  - Rank-ordering data
  - Examples: the first, second place in a race, scores of exams
- Interval
  - Equidistant data
  - Examples: temperature with Fahrenheit or Celsius scale
- Ratio
  - The ratio between the magnitude of a continuous value and the unit value of the same category
  - Examples: weight, length, time

2.2.2 Descriptive Statistics

Basics of Descriptive Statistics

- Descriptive statistics provides a quantitative summary for the data collected.

- It summarizes the main features of the collection of data.
  - Shape
  - Location
  - Spread

- It is a presentation of data collected and it does *not* provide any inferences about a more general condition.

Shape of the Data

- **Distribution** is used to describe the shape of the data.
- Distribution (also called frequency distribution) summarizes the frequency of an individual value or a range of values of a variable (either continuous or discrete).
- Distribution is depicted as a table or graph.

Shape of the Data

- Simple example of distribution
  - We are tossing a fair die. The possible value we obtain from each tossing is a value between 1 and 6.
  - Each value between 1 and 6 has a 1/6 chance to be hit for each tossing.
  - The distribution of this game describes the relationship between every possible value and the percentage of times the value is being hit (or count of times the value is being hit).

Shape of the Data

- Examples of continuous distribution
  - Normal Distribution
  - T distribution
  - Chi-square distribution
  - F distribution

- Examples of discrete distribution
  - Binomial distribution
  - Poisson distribution

Location of the Data

- The **location** (i.e. central tendency) of the data describes the value where the data tend to cluster around.

- There are multiple measurements to capture the location of the data:
  - Mean
  - Median
  - Mode.

Mean

- The **mean** is the arithmetic average of a data set.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

n is the number of values in the data set

- For example, we have a set of data: 2, 3, 5, 8, 5, and 9. The arithmetic mean of the data set is

$$\frac{2+3+5+8+5+9}{6} = 5.33$$

Median

- The **median** is the middle value of the data set in numeric order.

- It separates the finite set of data into two parts: one with values higher that the median and the other with values lower than the median.

- For example, we have a set of data: 45, 32, 67, 12, 37, 54, and 28. The median is 37 since it is the middle value of the sorted list of values (i.e., 12, 28, 32, 37, 45, 54, and 67).

Mode

- The **mode** is the value that occurs most often in the data set.
- If no number is repeated, there is no mode for the data set.
- For example, we have a data set: 55, 23, 45, 45, 68, 34, 45, 55. The mode is 45 since it occurs most frequently.

Spread of the Data

- The **spread** (i.e. variation) of the data describes the degree of data dispersing around the center value.

- There are multiple measurements to capture the spread of the data:
  - Range
  - Variance
  - Standard Deviation.

Range

- The **range** is the numeric difference between the greatest and smallest values in a data set.
- Only two data values (i.e., the greatest and the smallest values) are accounted for calculating the range.
- For example, we have a set of data: 34, 45, 23, 12, 32, 78, and 23. The range of the data is 78−12 = 66.

Variance

- The **variance** measures how far on average the data points spread out from the mean.
- It is the average squared deviation of each value from its mean.
- It is the square of the sample standard deviation.
- All the data points are accounted for calculating the variance.

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

where

n is the number of values in the data set

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Standard Deviation

- **Standard deviation** describes how far the data points spread away from the mean.
- It has the same measurement units as the data itself.
- It is simply the square root of the variance.
- All the data points are accounted for calculating the standard deviation.

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

where

n is the number of values in the data set

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

2.2.3 Normal Distribution & Normality

What is Normal Distribution?

- The **normal distribution** is a probability distribution of a continuous random variable whose values spread symmetrically around the mean.
- A normal distribution can be completely described by using its mean (μ) and variance (σ^2).
- When a variable x is normally distributed, we denote $x \sim N(\mu, \sigma^2)$.

Z Distribution

- The **Z distribution** is the simplest normal distribution with the mean equal to zero and the variance equal to one.
- Any normal distribution can be transferred to a Z distribution by applying

$$z = \frac{x - \mu}{\sigma}$$

where

$$x \sim N\left(\mu, \sigma^2\right) \qquad \sigma \neq 0$$

Z Score

- The **Z Score** is the measure of how many standard deviations an observation is above or below the mean.
- Positive Z Scores indicate the observation is above the mean or “right of the mean”.
- Negative Z Scores indicate the observation is below the mean of “left of the mean”
- Calculate Z Score using the formula below:

$$z = \frac{x - \mu}{\sigma}$$

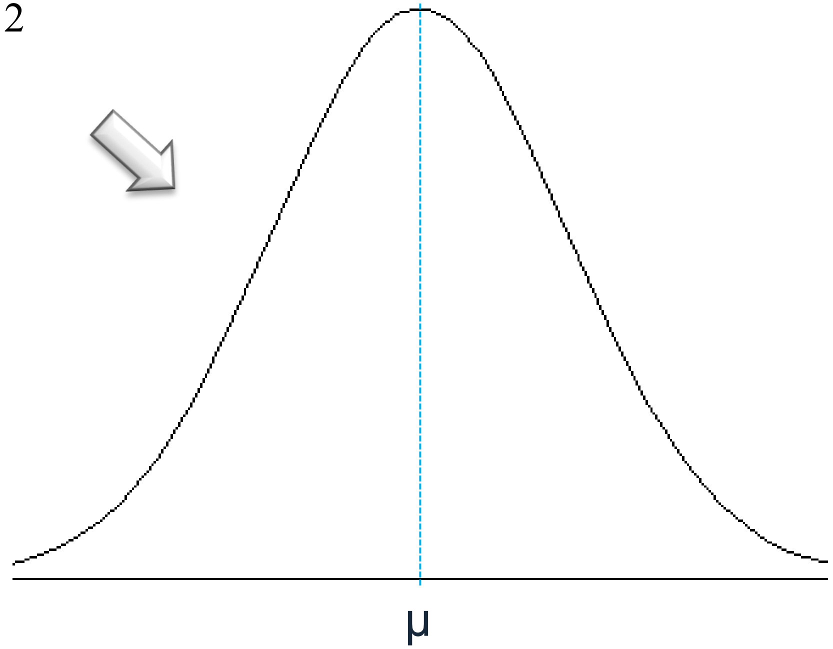
where

x is the observation
μ is the mean of the population
σ is the standard deviation of the population

Shape of Normal Distribution

- The probability density function curve of normal distribution is bell-shaped.
- Probability density function of normal distribution

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



Location of Normal Distribution

- If a variable is normally distributed, the mean, the median, and the mode have the same value.

- The probability density curve of normal distribution is symmetric around a center value which is the mean, the median, and the mode at the same time.

Spread of Normal Distribution

- The spread or variation of the normally-distributed data can be described using the variance or the standard deviation.

- The smaller the variance or the standard deviation, the less variability in the data set.

Empirical Rule

- The 68-95-99.7 rule or the *empirical rule in statistics* states that for a normal distribution:
  - About 68% of the data fall within one standard deviation of the mean
  - About 95% of the data fall within two standard deviations of the mean
  - About 99.7% of the data fall within three standard deviations of the mean.

Empirical Rule *(68-95-99.7 Rule)*

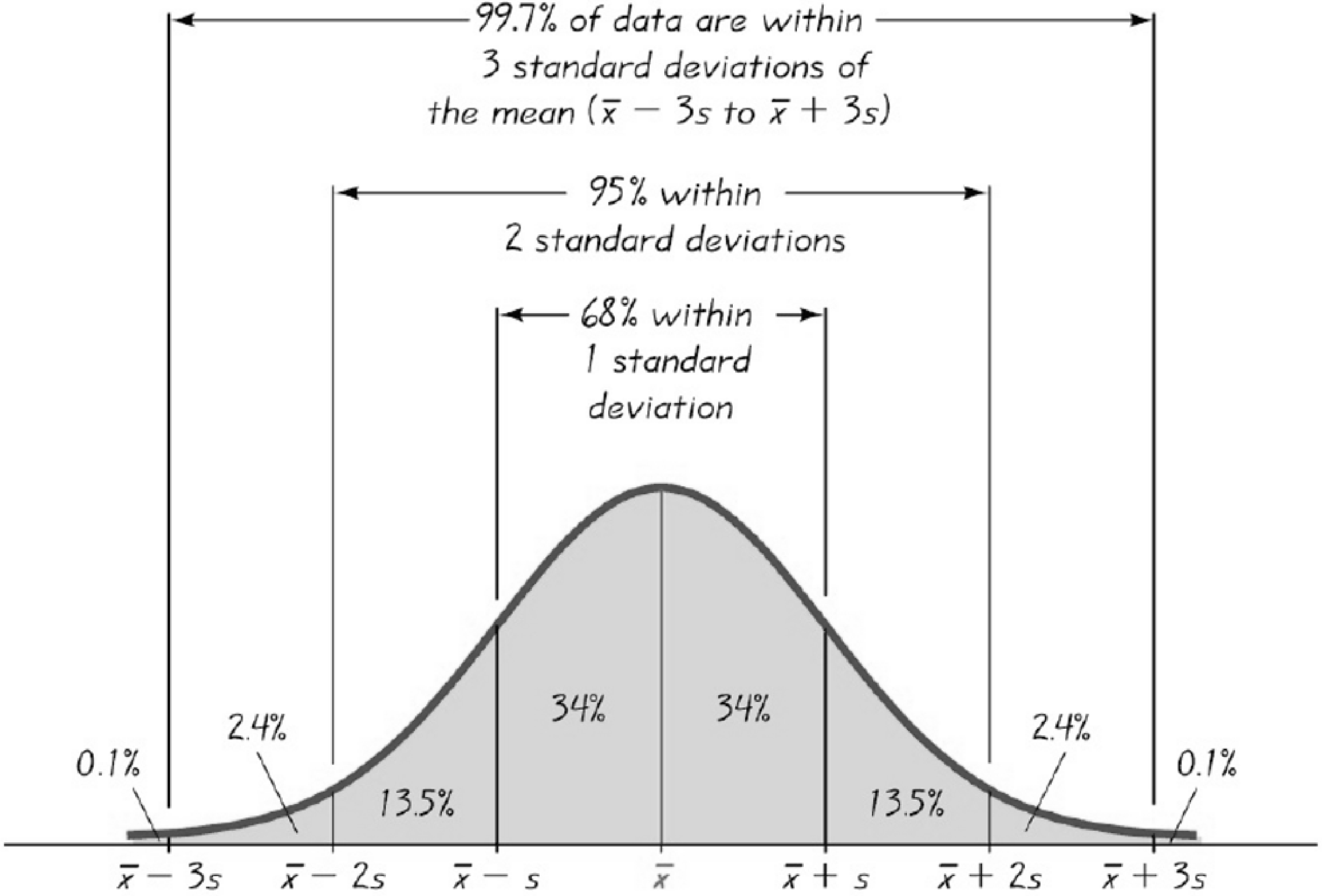


Image Attribution: Pearson Education Inc.

Normality

- Not all the distributions with a bell shape are normal distributions.

- To check whether a group of data points are normally distributed, we need to run a *normality test*.

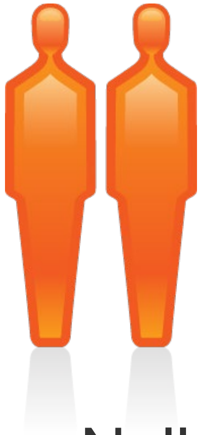
- There are different normality tests available:
  - Anderson-Darling test
  - Sharpiro-Wilk test
  - Jarque-Bera test.

- More details of normality test will be introduced in the Analyze module.

Normality Testing

- To check whether the population of our interest is normally distributed, we need to run normality test.
  - Null Hypothesis (H_0): The data points *are* normally distributed.
  - Alternative Hypothesis (H_a): The data points are *not* normally distributed.
- There are many normality tests available. For example, Anderson-Darling test, Sharpiro-Wilk test, Jarque-Bera test, and so on.

Use Minitab to Run a Normality Test

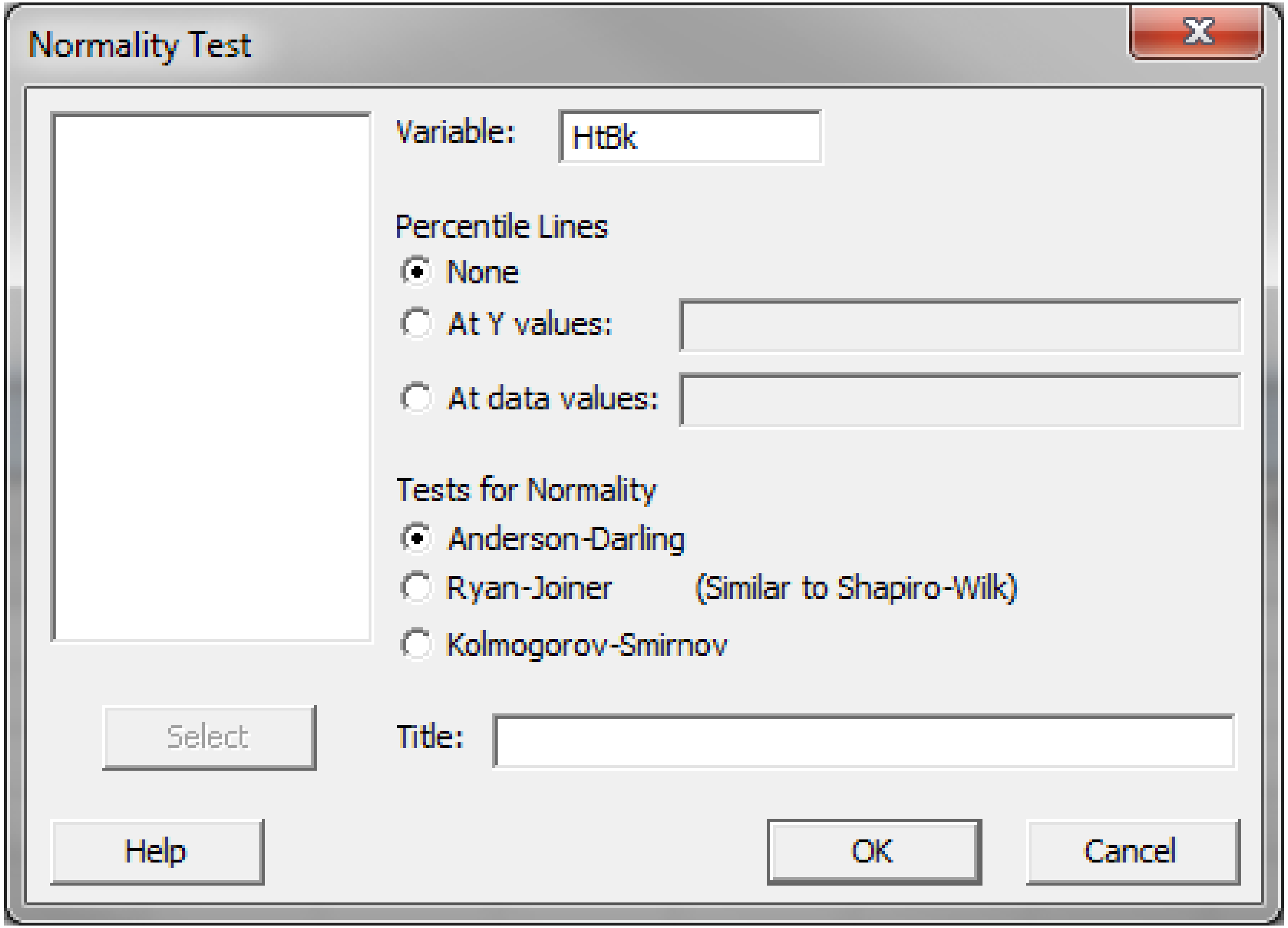


- Case study: we are interested to know whether the height of basketball players is normally distributed.
  - Data File: “One Sample T-Test” tab in “Sample Data.xlsx”
- Null Hypothesis (H0): the height of basketball players is normally distributed.
- Alternative Hypothesis (Ha): the height of basketball players is not normally distributed.

Use Minitab to Run a Normality Test

- Steps to run a normality test in Minitab
    1) Click Stat → Basic Statistics → Normality Test.
    2) A new window named "Normality Test" pops up.
    3) Select "HtBk" as the "Variable."
    4) Click "OK."
    5) The normality test results appear in the new window.

Use Minitab to Run a Normality Test



Normality Test

Variable: HtBk

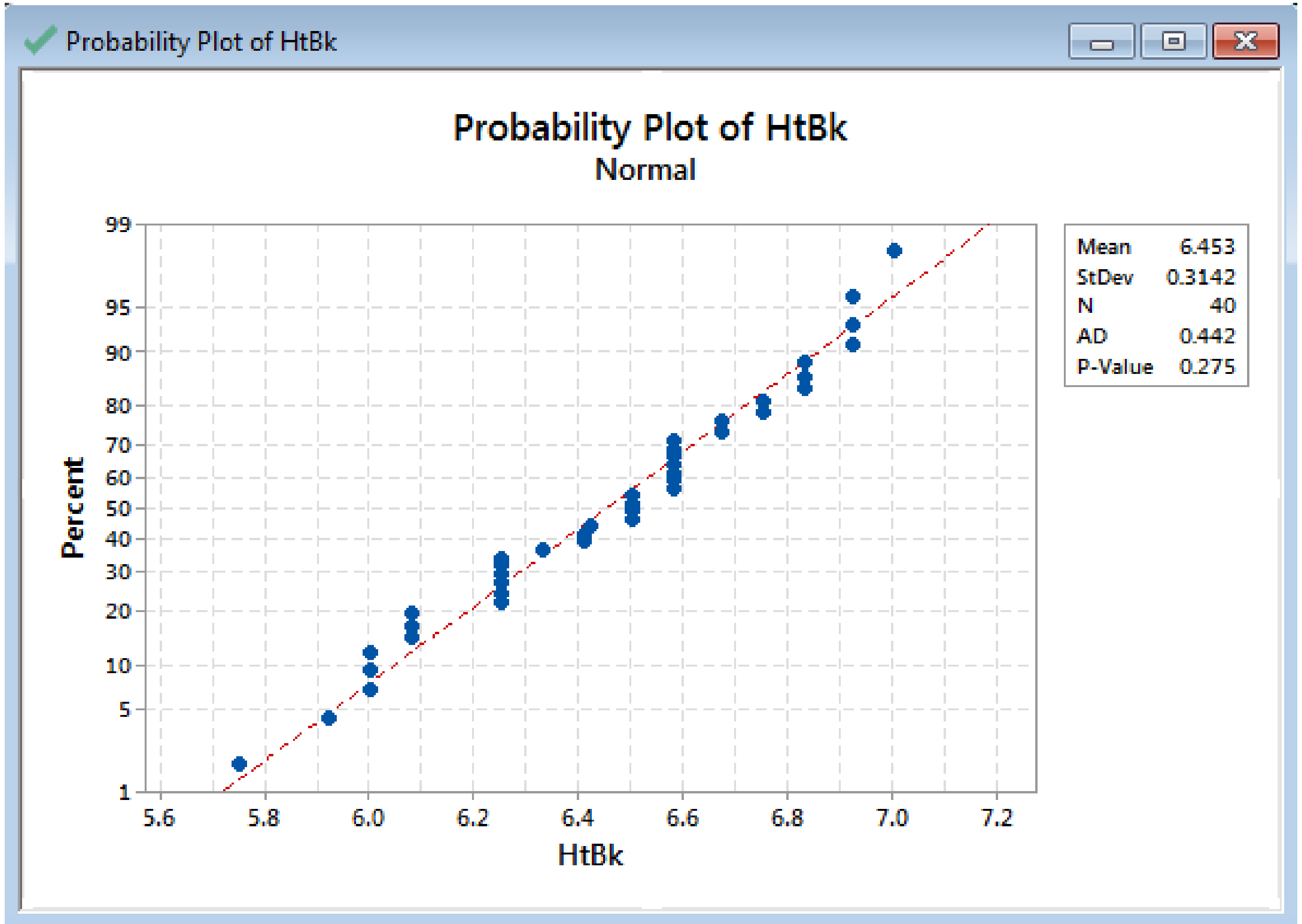
Percentile Lines
- ● None
- ○ At Y values:
- ○ At data values:

Tests for Normality
- ● Anderson-Darling
- ○ Ryan-Joiner (Similar to Shapiro-Wilk)
- ○ Kolmogorov-Smirnov

Title:

Use Minitab to Run a Normality Test

- Null Hypothesis (H_0): The data are normally distributed.
- Alternative Hypothesis (H_a): The data are not normally distributed.



Mean	6.453
StDev	0.3142
N	40
AD	0.442
P-Value	0.275

- Since the p-value of the normality is 0.275 greater than alpha level (0.05), we fail to reject the null and claim that the data are normally distributed.

2.2.4 Graphical Analysis

What is Graphical Analysis?

- In statistics, **graphical analysis** is a method to visualize the quantitative data.

- Graphical analysis is used to discover the structure and patterns in the data, explaining and presenting the statistical conclusions.

- A complete statistical analysis includes both quantitative analysis and graphical analysis.

Graphical Analysis Example

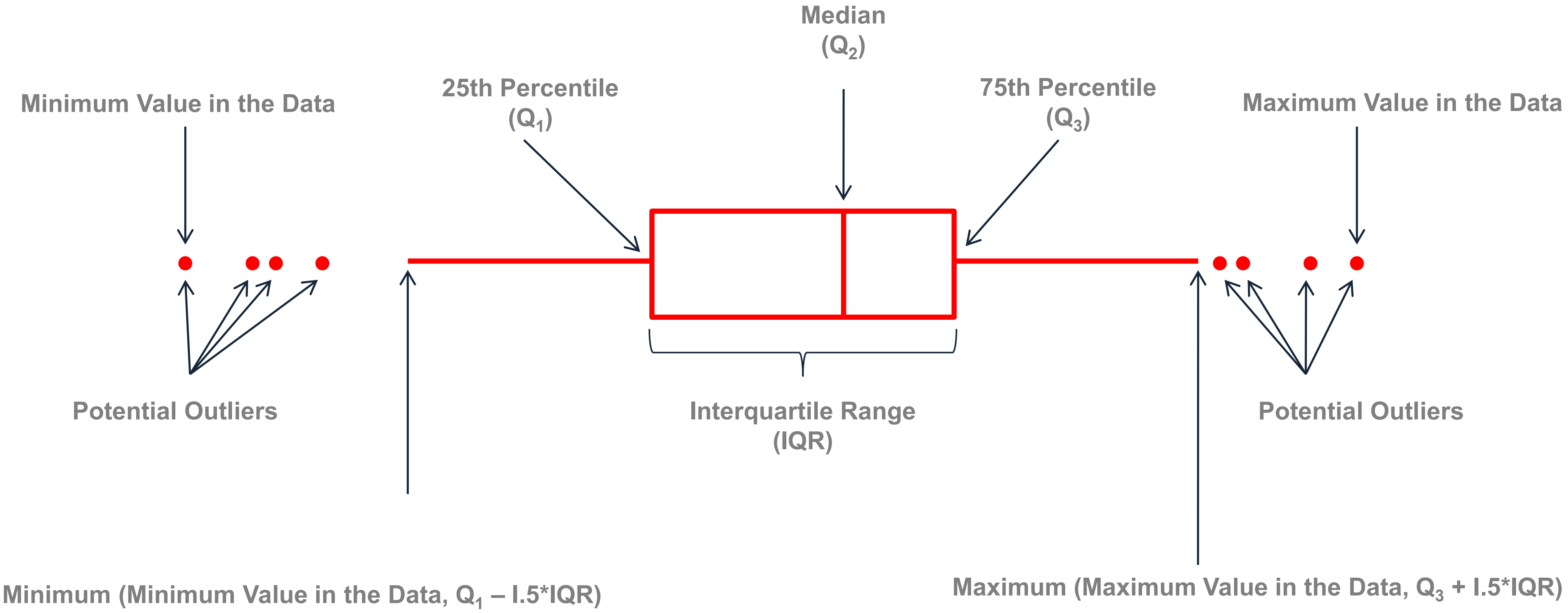
- There are various graphical analysis tools available. Here are four most commonly used examples:
  - Box Plot
  - Histogram
  - Scatter Plot
  - Run Chart.

Box Plot

- A **box plot** is a graphical method to summarize a data set by visualizing the minimum value, 25th percentile, median, 75th percentile, the maximum value, and potential outliers.

- A percentile is the value below which a certain percentage of data fall. For example, if 75% of the observations have values lower than 685 in a data set, then 685 is the 75th percentile of the data.

Box Plot

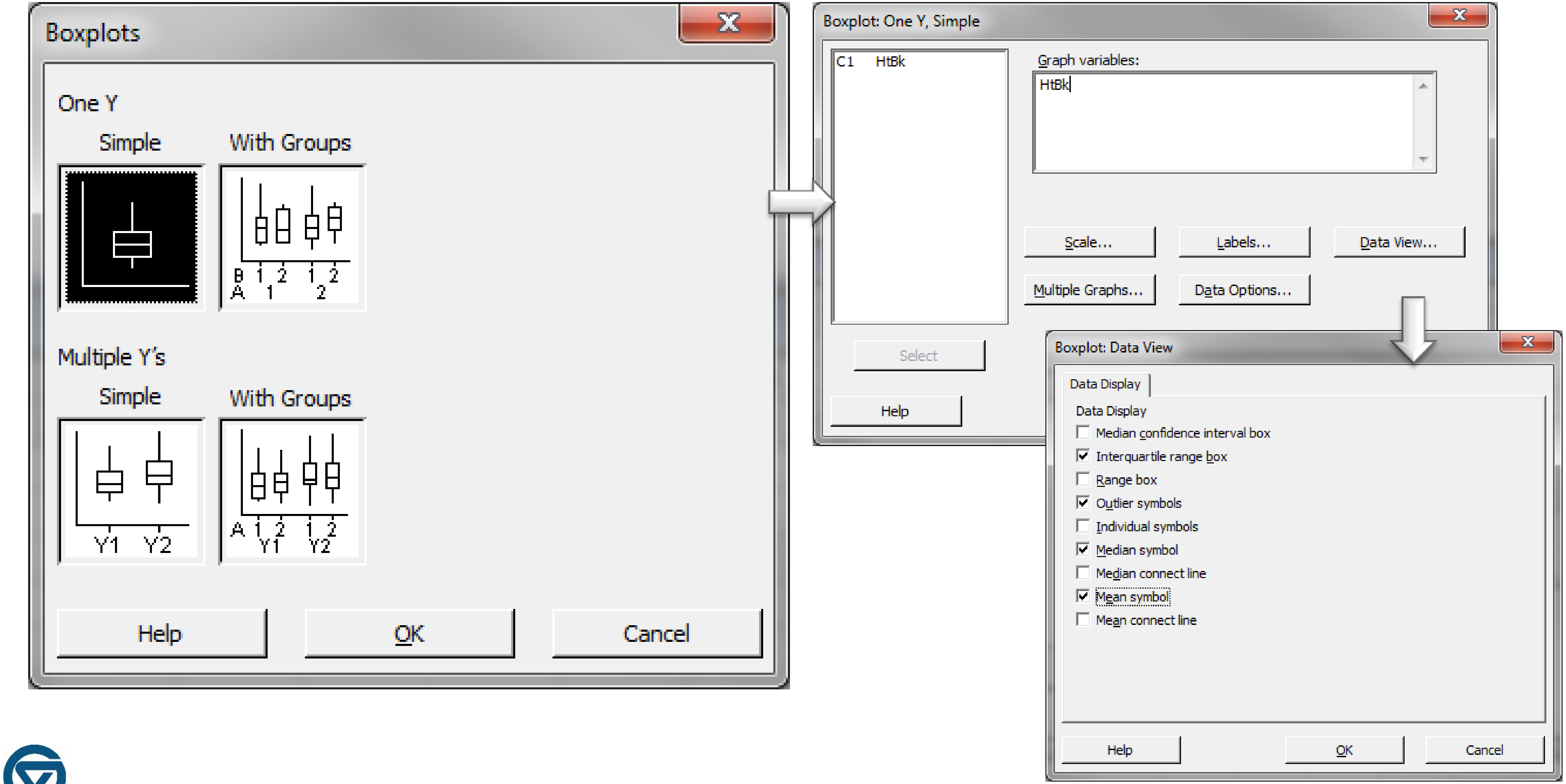


Interquartile Range = 75th Percentile – 25th Percentile

How to Use Minitab to Generate a Box Plot

- Data File: “Box Plot” tab in “Sample Data.xlsx”
- Steps to render a Box Plot in Minitab
    1) Click Graph → Boxplot.
    2) A new window named “Boxplots” pops up.
    3) Click “OK” in the window “Boxplots.”
    4) Another new window named “Boxplot – One Y, Simple” pops up.
    5) Select “HtBk” as the “Graph Variables.”
    6) Click the “Data View” button and a new window named “Boxplot – Data View” pops up.
    7) Check the boxes “Median symbol” and “Mean symbol.”
    8) Click “OK” in the window “Boxplot – Data View.”
    9) Click “OK” in the window “Boxplot – One Y, Simple.”
    10) The box plot appears automatically in the new window.

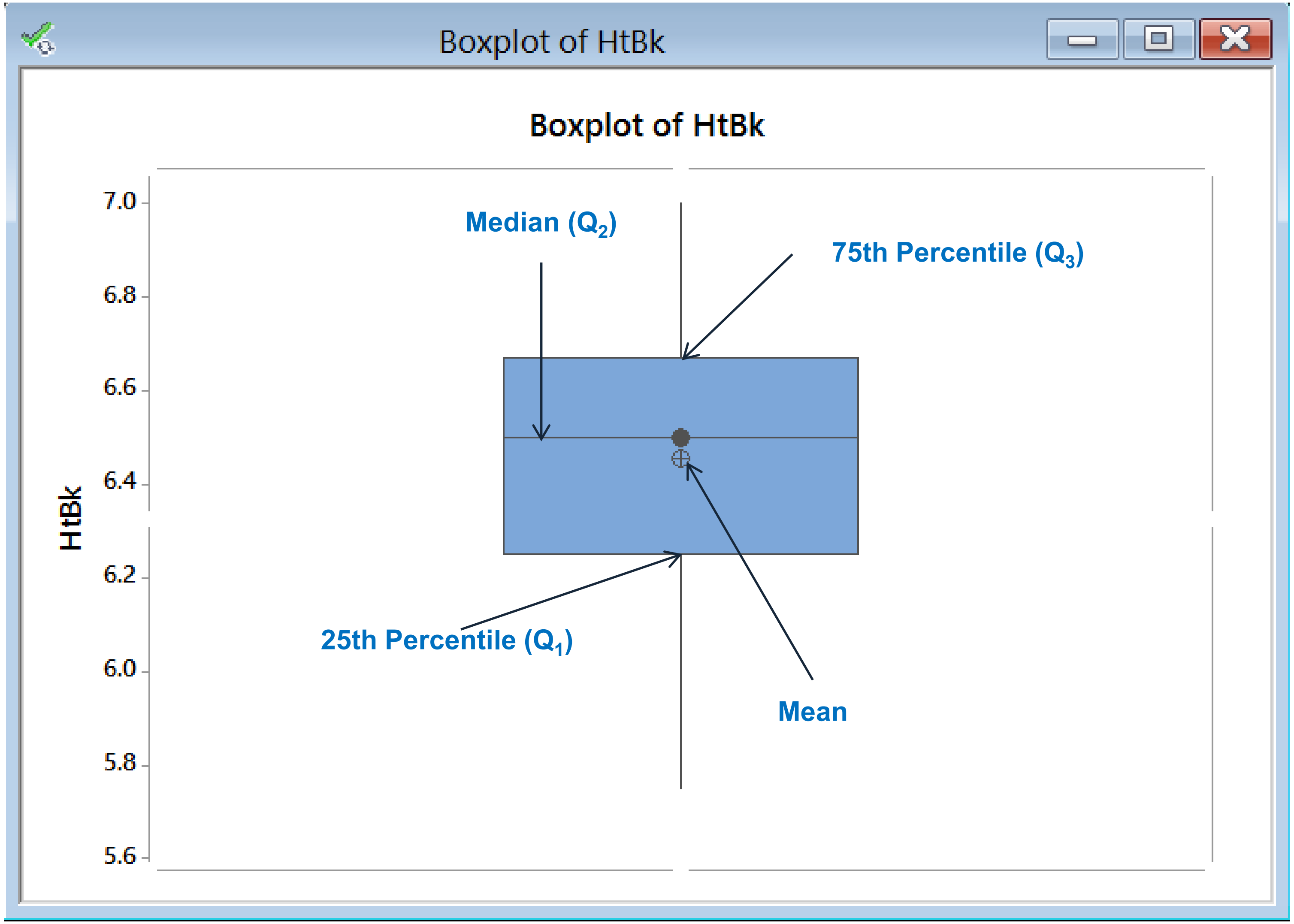
How to Use Minitab to Generate a Box Plot



Graph variables: HtBk

Data Display
- ☐ Median confidence interval box
- ☑ Interquartile range box
- ☐ Range box
- ☑ Outlier symbols
- ☐ Individual symbols
- ☑ Median symbol
- ☐ Median connect line
- ☑ Mean symbol
- ☐ Mean connect line

How to Use Minitab to Generate a Box Plot



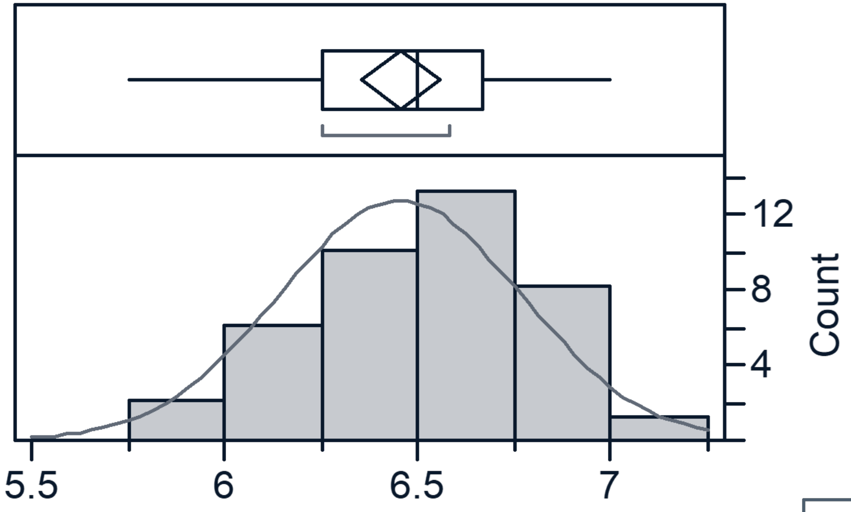
Histogram

- A **histogram** is a graphical tool to present the distribution of the data.
- The X axis represents the possible values of the variable and the Y axis represents the frequency of the value occurring.
- A histogram consists of adjacent rectangles erected over intervals with heights equal to the frequency density of the interval.
- The total area of all the rectangles in a histogram is the number of data values.

Histogram

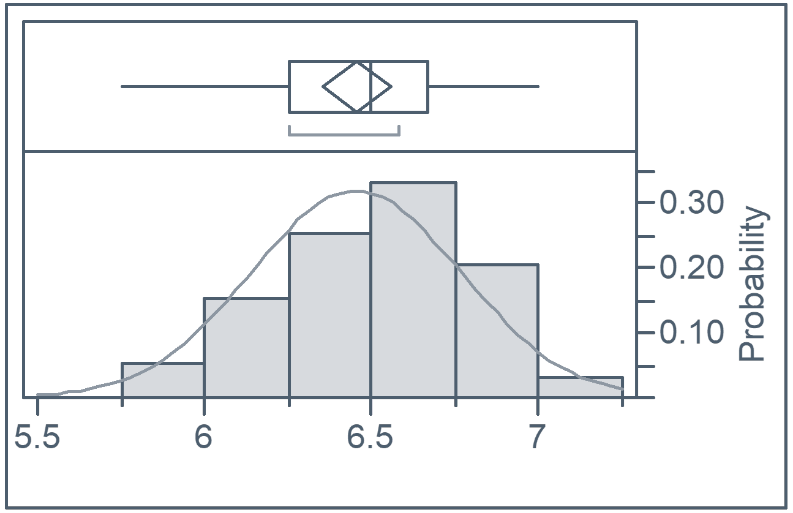
- A histogram can also be normalized. In this case, the X axis still represents the possible values of the variable, but the Y axis represents the percentage of observations that fall into each interval on the X axis.
- The total area of all the rectangles in a normalized histogram is 1.
- With the histogram, we have a better understanding of the shape, location, and spread of the data.

Histogram



Histogram with frequency (count) as the Y axis

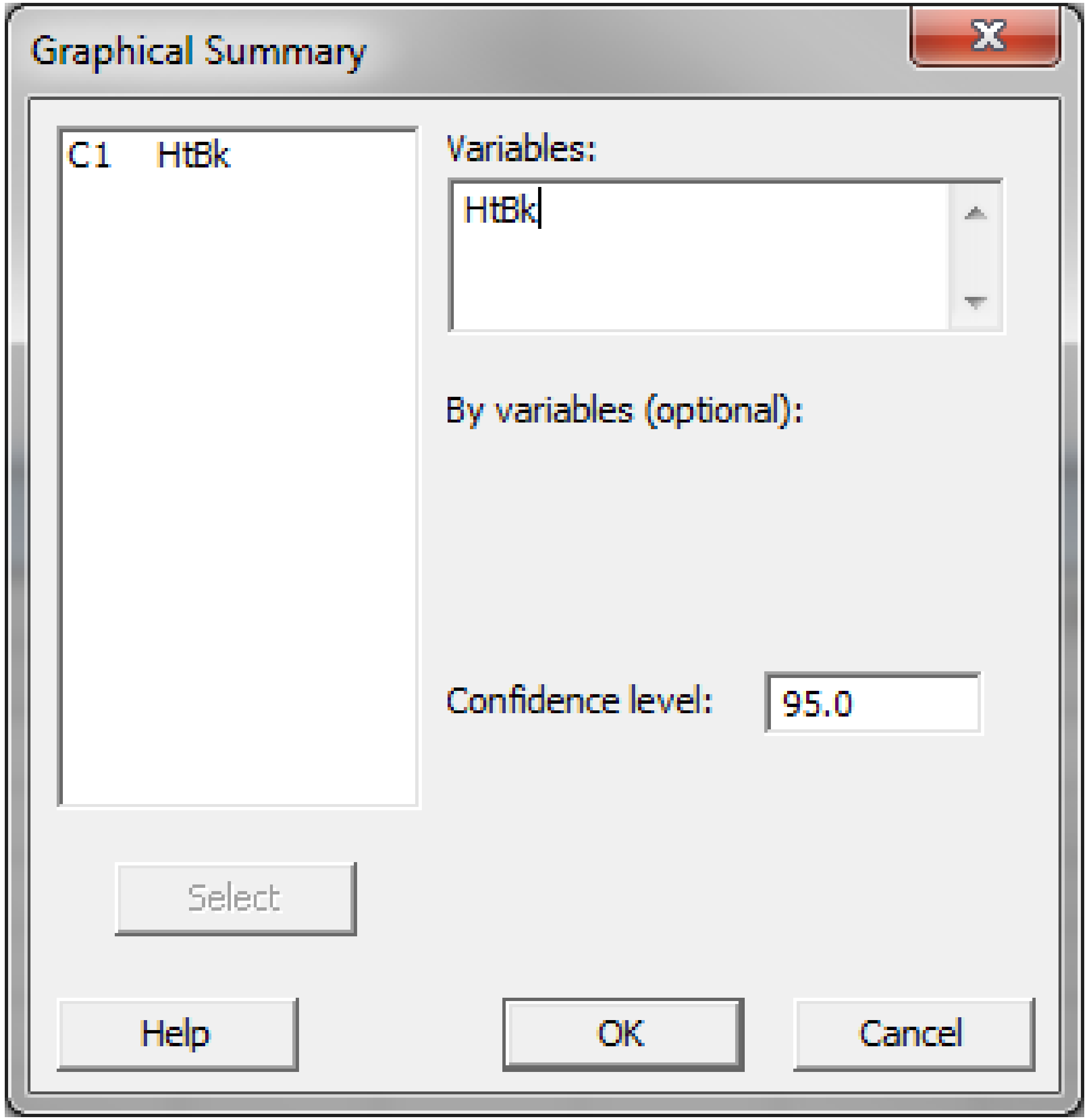
Normalized histogram with proportion (probability) as the Y axis



How to Use Minitab to Generate a Histogram

- Data File: “Histogram” tab in “Sample Data.xlsx”
- Steps to render a histogram in Minitab
  1) Click Stat → Basic Statistics → Graphical Summary.
  2) A new window named “Graphical Summary” pops up.
  3) Select “HtBk” as the “Variables.”
  4) Click “OK.”
  5) The histogram appears in the new window.

How to Use Minitab to Generate a Histogram

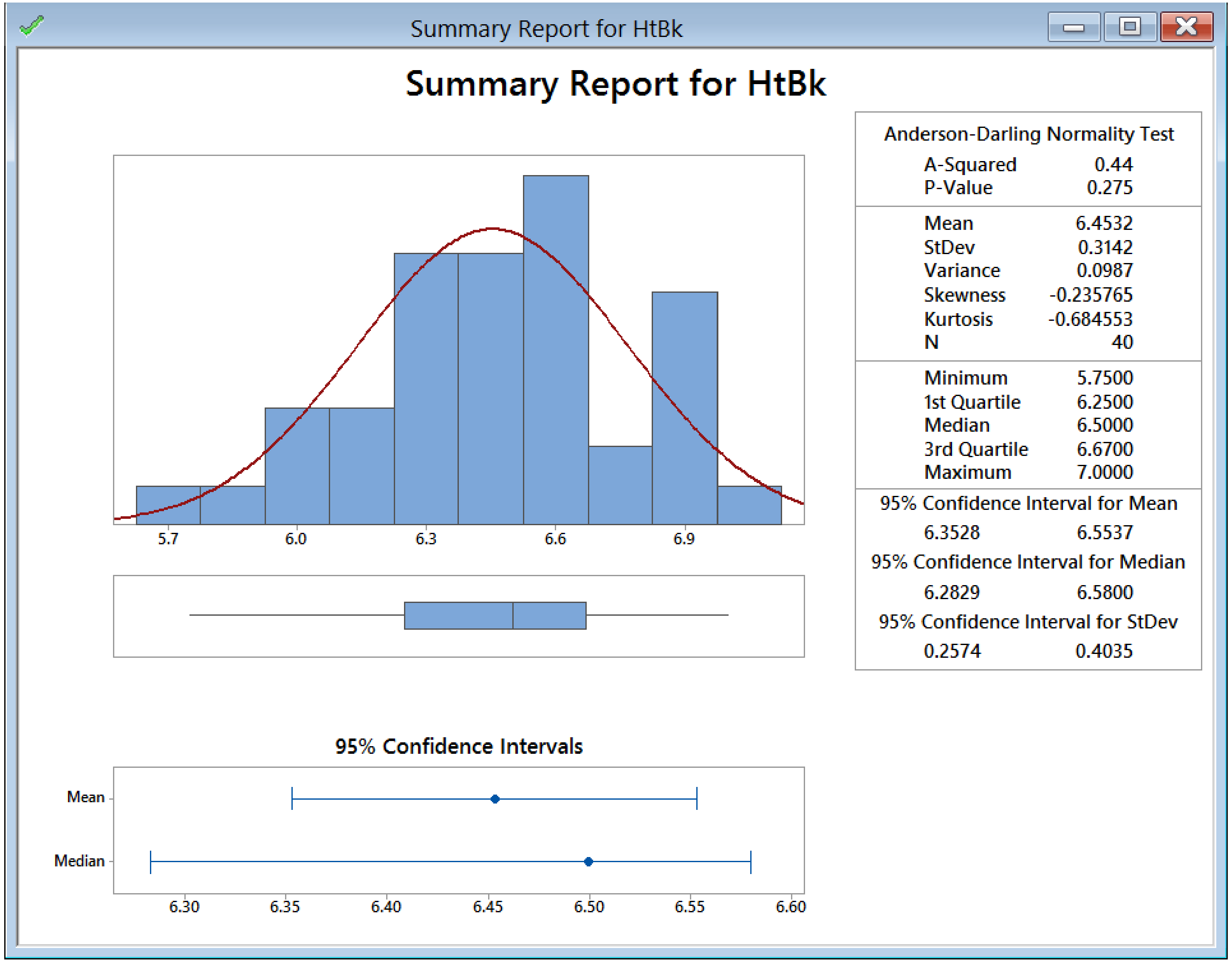


Variables: HtBk

By variables (optional):

Confidence level: 95.0

How to Use Minitab to Generate a Histogram



Summary Report for HtBk

Anderson-Darling Normality Test	
A-Squared	0.44
P-Value	0.275
Mean	6.4532
StDev	0.3142
Variance	0.0987
Skewness	-0.235765
Kurtosis	-0.684553
N	40
Minimum	5.7500
1st Quartile	6.2500
Median	6.5000
3rd Quartile	6.6700
Maximum	7.0000
95% Confidence Interval for Mean	
6.3528	6.5537
95% Confidence Interval for Median	
6.2829	6.5800
95% Confidence Interval for StDev	
0.2574	0.4035

Scatter Plot

- A **scatter plot** is a diagram to present the relationship between two variables of a data set.

- A scatter plot consists of a set of data points.

- On the scatter plot, a single observation is presented by a data point with its horizontal position equal to the value of one variable and its vertical position equal to the value of the other variable.

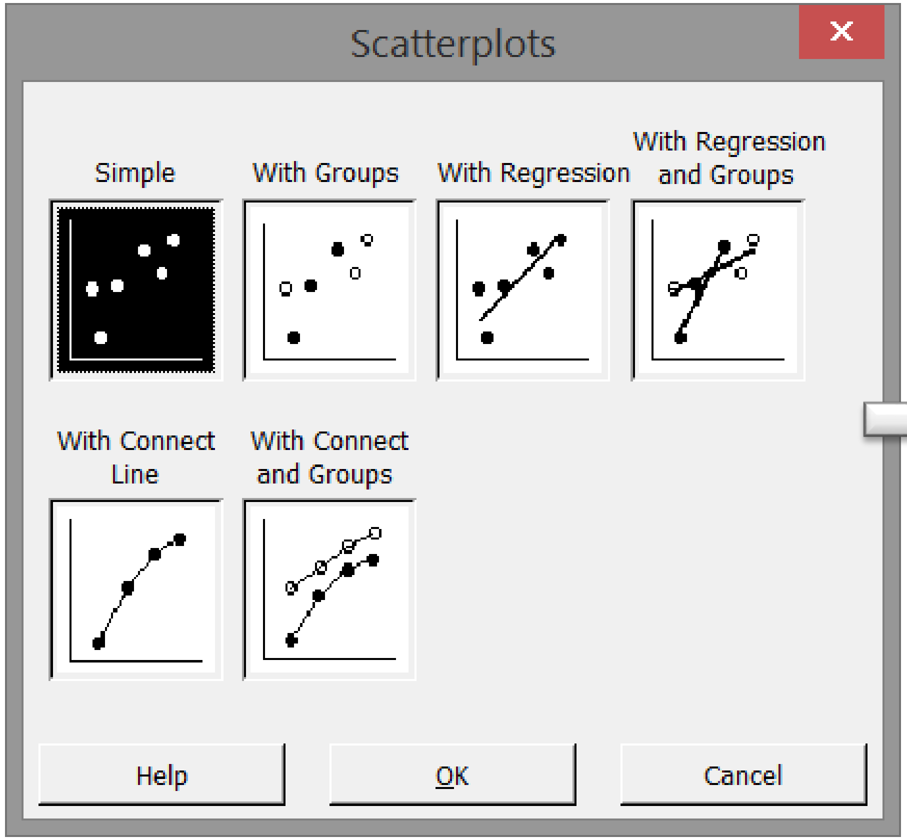
Scatter Plot

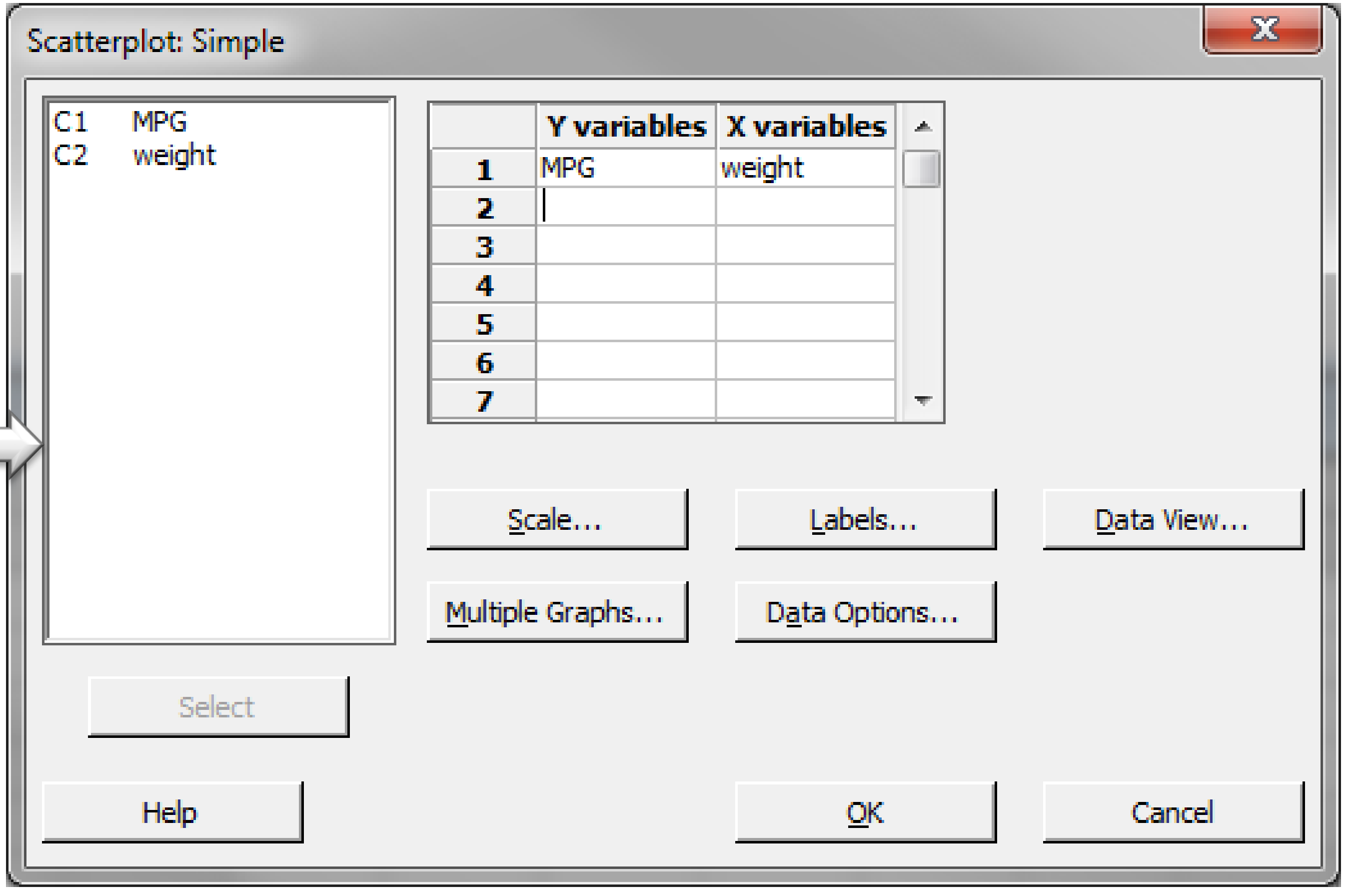
- A scatter plot helps to understand:
  - Whether the two variables are related to each other or not
  - How is the strength of their relationship
  - What is the shape of their relationship
  - What is the direction of their relationship
  - Whether outliers are present.

How to Use Minitab to Generate a Scatter Plot

- Data File: “Scatter Plot” tab in “Sample Data.xlsx”
- Steps to render a histogram in Minitab:
  1) Click Graph → Scatterplot.
  2) A new window named “Scatterplots” pops up.
  3) Click “OK” in the window “Scatterplots” and another window named “Scatterplot – Simple” pops up.
  4) Select “MPG” as the “Y variables.”
  5) Select “weight” as the “X variables.”
  6) Click “OK.”
  7) The scatter plot appears in the new window.

How to Use Minitab to Generate a Scatter Plot

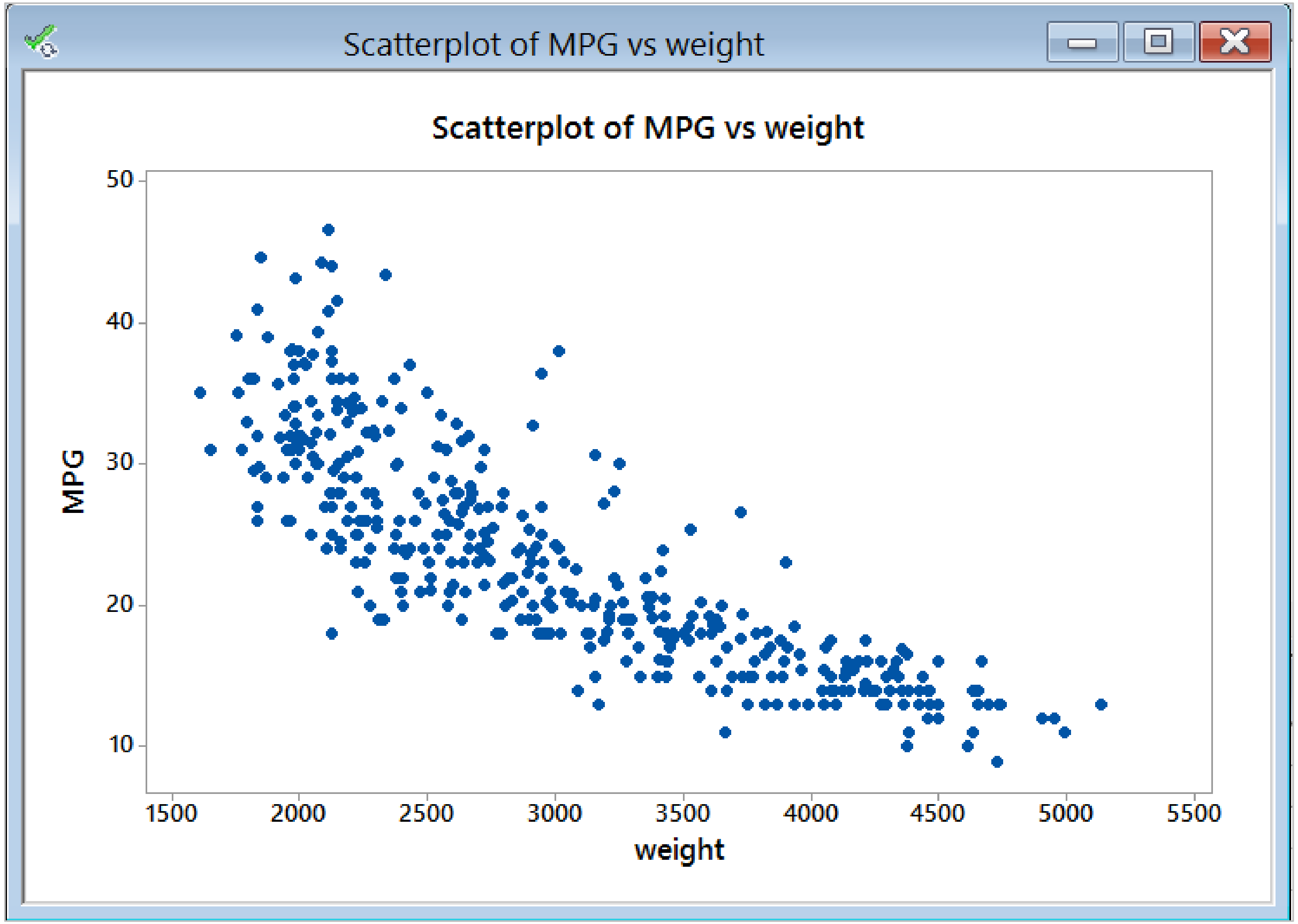




C1 MPG
C2 weight

	Y variables	X variables
1	MPG	weight
2		
3		
4		
5		
6		
7		

How to Use Minitab to Generate a Scatter Plot



Run Chart

- A **run chart** is a chart used to present the data in time order. It captures the process performance over time.

- The X axis of the run chart indicates the time and the Y axis indicates the observed values.

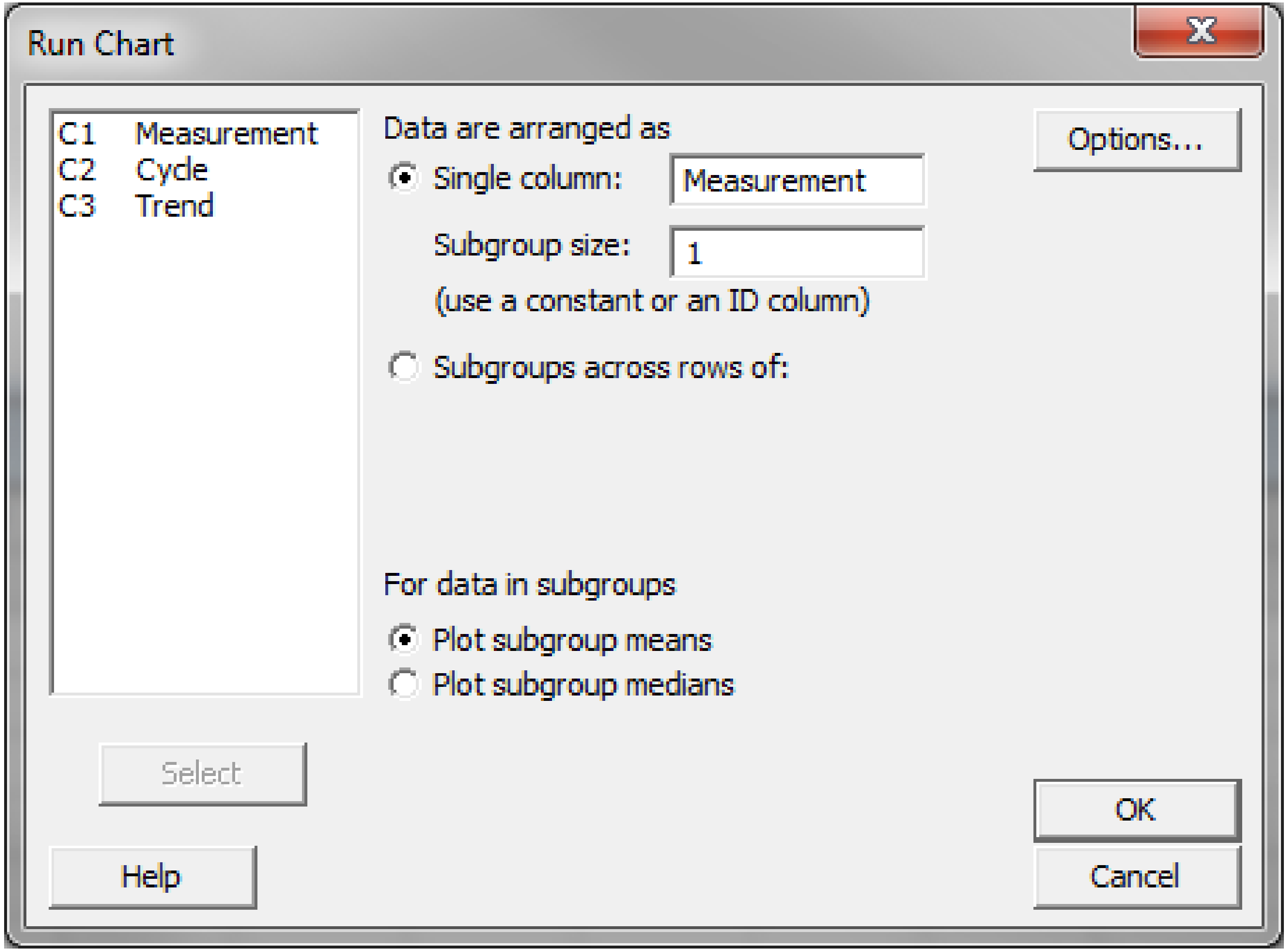
- Run chart looks similar to control charts except that a run chart does not have control limits plotted. It is easier to produce a run chart than a control chart.

- It is often used to identify the anomalies in the data and discover the pattern of data changing over time.

How to Plot a Run Chart in Minitab

- Steps to plot a run chart in Minitab:
  1) Data File: “Run Chart” tab in “Sample Data.xlsx.”
  2) Click Stat → Quality Tools → Run Chart.
  3) A new window named “Run Chart” pops up.
  4) Select “Measurement” as the “Single Column.”
  5) Enter “1” as the “Subgroup Size.”
  6) Click “OK.”
  7) The run chart appears automatically in the new window.

How to Plot a Run Chart in Minitab



Data are arranged as

● Single column: Measurement

Subgroup size: 1

(use a constant or an ID column)

○ Subgroups across rows of:

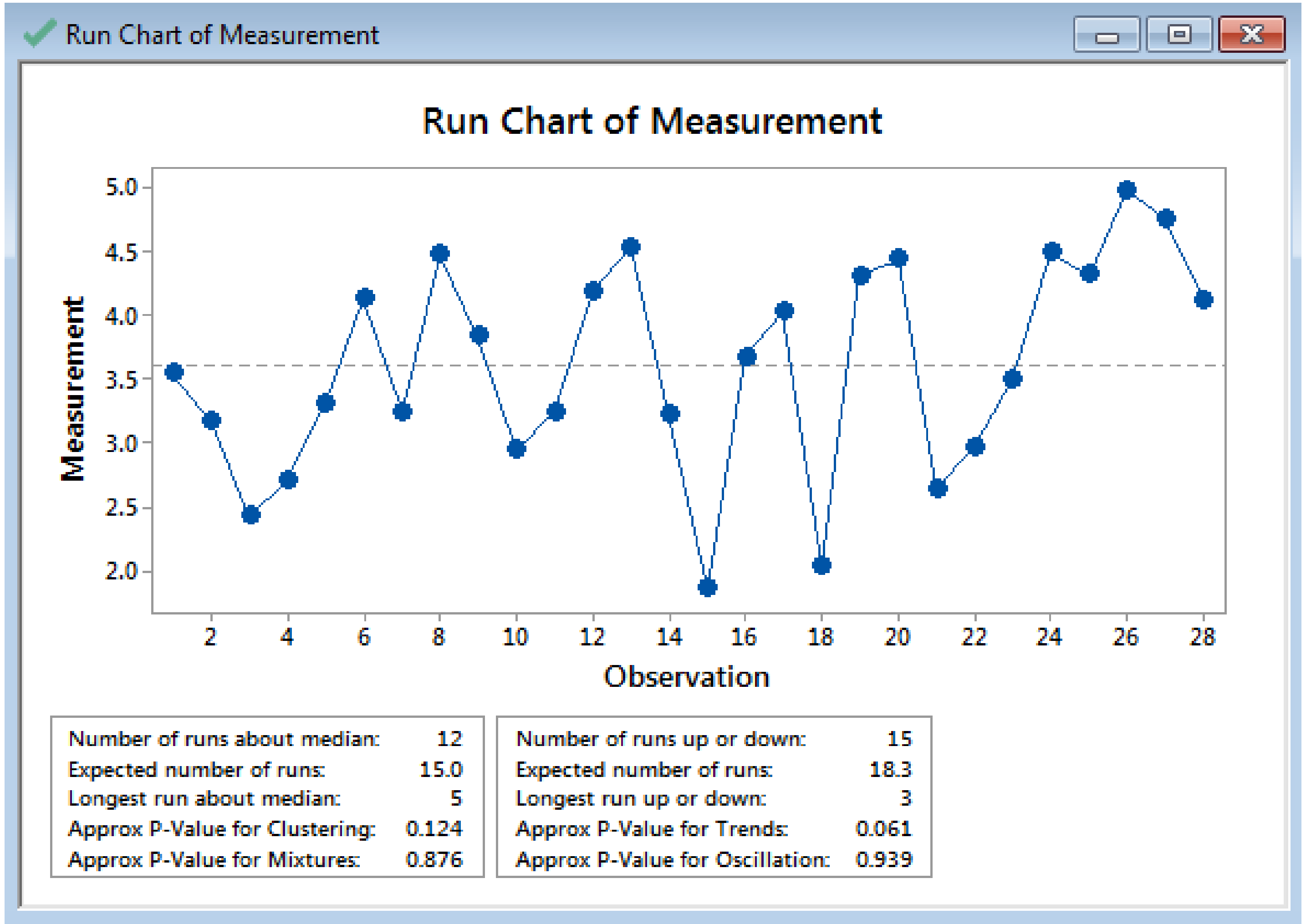
For data in subgroups

● Plot subgroup means

○ Plot subgroup medians

How to Plot a Run Chart in Minitab

- Run charts are used to identify the trend, cycle, seasonal pattern, or abnormality in the data.
  - The time series in this chart appears stable.
  - There are no extreme outliers, trending or seasonal patterns.



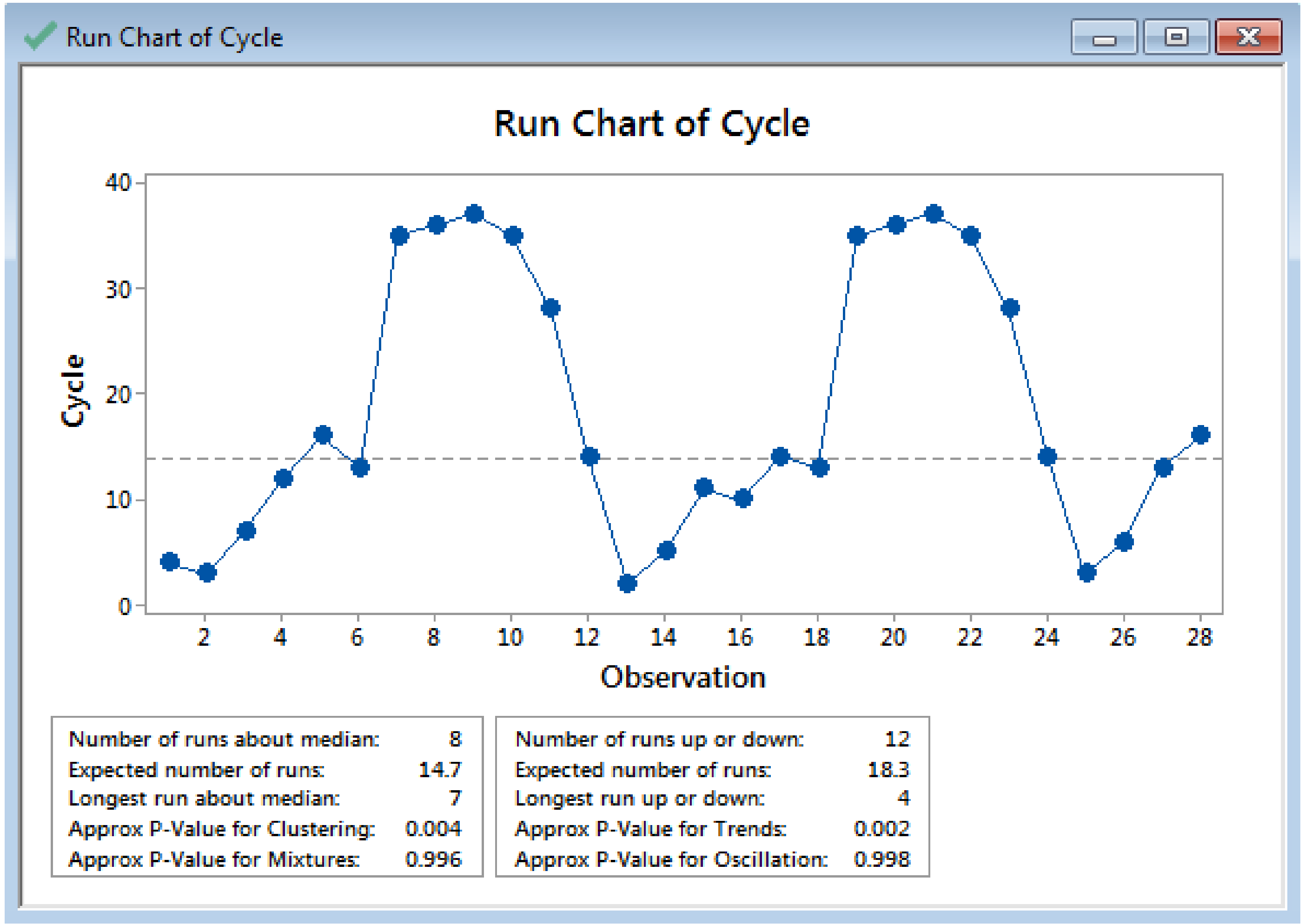
Number of runs about median:	12
Expected number of runs:	15.0
Longest run about median:	5
Approx P-Value for Clustering:	0.124
Approx P-Value for Mixtures:	0.876

Number of runs up or down:	15
Expected number of runs:	18.3
Longest run up or down:	3
Approx P-Value for Trends:	0.061
Approx P-Value for Oscillation:	0.939

Run Chart Example

- Following the previous "Run Chart" instructions, create a new chart using the data listed in the column "Cycle" in the "Run Chart" tab of "Sample Data.xlsx."

- Notice the repeating pattern of the data. This could suggest something cyclical or seasonal.

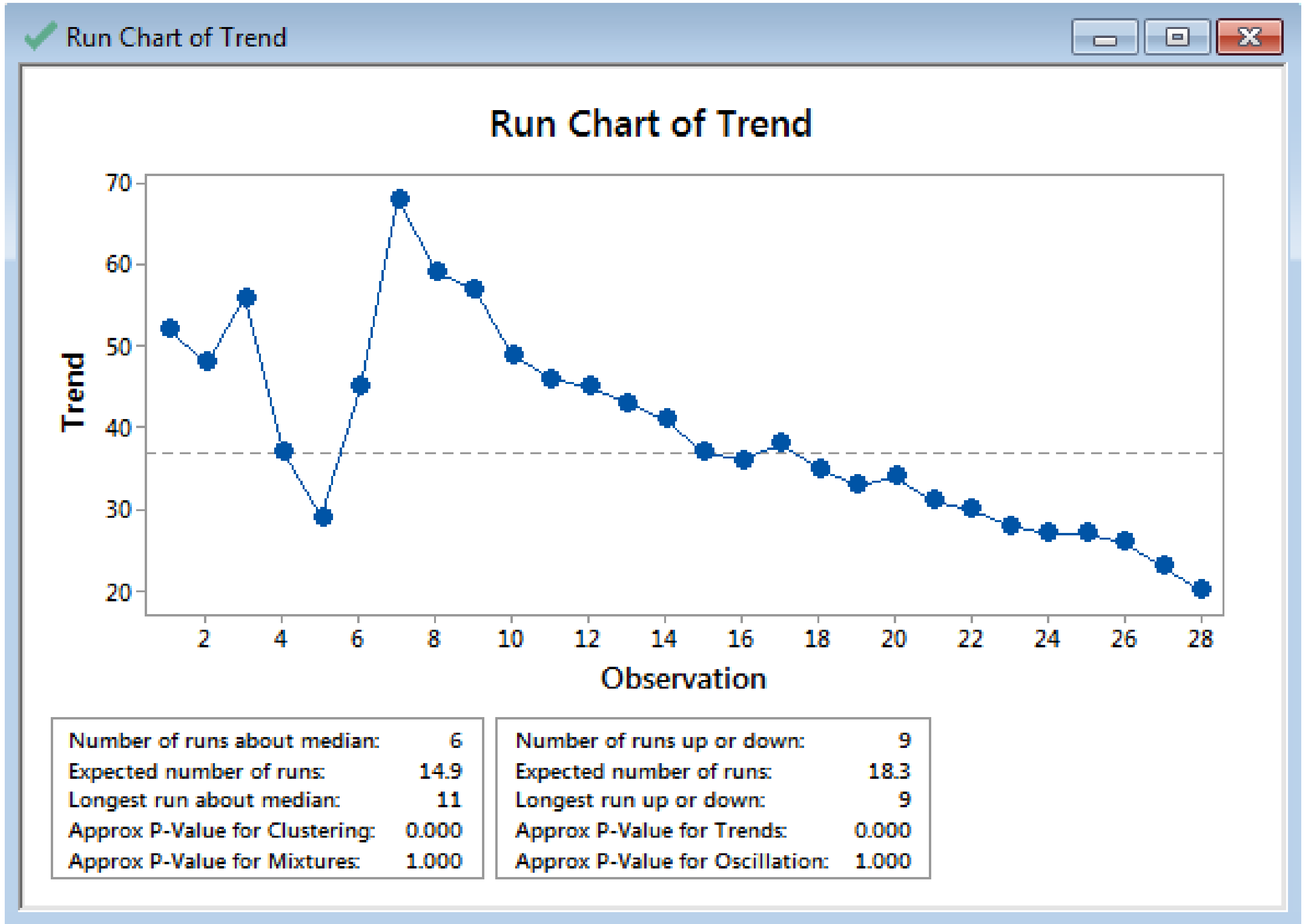


Number of runs about median:	8
Expected number of runs:	14.7
Longest run about median:	7
Approx P-Value for Clustering:	0.004
Approx P-Value for Mixtures:	0.996

Number of runs up or down:	12
Expected number of runs:	18.3
Longest run up or down:	4
Approx P-Value for Trends:	0.002
Approx P-Value for Oscillation:	0.998

Run Chart Example

- Create another run chart using the data listed in column “Trend” in the “Run Chart” tab of “Sample Data.xlsx.”
- Notice the trending pattern represented by the data points successively declining.



```
Number of runs about median:        6    Number of runs up or down:          9
Expected number of runs:         14.9    Expected number of runs:         18.3
Longest run about median:          11    Longest run up or down:             9
Approx P-Value for Clustering:  0.000    Approx P-Value for Trends:      0.000
Approx P-Value for Mixtures:    1.000    Approx P-Value for Oscillation: 1.000
```

2.3 MSA (Measurement System Analysis)

Black Belt Training: Measure Phase

2.1 Process Definition
- 2.1.1 Cause and Effect Diagrams
- 2.1.2 Cause and Effects Matrix
- 2.1.3 Process Mapping
- 2.1.4 FMEA: Failure Modes & Effects Analysis
- 2.1.5 Theory of Constraints

2.2 Six Sigma Statistics
- 2.2.1 Basic Statistics
- 2.2.2 Descriptive Statistics
- 2.2.3 Distributions and Normality
- 2.2.4 Graphical Analysis

2.3 Measurement System Analysis
- 2.3.1 Precision and Accuracy
- 2.3.2 Bias, Linearity, and Stability
- 2.3.3 Gage R&R
- 2.3.4 Variable and Attribute MSA

2.4 Process Capability
- 2.4.1 Capability Analysis
- 2.4.2 Concept of Stability
- 2.4.3 Attribute and Discrete Capability
- 2.4.4 Monitoring Techniques

2.3.1 Precision and Accuracy

What is Measurement System Analysis

- **Measurement System Analysis** (MSA) is a systematic method to identify and analyze the variation components in the measurement.
- It is a mandatory step in any Six Sigma project to ensure the data are reliable before making any data-based decisions.
- The MSA is the check point of data quality before we start any further analysis and draw any conclusions from the data.

Data-Based Analysis

- Here are some examples of data-based analysis where MSA is the prerequisite:
  - Correlation analysis
  - Regression analysis
  - Hypothesis testing
  - Analysis of variance
  - Design of experiments
  - Multivariate analysis
  - Statistical process control.

Measurement System

- A measurement system is a process to obtain data.

Inputs (Xs)

Measurement System

Output (Y)

- Y (output of the measurement system)
  - Observed values
- X’s (inputs of the measurement system)
  - True values
  - Measurement errors

Measurement Errors

Observed Value = True Value + Measurement Error

- **True Value**
  - The actual value we are interested to measure
  - It reflects the true performance of the process we are measuring

- **Measurement Error**
  - The errors brought in by measurement system

- **Observed Value**
  - The observed/measured value obtained by the measurement system

Measurement Errors

- Types of Observed Values:
  - Continuous measurements
    - Weight
    - Height
    - Money
  - Discrete measurements
    - Red/Yellow/Green
    - Yes/No
    - Ratings of 1–10

- A *variable MSA* is designed for continuous measurements and an *attribute MSA* is for discrete measurements.

Measurement Errors

- Sources of measurement errors:
  - Human
  - Environment
  - Equipment
  - Sample
  - Process
  - Material
  - Method.

- Fishbone diagrams can help to brainstorm the potential factors affecting the measurement system.

Measurement Errors

- The more errors the measurement system brings in, the less reliable the observed values are.
- A valid measurement system brings in minimum amount of measurement errors.
- The goal of MSA is to qualify the measurement system by quantitatively analyzing its characteristics.

Characteristics of a Measurement System

- Any measurement systems can be characterized by two aspects:
  - Accuracy (location related)
  - Precision (variation related).

- A valid measurement system is *both* accurate and precise.
  - Being accurate does not guarantee the measurement system is precise.
  - Being precise does not guarantee the measurement system is accurate.

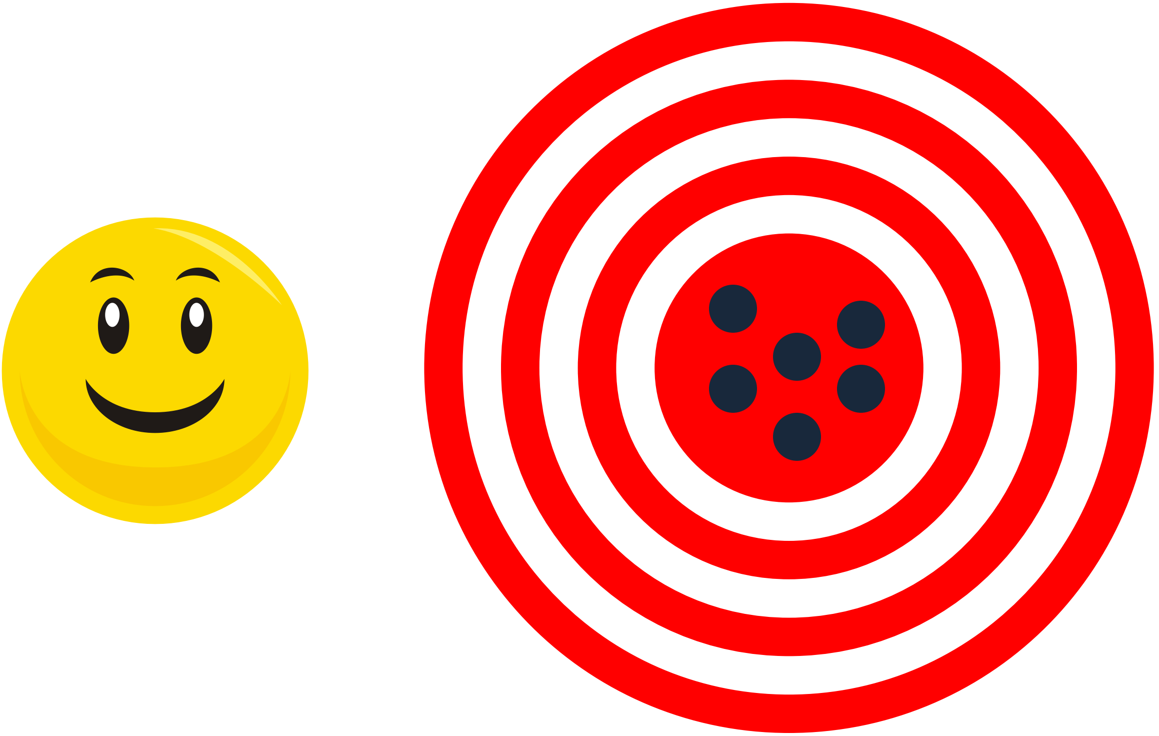
Accuracy vs. Precision

- Accuracy:
  - The level of closeness between the average observed value and the true value
  - How well the observed value reflects the true value.

- Precision:
  - The spread of measurement values
  - How consistent the repeated measurements deliver the same values under the same circumstances.

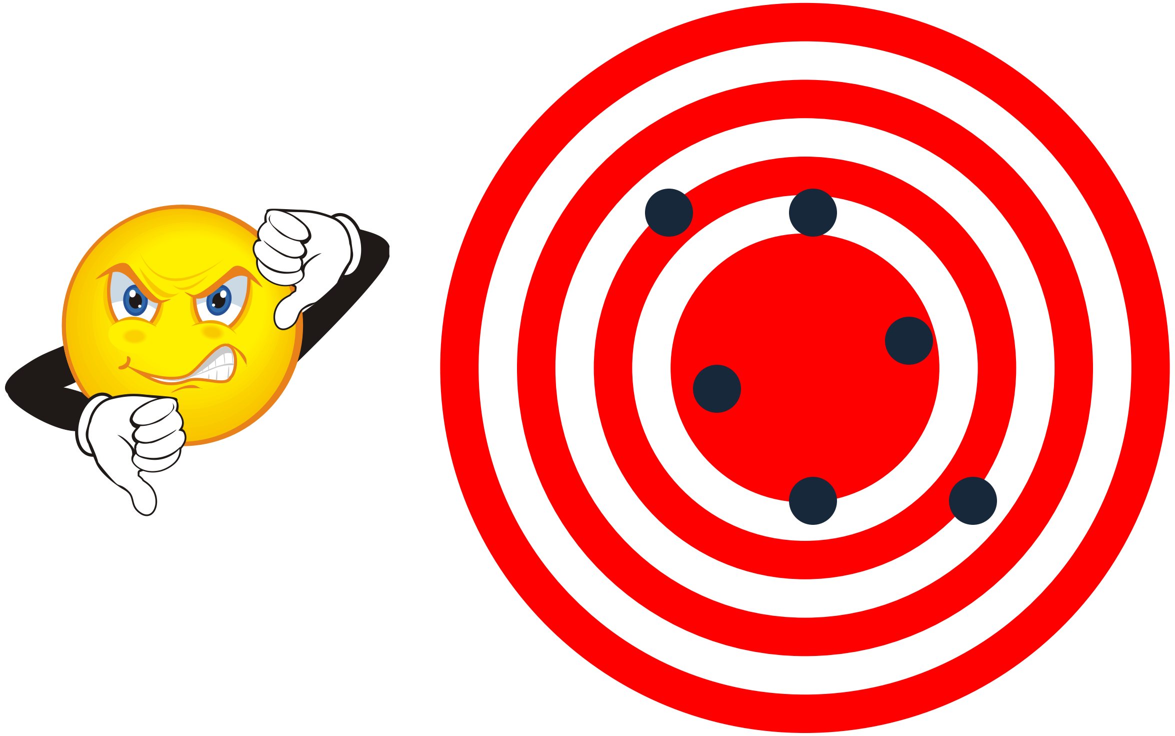
Accuracy vs. Precision

- Accurate and precise
  - high accuracy and high precision



Accuracy vs. Precision

- Accurate and not precise
  - high accuracy and low precision



Accuracy vs. Precision

- Precise and not accurate
  - high precision and low accuracy

Accuracy vs. Precision

- Not accurate and not precise
  - low accuracy and low precision

MSA Conclusions

- If the measurement system is considered *both* accurate and precise, we can start the data-based analysis or decision making.

- If the measurement system is either not accurate or not precise, we need to identify the factor(s) affecting it and calibrate the measurement system until it is both accurate and precise.

Stratification of Accuracy and Precision

- Accuracy
  - Bias
  - Linearity
  - Stability

- Precision
  - Repeatability
  - Reproducibility

2.3.2 Bias, Linearity and Stability

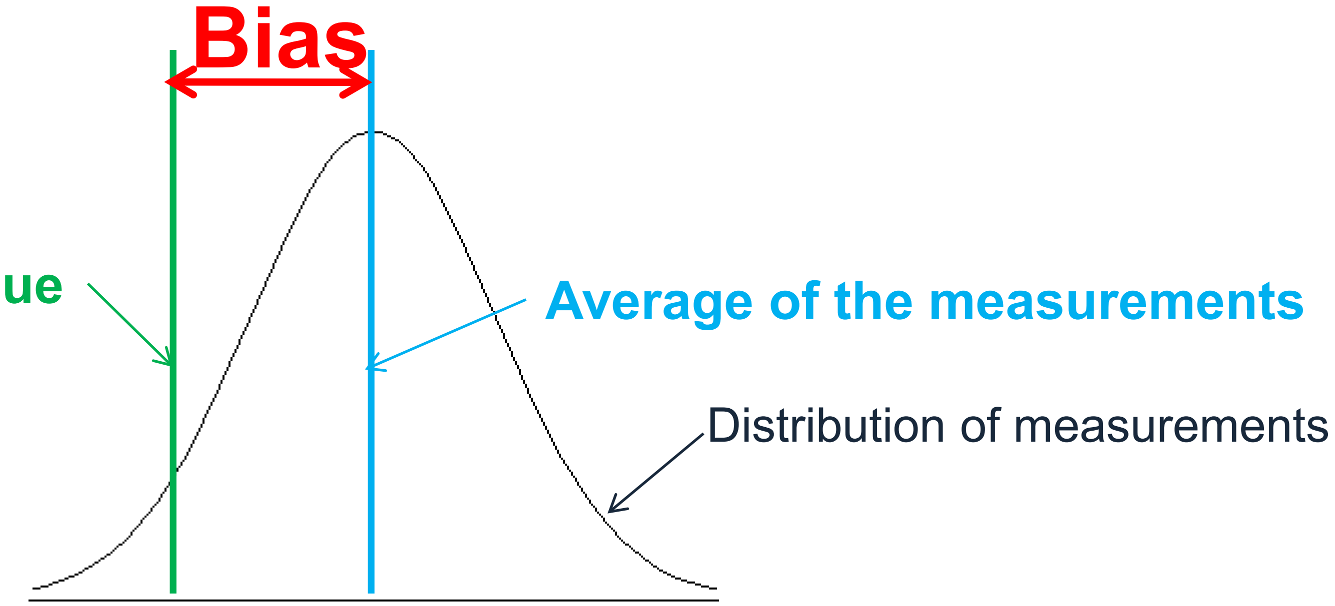
Bias

- **Bias** is the difference between the observed value and the true value of a parameter or metric being measured.
- It is calculated by subtracting the reference value from the average value of the measurements.

Bias = Grand Mean – Reference Value

where the reference value is a standard agreed upon

Bias



- The closer the average of all measurements is to the reference level, the smaller the bias.
- The reference level is the average of measurements of the same items using the master or standard instrument.

Bias

- To determine whether the difference between the average of observed measurement and the reference value is statistically significant (we will explain more details about statistical significance in the Analyze module), we can either conduct a *hypothesis testing* or *compare* the reference value against the confidence intervals of the average measurements.

- If the reference value falls into the confidence intervals, the bias is not statistically significant and can be ignored. Otherwise, the bias is statistically significant and must be fixed.

Bias

- Potential causes of bias:
  - Errors in measuring the reference value
  - Lack of proper training for appraisers
  - Damaged equipment or instrument
  - Measurement instrument not calibrated precisely
  - Appraisers read the data incorrectly.

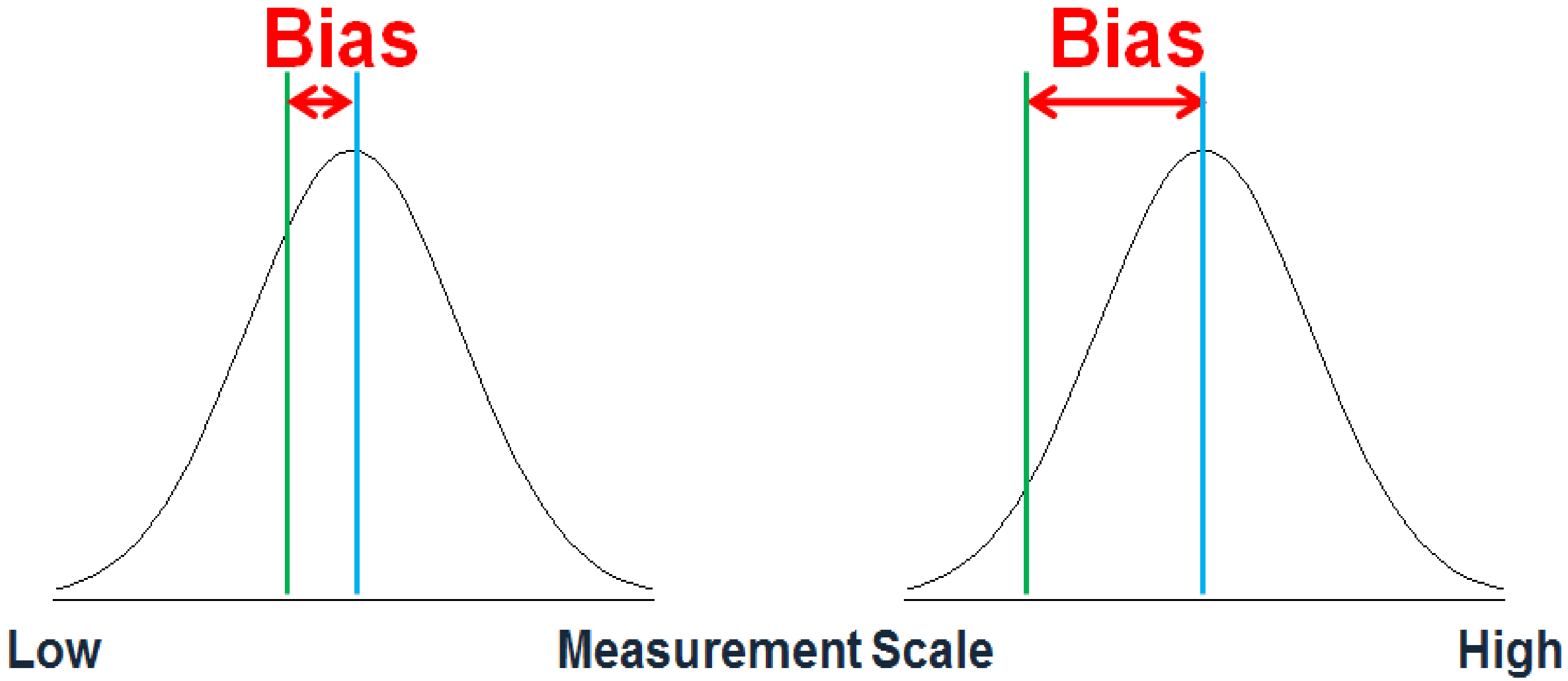
Linearity

- **Linearity** is the degree of the consistency of bias over the entire expected measurement range.

- It quantifies how the bias changes over the range of measurement.

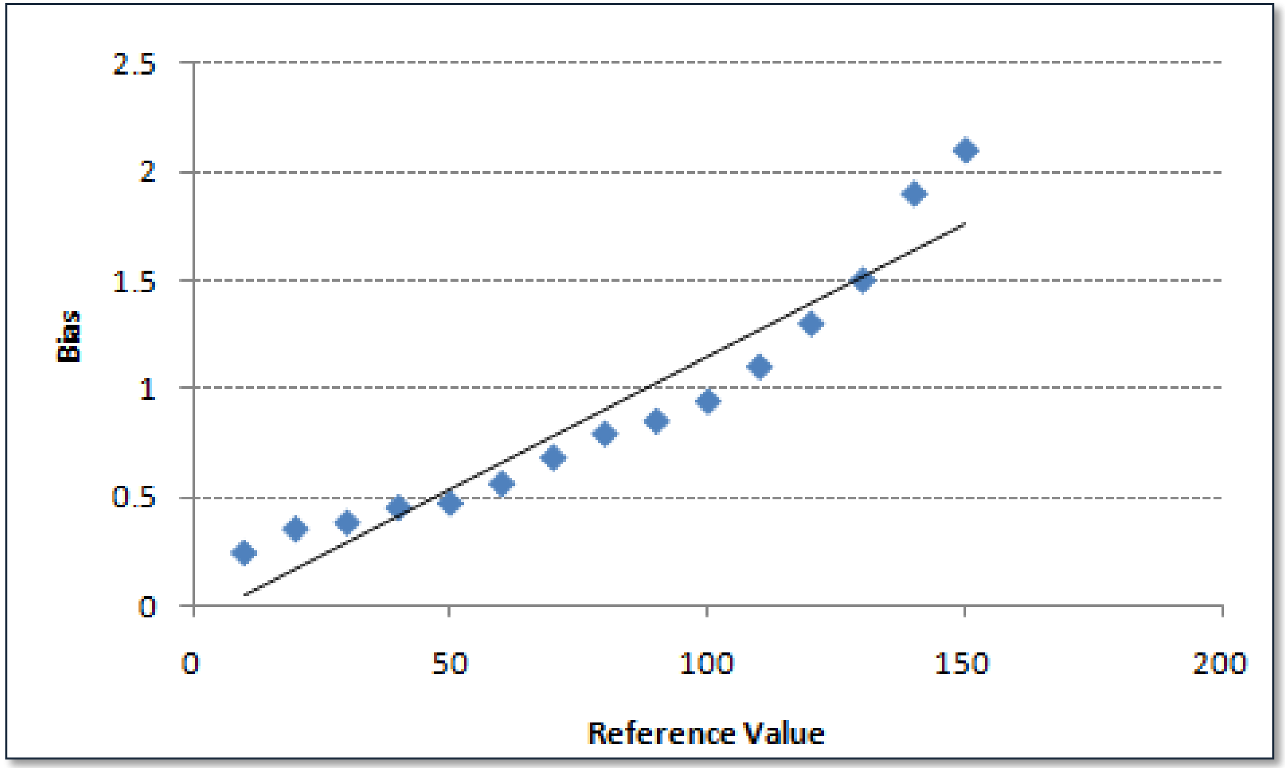
- For example, a scale is off by 0.01 pounds when measuring an object of 10 pounds. However, it is off by 10 pounds when measuring an object of 100 pounds. The scale’s bias is not linear.

Linearity



Linearity

- Create a scatter plot for bias (Y-axis) and reference level (X-axis).
- Find a best fit linear regression line and compute the slope of the line.
- The closer the slope is to zero, the better the measurement system performs.



Linearity

- Formula of the linearity of a measurement system:

Linearity = |Slope| × Process Variation

where

$$Slope = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n}\left(x_i^2\right) - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}$$

x_i is the reference value;
y_i is the bias at each reference level;
n is the sample size.

Linearity

- Potential causes of linearity:
    - Errors in measuring the lower end or higher end of the reference value
    - Lack of proper training for appraisers
    - Damaged equipment or instrument
    - Measurement instrument not calibrated correctly at the lower or higher end of the measurement scale
    - Innate nature of the instrument.

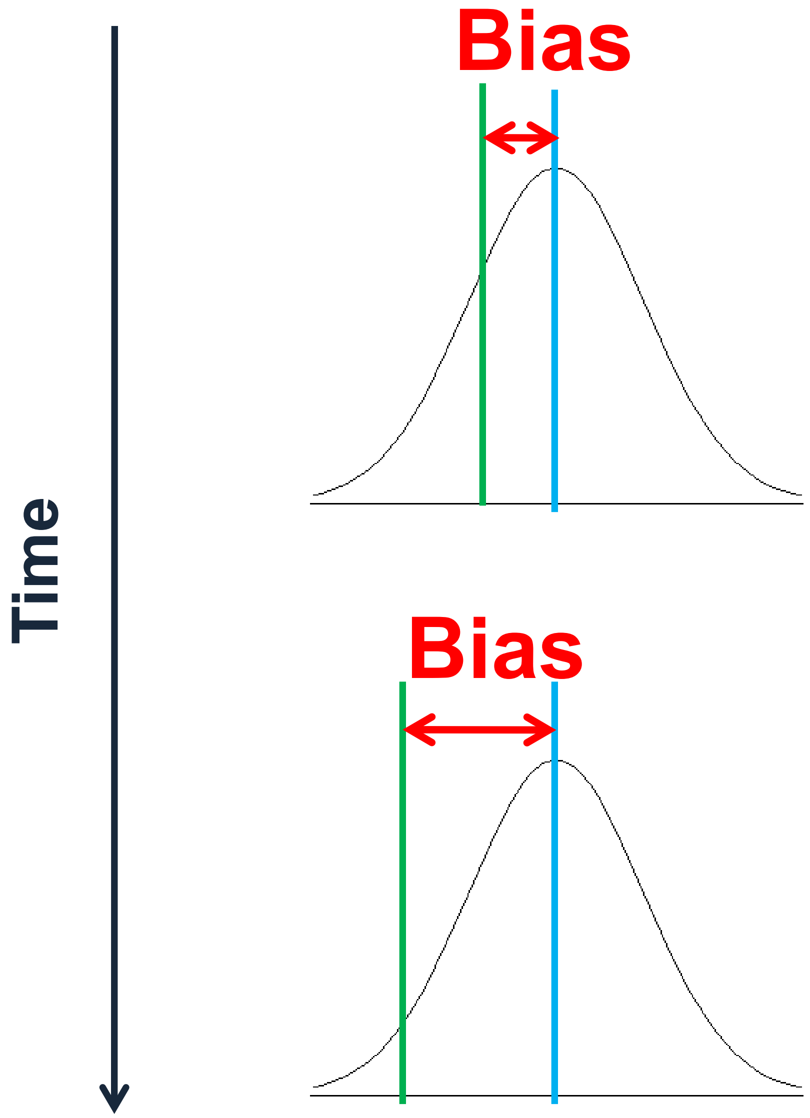
Stability

- **Stability** is the consistency level needed to obtain the same values when measuring the same objects over an extended period of time.

- A measurement system that has low bias and linearity close to zero but cannot consistently perform well would not deliver reliable data.

- Stability is evaluated using control charts.

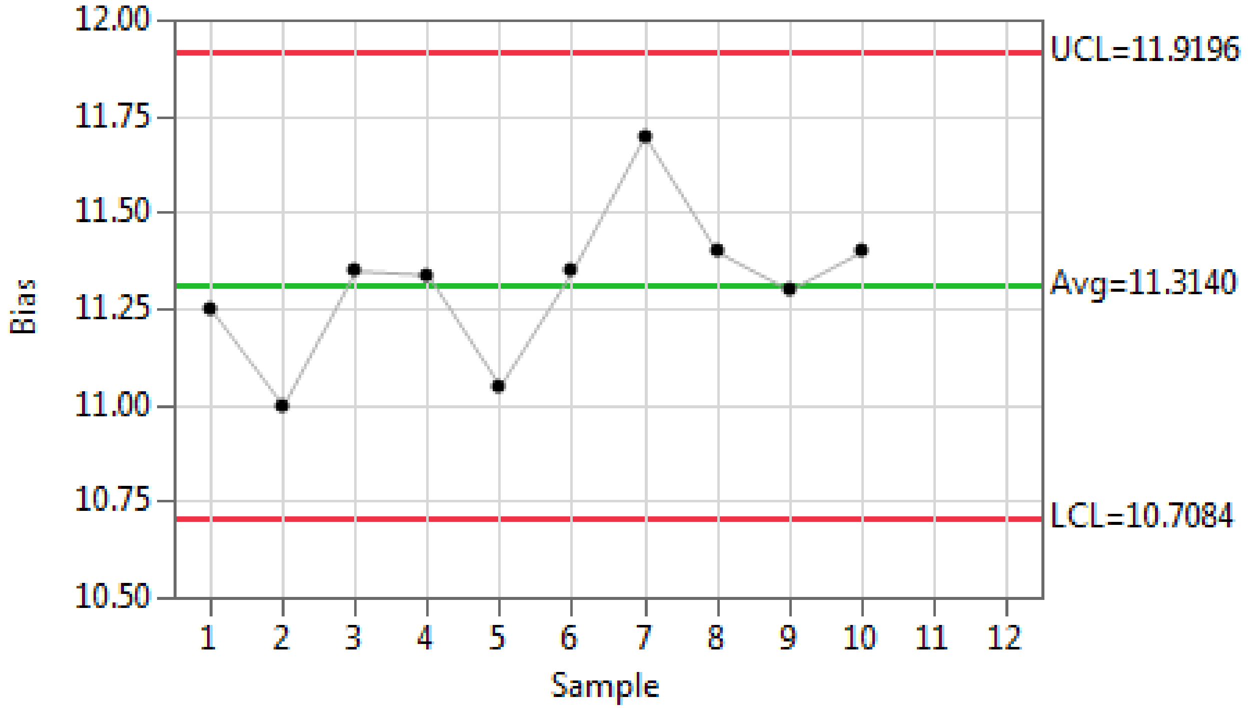
Stability



Stability

- Control charts are used to evaluate the stability of a measurement system.

- When there are no data points out of control, the measurement system is considered stable.



Stability

- Potential causes of instability:
  - Inconsistent training for appraisers
  - Damaged equipment or instrument
  - Worn equipment or instrument
  - Measurement instrument not calibrated
  - Appraisers do not follow the procedure consistently.

2.3.3 Gage R&R

Repeatability

- **Repeatability** evaluates whether the same appraiser can obtain the same value multiple times when measuring the same object using the same equipment under the same environment.

- It refers to the level of agreement between the repeated measurements of the same appraiser under the same condition.

- Repeatability measures the inherent variation of the measurement instrument.

Reproducibility

- **Reproducibility** evaluates whether different appraisers can obtain the same value when measuring the same object independently.

- It refers to the level of agreement between different appraisers.

- It is not caused by the inherent variation of the measurement instrument. It reflects the variability caused by different appraisers, locations, gauges, environments etc.

Gauge R&R

- Gauge R&R (i.e. Gauge Repeatability & Reproducibility) is a method to analyze the variability of a measurement system by partitioning the variation of the measurements using ANOVA (Analysis of Variance).
- Gauge R&R only addresses the precision of a measurement system.

Gauge R&R

- Data collection of a gauge R&R study:
    - let k appraisers measure n random samples independently and repeat the process p times.

- Different appraisers perform the measurement independently.

- The order of measurement (e.g., sequence of samples and sequence of appraisers) is randomized.

Gauge R&R

- The potential sources of variance in the measurement:
  - Appraisers: $\sigma^2_{appraisers}$
  - Parts: σ^2_{parts}
  - Appraisers × Parts: $\sigma^2_{appraisers \times parts}$
  - Repeatability: $\sigma^2_{repeatability}$

- Variance Components

$$\sigma^2_{total} = \sigma^2_{appraisers} + \sigma^2_{parts} + \sigma^2_{appraisers \times parts} + \sigma^2_{repeatability}$$

Gauge R&R

- A valid measurement system has low variability in both repeatability and reproducibility so that the total variability observed can reflect the true variability in the objects (parts) being measured.

$$\sigma^2_{total} = \sigma^2_{reproducibility} + \sigma^2_{repeatability} + \sigma^2_{parts}$$

where

$$\sigma^2_{reproducibility} = \sigma^2_{appraisers} + \sigma^2_{appraisers \times parts}$$

- Gauge R&R variance reflects the precision level of the measurement system.

$$\sigma^2_{R\&R} = \sigma^2_{repeatability} + \sigma^2_{reproducibility}$$

Gauge R&R

- Variation Components

$$Variation_{total} = Z_0 \times \sigma_{total}$$

$$Variation_{repeatability} = Z_0 \times \sigma_{repeatability}$$

$$Variation_{reproducibility} = Z_0 \times \sigma_{reproducibility}$$

$$Variation_{parts} = Z_0 \times \sigma_{parts}$$

where

$$\sigma^2_{total} = \sigma^2_{reproducibility} + \sigma^2_{repeatability} + \sigma^2_{parts}$$

Z_0 is a sigma multiplier that assumes a specific confidence level in the spread of the data.

Gauge R&R

- The percentage of variation R&R contributes to the total variation in the measurement:

$$Contribution\%_{R\&R} = \frac{Variation_{R\&R}}{Variation_{total}} \times 100\%$$

where $$Variation_{R\&R} = Z_0 \times \sqrt{\sigma^2_{repeatability} + \sigma^2_{reproducibility}}$$

Measurement System	% Study Var.	% Contribution	Distinct Categories
Acceptable	10% or less	1% or Less	5 or Greater
Marginal	10% - 30%	1% - 9%	
Unacceptable	30% or Greater	9% or Greater	Less than 5

2.3.4 Variable and Attribute MSA

Variable Gage R&R

- Whenever something is measured repeatedly or by different people or processes, the results of the measurements will vary. Variation comes from two primary sources:
  1. Differences between the parts being measured
  2. The measurement system.

- We can use a gage R&R to conduct a measurement system analysis to determine what portion of the variability comes from the parts and what portion comes from the measurement system.

- There are key study results that help us determine the components of variation within our measurement system.

Key Measures of a Variable Gage R&R

- %Contribution: The percent of contribution for a source is 100 times the variance component for that source divided by the total variance.

- %Study Var (6*SD): The percent of study variation for a source is 100 times the study variation for that source divided by the total variation.

- %Tolerance (SV/Tolerance): The percent of spec range taken up by the total width of the distribution of the data based on variation from that source.

- Distinct Categories: The number of distinct categories of parts that the measurement system is able to distinguish. If a measurement system is not capable of distinguishing at least five types of parts, it is probably not adequate.

Variable Gage R&R Guidelines (AIAG)

- **Percent Tolerance and Percent Study Variation**
  - 10% or less – Acceptable
  - 10% to 30% – Marginal
  - 30% or greater – Unacceptable

- **Percent Contribution**
  - 1% or less – Acceptable
  - 1% to 9% – Marginal
  - 9% or greater – Unacceptable

- **Distinct Categories**
  - Look for five or more distinct categories to indicate that your measurement system is acceptable.

Guidelines for Distinct Categories

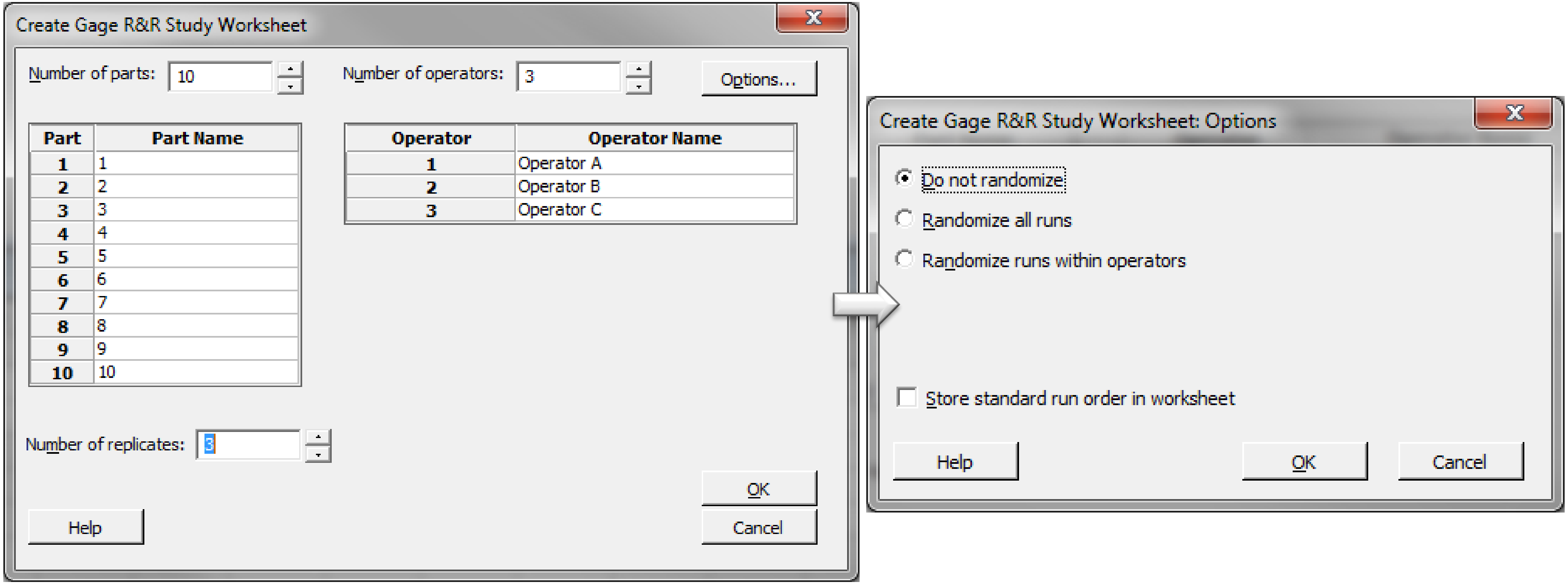
- Distinct categories is the number of categories of parts that your measurement system can distinguish. If it is below five, it is likely not able to distinguish between parts.

Number of Categories	Conclusion
Distinct Categories = 1	Measurement system cannot discriminate between parts
Distinct Categories = 2	Measurement system can only distinguish between high/low or big/small
Distinct Categories = 3 or 4	Measurement system is of little or no value
Distinct Categories = 5+	According to AIAG, the measurement system can acceptably discriminate parts

Use Minitab to Implement a Variable MSA

- Data File: “Variable MSA” tab in “Sample Data.xlsx”
- Step 1: Initiate the MSA study
    1) Click on Stat → Quality Tools → Gage Study → Create Gage R&R Study Worksheet.
    2) A new window named “Create Gage R&R Study Worksheet” pops up.
    3) Select 10 as the “Number of Parts.”
    4) Select 3 as the “Number of Operators.”
    5) Select 3 as the “Number of Replicates.”
    6) Enter the part name (e.g., Part 01, Part 02, Part 03, …).
    7) Enter the operator name (e.g., Operator A, Operator B, Operator C).
    8) Click on the “Options” button and another window named “Create Gage R&R Study Worksheet – Options” pops up.
    9) Select the radio button “Do not randomize.”
    10) Click “OK” in the window “Create Gage R&R Study Worksheet – Options.”
    11) Click “OK’ in the window “Create Gage R&R Study Worksheet.”
    12) A new data table is generated.

Use Minitab to Implement a Variable MSA



Number of parts: 10

Number of operators: 3

Part	Part Name
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10

Operator	Operator Name
1	Operator A
2	Operator B
3	Operator C

Number of replicates: 3

● Do not randomize

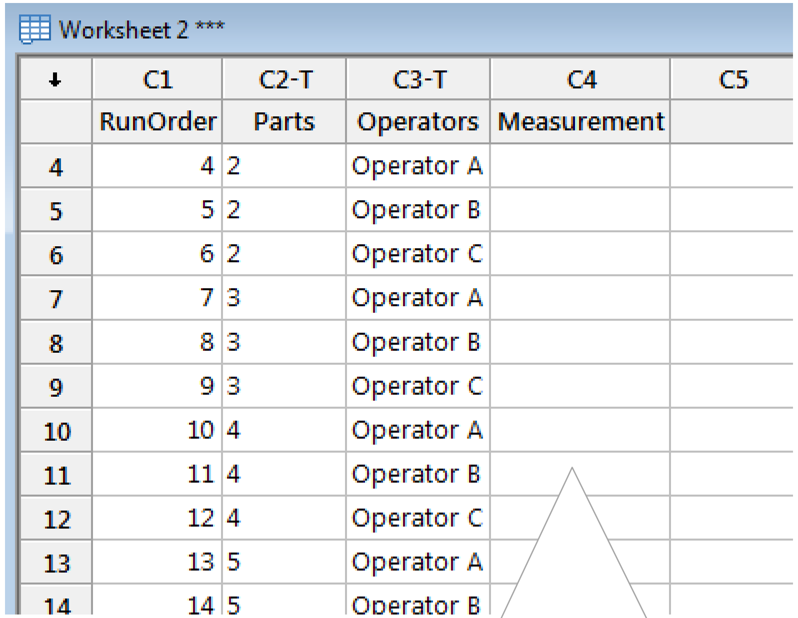
○ Randomize all runs

○ Randomize runs within operators

☐ Store standard run order in worksheet

Use Minitab to Implement a Variable MSA

- Step 2: Data collection
  - In the newly-generated data table, Minitab has provided the data layout where we can add our actual measurement data when it's collected.
  - When you actually conduct a variable MSA in your work environment it would be necessary to set up your study just as we have in the previous steps and then you could collect your measurement data properly.
  - However, for our purposes today, we have provided you with an MSA that is setup and data collected. We will use our "Variable MSA" tab in "Sample Data.xlsx," for the next steps.



↓	C1	C2-T	C3-T	C4	C5
	RunOrder	Parts	Operators	Measurement	
4	4	2	Operator A		
5	5	2	Operator B		
6	6	2	Operator C		
7	7	3	Operator A		
8	8	3	Operator B		
9	9	3	Operator C		
10	10	4	Operator A		
11	11	4	Operator B		
12	12	4	Operator C		
13	13	5	Operator A		
14	14	5	Operator B		

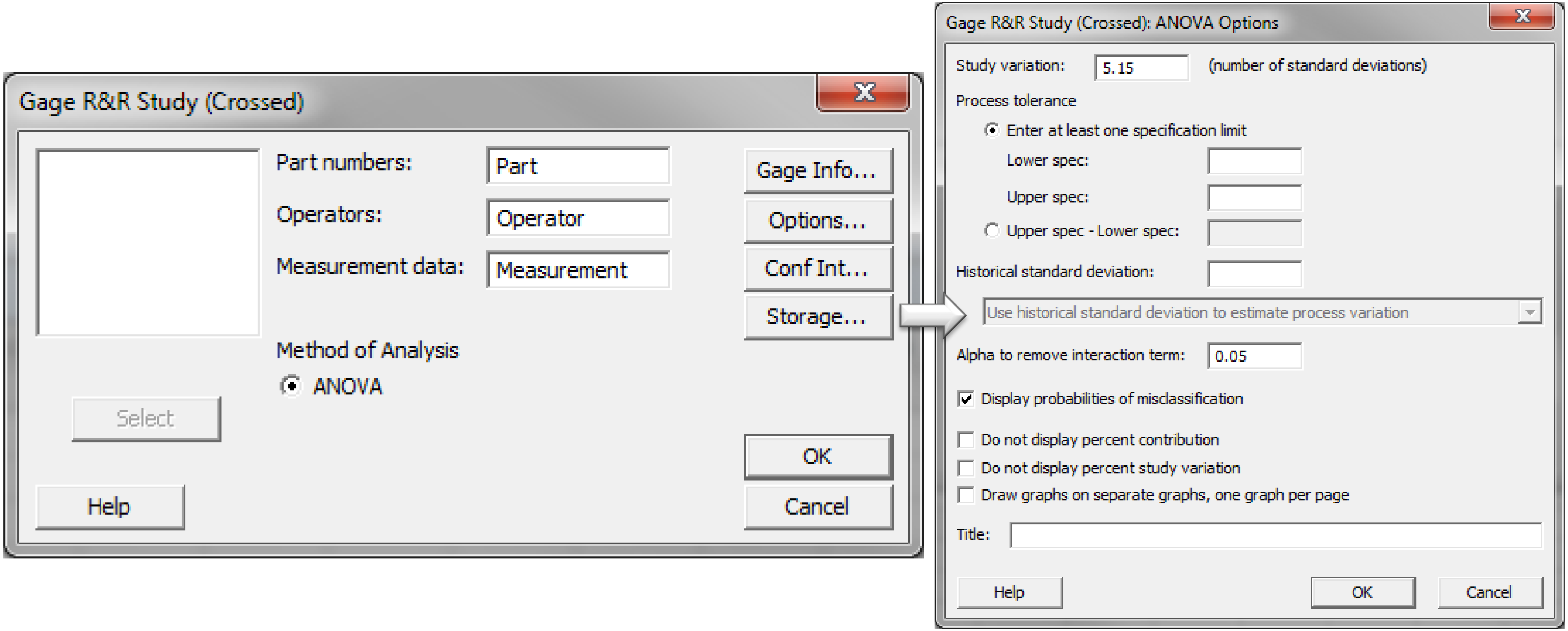
Your data collection plan is now setup for your variable MSA. In real life you would collect your measurement data adding it to this table.

For our purposes today, we have provided you with that data so that we can continue learning how to perform a variable MSA.

Use Minitab to Implement a Variable MSA

- Step 4: Implement Gauge R&R (using our “Variable MSA” data in “Sample Data.xlsx)
    1) Click Stat → Quality Tools → Gage Study → Gage R&R Study (Crossed).
    2) A new window named “Gage R&R Study (Crossed)” appears.
    3) Select “Part” as “Part numbers.”
    4) Select “Operator” as “Operators.”
    5) Select “Measurement” as “Measurement data.”
    6) Click on the “Options” button and another new window named “Gage R&R Study (Crossed) – ANOVA Options” pops up.
    7) Enter 5.15 as the “Study variation (number of standard deviations)”.
    8) Click “OK” in the window “Gage R&R Study (Crossed) – ANOVA Options.”
    9) Click “OK” in the window “Gage R&R Study (Crossed).”
    10) The MSA analysis results appear in the new window and the session window.

Use Minitab to Implement a Variable MSA



Gage R&R Study (Crossed)

Part numbers: Part

Operators: Operator

Measurement data: Measurement

Method of Analysis

● ANOVA

Gage R&R Study (Crossed): ANOVA Options

Study variation: 5.15 (number of standard deviations)

Process tolerance

● Enter at least one specification limit

Lower spec:

Upper spec:

○ Upper spec - Lower spec:

Historical standard deviation:

Use historical standard deviation to estimate process variation

Alpha to remove interaction term: 0.05

☑ Display probabilities of misclassification

☐ Do not display percent contribution

☐ Do not display percent study variation

☐ Draw graphs on separate graphs, one graph per page

Title:

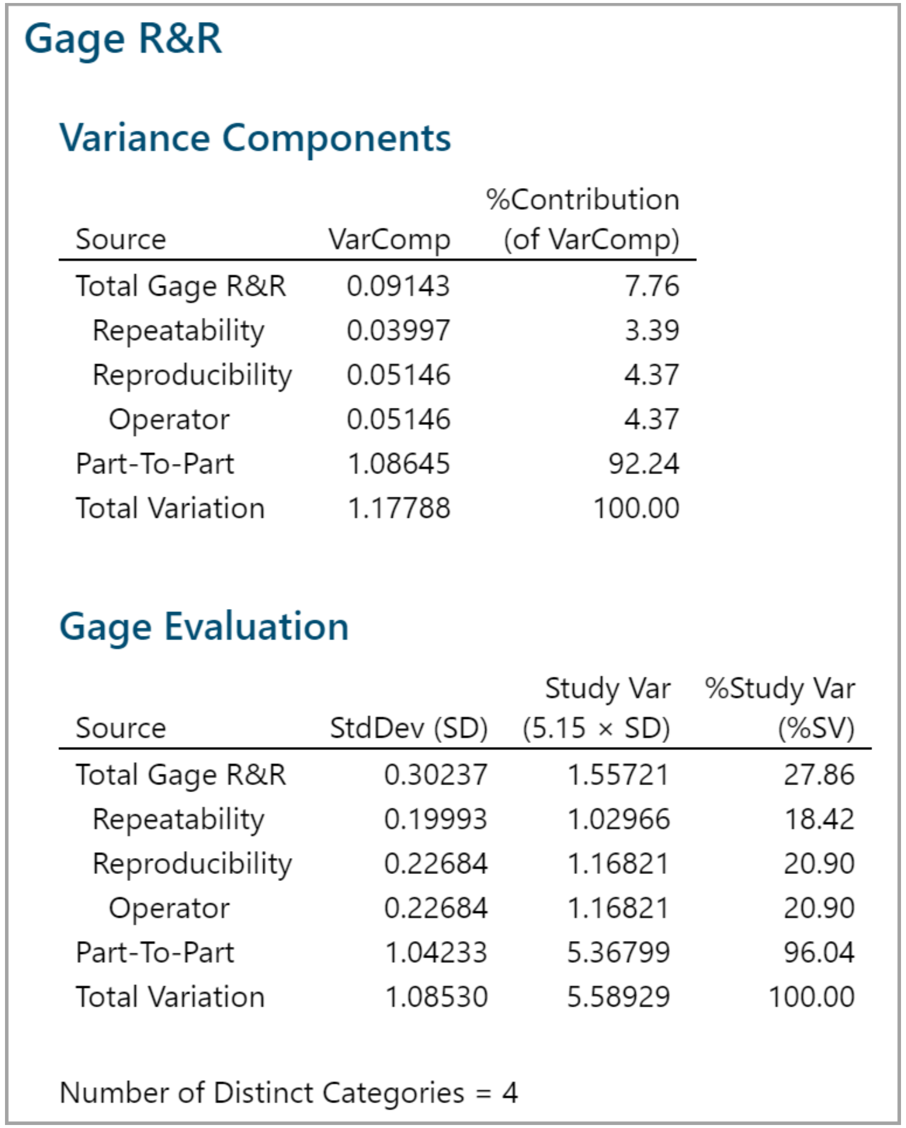
Use Minitab to Implement a Variable MSA

- 5.15 is the recommended standard deviation multiplier by the Automotive Industry Action Group (AIAG). It corresponds to 99% of data in the normal distribution. If we use 6 as the standard deviation multiplier, it corresponds to 99.73% of the data in the normal distribution.

Confidence Level	**Sigma Multiplier**
90%	3.29
95%	3.92
99%	5.15
99.73%	6

Use Minitab to Implement a Variable MSA

- Step 5: Analyze the MSA results
  - The percentage of variation that R&R contributes to the total variation is 27.86% and the precision level of this measurement system is at best “marginal” bordering on unacceptable.
  - Actions should be required to calibrate the measurement system.



Gage R&R

Variance Components

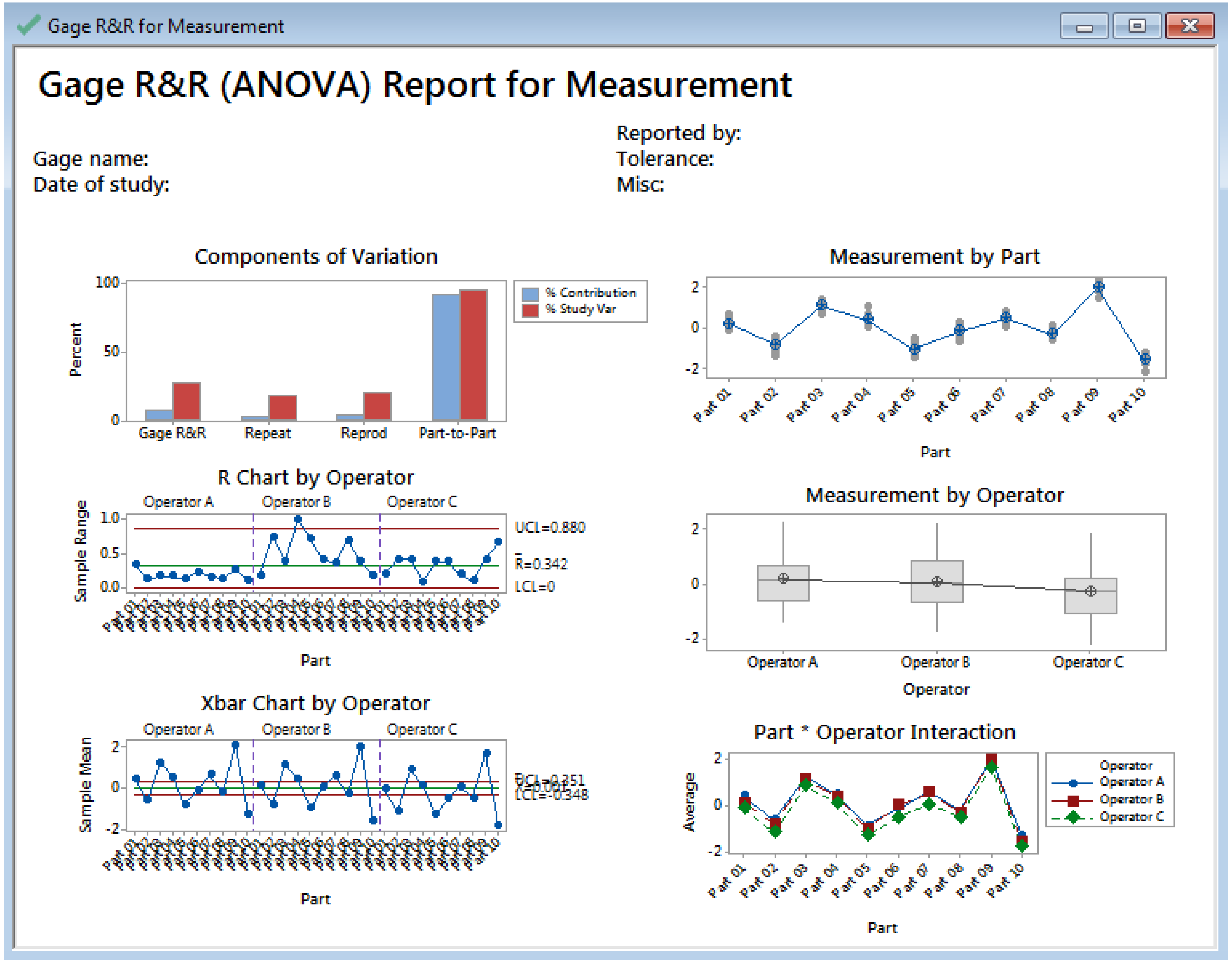
Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.09143	7.76
Repeatability	0.03997	3.39
Reproducibility	0.05146	4.37
Operator	0.05146	4.37
Part-To-Part	1.08645	92.24
Total Variation	1.17788	100.00

Gage Evaluation

Source	StdDev (SD)	Study Var (5.15 × SD)	%Study Var (%SV)
Total Gage R&R	0.30237	1.55721	27.86
Repeatability	0.19993	1.02966	18.42
Reproducibility	0.22684	1.16821	20.90
Operator	0.22684	1.16821	20.90
Part-To-Part	1.04233	5.36799	96.04
Total Variation	1.08530	5.58929	100.00

Number of Distinct Categories = 4

Use Minitab to Implement a Variable MSA



Gage R&R (ANOVA) Report for Measurement

Gage name:
Date of study:

Reported by:
Tolerance:
Misc:

Use Minitab to Implement an Attribute MSA

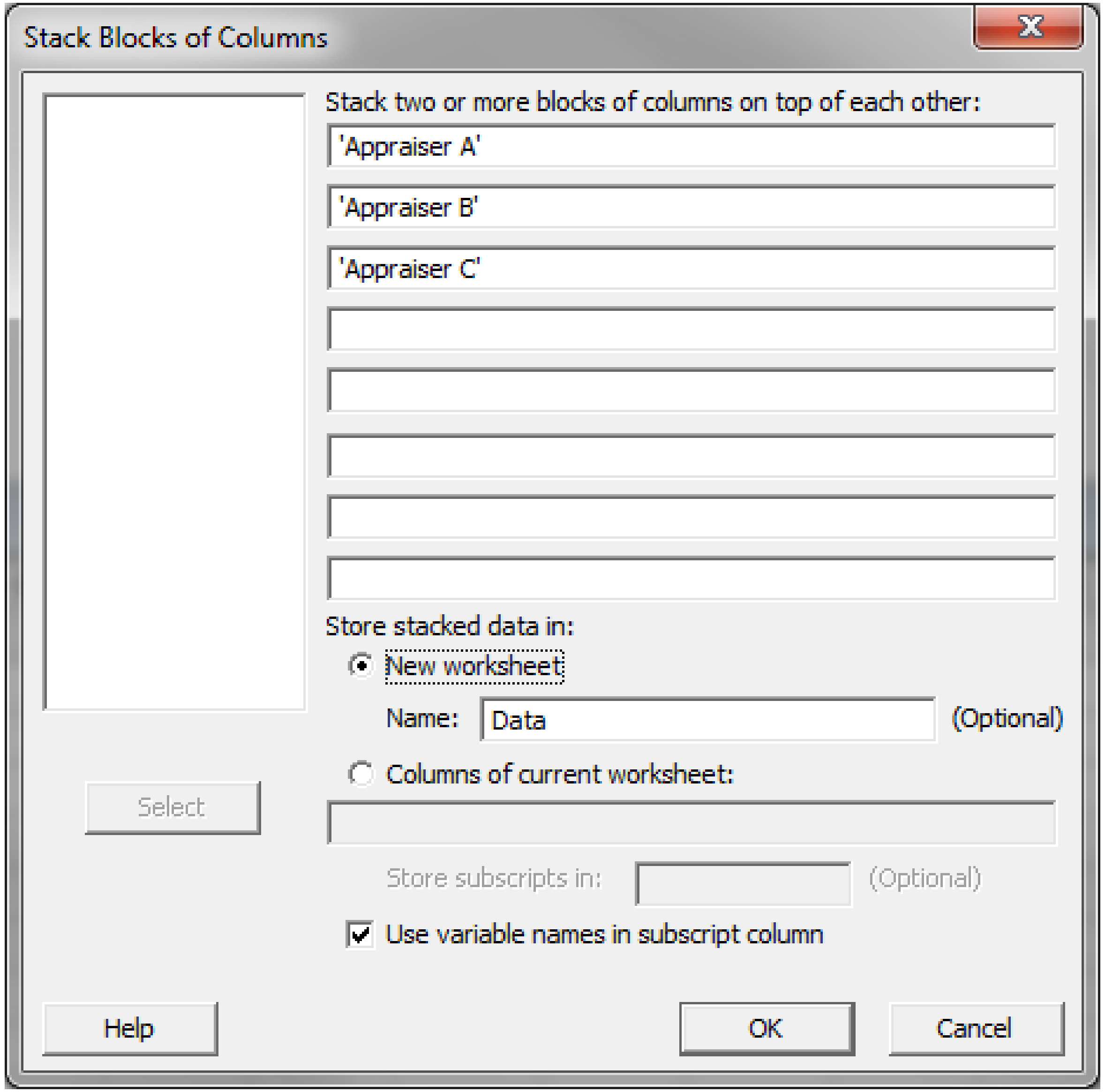
- Data File: “Attribute MSA” tab in “Sample Data.xlsx” (an example in the AIAG MSA Reference Manual, 3rd Edition).

- Steps in Minitab to run an attribute MSA

Use Minitab to Implement an Attribute MSA

- Step 1: Reorganize the original data into four new columns (i.e., Appraiser, Assessed Result, Part, and Reference).
  1) Click Data → Stack → Blocks of Columns.
  2) A new window named “Stack Blocks of Columns” pops up.
  3) Select “Appraiser A,” “Part,” and “Reference” as block one.
  4) Select “Appraiser B,” “Part,” and “Reference” as block two.
  5) Select “Appraiser C,” “Part,” and “Reference” as block three.
  6) Select the radio button of “New worksheet” and name the sheet “Data.”
  7) Check the box “Use variable names in subscript column.”
  8) Click “OK.”
  9) The stacked columns are created in the new worksheet named “Data.”
  10) Name the four columns from left to right in worksheet “Data”: Appraiser, Assessed Result, Part, and Reference.

Use Minitab to Implement an Attribute MSA



Stack Blocks of Columns

Stack two or more blocks of columns on top of each other:

'Appraiser A'

'Appraiser B'

'Appraiser C'

Store stacked data in:

● New worksheet

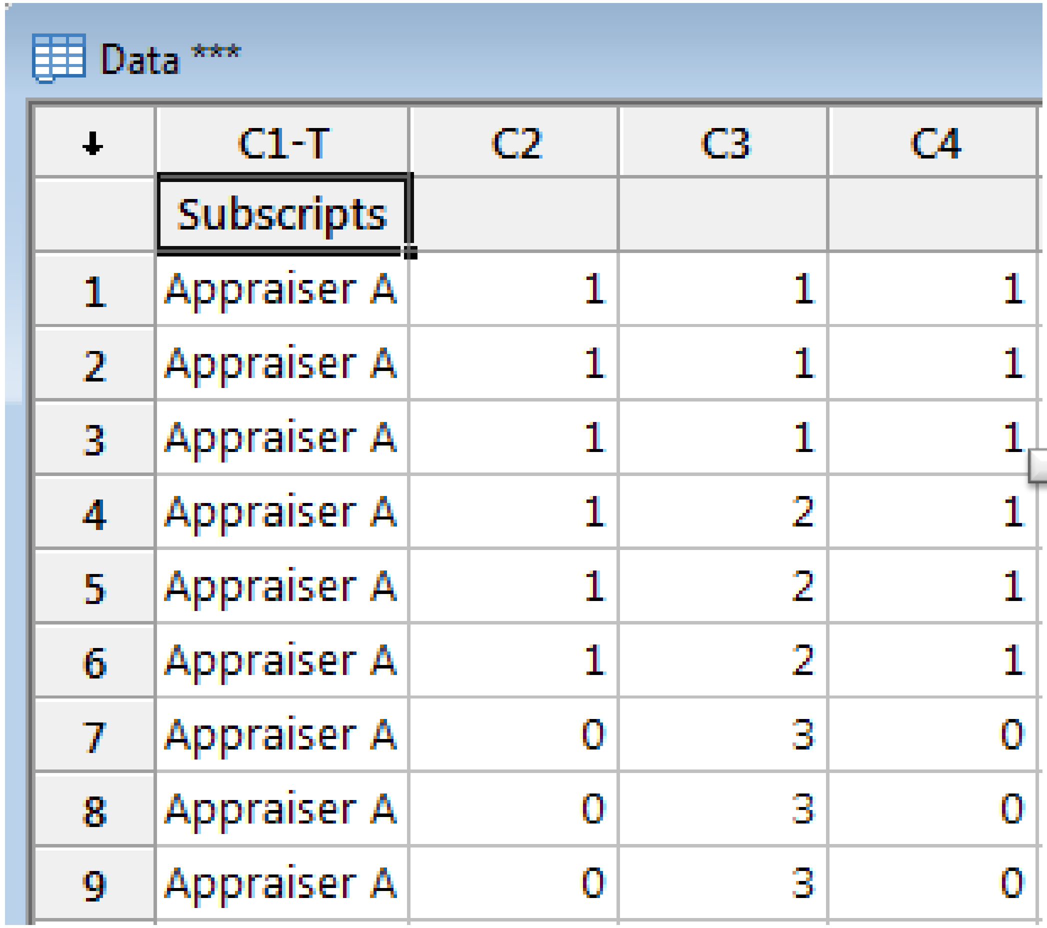
Name: Data (Optional)

○ Columns of current worksheet:

Store subscripts in: (Optional)

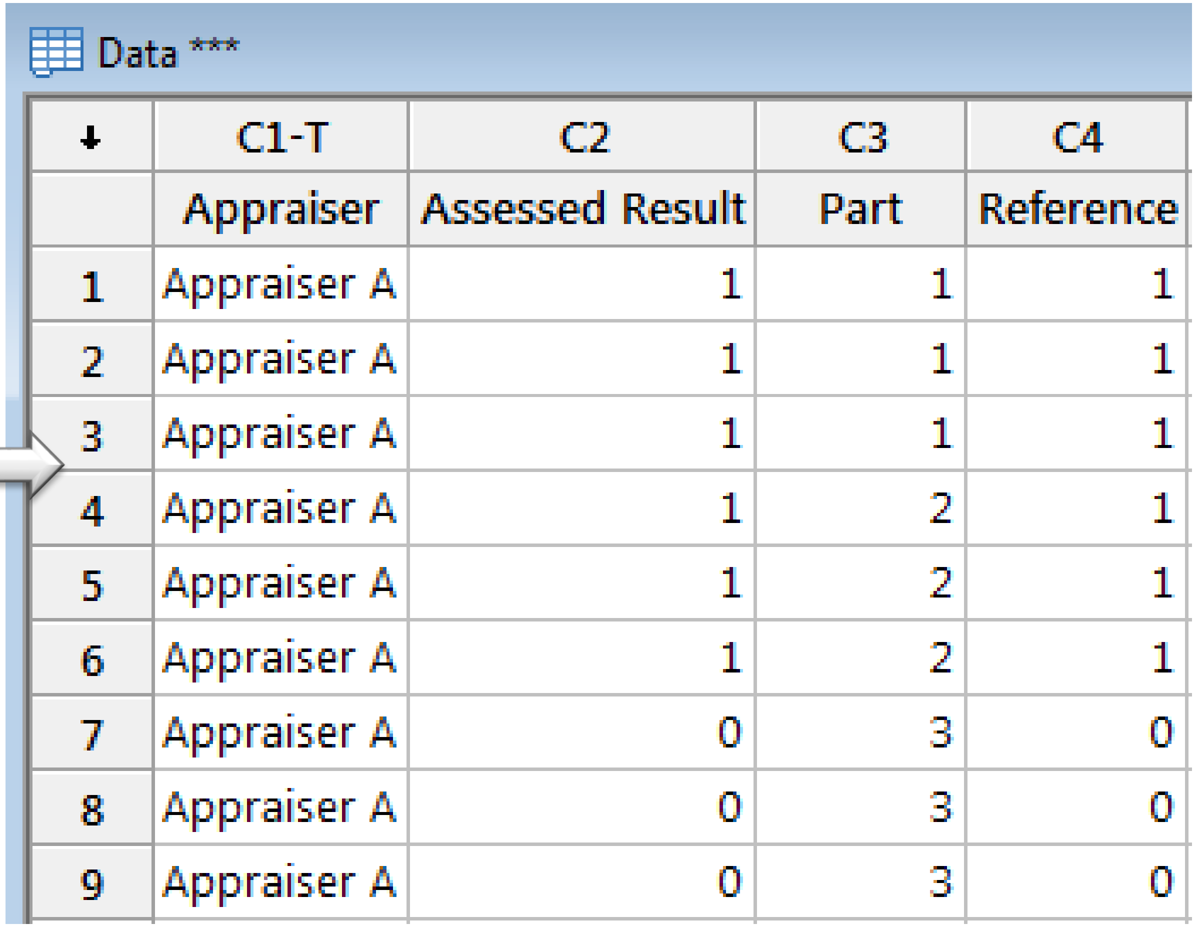
☑ Use variable names in subscript column

Use Minitab to Implement an Attribute MSA



Data ***

↓	C1-T	C2	C3	C4
	Subscripts			
1	Appraiser A	1	1	1
2	Appraiser A	1	1	1
3	Appraiser A	1	1	1
4	Appraiser A	1	2	1
5	Appraiser A	1	2	1
6	Appraiser A	1	2	1
7	Appraiser A	0	3	0
8	Appraiser A	0	3	0
9	Appraiser A	0	3	0



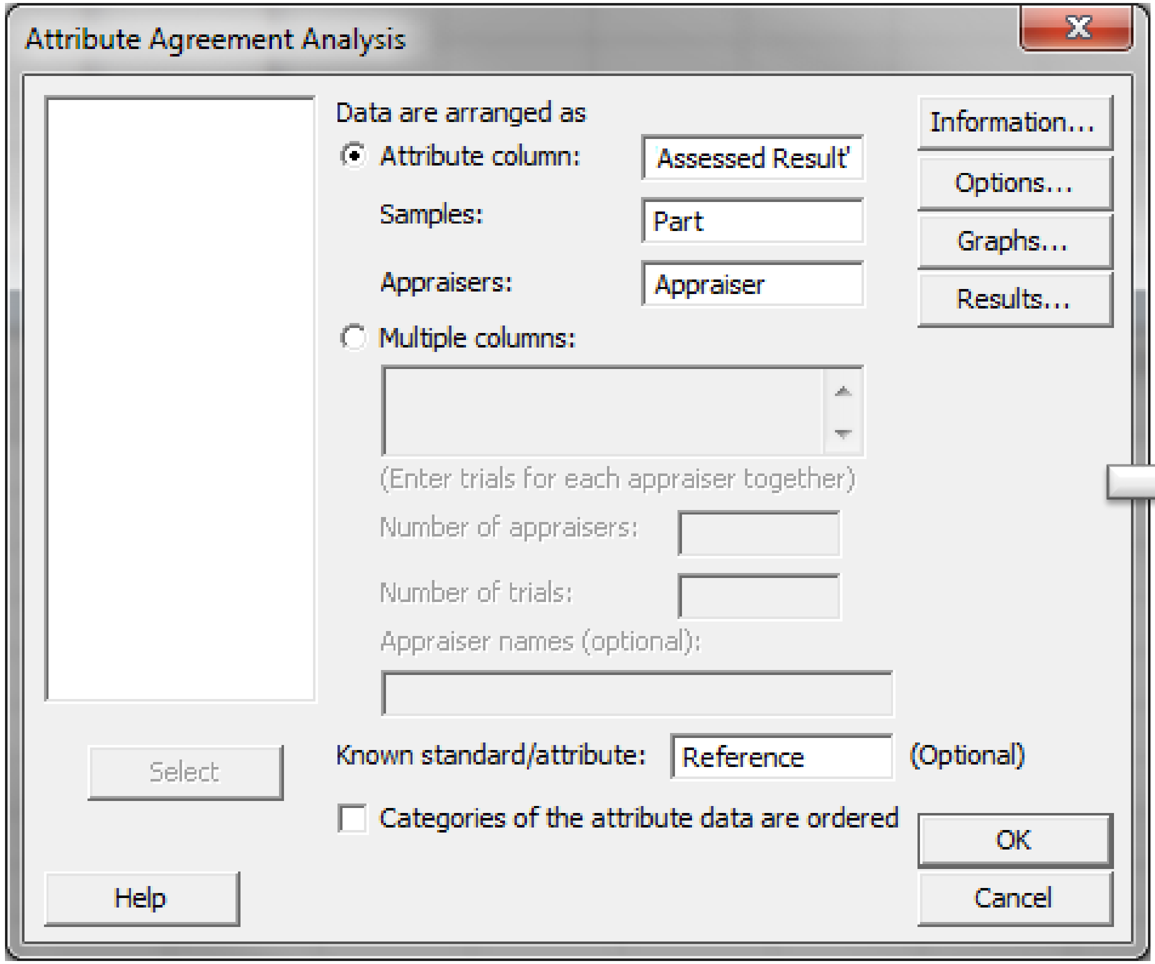
Data ***

↓	C1-T	C2	C3	C4
	Appraiser	Assessed Result	Part	Reference
1	Appraiser A	1	1	1
2	Appraiser A	1	1	1
3	Appraiser A	1	1	1
4	Appraiser A	1	2	1
5	Appraiser A	1	2	1
6	Appraiser A	1	2	1
7	Appraiser A	0	3	0
8	Appraiser A	0	3	0
9	Appraiser A	0	3	0

Use Minitab to Implement an Attribute MSA

- Step 2: Run MSA using Minitab
  1) Click Stat → Quality Tools → Attribute Agreement Analysis.
  2) A new window named “Attribute Agreement Analysis” pops up.
  3) Click in the blank box next to “Attribute column” and the variables appear in the list box on the left.
  4) Select “Assessed Result” as “Attribute column.”
  5) Select “Part” as “Samples.”
  6) Select “Appraiser” as “Appraisers.”
  7) Select “Reference” as “Known standard/attribute.”
  8) Click the “Options” button and another window named “Attribute Agreement Analysis – Options” pops up.
  9) Check the boxes of both “Calculate Cohen’s kappa if appropriate” and “Display disagreement table.”
  10) Click “OK” in the window “Attribute Agreement Analysis – Options.”
  11) Click “OK” in the window “Attribute Agreement Analysis.”
  12) The MSA results appear in the newly-generated window and the session window.

Use Minitab to Implement an Attribute MSA


Attribute Agreement Analysis

Data are arranged as

● Attribute column: Assessed Result'

Samples: Part

Appraisers: Appraiser

○ Multiple columns:

(Enter trials for each appraiser together)

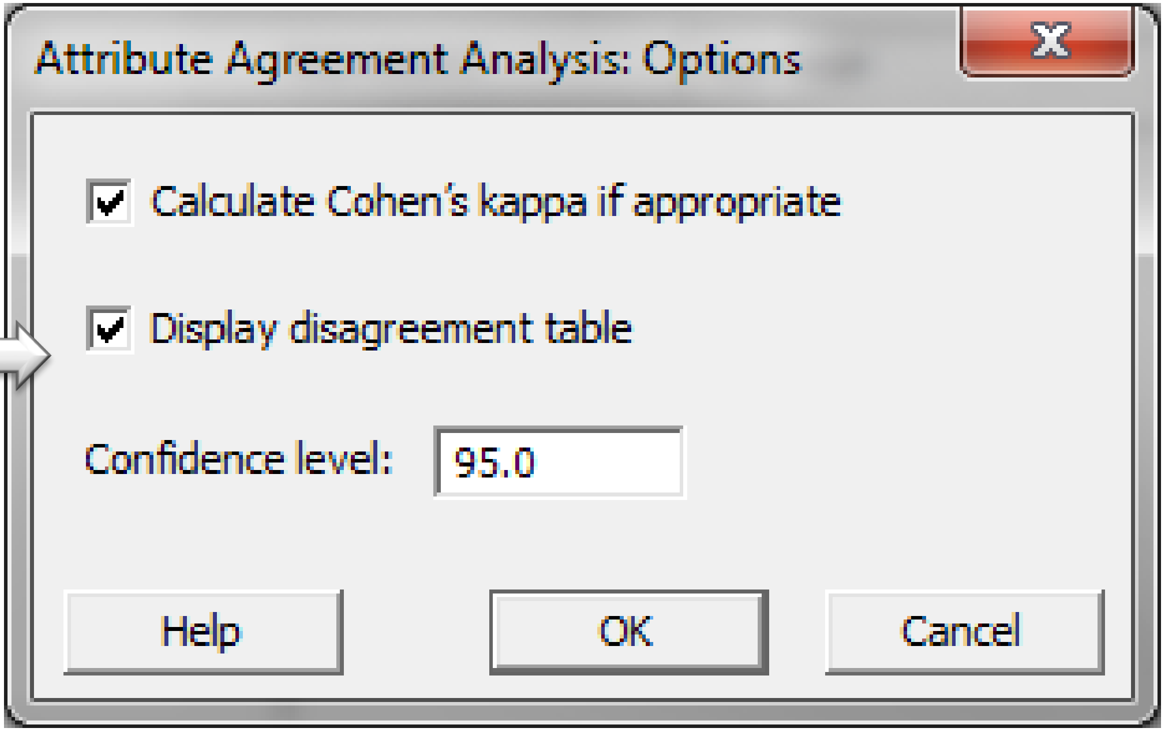
Number of appraisers:

Number of trials:

Appraiser names (optional):

Known standard/attribute: Reference (Optional)

☐ Categories of the attribute data are ordered

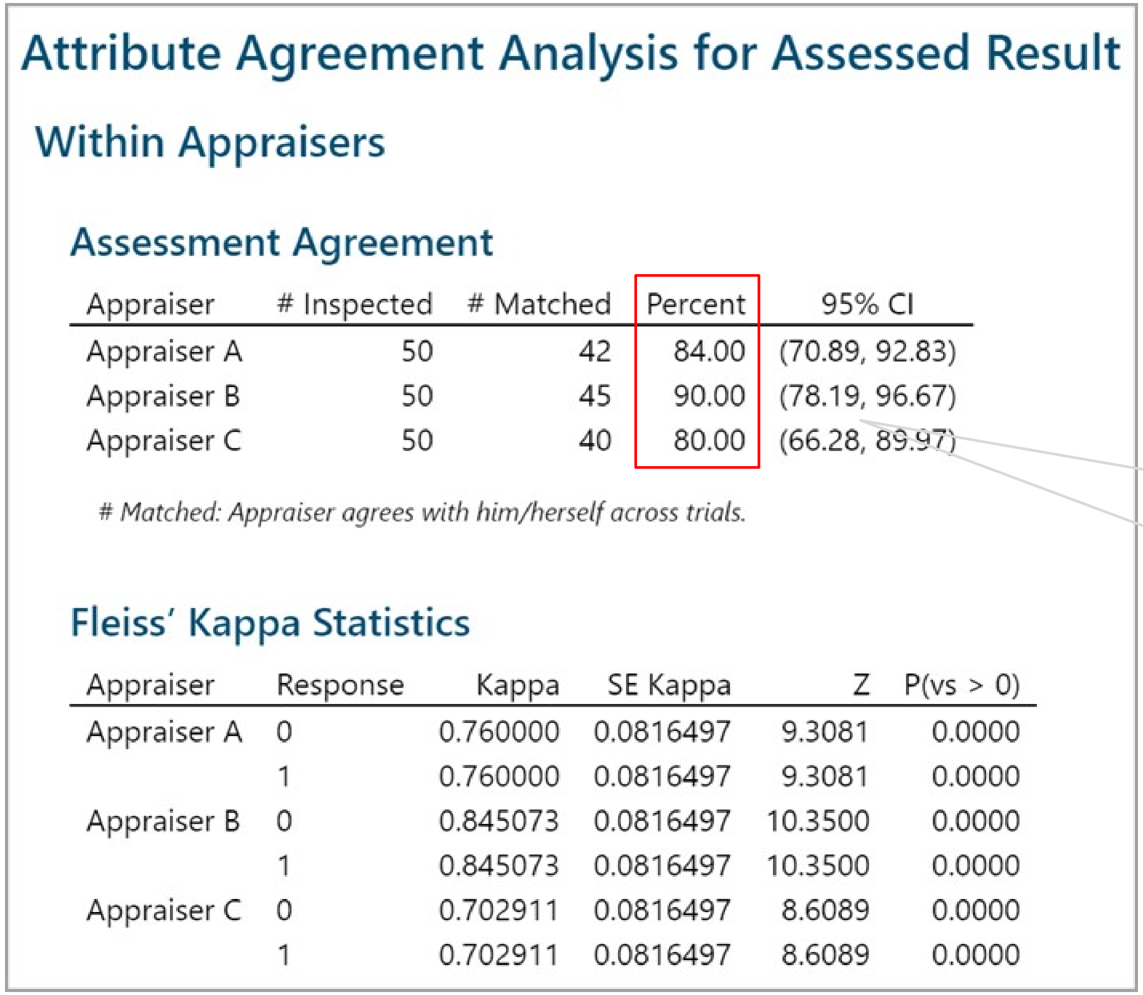

Attribute Agreement Analysis: Options

☑ Calculate Cohen's kappa if appropriate

☑ Display disagreement table

Confidence level: 95.0

Use Minitab to Implement an Attribute MSA



Attribute Agreement Analysis for Assessed Result

Within Appraisers

Assessment Agreement

Appraiser	# Inspected	# Matched	Percent	95% CI
Appraiser A	50	42	84.00	(70.89, 92.83)
Appraiser B	50	45	90.00	(78.19, 96.67)
Appraiser C	50	40	80.00	(66.28, 89.97)

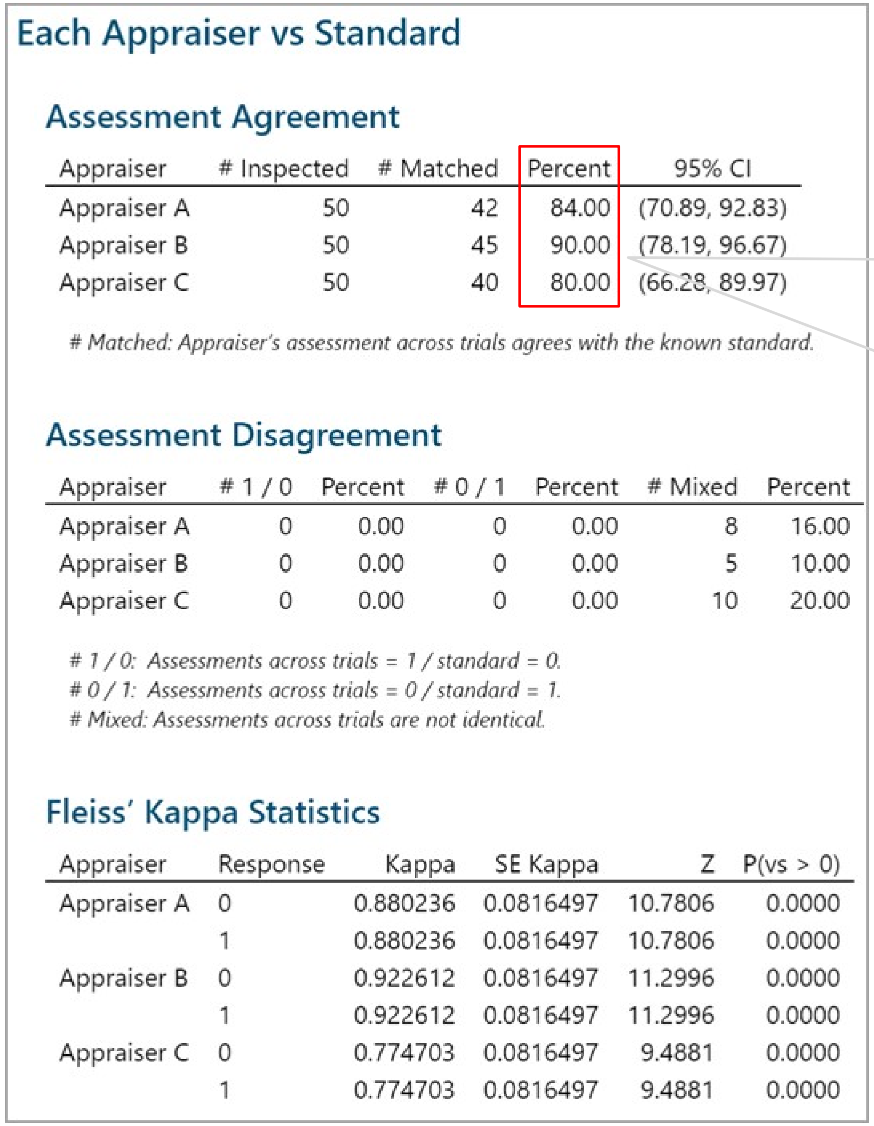
 # Matched: Appraiser agrees with him/herself across trials.

Fleiss' Kappa Statistics

Appraiser	Response	Kappa	SE Kappa	Z	P(vs > 0)
Appraiser A	0	0.760000	0.0816497	9.3081	0.0000
	1	0.760000	0.0816497	9.3081	0.0000
Appraiser B	0	0.845073	0.0816497	10.3500	0.0000
	1	0.845073	0.0816497	10.3500	0.0000
Appraiser C	0	0.702911	0.0816497	8.6089	0.0000
	1	0.702911	0.0816497	8.6089	0.0000

Within Appraiser Agreement Percent: the agreement percentage within each individual appraiser.

Use Minitab to Implement an Attribute MSA



Each Appraiser vs Standard

Assessment Agreement

Appraiser	# Inspected	# Matched	Percent	95% CI
Appraiser A	50	42	84.00	(70.89, 92.83)
Appraiser B	50	45	90.00	(78.19, 96.67)
Appraiser C	50	40	80.00	(66.28, 89.97)

 # Matched: Appraiser's assessment across trials agrees with the known standard.

Assessment Disagreement

Appraiser	# 1 / 0	Percent	# 0 / 1	Percent	# Mixed	Percent
Appraiser A	0	0.00	0	0.00	8	16.00
Appraiser B	0	0.00	0	0.00	5	10.00
Appraiser C	0	0.00	0	0.00	10	20.00

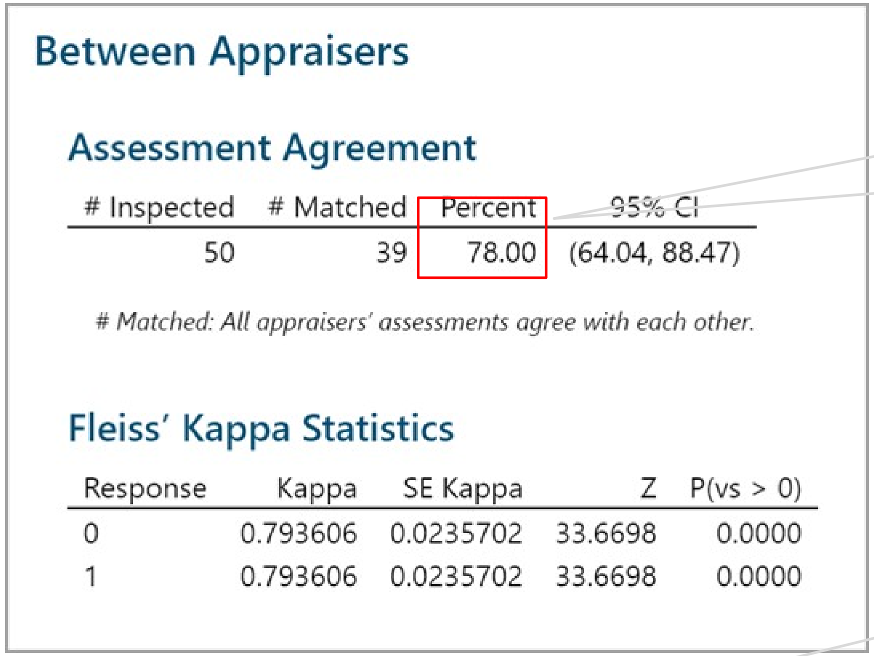
 # 1 / 0: Assessments across trials = 1 / standard = 0.
 # 0 / 1: Assessments across trials = 0 / standard = 1.
 # Mixed: Assessments across trials are not identical.

Fleiss' Kappa Statistics

Appraiser	Response	Kappa	SE Kappa	Z	P(vs > 0)
Appraiser A	0	0.880236	0.0816497	10.7806	0.0000
	1	0.880236	0.0816497	10.7806	0.0000
Appraiser B	0	0.922612	0.0816497	11.2996	0.0000
	1	0.922612	0.0816497	11.2996	0.0000
Appraiser C	0	0.774703	0.0816497	9.4881	0.0000
	1	0.774703	0.0816497	9.4881	0.0000

Each Appraiser vs. Standard Agreement Percent: the agreement percentage between each appraiser and the standard. It reflects the accuracy of the measurement system.

Use Minitab to Implement an Attribute MSA



Between Appraisers

Assessment Agreement

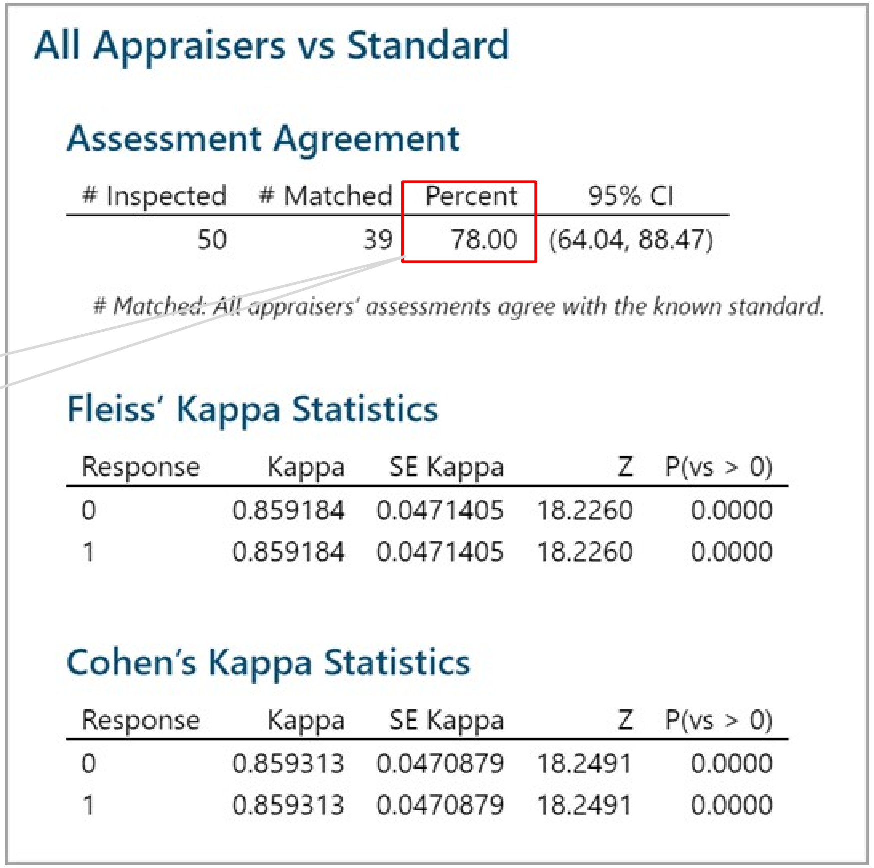
# Inspected	# Matched	Percent	95% CI
50	39	78.00	(64.04, 88.47)

 # Matched: All appraisers' assessments agree with each other.

Fleiss' Kappa Statistics

Response	Kappa	SE Kappa	Z	P(vs > 0)
0	0.793606	0.0235702	33.6698	0.0000
1	0.793606	0.0235702	33.6698	0.0000

Between Appraiser Agreement Percent: the agreement percentage between different appraisers.



All Appraisers vs Standard

Assessment Agreement

# Inspected	# Matched	Percent	95% CI
50	39	78.00	(64.04, 88.47)

 # Matched: All appraisers' assessments agree with the known standard.

Fleiss' Kappa Statistics

Response	Kappa	SE Kappa	Z	P(vs > 0)
0	0.859184	0.0471405	18.2260	0.0000
1	0.859184	0.0471405	18.2260	0.0000

Cohen's Kappa Statistics

Response	Kappa	SE Kappa	Z	P(vs > 0)
0	0.859313	0.0470879	18.2491	0.0000
1	0.859313	0.0470879	18.2491	0.0000

All Appraisers vs. Standard Agreement Percent: overall agreement percentage of both within and between appraisers. It reflects how precise the measurement system performs.

Use Minitab to Implement an Attribute MSA

- **Kappa statistic** is a coefficient indicating the agreement percentage above the expected agreement by chance.

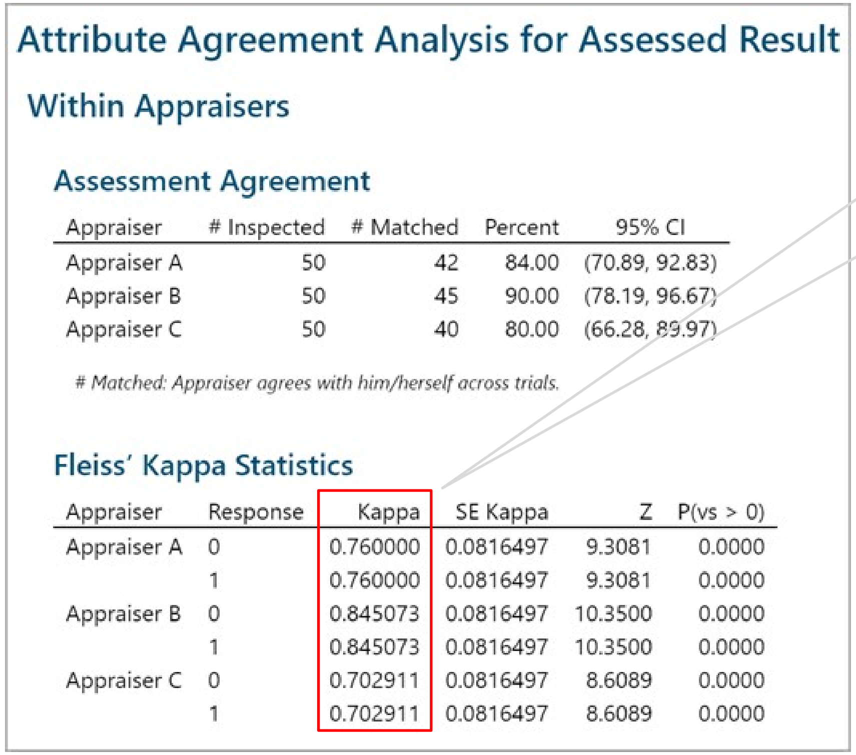
- Kappa ranges from -1 (perfect disagreement) to 1 (perfect agreement).

- When the observed agreement is less than the chance agreement, Kappa is negative.

- When the observed agreement is greater than the chance agreement, Kappa is positive.

- Rule of thumb: If Kappa is greater than 0.7, the measurement system is acceptable. If Kappa is greater than 0.9, the measurement system is excellent.

Use Minitab to Implement an Attribute MSA



Attribute Agreement Analysis for Assessed Result

Within Appraisers

Assessment Agreement

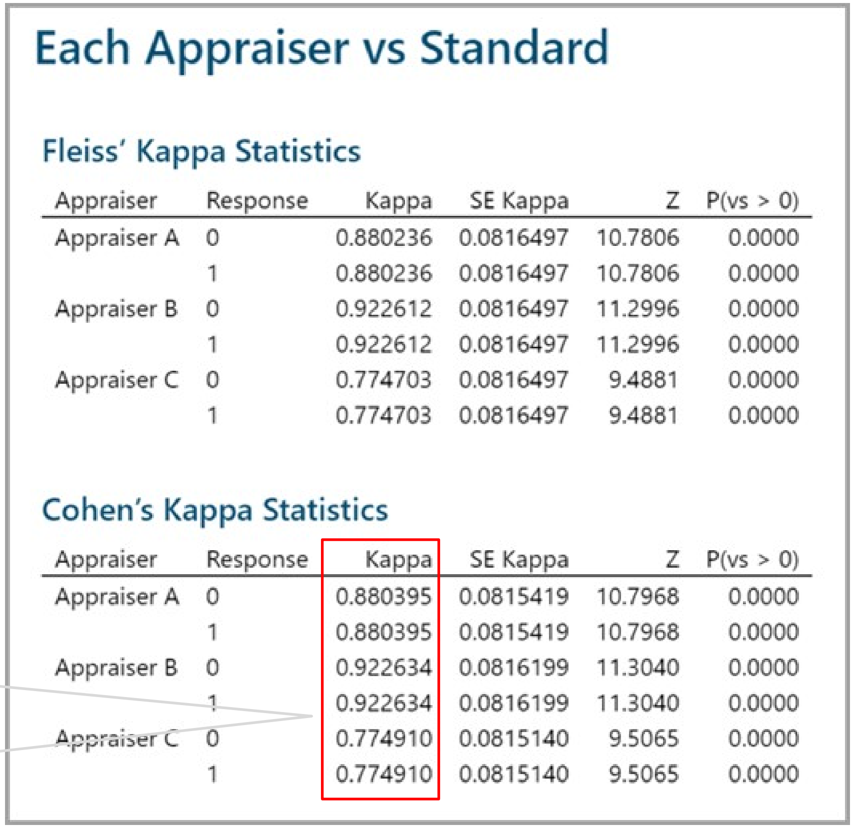
Appraiser	# Inspected	# Matched	Percent	95% CI
Appraiser A	50	42	84.00	(70.89, 92.83)
Appraiser B	50	45	90.00	(78.19, 96.67)
Appraiser C	50	40	80.00	(66.28, 89.97)

 # Matched: Appraiser agrees with him/herself across trials.

Fleiss' Kappa Statistics

Appraiser	Response	Kappa	SE Kappa	Z	P(vs > 0)
Appraiser A	0	0.760000	0.0816497	9.3081	0.0000
	1	0.760000	0.0816497	9.3081	0.0000
Appraiser B	0	0.845073	0.0816497	10.3500	0.0000
	1	0.845073	0.0816497	10.3500	0.0000
Appraiser C	0	0.702911	0.0816497	8.6089	0.0000
	1	0.702911	0.0816497	8.6089	0.0000

Kappa statistic of the agreement within each appraiser



Each Appraiser vs Standard

Fleiss' Kappa Statistics

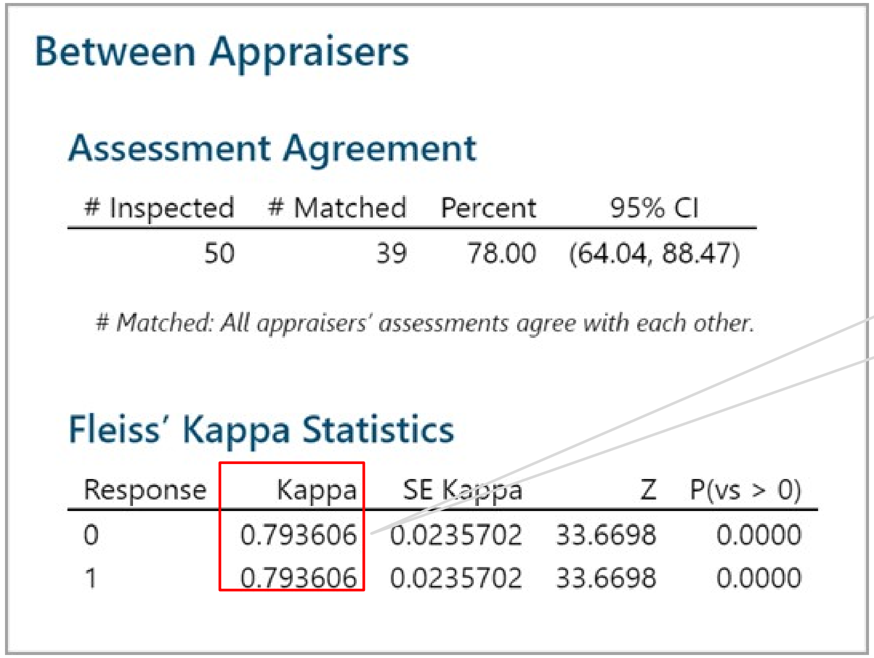
Appraiser	Response	Kappa	SE Kappa	Z	P(vs > 0)
Appraiser A	0	0.880236	0.0816497	10.7806	0.0000
	1	0.880236	0.0816497	10.7806	0.0000
Appraiser B	0	0.922612	0.0816497	11.2996	0.0000
	1	0.922612	0.0816497	11.2996	0.0000
Appraiser C	0	0.774703	0.0816497	9.4881	0.0000
	1	0.774703	0.0816497	9.4881	0.0000

Cohen's Kappa Statistics

Appraiser	Response	Kappa	SE Kappa	Z	P(vs > 0)
Appraiser A	0	0.880395	0.0815419	10.7968	0.0000
	1	0.880395	0.0815419	10.7968	0.0000
Appraiser B	0	0.922634	0.0816199	11.3040	0.0000
	1	0.922634	0.0816199	11.3040	0.0000
Appraiser C	0	0.774910	0.0815140	9.5065	0.0000
	1	0.774910	0.0815140	9.5065	0.0000

Kappa statistic of the agreement between individual appraiser and the standard

Use Minitab to Implement an Attribute MSA



Between Appraisers

Assessment Agreement

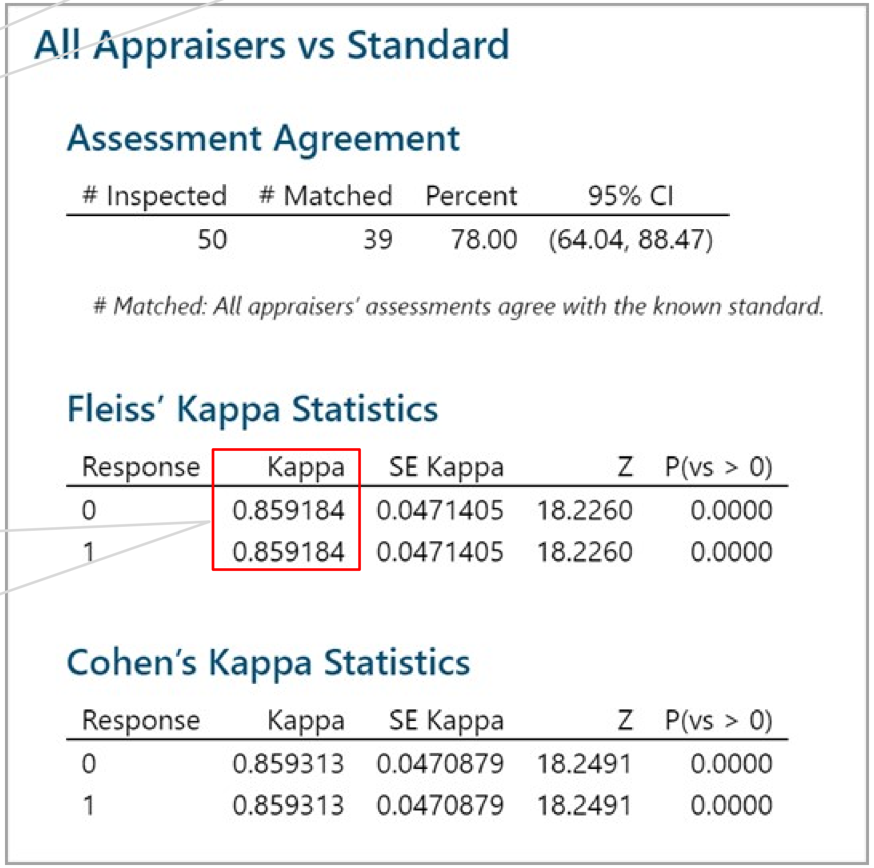
# Inspected	# Matched	Percent	95% CI
50	39	78.00	(64.04, 88.47)

 # Matched: All appraisers' assessments agree with each other.

Fleiss' Kappa Statistics

Response	Kappa	SE Kappa	Z	P(vs > 0)
0	0.793606	0.0235702	33.6698	0.0000
1	0.793606	0.0235702	33.6698	0.0000

Kappa statistic of the agreement between appraisers. They are greater than 0.7 and acceptable.



All Appraisers vs Standard

Assessment Agreement

# Inspected	# Matched	Percent	95% CI
50	39	78.00	(64.04, 88.47)

 # Matched: All appraisers' assessments agree with the known standard.

Fleiss' Kappa Statistics

Response	Kappa	SE Kappa	Z	P(vs > 0)
0	0.859184	0.0471405	18.2260	0.0000
1	0.859184	0.0471405	18.2260	0.0000

Cohen's Kappa Statistics

Response	Kappa	SE Kappa	Z	P(vs > 0)
0	0.859313	0.0470879	18.2491	0.0000
1	0.859313	0.0470879	18.2491	0.0000

Kappa statistic of the overall agreement between appraisers and the standard. Again, greater than 0.7 suggests an acceptable measurement system.

2.4 Process Capability

Black Belt Training: Measure Phase

2.1 Process Definition
- 2.1.1 Cause and Effect Diagrams
- 2.1.2 Cause and Effects Matrix
- 2.1.3 Process Mapping
- 2.1.4 FMEA: Failure Modes & Effects Analysis
- 2.1.5 Theory of Constraints

2.2 Six Sigma Statistics
- 2.2.1 Basic Statistics
- 2.2.2 Descriptive Statistics
- 2.2.3 Distributions and Normality
- 2.2.4 Graphical Analysis

2.3 Measurement System Analysis
- 2.3.1 Precision and Accuracy
- 2.3.2 Bias, Linearity, and Stability
- 2.3.3 Gage R&R
- 2.3.4 Variable and Attribute MSA

2.4 Process Capability
- 2.4.1 Capability Analysis
- 2.4.2 Concept of Stability
- 2.4.3 Attribute and Discrete Capability
- 2.4.4 Monitoring Techniques

2.4.1 Capability Analysis

What is Process Capability?

- The **process capability** measures how well the process performs to meet given specified outcome.

- It indicates the conformance of a process to meet given requirements or specifications.

- **Capability analysis** helps to better understand the performance of the process with respect to meeting customer's specifications and identify the process improvement opportunities.

Process Capability Analysis Steps

- Step 1: Determine the metric or parameter to measure and analyze.
- Step 2: Collect the historical data for the parameter of interest.
- Step 3: Prove the process is statistically stable (i.e., in control).
- Step 4: Calculate the process capability indices.
- Step 5: Monitor the process and ensure it remains in control over time. Update the process capability indices if needed.

Process Capability Indices

- Process capability can be presented using various indices depending on the nature of the process and the goal of the analysis.

- Popular process capability indices:
  - C_p
  - P_p
  - C_{pk}
  - P_{pk}
  - C_{pm}

C_p

- **C_p** stands for **capability of the process**.

$$C_p = \frac{USL - LSL}{6 \times \sigma_{within}}$$

where

$$\sigma_{within} = \frac{s_p}{c_4(d+1)} \qquad s_p = \sqrt{\frac{\sum_i \sum_j (x_{ij} - \bar{x}_i)}{\sum_i (n_i - 1)}}$$

$$d = \sum_i (n_i - 1) \qquad c_4 = \frac{4(n-1)}{(4n-3)}$$

USL and ***LSL*** are the upper and lower specification limits.
n is the sample size.

C_p

- C_p measures the process' potential capability to meet the two-sided specifications.

- It does not take the process average into consideration.

- High C_p indicates the small spread of the process with respect to the spread of the customer specifications.

- C_p is recommended when the process is centered between the specification limits.

- C_p works when there are both upper and lower specification limits.

P_p

- **P_p** stands for **performance of the process**.

$$P_p = \frac{USL - LSL}{6 \times \sigma_{overall}}$$

where

$$\sigma_{overall} = \frac{s}{c_4(n)} \qquad s = \sqrt{\sum_i \sum_j \frac{(x_{ij} - \bar{x})^2}{n-1}}$$

$$c_4 = \frac{4(n-1)}{(4n-3)}$$

USL and ***LSL*** are the upper and lower specification limits.
n is the sample size.

P_p

- Similar to C_p, P_p measures the capability of the process to meet the two-sided specifications.
- It only focuses on the spread and does not take the process centralization into consideration.
- It is recommended when the process is centered between the specification limits.
- C_p considers the within-subgroup standard deviation and P_p considers the total standard deviation from the sample data.
- P_p works when there are both upper and lower specification limits.

C_{pk}

- **C_{pk}** stands for the capability of the process with a k factor adjustment.

$$C_{pk} = (1-k) \times C_p$$

where

$$k = \frac{|m-\mu|}{\frac{USL-LSL}{2}} \qquad m = \frac{USL+LSL}{2}$$

μ is the process mean; *n* is the sample size.

USL and ***LSL*** are the upper and lower specification limits.

C_{pk}

- The formulas to calculate C_{pk} can also be expressed as follows:

$$C_{pk} = \min\left(\frac{USL - \mu}{3 \times \sigma_{within}}, \frac{\mu - LSL}{3 \times \sigma_{within}} \right)$$

where

$$\sigma_{within} = \frac{s_p}{c_4(d+1)} \qquad s_p = \sqrt{\frac{\sum_i \sum_j (x_{ij} - \bar{x}_i)}{\sum_i (n_i - 1)}}$$

$$d = \sum_i (n_i - 1) \qquad c_4 = \frac{4(n-1)}{(4n-3)}$$

USL and ***LSL*** are the upper and lower specification limits.

C_{pk}

- C_{pk} measures the process' actual capability by taking both the variation and average of the process into consideration.

- The process does not need to be centered between the specification limits to make the index meaningful.

- C_{pk} is recommended when the process is not in the center between the specification limits.

- When there is only a one-sided limit, C_{pk} is calculated using C_{pu} or C_{pl}.

C_{pk}

- C_{pk} for upper specification limit:

$$C_{pu} = \frac{USL - \mu}{3 \times \sigma_{within}}$$

- C_{pk} for lower specification limit:

$$C_{pl} = \frac{\mu - LSL}{3 \times \sigma_{within}}$$

USL and ***LSL*** are the upper and lower specification limits.
μ is the process mean.

P_{pk}

- **P_{pk}** stands for the performance of the process with a k factor adjustment.

$$P_{pk} = (1-k) \times P_p$$

where

$$k = \frac{|m-\mu|}{\frac{USL-LSL}{2}}$$

$$m = \frac{USL+LSL}{2}$$

USL and ***LSL*** are the upper and lower specification limits.
μ is the process mean.

P$_{pk}$

- The formulas to calculate P$_{pk}$ can also be expressed as follows:

$$P_{pk} = \min\left(\frac{USL-\mu}{3\times\sigma_{overall}}, \frac{\mu-LSL}{3\times\sigma_{overall}}\right)$$

$$\sigma_{overall} = \frac{s}{c_4(n)}$$

$$s = \sqrt{\sum_i \sum_j \frac{(x_{ij}-\bar{x})^2}{n-1}}$$

$$c_4 = \frac{4(n-1)}{(4n-3)}$$

USL and ***LSL*** are the upper and lower specification limits.
μ is the process mean. *n* is the sample size.

P_{pk}

- Similar to C_{pk}, P_{pk} measures the process capability by taking both the variation and the average of the process into consideration.
- P_{pk} solves the decentralization problem P_p cannot overcome.
- C_{pk} considers the within-subgroup standard deviation, while P_{pk} considers the total standard deviation from the sample data.
- When there is only a one-sided specification limit, P_{pk} is calculated using P_{pu} or P_{pl}.

P_{pk}

- P_{pk} for upper specification limit:

$$P_{pu} = \frac{USL - \mu}{3 \times \sigma_{overall}}$$

- P_{pk} for lower specification limit:

$$P_{pl} = \frac{\mu - LSL}{3 \times \sigma_{overall}}$$

USL and ***LSL*** are the upper and lower specification limits.

C_{pm}

- C_p, P_p, C_{pk}, and P_{pk} all consider the variation of the process. C_{pk} and P_{pk} take both the variation and the average of the process into consideration when measuring the process capability.

- It is possible that the process average fails to meet the target customers require while the process still remains between the specification limits. C_{pm} (Taguchi's capability index) helps to capture the variation from the specified target.

C_{pm}

- Formula to calculate C_{pm}

$$C_{pm} = \frac{\min(T - LSL, USL - T)}{3 \times \sqrt{s^2 + (\mu - T)^2}}$$

USL and ***LSL*** are the upper and lower specification limits.
T is the specified target.
μ is the process mean.

Note: Cpm can work only if there is a target value specified.

Interpreting Capability

- Interpreting the results of a capability analysis is very dependent upon the nature of the business, product and/or customer. The Automotive Industry Action Group (AIAG) suggests that:
  - P_p and P_{pk} should be > 1.67
  - C_p and C_{pk} should be > 1.33.
- The reality is that anything greater than 1.0 is a fairly capable process but your business needs to assess the costs vs. benefits of achieving capability greater than 1.67 or even higher. Below is a simple table for quick interpretation

Index	Value	Interpretation	Sigma Level
C_{pk}	<1.0	Not Very Capable	<3
C_{pk}	1.0-1.99	Capable	3-6
C_{pk}	>2.0	Very Capable	>6

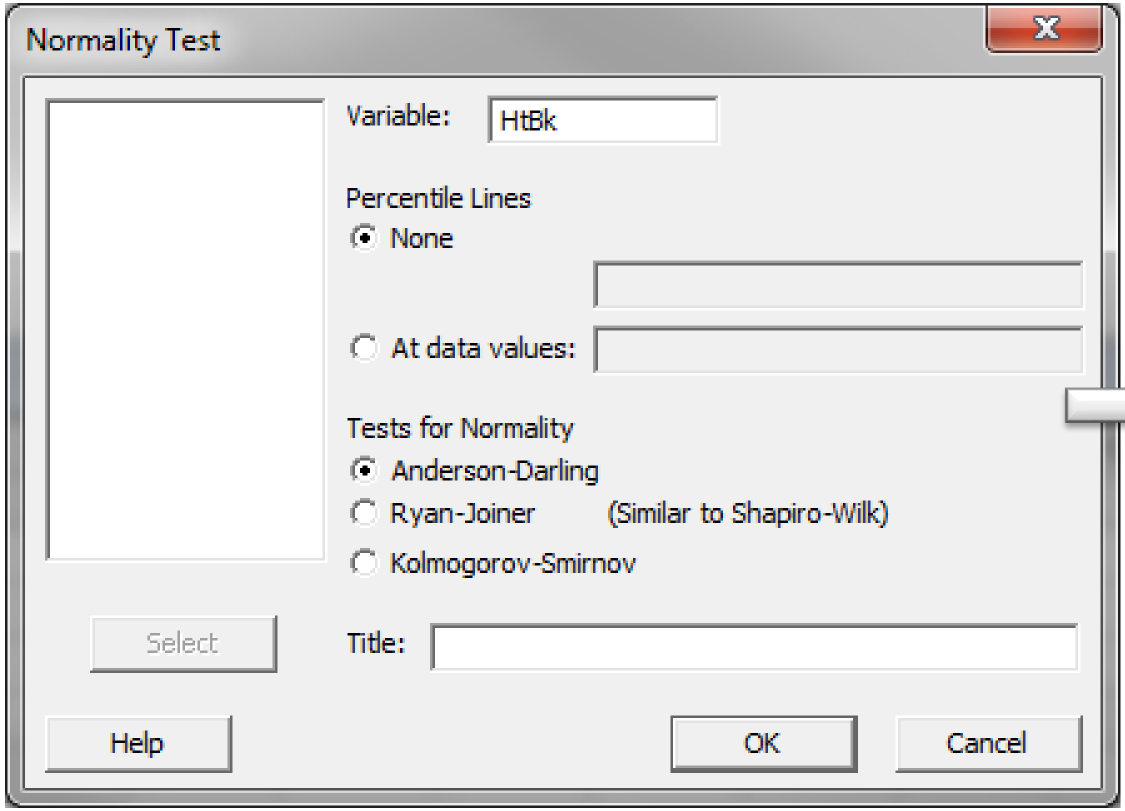
- If you're dealing with customer safety or life and death influences then obviously the product necessitates a capability greater than 2.0.

Use Minitab to Run a Process Capability Analysis

- Data File: “Capability Analysis” tab in “Sample Data.xlsx”
- Steps in Minitab to run a process capability analysis:

  1) Click Stat → Basic Statistics → Normality Test.
  2) A new window named “Normality Test” pops up.
  3) Select “HtBk” as the variable and Click “OK.”
  4) The histogram and the normality test results appear in the new window.
  5) In this example, the p-value is 0.275, greater than the alpha level (0.05). We fail to reject the hypothesis and conclude that the data are normally distributed.
  6) Click Stat → Quality Tools → Capability Analysis → Normal.
  7) A new window named “Capability Analysis (Normal Distribution)” pops up.
  8) Select “HtBk” as the single column and enter “1” as the subgroup size.
  9) Enter “6” as the “Lower spec” and “7” as the “Upper spec”
  10) Click “Options” button and another new window named “Capability Analysis (Normal Distribution) – Options” pops up.
  11) Enter “6.5” as the target and click “OK.”
  12) Click “OK” in the “Capability Analysis (Normal Distribution)” window. The capability analysis results are displayed.

Use Minitab to Run a Process Capability Analysis


Normality Test

Variable: HtBk

Percentile Lines

● None

○ At data values:

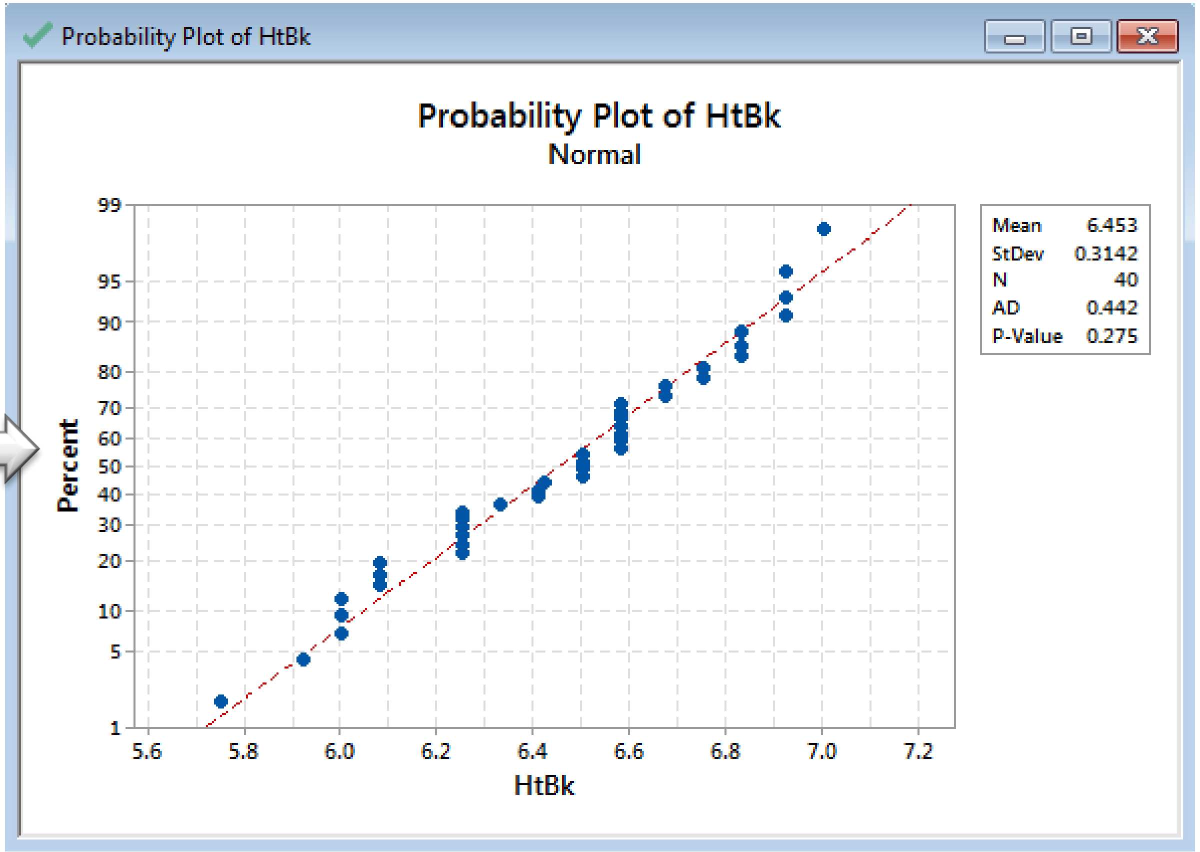
Tests for Normality

● Anderson-Darling

○ Ryan-Joiner (Similar to Shapiro-Wilk)

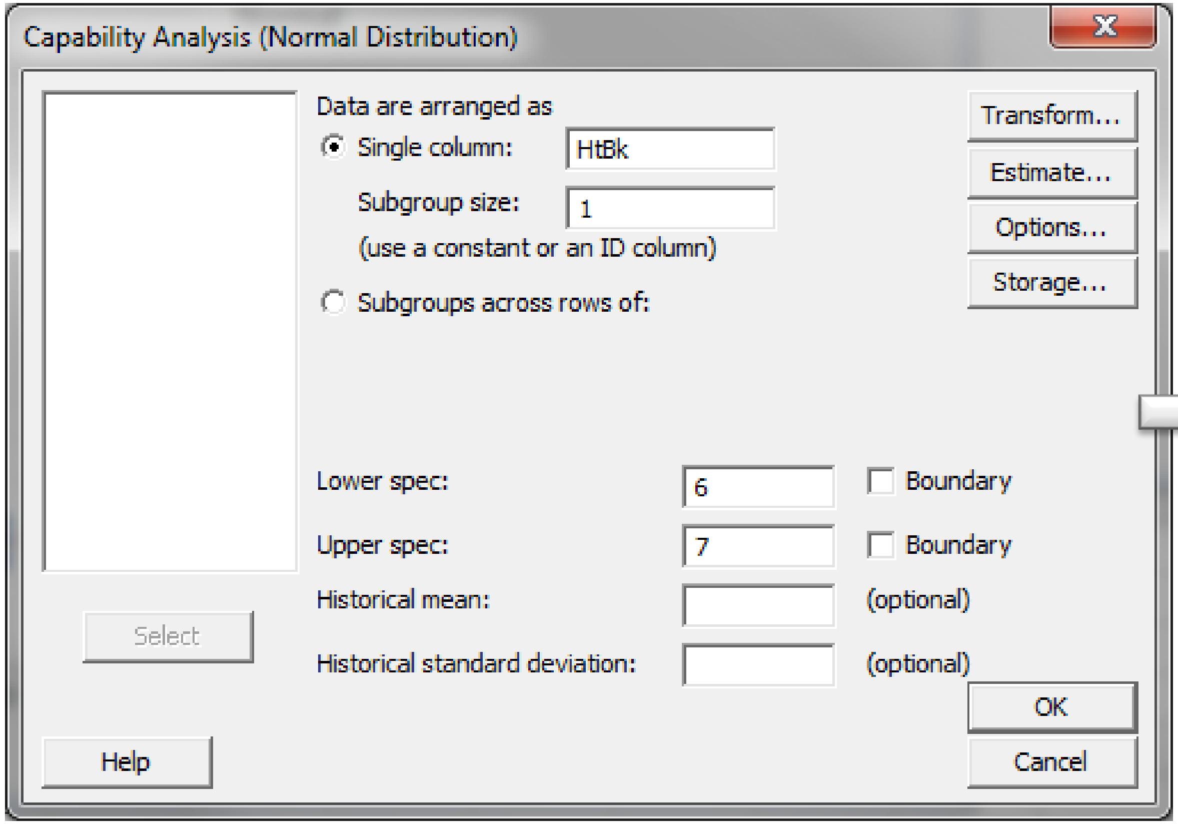
○ Kolmogorov-Smirnov

Title:



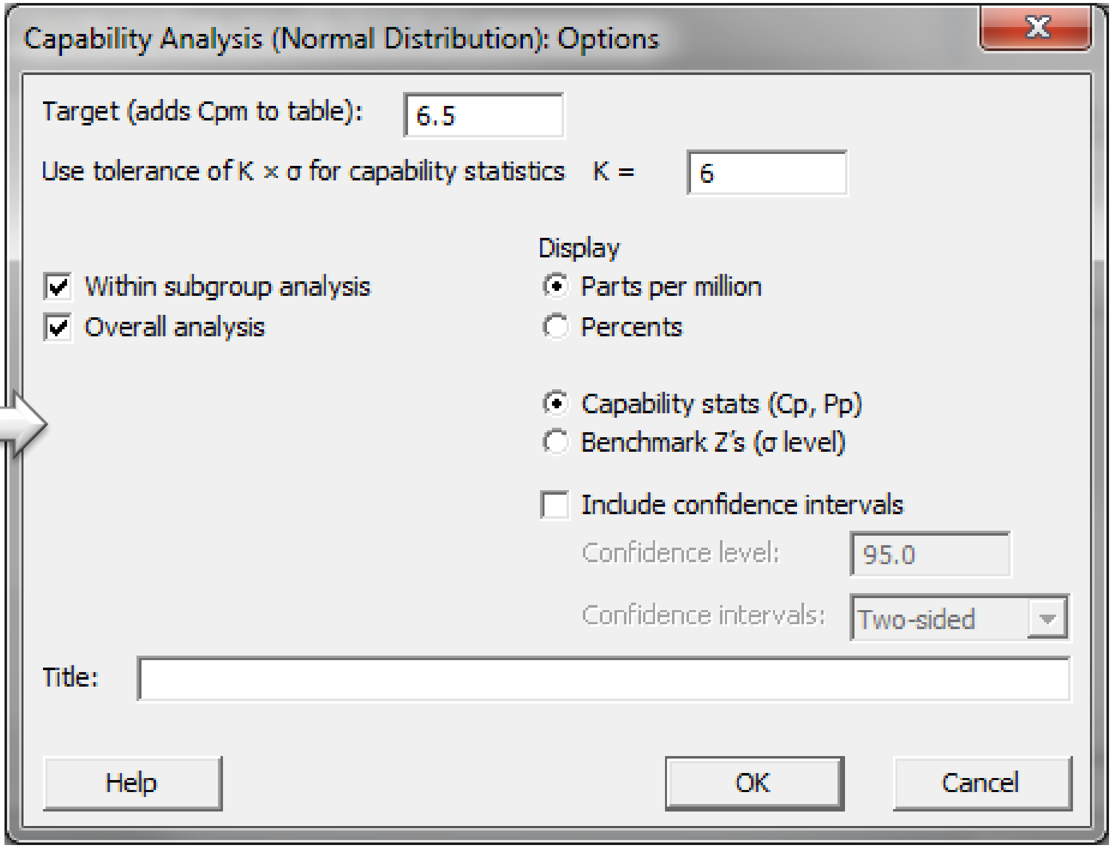
Mean	6.453
StDev	0.3142
N	40
AD	0.442
P-Value	0.275

Use Minitab to Run a Process Capability Analysis


Capability Analysis (Normal Distribution)

Data are arranged as
● Single column: HtBk
Subgroup size: 1
(use a constant or an ID column)
○ Subgroups across rows of:

Lower spec: 6 ☐ Boundary
Upper spec: 7 ☐ Boundary
Historical mean: (optional)
Historical standard deviation: (optional)


Capability Analysis (Normal Distribution): Options

Target (adds Cpm to table): 6.5
Use tolerance of K × σ for capability statistics K = 6

☑ Within subgroup analysis
☑ Overall analysis

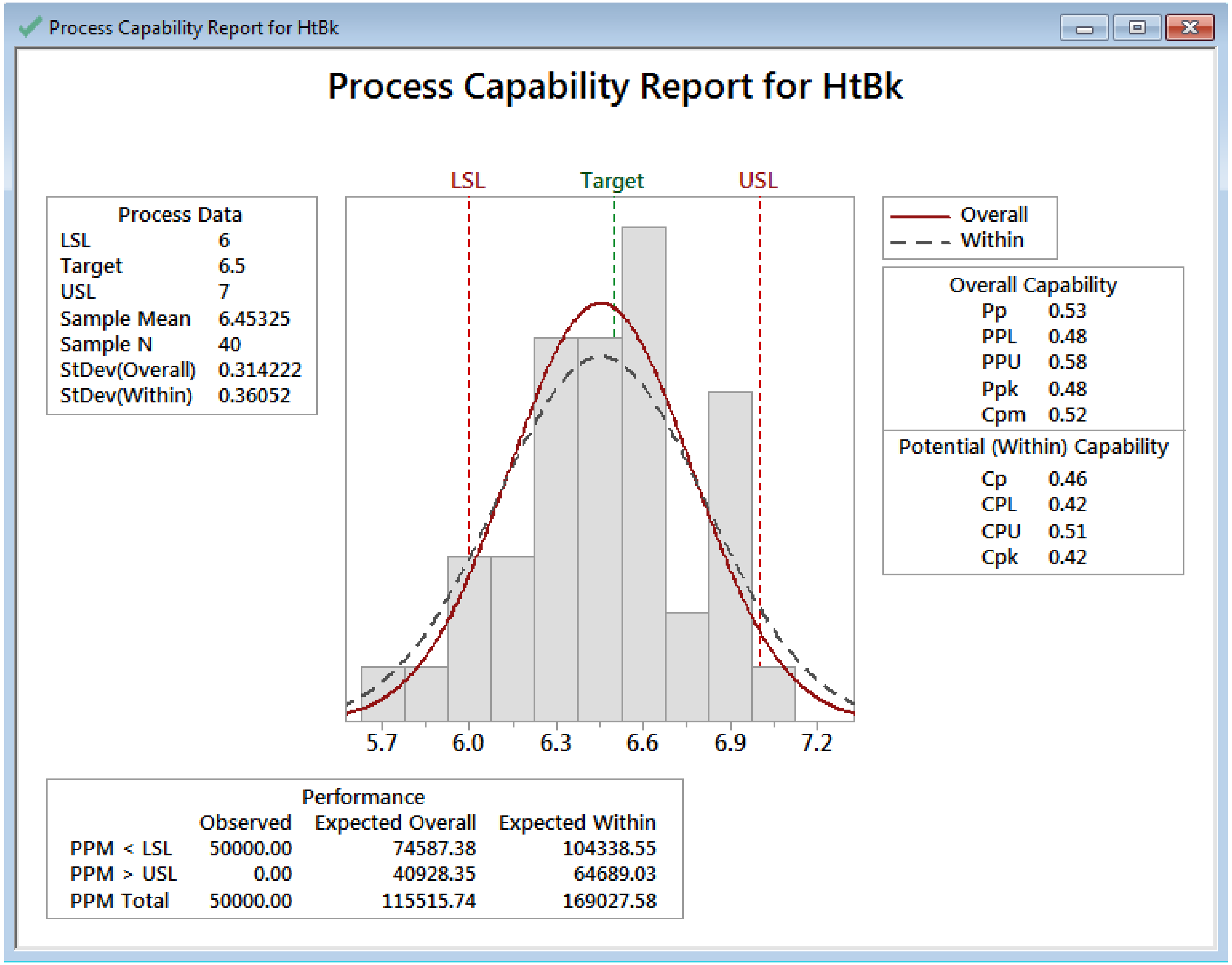
Display
● Parts per million
○ Percents

● Capability stats (Cp, Pp)
○ Benchmark Z's (σ level)

☐ Include confidence intervals
Confidence level: 95.0
Confidence intervals: Two-sided

Title:

Use Minitab to Run a Process Capability Analysis



Process Capability Report for HtBk

Process Data	
LSL	6
Target	6.5
USL	7
Sample Mean	6.45325
Sample N	40
StDev(Overall)	0.314222
StDev(Within)	0.36052

Overall Capability	
Pp	0.53
PPL	0.48
PPU	0.58
Ppk	0.48
Cpm	0.52

Potential (Within) Capability	
Cp	0.46
CPL	0.42
CPU	0.51
Cpk	0.42

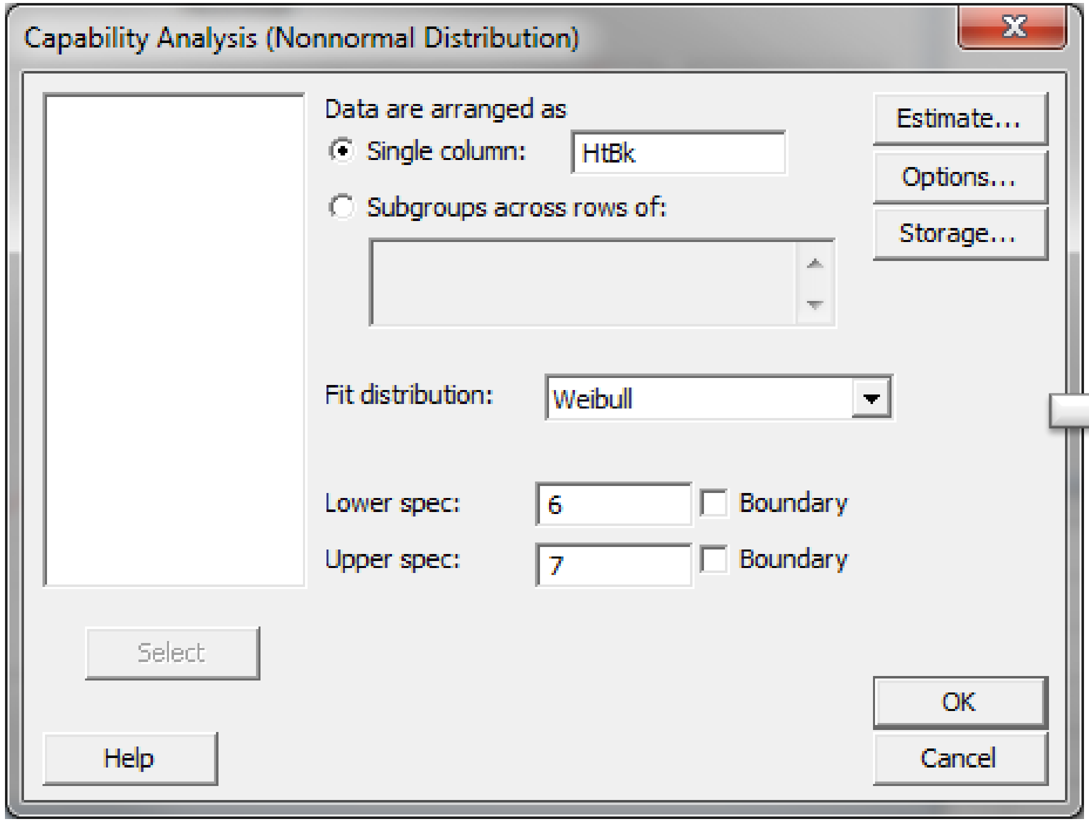
Performance

	Observed	Expected Overall	Expected Within
PPM < LSL	50000.00	74587.38	104338.55
PPM > USL	0.00	40928.35	64689.03
PPM Total	50000.00	115515.74	169027.58

Use Minitab to Run a Process Capability Analysis

- If the p-value of the normality test is smaller than the alpha level (0.05), we reject the null hypothesis and conclude that the data are not normally distributed. In this case we run a non-normal analysis.
  1) Click Stat → Quality Tools → Capability Analysis → Nonnormal
  2) A new window named "Capability Analysis (Nonnormal Distribution)" pops up.
  3) Select "HtBk" as the single column.
  4) Enter "6" as the "Lower spec" and "7" as the "Upper spec."
  5) Click "Options" button and another new window named "Capability Analysis (Nonnormal Distribution) – Options" pops up.
  6) Enter "6.5" as the target and click "OK."
  7) Click "OK" in the "Capability Analysis (Nonnormal Distribution)" window.
  8) The capability analysis results appear in the new window.

Use Minitab to Run a Process Capability Analysis



Capability Analysis (Nonnormal Distribution)

Data are arranged as

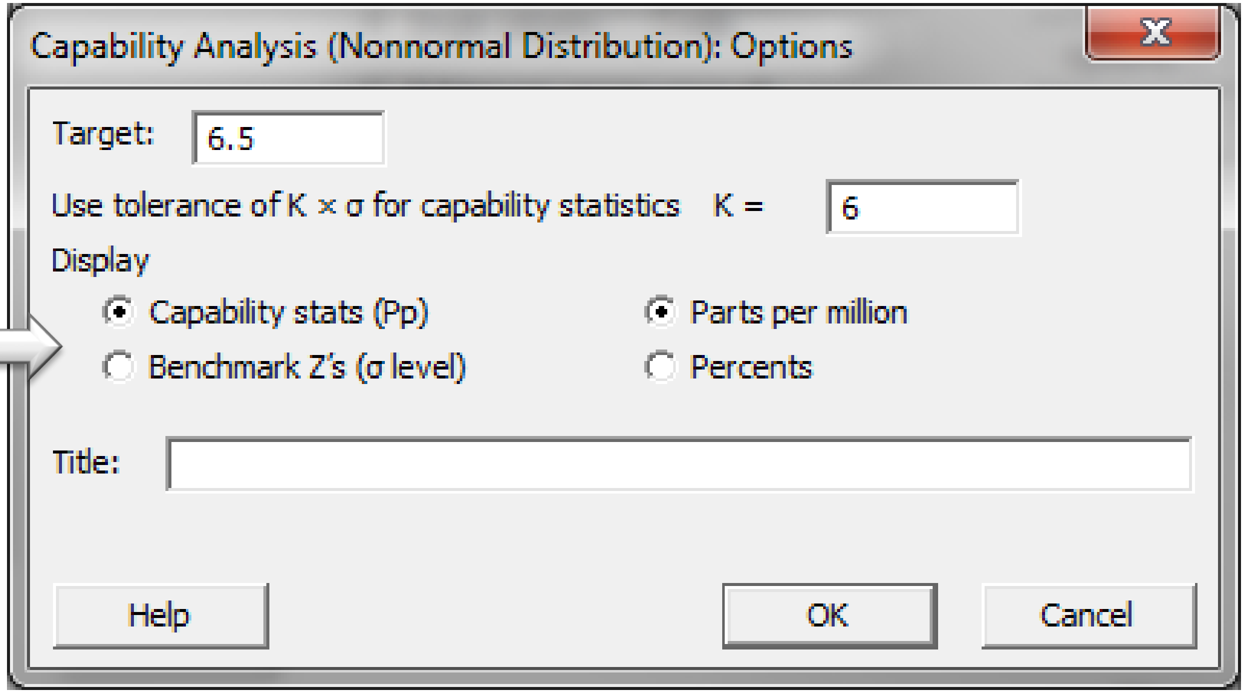
● Single column: HtBk

○ Subgroups across rows of:

Fit distribution: Weibull

Lower spec: 6 ☐ Boundary

Upper spec: 7 ☐ Boundary



Capability Analysis (Nonnormal Distribution): Options

Target: 6.5

Use tolerance of K × σ for capability statistics K = 6

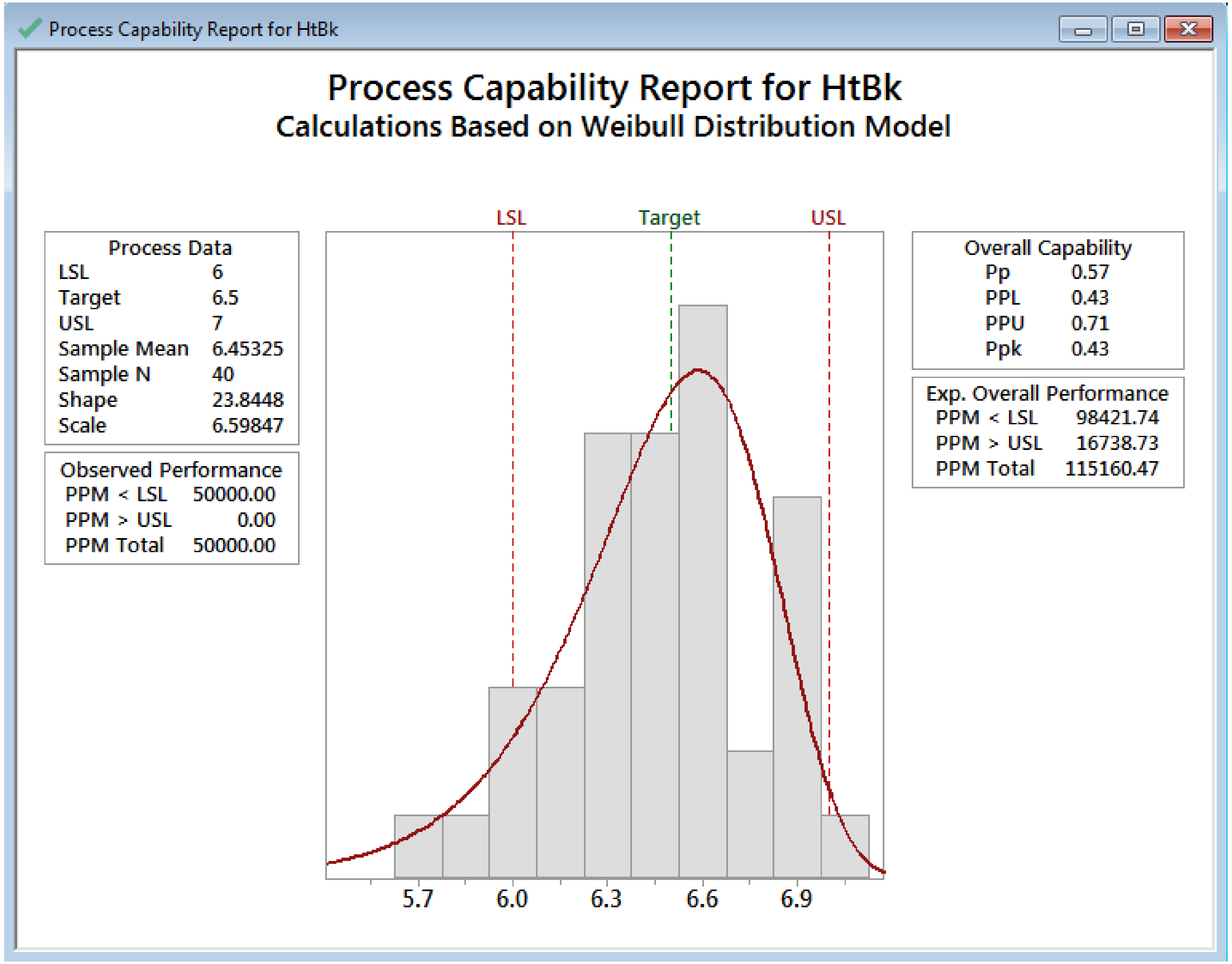
Display

● Capability stats (Pp) ● Parts per million

○ Benchmark Z's (σ level) ○ Percents

Title:

Use Minitab to Run a Process Capability Analysis



Process Capability Report for HtBk

Process Capability Report for HtBk

Calculations Based on Weibull Distribution Model

Process Data	
LSL	6
Target	6.5
USL	7
Sample Mean	6.45325
Sample N	40
Shape	23.8448
Scale	6.59847

Observed Performance	
PPM < LSL	50000.00
PPM > USL	0.00
PPM Total	50000.00

Overall Capability	
Pp	0.57
PPL	0.43
PPU	0.71
Ppk	0.43

Exp. Overall Performance	
PPM < LSL	98421.74
PPM > USL	16738.73
PPM Total	115160.47

2.4.2 Concept of Stability

What is Process Stability?

- A process is said to be stable when:
  - the process is in control
  - the future behavior of the process is predictable at least between some limits
  - there is only random variation involved in the process.
  - the causes of variation in the process are only due to chance or common causes
  - there are not any trends, patterns, or outliers in the control chart of the process.

Root Causes of Variation in the Process

- **Common Cause:**
  - Chance
  - Random and anticipated
  - Natural noise
  - Inherent in the process
  - Unable to be eliminated from the process.

- **Special Cause:**
  - Assignable cause
  - Unanticipated
  - Unnatural pattern
  - Signal of changes in the process
  - Able to be eliminated from the process.

Control Charts

- **Control charts** are the graphical tools to analyze the stability of a process.

- A control chart is used to identify the presence of potential special causes in the process and to determine whether the process is statistically in control.

- If the samples or calculations of samples are all in control, the process is stable and the data from the process can be used to predict the future performance of the process.

Popular Control Charts

- I-MR Chart
- Xbar-R Chart
- Xbar-S Chart
- C Chart
- U Chart
- P Chart
- NP Chart
- EWMA Chart
- CUSUM Chart

Note: More details of the control charts will be introduced in the Control module.

Process Stability vs. Process Capability

- Process stability indicates how stable a process performed in the past.
- When the process is stable, we can use the data from the process to predict its future behavior.
- Process capability indicates how well a process performs with respect to meeting the customer's specifications.
- The process capability analysis is valid only if the process is statistically stable (i.e., in control, predictable).
- Being stable does *not* guarantee that the process is also capable. However, being stable is the prerequisite to determine whether a process is capable.

2.4.3 Attribute & Discrete Capability

Process Capability Analysis for Binomial Data

- If we are measuring the count of defectives in each sample set to assess the process performance of meeting the customer specifications, we use “%Defective” (percentage of items in the samples that are defective) as the process capability index.

$$\%\text{Defective} = \frac{N_{defectives}}{N_{overall}}$$

where $N_{\text{defectives}}$ is the total count of defectives in the samples and N_{overall} is the sum of all the sample sizes.

Process Capability Analysis for Poisson Data

- If we are measuring the count of defects in each sample set to assess the process performance of meeting the customer specifications, we use Mean DPU (defects per unit of measurement) as the process capability index.

$$DPU = \frac{N_{defects}}{N_{overall}}$$

where $N_{defects}$ is the total count of defects in the samples and $N_{overall}$ is the sum of all the units in the samples.

2.4.4 Monitoring Techniques

Capability and Monitoring

- In the Measure phase of the project, process stability analysis and process capability analysis are used to baseline the performance of current process.

- A stable process means we can expect the same process capability in the future.

- In the Control phase of the project, process stability analysis and process capability analysis are combined to monitor whether the improved process is maintained consistently as expected.

3.0 Analyze Phase

Black Belt Training: Analyze Phase

3.1 Patterns of Variation
- 3.1.1 Multi-Vari Analysis
- 3.1.2 Classes of Distributions

3.2 Inferential Statistics
- 3.2.1 Understanding Inference
- 3.2.2 Sampling Techniques and Uses
- 3.2.3 Sample Size
- 3.2.4 Central Limit Theorem

3.3 Hypothesis Testing
- 3.3.1 Goals of Hypothesis Testing
- 3.3.2 Statistical Significance
- 3.3.3 Risk; Alpha and Beta
- 3.3.4 Types of Hypothesis Tests

3.4 Hypothesis Testing: Normal Data
- 3.4.1 One and Two Sample T-Tests
- 3.4.2 One sample variance
- 3.4.3 One Way ANOVA

3.5 Hypothesis Testing: Non-Normal Data
- 3.5.1 Mann-Whitney
- 3.5.2 Kruskal-Wallis
- 3.5.3 Moods Median
- 3.5.4 Friedman
- 3.5.5 One Sample Sign
- 3.5.6 One Sample Wilcoxon
- 3.5.7 One and Two Sample Proportion
- 3.5.8 Chi-Squared (Contingency Tables)
- 3.5.9 Test of Equal Variances

3.1 Patterns of Variation

Black Belt Training: Analyze Phase

3.1 Patterns of Variation
- 3.1.1 Multi-Vari Analysis
- 3.1.2 Classes of Distributions

3.2 Inferential Statistics
- 3.2.1 Understanding Inference
- 3.2.2 Sampling Techniques and Uses
- 3.2.3 Sample Size
- 3.2.4 Central Limit Theorem

3.3 Hypothesis Testing
- 3.3.1 Goals of Hypothesis Testing
- 3.3.2 Statistical Significance
- 3.3.3 Risk; Alpha and Beta
- 3.3.4 Types of Hypothesis Tests

3.4 Hypothesis Testing: Normal Data
- 3.4.1 One and Two Sample T-Tests
- 3.4.2 One sample variance
- 3.4.3 One Way ANOVA

3.5 Hypothesis Testing: Non-Normal Data
- 3.5.1 Mann-Whitney
- 3.5.2 Kruskal-Wallis
- 3.5.3 Moods Median
- 3.5.4 Friedman
- 3.5.5 One Sample Sign
- 3.5.6 One Sample Wilcoxon
- 3.5.7 One and Two Sample Proportion
- 3.5.8 Chi-Squared (Contingency Tables)
- 3.5.9 Test of Equal Variances

3.1.1 Multi-Vari Analysis

What is Multi-Vari Analysis?

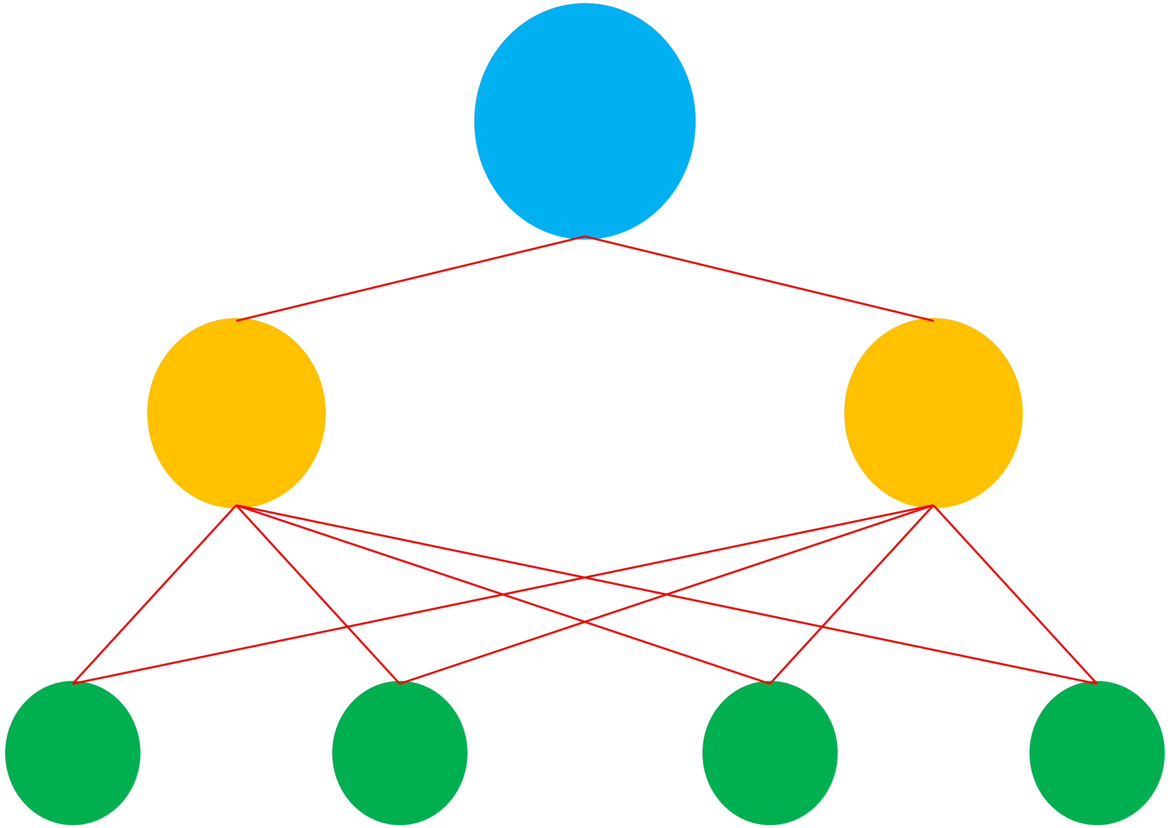
- **Multi-Vari analysis** is a graphic-driven method to analyze the effects of categorical inputs on a continuous output.
  - Y: continuous variable
  - X's: discrete categorical variables. One X may have multiple levels.
- It studies how the variation in the output changes across different inputs and helps us quantitatively determine the major source of variability in the output.
- Multi-Vari charts are used to visualize the source of variation. They work for both crossed and nested hierarchies.

Hierarchy

- **Hierarchy** is a structure of objects in which the relationship of objects can be expressed similar to an organization tree.
- Each object in the hierarchy is described as above, below, or at the same level as another one.
  - If object A is above object B and they are directly connected to each other in the hierarchy tree, A is B's parent and B is A's child.
- In Multi-Vari analysis, we use the hierarchy to present the relationship between categorical factors (inputs).
- Each object in the hierarchy tree indicates a specific level of a factor (input).
- There are generally two types of hierarchies: crossed and nested.

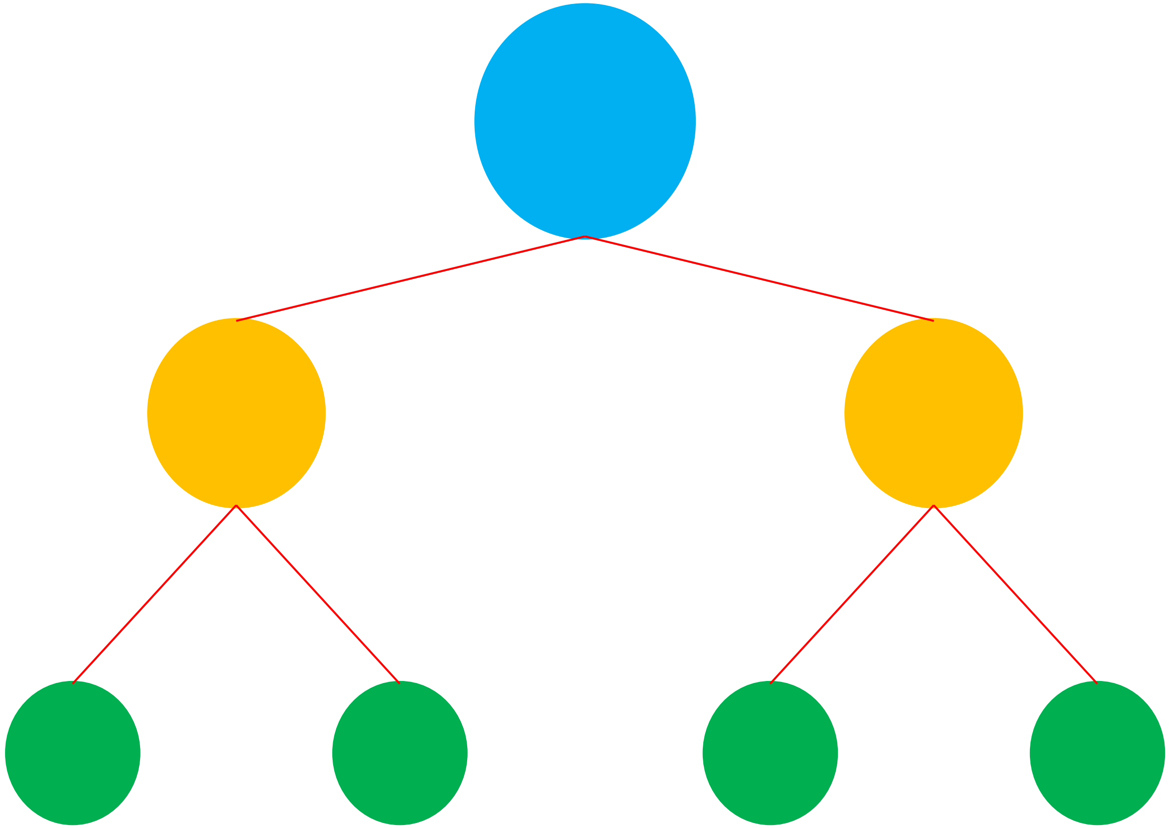
Crossed Hierarchy

- In the hierarchy tree, if one child item has more than one parent item at the higher level, it is a crossed relationship.



Nested Hierarchy

- In the hierarchy tree, if one child item only has one parent item at the higher level, it is a nested relationship.



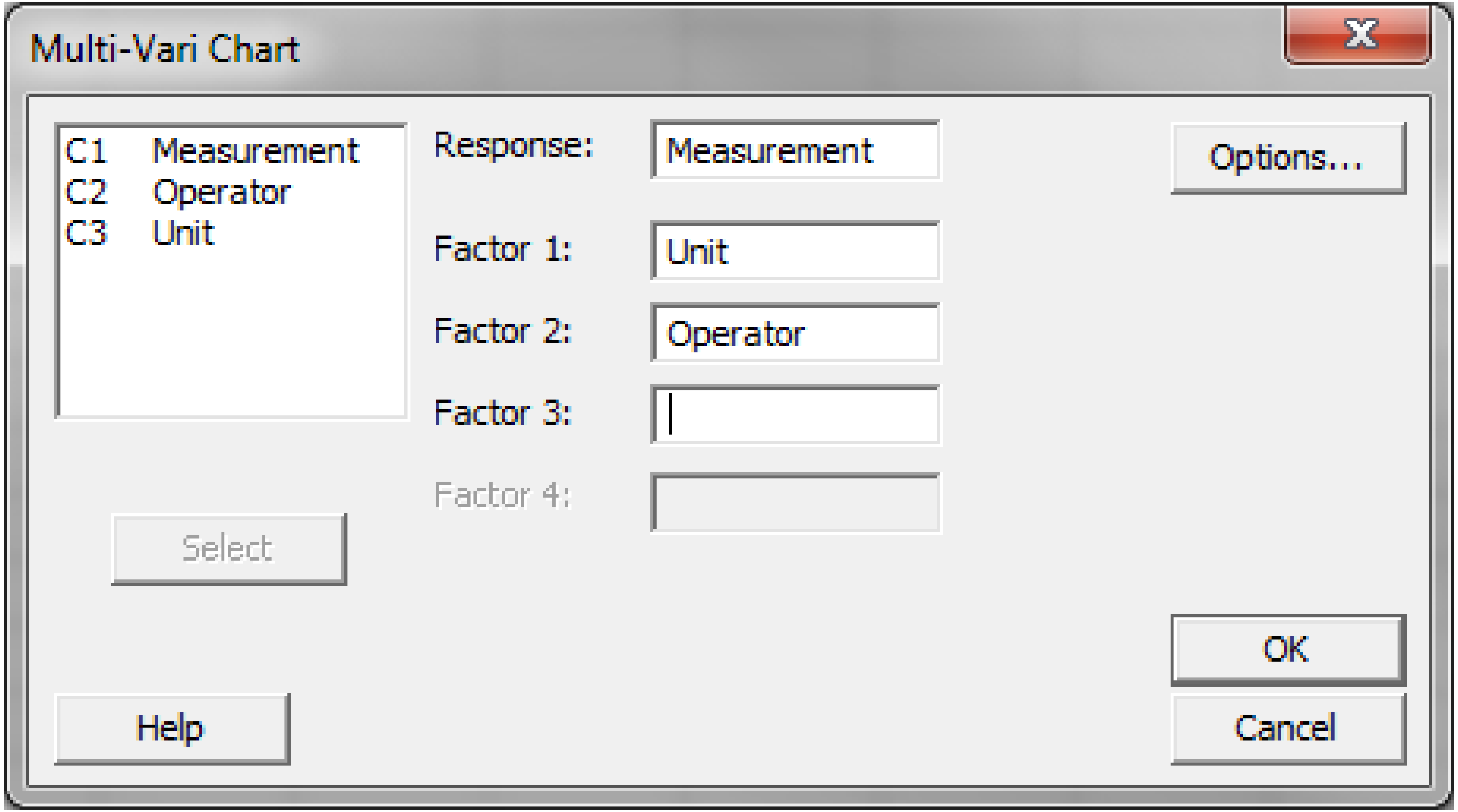
Use Minitab to Perform Multi-Vari Analysis

- Data File: “Multi-Vari” tab in “Sample Data.xlsx.”
  - *Case study:*
    - ABC company produces 10 kinds of units with different weights. Operators measure the weights of the units before sending them out to customers.
    - Multiple factors could have an impact on the weight measurements. The ABC company wants to have a better understanding of the main source of variability existing in the weight measurement.
    - The ABC company randomly selects 3 operators (Joe, John, and Jack) each of whom measures the weights of 10 different units. For each unit, there are 3 items sampled.

Use Minitab to Perform Multi-Vari Analysis

- Step 1:
    1. Click Stat → Quality Tools → Multi-Vari Chart.
- Step 2:
    1. A window named “Multi-Vari Chart” pops up.
    2. Select “Measurement” as the “Response,” “Unit” as “Factor 1,” and “Operator” as “Factor 2.”

Use Minitab to Perform Multi-Vari Analysis



Response: Measurement

Factor 1: Unit

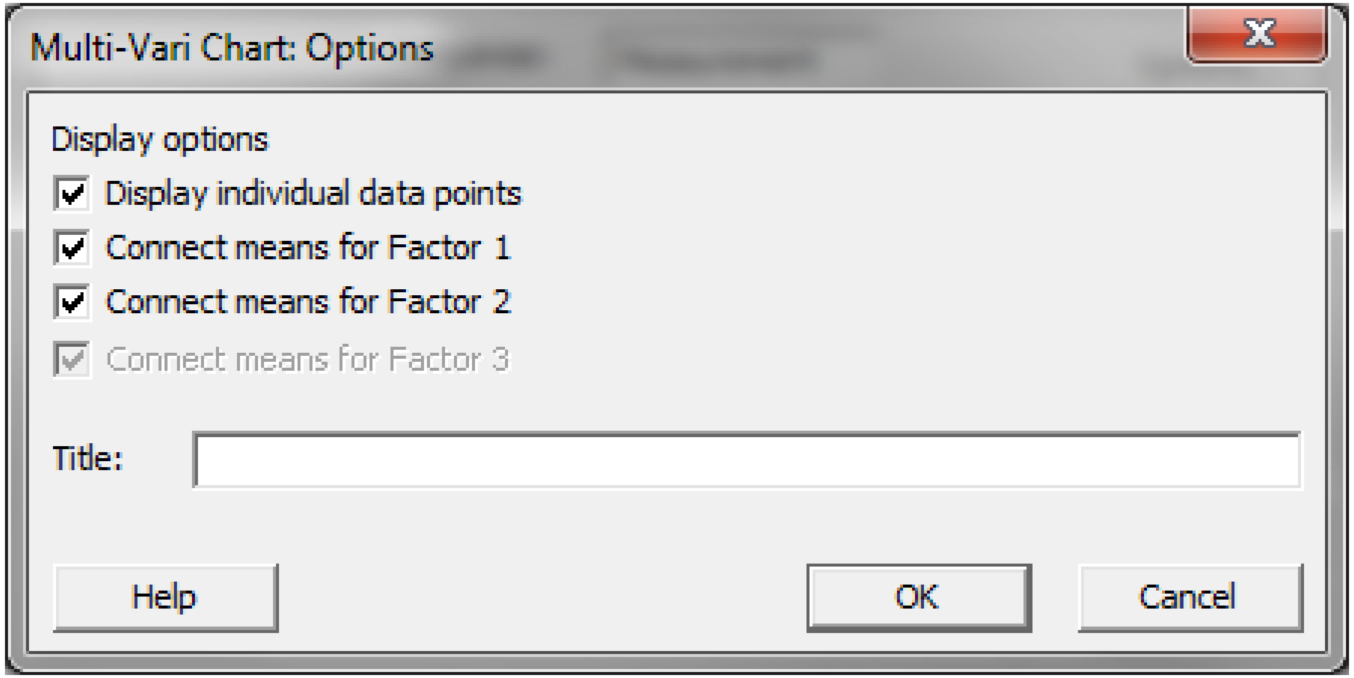
Factor 2: Operator

Factor 3:

Factor 4:

Use Minitab to Perform Multi-Vari Analysis

- Step 3:
    1) Click "Options" button.
    2) A new window named "Multi-Vari Chart" appears.
    3) Check the box of "Display individual data points."
    4) Click "OK."
    5) Click "OK again."


Multi-Vari Chart: Options

Display options

☑ Display individual data points

☑ Connect means for Factor 1

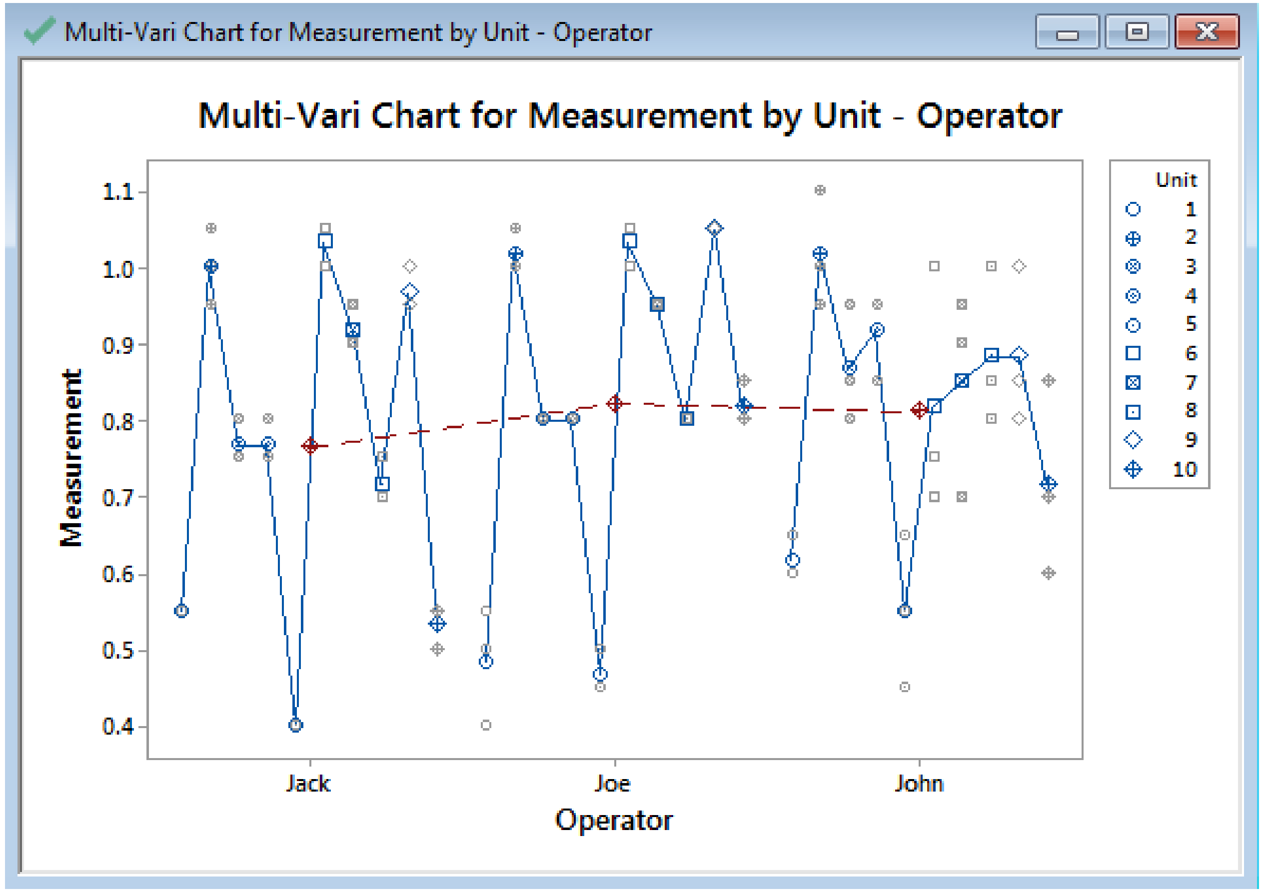
☑ Connect means for Factor 2

☑ Connect means for Factor 3

Title:

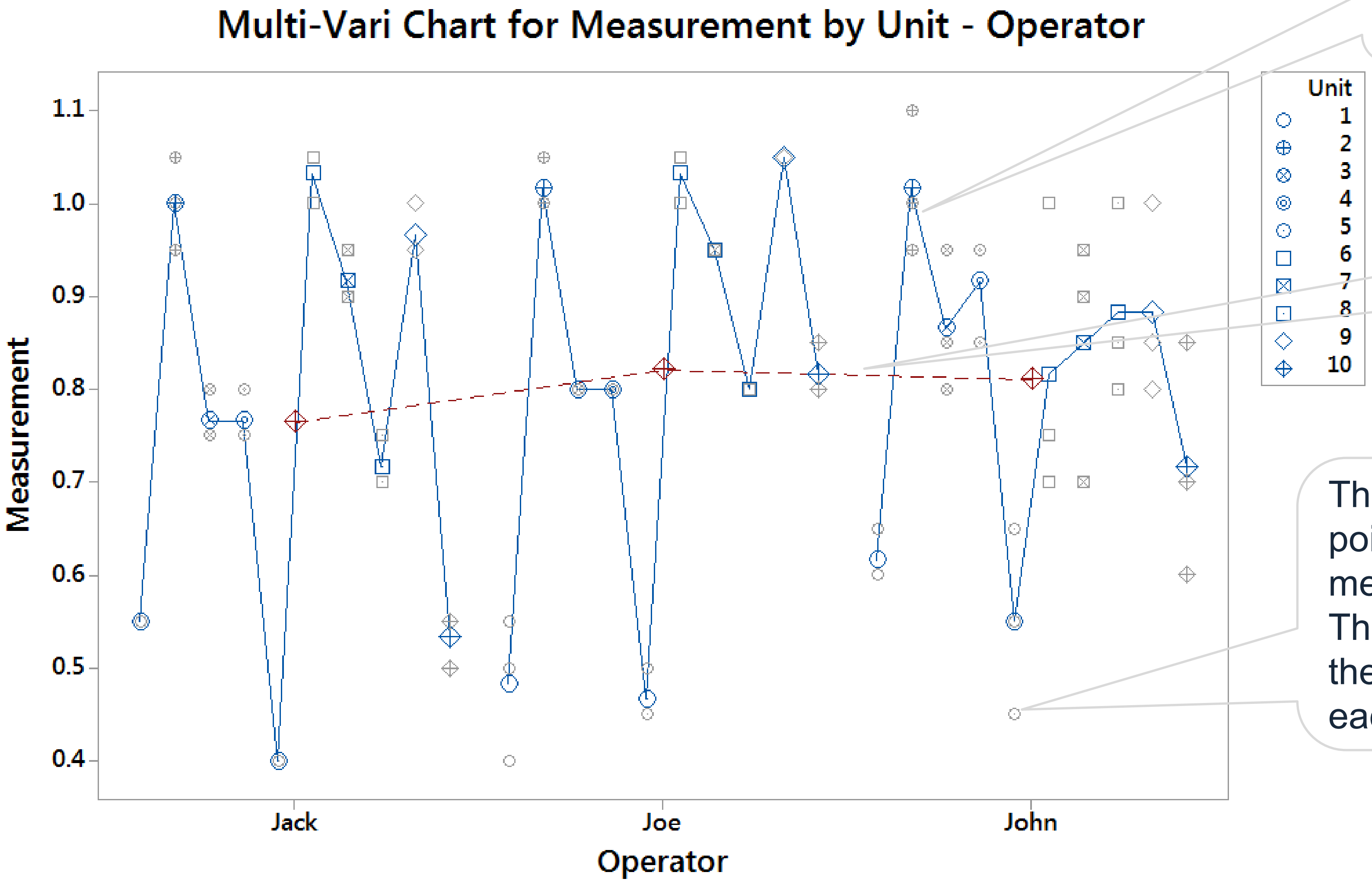
Use Minitab to Perform Multi-Vari Analysis

- Step 4:
  1) A Multi-Vari chart is generated automatically.
  2) Analyze the chart to determine the major source of the variation in the output.



Interpreting Minitab’s Multi-Vari Analysis

- Here is the Multi-Vari chart.



The blue lines connect all the mean measurements of individual units.

The red dotted line connects the means of individual operators.

The maximum and minimum points present the range of measurements for each unit. The center point presents the mean measurement for each unit.

Interpreting Minitab’s Multi-Vari Analysis

- Based on the Multi-Vari chart, the measurement of units range from 0.4 to 1.1.
- Joe’s and John’s mean measurements stay between 0.8 and 0.9. Jack’s mean is slightly lower than both Joe’s and John’s.
- John has the worst variation when measuring the same kind of unit because John has the highest difference between the maximum and minimum bars for any kind of unit.
- By observing the black lines of three operators, it seems like all three operators’ measurements follow the same pattern. The operator-to-operator variability is not large.
- The unit-to-unit variability is large and it could be the main source of variation in measurements.

3.1.2 Classes of Distribution

What is Probability?

- **Probability** is the likelihood of an event occurring.
- The probability of event A occurring is written as P(A).
- Probability is a percentage between 0% and 100%. The higher the probability, the more likely the event will happen.
- When probability is equal to 0%, it indicates that the event will take place with no chance.
- When probability is equal to 100%, it indicates the event will definitely take place without any uncertainty.
- Example of probability: when tossing a coin, there is 50% chance that the head lands face up and 50% chance that the head lands face down.

Probability Property

- The probability of event A not occurring:

$$P(\overline{A}) = 1 - P(A)$$

To be, or not to be: that is the question.

William Shakespeare

Probability Property

- The probability of event A OR B OR both events occurring:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- If event A and B are mutually exclusive:

$$P(A \cap B) = 0$$

and

$$P(A \cup B) = P(A) + P(B)$$

Probability Property

- The probability of event A AND B occurring together:

$$P(A \cap B) = P(A \mid B)P(B)$$

- If event A and B are independent of each other:

$$P(A \mid B) = P(A)$$

and

$$P(A \cap B) = P(A)P(B)$$

Probability Property

- The probability of event A occurring given event B taking place:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{where} \quad P(B) \neq 0$$

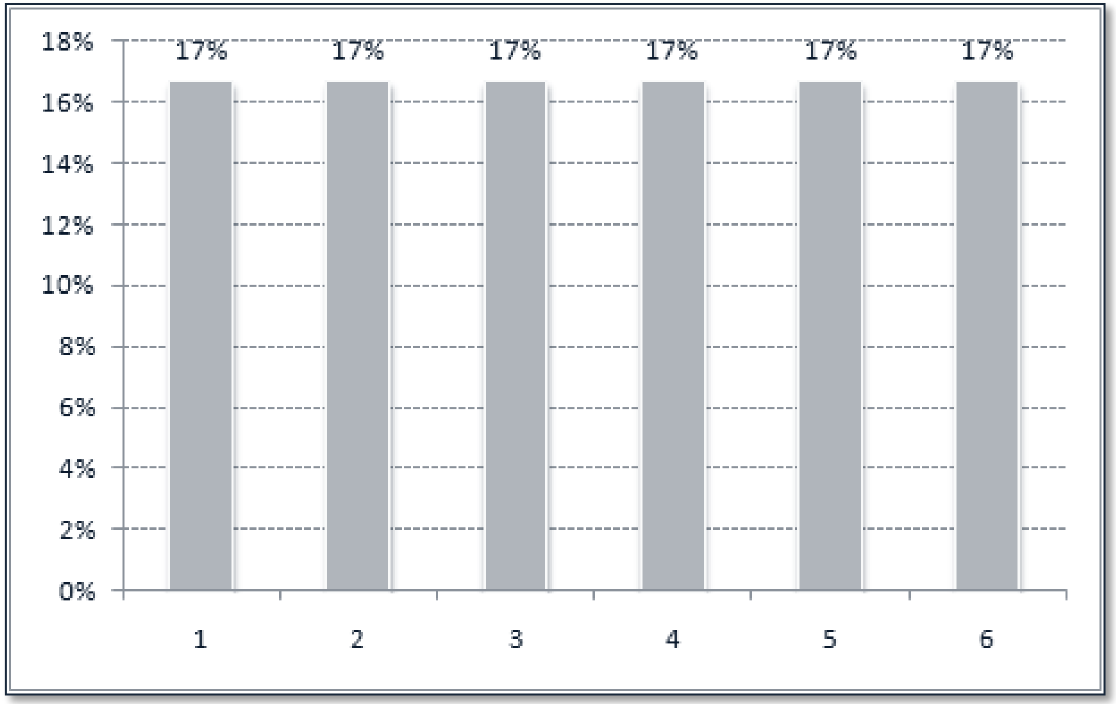
What is Distribution?

- **Distribution** (or probability distribution) describes the range of possible events and the possibility of each event occurring.

- In statistical terms, distribution is the probability of each possible value of a random variable when the variable is discrete, or the probability of a value falling in a specific interval when the variable is continuous.

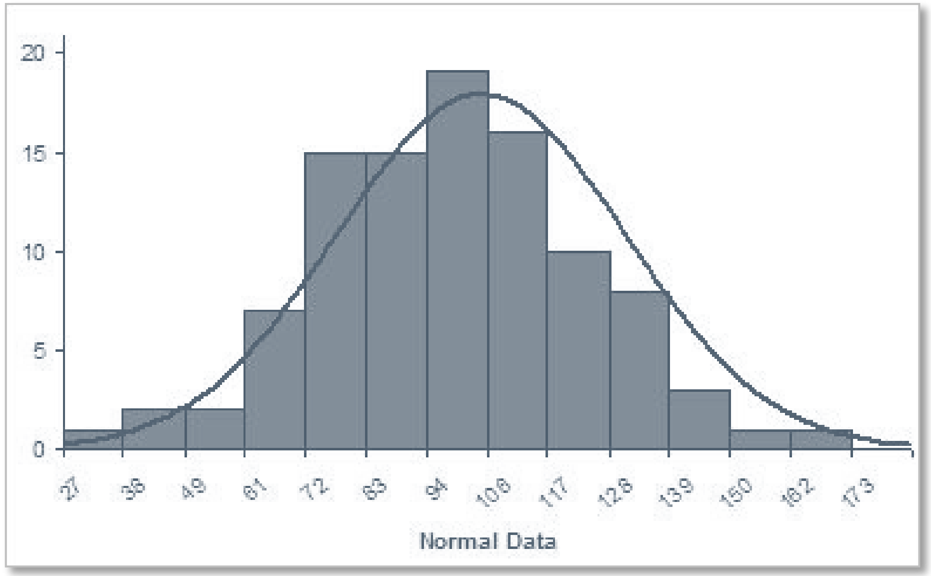
Probability Mass Function

- **PMF** (Probability Mass Function) is a function describing the probability of a discrete random variable equal to a particular value.
- Here is the probability mass function chart for a fair die.



Probability Density Function

- **PDF (Probability Density Function)** is a function used to retrieve the probability of a continuous random variable falling within a particular interval.
- The probability of a variable falling within a particular region is the integral of the probability density function over the region.
- Here is the probability density function chart of a standard normal distribution.



Advantages of Distribution

- Distributions capture the most basic features of data
  - Shape
    - Which distribution family may the data belong to?
    - Is the data symmetric?
    - Are the tails of the data flat?
  - Center
    - Where is the midpoint of the data?
  - Scale
    - What is the range of data?

Measures of Skewness and Kurtosis

- The shape, center (i.e., location), and scale (i.e., variability) are three basic data characteristics that probability distributions capture.
- Skewness and kurtosis are the two more advanced characteristics of the probability distribution.

Skewness

- **Skewness** is a measure of the asymmetry degree of the probability distribution.
- The mathematical definition of the skewness is

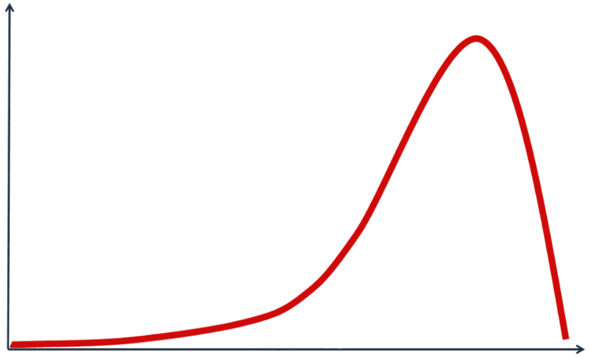
$$\gamma_1 = \mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3}{\sigma^3} = \frac{\mathrm{E}\left[(X-\mu)^3\right]}{\left(\mathrm{E}\left[(X-\mu)^2\right]\right)^{3/2}} = \frac{\kappa_3}{\kappa_2^{3/2}}$$

where μ_3 is the third moment about the mean and σ is the standard deviation.

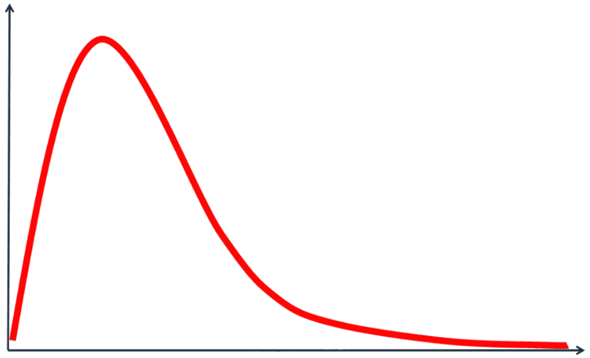
- The skewness of a normal distribution is zero.
- When the distribution looks symmetric to the left side and the right side of the center point, the skewness is close to zero.
- Potential causes of skewness:
  - Extreme values exist in the data
  - Data have a lower or higher bound.

Skewness

- Left Skew
  - Skewness < 0
  - The left tail of the distribution is longer than the right tail.



- Right Skew
  - Skewness > 0
  - The right tail of the distribution is longer than the left tail.



Kurtosis

- **Kurtosis** is a measure of the peakedness of the probability distribution.
- The mathematical definition of the excess kurtosis is

$$\gamma_2 = \frac{\kappa_4}{\kappa_2^2} = \frac{\mu_4}{\sigma^4} - 3$$

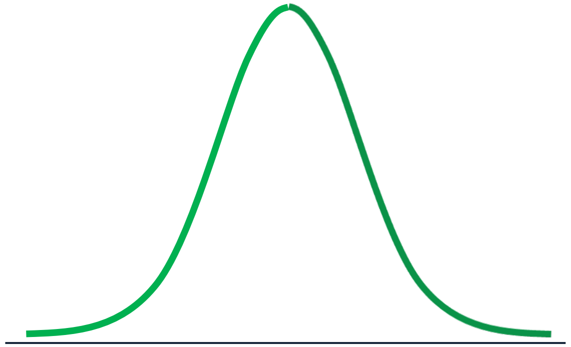
where μ_4 is the fourth moment about the mean and σ is the standard deviation.

- The kurtosis of a normal distribution is zero.
- Distributions with a zero excess kurtosis are called mesokurtic.

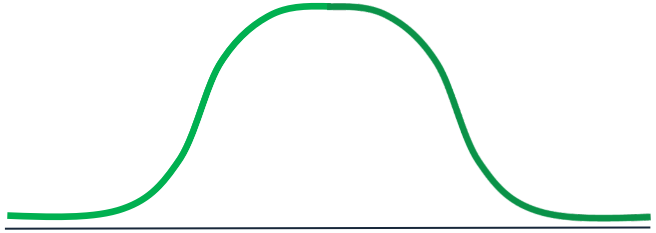
- The distribution with a high kurtosis has a more distinct peak and fatter, longer tails.
- The distribution with a low kurtosis has a more rounded peak and thinner, shorter tails.

Kurtosis

- Leptokurtic
  - Excess Kurtosis > 0
  - A sharper peak near the mean, declining rapidly, and fatter tails.



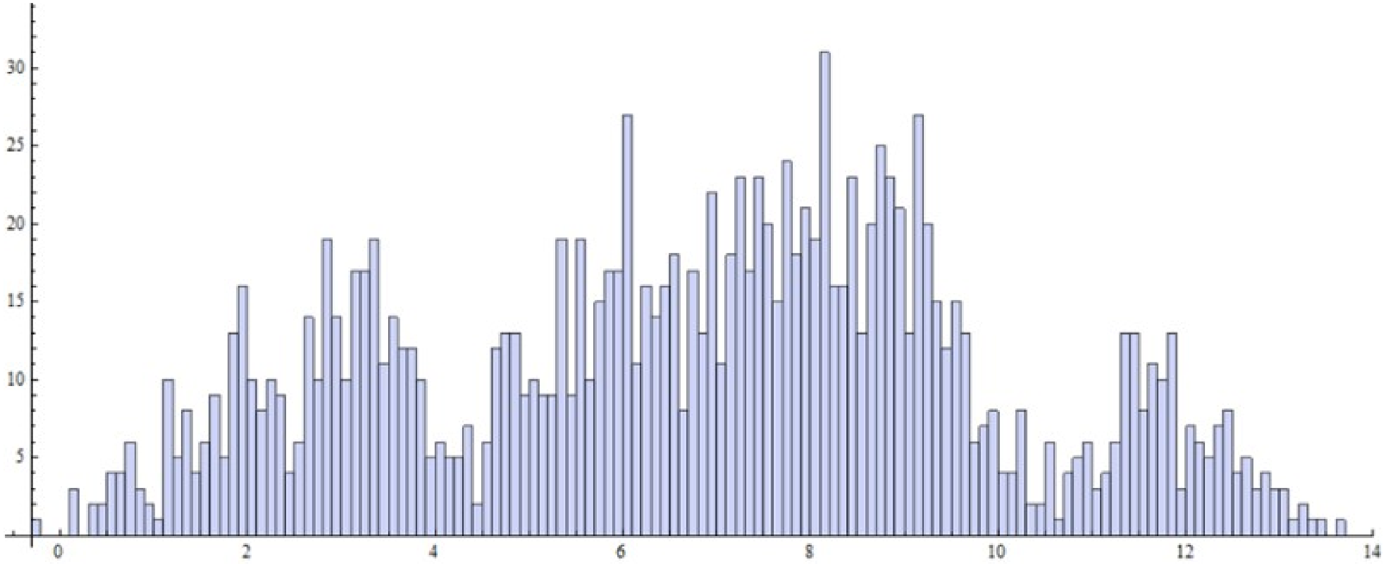
- Platykurtic
  - Excess Kurtosis < 0
  - A flatter top near the mean, declining slowly, and thinner tails.



Multimodal Distribution

- In statistics, a continuous probability distribution with two or more different unimodal distributions is called a **multimodal distribution**.
- A **unimodal distribution** has only one local maximum which is equal to its global maximum.
- A multimodal probability distribution has more than one local maximum.
  - Example: a bivariate, multimodal distribution

(Image source: S&P 500 Composite 20-year total real returns)



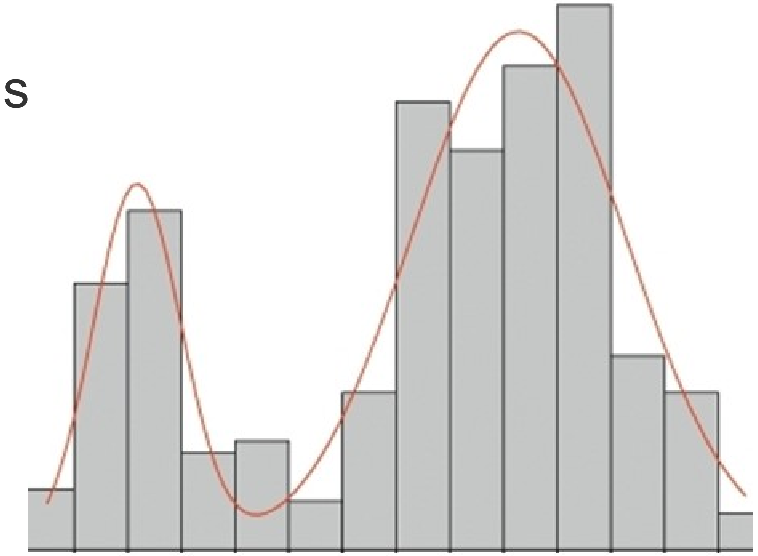
Bimodal Distribution

- In statistics, a continuous probability distribution with two different unimodal distributions is called a **bimodal distribution**.

- A bimodal probability distribution has two local maxima.

- By mixing two unimodal distributions with different means the mixture distribution is not necessarily a bimodal distribution.

- The mixture of two normal distributions with the same standard deviation is a bimodal distribution only if the difference between their means is at least twice their common standard deviation.



(Image source: https://openi.nlm.nih.gov)

Examples of Discrete Distributions

- Binomial Distribution
- Poisson Distribution

Binomial Distribution

- Assume there is an experiment with *n* independent trials, each of which only has two potential outcomes (i.e., yes/no, fail/pass). *p* is the probability of one outcome and (1 – *p*) is the probability of the other.

- **Binomial distribution** is a discrete probability distribution describing the probability of any outcome of the experiment.

- Excel formula for binomial distribution:
  - **BINOMDIST**(**number_s**, **trials**, **probability_s**, **cumulative**)

Binomial Distribution

- PMF of Binomial Distribution

  the probability of getting *k* successes in *n* trials:

$$f(k;n,p)=\binom{n}{k}p^{k}(1-p)^{n-k} \quad \text{where} \quad \binom{n}{k}=\frac{n!}{k!(n-k)!}$$

- Mean of Binomial Distribution

$$n\times p$$

- Variance of Binomial Distribution

$$n\times p\times(1-p)$$

Poisson Distribution

- **Poisson distribution** is a discrete probability distribution describing the probability of a number of events occurring at a known average rate and in a fixed period of time. The expected number of occurrences in this fixed period of time is λ.

- Excel formula for Poisson distribution
  - **POISSON**(**x, mean**, **cumulative**)

Poisson Distribution

- PMF of Poisson Distribution

  the probability of getting k occurrences in *n* trials

$$f(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

- Mean of Poisson Distribution

$$\lambda$$

- Variance of Poisson Distribution

$$\lambda$$

Examples of Continuous Distributions

- Normal Distribution
- Exponential Distribution
- Weibull Distribution
- Student’s t distribution
- Chi-square distribution
- F distribution

Normal Distribution

- **Normal distribution** is a continuous probability distribution describing random variables which cluster around the mean.

- Its probability density function is a bell-shaped curve. But not all the bell-shaped PDFs are from normal distributions.

- It is the most extensively used distribution.

- Excel formula for normal distribution
  - **NORMDIST**(**x**, **mean**, **standard_dev**, **cumulative**)

- The Empirical Rule (68-95-99.7 rule) for normal distribution:
  - about 68% of the data stay within σ from the mean
  - about 95% of the data stay within 2σ from the mean
  - about 99.7% of the data stay within 3σ from the mean.

Normal Distribution

- PDF of Normal Distribution

$$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- Mean of Normal Distribution

$$\mu$$

- Variance of Normal Distribution

$$\sigma^2$$

Exponential Distribution

- **Exponential distribution** is a continuous probability distribution describing the probability of events occurring at a known constant average rate between points of time.

- The exponential distribution is very similar to the Poisson distribution, except that the former is built on continuous variables and the latter is built on discrete variables.

- Excel formula for exponential distribution
  - **EXPONDIST**(**x**, **lambda**, **cumulative**)

Exponential Distribution

- PDF of Exponential Distribution

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq o \\ 0 & x < o \end{cases}$$

- Mean of Exponential Distribution

$$\frac{1}{\lambda}$$

- Variance of Exponential Distribution

$$\frac{1}{\lambda^2}$$

Weibull Distribution

- **Weibull distribution** is a continuous probability distribution which is widely used to model a great variety of data due to its flexibility.

- *K* is the shape parameter and *λ* is the scale parameter. Both are positive numbers.

- Excel formula for Weibull distribution
  - **WEIBULL**(**x**, **alpha**, **beta**, **cumulative**)

Weibull Distribution

- PDF of Weibull Distribution

$$f(x) = \begin{cases} \dfrac{k}{\lambda}\left(\dfrac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)k} & x \geq o \\ 0 & x < o \end{cases}$$

- Mean of Weibull Distribution

$$\lambda\Gamma\left(1+\frac{1}{k}\right)$$

- Variance of Weibull Distribution

$$\lambda^2\Gamma\left(1+\frac{2}{k}\right)-\mu^2$$

Student’s T Distribution

- **Student’s t distribution** is a continuous probability distribution that resembles a normal distribution.

- When *n* random samples were drawn from a normally-distributed population with the mean *μ* and the standard deviation *σ*, then

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

follows a t distribution with (*n* – 1) degrees of freedom.

- Its probability density function is a bell-shaped curve but with heavier tails than those of the normal distribution.

Student’s T Distribution

- PDF of Student’s T Distribution

$$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$$

- Mean of Student’s T Distribution

0 for degrees of freedom greater than 1, otherwise undefined

- Variance of Student’s T Distribution

$$\begin{cases} \nu/(\nu-2) & \nu>2 \\ \infty & 1<\nu\le 2 \\ \text{undefined} & \text{otherwise} \end{cases}$$

Chi-Square Distribution

- **Chi-square distribution** is a continuous probability distribution of the sum of squares of multiple independent standard normal random variables.

- If $Y_1, Y_2, \ldots, Y_k$ are a group of independent normal distributed variables, each with mean 0 and standard deviation 1, then

$$\chi^2 = \sum_{i=1}^{k} Y_i^2$$

follows a chi-square distribution with *k* degrees of freedom.

Chi-Square Distribution

- PDF of Chi-Square Distribution

$$\frac{1}{2^{k/2}\Gamma(k/2)}x^{k/2-1}e^{-x/2}$$

- Mean of Chi-Square Distribution

$$k$$

- Variance of Chi-Square Distribution

$$2k$$

F Distribution

- **F distribution** is a continuous probability distribution which arises in the analysis of variance or test of equality between two variances.

- If $\chi^2_{\nu_1}$ and $\chi^2_{\nu_2}$ are independent variables with chi-square distributions with ν_1 and ν_2 degrees of freedom,

$$F = \frac{\chi^2_{\nu_1} / \nu_1}{\chi^2_{\nu_2} / \nu_2}$$

follows a F distribution.

F Distribution

- PDF of F Distribution

$$\frac{\sqrt{\frac{(\nu_1 x)^{\nu_1}\nu_2{}^{\nu_2}}{(\nu_1 x+\nu_2)^{\nu_1+\nu_2}}}}{xB\left(\frac{\nu_1}{2},\frac{\nu_2}{2}\right)}$$

- Mean of F Distribution

$$\frac{\nu_2}{\nu_2-2}$$

- Variance of F Distribution

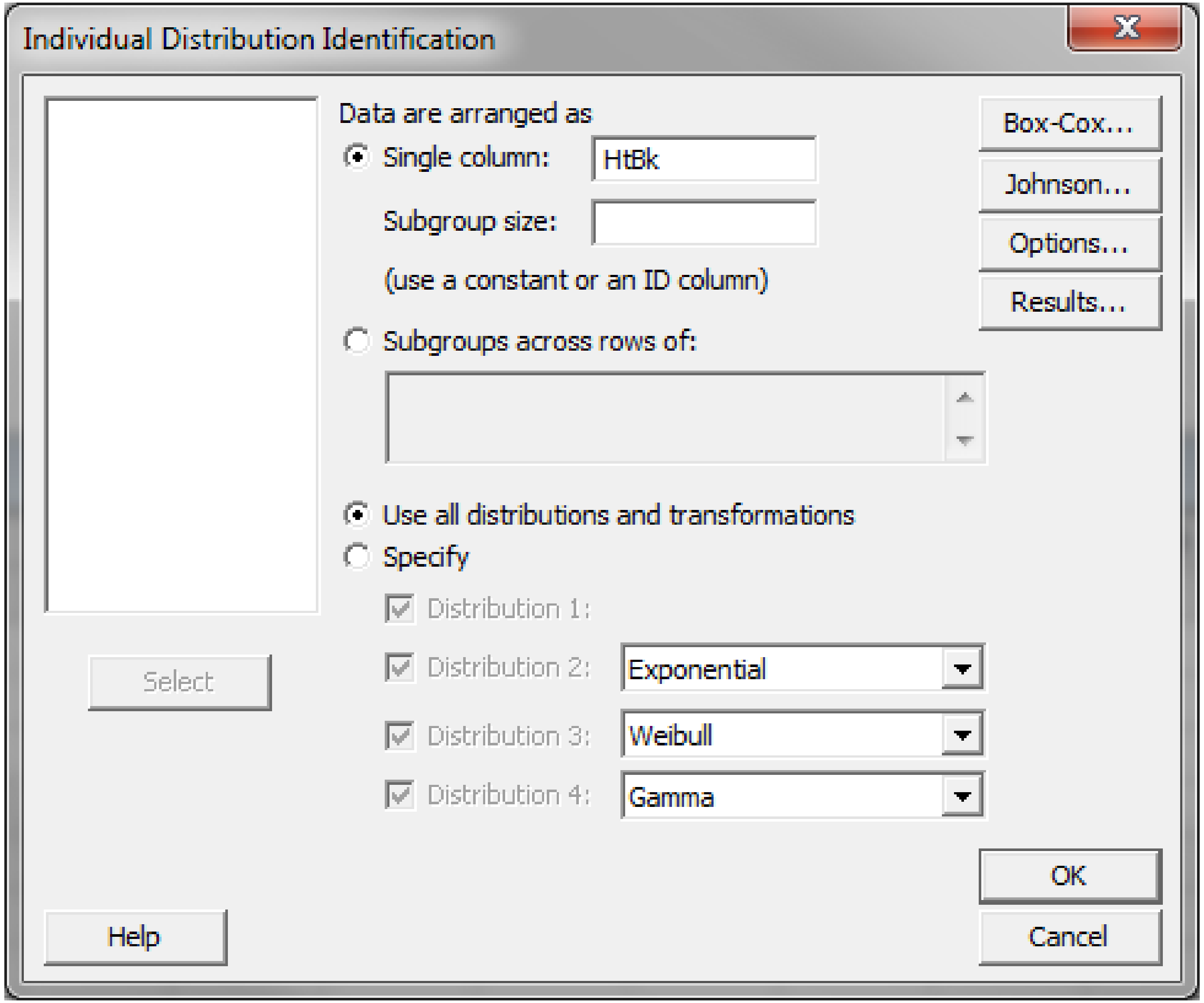
$$\frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}$$

Use Minitab to Fit a Distribution

- *Case study:* we are interested to fit the height data of basketball players into a distribution.
  - Data File: "One Sample T-Test" tab in "Sample Data.xlsx"
- Steps to fit a distribution in Minitab:

1) Click Stat → Quality Tools → Individual Distribution Identification.
2) A new window named "Individual Distribution Identification" pops up.
3) Click in the box next to "Single Column" and the name of the series of our interest appears in the list box on the left.
4) Select "HtBk" as the "Single Column."
5) Click "OK."
6) Multiple windows appear displaying the distribution fitting results.

Use Minitab to Fit a Distribution



Individual Distribution Identification

Data are arranged as

● Single column: HtBk

Subgroup size:

(use a constant or an ID column)

○ Subgroups across rows of:

● Use all distributions and transformations

○ Specify

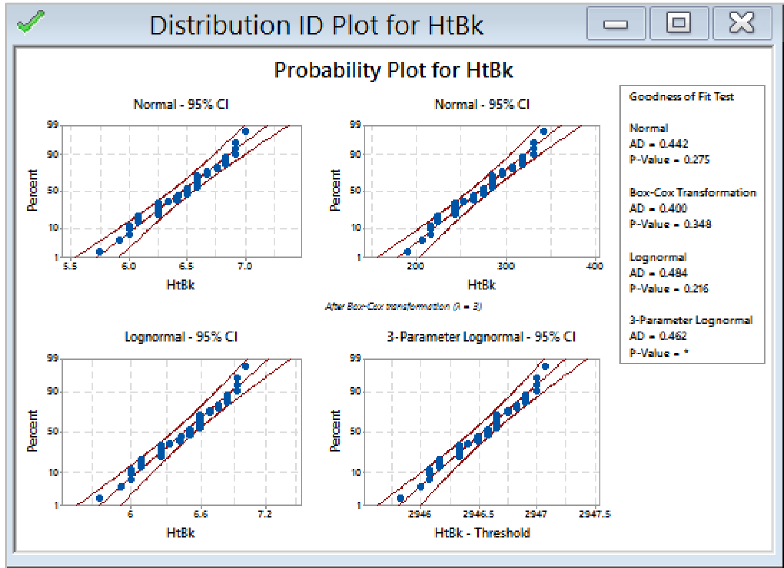
☑ Distribution 1:

☑ Distribution 2: Exponential

☑ Distribution 3: Weibull

☑ Distribution 4: Gamma

Use Minitab to Fit a Distribution


Distribution ID Plot for HtBk

Probability Plot for HtBk

Normal - 95% CI; Normal - 95% CI; Lognormal - 95% CI; 3-Parameter Lognormal - 95% CI

After Box-Cox transformation (λ = 3)

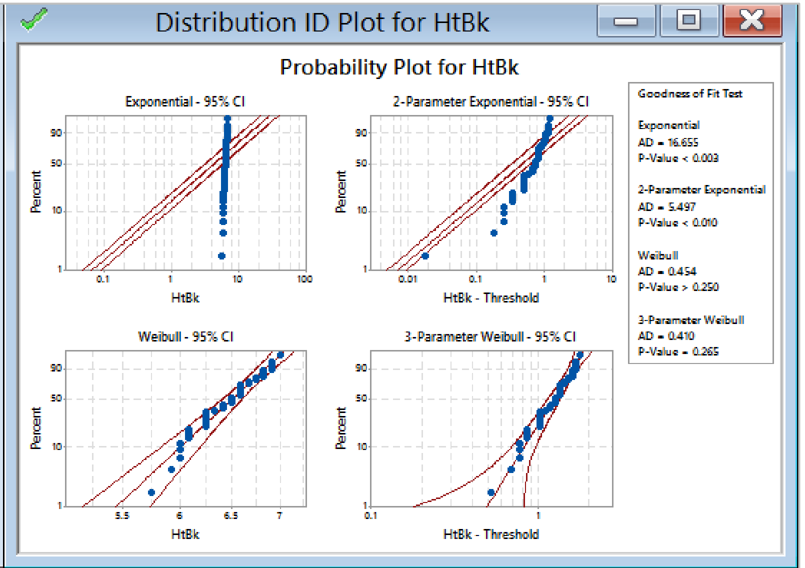
Goodness of Fit Test

Normal
AD = 0.442
P-Value = 0.275

Box-Cox Transformation
AD = 0.400
P-Value = 0.348

Lognormal
AD = 0.484
P-Value = 0.216

3-Parameter Lognormal
AD = 0.462
P-Value = *


Distribution ID Plot for HtBk

Probability Plot for HtBk

Exponential - 95% CI; 2-Parameter Exponential - 95% CI; Weibull - 95% CI; 3-Parameter Weibull - 95% CI

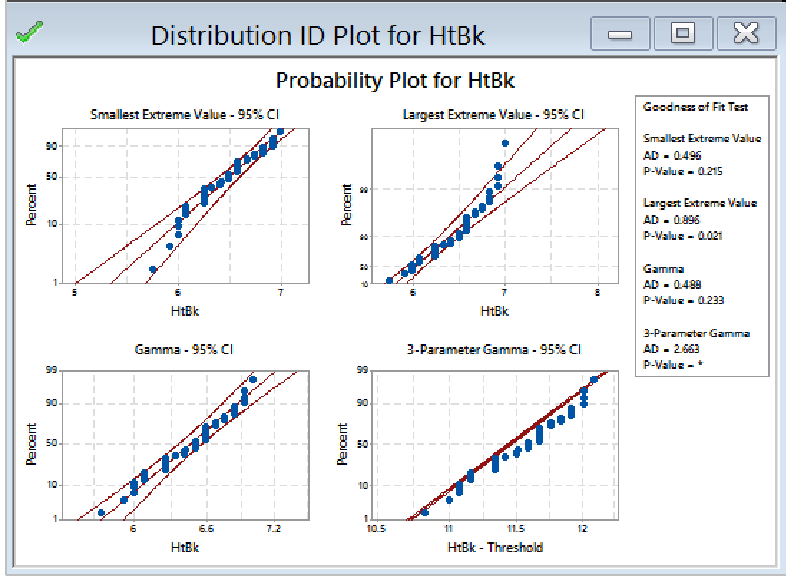
Goodness of Fit Test

Exponential
AD = 16.655
P-Value < 0.003

2-Parameter Exponential
AD = 5.497
P-Value < 0.010

Weibull
AD = 0.454
P-Value > 0.250

3-Parameter Weibull
AD = 0.410
P-Value = 0.265


Distribution ID Plot for HtBk

Probability Plot for HtBk

Smallest Extreme Value - 95% CI; Largest Extreme Value - 95% CI; Gamma - 95% CI; 3-Parameter Gamma - 95% CI

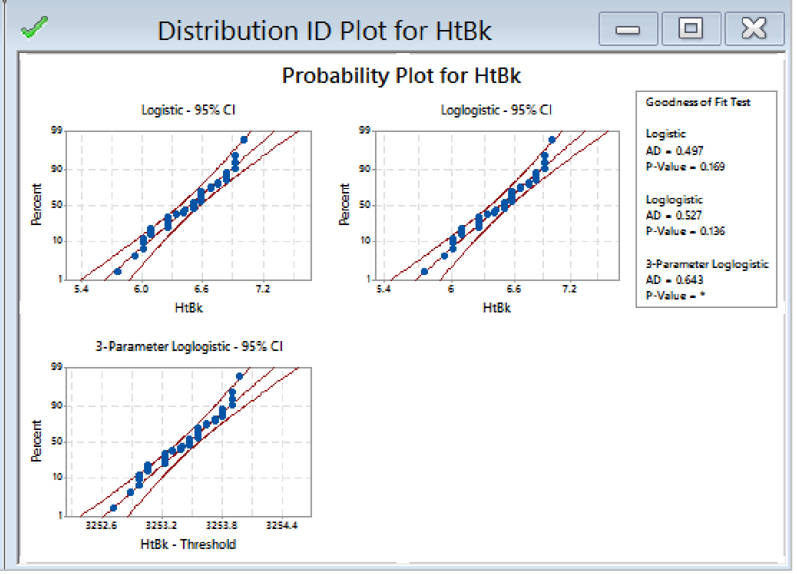
Goodness of Fit Test

Smallest Extreme Value
AD = 0.496
P-Value = 0.215

Largest Extreme Value
AD = 0.896
P-Value = 0.021

Gamma
AD = 0.488
P-Value = 0.233

3-Parameter Gamma
AD = 2.663
P-Value = *


Distribution ID Plot for HtBk

Probability Plot for HtBk

Logistic - 95% CI; Loglogistic - 95% CI; 3-Parameter Loglogistic - 95% CI

Goodness of Fit Test

Logistic
AD = 0.497
P-Value = 0.169

Loglogistic
AD = 0.527
P-Value = 0.136

3-Parameter Loglogistic
AD = 0.643
P-Value = *

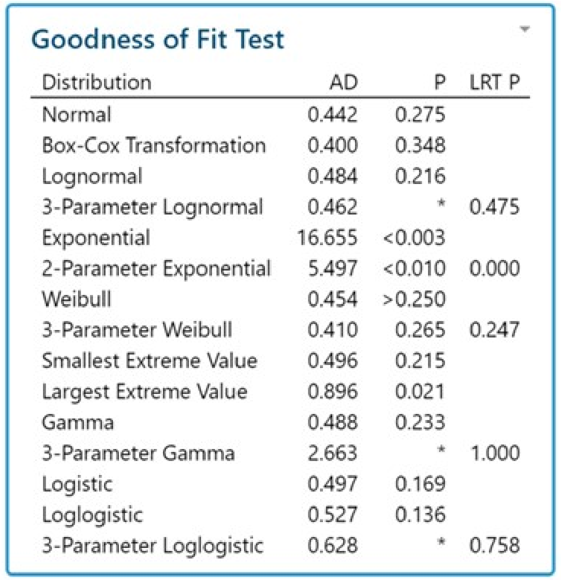
Use Minitab to Fit a Distribution

- The distribution fitting report is displayed in the session window.
- When the p-value of goodness-of-fit test is higher than the alpha level (0.05), the specific distribution fits the data well.

Descriptive Statistics

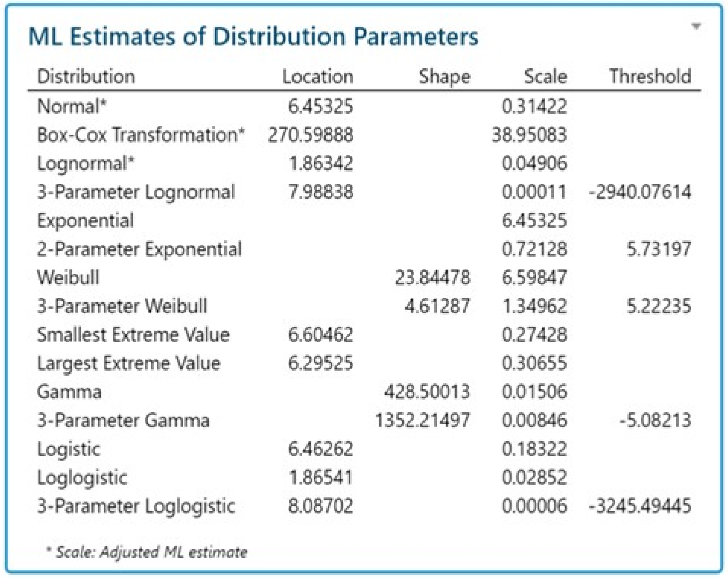
N	N*	Mean	StDev	Median	Minimum	Maximum	Skewness	Kurtosis
40	0	6.45325	0.314222	6.5	5.75	7	-0.235765	-0.684553

Box-Cox transformation: λ = 3



Goodness of Fit Test

Distribution	AD	P	LRT P
Normal	0.442	0.275	
Box-Cox Transformation	0.400	0.348	
Lognormal	0.484	0.216	
3-Parameter Lognormal	0.462	*	0.475
Exponential	16.655	<0.003	
2-Parameter Exponential	5.497	<0.010	0.000
Weibull	0.454	>0.250	
3-Parameter Weibull	0.410	0.265	0.247
Smallest Extreme Value	0.496	0.215	
Largest Extreme Value	0.896	0.021	
Gamma	0.488	0.233	
3-Parameter Gamma	2.663	*	1.000
Logistic	0.497	0.169	
Loglogistic	0.527	0.136	
3-Parameter Loglogistic	0.628	*	0.758



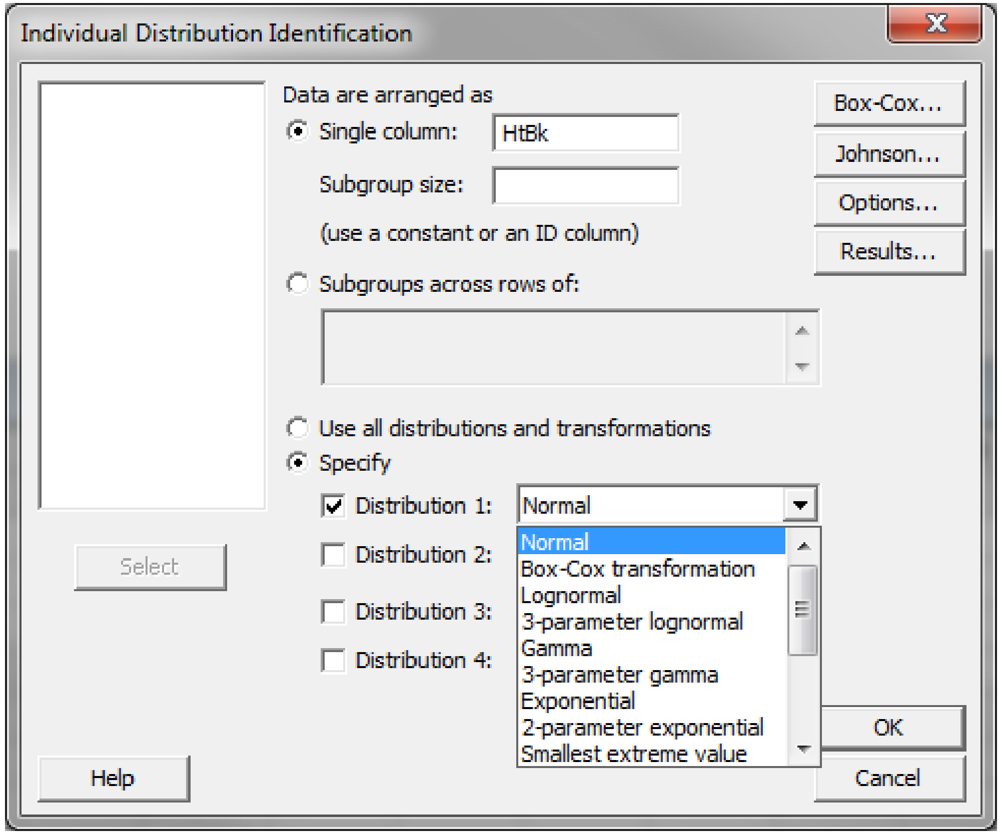
ML Estimates of Distribution Parameters

Distribution	Location	Shape	Scale	Threshold
Normal*	6.45325		0.31422	
Box-Cox Transformation*	270.59888		38.95083	
Lognormal*	1.86342		0.04906	
3-Parameter Lognormal	7.98838		0.00011	-2940.07614
Exponential			6.45325	
2-Parameter Exponential			0.72128	5.73197
Weibull		23.84478	6.59847	
3-Parameter Weibull		4.61287	1.34962	5.22235
Smallest Extreme Value	6.60462		0.27428	
Largest Extreme Value	6.29525		0.30655	
Gamma		428.50013	0.01506	
3-Parameter Gamma		1352.21497	0.00846	-5.08213
Logistic	6.46262		0.18322	
Loglogistic	1.86541		0.02852	
3-Parameter Loglogistic	8.08702		0.00006	-3245.49445

** Scale: Adjusted ML estimate*

Use Minitab to Fit a Distribution

- To fit a specific distribution, in the window "Individual Distribution Identification", click on the radio button "Specify."
- Uncheck the boxes of "Distribution 2," "Distribution 3," and "Distribution 4."
- In the dropdown menu next to "Distribution 1," select the distribution of interest and click "OK".
- In this example, we select "Normal" (i.e., normal distribution).


Individual Distribution Identification

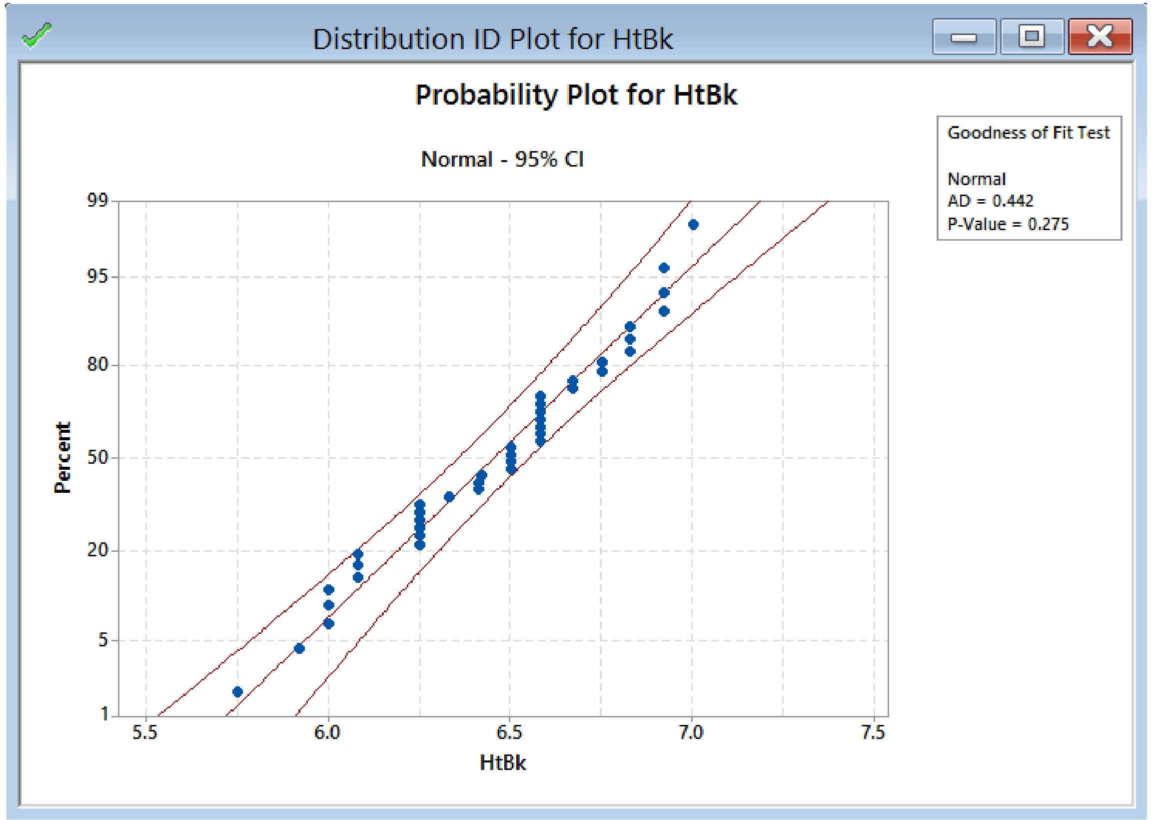
Data are arranged as
● Single column: HtBk
Subgroup size:
(use a constant or an ID column)
○ Subgroups across rows of:

○ Use all distributions and transformations
● Specify
☑ Distribution 1: Normal
☐ Distribution 2:
☐ Distribution 3:
☐ Distribution 4:

Use Minitab to Fit a Distribution

- The result of fitting the specified distribution (normal distribution in this case) is generated in a new window.

- Since the p-value of the goodness-of-fit test is 0.275, higher than the alpha level (0.05), the variable HtBk is normally distributed.



Goodness of Fit Test

Normal
AD = 0.442
P-Value = 0.275

3.2 Inferential Statistics

Black Belt Training: Analyze Phase

3.1 Patterns of Variation
- 3.1.1 Multi-Vari Analysis
- 3.1.2 Classes of Distributions

3.2 Inferential Statistics
- 3.2.1 Understanding Inference
- 3.2.2 Sampling Techniques and Uses
- 3.2.3 Sample Size
- 3.2.4 Central Limit Theorem

3.3 Hypothesis Testing
- 3.3.1 Goals of Hypothesis Testing
- 3.3.2 Statistical Significance
- 3.3.3 Risk; Alpha and Beta
- 3.3.4 Types of Hypothesis Tests

3.4 Hypothesis Testing: Normal Data
- 3.4.1 One and Two Sample T-Tests
- 3.4.2 One sample variance
- 3.4.3 One Way ANOVA

3.5 Hypothesis Testing: Non-Normal Data
- 3.5.1 Mann-Whitney
- 3.5.2 Kruskal-Wallis
- 3.5.3 Moods Median
- 3.5.4 Friedman
- 3.5.5 One Sample Sign
- 3.5.6 One Sample Wilcoxon
- 3.5.7 One and Two Sample Proportion
- 3.5.8 Chi-Squared (Contingency Tables)
- 3.5.9 Test of Equal Variances

3.2.1 Understanding Inference

What is Statistical Inference?

- **Statistical inference** is the process of making inferences regarding the characteristics of an unobservable population based on the characteristics of an observed sample.

- We rely on sample data to draw conclusions about the population from which the sample is drawn.

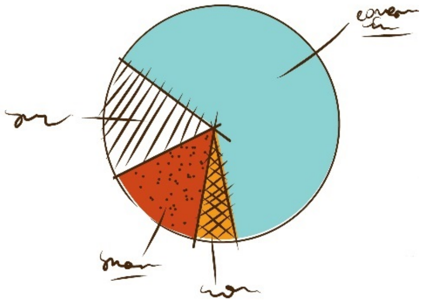
- Statistical inference is widely used since it is difficult or sometimes impossible to collect the entire population data.

Outcome of Statistical Inference

- The outcome or conclusion of statistical inference is a **statistical proposition** about the population.

- Examples of statistical propositions:
    - Estimating a population parameter
    - Identifying an interval or a region where the true population parameter would fall with some certainty
    - Deciding whether to reject a hypothesis made on characteristics of the population of interest
    - Making predictions
    - Clustering or partitioning data into different groups.

Population and Sample



- A **statistical population** is an entire set of objects or observations about which statistical inferences are to be drawn based on its sample.

- It is usually impractical or impossible to obtain the data for the entire population. For example, if we are interested in analyzing the population of all the trees, it is extremely difficult to collect the data for all the trees that existed in the past, exist now, and will exist in the future.

- A **sample** is a subset of the population (like a piece of the pie above). It is necessary for samples to be *representative* of the population.

- The process of selecting a subset of observations within a population is referred to as **sampling**.

Population and Sample

- Population Parameters (Greek letters)
  - Mean: μ
  - Standard deviation: σ
  - Variance: σ^2
  - Median: η

- Sample Statistics (Roman letters)
  - Sample Mean: $\bar{X}$
  - Standard deviation: S
  - Variance: S^2
  - Median: $\tilde{X}$

- The **population parameter** is the numerical summary of a population.
- The **sample statistic** is the numerical measurement calculated based on a sample of that population. It is used to estimate the true population parameter.

Descriptive Statistics vs. Statistical Inference

- **Descriptive Statistics**
  - Descriptive statistics summarize the characteristics of a collection of data.
  - Descriptive statistics are descriptive only and they do not make any generalizations beyond the data at hand.
  - Data used for descriptive statistics are for the purpose of representing or reporting.
- **Statistical Inference**
  - Statistical inference makes generalizations from a sample at hand of a population.
  - Data used for statistical inference are for the purpose of making inferences on the entire population of interest.
- A complete statistical analysis includes *both* descriptive statistics and statistical inference.

Error Sources of Statistical Inference

- Statistical inference uses sample data to best approximate the true features of the population.
- A valid sample must be unbiased and representative of the population.
- Two sources of error in statistical inference:
  - Random sampling error
  - Selection bias.

Error Sources of Statistical Inference

- Random Sampling Error
  - Random variation due to observations being selected randomly
  - It is inherent to the sampling process and beyond one's control.

- Selection Bias
  - Non-random variation due to inadequate design of sampling
  - It can be improved by adjusting the sampling size and sampling strategy.

3.2.2 Sampling Techniques

What is Sampling?

- **Sampling** is the process of selecting objects or observations from a population of interest about which we wish to make a statistical inference.
- It is extensively used to collect information about a population, which can represent physical or intangible objects.

Advantages of Sampling

- It is usually impractical or impossible to collect the data of an entire population:
  - High cost
  - Time consuming
  - Unavailability of historical records
  - Dynamic nature of the population.

- Advantages of sampling a representative subset of the population:
  - Lower cost
  - Faster data collection
  - Easier to manipulate.

Uses of Sample

- With valid samples collected, we can draw statistical conclusions about the population of our interest.

- There are two major uses of samples in making statistical inference:
  - *Estimation*: estimating the population parameters using the sample statistics.
  - *Hypothesis testing*: testing a statement about the population characteristics using sample data.

- This module covers sample size calculation for estimation purposes only. Sample size calculations for hypothesis testing purposes will be covered in the Hypothesis Testing module.

Basic Sampling Steps

1) Determine the population of interest
2) Determine the sampling frame
3) Determine the sampling strategy
4) Calculate the sample size
5) Conduct sampling

Population

- **Population** in statistics is the entire set of objects or observations about which we are interested to draw conclusions or make generalizations based on some representative sample data.
- A population can be either physical or intangible.
  - Physical: trees, customers, monitors etc.
  - Intangible: credit score, pass/fail decisions etc.
- A population can be static or dynamic
  - Characteristics of individuals are relatively static over time
  - Items making up the population continue to change or be generated over time
- The population covers all the items with characteristics we are interested to analyze.

Sampling Frame

- In the ideal situation, the scope of a population could be defined.
  - Example: Mary is interested to know how the soup she is cooking tastes. The population is simply the pot of soup.
- However, in some other situations, the population cannot be identified or defined precisely.
  - Example: to collect the information for an opinion poll, we do not have a list of all the people in the world at hand.

- A **sampling frame** is a list of items of the population (preferably the entire population; if not, approximate population).
  - Example: The telephone directory would be a sampling frame for opinion poll data collection.

Basic Sampling Strategies

- Simple Random Sampling
- Matched Random Sampling
- Stratified Sampling
- Systematic Sampling
- Proportional to Size Sampling
- Cluster Sampling

Simple Random Sampling

- **Simple random samples** are selected in such a way that each item in the population has an equal chance of being selected.

- There is no bias involved in the sample selection. Such selection minimizes the variation between the characteristics of samples and the population.

- It is the basic sampling strategy.

Matched Random Sampling

- **Matched random samples** are samples randomly selected in pairs, each of which has the same attribute.

- Example:
  - Researchers are interested in understanding the weight of twins.
  - Researchers are interested in understanding the patients' blood pressure before and after taking some medicine.

Stratified Sampling

- A population can be grouped or “stratified” into distinct and independent categories. An individual category can be considered as a sub-population. **Stratified samples** are randomly selected in each category of the population.

- The categories can be gender, region, income level etc.

- Stratified sampling requires advanced knowledge of the population characteristics.

- Example: A fruit store wants to measure the quality of all their oranges. They decide to use stratified sampling by region to collect sample data. Since about 40% of their oranges are from California, 40% of the sample is selected from the California oranges sub-population.

Systematic Sampling

- **Systematic samples** are selected at regular intervals based on an ordered list where items in the population are arranged according to a certain criterion.

- Systematic sampling has a random start and then every i^{th} item is selected going forward.

- For example, we are sampling the every 5th unit produced on the production line.

Cluster Sampling

- **Cluster sampling** is a sampling method in which samples are only selected from certain clusters or groups of the population.
- It reduces the cost and time spent on the sampling but bears the risk that the selected clusters are biased.
- For example, selecting samples from the region where researchers are located so that the cost and time spent on travelling is reduced.

Sampling Strategy Decision Factors

- When determining the sampling strategy, we need to consider the following factors:
  - Cost and time constraints
  - Nature of the population of interest
  - Availability of advanced knowledge of the population
  - Accuracy requirement.

Sample Size

- The **sample size** is a critical element that can influence the results of statistical inference.
- The smaller the sample size, the higher the risk that the sample statistic will not reflect the true population parameter.
- The greater the sample size, the more time and money we will spend on collecting the samples.

Sample Size Factors

- Is the variable of interest continuous or discrete?
- How large is the population size?
- How much risk do you want to take regarding missing the true population parameters?
- What is the acceptable margin of error you want to detect?
- How much is the variation in the population?

Sample Size Calculation for Continuous Data

- Sample size equation for continuous data

$$n_0 = \left(\frac{Z_{\alpha/2} \times s}{d} \right)^2$$

where

n_0 is the number of samples.

$Z_{\alpha/2}$ is the Z score when risk level is $\alpha/2$.

  - When α is 0.05, it is 1.96.
  - When α is 0.10, it is 1.65.

s is the estimation of standard deviation in the population

d is the acceptable margin of error.

Sample Size Calculation for Continuous Data

- When the sample size calculated using the formula

$$n_0 = \left(\frac{Z_{\alpha/2} \times s}{d} \right)^2$$

exceeds 5% of the population size, we use a correction formula to calculate the final sample size.

$$n = \frac{n_0}{\left(1 + \frac{n_0}{N}\right)}$$

where
n_0 is the sample size calculated using equation $n_0 = \left(\frac{Z_{\alpha/2} \times s}{d} \right)^2$
N is the population size.

Sample Size Calculation for Discrete Data

- Sample size equation for discrete data

$$n_0 = \left(\frac{Z_{\alpha/2}}{d} \right)^2 \times p \times (1 - p)$$

where

n_0 is the number of samples.

$Z_{\alpha/2}$ is the Z score when risk level is $\alpha/2$.

- When α is 0.05, it is 1.96.
- When α is 0.10, it is 1.65.

s is the estimation of standard deviation in the population.

d is the acceptable margin of error.

p is the proportion of one type of event occurring (e.g., proportion of passes).

$p \times (1 - p)$ is the estimate of variance.

Sample Size Calculation for Discrete Data

- When the sample size calculated using the formula

$$n_0 = \left(\frac{Z_{\alpha/2}}{d} \right)^2 \times p \times (1-p)$$

exceeds 5% of the population size, we use a correction formula to calculate the final sample size.

$$n = \frac{n_0}{\left(1 + \frac{n_0}{N}\right)}$$

where
n_0 is the sample size calculated using equation $n_0 = \left(\frac{Z_{\alpha/2}}{d} \right)^2 \times p \times (1-p)$
N is the population size.

Sampling Errors

- Random Sampling Error
  - Random variation due to observations being selected randomly
  - It is inherent in the sampling process and beyond one’s control.

- Selection Bias
  - Non-random variation due to inadequate design of sampling
  - It can be improved by adjusting the sampling size and sampling strategy.

3.2.3 Sample Size

Sample Size

- The sample size is a critical element that can influence the results of hypothesis testing.
- The smaller the sample size, the higher the risk that the statistical conclusions will not reflect the population relationship.
- The greater the sample size, the more time and money we will spend on collecting the samples.

Sample Size Calculation

- General sample size formula for *continuous data*

$$n = \left(\frac{\left(Z_{\alpha/2} + Z_{\beta} \right) \times s}{d} \right)^2$$

- General sample size formula for *discrete data*

$$n = \left(\frac{Z_{\alpha/2} + Z_{\beta}}{d} \right)^2 \times p \times (1 - p)$$

Sample Size Calculation

- n is the number of observations in the sample.
- α is the risk of committing a false positive error.
- β is the risk of committing a false negative error.
- s is the estimation of standard deviation in the population
- d is the size of effect you want to be able to detect.
- p is the proportion of one type of event occurring (e.g. proportion of passes).

Use Minitab to Calculate the Sample Size

- *Case study:*
  - We are interested in comparing the average retail price of a product between two states.

  - We will run a hypothesis test on the two sample means to determine whether there is a statistically significant difference between the retail price in the two states.

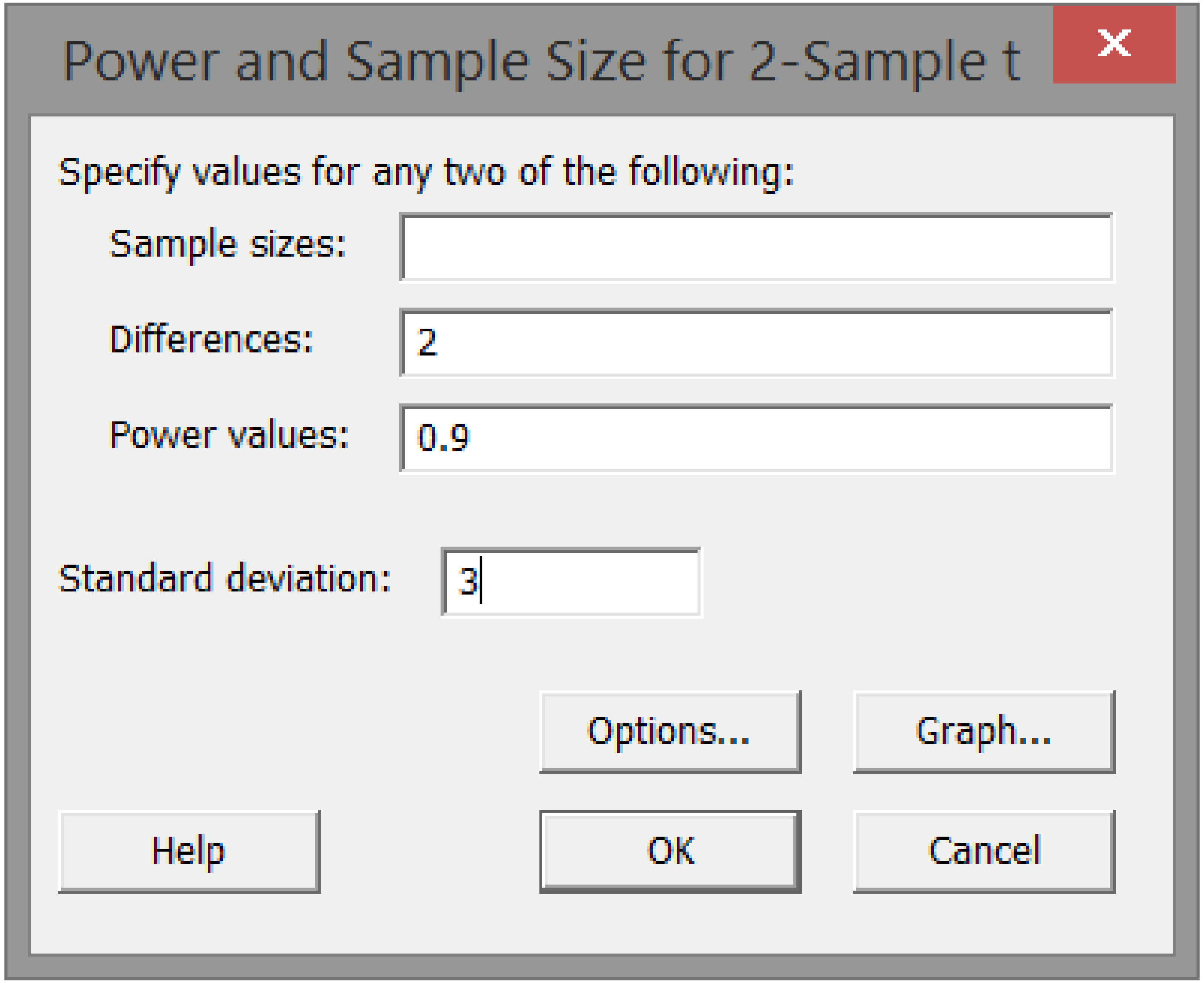
  - The average retail price of the product is $23 based on our estimation and the standard deviation is 3. We want to detect at least a $2 difference with a 90% chance when it is true and we can tolerate the alpha risk at 5%

  - **What should the sample size be?**

Use Minitab to Calculate the Sample Size

- Steps to calculate the sample size in Minitab
  1) Click Stat → Power and Sample Size → 2-Sample t.
  2) A window named “Power and Sample Size for 2-Sample t” appears.
  3) Enter “2” as “Differences.”
  4) Enter “0.9” as “Power values.”
  5) Enter “3” as “Standard deviation.”
  6) Click “OK.”
  7) The sample size calculation results appear in a new window.

Use Minitab to Calculate the Sample Size



Specify values for any two of the following:

Sample sizes:

Differences: 2

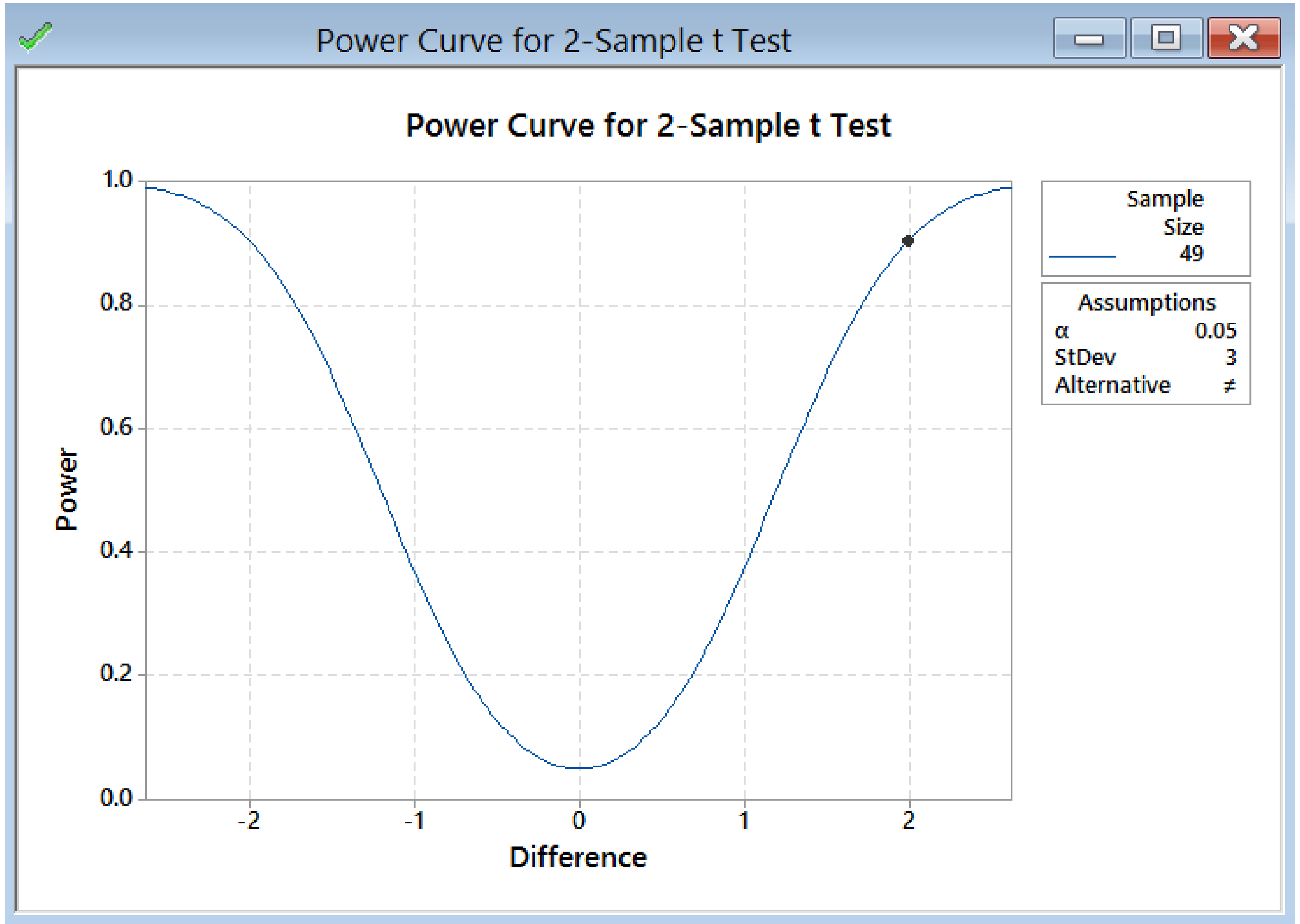
Power values: 0.9

Standard deviation: 3

Use Minitab to Calculate the Sample Size

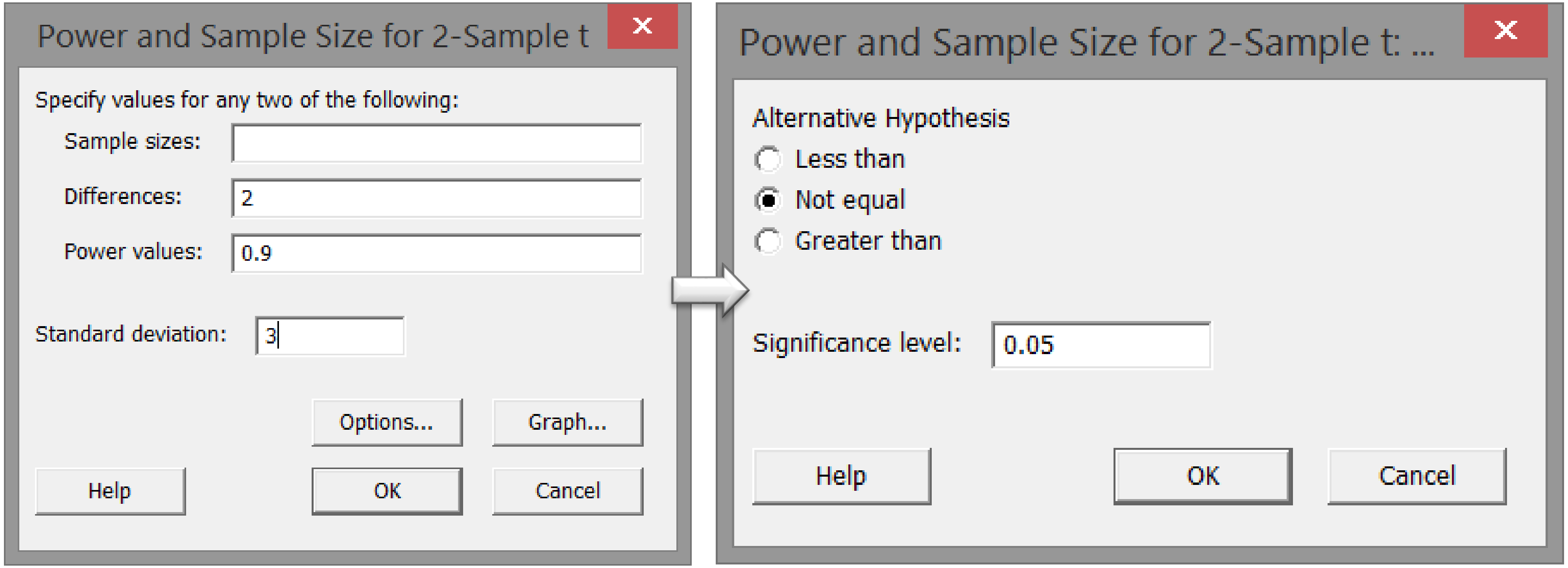
- The sample size for each group is 49 based on the sample size calculator.

- When the absolute value of the difference to detect decreases, the power decreases accordingly if the sample size is kept the same.



Use Minitab to Calculate the Sample Size

- 0.05 is the significance level by default.
- To customize the significance level, click on the “Options” button in the window “Power and Sample Size for 2-Sample t,” specify the significance value in the box, and click “OK.”



Power and Sample Size for 2-Sample t

Specify values for any two of the following:

Sample sizes:

Differences: 2

Power values: 0.9

Standard deviation: 3

Power and Sample Size for 2-Sample t: ...

Alternative Hypothesis

○ Less than
● Not equal
○ Greater than

Significance level: 0.05

3.2.4 Central Limit Theorem

What is Central Limit Theorem?

- The **Central Limit Theorem** is one of the fundamental theorems of probability theory.

- It states a condition under which the mean of a large number of independent and identically-distributed random variables, each of which has a finite mean and variance, would be approximately normally distributed.

What is Central Limit Theorem?

- Let us assume Y_1, Y_2, …, Y_n is a sequence of *n* i.i.d. random variables, each of which has finite mean μ and variance σ^2, where $\sigma^2 > 0$.

- When *n* increases, the sample average of the *n* random variables is approximately normally distributed, with the mean equal to μ and variance equal to σ^2/n, regardless of the common distribution Y_i follows where $i = 1, 2, \ldots, n$.

Independent and Identically Distributed

- A sequence of random variables is **independent and identically distributed** (i.i.d.) if each random variable is independent of others and has the same probability distribution as others.

- It is one of the basic assumptions in Central Limit Theorem.

Central Limit Theorem Example

- Let us assume we have10 fair die at hand.
- Each time we roll all 10 die together we record the average of the 10 die.
- We repeat rolling the die 50 times until we will have 50 data points.
- Upon doing so, we will discover that the probability distribution of the sample average approximates the normal distribution even though a single roll of a fair die follows a discrete uniform distribution.

Central Limit Theorem Explained in Formulas

- Let us assume Y_1, Y_2, …, Y_n are i.i.d. random variables with

$$E(Y_i) = \mu_Y \quad \text{where} \quad -\infty < \mu_Y < \infty$$

$$\operatorname{var}(Y_i) = \sigma_Y^2 \quad \text{where} \quad 0 < \sigma_Y^2 < \infty$$

As $n \rightarrow \infty$, the distribution of $\overline{Y}$ becomes approximately normally distributed with

$$E(\overline{Y}) = \mu_Y$$

$$\operatorname{var}(\overline{Y}) = \sigma_{\overline{Y}}^2 = \frac{\sigma_Y^2}{n}$$

Central Limit Theorem Application

- Use the sample mean to estimate the population mean.
- If the assumptions of Central Limit Theorem are met,

$$E(Y_i) = E(\overline{Y}) \quad \text{where} \quad i = 1,2,\ldots,n$$

Central Limit Theorem Application

- Use standard error of the mean to measure the standard deviation of the sample mean estimate of a population mean.

$$SE_{\overline{Y}} = \frac{s}{\sqrt{n}}$$

where *s* is the standard deviation of the sample and *n* is the sample size.

- Standard deviation of the population mean

$$SD_{\overline{Y}} = \frac{\sigma}{\sqrt{n}}$$

where *σ* is the standard deviation of the population and *n* is the sample size.

Central Limit Theorem Application

- Use a larger sample size, if economically feasible, to decrease the variance of the sampling distribution.
- The larger the sample size, the more precise the estimation of the population parameter.
- Use a confidence interval to describe the region which the population parameter would fall in.
- The sample distribution approximates the normal distribution in which 95% of the data stays within two standard deviations from the center.
- Population mean would fall in the interval of two standard errors of the mean away from the sample mean, 95% of the time.

Confidence Interval

- The **confidence interval** is an interval where the true population parameter would fall within a certain confidence level.
- A 95% confidence interval indicates that the population parameter would fall in that region 95% of the time or we are 95% confident that the population parameter would fall in that region.
- 95% is the most commonly used confidence level.
- Confidence interval is used to describe the reliability of a statistical estimate of a population parameter.

Confidence Interval

- The width of a confidence interval depends on:
  - Confidence level
  - Sample size
  - Variability in the data.

- The higher the confidence level, the wider the confidence interval.

- The smaller the sample size, the wider the confidence interval.

- The more variability, the wider the confidence interval.

Confidence Interval of the Mean

- Confidence interval of the population mean μ_Y of a continuous variable Y is

$$\left[\overline{Y} - Z_{\alpha/2} \times \left(\frac{\sigma}{\sqrt{n}}\right), \overline{Y} + Z_{\alpha/2} \times \left(\frac{\sigma}{\sqrt{n}}\right)\right]$$

where

$\overline{Y}$ is the sample mean

σ is the standard deviation of the population

n is the sample size

$Z_{\alpha/2}$ is the Z score when risk level is *α/2*.

  - When *α* is 0.05, it is 1.96.
  - When *α* is 0.10, it is 1.65.

α is (1 – confidence level).

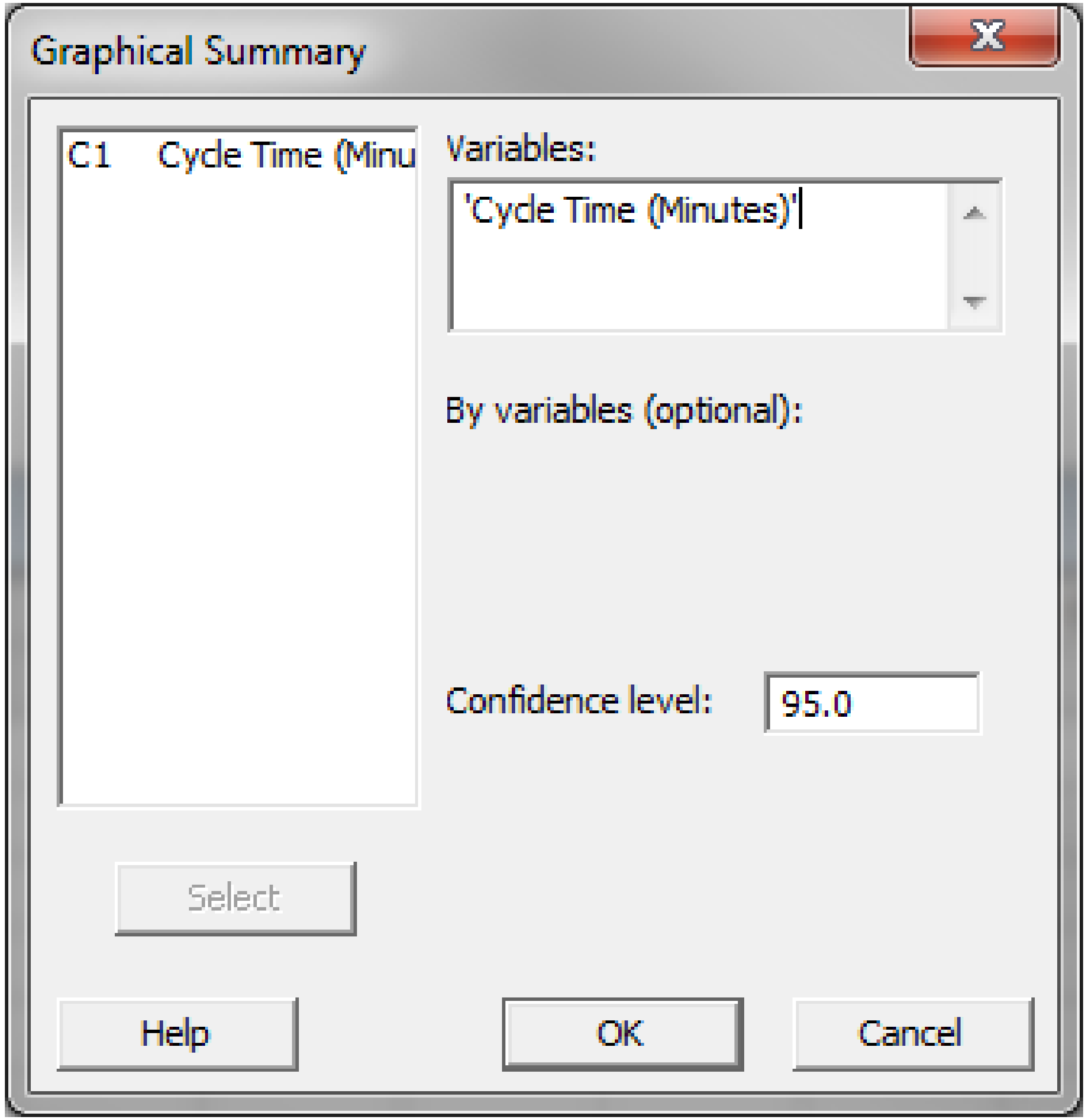
  - When confidence level is 95%, *α* is 5%.
  - When the confidence level is 90%, *α* is 10%.

Special Note: Since 95% is the most commonly used confidence level, 0.05 is the most commonly used α (also called alpha level).

Minitab: Calculate the Confidence Interval of the Mean

- Data File: “Central Limit Theorem” tab in “Sample Data.xlsx”
  1) Click Stat → Basic Statistics → Graphical Summary.
  2) A new window named “Graphical Summary” pops up.
  3) Select “Cycle Time (Minutes)” as the variable.
  4) The confidence level is 0.95 by default.
  5) Click “OK.”
  6) A new window named “Summary for Cycle Time (Minutes)” pops up.

Minitab: Calculate the Confidence Interval of the Mean



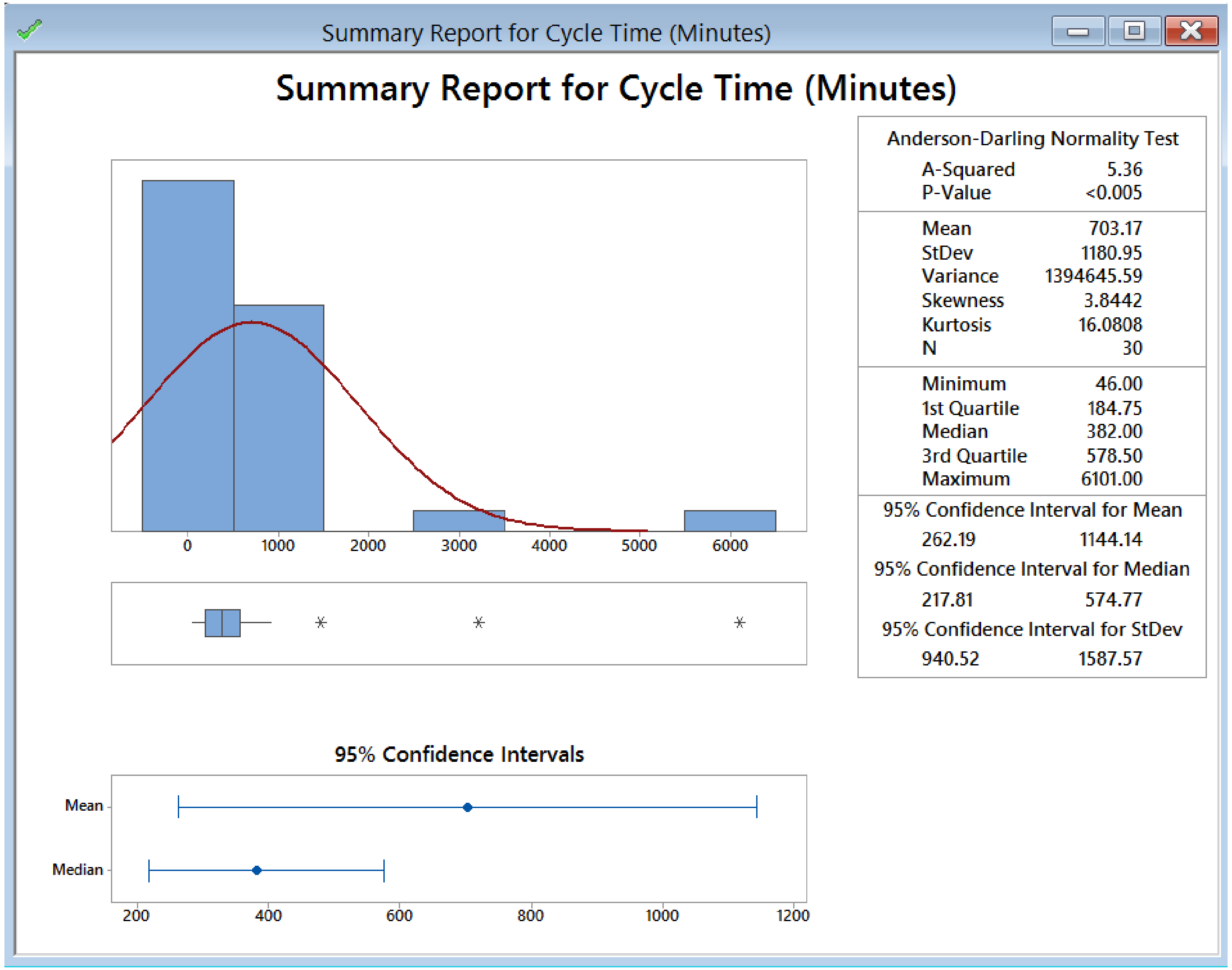
Variables: 'Cycle Time (Minutes)'

By variables (optional):

Confidence level: 95.0

Minitab: Calculate the Confidence Interval of the Mean

The 95% confidence interval of the mean is shown in the newly-generated “Summary for Cycle Time (Minutes).”



Summary Report for Cycle Time (Minutes)

Anderson-Darling Normality Test

A-Squared	5.36
P-Value	<0.005
Mean	703.17
StDev	1180.95
Variance	1394645.59
Skewness	3.8442
Kurtosis	16.0808
N	30
Minimum	46.00
1st Quartile	184.75
Median	382.00
3rd Quartile	578.50
Maximum	6101.00

95% Confidence Interval for Mean

262.19 1144.14

95% Confidence Interval for Median

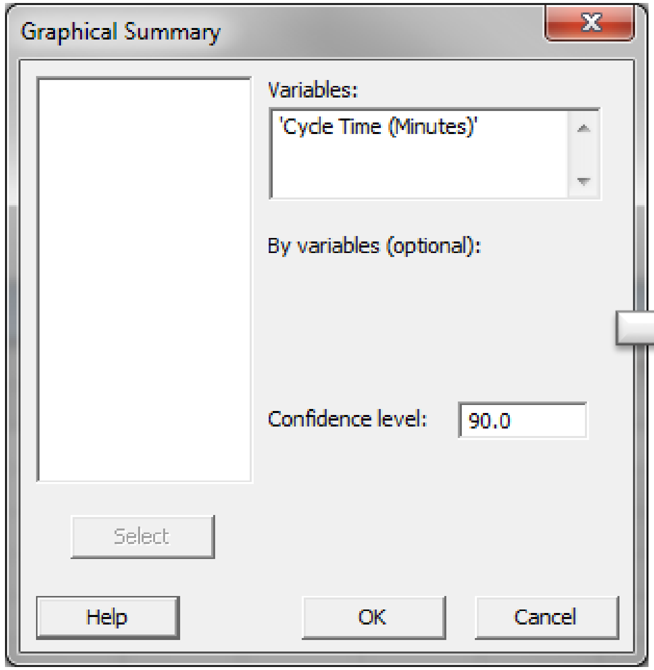
217.81 574.77

95% Confidence Interval for StDev

940.52 1587.57

Minitab: Calculate the Confidence Interval of the Mean

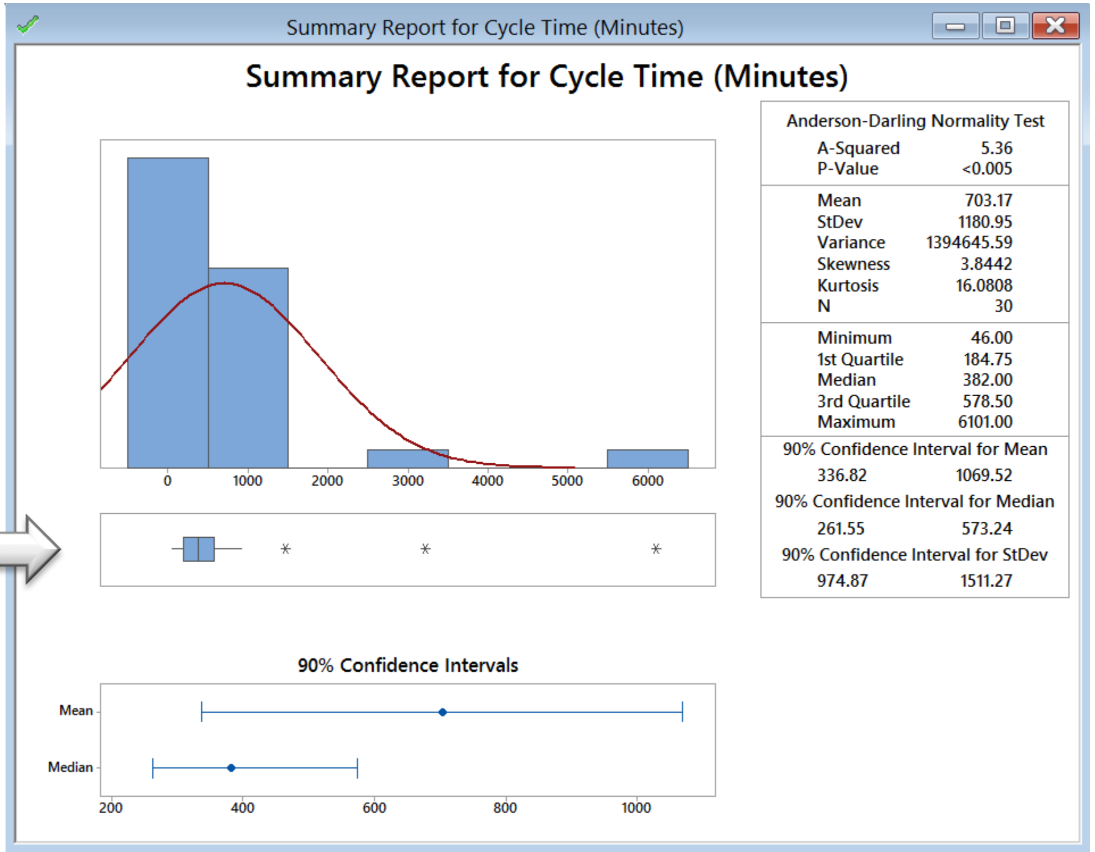
- The confidence level is 95% by default.
- In order to see the confidence interval of “Cycle Time (Minutes)” at other confidence levels, we need to enter the confidence level of our interest in the window “Graphical Summary” and click “OK.”
- The following example shows how to generate 90% confidence interval of the mean.



Variables: 'Cycle Time (Minutes)'

By variables (optional):

Confidence level: 90.0



Summary Report for Cycle Time (Minutes)

Anderson-Darling Normality Test	
A-Squared	5.36
P-Value	<0.005
Mean	703.17
StDev	1180.95
Variance	1394645.59
Skewness	3.8442
Kurtosis	16.0808
N	30
Minimum	46.00
1st Quartile	184.75
Median	382.00
3rd Quartile	578.50
Maximum	6101.00
90% Confidence Interval for Mean	
336.82	1069.52
90% Confidence Interval for Median	
261.55	573.24
90% Confidence Interval for StDev	
974.87	1511.27

3.3 Hypothesis Testing

Black Belt Training: Analyze Phase

3.1 Patterns of Variation
- 3.1.1 Multi-Vari Analysis
- 3.1.2 Classes of Distributions

3.2 Inferential Statistics
- 3.2.1 Understanding Inference
- 3.2.2 Sampling Techniques and Uses
- 3.2.3 Sample Size
- 3.2.4 Central Limit Theorem

3.3 Hypothesis Testing
- 3.3.1 Goals of Hypothesis Testing
- 3.3.2 Statistical Significance
- 3.3.3 Risk; Alpha and Beta
- 3.3.4 Types of Hypothesis Tests

3.4 Hypothesis Testing: Normal Data
- 3.4.1 One and Two Sample T-Tests
- 3.4.2 One sample variance
- 3.4.3 One Way ANOVA

3.5 Hypothesis Testing: Non-Normal Data
- 3.5.1 Mann-Whitney
- 3.5.2 Kruskal-Wallis
- 3.5.3 Moods Median
- 3.5.4 Friedman
- 3.5.5 One Sample Sign
- 3.5.6 One Sample Wilcoxon
- 3.5.7 One and Two Sample Proportion
- 3.5.8 Chi-Squared (Contingency Tables)
- 3.5.9 Test of Equal Variances

3.3.1 Goals of Hypothesis Testing

What is Hypothesis Testing?

- A **hypothesis test** is a statistical method in which a specific hypothesis is formulated about a population, and the decision of whether to reject the hypothesis is made based on sample data.
- Hypothesis tests help to determine whether a hypothesis about a population or multiple populations is true with certain confidence level based on sample data.

Hypothesis Testing Examples

- Hypothesis testing tries to answer whether there is a difference between different groups or there is some change occurring.
    - Are the average SAT scores of graduates from high school A and B the same?
    - Is the error rate of one group of operators higher than that of another group?
    - Are there any non-random causes influencing the height of kids in one state?

What is Statistical Hypothesis?

- A **statistical hypothesis** is an assumption about one or multiple population.
- It is a statement about whether there is any difference between different groups.
- It can be a conjecture about the population parameters or the nature of the population distributions.
- A statistical hypothesis is formulated in pairs:
  - Null Hypothesis
  - Alternative Hypothesis.

Null and Alternative Hypotheses

- *Null Hypothesis* (H_0) states that:
  - there is no difference in the measurement of different groups
  - no changes occurred
  - sample observations result from random chance.

- *Alternative Hypothesis* (H_1 or H_a) states that:
  - there is a difference in the measurement of different groups
  - some changes occurred
  - sample observations are affected by non-random causes.

Null and Alternative Hypotheses

- A statistical hypothesis can be expressed in mathematical language by using population parameters (Greek letters) and mathematical symbols.

- Population Parameters (Greek letters)
  - Mean: μ
  - Standard deviation: σ
  - Variance: σ^2
  - Median: η

- Mathematical Symbols
  - Equal: =
  - Not equal: ≠
  - Greater than: >
  - Smaller than: <

Null and Alternative Hypotheses

- Examples of null and alternative hypotheses written in mathematical language.

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

$$\begin{cases} H_0: \sigma_1 = 0 \\ H_1: \sigma_1 \neq 0 \end{cases}$$

$$\begin{cases} H_0: \eta_1 = 10 \\ H_1: \eta_1 > 10 \end{cases}$$

$$\begin{cases} H_0: \mu_1 = 10 \\ H_1: \mu_1 < 10 \end{cases}$$

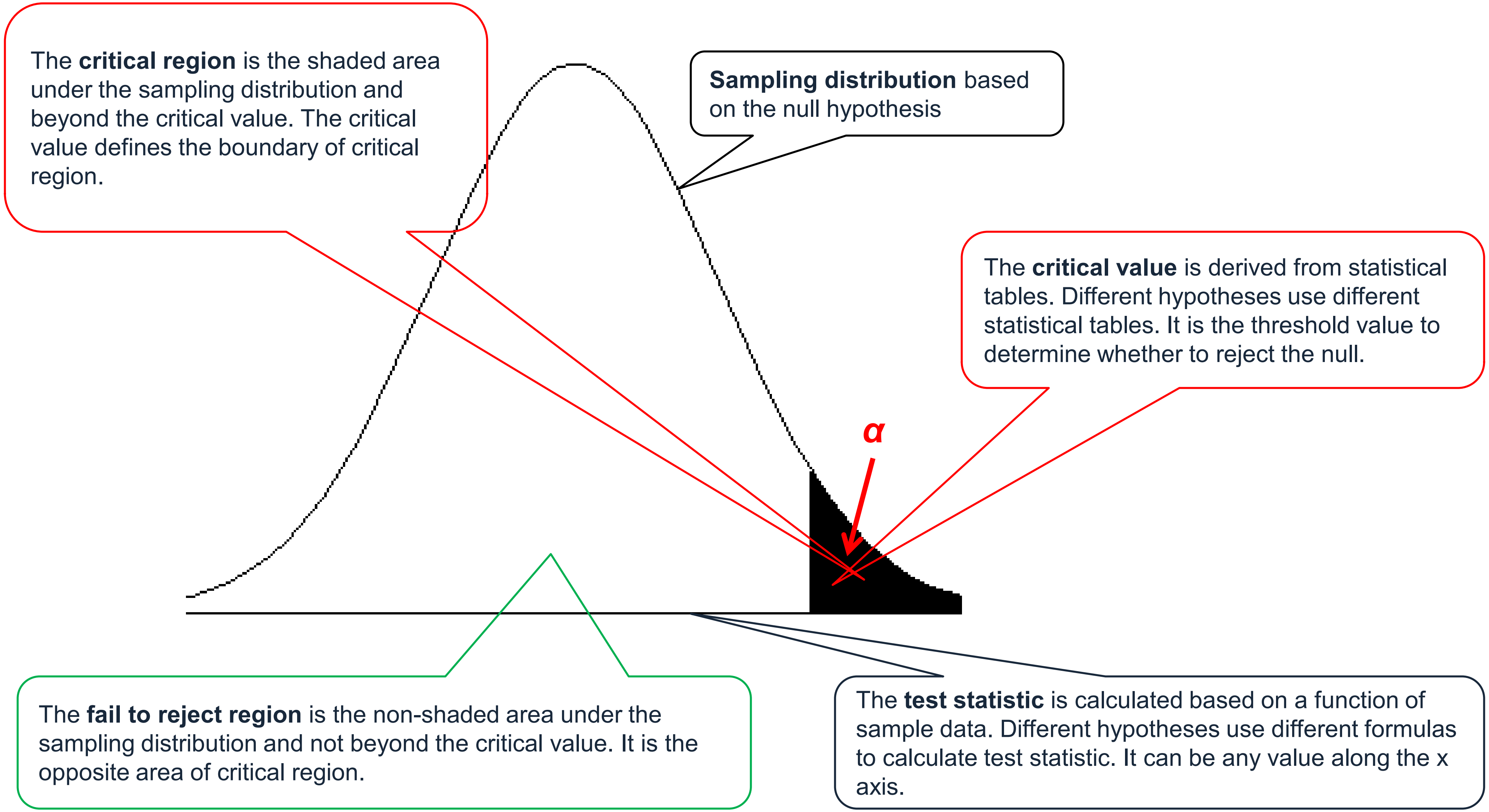
Hypothesis Testing Conclusion

- There are two possible conclusions of hypothesis testing:
    - Reject the null
    - Fail to reject the null.

- When there is enough evidence based on the sample information to prove the alternative hypothesis, we reject the null.

- When there is *not* enough evidence or the sample information is *not* sufficiently persuasive, we fail to reject the null.

Decision Rules in Hypothesis Testing



The **critical region** is the shaded area under the sampling distribution and beyond the critical value. The critical value defines the boundary of critical region.

Sampling distribution based on the null hypothesis

The **critical value** is derived from statistical tables. Different hypotheses use different statistical tables. It is the threshold value to determine whether to reject the null.

The **fail to reject region** is the non-shaded area under the sampling distribution and not beyond the critical value. It is the opposite area of critical region.

The **test statistic** is calculated based on a function of sample data. Different hypotheses use different formulas to calculate test statistic. It can be any value along the x axis.

Decision Rules in Hypothesis Testing

- The **test statistic** in hypothesis testing is a value calculated using a function of the sample.
- Test statistics are considered the sample data's numerical summary that can be used in hypothesis testing.
- Different hypothesis tests have different formulas to calculate the test statistic.
- The **critical value** in hypothesis testing is a threshold value to which the test statistic is compared in order to determine whether the null hypothesis is rejected.
- The critical value is obtained from statistical tables.
- Different hypothesis tests need different statistical tables for critical values.

Decision Rules in Hypothesis Testing

- When the test statistic falls into the fail to reject region, we fail to reject the null and claim that there is no statistically significant difference between the groups.

- When the test statistic falls into the critical region, we reject the null and claim that there is a statistically significant difference between the groups.

Decision Rules in Hypothesis Testing

- The proportion of the area under the sampling distribution and beyond the critical value indicates α risk (also called *α level*). The most commonly selected α level is 5%.

- The proportion of the area under the sampling distribution and beyond the test statistic is the *p-value*. It is the probability of getting a test statistic at least as extreme as the observed one, given the null is true.

Decision Rules in Hypothesis Testing

- When the p-value is smaller than the α level, we reject the null and claim that there is a statistically significant difference between different groups.

- When the p-value is higher than the α level, we fail to reject the null and claim that there is no statistically significant difference between different groups.

Steps in Hypothesis Testing

- Step 1: State the null and alternative hypothesis.
- Step 2: Determine α level.
- Step 3: Collect sample data.
- Step 4: Select a proper hypothesis test.
- Step 5: Run the hypothesis test.
- Step 6: Determine whether to reject the null.

3.3.2 Statistical Significance

Statistical Significance

- In statistics, an observed difference is *statistically significant* if it is unlikely that the difference occurred by pure chance, given a predetermined probability threshold.

- Statistical significance indicates that there are some non-random factors causing the result to take place.

- The statistical significance level in hypothesis testing indicates the amount of evidence which is sufficiently persuasive to prove that a difference between groups exists not due to random chance alone.

Practical Significance

- An observed difference is *practically significant* when it is large enough to make a practical difference.
- A difference between groups that is *statistically significant* might not be large enough to be practically significant.
- In some business situations, statistical differences can have little to no meaning because the difference is not large enough to be practical for a business to act upon.

Example

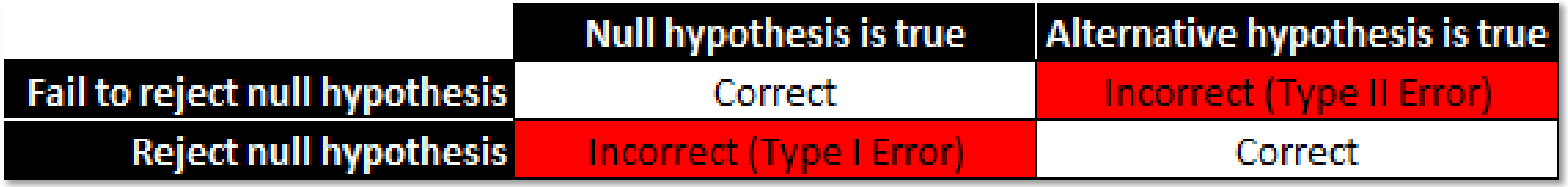
- You started to use the premium gas recently, which was supposed to make your car run better.
- After running a controlled experiment to measure the performance of the car before and after using the premium gas, you performed a statistical hypothesis test and found that the difference before and after was statistically significant.
- Using premium gas did improve the performance.

- However, due to the high cost of the premium gas, you decided that the difference was not large enough to make you pay extra money for it. In other words, the difference is not practically significant.

3.3.3 Risk; Alpha & Beta

Errors in Hypothesis Testing

- In statistical hypothesis testing, there are two types of errors:
  - *Type I Error*
    - a null hypothesis is rejected when it is true in fact.
  - *Type II Error*
    - a null hypothesis is not rejected when it is not true in fact.



	Null hypothesis is true	Alternative hypothesis is true
Fail to reject null hypothesis	Correct	Incorrect (Type II Error)
Reject null hypothesis	Incorrect (Type I Error)	Correct

Type I Error

- Type I error is also called false positive, false alarm, or alpha (*α)* error.
- Type I error is associated with the risk of accepting false positives.
- It occurs when we think there is a difference between groups but in fact there is none.
- Example: telling a patient he is sick and in fact he is not.

Alpha (α)

- *α* indicates the probability of making a type I error. It ranges from 0 to 1.
- *α* risk is the risk of making a type I error.
- *5%* is the most commonly used *α.*
- (*100% – α*) is the confidence level which is used to calculate the confidence intervals.
- When making a decision on whether to reject the null, we compare the p-value against *α:*
  - If p-value is smaller than *α*, we reject the null
  - If p-value is greater than *α*, we fail to reject the null.
- To reduce the *α* risk, we decrease the *α* value to which the p-value is compared.

Type II Error

- Type II error is also called false negative, oversight, or beta (β) error.
- Type II error is associated with the risk of accepting false negatives.
- It occurs when we think there is not any difference between groups but in fact there is.
- Example: telling a patient he is not sick and in fact he is.

Beta (β)

- β indicates the probability of making a type II error. It ranges from 0 to 1.
- β risk is the risk of making a type II error.
- *10%* is the most commonly used β*.*
- (*100%* – β) is called *power,* which denotes the probability of detecting a difference between groups when in fact the difference truly exists.
- To reduce the β risk, we increase the sample size.
- When holding other factors constant, β is inversely related to *α.*

3.3.4 Types of Hypothesis Tests

Two Types of Hypothesis Tests

- **Two-tailed hypothesis test**
  - It is also called two-sided hypothesis test
  - It is a statistical hypothesis test in which the critical region is split into two equal areas, each of which stays on one side of the sampling distribution.

- **One-tailed hypothesis test**
  - It is also called one-sided hypothesis test
  - It is a statistical hypothesis test in which the critical region is only on one side of the sampling distribution.

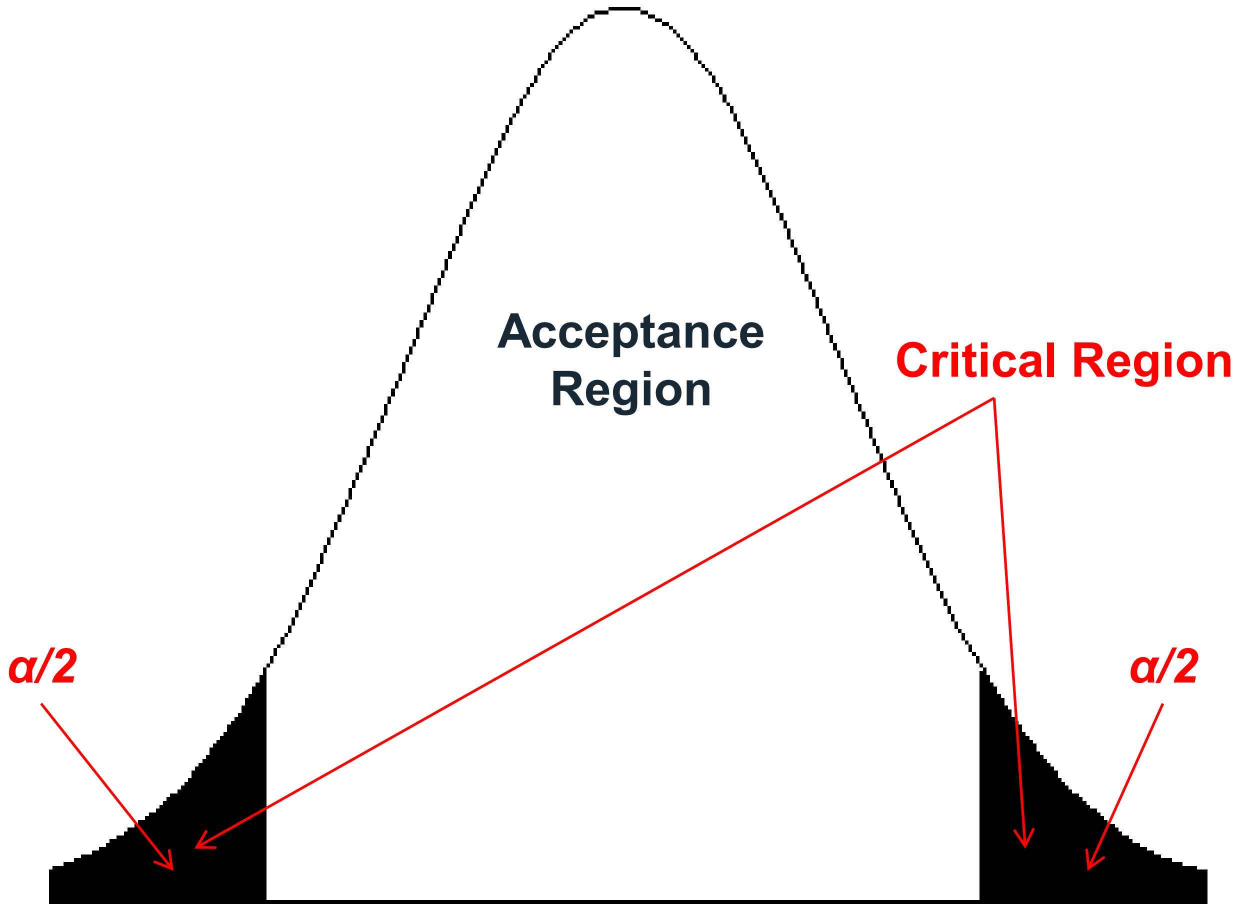
Two-Tailed Hypothesis Test

- A two-tailed hypothesis test is used when we care about whether there is a difference between groups and we do not care about the direction of the difference.
- Examples of a two-tailed hypothesis test:

  H_0: $\mu_1 = 10$
  H_a: $\mu_1 \neq 10$

  The null hypothesis (H_0) is rejected when:
  - the test statistic falls into either half of the critical region
  - the test statistic is either sufficiently small or sufficiently large
  - the absolute value of the test statistic is greater than the absolute value of the critical value.
- A two-tailed hypothesis test is the most commonly used hypothesis test. In the next modules we will cover more details about it.

Two-Tailed Hypothesis Test



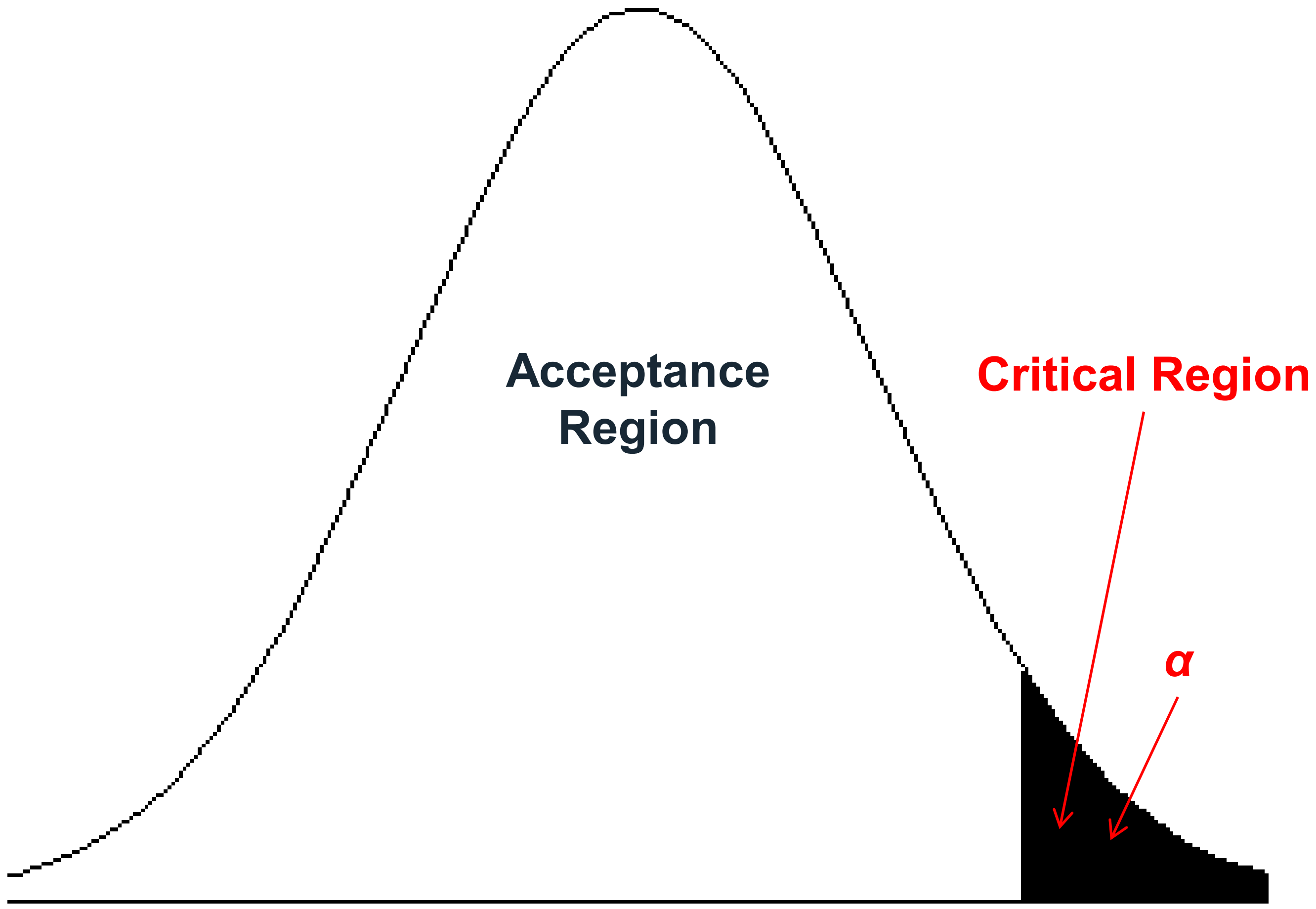
One-Tailed Hypothesis Test

- A one-tailed hypothesis test is used when we care about one direction of the difference between groups.
- Examples of one-tailed hypothesis tests:

$$H_0: \mu_1 = 10$$
$$H_a: \mu_1 > 10$$

- The null hypothesis is rejected when the test statistic:
  - falls in the critical region which exists on the right side of the sampling distribution
  - is sufficiently large
  - is greater than the critical value.

One-Tailed Hypothesis Test



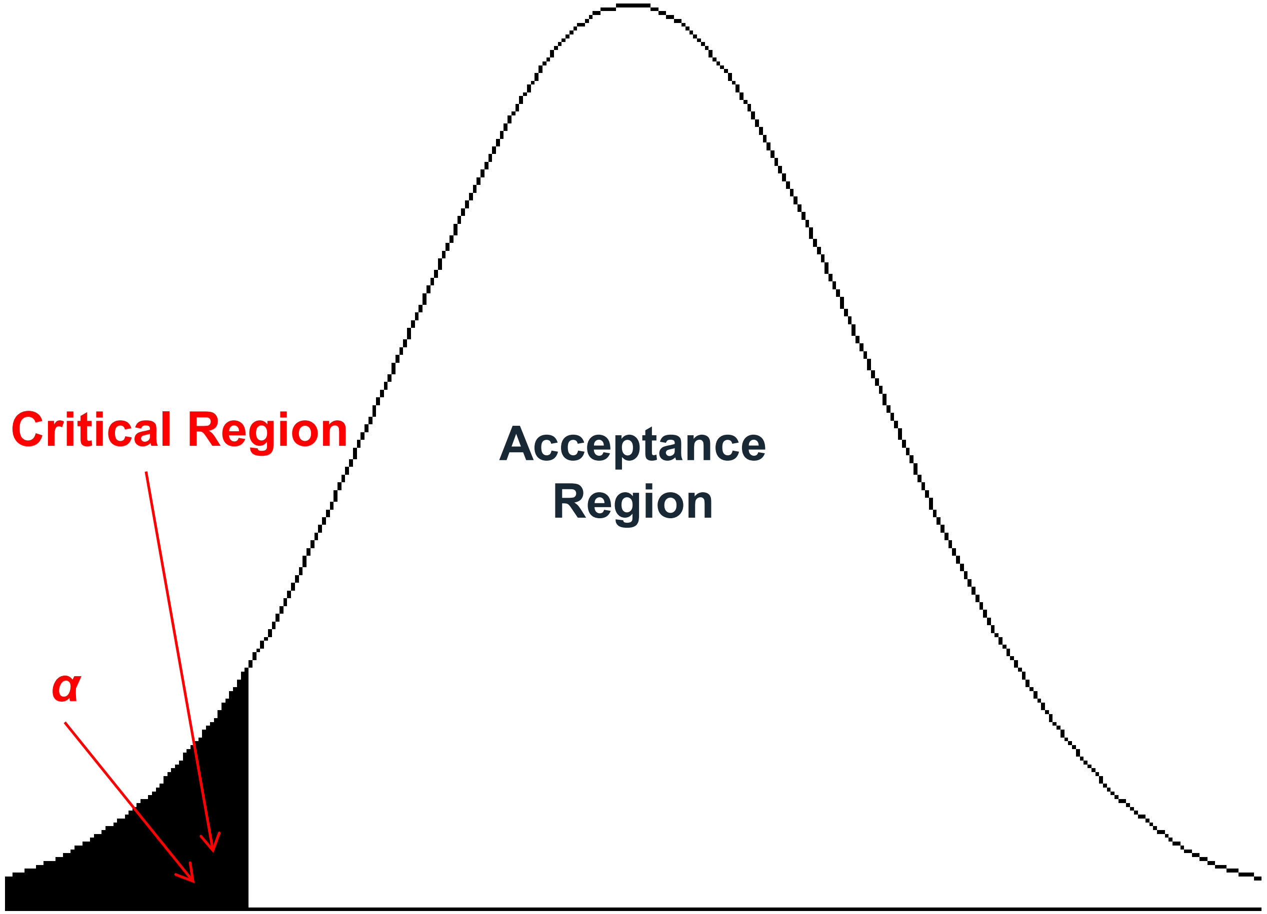
One-Tailed Hypothesis Test

- Examples of one-tailed hypothesis test

$H_0: \mu_1 = 10$
$H_a: \mu_1 < 10$

- The null hypothesis is rejected when the test statistic:
    - falls into the critical region which only exists on the left side of the sampling distribution
    - is sufficiently small
    - is smaller than the critical value.

One-Tailed Hypothesis Test



3.4 Hypothesis Tests: Normal Data

Black Belt Training: Analyze Phase

3.1 Patterns of Variation
- 3.1.1 Multi-Vari Analysis
- 3.1.2 Classes of Distributions

3.2 Inferential Statistics
- 3.2.1 Understanding Inference
- 3.2.2 Sampling Techniques and Uses
- 3.2.3 Sample Size
- 3.2.4 Central Limit Theorem

3.3 Hypothesis Testing
- 3.3.1 Goals of Hypothesis Testing
- 3.3.2 Statistical Significance
- 3.3.3 Risk; Alpha and Beta
- 3.3.4 Types of Hypothesis Tests

3.4 Hypothesis Testing: Normal Data
- 3.4.1 One and Two Sample T-Tests
- 3.4.2 One sample variance
- 3.4.3 One Way ANOVA

3.5 Hypothesis Testing: Non-Normal Data
- 3.5.1 Mann-Whitney
- 3.5.2 Kruskal-Wallis
- 3.5.3 Moods Median
- 3.5.4 Friedman
- 3.5.5 One Sample Sign
- 3.5.6 One Sample Wilcoxon
- 3.5.7 One and Two Sample Proportion
- 3.5.8 Chi-Squared (Contingency Tables)
- 3.5.9 Test of Equal Variances

3.4.1 One & Two Sample T-Tests

What is a T-Test?

- In statistics, a **t-test** is a hypothesis test in which the test statistic follows a *Student t* distribution if the null hypothesis is true.

- We apply a t-test when the population variance (σ) is unknown and we use the sample standard deviation (*s*) instead.

What is One Sample T-Test?

- **One sample t-test** is a hypothesis test to study whether there is a statistically significant difference between a population mean and a specified value.
  - Null Hypothesis (H_0): $\mu = \mu_0$
  - Alternative Hypothesis (H_a): $\mu \neq \mu_0$

where μ is the mean of a population of our interest and μ_0 is the specific value we want to compare against.

Assumptions of One Sample T-Test

- The sample data drawn from the population of interest are unbiased and representative.
- The data of the population are continuous.
- The data of the population are normally distributed.
- The variance of the population of our interest is unknown.
- One sample t-test is more robust than the z-test when the sample size is small (< 30).

Normality Test

- To check whether the population of our interest is normally distributed, we need to run normality test.

  - Null Hypothesis (H_0): The data are normally distributed.
  - Alternative Hypothesis (H_a): The data are not normally distributed.

- There are a lot of normality tests available:
  - Anderson-Darling
  - Sharpiro-Wilk
  - Jarque-Bera etc.

Test Statistic and Critical Value of One Sample T-Test

- Test Statistic

$$t_{calc} = \frac{\bar{Y}}{s/\sqrt{n}}$$, where

$\bar{Y}$ is the sample mean, *n* is the sample size, and s is the sample standard deviation

$$s = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n-1}}$$

- Critical Value
  - t_{crit} is the t-value in a Student t distribution with the predetermined significance level α and degrees of freedom (*n* –1).
  - t_{crit} values for a two-sided and a one-sided hypothesis test with the same significance level α and degrees of freedom (*n* – 1) are different.

Decision Rules of One Sample T-Test

- Based on the sample data, we calculated the test statistic t_{calc}, which is compared against t_{crit} to make a decision of whether to reject the null.

  - Null Hypothesis (H_0): $\mu = \mu_0$
  - Alternative Hypothesis (H_a): $\mu \neq \mu_0$

- If $|t_{calc}| > t_{crit}$, we reject the null and claim there is a statistically significant difference between the population mean μ and the specified value μ_0.

- If $|t_{calc}| < t_{crit}$, we fail to reject the null and claim there is not any statistically significant difference between the population mean μ and the specified value μ_0.

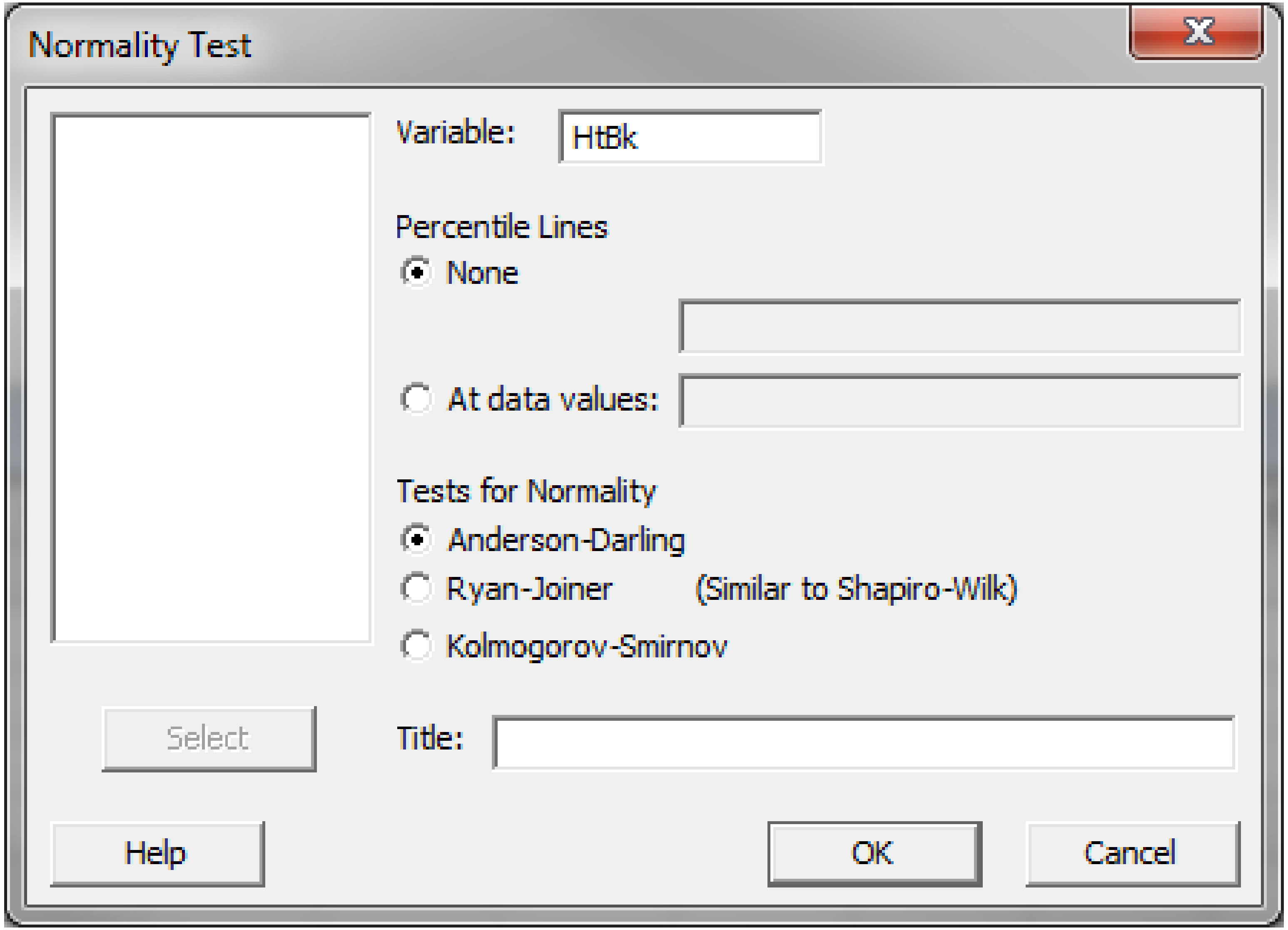
Use Minitab to Run a One-Sample T-Test

- *Case study:* We want to compare the average height of basketball players against 7 feet.
  - Data File: “One Sample T-Test” tab in “Sample Data.xlsx”
  - Null Hypothesis (H_0): $\mu = 7$
  - Alternative Hypothesis (H_a): $\mu \neq 7$

Use Minitab to Run a One-Sample T-Test

- Step 1: Test whether the data are normally distributed
  1) Click Stat → Basic Statistics → Normality Test.
  2) A new window named “Normality Test” pops up.
  3) Select “HtBk” as the variable.
  4) Click “OK.”
  5) A new window named “Probability Plot of HtBk” appears, which covers the results of the normality test.

Use Minitab to Run a One-Sample T-Test



Normality Test

Variable: HtBk

Percentile Lines

● None

○ At data values:

Tests for Normality

● Anderson-Darling

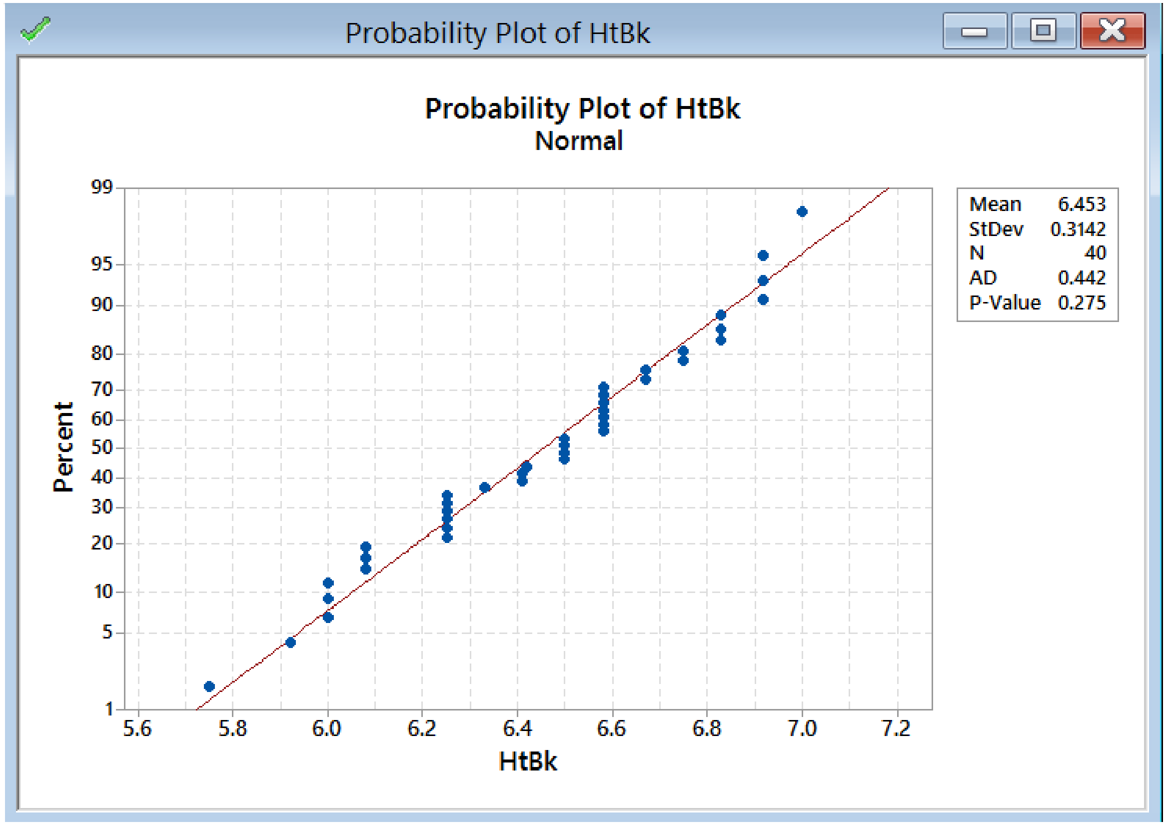
○ Ryan-Joiner (Similar to Shapiro-Wilk)

○ Kolmogorov-Smirnov

Title:

Use Minitab to Run a One-Sample T-Test

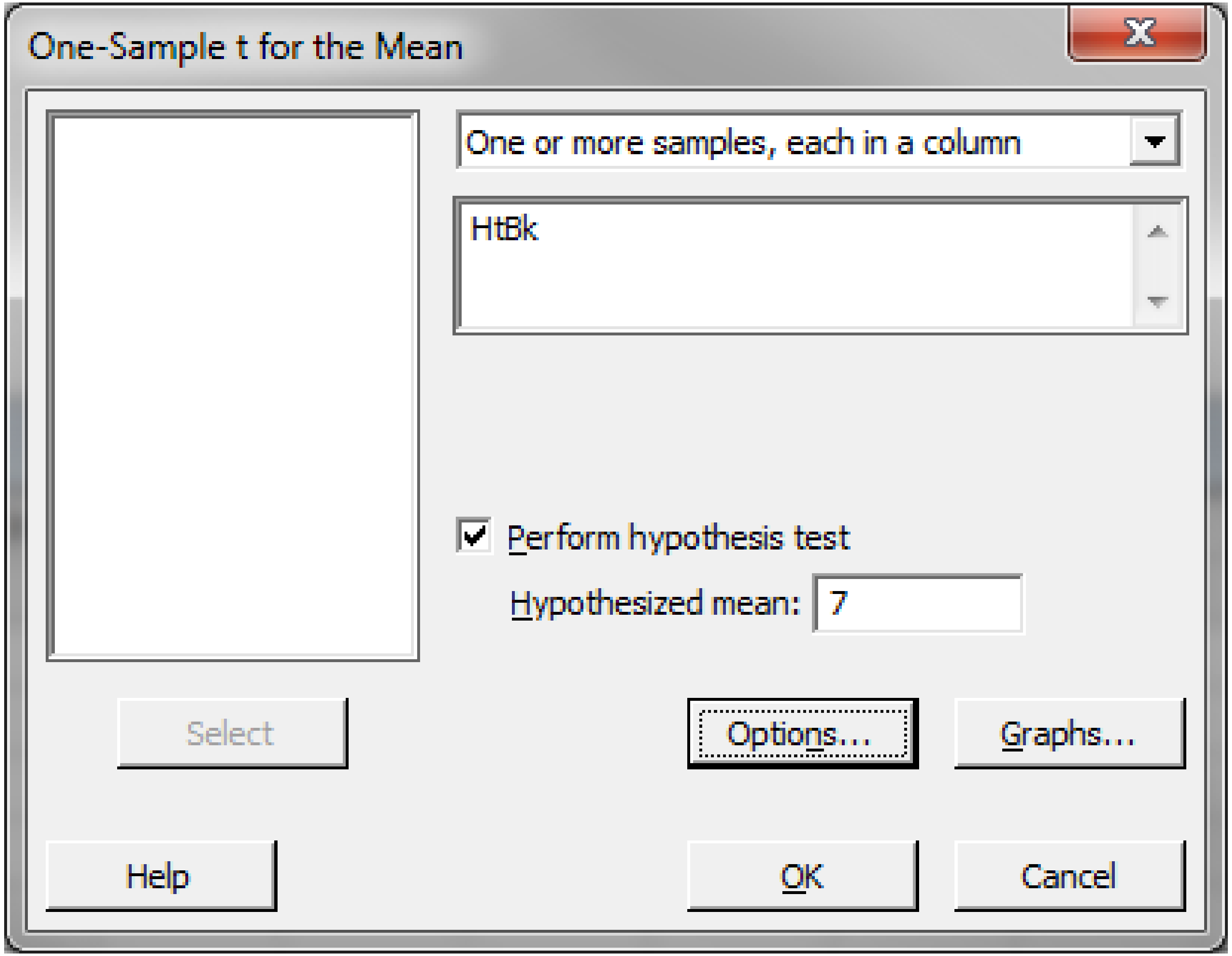
- Null Hypothesis (H_0): The data are normally distributed.
- Alternative Hypothesis (H_a): The data are not normally distributed.
- Since the p-value of the normality is 0.275, which is greater than alpha level (0.05), we fail to reject the null and claim that the data are normally distributed.
- If the data are not normally distributed, you need to use hypothesis tests other than the one sample t-test.



Use Minitab to Run a One-Sample T-Test

- Step 2: Run the one-sample t-test
  1) Click Stat → Basic Statistics → 1 Sample t.
  2) A new window named “1 Sample t (Test and Confidence Interval)” pops up.
  3) Click in the blank box under “Samples in columns” and “HtBk” appears in the list box on the left.
  4) Select “HtBk” as the “Samples in columns.”
  5) Check the box of “Perform hypothesis test.”
  6) Enter the hypothesized value “7” into the box next to “Perform hypothesis test.”
  7) Click “OK.”
  8) The one-sample t-test result appears automatically in the session window.

Use Minitab to Run a One-Sample T-Test



One or more samples, each in a column

HtBk

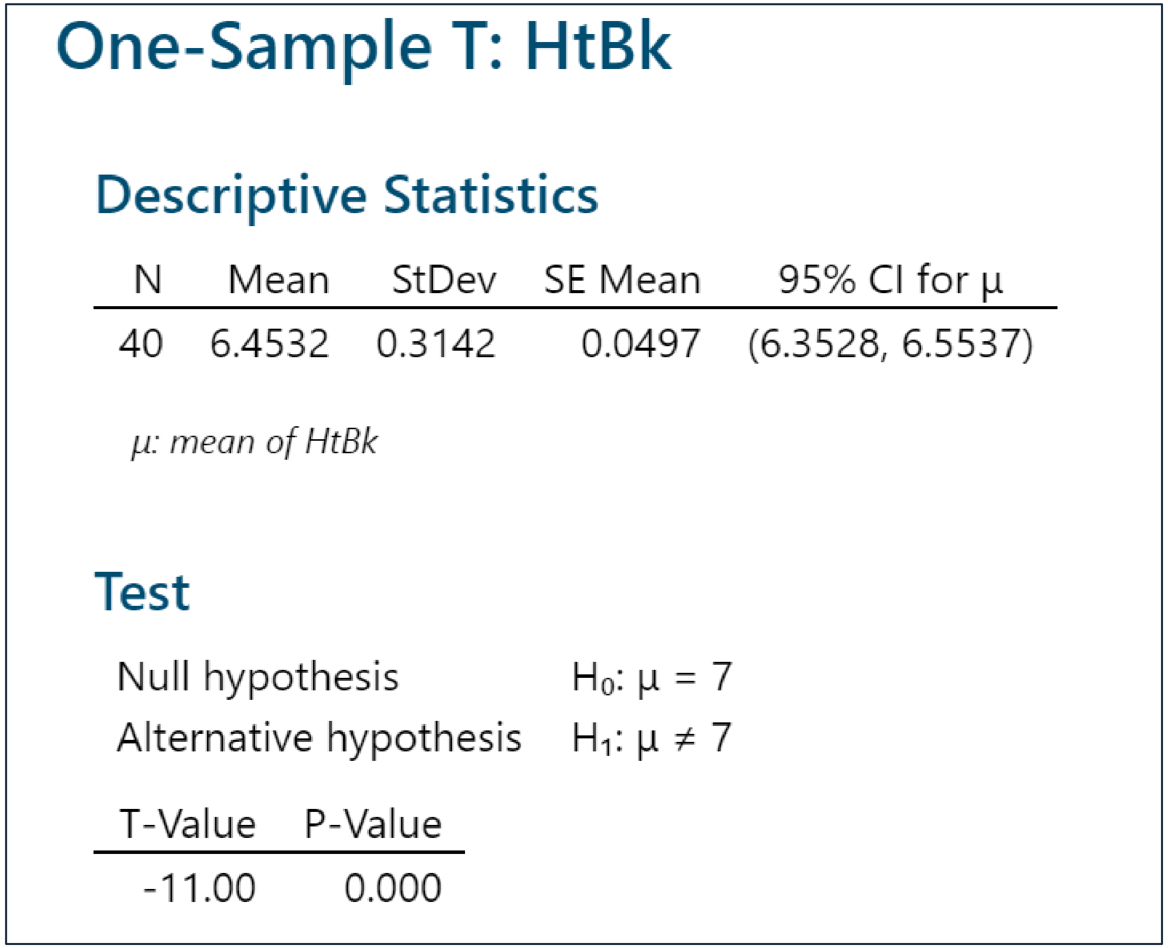
☑ Perform hypothesis test

Hypothesized mean: 7

Use Minitab to Run a One-Sample T-Test

- Null Hypothesis (H_0): $\mu = 7$
- Alternative Hypothesis (H_a): $\mu \neq 7$

- Since the p-value is smaller than alpha level (0.05), we reject the null hypothesis and claim that average of basketball players is statistically different from 7 feet.



```
One-Sample T: HtBk

Descriptive Statistics

 N     Mean    StDev   SE Mean    95% CI for μ
40   6.4532   0.3142    0.0497   (6.3528, 6.5537)

  μ: mean of HtBk

Test

 Null hypothesis          H₀: μ = 7
 Alternative hypothesis   H₁: μ ≠ 7

 T-Value   P-Value
  -11.00     0.000
```

What is Two Sample T-Test?

- **Two sample t-test** is a hypothesis test to study whether there is a statistically significant difference between the means of two populations.
  - Null Hypothesis (H_0): $\mu_1 = \mu_2$
  - Alternative Hypothesis (H_a): $\mu_1 \neq \mu_2$

where μ_1 is the mean of one population and μ_2 is the mean of the other population of our interest.

Assumptions of Two Sample T-Tests

- The sample data drawn from both populations are unbiased and representative.
- The data of both populations are continuous.
- The data of both populations are normally distributed.
- The variances of both populations are unknown.
- Two sample t-test is more robust than a z-test when the sample size is small (< 30).

Three Types of Two Sample T-Tests

1. Two sample t-test when the variances of two populations are unknown but equal
   - Two sample t-test (when $\sigma^2_1 = \sigma^2_2$)
2. Two sample t-test when the variances of the two population are unknown and unequal
   - Two sample t-test (when $\sigma^2_1 \neq \sigma^2_2$)
3. Paired t-test when the two populations are dependent of each other

Test of Equal Variance

- To check whether the variances of two populations of interest are statistically significant different, we use the test of equal variance.
  - Null Hypothesis (H_0): $\sigma_1^2 = \sigma_2^2$
  - Alternative Hypothesis (H_1): $\sigma_1^2 \neq \sigma_2^2$
- An F-test is used to test the equality of variances between two normally distributed populations.

Test of Equal Variance

- An **F-test** is a statistic hypothesis test in which the test statistic follows an F-distribution when the null hypothesis is true.

- The most known F-test is the test of equal variance for two normally distributed populations.

- The F-test is very sensitive to non-normality. When any one of the two populations is not normal, we use the Brown-Forsythe test for checking the equality of variances.

Test of Equal Variance

- Test Statistic

$$F_{calc} = \frac{s_1^{\ 2}}{s_2^{\ 2}}$$

where

s_1 and s_2 are the sample standard deviations.

- Critical Value
  - F_{crit} is the F value in a F distribution with the predetermined significance level α and degrees of freedom ($n_1 - 1$) and ($n_2 - 1$).
  - F_{crit} values for a two-sided and a one-sided F-test with the same significance level α and degrees of freedom ($n_1 - 1$) and ($n_2 - 1$) are different.

Test of Equal Variance

- Based on the sample data, we calculated the test statistic F_{calc}, which is compared against F_{crit} to make a decision of whether to reject the null.
  - Null Hypothesis (H_0): $\sigma_1^2 = \sigma_2^2$
  - Alternative Hypothesis (H_a): $\sigma_1^2 \neq \sigma_2^2$

- If $F_{calc} > F_{crit}$, we reject the null and claim there is a statistically significant difference between the variances of the two populations.

- If $F_{calc} < F_{crit}$, we fail to reject the null and claim there is not any statistically significant difference between the variances of the two populations.

Test Statistic & Critical Value of a Two Sample T-Test *when* $\sigma^2_1 = \sigma^2_2$

- Test Statistic

$$t_{calc} = \frac{\overline{Y}_1 - \overline{Y}_2}{s_{pooled}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1{}^2 + (n_2 - 1)s_2{}^2}{(n_1 + n_2 - 2)}}$$

where

$\overline{Y}_1$ and $\overline{Y}_2$ are the sample means of the two populations of our interest.

n_1 and n_2 are the sample sizes. n_1 is not necessarily equal to n_2.

s_{pooled} is a pooled estimate of variance. s_1 and s_2 are the sample standard deviations.

- Critical Value

t_{crit} is the *t* value in a Student t distribution with the predetermined significance level *α* and degrees of freedom ($n_1 + n_2 - 2$).

t_{crit} values for a two-sided and a one-sided t-test with the same significance level *α* and degrees of freedom ($n_1 + n_2 - 2$) are different.

Test Statistic & Critical Value of a Two Sample T-Test *when* $\sigma^2_1 \neq \sigma^2_2$

- Test Statistic

$$t_{calc} = \frac{\overline{Y}_1 - \overline{Y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$

where

$\overline{Y}_1$ and $\overline{Y}_2$ are the sample means of the two populations of our interest.

n_1 and n_2 are the sample sizes. n_1 is not necessarily equal to n_2.

s_1 and s_2 are the sample standard deviations.

- Critical Value

t_{crit} is the t value in a Student t distribution with the predetermined significance level *α* and degrees of freedom *df* calculated using the formula above.

t_{crit} values for a two-sided and a one-sided t-test with the same significance level *α* and degrees of freedom *df are different*.

Test Statistic & Critical Value of a Paired T-Test

- Test Statistic

$$t_{calc} = \frac{\bar{d}}{s_d/\sqrt{n}}$$

where

d is the difference between each pair of data.

$\bar{d}$ is the average of d.

n is the sample size of either population of interest.

s_d is standard deviation of *d*.

- Critical Value

t_{crit} is the t value in a Student t distribution with the predetermined significance level *α* and degrees of freedom (*n* – 1).

t_{crit} values for a two-sided and a one-sided t-test with the same significance level α and degrees of freedom (*n* – 1) are different.

Decision Rules of a Two Sample T-Test

- Based on the sample data, we calculated the test statistic t_{calc}, which is compared against t_{crit} to make a decision of whether to reject the null.
  - Null Hypothesis (H_0): $\mu_1 = \mu_2$
  - Alternative Hypothesis (H_a): $\mu_1 \neq \mu_2$

- If $|t_{calc}| > t_{crit}$, we reject the null and claim there is a statistically significant difference between the means of the two populations.

- If $|t_{calc}| < t_{crit}$, we fail to reject the null and claim there is not any statistically significant difference between the means of the two populations.

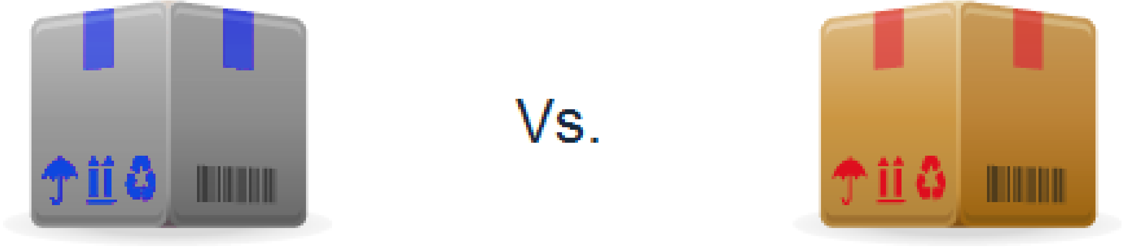
Use Minitab to Run a Two-Sample T-Test

- *Case study:* We are trying to compare the average retail price of a product in state A and state B.
  - Data File: “Two-Sample T-Test” tab in “Sample Data.xlsx”

Avg. Product Price State “A”

Vs.

Avg. Product Price State “B”

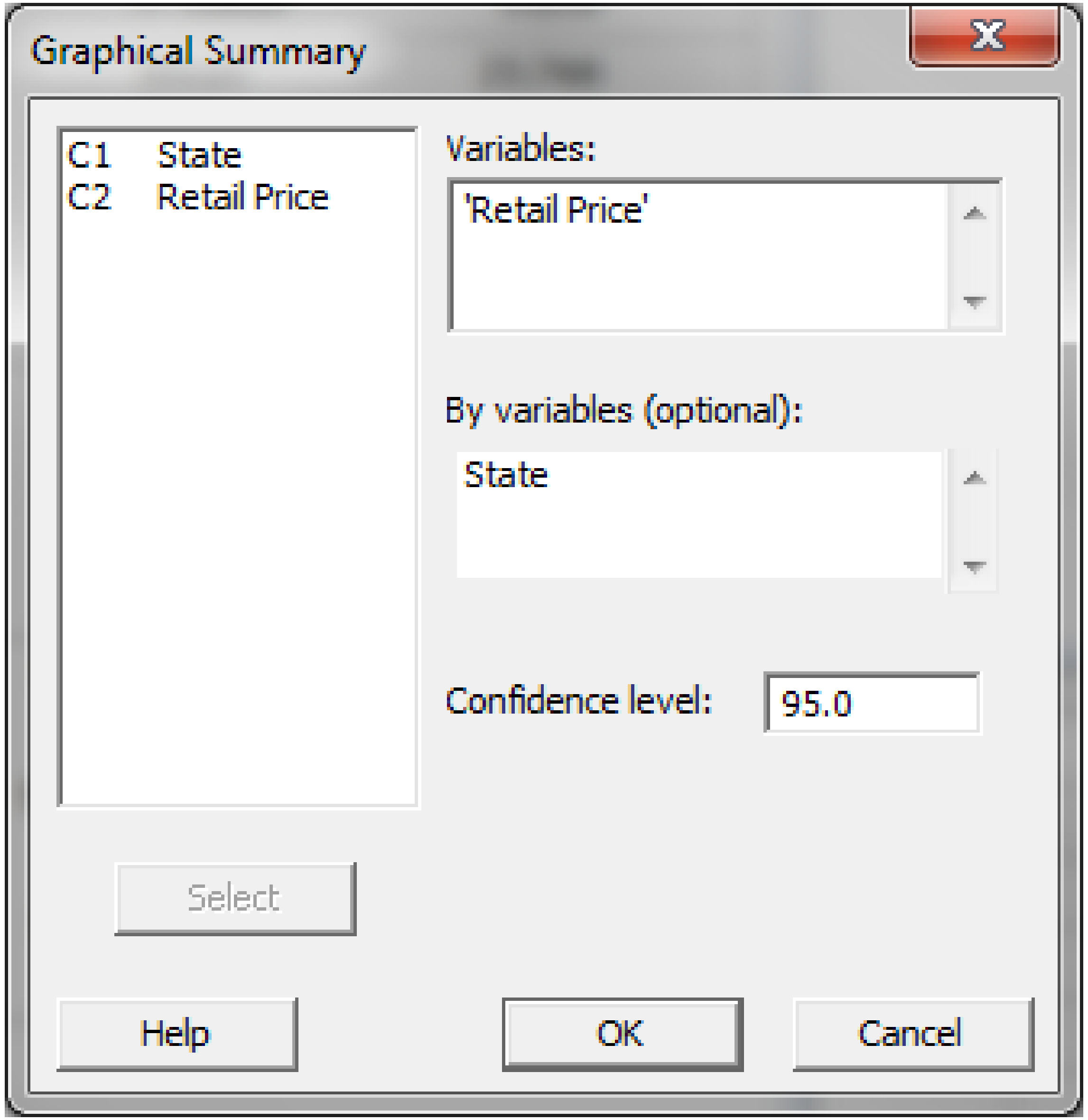


- Null Hypothesis (H_0): $\mu_1 = \mu_2$
- Alternative Hypothesis (H_a): $\mu_1 \neq \mu_2$

Use Minitab to Run a Two-Sample T-Test

- Step 1: Test the normality of the retail price for both state A and B.
    1) Click Stat → Basic Statistics → Graphical Summary.
    2) A window named “Graphical Summary” pops up.
    3) Select “Retail Price” as the “Variables.”
    4) Click in the blank box right below “By variables (optional)” and “State” appears in the list box on the left.
    5) Select the “State” as the “By variables (optional).”
    6) Click “OK.”
    7) The normality test results would appear in the new windows.

Use Minitab to Run a Two-Sample T-Test



Graphical Summary

C1 State
C2 Retail Price

Variables: 'Retail Price'

By variables (optional): State

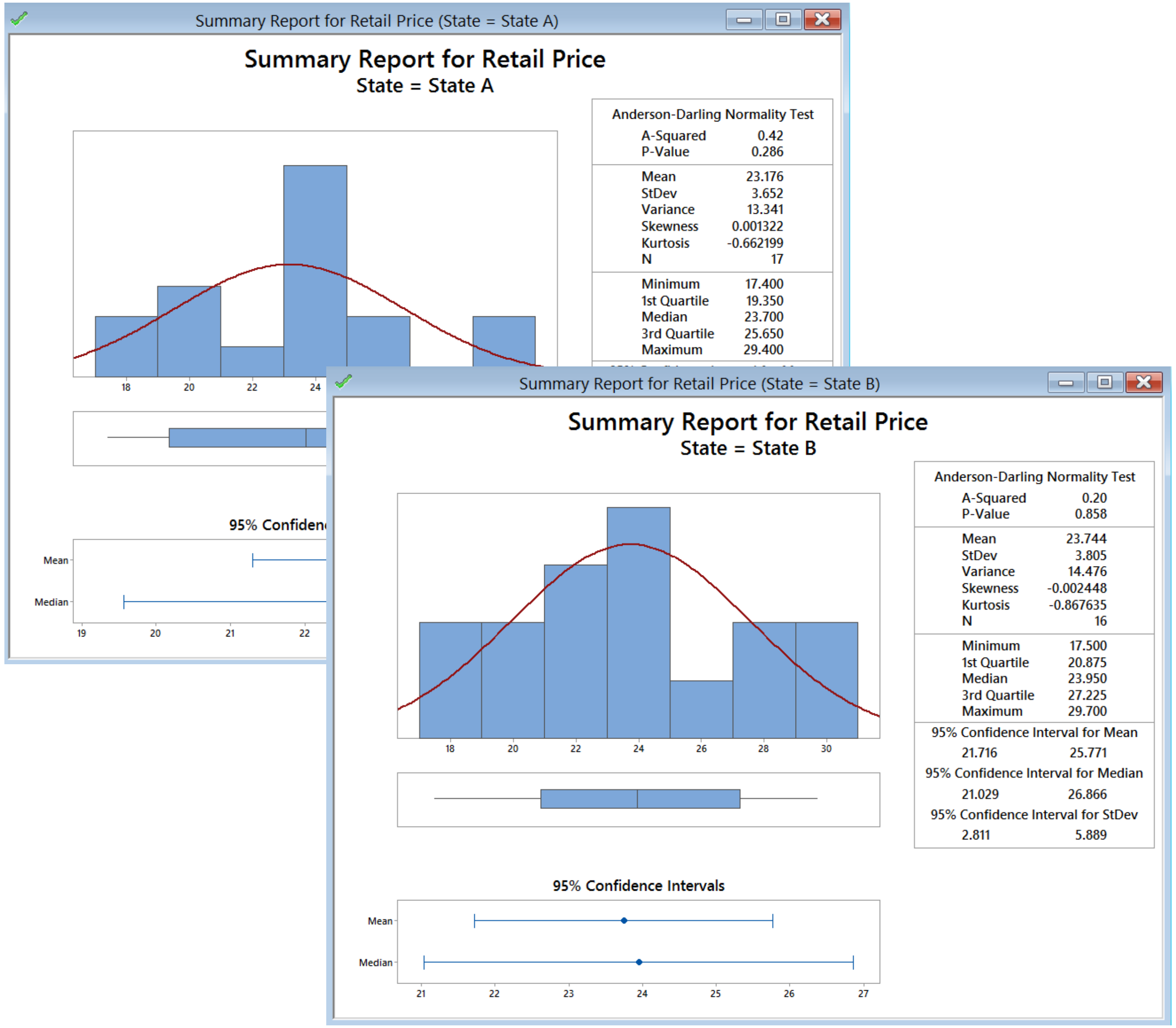
Confidence level: 95.0

Use Minitab to Run a Two-Sample T-Test

- Null Hypothesis (H0): The data are normally distributed.
- Alternative Hypothesis (Ha): The data are not normally distributed.

- Both retail price data of state A and B are normally distributed since the p-values are both greater than alpha level (0.05).

- If any of the data series is not normally distributed, we need to use other hypothesis testing methods other than the two sample t-test.



Summary Report for Retail Price
State = State A

```
Anderson-Darling Normality Test
   A-Squared          0.42
   P-Value           0.286

   Mean             23.176
   StDev             3.652
   Variance         13.341
   Skewness       0.001322
   Kurtosis      -0.662199
   N                    17

   Minimum          17.400
   1st Quartile     19.350
   Median           23.700
   3rd Quartile     25.650
   Maximum          29.400
```

Summary Report for Retail Price
State = State B

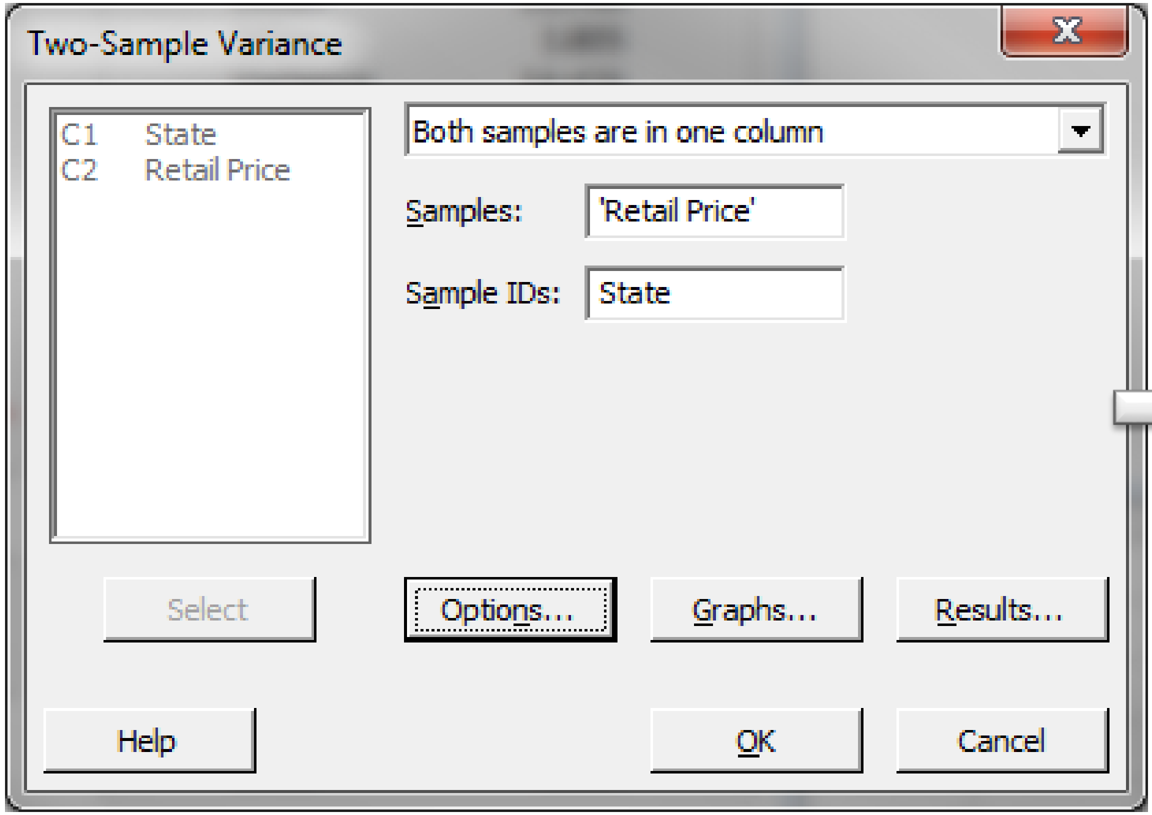
```
Anderson-Darling Normality Test
   A-Squared          0.20
   P-Value           0.858

   Mean             23.744
   StDev             3.805
   Variance         14.476
   Skewness      -0.002448
   Kurtosis      -0.867635
   N                    16

   Minimum          17.500
   1st Quartile     20.875
   Median           23.950
   3rd Quartile     27.225
   Maximum          29.700
95% Confidence Interval for Mean
   21.716           25.771
95% Confidence Interval for Median
   21.029           26.866
95% Confidence Interval for StDev
   2.811             5.889
```

Use Minitab to Run a Two-Sample T-Test

- Step 2: Test whether the variances of the two data sets are equal.
  - Null Hypothesis (H_0): $\sigma_1^2 = \sigma_2^2$
  - Alternative Hypothesis (H_a): $\sigma_1^2 \neq \sigma_2^2$

1) Click Basic Statistics → 2 Variances.
2) A new window named “2 Sample Variance” pops up.
3) Click in the blank box of “Samples” and “Retail Price” appears in the list box on the left.
4) Select “Retail Price” as the “Samples.”
5) Click in the blank box of “Sample ID” and “State” appears in the list box on the left.
6) Select “State” as the “Sample ID.”
7) Click “Options” and select “Use test and confidence intervals based on normal distribution”
8) Click “OK.” to the “Options” dialogue box
9) Click “OK.” to the 2 Variances (Test Confidence Intervals) window.
10) The results show up in the session window.

Use Minitab to Run a Two-Sample T-Test

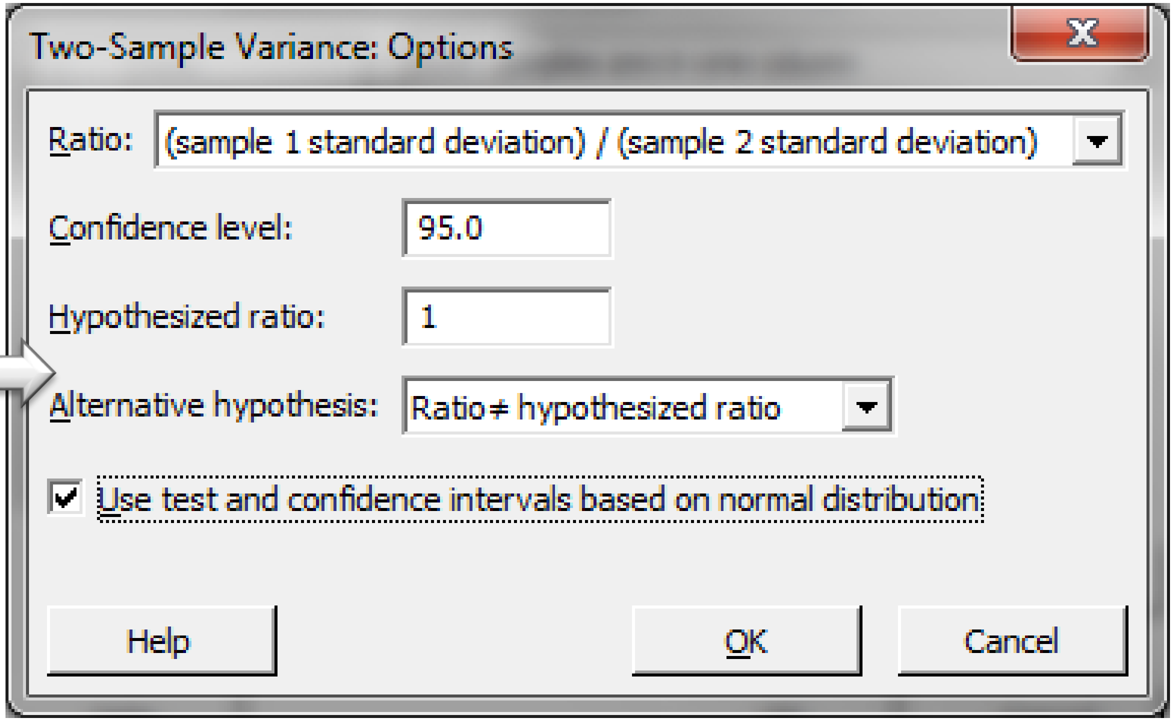

Two-Sample Variance

C1 State
C2 Retail Price

Both samples are in one column

Samples: 'Retail Price'

Sample IDs: State


Two-Sample Variance: Options

Ratio: (sample 1 standard deviation) / (sample 2 standard deviation)

Confidence level: 95.0

Hypothesized ratio: 1

Alternative hypothesis: Ratio ≠ hypothesized ratio

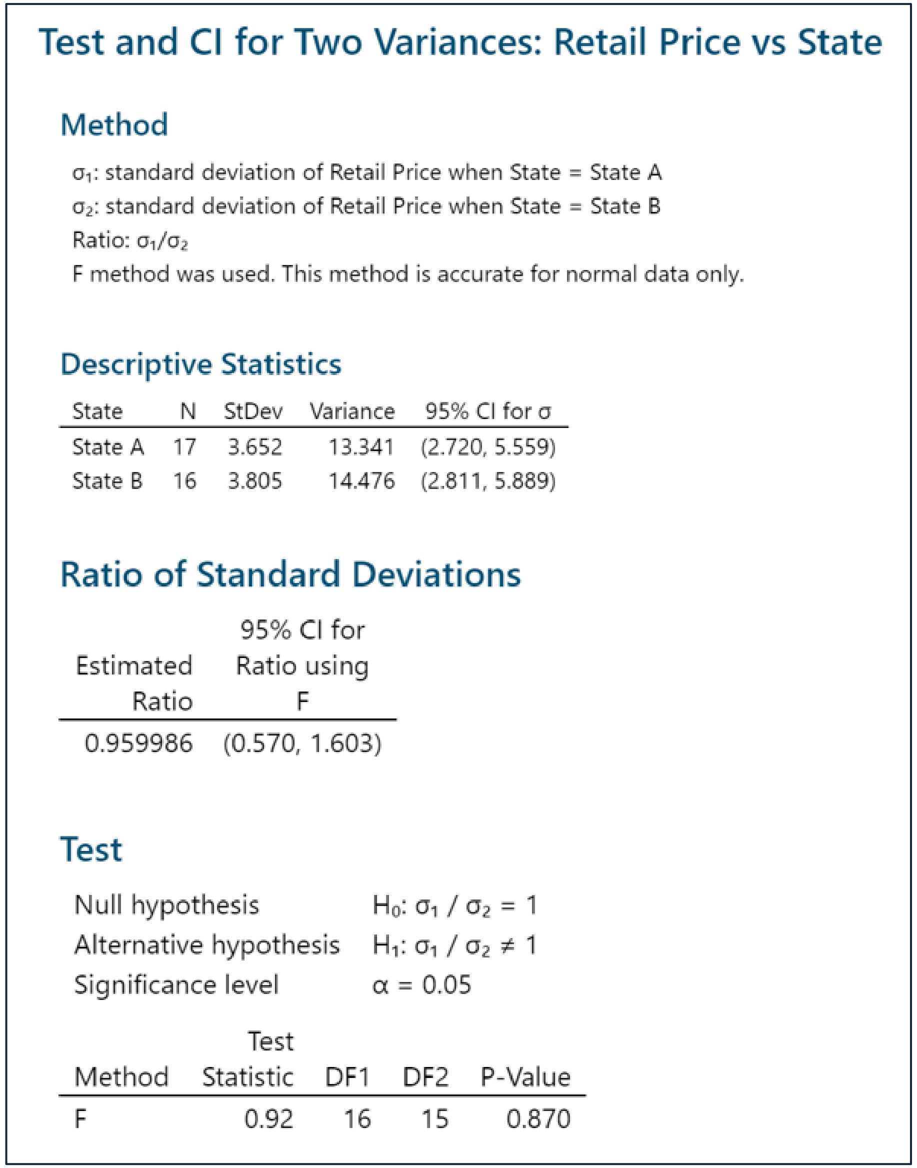
☑ Use test and confidence intervals based on normal distribution

Use Minitab to Run a Two-Sample T-Test

- Because the retail prices at state A and state B are both normally distributed, an F test is used to test their variance equality.

- The p-value of F test is 0.870, greater than the alpha level (0.05), so we fail to reject the null hypothesis and we claim that the variances of the two data sets are equal. We will use the two sample t-test (when $\sigma^2_1 = \sigma^2_2$) to compare the means of the two groups.

- If $\sigma^2_1 \neq \sigma^2_2$, we will use the two sample t-test (when $\sigma^2_1 \neq \sigma^2_2$) to compare the means of the two groups.



Test and CI for Two Variances: Retail Price vs State

Method

σ_1: standard deviation of Retail Price when State = State A
σ_2: standard deviation of Retail Price when State = State B
Ratio: σ_1/σ_2
F method was used. This method is accurate for normal data only.

Descriptive Statistics

State	N	StDev	Variance	95% CI for σ
State A	17	3.652	13.341	(2.720, 5.559)
State B	16	3.805	14.476	(2.811, 5.889)

Ratio of Standard Deviations

Estimated Ratio	95% CI for Ratio using F
0.959986	(0.570, 1.603)

Test

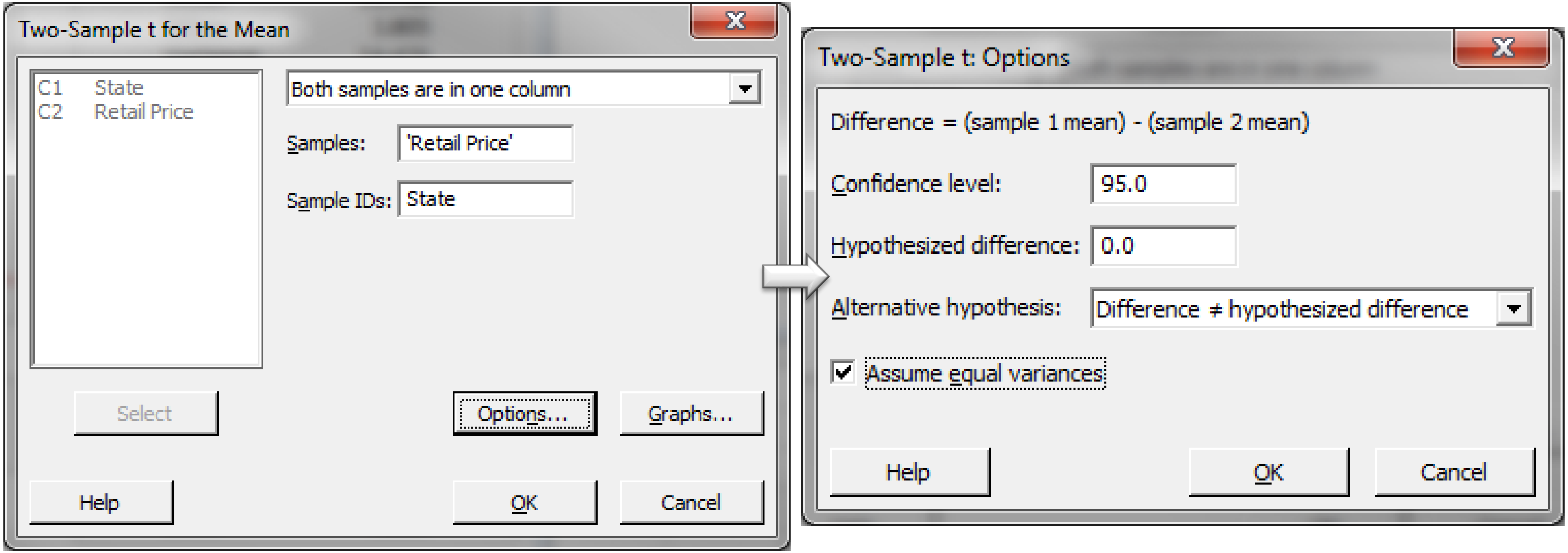
Null hypothesis H_0: $\sigma_1 / \sigma_2 = 1$
Alternative hypothesis H_1: $\sigma_1 / \sigma_2 \neq 1$
Significance level $\alpha = 0.05$

Method	Test Statistic	DF1	DF2	P-Value
F	0.92	16	15	0.870

Use Minitab to Run a Two-Sample T-Test

- Step 3: Run two-sample t-test to compare the means of two groups.
    1) Click Stat → Basic Statistics → 2-Sample t.
    2) A new window named “2-Sample t for the Mean” pops up.
    3) Click in the blank box right next to “Samples” and the “Retail Price” appears in the list box on the left.
    4) Select “Retail Price” as the “Samples.”
    5) Click in the blank box right next to “Sample ID” and the “State” appears in the list box on the left.
    6) Select “State” as the “Sample ID.”
    7) Click “Options”
    8) Check the box “Assume equal variances.”
    9) Click “OK.” in the “Options” window.
    10) Click “OK” in the 2-Sample t window
    11) The results of the two-sample t-test (when $\sigma_1 = \sigma_2$) appear automatically in the session window.

Use Minitab to Run a Two-Sample T-Test



C1 State
C2 Retail Price

Both samples are in one column

Samples: 'Retail Price'

Sample IDs: State

Difference = (sample 1 mean) - (sample 2 mean)

Confidence level: 95.0

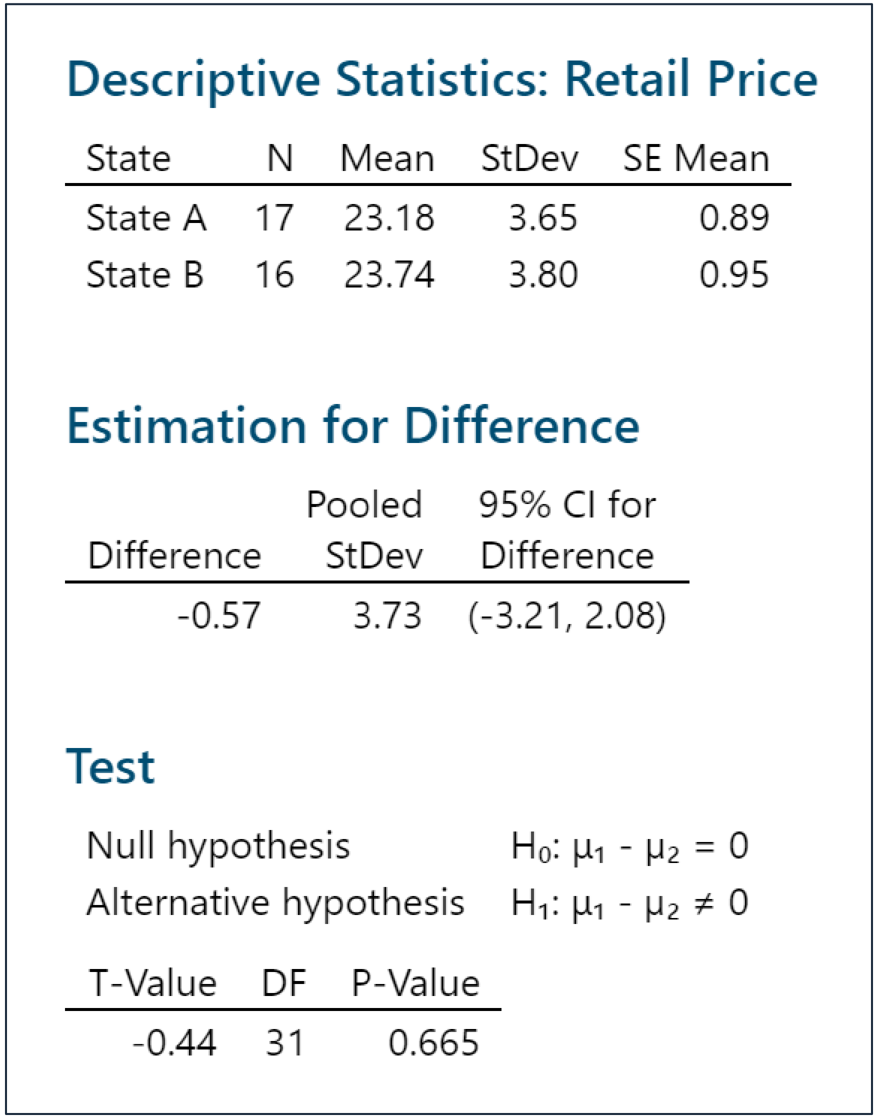
Hypothesized difference: 0.0

Alternative hypothesis: Difference ≠ hypothesized difference

☑ Assume equal variances

Use Minitab to Run a Two-Sample T-Test

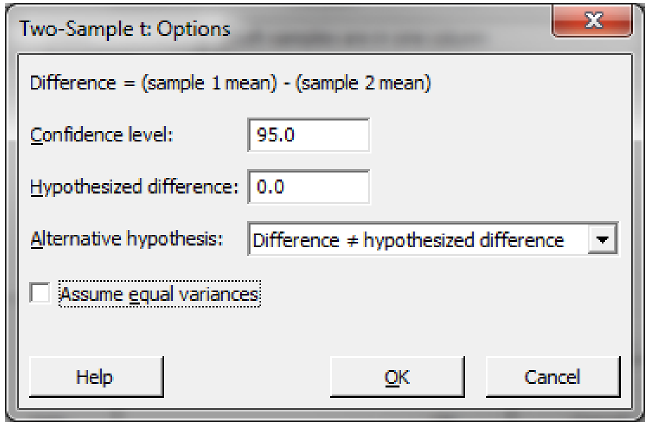
Since the p-value of the t-test (assuming equal variance) is 0.665, greater than the alpha level (0.05), we fail to reject the null hypothesis and we claim that the means of the two data sets are equal.



```
Descriptive Statistics: Retail Price

State     N   Mean  StDev  SE Mean
State A  17  23.18   3.65     0.89
State B  16  23.74   3.80     0.95

Estimation for Difference

             Pooled  95% CI for
Difference    StDev  Difference
     -0.57     3.73  (-3.21, 2.08)

Test

Null hypothesis         H₀: μ₁ - μ₂ = 0
Alternative hypothesis  H₁: μ₁ - μ₂ ≠ 0

T-Value  DF  P-Value
  -0.44  31    0.665
```

Use Minitab to Run a Two-Sample T-Test

- If the variances of the two groups do not equal, we will need to use the two-sample t-test (when $\sigma_1 \neq \sigma_2$) to compare the means of the two groups.
  - In the window of “2-Sample (Test and Confidence Interval),” uncheck the box next to “Assume equal variances” and run the 2-sample t-test again.


Two-Sample t: Options

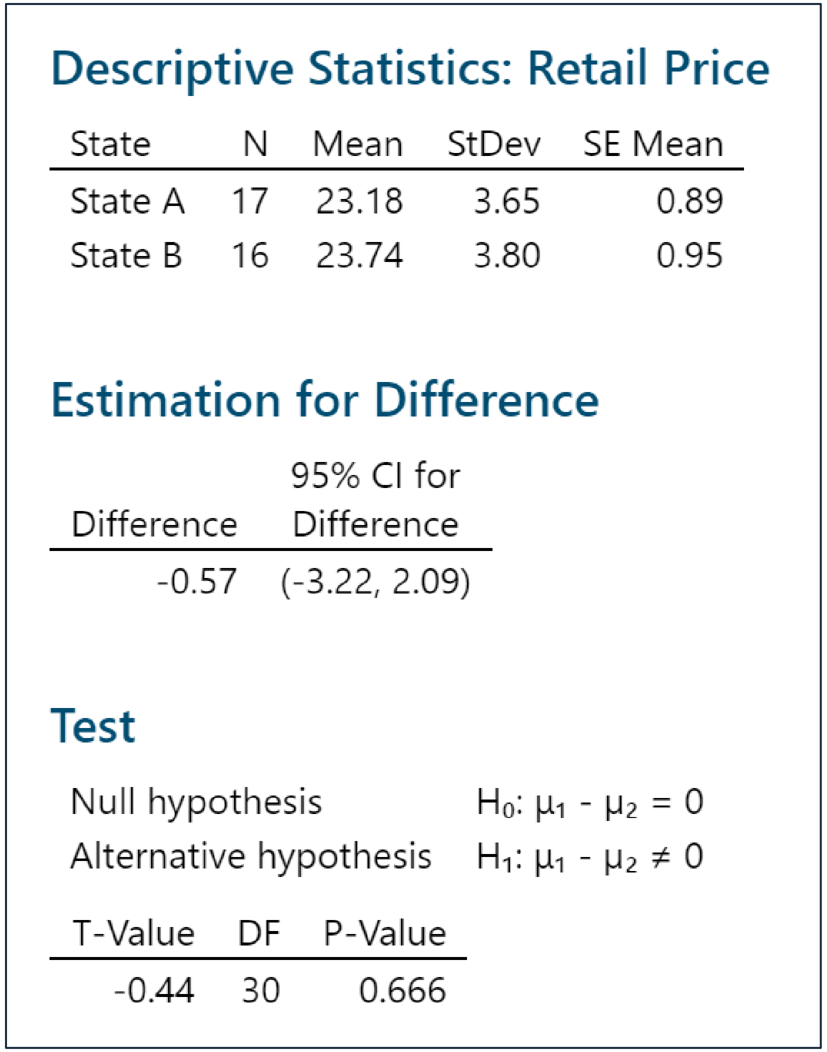
Difference = (sample 1 mean) - (sample 2 mean)

Confidence level: 95.0

Hypothesized difference: 0.0

Alternative hypothesis: Difference ≠ hypothesized difference

☐ Assume equal variances


```
Descriptive Statistics: Retail Price

State      N    Mean   StDev   SE Mean
State A   17   23.18    3.65      0.89
State B   16   23.74    3.80      0.95

Estimation for Difference

              95% CI for
Difference    Difference
     -0.57    (-3.22, 2.09)

Test

Null hypothesis          H₀: μ₁ - μ₂ = 0
Alternative hypothesis   H₁: μ₁ - μ₂ ≠ 0

T-Value   DF   P-Value
  -0.44   30     0.666
```

- Since the p-value of the t-test (assuming unequal variance) is 0.666, greater than the alpha level (0.05), we fail to reject the null hypothesis and we claim that the means of two groups are equal.

Use Minitab to Run a Paired T-Test

Case study:

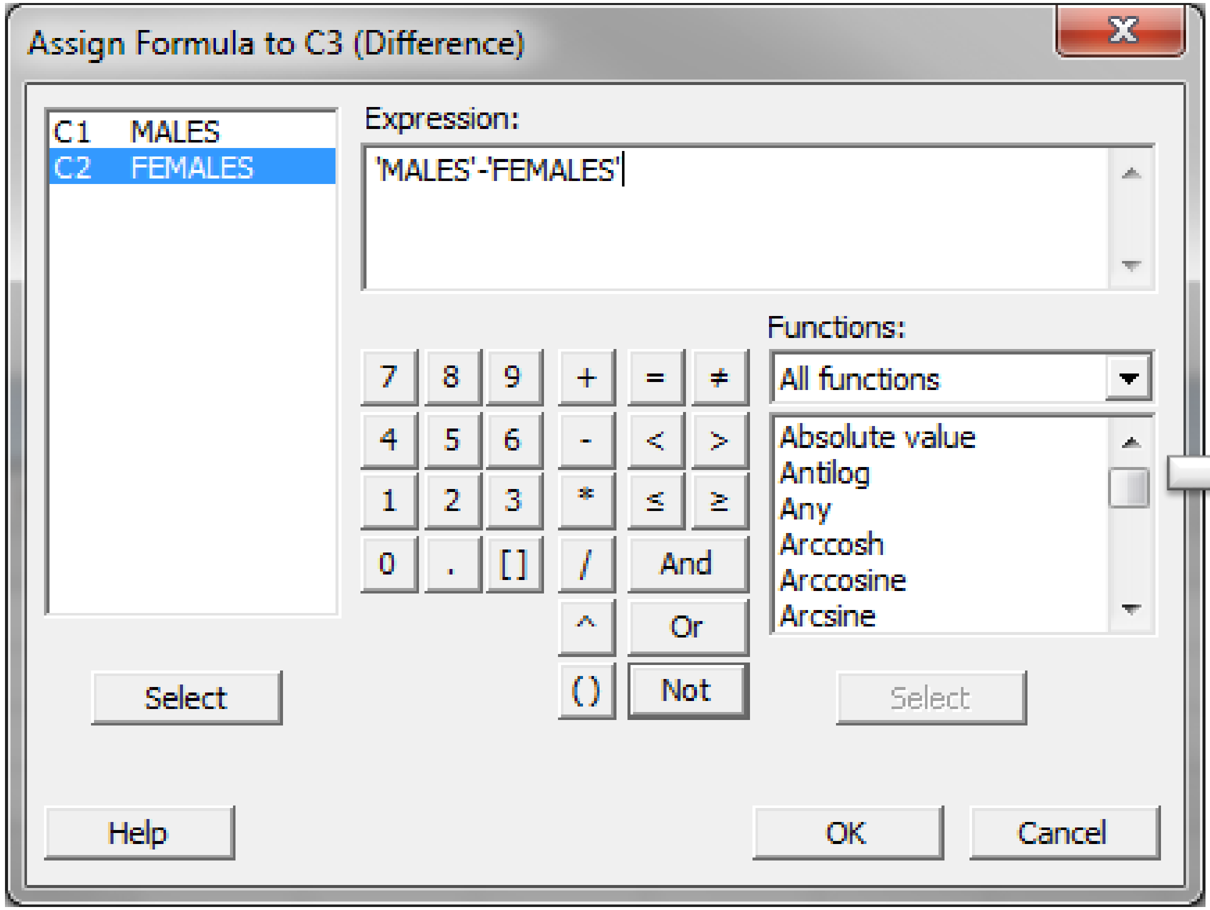
- We are interested to know whether the average salaries ($1000/yr.) of male and female professors at the same university are the same.
  - Data File: “Paired T-Test” tab in “Sample Data.xlsx”
  - The data were randomly collected from 22 universities. For each university, the salaries of a male and female professors were randomly selected.

- Null Hypothesis (H_0): $\mu_{male} - \mu_{female} = 0$
- Alternative Hypothesis (H_a): $\mu_{male} - \mu_{female} \neq 0$

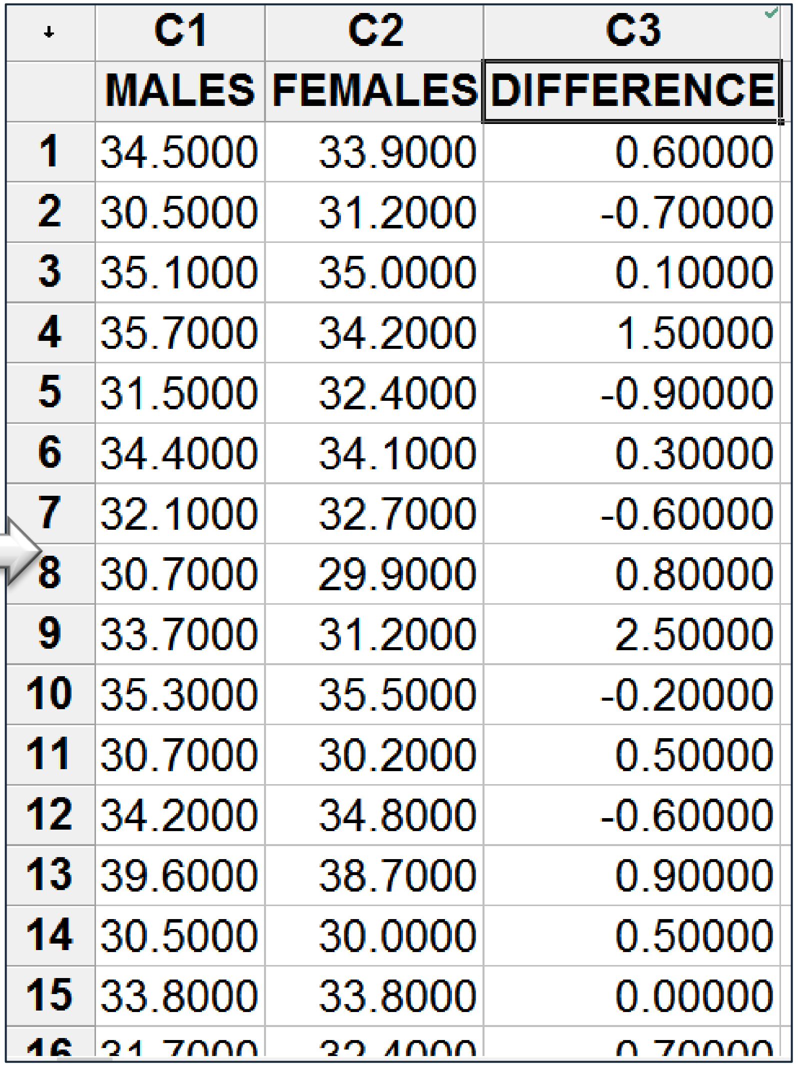
Use Minitab to Run a Paired T-Test

- Step 1: Create a new column of the difference between the two data sets of interest and name it “Difference.”
  (Note: if your data already has a column for “Difference” with data in it, this step will not be necessary.)
  1) Right click on the newly-generated column, “Difference.”
  2) Select “Formulas” and then “Assign Formula to Column.”
  3) A new window named “Assign Formula to C3 (Difference)” pops up.
  4) Enter “ ‘MALES’-‘ FEMALES’ ” into the box right below “Expression.”
  5) Click “OK.”
  6) The values would appear in the column “Difference.”

Use Minitab to Run a Paired T-Test



Expression: 'MALES'-'FEMALES'



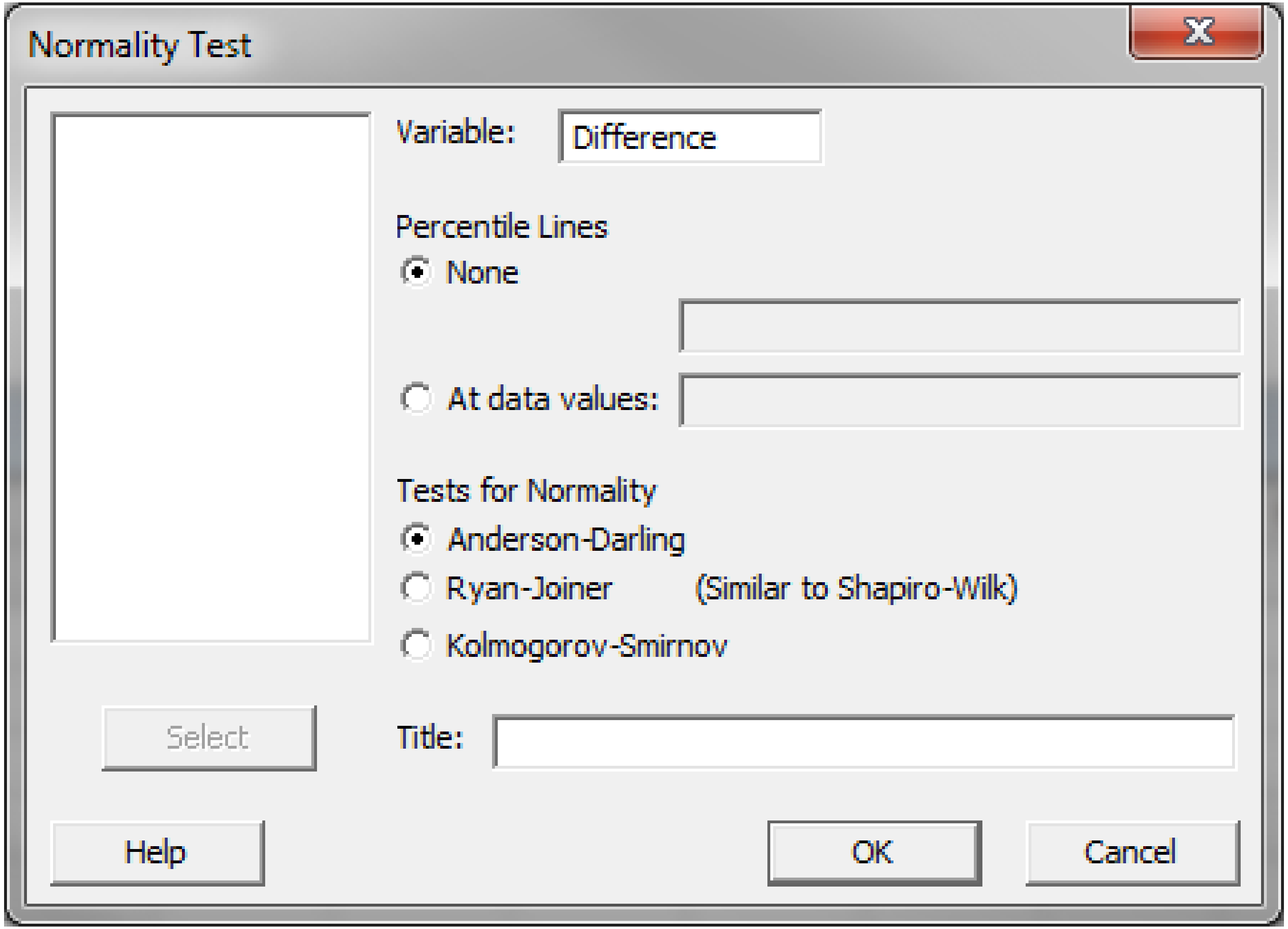
	C1	C2	C3
	MALES	FEMALES	DIFFERENCE
1	34.5000	33.9000	0.60000
2	30.5000	31.2000	-0.70000
3	35.1000	35.0000	0.10000
4	35.7000	34.2000	1.50000
5	31.5000	32.4000	-0.90000
6	34.4000	34.1000	0.30000
7	32.1000	32.7000	-0.60000
8	30.7000	29.9000	0.80000
9	33.7000	31.2000	2.50000
10	35.3000	35.5000	-0.20000
11	30.7000	30.2000	0.50000
12	34.2000	34.8000	-0.60000
13	39.6000	38.7000	0.90000
14	30.5000	30.0000	0.50000
15	33.8000	33.8000	0.00000

Use Minitab to Run a Paired T-Test

- Step 2: Test whether the difference is normally distributed.
  - Null Hypothesis (H_0): The difference between two data sets is normally distributed.
  - Alternative Hypothesis (H_a): The difference between two data sets is not normally distributed.

1) Click Stat → Basic Statistics → Normality Test.
2) A new window named “Normality Test” pops up.
3) Select “Difference” as the “Variable.”
4) Click “OK.”
5) The normality test results would appear in the new window.

Use Minitab to Run a Paired T-Test



Normality Test

Variable: Difference

Percentile Lines

● None

○ At data values:

Tests for Normality

● Anderson-Darling

○ Ryan-Joiner (Similar to Shapiro-Wilk)

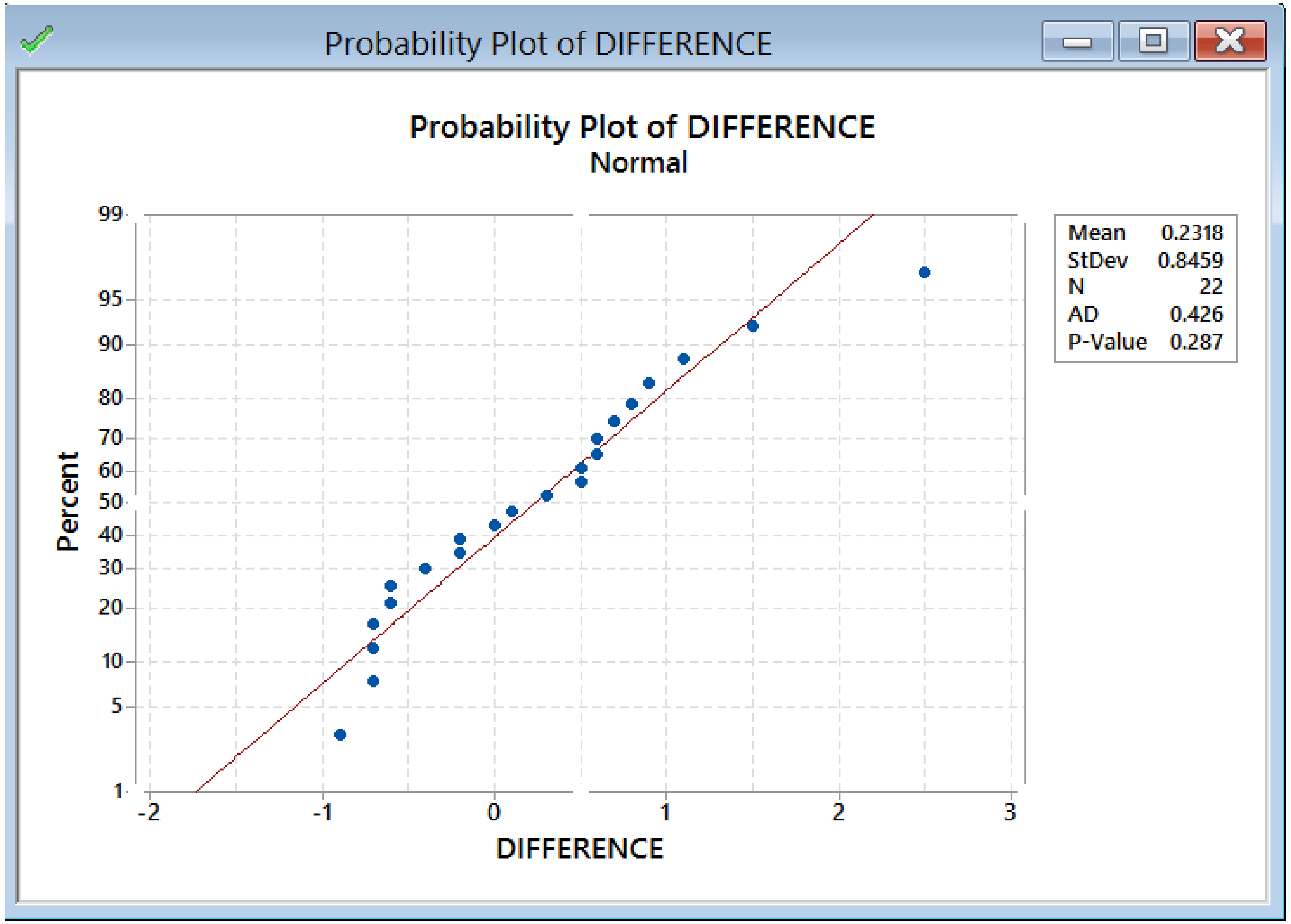
○ Kolmogorov-Smirnov

Title:

Use Minitab to Run a Paired T-Test

- The p-value of the normality test is 0.287, greater than the alpha level (0.05), so we fail to reject the null hypothesis and we claim that the difference is normally distributed.

- When the difference is not normally distributed, we need other hypothesis testing methods.

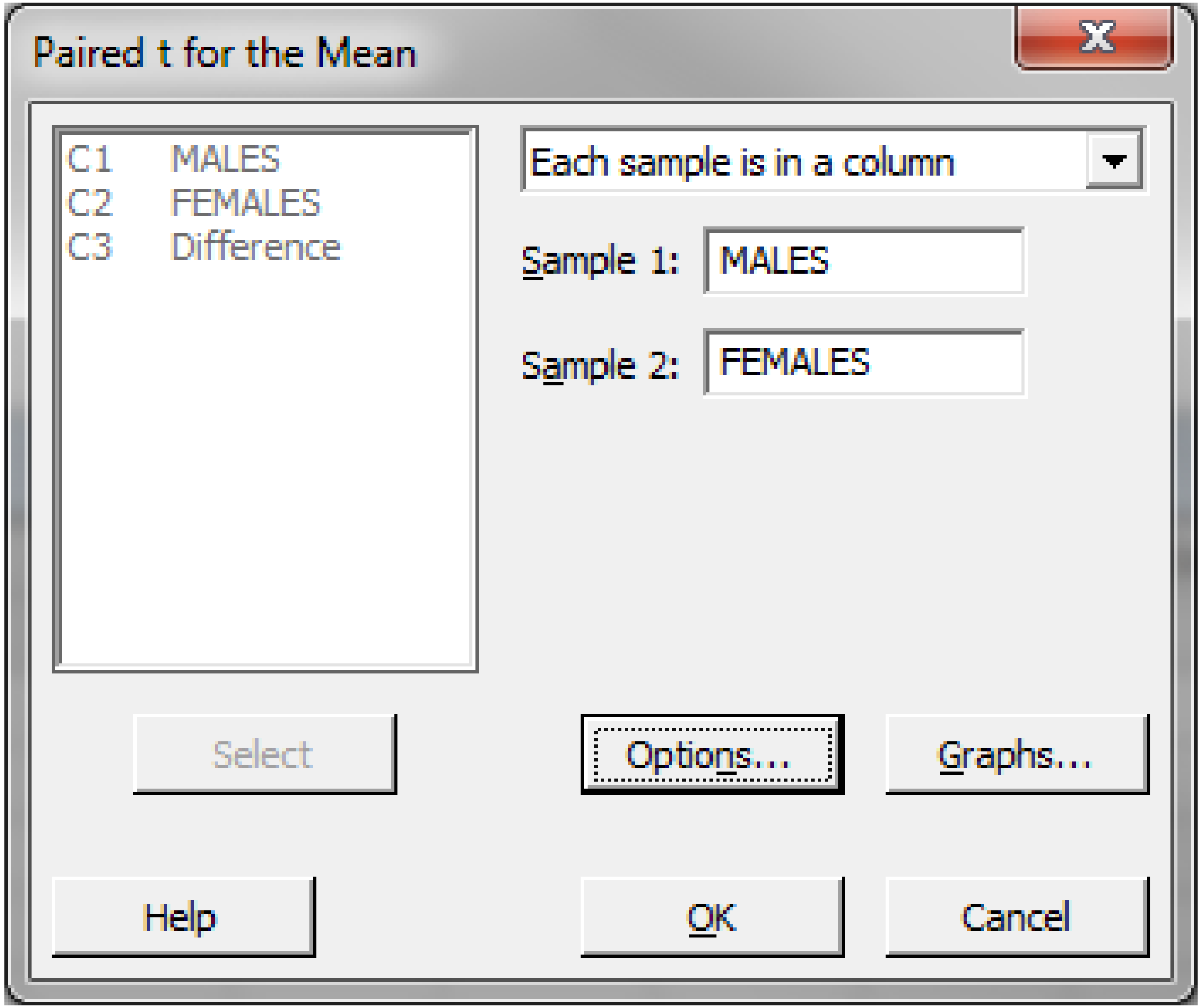


Mean	0.2318
StDev	0.8459
N	22
AD	0.426
P-Value	0.287

Use Minitab to Run a Paired T-Test

- Step 3: Run the paired t-test to compare the means of two dependent data sets.
  1) Click Stat → Basic Statistics → Paired t.
  2) A new window named “Paired t (Test and Confidence Interval)” pops up.
  3) Click in the blank box right next to “First Sample” and the three columns names appear in the list box on the left.
  4) Select “MALES” as “First Sample.”
  5) Select “FEMALES” as “Second Sample.”
  6) Click “OK.”
  7) The paired t-test results appears in the session window.

Use Minitab to Run a Paired T-Test



Paired t for the Mean

C1 MALES
C2 FEMALES
C3 Difference

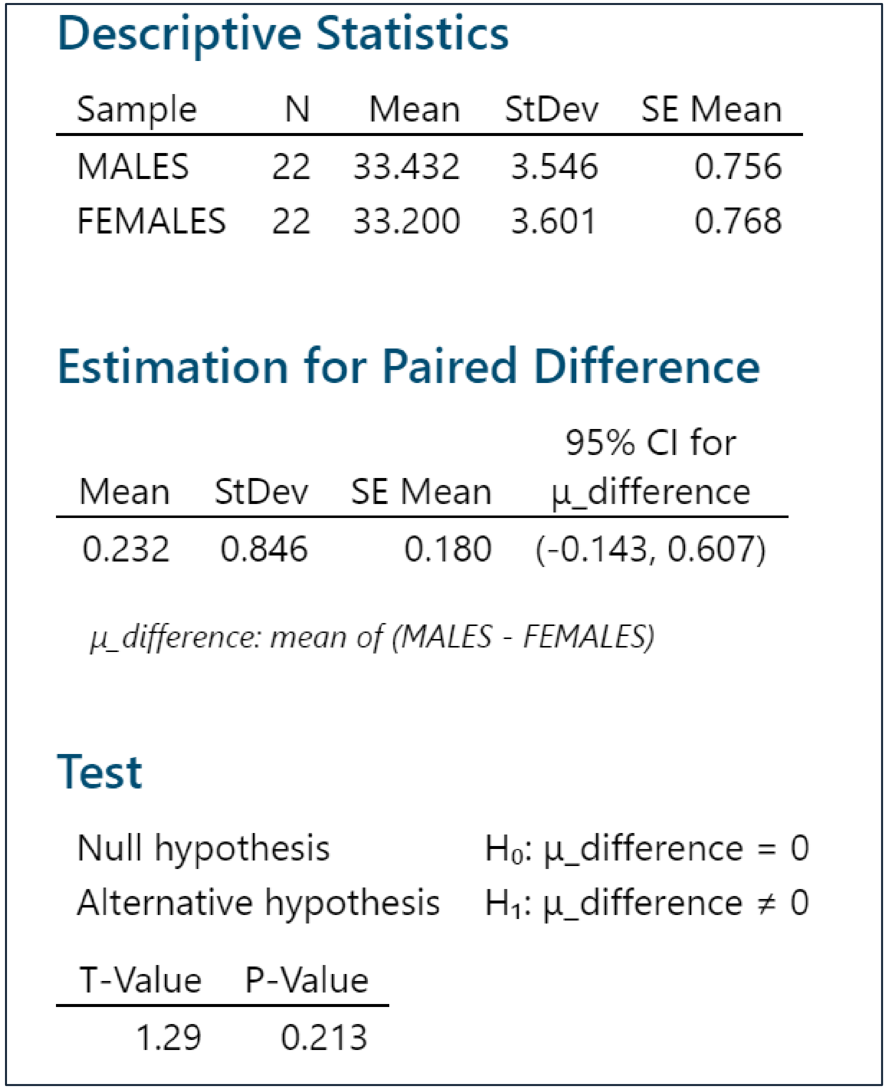
Each sample is in a column

Sample 1: MALES

Sample 2: FEMALES

Use Minitab to Run a Paired T-Test

- The p-value of the paired t-test is 0.213, greater than the alpha level (0.05), so we fail to reject the null hypothesis and we claim that there is no statistically significant difference between the salaries of male and those of female professors.



```
Descriptive Statistics

Sample     N    Mean   StDev  SE Mean
MALES     22  33.432   3.546    0.756
FEMALES   22  33.200   3.601    0.768


Estimation for Paired Difference

                            95% CI for
 Mean   StDev  SE Mean    μ_difference
0.232   0.846    0.180  (-0.143, 0.607)

μ_difference: mean of (MALES - FEMALES)


Test

Null hypothesis         H₀: μ_difference = 0
Alternative hypothesis  H₁: μ_difference ≠ 0

T-Value  P-Value
   1.29    0.213
```

3.4.2 One Sample Variance

What is One Sample Variance Test?

- **One sample variance test** is a hypothesis testing method to study whether there is a statistically significant difference between a population variance and a specified value.

  - Null Hypothesis (H_0): $\sigma^2 = \sigma_0{}^2$
  - Alternative Hypothesis (H_a): $\sigma^2 \neq \sigma_0{}^2$

where σ^2 is the variance of a population of our interest and $\sigma_0{}^2$ is the specific value we want to compare against.

Chi-Square Test

- A **chi-square test** is a statistical hypothesis test in which the test statistic follows a chi-square distribution when the null hypothesis is true.

- A chi-square test can be used to test the equality between the variance of a normally distributed population and a specified value.

Chi-Square Test

- Test Statistic

$$\chi_{calc}^{2} = \frac{(n-1)s^{2}}{\sigma_{0}^{2}}$$

where

s^2 is the observed variance and n is the sample size.

σ_0^{2} is the specified value we compare against.

- Critical Value
  - χ_{crit}^{2} is the χ^2 value in a χ^2 distribution with the predetermined significance level α and degrees of freedom (n – 1).
  - χ_{crit}^{2} values for a two-sided and a one-sided χ^2-test with the same significance level α and degrees of freedom (n –1) *are different.*

Chi-Square Test

- Based on the sample data, we calculated the test statistic $\chi_{crit}{}^{2}$, which is compared against $\chi_{calc}{}^{2}$ to make a decision of whether to reject the null.
  - Null Hypothesis (H_0): $\sigma^2 = \sigma_0{}^2$
  - Alternative Hypothesis (H_a): $\sigma^2 \neq \sigma_0{}^2$

- If $\left|\chi_{calc}{}^{2}\right| > \chi_{crit}{}^{2}$, we reject the null and claim there is a statistically significant difference between the population variance and the specified value.

- If $\left|\chi_{calc}{}^{2}\right| < \chi_{crit}{}^{2}$, we fail to reject the null and claim there is not any statistically significant difference between the population variance and specified value.

Use Minitab to Run a One Sample Variance Test

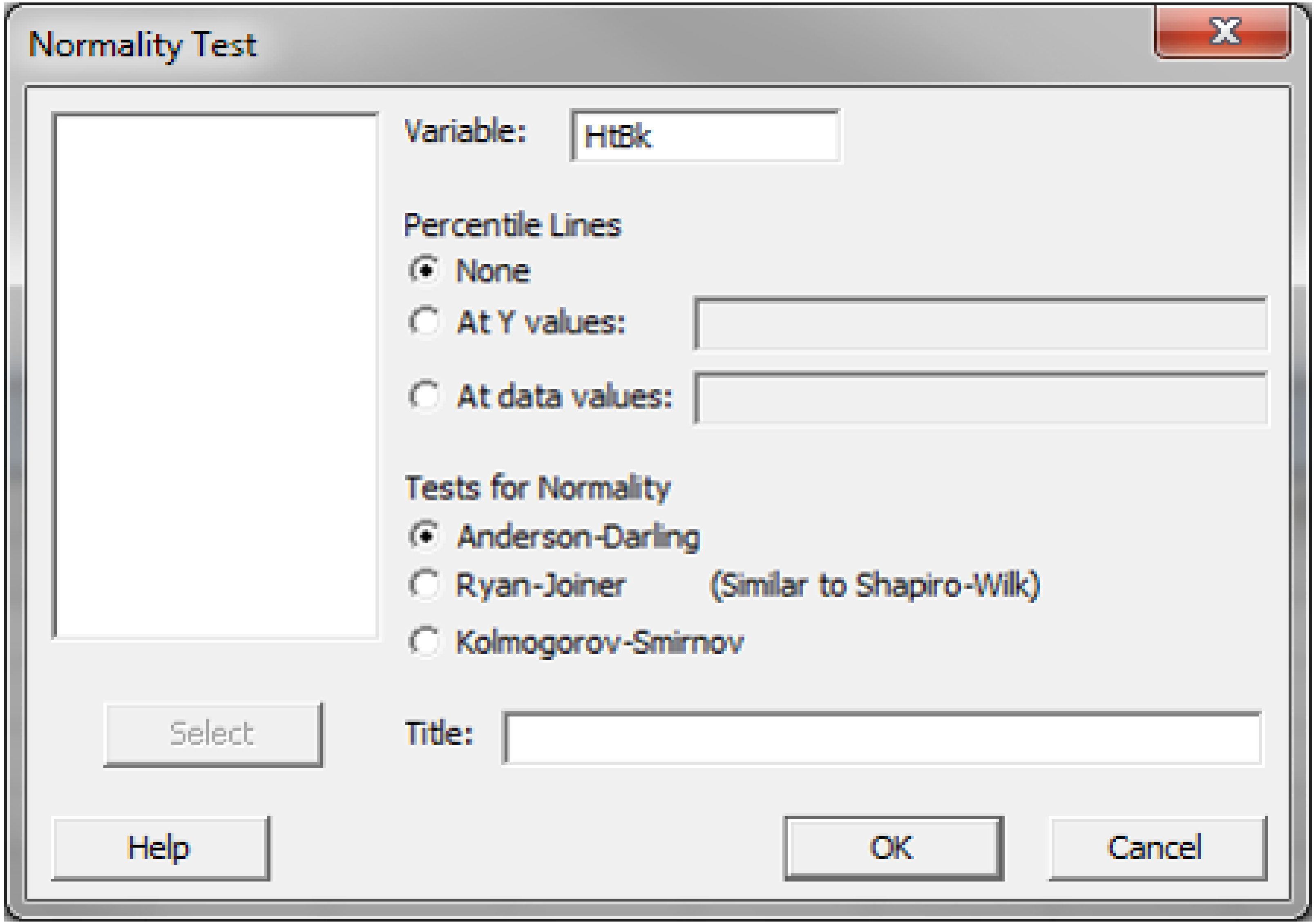
- *Case study:* We are interested in comparing the variance of the height of basketball players with zero.
  - Data File: “One Sample T-Test” tab in “Sample Data.xlsx”

  - Null Hypothesis (H_0): $\sigma^2 = 0$
  - Alternative Hypothesis (H_a): $\sigma^2 \neq 0$

Use Minitab to Run a One Sample Variance Test

- Step 1: Test whether the data are normally distributed
    1. Click Stat → Basic Statistics → Normality Test.
    2. A new window named “Normality Test” pops up.
    3. Select “HtBk” as the “Variable.”
    4. Click “OK.”
    5. The normality test result appears in the new window.

Use Minitab to Run a One Sample Variance Test



Variable: HtBk

Percentile Lines

● None

○ At Y values:

○ At data values:

Tests for Normality

● Anderson-Darling

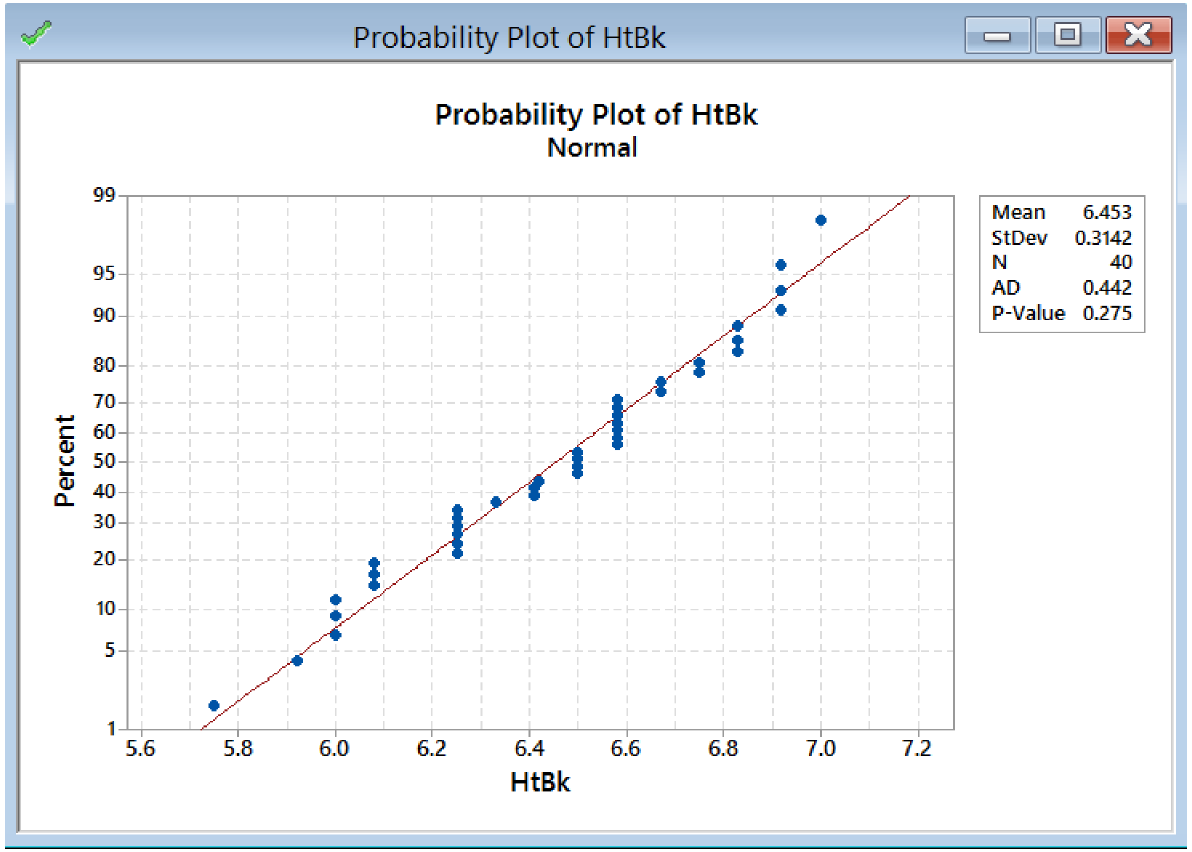
○ Ryan-Joiner (Similar to Shapiro-Wilk)

○ Kolmogorov-Smirnov

Title:

Use Minitab to Run a One Sample Variance Test

- Null Hypothesis (H_0): The data are normally distributed.
- Alternative Hypothesis (H_a): The data are not normally distributed.



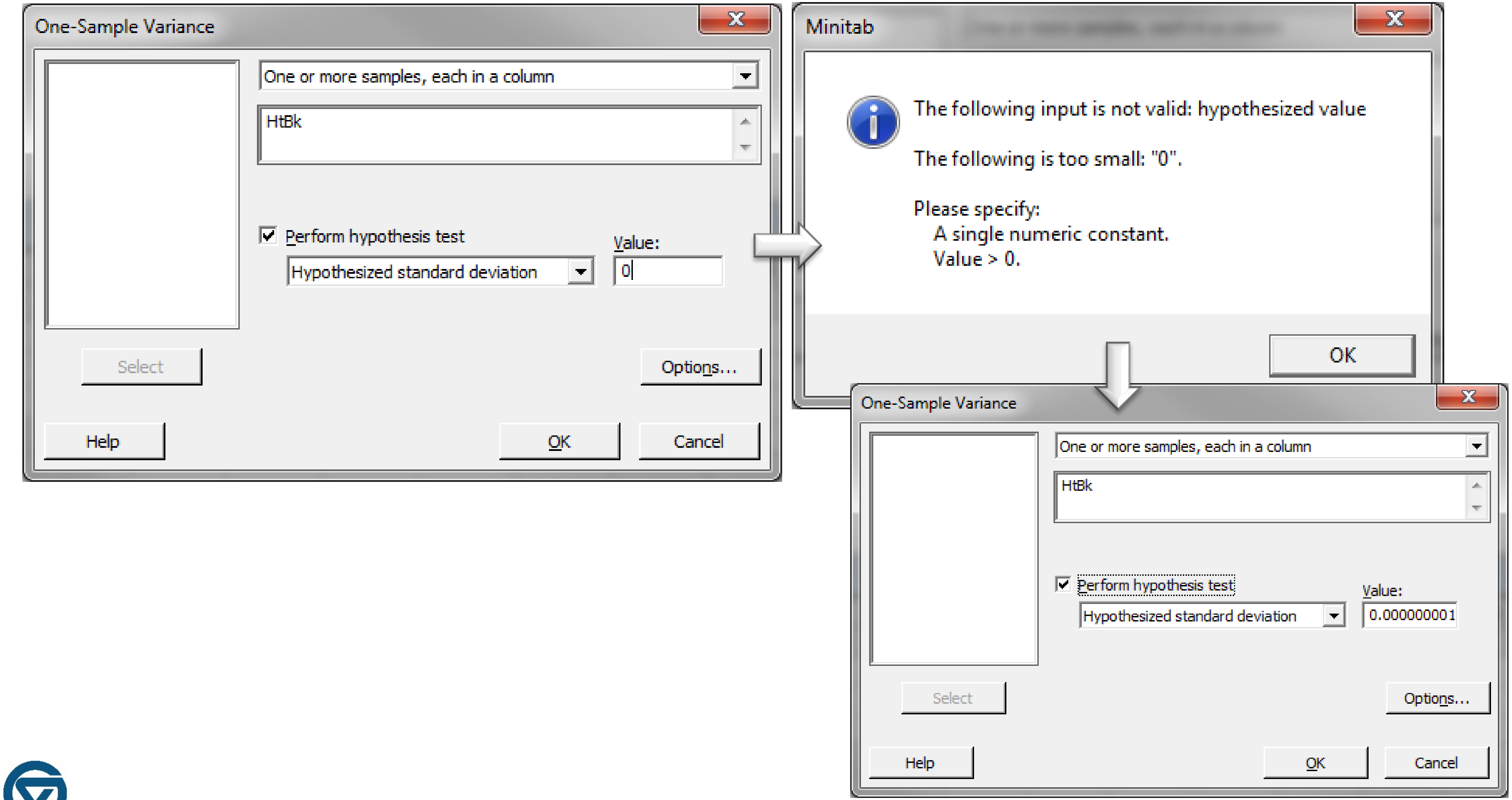
Mean	6.453
StDev	0.3142
N	40
AD	0.442
P-Value	0.275

- Since the p-value of the normality is 0.275, greater than alpha level (0.05), we fail to reject the null and claim that the data are normally distributed.
- If the data are not normally distributed, you need to use other hypothesis tests.

Use Minitab to Run a One Sample Variance Test

- Step 2: Check if the population variance is equal to the specified value (zero).
  1) Click Stat → Basic Statistics → 1 Variance.
  2) A new window named “1 Variance” appears.
  3) Click in the blank box below “Columns” and “HtBk” appears in the list box on the left.
  4) Check the box of “Perform hypothesis test.”
  5) Select “Hypothesized variance” from the dropdown menu.
  6) Enter “0” as the “Hypothesized variance.”
  7) Click “OK.”
  8) An alert window pops up saying “Invalid hypothesized value. The following value is too small: “0”. Please specify: A single numeric constant. Value > 0.”
  9) Click “OK” in the alert message window.
  10) Enter “0.0000000001” as the “Hypothesized variance” instead.
  11) Click “OK.”
  12) The one sample variance test appears in new window.

Use Minitab to Run a One Sample Variance Test



One or more samples, each in a column

HtBk

☑ Perform hypothesis test

Hypothesized standard deviation

Value: 0

The following input is not valid: hypothesized value

The following is too small: "0".

Please specify:
A single numeric constant.
Value > 0.

One or more samples, each in a column

HtBk

☑ Perform hypothesis test

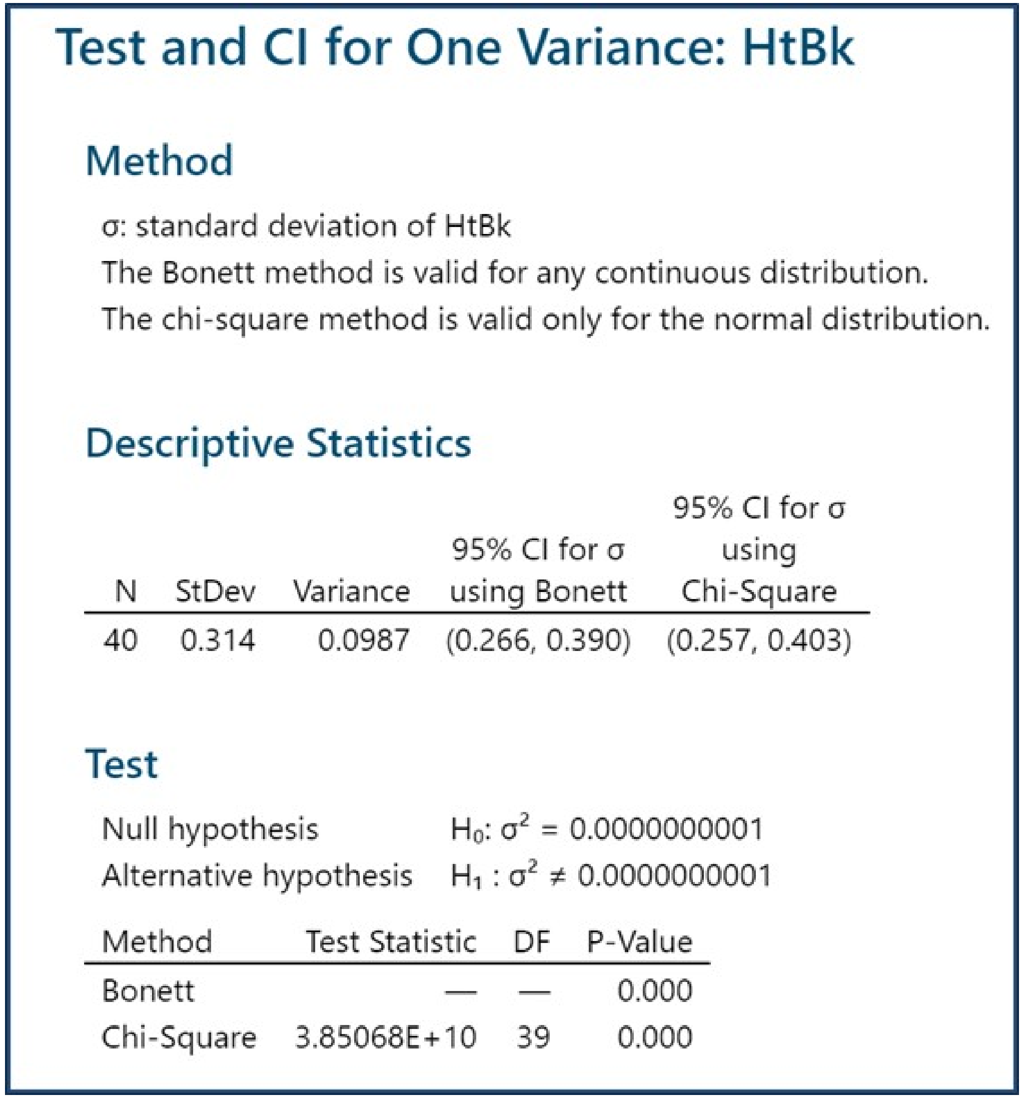
Hypothesized standard deviation

Value: 0.000000001

Use Minitab to Run a One Sample Variance Test

Since the p-value of the one sample variance test is smaller than alpha level (0.05), we reject the null and claim that the population variance is statistically different from zero.

Another way to see whether the population variance is statistically different from zero is to check whether zero stays between the upper and lower confidence interval boundaries of the variance. If yes, we fail to reject the null hypothesis and claim that the population variance is not statistically different from zero.



```
Test and CI for One Variance: HtBk

Method

  σ: standard deviation of HtBk
  The Bonett method is valid for any continuous distribution.
  The chi-square method is valid only for the normal distribution.

Descriptive Statistics

                                                95% CI for σ
                                 95% CI for σ       using
   N   StDev   Variance         using Bonett    Chi-Square
  40   0.314     0.0987   (0.266, 0.390)   (0.257, 0.403)

Test

  Null hypothesis          H₀: σ² = 0.0000000001
  Alternative hypothesis   H₁ : σ² ≠ 0.0000000001

  Method       Test Statistic   DF   P-Value
  Bonett                    —    —     0.000
  Chi-Square      3.85068E+10   39     0.000
```

3.4.3 One Way ANOVA

What is One-Way ANOVA?

- **One-way ANOVA** (one-way analysis of variance) is a statistical method to compare means of two or more populations.
  - Null Hypothesis (H_0): $\mu_1 = \mu_2 = \ldots = \mu_k$
  - Alternative Hypothesis (H_a): at least one μ_i is different, where *i* is any value from 1 to k.
- It is a generalized form of the two sample t-test since a two sample t-test compares two population means and one-way ANOVA compares *k* population means where $k \geq 2$.

Assumptions of One-Way ANOVA

- The sample data drawn from *k* populations are unbiased and representative.
- The data of *k* populations are continuous.
- The data of *k* populations are normally distributed.
- The variances of *k* populations are equal.

How ANOVA Works

- ANOVA compares the means of different groups by analyzing the variances between and within groups.

- Let us say we are interested in comparing the means of three normally distributed populations. We randomly collected one sample for each population of our interest.

  - Null Hypothesis (H_0): $\mu_1 = \mu_2 = \mu_3$
  - Alternative Hypothesis (H_a): one of the μ is different from the others.

How ANOVA Works

- Based on the sample data, the means of the three populations might look different because of two variation sources.
  1) Variation between groups

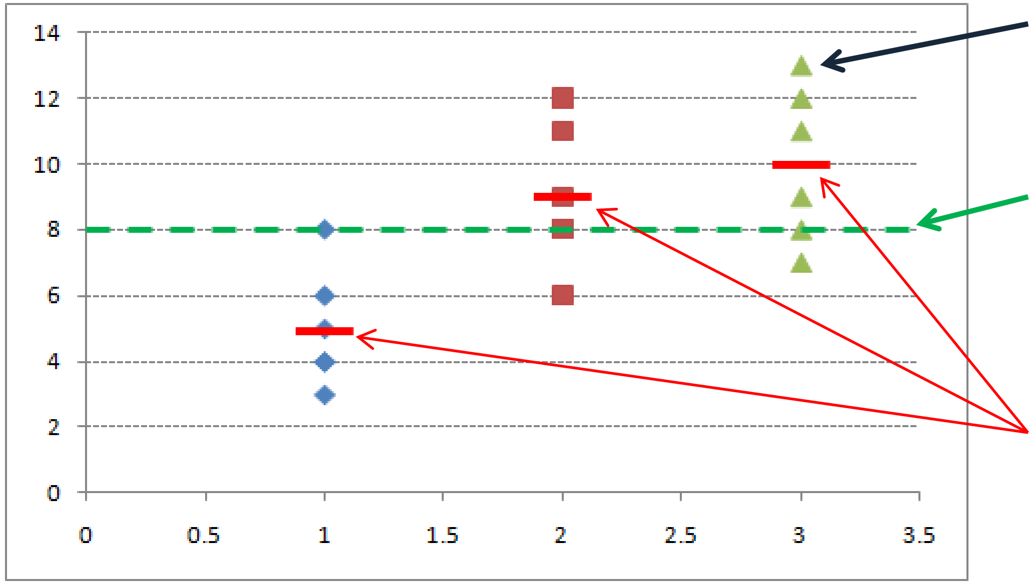
     There are non-random factors leading to the variation between groups.
  2) Variation within groups

     There are random errors resulting in the variation within each individual group.

- What we care about the most is the variation between groups since we are interested in whether the groups are statistically different from each other.

- Variation between groups is the *signal* we want to detect and variation within groups is the *noise* which corrupts the signal.

How ANOVA Works



Individual observation: Y_{ij}

Grand mean of all observations: $\bar{\bar{Y}}$

Group means of each individual sample: $\bar{Y}_j$

$$\text{Total Variation} = \text{SS(Total)} = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(Y_{ij} - \bar{\bar{Y}})^2$$

$$\text{Between Variation} = \text{SS(Between)} = \sum_{j=1}^{k} n_j(\bar{Y}_j - \bar{\bar{Y}})^2$$

$$\text{Within Variation} = \text{SS(Within)} = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(Y_{ij} - \bar{Y}_j)^2$$

How ANOVA Works

- Variation Components
  - Total Variation = Variation Between Groups + Variation Within Groups

$$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(Y_{ij}-\bar{\bar{Y}})^2 = \sum_{j=1}^{k} n_j(\bar{Y}_j-\bar{\bar{Y}})^2 + \sum_{j=1}^{k}\sum_{i=1}^{n_j}(Y_{ij}-\bar{Y}_j)^2$$

    - Total Variation = sums of squares of the vertical difference between the individual observation and the grand mean
    - Variation Between Groups = sums of squares of the vertical difference between the group mean and the grand mean
    - Variation Within Groups = sum of squares of the vertical difference between the individual observation and the group mean

How ANOVA Works

- Degrees of Freedom (DF)
  - In statistics, the degrees of freedom is the number of unrestricted values in the calculation of a statistic.

- Degrees of Freedom Components
  - $DF_{total} = DF_{between} + DF_{within}$
  - $DF_{total} = n - 1$
  - $DF_{between} = k - 1$
  - $DF_{within} = n - k$

where

n is the total number of observations
k is the number of groups.

How ANOVA Works

- Signal-to-Noise Ratio (SNR)
  - SNR denotes the ratio of a signal to the noise corrupting the signal.
  - It measures how much a signal has been corrupted by the noise.
  - When it is higher than 1, there is more signal than noise.
  - The higher the SNR, the less the signal has been corrupted by the noise.
- F-ratio is the SNR in ANOVA

$$F = \frac{MS_{between}}{MS_{within}} = \frac{SS_{between} \big/ DF_{between}}{SS_{within} \big/ DF_{within}} = \frac{\sum_{j=1}^{k} (\overline{Y}_j - \overline{\overline{Y}})^2 \Big/ (k-1)}{\sum_{j=1}^{k} \sum_{i=1}^{n_j} \left(Y_{ij} - \overline{Y}_j\right)^2 \Big/ (n-k)}$$

  - In ANOVA, we use the F-test to compare the means of different groups. The F-ratio calculated as above is the test statistic F_{calc}.
  - The critical value (F_{cri}) in an F-test can be derived from the F table with predetermined significance level (α) and with (k –1) degrees of freedom in the numerator and (n – k) degrees of freedom in the denominator.

How ANOVA Works

- Null Hypothesis (H_0): $\mu_1 = \mu_2 = \ldots = \mu_k$
- Alternative Hypothesis (H_a): at least one μ_i is different, where *i* is any value from 1 to *k*.
- If $|F_{calc}| < F_{crit}$, we fail to reject the null and claim that the means of all the populations of our interest are the same.
- If $|F_{calc}| > F_{crit}$, we reject the null and claim that there is at least one mean different from the others.

Model Validation

- ANOVA is a modeling procedure. To make sure the conclusions made in ANOVA are reliable, we need to perform residuals analysis.

- Good residuals:
  - Have a mean of zero
  - Are normally distributed
  - Are independent of each other
  - Have equal variance.

Use Minitab to Run an ANOVA

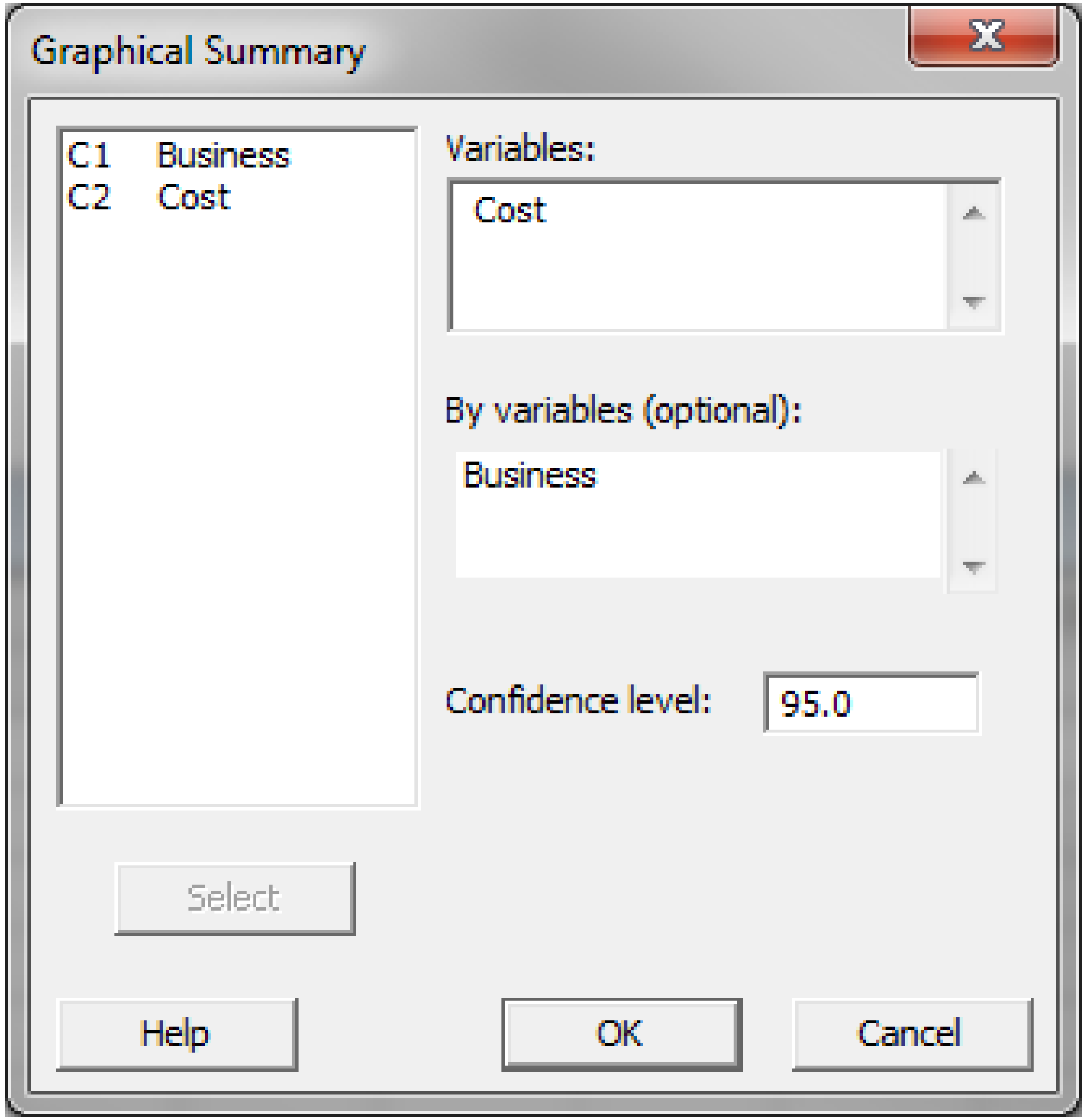
- *Case study:* We are interested in comparing the average startup costs of five kinds of business.
  - Data File: “One-Way ANOVA” tab in “Sample Data.xlsx”

- Null Hypothesis (H_0): $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$
- Alternative Hypothesis (H_a): at least one of the five means is different from others.

Use Minitab to Run an ANOVA

- Step 1: Test whether the data for each level are normally distributed.
  1) Click Stat → Basic Statistics → Graphical Summary.
  2) A new window named "Graphical Summary" pops up.
  3) Select the "Cost" as the variable.
  4) Click in the blank box right next to "By variables (optional)" and the "Business" appears in the list box on the left.
  5) Select the "Business" as the "By variables (optional)."
  6) Click "OK."
  7) The normality results appear in the new window.

Use Minitab to Run an ANOVA



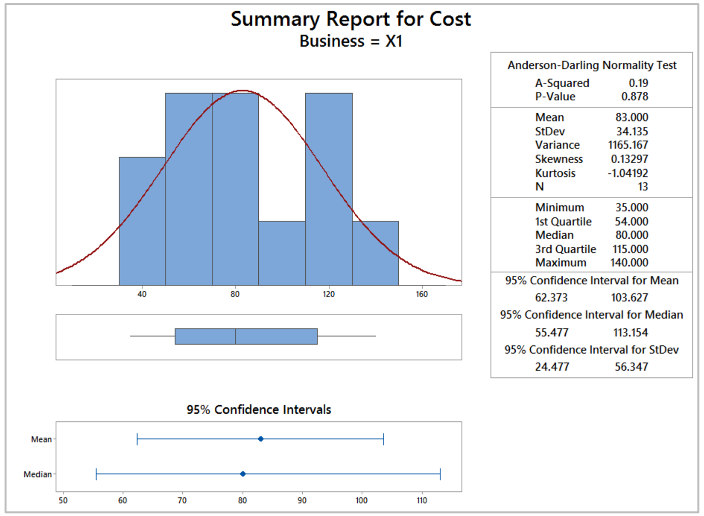
Variables: Cost

By variables (optional): Business

Confidence level: 95.0

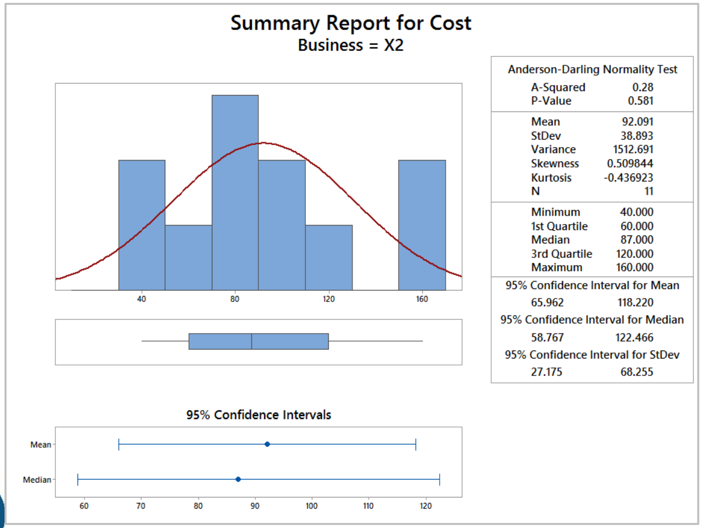
Use Minitab to Run an ANOVA

Business = X1



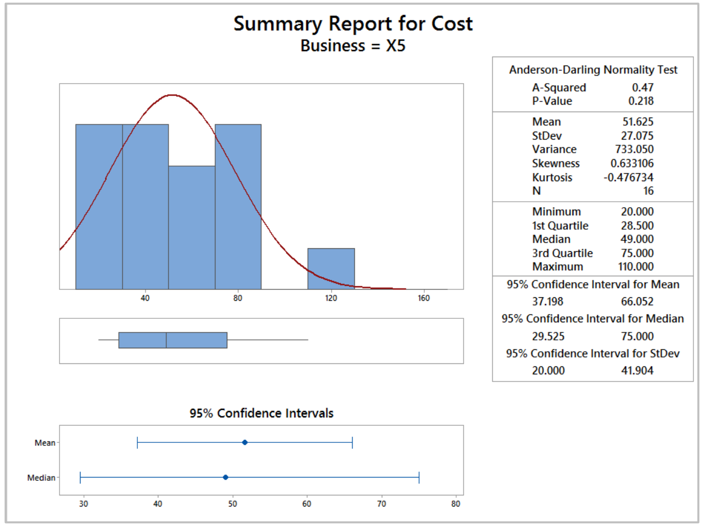
Anderson-Darling Normality Test	
A-Squared	0.19
P-Value	0.878
Mean	83.000
StDev	34.135
Variance	1165.167
Skewness	0.13297
Kurtosis	-1.04192
N	13
Minimum	35.000
1st Quartile	54.000
Median	80.000
3rd Quartile	115.000
Maximum	140.000
95% Confidence Interval for Mean	
62.373	103.627
95% Confidence Interval for Median	
55.477	113.154
95% Confidence Interval for StDev	
24.477	56.347

Business = X2



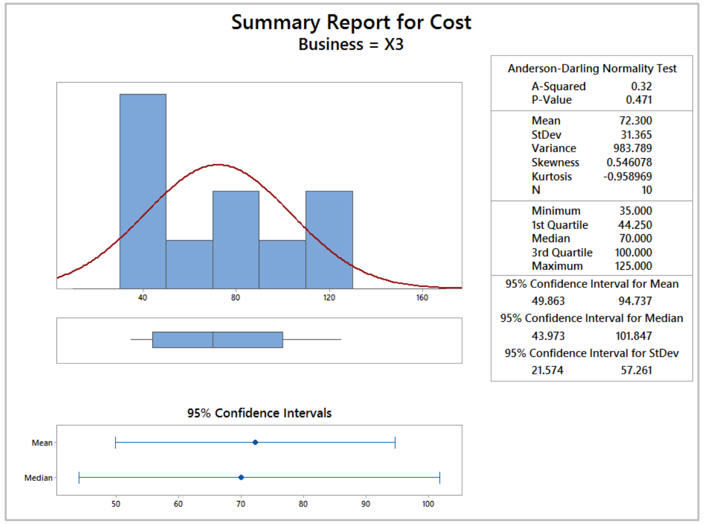
Anderson-Darling Normality Test	
A-Squared	0.28
P-Value	0.581
Mean	92.091
StDev	38.893
Variance	1512.691
Skewness	0.509844
Kurtosis	-0.436923
N	11
Minimum	40.000
1st Quartile	60.000
Median	87.000
3rd Quartile	120.000
Maximum	160.000
95% Confidence Interval for Mean	
65.962	118.220
95% Confidence Interval for Median	
58.767	122.466
95% Confidence Interval for StDev	
27.175	68.255

Business = X5



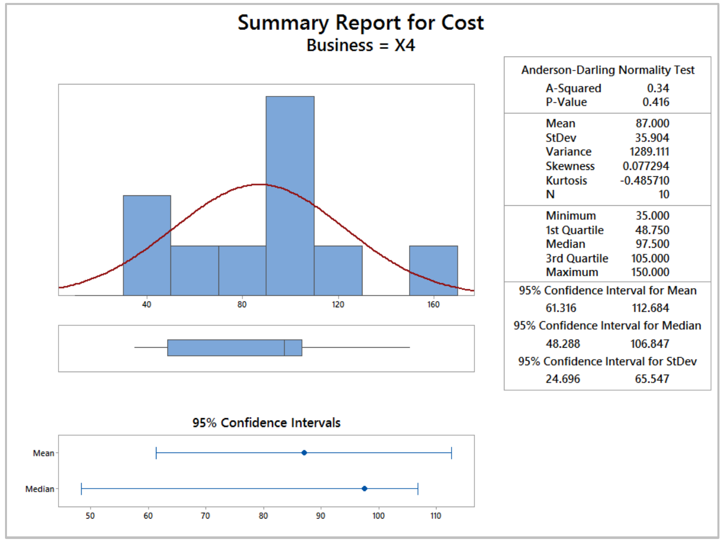
Anderson-Darling Normality Test	
A-Squared	0.47
P-Value	0.218
Mean	51.625
StDev	27.075
Variance	733.050
Skewness	0.633106
Kurtosis	-0.476734
N	16
Minimum	20.000
1st Quartile	28.500
Median	49.000
3rd Quartile	75.000
Maximum	110.000
95% Confidence Interval for Mean	
37.198	66.052
95% Confidence Interval for Median	
29.525	75.000
95% Confidence Interval for StDev	
20.000	41.904

Business = X3



Anderson-Darling Normality Test	
A-Squared	0.32
P-Value	0.471
Mean	72.300
StDev	31.365
Variance	983.789
Skewness	0.546078
Kurtosis	-0.958969
N	10
Minimum	35.000
1st Quartile	44.250
Median	70.000
3rd Quartile	100.000
Maximum	125.000
95% Confidence Interval for Mean	
49.863	94.737
95% Confidence Interval for Median	
43.973	101.847
95% Confidence Interval for StDev	
21.574	57.261

Business = X4



Anderson-Darling Normality Test	
A-Squared	0.34
P-Value	0.416
Mean	87.000
StDev	35.904
Variance	1289.111
Skewness	0.077294
Kurtosis	-0.485710
N	10
Minimum	35.000
1st Quartile	48.750
Median	97.500
3rd Quartile	105.000
Maximum	150.000
95% Confidence Interval for Mean	
61.316	112.684
95% Confidence Interval for Median	
48.288	106.847
95% Confidence Interval for StDev	
24.696	65.547

Use Minitab to Run an ANOVA

Null Hypothesis (H_0): The data are normally distributed.

Alternative Hypothesis (H_a): The data are not normally distributed.

- Since the p-values of normality tests for the five data sets are higher than alpha level (0.05), we fail to reject the null hypothesis and claim that the startup costs for any of the five businesses are normally distributed.

- If any of the five data sets are not normally distributed, we need to use other hypothesis testing methods other than one-way ANOVA.

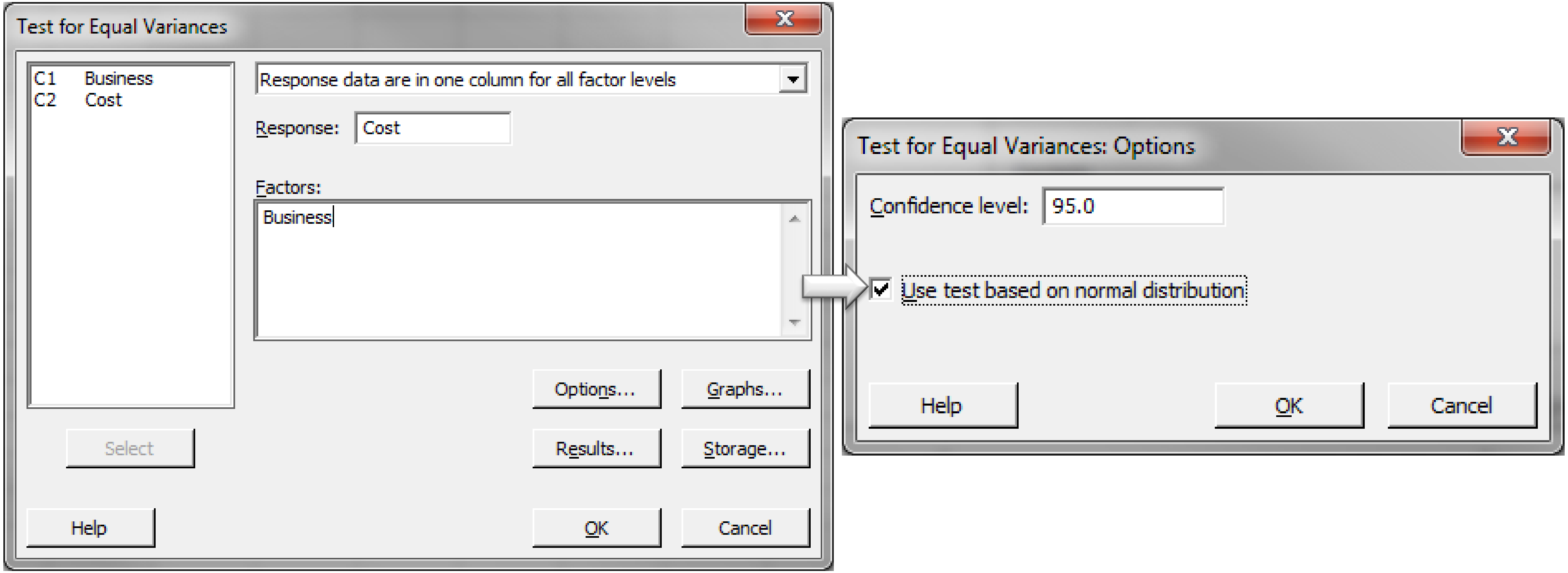
Use Minitab to Run an ANOVA

- Step 2: Test whether the variance of the data for each level is equal to the variance of other levels.

Null Hypothesis (H_0): $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2$
Alternative Hypothesis (H_a): at least one of the variances is different from others.

1) Click Stat → ANOVA → Test for equal variances.
2) A new window named "Test for Equal Variances" pops up.
3) Select the "Cost" as the "Response."
4) Select the "Business" as the "Factors."
5) Click "Options."
6) Select "Use test based on normal distribution"
7) Click "OK" to close the "Options" window
8) Click "OK" to run the test
9) The results shows up in a new window.
10) Use the Bartlett's test for testing the equal variances between five levels in this case since there are more than two levels in the data and the data of each level are normally distributed.

Use Minitab to Run an ANOVA



Test for Equal Variances

C1 Business
C2 Cost

Response data are in one column for all factor levels

Response: Cost

Factors: Business

Test for Equal Variances: Options

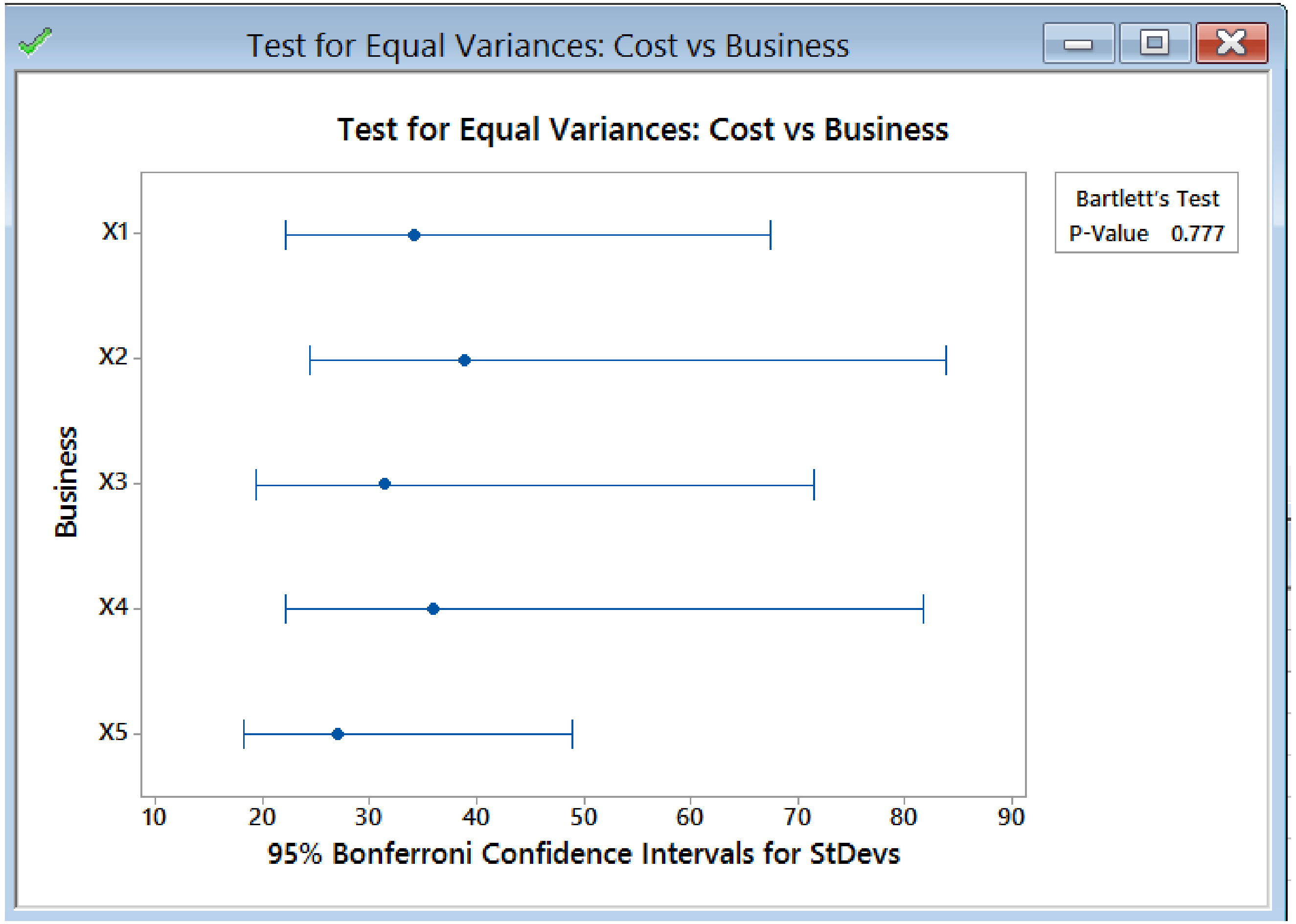
Confidence level: 95.0

☑ Use test based on normal distribution

Use Minitab to Run an ANOVA

- The p-value of Barlett's test is 0.777, greater than the alpha level (0.05), so we fail to reject the null hypothesis and we claim that the variances of five groups are equal.

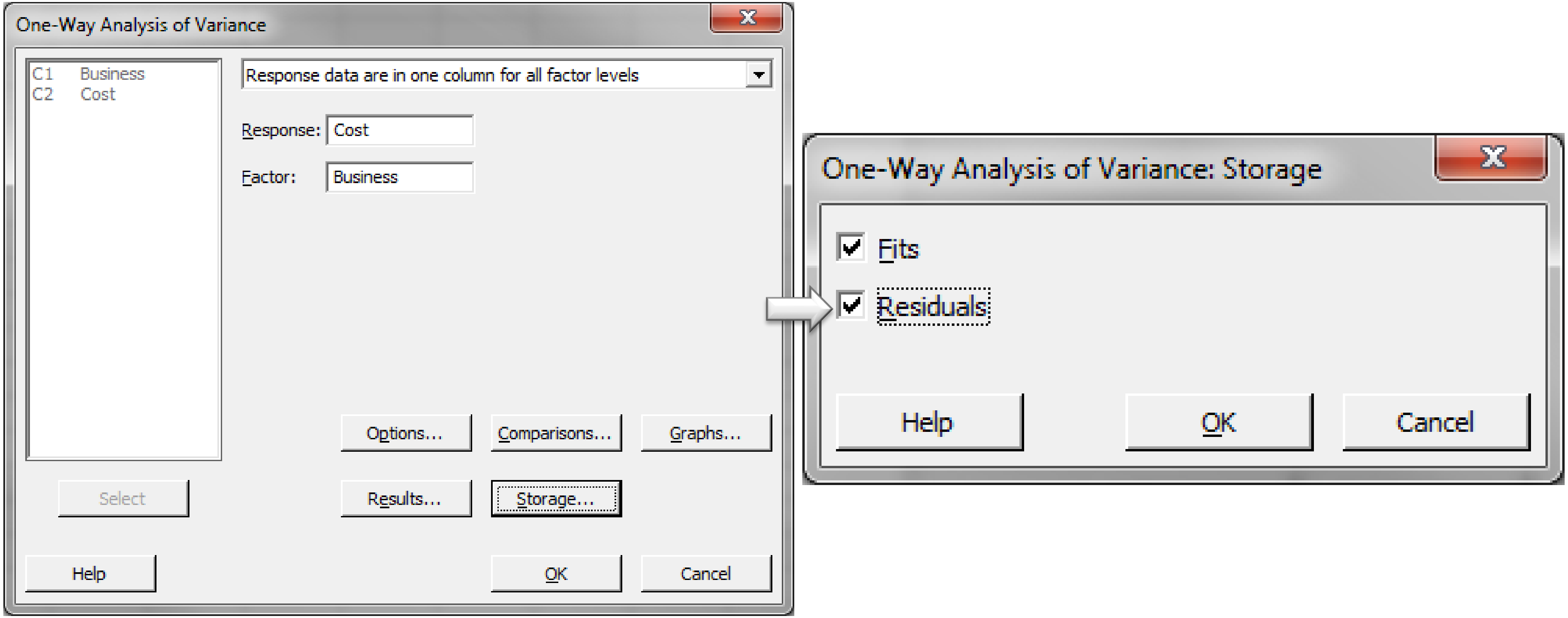
- If the variances are not all equal, we need to use other hypothesis testing methods other than one-way ANOVA.



Use Minitab to Run an ANOVA

- Step 3: Test whether the mean of the data for each level is equal to the means of other levels.
  - Null Hypothesis (H_0): $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$
  - Alternative Hypothesis (H_a): at lease one of the means is different from others.
    1) Click Stat → ANOVA → One-way.
    2) A new window named “One-Way Analysis of Variance” pops up.
    3) Select “Cost” as “Response.”
    4) Select “Business” as “Factor.”
    5) Click “Storage.”
    6) Check the boxes next to “Fits” and “Residuals”
    7) Click “OK” to close the “Storage” window
    8) Click “OK.”
    9) The ANOVA results appear in the session window. The fitted response and the residuals are stored in the data table.

Use Minitab to Run an ANOVA



One-Way Analysis of Variance

C1 Business
C2 Cost

Response data are in one column for all factor levels

Response: Cost

Factor: Business

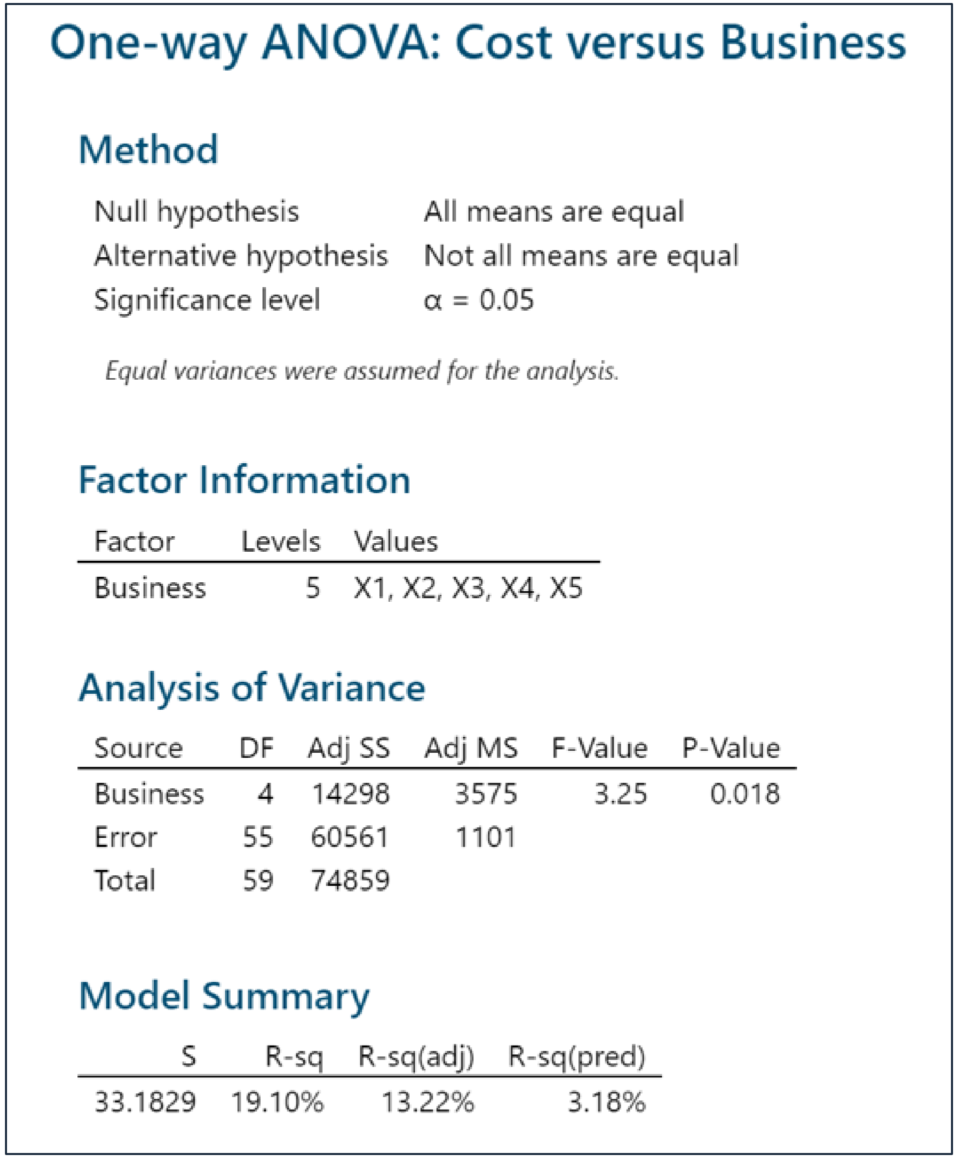
One-Way Analysis of Variance: Storage

☑ Fits

☑ Residuals

Use Minitab to Run an ANOVA

- Since the p-value of the F test is 0.018, lower than the alpha level (0.05), the null hypothesis is rejected and we conclude that the at least one of the means of the five groups is different from others.

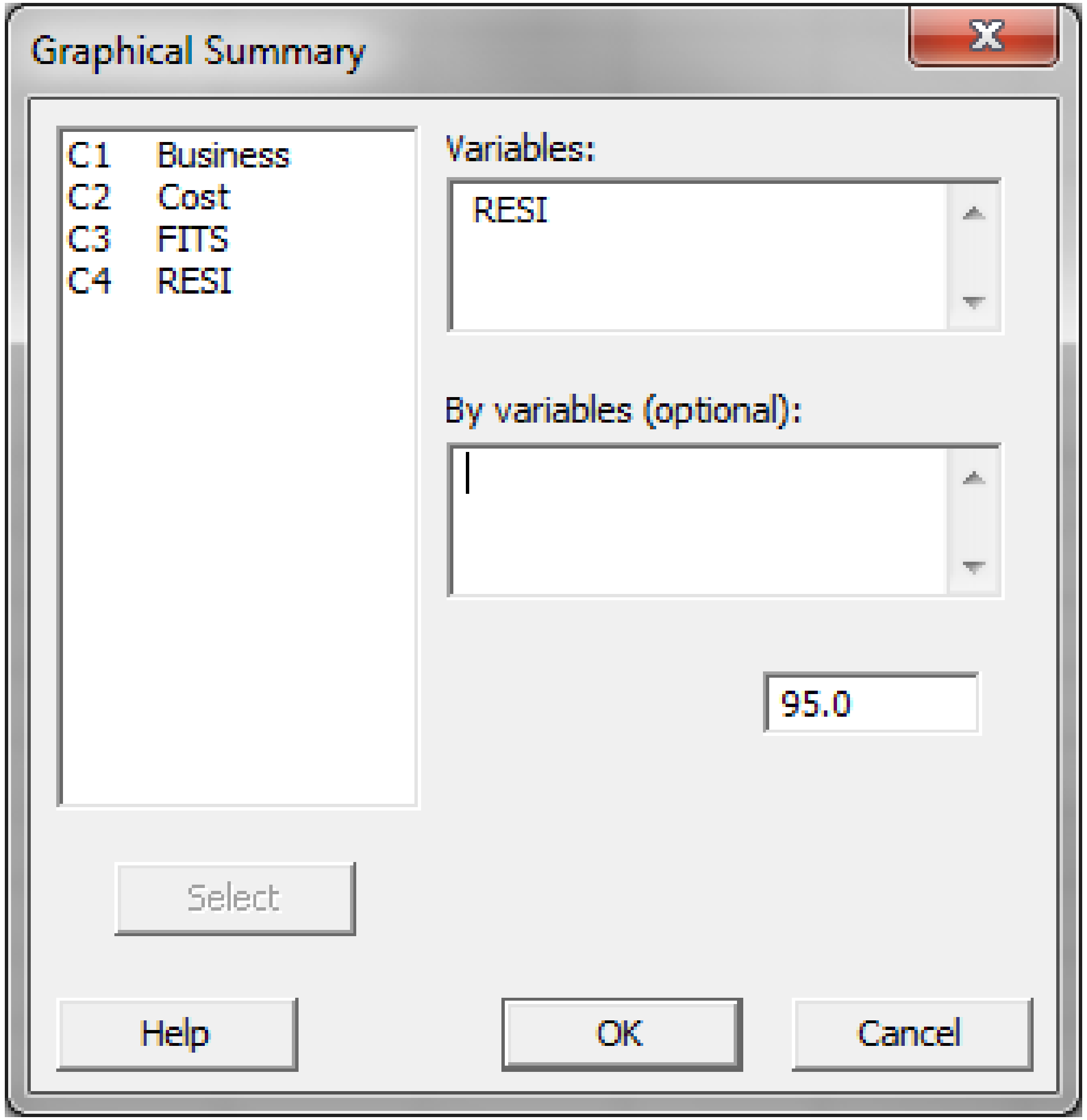


```
One-way ANOVA: Cost versus Business

Method

  Null hypothesis          All means are equal
  Alternative hypothesis   Not all means are equal
  Significance level       α = 0.05

  Equal variances were assumed for the analysis.

Factor Information

  Factor     Levels  Values
  Business        5  X1, X2, X3, X4, X5

Analysis of Variance

  Source    DF  Adj SS  Adj MS  F-Value  P-Value
  Business   4   14298    3575     3.25    0.018
  Error     55   60561    1101
  Total     59   74859

Model Summary

        S    R-sq  R-sq(adj)  R-sq(pred)
  33.1829  19.10%     13.22%       3.18%
```

Use Minitab to Run an ANOVA

- Step 4: Test whether the residuals are normally distributed with mean equal zero
    1) The residuals have been stored in the data table in step 4.
    2) Click Stat → Basic Statistics → Graphical Summary.
    3) A new window named “Graphical Summary” appears.
    4) Select “RESI” as the “Variables.”
    5) Click “OK.”
    6) The normality test results show up in a new window.

Use Minitab to Run an ANOVA



Variables: RESI

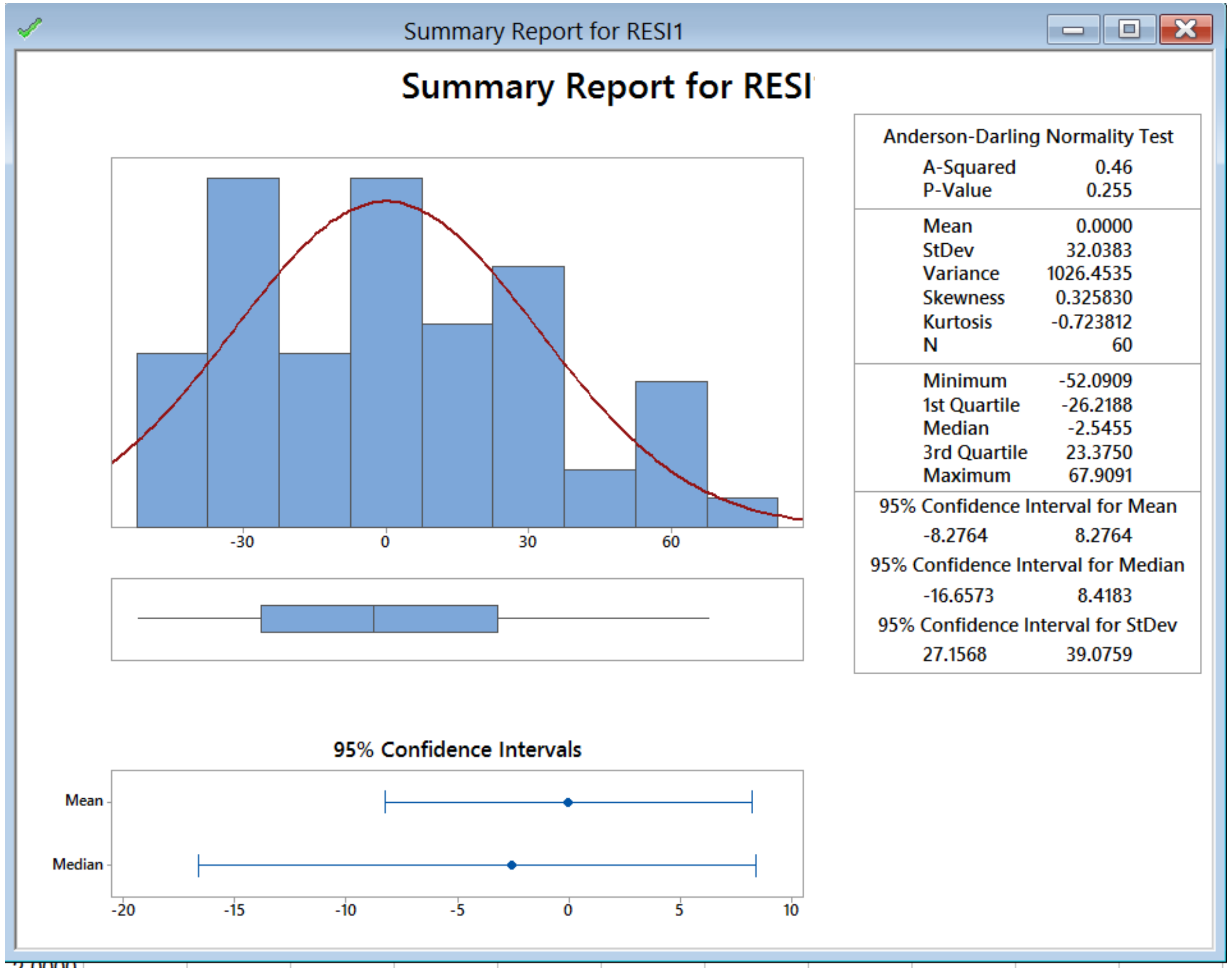
By variables (optional):

95.0

Use Minitab to Run an ANOVA

- The p-value of the normality test is 0.255, greater than the alpha level (0.05), and we conclude that the residuals are normally distributed.

- The mean of the residuals are 0.0000.



Summary Report for RESI

Anderson-Darling Normality Test

A-Squared	0.46
P-Value	0.255
Mean	0.0000
StDev	32.0383
Variance	1026.4535
Skewness	0.325830
Kurtosis	-0.723812
N	60
Minimum	-52.0909
1st Quartile	-26.2188
Median	-2.5455
3rd Quartile	23.3750
Maximum	67.9091

95% Confidence Interval for Mean

-8.2764	8.2764

95% Confidence Interval for Median

-16.6573	8.4183

95% Confidence Interval for StDev

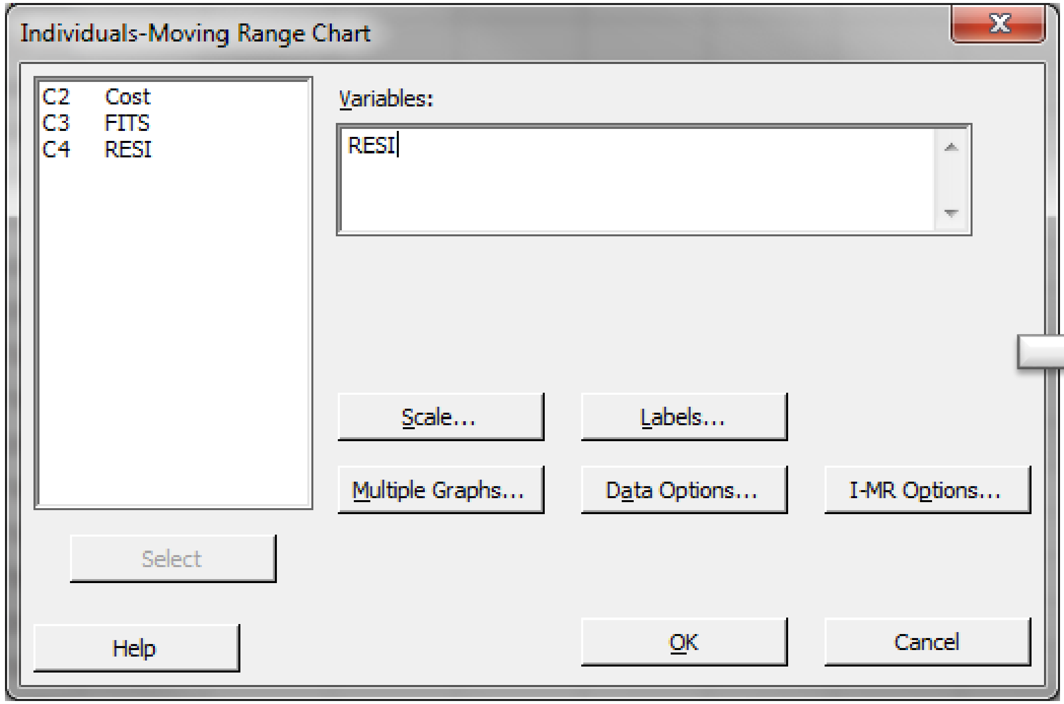
27.1568	39.0759

Use Minitab to Run an ANOVA

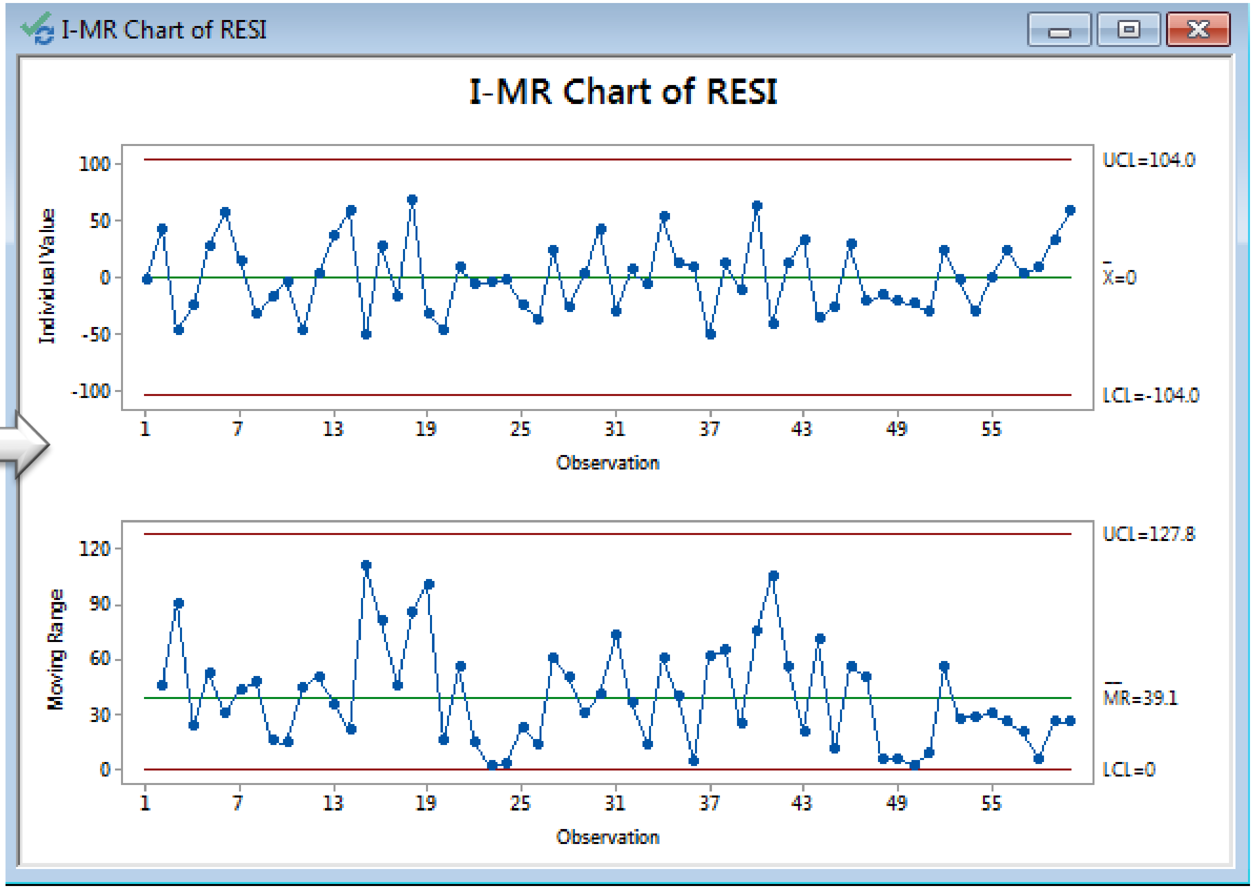
- Step 5: Check whether the residuals are independent of each other.
  - If the residuals are in time order, we can plot I-MR charts to check the independence. When no data points on the I-MR charts fail any tests, the residuals are independent of each other.
  - If the residuals are not in time order, the I-MR charts cannot deliver reliable conclusion on the independence.

    1) Click Stat → Control Charts → Variables Charts for Individuals → I-MR.
    2) A new window named “Individuals – Moving Range Chart” pops up.
    3) Select “RESI1” as the “Variables.”
    4) Click “OK.”

Use Minitab to Run an ANOVA



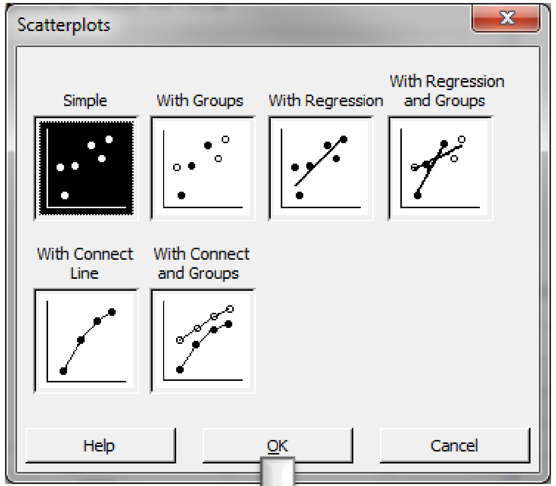
Variables: RESI

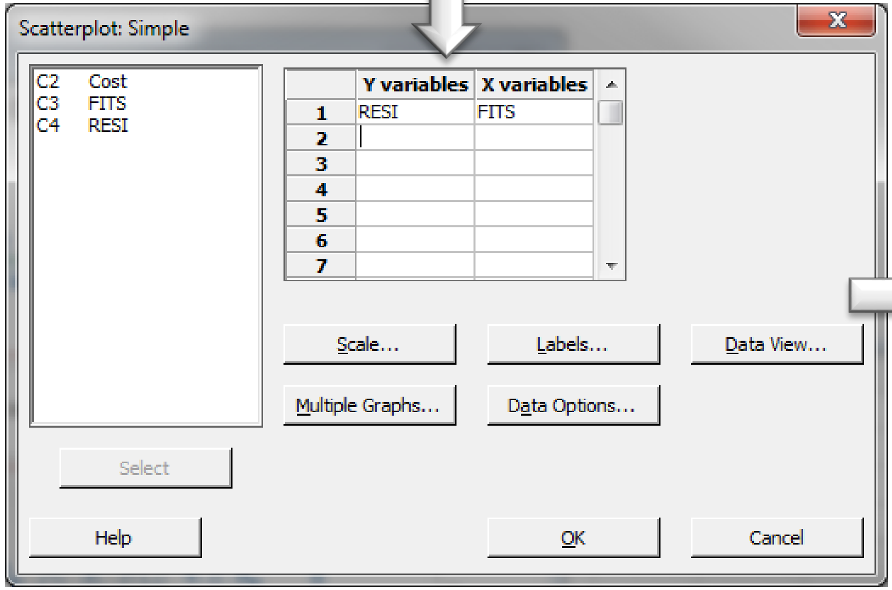


Use Minitab to Run an ANOVA

- Step 6: Plot residuals versus fitted values and check whether there is any systematic pattern.
    1) Click Graph → Scatter Plot.
    2) A new window named “Scatterplots” pops up.
    3) Click “OK.”
    4) A new window named “Scatterplot – Simple” pops up.
    5) Select the “RESI1” as the “Y variables.”
    6) Select the “FITS1” as “X variables.”
    7) Click “OK.”
    8) The charts appears in a new window.
    9) If the data points spread out evenly at any of the five levels, we claim that the residuals have equal variances across the five levels.

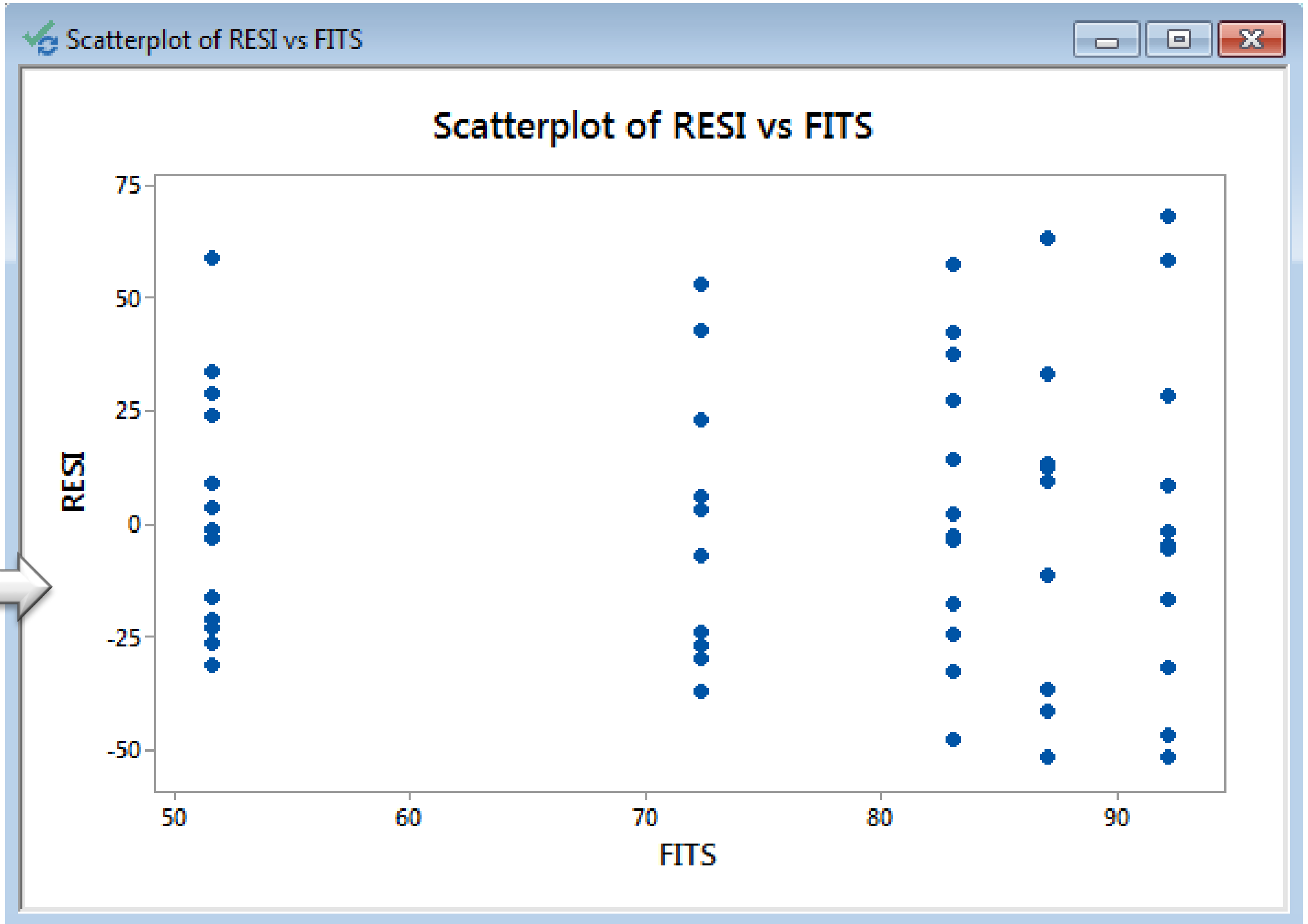
Use Minitab to Run an ANOVA





C2 Cost
C3 FITS
C4 RESI

	Y variables	X variables
1	RESI	FITS
2		
3		
4		
5		
6		
7		



3.5 Hypothesis Testing: Non-Normal Data

Black Belt Training: Analyze Phase

3.5.1 Mann-Whitney

What is the Mann-Whitney Test?

- The **Mann-Whitney test** (also called Mann-Whitney U test or Wilcoxon rank-sum test) is a statistical hypothesis test to compare the medians of two populations that are not normally distributed.
  - Null Hypothesis (H_0): $\eta_1 = \eta_2$
  - Alternative Hypothesis (H_a): $\eta_1 \neq \eta_2$

where η_1 is the median of one population and η_1 is the median of the other population.

Mann-Whitney Test Assumptions

- The sample data drawn from the populations of interest are unbiased and representative.
- The data of both populations are continuous or ordinal when the spacing between adjacent values is not constant.
- The two populations are independent to each other.
- The Mann-Whitney test is robust for the non-normally distributed population.
- The Mann-Whitney test can be used when shapes of the two populations' distributions are different.

How Mann-Whitney Test Works

- Step 1: Group the two samples from two populations (sample 1 is from population 1 and sample 2 is from population 2) into a single data set and then sort the data in ascending order ranked from 1 to n, where n is the total number of observations.

- Step 2: Add up the ranks for all the observations from sample 1 and call it $\mathbf{R_1}$. Add up the ranks for all the observations from sample 2 and call it $\mathbf{R_2}$.

How Mann-Whitney Test Works

- Step 3: Calculate the test statistics

$$U = \min(U_1, U_2)$$

where $$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

n_1 and n_2 are the sample sizes.

R_1 and R_2 are the sum of ranks for observations from sample 1 and 2 respectively.

How Mann-Whitney Test Works

- Step 4: Make a decision on whether to reject the null hypothesis.
  - Null Hypothesis (H_0): $\eta_1 = \eta_2$
  - Alternative Hypothesis (H_a): $\eta_1 \neq \eta_2$

- If both of the sample sizes are smaller than 10, the distribution of U under the null hypothesis is tabulated.
  - The test statistic is U and, by using the Mann-Whitney table, we would find the p-value.
  - If the p-value is smaller than alpha level (0.05), we reject the null hypothesis.
  - If the p-value is greater than alpha level (0.05), we fail to reject the null hypothesis.

How Mann-Whitney Test Works

- If both sample sizes are greater than 10, the distribution of U can be approximated by a normal distribution. In other words, $\frac{U-\mu}{\sigma}$ follows a standard normal distribution.

$$Z_{calc} = \frac{U-\mu}{\sigma}$$

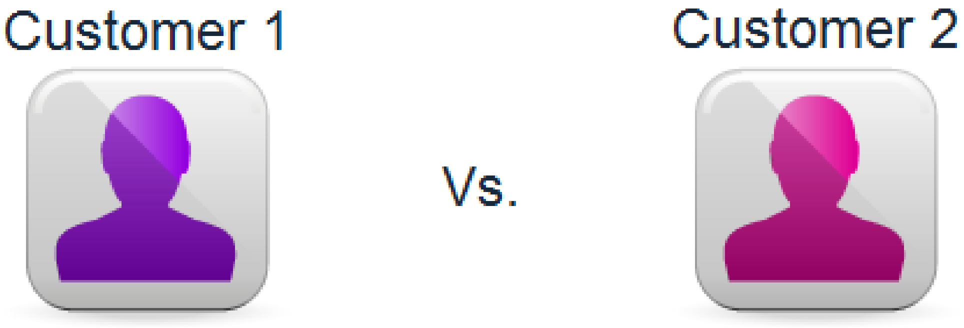
where

$$\mu = \frac{n_1 n_2}{2} \qquad \sigma = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

When $|Z_{calc}|$ is greater than Z value at α/2 level (e.g., when α = 5%, the z value we compare $|Z_{calc}|$ to is 1.96), we reject the null hypothesis.

Use Minitab to Run a Mann-Whitney Test

- *Case study:* We are interested in comparing customer satisfaction between two types of customers using a nonparametric (i.e., distribution-free) hypothesis test: Mann-Whitney test.
  - Data File: “Mann-Whitney” tab in “Sample Data.xlsx”


Customer 1

Vs.

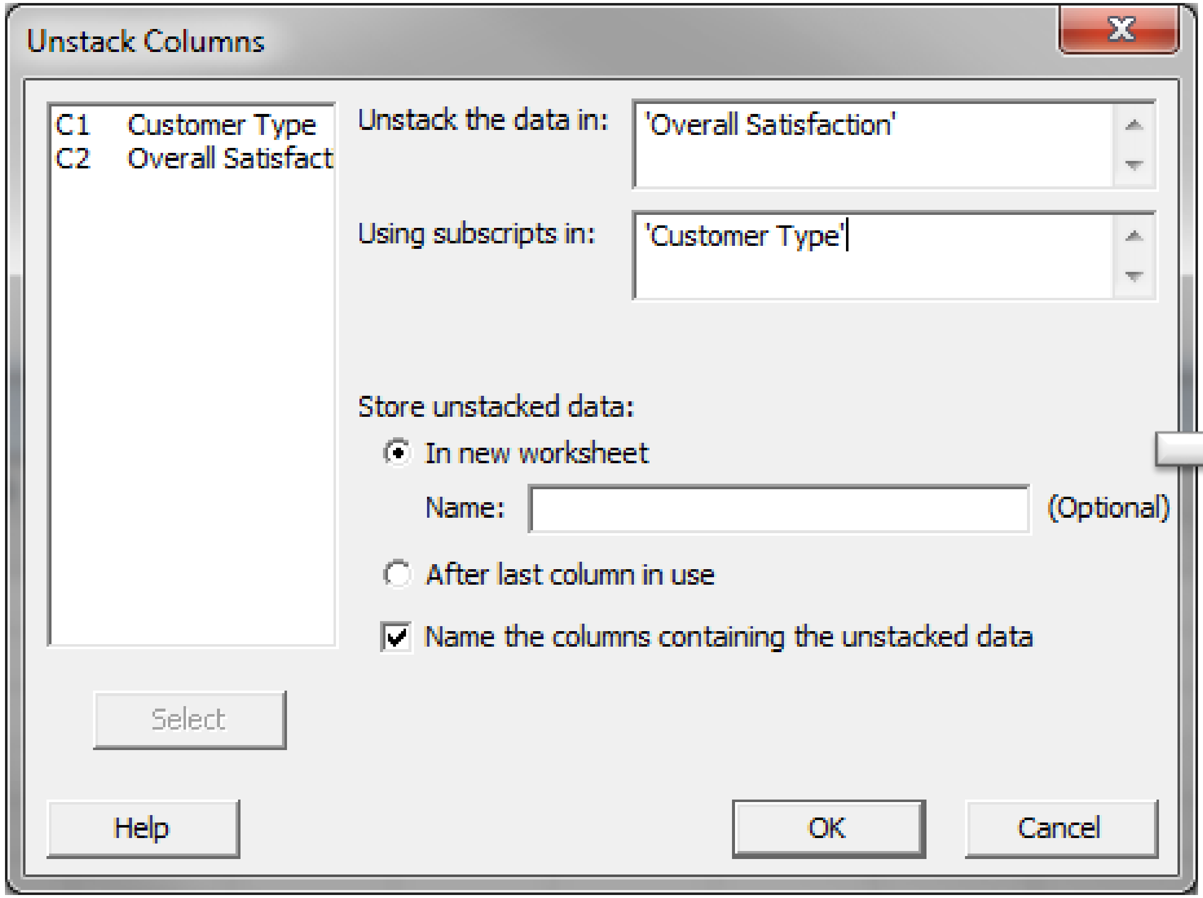
Customer 2

- Null Hypothesis (H_0): $\eta_1 = \eta_2$
- Alternative Hypothesis (H_a): $\eta_1 \neq \eta_2$

Use Minitab to Run a Mann-Whitney Test

- Steps to run a Mann-Whitney Test in Minitab:
  1) Un-stack the data into two separate columns:
     - “Overall Satisfaction 1” for customer type = 1
     - “Overall Satisfaction 2” for customer type = 2
  2) Click Stat → Nonparametrics → Mann-Whitney.
  3) A new window named “Mann-Whitney” pops up.
  4) Select “Overall Satisfaction 1” as the “First Sample.”
  5) Select “Overall Satisfaction 2” as the “Second Sample.”
  6) Click “OK.”
  7) The Mann-Whitney test results appear in the session window.

Use Minitab to Run a Mann-Whitney Test



C1 Customer Type
C2 Overall Satisfact

Unstack the data in: 'Overall Satisfaction'

Using subscripts in: 'Customer Type'

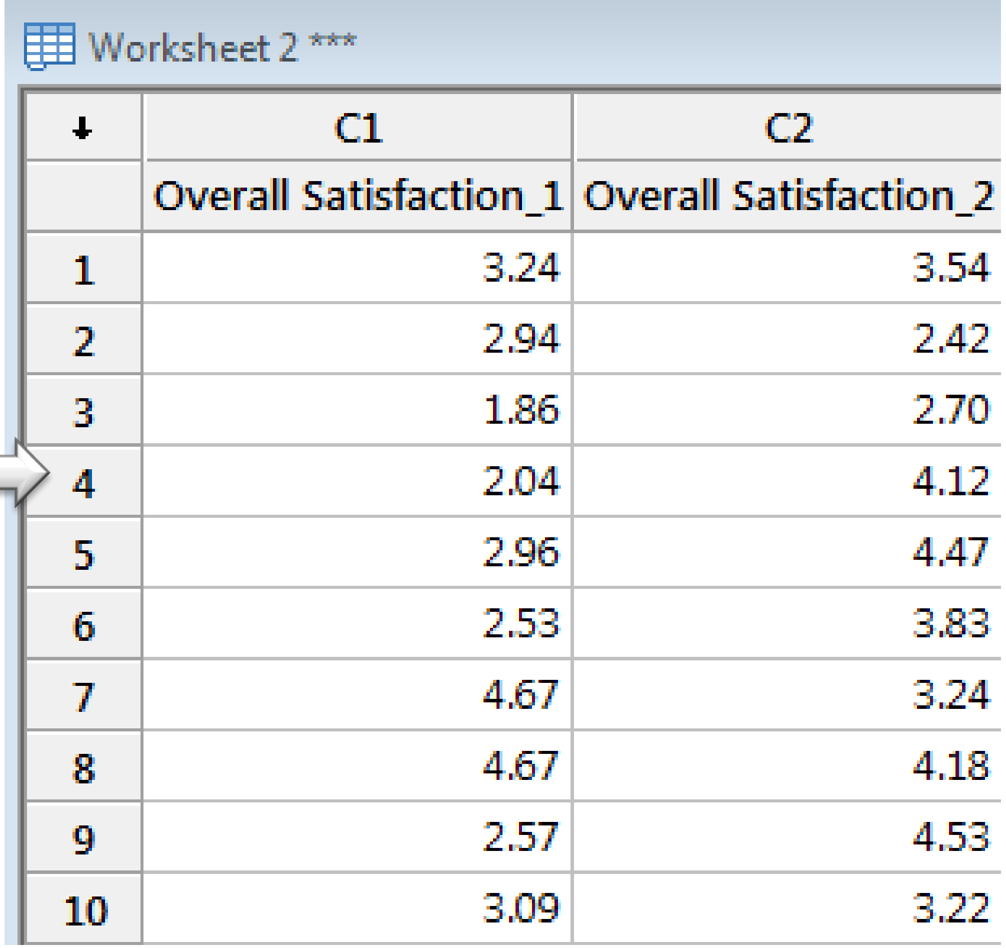
Store unstacked data:

● In new worksheet

Name: (Optional)

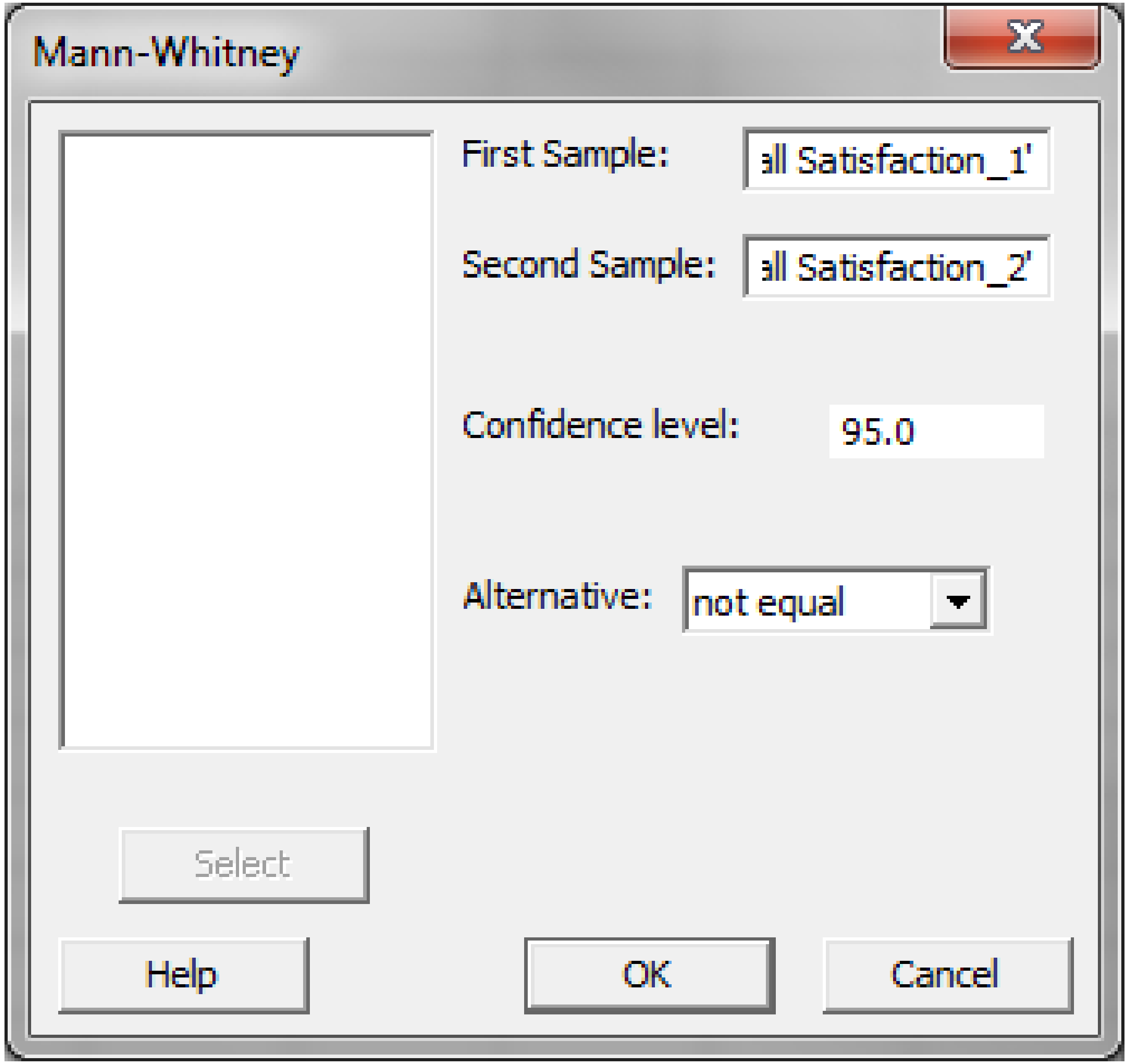
○ After last column in use

☑ Name the columns containing the unstacked data



↓	C1	C2
	Overall Satisfaction_1	Overall Satisfaction_2
1	3.24	3.54
2	2.94	2.42
3	1.86	2.70
4	2.04	4.12
5	2.96	4.47
6	2.53	3.83
7	4.67	3.24
8	4.67	4.18
9	2.57	4.53
10	3.09	3.22

Use Minitab to Run a Mann-Whitney Test



Mann-Whitney

First Sample: all Satisfaction_1'

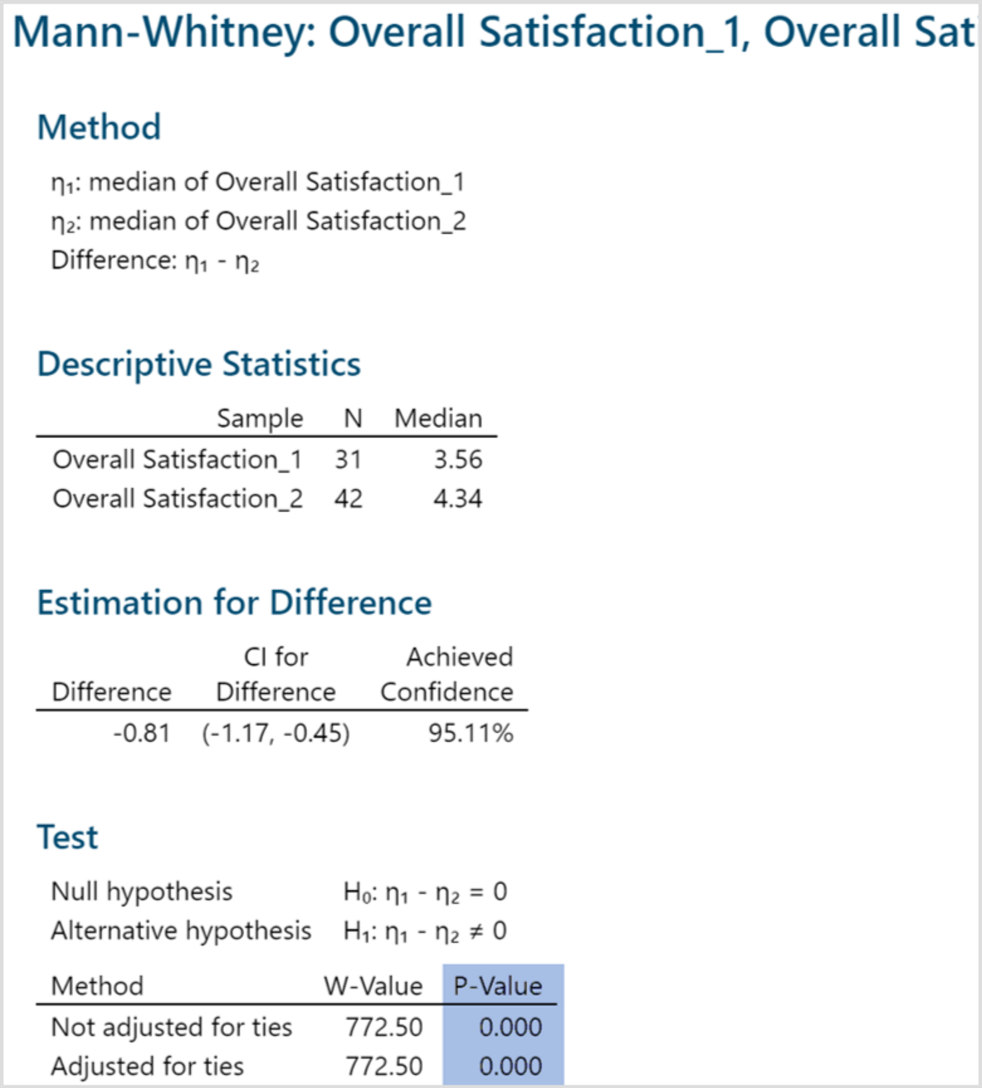
Second Sample: all Satisfaction_2'

Confidence level: 95.0

Alternative: not equal

Use Minitab to Run a Mann-Whitney Test

- The p-value of the test is lower than alpha level (0.05); so we reject the null hypothesis and conclude that there is a statistically significant difference between the overall satisfaction medians of the two customer types.



```
Mann-Whitney: Overall Satisfaction_1, Overall Sat

Method

 η₁: median of Overall Satisfaction_1
 η₂: median of Overall Satisfaction_2
 Difference: η₁ - η₂

Descriptive Statistics

                  Sample   N  Median
 Overall Satisfaction_1   31    3.56
 Overall Satisfaction_2   42    4.34

Estimation for Difference

                  CI for       Achieved
 Difference    Difference     Confidence
      -0.81  (-1.17, -0.45)       95.11%

Test

 Null hypothesis         H₀: η₁ - η₂ = 0
 Alternative hypothesis  H₁: η₁ - η₂ ≠ 0

 Method                 W-Value  P-Value
 Not adjusted for ties   772.50    0.000
 Adjusted for ties       772.50    0.000
```

3.5.2 Kruskal-Wallis

Kruskal-Wallis One-Way ANOVA

- The **Kruskal-Wallis one-way analysis of variance** is a statistical hypothesis test to compare the medians among more than two groups.
  - Null Hypothesis (H_0): $\eta_1 = \eta_2 = \ldots = \eta_k$
  - Alternative Hypothesis (H_a): at least one of the medians is different from others.

  η_i is the median of population i, and k is the number of groups of our interest.

- It is an extension of Mann-Whitney test.
- If the distributions of k populations are not identically shaped or there are outliers in the distribution, Mood's median test is a more robust than Kruskal-Wallis.

Kruskal-Wallis One-Way ANOVA: Assumptions

- The sample data drawn from the populations of interest are unbiased and representative.

- The data of k populations are continuous or ordinal when the spacing between adjacent values is not constant.

- The k populations are independent to each other.

- The Kruskal-Wallis test is robust for the non-normally distributed population.

How Kruskal-Wallis One-Way ANOVA Works

- Step 1: Group the k samples from k populations (sample i is from population i) into one single data set and then sort the data in ascending order ranked from 1 to N, where N is the total number of observations across k groups.

- Step 2: Add up the ranks for all the observations from sample i and call it $\mathbf{r}_i$, where i can be any integer between 1 and k.

How Kruskal-Wallis One-Way ANOVA Works

- Step 3: Calculate the test statistic

$$T = (N-1)\frac{\sum_{i=1}^{k} n_i(\bar{r}_i - \bar{r})^2}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(r_{ij} - \bar{r})^2}$$

where

k is the number of groups.
n_i is the sample size of sample i.
N is the total number of all the observations across k groups.
r_{ij} is the rank (among all the observations) of observation j from group i.

$$\bar{r}_i = \frac{\sum_{j=1}^{n_i} r_{ij}}{n_i} \qquad \bar{r} = \frac{1}{2}(N+1)$$

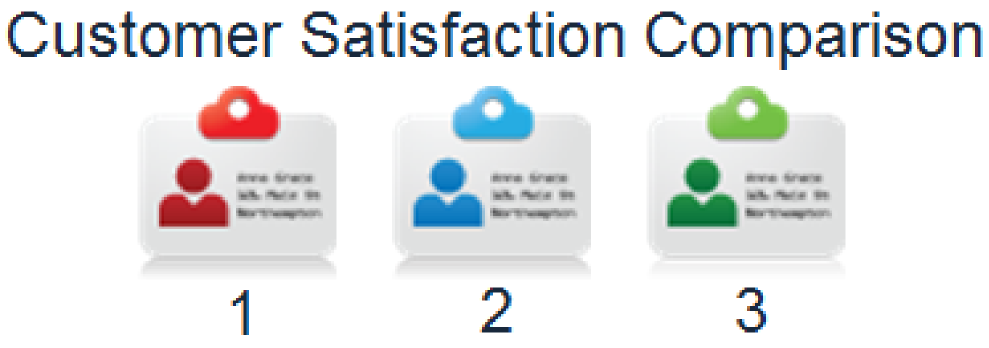
How Kruskal-Wallis One-Way ANOVA Works

- Step 4: Make a decision of whether to reject the null hypothesis.
  - Null Hypothesis (H_0): $\eta_1 = \eta_2 = \ldots = \eta_k$
  - Alternative Hypothesis (H_a): at least one of the medians is different from others.

- The test statistic follows chi-square distribution when the null hypothesis is true.
  - If T is greater than χ^2_{k-1} , we reject the null and claim there is at least one median statistically different from other medians.
  - If T is smaller than χ^2_{k-1}, we fail to reject the null and claim the medians of k groups are equal.

Use Minitab to Run a Kruskal-Wallis One-Way ANOVA

- *Case study:* We are interested in comparing customer satisfaction among three types of customers using a nonparametric (i.e., distribution-free) hypothesis test: Kruskal-Wallis one-way ANOVA.
  - Data File: “Kruskal-Wallis” tab in “Sample Data.xlsx”

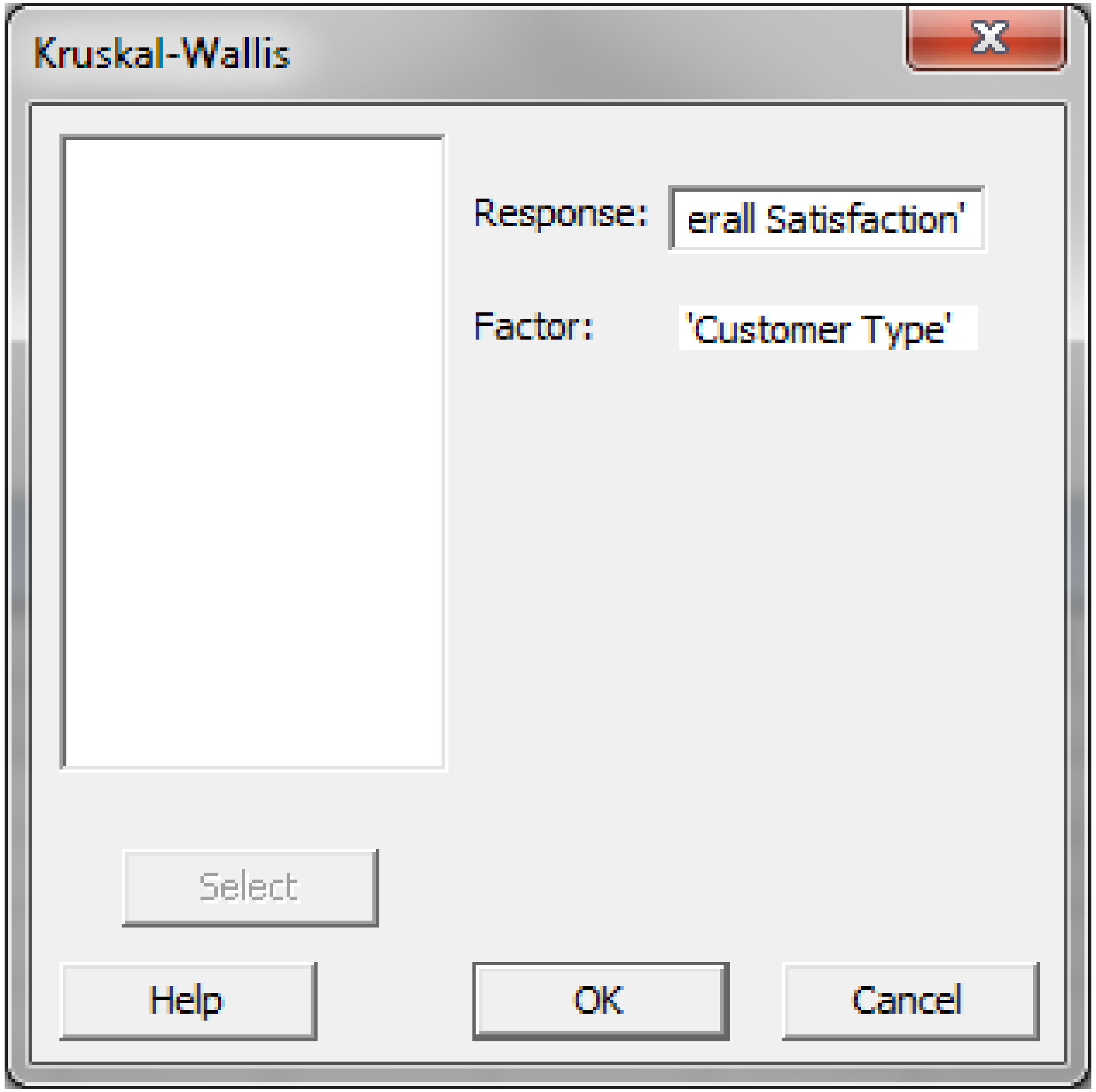


- Null Hypothesis (H_0): $\eta_1 = \eta_2 = \eta_3$
- Alternative Hypothesis (H_a): at least one of the customer types has different overall satisfaction levels from the others.

Use Minitab to Run a Kruskal-Wallis One-Way ANOVA

- Steps to run a Kruskal-Wallis One-Way ANOVA in Minitab
  1) Click Stat → Nonparametrics → Kruskal-Wallis.
  2) A new window named “Kruskal-Wallis” pops up.
  3) Select “Overall Satisfaction” as the “Response.”
  4) Select “Customer Type” as the “Factor.”
  5) Click “OK.”
  6) The Kruskal-Wallis test results appear in the session window.

Use Minitab to Run a Kruskal-Wallis One-Way ANOVA

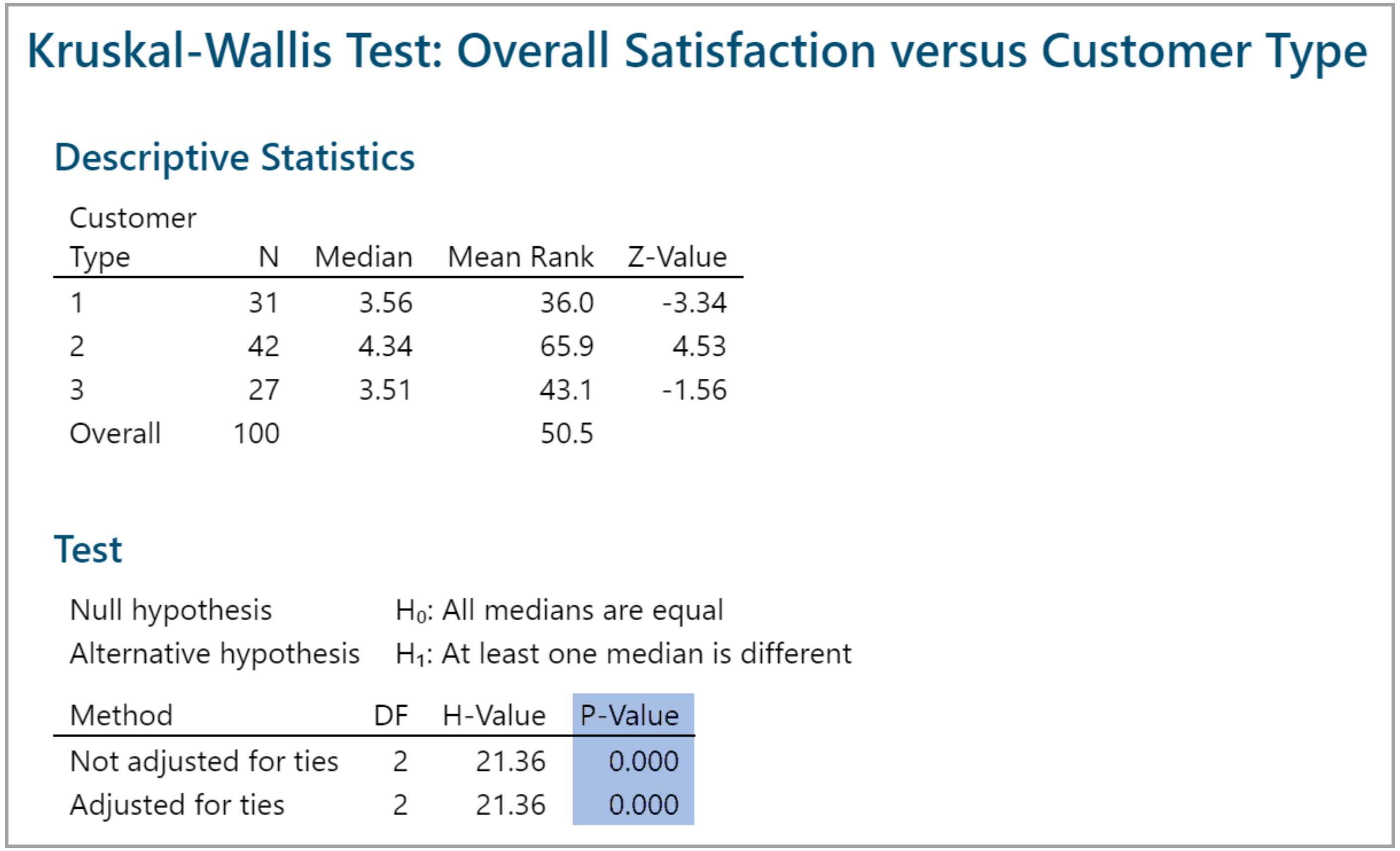


Response: erall Satisfaction'

Factor: 'Customer Type'

Use Minitab to Run a Kruskal-Wallis One-Way ANOVA

- The p-value of the test is lower than alpha level (0.05), and we reject the null hypothesis and conclude that at least the overall satisfaction median of one customer type is statistically different from the others.



Kruskal-Wallis Test: Overall Satisfaction versus Customer Type

Descriptive Statistics

Customer Type	N	Median	Mean Rank	Z-Value
1	31	3.56	36.0	-3.34
2	42	4.34	65.9	4.53
3	27	3.51	43.1	-1.56
Overall	100		50.5	

Test

Null hypothesis	H_0: All medians are equal
Alternative hypothesis	H_1: At least one median is different

Method	DF	H-Value	P-Value
Not adjusted for ties	2	21.36	0.000
Adjusted for ties	2	21.36	0.000

3.5.3 Mood’s Median

What is Mood’s Median Test?

- **Mood’s median test** is a statistical test to compare the medians of two or more populations.
  - Null Hypothesis (H_0): $\eta_1 = \ldots = \eta_k$
  - Alternative Hypothesis (H_a): at least one of the medians is different from the others.
  - k is the number of groups of our interest and is equal to or greater than two.

- Mood’s median is an alternative to Kruskal-Wallis.

- It is the extension of one sample sign test.

- For the data with outliers, Mood’s median test is more robust than the Kruskal-Wallis test.

Mood’s Median Test Assumptions

- The sample data drawn from the populations of interest are unbiased and representative.

- The data of k populations are continuous or ordinal when the spacing between adjacent values is not constant.

- The k populations are independent to each other.

- The distributions of k populations have the same shape.

- Mood’s median test is robust for non-normally distributed populations.

- Mood’s median test is robust for data with outliers.

How Mood’s Median Test Works

- Step 1: Group the k samples from k populations (sample i is from population i) into one single data set and get the median of this combined data set.

- Step 2: Separate the data in each sample into two groups. One consists of all the observations with values higher than the grand median and the other consists of all the observations with values lower than the grand median.

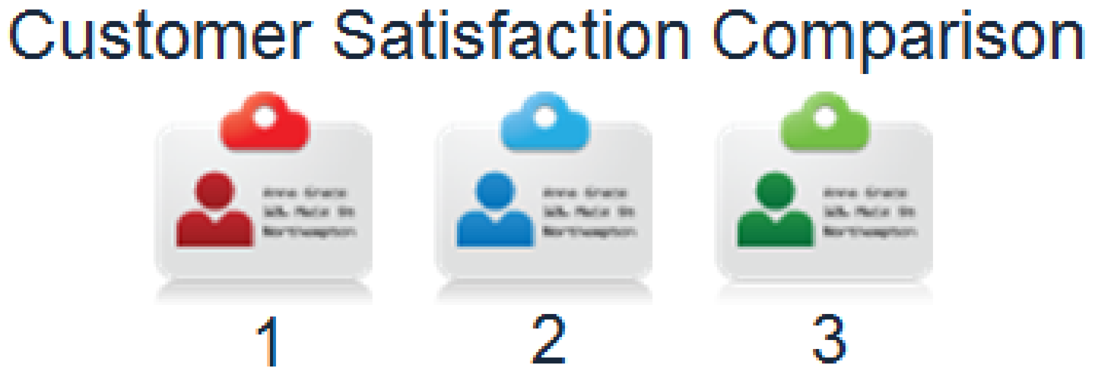
How Mood's Median Test Works

- Step 3: Run a Pearson's chi-square test to determine whether to reject the null hypothesis.
  - Null Hypothesis (H_0): $\eta_1 = \ldots = \eta_k$
  - Alternative Hypothesis (H_a): at least one of the medians is different from the others.

- If χ^2_{calc} is greater than χ^2_{crit}, we reject the null hypothesis and claim that at least one median is different from the others.
- If χ^2_{calc} is smaller than χ^2_{crit}, we fail to reject the null hypothesis and claim that the medians of k populations are not statistically different.

Use Minitab to Run a Mood’s Median Test

- *Case study:* We are interested in comparing customer satisfaction among three types of customers using a nonparametric (i.e., distribution-free) hypothesis test: Mood’s median test.
  - Data File: “Median Test” tab in “Sample Data.xlsx”

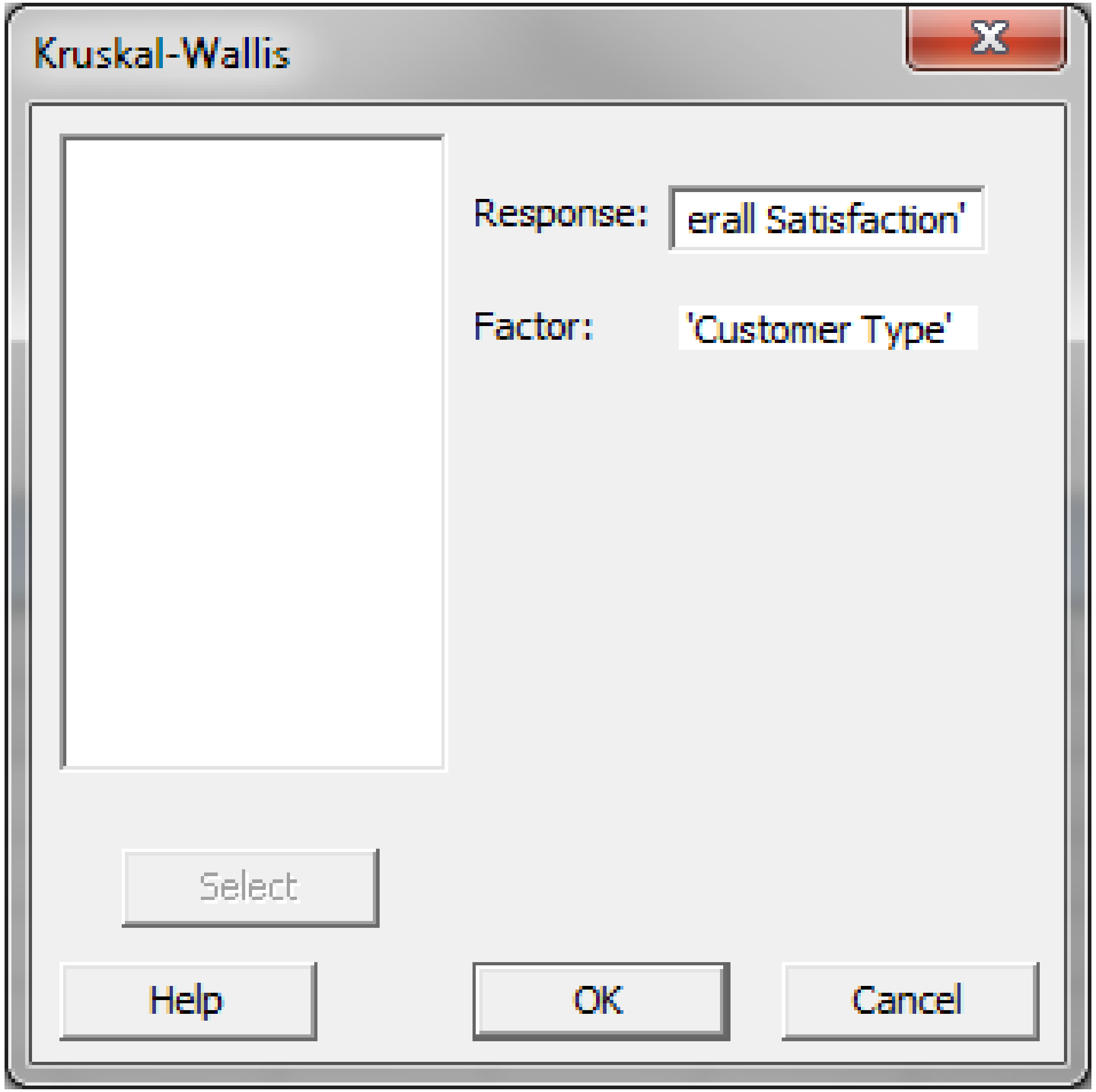


- Null Hypothesis (H_0): $\eta_1 = \eta_2 = \eta_3$
- Alternative Hypothesis (H_a): at least one customer type has different overall satisfaction from the others.

Use Minitab to Run a Mood’s Median Test

- Steps to run a Mood’s median test in Minitab
  1) Click Stat → Nonparametrics → Mood’s Median Test.
  2) A new window named “Mood’s Median Test” pops up.
  3) Select “Overall Satisfaction” as the “Response.”
  4) Select “Customer Type” as the “Factor.”
  5) Click “OK.”
  6) The Mood’s median test results appear in the session window.

Use Minitab to Run a Mood’s Median Test

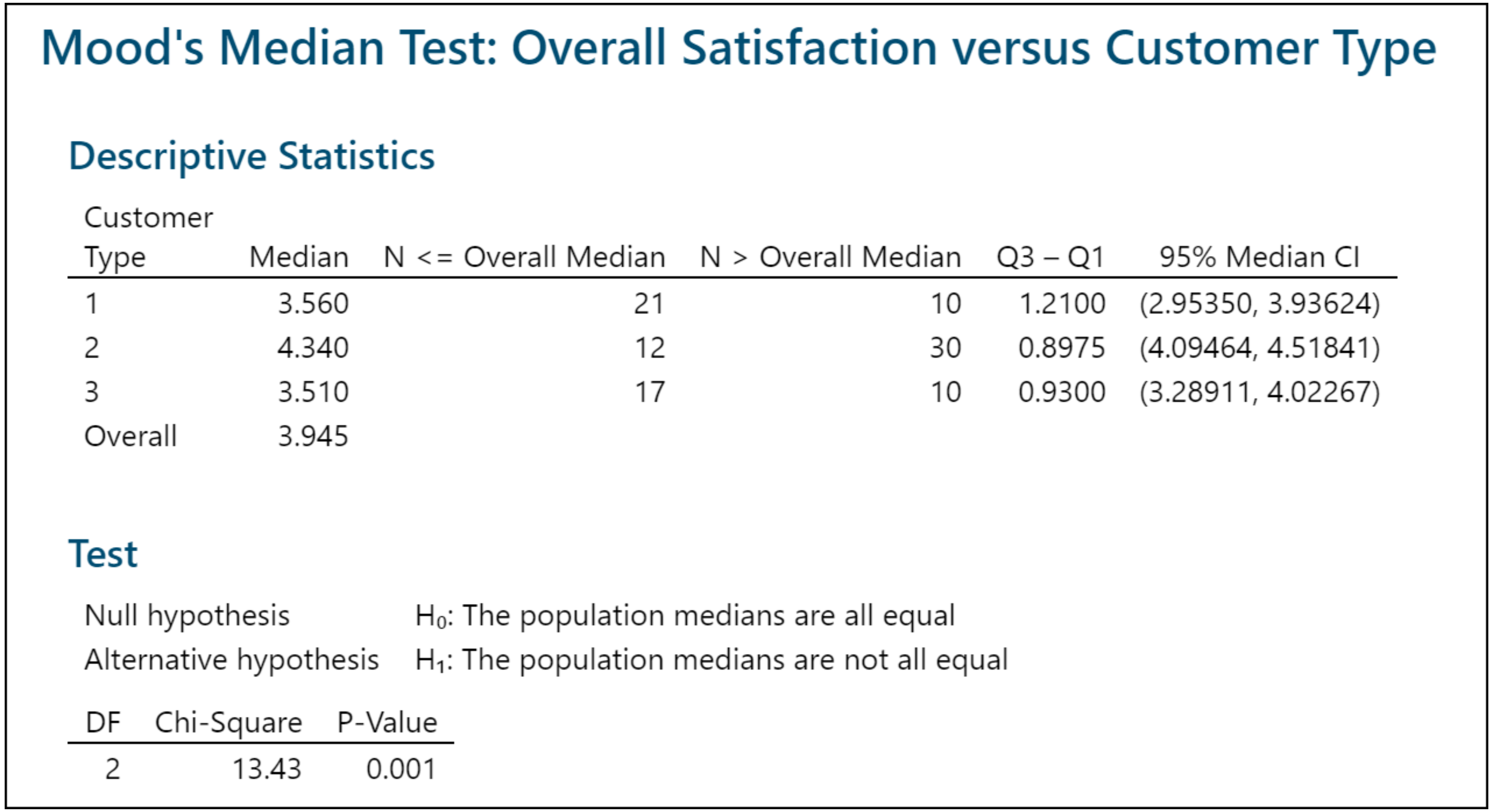

Kruskal-Wallis

Response: erall Satisfaction'

Factor: 'Customer Type'

Use Minitab to Run a Mood’s Median Test

- The p-value of the test is lower than alpha level (0.05), and we reject the null hypothesis and conclude that at least the overall satisfaction median of one customer type is statistically different from the others.



Mood's Median Test: Overall Satisfaction versus Customer Type

Descriptive Statistics

Customer Type	Median	N <= Overall Median	N > Overall Median	Q3 – Q1	95% Median CI
1	3.560	21	10	1.2100	(2.95350, 3.93624)
2	4.340	12	30	0.8975	(4.09464, 4.51841)
3	3.510	17	10	0.9300	(3.28911, 4.02267)
Overall	3.945				

Test

Null hypothesis	H_0: The population medians are all equal
Alternative hypothesis	H_1: The population medians are not all equal

DF	Chi-Square	P-Value
2	13.43	0.001

3.5.4 Friedman

What is the Friedman Test?

- The **Friedman test** is a hypothesis test used to detect the differences in the medians of various groups across multiple attempts.
    - Null Hypothesis (H_0): The treatments have identical effects (i.e., $\eta_1 = \ldots = \eta_k$).
    - Alternative Hypothesis (H_a): At least one of the treatments has different effects from the others (i.e., at least one of the medians is statistically different from others).

- It is used as an alternative of the parametric repeated measures ANOVA when the assumption of normality or variance equality is not met.

Friedman Test Assumptions

- The sample data drawn from the populations of interest are unbiased and representative.
- The data are continuous or ordinal when the spacing between adjacent values is not constant.
- Results in one block are independent of the results in another.
- The Friedman test is robust for the non-normally distributed population.
- The Friedman test is robust for populations with unequal variances.

How Friedman Test Works

- Step 1: Organize the data into a tabular view with n rows indicating the blocks and k columns indicating the treatments. Each observation x_{ij} is filled into the intersection of a specific block i and a specific treatment j.

- Step 2: Calculate the ranks of each observation within each block.

- Step 3: Replace the values in the table created in step 1 with the order r_{ij} (within a block) of the corresponding observation x_{ij}.

How Friedman Test Works

- Step 4: Calculate the test statistic

$$Q = \frac{n\sum_{j=1}^{k}(\bar{r}_j - \bar{r})^2}{\frac{1}{n(k-1)}\sum_{i=1}^{n}\sum_{j=1}^{k}(r_{ij} - \bar{r})^2}$$

where

$$\bar{r}_j = \frac{1}{n}\sum_{i=1}^{n} r_{ij}$$

$$\bar{r} = \frac{1}{nk}\sum_{i=1}^{n}\sum_{j=1}^{k} r_{ij}$$

How Friedman Test Works

- Step 5: Make a decision of whether to reject the null hypothesis.
  - When n > 15 or k > 4, the test statistic Q follows a chi-square distribution if the null hypothesis is true and the p-value is $P\left(\chi^2_{k-1} \geq Q\right)$
  - When n < 15 and k < 4, the test statistic Q does not approximate a chi-square distribution and the p-value can be obtained from tables of Q for the Friedman test.
  - If p-value > alpha level (0.05), we fail to reject the null.

Friedman Test Examples

- A number of n water testers judge the quality of k different water samples, each of which is from a distinct water source. We will apply a Friedman Test to determine whether the water qualities of the k sources are the same.

- There are n blocks and k treatments in this experiment.

- One tester's decision would not have any influence on other testers.

- When running the experiment, each tester judges the water in a random sequence.

Use Minitab to Run a Friedman Test

- *Case study:* We are interested in comparing the effect of a drug treatment on enzyme activity. Three different drug therapies were given to four animals and each animal belongs to a different litter.
  - Data File: “Friedman.MTW,”
    (Note: this is a different data source than what you have been using to this point)
- Null Hypothesis (H_0): $\eta_1 = \eta_2 = \eta_3$
- Alternative Hypothesis (H_a): at least one treatment effect is statistically different from the others.

Use Minitab to Run a Friedman Test

- Steps to run a Friedman test in Minitab
  1) Click Stat → Nonparametrics → Friedman.
  2) A new window named “Friedman” pops up.
  3) Select “EnzymeActivity” as the “Response.”
  4) Select “Therapy” as the “Treatment.”
  5) Select “Litter” as the “Blocks.”
  6) Click “OK.”
  7) The Friedman test results appear in the session window.

Use Minitab to Run a Friedman Test

Response: EnzymeActivity

Treatment: Therapy

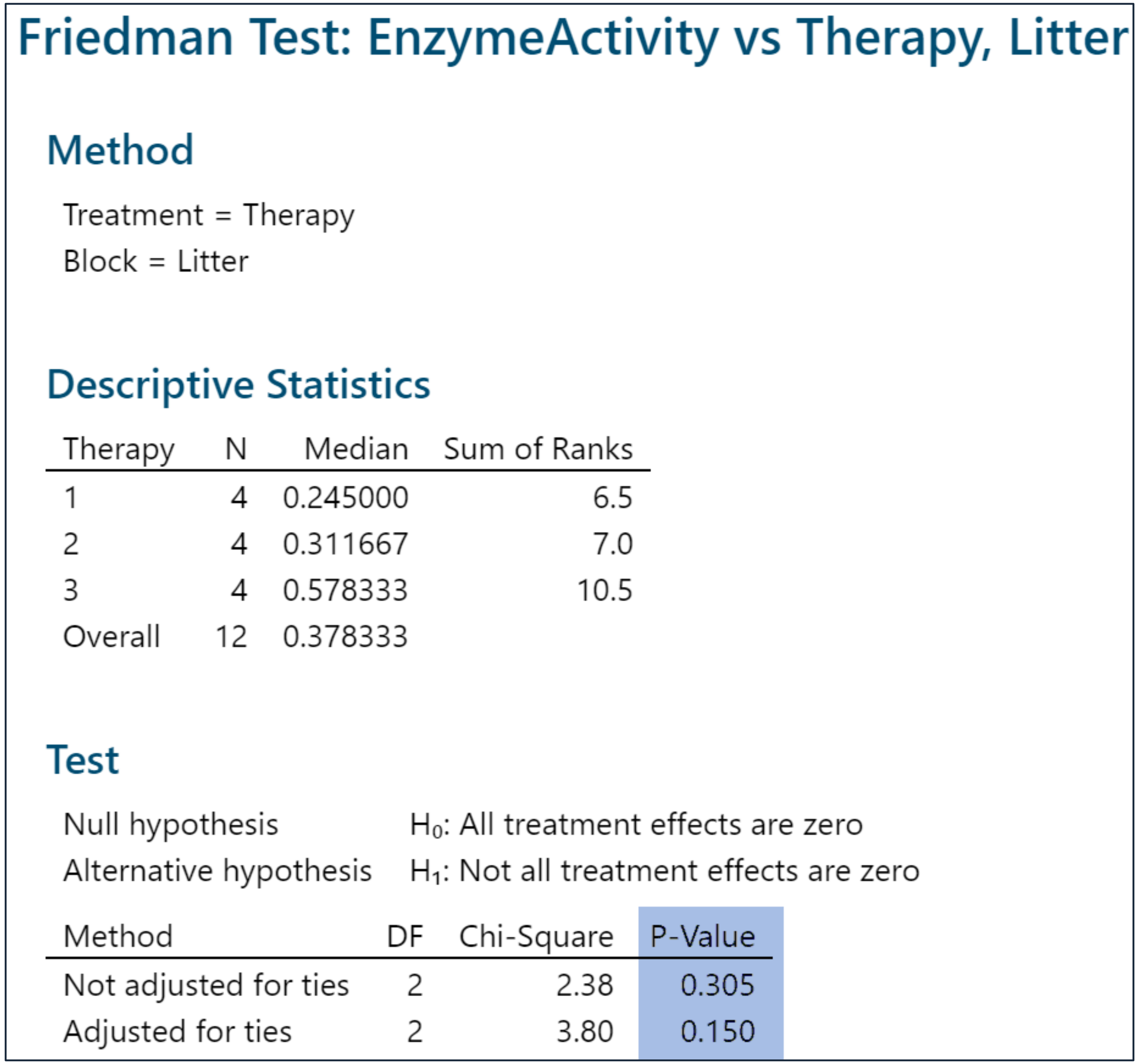
Blocks: Litter

☐ Store residuals

☐ Store fits

Use Minitab to Run a Friedman Test

- The p-value of the Friedman test is greater than alpha level (0.05). We fail to reject the null hypothesis and conclude that there is not any statistically significant difference between the treatment effects.



```
Friedman Test: EnzymeActivity vs Therapy, Litter

Method

  Treatment = Therapy
  Block = Litter

Descriptive Statistics

  Therapy   N    Median   Sum of Ranks
  1         4  0.245000            6.5
  2         4  0.311667            7.0
  3         4  0.578333           10.5
  Overall  12  0.378333

Test

  Null hypothesis          H₀: All treatment effects are zero
  Alternative hypothesis   H₁: Not all treatment effects are zero

  Method                  DF  Chi-Square  P-Value
  Not adjusted for ties    2        2.38    0.305
  Adjusted for ties        2        3.80    0.150
```

3.5.5 One Sample Sign

What is the One Sample Sign Test?

- The **one sample sign test** is a hypothesis test to compare the median of a population with a specified value.
  - Null Hypothesis (H_0): $\eta = \eta_0$
  - Alternative Hypothesis (H_a): $\eta \neq \eta_0$

- It is an alternative test of one sample t-test when the distribution of the data is non-normal. It is robust for the data with non-symmetric distribution.

One Sample Sign Test Assumptions

- The sample data drawn from the population of interest are unbiased and representative.
- The data are continuous or ordinal when the spacing between adjacent values is not constant.
- The one sample sign test is robust for the non-normally distributed population.
- The one sample sign test does not have any assumptions on the distribution. It is a distribution-free test.

How the One Sample Sign Test Works

- Step 1: Separate the sample set of data into two groups: one with values greater than and the other with values less than the hypothesized median η_0. Count the number of observations in each group.

- Step 2: Calculate the test statistic.
  - If the null hypothesis is true, the number of observations in each group should not be significantly different from half of the total sample size.
  - The test statistic follows a binomial distribution when the null is true. When n is large, we use the normal distribution to approximate the binomial distribution.

How the One Sample Sign Test Works

Test Statistic:

$$Z_{calc} = \frac{n_+ - np}{\sqrt{np(1-p)}}$$

where

n_+ is the number of observations with values greater than the hypothesized median.

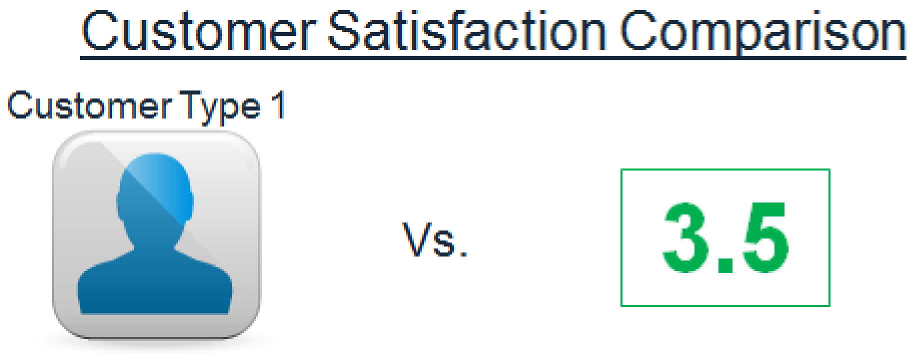
n is the total number of observations .

p is 0.5.

- Step 3: Make a decision on whether to reject the null hypothesis. If the $|Z_{calc}|$ is smaller than Z_{crit}, we fail to reject the null hypothesis and claim that there is no significant difference between the population median and the hypothesized median.

Use Minitab to Run a One Sample Sign Test

- *Case study:* We are interested in comparing the overall satisfaction of customer type 1 against a specified benchmark satisfaction (3.5) using a nonparametric (i.e., distribution-free) hypothesis test: one sample sign test.
  - Data File: “One Sample Wilcoxon” tab in “Sample Data.xlsx”



Customer Satisfaction Comparison

Customer Type 1

Vs.

3.5

- Null Hypothesis (H_0): $\eta_1 = 3.5$
- Alternative Hypothesis (H_a): $\eta_1 \neq 3.5$

Use Minitab to Run a One Sample Sign Test

- Steps to run a one sample sign test in Minitab
  1) Click Stat → Nonparametrics → 1-Sample Sign.
  2) A new window named “1-Sample Sign” pops up.
  3) Select “Overall Satisfaction” as the “Variables.”
  4) Click the box of “Test median.”
  5) Enter “3.5” in the box next to “Test median.”
  6) Click “OK.”
  7) The one sample sign test results appear in session window.

Use Minitab to Run a One Sample Sign Test

1-Sample Sign

Variables: 'Overall Satisfaction'

○ Confidence interval

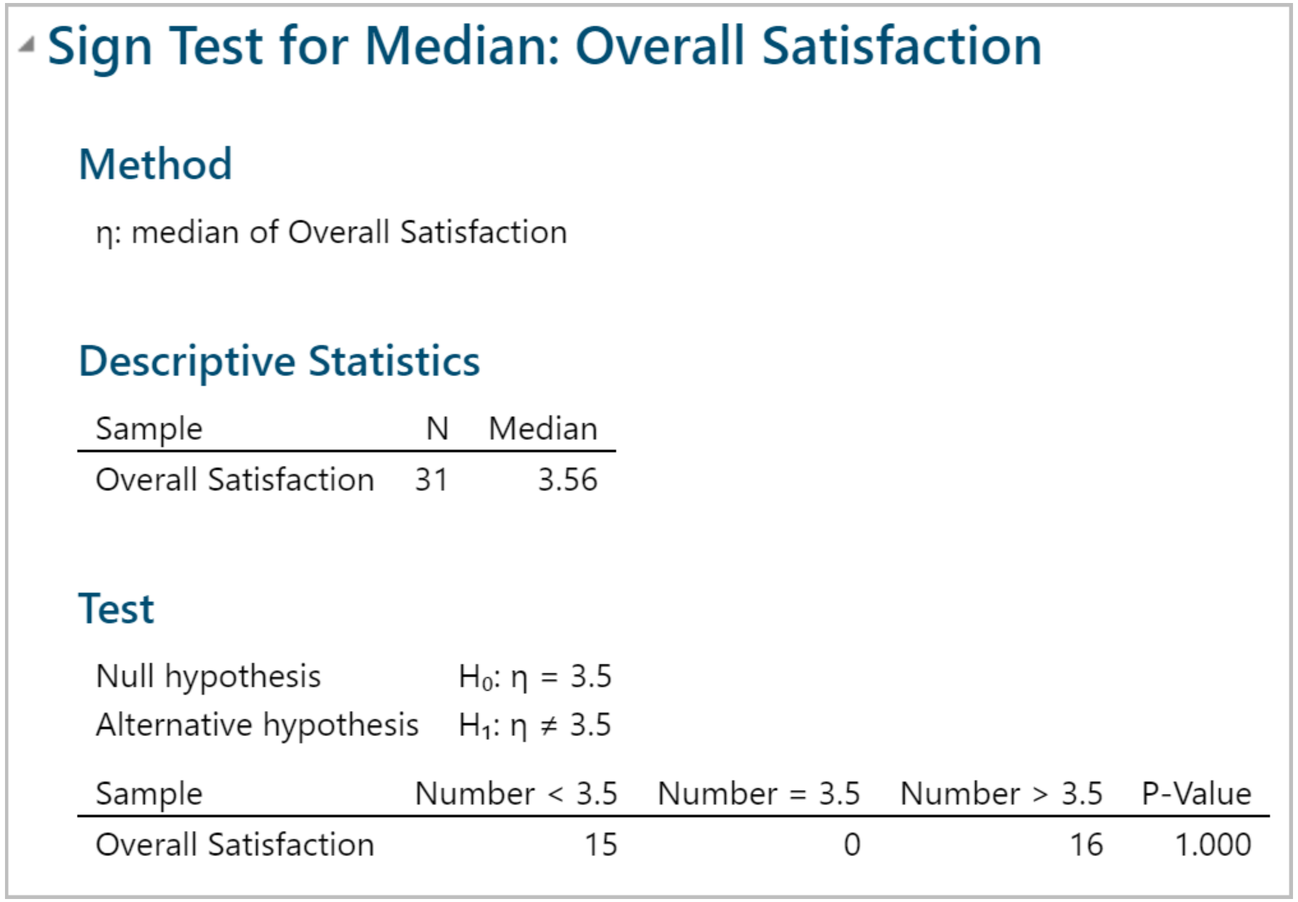
Level: 95.0

● Test median: 3.5

Alternative: not equal

Use Minitab to Run a One Sample Sign Test

- The p-value of the one sample sign test is 1.0000, higher than the alpha level (0.05), and we fail to reject the null hypothesis. There is not any statistically significant difference between the overall satisfaction of customer type 1 and the benchmark satisfaction level.



```
Sign Test for Median: Overall Satisfaction

Method

η: median of Overall Satisfaction

Descriptive Statistics

Sample                 N   Median
Overall Satisfaction  31     3.56

Test

Null hypothesis         H₀: η = 3.5
Alternative hypothesis  H₁: η ≠ 3.5

Sample                Number < 3.5  Number = 3.5  Number > 3.5  P-Value
Overall Satisfaction            15             0            16    1.000
```

3.5.6 One Sample Wilcoxon

What is the One Sample Wilcoxon Test?

- The **one sample Wilcoxon test** is a hypothesis test to compare the median of one population with a specified value.
  - Null Hypothesis (H_0): $\eta = \eta_0$
  - Alternative Hypothesis (H_a): $\eta \neq \eta_0$
- It is an alternative test of one sample t-test when the distribution of the data is non-normal.
- It is more powerful than one sample sign test but it assumes the distribution of the data is symmetric.

One Sample Wilcoxon Test Assumptions

- The sample data drawn from the population of interest are unbiased and representative.
- The data are continuous or ordinal when the spacing between adjacent values is not constant.
- The distribution of the data is symmetric about a median.
- The one sample Wilcoxon test is robust for the non-normally distributed population.

How the One Sample Wilcoxon Test Works

- Step 1: Create the following columns one by one:
  - Column 1: all the raw observations (X)
  - Column 2: the differences between each observation value and the hypothesized median ($X - \eta_0$)
  - Column 3: the signs (+ or –) of column 2
  - Column 4: the absolute value of column 2
  - Column 5: the ranks of each item in column 4 in ascending order
  - Column 6: the product of column 3 and column 5

How the One Sample Wilcoxon Test Works

- Step 2: Calculate the test statistic W_{calc} , which is the sum of all the non-negative values in column 6.

- Step 3: Make a decision on whether to reject the null hypothesis. Use the table of critical values for the Wilcoxon test to get the W_{crit} with predetermined alpha level and number of observations.

  - If the W_{calc} is smaller than the W_{crit}, we fail to reject the null hypothesis and claim that there is no significant difference between the population median and the hypothesized median.

Use Minitab to Run a One Sample Wilcoxon Test

- *Case study:* We are interested in comparing the overall satisfaction of customer type 1 against a specified benchmark satisfaction (3.5) using a nonparametric (i.e., distribution-free) hypothesis test: one sample Wilcoxon test.
  - Data File: “One Sample Wilcoxon” tab in “Sample Data.xlsx”

Customer Satisfaction Comparison

Customer Type 1

Vs.

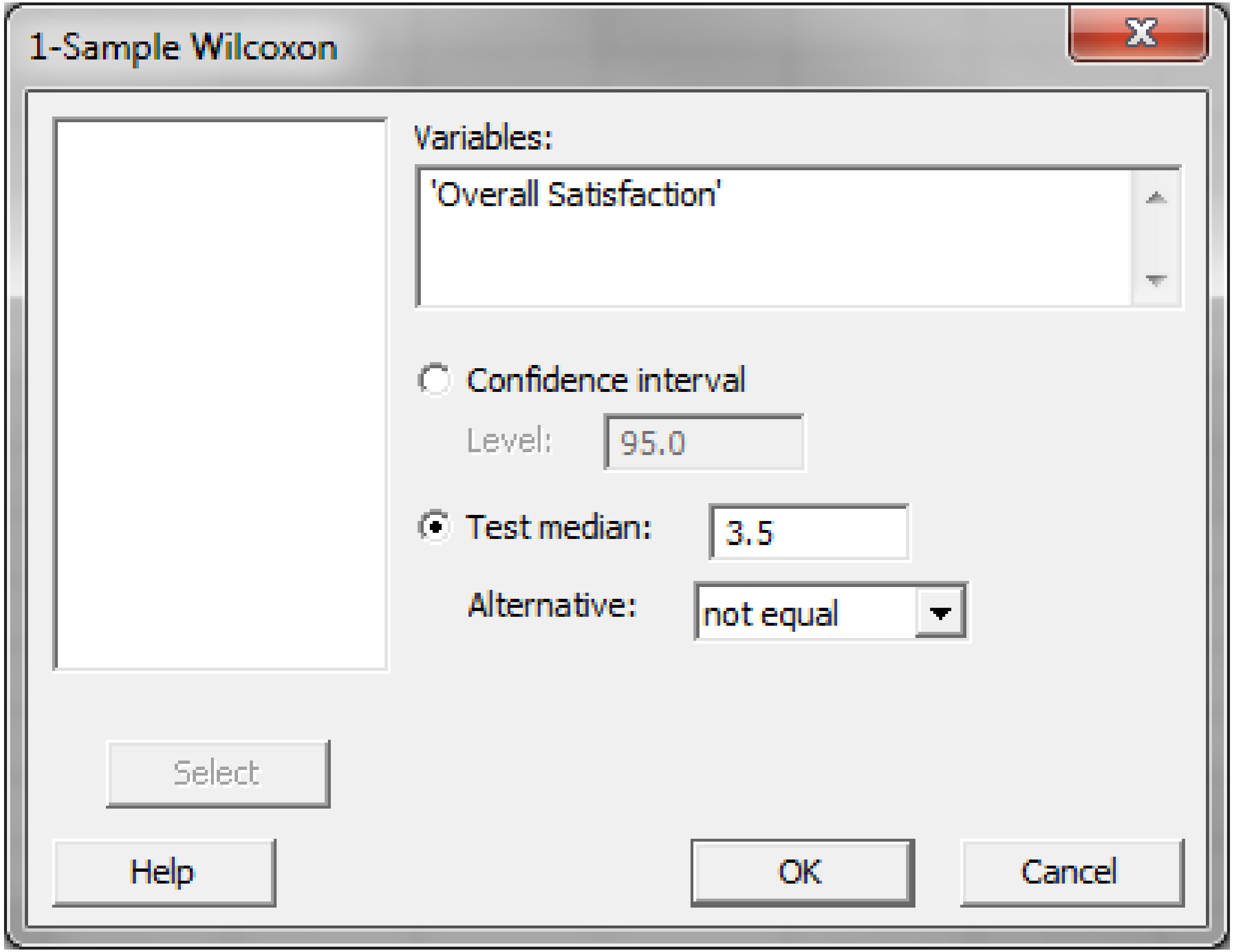
3.5

- Null Hypothesis (H_0): $\eta_1 = 3.5$
- Alternative Hypothesis (H_a): $\eta_1 \neq 3.5$

Use Minitab to Run a One Sample Wilcoxon Test

- Steps to run a one sample Wilcoxon test in Minitab
  1) Click Stat → Nonparametrics → 1-Sample Wilcoxon.
  2) A new window named “1-Sample Wilcoxon” pops up.
  3) Select “Overall Satisfaction” as the “Variables.”
  4) Click the radio button “Test median.”
  5) Enter “3.5” in the box next to “Test median.”
  6) Click “OK.”
  7) The one sample Wilcoxon test results appear in the session window.

Use Minitab to Run a One Sample Wilcoxon Test



1-Sample Wilcoxon

Variables: 'Overall Satisfaction'

○ Confidence interval

Level: 95.0

● Test median: 3.5

Alternative: not equal

Use Minitab to Run a One Sample Wilcoxon Test

- The p-value of the one sample Wilcoxon test is 0.557, higher than the alpha level (0.05), and we fail to reject the null hypothesis. There is not any statistically significant difference between the overall satisfaction of customer type 1 and the benchmark satisfaction level.

Wilcoxon Signed Rank Test: Overall Satisfaction

Method

η: median of Overall Satisfaction

Descriptive Statistics

Sample	N	Median
Overall Satisfaction	31	3.4075

Test

Null hypothesis H_0: $\eta = 3.5$
Alternative hypothesis H_1: $\eta \neq 3.5$

Sample	N for Test	Wilcoxon Statistic	P-Value
Overall Satisfaction	31	217.50	0.557

3.5.7 One & Two Sample Proportion

What is the One Sample Proportion Test?

- **One sample proportion test** is a hypothesis test to compare the proportion of one certain outcome occurring in a population following the binomial distribution with a specified proportion.
  - Null Hypothesis (H_0): $p = p_0$
  - Alternative Hypothesis (H_a): $p \neq p_0$

One Sample Proportion Test Assumptions

- The sample data drawn from the population of interest are unbiased and representative.

- There are only two possible outcomes in each trial: success/failure, yes/no, and defective/non-defective etc.

- The underlying distribution of the population is binomial distribution.

- When $np \geq 5$ and $np(1 – p) \geq 5$, the binomial distribution can be approximated by the normal distribution.

How the One Sample Proportion Test Works

- When np ≥ 5 and np(1 – p) ≥ 5, we use normal distribution to approximate the underlying binomial distribution of the population.

Test Statistic:

$$Z_{calc} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

where

$\hat{p}$ is the observed probability of one certain outcome occurring.

p_0 is the hypothesized probability.

n is number of trials.

When $|Z_{calc}|$ is smaller than Z_{crit}, we fail to reject the null hypothesis and claim that there is no statistically significant difference between the population proportion and the hypothesized proportion.

Use Minitab to Run a One Sample Proportion Test

- *Case study:* We are interested in comparing the exam pass rate of a high school this month against a specified rate (70%) using a nonparametric (i.e., distribution-free) hypothesis test: one sample proportion test.
  - Data File: “One Sample Proportion” tab in “Sample Data.xlsx”

Exam Pass Rate

Vs.

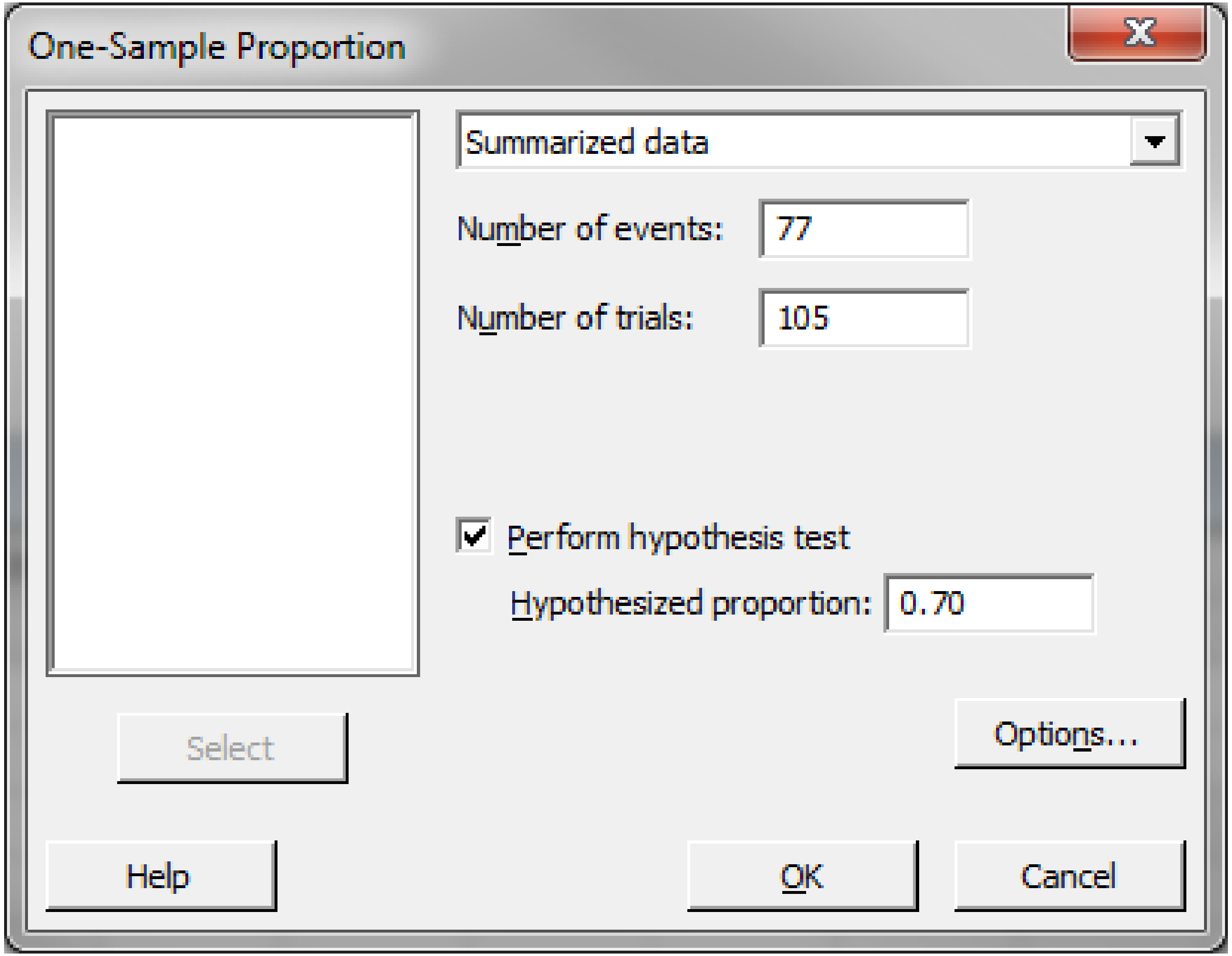
70%

- Null Hypothesis (H_0): $p = 70\%$
- Alternative Hypothesis (H_a): $p \neq 70\%$

Use Minitab to Run a One Sample Proportion Test

- Steps to run a one sample proportion test in Minitab.
    1) Click Stat → Basic Statistics → 1 Proportion.
    2) A new window named “1 Proportion (Test and Confidence Interval)” appears.
    3) Click on the drop down menu and select “Summarized data.”
    4) Enter “77” in the box of “Number of events.”
    5) Enter “105” in the box of “Number of trials.”
    6) Check the box “Perform hypothesis test.”
    7) Enter “0.70” as the “Hypothesized proportion.”
    8) Click “OK.”
    9) The one sample proportion test results appear in the session window.

Use Minitab to Run a One Sample Proportion Test



Summarized data

Number of events: 77

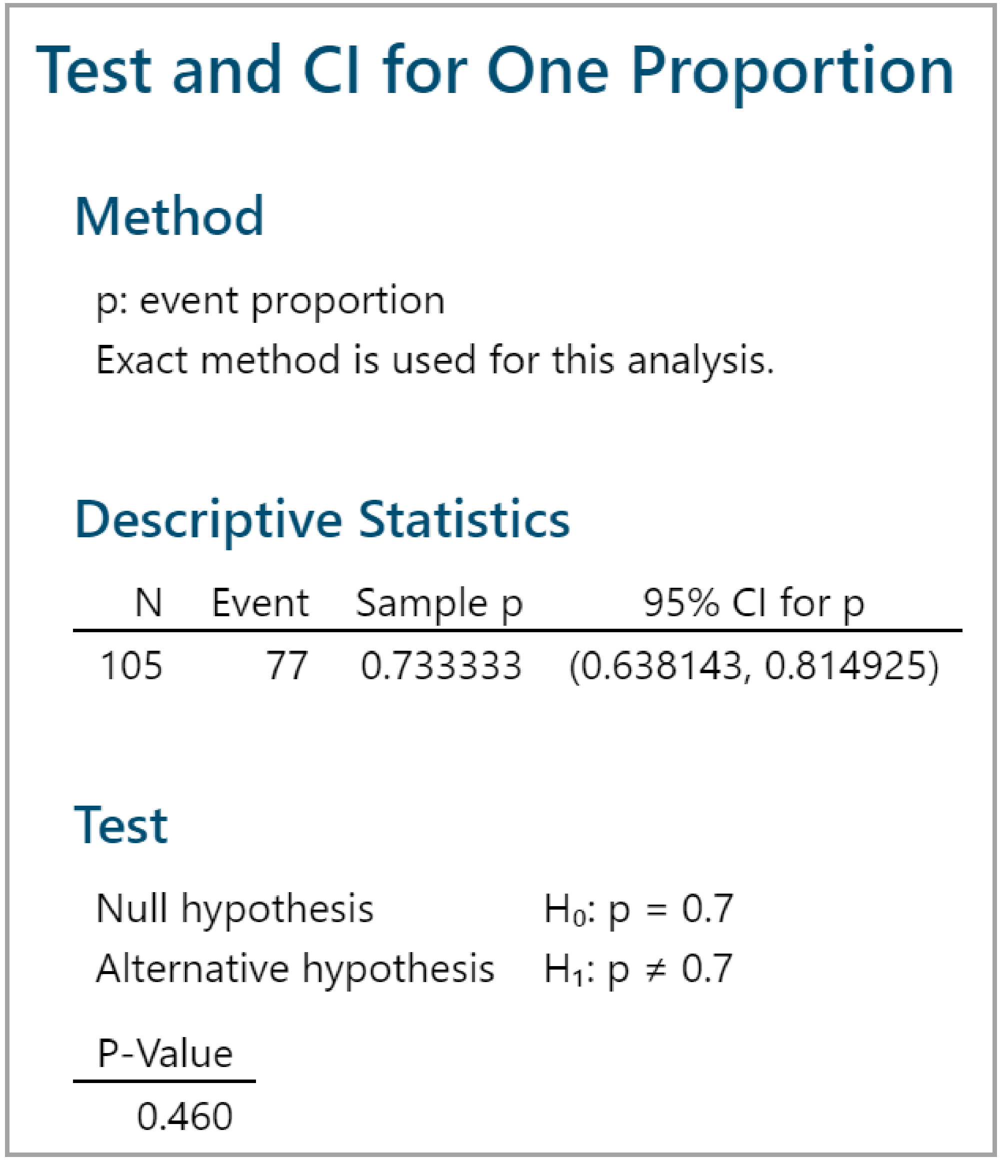
Number of trials: 105

☑ Perform hypothesis test

Hypothesized proportion: 0.70

Use Minitab to Run a One Sample Proportion Test

- The p-value of the one sample proportion test is 0.460, greater than the alpha level (0.05), and we fail to reject the null hypothesis.
- We conclude that the exam pass rate of the high school this month is not statistically different from 70%.



```
Test and CI for One Proportion

Method

  p: event proportion
  Exact method is used for this analysis.

Descriptive Statistics

    N   Event   Sample p        95% CI for p
  105      77   0.733333   (0.638143, 0.814925)

Test

  Null hypothesis          H₀: p = 0.7
  Alternative hypothesis   H₁: p ≠ 0.7

  P-Value
    0.460
```

What is the Two Sample Proportion Test?

- The **two sample proportion test** is a hypothesis test to compare the proportions of one certain event occurring in two populations following the binomial distribution.
  - Null Hypothesis (H_0): $p_1 = p_2$
  - Alternative Hypothesis (H_a): $p_1 \neq p_2$

Two Sample Proportion Test Assumptions

- The sample data drawn from the populations of interest are unbiased and representative.
- There are only two possible outcomes in each trial for both populations: success/failure, yes/no, and defective/non-defective etc.
- The underlying distributions of both populations are binomial distribution.
- When $np \geq 5$ and $np(1 - p) \geq 5$, the binomial distribution can be approximated by the normal distribution.

How the Two Sample Proportion Test Works

- When np ≥ 5 and np(1 – p) ≥ 5, we use normal distribution to approximate the underlying binomial distributions of the populations.

Test Statistic:

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_0\left(1-\hat{p}_0\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

where $\hat{p}_0 = \frac{x_1 + x_2}{n_1 + n_2}$

$\hat{p}_1$ and $\hat{p}_2$ are the observed proportions of events in the two samples.

n_1 and n_2 are the number of trials in the two samples respectively.

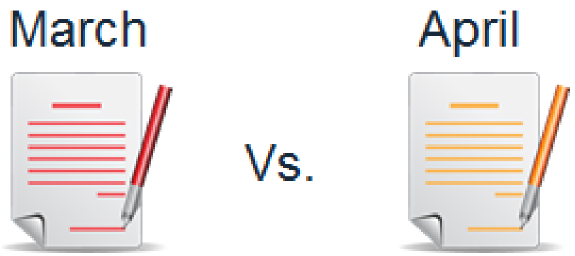
x_1 and x_2 are the number of events in the two samples respectively.

When $|Z_{calc}|$ is smaller than Z_{crit}, we fail to reject the null hypothesis.

Use Minitab to Run a Two Sample Proportion Test

- *Case study:* We are interested in comparing the exam pass rates of a high school in March and April using a nonparametric (i.e., distribution-free) hypothesis test: two sample proportion test.
  - Data File: “Two Sample Proportion” tab in “Sample Data.xlsx”

Exam Pass Rate Comparison

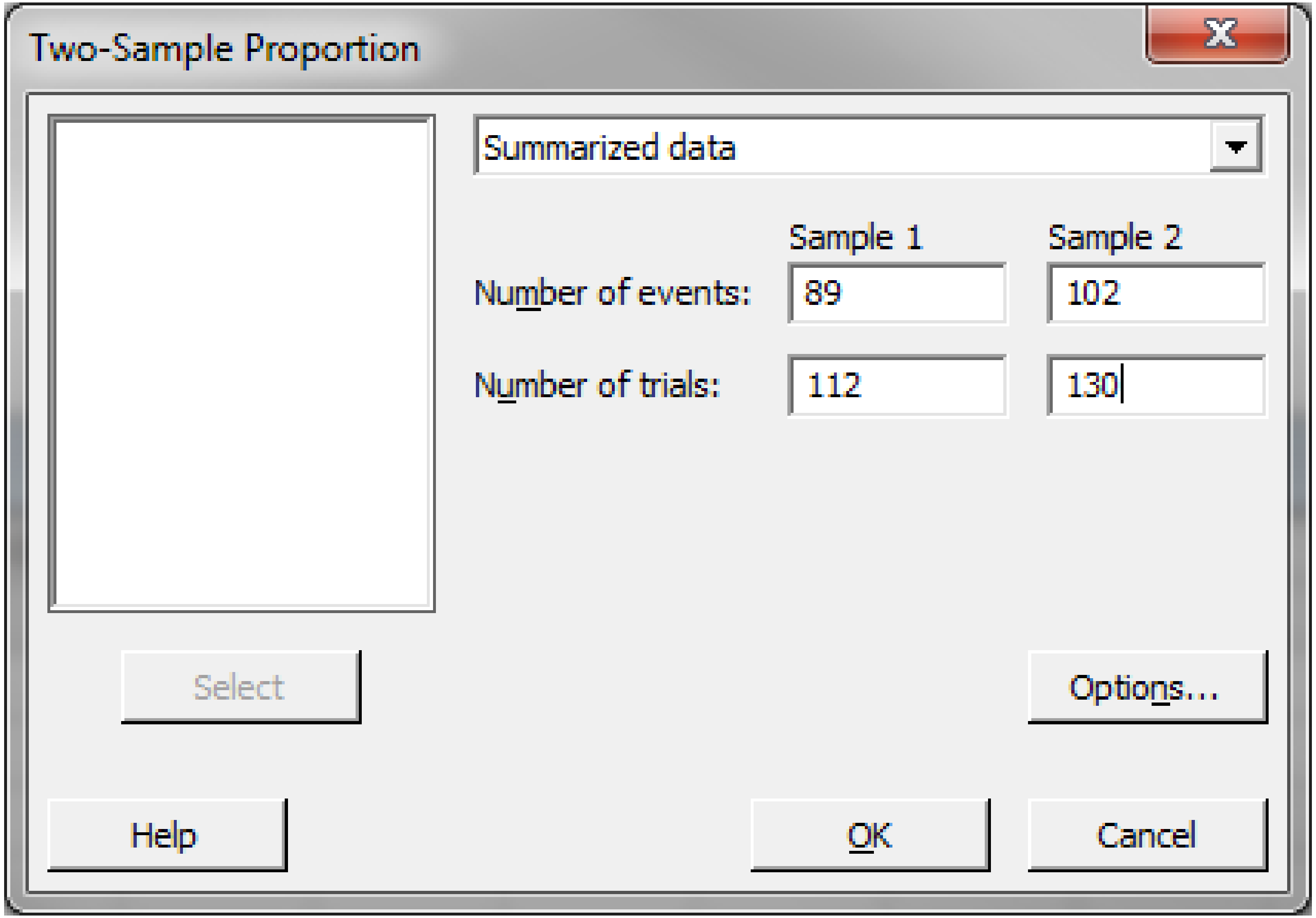


- Null Hypothesis (H_0): $p_{March} = p_{April}$
- Alternative Hypothesis (H_a): $p_{March} \neq p_{April}$

Use Minitab to Run a Two Sample Proportion Test

- Steps to run a two sample proportion test in Minitab
  1) Click Stat → Basic Statistics → 2 Proportions.
  2) A new window named “2 Proportions (Test and Confidence Interval)” pops up.
  3) Choose “Summarized data” from the dropdown menu
  4) Enter “89” in the box intersecting “First” and “Events.”
  5) Enter “112” in the box intersecting “First” and “Trials.”
  6) Enter “102” in the box intersecting “Second” and “Events.”
  7) Enter “130” in the box intersecting “Second” and “Trials.”
  8) Click “OK.”
  9) The two sample proportion test results appear in the session window.

Use Minitab to Run a Two Sample Proportion Test



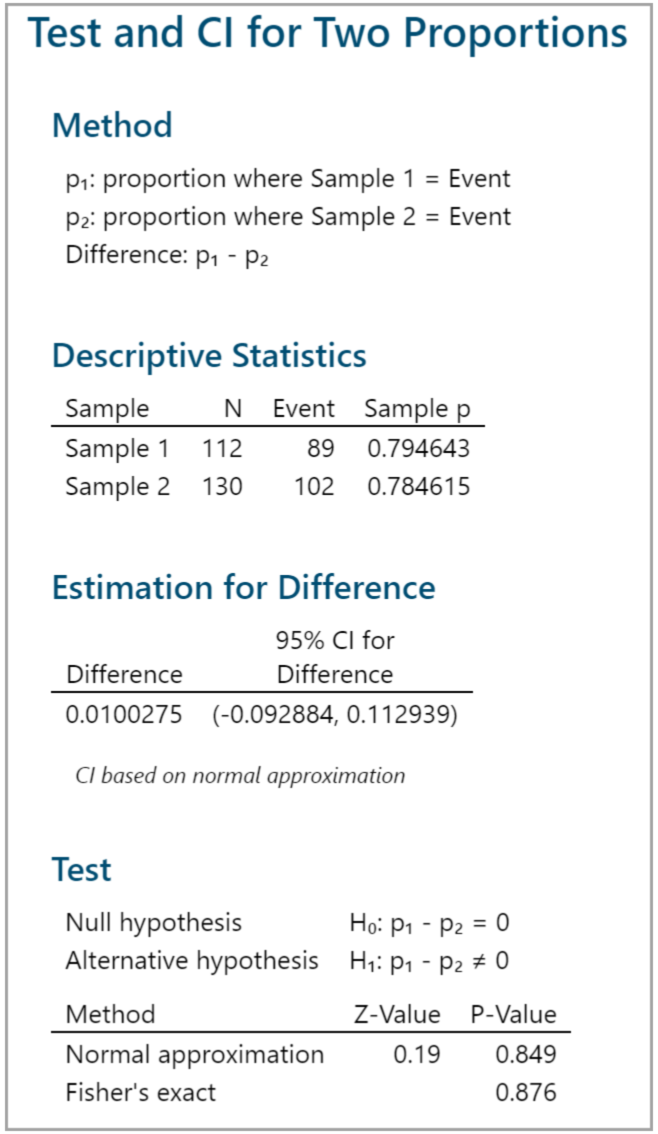
Two-Sample Proportion

Summarized data

	Sample 1	Sample 2
Number of events:	89	102
Number of trials:	112	130

Use Minitab to Run a Two Sample Proportion Test

- The p-value of the two sample proportion test is 0.849, greater than the alpha level (0.05), and we fail to reject the null hypothesis.
- We conclude that the exam pass rates of the high school in March and April are not statistically different.



```
Test and CI for Two Proportions

Method

p₁: proportion where Sample 1 = Event
p₂: proportion where Sample 2 = Event
Difference: p₁ - p₂

Descriptive Statistics

Sample      N   Event  Sample p
Sample 1  112      89  0.794643
Sample 2  130     102  0.784615

Estimation for Difference

                95% CI for
Difference      Difference
0.0100275   (-0.092884, 0.112939)

 CI based on normal approximation

Test

Null hypothesis          H₀: p₁ - p₂ = 0
Alternative hypothesis   H₁: p₁ - p₂ ≠ 0

Method                   Z-Value  P-Value
Normal approximation        0.19    0.849
Fisher's exact                      0.876
```

3.5.8 Chi-Squared (Contingency Tables)

What is the Chi-Square Test?

- **A chi-square test** is a hypothesis test in which the sampling distribution of the test statistic follows a chi-square distribution when the null hypothesis is true.

- There are multiple chi-square tests available and in this module we will cover the Pearson's chi-square test used in contingency analysis.
  - Null Hypothesis (H_0): $p_1 = p_2 = \ldots = p_k$
  - Alternative Hypothesis (H_a): at least on of the proportions is different from others.
  - k is the number of populations of our interest. $k \geq 2$.

What is the Chi-Square Test?

- The chi-square test can also be used to test whether two factors are independent of each other. In other words, it can be used to test whether there is any statistically significant relationship between two discrete factors.
  - Null Hypothesis (H_0): Factor 1 is independent of factor 2.
  - Alternative Hypothesis (H_a): Factor 1 is not independent of factor 2.

Chi-Square Test Assumptions

- The sample data drawn from the populations of interest are unbiased and representative.
- There are only two possible outcomes in each trial for an individual population: success/failure, yes/no, and defective/non-defective etc.
- The underlying distribution of each population is binomial distribution.
- When $np \geq 5$ and $np(1 – p) \geq 5$, the binomial distribution can be approximated by the normal distribution.

How Chi-Square Test Works

- Test Statistic

$$\chi^2_{calc} = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

where

O_i is an observed frequency.

E_i is an expected frequency.

N is the number of cells in the contingency table.

If χ^2_{calc} is smaller than χ^2_{crit} , we fail to reject the null hypothesis.

Use Minitab to Run a Chi-Square Test

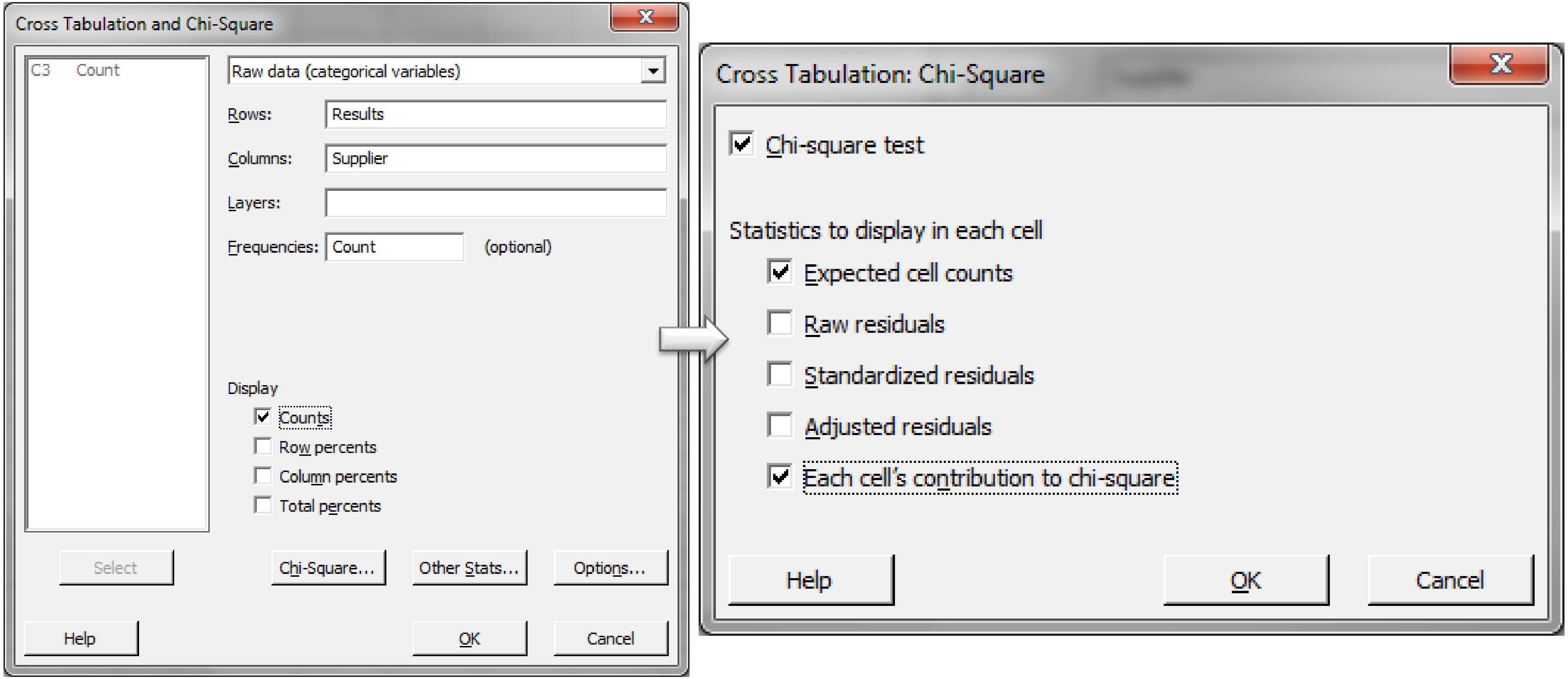
- *Case study 1:* We are interested in comparing the product quality exam pass rates of three suppliers A, B, and C using a nonparametric (i.e., distribution-free) hypothesis test: chi-square test.
  - Data File: “Chi-Square Test1” tab in “Sample Data.xlsx”

- Null Hypothesis (H_0): $p_A = p_B = p_C$
- Alternative Hypothesis (H_a): at least one of the suppliers has different pass rates from the others.

Use Minitab to Run a Chi-Square Test

- Steps to run a chi-square test in Minitab
  1) Click Stat → Tables → Cross Tabulation and Chi-Square.
  2) A new window named "Cross Tabulation and Chi-Square" pops up.
  3) Select "Results" as "For rows."
  4) Select "Supplier" as "For columns."
  5) Select "Count" as "Frequencies are in."
  6) Click the "Chi-Square" button.
  7) A new window named "Cross Tabulation – Chi-Square" pops up.
  8) Check the boxes of "Chi-square test", "Expected cell counts," and "Each cell's contribution to the Chi-Square statistic."
  9) Click "OK" in the window named "Cross Tabulation – Chi-Square."
  10) Click "OK" in the window named "Cross Tabulation and Chi-Square."
  11) The Chi-square test results appear in the session window.

Use Minitab to Run a Chi-Square Test



Cross Tabulation and Chi-Square

C3 Count

Raw data (categorical variables)

Rows: Results

Columns: Supplier

Layers:

Frequencies: Count (optional)

Display

☑ Counts
☐ Row percents
☐ Column percents
☐ Total percents

Cross Tabulation: Chi-Square

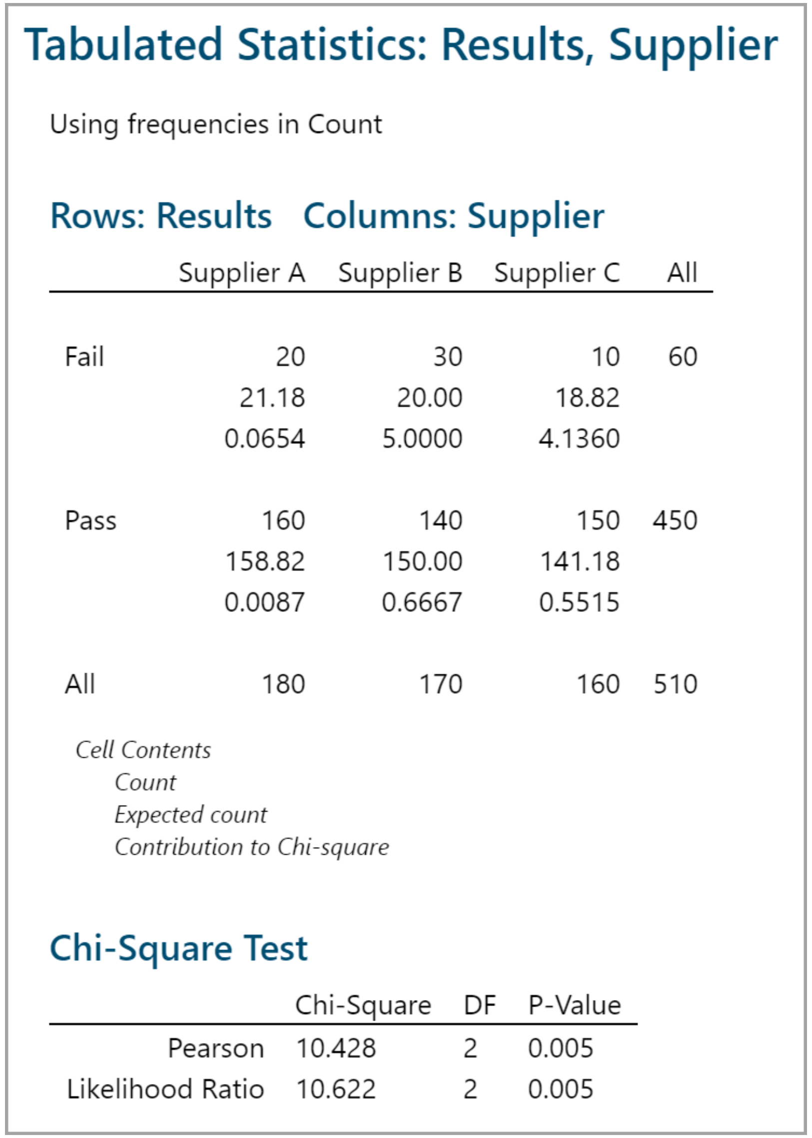
☑ Chi-square test

Statistics to display in each cell

☑ Expected cell counts
☐ Raw residuals
☐ Standardized residuals
☐ Adjusted residuals
☑ Each cell's contribution to chi-square

Use Minitab to Run a Chi-Square Test

- Counts are based on the sample observation.
- Expected counts are based on the assumption that the null hypothesis is true.
- Since the p-value is smaller than alpha level (0.05), we reject the null hypothesis and claim that at least one supplier has different pass rate from others.



```
Tabulated Statistics: Results, Supplier

Using frequencies in Count

Rows: Results   Columns: Supplier

          Supplier A  Supplier B  Supplier C  All
-------------------------------------------------
Fail              20          30          10   60
               21.18       20.00       18.82
              0.0654      5.0000      4.1360

Pass             160         140         150  450
              158.82      150.00      141.18
              0.0087      0.6667      0.5515

All              180         170         160  510

 Cell Contents
    Count
    Expected count
    Contribution to Chi-square


Chi-Square Test

                  Chi-Square  DF  P-Value
-----------------------------------------
          Pearson     10.428   2    0.005
 Likelihood Ratio     10.622   2    0.005
```

Use Minitab to Run a Chi-Square Test

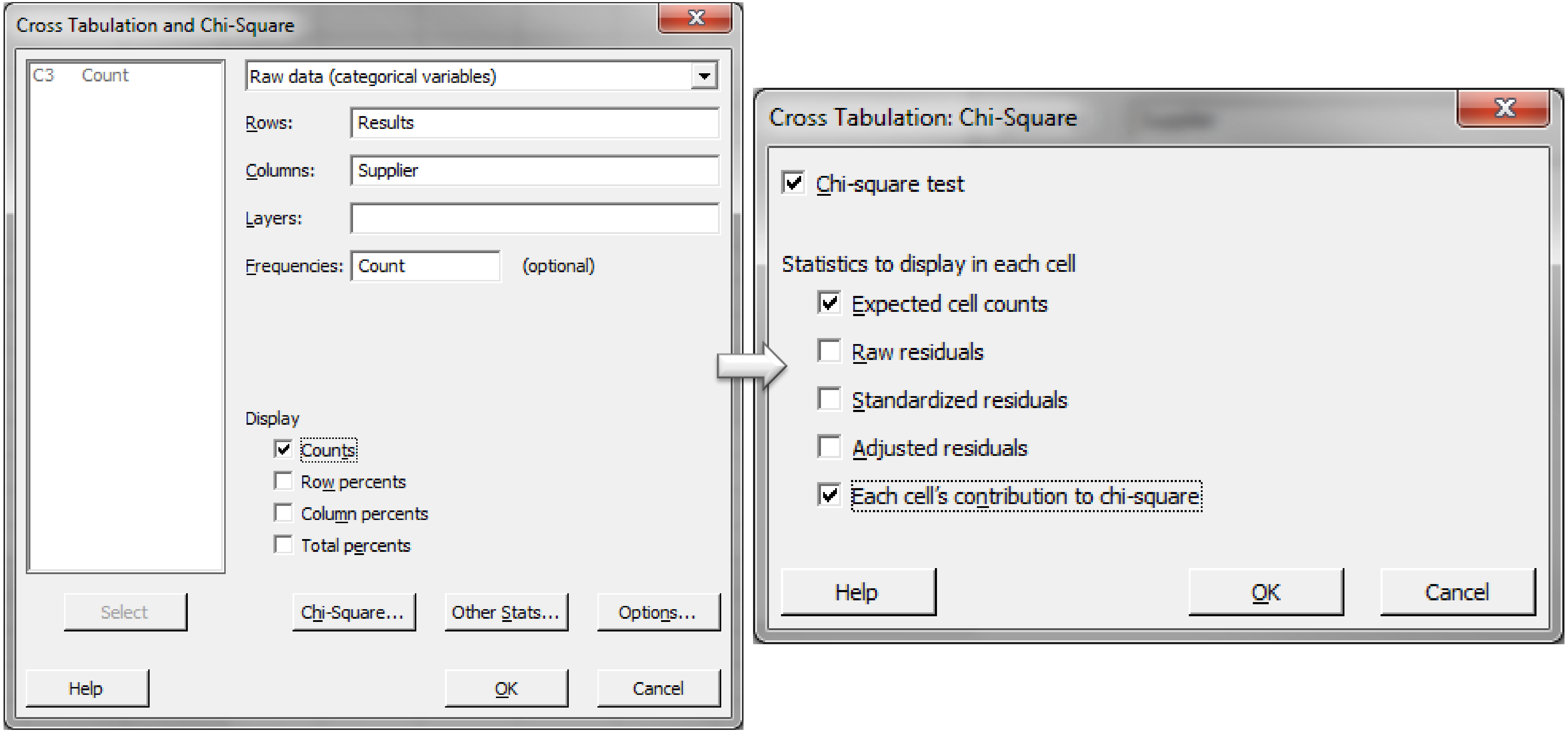
- *Case study 2:* We are trying to check whether there is a relationship between the suppliers and the results of the product quality exam using nonparametric (i.e., distribution-free) hypothesis test: chi-square test.
    - Data File: “Chi-Square Test2” tab in “Sample Data.xlsx”

- Null Hypothesis (H_0): product quality exam results are independent of the suppliers.
- Alternative Hypothesis (H_a): product quality exam results depend on the suppliers.

Use Minitab to Run a Chi-Square Test

- Steps to run a chi-square test in Minitab
  1) Click Stat → Tables → Cross Tabulation and Chi-Square.
  2) A new window named “Cross Tabulation and Chi-Square” pops up.
  3) Select “Results” as “For rows.”
  4) Select “Supplier” as “For columns.”
  5) Select “Count” as “Frequencies are in.”
  6) Click the “Chi-Square” button.
  7) Check the boxes of “Chi-square analysis”, “Expected cell counts,” and “Each cell’s contribution to the Chi-Square statistic.”
  8) Click “OK” in the window named “Cross Tabulation – Chi-Square.”
  9) Click “OK” in the window named “Cross Tabulation and Chi-Square.”
  10) The Chi-square test results appear in the session window.

Use Minitab to Run a Chi-Square Test



Cross Tabulation and Chi-Square

C3 Count

Raw data (categorical variables)

Rows: Results

Columns: Supplier

Layers:

Frequencies: Count (optional)

Display

☑ Counts
☐ Row percents
☐ Column percents
☐ Total percents

Cross Tabulation: Chi-Square

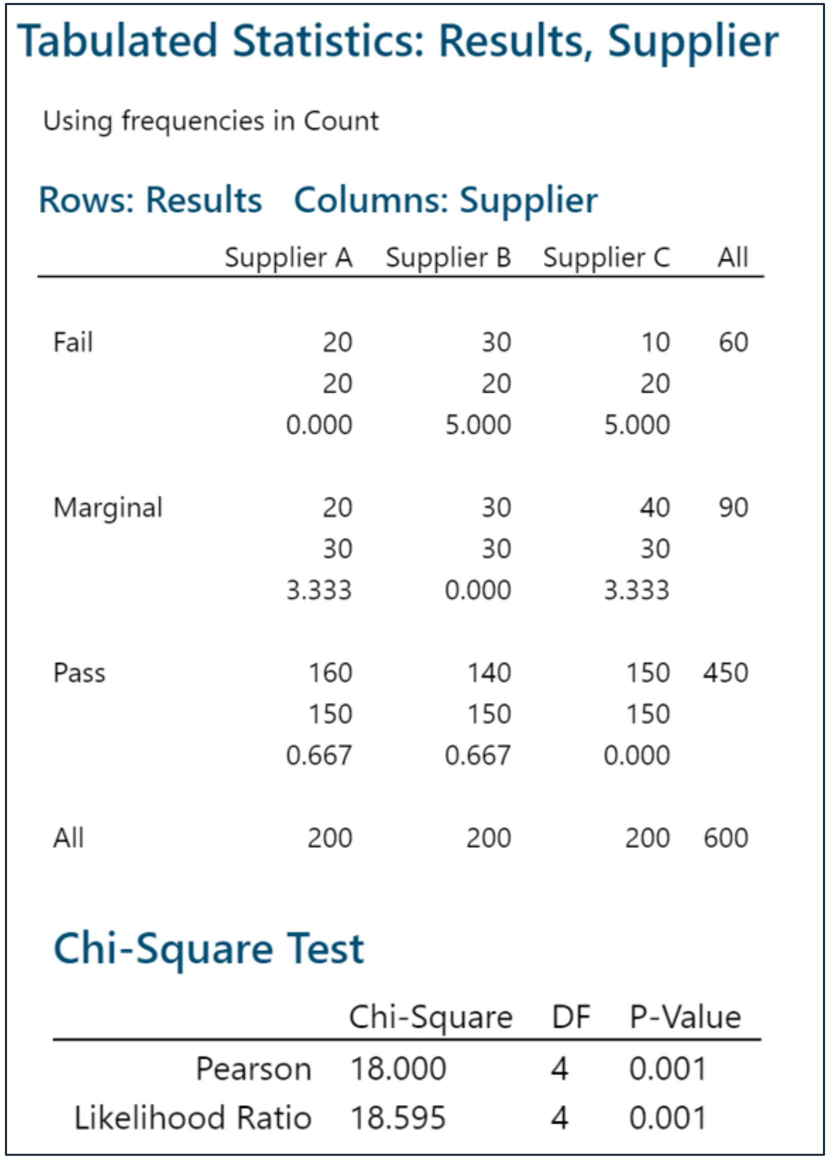
☑ Chi-square test

Statistics to display in each cell

☑ Expected cell counts
☐ Raw residuals
☐ Standardized residuals
☐ Adjusted residuals
☑ Each cell's contribution to chi-square

Use Minitab to Run a Chi-Square Test

- The p-value is smaller than the alpha level (0.05) and we reject the null hypothesis.
- The product quality exam results are not independent of the suppliers.



```
Tabulated Statistics: Results, Supplier

 Using frequencies in Count

 Rows: Results   Columns: Supplier

                   Supplier A   Supplier B   Supplier C   All

  Fail                     20           30           10    60
                           20           20           20
                        0.000        5.000        5.000

  Marginal                 20           30           40    90
                           30           30           30
                        3.333        0.000        3.333

  Pass                    160          140          150   450
                          150          150          150
                        0.667        0.667        0.000

  All                     200          200          200   600


  Chi-Square Test

                        Chi-Square   DF   P-Value
          Pearson       18.000       4    0.001
    Likelihood Ratio    18.595       4    0.001
```

3.5.9 Tests of Equal Variance

What are Tests of Equal Variance?

- **Tests of equal variance** are a family of hypothesis tests used to check whether there is a statistically significant difference between the variances of two or more populations.
  - Null Hypothesis (H_0): $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$
  - Alternative Hypothesis (H_a): at least the variance of one population is different from others.
  - k is the number of populations of interest. $k \geq 2$.
- A tests of equal variance can be used alone but most of the time it is used with other statistical methods to verify or support the assumption about the variance equality.

F-Test

- The **F-test** is used to compare the variances between two normally distributed populations.
- It is extremely sensitive to non-normality and serves as a preliminary step for two sample t-test.
- Test Statistic:

$$F_{calc} = \frac{s_1^{\,2}}{s_2^{\,2}}$$, where $\mathbf{s_1}$ and s_2 are the sample standard deviations**.**

The sampling distribution of the test statistic follows F distribution when the null is true.

Bartlett's Test

- **Bartlett's test** is used to compare the variances among two or more normally distributed populations.
- It is sensitive to non-normality and it serves as a preliminary step for ANOVA.

- Test Statistic:

$$\chi^2 = \frac{(N-k)\ln(S_p^2) - \sum_{i=1}^{k}(n_i-1)\ln(S_i^2)}{1+\frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\left(\frac{1}{n_i-1}\right)-\frac{1}{N-k}\right)}$$

, where $N = \sum_{i=1}^{k} n_i$ and $S_p^2 = \frac{1}{N-k}\sum_{i}(n_i-1)S_i^2$

The sampling distribution of test statistic follows chi^2 distribution when the null is true.

Brown-Forsythe Test

- The **Brown-Forsythe test** is used to compare the variances between two or more populations with any distributions.

- It is not so sensitive to non-normality as Bartlett's test.

- The Test statistic is the model F statistic from the ANOVA on the transformed response $z_{ij} = \left| y_{ij} - \tilde{y}_i \right|$ where $\tilde{y}_i$ is the median response at ith level.

Levene's Test

- **Levene's test** is used to compare the variances between two or more populations with any distributions.

- It is not so sensitive to non-normality as Bartlett's test.

- The test statistic is the model F statistic from the ANOVA on the transformed response $z_{ij} = \left| y_{ij} - \bar{y}_i \right|$ where $\bar{y}_i$ is the mean response at i^{th} level.

Brown-Forsythe Test vs. Levene’s Test

$$F = \frac{(N-k)}{(k-1)} \frac{\sum_{i=1}^{k} N_i (Z_{i.} - Z_{..})^2}{\sum_{i=1}^{k} \sum_{j=1}^{N_i} (Z_{ij} - Z_{i.})^2}$$

where

N is the total number of observations.
k is the number of groups.
N_i is the number of observations in the ith group.
$Z_{i.}$ is the group mean of the ith group.
$Z_{..}$ is the grand mean of all the observations.

In Brown-Forsythe Test, $Z_{ij} = \left| Y_{ij} - \widetilde{Y}_{ij} \right|$, where $\widetilde{Y}_{ij}$ is the group median of the ith group.

In Levene’s Test, $Z_{ij} = \left| Y_{ij} - \overline{Y}_{ij} \right|$, where $\overline{Y}_{ij}$ is the group mean of the ith group.

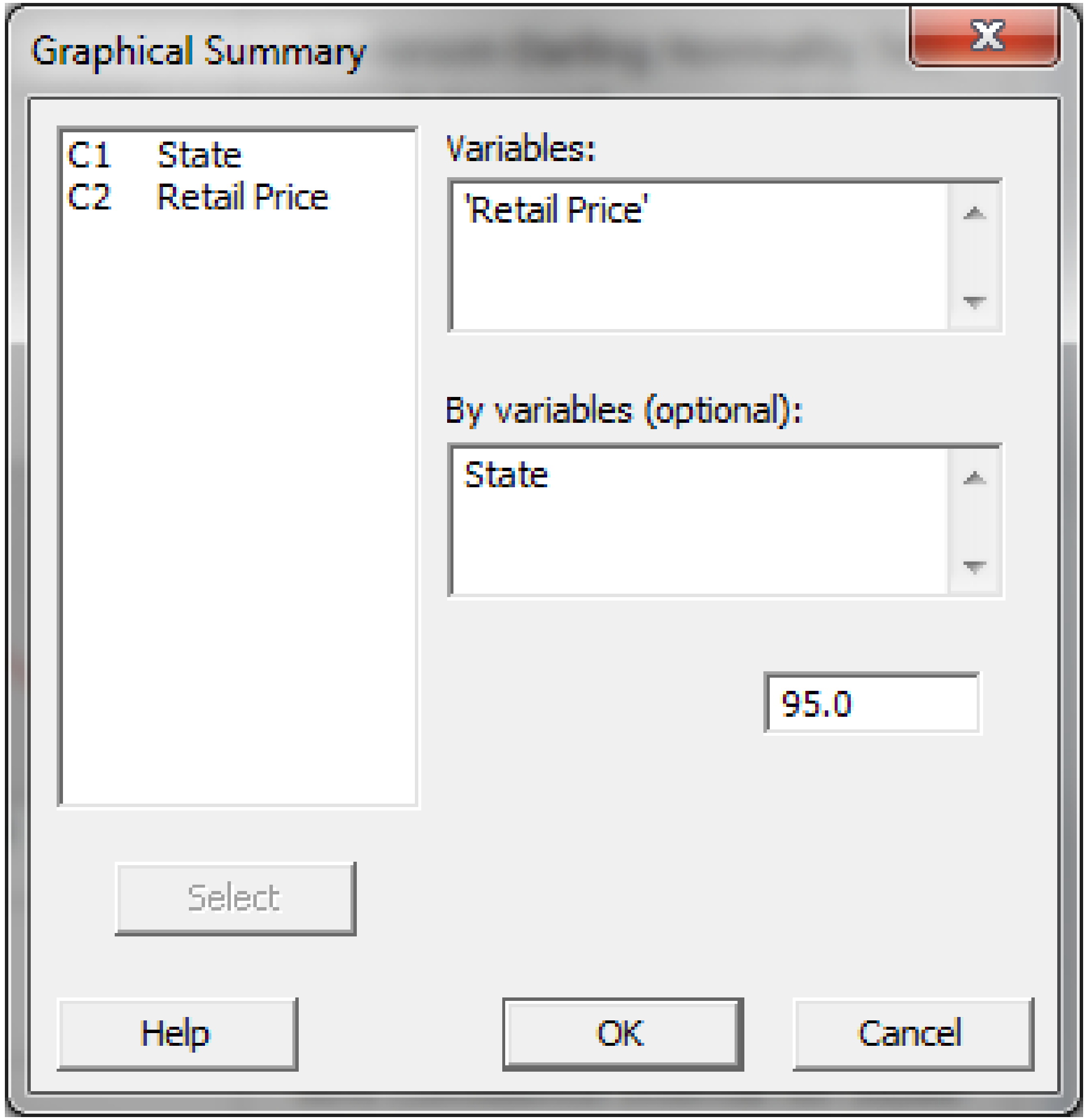
Use Minitab to Run Tests of Equal Variance

- *Case study:* We are interested in comparing the variances of the retail price of a product in state A and state B.
  - Data File: “Two-Sample T-Test” tab in “Sample Data.xlsx”
- Null Hypothesis (H_0): $\sigma_A^2 = \sigma_B^2$
- Alternative Hypothesis (H_a): $\sigma_A^2 \neq \sigma_B^2$

Use Minitab to Run Tests of Equal Variance

- Step 1: Run the normality test to check whether all levels of data are normally distributed
    1) Click Stat → Basic Statistics → Graphical Summary.
    2) A new window named “Graphical Summary” pops up.
    3) Select “Retail Price” as the “Variables.”
    4) Click in the blank box right below “By variables” and the “State” appears in the list box on the left.
    5) Select “State” as “By variables.”
    6) Click “OK.”
    7) The normality test results appear in the new windows.

Use Minitab to Run Tests of Equal Variance



Graphical Summary

C1 State
C2 Retail Price

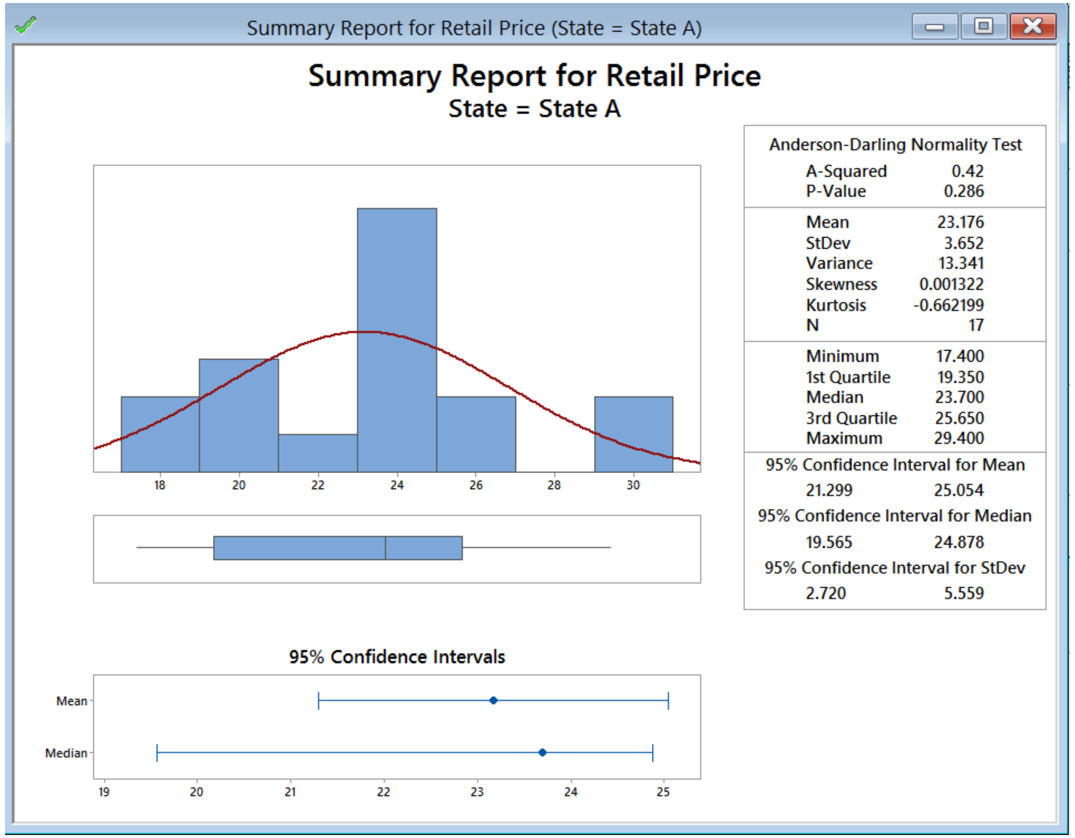
Variables: 'Retail Price'

By variables (optional): State

95.0

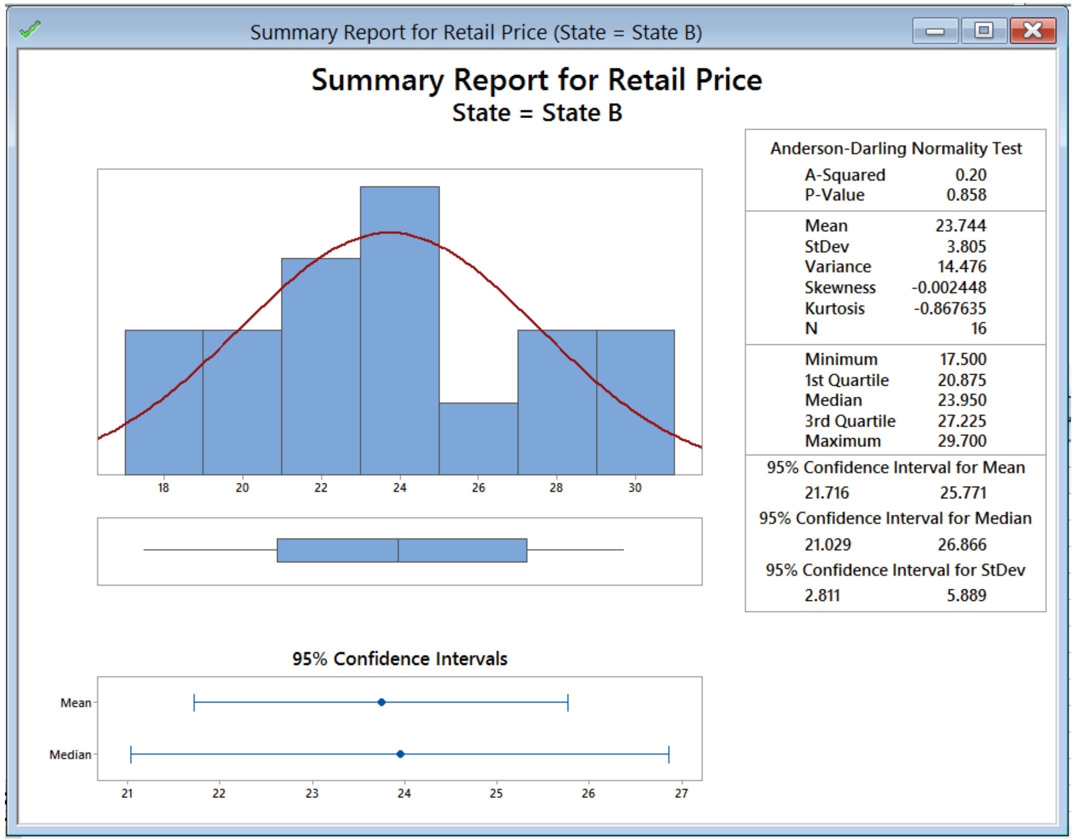
Use Minitab to Run Tests of Equal Variance

- Null Hypothesis (H0): The data are normally distributed.
- Alternative Hypothesis (Ha): The data are not normally distributed.
- Both retail price data of state A and B are normally distributed since the p-values are both greater than alpha level (0.05).



Summary Report for Retail Price
State = State A

Anderson-Darling Normality Test	
A-Squared	0.42
P-Value	0.286
Mean	23.176
StDev	3.652
Variance	13.341
Skewness	0.001322
Kurtosis	-0.662199
N	17
Minimum	17.400
1st Quartile	19.350
Median	23.700
3rd Quartile	25.650
Maximum	29.400
95% Confidence Interval for Mean	
21.299	25.054
95% Confidence Interval for Median	
19.565	24.878
95% Confidence Interval for StDev	
2.720	5.559



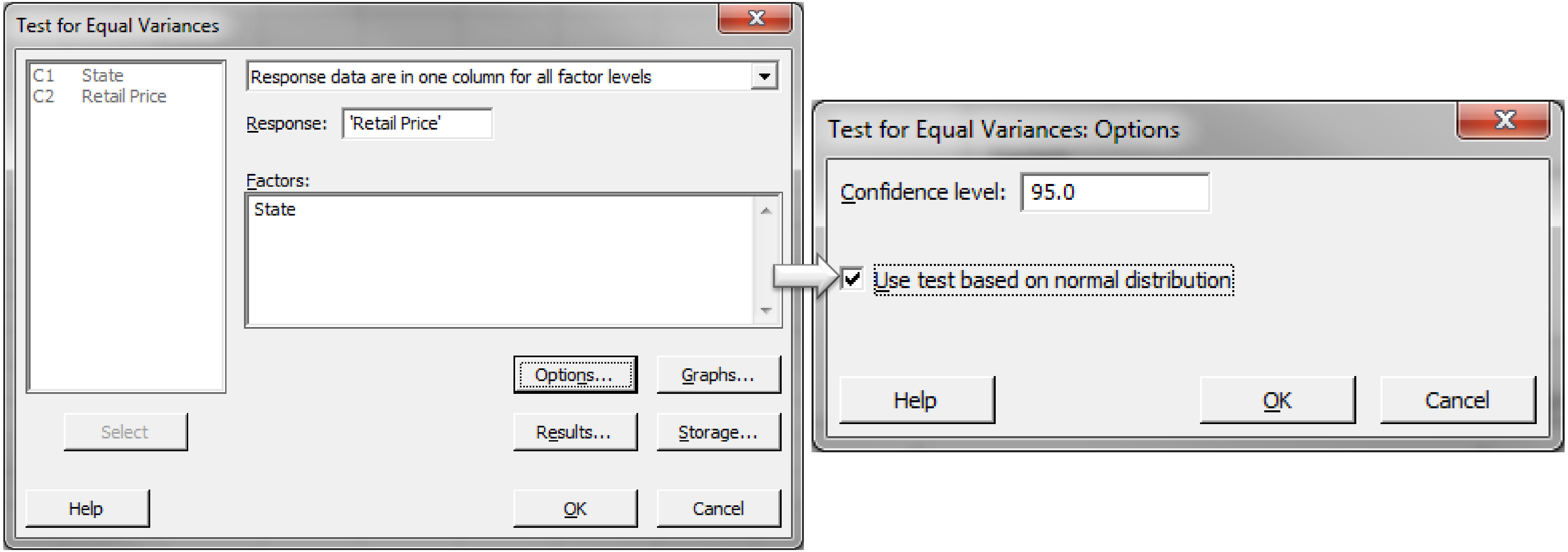
Summary Report for Retail Price
State = State B

Anderson-Darling Normality Test	
A-Squared	0.20
P-Value	0.858
Mean	23.744
StDev	3.805
Variance	14.476
Skewness	-0.002448
Kurtosis	-0.867635
N	16
Minimum	17.500
1st Quartile	20.875
Median	23.950
3rd Quartile	27.225
Maximum	29.700
95% Confidence Interval for Mean	
21.716	25.771
95% Confidence Interval for Median	
21.029	26.866
95% Confidence Interval for StDev	
2.811	5.889

Use Minitab to Run Tests of Equal Variance

- Step 2: Run tests of equal variance in Minitab
  1) Click Stat → ANOVA → Test for Equal Variances.
  2) A new window named “Test for Equal Variances” pops up.
  3) Select “Retail Price” as the “Response.”
  4) Select “State” as “Factors.”
  5) Click “Options”
  6) Check “Use test based on normal distribution”
  7) Click “OK” to close the Options window
  8) Click “OK.”
  9) The results of variances equality tests appear in the new window.

Use Minitab to Run Tests of Equal Variance



Test for Equal Variances

C1 State
C2 Retail Price

Response data are in one column for all factor levels

Response: 'Retail Price'

Factors: State

Test for Equal Variances: Options

Confidence level: 95.0

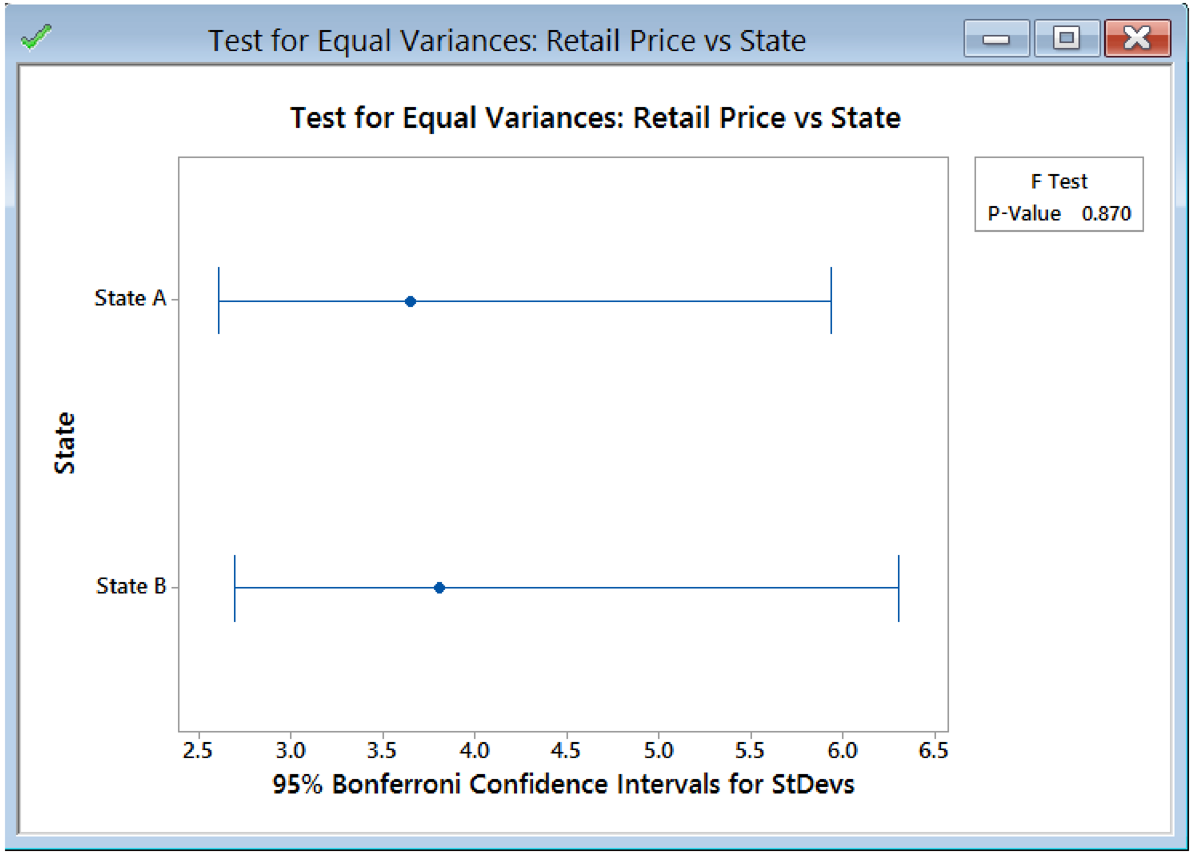
☑ Use test based on normal distribution

Use Minitab to Run Tests of Equal Variance

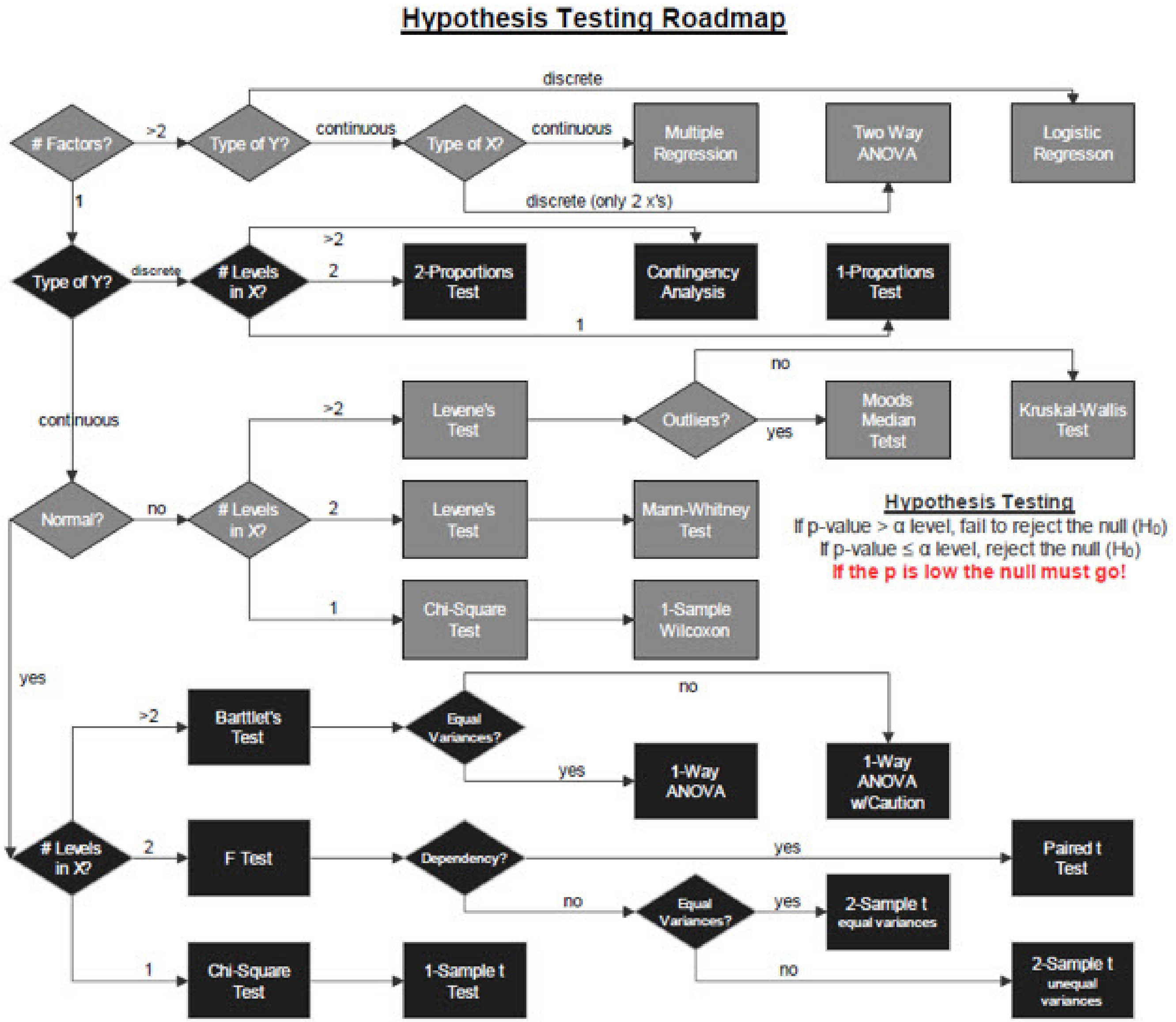
- If all the groups of data are normally distributed, use the F-test or Bartlett’s test in Minitab to test the equality of the variances.

- If at least one of the groups is not normally distributed, use Levene’s test in Minitab to test the equality of the variances.

- If the p-value of the variances equality test is greater than the alpha level (0.05), we fail to reject the null hypothesis and conclude that the variances of different groups are identical.



Hypothesis Testing Roadmap: Putting it all together



Hypothesis Testing

If p-value > α level, fail to reject the null (H_0)

If p-value ≤ α level, reject the null (H_0)

If the p is low the null must go!

4.0 Improve Phase

Black Belt Training: Improve Phase

4.1 Simple Linear Regression
- 4.1.1 Correlation
- 4.1.2 X-Y Diagram
- 4.1.3 Regression Equations
- 4.1.4 Residuals Analysis

4.2 Multiple Regression Analysis
- 4.2.1 Non-Linear Regression
- 4.2.2 Multiple Linear Regression
- 4.2.3 Confidence Intervals
- 4.2.4 Residuals Analysis
- 4.2.5 Data Transformation, Box Cox
- 4.2.6 Stepwise Regression
- 4.2.7 Logistic Regression

4.3 Designed Experiments
- 4.3.1 Experiment Objectives
- 4.3.2 Experimental Methods
- 4.3.3 DOE Design Considerations

4.4 Full Factorial Experiments
- 4.4.1 2k Full Factorial Designs
- 4.4.2 Linear and Quadratic Models
- 4.4.3 Balanced and Orthogonal Designs
- 4.4.4 Fit, Model, and Center Points

4.5 Fractional Factorial Experiments
- 4.5.1 Designs
- 4.5.2 Confounding Effects
- 4.5.3 Experimental Resolution

4.1 Simple Linear Regression

Black Belt Training: Improve Phase

4.1 Simple Linear Regression
  - 4.1.1 Correlation
  - 4.1.2 X-Y Diagram
  - 4.1.3 Regression Equations
  - 4.1.4 Residuals Analysis

4.2 Multiple Regression Analysis
  - 4.2.1 Non-Linear Regression
  - 4.2.2 Multiple Linear Regression
  - 4.2.3 Confidence Intervals
  - 4.2.4 Residuals Analysis
  - 4.2.5 Data Transformation, Box Cox
  - 4.2.6 Stepwise Regression
  - 4.2.7 Logistic Regression

4.3 Designed Experiments
  - 4.3.1 Experiment Objectives
  - 4.3.2 Experimental Methods
  - 4.3.3 DOE Design Considerations

4.4 Full Factorial Experiments
  - 4.4.1 2k Full Factorial Designs
  - 4.4.2 Linear and Quadratic Models
  - 4.4.3 Balanced and Orthogonal Designs
  - 4.4.4 Fit, Model, and Center Points

4.5 Fractional Factorial Experiments
  - 4.5.1 Designs
  - 4.5.2 Confounding Effects
  - 4.5.3 Experimental Resolution

4.1.1 Correlation

What is Correlation?

- **Correlation** is a statistical technique that describes whether and how strongly two or more variables are related.

- **Correlation analysis** helps to understand the direction and degree of association between variables, and it suggests whether one variable can be used to predict another.

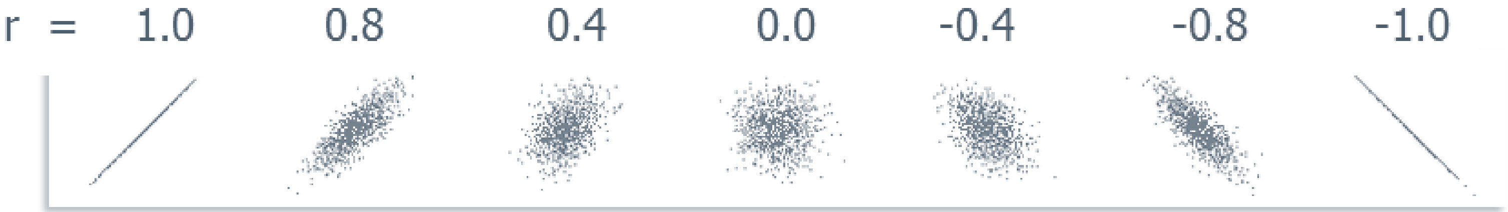
- Of the different metrics to measure correlation, Pearson's correlation coefficient is the most popular. It measures the linear relationship between two variables.

Pearson's Correlation Coefficient

- **Pearson's correlation coefficient** is also called:
  - Pearson's r or coefficient of correlation
  - Pearson's product moment correlation coefficient (r)
- "r" is a statistic measuring the linear relationship between two variables.
- Correlation coefficients range from -1 to 1.
  - If r = 0, there is no linear relationship between the variables.
  - The *sign* of r indicates the *direction* of the relationship:
    - If $r < 0$, there is a negative linear correlation.
    - If $r > 0$, there is a positive linear correlation.
  - The *absolute value* of r describes the *strength* of the relationship:
    - If $|r| \leq 0.5$, there is a weak linear correlation.
    - If $|r| > 0.5$, there is a strong linear correlation.
    - If $|r| = 1$, there is a perfect linear correlation.

Pearson's Correlation Coefficient

- When the correlation is *strong*, the data points on a scatter plot will be close together (tight).
  - The closer "r" is to -1 or 1, the stronger the relationship.
    - -1 Strong inverse relationship
    - +1 Strong direct relationship
- When the correlation is *weak*, the data points are spread apart more (loose).
  - The closer the correlation is to 0 the weaker the relationship.

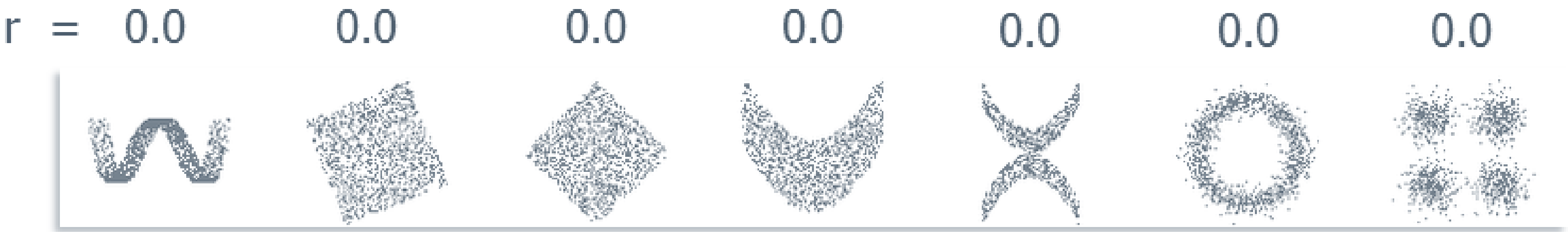


Pearson’s Correlation Coefficient

- Pearson’s correlation coefficient is only sensitive to the *linear* dependence between two variables.

- It is possible that two variables have a perfect non-linear relationship when the correlation coefficient is low.

- Notice the scatter plots below with correlation equal to 0. There are clearly *relationships* but they are not linear and therefore can not be determined with Pearson’s correlation coefficient.



Correlation and Causation

- Correlation *does not* imply causation.

- If variable A is highly correlated with variable B, it does not necessarily mean A causes B or vice versa. It is possible that an unknown third variable C is causing both A and B to change.

- For example, if ice cream sales at the beach are highly correlated with the number of shark attacks, it does not imply that increased ice cream sales causes increased shark attacks. They are triggered by a third factor: summer.

Correlation and Dependence

- If two variables are independent, the correlation coefficient is zero.
- WARNING! If the correlation coefficient of two variables is zero, it does not imply they are independent.
- The correlation coefficient only indicates the linear dependence between two variables. When variables are non-linearly related, they are not independent of each other but their correlation coefficient could be zero.

Correlation Coefficient and X-Y Diagram

- The correlation coefficient indicates the direction and strength of the linear dependence between two variables but it does not cover all the existing relationship patterns.

- With the same correlation coefficient, two variables might have completely different dependence patterns.

- A scatter plot or X-Y diagram can help to discover and understand additional characteristics of the relationship between variables.

- Correlation coefficient is not a replacement for examining the scatter plot to study the variables' relationship.

Statistical Significance of the Correlation Coefficient

- The correlation coefficient could be high or low by chance (randomness). It may have been calculated based on two small samples that do not provide good inference on the correlation between two populations.

- In order to test whether there is a statistically significant relationship between two variables, we need to run a hypothesis test to determine whether the correlation coefficient is statistically different from zero.
  - Hypothesis Test Statements
    - H_0: $r = 0$: Null Hypothesis: There is *no* correlation.
    - H_1: $r \neq 0$: Alternate Hypothesis: There is a correlation.

Statistical Significance of the Correlation Coefficient

- Hypothesis tests will produce p-values as a result of the statistical significance test on r.
  - When the p-value for a test is low (less than 0.05), we can reject the null hypothesis and conclude that “r” is significant; there is a correlation.
  - When the p-value for a test is > 0.05, then we fail to reject the null hypothesis; there is no correlation.

- We can also use the t statistic to draw the same conclusions regarding our test for significance of the correlation coefficient.

Statistical Significance of the Correlation Coefficient

- Test Statistic: $$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$
- Critical Statistic: t-value in t-table with (n – 2) degrees of freedom
- If $|t| \leq t_{critical}$, we fail to reject the null. There is no statistically significant linear relationship between X and Y.
- If $|t| > t_{critical}$, we reject the null. There is a statistically significant linear relationship between X and Y.

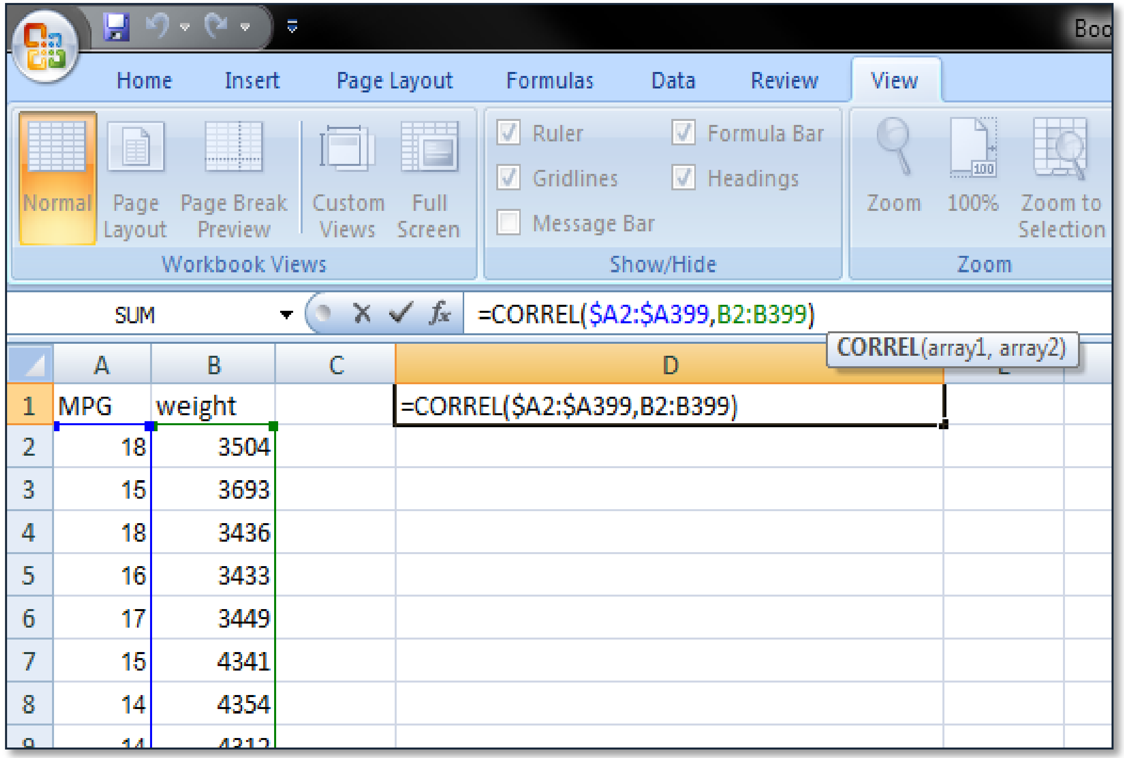
Using Software to Calculate the Correlation Coefficient

- We are interested in understanding whether there is linear dependence between a car’s MPG and its weight and if so, how they are related.

- The MPG and weight data are stored in the “Correlation Coefficient” tab in “Sample Data.xlsx.” We will discuss three ways to get the results.

Use Excel to Calculate the Correlation Coefficient

- The formula CORREL in Excel calculates the sample correlation coefficient of two data series.
- The correlation coefficient between the two data series is -0.83, which indicates a strong negative linear relationship between MPG and weight.



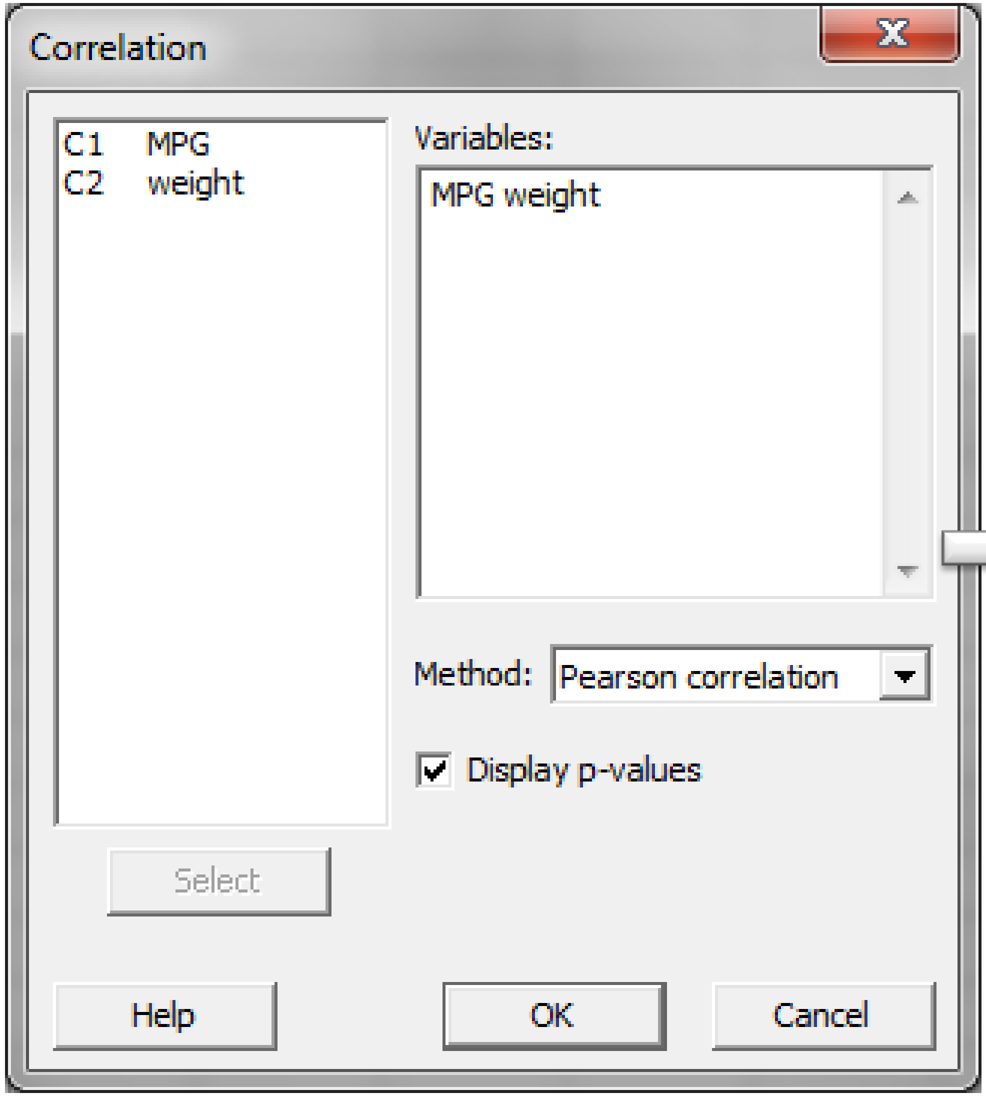
=CORREL($A2:$A399,B2:B399)

	A	B	C	D
1	MPG	weight		=CORREL($A2:$A399,B2:B399)
2	18	3504		
3	15	3693		
4	18	3436		
5	16	3433		
6	17	3449		
7	15	4341		
8	14	4354		

Use Minitab to Calculate the Correlation Coefficient

- Step 1: Stat → Basic Statistics → Correlation.
- Step 2: Select the two variables of interest in the pop-up window "Correlation" and click "OK."
- The correlation coefficient result (-0.832) appears in the session window. The p-value (0.000) is lower than the alpha level (0.05), indicating the linear correlation is statistically significant.

Use Minitab to Calculate the Correlation Coefficient

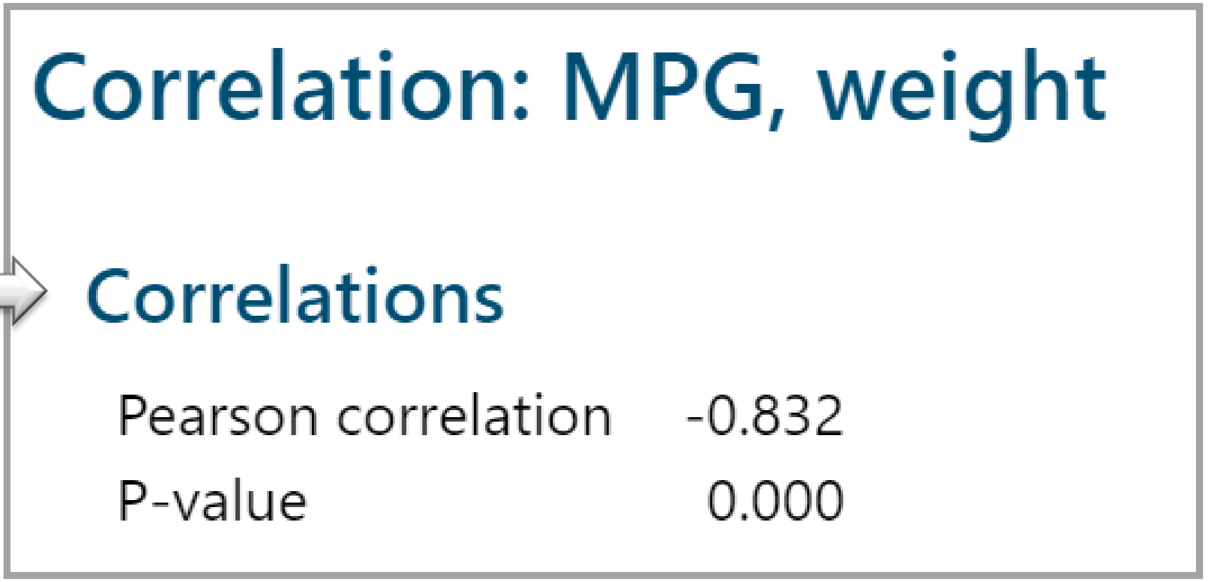


C1 MPG
C2 weight

Variables: MPG weight

Method: Pearson correlation

☑ Display p-values



Correlation: MPG, weight

Correlations

```
Pearson correlation    -0.832
P-value                 0.000
```

Interpreting Results

- How do we interpret results and make decisions based Pearson’s correlation coefficient (r) and p-values?
  - Let us look at a few examples:
    - r = -0.832, p = 0.000 (previous example). The two variables are inversely related and the linear relationship is strong. Also, this conclusion is significant as supported by p-value of 0.00.
    - r = -0.832, p = 0.71. Based on r, you should conclude the linear relationship between the two variables is strong and inversely related. However, with a p-value of 0.71, you should then conclude that r is not significant and that your sample size may be too small to accurately characterize the relationship.
    - r = 0.5, p = 0.00. Moderately positive linear relationship, r is statistically significant.
    - r = 0.92, p = 0.61. Strong positive linear relationship but r is not statistically significant. Get more data.
    - r = 1.0, p = 0.00. The two variables have a perfect linear relationship and r is significant.

Correlation Coefficient Calculation

- Population Correlation Coefficient (ρ)

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

- Sample Correlation Coefficient (r)

$$r = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i-\bar{Y})^2}}$$

- It is only defined when the standard deviations of both X and Y are non-zero and finite.
- When covariance of X and Y is zero, the correlation coefficient is zero.

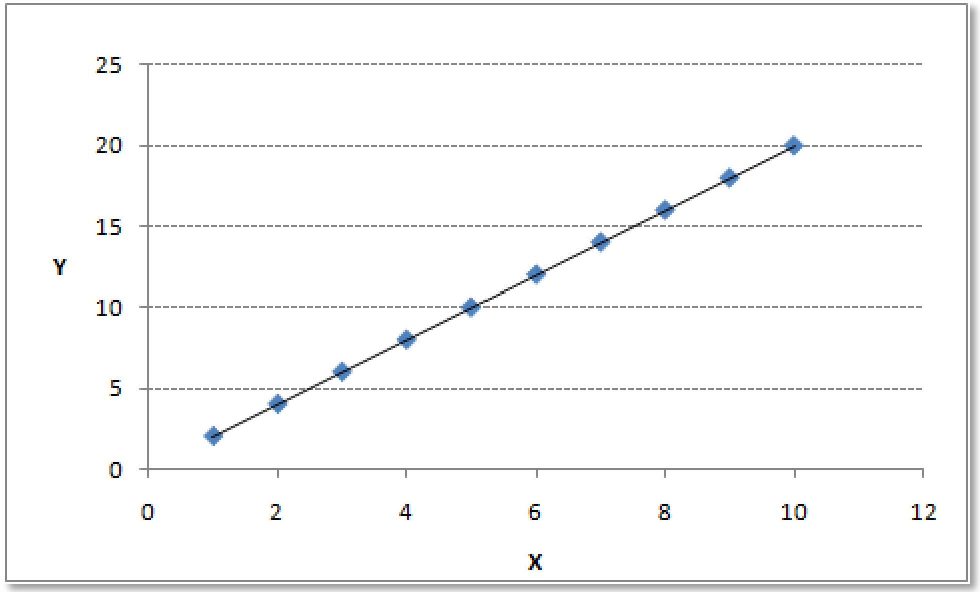
4.1.2 X-Y Diagram

What is an X-Y Diagram?

- An **X-Y diagram** is a scatter plot depicting the relationship between two variables (i.e., X and Y).
- Each point on the X-Y diagram represents a pair of X and Y values, with X plotted on the horizontal axis and Y plotted on the vertical axis.
- With an X-Y diagram, you can qualitatively assess both the strength and direction of the relationship between X and Y.
- To quantitatively measure the relationship between X and Y, you may need to calculate the correlation coefficient.

Example 1: Perfect Linear Correlation

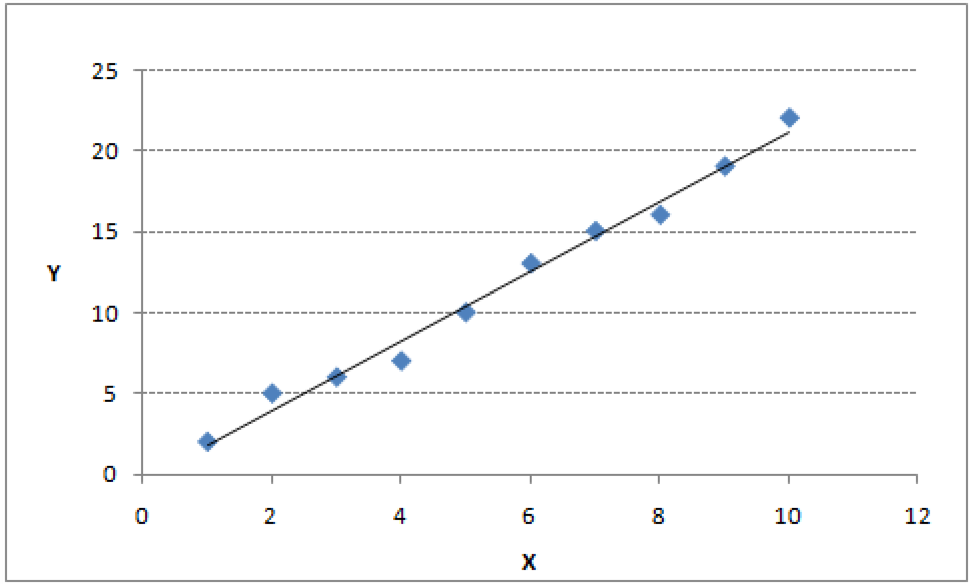
- In the chart below, there are 10 data points depicted (10 pairs of X and Y values), and they were created using the equation Y = 2X.
- As a result, all the data points fall onto the straight line of Y = 2X.
- The chart demonstrates a perfect positive linear correlation between X and Y since the relationship between X and Y can be perfectly described by a linear equation in a format of Y = a × X + b where a ≠ 0.



X	Y
1	2
2	4
3	6
4	8
5	10
6	12
7	14
8	16
9	18
10	20

Example 2: Strong Linear Correlation

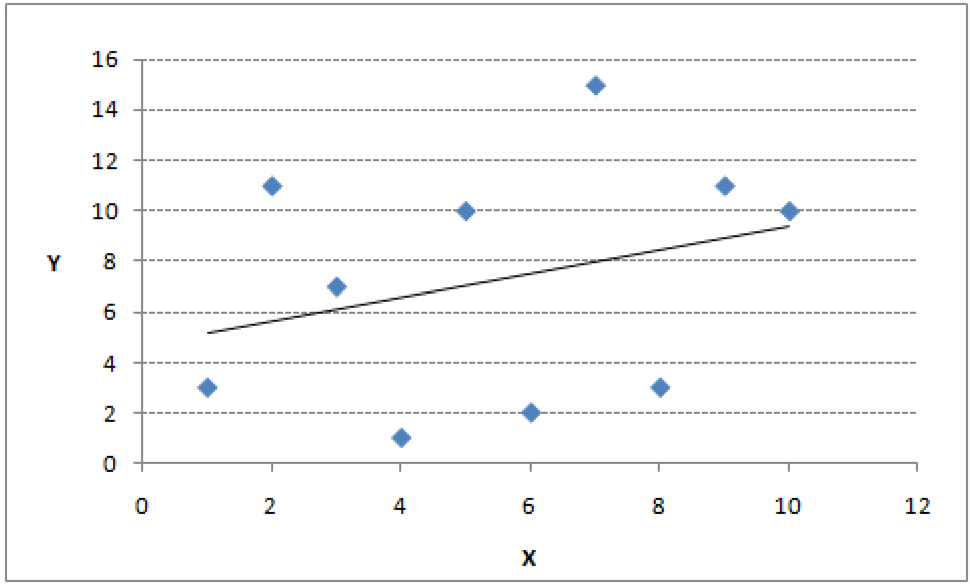
- In this chart, the data points scatter closely around a straight line.
- When X increases, Y increases accordingly.
- This chart demonstrates a strong positive linear correlation between X and Y.
- The straight line is the trend line showing how Y's trend goes with changes in X.



X	Y
1	2
2	5
3	6
4	7
5	10
6	13
7	15
8	16
9	19
10	22

Example 3: Weak Linear Correlation

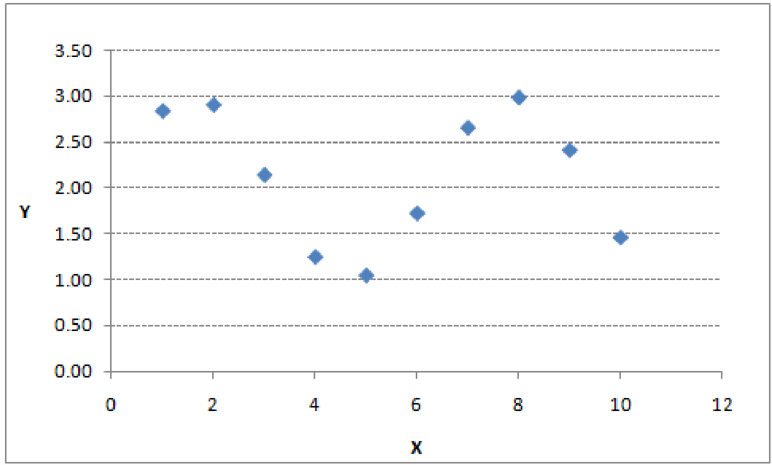
- In this chart, the data points scatter remotely around a straight line.
- When X increases, Y increases accordingly.
- This chart demonstrates a weak positive linear correlation between X and Y since the distance between the data points and the trend line is relatively far on average.



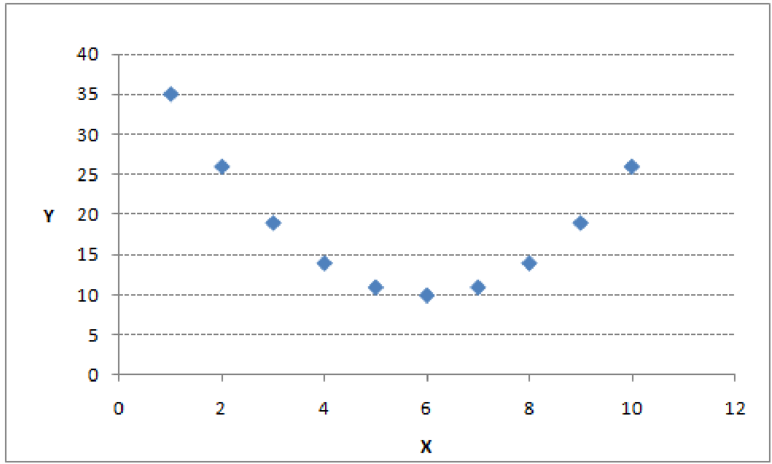
X	Y
1	3
2	11
3	7
4	1
5	10
6	2
7	15
8	3
9	11
10	10

Example 4: Non-Linear Correlation

- The X-Y diagram also helps to identify any nonlinear relationship between X and Y.



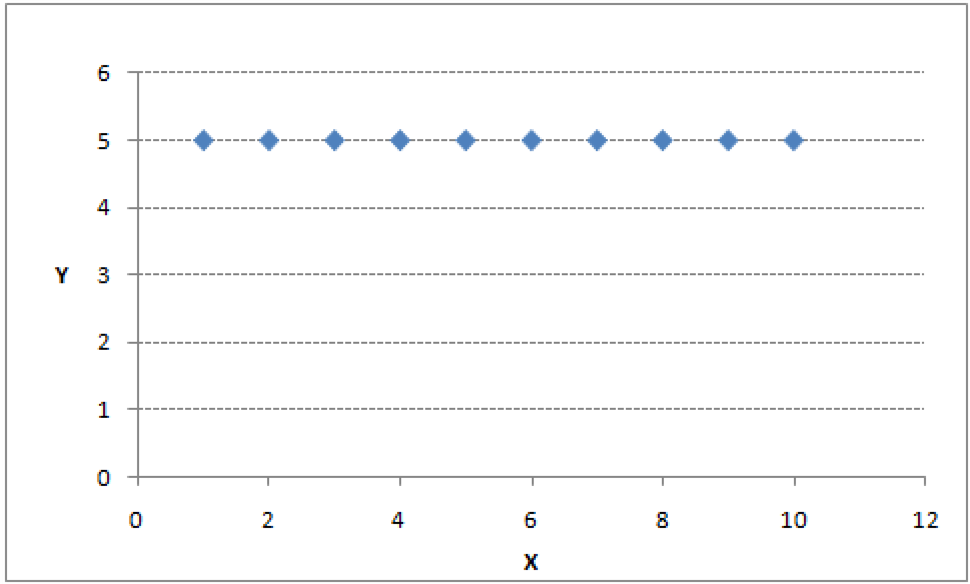
X	Y
1	2.84
2	2.91
3	2.14
4	1.24
5	1.04
6	1.72
7	2.66
8	2.99
9	2.41
10	1.46



X	Y
1	35
2	26
3	19
4	14
5	11
6	10
7	11
8	14
9	19
10	26

Example 5: Uncorrelated

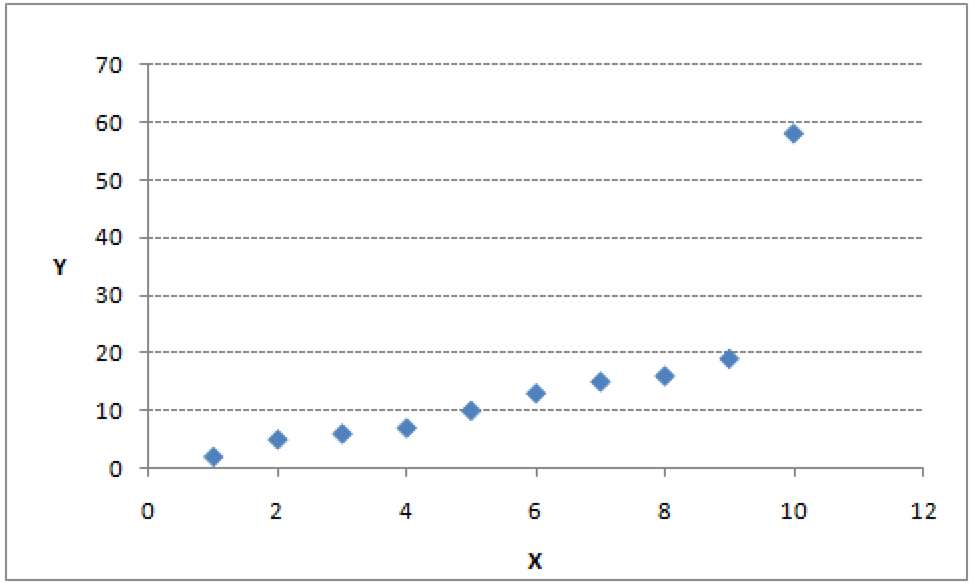
- In this chart, the Y value of each data point is a constant regardless of what the X value is.
- Changes in X do not show any relative impact on Y. As a result, there is no correlation between X and Y.



X	Y
1	5
2	5
3	5
4	5
5	5
6	5
7	5
8	5
9	5
10	5

Example 6: Outlier Identification

- Using X-Y diagram, you may identify outliers in the data.
- In this chart, the last data point does not seem to follow the trend of other data points.
- This should require further investigation on the last data point to determine whether it is an outlier.



X	Y
1	2
2	5
3	6
4	7
5	10
6	13
7	15
8	16
9	19
10	58

Benefits of Using an X-Y Diagram

- An X-Y diagram graphically demonstrates the relationship between two variables.
- It suggests whether two variables are associated and helps to identify the linear or nonlinear correlation between X and Y.
- It captures the strength and direction of the relationship between X and Y.
- It helps identify any outliers in the data.

Limitations of the X-Y Diagram

- Although the X-Y diagram helps to “spot” interesting features in the data, it does not provide any quantitative conclusions about the data and further statistical analysis is needed to:
  - Assess whether the association between variables is statistically significant.
  - Measure the strength of the relationship between variables.
  - Determine whether outliers exist in the data.
  - Quantitatively describe the pattern of the data.

4.1.3 Regression Equations

Correlation and Regression Analysis

- The correlation coefficient answers the following questions:
  - Are two variables correlated?
  - How strong is the relationship between two variables?
  - When one variable increases, does the other variable increase or decrease?
- The correlation coefficient *cannot* address the following questions:
  - How much does one variable changes when the other variable changes by one unit?
  - How can we set the value of one variable to obtain a targeted value of the other variable?
  - How can we use the relationship between two variables to make predictions?
- The simple linear regression analysis helps to answer these questions.

What is Simple Linear Regression?

- **Simple linear regression** is a statistical technique to fit a straight line through the data points.
- It models the quantitative relationship between two variables.
- It describes how one variable changes according to the change of another variable.
- Both variables need to be continuous.
- It is simple because only one predictor variable is involved.

Simple Linear Regression Equation

- The simple linear regression analysis fits the data to a regression equation in the form

$$Y = \alpha \times X + \beta + e$$

where:

- Y is the dependent variable (the response) and X is the single independent variable (the predictor).
- α is the slope describing the steepness of the fitting line. β is the intercept indicating the Y value when X is equal to 0.
- e stands for error (residual). It is the difference between the actual Y and the fitted Y (i.e., the vertical difference between the data point and the fitting line).

Ordinary Least Squares

- The **ordinary least square** is a statistical method used in linear regression analysis to find the best fitting line for the data points.

- It estimates the unknown parameters of the regression equation by minimizing the sum of squared residuals (i.e. the vertical difference between the data point and the fitting line).

- In mathematical language, we look for α and β that satisfy the following criteria:

$$\min_{\alpha,\beta} Q(\alpha,\beta) \text{ where } Q(\alpha,\beta)=\sum_{i=1}^{n} e_i^{\,2}=\sum_{i=1}^{n}(y_i-\alpha-\beta x_i)^2$$

Ordinary Least Squares

- The actual value of the dependent variable:

$$Y_i = \alpha * X_i + \beta + e_i$$ where *i* = 1, 2,…, n

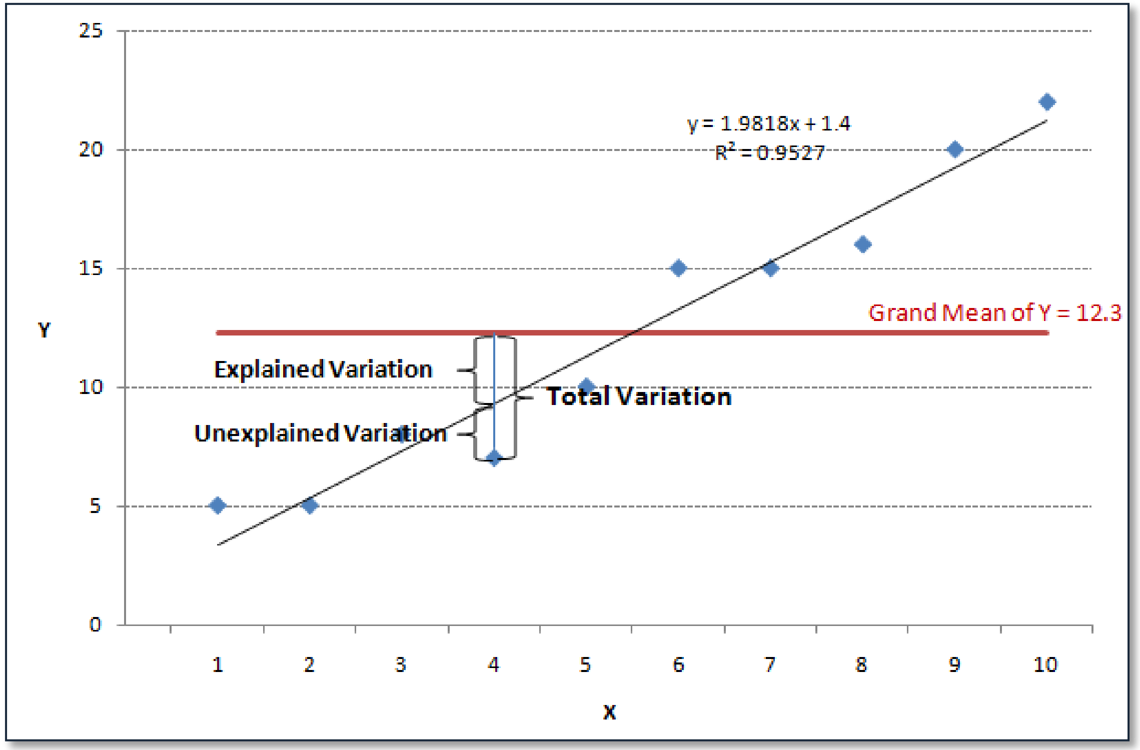
- The fitted value of the dependant variable:

$$\hat{Y}_i = \alpha * X_i + \beta$$ where *i* = 1, 2,…, n

- By using calculus, it can be shown the sum of squared error is minimal when

$$\beta = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \qquad \alpha = \bar{Y} - \beta \bar{X}$$

ANOVA in Simple Linear Regression



X: the independent variable that we use to predict;
Y: the dependent variable that we want to predict.

X	Y
1	5
2	5
3	8
4	7
5	10
6	15
7	15
8	16
9	20
10	22

Total Variation = Total Sums of Squares = $\sum_{i=1}^{n}(Y_i - \bar{Y})^2$

Explained Variation = Regression Sums of Squares = $\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$

Unexplained Variation = Error Sums of Squares = $\sum_{i=1}^{n}(Y_i - \hat{Y})^2$

ANOVA in Simple Linear Regression

- Linear regression is also analysis of variance (ANOVA).
  - Variation Components:
    - Total Variation = Explained Variation + Unexplained Variation

      i.e., Total Sums of Squares = Regression Sums of Squares + Error Sums of Squares
  - Degrees of Freedom Components
    - Total Degrees of Freedom = Regression Degrees of Freedom + Residual Degrees of Freedom

      i.e., n – 1 = (k – 1) + (n – k), where n is the number of data points, k is the number of predictors

ANOVA in Simple Linear Regression

- Whether the overall model is statistically significant can be tested by using F-test of ANOVA.
  - H_0: The model is not statistically significant.
  - H_a: The model is statistically significant.
- Test Statistic:

$$F = \frac{MSR}{MSE} = \frac{\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2 \Big/ (k-1)}{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \Big/ (n-k)}$$

- Critical Statistic: F value in F table with (k – 1) degrees of freedom in the numerator and (n – k) degrees of freedom in the denominator.

- If $F \leq F_{critical}$, we fail to reject the null. There is no statistically significant relationship between X and Y.
- If $F > F_{critical}$, we reject the null. There is a statistically significant relationship between X and Y.

Coefficient of Determination

- R^2 (also called coefficient of determination) measures the proportion of variability in the data that can be explained by the model.

- R^2 ranges from 0 to 1. The higher R^2 is, the better the model can fit the actual data.

- How to calculate R^2:

$$R^2 = \frac{SS_{regression}}{SS_{total}} = 1 - \frac{SS_{error}}{SS_{total}} = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

Use Minitab to Run a Simple Linear Regression

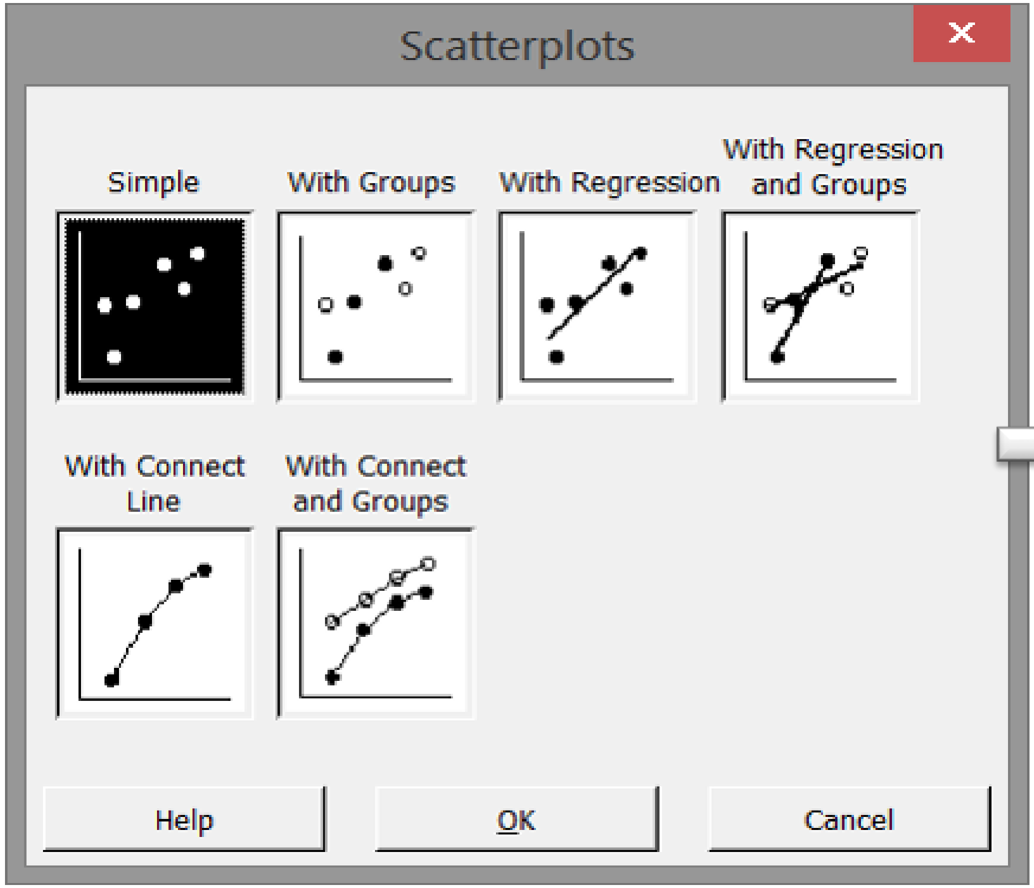
- *Case study:*
  - We want to see whether the score on exam one has any statistically significant relationship with the score on the final exam. If yes, how much impact does exam one have on the final exam?
  - Data File: “Simple Linear Regression” tab in “Sample Data.xlsx”

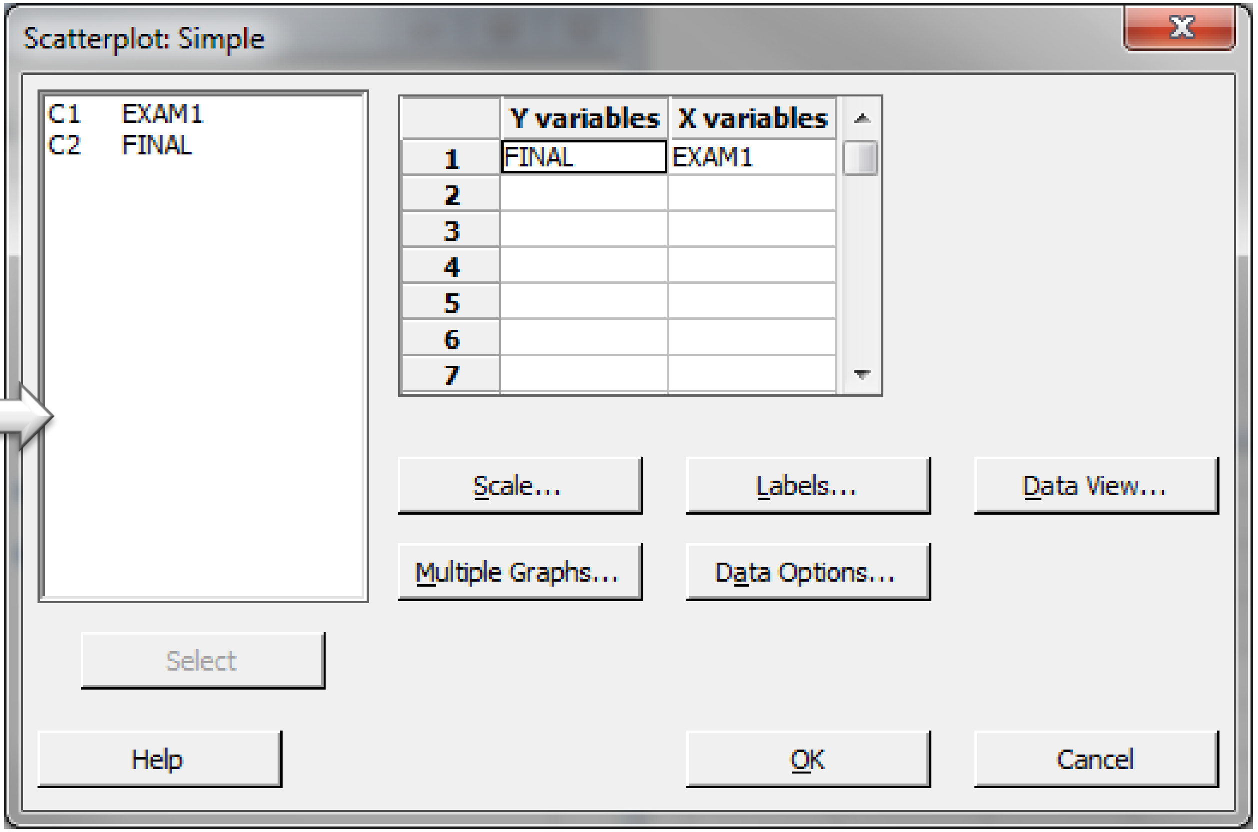
- Step 1: Determine the dependent and independent variables. Both should be continuous variables.
  - Y (dependent variable) is the score of final exam.
  - X (independent variable) is the score of exam one.
  - Both variables are continuous.

Use Minitab to Run a Simple Linear Regression

- Step 2: Create a scatter plot to visualize whether there seems to be a linear relationship between X and Y.
  1) Click Graph → Scatterplot.
  2) A new window named "Scatterplots" pops up.
  3) Click "OK."
  4) A new window named "Scatterplot - Simple" pops up.
  5) Select "FINAL" as "Y variables" and "EXAM1" as "X variables."
  6) Click "OK."
  7) A scatter plot is generated in a new window.

Use Minitab to Run a Simple Linear Regression



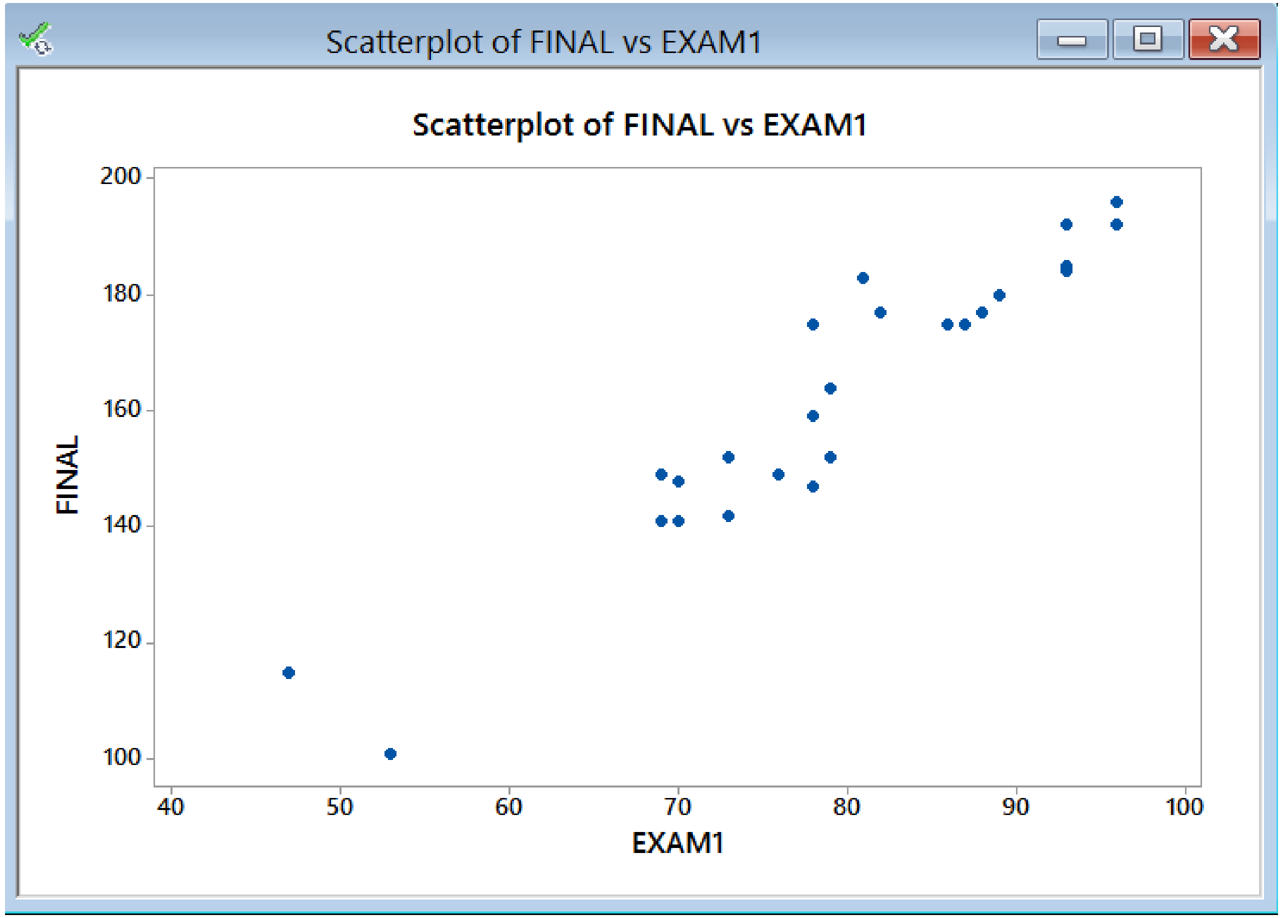


C1 EXAM1
C2 FINAL

	Y variables	X variables
1	FINAL	EXAM1
2		
3		
4		
5		
6		
7		

Use Minitab to Run a Simple Linear Regression

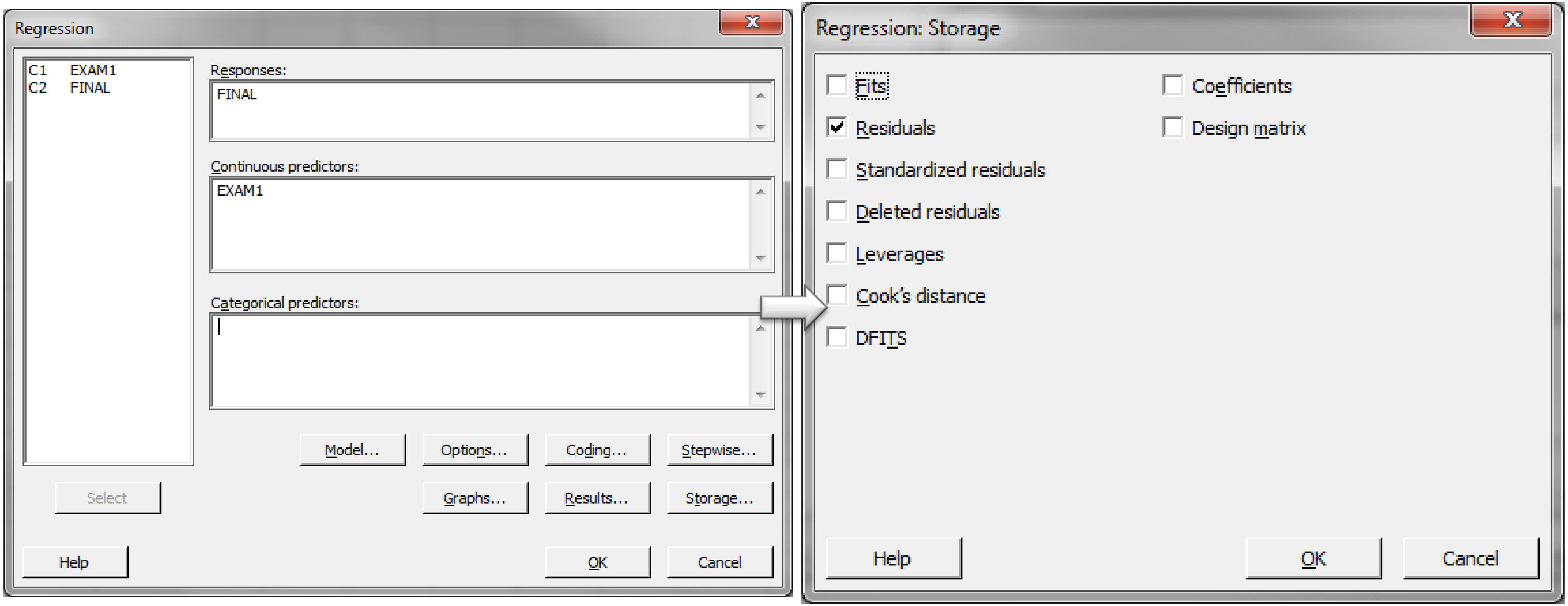
- Based on the scatter plot, the relationship between exam one and final seems linear. The higher the score on exam one, the higher the score on the final.



Use Minitab to Run a Simple Linear Regression

- Step 3: Run the simple linear regression analysis.
  1) Click Stat → Regression → Regression → Fit Regression Model
  2) A new window named “Regression” pops up.
  3) Select “FINAL” as “Response” and “EXAM1” as “Continuous Predictors.”
  4) Click the “Storage” button.
  5) Check the box of “Residuals” so that the residuals can be saved automatically in the last column of the data table.
  6) Click “OK.”
  7) The regression analysis results appear in the new window.

Use Minitab to Run a Simple Linear Regression



Regression

C1 EXAM1
C2 FINAL

Responses: FINAL

Continuous predictors: EXAM1

Categorical predictors:

Regression: Storage

☐ Fits
☑ Residuals
☐ Standardized residuals
☐ Deleted residuals
☐ Leverages
☐ Cook's distance
☐ DFITS
☐ Coefficients
☐ Design matrix

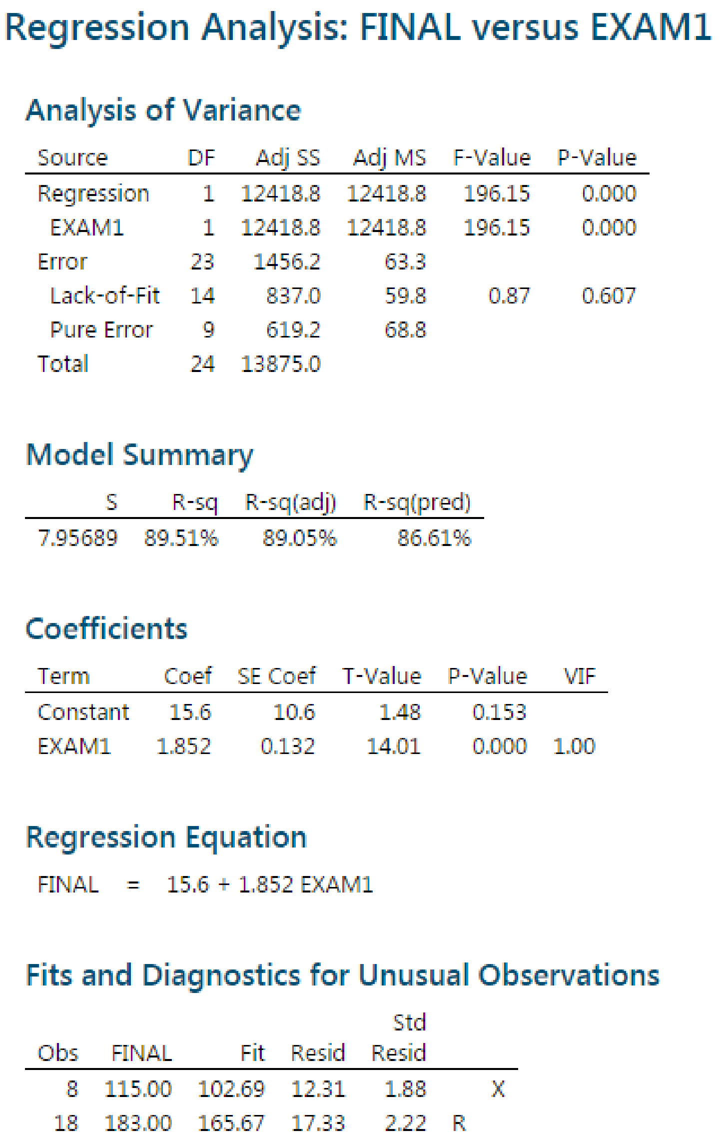
Use Minitab to Run a Simple Linear Regression

- Step 4: Check whether the model is statistically significant. If not significant, we will need to re-examine the predictor or look for new predictors before continuing.

- R2 measures the percentage of variation in the data set that can be explained by the model. 89.5% of the variability in the data can be accounted for by this linear regression model.

- “Analysis of Variance” section provides a ANOVA table covering degrees of freedom, sum of squares, and mean square information for total, regression and error.

- The p-value of the F-test is lower than the α level (0.05), indicating that the model is statistically significant.



Regression Analysis: FINAL versus EXAM1

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	12418.8	12418.8	196.15	0.000
EXAM1	1	12418.8	12418.8	196.15	0.000
Error	23	1456.2	63.3		
Lack-of-Fit	14	837.0	59.8	0.87	0.607
Pure Error	9	619.2	68.8		
Total	24	13875.0			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
7.95689	89.51%	89.05%	86.61%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	15.6	10.6	1.48	0.153	
EXAM1	1.852	0.132	14.01	0.000	1.00

Regression Equation

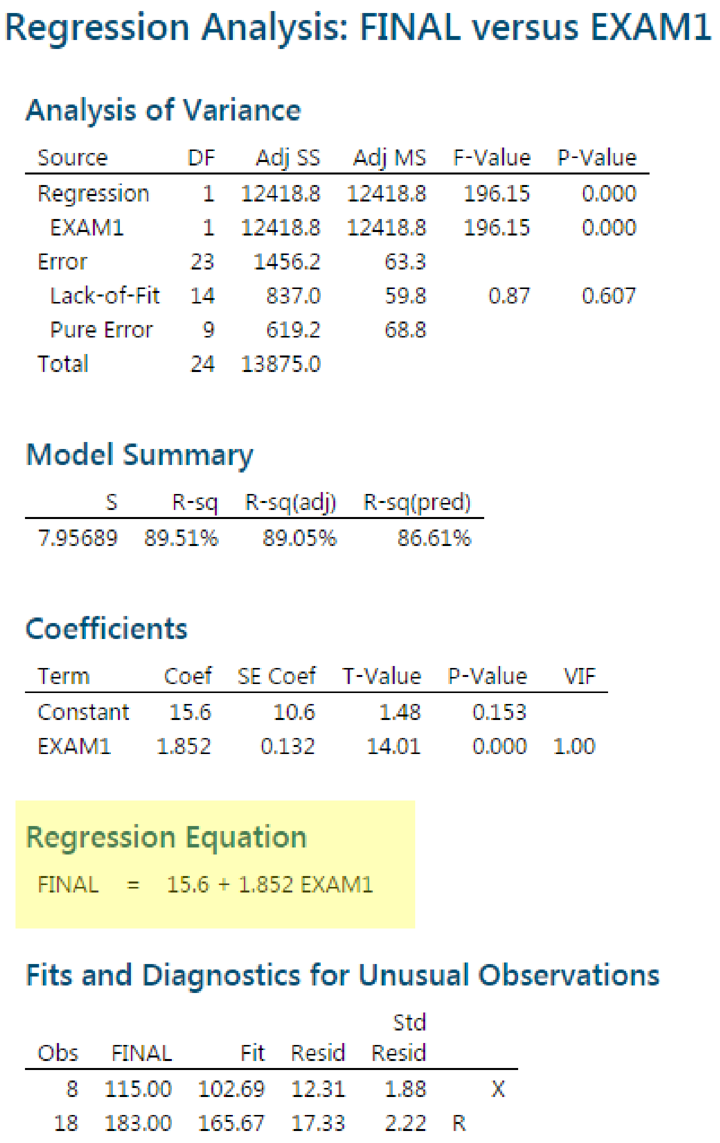
FINAL = 15.6 + 1.852 EXAM1

Fits and Diagnostics for Unusual Observations

Obs	FINAL	Fit	Resid	Std Resid	
8	115.00	102.69	12.31	1.88	X
18	183.00	165.67	17.33	2.22	R

Use Minitab to Run a Simple Linear Regression

- Step 5: Understand regression equation
  - The estimates of slope and intercept are shown in the “Parameter Estimate” section.
  - In this example, Y = 15.6 + 1.85 × X.
  - One unit increase in the score of Exam1 would increase the final score by 1.85.



```
Regression Analysis: FINAL versus EXAM1

Analysis of Variance

Source        DF   Adj SS   Adj MS  F-Value  P-Value
Regression     1  12418.8  12418.8   196.15    0.000
  EXAM1        1  12418.8  12418.8   196.15    0.000
Error         23   1456.2     63.3
  Lack-of-Fit 14    837.0     59.8     0.87    0.607
  Pure Error   9    619.2     68.8
Total         24  13875.0

Model Summary

      S    R-sq  R-sq(adj)  R-sq(pred)
7.95689  89.51%     89.05%      86.61%

Coefficients

Term       Coef  SE Coef  T-Value  P-Value   VIF
Constant   15.6     10.6     1.48    0.153
EXAM1     1.852    0.132    14.01    0.000  1.00

Regression Equation

FINAL  =  15.6 + 1.852 EXAM1

Fits and Diagnostics for Unusual Observations

                            Std
Obs   FINAL     Fit  Resid  Resid
  8  115.00  102.69  12.31   1.88     X
 18  183.00  165.67  17.33   2.22  R
```

Interpreting the Results

- Rsquare Adj = 89.05%
  - 89% of the variation in FINAL can be explained by EXAM1

- P-value of the F-test = 0.000
  - We have a statistically significant model

- Prediction Equation: 15.6 + 1.85 × EXAM1
  - 15.6 is the Y intercept, all equations will start with 15.6
  - 1.85 is the EXAM1 Coefficient: multiply it by EXAM1 score

- Let us say you are the professor and you want to use this prediction equation to estimate what two of your students might get on their final exam.

Interpreting the Results

- Let us assume the following:
  - Student “A” exam 1 results were: 79
  - Student “B” exam 1 results were: 94.

- Remember our prediction equation?
  - 15.6 + 1.85 × Exam1
  - Now apply the equation to each student
    - Student “A” Estimate: 15.6 + (1.85 × 79) = 161.8
    - Student “B” Estimate: 15.6 + (1.85 × 94) = 189.5

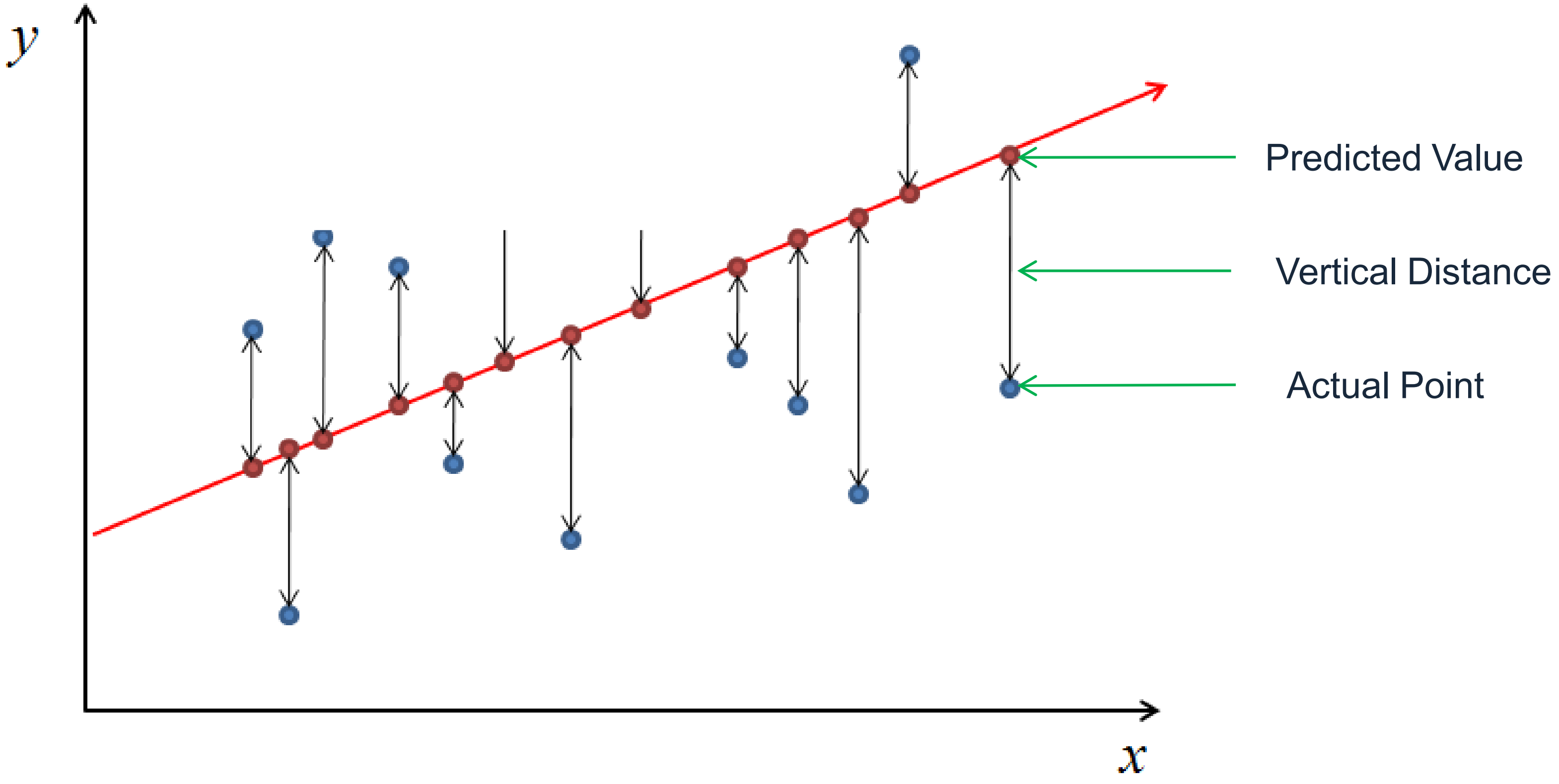
Now you can use your “magic” as the professor and allocate your time appropriately to the student(s) who you predict need the most help

Nice Work!

4.1.4 Residuals Analysis

What are Residuals?

- **Residuals** are the vertical differences between actual values and the predicted values or the “fitted line” created by the regression model.



Why Perform Residuals Analysis?

- *Regression equations* are generated on the basis of certain statistical assumptions.

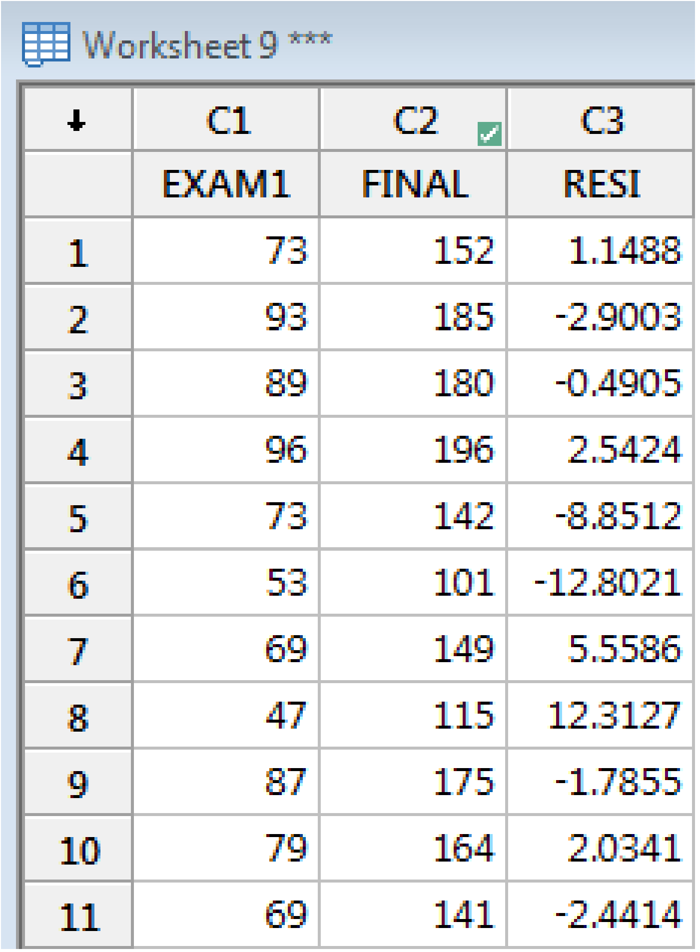
- *Residuals analysis* helps to determine the validity of these assumptions.

- The *assumptions* are:
  - The residuals are normally distributed, mean equal to zero.
  - The residuals are independent.
  - The residuals have a constant variance.
  - The underlying population relationship is linear.

- If residuals performance does not meet the requirements, we will need to rebuild the model by replacing the predictor with a new one, adding new predictors, building non-linear models, and so on.

Use Minitab to Perform Residuals Analysis

- The residuals of the model are saved in a column of your data table



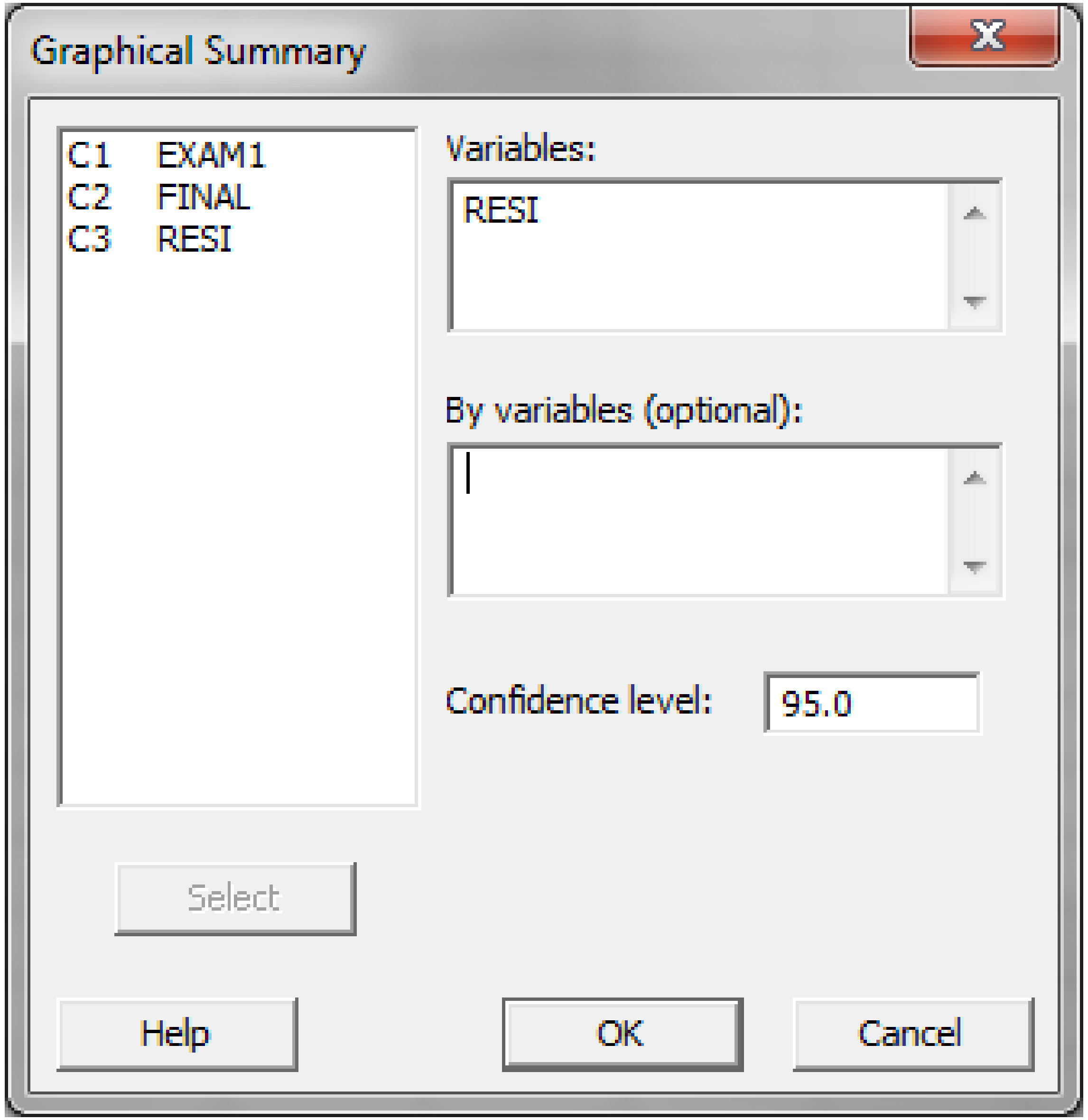
Worksheet 9 ***

↓	C1	C2	C3
	EXAM1	FINAL	RESI
1	73	152	1.1488
2	93	185	-2.9003
3	89	180	-0.4905
4	96	196	2.5424
5	73	142	-8.8512
6	53	101	-12.8021
7	69	149	5.5586
8	47	115	12.3127
9	87	175	-1.7855
10	79	164	2.0341
11	69	141	-2.4414

Use Minitab to Perform Residuals Analysis

- Step 1: Check whether residuals are normally distributed around the mean of zero.
  1) Click Stat → Basic Statistics → Graphical Summary.
  2) A new window named “Graphical Summary” pops up.
  3) Select “RESI” as “Variables.”
  4) Click “OK.”
  5) The histogram and the normality test of the residuals are displayed in the new window.

Use Minitab to Perform Residuals Analysis



Graphical Summary

C1 EXAM1
C2 FINAL
C3 RESI

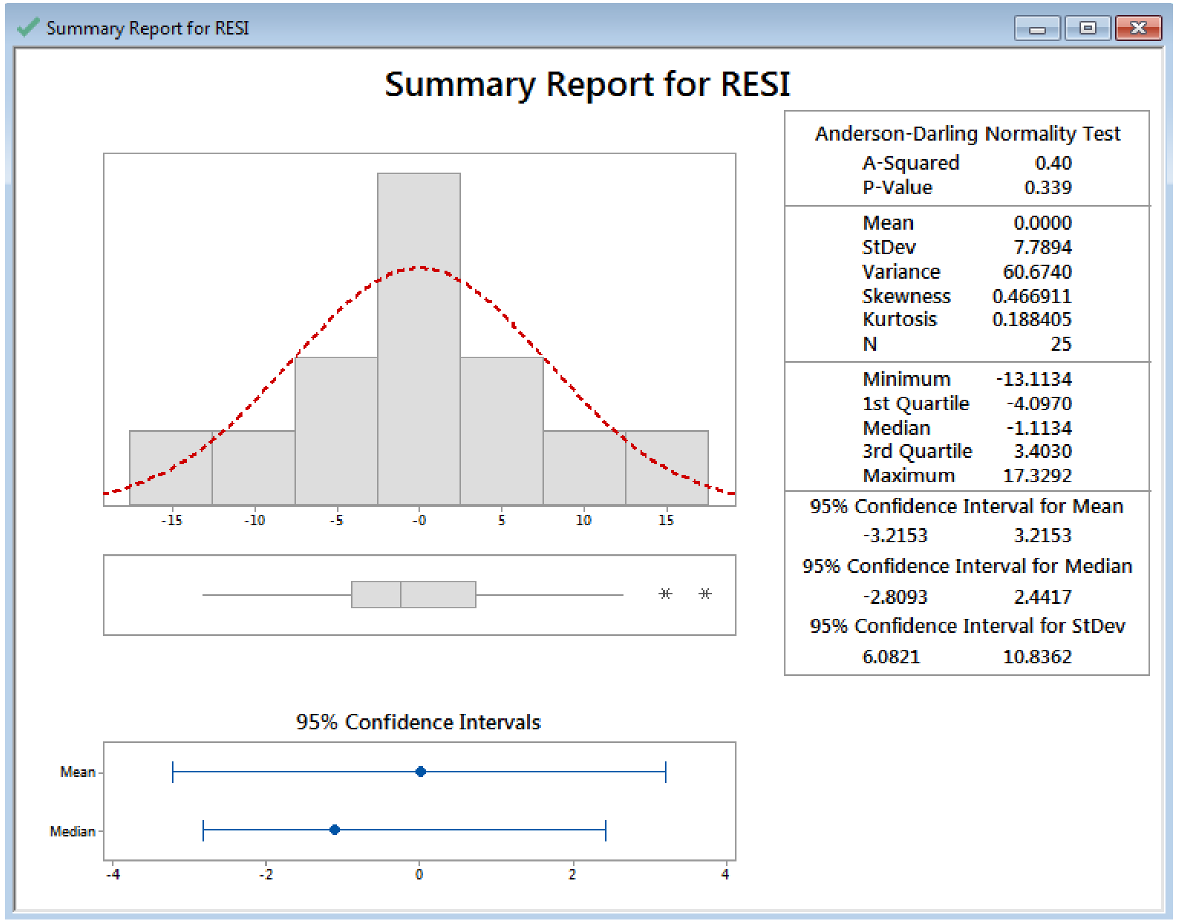
Variables: RESI

By variables (optional):

Confidence level: 95.0

Use Minitab to Perform Residuals Analysis

- The mean of residuals is -0.0000.
- The Anderson-Darling test is used to test the normality. Since the p-value (0.339) is greater than the alpha level (0.05), we fail to reject the null hypothesis; the residuals are normally distributed.
  - H_0: The residuals are normally distributed.
  - H_1: The residuals are not normally distributed.



Summary Report for RESI

Anderson-Darling Normality Test

A-Squared	0.40
P-Value	0.339
Mean	0.0000
StDev	7.7894
Variance	60.6740
Skewness	0.466911
Kurtosis	0.188405
N	25
Minimum	-13.1134
1st Quartile	-4.0970
Median	-1.1134
3rd Quartile	3.4030
Maximum	17.3292

95% Confidence Interval for Mean

-3.2153 3.2153

95% Confidence Interval for Median

-2.8093 2.4417

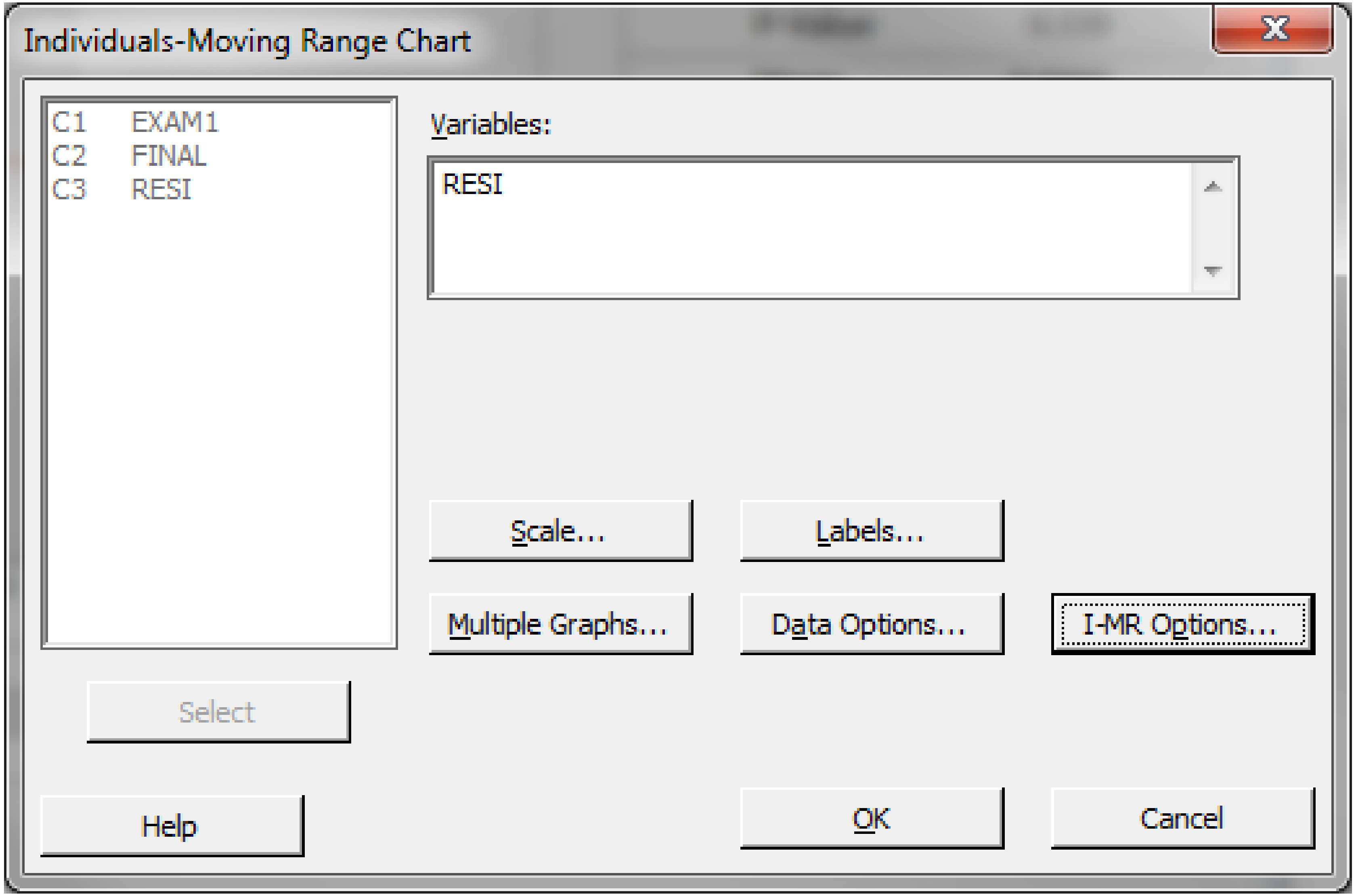
95% Confidence Interval for StDev

6.0821 10.8362

Use Minitab to Perform Residuals Analysis

- Step 2: If the data are in time order, run the I-MR chart to check whether residuals are independent.
  1) Click Stat → Control Charts → Variable Charts for Individuals → I-MR.
  2) A new window named “Individuals – Moving Range Chart” pops up.
  3) Select “Residuals” as “Variables” and check the box of “Test for special causes.”
  4) Click “OK.”
  5) The control charts are shown automatically in the new window.

Use Minitab to Perform Residuals Analysis



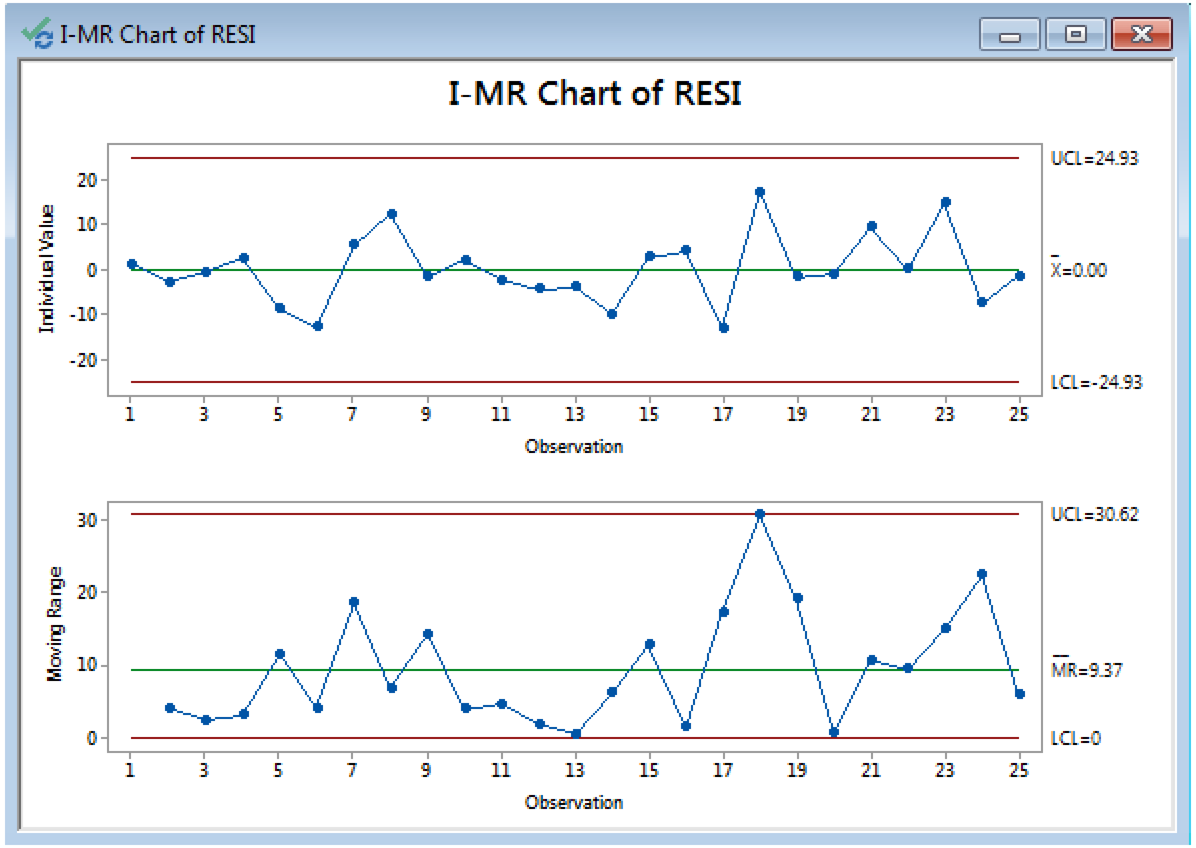
RESI

Use Minitab to Perform Residuals Analysis

- If no data points are out of control in both the I-chart and MR chart, the residuals are independent of each other.

- If the residuals are not independent, it is possible that some important predictors are not included in the model.

- In this example, since the I-MR chart is in control, residuals are independent.



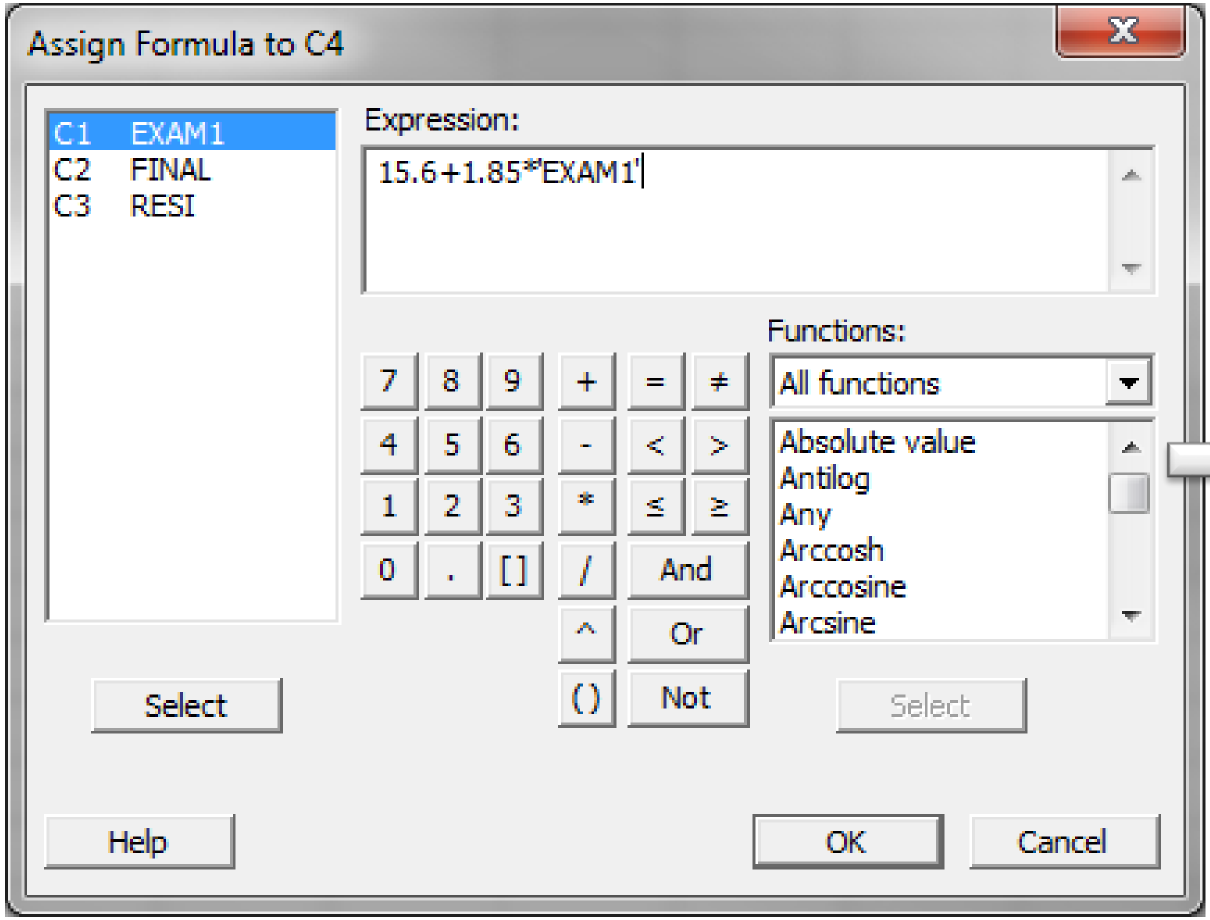
Use Minitab to Perform Residuals Analysis

- Step 3: Check whether residuals have equal variance across the predicted responses.
  - Create a column of the fitted response using the regression formula returned in the session window.
    1) Right click on the last column in the data table.
    2) Select Formulas → Assign Formula to Column.
    3) A new window named "Assign Formula to C4" appears.
    4) Enter the regression equation "15.6+1.85*EXAM1" into the "Expression" box.
    5) Click "OK."
    6) The column of the fitted values is created in the data table.

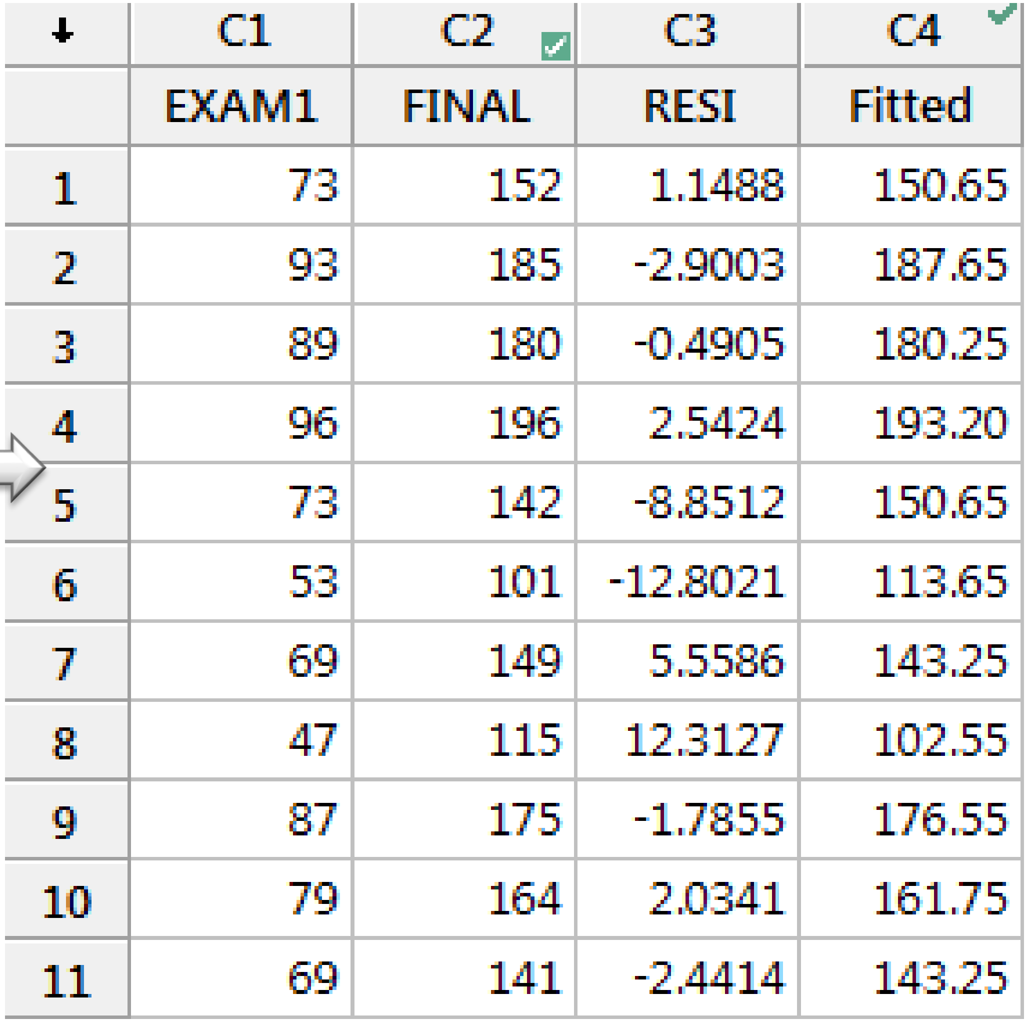
  - Create a scatter plot with Y being the residuals and the X being the fitted values.
    1) Click Graph → Scatterplot.
    2) A new window named "Scatterplots" appears.
    3) Select "RESI" as the "Y variables" and "Fitted" as the "X variables."
    4) Click "OK."
    5) The scatter plot appears in a new window.

  - We are looking for the pattern in which residuals spread out evenly around zero from the top to the bottom.

Use Minitab to Perform Residuals Analysis

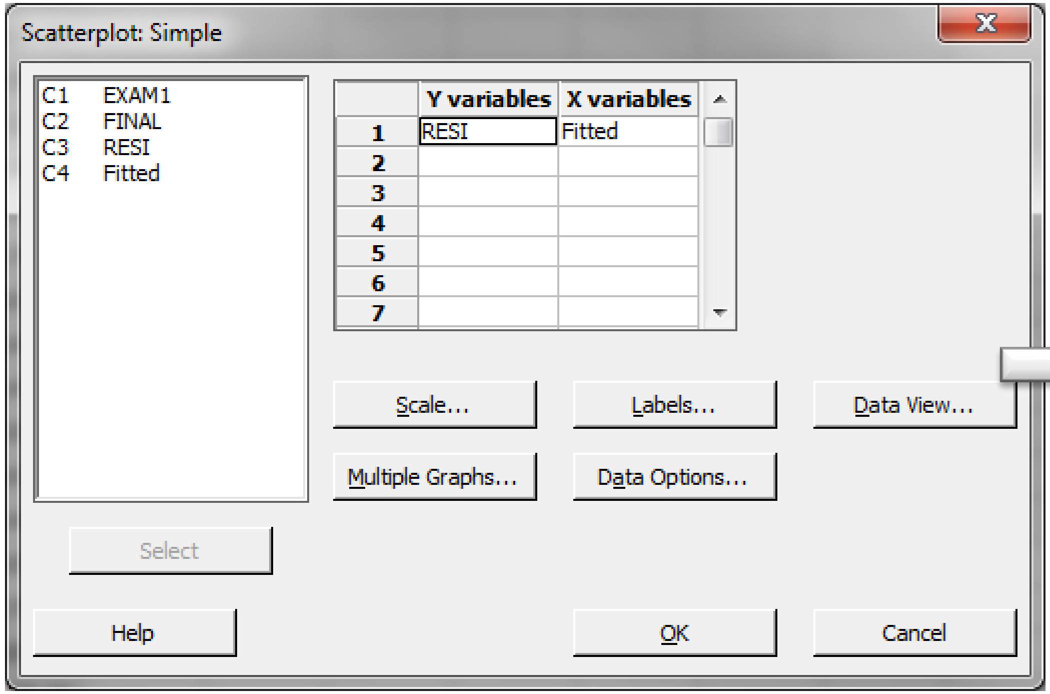


Expression: 15.6+1.85*'EXAM1'



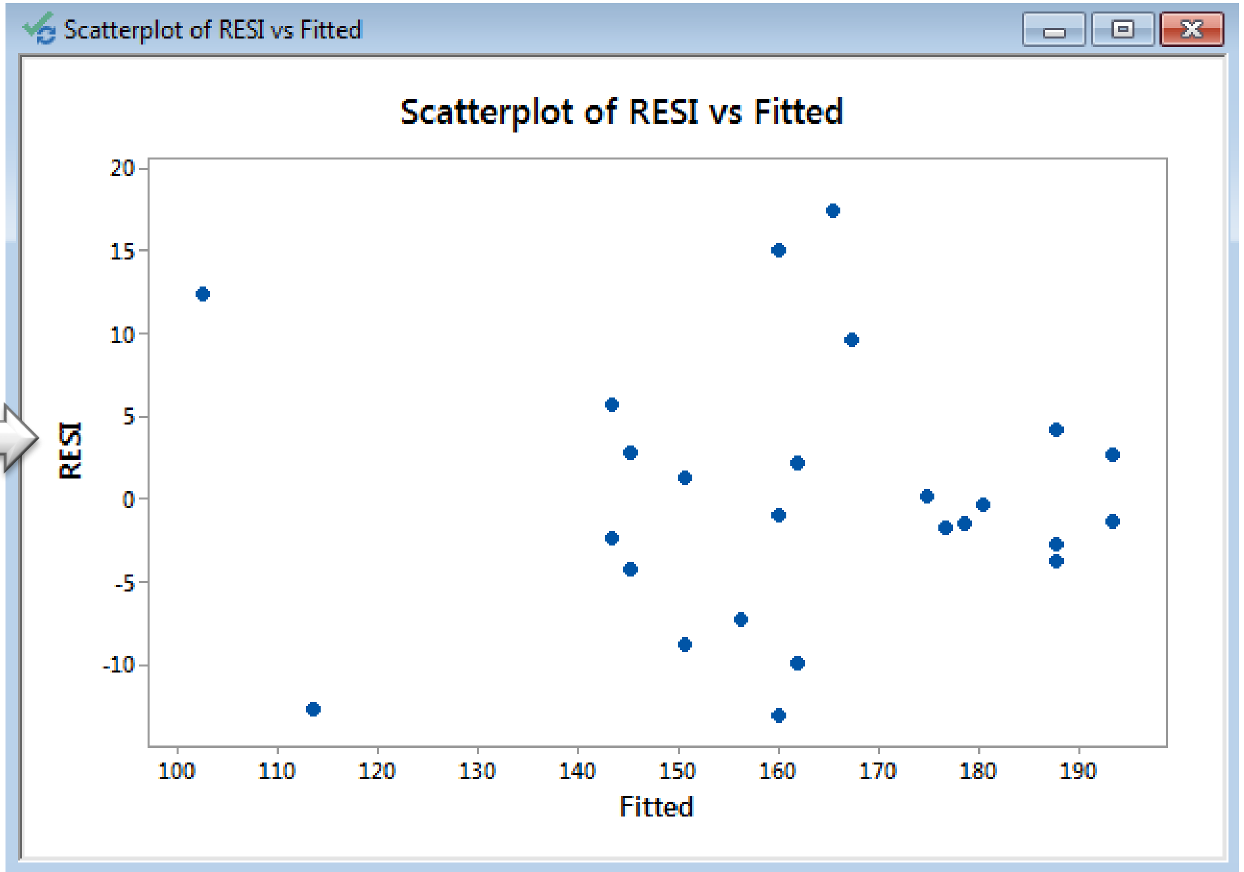
↓	C1	C2	C3	C4
	EXAM1	FINAL	RESI	Fitted
1	73	152	1.1488	150.65
2	93	185	-2.9003	187.65
3	89	180	-0.4905	180.25
4	96	196	2.5424	193.20
5	73	142	-8.8512	150.65
6	53	101	-12.8021	113.65
7	69	149	5.5586	143.25
8	47	115	12.3127	102.55
9	87	175	-1.7855	176.55
10	79	164	2.0341	161.75
11	69	141	-2.4414	143.25

Use Minitab to Perform Residuals Analysis


Scatterplot: Simple

C1 EXAM1
C2 FINAL
C3 RESI
C4 Fitted

	Y variables	X variables
1	RESI	Fitted
2		
3		
4		
5		
6		
7		



4.2 Multiple Regression Analysis

Black Belt Training: Improve Phase

4.2.1 Non-Linear Regression

Linear and Non-Linear

- The word *linear* originally comes from Latin word *linearis* meaning "created by lines."

- A linear function in mathematics follows the following pattern (i.e., the output is proportional to its input):

$$f(x) = \alpha * x + \beta$$

$$f(x_1, x_2, \ldots, x_n) = \alpha_1 * x_1 + \alpha_2 * x_2 + \ldots + \alpha_n * x_n + \beta$$

- A non-linear function does not follow the above pattern. There are usually exponents, logarithms, power, polynomial components, and other non-linear functions of the independent variables and parameters.

Non-Linear Relationships Using Linear Models

- Many non-linear relationships can be transformed into linear relationships, and from there we can use linear regression methods to model the relationship.

- Some non-linear relationships cannot be transformed to linear ones and we need to apply other methods to build the non-linear models.

- In this section, we will focus on building non-linear regression models using linear transformation (i.e., transforming the independent or dependent variables or parameters to generate a linear function).

Assumptions in Using Non-Linear Regression

- The population relationship is non-linear based on a reliable underlying theory.

- Across the range of all the possible values of the independent variables, the non-linear relationship applies. It is possible that at some extreme values the relationship between variables changes dramatically.

Non-Linear Functions: Transforming to Linear

- Examples of non-linear functions that can be transformed to linear functions:
  - Exponential Function
  - Inverse Function
  - Polynomial Function
  - Power Function.

Exponential Function

- Exponential Function

$$Y = a \times b^X$$

- Transformation

$$\log Y = \log a + X \times \log b$$

Inverse Function

- Inverse Function

$$Y = a + b \times \frac{1}{X}$$

- Transformation

$$Y = a + b \times Z \quad \textbf{where} \quad Z = \frac{1}{X}$$

Polynomial Function

- Polynomial Function

$$Y = a + b \times X + c \times X^2$$

- Transformation

$$Y = a + b \times X + c \times Z \quad \textbf{where} \quad Z = X^2$$

Power Function

- Power Function

$$Y = a \times X^b$$

- Transformation

$$\log Y = \log a + b \times \log X$$

4.2.2 Multiple Linear Regression

What is Multiple Linear Regression?

- **Multiple linear regression** is a statistical technique to model the relationship between one dependent variable and two or more independent variables by fitting the data set into a linear equation.

- The difference between simple linear regression and multiple linear regression:
  - Simple linear regression only has one predictor.
  - Multiple linear regression has two or more predictors.

Multiple Linear Regression Equation

$$Y = \alpha_1 * X_1 + \alpha_2 * X_2 + \ldots + \alpha_p * X_p + \beta + e$$

- Y is the dependent variable (response).
- X_1, X_2, … , X_p are the independent variables (predictors). There are p predictors in total.
- Both dependent and independent variables are continuous.
- β is the intercept indicating the Y value when all the predictors are zeros.
- α_1, α_2, … , α_p are the coefficients of predictors. They reflect the contribution of each independent variable in predicting the dependent variable.
- e is the residual term indicating the difference between the actual and the fitted response value.

Use Minitab to Run a Multiple Linear Regression

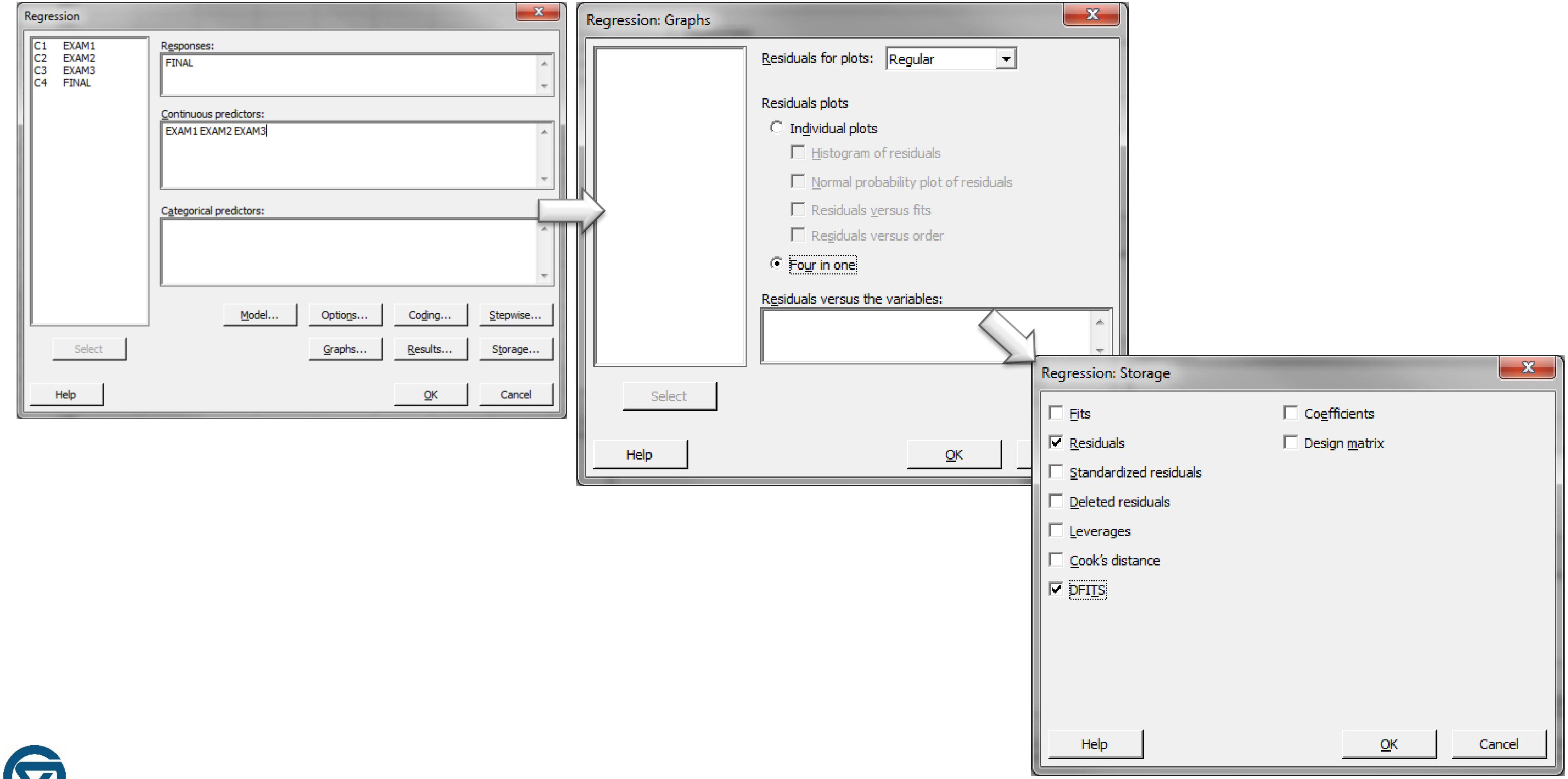
- *Case study:*
  - We want to see whether the scores in exam one, two, and three have any statistically significant relationship with the score in final exam. If so, how are they related to final exam score? Can we use the scores in exam one, two, and three to predict the score in final exam?
  - Data File: “Multiple Regression Analysis” tab in “Sample Data.xlsx.”

- Step 1: Determine the dependent and independent variables. All should be continuous.
  - Y (dependent variable) is the score of final exam.
  - X_1, X_2, and X_3 (independent variables) are the scores of exam one, two, and three respectively.
  - All the variables are continuous.

Use Minitab to Run a Multiple Linear Regression

- Step 2: Start building the multiple linear regression model
  1) Click Stat → Regression → Regression → Fit Regression Model
  2) A new window named “Regression” pops up.
  3) Select “FINAL” as “Response” and “EXAM1”, “EXAM2” and “EXAM3” as “Continuous Predictors.”
  4) Click the “Graph” button, select the radio button “Four in one” and click “OK.”
  5) Click the “Storage” button, check the boxes of “Residuals” and “DFITS” and click “OK.”
  6) Click “OK” in the window named “Regression.”
  7) The regression analysis results appear in the session window and the four residual plots appear in another window named “Residual Plots for FINAL.”

Use Minitab to Run a Multiple Linear Regression



Regression

C1 EXAM1
C2 EXAM2
C3 EXAM3
C4 FINAL

Responses: FINAL

Continuous predictors: EXAM1 EXAM2 EXAM3

Categorical predictors:

Regression: Graphs

Residuals for plots: Regular

Residuals plots

○ Individual plots

☐ Histogram of residuals

☐ Normal probability plot of residuals

☐ Residuals versus fits

☐ Residuals versus order

● Four in one

Residuals versus the variables:

Regression: Storage

☐ Fits

☑ Residuals

☐ Standardized residuals

☐ Deleted residuals

☐ Leverages

☐ Cook's distance

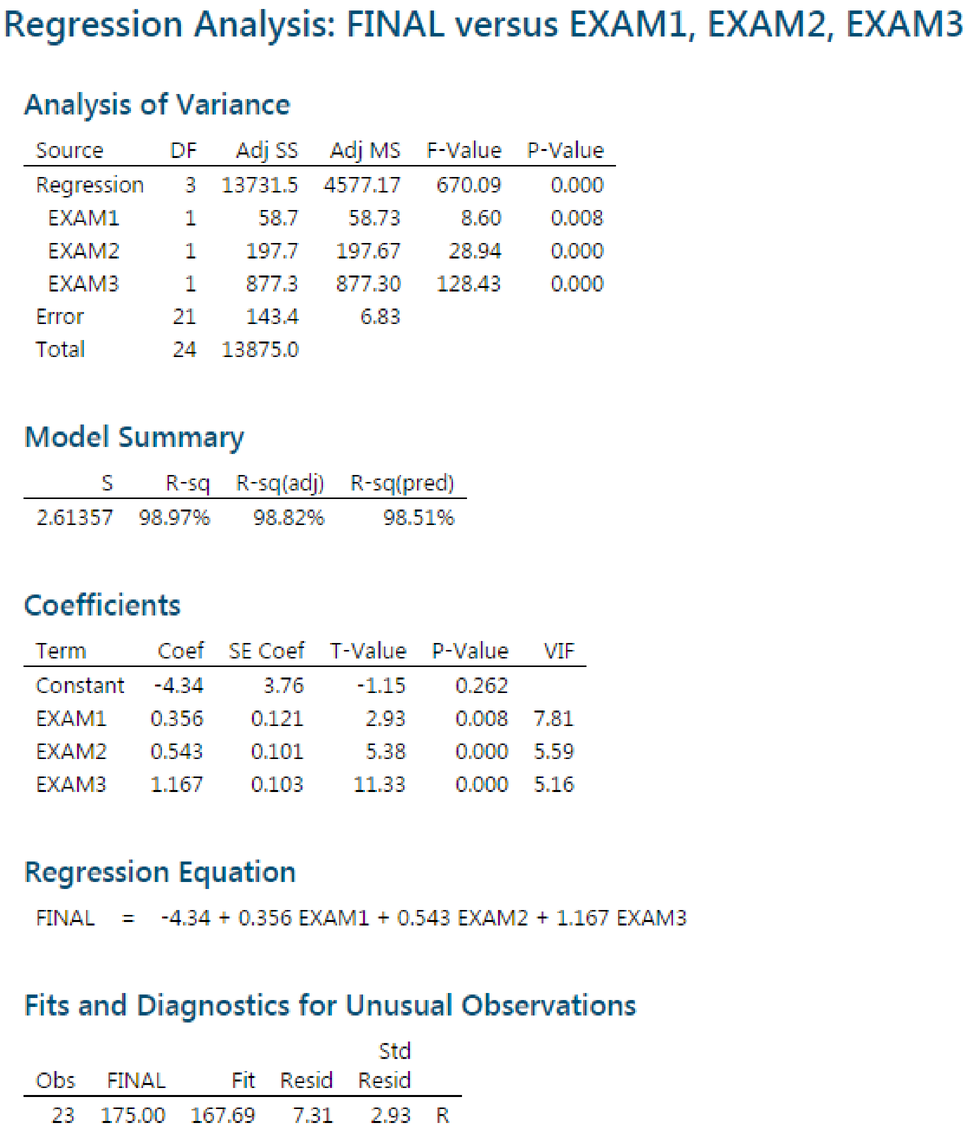
☑ DFITS

☐ Coefficients

☐ Design matrix

Use Minitab to Run a Multiple Linear Regression

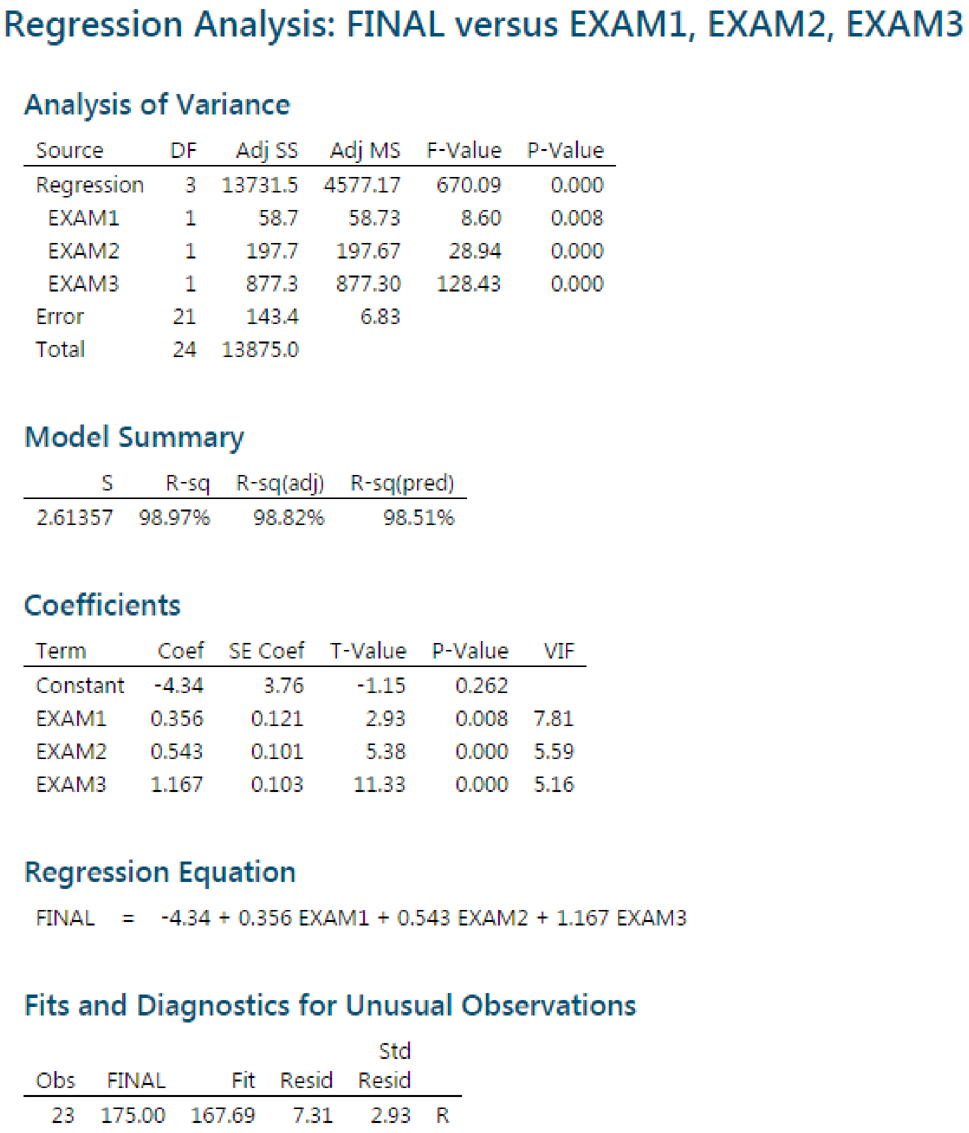
- Step 3: Check whether the whole model is statistically significant. If not, we need to re-examine the predictors or look for new predictors before continuing.
  - H_0: The model is not statistically significant (i.e., all the parameters of predictors are not significantly different from zeros).
  - H_1: The model is statistically significant (i.e., at least one predictor parameter is significantly different from zero).
  - In this example, p-value is much smaller than alpha level (0.05), hence we reject the null hypothesis; the model is statistically significant.



```
Regression Analysis: FINAL versus EXAM1, EXAM2, EXAM3

Analysis of Variance

Source      DF    Adj SS   Adj MS  F-Value  P-Value
Regression   3   13731.5  4577.17   670.09    0.000
  EXAM1      1      58.7    58.73     8.60    0.008
  EXAM2      1     197.7   197.67    28.94    0.000
  EXAM3      1     877.3   877.30   128.43    0.000
Error       21     143.4     6.83
Total       24   13875.0


Model Summary

      S    R-sq  R-sq(adj)  R-sq(pred)
2.61357  98.97%     98.82%      98.51%


Coefficients

Term       Coef  SE Coef  T-Value  P-Value   VIF
Constant  -4.34     3.76    -1.15    0.262
EXAM1     0.356    0.121     2.93    0.008  7.81
EXAM2     0.543    0.101     5.38    0.000  5.59
EXAM3     1.167    0.103    11.33    0.000  5.16


Regression Equation

FINAL  =  -4.34 + 0.356 EXAM1 + 0.543 EXAM2 + 1.167 EXAM3


Fits and Diagnostics for Unusual Observations

                                Std
Obs   FINAL     Fit  Resid    Resid
 23  175.00  167.69   7.31     2.93  R
```

Use Minitab to Run a Multiple Linear Regression

- Step 4: Check whether multicollinearity exists in the model.
  - The VIF information is automatically generated in the table Coefficients.



```
Regression Analysis: FINAL versus EXAM1, EXAM2, EXAM3

Analysis of Variance

Source       DF    Adj SS    Adj MS   F-Value   P-Value
Regression    3   13731.5   4577.17    670.09     0.000
  EXAM1       1      58.7     58.73      8.60     0.008
  EXAM2       1     197.7    197.67     28.94     0.000
  EXAM3       1     877.3    877.30    128.43     0.000
Error        21     143.4      6.83
Total        24   13875.0

Model Summary

      S     R-sq   R-sq(adj)   R-sq(pred)
2.61357   98.97%      98.82%       98.51%

Coefficients

Term        Coef   SE Coef   T-Value   P-Value    VIF
Constant   -4.34      3.76     -1.15     0.262
EXAM1      0.356     0.121      2.93     0.008   7.81
EXAM2      0.543     0.101      5.38     0.000   5.59
EXAM3      1.167     0.103     11.33     0.000   5.16

Regression Equation

FINAL  =  -4.34 + 0.356 EXAM1 + 0.543 EXAM2 + 1.167 EXAM3

Fits and Diagnostics for Unusual Observations

                                  Std
Obs    FINAL      Fit   Resid   Resid
 23   175.00   167.69    7.31    2.93  R
```

Multicollinearity

- **Multicollinearity** is the situation when two or more independent variables in a multiple regression model are correlated with each other.
- Although multicollinearity does not necessarily reduce the predictability for the model as a whole, it may mislead the calculation for individual independent variables.
- To detect multicollinearity, we use VIF (Variance Inflation Factor) to quantify its severity in the model.

Variance Inflation Factor

- VIF quantifies the degree of multicollinearity for each individual independent variable in the model.
- VIF calculation:
  - Assume we are building a multiple linear regression model using p predictors.

$$Y = \alpha_1 \times X_1 + \alpha_2 \times X_2 + \ldots + \alpha_p \times X_p + \beta$$

  - Two steps are needed to calculate VIF for X_1.
    - Step 1: Build a multiple linear regression model for X_1 by using X_2, X_3, …, X_p as predictors.

$$X_1 = a_2 \times X_2 + a_3 \times X_3 + \ldots + a_p \times X_p + b$$

    - Step 2: Use the R^2 generated by the linear model in step 1 to calculate the VIF for X_1.

$$VIF = \frac{1}{1 - R^2}$$

    - Apply the same methods to obtain the VIFs for other Xs. The VIF value ranges from one to positive infinity.

Variance Inflation Factor

- Rules of thumb to analyze variance inflation factor (VIF):
  - If VIF = 1, there is no multicollinearity.
  - If 1 < VIF < 5, there is small multicollinearity.
  - If VIF ≥ 5, there is medium multicollinearity.
  - If VIF ≥ 10, there is large multicollinearity.

How to Deal With Multicollinearity

- Increase the sample size.
- Collect samples with a broader range for some predictors.
- Remove the variable with high multicollinearity and high p-value.
- Remove variables that are included more than once.
- Combine correlated variables to create a new one.
- In this section, we will focus on removing variables with high VIF and high p-value.

Use Minitab to Run a Multiple Linear Regression

- Step 5: Deal with multicollinearity:
  - Step 5.1: Identify a list of independent variables with VIF higher than 5. If no variable has VIF higher than 5, go to Step 6 directly.
  - Step 5.2: Among variables identified in Step 5.1, remove the one with the highest p-value.
  - Step 5.3: Run the model again, check the VIFs and repeat Step 5.1.
  - Note: we only remove one independent variable at a time.

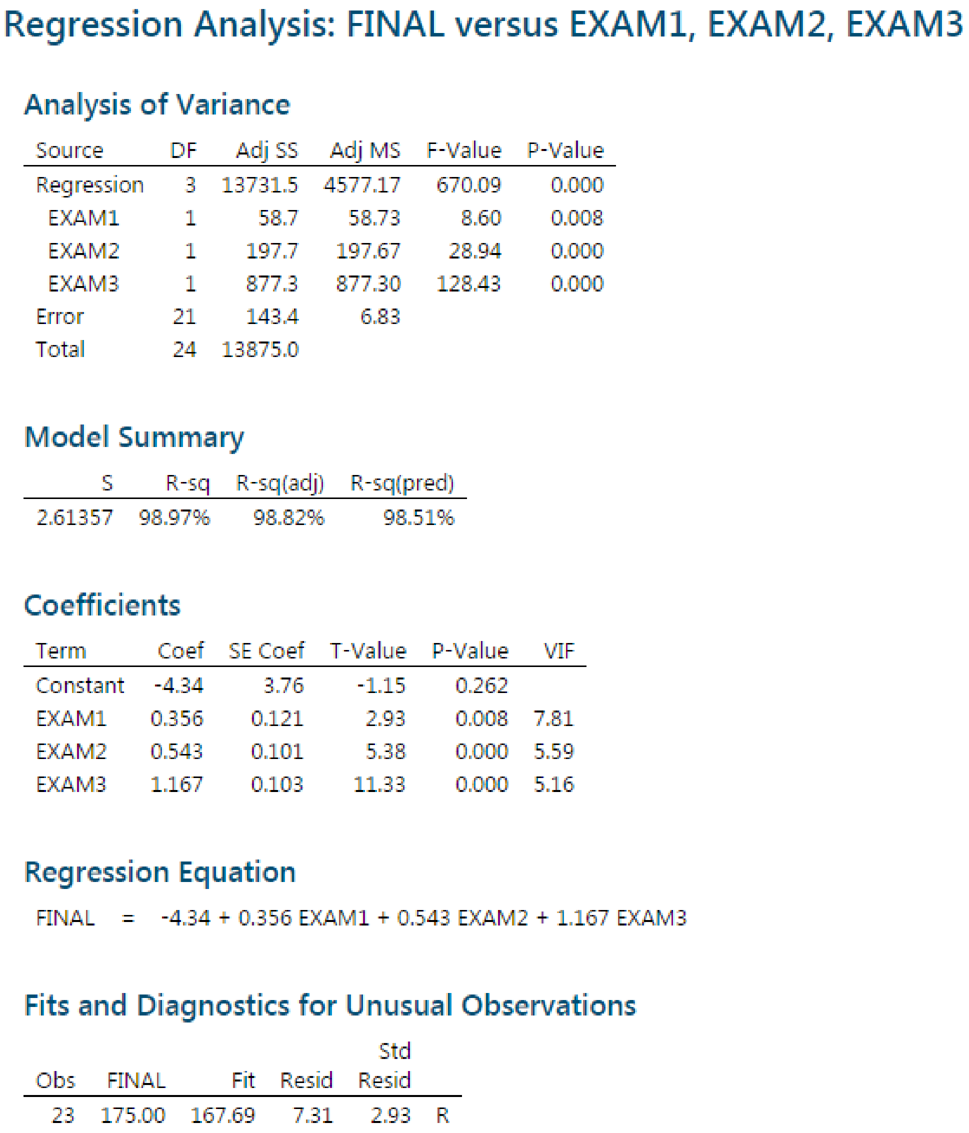
Use Minitab to Run a Multiple Linear Regression

- Step 6: Identify the statistically insignificant predictors. Remove one insignificant predictor at a time and run the model again. Repeat this step until all the predictors in the model are statistically significant.
  - Insignificant predictors are the ones with p-value higher than alpha level (0.05). When p > alpha level, we fail to reject the null hypothesis; the predictor is not significant.
    - H_0: The predictor is not statistically significant.
    - H_1: The predictor is statistically significant.
- As long as the p-value is greater than 0.05, remove the insignificant variables one at a time in the order of the highest p-value.
- Once one insignificant variable is eliminated from the model, we need to run the model again to obtain new p-values for other predictors left in the new model.

Use Minitab to Run a Multiple Linear Regression

- In this example, all three predictors have VIF higher than 5. Among them, EXAM1 has the highest p-value.

- We will remove EXAM1 from the equation and run the model again.



```
Regression Analysis: FINAL versus EXAM1, EXAM2, EXAM3

Analysis of Variance

Source       DF    Adj SS   Adj MS   F-Value   P-Value
Regression    3   13731.5  4577.17    670.09     0.000
  EXAM1       1      58.7    58.73      8.60     0.008
  EXAM2       1     197.7   197.67     28.94     0.000
  EXAM3       1     877.3   877.30    128.43     0.000
Error        21     143.4     6.83
Total        24   13875.0

Model Summary

      S    R-sq  R-sq(adj)  R-sq(pred)
2.61357  98.97%     98.82%      98.51%

Coefficients

Term       Coef  SE Coef  T-Value  P-Value   VIF
Constant  -4.34     3.76    -1.15    0.262
EXAM1     0.356    0.121     2.93    0.008  7.81
EXAM2     0.543    0.101     5.38    0.000  5.59
EXAM3     1.167    0.103    11.33    0.000  5.16

Regression Equation

FINAL  =  -4.34 + 0.356 EXAM1 + 0.543 EXAM2 + 1.167 EXAM3

Fits and Diagnostics for Unusual Observations

                                 Std
Obs   FINAL      Fit   Resid   Resid
 23  175.00   167.69    7.31    2.93  R
```

Use Minitab to Run a Multiple Linear Regression

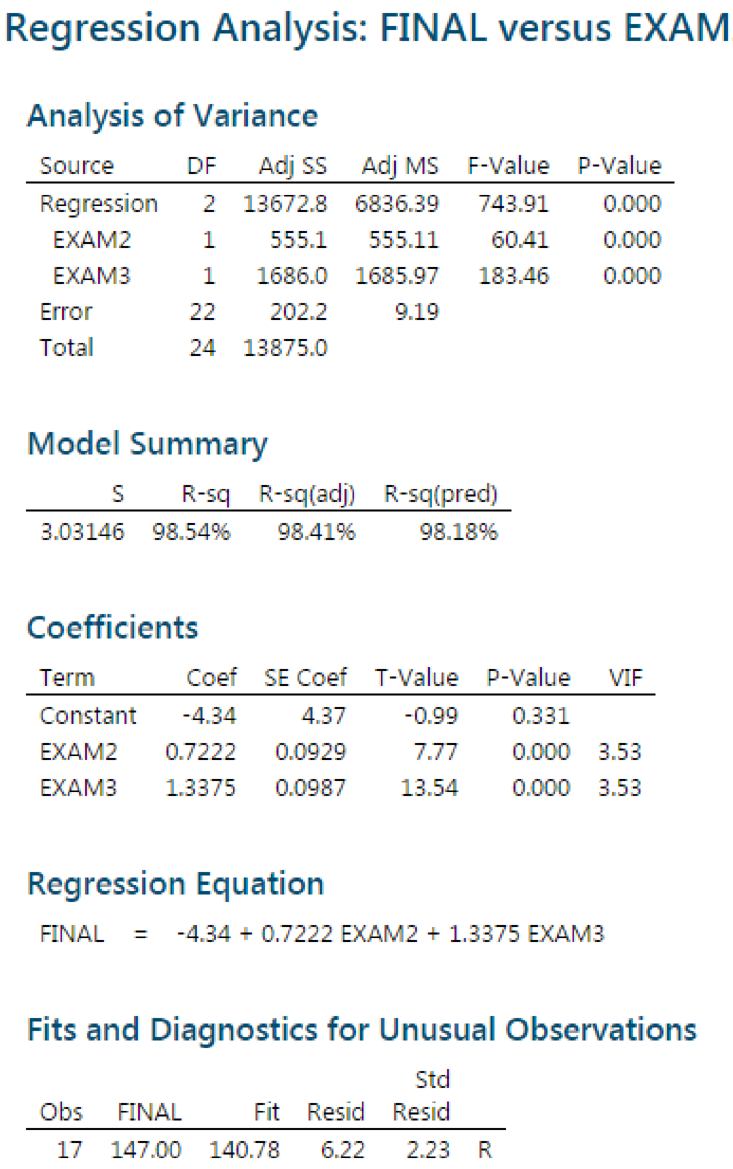
- Run the new multiple linear regression with only two predictors (i.e., EXAM2 and EXAM3).

- Check the VIFs of EXAM2 AND EXAM3. They are both smaller than 5; hence, there is little multicollinearity existing in the model.

Use Minitab to Run a Multiple Linear Regression

- In this example, both predictors' p-values are smaller than alpha level (0.05).

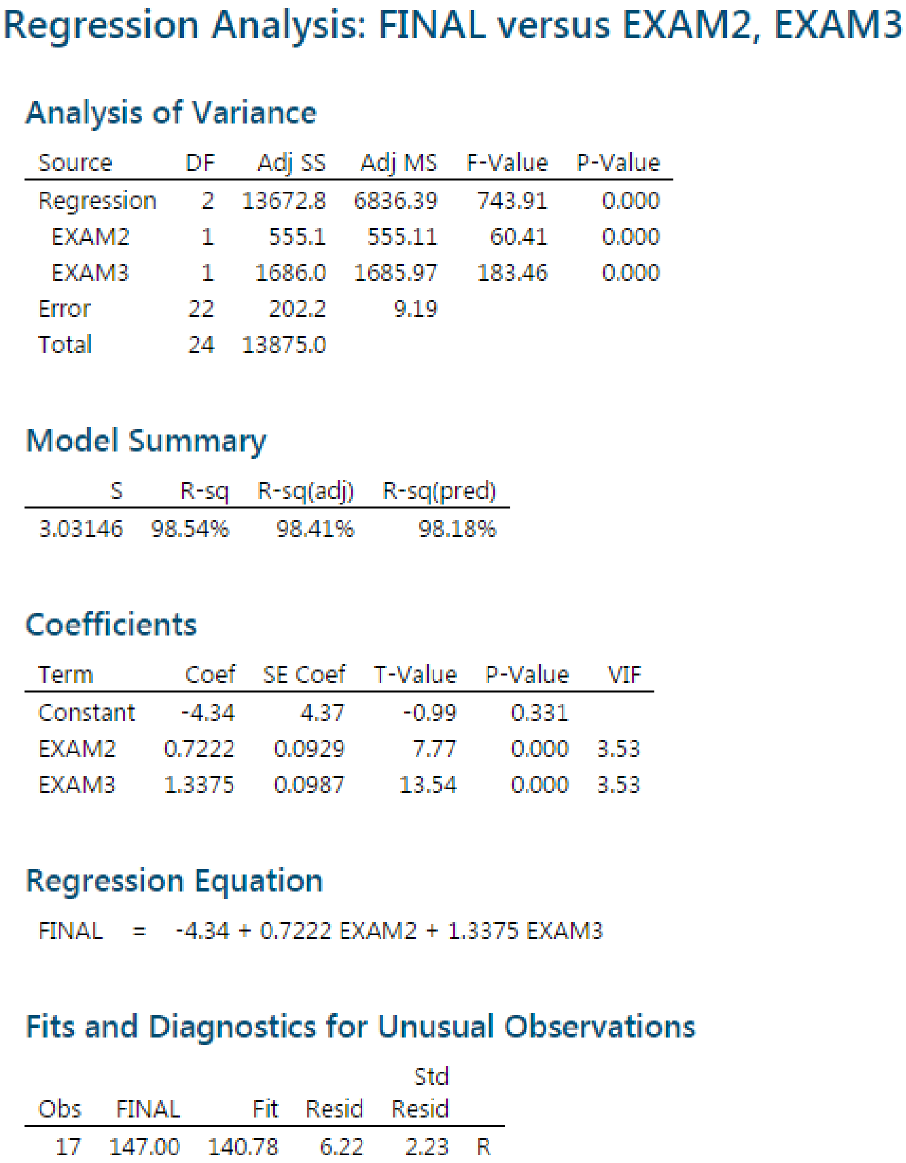
- As a result, we do not need to eliminate any variables from the model.



```
Regression Analysis: FINAL versus EXAM2, EXAM3

Analysis of Variance

Source       DF   Adj SS   Adj MS  F-Value  P-Value
Regression    2  13672.8  6836.39   743.91    0.000
  EXAM2       1    555.1   555.11    60.41    0.000
  EXAM3       1   1686.0  1685.97   183.46    0.000
Error        22    202.2     9.19
Total        24  13875.0

Model Summary

      S    R-sq  R-sq(adj)  R-sq(pred)
3.03146  98.54%     98.41%      98.18%

Coefficients

Term        Coef  SE Coef  T-Value  P-Value   VIF
Constant   -4.34     4.37    -0.99    0.331
EXAM2     0.7222   0.0929     7.77    0.000  3.53
EXAM3     1.3375   0.0987    13.54    0.000  3.53

Regression Equation

FINAL  =  -4.34 + 0.7222 EXAM2 + 1.3375 EXAM3

Fits and Diagnostics for Unusual Observations

                               Std
Obs   FINAL     Fit  Resid   Resid
 17  147.00  140.78   6.22    2.23  R
```

Use Minitab to Run a Multiple Linear Regression

- Step 7: Interpret the regression equation
  - The multiple linear regression equation appears automatically at the top of the session window.
  - “Coefficients” section provides the estimates of parameters in the linear regression equation.



```
Regression Analysis: FINAL versus EXAM2, EXAM3

Analysis of Variance

Source      DF   Adj SS   Adj MS  F-Value  P-Value
Regression   2  13672.8  6836.39   743.91    0.000
  EXAM2      1    555.1   555.11    60.41    0.000
  EXAM3      1   1686.0  1685.97   183.46    0.000
Error       22    202.2     9.19
Total       24  13875.0

Model Summary

      S    R-sq  R-sq(adj)  R-sq(pred)
3.03146  98.54%     98.41%      98.18%

Coefficients

Term        Coef  SE Coef  T-Value  P-Value   VIF
Constant   -4.34     4.37    -0.99    0.331
EXAM2     0.7222   0.0929     7.77    0.000  3.53
EXAM3     1.3375   0.0987    13.54    0.000  3.53

Regression Equation

FINAL  =  -4.34 + 0.7222 EXAM2 + 1.3375 EXAM3

Fits and Diagnostics for Unusual Observations

                             Std
Obs   FINAL     Fit  Resid  Resid
 17  147.00  140.78   6.22   2.23  R
```

Interpreting the Results

- Rsquare Adj = 98.4%
  - 98% of the variation in FINAL can be explained by the predictor variables EXAM2 & EXAM3.
- P-value of the F-test = 0.000
  - We have a statistically significant model.
- Variables p-value:
  - Both are significant (less than 0.05).
- VIF
  - EXAM2 & EXAM3 are both below 5; we’re in good shape!
- Equation: -4.34 + 0.722*EXAM2 + 1.34*EXAM3
  - -4.34 is the Y intercept, all equations will start with -4.34.
  - 0.722 is the EXAM2 coefficient; multiply it by EXAM2 score.
  - 1.34 is the EXAM3 coefficient; multiply it by EXAM3 score.

Interpreting the Results

- Let us say you are the professor again, and this time you want to use your prediction equation to estimate what one of your students might get on their final exam.

- Assume the following:
    - Exam 2 results were: 84
    - Exam 3 results were: 102.

- Use your equation: -4.34 + 0.722*EXAM2 + 1.34*EXAM3

- Predict your student's final exam score:
    - -4.34 + (0.722*84) + (1.34*102) =-4.34 + 60.648 + 136.68 = **192.988**

Nice work again! Now you can use your "magic" as the smart and efficient professor and allocate your time to other students because this one projects to perform much better than the average score of 162.

4.2.3 Confidence & Prediction Intervals

Prediction

- The purpose of building a regression model is not only to understand what happened in the past but more importantly to *predict* the future based on the past.

- By plugging the values of independent variables into the regression equation, we obtain the estimation/prediction of the dependent variable.

Uncertainty of Prediction

- We build the regression model using the sample data to describe as close as possible the true population relationship between dependent and independent variables.
- Due to noise in the data, the prediction will probably differ from the true response value.
- However, the true response value might fall in a range around the prediction with some certainty.
- To measure the uncertainty of the prediction, we need confidence interval and prediction interval.

Confidence Interval

- The **confidence interval** of the prediction is a range in which the population mean of the dependent variable would fall with some certainty, given specified values of the independent variables.
- The width of confidence interval is related to:
    - Sample size
    - Confidence level
    - Variation in the data.
- We build the model based on a sample set $\{y_1, y_2, \ldots, y_n\}$. The confidence interval is used to estimate the value of the population mean μ of the underlying population.
- The focus of the confidence interval are the unobservable population parameters.
- The confidence interval accounts for the uncertainty in the estimates of regression parameters.

Prediction Interval

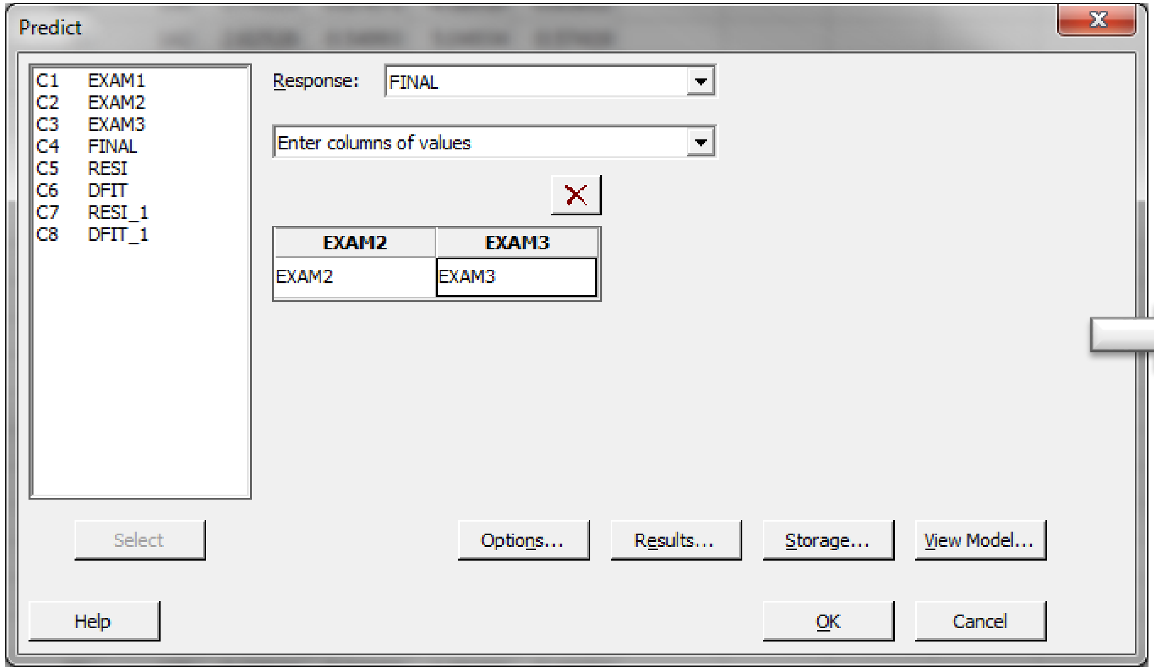
- The prediction interval is a range in which future values of the dependent variable would fall with some certainty, given specified values of the independent variables.
- We build the model based on a sample set {y1, y2,…, yn}. The prediction interval is used to estimate the value of future observation y_{n+1}.
- The focus of the prediction interval are the future observations.
- Prediction interval is wider than confidence interval because it accounts for the uncertainty in the estimates of regression parameters and the uncertainty of the new measurement.

Use Minitab to Obtain Prediction & Confidence Interval

- Steps in Minitab to obtain prediction, confidence interval and prediction interval after running Regression tests.

  1. Click Stat → Regression → Predict
  2. Choose “FINAL” as Response
  3. Select “Enter columns of values” from second dropdown menu
  4. Click on “Exam 2” for the first column and “Exam 3” for the second column
  5. Click “OK”
  6. The fitted response (PFIT), the confidence interval (CLIM), and the prediction interval (PLIM) are automatically added to your data table.

Use Minitab to Obtain Prediction & Confidence Interval



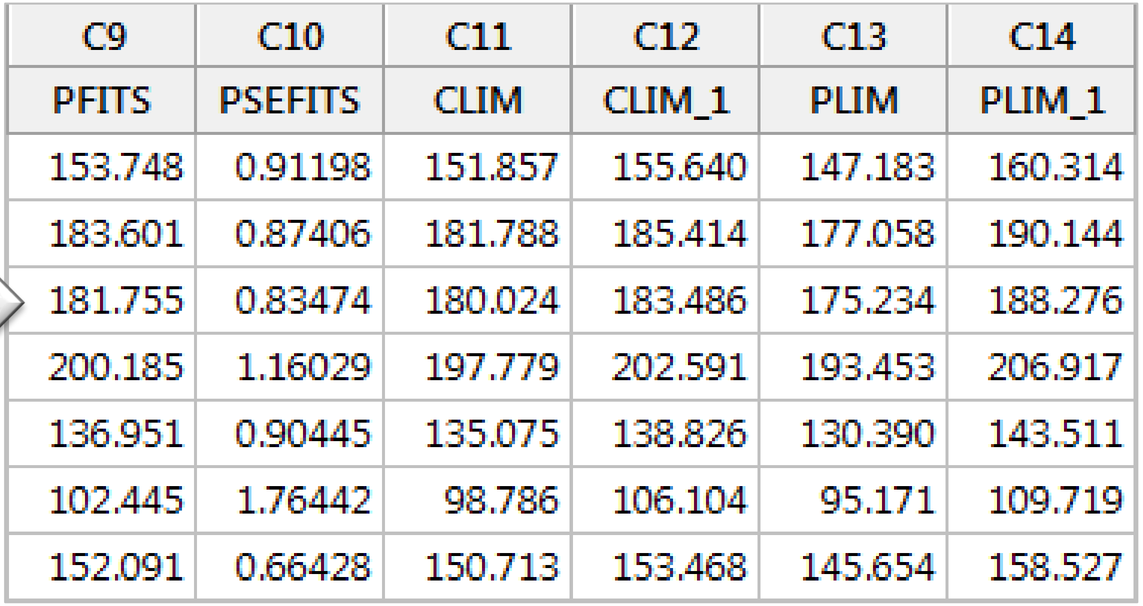
Predict

C1 EXAM1
C2 EXAM2
C3 EXAM3
C4 FINAL
C5 RESI
C6 DFIT
C7 RESI_1
C8 DFIT_1

Response: FINAL

Enter columns of values

EXAM2	EXAM3
EXAM2	EXAM3

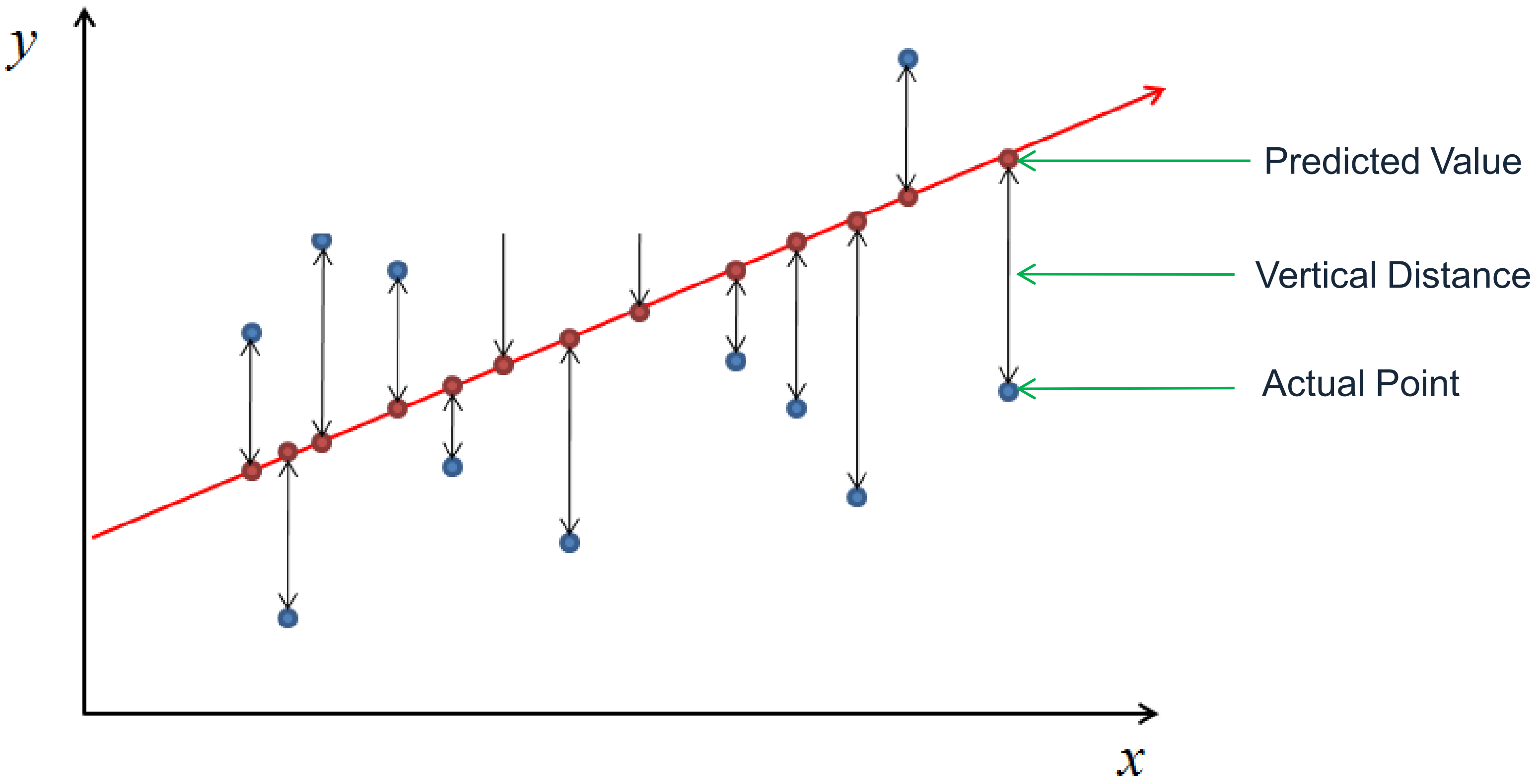


C9	C10	C11	C12	C13	C14
PFITS	PSEFITS	CLIM	CLIM_1	PLIM	PLIM_1
153.748	0.91198	151.857	155.640	147.183	160.314
183.601	0.87406	181.788	185.414	177.058	190.144
181.755	0.83474	180.024	183.486	175.234	188.276
200.185	1.16029	197.779	202.591	193.453	206.917
136.951	0.90445	135.075	138.826	130.390	143.511
102.445	1.76442	98.786	106.104	95.171	109.719
152.091	0.66428	150.713	153.468	145.654	158.527

4.2.4 Residuals Analysis

Remember what Residuals Are?

- **Residuals** are the vertical difference between actual values and the predicted values or the “fitted line” created by the regression model.



Why Perform Residuals Analysis?

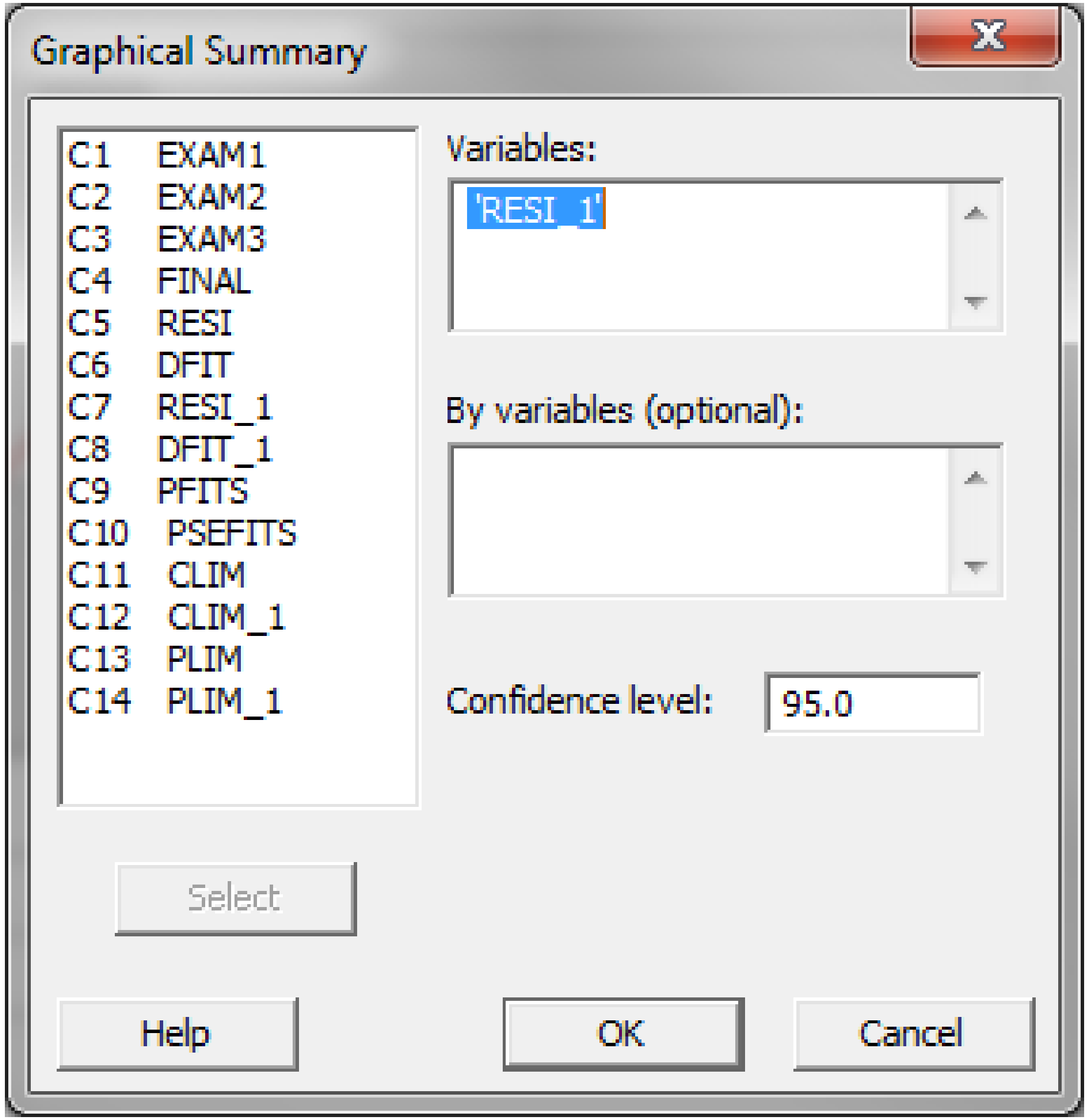
- The *regression equation* generated based on the sample data can make accurate statistical inference only if certain assumptions are met. Residuals analysis can help to validate these assumptions. The following assumptions must be met to ensure the reliability of the linear regression model:
  - The errors are normally distributed with mean equal to zero.
  - The errors are independent.
  - The errors have a constant variance.
  - The underlying population relationship is linear.

- If the residuals performance does not meet the requirement, we will need to rebuild the model by replacing the predictors with new ones, adding new predictors, building non-linear models, and so on.

Use Minitab to Perform Residuals Analysis

- Step 1: Check whether residuals are normally distributed around the mean of zero.
  1) Click Stat → Basic Statistics → Graphical Summary.
  2) A new window named "Graphical Summary" pops up.
  3) Select "RESI_1" as the "Variables." *We are selecting RESI_1 because we ran our regression two times, the first with all Exams which gave us our first set of residuals (RESI), the second model we ran without Exam1 and then we got residuals added to our data table and that column was automatically labeled RESI_1.*
  4) Click "OK."
  5) The histogram and analysis results are shown automatically in the new window.

Use Minitab to Perform Residuals Analysis



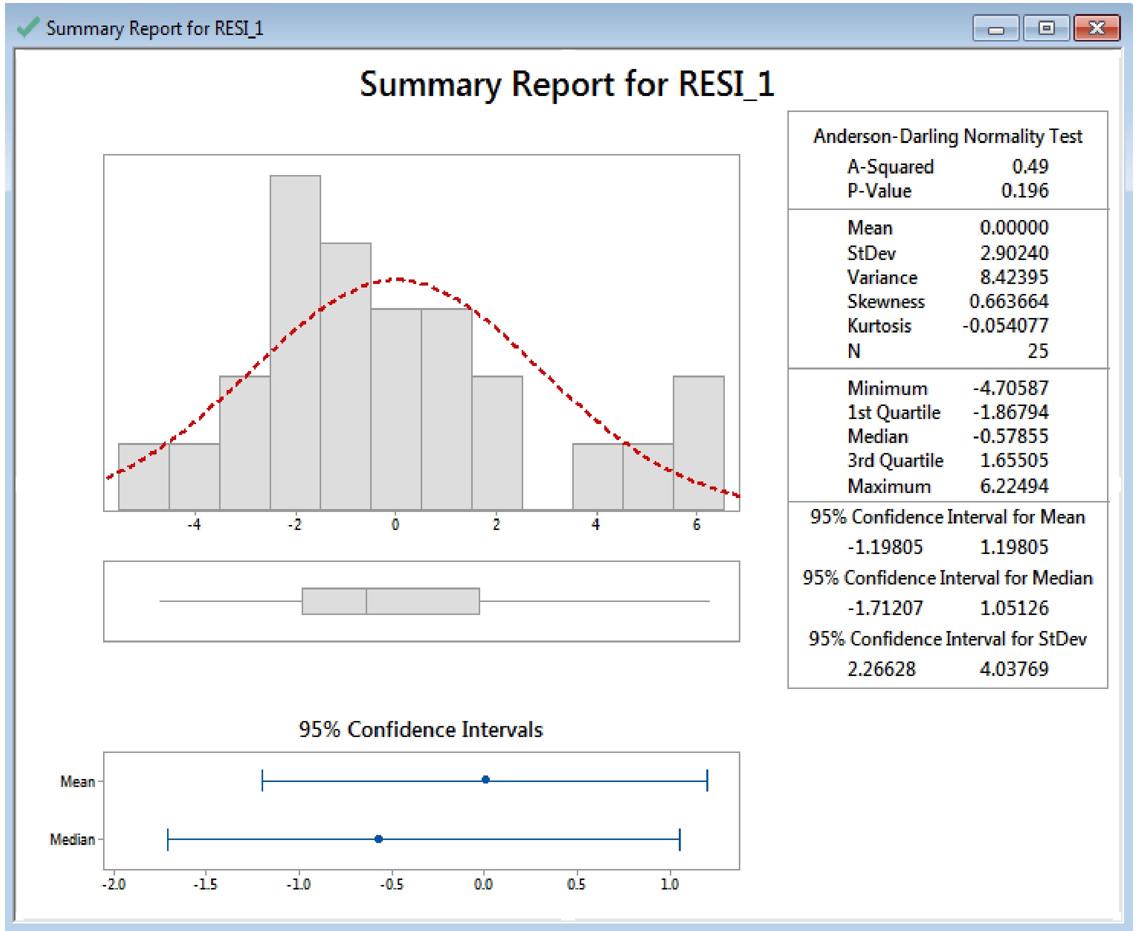
Variables: 'RESI_1'

By variables (optional):

Confidence level: 95.0

Use Minitab to Perform Residuals Analysis

- The mean of residuals is -0.0000.
- The Anderson-Darling test is used to test the normality. Since the p-value (0.196) is greater than the alpha level (0.05), we fail to reject the null hypothesis; the residuals are normally distributed.
  - H_0: The residuals are normally distributed.
  - H_1: The residuals are not normally distributed.



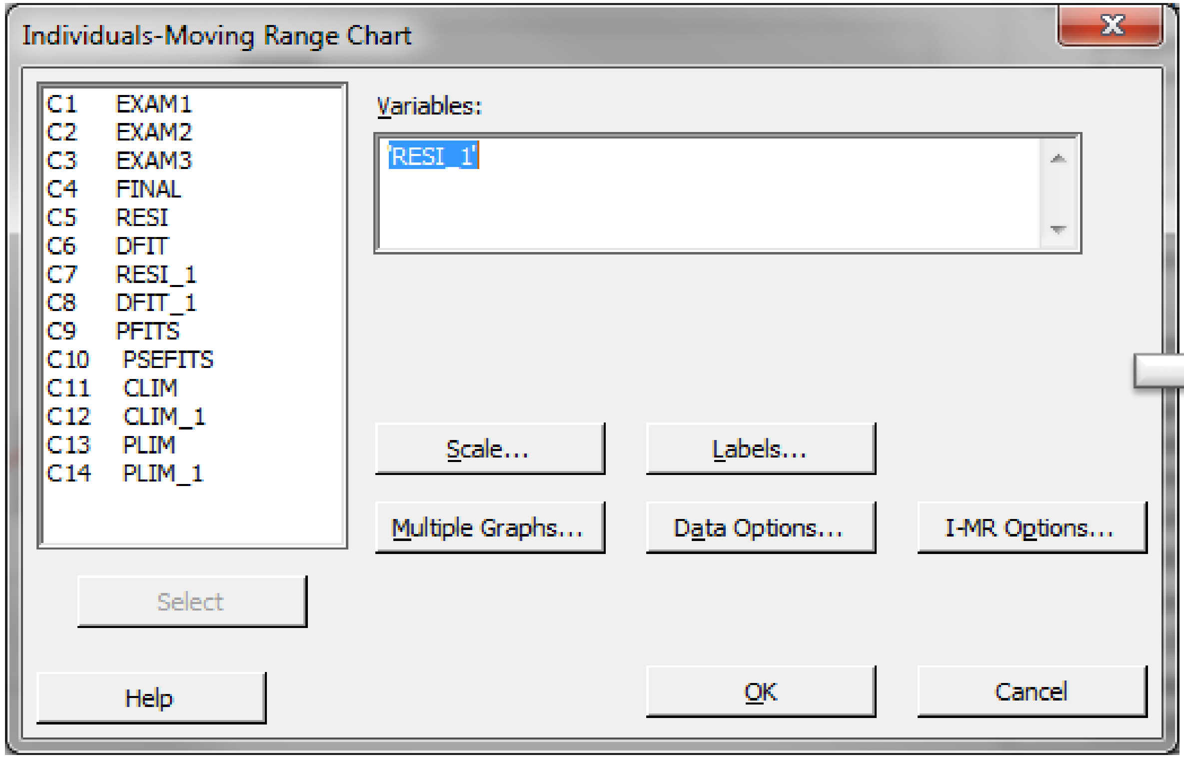
Summary Report for RESI_1

```
Anderson-Darling Normality Test
   A-Squared          0.49
   P-Value           0.196

   Mean            0.00000
   StDev           2.90240
   Variance        8.42395
   Skewness       0.663664
   Kurtosis      -0.054077
   N                    25

   Minimum        -4.70587
   1st Quartile   -1.86794
   Median         -0.57855
   3rd Quartile    1.65505
   Maximum         6.22494
95% Confidence Interval for Mean
   -1.19805        1.19805
95% Confidence Interval for Median
   -1.71207        1.05126
95% Confidence Interval for StDev
    2.26628        4.03769
```

Use Minitab to Perform Residuals Analysis

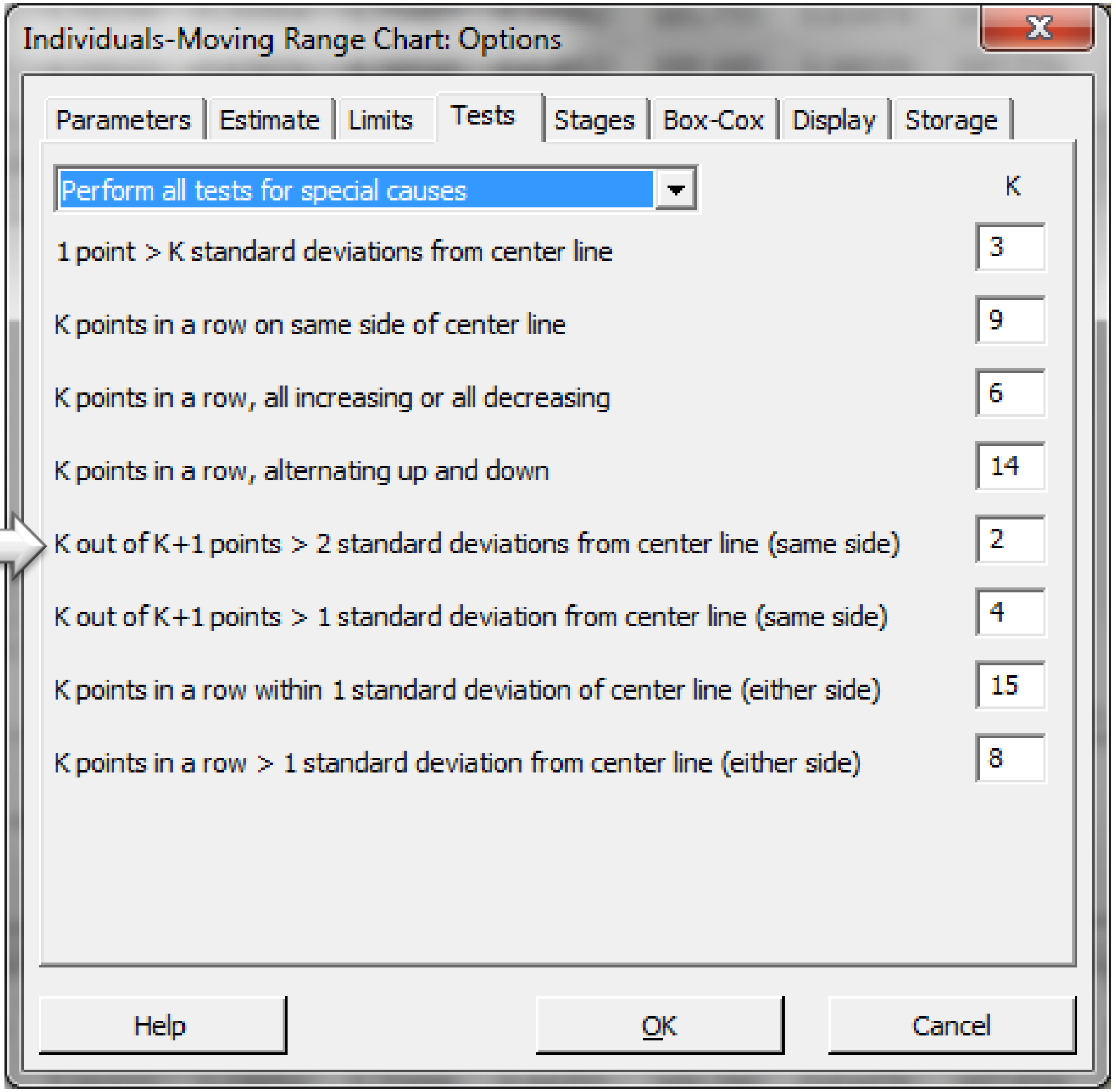
- Step 2: If the data are in time order, run the I-MR chart to check whether residuals are independent.
    1) Click Stat → Control Charts → Variable Charts for Individuals → I-MR.
    2) A new window named “Individuals – Moving Range Chart” pops up.
    3) Select “RESI1” as “Variables
    4) Click “I-MR Options”
    5) Click on “Tests” tab and select “Perform all tests for special causes”
    6) Click “OK.” to close “I-MR” Options
    7) Click OK
    8) The control charts are shown automatically in the new window.

Use Minitab to Perform Residuals Analysis


Individuals-Moving Range Chart

C1 EXAM1
C2 EXAM2
C3 EXAM3
C4 FINAL
C5 RESI
C6 DFIT
C7 RESI_1
C8 DFIT_1
C9 PFITS
C10 PSEFITS
C11 CLIM
C12 CLIM_1
C13 PLIM
C14 PLIM_1

Variables: 'RESI_1'

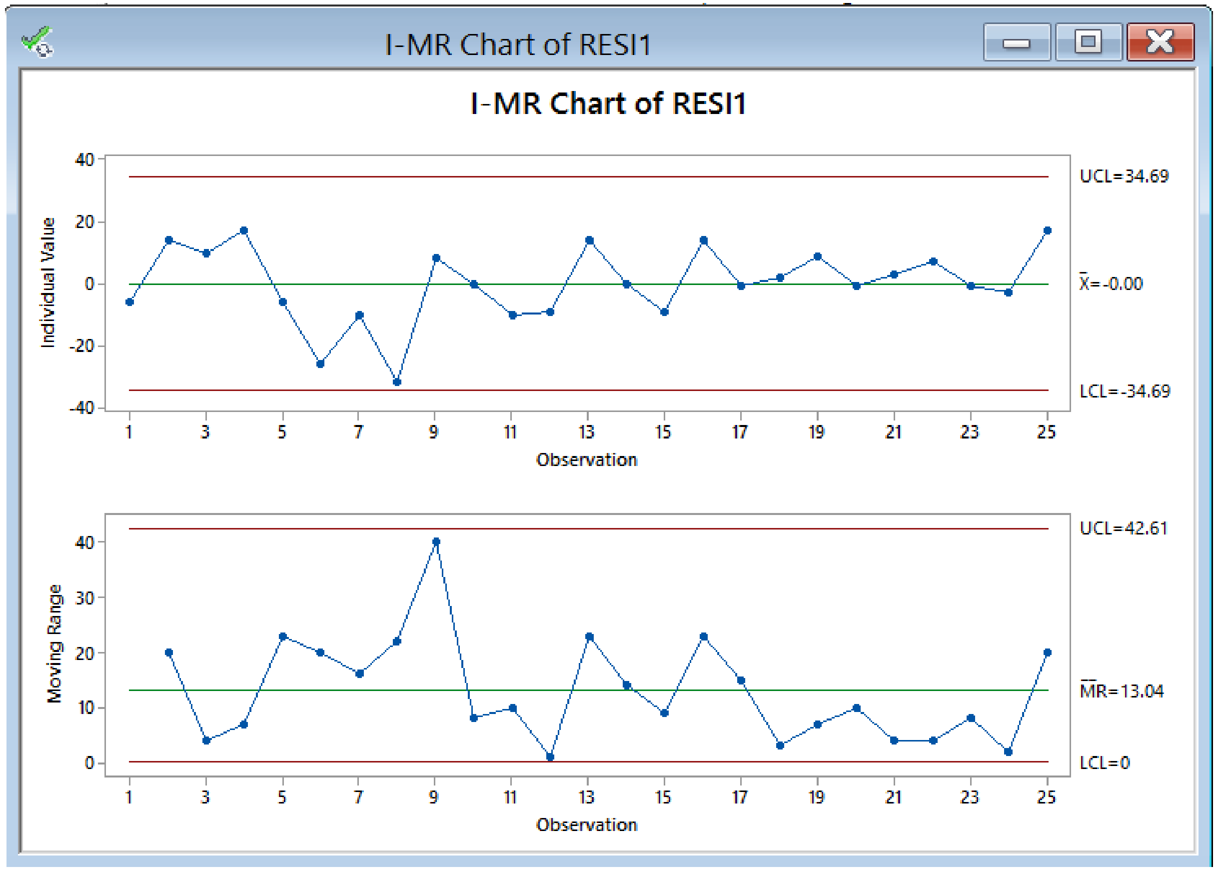

Individuals-Moving Range Chart: Options

Perform all tests for special causes

	K
1 point > K standard deviations from center line	3
K points in a row on same side of center line	9
K points in a row, all increasing or all decreasing	6
K points in a row, alternating up and down	14
K out of K+1 points > 2 standard deviations from center line (same side)	2
K out of K+1 points > 1 standard deviation from center line (same side)	4
K points in a row within 1 standard deviation of center line (either side)	15
K points in a row > 1 standard deviation from center line (either side)	8

Use Minitab to Perform Residuals Analysis

- If no data points are out of control in both the I-chart and MR chart, the residuals are independent of each other.
- If the residuals are not independent, it is possible that some important predictors are not included in the model.
- In this example, since the IMR chart is in control, residuals are independent.

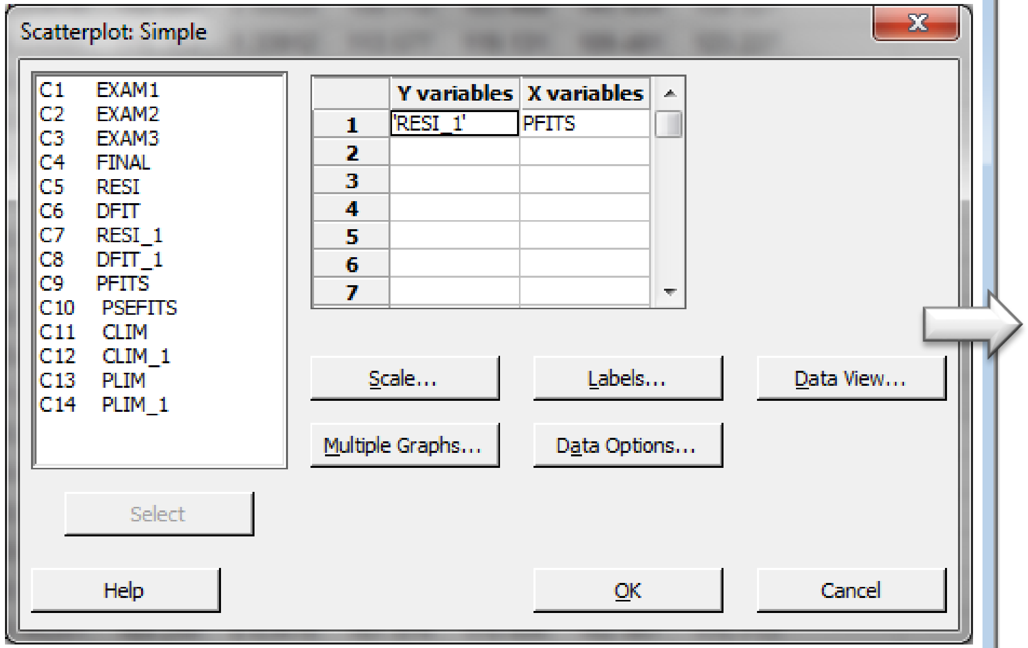


Use Minitab to Perform Residuals Analysis

- Step 3: Check whether residuals have equal variance across the predicted responses.
  - Create a scatterplot with Y being the residuals and the X being the fitted values.
    1. Click Graph → Scatterplot.
    2. A new window named “Scatterplots” appears.
    3. Select “RESI1” as the “Y variables” and “Fitted” as the “X variables.”
    4. Click “OK.”
    5. The scatterplot appears in a new window.

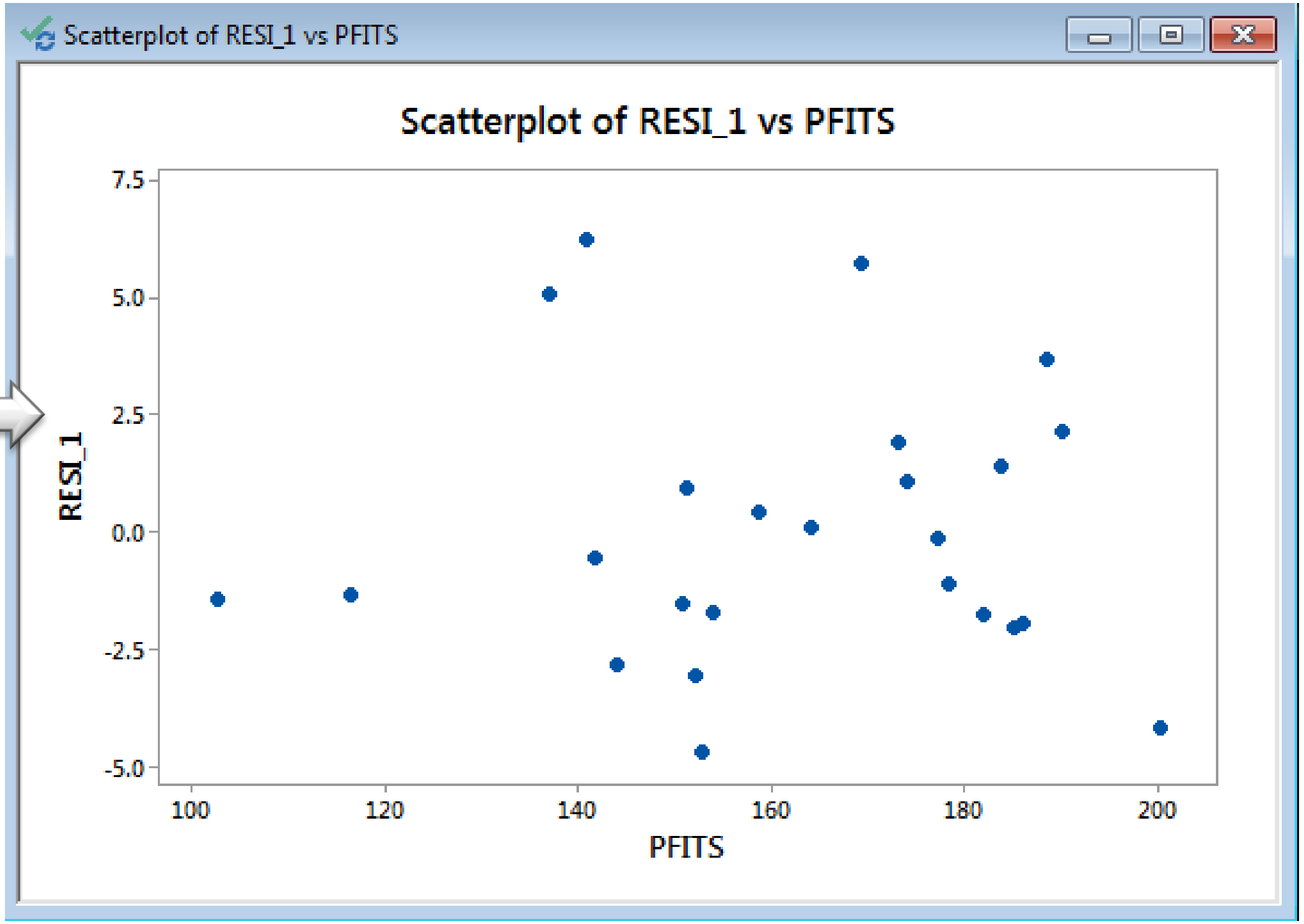
  - We are looking for the pattern in which residuals spread out evenly around zero from the top to the bottom.

Use Minitab to Perform Residuals Analysis


Scatterplot: Simple

C1 EXAM1
C2 EXAM2
C3 EXAM3
C4 FINAL
C5 RESI
C6 DFIT
C7 RESI_1
C8 DFIT_1
C9 PFITS
C10 PSEFITS
C11 CLIM
C12 CLIM_1
C13 PLIM
C14 PLIM_1

	Y variables	X variables
1	'RESI_1'	PFITS
2		
3		
4		
5		
6		
7		


Scatterplot of RESI_1 vs PFITS

4.2.5 Data Transformation

What is the Box-Cox Transformation?

- When a response does not fit the model well, sometimes using a transformation of the response can improve the fit.
- **Power transformation** is a class of transformation functions that raise the response to some power. For example, a square root transformation converts X to $X^{1/2}$.
- **Box-Cox transformation** is a popular power transformation method developed by George E. P. Box and David Cox.

Box-Cox Transformation Formula

- The formula of the Box-Cox transformation is:

$$\begin{cases} y = \dfrac{x^{\lambda} - 1}{\lambda} & \text{where } \lambda \neq 0 \\ y = \ln x & \text{where } \lambda = 0 \end{cases}$$

y is the transformation result.
x is the variable under transformation.
λ is the transformation parameter.

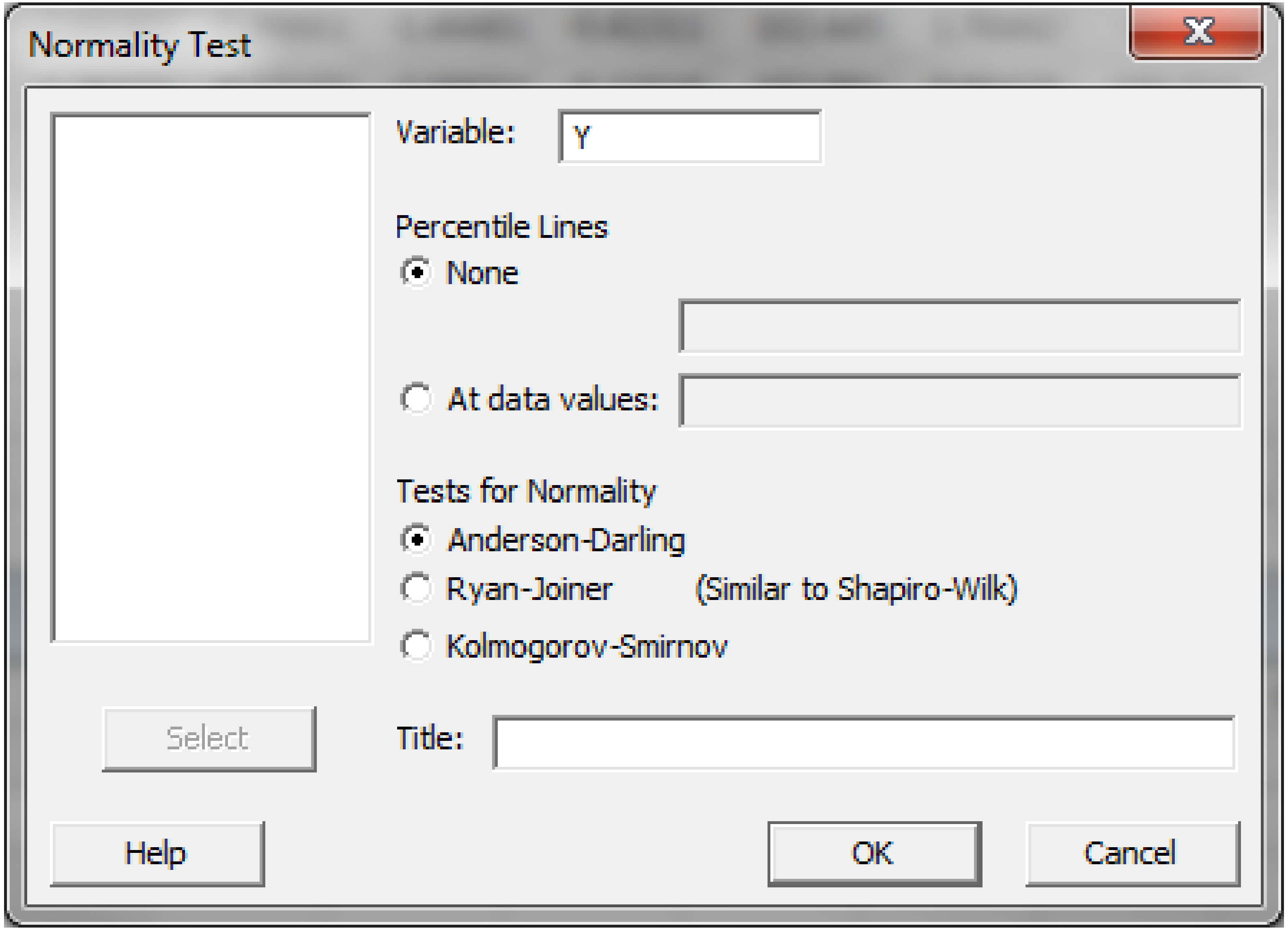
Use Minitab to Perform a Box-Cox Transformation

- Minitab provides the best Box-Cox transformation with an optimal λ that minimizes the model SSE (sum of squared error).
- Here is an example of how we transform the non-normally distributed response to normal data using Box-Cox method.
- Data File: “Box-Cox” tab in “Sample Data.xlsx”

Use Minitab to Perform a Box-Cox Transformation

- Step 1: Test the normality of the original data set.
  1) Click Stat → Basic Statistics → Normality Test.
  2) A new window named “Normality Test” pops up.
  3) Select “Y” as “Variable.”
  4) Click “OK.”
  5) The normality test results are shown automatically in the new window.

Use Minitab to Perform a Box-Cox Transformation



Variable: Y

Percentile Lines

● None

○ At data values:

Tests for Normality

● Anderson-Darling

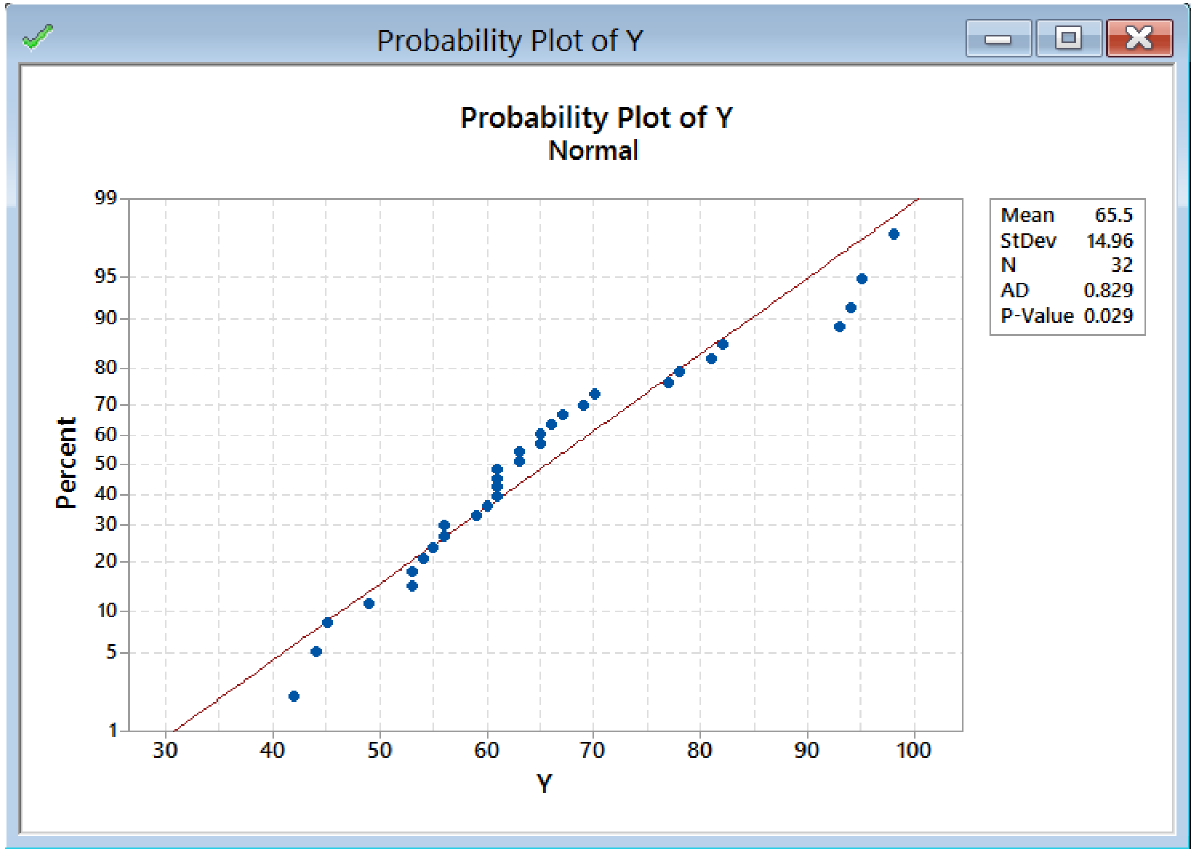
○ Ryan-Joiner (Similar to Shapiro-Wilk)

○ Kolmogorov-Smirnov

Title:

Use Minitab to Perform a Box-Cox Transformation

- Normality Test:
  - H0: The data are normally distributed.
  - H1: The data are not normally distributed.
- If p-value > alpha level (0.05), we fail to reject the null hypothesis. Otherwise, we reject the null.
- In this example, p-value = 0.029 < alpha level (0.05). The data are not normally distributed.

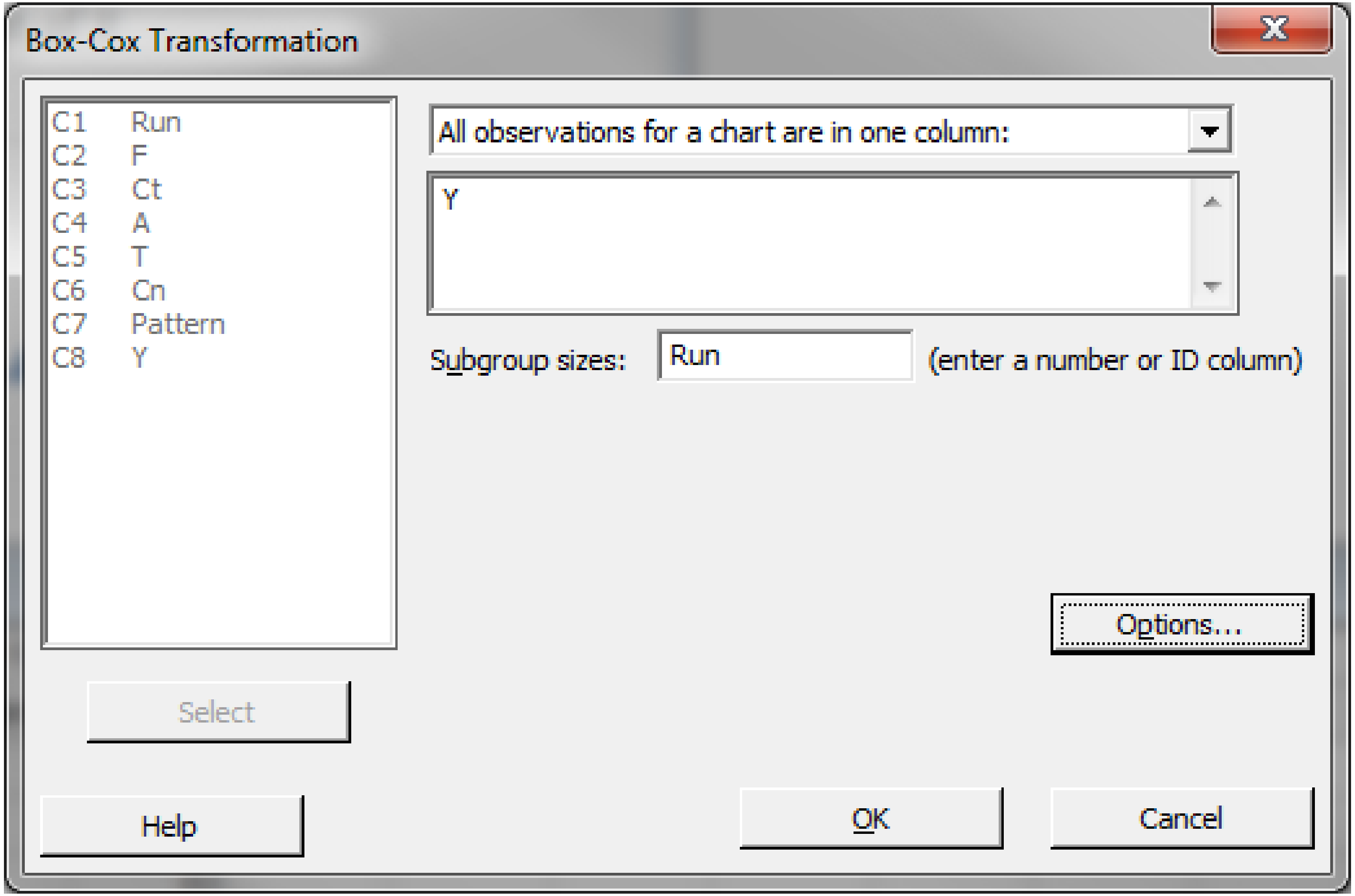


Mean 65.5
StDev 14.96
N 32
AD 0.829
P-Value 0.029

Use Minitab to Perform a Box-Cox Transformation

- Step 2: Run the Box-Cox Transformation:
  1) Click Stat → Control Charts → Box-Cox Transformation.
  2) A new window named “Box-Cox Transformation” pops up.
  3) Click into the blank list box below “All observations for a chart are in one column.”
  4) Select “Y” as the variable.
  5) Select “Run” into the box next to “Subgroup sizes (enter a number or ID column).”
  6) Click “OK.”
  7) The analysis results are shown automatically in the new window.

Use Minitab to Perform a Box-Cox Transformation



Box-Cox Transformation

C1 Run
C2 F
C3 Ct
C4 A
C5 T
C6 Cn
C7 Pattern
C8 Y

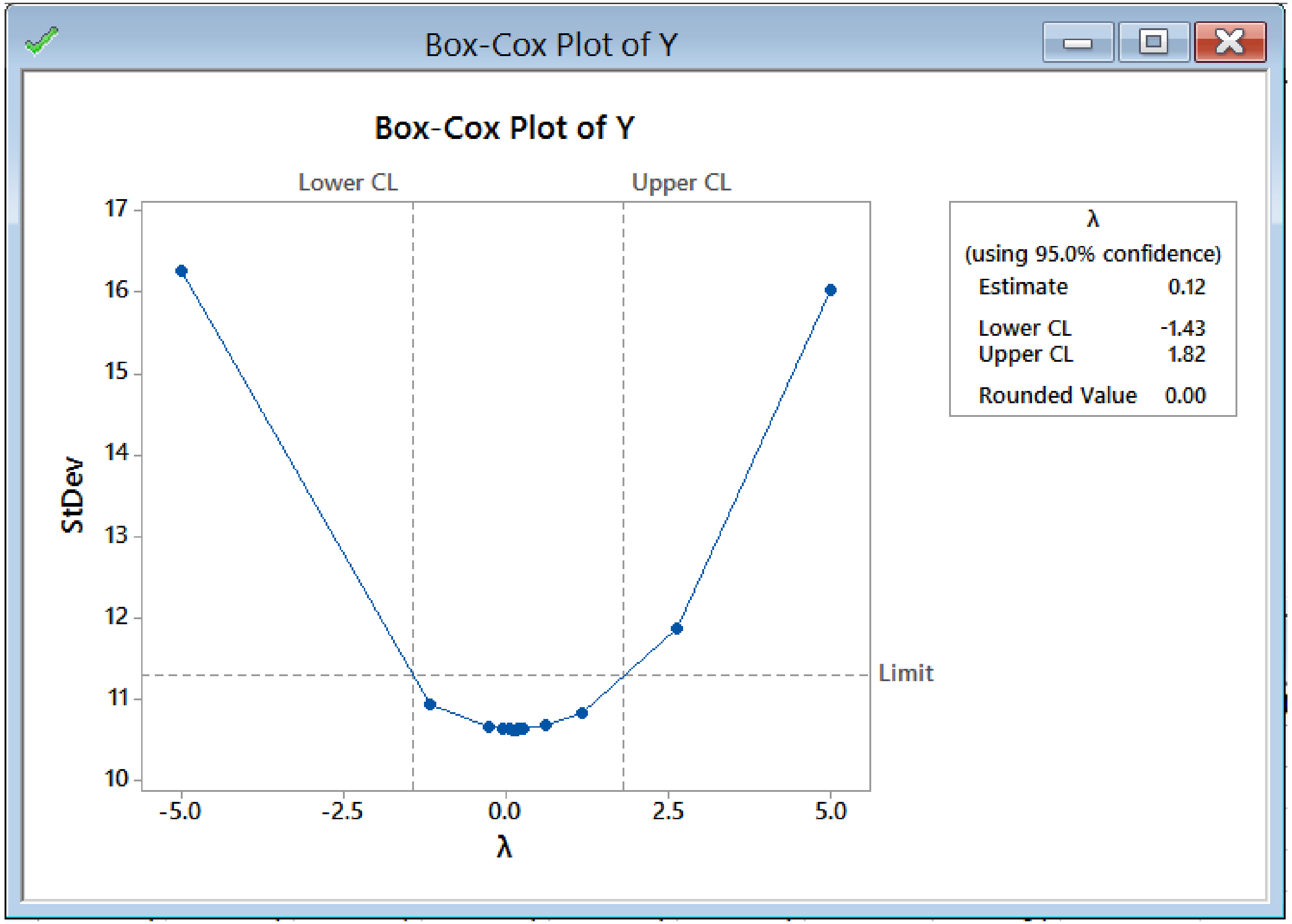
All observations for a chart are in one column:

Y

Subgroup sizes: Run (enter a number or ID column)

Use Minitab to Perform a Box-Cox Transformation

- When λ is 0.12, the transformation is the optimal with minimum SSE.



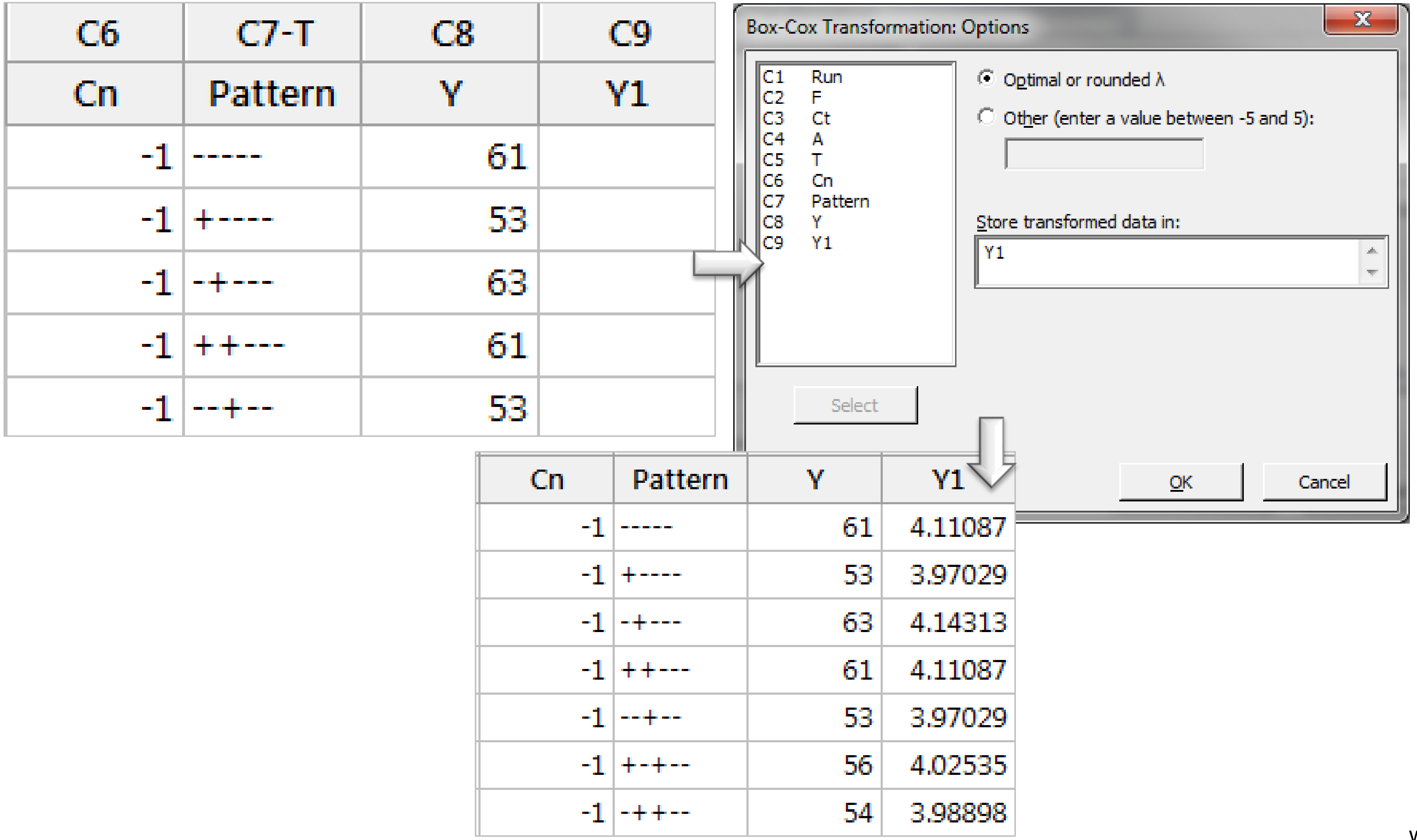
λ
(using 95.0% confidence)

Estimate	0.12
Lower CL	-1.43
Upper CL	1.82
Rounded Value	0.00

Use Minitab to Perform a Box-Cox Transformation

- The transformed Y can also be saved in another column.
  1) Create a new column named “Y1” in the data table.
  2) Click on the “Options” button in the “Box-Cox Transformation” window.
  3) A new window named “Box-Cox Transformation – Options” appears.
  4) Click in the blank box under “Store transformed data in” and all the columns pop up in the list box on the left.
  5) Select “Y1” in “Store transformed data in.”
  6) Click “OK” in the window “Box-Cox Transformation – Options.”
  7) Click “OK” in the window “Box-Cox Transformation.”
  8) The transformed column is stored in the column “Y1.”

Use Minitab to Perform a Box-Cox Transformation



C6	C7-T	C8	C9
Cn	Pattern	Y	Y1
-1	-----	61	
-1	+----	53	
-1	-+---	63	
-1	++---	61	
-1	--+--	53	

Box-Cox Transformation: Options

C1	Run
C2	F
C3	Ct
C4	A
C5	T
C6	Cn
C7	Pattern
C8	Y
C9	Y1

● Optimal or rounded λ

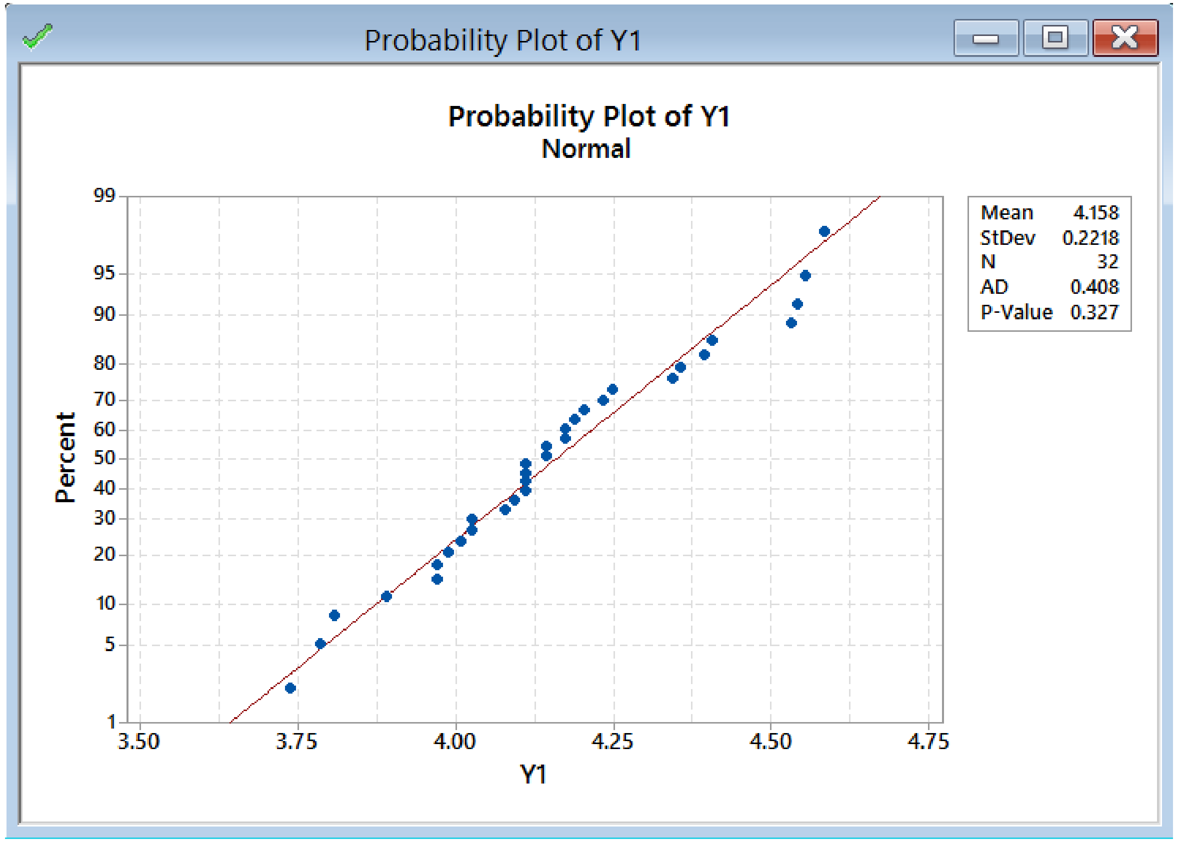
○ Other (enter a value between -5 and 5):

Store transformed data in: Y1

Cn	Pattern	Y	Y1
-1	-----	61	4.11087
-1	+----	53	3.97029
-1	-+---	63	4.14313
-1	++---	61	4.11087
-1	--+--	53	3.97029
-1	+-+--	56	4.02535
-1	-++--	54	3.98898

Use Minitab to Perform a Box-Cox Transformation

- Run the normality test to check whether the transformed data are normally distributed.
- Use the Anderson-Darling test to test the normality of the transformed data
  - H0: The data are normally distributed.
  - H1: The data are not normally distributed.
- If p-value > alpha level (0.05), we fail to reject the null. Otherwise, we reject the null.
- In this example, p-value = 0.327 > alpha level (0.05). The data are normally distributed.



Mean	4.158
StDev	0.2218
N	32
AD	0.408
P-Value	0.327

4.2.6 Stepwise Regression

What is Stepwise Regression?

- **Stepwise regression** is a statistical method to automatically select regression models with the best sets of predictive variables from a large set of potential variables.

- There are different statistical methods used in stepwise regression to evaluate the potential variables in the model:
  - F-test
  - T-test
  - R-square
  - AIC.

Three Approaches to Stepwise Regression

- Forward Selection
  - Bring in potential predictors one by one and keep them if they have significant impact on improving the model.
- Backward Selection
  - Try out potential predictors one by one and eliminate them if they are insignificant to improve the fit.
- Mixed Selection
  - Is a combination of both forward selection and backward selection. Add and remove variables based on pre-defined significance threshold levels.

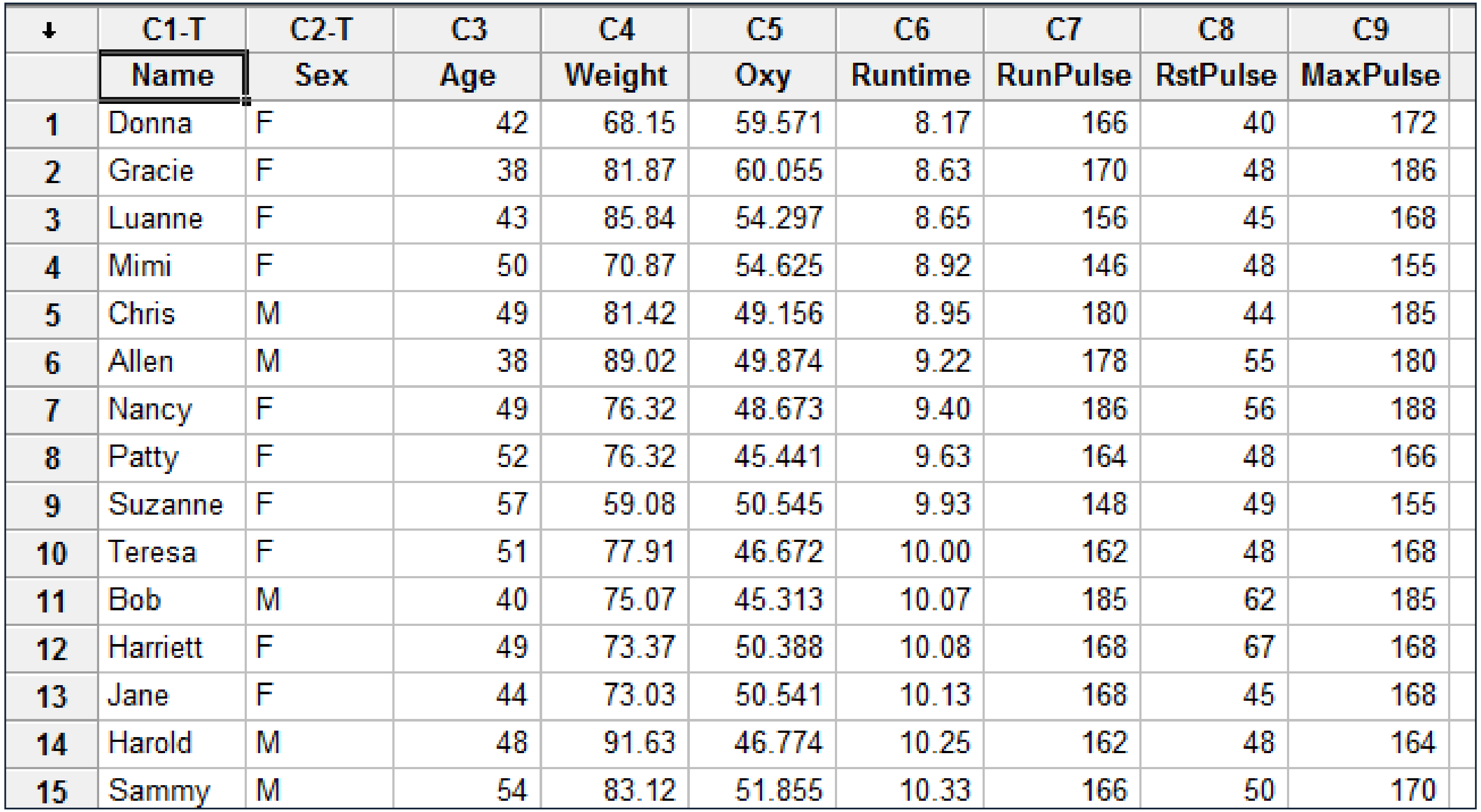
How to Use Minitab to Run a Stepwise Regression

- *Case study:* We want to build a regression model to predict the oxygen uptake of a person who runs 1.5 miles. The potential predictors are:
  - Age
  - Weight
  - Runtime
  - Runpulse
  - RstPulse
  - MaxPulse.

- Data File: “Stepwise Regression” tab in “Sample Data.xlsx”

How to Use Minitab to Run a Stepwise Regression

- Sample Data Glance

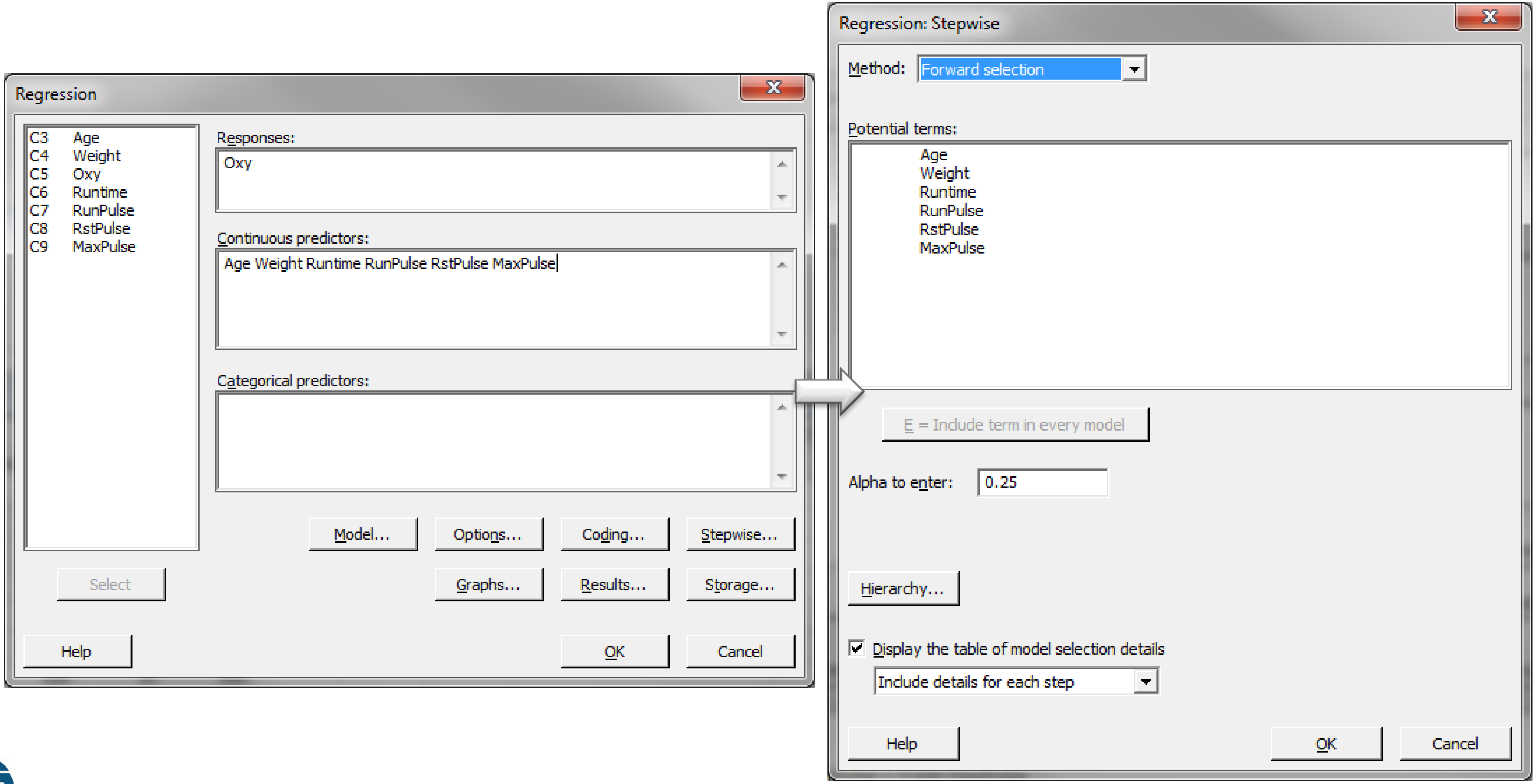


↓	C1-T	C2-T	C3	C4	C5	C6	C7	C8	C9
	Name	Sex	Age	Weight	Oxy	Runtime	RunPulse	RstPulse	MaxPulse
1	Donna	F	42	68.15	59.571	8.17	166	40	172
2	Gracie	F	38	81.87	60.055	8.63	170	48	186
3	Luanne	F	43	85.84	54.297	8.65	156	45	168
4	Mimi	F	50	70.87	54.625	8.92	146	48	155
5	Chris	M	49	81.42	49.156	8.95	180	44	185
6	Allen	M	38	89.02	49.874	9.22	178	55	180
7	Nancy	F	49	76.32	48.673	9.40	186	56	188
8	Patty	F	52	76.32	45.441	9.63	164	48	166
9	Suzanne	F	57	59.08	50.545	9.93	148	49	155
10	Teresa	F	51	77.91	46.672	10.00	162	48	168
11	Bob	M	40	75.07	45.313	10.07	185	62	185
12	Harriett	F	49	73.37	50.388	10.08	168	67	168
13	Jane	F	44	73.03	50.541	10.13	168	45	168
14	Harold	M	48	91.63	46.774	10.25	162	48	164
15	Sammy	M	54	83.12	51.855	10.33	166	50	170

How to Use Minitab to Run a Stepwise Regression

- Steps to run stepwise regression in Minitab
    1) Click Stat → Regression → Regression → Fit Regression Model.
    2) Select "Oxy" as the Response and select all the other variables into the "Continuous Predictors" box.
    3) Click the "Stepwise" button and a new window named "Regression: Stepwise" pops up.
    4) Select the method of stepwise regression and enter the alphas to enter/remove. In this example, we use the "Forward selection" method and the alpha to enter is 0.25.
    5) Select the option to "Include details for each step"
    6) Click "OK" in the window "Regression: Stepwise."
    7) Click "OK" in the window "Stepwise Regression."
    8) The results appear in the session window.

How to Use Minitab to Run a Stepwise Regression



Regression

C3 Age
C4 Weight
C5 Oxy
C6 Runtime
C7 RunPulse
C8 RstPulse
C9 MaxPulse

Responses: Oxy

Continuous predictors: Age Weight Runtime RunPulse RstPulse MaxPulse

Categorical predictors:

Regression: Stepwise

Method: Forward selection

Potential terms:
Age
Weight
Runtime
RunPulse
RstPulse
MaxPulse

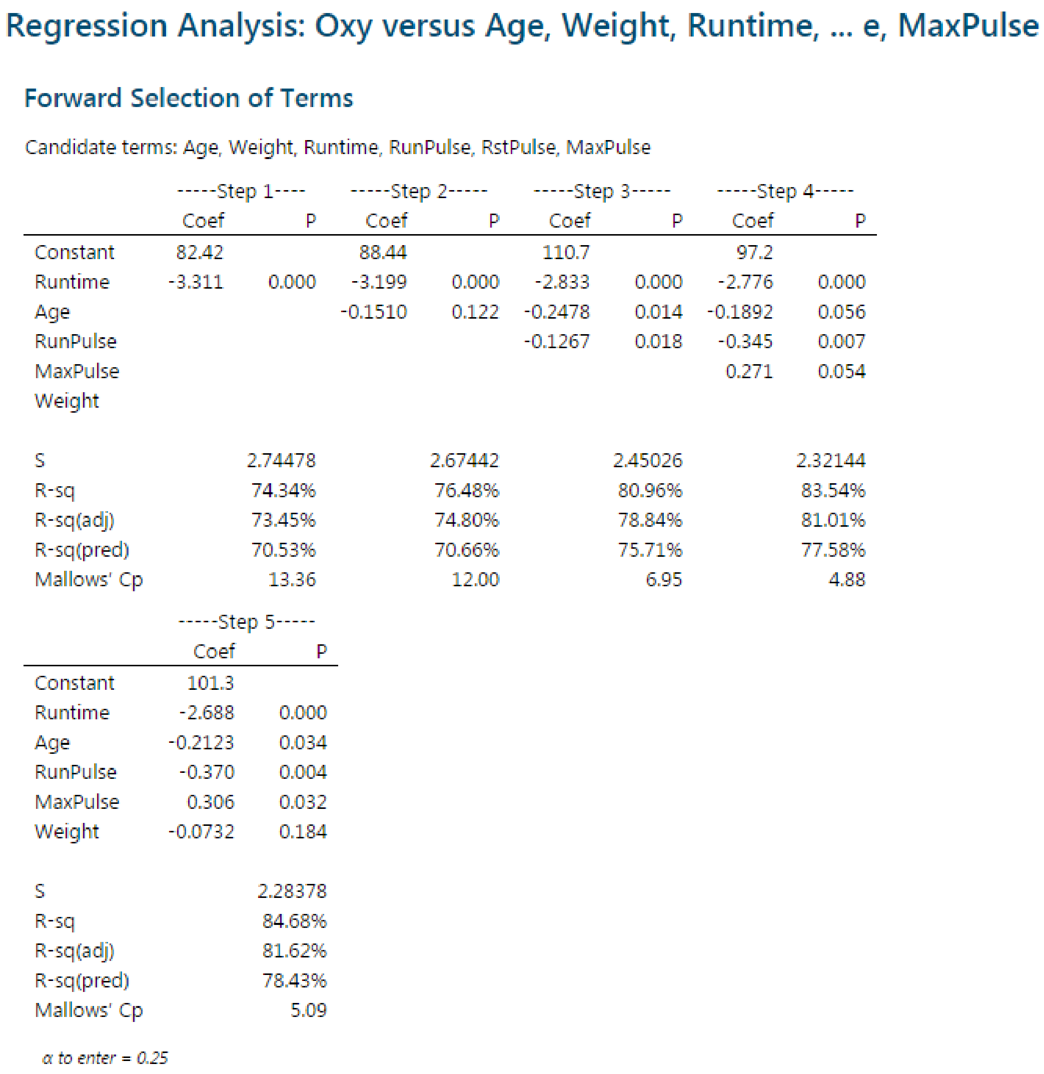
Alpha to enter: 0.25

☑ Display the table of model selection details

Include details for each step

How to Use Minitab to Run a Stepwise Regression

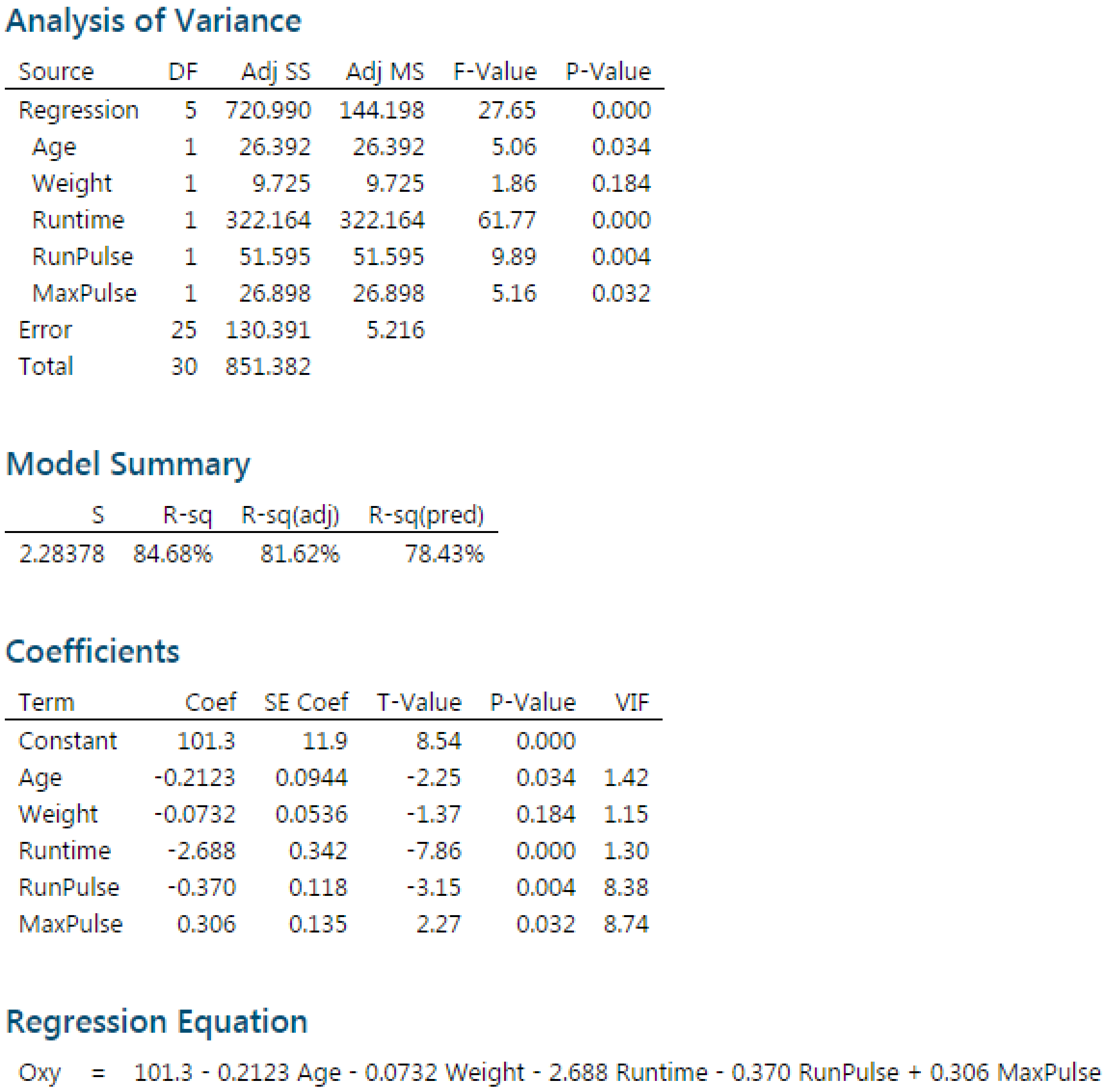
- **Step History:**
  - Step-by-step records on how to come up with the final model. Each column indicates the model built in each step.



```
Regression Analysis: Oxy versus Age, Weight, Runtime, ... e, MaxPulse

Forward Selection of Terms

Candidate terms: Age, Weight, Runtime, RunPulse, RstPulse, MaxPulse

              ----Step 1----    -----Step 2-----    -----Step 3-----    -----Step 4-----
                 Coef      P       Coef       P       Coef       P       Coef       P
Constant        82.42              88.44              110.7              97.2
Runtime        -3.311  0.000     -3.199   0.000     -2.833   0.000     -2.776   0.000
Age                              -0.1510  0.122     -0.2478  0.014     -0.1892  0.056
RunPulse                                            -0.1267  0.018     -0.345   0.007
MaxPulse                                                               0.271    0.054
Weight

S                      2.74478           2.67442            2.45026            2.32144
R-sq                    74.34%            76.48%             80.96%             83.54%
R-sq(adj)               73.45%            74.80%             78.84%             81.01%
R-sq(pred)              70.53%            70.66%             75.71%             77.58%
Mallows' Cp              13.36             12.00               6.95               4.88

              -----Step 5-----
                 Coef       P
Constant        101.3
Runtime        -2.688   0.000
Age            -0.2123  0.034
RunPulse       -0.370   0.004
MaxPulse        0.306   0.032
Weight         -0.0732  0.184

S                       2.28378
R-sq                     84.68%
R-sq(adj)                81.62%
R-sq(pred)               78.43%
Mallows' Cp                5.09

α to enter = 0.25
```

How to Use Minitab to Run a Stepwise Regression

- **Model summary:**
  - One out of six potential factors is not statistically significant since its p-value is higher than the alpha to enter.
  - We also have multi-collinearity with two factors with VIF’s greater than 5



```
Analysis of Variance

Source       DF   Adj SS   Adj MS  F-Value  P-Value
Regression    5  720.990  144.198    27.65    0.000
  Age         1   26.392   26.392     5.06    0.034
  Weight      1    9.725    9.725     1.86    0.184
  Runtime     1  322.164  322.164    61.77    0.000
  RunPulse    1   51.595   51.595     9.89    0.004
  MaxPulse    1   26.898   26.898     5.16    0.032
Error        25  130.391    5.216
Total        30  851.382

Model Summary

      S    R-sq  R-sq(adj)  R-sq(pred)
2.28378  84.68%     81.62%      78.43%

Coefficients

Term         Coef  SE Coef  T-Value  P-Value   VIF
Constant    101.3     11.9     8.54    0.000
Age       -0.2123   0.0944    -2.25    0.034  1.42
Weight    -0.0732   0.0536    -1.37    0.184  1.15
Runtime    -2.688    0.342    -7.86    0.000  1.30
RunPulse   -0.370    0.118    -3.15    0.004  8.38
MaxPulse    0.306    0.135     2.27    0.032  8.74

Regression Equation

Oxy  =  101.3 - 0.2123 Age - 0.0732 Weight - 2.688 Runtime - 0.370 RunPulse + 0.306 MaxPulse
```

4.2.7 Logistic Regression

What is Logistic Regression?

- **Logistic regression** is a statistical method to predict the probability of an event occurring by fitting the data to a logistic curve using logistic function.
- The dependent variable in a logistic regression can be binary (e.g., 1/0, yes/no, pass/fail), nominal (blue/yellow/green), or ordinal (satisfied/neutral/dissatisfied).
- The independent variables can be either continuous or discrete.

Logistic Function

$$f(z) = \frac{1}{1+e^{-z}}$$

where *z* can be any value ranging from negative infinity to positive infinity.

The value of *f(z)* ranges from 0 to 1, which matches exactly the nature of probability (i.e., 0 ≤ P ≤ 1).

Logistic Regression Equation

- Based on the logistic function

$$f(z) = \frac{1}{1+e^{-z}}$$

we define f(z) as the **probability** of an event occurring and *z* is the weighted sum of the significant predictive variables.

$$z = \beta_0 + \beta_1 \times x_1 + \beta_2 \times x_2 + \cdots + \beta_k \times x_k$$

Logistic Regression Equation

- Logistic Regression Y = F(x)

$$Y = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \times x_1 + \beta_2 \times x_2 + \cdots + \beta_k \times x_k)}}$$

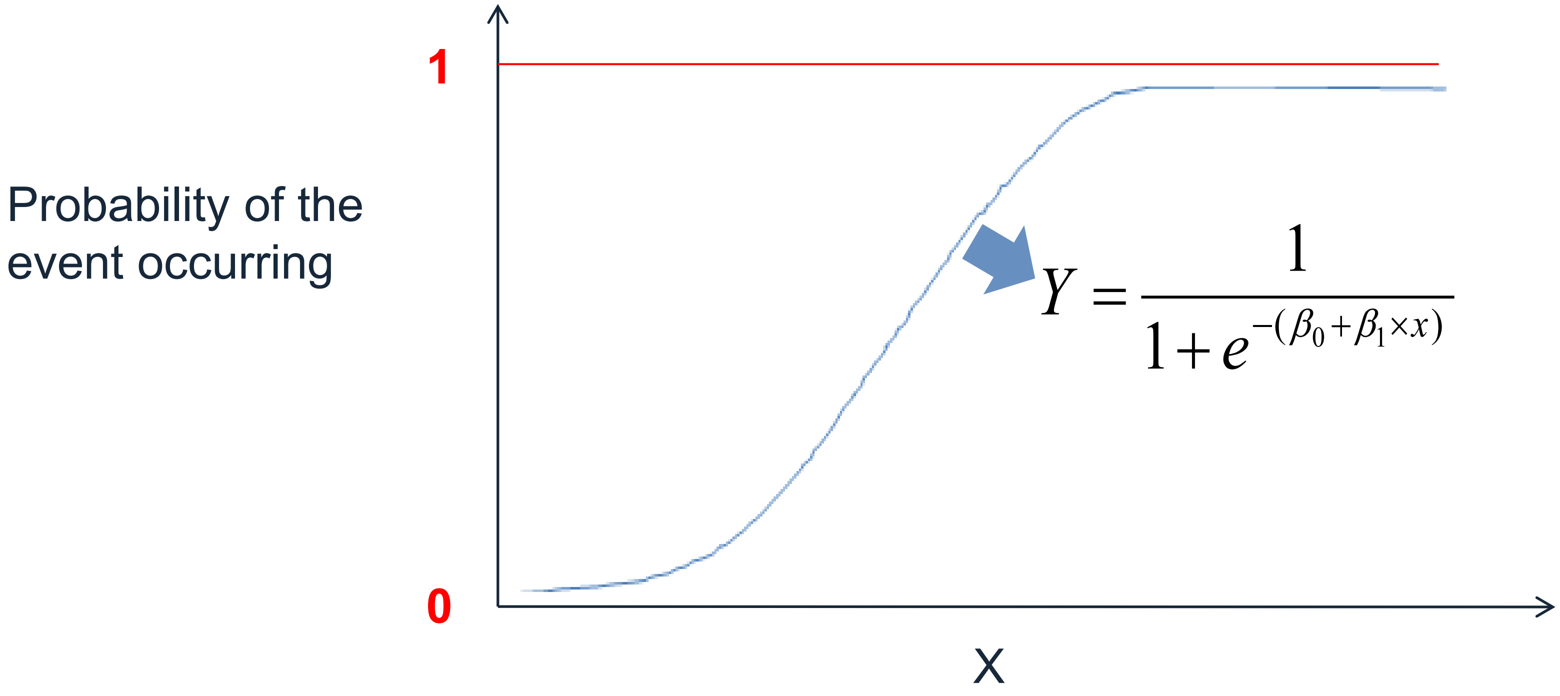
where *Y* is the probability of an event occurring and *x's* are the significant predictors.

Note:

- When building the regression model, we use the actual *Y,* which is discrete (e.g. binary, nominal, ordinal).
- After completing building the model, the fitted *Y* calculated using the logistic regression equation is the probability ranging from 0 to 1. To transfer the probability back to the discrete value, we need SMEs' inputs to select the probability cut point.

Logistic Curve

- Logistic curve for binary logistic regression with one continuous predictor:



$$Y = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \times x)}}$$

Odds

- **Odds** is the probability of an event occurring divided by the probability of the event not occurring.

$$Odds = \frac{P}{1-P}$$

- Odds range from 0 to positive infinity.

Odds

- Probability can be calculated using odds.

$$P = \frac{odds}{1+odds} = \frac{e^{\ln(odds)}}{1+e^{\ln(odds)}} = \frac{1}{1+e^{-\ln(odds)}}$$

- Since in logistic regression model

$$P = \frac{1}{1+e^{-(\beta_0+\beta_1\times x_1+\beta_2\times x_2+\cdots+\beta_k\times x_k)}}$$

therefore

$$\ln(odds) = \beta_0 + \beta_1 \times x_1 + \beta_2 \times x_2 + \cdots + \beta_k \times x_k$$

Three Types of Logistic Regression

- Binary Logistic Regression
  - Binary response variable
  - Example: yes/no, pass/fail, female/male
- Nominal Logistic Regression
  - Nominal response variable
  - Example: set of colors, set of countries
- Ordinal Logistic Regression
  - Ordinal response variable
  - Example: satisfied/neutral/dissatisfied

- All three logistic regression models can use multiple continuous or discrete independent variables.
- All three logistic regression models can be developed in Minitab using the same steps.

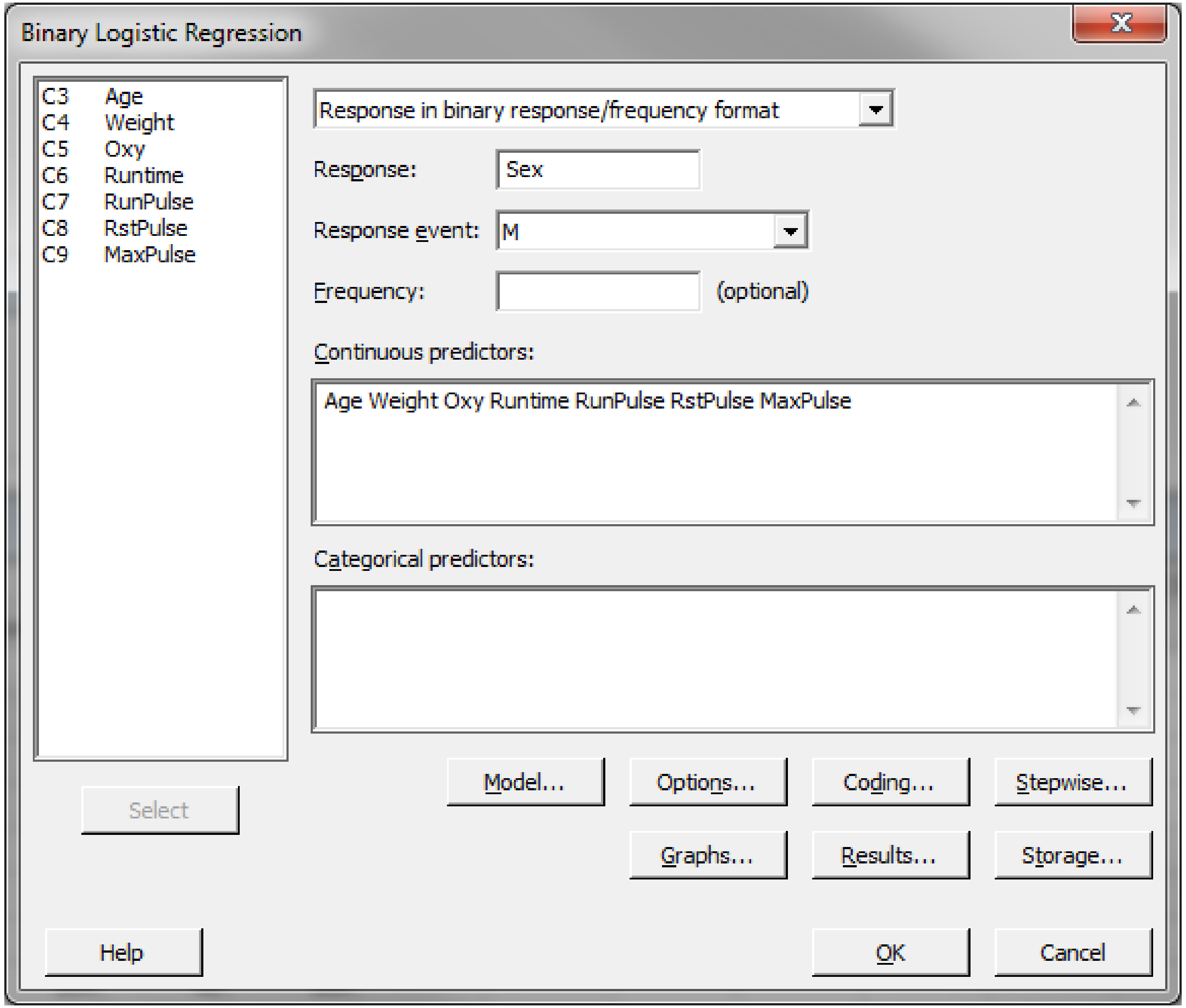
How to Run a Logistic Regression in Minitab

- Data File: “Logistic Regression” tab in “Sample Data.xlsx”
- Response and potential factors
  - Response (Y): Female/Male
  - Potential Factors (Xs):
    - Age
    - Weight
    - Oxy
    - Runtime
    - RunPulse
    - RstPulse
    - MaxPulse.
- We want to build a logistic regression model using the potential factors to predict the probability that the person measured is female or male.

How to Run a Logistic Regression in Minitab

- Step 1:
  1) Click Stat → Regression → Binary Logistic Regression → Fit Binary Logistic Model
  2) A new window named “Binary Logistic Regression” appears.
  3) Click into the blank box next to “Response” and all the variables pop up in the list box on the left.
  4) Select “Sex” as the “Response.”
  5) Select “Age”, “Weight”, “Oxy”, “Runtime”, “RunPulse”, “RstPulse”, “MaxPulse” into the “Continuous Predictors.”
  6) Click “OK.”

How to Run a Logistic Regression in Minitab



Binary Logistic Regression

C3 Age
C4 Weight
C5 Oxy
C6 Runtime
C7 RunPulse
C8 RstPulse
C9 MaxPulse

Response in binary response/frequency format

Response: Sex

Response event: M

Frequency: (optional)

Continuous predictors:

Age Weight Oxy Runtime RunPulse RstPulse MaxPulse

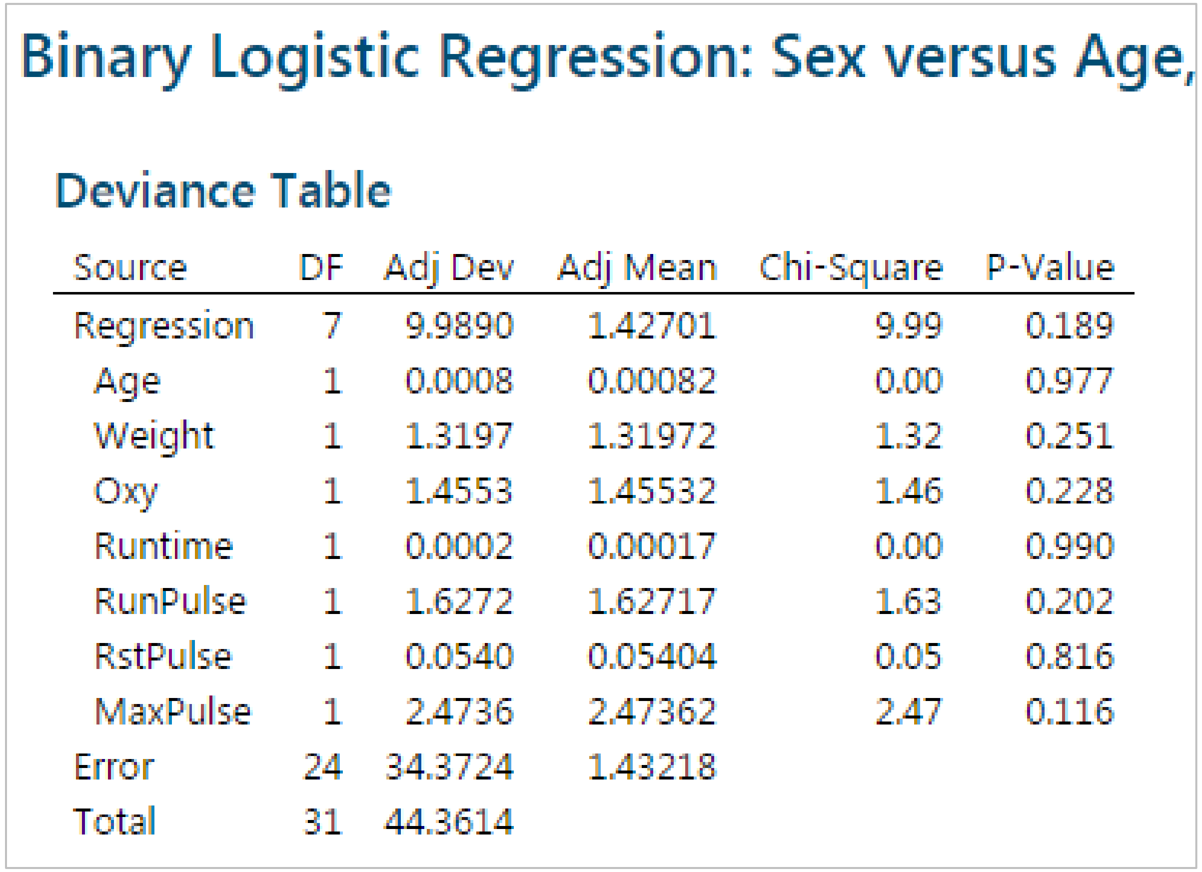
Categorical predictors:

How to Run a Logistic Regression in Minitab

- Step 2:
  1) Step 2.1: The results of the logistic regression model appear in session window.
  2) Step 2.2: Check the p-values of all the independent variables in the model.
  3) Step 2.3: Remove the insignificant independent variable one at a time from the model and rerun the model.
  4) Step 2.4: Repeat step 2.1 until all of the independent variables in the model are statistically significant.

How to Run a Logistic Regression in Minitab

- Since the p-values of all the independent variables are higher than the alpha level (0.05), we need to remove the insignificant independent variables one at a time from the model, starting from the one with the highest p-value.
- Runtime has the highest p-value (0.990), so it would be removed from the model first.

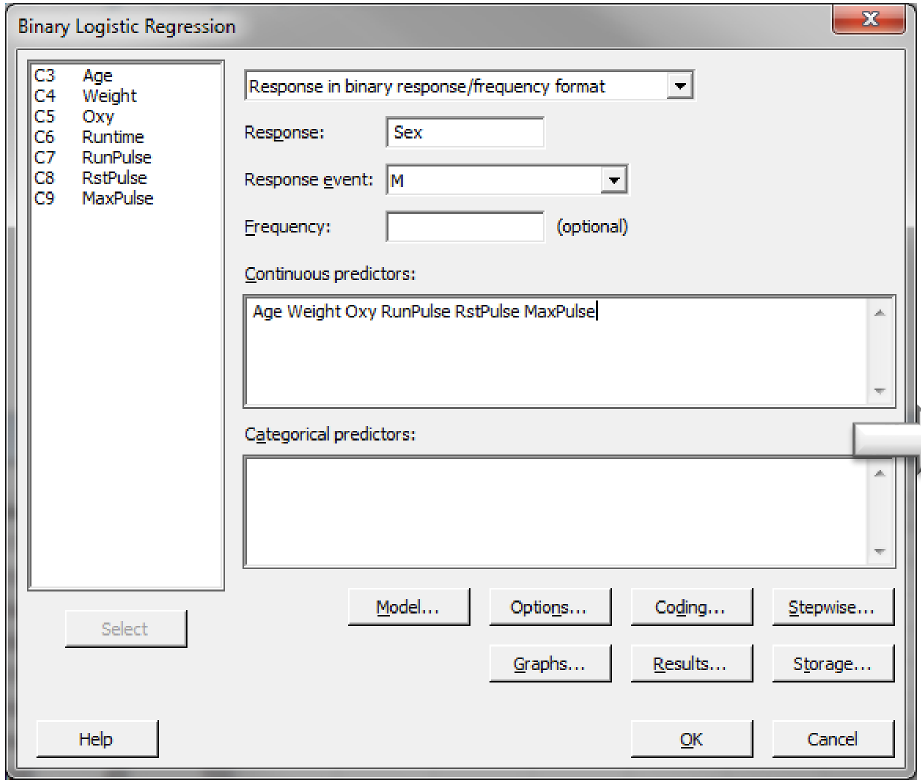


Binary Logistic Regression: Sex versus Age,

Deviance Table

Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	7	9.9890	1.42701	9.99	0.189
Age	1	0.0008	0.00082	0.00	0.977
Weight	1	1.3197	1.31972	1.32	0.251
Oxy	1	1.4553	1.45532	1.46	0.228
Runtime	1	0.0002	0.00017	0.00	0.990
RunPulse	1	1.6272	1.62717	1.63	0.202
RstPulse	1	0.0540	0.05404	0.05	0.816
MaxPulse	1	2.4736	2.47362	2.47	0.116
Error	24	34.3724	1.43218		
Total	31	44.3614			

How to Run a Logistic Regression in Minitab



Response in binary response/frequency format

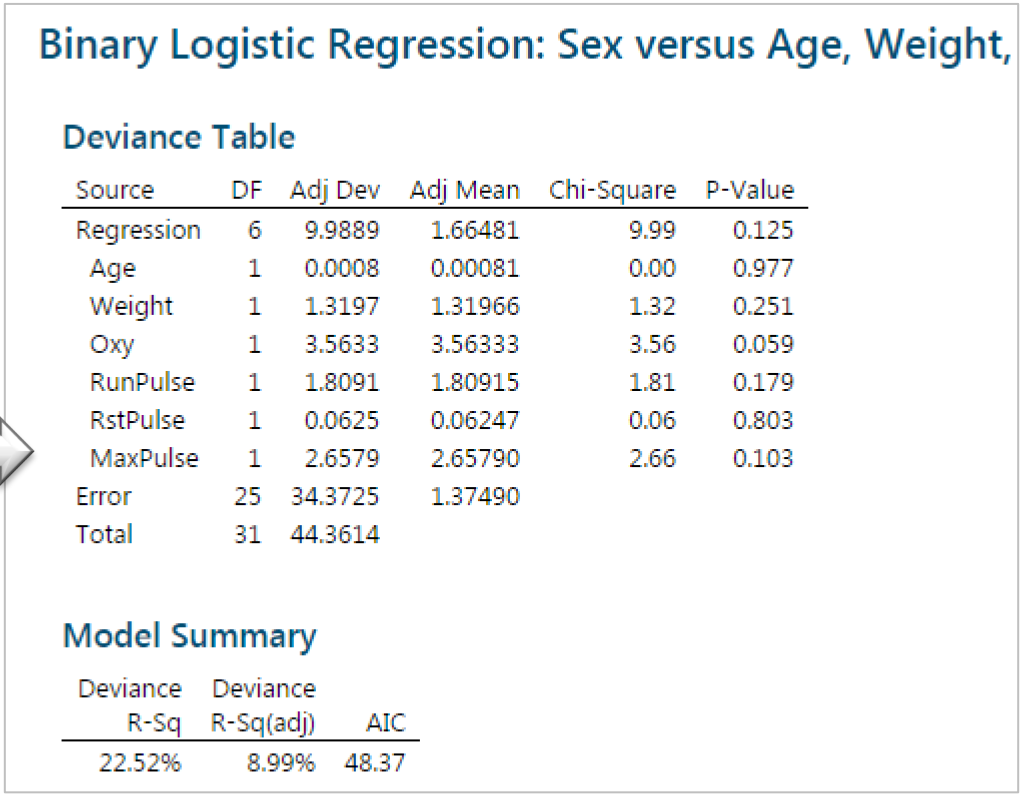
Response: Sex

Response event: M

Frequency: (optional)

Continuous predictors: Age Weight Oxy RunPulse RstPulse MaxPulse

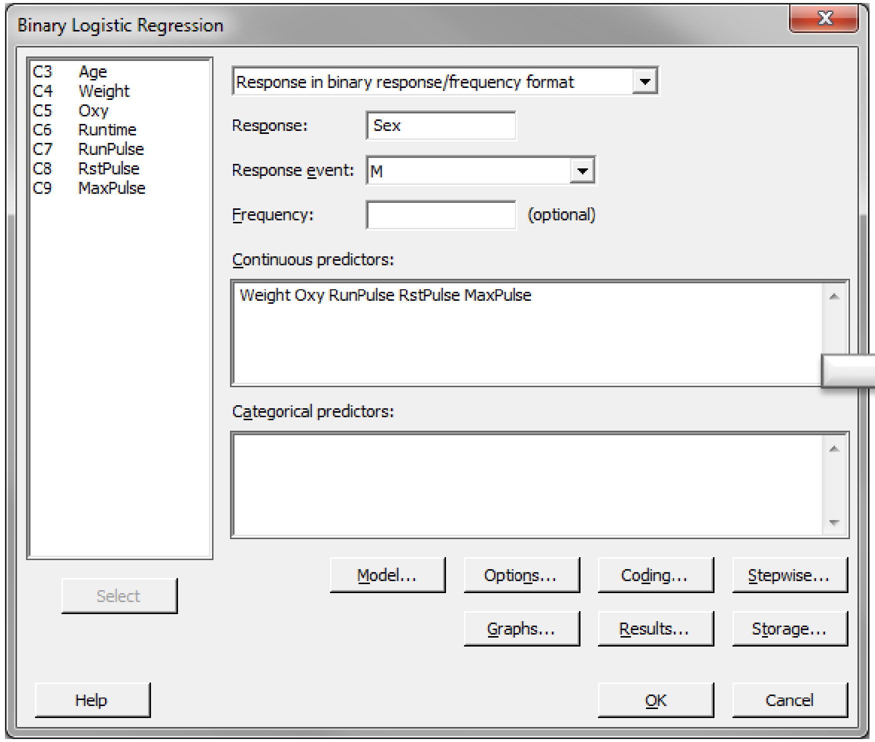
Categorical predictors:



```
Binary Logistic Regression: Sex versus Age, Weight,

Deviance Table

Source      DF   Adj Dev   Adj Mean  Chi-Square  P-Value
Regression   6    9.9889    1.66481        9.99    0.125
  Age        1    0.0008    0.00081        0.00    0.977
  Weight     1    1.3197    1.31966        1.32    0.251
  Oxy        1    3.5633    3.56333        3.56    0.059
  RunPulse   1    1.8091    1.80915        1.81    0.179
  RstPulse   1    0.0625    0.06247        0.06    0.803
  MaxPulse   1    2.6579    2.65790        2.66    0.103
Error       25   34.3725    1.37490
Total       31   44.3614


Model Summary

Deviance   Deviance
    R-Sq  R-Sq(adj)    AIC
  22.52%      8.99%  48.37
```

- After removing Runtime from the model, the p-values of all the independent variables are still higher than the alpha level (0.05).
- We need to continue removing the insignificant independent variables one at a time from the model, starting from the one with the highest p-value.
- Age has the highest p-value (0.977), it will be removed from the model next.

How to Run a Logistic Regression in Minitab



Response in binary response/frequency format

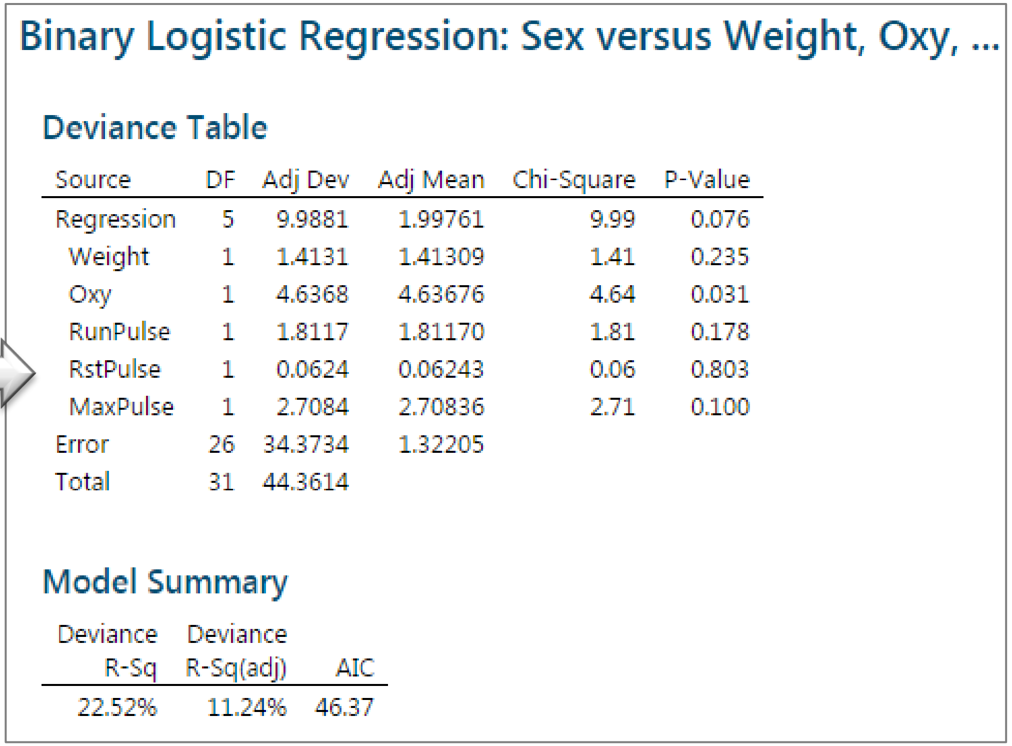
Response: Sex

Response event: M

Frequency: (optional)

Continuous predictors: Weight Oxy RunPulse RstPulse MaxPulse

Categorical predictors:



Binary Logistic Regression: Sex versus Weight, Oxy, ...

Deviance Table

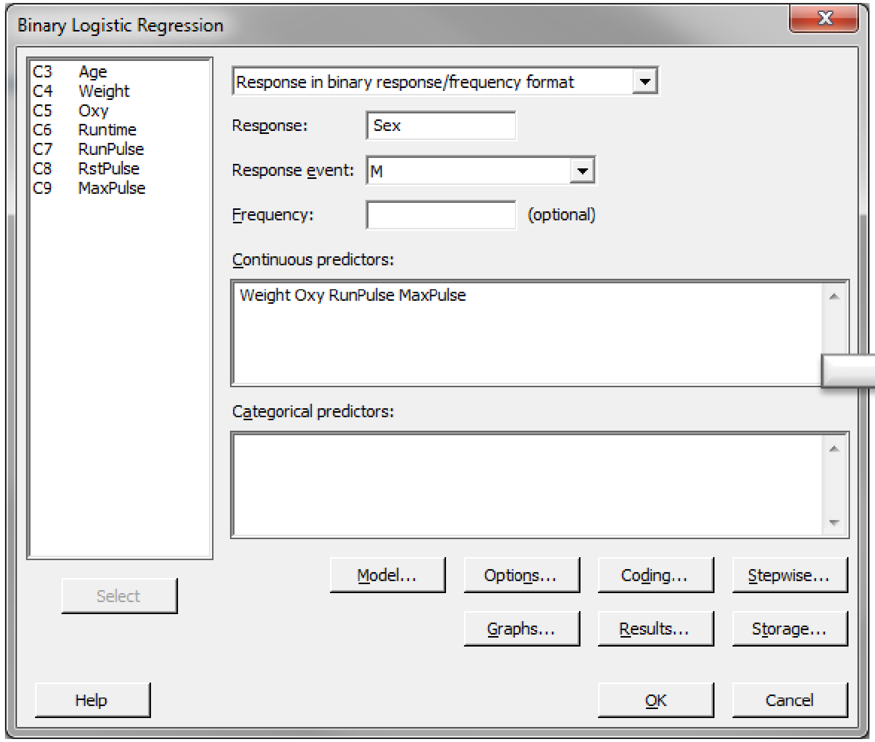
Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	5	9.9881	1.99761	9.99	0.076
Weight	1	1.4131	1.41309	1.41	0.235
Oxy	1	4.6368	4.63676	4.64	0.031
RunPulse	1	1.8117	1.81170	1.81	0.178
RstPulse	1	0.0624	0.06243	0.06	0.803
MaxPulse	1	2.7084	2.70836	2.71	0.100
Error	26	34.3734	1.32205		
Total	31	44.3614			

Model Summary

Deviance R-Sq	Deviance R-Sq(adj)	AIC
22.52%	11.24%	46.37

- After removing Age from the model, the p-values of all the independent variables are still higher than the alpha level (0.05).
- We need to continue removing the insignificant independent variables one at a time from the model, starting from the one with the highest p-value.
- RstPulse has the highest p-value (0.803), it will be removed from the model next.

How to Run a Logistic Regression in Minitab


Binary Logistic Regression

C3 Age
C4 Weight
C5 Oxy
C6 Runtime
C7 RunPulse
C8 RstPulse
C9 MaxPulse

Response in binary response/frequency format

Response: Sex

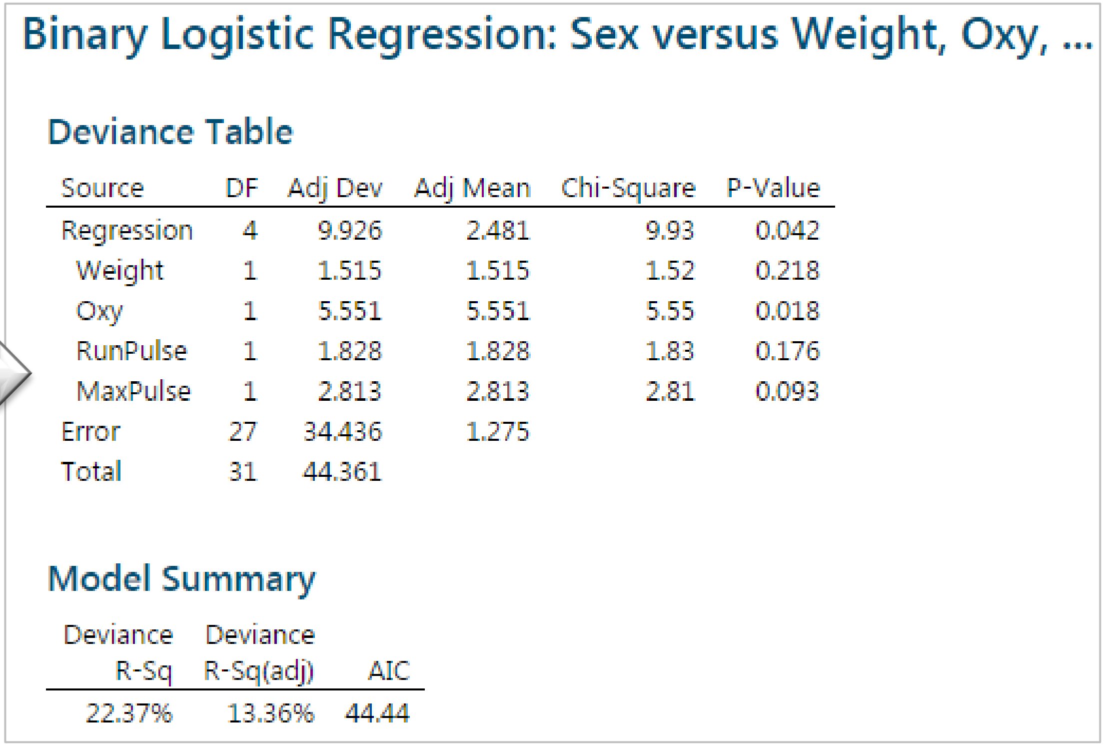
Response event: M

Frequency: (optional)

Continuous predictors:

Weight Oxy RunPulse MaxPulse

Categorical predictors:


Binary Logistic Regression: Sex versus Weight, Oxy, ...

Deviance Table

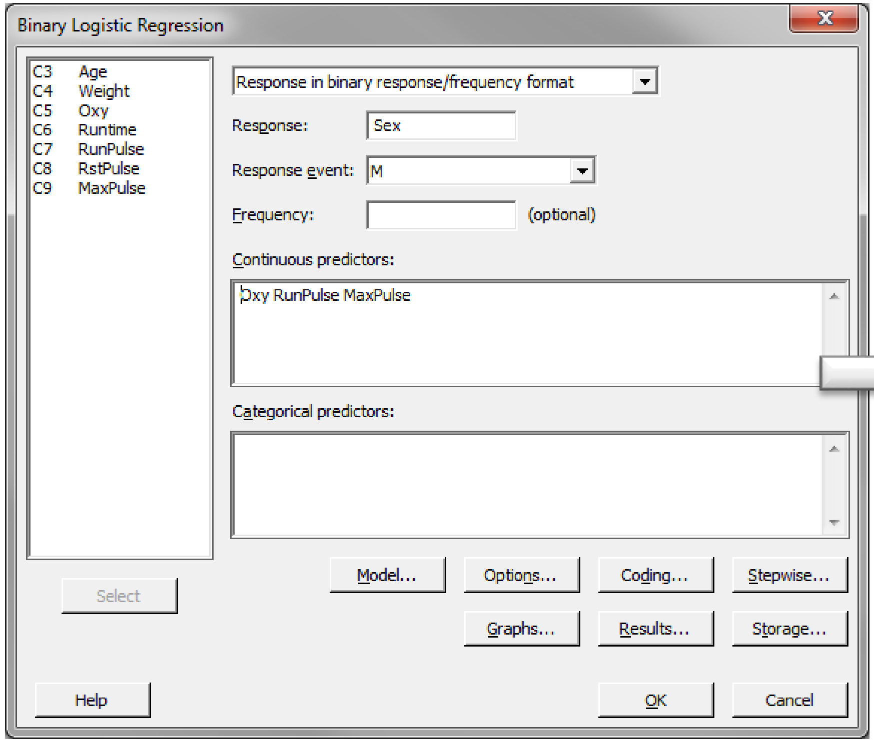
Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	4	9.926	2.481	9.93	0.042
Weight	1	1.515	1.515	1.52	0.218
Oxy	1	5.551	5.551	5.55	0.018
RunPulse	1	1.828	1.828	1.83	0.176
MaxPulse	1	2.813	2.813	2.81	0.093
Error	27	34.436	1.275		
Total	31	44.361			

Model Summary

Deviance R-Sq	Deviance R-Sq(adj)	AIC
22.37%	13.36%	44.44

- After removing RstPulse from the model, the p-values of all the independent variables are still higher than the alpha level (0.05).
- We need to continue removing the insignificant independent variables one at a time from the model, starting from the one with the highest p-value.
- Weight has the highest p-value (0.218), so it would be removed from the model next.

How to Run a Logistic Regression in Minitab



C3 Age
C4 Weight
C5 Oxy
C6 Runtime
C7 RunPulse
C8 RstPulse
C9 MaxPulse

Response in binary response/frequency format

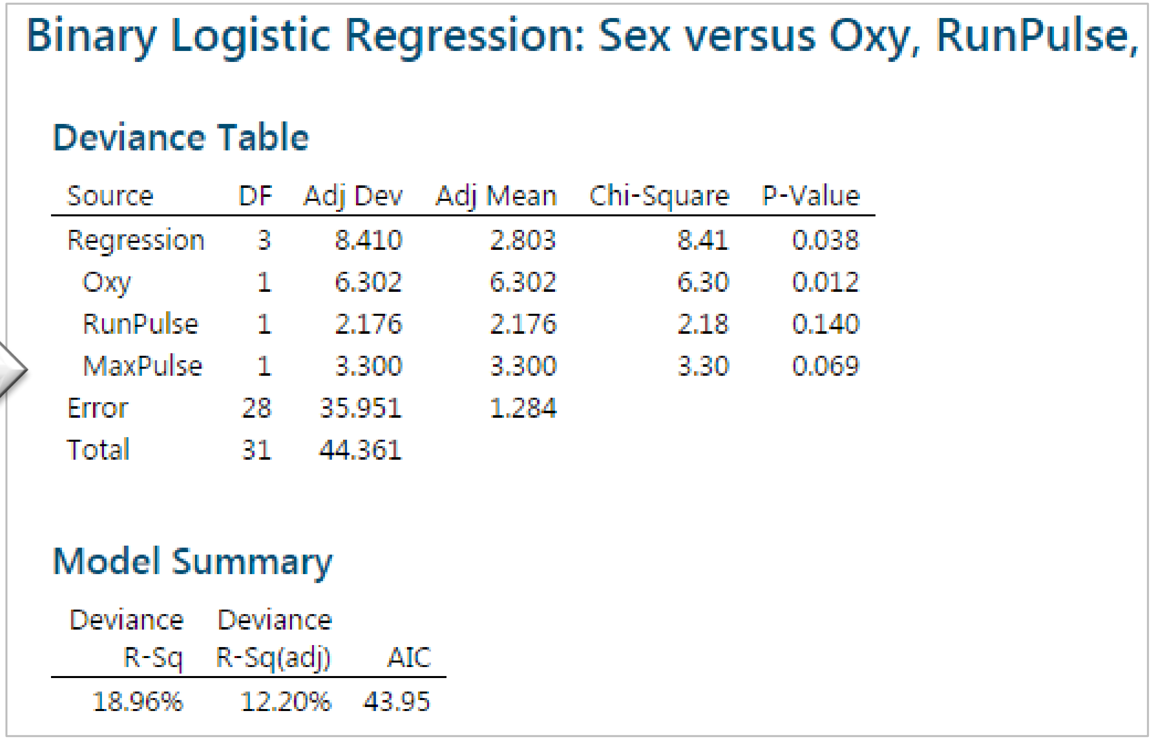
Response: Sex

Response event: M

Frequency: (optional)

Continuous predictors: Oxy RunPulse MaxPulse

Categorical predictors:



Binary Logistic Regression: Sex versus Oxy, RunPulse,

Deviance Table

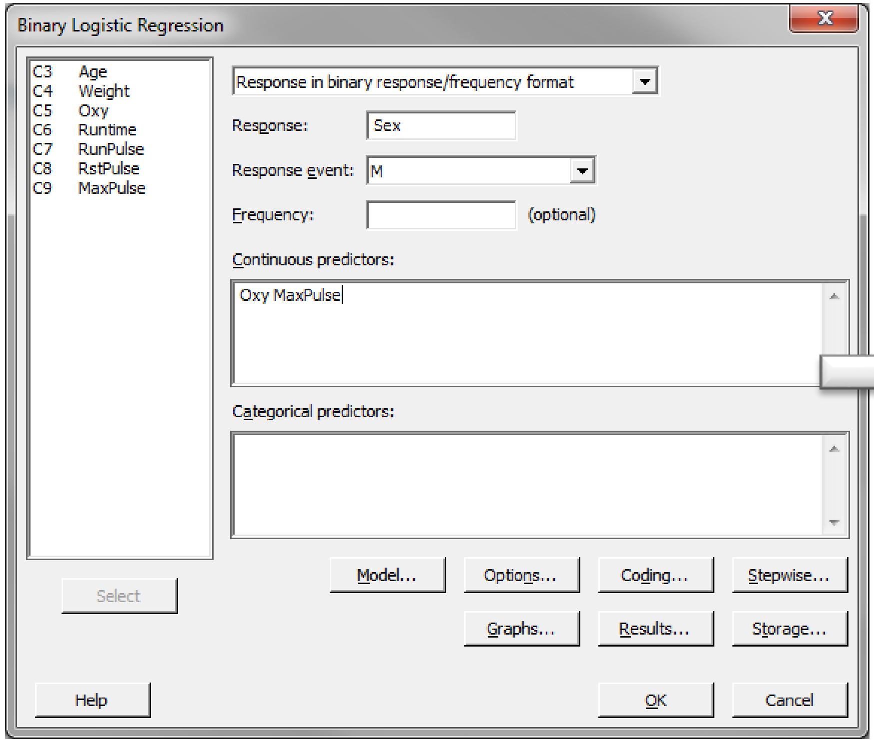
Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	3	8.410	2.803	8.41	0.038
Oxy	1	6.302	6.302	6.30	0.012
RunPulse	1	2.176	2.176	2.18	0.140
MaxPulse	1	3.300	3.300	3.30	0.069
Error	28	35.951	1.284		
Total	31	44.361			

Model Summary

Deviance R-Sq	Deviance R-Sq(adj)	AIC
18.96%	12.20%	43.95

- After removing Weight from the model, the p-values of all the independent variables are still higher than the alpha level (0.05).
- We need to continue removing the insignificant independent variables one at a time from the model, starting from the one with the highest p-value.
- RunPulse has the highest p-value (0.140), it will be removed from the model next.

How to Run a Logistic Regression in Minitab



Response in binary response/frequency format

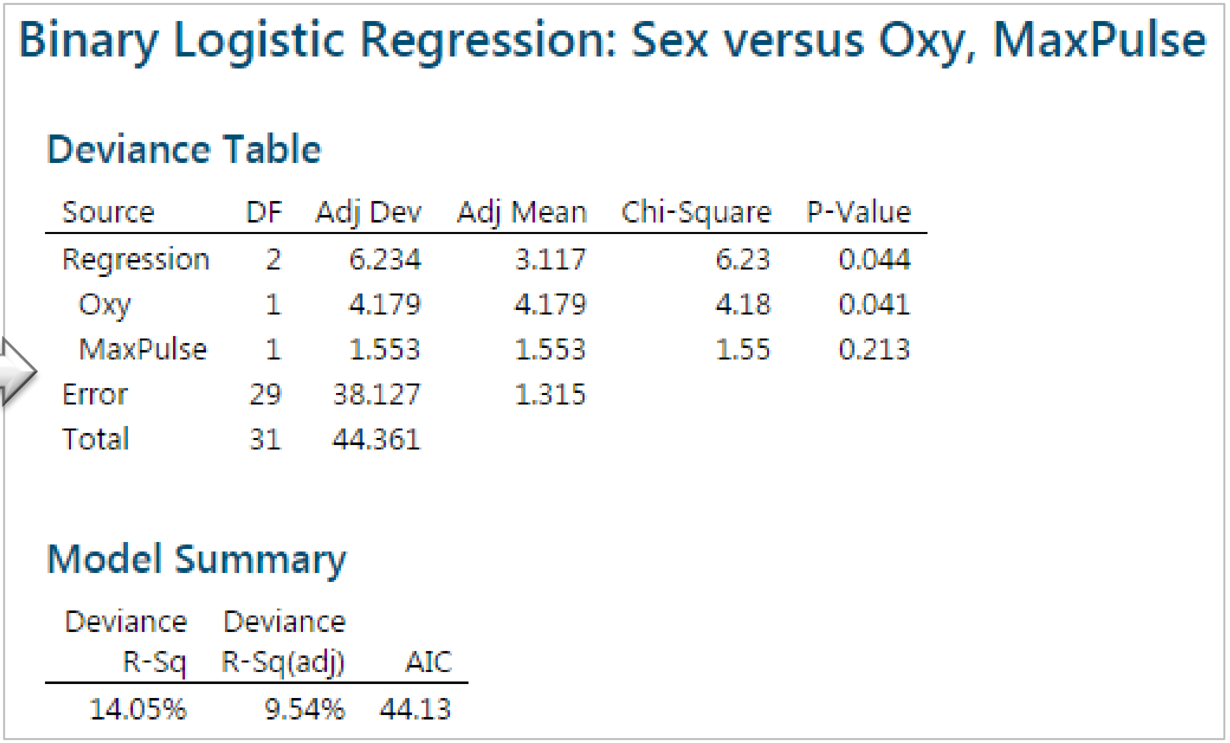
Response: Sex

Response event: M

Frequency: (optional)

Continuous predictors: Oxy MaxPulse

Categorical predictors:



Binary Logistic Regression: Sex versus Oxy, MaxPulse

Deviance Table

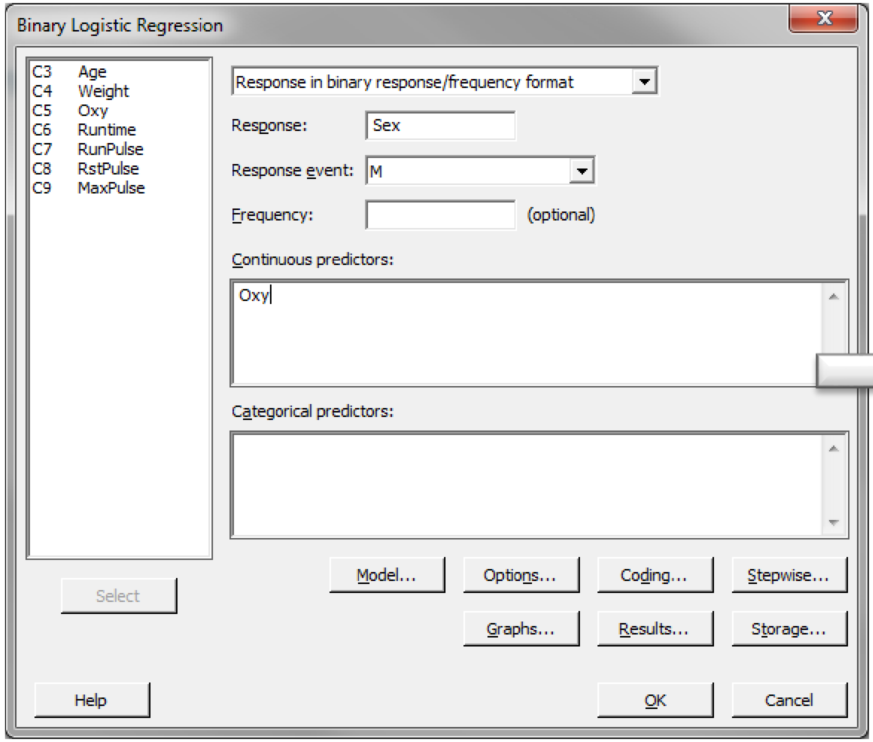
Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	2	6.234	3.117	6.23	0.044
Oxy	1	4.179	4.179	4.18	0.041
MaxPulse	1	1.553	1.553	1.55	0.213
Error	29	38.127	1.315		
Total	31	44.361			

Model Summary

Deviance R-Sq	Deviance R-Sq(adj)	AIC
14.05%	9.54%	44.13

- After removing RunPulse from the model, the p-values of all the independent variables are still higher than the alpha level (0.05).
- We need to continue removing the insignificant independent variables one at a time from the model, starting from the one with the highest p-value.
- MaxPulse has the highest p-value (0.213), it will be removed from the model next.

How to Run a Logistic Regression in Minitab



Response in binary response/frequency format

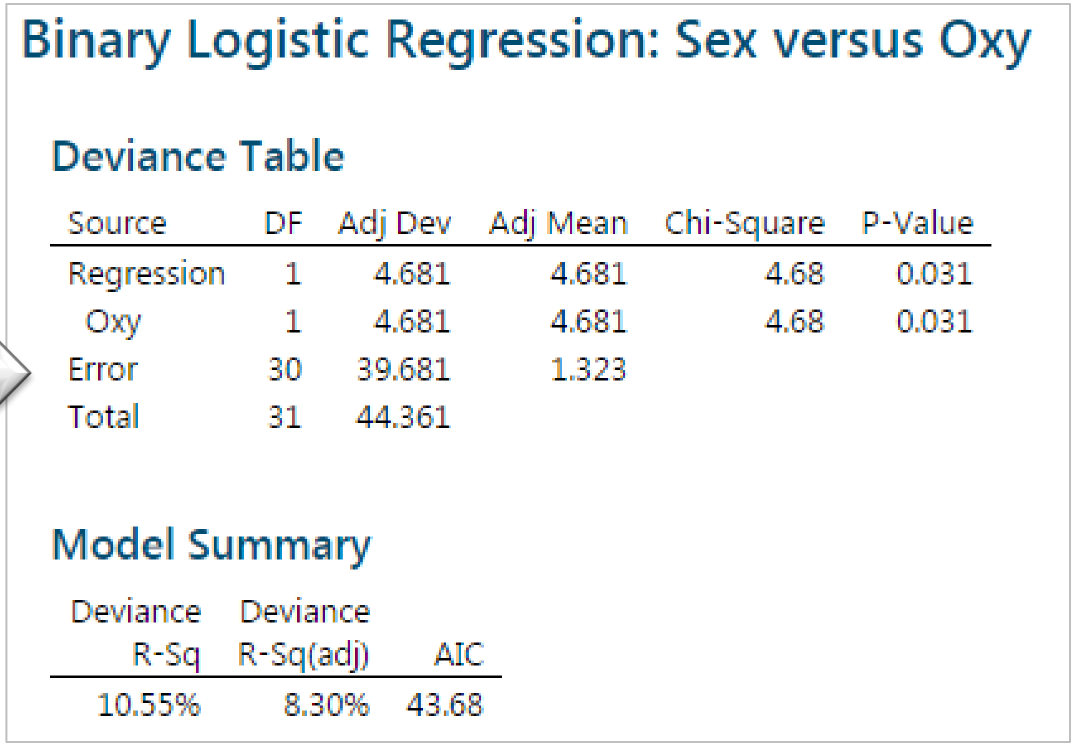
Response: Sex

Response event: M

Frequency: (optional)

Continuous predictors: Oxy

Categorical predictors:



Binary Logistic Regression: Sex versus Oxy

Deviance Table

Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	1	4.681	4.681	4.68	0.031
Oxy	1	4.681	4.681	4.68	0.031
Error	30	39.681	1.323		
Total	31	44.361			

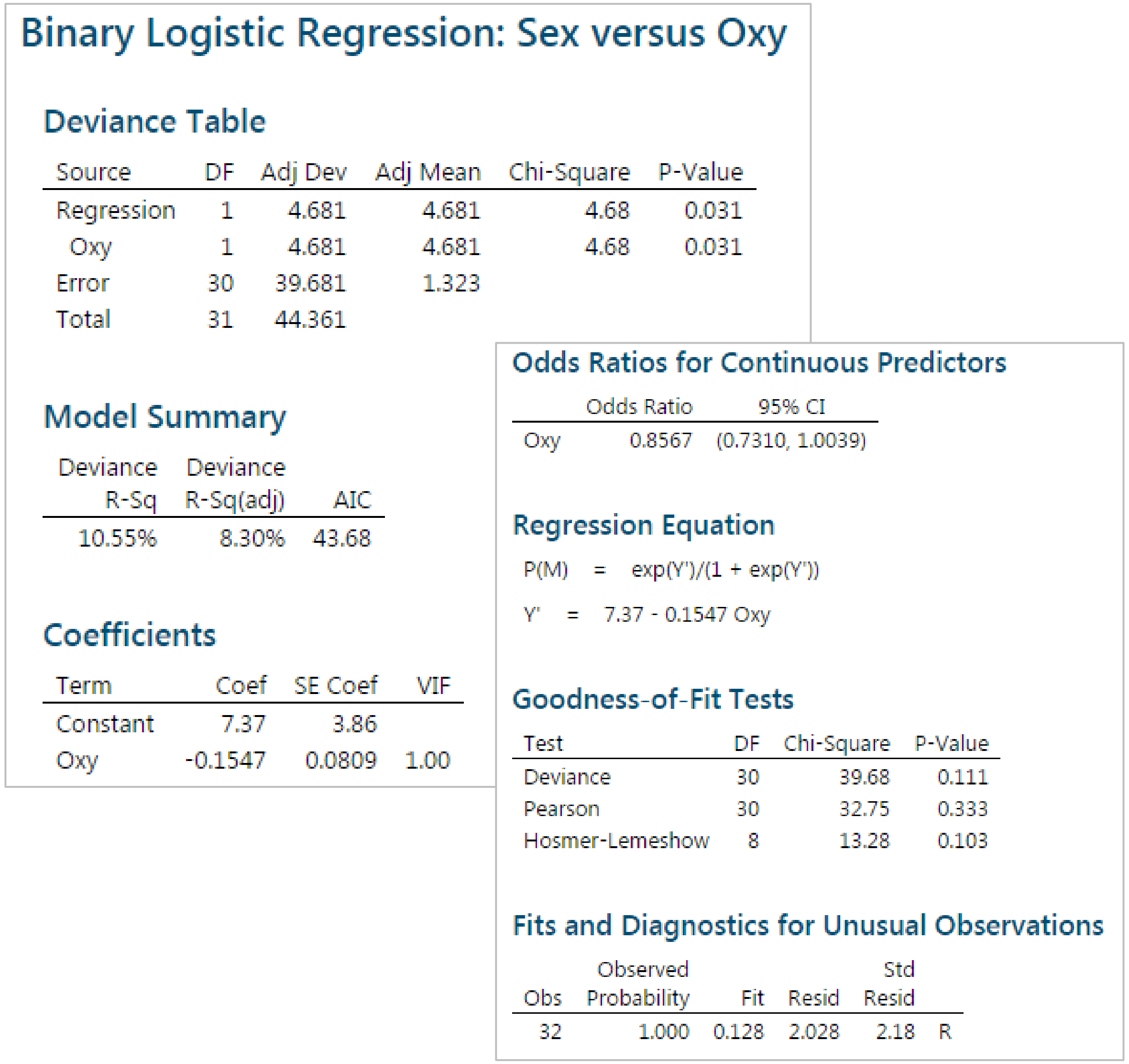
Model Summary

Deviance R-Sq	Deviance R-Sq(adj)	AIC
10.55%	8.30%	43.68

- **After removing MaxPulse from the model, the p-value of the only independent variable “Oxy” is lower than the alpha level (0.05). There is no need to remove “Oxy” from the model.**

How to Run a Logistic Regression in Minitab

- Step 3:
  - Analyze the binary logistic report in the session window and check the performance of the logistic regression model.
  - The p-value here is 0.031, smaller than alpha level (0.05). We conclude that at least one of the slope coefficients is not equal to zero.
  - The p-value of goodness of fit tests are all higher than alpha level (0.05). We conclude that the model fits the data.



Binary Logistic Regression: Sex versus Oxy

Deviance Table

Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	1	4.681	4.681	4.68	0.031
Oxy	1	4.681	4.681	4.68	0.031
Error	30	39.681	1.323		
Total	31	44.361			

Model Summary

Deviance R-Sq	Deviance R-Sq(adj)	AIC
10.55%	8.30%	43.68

Coefficients

Term	Coef	SE Coef	VIF
Constant	7.37	3.86	
Oxy	-0.1547	0.0809	1.00

Odds Ratios for Continuous Predictors

	Odds Ratio	95% CI
Oxy	0.8567	(0.7310, 1.0039)

Regression Equation

P(M) = exp(Y')/(1 + exp(Y'))

Y' = 7.37 - 0.1547 Oxy

Goodness-of-Fit Tests

Test	DF	Chi-Square	P-Value
Deviance	30	39.68	0.111
Pearson	30	32.75	0.333
Hosmer-Lemeshow	8	13.28	0.103

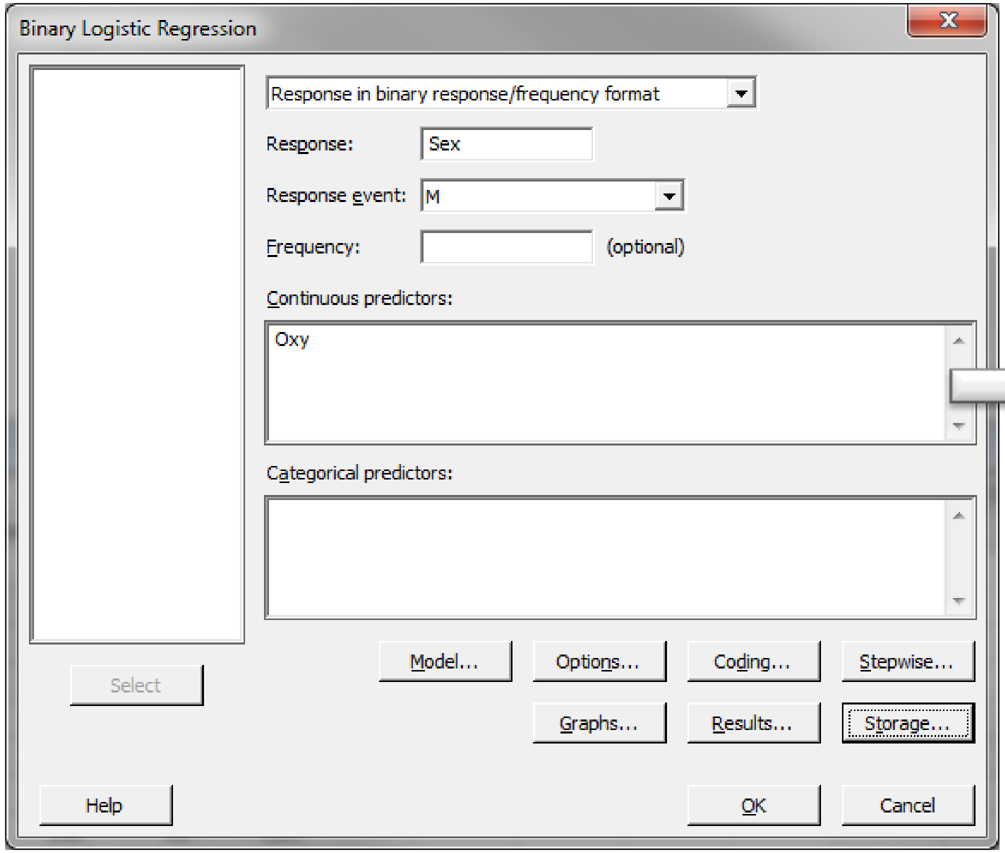
Fits and Diagnostics for Unusual Observations

Obs	Observed Probability	Fit	Resid	Std Resid	
32	1.000	0.128	2.028	2.18	R

How to Run a Logistic Regression in Minitab

- Step 4: get the predicted probabilities of the event (i.e., Sex = M) occurring using the logistic regression model.
    1. Click the “Storage” button in the window named “Binary Logistic Regression” and a new window named “Binary Logistic Regression – Storage” pops up.
    2. Check the box “Fits (event probability).”
    3. Click “OK” in the window of “Binary Logistic Regression – Storage.”
    4. Click “OK” in the window of “Binary Logistic Regression.”
    5. A column of the predicted event probability is added to the data table with the heading “FITS1”.

How to Run a Logistic Regression in Minitab



Response in binary response/frequency format

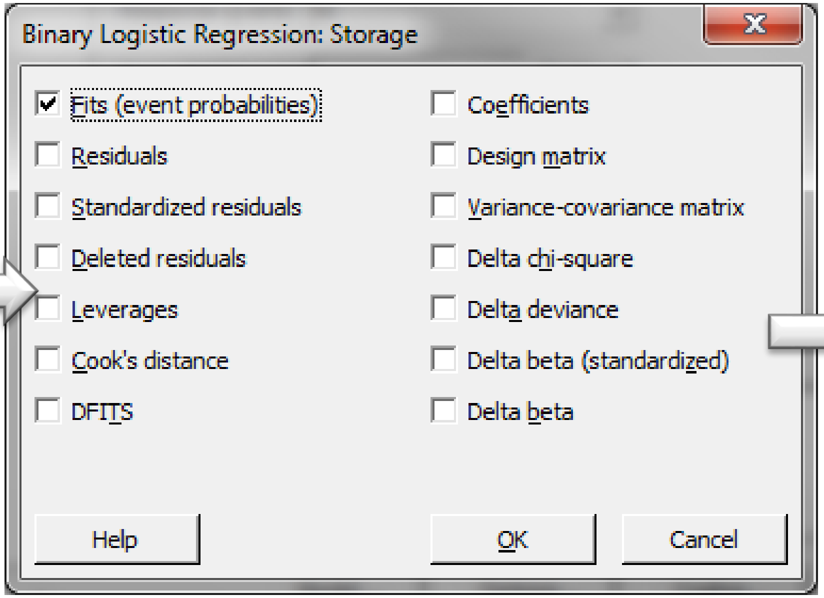
Response: Sex

Response event: M

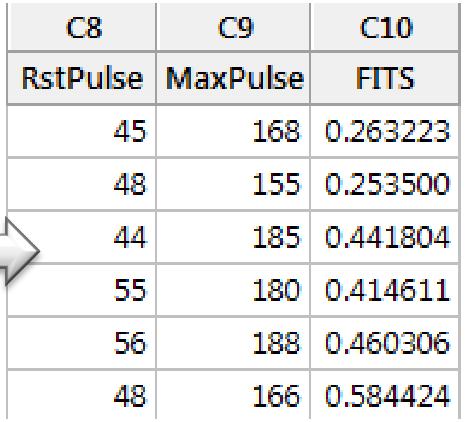
Frequency: (optional)

Continuous predictors: Oxy

Categorical predictors:



☑ Fits (event probabilities)
☐ Residuals
☐ Standardized residuals
☐ Deleted residuals
☐ Leverages
☐ Cook's distance
☐ DFITS
☐ Coefficients
☐ Design matrix
☐ Variance-covariance matrix
☐ Delta chi-square
☐ Delta deviance
☐ Delta beta (standardized)
☐ Delta beta



C8	C9	C10
RstPulse	MaxPulse	FITS
45	168	0.263223
48	155	0.253500
44	185	0.441804
55	180	0.414611
56	188	0.460306
48	166	0.584424

4.3 Designed Experiments

Black Belt Training: Improve Phase

4.1 Simple Linear Regression
- 4.1.1 Correlation
- 4.1.2 X-Y Diagram
- 4.1.3 Regression Equations
- 4.1.4 Residuals Analysis

4.2 Multiple Regression Analysis
- 4.2.1 Non-Linear Regression
- 4.2.2 Multiple Linear Regression
- 4.2.3 Confidence Intervals
- 4.2.4 Residuals Analysis
- 4.2.5 Data Transformation, Box Cox
- 4.2.6 Stepwise Regression
- 4.2.7 Logistic Regression

4.3 Designed Experiments
- 4.3.1 Experiment Objectives
- 4.3.2 Experimental Methods
- 4.3.3 DOE Design Considerations

4.4 Full Factorial Experiments
- 4.4.1 2k Full Factorial Designs
- 4.4.2 Linear and Quadratic Models
- 4.4.3 Balanced and Orthogonal Designs
- 4.4.4 Fit, Model, and Center Points

4.5 Fractional Factorial Experiments
- 4.5.1 Designs
- 4.5.2 Confounding Effects
- 4.5.3 Experimental Resolution

4.3.1 Experiment Objectives

What is an Experiment?

- An **experiment** is a scientific exercise to gather data to test a hypothesis, theory, or previous results.

- Experiments are planned studies in which data are collected actively and purposefully.

- A typical experiment follows this sequence:
    - A problem arises.
    - A hypothesis is stated.
    - An experiment is designed and implemented to collect data.
    - The analysis is performed to test the hypothesis.
    - Conclusions are drawn based on the analysis results.

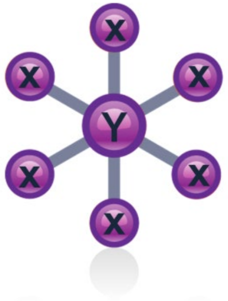
Why Use Experiments?

- Resolve Problems
  - Eliminate defects or defectives.
  - Shift performance means to meet customer expectations.
  - Squeeze variation, making process more predictable.
- Optimize Performance
  - Use statistical methods to get desired process response.
  - Minimize undesirable conditions, waste, or costs.
  - Reduce variation to eliminate out-of-spec conditions.
- Intelligent Design
  - Design processes around variables that have significant impact on process outputs.
  - Design processes around variables that are easier or more cost effective to manage.

Traditional Methods of Learning

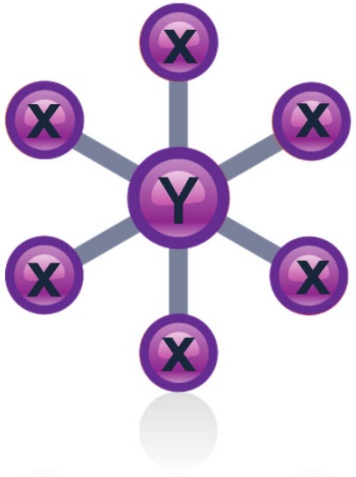
- Passive Learning
  - Study or observe events while they occur.
  - Study or analyze events after they occur.

- OFAT Experiment (One Factor at a Time)
  - An experimental style limited to controlling one factor independently while others remain either static or are allowed to vary.



- Design of Experiments
  - Purposefully and proactively manipulate variables so their effect on the dependent variable can be studied.
  - Encourages an informative event to occur within specific planned parameters.

What is the Design of Experiment (DOE)?



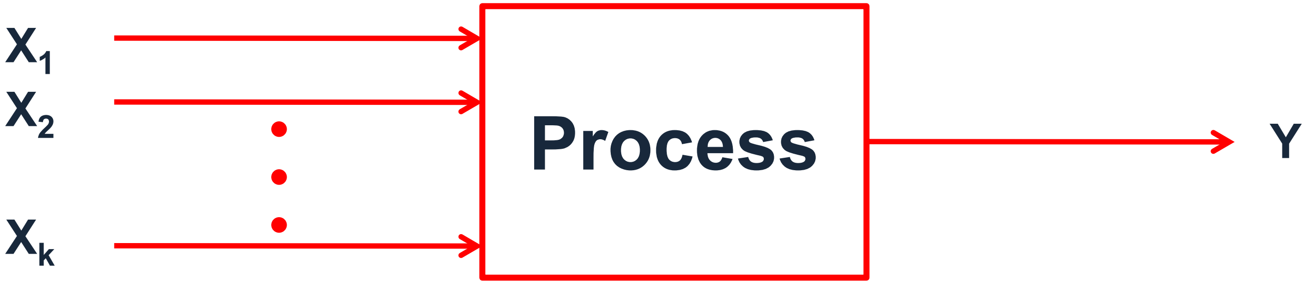
- **Design of experiment** (DOE) is a systematic and cost effective approach to plan the exercises necessary to collect data and quantify the cause-and-effect relationship between the response and different factors.

- In DOE, we observe the changes occurring in the output (i.e., response, dependent variable) of a process by changing one or multiple inputs (i.e., factors, independent variables).
- Using DOE we are able to manipulate several factors with every run and determine their effect on the “Y” or output.

Common DOE Terminology

- **Response**: Y, dependent variable, process output measurement
- **Effect**: The change in the average response across two levels of a factor or between experimental conditions.
- **Factor:** X's, inputs, independent variables
- **Fixed Factor**: Factor that can be controlled during the study
- **Random Factor**: Factor that cannot be controlled during the study
- **Level**: factor settings, usually high and low, + and -, 1 and -1
- **Treatment Combination**: setting of all factors to obtain one response measurement (also referred to as a "run")
- **Replication**: Running the same treatment combination more than once (sequence of runs is typically randomized)
- **Repeat**: Non-randomized replicate of all treatment combinations
- **Inference Space**: Operating range of factors under study

Y = f (Xs)



- In Design of Experiments (DOEs);
  - X_1, X_2, …, X_k are the inputs of the process. They can be either continuous or discrete variables.
  - Y is the output of the process. Y is a continuous variable.
  - A single X or a group of Xs potentially have statistically significant impact on the Y with different levels of influence.

Note: DOE for discrete Y is not covered in this module.

Objectives of Experiments

- Active Learning:
  - Learn as much as possible with as little resources as possible. Efficiency is the name of the game.
- Identify Critical X’s:
  - Identify the critical factors that drive the output or dependent variable (“Y”).
- Quantify Relationships:
  - Generate an equation that characterizes the relationship between the X’s and the Y.
- Optimize:
  - Determine the required settings of all the X’s to achieve the optimal output or response.
- Validate:
  - Prove results through confirmation or pilot implementation before fully implementing solutions.

Principle of Parsimony

- The ultimate goal of DOE is to use the minimum amount of data to discover the maximum amount of information.

- The experiment is designed to obtain data as parsimoniously as possible to draw as much information about the relationship between the response and factors as possible.

- DOE is often used when the resources to collect the data and build the model are limited.

Tradeoffs

- The objective of a specific DOE study should be determined by the team.

- The more complex the objective is, the more data are required. As a result, more time and money are needed to collect the data.

- By prioritizing the objectives and filtering factors which obviously do not have any effect on the response, a considerable amount of cost of running the DOE would be reduced.

4.3.2 Experimental Methods

Planning a DOE

1. **Problem Statement**: Quantifiably define the problem
2. **Objective**: State the objective of the experiment
3. **Primary Metric**: Determine the response variable or “Y”
4. **Key Process Input Variables (KPIV)**: Select the input variable(s) or X’s
5. **Factor Levels**: Determine level settings for all factors
6. **Design**: Select the experimental design for your DOE
7. **Project Plan**: Prepare a project plan for your experiment – human and capital resource allocation, timing, duration etc.
8. **Gain Buy-in**: Your experiment will need support and buy-in from several places, get it!
9. **Run the experiment**
10. **Analyze, Interpret, and Share Results:**

DOE Problem Statement

- Much like the overall problem statement discussed in the Define phase, a DOE problem statement should specific and have quantifiable elements.

- Make sure your problem statement ties to a business goal or business performance indicator.

- Problem statements should *not* include:
  - Conclusions
  - Solutions
  - Causes.

DOE Objectives

State the objective of the experiment

- DOE Objective Examples:
  - “To determine the effects of tire pressure on gas mileage”
  - “To determine which factors have the most impact on gas mileage”
  - “To determine the effects of a pick-up truck’s tailgate (up vs. down) on the truck’s gas mileage”
  - “To determine which HVAC and Lighting factors are contributing most to Kwh consumption”

- DOE objective statements should:
  - Clearly define what you want to learn from the experiment
  - Declare study parameters or scope.

DOE Primary Metric

- Recall our discussion of primary metrics in the Define phase?
  - We said something to this effect: “the primary metric is the most important measure of success.”
- This statement still holds true when discussing the primary metric for a DOE.
- The primary metric for a DOE is your experiment’s output or response under study.
- It is the “Y” in your transfer function $Y = f(x)$.
- It is critical to understand the following things about your metric:
  - Measurement accuracy, R&R, and/or bias
  - Measurement frequency
  - Measurement resolution.

DOE Primary Metric

- Before embarking on a design of experiment you should have validated your measurement system.
- If you have not, now would be a wise time to back track and do so.
- DOEs can be costly on a firm's time and resources.

- DOEs not planned and run properly will create waste and defective products because treatment combinations should "test" boundaries.

- Do not allow your entire study to be wasted because of a poor measurement system or one that can not respond frequently enough for your test.

- Measurement resolution is another key factor to consider for your DOE. Make sure your primary metric can decipher the size of the main effects and interactions you are looking for.

DOE Primary Metric

- There is still even more to understand about your primary metric before jumping into a DOE.
- Be sure to answer these questions before proceeding:

- Is it discrete or continuous?
- Do you want to shift its mean?
- Do you want to squeeze variation?
- What is the baseline?
- Is it in control?
- Is there seasonality, trending, or any cycle?
- How much change do you want to detect?
- Is your metric normally distributed?

DOE Input Variables – KPIVs (key process input variables)

- Factor selection is an important element in planning for a design of experiment.
- Factors are the “things” or metrics that your study responds to.
- Factors can be discrete or continuous.
- Factors should have largely been determined through the use of tools and analytics used throughout the DMAIC roadmap.
    - Failure modes and effects analysis
    - Cause and effect diagram (fishbone)
    - X-Y matrix
    - Process mapping
    - Planned studies
    - Passive analytics
    - etc.

DOE Factor Settings

- **Factor settings** are the range of values selected for each factor.
- Factor settings need to be predetermined before every experiment.
  - Let us use this paper airplane as an example. Our factors under study for max distance as a response might be:
    1. Number of folds
    2. Paper weight
    3. Paperclip (yes or no)
    4. Thrust
    5. Launch height
    6. Launch angle etc.

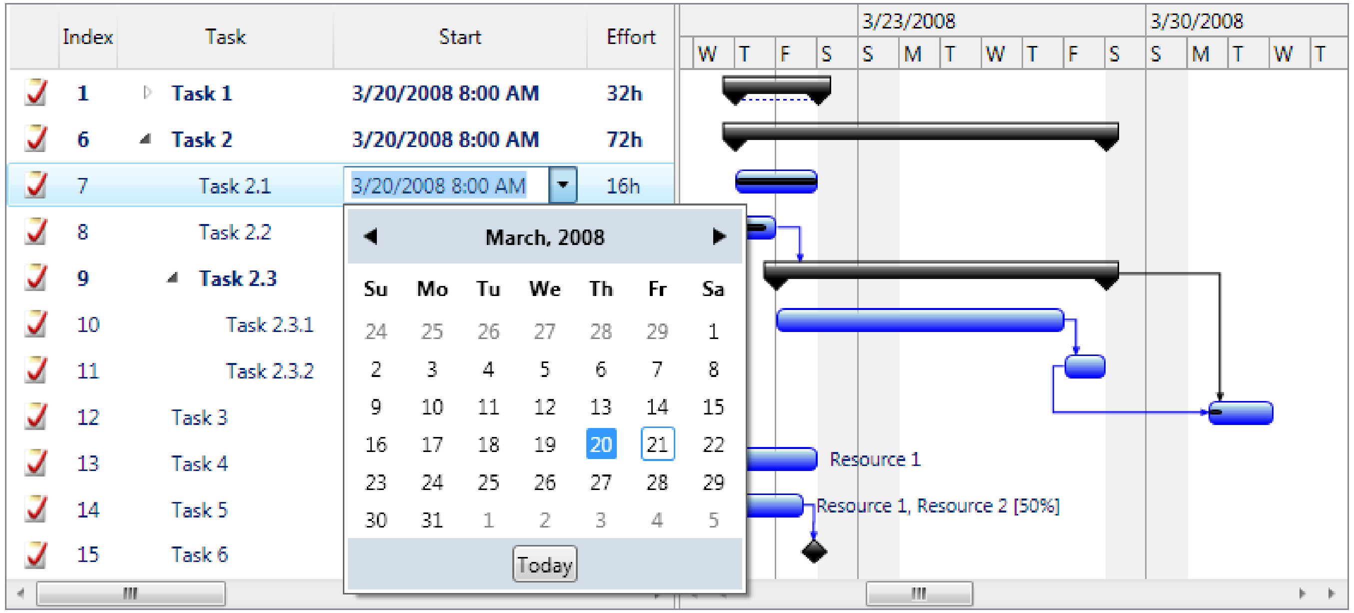
- Factor settings are important because each setting selection should provide value to the study.
  - Understanding the “margin” where that value can be found is where an experienced Black Belt will earn their keep.
  - For example, what would we learn if we chose the test two levels for “number of folds” and we tested zero folds and 80 folds as our two settings for that factor? Right, nothing!

DOE – Design Considerations

- **Design selection** is an aspect of properly planned design of experiments.
- High-level overview of design types (you will learn more about these in later modules)
  - Screening Designs:
    - Fractional factorials
    - Intended to narrow down the many factors to the vital few
    - Screening designs enable many factors to be evaluated at a lower cost because of the nature of the design (main effects)
  - Characterization Designs:
    - Higher resolution fractional factorials
    - Full factorials with fewer variables
    - Main effect and interactions
  - Optimization Designs:
    - Full factorials
    - Response surface designs.

DOE Project Plan

- DOEs can require a fair amount of coordination between human and capital resources.
- Preparing a DOE project plan is a wise step.
- Make sure you identify critical dependencies like interfering with production or operating hours etc.
- Use project management software or, at a minimum, a Gantt chart to help you identify *critical path items*.



Index	Task	Start	Effort
1	**Task 1**	**3/20/2008 8:00 AM**	**32h**
6	**Task 2**	**3/20/2008 8:00 AM**	**72h**
7	Task 2.1	3/20/2008 8:00 AM	16h
8	Task 2.2		
9	**Task 2.3**		
10	Task 2.3.1		
11	Task 2.3.2		
12	Task 3		
13	Task 4		
14	Task 5		
15	Task 6		

March, 2008

Su	Mo	Tu	We	Th	Fr	Sa
24	25	26	27	28	29	1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31	1	2	3	4	5

DOE – Getting Support

- By running a DOE you are about to “discover” something new and better or validate an existing belief or assumption.

(If you do not believe this you should not be running your DOE)

- You are also about to cost your firm money or disrupt someone’ s routine…Get everyone on the same page before moving forward!
- Hopefully, you have already established the framework for the following items. If not, follow these steps:
  1. **Business Case and Justification**
  2. **Stakeholder Analysis** – figure out who is with you and your project and who is not
  3. **Communication Plan** – build and execute your communication plan based on the first two items.

Run Your Experiment

Analyze, Interpret, and Share Your DOE Results

DOE Iterations

- One single DOE might not be enough to completely answer the questions due to the following reasons:
  - None of the potential factors identified are statistically significant to the response.
  - More factors need to be included in the model.
  - Residuals of the model are not performing well.
  - The objectives of DOE changes due to business reasons.

4.3.3 Experiment Design Considerations

Considerations of Experiments

- Objectives of the experiments
- Resource availability to run the experiments
- Potential costs to implement the experiments
- Stakeholders’ support for successful experiments
- Accuracy and precision of the measurement system
- Statistical stability of the process
- Unexpected plans or changes
- Sequence of running the experiments
- Simplicity of the model
- Inclusion of potential significant factors
- Settings of factors
- Well-behaved residuals

4.4 Full Factorial Experiments

Black Belt Training: Improve Phase

4.1 Simple Linear Regression
- 4.1.1 Correlation
- 4.1.2 X-Y Diagram
- 4.1.3 Regression Equations
- 4.1.4 Residuals Analysis

4.2 Multiple Regression Analysis
- 4.2.1 Non-Linear Regression
- 4.2.2 Multiple Linear Regression
- 4.2.3 Confidence Intervals
- 4.2.4 Residuals Analysis
- 4.2.5 Data Transformation, Box Cox
- 4.2.6 Stepwise Regression
- 4.2.7 Logistic Regression

4.3 Designed Experiments
- 4.3.1 Experiment Objectives
- 4.3.2 Experimental Methods
- 4.3.3 DOE Design Considerations

4.4 Full Factorial Experiments
- 4.4.1 2k Full Factorial Designs
- 4.4.2 Linear and Quadratic Models
- 4.4.3 Balanced and Orthogonal Designs
- 4.4.4 Fit, Model, and Center Points

4.5 Fractional Factorial Experiments
- 4.5.1 Designs
- 4.5.2 Confounding Effects
- 4.5.3 Experimental Resolution

4.4.1 2k Full Factorial Designs

DOE Key Terms

- Response: Y, dependent variable, process output measurement
- Effect: The change in the average response across two levels of a factor or between experimental conditions
- Factor: X’s, inputs, independent variables
- Fixed Factor: Factor that can be controlled during the study
- Random Factor: Factor that cannot be controlled during the study
- Factor Levels: Factor settings, usually high and low, + and -, 1 and -1
- Treatment Combination: setting of all factors to obtain one response measurement (also referred to as a “run”)
- Replication: Running the same treatment combination more than once (sequence of runs is typically randomized)
- Repeat: Non-randomized replicate of all treatment combinations
- Inference Space: Operating range of factors under study
- Main Effect: The average change from one level setting to another for a single factor
- Interaction: The combined effect of two factors independent of the main effect of each factor

More About Factors

- Factors in DOE are the potential causes that might have significant impact on the outcome of a process or a product.
- Factors are the inputs of a process or system and the response is the output of it.
- Factors can be continuous or discrete variables.
- Most experiments use two or more factors. Ask subject matter experts' opinions for main factors selection. The more factors in DOE, the more costs triggered.

More About Factors

- The goal of a DOE is to measure the effects of factors and interactions of factors on the outcome.

- In DOE:
  - First, we create a set of different combinations of factors each at different levels.
  - Second, we run the experiment and collect the response values in different scenarios.
  - Third, we analyze the results and test how the response changes accordingly when factors change.

Factor Levels

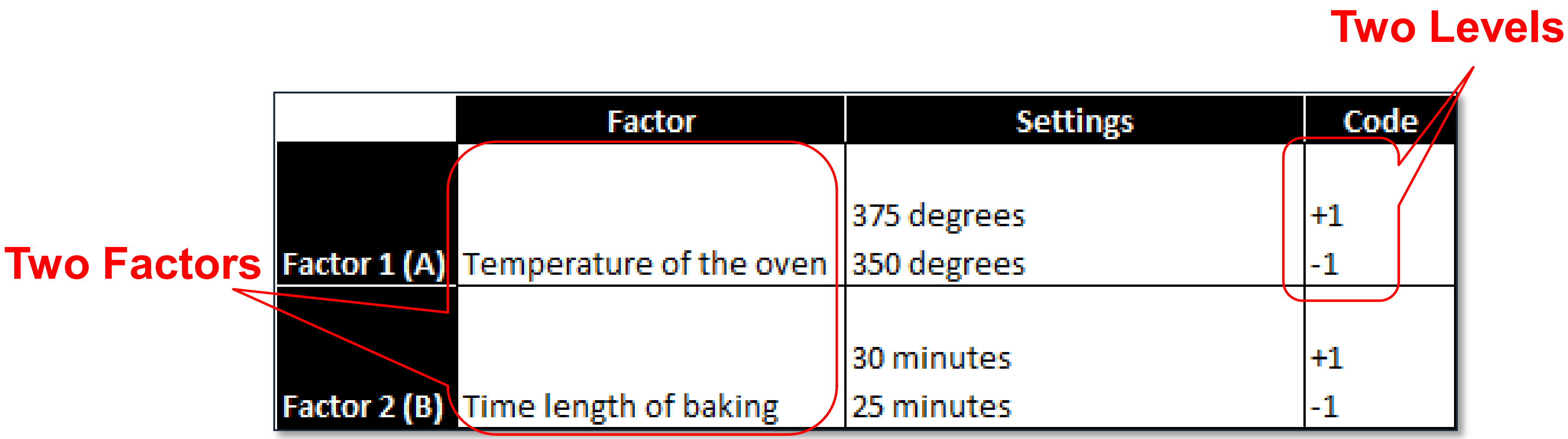
- Factor levels are the selected settings of a factor we are testing in the experiment.

- The more levels a factor has, the more scenarios we have to create and test. As a result, more resources are required to collect samples and the model is more complicated.

- The most popular DOE is two-level design, meaning only two levels for each factor.

Factor Levels

- Code of factor levels
  - High vs. Low
  - (+1) vs. (-1)
  - (+) vs. (-)
- Example of factor levels
  - A study on how two factors affect the taste of cakes.



	Factor	Settings	Code
Factor 1 (A)	Temperature of the oven	375 degrees 350 degrees	+1 -1
Factor 2 (B)	Time length of baking	30 minutes 25 minutes	+1 -1

Two Levels

Two Factors

Treatment Combination

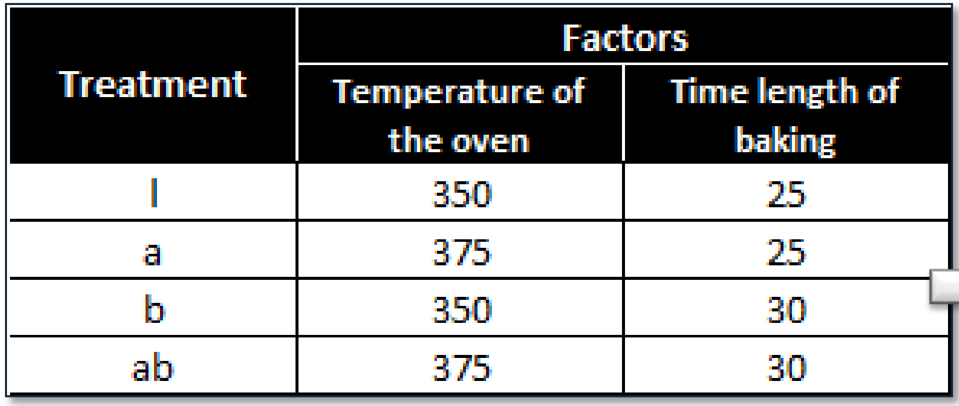
- A treatment is a combination of different factors at different levels.

- It is a unique scenario setting in an experiment.

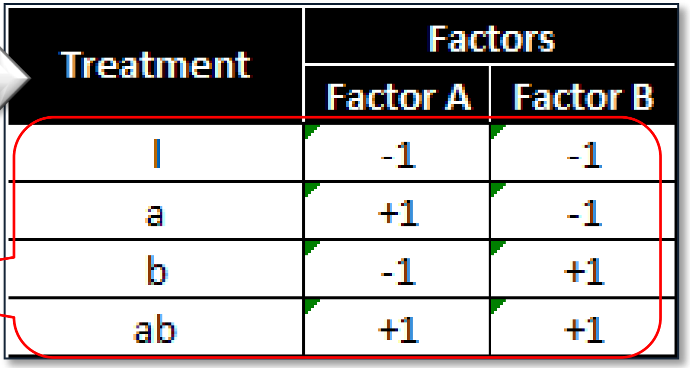
- Coding treatments:
  - If there are two factors in an experiment, we name one factor A and the other factor B.
  - A single treatment in an experiment is named for each factor:
    - --, +-, -+ and ++
    - I, a, b and ab

Treatment Combination

- Example of treatment
  - A study on how two factors affect the taste of cakes.



Treatment	Factors	
	Temperature of the oven	Time length of baking
I	350	25
a	375	25
b	350	30
ab	375	30



Treatment	Factors	
	Factor A	Factor B
I	-1	-1
a	+1	-1
b	-1	+1
ab	+1	+1

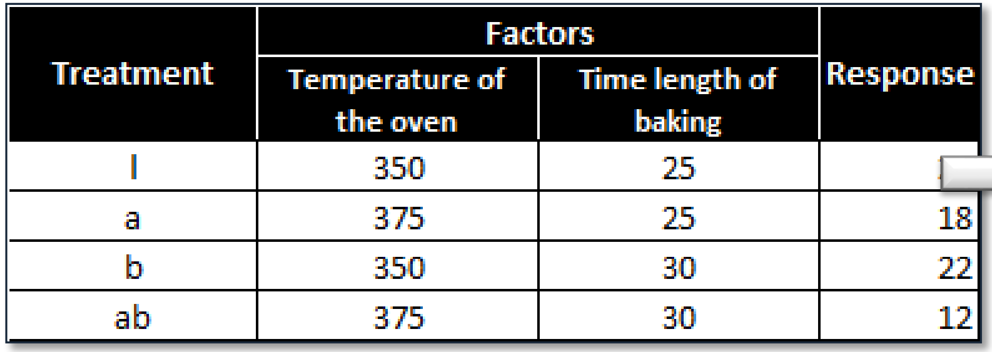
Four treatments

Response

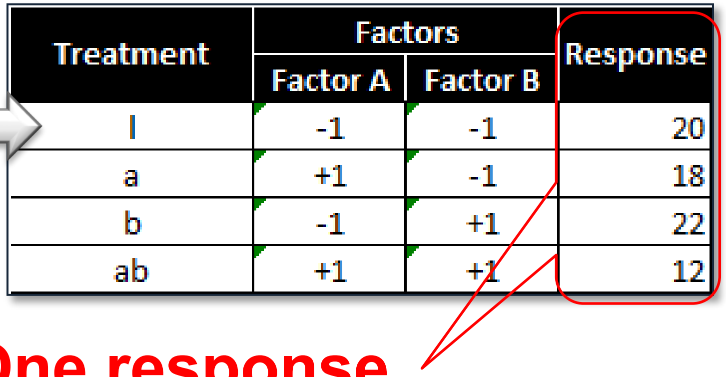
- A response is the output of a process, a system, or a product.
- It is the outcome we are interested to improve by optimizing the settings of factors.
- In an experiment, we observe, record, and analyze the corresponding changes occurring in the response after changing the settings of different factors.

Response

- Example of treatment
  - A study on how two factors affect the taste of cakes. We use a predefined metric to measure the cake's tastiness.
  - After running each treatment, we obtain and record the resulting response value.



Treatment	Factors		Response
	Temperature of the oven	Time length of baking	
I	350	25	[illegible]
a	375	25	18
b	350	30	22
ab	375	30	12



Treatment	Factors		Response
	Factor A	Factor B	
I	-1	-1	20
a	+1	-1	18
b	-1	+1	22
ab	+1	+1	12

One response

Main Effect

- Main effect is the average change in the response resulting from the change in the levels of a factor.
- It captures the effect of a factor alone on the response without taking into consideration other factors.
- In the baking cake example, the main effects of temperature and time length are

$$MainEffect_A = \frac{18+12}{2} - \frac{20+22}{2} = -6$$

$$MainEffect_B = \frac{22+12}{2} - \frac{20+18}{2} = -2$$

Main Effect

- Interpretation of main effect in the example:
  - By changing the temperature level of the oven from high to low, the response (i.e., tastiness of the cake) increases by 6 measurement units.
  - By changing the time length of baking from high to low, the response (i.e., tastiness of the cake) increases by 2 measurement units.

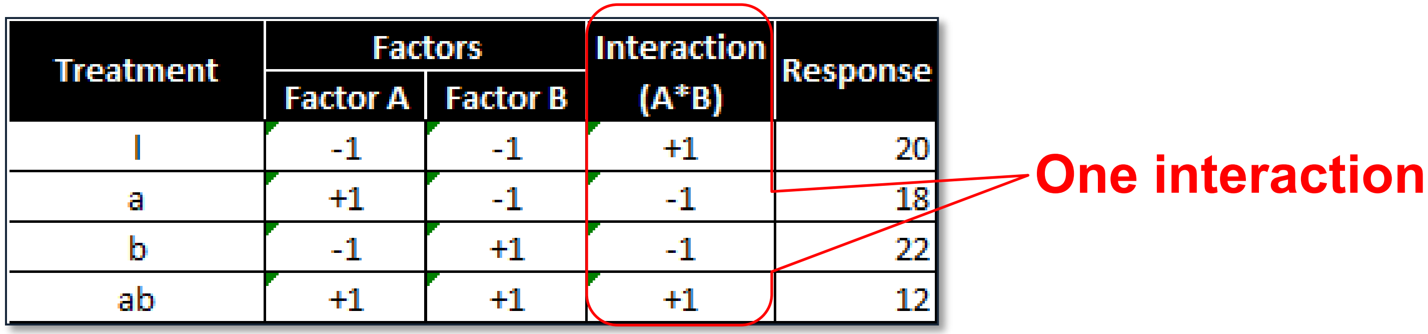
Interaction Effect

- Interaction effect is the average change in the response resulting from the change in the interaction of multiple factors.

- It occurs when the change in the response triggered by one factor depends on the change in other factor(s).

Interaction Effect

- Example of interaction effect



Treatment	Factors		Interaction	Response
	Factor A	Factor B	(A*B)	
I	-1	-1	+1	20
a	+1	-1	-1	18
b	-1	+1	-1	22
ab	+1	+1	+1	12

One interaction

$$InteractionEffect = \frac{20+12}{2} - \frac{22+18}{2} = -4$$

- By changing the interaction from high to low, the response (i.e., tastiness of the cake) increases by 4 measurement units.

2k Full Factorial DOE

- In a full factorial experiment, all of the possible combinations of factors and levels are created and tested.

- For example, for two-level design (i.e., each factor has two levels) with k factors, there are 2k possible scenarios or treatments.
  - 2 factors, each with 2 levels, we have 22 = 4 treatments
  - 3 factors, each with 2 levels, we have 23 = 8 treatments
  - k factors, each with 2 levels, we have 2k treatments

2k Full Factorial DOE

- Full factorial DOE is used to discover the cause-and-effect relationship between the response and both individual factors and the interaction of factors.

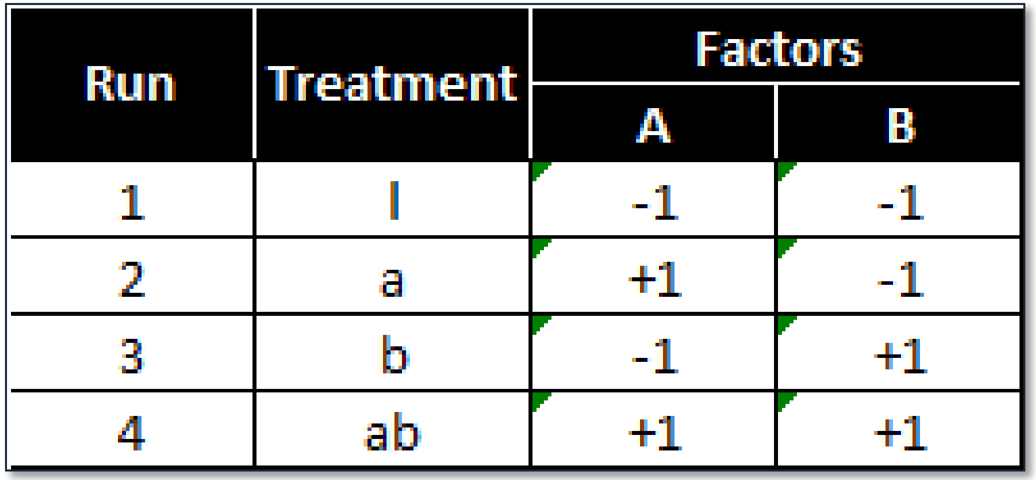
- Generate an equation to describe the relationship between Y and the important Xs:

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_p X_1 X_2 \cdots X_k + \varepsilon$$

where
Y is the response and $X_1, X_2, \ldots, X_k$ are the factors.
α_0 is the intercept and $\alpha_1, \alpha_2, \ldots, \alpha_p$ are the coefficients of the factors and interactions.
ε is the error of the model.

Two-Level Two-Factor Full Factorial

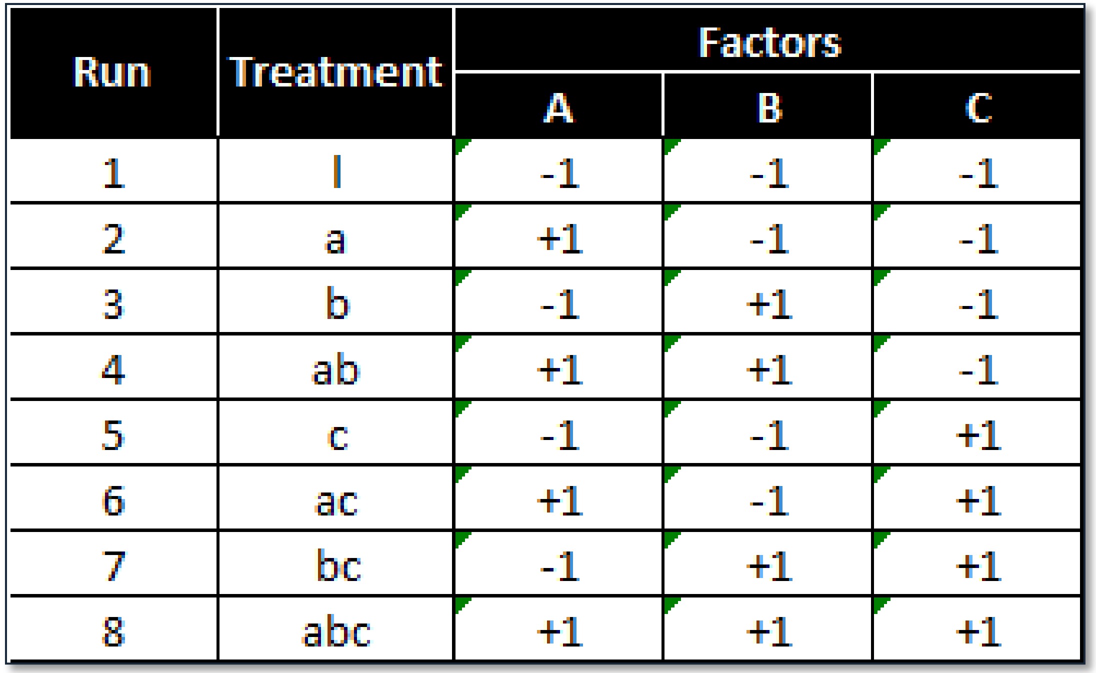
- Below is a design pattern of a two-level two-factor full factorial experiment:



Run	Treatment	Factors	
		A	B
1	l	-1	-1
2	a	+1	-1
3	b	-1	+1
4	ab	+1	+1

Two-Level Three-Factor Full Factorial

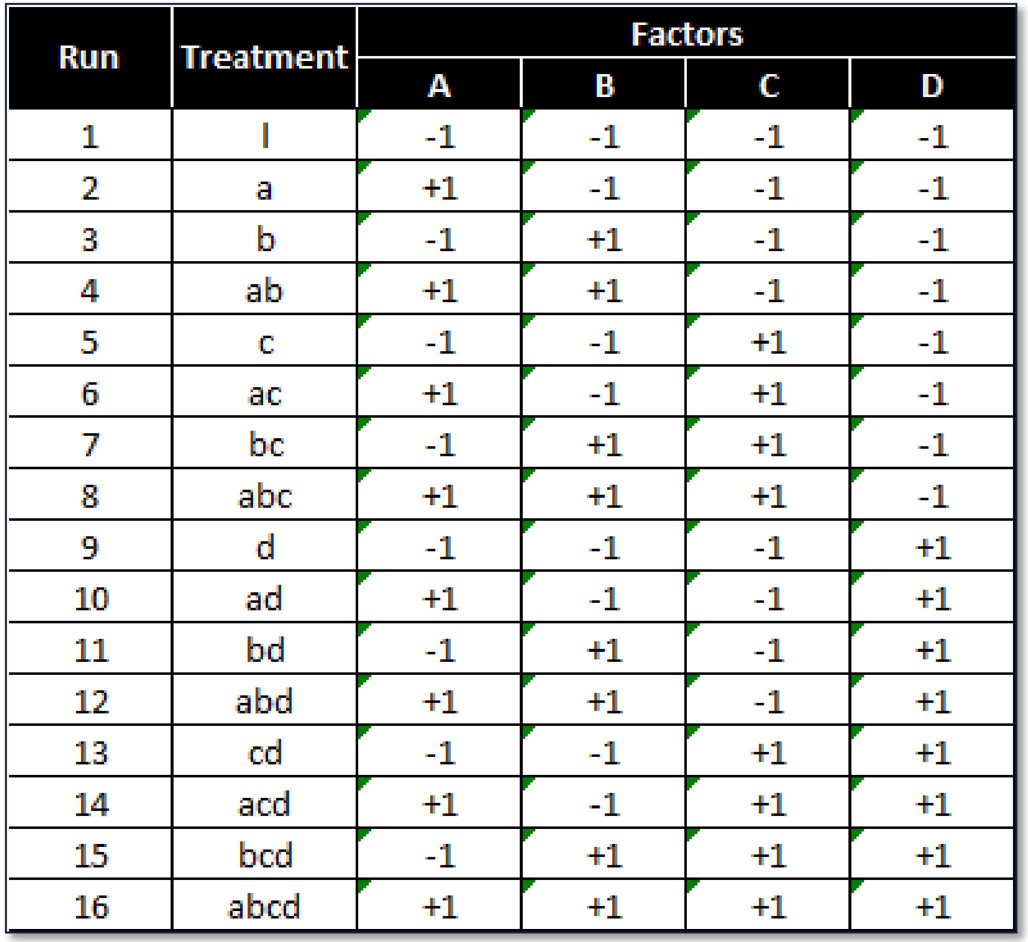
- Below is a design pattern of a two-level three-factor full factorial experiment



Run	Treatment	Factors		
		A	B	C
1	l	-1	-1	-1
2	a	+1	-1	-1
3	b	-1	+1	-1
4	ab	+1	+1	-1
5	c	-1	-1	+1
6	ac	+1	-1	+1
7	bc	-1	+1	+1
8	abc	+1	+1	+1

Two-Level Four-Factor Full Factorial

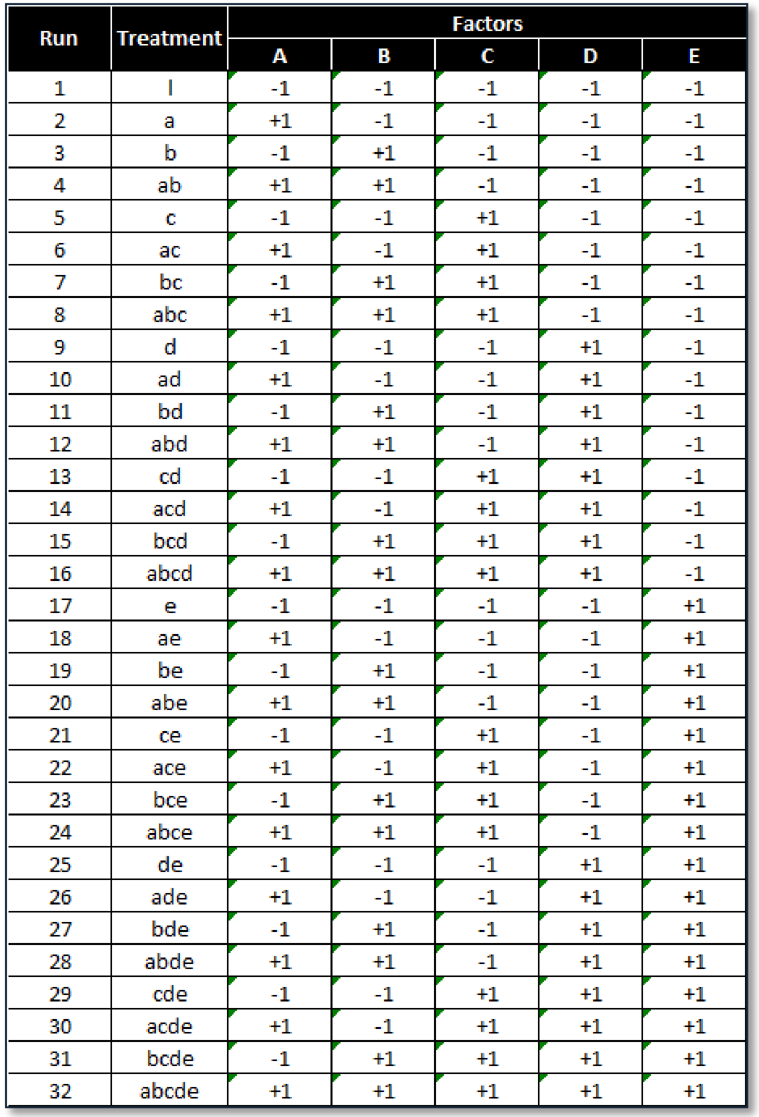
- Below is a design pattern of a two-level four-factor full factorial experiment



Run	Treatment	Factors			
		A	B	C	D
1	l	-1	-1	-1	-1
2	a	+1	-1	-1	-1
3	b	-1	+1	-1	-1
4	ab	+1	+1	-1	-1
5	c	-1	-1	+1	-1
6	ac	+1	-1	+1	-1
7	bc	-1	+1	+1	-1
8	abc	+1	+1	+1	-1
9	d	-1	-1	-1	+1
10	ad	+1	-1	-1	+1
11	bd	-1	+1	-1	+1
12	abd	+1	+1	-1	+1
13	cd	-1	-1	+1	+1
14	acd	+1	-1	+1	+1
15	bcd	-1	+1	+1	+1
16	abcd	+1	+1	+1	+1

Two-Level Five-Factor Full Factorial

- At right is a design pattern of a two-level five-factor full factorial experiment



Run	Treatment	Factors				
		A	B	C	D	E
1	l	-1	-1	-1	-1	-1
2	a	+1	-1	-1	-1	-1
3	b	-1	+1	-1	-1	-1
4	ab	+1	+1	-1	-1	-1
5	c	-1	-1	+1	-1	-1
6	ac	+1	-1	+1	-1	-1
7	bc	-1	+1	+1	-1	-1
8	abc	+1	+1	+1	-1	-1
9	d	-1	-1	-1	+1	-1
10	ad	+1	-1	-1	+1	-1
11	bd	-1	+1	-1	+1	-1
12	abd	+1	+1	-1	+1	-1
13	cd	-1	-1	+1	+1	-1
14	acd	+1	-1	+1	+1	-1
15	bcd	-1	+1	+1	+1	-1
16	abcd	+1	+1	+1	+1	-1
17	e	-1	-1	-1	-1	+1
18	ae	+1	-1	-1	-1	+1
19	be	-1	+1	-1	-1	+1
20	abe	+1	+1	-1	-1	+1
21	ce	-1	-1	+1	-1	+1
22	ace	+1	-1	+1	-1	+1
23	bce	-1	+1	+1	-1	+1
24	abce	+1	+1	+1	-1	+1
25	de	-1	-1	-1	+1	+1
26	ade	+1	-1	-1	+1	+1
27	bde	-1	+1	-1	+1	+1
28	abde	+1	+1	-1	+1	+1
29	cde	-1	-1	+1	+1	+1
30	acde	+1	-1	+1	+1	+1
31	bcde	-1	+1	+1	+1	+1
32	abcde	+1	+1	+1	+1	+1

Order to Run Experiments

- The four design patterns shown earlier are listed in the standard order.
- Standard order is used to design the combinations/treatments before experiments start.
- When actually running the experiments, randomizing the standard order is recommended to minimize the noise.

Replication in Experiments

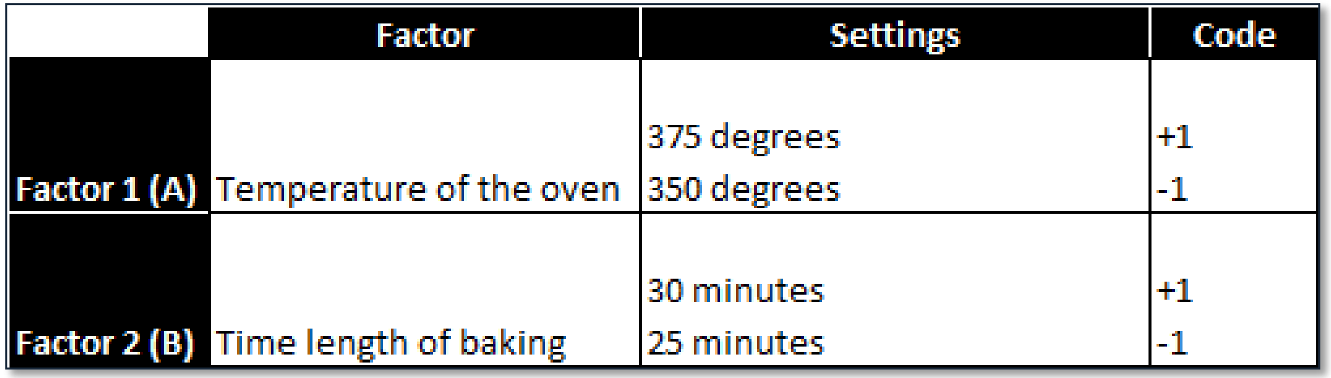
- Each treatment can be tested multiple times in an experiment in order to increase the degrees of freedom and improve the capability of analysis. We call this method replication.

- Replicates are the number of repetitions of running an individual treatment.

- The order to run the treatments in an experiment should be randomized to minimize the noise.

2^2 Full Factorial DOE

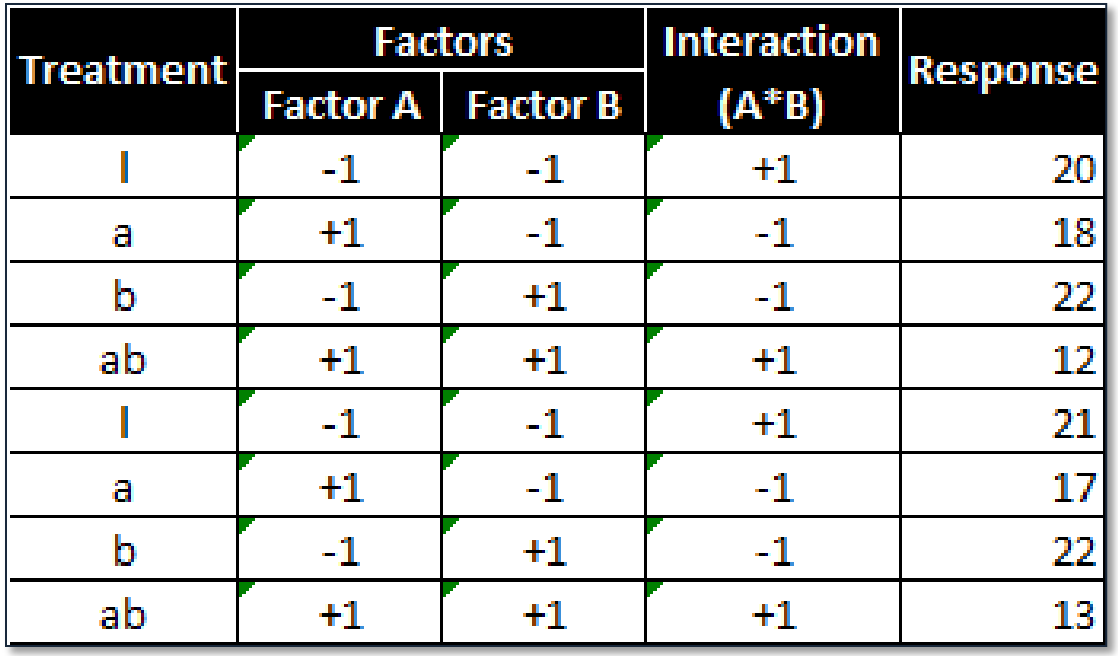
- Case study: we are running a 22 full factorial DOE to discover the cause-and-effect relationship between the cake tastiness and two factors: temperature of the oven and time length of baking.
  - Each factor has two levels and there are four treatments in total.
  - We decide to run each treatment twice so that we have enough degrees of freedom to measure the impact of two factors and the interaction between two factors. Therefore, there are eight observations in response eventually.



	Factor	Settings	Code
Factor 1 (A)	Temperature of the oven	375 degrees 350 degrees	+1 -1
Factor 2 (B)	Time length of baking	30 minutes 25 minutes	+1 -1

2^2 Full Factorial DOE

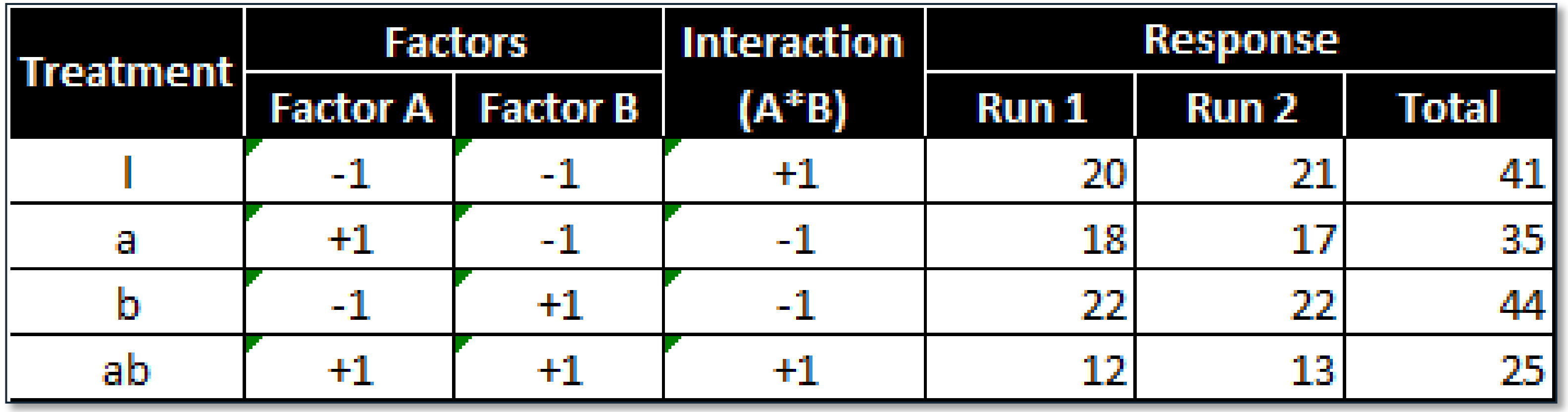
- After running the four treatments twice in a random order, we obtain the following results



Treatment	Factors		Interaction	Response
	Factor A	Factor B	(A*B)	
I	-1	-1	+1	20
a	+1	-1	-1	18
b	-1	+1	-1	22
ab	+1	+1	+1	12
I	-1	-1	+1	21
a	+1	-1	-1	17
b	-1	+1	-1	22
ab	+1	+1	+1	13

2^2 Full Factorial DOE

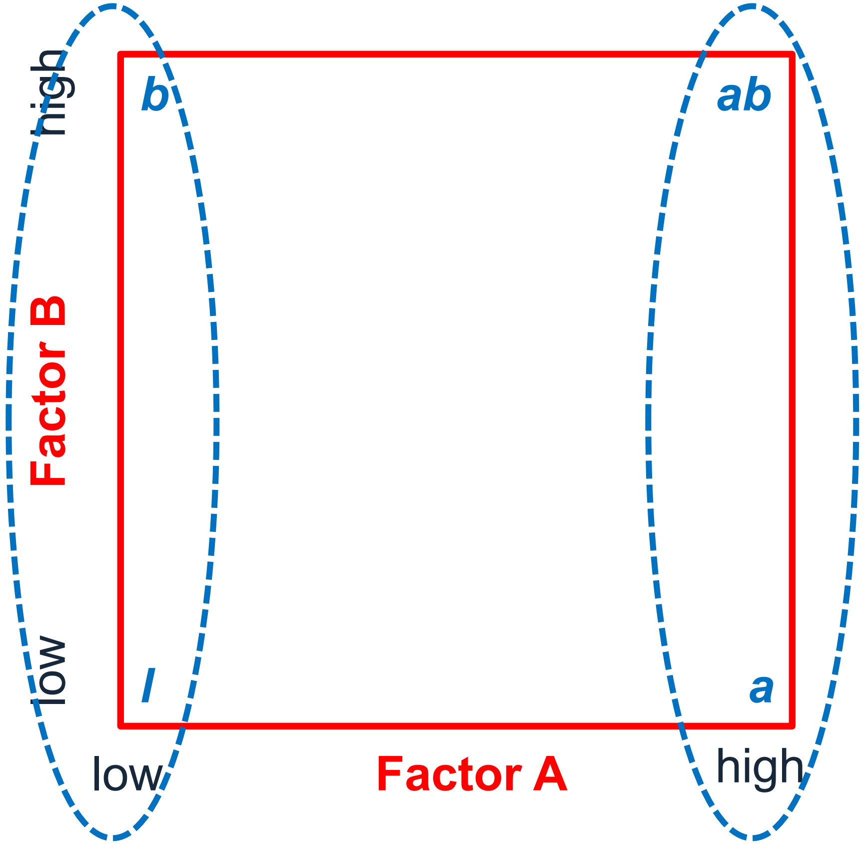
- The experiment results are consolidated into the following table



Treatment	Factors		Interaction	Response		
	Factor A	Factor B	(A*B)	Run 1	Run 2	Total
I	-1	-1	+1	20	21	41
a	+1	-1	-1	18	17	35
b	-1	+1	-1	22	22	44
ab	+1	+1	+1	12	13	25

2^2 Full Factorial DOE

- Main effect of factor A (temperature of the oven):



$$MainEffect_A = \frac{(a+ab)-(b+l)}{2^{k-1}r}$$

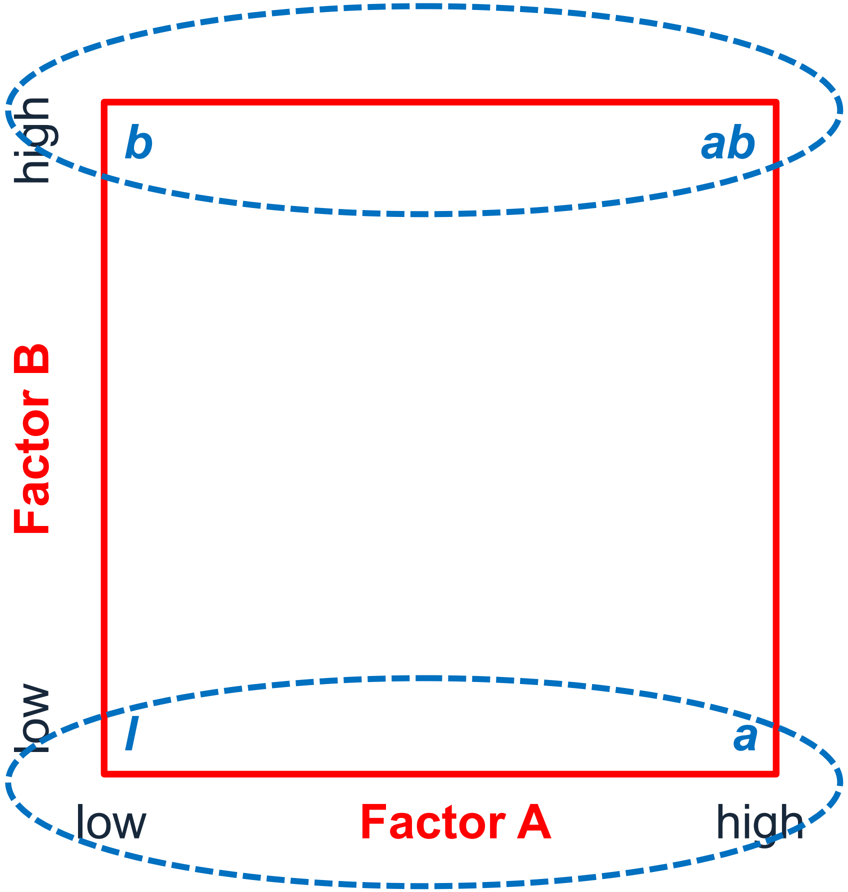
$$= \frac{(35+25)-(44+41)}{2^{2-1}\times 2}$$

$$= -6.25$$

where
k is the number of factors,
r is the number of times individual treatments are being run.

2^2 Full Factorial DOE

- Main effect of factor B (time length of baking):



$$MainEffect_B = \frac{(b+ab)-(a+l)}{2^{k-1}r}$$

$$= \frac{(44+25)-(35+41)}{2^{2-1}\times 2}$$

$$= -1.75$$

where
k is the number of factors,
r is the number of times individual treatments are being run.

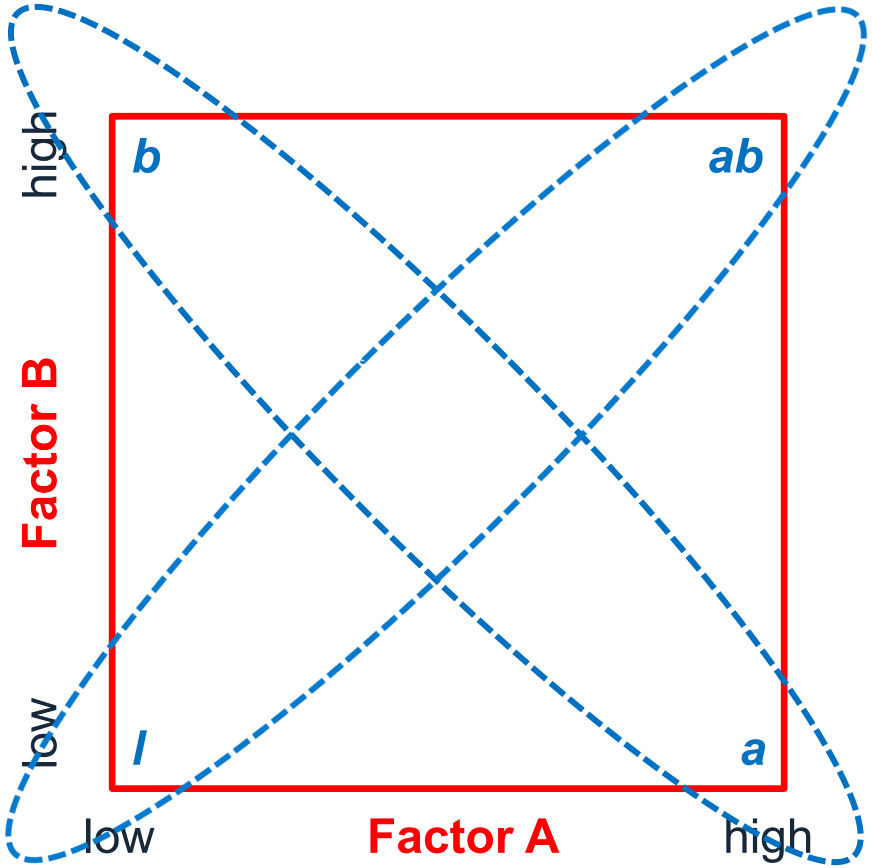
2^2 Full Factorial DOE

- Interaction (i.e., A*B) effect

$$InteractionEffect = \frac{(l+ab)-(a+b)}{2^{k-1}r}$$

$$= \frac{(41+25)-(35+44)}{2^{2-1}\times 2}$$

$$= -3.25$$



where
k is the number of factors,
r is the number of times individual treatments are being run.

2^2 Full Factorial DOE

- Sum of squares of factors and interaction

$$SS_A = \frac{(a+ab-b-l)^2}{2^k r} = \frac{(35+25-44-41)^2}{2^2 \times 2} = 78.125$$

$$SS_B = \frac{(b+ab-a-l)^2}{2^k r} = \frac{(44+25-35-41)^2}{2^2 \times 2} = 6.125$$

$$SS_{Interaction} = \frac{(l+ab-a-b)^2}{2^k r} = \frac{(41+25-35-44)^2}{2^2 \times 2} = 21.125$$

where
k is the number of factors,
r is the number of times individual treatments are being run.

2^2 Full Factorial DOE

- Degrees of freedom of factors and interaction

$$df_A = 1$$

$$df_B = 1$$

$$df_{Interaction} = 1$$

$$df_{error} = 1$$

- Four degrees of freedom are required in the model because there are three independent variables (a, b, ab interaction) in the equation and one degree of freedom is required for the error:
  - $Y = \alpha 0 + \alpha 1 * A + \alpha 2 * B + \alpha 3 * A * B + Error$

2^2 Full Factorial DOE

- Mean squares of factors and interaction

$$MS_A = \frac{SS_A}{df_A} = \frac{78.125}{1} = 78.125$$

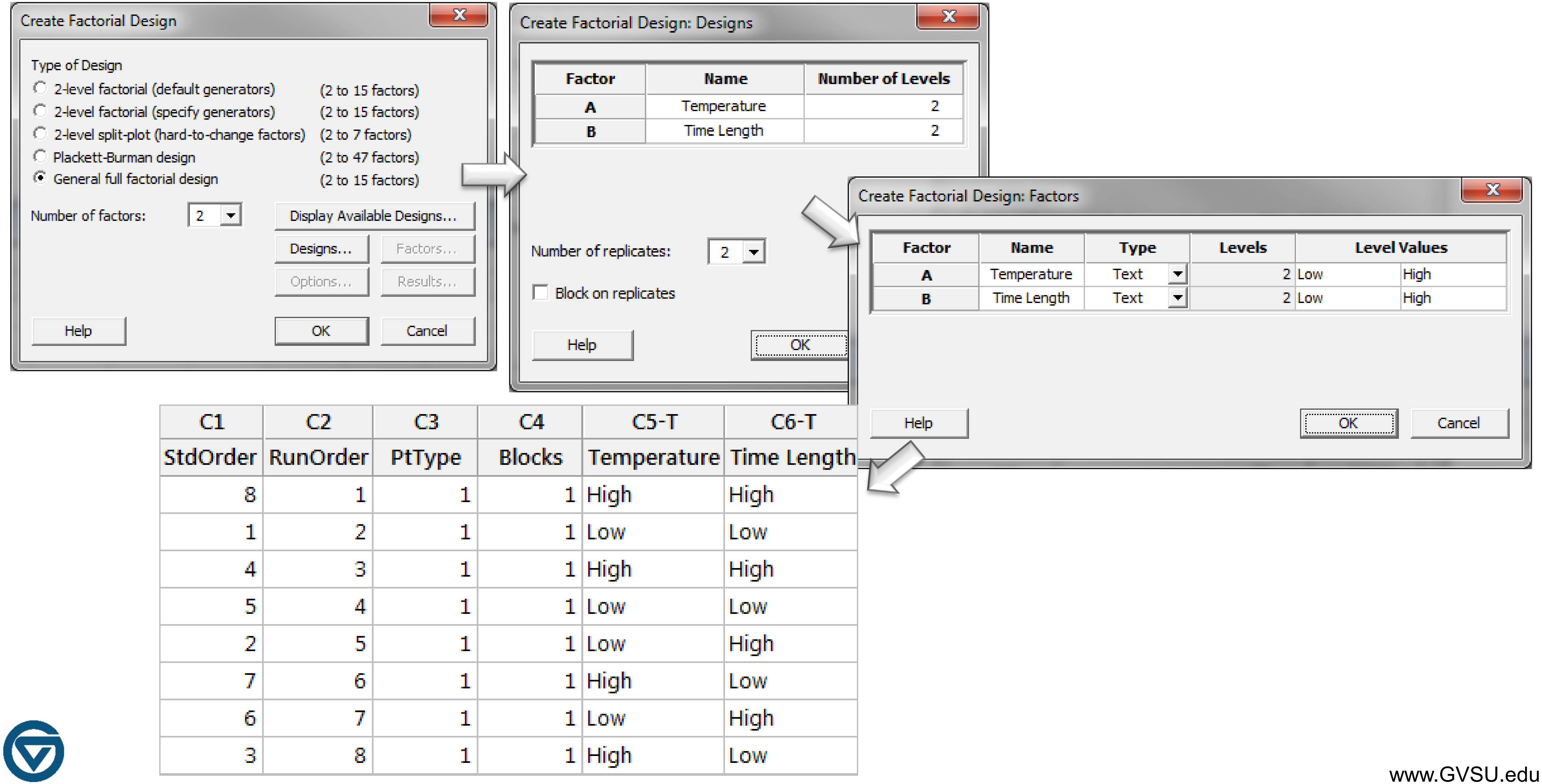
$$MS_B = \frac{SS_B}{df_B} = \frac{6.125}{1} = 6.125$$

$$MS_{Interaction} = \frac{SS_{Interaction}}{df_{Interaction}} = \frac{21.125}{1} = 21.125$$

Use Minitab to Run a 2k Full Factorial DOE

- Step 1: Initiate the experiment design
  - Click Stat → DOE → Factorial → Create Factorial Design to open the dialog box "Create Factorial Design".
  - Select the button of "General full factorial design."
  - Select "2" as the "Number of factors."
  - Click the "Designs" button and a new window named "Create Factorial Design – Designs" pops up.
  - Enter the factor name "Temperature" for factor A and 2 for number of levels.
  - Enter the factor name "Time Length" for factor B and 2 for number of levels.
  - Enter "2" as the "Number of replicates.
  - Click "OK" button in the window "Create Factorial Design – Designs."
  - Click on the "Factors" button in the window "Create Factorial Design" and a new window named "Create Factorial Design – Factors" pops up.
  - Select "Text" as the "Type" of both factor A and B.
  - Enter "Low" and "High" as the two levels for both factor A and B.
  - Click "OK" in the window "Create Factorial Design – Factors."
  - Click "OK" in the window "Create Factorial Design."
  - The design table is created in the data table.

Use Minitab to Run a 2^k Full Factorial DOE



Create Factorial Design

Type of Design
- ○ 2-level factorial (default generators) (2 to 15 factors)
- ○ 2-level factorial (specify generators) (2 to 15 factors)
- ○ 2-level split-plot (hard-to-change factors) (2 to 7 factors)
- ○ Plackett-Burman design (2 to 47 factors)
- ● General full factorial design (2 to 15 factors)

Number of factors: 2

Create Factorial Design: Designs

Factor	Name	Number of Levels
A	Temperature	2
B	Time Length	2

Number of replicates: 2

☐ Block on replicates

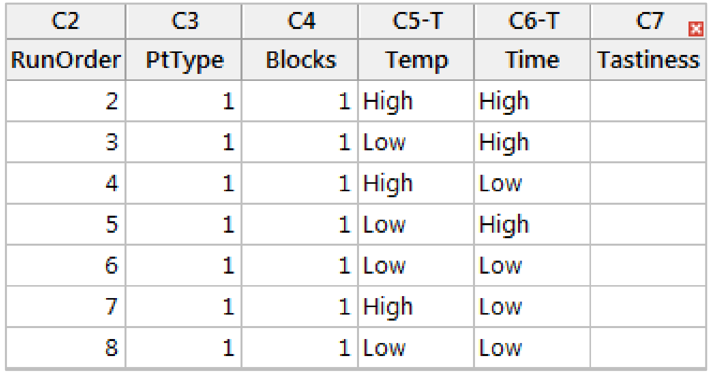
Create Factorial Design: Factors

Factor	Name	Type	Levels	Level Values	
A	Temperature	Text	2	Low	High
B	Time Length	Text	2	Low	High

C1	C2	C3	C4	C5-T	C6-T
StdOrder	RunOrder	PtType	Blocks	Temperature	Time Length
8	1	1	1	High	High
1	2	1	1	Low	Low
4	3	1	1	High	High
5	4	1	1	Low	Low
2	5	1	1	Low	High
7	6	1	1	High	Low
6	7	1	1	Low	High
3	8	1	1	High	Low

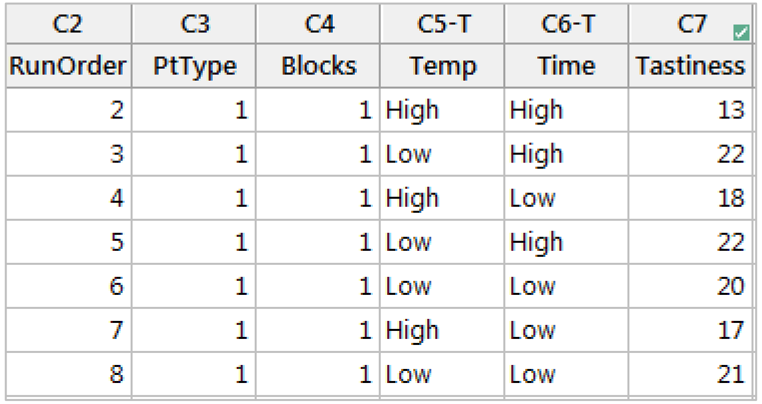
Use Minitab to Run a 2k Full Factorial DOE

- Step 2: Run the experiment in Run Order and record the response in a new column labeled “Tastiness” in the table created by Minitab.



C2	C3	C4	C5-T	C6-T	C7 ☒
RunOrder	PtType	Blocks	Temp	Time	Tastiness
2	1	1	High	High	
3	1	1	Low	High	
4	1	1	High	Low	
5	1	1	Low	High	
6	1	1	Low	Low	
7	1	1	High	Low	
8	1	1	Low	Low	

- We have run the experiment for you. Open data file “DOE Full.mtw”. The results of the experiment have been recorded in this file.

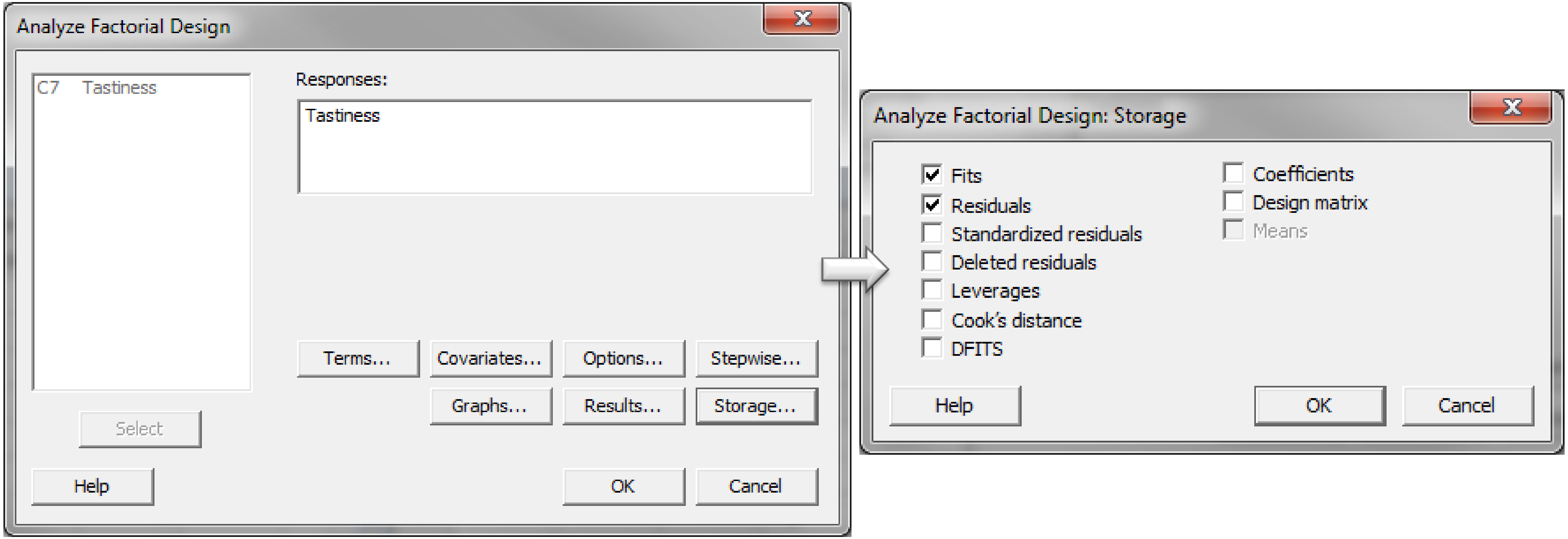


C2	C3	C4	C5-T	C6-T	C7 ✓
RunOrder	PtType	Blocks	Temp	Time	Tastiness
2	1	1	High	High	13
3	1	1	Low	High	22
4	1	1	High	Low	18
5	1	1	Low	High	22
6	1	1	Low	Low	20
7	1	1	High	Low	17
8	1	1	Low	Low	21

Use Minitab to Run a 2k Full Factorial DOE

- Step 4: Analyze the experiment results
  - Click Stat → DOE → Factorial → Analyze Factorial Design.
  - A new window named “Analyze Factorial Design” appears.
  - Select “Tastiness” in the list box of “Responses.”
  - Click on the “Storage” button and a new window named “Analyze Factorial Design – Storage” pops up.
  - Check the boxes “Fits” and “Residuals” in the window “Analyze Factorial Design – Storage.”
  - Click “OK” in the window “Analyze Factorial Design – Storage.”
  - Click “OK” in the window “Analyze Factorial Design.”
  - The DOE analysis results appear in the session window.

Use Minitab to Run a 2^k Full Factorial DOE



Analyze Factorial Design

C7 Tastiness

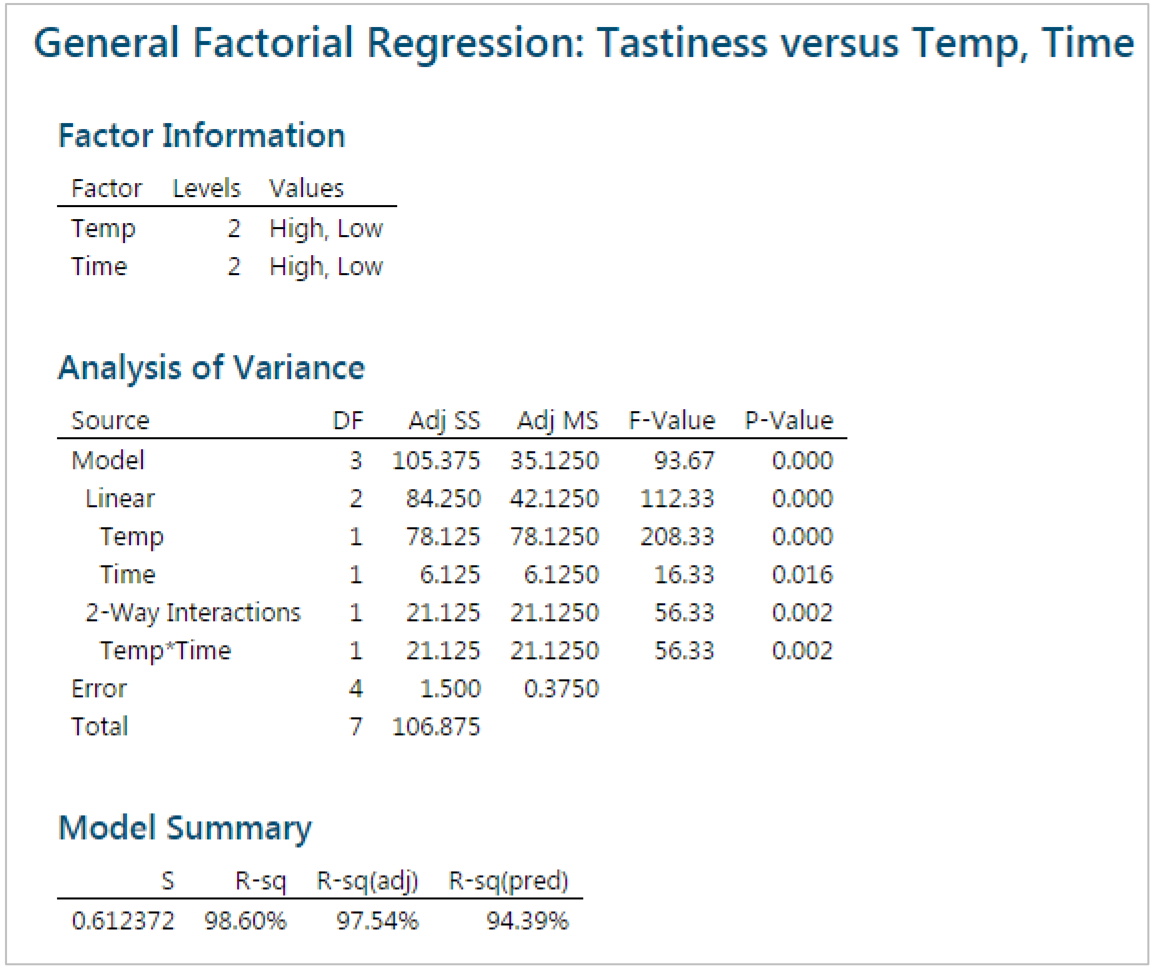
Responses: Tastiness

Analyze Factorial Design: Storage

☑ Fits
☑ Residuals
☐ Standardized residuals
☐ Deleted residuals
☐ Leverages
☐ Cook's distance
☐ DFITS
☐ Coefficients
☐ Design matrix
☐ Means

Use Minitab to Run a 2k Full Factorial DOE

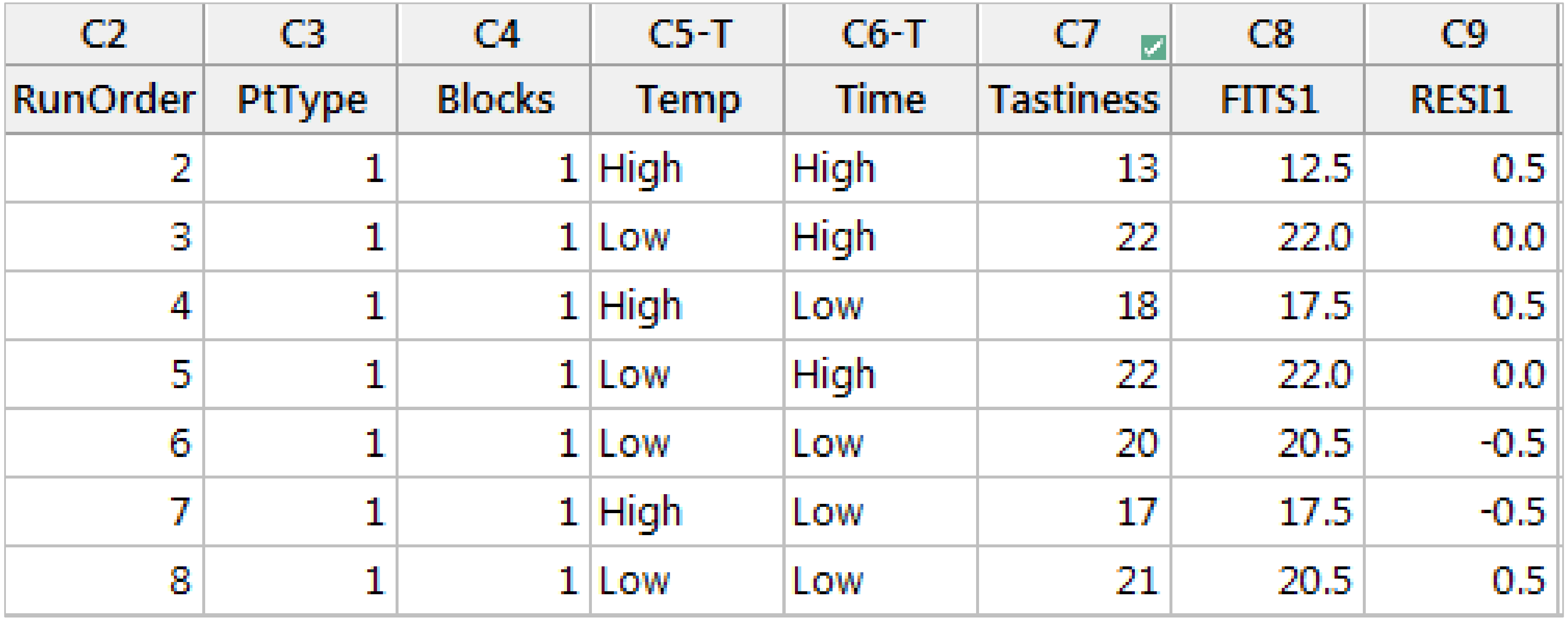
- Since the p values of all the independent variables in the model are smaller than the alpha level (0.05), both factors and their interactions have statistically significant impact on the response.
- The p-value for the overall model indicate that the model is significant and at least one of the factors is statistically significant.
- A high R2 value shows around 98% of the variation in the response can be explained by the model.



```
General Factorial Regression: Tastiness versus Temp, Time

Factor Information

Factor  Levels  Values
Temp         2  High, Low
Time         2  High, Low

Analysis of Variance

Source              DF   Adj SS   Adj MS  F-Value  P-Value
Model                3  105.375  35.1250    93.67    0.000
  Linear             2   84.250  42.1250   112.33    0.000
    Temp             1   78.125  78.1250   208.33    0.000
    Time             1    6.125   6.1250    16.33    0.016
  2-Way Interactions 1   21.125  21.1250    56.33    0.002
    Temp*Time        1   21.125  21.1250    56.33    0.002
Error                4    1.500   0.3750
Total                7  106.875

Model Summary

       S    R-sq  R-sq(adj)  R-sq(pred)
0.612372  98.60%     97.54%      94.39%
```

Use Minitab to Run a 2^k Full Factorial DOE

- The fitted response and the residuals of the DOE model are stored in the last two columns of the data table.



C2	C3	C4	C5-T	C6-T	C7 ✓	C8	C9
RunOrder	PtType	Blocks	Temp	Time	Tastiness	FITS1	RESI1
2	1	1	High	High	13	12.5	0.5
3	1	1	Low	High	22	22.0	0.0
4	1	1	High	Low	18	17.5	0.5
5	1	1	Low	High	22	22.0	0.0
6	1	1	Low	Low	20	20.5	-0.5
7	1	1	High	Low	17	17.5	-0.5
8	1	1	Low	Low	21	20.5	0.5

4.4.2 Linear & Quadratic Models

Linear DOE

- If we assume the relationship between the response and the factors is linear, we use a two-level design experiment in which there are two settings for individual factors.
- The linear model of a two-level design with three factors:

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_{12} X_1 X_2 + \alpha_{13} X_1 X_3$$
$$+ \alpha_{23} X_2 X_3 + \alpha_{123} X_1 X_2 X_3 + \varepsilon$$

where
Y is the response.
X_1, X_2, X_3 are the three two factors.
α_0 is the intercept.
α_1, α_2, …, α_{123} are the coefficients of the factors and interactions.
ε is the error of the model.

Quadratic DOE

- It is possible that the relationship between the response and the factors is non-linear.

- To test the non-linear relationship between the response and the factors, we need at least a three-level design, i.e., quadratic design of experiment.

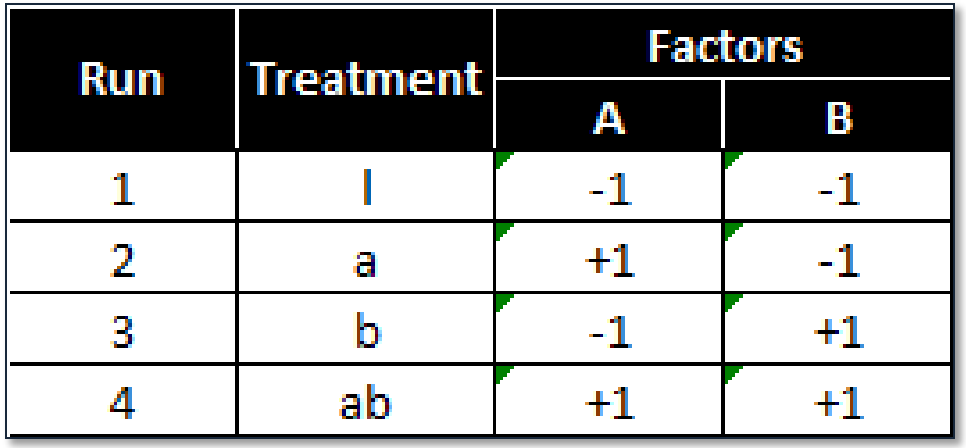
- The quadratic DOE adds the X_1^2 , X_2^2 ,…, X_k^2 and corresponding interactions into the model as potential significant factor of the response.

4.4.3 Balanced & Orthogonal Designs

Balanced Design

- In DOE, an experiment design is balanced if individual treatments have the same number of observations.
- In other words, if an experiment is balanced, each level of each factor is run the same number of times.

- The following design is balanced since the two factors are both run twice at each level.



Run	Treatment	Factors	
		A	B
1	I	-1	-1
2	a	+1	-1
3	b	-1	+1
4	ab	+1	+1

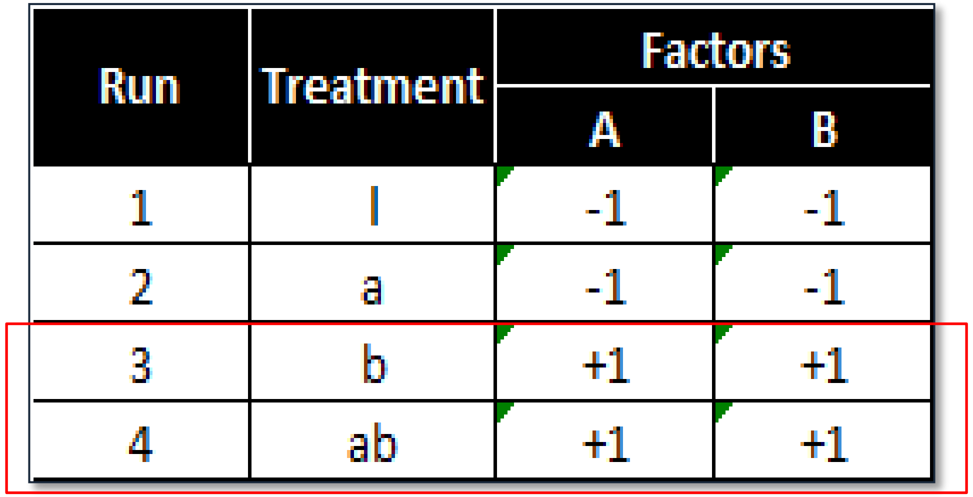
Orthogonal Design

- In DOE, an experiment design is orthogonal if the main effect of one factor can be evaluated independently of other factors.

- In other words, for each level (high setting or low setting) of a factor, the number of high setting and low setting in any other factors must be the same.

Orthogonal Design

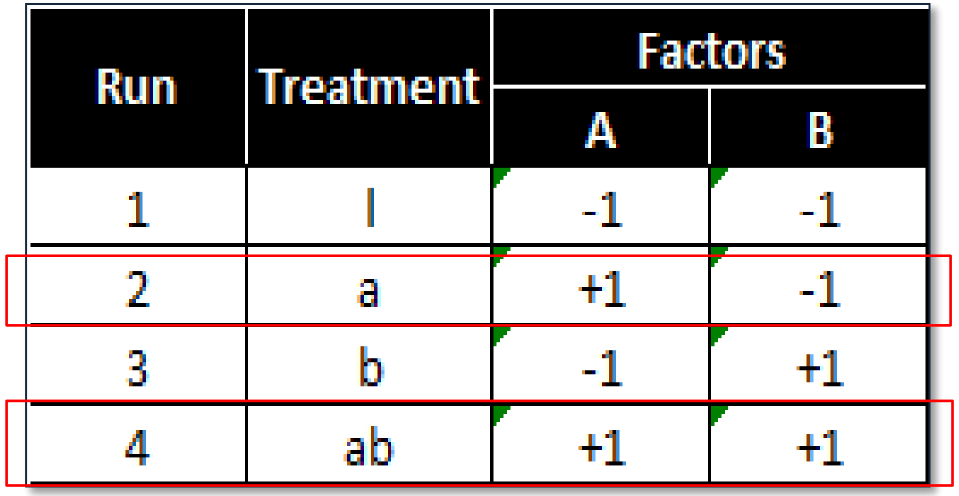
- In the following design, for the higher level (+1) of factor A, the number of “+1” in factor B is 2 but the number of “-1” in factor B is 0. Therefore, this is not an orthogonal design.



Run	Treatment	Factors	
		A	B
1	l	-1	-1
2	a	-1	-1
3	b	+1	+1
4	ab	+1	+1

Orthogonal Design

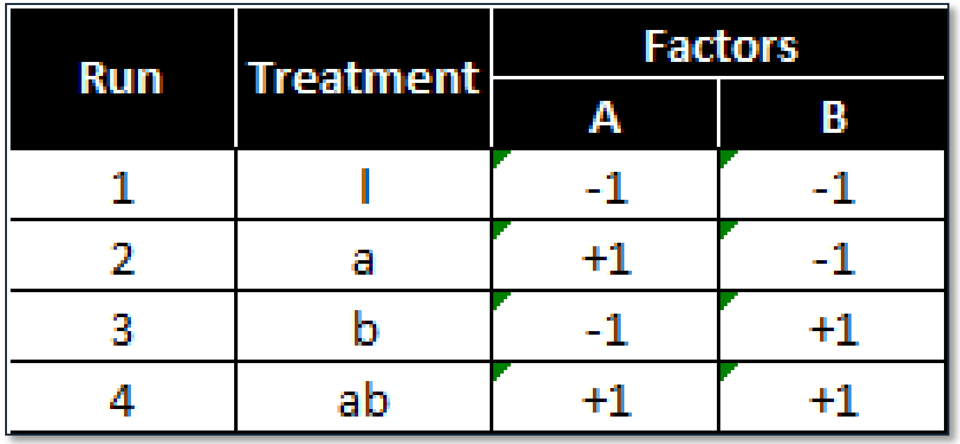
- In the following design, for the higher level (+1) of factor A, the number of “+1” in factor B is 1 and the number of “-1” in factor B is 1 as well. Therefore, this is an orthogonal design.



Run	Treatment	Factors	
		A	B
1	I	-1	-1
2	a	+1	-1
3	b	-1	+1
4	ab	+1	+1

Checking for an Orthogonal Design

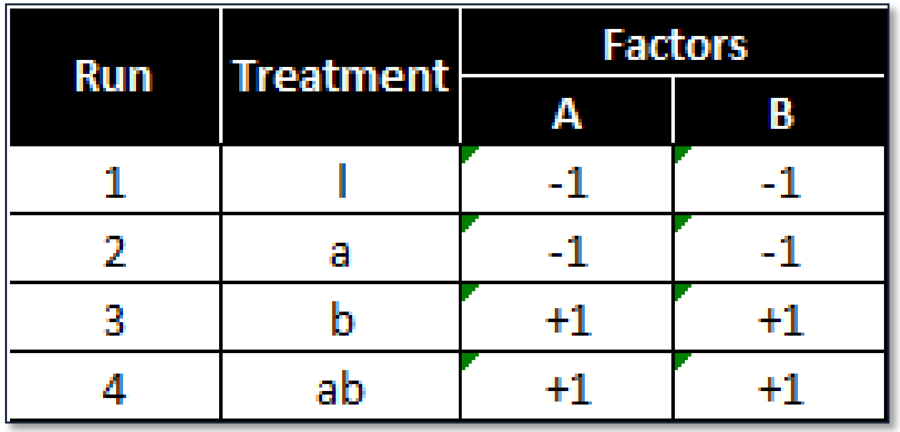
- We can check to see if our design is orthogonal with some very simple math.
- In an orthogonal design, the sum of products for each run should equal zero.
- Let us check the earlier design that we said was orthogonal:
  - Run 1: -1 * -1 = 1
  - Run 2: +1 * -1 = -1
  - Run 3: -1 * +1 = -1
  - Run 4: +1 * +1 = 1
  - Sum = 0
  - Design is orthogonal



Run	Treatment	Factors	
		A	B
1	l	-1	-1
2	a	+1	-1
3	b	-1	+1
4	ab	+1	+1

Checking for an Orthogonal Design

- Now let us check the earlier design that we said was not orthogonal:
  - Run 1: -1 * -1 = 1
  - Run 2: -1 * -1 = 1
  - Run 3: +1 * +1 = 1
  - Run 4: +1 * +1 = 1
  - Sum = 4
  - Design is not orthogonal



Run	Treatment	Factors	
		A	B
1	I	-1	-1
2	a	-1	-1
3	b	+1	+1
4	ab	+1	+1

4.4.4 Fit, Diagnosis and Center Points

Use Minitab to Fit the Model of DOE

- If one of the factors or interactions has a p-value larger than alpha level (0.05), it indicates that the particular factor or interaction does *not* have statistically significant impact on the response.

- To fit the model better, we need to remove the insignificant independent variables one at a time and run the model again until all the independent variables in the model are statistically significant (i.e. p-value smaller than alpha level).

Use Minitab to Diagnose the Model

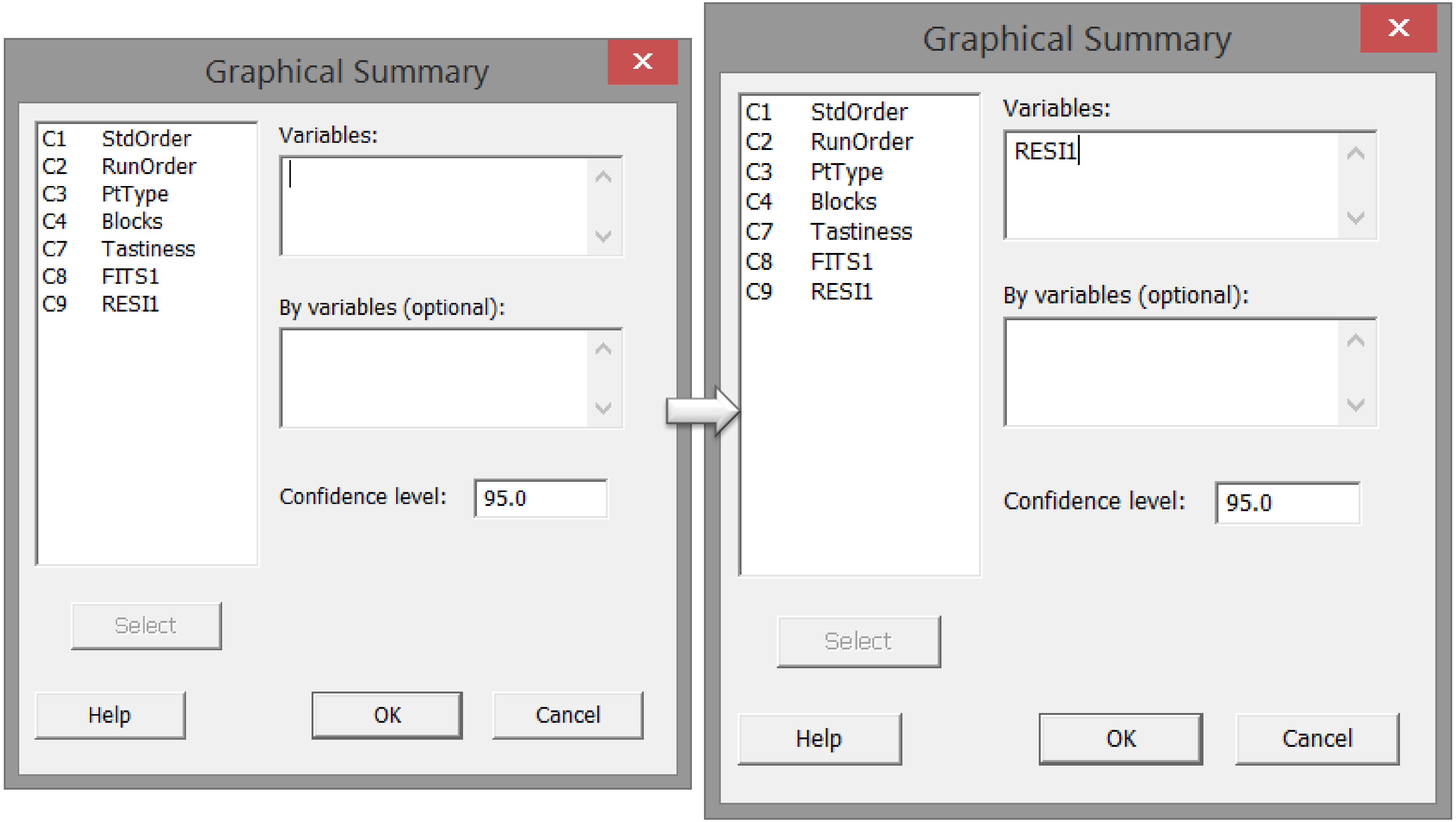
- To ensure that the quality of the model is acceptable, we need to conduct the residuals analysis.
  - Residual is the difference between the actual response value and the fitted response value.

- Well-performing residuals:
  - Are normally distributed
  - Have a mean equal to zero
  - Are independent
  - Have equal variance across the fitted values.

Use Minitab to Diagnose the Model

- Step 1: Check whether the mean of residuals is zero and residuals are normally distributed.
  1) Click Stat → Basic Statistics → Graphical Summary.
  2) A new window named "Graphical Summary" appears.
  3) Select "RESI1" as the "Variables."
  4) Click "OK."
  5) The histogram and normality test results appear in the new window.

Use Minitab to Diagnose the Model



Graphical Summary

C1 StdOrder
C2 RunOrder
C3 PtType
C4 Blocks
C7 Tastiness
C8 FITS1
C9 RESI1

Variables:

By variables (optional):

Confidence level: 95.0

Graphical Summary

C1 StdOrder
C2 RunOrder
C3 PtType
C4 Blocks
C7 Tastiness
C8 FITS1
C9 RESI1

Variables: RESI1

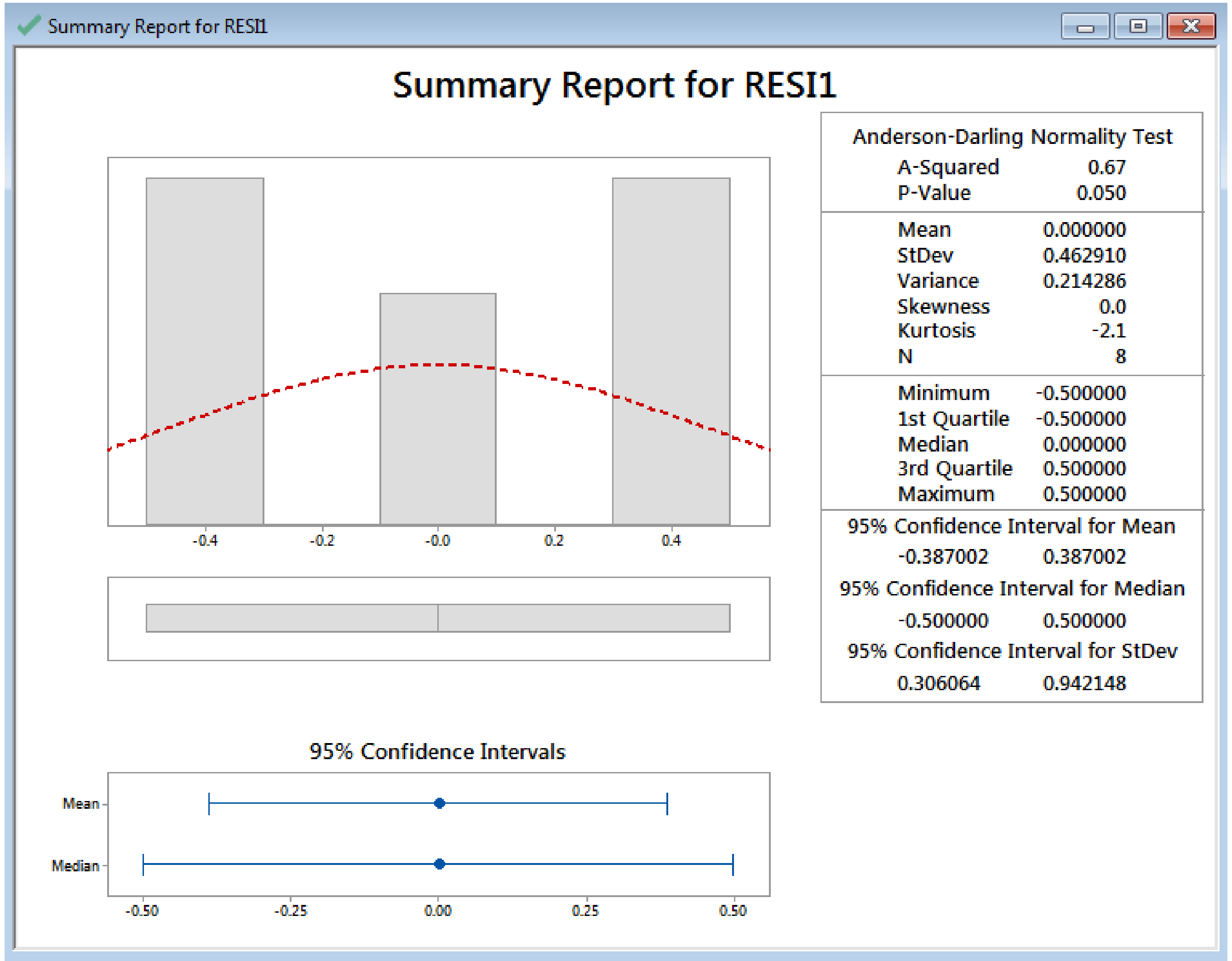
By variables (optional):

Confidence level: 95.0

Use Minitab to Diagnose the Model

- Since the p-value of the normality test is not less the alpha level, we fail to reject the null and conclude that the residuals are normally distributed

- The mean of the residuals is approximately zero.



Summary Report for RESI1

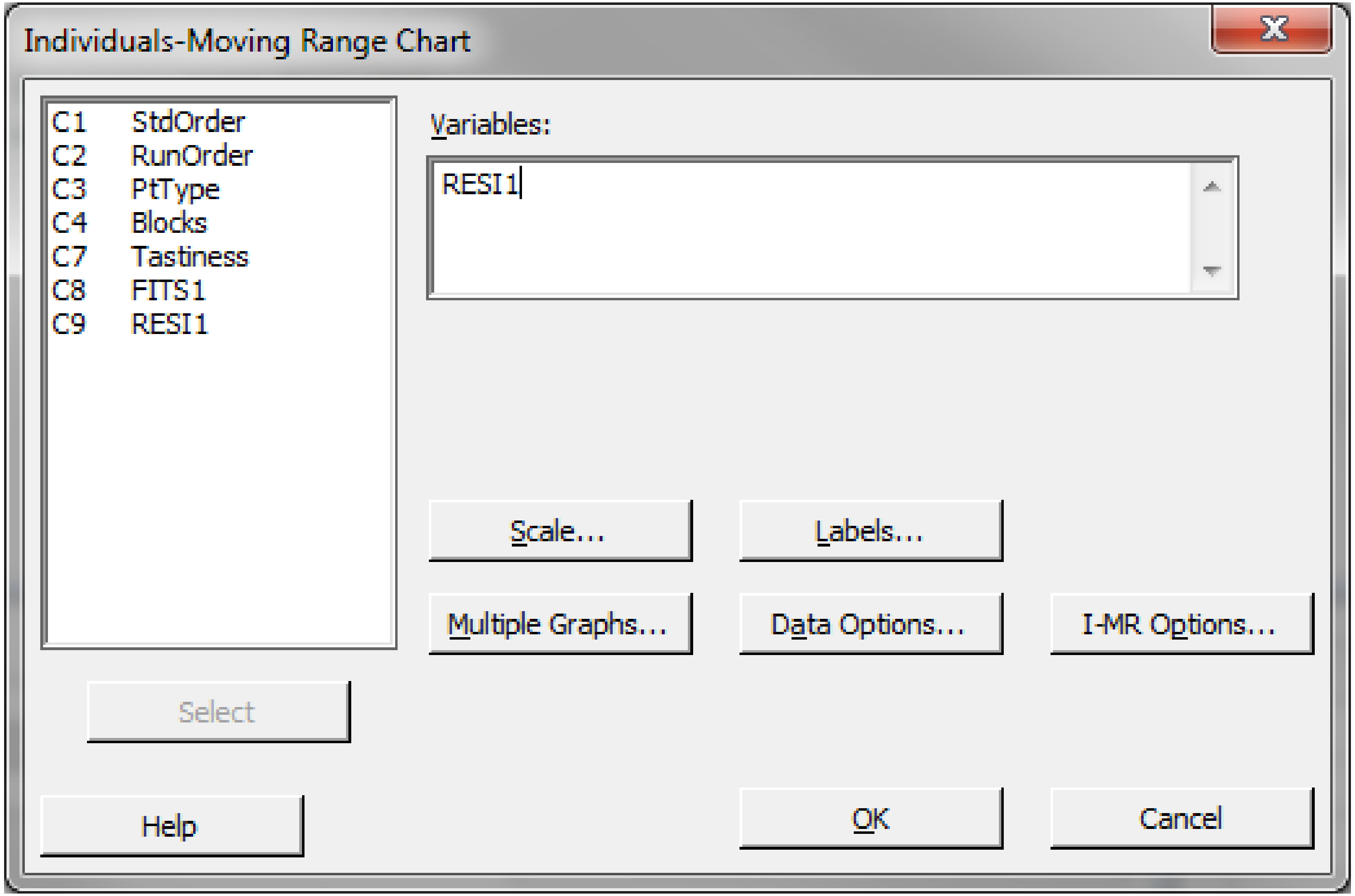
```
Anderson-Darling Normality Test
  A-Squared          0.67
  P-Value           0.050

  Mean          0.000000
  StDev         0.462910
  Variance      0.214286
  Skewness           0.0
  Kurtosis          -2.1
  N                    8

  Minimum      -0.500000
  1st Quartile -0.500000
  Median        0.000000
  3rd Quartile  0.500000
  Maximum       0.500000
95% Confidence Interval for Mean
  -0.387002     0.387002
95% Confidence Interval for Median
  -0.500000     0.500000
95% Confidence Interval for StDev
  0.306064      0.942148
```

Use Minitab to Diagnose the Model

- Step 2: Check whether the residuals are independent.
  1) Click Stat → Control Charts → Variable Charts for Individuals → I-MR.
  2) A new window named "Individuals-Moving Range Chart" appears.
  3) Select "RESI1" as the "Variables."
  4) Click "OK."
  5) The I-MR charts appear in the new window.
  6) If no data points on the control charts fail any tests, the residuals are in control and independent of each other.

Note: The prerequisite of plotting IMR chart for residuals is that the residuals are in time order.

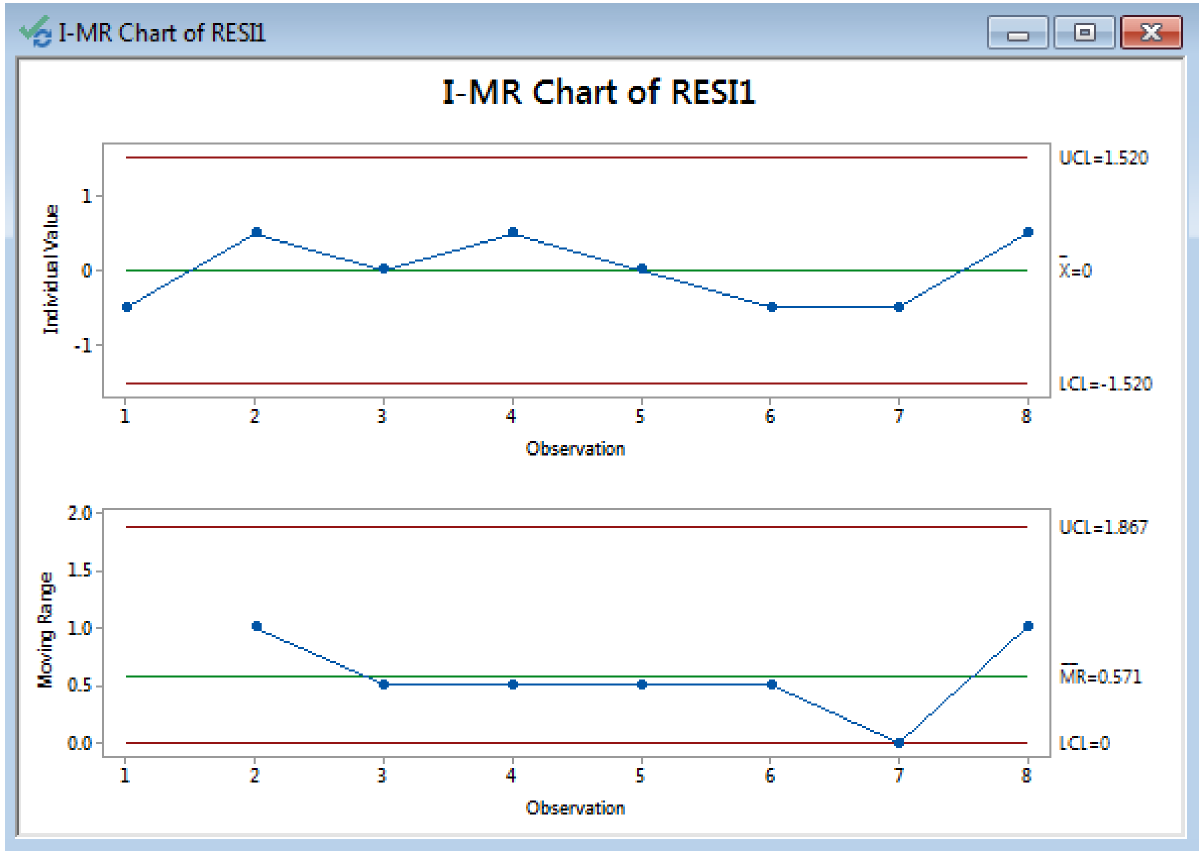
Use Minitab to Diagnose the Model



Variables: RESI1

Use Minitab to Diagnose the Model

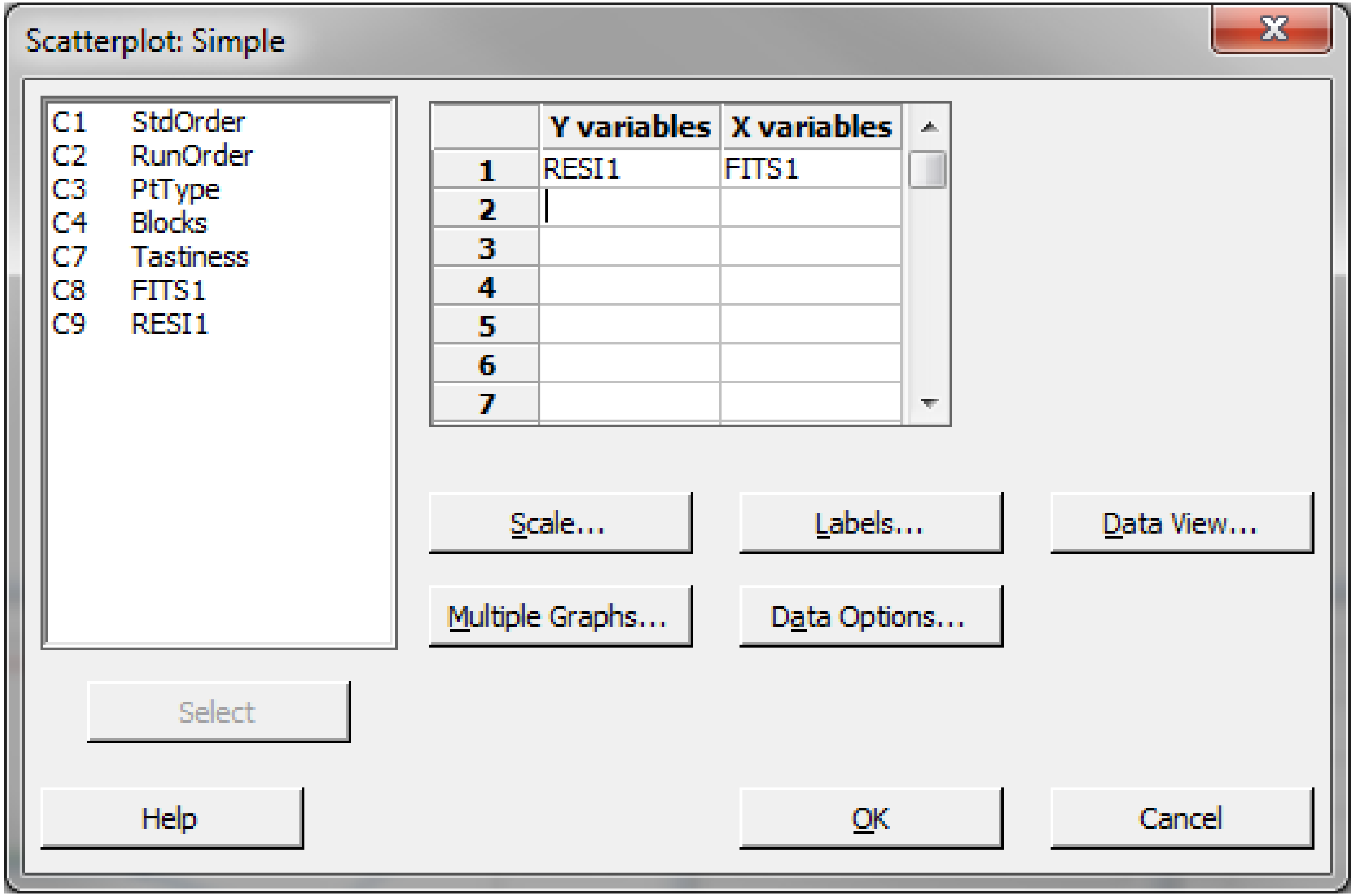
- If no data points on the control charts fail any tests, the residuals are in control and independent of each other.



Use Minitab to Diagnose the Model

- Step 3: Check whether the residuals have equal variance across the predicted response values.
  1) Click Graph → Scatterplot.
  2) A new window named “Scatterplots” appears.
  3) Click “OK” in the window “Scatterplots” and a new window named “Scatterplot – Simple” appears.
  4) Select “RESI1” as the “Y variables” and “FITS1” as the “X variables.”
  5) Click “OK.”
  6) The scatter plot between the residuals and the fitted response appears in the new window.

Use Minitab to Diagnose the Model



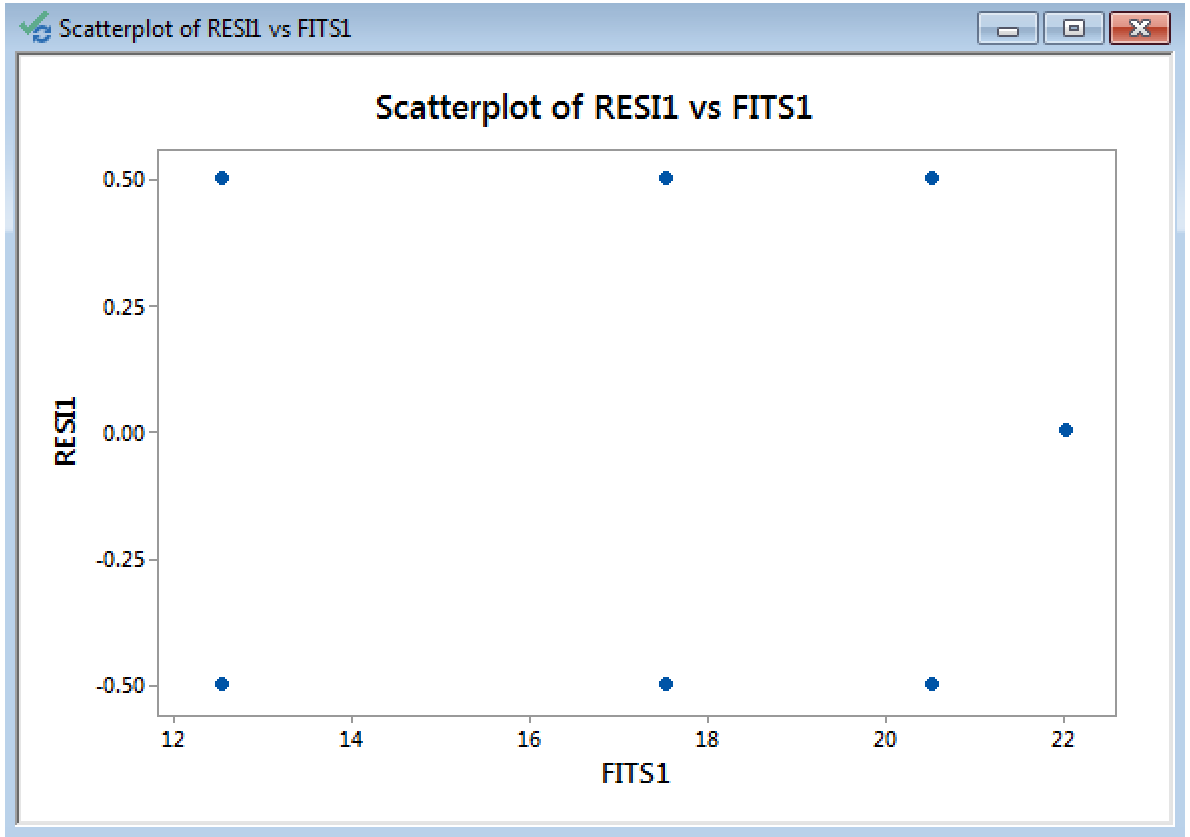
Scatterplot: Simple

C1 StdOrder
C2 RunOrder
C3 PtType
C4 Blocks
C7 Tastiness
C8 FITS1
C9 RESI1

	Y variables	X variables
1	RESI1	FITS1
2		
3		
4		
5		
6		
7		

Use Minitab to Diagnose the Model

- We look for patterns in which the residuals tend to have even variation across the entire range of the fitted response values.



Center Points in DOE

- The two-level design of experiment assumes that the relationship between the response and the factors is linear.
- We can use center points to check whether the assumption is true by adding center point runs to the experiment.

- In each center point run, the factors are all set to be in the center setting (zero) between high (+1) and low (-1) settings.
  - Center points do not change the model to quadratic.
  - They allow a check for adequacy of linear model.
  - If a line connecting factor settings passes through the center of the design, the model is adequate to predict within the inference space.

4.5 Fractional Factorial Experiments

Black Belt Training: Improve Phase

4.1 Simple Linear Regression
- 4.1.1 Correlation
- 4.1.2 X-Y Diagram
- 4.1.3 Regression Equations
- 4.1.4 Residuals Analysis

4.2 Multiple Regression Analysis
- 4.2.1 Non-Linear Regression
- 4.2.2 Multiple Linear Regression
- 4.2.3 Confidence Intervals
- 4.2.4 Residuals Analysis
- 4.2.5 Data Transformation, Box Cox
- 4.2.6 Stepwise Regression
- 4.2.7 Logistic Regression

4.3 Designed Experiments
- 4.3.1 Experiment Objectives
- 4.3.2 Experimental Methods
- 4.3.3 DOE Design Considerations

4.4 Full Factorial Experiments
- 4.4.1 2k Full Factorial Designs
- 4.4.2 Linear and Quadratic Models
- 4.4.3 Balanced and Orthogonal Designs
- 4.4.4 Fit, Model, and Center Points

4.5 Fractional Factorial Experiments
- 4.5.1 Designs
- 4.5.2 Confounding Effects
- 4.5.3 Experimental Resolution

4.5.1 Fractional Designs

What Are Fractional Factorial Experiments?

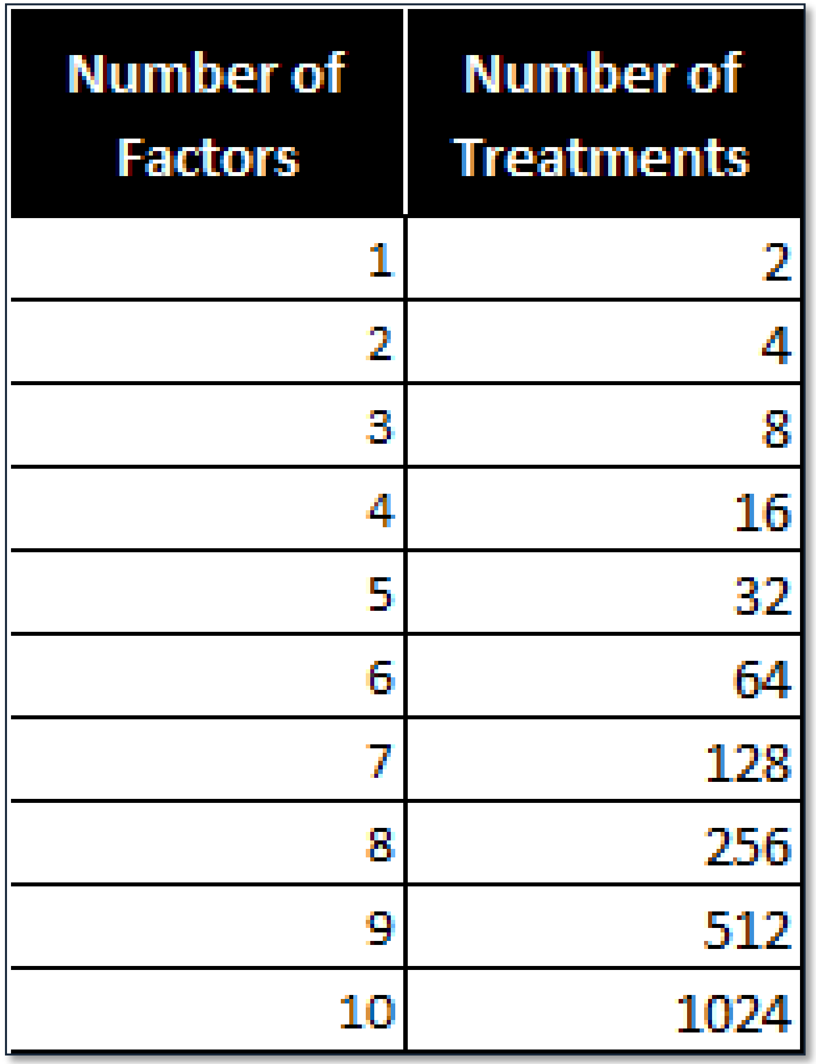
- **In** simple terms, a **fractional factorial experiment** is a subset of a full factorial experiment.
- Fractional factorials use fewer treatment combinations and runs.

- Fractional factorials are less able to determine effects because of fewer degrees of freedom available to evaluate higher order interactions.
- Fractional factorials can be used to screen a larger number of factors.
- Fractional factorials can also be used for optimization.

Why Fractional Factorial Experiments?

- To run a full factorial experiment for k factors, we need 2^k unique treatments. In other words, we need resources that can afford at least 2^k runs.
- With k increasing, the number of runs required in full factorial experiments rises dramatically even without any replications, and the percentage of degrees of freedom spent on the main effects decreases.
- The main effects and two-way interaction are the key effects we need to evaluate. The higher order the interaction is, the more we can ignore it.

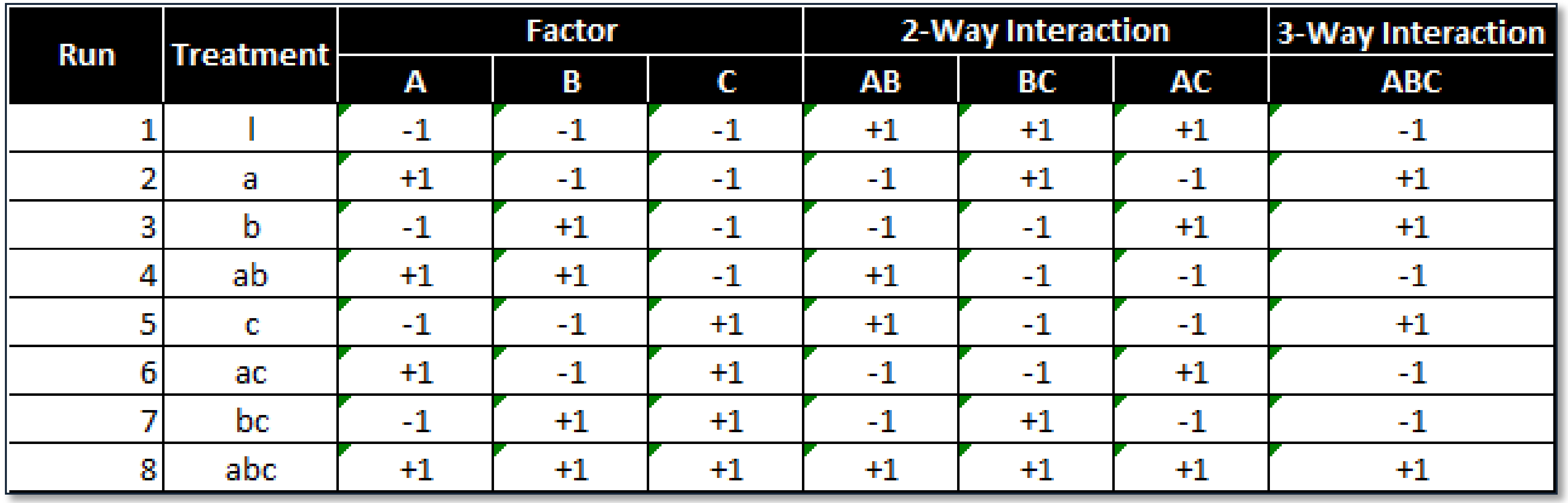
Why Fractional Factorial Experiments?



Number of Factors	Number of Treatments
1	2
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024

How Does a Fractional Factorial Work?

- We are trying to find the cause-and-effect relationship between a response (Y) and three factors (factor A, B, and C) and their interactions (AB, BC, AC, and ABC).
  - As follows is the 2^3 full factorial design (2 level 3 factor).
  - There are eight treatment combinations (2 * 2 * 2).



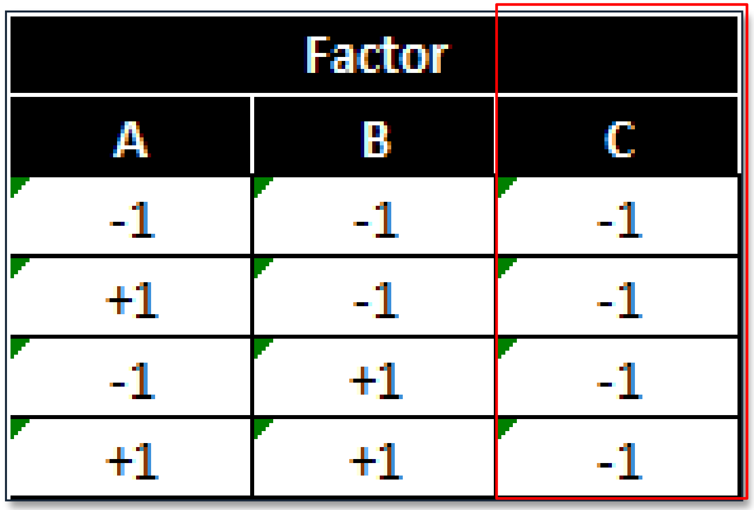
Run	Treatment	Factor			2-Way Interaction			3-Way Interaction
		A	B	C	AB	BC	AC	ABC
1	I	-1	-1	-1	+1	+1	+1	-1
2	a	+1	-1	-1	-1	+1	-1	+1
3	b	-1	+1	-1	-1	-1	+1	+1
4	ab	+1	+1	-1	+1	-1	-1	-1
5	c	-1	-1	+1	+1	-1	-1	+1
6	ac	+1	-1	+1	-1	-1	+1	-1
7	bc	-1	+1	+1	-1	+1	-1	-1
8	abc	+1	+1	+1	+1	+1	+1	+1

How Does a Fractional Factorial Work?

- To perform a 2^3 full factorial experiment, we need to run at least eight unique treatments (2 * 2 * 2).
- What if we only have enough resources to run four treatments?
- As a result, we need to carefully select a subset from the eight treatments so that all of our main effects can be evaluated and the design can be kept balanced and orthogonal.

How Does a Fractional Factorial Work?

- Example of an invalid design



Factor		
A	B	C
-1	-1	-1
+1	-1	-1
-1	+1	-1
+1	+1	-1

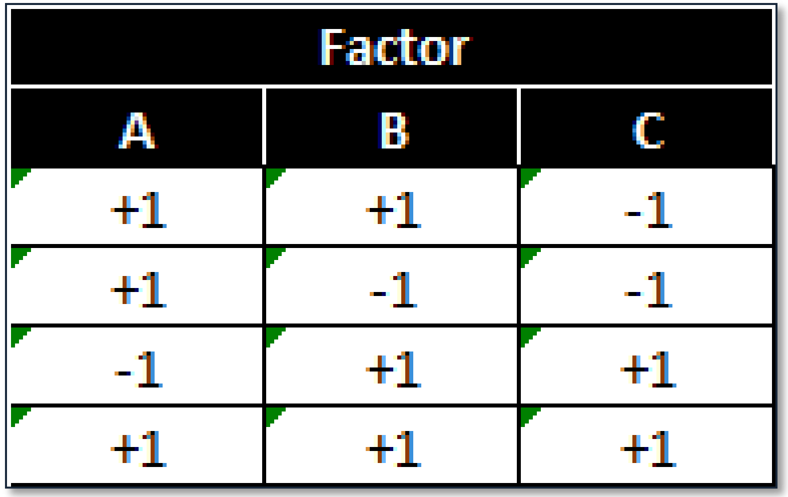
This design is invalid because only the low setting of factor C is tested.

We cannot evaluate the main effect of factor C using this design.

- Remember orthogonality?

How Does a Fractional Factorial Work?

- Example of an invalid design



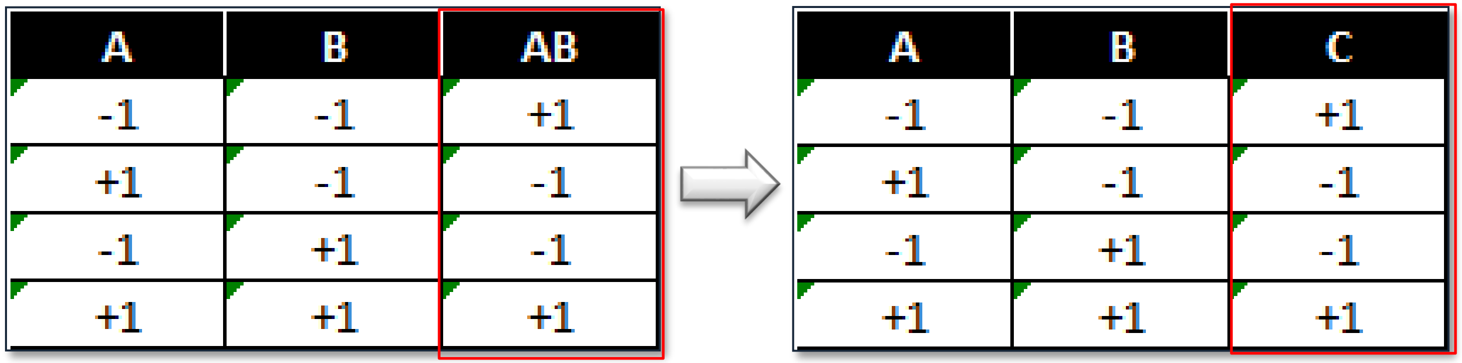
Factor		
A	**B**	**C**
+1	+1	-1
+1	-1	-1
-1	+1	+1
+1	+1	+1

This design is also invalid because it is neither balanced nor orthogonal.

- Checking orthogonality: the sum of AC interaction signs should equal zero (0).
  - Run 1 (-)
  - Run 2 (-)
  - Run 3 (-)
  - Run 4 (+)
  - **Sum (-1)**

How Does Fractional Factorial Work?

- To select the four treatments run in the 2^{3-1} fractional factorial experiment, we start from the 2^2 full factorial design of experiment.
- If we replace the two-way interaction (AB) column with the factor C column, the design will be valid.

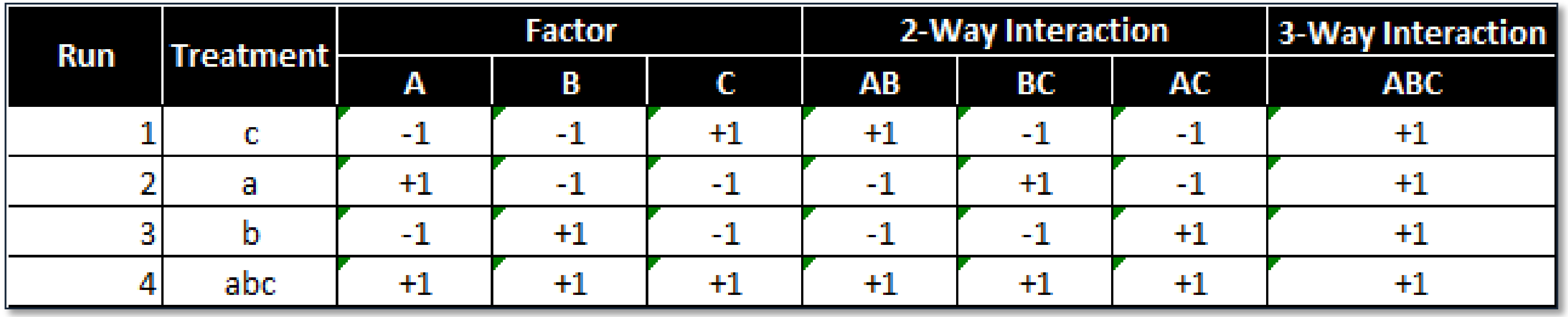


A	B	AB
-1	-1	+1
+1	-1	-1
-1	+1	-1
+1	+1	+1

A	B	C
-1	-1	+1
+1	-1	-1
-1	+1	-1
+1	+1	+1

How Does a Fractional Factorial Work?

- 2^{3-1} Fractional Factorial Design Pattern
  - Three factors and four treatments



Run	Treatment	Factor			2-Way Interaction			3-Way Interaction
		A	B	C	AB	BC	AC	ABC
1	c	-1	-1	+1	+1	-1	-1	+1
2	a	+1	-1	-1	-1	+1	-1	+1
3	b	-1	+1	-1	-1	-1	+1	+1
4	abc	+1	+1	+1	+1	+1	+1	+1

- Note: We also call this kind of design a half-factorial design since we only have half of the treatments that we would have in a full factorial design.
- In 2^{3-1} fractional factorial design of experiment, the effect of three-way interaction (ABC) is not measurable since it only has “+1”.

How Does a Fractional Factorial Work?

- In the 2^{3-1} fractional factorial design, we notice that the column of each main effect has identical “+1” and “-1” values with one two-way interaction column.
  - A and BC
  - B and AC
  - C and AB
- In this situation, we say that A is aliased with BC or A is the alias of BC.

How Does a Fractional Factorial Work?

- By multiplying any column with itself, we obtain the identity (I).
  - A * A = I
  - The product of any column and the identity is the column itself.
  - A * I = A

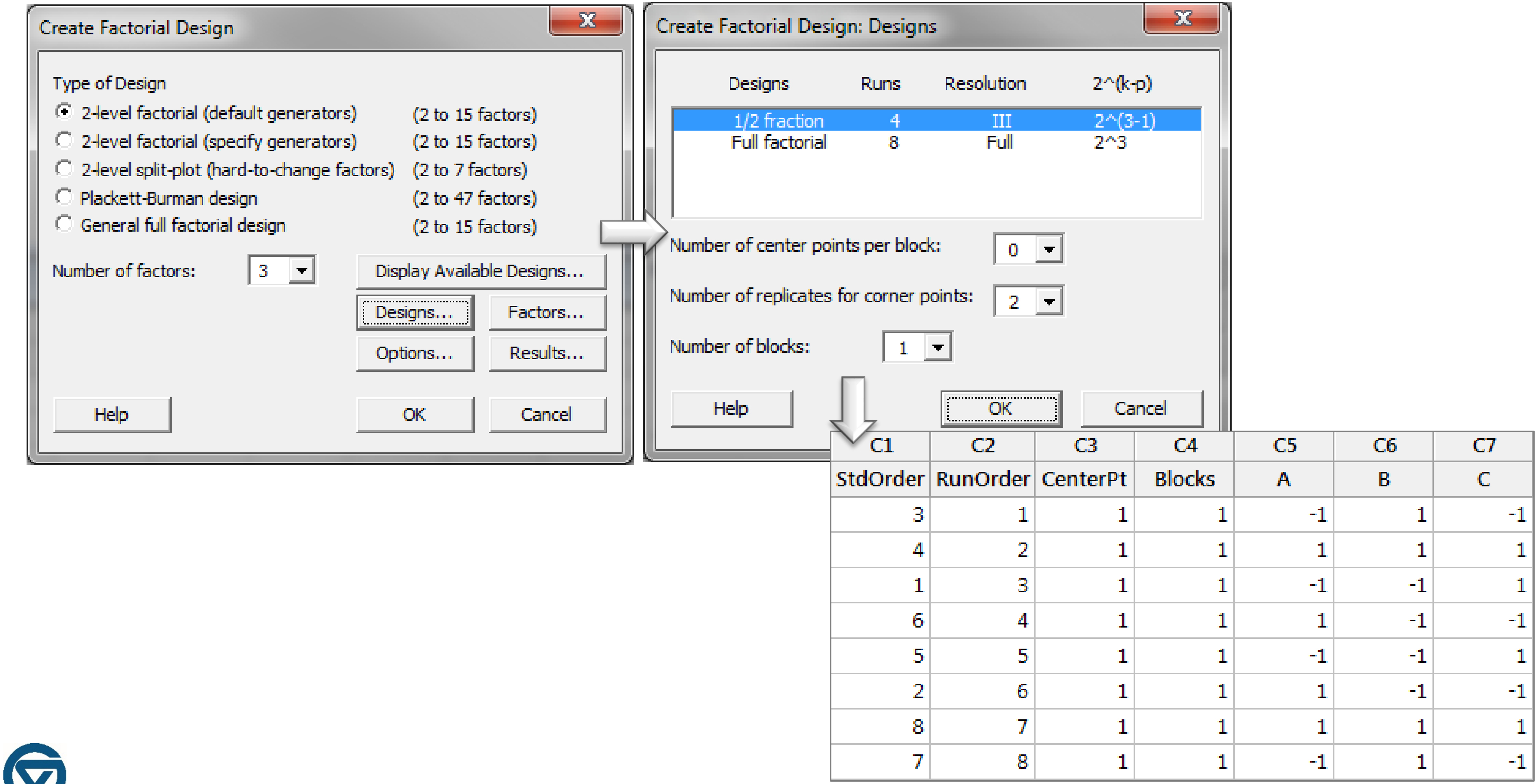
- Column ABC is called the generator.
  - By multiplying any column with the generator, we obtain its alias.
  - A * ABC = (A*A) * BC = I * BC = BC

Use Minitab to Run a Fractional Factorial Experiment

- Step 1: Initiate the experiment design

1) Click Stat → DOE → Factorial → Create Factorial Design.
2) A new window named “Create Factorial Design” pops up.
3) Select the radio button “2-level factorial (default generators).”
4) Enter “3” as the “Number of factors.”
5) Click on the “Design” button in the window “Create Factorial Design” and another new window named “Create Factorial Design – Designs” pops up.
6) Highlight the “1/2 fraction” design in the box.
7) Select “2” as the “Number of replicates for corner points.”
8) Click “OK” in the window “Create Factorial Design – Designs.”
9) Click “OK” in the window “Create Factorial Design.”
10) The design table is created in the data table.

Use Minitab to Run a Fractional Factorial Experiment



Create Factorial Design

Type of Design

- ● 2-level factorial (default generators) (2 to 15 factors)
- ○ 2-level factorial (specify generators) (2 to 15 factors)
- ○ 2-level split-plot (hard-to-change factors) (2 to 7 factors)
- ○ Plackett-Burman design (2 to 47 factors)
- ○ General full factorial design (2 to 15 factors)

Number of factors: 3

Create Factorial Design: Designs

Designs	Runs	Resolution	2^(k-p)
1/2 fraction	4	III	2^(3-1)
Full factorial	8	Full	2^3

Number of center points per block: 0

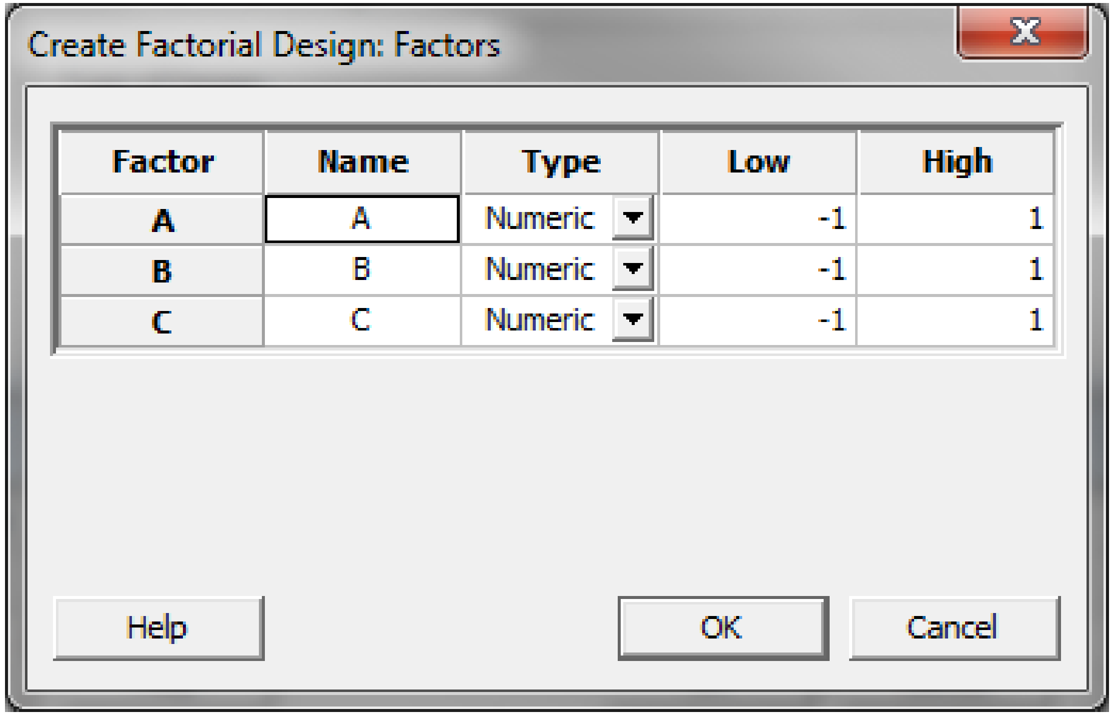
Number of replicates for corner points: 2

Number of blocks: 1

C1 StdOrder	C2 RunOrder	C3 CenterPt	C4 Blocks	C5 A	C6 B	C7 C
3	1	1	1	-1	1	-1
4	2	1	1	1	1	1
1	3	1	1	-1	-1	1
6	4	1	1	1	-1	-1
5	5	1	1	-1	-1	1
2	6	1	1	1	-1	-1
8	7	1	1	1	1	1
7	8	1	1	-1	1	-1

Use Minitab to Run a Fractional Factorial Experiment

- Step 1.2 (optional): Define the names and the levels for individual factors
  1) Click on the “Factors” button in the window “Create Factorial Design” and a new window named “Create Factorial Design – Factors” appears.
  2) Enter the names and the levels for individual factors. In this example, we keep the default names and levels for factors A, B, and C.

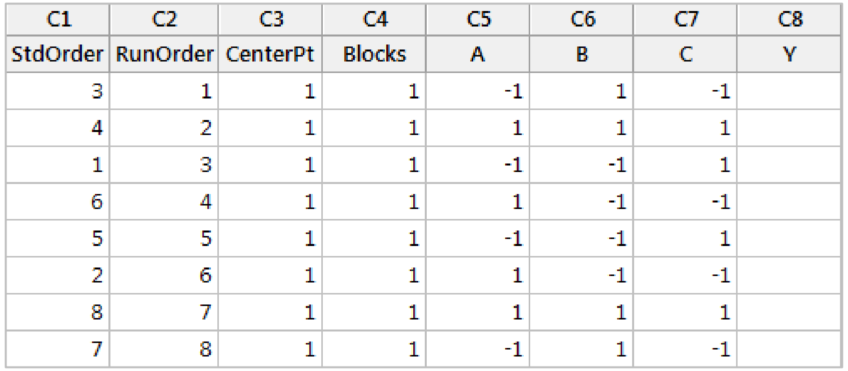


Create Factorial Design: Factors

Factor	Name	Type	Low	High
A	A	Numeric	-1	1
B	B	Numeric	-1	1
C	C	Numeric	-1	1

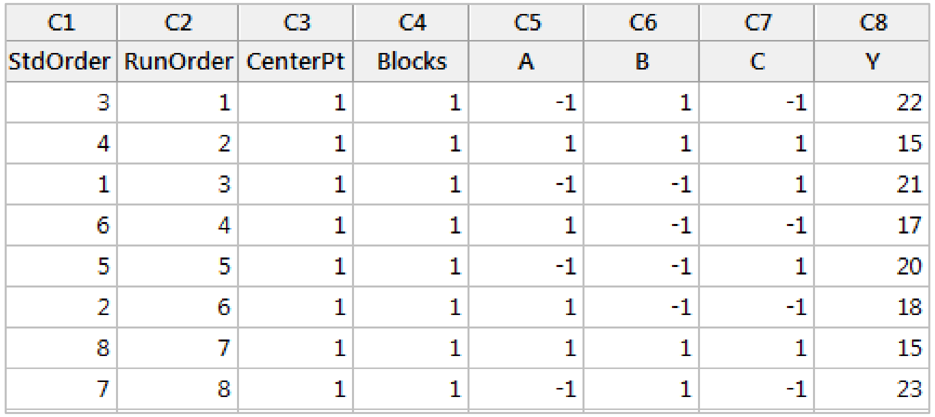
Use Minitab to Run a Fractional Factorial Experiment

- Step 2: Run the experiment in Run Order and record the response in a new column labeled “Y” in the table created by Minitab.



C1	C2	C3	C4	C5	C6	C7	C8
StdOrder	RunOrder	CenterPt	Blocks	A	B	C	Y
3	1	1	1	-1	1	-1	
4	2	1	1	1	1	1	
1	3	1	1	-1	-1	1	
6	4	1	1	1	-1	-1	
5	5	1	1	-1	-1	1	
2	6	1	1	1	-1	-1	
8	7	1	1	1	1	1	
7	8	1	1	-1	1	-1	

- We have run the experiment for you. Open data file “DOE Fractional.mtw”. The results of the experiment have been recorded in this file.

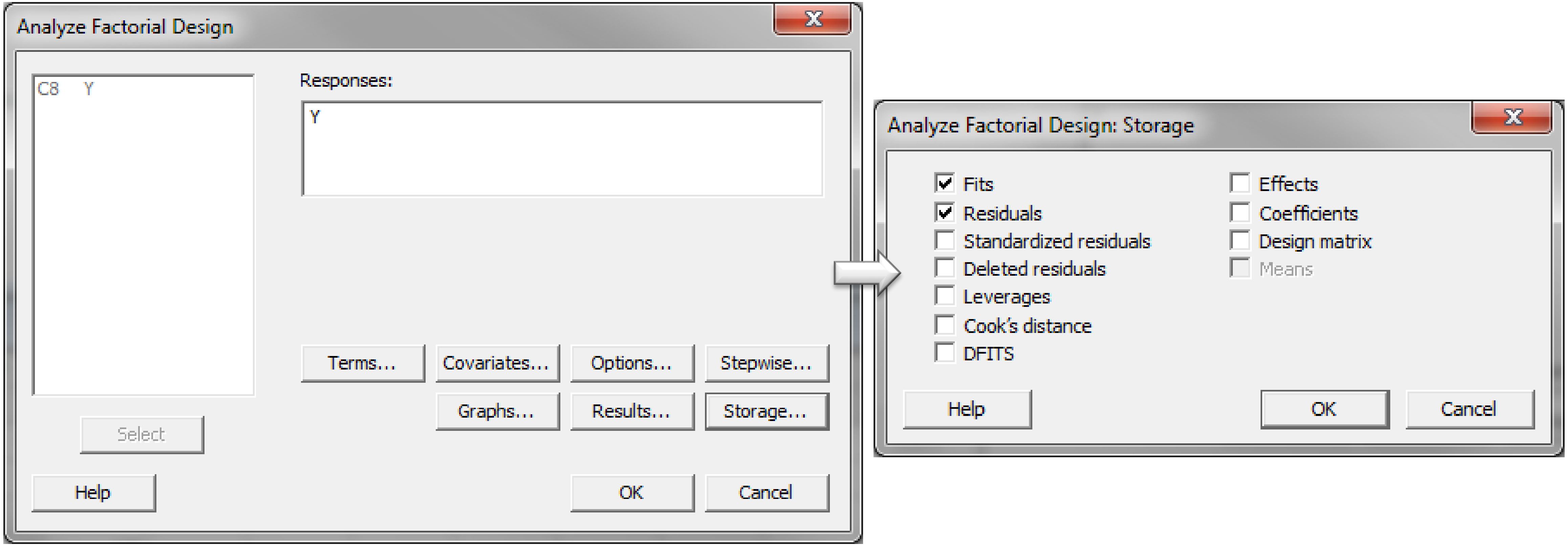


C1	C2	C3	C4	C5	C6	C7	C8
StdOrder	RunOrder	CenterPt	Blocks	A	B	C	Y
3	1	1	1	-1	1	-1	22
4	2	1	1	1	1	1	15
1	3	1	1	-1	-1	1	21
6	4	1	1	1	-1	-1	17
5	5	1	1	-1	-1	1	20
2	6	1	1	1	-1	-1	18
8	7	1	1	1	1	1	15
7	8	1	1	-1	1	-1	23

Use Minitab to Run a Fractional Factorial Experiment

- Step 3: Fit the model using the experiment results
  - Click Stat → DOE → Factorial → Analyze Factorial Design.
  - A new window named "Analyze Factorial Design" appears.
  - Select "Y" as the "Responses."
  - Click on the "Storage" button and another new window named "Analyze Factorial Design – Storage" pops up.
  - Check the box "Fits" and "Residuals" so that both fitted responses and the residuals would be saved in the data table.
  - Click "OK" in the window "Analyze Factorial Design – Storage."
  - Click "OK" in the window "Analyze Factorial Design."
  - The DOE analysis results appear in the session window.

Use Minitab to Run a Fractional Factorial Experiment



Analyze Factorial Design

C8 Y

Responses: Y

Analyze Factorial Design: Storage

☑ Fits
☑ Residuals
☐ Standardized residuals
☐ Deleted residuals
☐ Leverages
☐ Cook's distance
☐ DFITS
☐ Effects
☐ Coefficients
☐ Design matrix
☐ Means

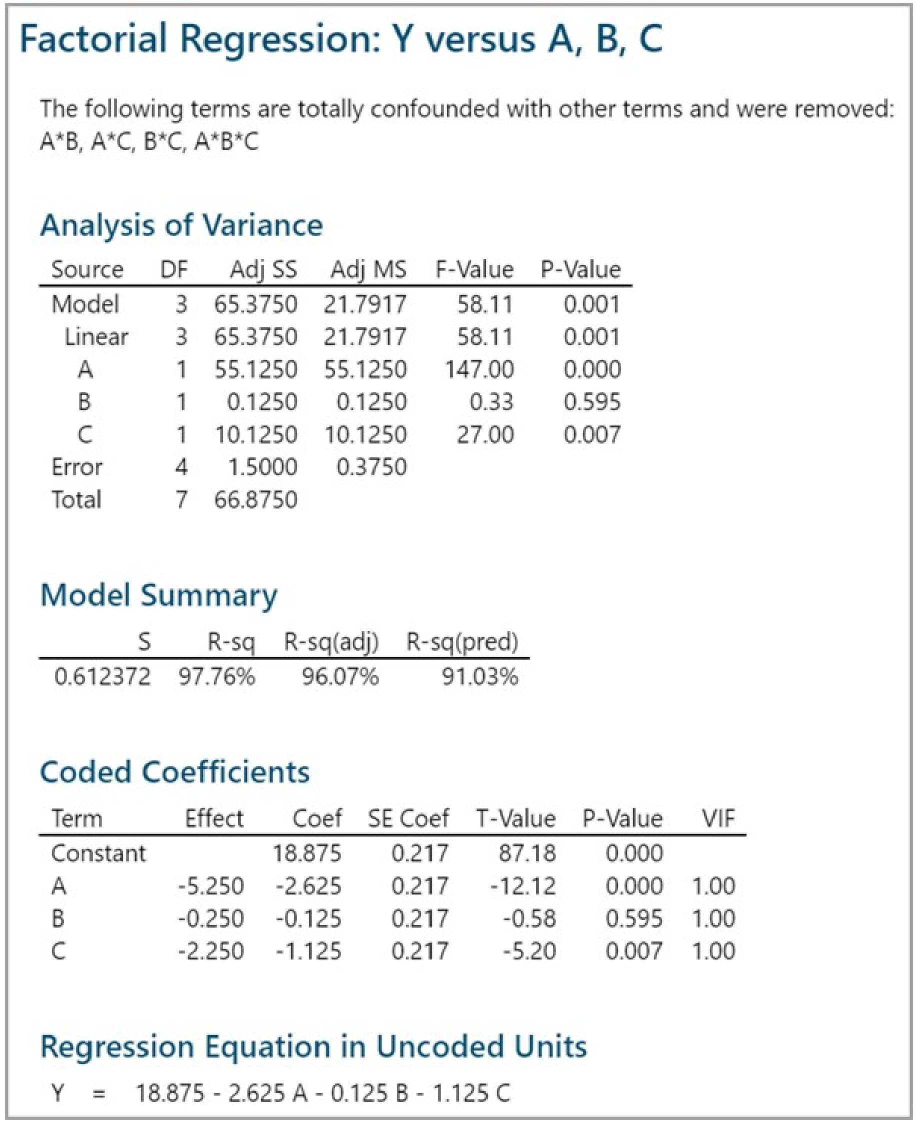
Use Minitab to Run a Fractional Factorial Experiment

- Step 4: Analyze the model results
  1) Check whether the model is statistically significant.
  2) Check which factors are insignificant.
  3) If any independent variables are not significant, remove them one at a time and rerun the model until all the independent variables in the model are significant.

Use Minitab to Run a Fractional Factorial Experiment

- The p-value of factor B is greater than the alpha level (0.05), so it is not statistically significant.

- Next, we need to remove factor B and rerun the model.

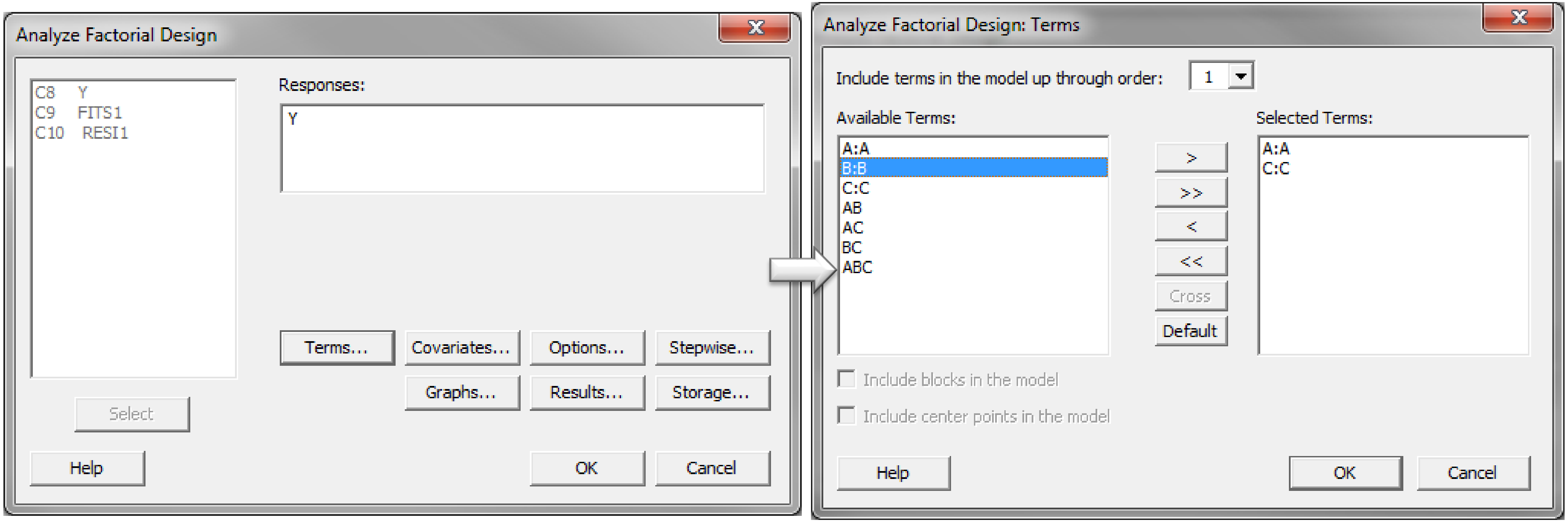


```
Factorial Regression: Y versus A, B, C

The following terms are totally confounded with other terms and were removed:
A*B, A*C, B*C, A*B*C

Analysis of Variance

Source   DF   Adj SS   Adj MS  F-Value  P-Value
Model     3  65.3750  21.7917    58.11    0.001
  Linear  3  65.3750  21.7917    58.11    0.001
    A     1  55.1250  55.1250   147.00    0.000
    B     1   0.1250   0.1250     0.33    0.595
    C     1  10.1250  10.1250    27.00    0.007
Error     4   1.5000   0.3750
Total     7  66.8750

Model Summary

       S    R-sq  R-sq(adj)  R-sq(pred)
0.612372  97.76%     96.07%      91.03%

Coded Coefficients

Term      Effect    Coef  SE Coef  T-Value  P-Value   VIF
Constant          18.875    0.217    87.18    0.000
A         -5.250  -2.625    0.217   -12.12    0.000  1.00
B         -0.250  -0.125    0.217    -0.58    0.595  1.00
C         -2.250  -1.125    0.217    -5.20    0.007  1.00

Regression Equation in Uncoded Units

Y  =  18.875 - 2.625 A - 0.125 B - 1.125 C
```

Use Minitab to Run a Fractional Factorial Experiment

- In this example, since factor B is not statistically significant, it needs to be removed from the model.
  - Click Stat → DOE → Factorial → Analyze Factorial Design.
  - A new window named “Analyze Factorial Design” appears.
  - Select “Y” as the “Responses.”
  - Click on the “Terms” button and another new window named “Analyze Factorial Design – Terms” pops up.
  - Deselect the factor “B:B” from the box “Selected Terms.”
  - Click “OK” in the window “Analyze Factorial Design – Terms.”
  - Click “OK” in the window “Analyze Factorial Design.”
  - The DOE analysis results appear in the session window.

Use Minitab to Run a Fractional Factorial Experiment



Analyze Factorial Design

C8 Y
C9 FITS1
C10 RESI1

Responses: Y

Analyze Factorial Design: Terms

Include terms in the model up through order: 1

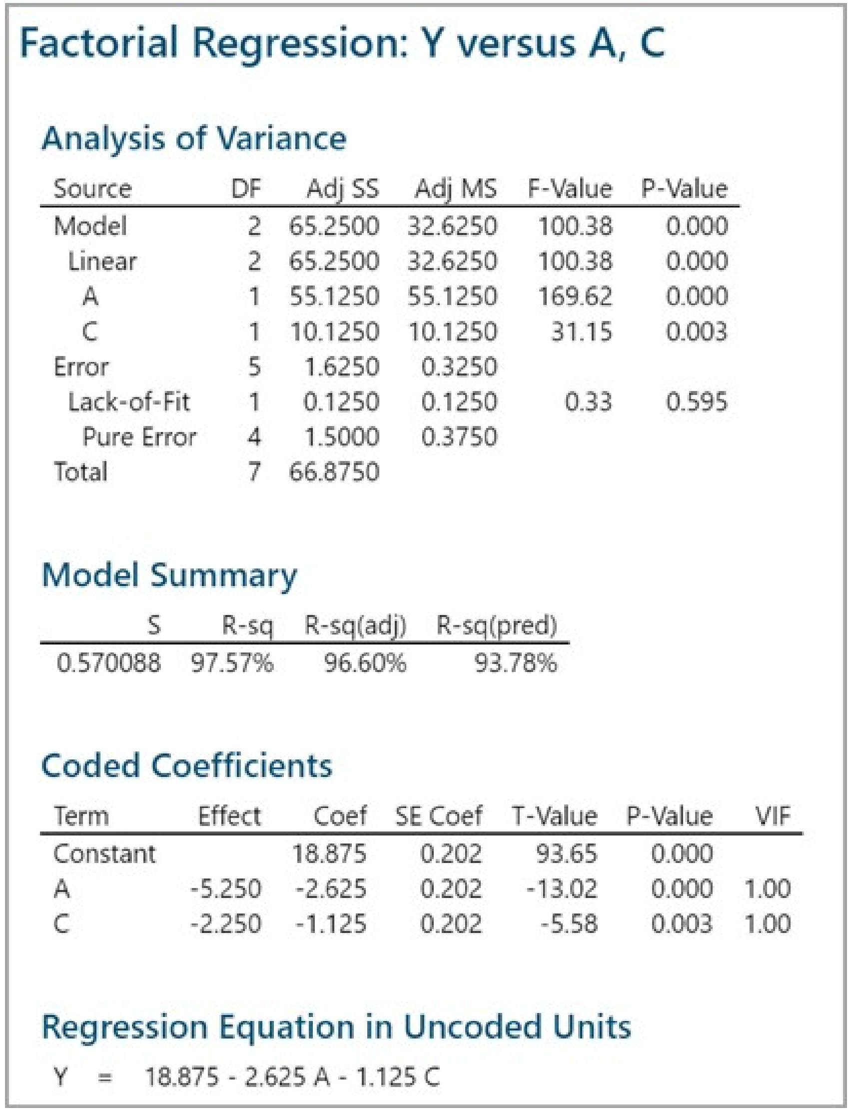
Available Terms:
A:A
B:B
C:C
AB
AC
BC
ABC

Selected Terms:
A:A
C:C

☐ Include blocks in the model
☐ Include center points in the model

Use Minitab to Run a Fractional Factorial Experiment

- The p-values of all the independent variables are smaller than 0.05. There is no need to remove any independent variables from the model.



```
Factorial Regression: Y versus A, C

Analysis of Variance

Source          DF   Adj SS    Adj MS   F-Value   P-Value
Model            2  65.2500   32.6250    100.38     0.000
  Linear         2  65.2500   32.6250    100.38     0.000
    A            1  55.1250   55.1250    169.62     0.000
    C            1  10.1250   10.1250     31.15     0.003
Error            5   1.6250    0.3250
  Lack-of-Fit    1   0.1250    0.1250      0.33     0.595
    Pure Error   4   1.5000    0.3750
Total            7  66.8750

Model Summary

       S    R-sq   R-sq(adj)   R-sq(pred)
0.570088  97.57%      96.60%       93.78%

Coded Coefficients

Term        Effect     Coef  SE Coef  T-Value  P-Value   VIF
Constant             18.875    0.202    93.65    0.000
A           -5.250   -2.625    0.202   -13.02    0.000  1.00
C           -2.250   -1.125    0.202    -5.58    0.003  1.00

Regression Equation in Uncoded Units

Y  =  18.875 - 2.625 A - 1.125 C
```

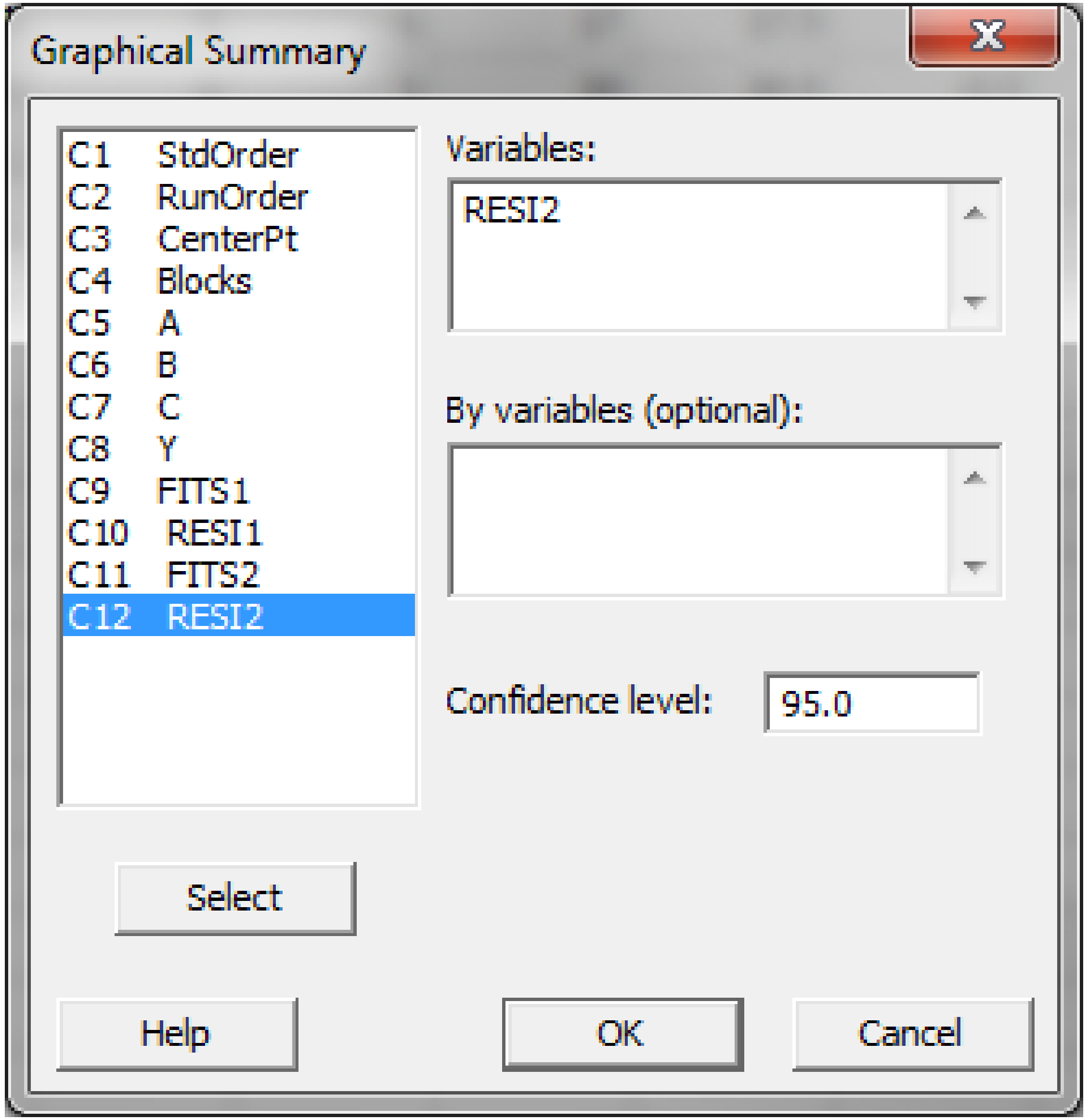
Use Minitab to Run a Fractional Factorial Experiment

- Step 5: Conduct residuals analysis to ensure that the residuals of the model satisfy the following criteria.
    - Their mean is equal to zero.
    - They are normally distributed.
    - They are independent.
    - There is equal variance across the fitted response values.

Use Minitab to Run a Fractional Factorial Experiment

- Step 5.1: Check whether residuals are normally distributed with mean equal to zero.
    - Click Stat → Basic Statistics → Graphical Summary.
    - A new window named “Graphical Summary” appears.
    - Select “RESI2” as the “Variables.”
    - Click “OK.”
    - The histogram and the normality test of the residuals appear in the newly-generated window.
    - If the p-value of the normality test is greater than the alpha level (0.05), the residuals are normally distributed.

Use Minitab to Run a Fractional Factorial Experiment



Variables: RESI2

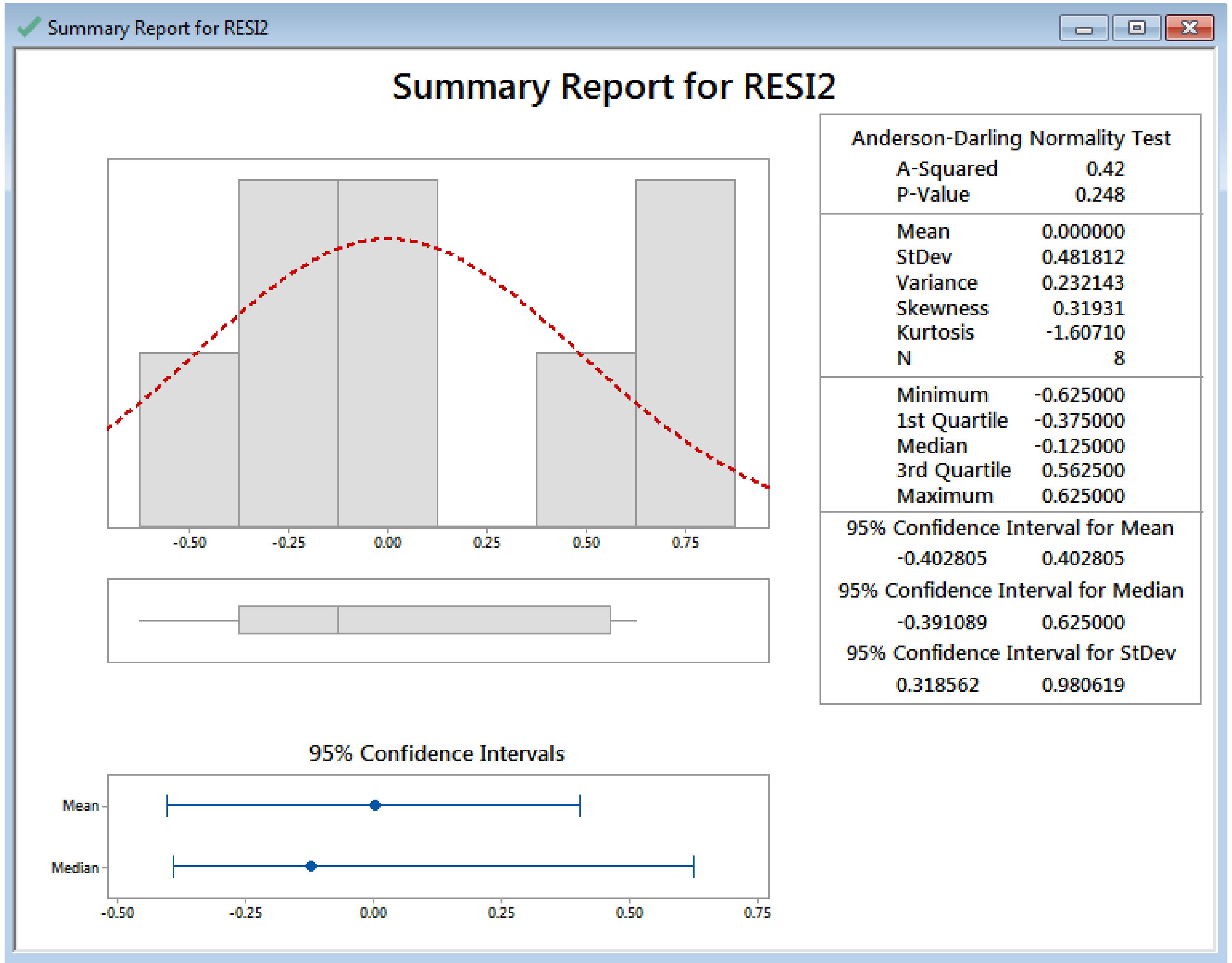
By variables (optional):

Confidence level: 95.0

Use Minitab to Run a Fractional Factorial Experiment

- The p-value of the normality test is larger than alpha level (0.05). The residuals are normally distributed.

- Residuals’ mean is zero.



Summary Report for RESI2

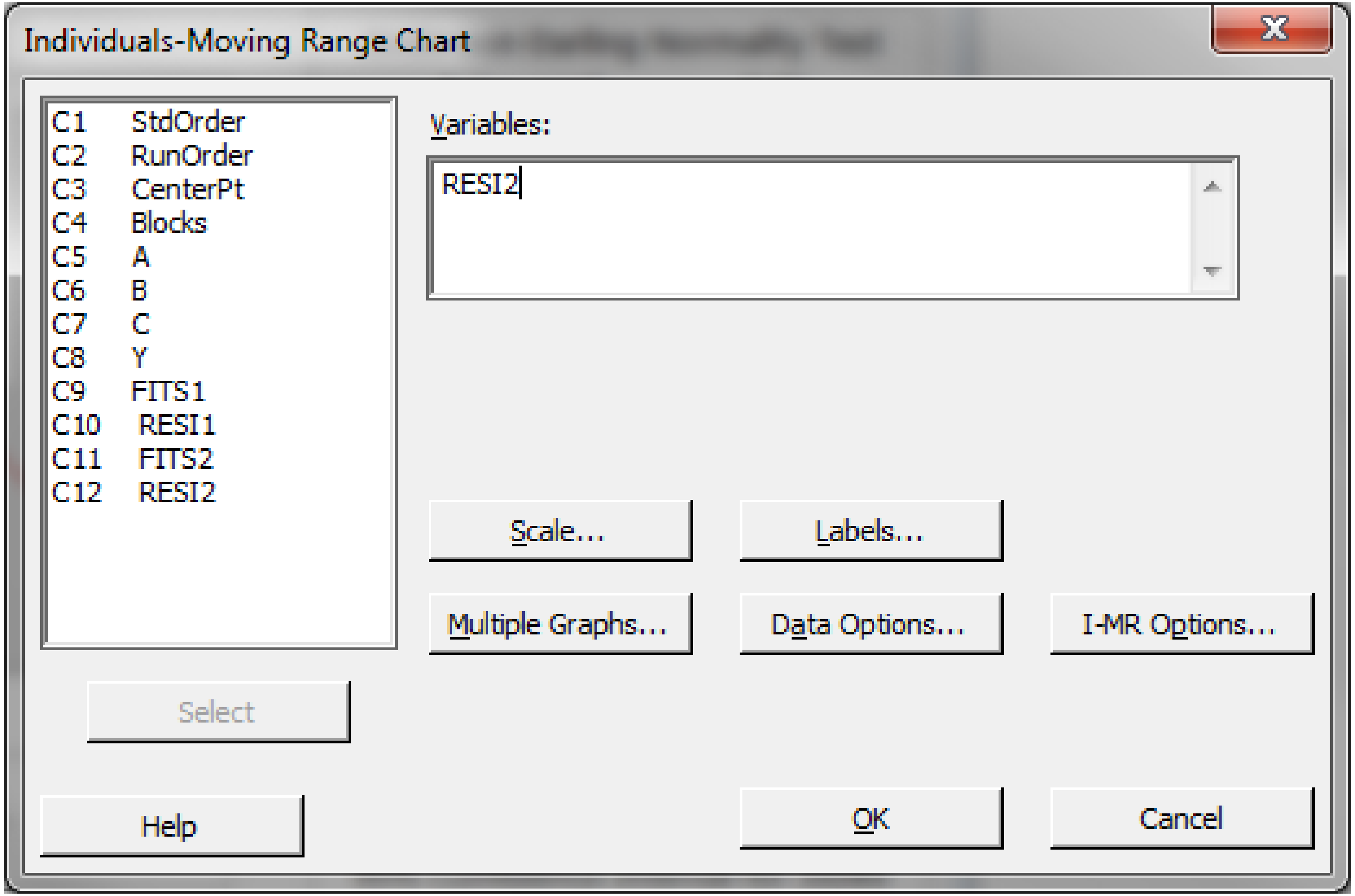
```
Anderson-Darling Normality Test
   A-Squared            0.42
   P-Value             0.248

   Mean            0.000000
   StDev           0.481812
   Variance        0.232143
   Skewness         0.31931
   Kurtosis        -1.60710
   N                      8

   Minimum        -0.625000
   1st Quartile   -0.375000
   Median         -0.125000
   3rd Quartile    0.562500
   Maximum         0.625000
95% Confidence Interval for Mean
   -0.402805       0.402805
95% Confidence Interval for Median
   -0.391089       0.625000
95% Confidence Interval for StDev
   0.318562        0.980619
```

Use Minitab to Run a Fractional Factorial Experiment

- Step 5.2: Check whether the residuals are independent.
  - Click Stat → Control Charts → Variables Charts for Individuals → I-MR.
  - A new window named “Individuals – Moving Range Chart” appears.
  - Select the “RESI2” as the “Variables.”
  - Click “OK.”
  - The control charts appear in the newly-generated window.
  - If no data points on the control charts fail any tests, the residuals are in control and independent of each other.

Note: The prerequisite of plotting I-MR chart for residuals: the residuals are in the time order.

Use Minitab to Run a Fractional Factorial Experiment

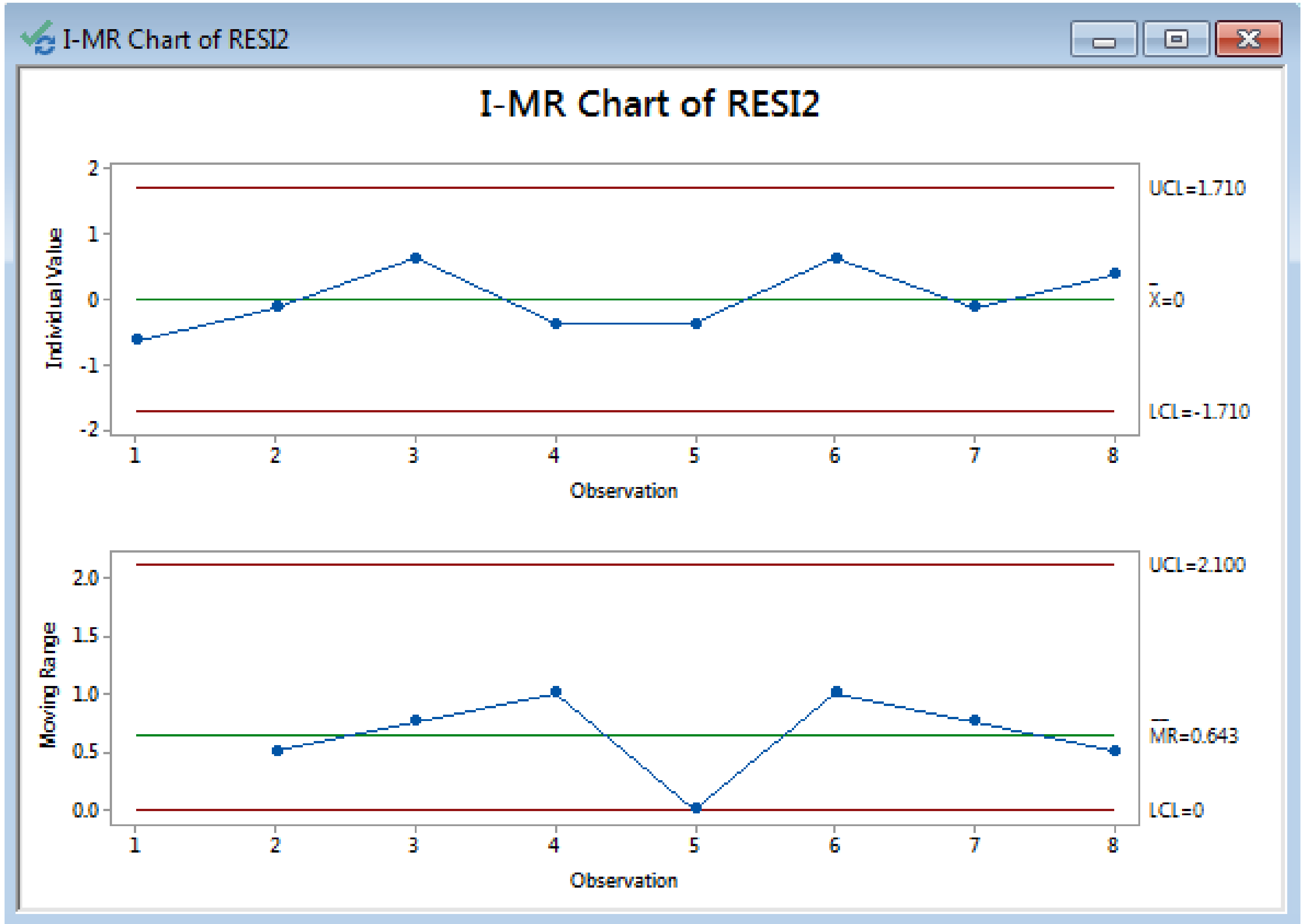


Individuals-Moving Range Chart

C1	StdOrder
C2	RunOrder
C3	CenterPt
C4	Blocks
C5	A
C6	B
C7	C
C8	Y
C9	FITS1
C10	RESI1
C11	FITS2
C12	RESI2

Variables: RESI2

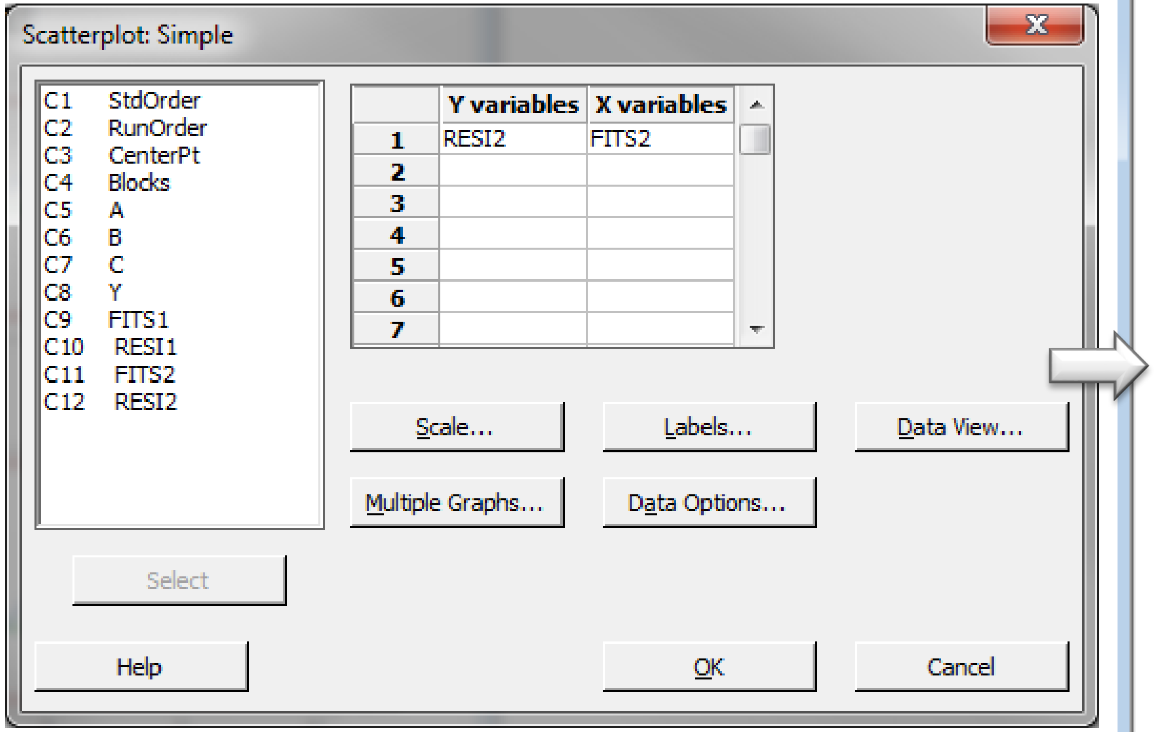
Use Minitab to Run a Fractional Factorial Experiment



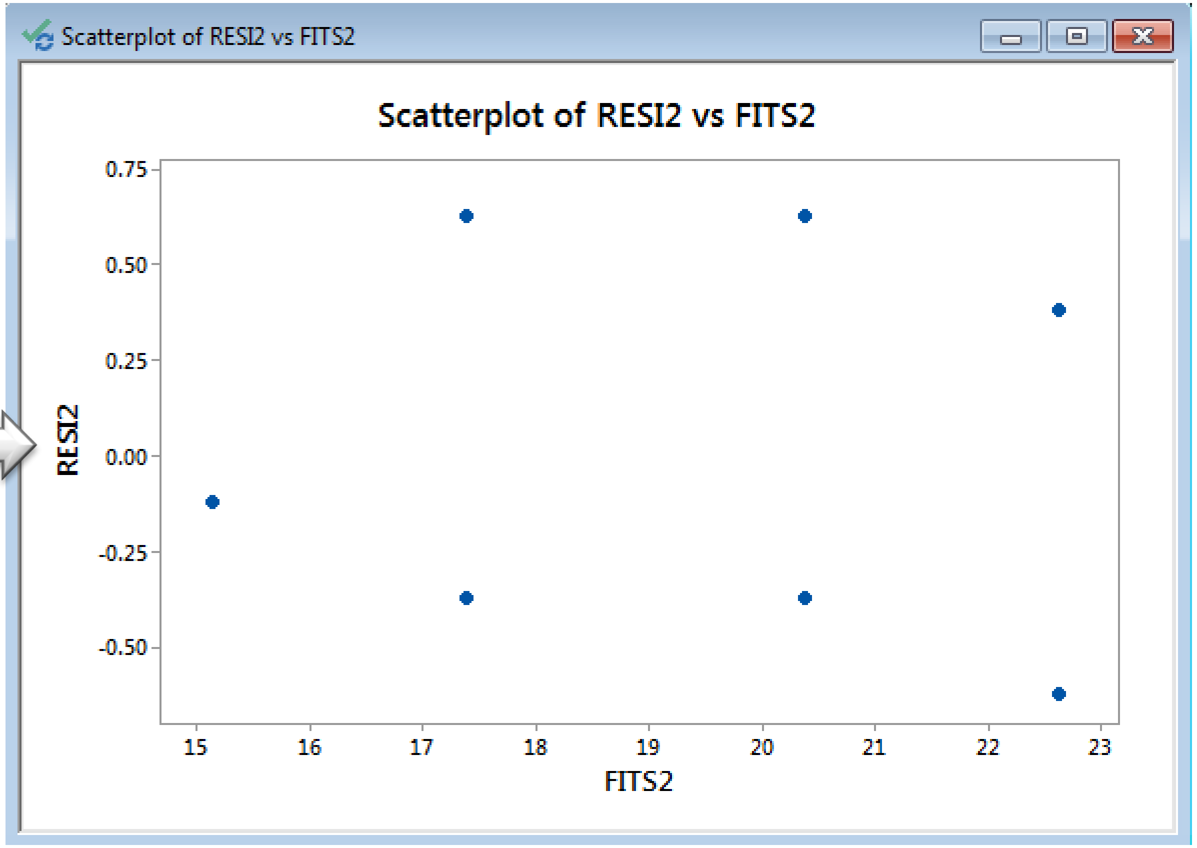
Use Minitab to Run a Fractional Factorial Experiment

- Step 5.3: Check whether the residuals have equal variance across the predicted response values.
  - Click Graph → Scatterplot.
  - A new window named “Scatterplots” pops up.
  - Click “OK” in the window “Scatterplots.”
  - Another window named “Scatterplot – Simple” pops up.
  - Select “RESI2” as “Y variables” and “FITS2” as “X variables.”
  - Click “OK” in the window “Scatterplot – Simple.”
  - The scatterplot between the fitted responses and the residuals appears in a new window.
  - We look for patterns in which the residuals tend to have even variation across the entire range of the fitted response values.

Use Minitab to Run a Fractional Factorial Experiment


Scatterplot: Simple

	Y variables	X variables
1	RESI2	FITS2
2		
3		
4		
5		
6		
7		



4.5.2 Confounding Effects

Confounding Effects

- In full factorial experiments the effects of every factor and their interactions can be calculated because there is a degree of freedom for every treatment combination as well as the error term.
- Fractional factorial experiments, however, are designed with fewer runs or treatment combinations but will have the same number of input factors.
- As a result, the experimenter must understand the impact of these consequences.
- **Confounding** (or aliased factors) is that consequence.
- **Resolution** is a quantification or degree of confounding.

Confounding Effects

- When two input factors are aliases of each other, the effects they each have on the response cannot be separated.
    - In the 2^{3-1} fractional factorial design, A and BC are aliases.
    - The main effect that A has on the response would be exactly the same as the interaction effect of BC. We cannot tell which one is truly significant.

- The effects of aliased factors are what is referred to as confounded.
- Due to the confounding nature of aliased factors, the effect we observe is the mixed effect of aliased factors.
    - A + BC
    - B + AC
    - C + AB

Resolution

- **Resolution** is the measure or degree of confounding.
- Higher resolution means less confounding or merely confounding main effects with higher order interactions.
- The resolution of a fractional factorial experiment will always be one number higher than the order of interactions that are confounded with main effects.
- **Resolution III**: Main effects are confounded with two-way interactions. If you are only interested in main effects, then resolution III is acceptable. Otherwise, resolution III is typically undesirable.
- **Resolution IV**: Main effect are confounded with three-way interactions.
- **Resolution V**: Main effects are confounded with four-way interactions.

Consequences of Fractional Factorials

- Benefits:
  - The primary benefits of properly designed fractional factorial experiments are achieved because of fewer runs.
  - You may have limited time, resources, or capital and cannot run a full factorial.
- Disadvantages:
  - Confounding (a.k.a. aliasing)

- Mitigation Option
  - **Replications**: by adding replicates you can introduce more study runs and increase the resolution of your experiment.
  - Replications, however, will also increase the number of trials or runs sometimes, defeating the purpose of running a fractional factorial experiment.

4.5.3 Experimental Resolution

Alias in 2^{3-1} Fractional Factorial Experiment

- In the 2^{3-1} fractional factorial design of experiment, the main factor is aliased or confounded with two-way interactions.
  - A is the alias of BC.
  - B is the alias of AC.
  - C is the alias of AB.

Alias in 2^{4-1} Fractional Factorial Experiment

- In the 2^{4-1} fractional factorial design of experiment, the main factor is aliased with three-way interaction and two-way interaction is aliased with two-way interaction.
  - A is the alias of BCD.
  - B is the alias of ACD.
  - C is the alias of ABD.
  - D is the alias of ABC.
  - AB is the alias of CD.
  - AC is the alias of BD.
  - AD is the alias of BC.

Design Resolution

- Design resolution (i.e., experimental resolution) refers to the confounding patterns indicating how the effects are confounded with others.
- If two factors are aliases, they have confounded effect on the response. There is no need to consider both factors in the model due to the redundancy. Only one factor needs to be included in the model for simplicity.
- Most popular resolutions are resolution III, IV, and V.
  - Resolution III: Main effects are aliased with two-way interactions.
  - Resolution IV: Main effects are aliased with three-way interactions. Two-way interactions are aliased with other two-way interactions.
  - Resolution V: Main effects are aliased with four-way interactions. Two-way interactions are aliased with three-way interactions.

Determine Aliases

- To find the alias of a particular factor, we only need to multiply the factor of interest with the identity (I) of the design.
  - In 2^{3-1} fractional factorial design, ABC is the identity. To find the alias of factor A, we use A * ABC = BC.
  - In 2^{4-1} fractional factorial design, ABCD is the identity. To find the alias of factor A, we use A * ABCD = BCD.
  - In 2^{5-1} fractional factorial design, ABCDE is the identity. To find the alias of factor A, we use A * ABCDE = BCDE.

5.0 Control Phase

Black Belt Training: Control Phase

5.1 Lean Controls
  5.1.1 Control Methods for 5S
  5.1.2 Kanban
  5.1.3 Poka-Yoke (Mistake Proofing)

5.2 Statistical Process Control (SPC)
  5.2.1 Data Collection for SPC
  5.2.2 I-MR Chart
  5.2.3 Xbar-R Chart
  5.2.4 U Chart
  5.2.5 P Chart
  5.2.6 NP Chart
  5.2.7 X-S chart
  5.2.8 CumSum Chart
  5.2.9 EWMA Chart
  5.2.10 Control Methods
  5.2.11 Control Chart Anatomy
  5.2.12 Subgroups, Variation, Sampling

5.3 Six Sigma Control Plans
  5.3.1 Cost Benefit Analysis
  5.3.2 Elements of the Control Plan
  5.3.3 Elements of the Response Plan

5.1 Lean Controls

Black Belt Training: Control Phase

5.1.1 Control Methods for 5S

What is 5S?

- 5S is a systematic method to organize, order, clean, and standardize a workplace…and keep it that way!
  - 5S is a methodology of organizing and improving the work environment.

- 5S is summarized in five Japanese words, all starting with the letter S:
  - **Seiri** (sorting)
  - **Seiton** (straightening)
  - **Seiso** (shining)
  - **Seiketsu** (standardizing)
  - **Shisuke** (sustaining).

- 5S was originally developed in Japan, and is widely used to optimize the workplace to increase productivity and efficiency.

5S Goals

- Reduced waste
- Reduced cost
- Establish a work environment that is:
  - self-explaining
  - self-ordering
  - self-regulating
  - self improving.
  - Where there is/are **no more:**
    - Wandering and/or searching
    - Waiting or delays
    - Secrets hiding spots for tools
    - Obstacles or detours
    - Extra pieces, parts, materials, etc.
    - Injuries
    - Waste.

5S Benefits

- Reduced changeovers
- Reduced defects
- Reduced waste
- Reduced delays
- Reduced injuries
- Reduced breakdowns
- Reduced complaints
- Reduced red ink
- Higher quality
- Lower costs
- Safer work environment
- Greater associate and equipment capacity

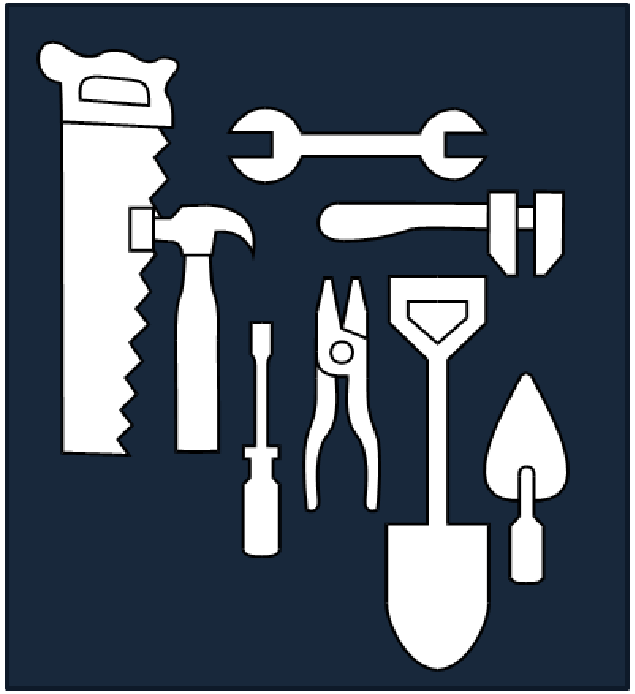
5S Systems Reported Results

- Cut in floor space: 60%
- Cut in flow distance: 80%
- Cut in accidents: 70%
- Cut in rack storage: 68%
- Cut in number of forklifts: 45%
- Cut in machine changeover time: 62%
- Cut in annual physical inventory time: 50%
- Cut in classroom training requirements: 55%
- Cut in nonconformance in assembly: 96%
- Increase in test yields: 50%
- Late deliveries: 0%
- Increase in throughput: 15%

Sorting (Seiri)

- Go through all tools, parts, equipment, supply, and materials in the workplace.
- Categorize into two major groups: needed and unneeded.
- Eliminate unneeded items from the workplace. Dispose of or recycle those items.
- Keep need items and sort in order of priority. **When in doubt…throw it out!**

Straightening (Seiton)



- **Straightening** in 5S is also called **setting in order.**
  - Label each needed item.
  - Store items at their best locations so that the workers can find them easily whenever they needed any item.
  - Reduce the motion and time required to locate and obtain any item whenever it is needed.
  - Promote an efficient work flow path.
  - Use visual aids like the tool board image on this page.

Shining (Seiso)

- **Shining** in 5S is also called **sweeping**.
- Clean the workplace thoroughly.
- Maintain the tidiness of the workplace.
- Make sure every item is located at the specific location where it should be.
- Create the ownership in the team to keep the work area clean and organized.

Standardizing (Seiketsu)

- **Standardize** the workstation and the layout of tools, equipment, and parts.

- Create identical workstations with a consistent way of storing the items at their specific locations so that workers can be moved around to any workstation any time and perform the same task.

Sustaining (Shisuke)

- Sustaining in 5S is also called self-discipline.
- Create the culture in the team to follow the first four S’s consistently.
- Avoid falling back to the old ways of cluttered and unorganized work environment.
- Keep the momentum of optimizing the workplace.
- Promote innovations of workplace improvement.
- Sustain the first fours S’s using:
  - 5S Maps
  - 5S Schedules
  - 5S Job cycle charts
  - Integration of regular work duties
  - 5S Blitz schedules
  - Daily workplace scans.

Simplified Summary of 5S

- **Sort** – "when in doubt, move it out."
- **Set in Order** – Organize all necessary tools, parts, and components of production. Use visual ordering techniques wherever possible.
- **Shine** – Clean machines and/or work areas. Set regular cleaning schedules and responsibilities.
- **Standardize** – Solidify previous three steps, make 5S a regular part of the work environment and everyday life.
- **Sustain** – Audit, manage, and comply with established five-s guidelines for your business or facility

5.1.2 Kanban

What is Kanban?

- The Japanese word “Kanban” means “signboard.”
- **Kanban system** is a “pull” production scheduling system to determine when to produce, what to produce, and how much to produce based on the demand.
- It was originally developed by Taiichi Ohno in order to reduce the waste in inventory and increase the speed of responding to the immediate demand.

Kanban System

- Kanban system is a demand-driven system.
- The customer demand is the signal to trigger or pull the production.
- Products are made only to meet the immediate demand. When there is no demand, there is no production.
- It is designed to minimize the in-process inventory and to have the right material with the right amount at the right location at the right time.

Kanban System

- Principles of the Kanban System:
  - Only produce products with exactly the same amount that customers consume.
  - Only produce products when customers consume.
- The production is driven by the actual demand from the customer side instead of the forecasted demand planned by the staff.

Kanban Card

- The Kanban card is the ticket or signal to authorize the production or movement of materials. It is the message of asking for more.

- It is sent from the end customer up to the chain of production.

- Upon receiving of a Kanban card, the production station would start to produce goods.

- The Kanban card can be a physical card or an electronic signal.

Kanban System Example

- The simplest example of a Kanban system is the supermarket operation.
- Customers visit the supermarkets and buy what they need.
- The checkout scanners send electronic Kanban cards to the local warehouse asking for more when the items are sold to customers.
- When the warehouse receives the Kanban cards, it starts to replenish the exact goods being sold.
- It the warehouse prepares more than what Kanban cards require, the goods would become obsolete. If it prepares less, the supermarket would not have the goods available when customers need them.

Kanban System Benefits

- Minimize in-process inventory
- Free up space occupied by unnecessary inventory
- Prevent overproduction
- Improve responsiveness to dynamic demand
- Avoid the risk of inaccurate demand forecast
- Streamline the production flow
- Visualize the work flow.

5.1.3 Poka-Yoke

What is Poka-Yoke?

- The Japanese term “poka-yoke” means “mistake-proofing.”
- It is a mechanism to eliminate defects as early as possible in the process.
- It was originally developed by Shigeo Shingo and was initially called “baka-yoke” (fool-proofing).

Two Types of Poka-Yoke

- Prevention
  - Preventing defects from occurring
  - Removing the possibility that an error could occur
  - Making the occurrence of an error impossible.

- Detection
  - Detecting defects once they occur
  - Highlighting defects to draw workers’ attention immediately
  - Correcting defects so that they would not reach the next stage.

Three Methods of Poka-Yoke

- Contact Method
  - Use of shape, color, size, or any other physical attributes of the items.
- Constant Number Method
  - Use of a fixed number to make sure a certain number of motions are completed.
- Sequence Method
  - Use of a checklist to make sure all the prescribed process steps are followed in the right order.

Poka-Yoke Devices

- We are surrounded by poka-yoke devices daily.
  - Prevention Devices
    - Example: the dishwasher does not start to run when the door is open.
  - Detection Devices
    - Example: the car starts to beep when the passengers do not buckle their seatbelts.

- Poka-yoke devices can be in any format that can quickly and effectively prevent or detect mistakes.
  - Visual, electrical, mechanical, procedural, human etc.

Steps to Apply Poka-Yoke

- Step 1: Identify the process steps in need of mistake proofing.
- Step 2: Use the 5-why's method to analyze the possible mistakes or failures for the process step.
- Step 3: Determine the type of poka-yoke: prevention or detection.
- Step 4: Determine the method of poka-yoke: contact, constant number, or sequence.
- Step 5: Pilot the poka-yoke approach and make any adjustments if needed.
- Step 6: Implement poka-yoke in the operating process and maintain the performance.

5.2 Statistical Process Control

Black Belt Training: Control Phase

5.1 Lean Controls
  5.1.1 Control Methods for 5S
  5.1.2 Kanban
  5.1.3 Poka-Yoke (Mistake Proofing)

5.2 Statistical Process Control (SPC)
  5.2.1 Data Collection for SPC
  5.2.2 I-MR Chart
  5.2.3 Xbar-R Chart
  5.2.4 U Chart
  5.2.5 P Chart
  5.2.6 NP Chart
  5.2.7 X-S chart
  5.2.8 CumSum Chart
  5.2.9 EWMA Chart
  5.2.10 Control Methods
  5.2.11 Control Chart Anatomy
  5.2.12 Subgroups, Variation, Sampling

5.3 Six Sigma Control Plans
  5.3.1 Cost Benefit Analysis
  5.3.2 Elements of the Control Plan
  5.3.3 Elements of the Response Plan

5.2.1 Data Collection for SPC

What is SPC?

- Statistical process control (SPC) is a statistical method to monitor the performance of a process using control charts in order to keep the process in statistical control.
- Statistical process control can be used to distinguish between special cause variation and common cause variation in the process.
- It presents the voice of the process.

Common Cause Variation

- Common cause variation (also called chance variation) is the inherent natural variation in any processes.
- It is the random background noise, which cannot be controlled or eliminated from the process.
- Its presence in the process is expected and acceptable due to its relatively small influence on the process.

Special Cause Variation

- **Special cause variation** (also called assignable cause variation) is the unnatural variation in the process.

- It is the cause of process instability and leads to defects of the products or services.

- It is the signal of unanticipated change (either positive or negative) in the process.

- It is possible to eliminate the special cause variation from the process.

Process Stability

- A process is **stable** when:
  - There is not any special cause variation involved in the process
  - The process is in statistical control
  - The future performance of the process is predictable within certain limits
  - The changes happening in the process are all due to random inherent variation
  - There are not any trends, unnatural patterns, and outliers in the control chart of the process.

SPC Benefits

- Statistical process control can be used in different phases of Six Sigma projects to:
    - Understand the stability of a process
    - Detect the special cause variation in the process
    - Identify the statistical difference between two phases
    - Eliminate or apply the unnatural change in the process
    - Improve the quality and productivity.

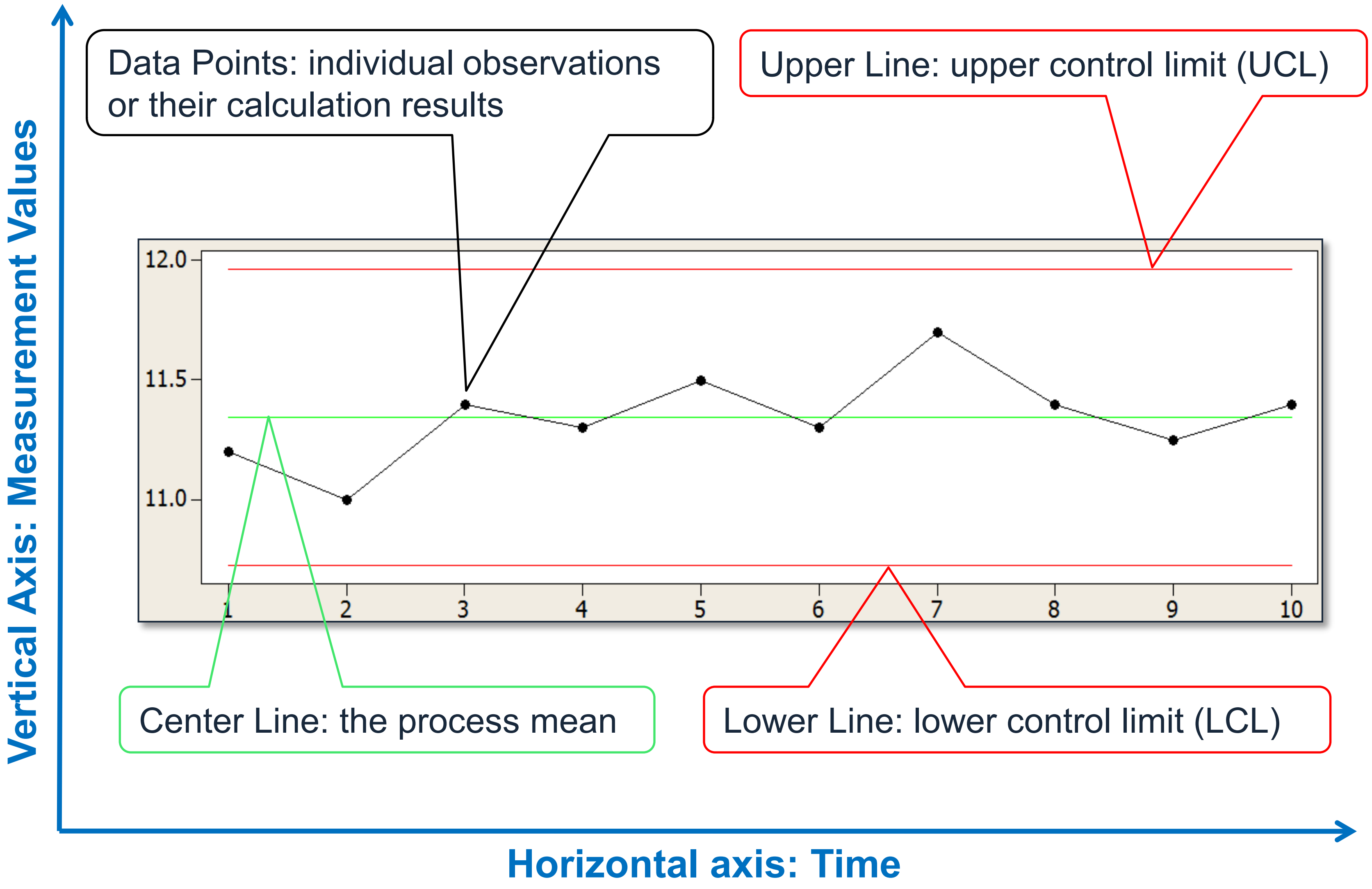
Control Charts

- **Control charts** are graphical tools to present and analyze the process performance in statistical process control.

- Control charts are used to detect special cause variation and determine whether the process is in statistical control (stable).

- Variation solutions:
  - Minimize the common cause variation
  - Eliminate the special cause variation when it leads to unanticipated negative changes in the outcome
  - Implement the special cause variation when it leads to unanticipated positive changes in the outcome.

Control Charts Elements



Data Points: individual observations or their calculation results

Upper Line: upper control limit (UCL)

Center Line: the process mean

Lower Line: lower control limit (LCL)

Vertical Axis: Measurement Values

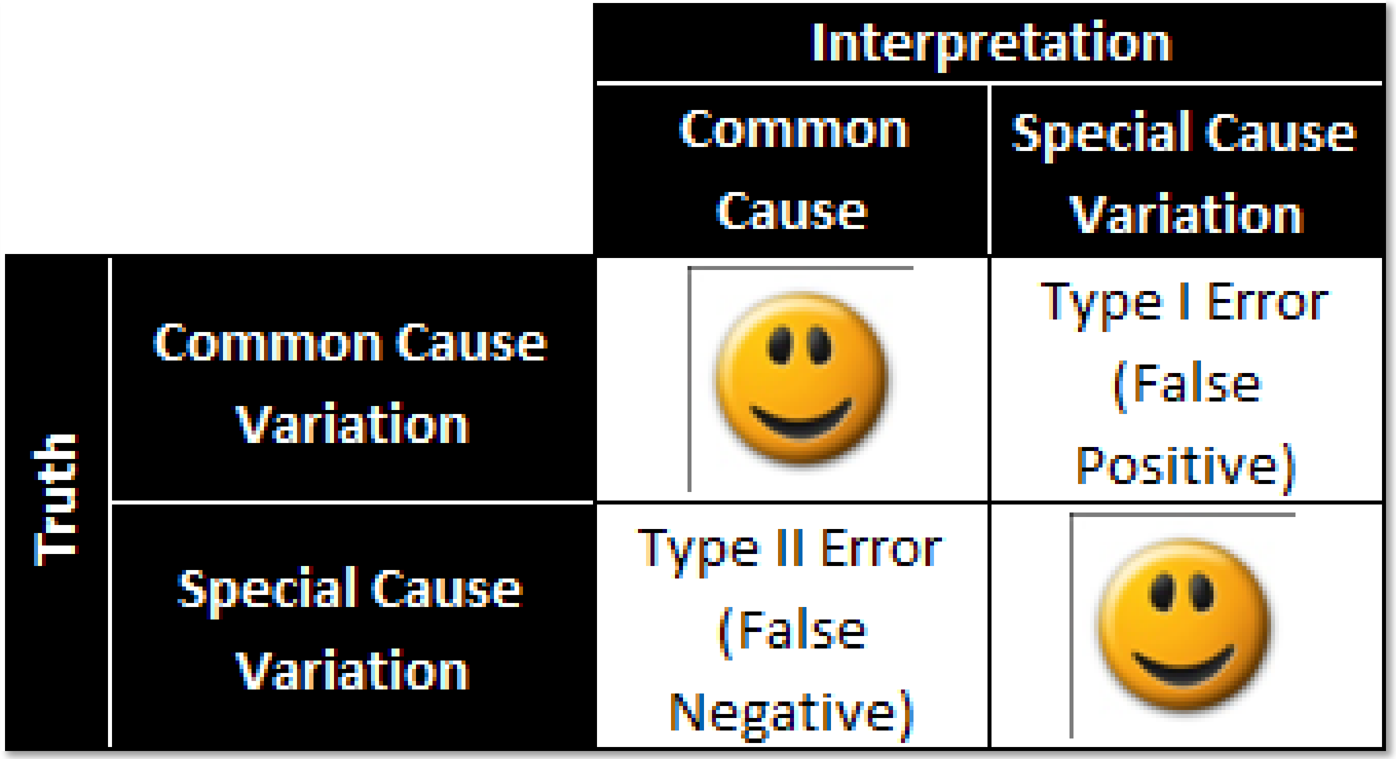
Horizontal axis: Time

Control Charts Elements

- Control charts can work for both continuous data and discrete or count data.
- Control limits are approximately three sigma away from the process mean.
- A process is in statistical control when all the data points on the control charts fall within the control limits and have random patterns only.
- Otherwise, the process is out of control and we need to investigate the special cause variation in the process.

Possible Errors in SPC

- There are two types of possible errors in interpreting controls charts.



		Interpretation	
		Common Cause	Special Cause Variation
Truth	Common Cause Variation	☺	Type I Error (False Positive)
	Special Cause Variation	Type II Error (False Negative)	☺

Possible Errors in SPC

- It is similar to the way of defining the type I and type II errors in hypothesis testing.

- Control charts can be interpreted as a way of testing the hypothesis about the process stability.
  - Null Hypothesis (H_0): The process is stable (i.e., in statistical control).
  - Alternative Hypothesis (H_A): The process is unstable (i.e., out of statistical control).

Possible Errors in SPC

- Type I Error
  - False positive
  - False alarm
  - Considering true common cause variation as special cause variation
  - Type I errors waste resources spent on investigation.

- Type II Error
  - False negative
  - Miss
  - Considering true special cause variation as common cause variation
  - Type II errors neglect the need to investigate critical changes in the process.

Data Collection Considerations

- To collect data for plotting control charts, we need to consider:
  - What is the measurement of interest?
  - Are the data discrete or continuous?
  - How many samples do we need?
  - How often do we sample?
  - Where do we sample?
  - What is the sampling strategy?
  - Do we use the raw data collected or transfer them to percentages, proportions, rates, etc.?

Subgroups and Rational Subgrouping

- When sampling, we randomly select a group of items (i.e., a subgroup) from the population of interest.
- The *subgroup size* is the count of samples in a subgroup. It can be constant or variable.
- Depending on the subgroup sizes, we select different control charts accordingly.
- **Rational subgrouping** is the basic sampling scheme in SPC.
- The goal of rational subgrouping is to maximize the likelihood of detecting special cause variation. In other words, the control limits should only reflect the variation between subgroups.
- The number of subgroups, subgroup size, and frequency of sampling have great impact on the quality of control charts.

Impact of Variation

- The rational subgrouping strategy is designed to minimize the opportunity of having special cause variation *within* subgroups.

- If there is only random variation (background noise) within subgroups, all the special cause variation would be reflected between subgroups. It is easier to detect an out-of-control situation.

- Random variation is inherent and indelible in the process. We are more interested in identifying and taking actions on special cause variation.

Frequency of Sampling

- The frequency of sampling in SPC depends on whether we have sufficient data to signal the changes in a process with reasonable time and costs.
- The more frequently we sample, the higher costs it may trigger.
- We need the subject matter experts' knowledge on the nature and characteristics of the process to make good decisions on sampling frequency.

5.2.2 I-MR Chart

I-MR Chart

- The **I-MR chart** (also called individual-moving range chart or I-MR chart) is a popular control chart for continuous data with subgroup size equal to one.
- The I chart plots an individual observation as a data point.
- The MR chart plots the absolute value of the difference between two consecutive observations in individual charts as a data point.
- If there are *n* data points in the I chart, there are *n –1* data points in and MR chart with a moving range of 2.
- The I chart is valid only if the MR chart is in control.
- The underlying distribution of the I-MR chart is normal distribution.

I Chart Equations

- I Chart (Individuals Chart)
  - Data Point: x_i
  - Center Line: $\dfrac{\sum_{i=1}^{n} x_i}{n}$
  - Control Limits: $\dfrac{\sum_{i=1}^{n} x_i}{n} \pm 2.66 \times \overline{MR}$

where *n* is the number of observations

MR-Chart Equations

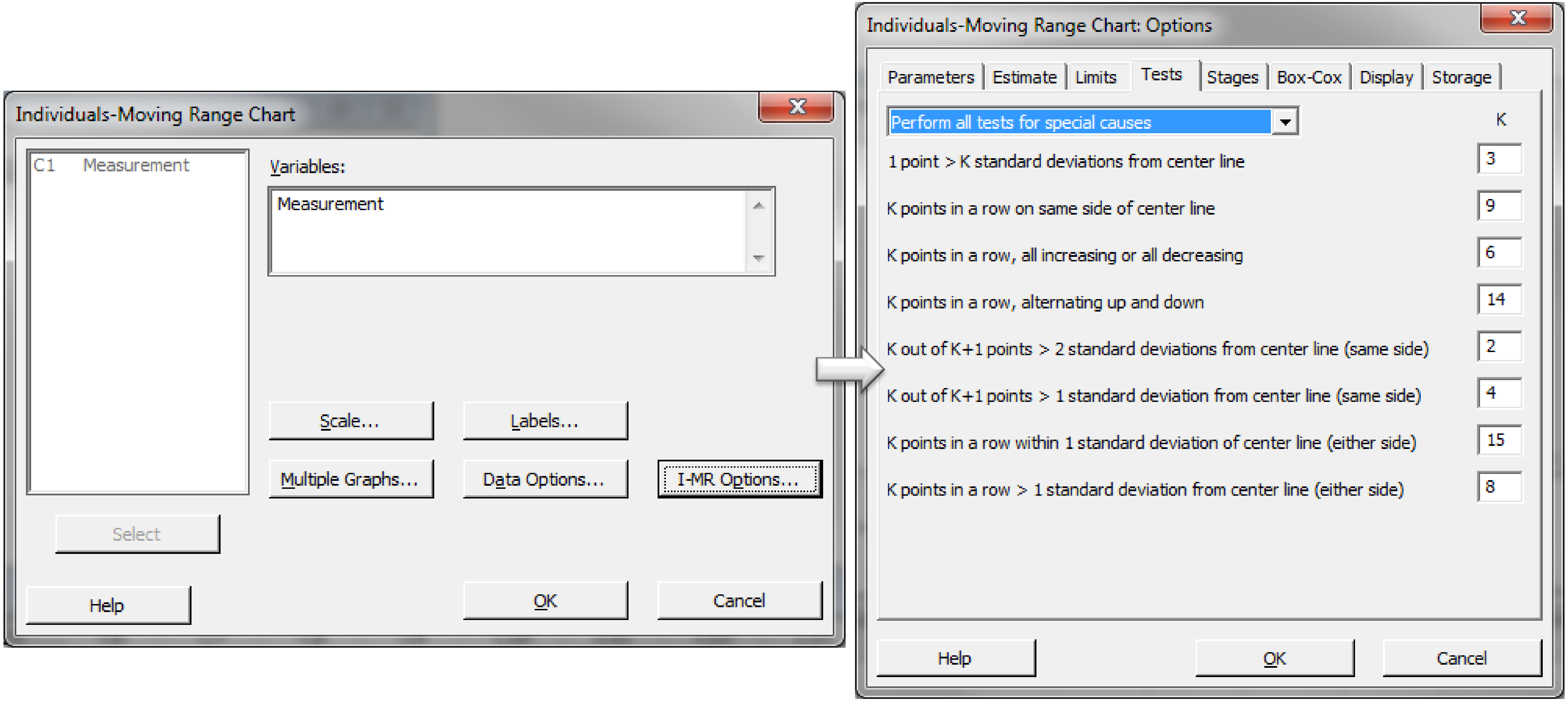
- MR Chart (Moving Range Chart)
  - Data Point: $\left|x_{i+1} - x_i\right|$
  - Center Line: $\dfrac{\Sigma \mid x_{i+1} - x_i \mid}{n-1}$
  - Upper Control Limit: $3.267 \times \dfrac{\Sigma \mid x_{i+1} - x_i \mid}{n-1}$
  - Lower Control Limit: 0

where *n* is the number of observations

Use Minitab to Plot I-MR Charts

- Data File: “IR” tab in “Sample Data.xlsx”
- Steps to plot I-MR charts in Minitab
    1) Click Stat → Control Charts → Variable Charts for Individuals → I-MR.
    2) A new window named “Individuals – Moving Range Chart” appears.
    3) Select “Measurement” as the “Variables.”
    4) Click “I-MR Options” button and a new window “Individual – Moving Range Chart – Options” appears.
    5) Click on the tab “Tests.”
    6) Select the item “Perform all tests for special causes” in the dropdown box.
    7) Click “OK” in the window “Individuals – Moving Range Chart Options.”
    8) Click “OK.”
    9) The I-MR charts appear in the newly-generated window.

Use Minitab to Plot I-MR Charts



Individuals-Moving Range Chart

C1 Measurement

Variables: Measurement

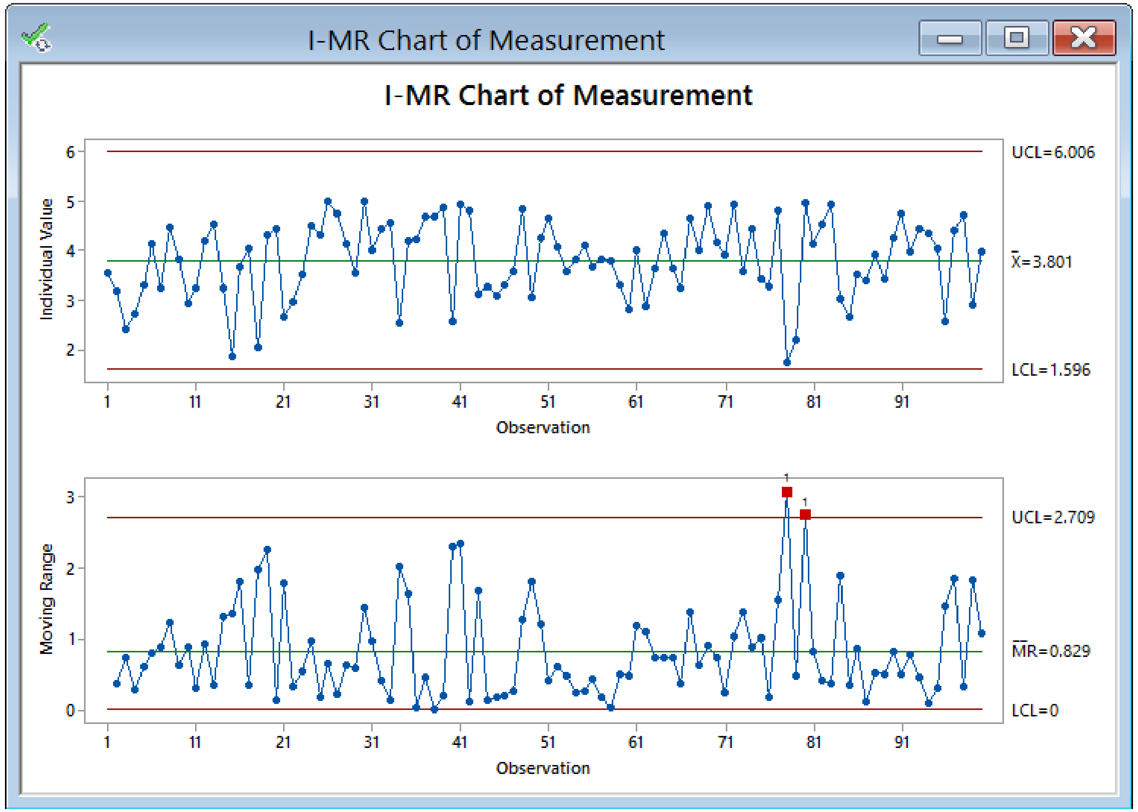
Individuals-Moving Range Chart: Options

Perform all tests for special causes

	K
1 point > K standard deviations from center line	3
K points in a row on same side of center line	9
K points in a row, all increasing or all decreasing	6
K points in a row, alternating up and down	14
K out of K+1 points > 2 standard deviations from center line (same side)	2
K out of K+1 points > 1 standard deviation from center line (same side)	4
K points in a row within 1 standard deviation of center line (either side)	15
K points in a row > 1 standard deviation from center line (either side)	8

I-MR Charts Diagnosis

- **I Chart (Individuals' Chart):**
  - Since the MR chart is out of control, the I chart is invalid.
- **MR Chart (Moving Range Chart):**
  - Two data points fall beyond the upper control limit. This indicates the MR chart is out of control (i.e., the variations between every two contiguous individual samples are not stable over time).
  - We need to further investigate the process, identify the root causes that trigger the outliers, and correct them to bring the process back in control.



5.2.3 Xbar-R Chart

Xbar-R Chart

- The **Xbar-R** chart is a control chart for continuous data with a constant subgroup size between two and ten.
- The Xbar chart plots the average of a subgroup as a data point.
- The R chart plots the difference between the highest and lowest values within a subgroup as a data point.
- The Xbar chart monitors the process mean and the R chart monitors the variation within subgroups.
- The Xbar is valid only if the R chart is in control.
- The underlying distribution of the Xbar-R chart is normal distribution.

Xbar Chart Equations

- Xbar chart
  - Data Point: $\overline{X}_i = \frac{\sum_{j=1}^{m} x_{ij}}{m}$
  - Center Line: $\overline{\overline{X}} = \frac{\sum_{i=1}^{k} \overline{X}_i}{k}$
  - Control Limits: $\overline{\overline{X}} \pm A_2 \overline{R}$

where *m* is the subgroup size and *k* is the number of subgroups. A_2 is a constant depending on the subgroup size.

R Chart Equations

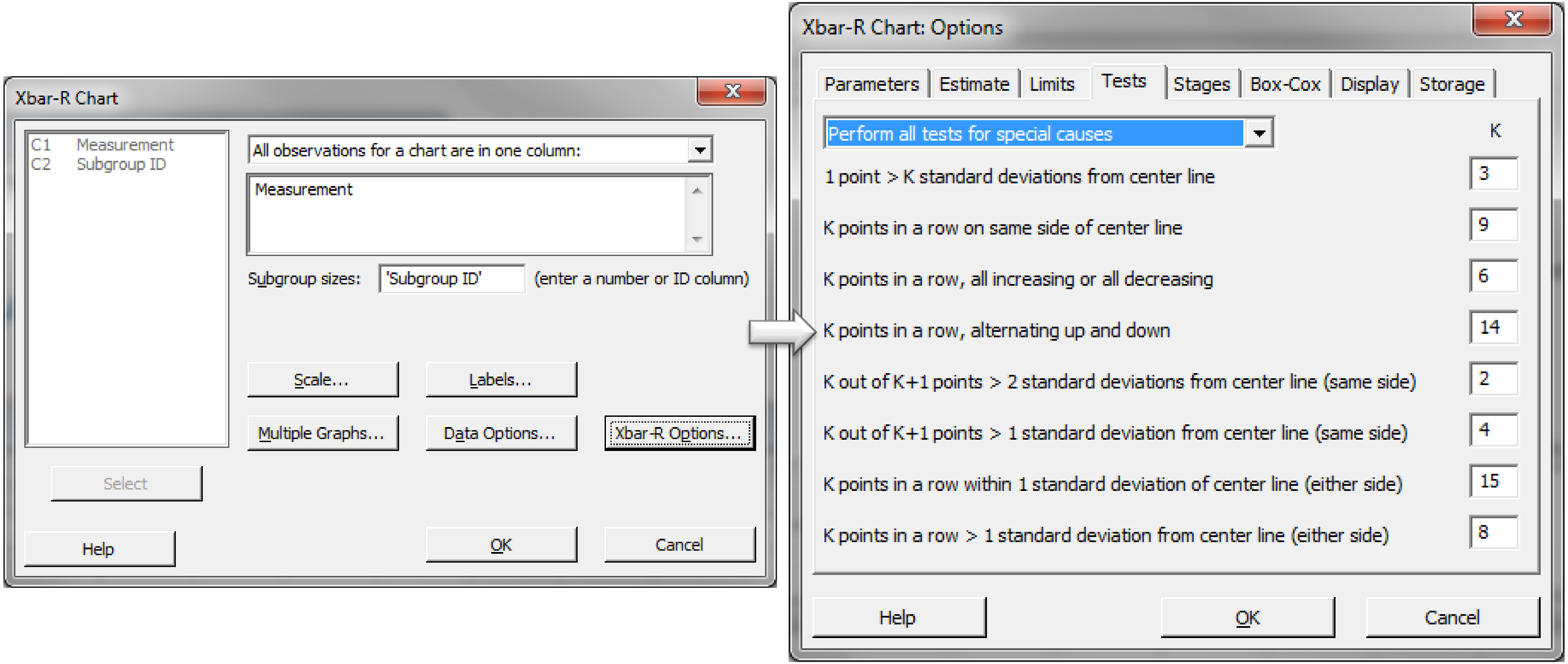
- R chart (Rage Chart)
  - Data Point: $R_i = \underset{j \in [1,m]}{Max}(x_{ij}) - \underset{j \in [1,m]}{Min}(x_{ij})$
  - Center Line: $\overline{R} = \dfrac{\sum_{i=1}^{k} R_i}{k}$
  - Upper Control Limit: $D_4 \times \overline{R}$
  - Lower Control Limit: $D_3 \times \overline{R}$

where *m* is the subgroup size and *k* is the number of subgroups. D_3 and D_4 are constants depending on the subgroup size.

Use Minitab to Plot Xbar-R Charts

- Data File: “Xbar-R” tab in “Sample Data.xlsx”
- Steps to plot Xbar-R charts in Minitab
  - Click Stat → Control Charts → Variable Charts for Subgroups → Xbar-R.
  - A new window named “Xbar-R Chart” appears.
  - Click in the blank box right below “All observations for a chart are in one column” and the variables appear in the list box on the left.
  - Select the “Measurement” into the box below “All observations for a chart are in one column.”
  - Select the “Subgroup ID” as the “Subgroup size (enter a number or ID column).
  - Click “Xbar-R Options” button and a new window “Xbar-R Chart – Options” appears.
  - Click on the tab “Tests.”
  - Select the item “Perform all tests for special causes” in the dropdown menu.
  - Click “OK” in the window “Xbar-R Chart – Options.”
  - Click “OK.”
  - The Xbar-R charts appear in the newly-generated window.

Use Minitab to Plot Xbar-R Charts



All observations for a chart are in one column: Measurement

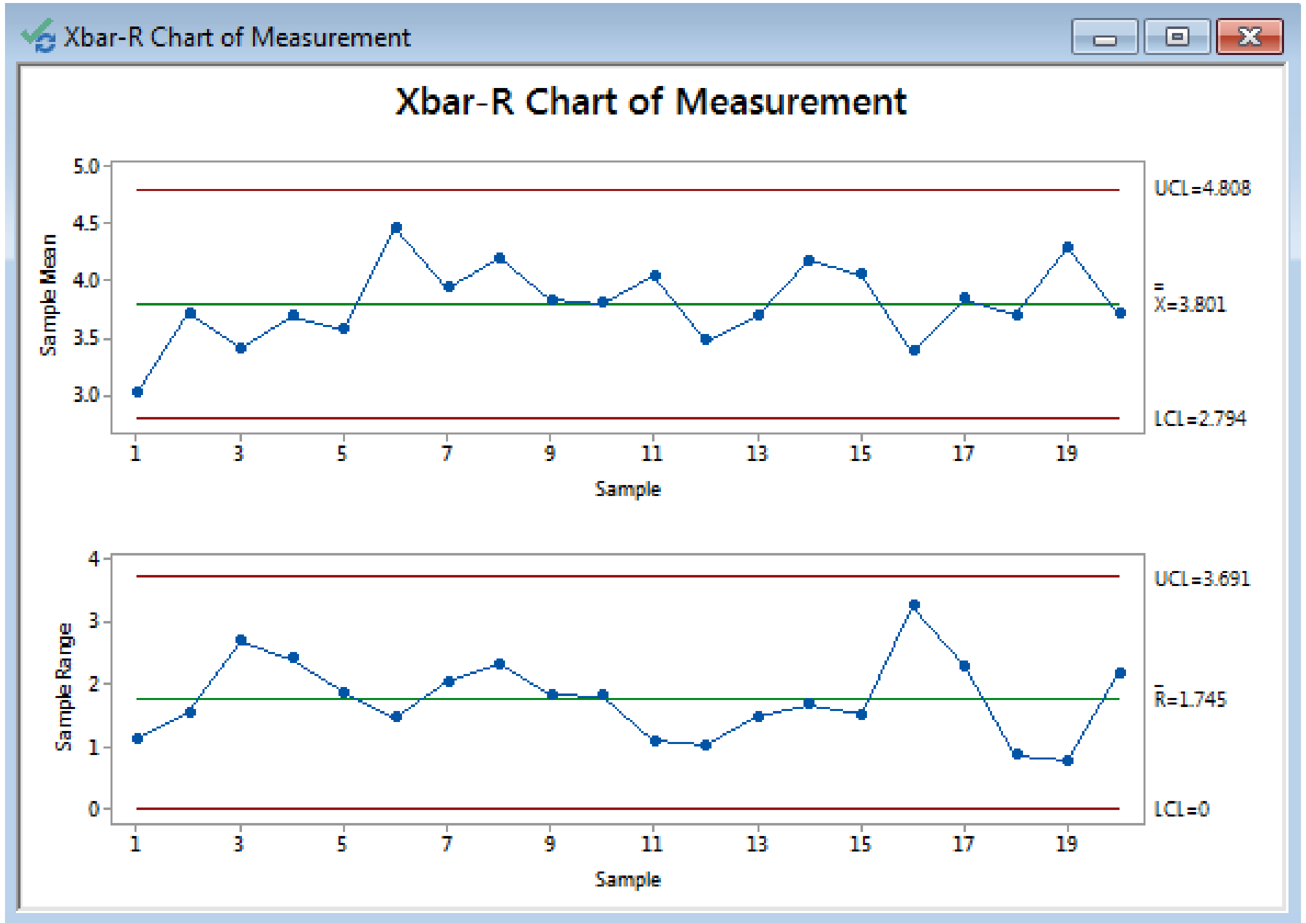
Subgroup sizes: 'Subgroup ID' (enter a number or ID column)

Perform all tests for special causes

	K
1 point > K standard deviations from center line	3
K points in a row on same side of center line	9
K points in a row, all increasing or all decreasing	6
K points in a row, alternating up and down	14
K out of K+1 points > 2 standard deviations from center line (same side)	2
K out of K+1 points > 1 standard deviation from center line (same side)	4
K points in a row within 1 standard deviation of center line (either side)	15
K points in a row > 1 standard deviation from center line (either side)	8

Xbar-R Charts Diagnosis

- **Xbar-R Chart Interpretation:**
  - Since the R chart is in control, the Xbar chart is valid.
  - In both charts, there are not any data points failing any tests for special causes (i.e., all the data points fall between the control limits and spread around the center line with a random pattern).
  - We conclude that the process is in control.



5.2.4 U Chart

Defect vs. Defective

- A **defect** of a unit is the unit’s characteristic that does not meet the customers’ requirements.
- A **defective** is a unit that is not acceptable to the customers.
- One defective might have multiple defects.
- One unit might have multiple defects but be still usable to the customers.

U Chart

- The **U chart** is a control chart monitoring the average defects per unit.
- The U chart plots the count of defects per unit of a subgroup as a data point.
- It considers the situation when the subgroup size of inspected units for which the defects would be counted is not constant.
- The underlying distribution of the U chart is Poisson distribution.

U Chart Equations

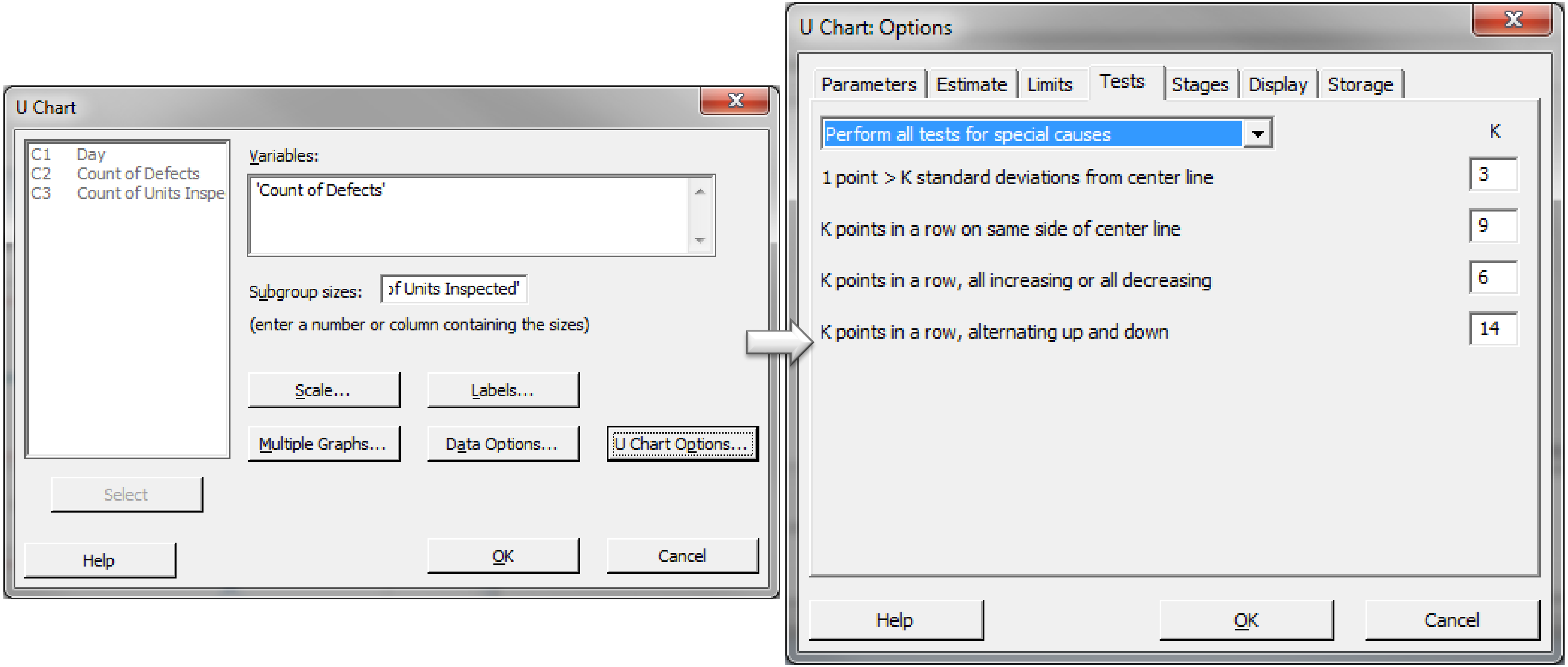
- U chart
  - Data Point: $u_i = \frac{x_i}{n_i}$
  - Center Line: $\bar{u} = \frac{\sum_{i=1}^{k} u_i}{k}$
  - Control Limits: $\bar{u} \pm 3 \times \sqrt{\frac{\bar{u}}{n_i}}$

where n_i is the subgroup size for the ith subgroup;
k is the number of subgroups;
x_i is the number of defects in the ith subgroup.

Use Minitab to Plot a U Chart

- Data File: “U” tab in “Sample Data.xlsx”
- Steps to plot a U chart in Minitab
    1) Click Stat → Control Charts → Attributes Charts → U.
    2) A new window named “U Chart” appears.
    3) Select “Count of Defects” as the “Variables.”
    4) Select “Count of Units Inspected” as the “Subgroup Sizes.”
    5) Click the button “U Chart Options” and another window named “U Chart Options” pops up.
    6) Click the tab “Tests.”
    7) Select the item “Perform all tests for special causes” in the dropdown menu.
    8) Click “OK” in the window “U Chart Options.”
    9) Click “OK” in the window “U Chart.”
    10) The U chart appears in the newly-generated window.

Use Minitab to Plot a U Chart



U Chart

C1 Day
C2 Count of Defects
C3 Count of Units Inspe

Variables: 'Count of Defects'

Subgroup sizes: ›f Units Inspected'

(enter a number or column containing the sizes)

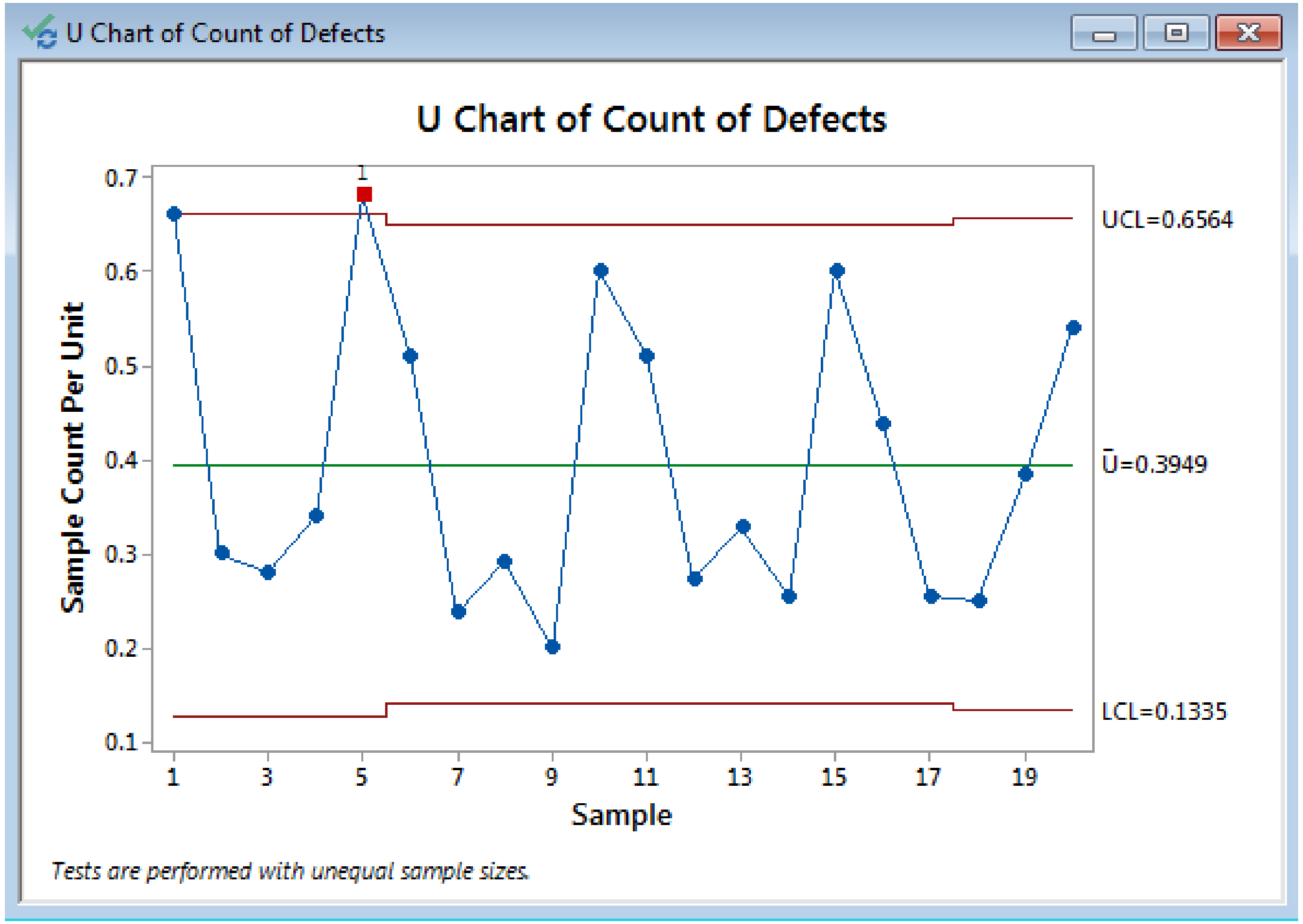
U Chart: Options

Perform all tests for special causes

	K
1 point > K standard deviations from center line	3
K points in a row on same side of center line	9
K points in a row, all increasing or all decreasing	6
K points in a row, alternating up and down	14

Use Minitab to Plot a U Chart

- **U Chart Interpretation:**
  - Since the sample sizes are not constant over time, the control limits are adjusted to different values accordingly.
  - The data point circled in red falls beyond the upper control limit. We conclude that the process is out of control.
  - Further investigation is needed to determine the special causes that triggered the unnatural pattern of the process.



Tests are performed with unequal sample sizes.

5.2.5 P Chart

P Chart

- The **P chart** is a control chart monitoring the percentages of defectives.
- The P chart plots the percentage of defectives in one subgroup as a data point.
- It considers the situation when the subgroup size of inspected units is not constant.
- The underlying distribution of the P chart is binomial distribution.

P Chart Equations

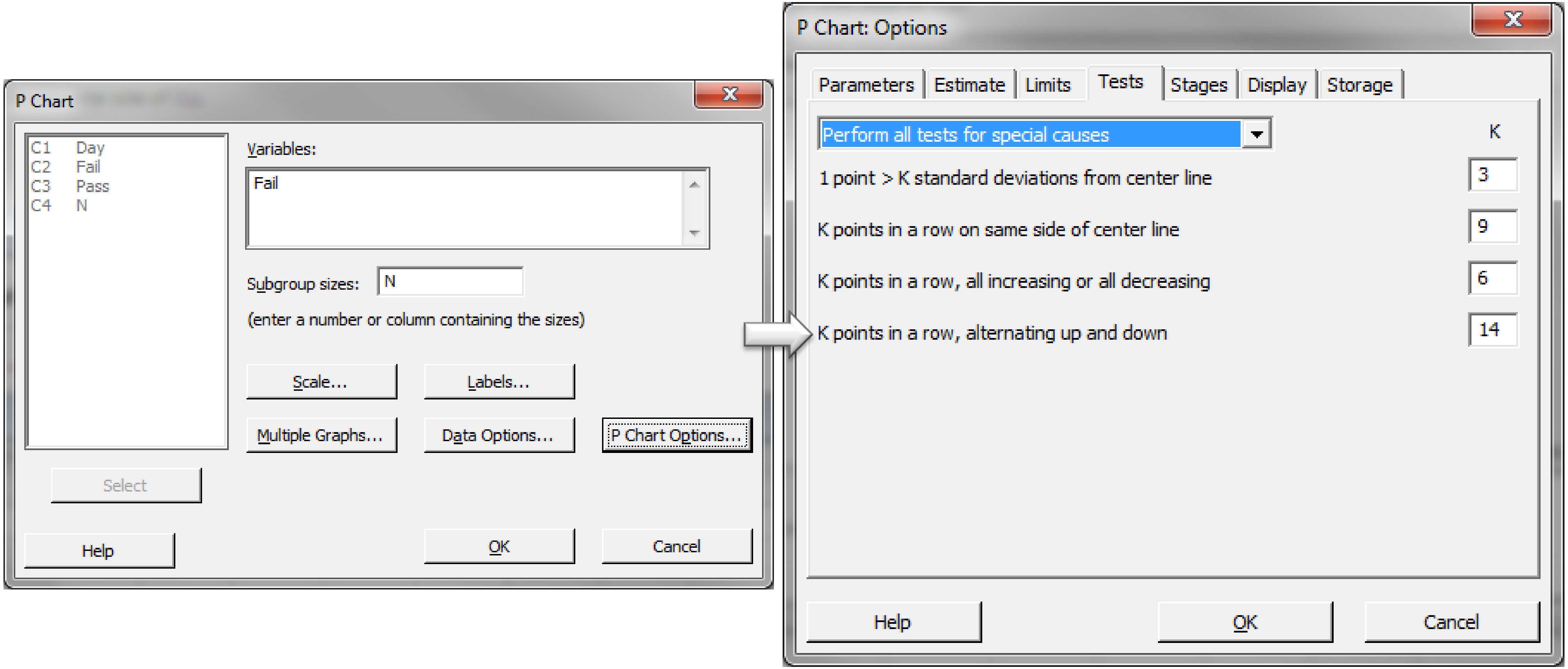
- P chart
  - Data Point: $p_i = \frac{x_i}{n_i}$
  - Center Line: $\bar{p} = \frac{\sum_{i=1}^{k} x_i}{\sum_{i=1}^{k} n_i}$
  - Control Limits: $\bar{p} \pm 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n_i}}$

where n_i is the subgroup size for the ith subgroup;
k is the number of subgroups;
x_i is the number of defectives in the ith subgroup.

Use Minitab to Plot a P Chart

- Data File: “P” tab in “Sample Data.xlsx”
- Steps to plot a P chart in Minitab
    1) Click Stat → Control Charts → Attributes Charts → P.
    2) A new window named “P Chart” appears.
    3) Select “Fail” as the “Variables.”
    4) Select “N” as the “Subgroup Sizes.”
    5) Click the button “P Chart Options” and another window named “P Chart Options” pops up.
    6) Click the tab “Tests.”
    7) Select the item “Perform all tests for special causes” in the dropdown menu.
    8) Click “OK” in the window “P Chart Options.”
    9) Click “OK.”
    10) The P chart appears in the newly-generated window.

Use Minitab to Plot a P Chart



P Chart

C1 Day
C2 Fail
C3 Pass
C4 N

Variables: Fail

Subgroup sizes: N

(enter a number or column containing the sizes)

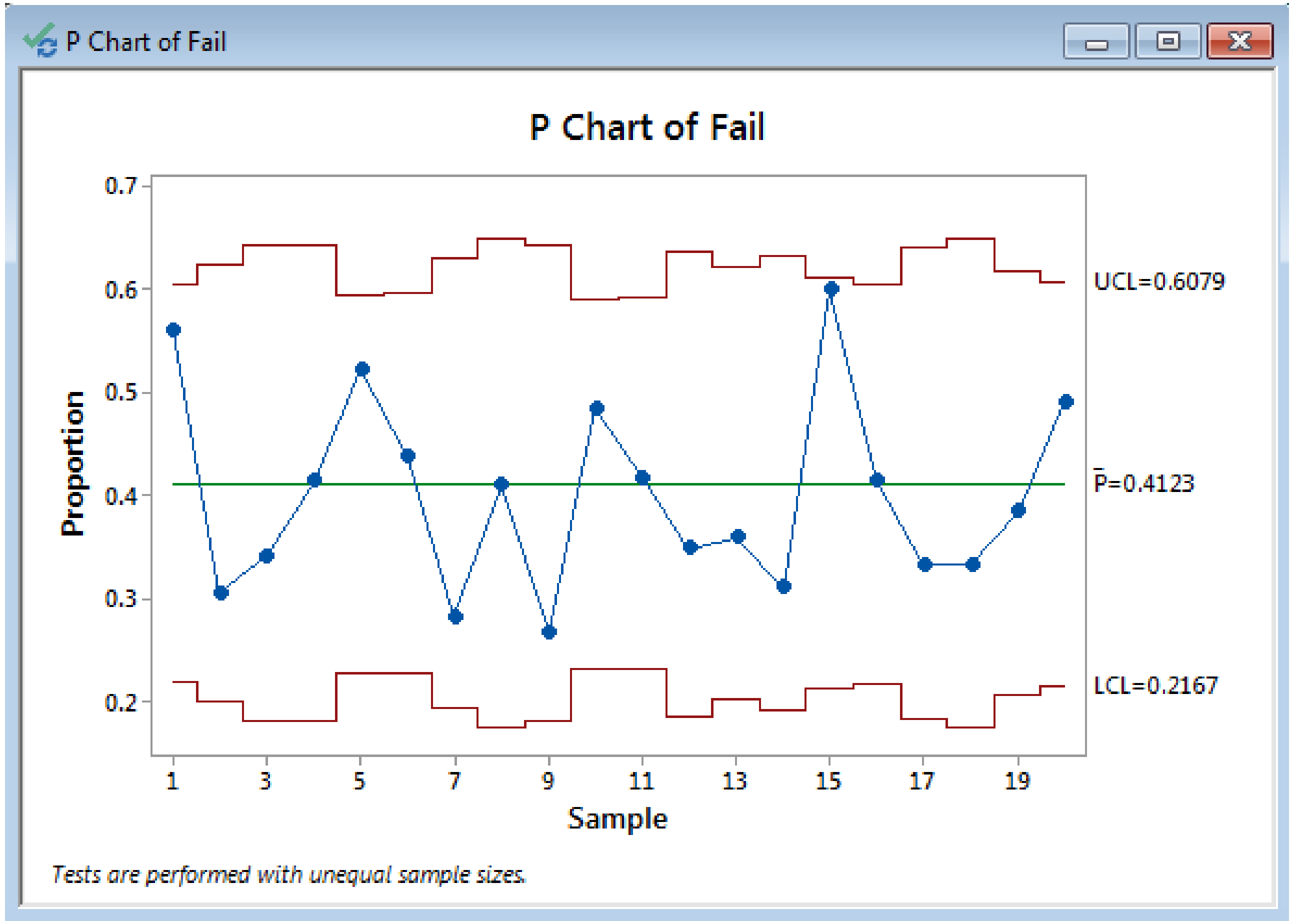
P Chart: Options

Perform all tests for special causes

	K
1 point > K standard deviations from center line	3
K points in a row on same side of center line	9
K points in a row, all increasing or all decreasing	6
K points in a row, alternating up and down	14

P Chart Diagnosis

- **P Chart Interpretation:**
  - Since the sample sizes are not constant over time, the control limits are adjusted to different values accordingly.
  - All the data points fall within the control limits and spread randomly around the mean. We conclude that the process is in control.



Tests are performed with unequal sample sizes.

5.2.6 NP Chart

NP Chart

- The **NP chart** is a control chart monitoring the count of defectives.
- The NP chart plots the number of defectives in one subgroup as a data point.
- The subgroup size of the NP chart is constant.
- The underlying distribution of the NP chart is binomial distribution.

NP Chart Equations

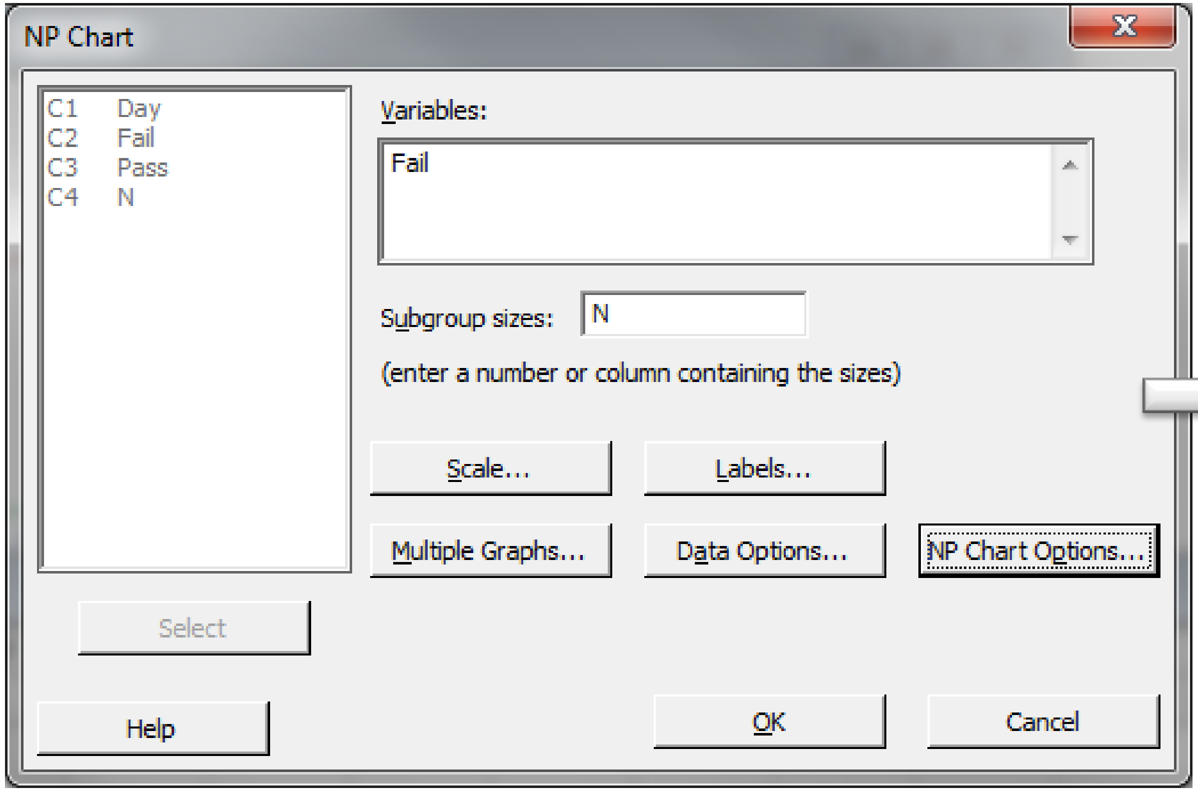
- NP chart
  - Data Point: x_i
  - Center Line: $n\overline{p} = \frac{\sum_{i=1}^{k} x_i}{k}$
  - Control Limits: $n\overline{p} \pm 3 \times \sqrt{n\overline{p}(1-\overline{p})}$

where *n* is the constant subgroup size;
k is the number of subgroups;
x_i is the number of defectives in the ith subgroup.

Use Minitab to Plot an NP Chart

- Data File: “NP” tab in “Sample Data.xlsx”
- Steps to plot a NP chart in Minitab
    1. Click Stat → Control Charts → Attributes Charts → NP.
    2. A new window named “NP Chart” appears.
    3. Select “Fail” as the “Variables.”
    4. Enter “50” as the “Subgroup Sizes.”
    5. Click the button “NP Chart Options” and another window named “NP Chart Options” pops up.
    6. Click the tab “Tests.”
    7. Select the item “Perform all tests for special causes” in the dropdown menu.
    8. Click “OK” in the window “NP Chart Options.”
    9. Click “OK.”
    10. The NP chart appears in the newly-generated window.

Use Minitab to Plot an NP Chart

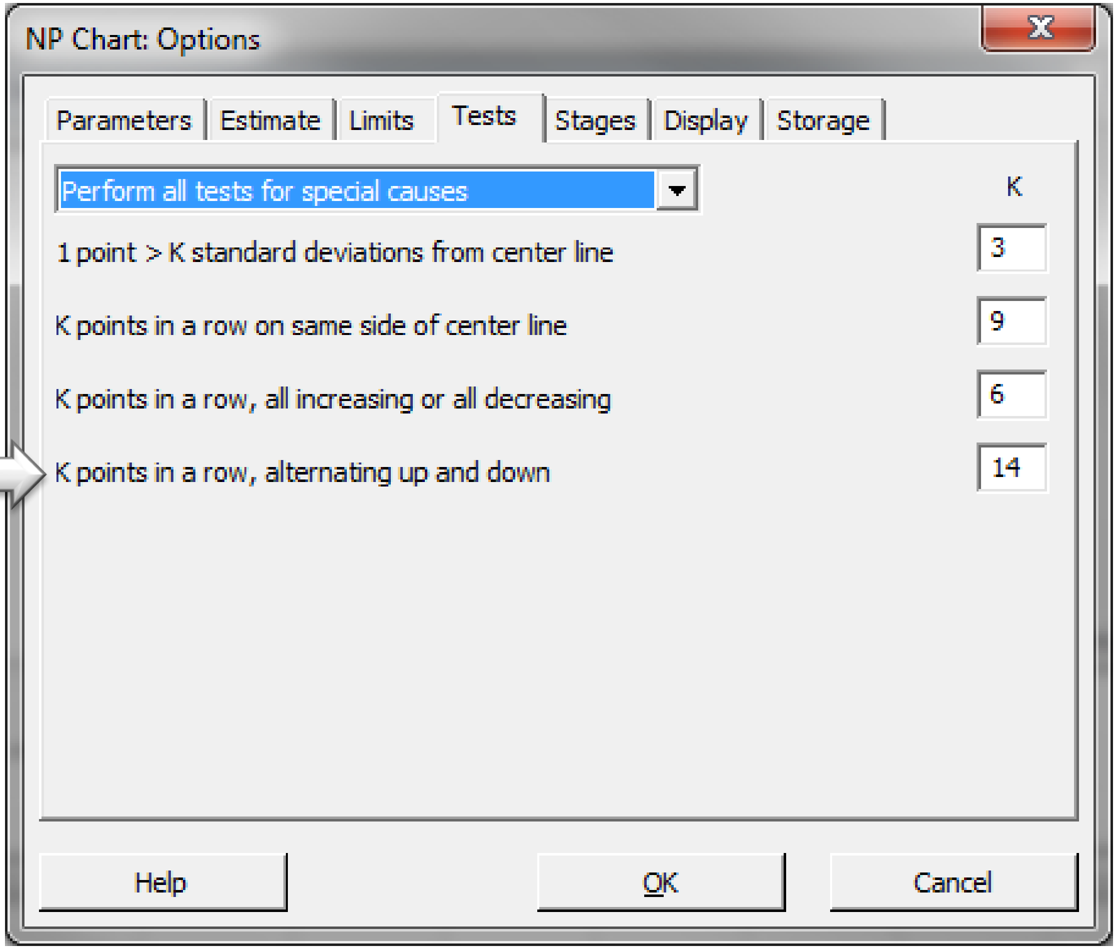


C1 Day
C2 Fail
C3 Pass
C4 N

Variables: Fail

Subgroup sizes: N

(enter a number or column containing the sizes)

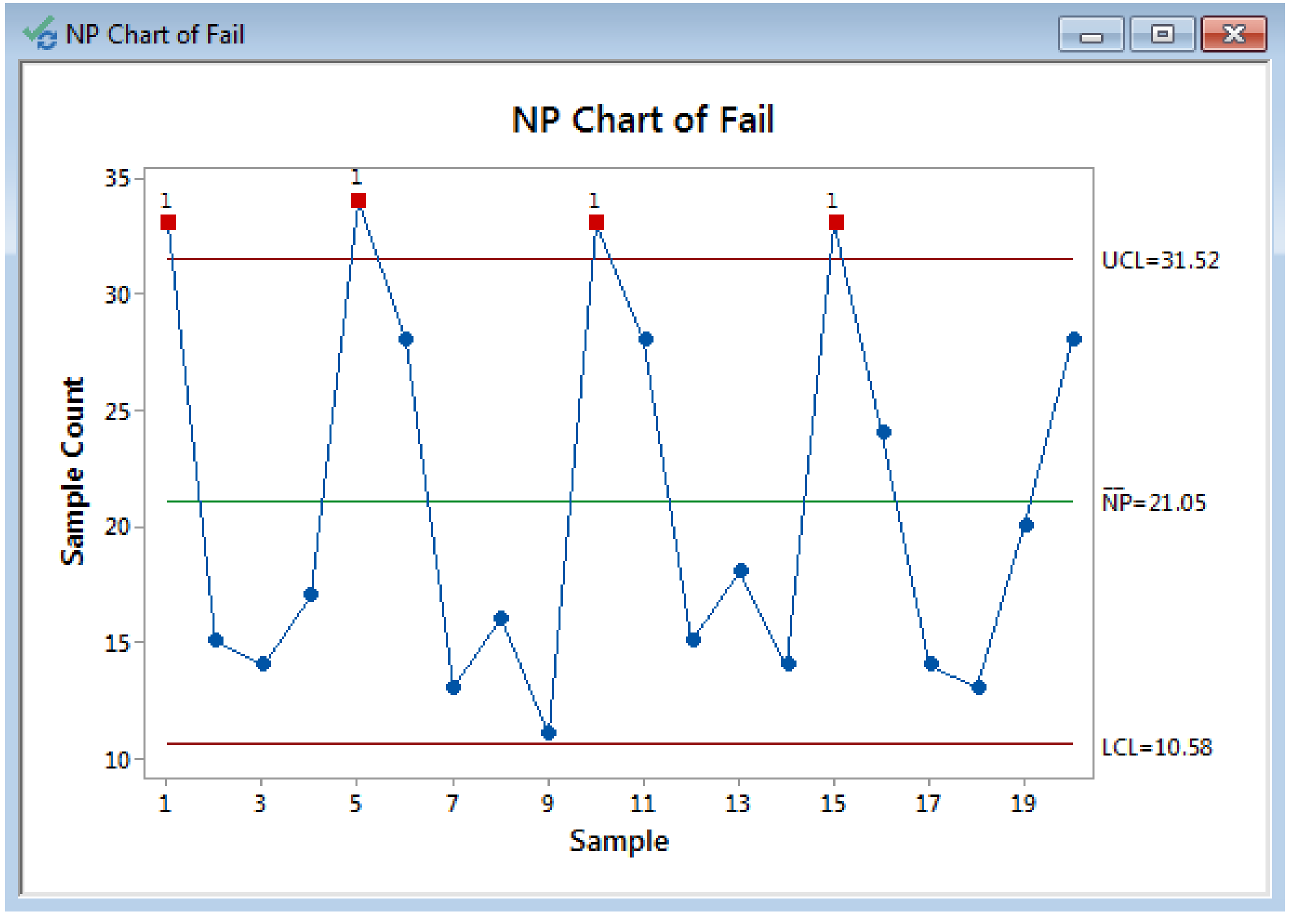


Perform all tests for special causes

	K
1 point > K standard deviations from center line	3
K points in a row on same side of center line	9
K points in a row, all increasing or all decreasing	6
K points in a row, alternating up and down	14

NP Chart Diagnosis

- **NP Chart Interpretation:**
  - Four data points, circled in red, fall beyond the upper control limit.
  - We conclude that the NP chart is out of control.
  - Further investigation is needed to determine the special causes that triggered the unnatural pattern of the process.



5.2.7 X-S Chart

X-S Chart

- The **X-S chart** (also called Xbar-S chart) is a control chart for continuous data with a constant subgroup size greater than ten.
- The Xbar chart plots the average of a subgroup as a data point.
- The S chart plots the standard deviation within a subgroup as a data point.
- The Xbar chart monitors the process mean and the S chart monitors the variability within subgroups.
- The Xbar is valid only if the S chart is in control.
- The underlying distribution of the Xbar-S chart is normal distribution.

Xbar Chart Equations

- Xbar Chart
  - Data Point: $\overline{X}_i = \dfrac{\sum_{j=1}^{m} x_{ij}}{m}$
  - Center Line: $\overline{\overline{X}} = \dfrac{\sum_{i=1}^{k} \overline{X}_i}{k}$
  - Control Limits: $\overline{\overline{X}} \pm A_3 \overline{s}$

where *m* is the subgroup size and *k* is the number of subgroups. A_3 is a constant depending on the subgroup size.

S Chart Equations

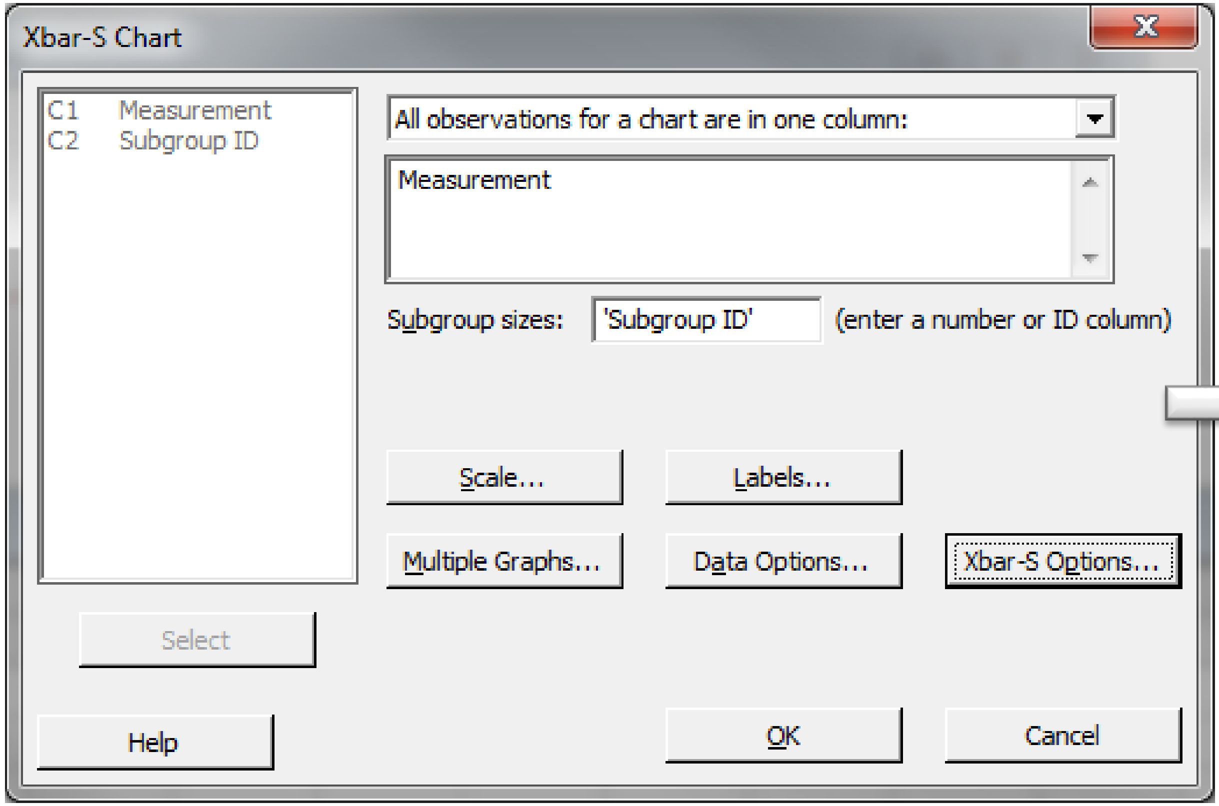
- S chart
  - Data Point: $s_i = \sqrt{\frac{\sum_{j=1}^{m} (x_{ij} - \bar{x}_i)^2}{m-1}}$
  - Center Line: $\bar{s} = \frac{\sum_{i=1}^{k} s_i}{k}$
  - Upper Control Limit: $B_4 \times \bar{s}$
  - Lower Control Limit: $B_3 \times \bar{s}$

where *m* is the subgroup size and *k* is the number of subgroups.
B_3 and B_4 are constants depending on the subgroup size.

Use Minitab to Plot Xbar-S Charts

- Data File: “Xbar-S” tab in “Sample Data.xlsx”
- Steps to plot Xbar-S charts in Minitab
    1) Click Stat → Control Charts → Variables Charts for Subgroups → Xbar-S.
    2) A new window named “Xbar-S Chart” appears.
    3) Select “Measurement” by ticking the box below “All observations for a chart are in one column.”
    4) Select “Subgroup ID” as the “Subgroup Sizes.”
    5) Click the button “Xbar-S Options” and another window named “Xbar-S Chart Options” pops up.
    6) Click the tab “Tests.”
    7) Select the item “Perform all tests for special causes” in the dropdown menu.
    8) Click “OK” in the window “Xbar-S Chart Options.”
    9) Click “OK.”
    10) The Xbar-S charts appear in the newly-generated window.

Use Minitab to Plot Xbar-S Charts



Xbar-S Chart

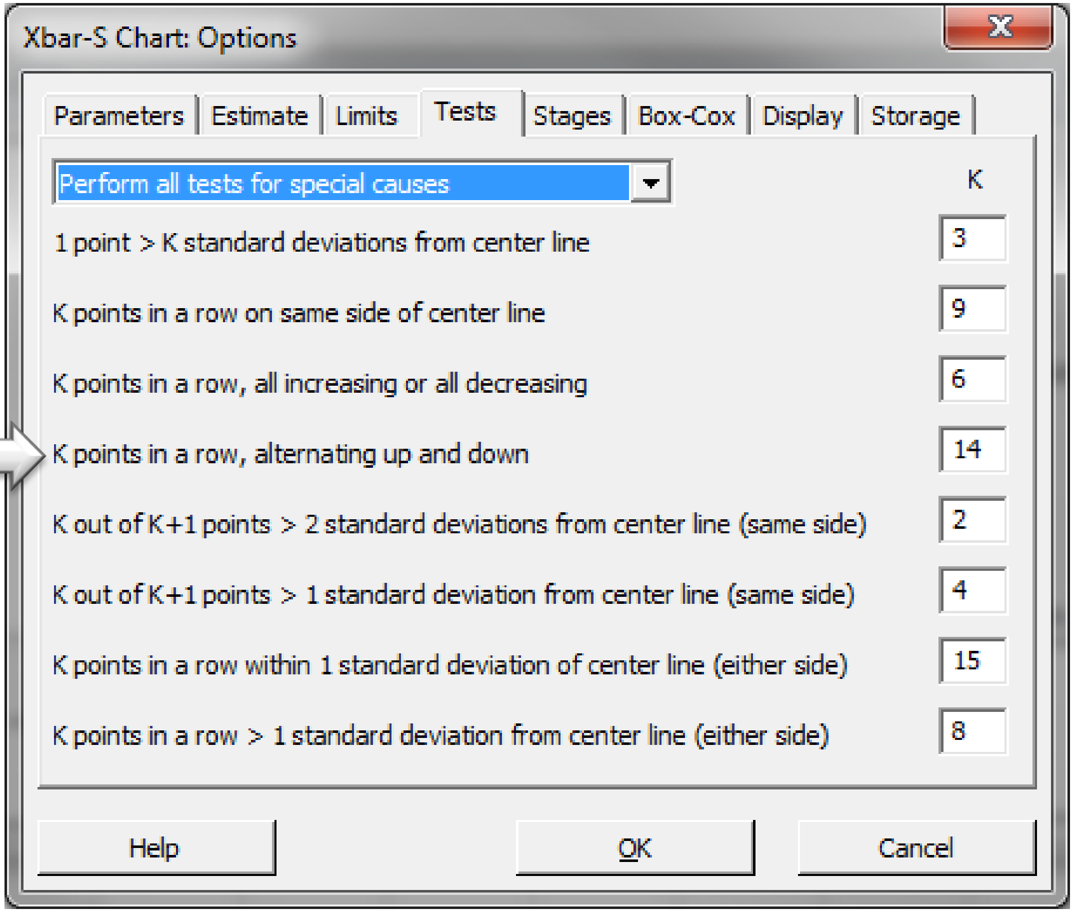
C1 Measurement

C2 Subgroup ID

All observations for a chart are in one column:

Measurement

Subgroup sizes: 'Subgroup ID' (enter a number or ID column)

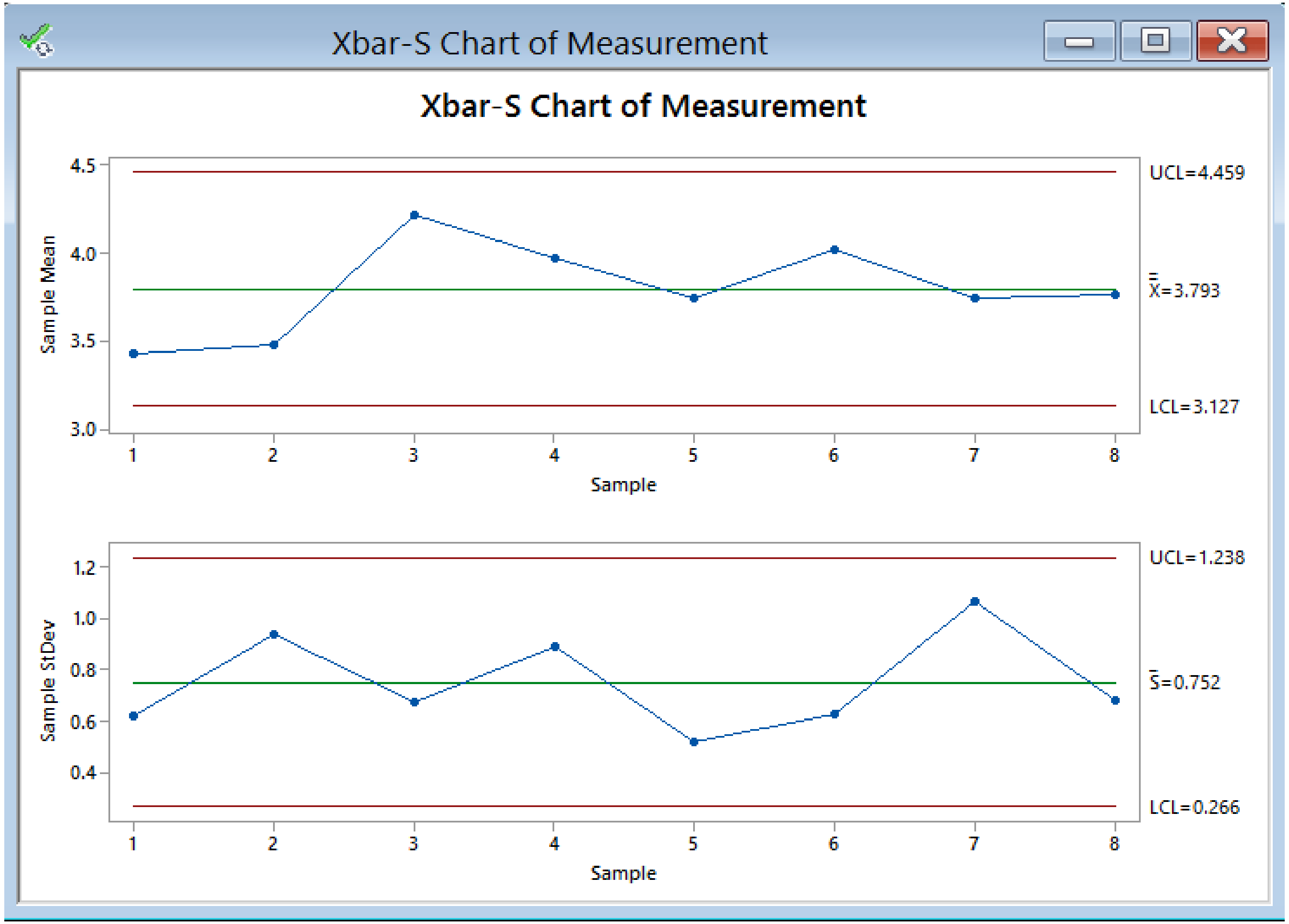


Xbar-S Chart: Options

Perform all tests for special causes

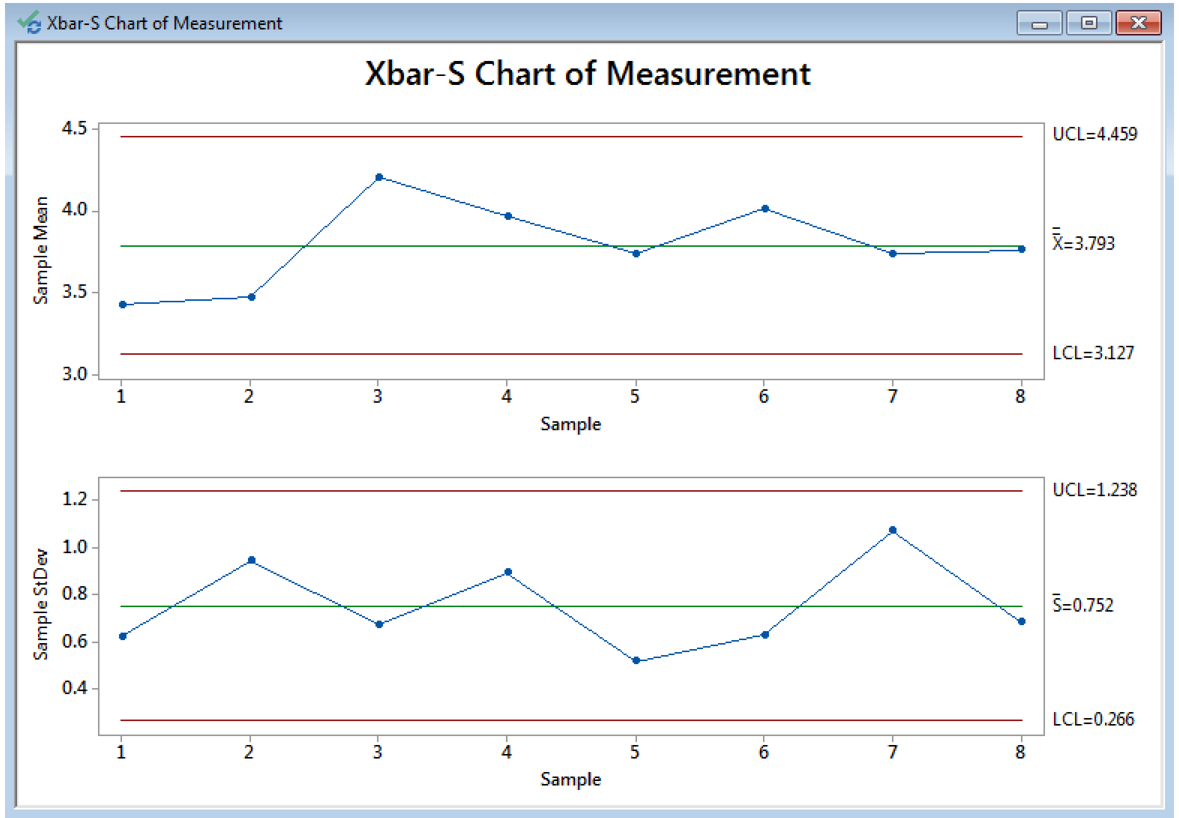
	K
1 point > K standard deviations from center line	3
K points in a row on same side of center line	9
K points in a row, all increasing or all decreasing	6
K points in a row, alternating up and down	14
K out of K+1 points > 2 standard deviations from center line (same side)	2
K out of K+1 points > 1 standard deviation from center line (same side)	4
K points in a row within 1 standard deviation of center line (either side)	15
K points in a row > 1 standard deviation from center line (either side)	8

Use Minitab to Plot Xbar-S Charts



Xbar-S Charts Diagnosis

- Xbar-S Chart Interpretation:
  - Since the S chart is in control, the Xbar chart is valid.
  - In both charts, there are not any data points failing any tests for special cause (i.e., all the data points fall between the control limits and spread around the center line with a random pattern).
  - We conclude that the process is in control.



5.2.8 CumSum Chart

CumSum Chart

- The **CumSum chart** (also called cumulative sum control chart or CUSUM chart) is a control chart of monitoring the cumulative sum of the subgroup mean deviations from the process target.

- It detects the shift of the process mean from the process target over time.

- The underlying distribution of the CumSum chart is normal distribution.

- There are two types of CumSum charts:
  - One two-sided CumSum charts
  - Two one-sided CumSum charts.

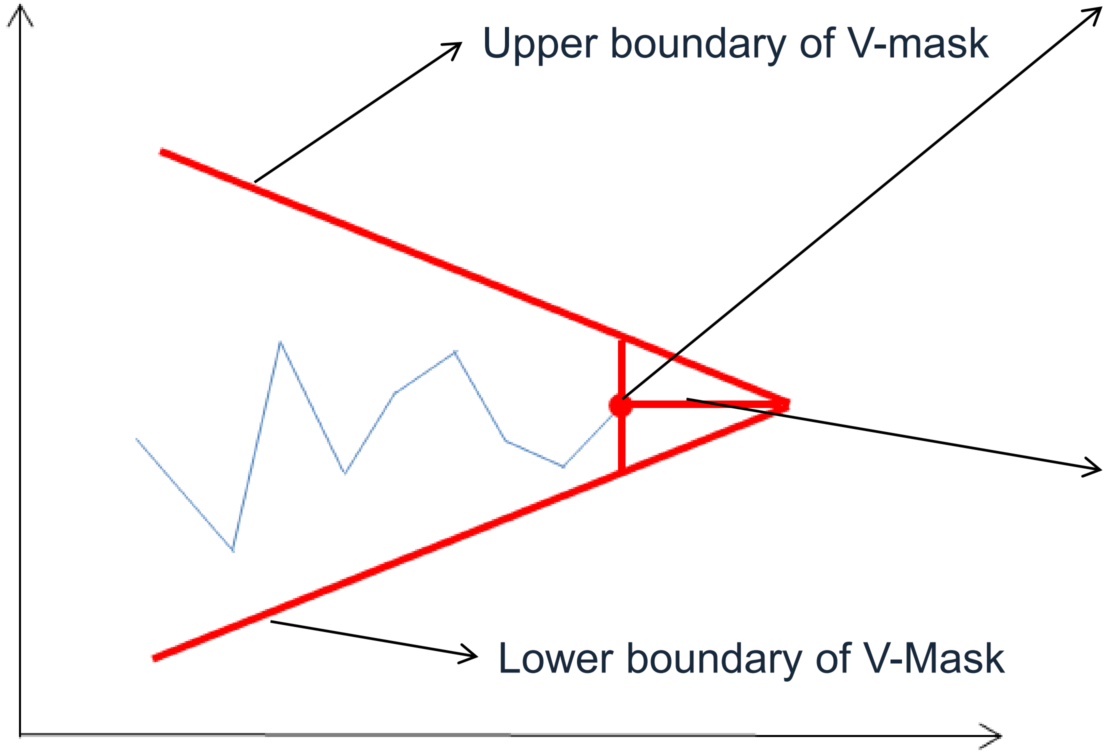
Two-Sided CumSum

- In two-sided CumSum charts, we use a V-mask to identify out-of-control data points.

- The V-mask is a transparent overlay shape of a horizontal “V” applied on the top of the CumSum chart. Its origin point is placed on the last data point of the CumSum chart and its center line is horizontal.

- If all of the data points stay inside the V-mask, we consider the process is in statistical control.

V-Mask



Upper boundary of V-mask

Origin point on the V-mask, which is applied right on the last data point of the CumSum chart.

The center line of the V-mask, which should be parallel to the x axis (time).

Lower boundary of V-Mask

If any data points in the CumSum chart to the left of the origin point are outside the V-mask, the process is considered out of statistical control.

Two-Sided CumSum Equations

- Two-Sided CumSum
  - Data Point: $\begin{cases} c_i = c_{i-1} + (\bar{x}_i - T) & i > 0 \\ c_i = 0 & i = 0 \end{cases}$
  - V-Mask Slope: $k\dfrac{\sigma}{\sqrt{m}}$
  - V-Mask Width at the Origin Point: $2h\dfrac{\sigma}{\sqrt{m}}$

where $\bar{x}_i$ is the mean of the ith subgroup;
T is the process target;
σ is the estimation of process standard deviation;
m is the subgroup size.

One-Sided CumSum

- We can also use two one-sided CumSum charts to detect the shift of the process mean from the process target.

- The upper one-sided CumSum detects the upward shifts of the process mean.

- The lower one-sided CumSum detects the downward shifts of the process mean.

One-Sided CumSum Equations

- One-Sided CumSum
  - Data Point: $\begin{cases} c_i = c_{i-1} + (\bar{x}_i - T) & i > 0 \\ c_i = 0 & i = 0 \end{cases}$
  - Center Line: T
  - Upper Control Limit: $c_i^+ = \max\left(0, \bar{x}_i - (T+k) + c_{i-1}^+\right)$
  - Lower Control Limit: $c_i^- = \max\left(0, (T-k) - \bar{x}_i + c_{i-1}^-\right)$

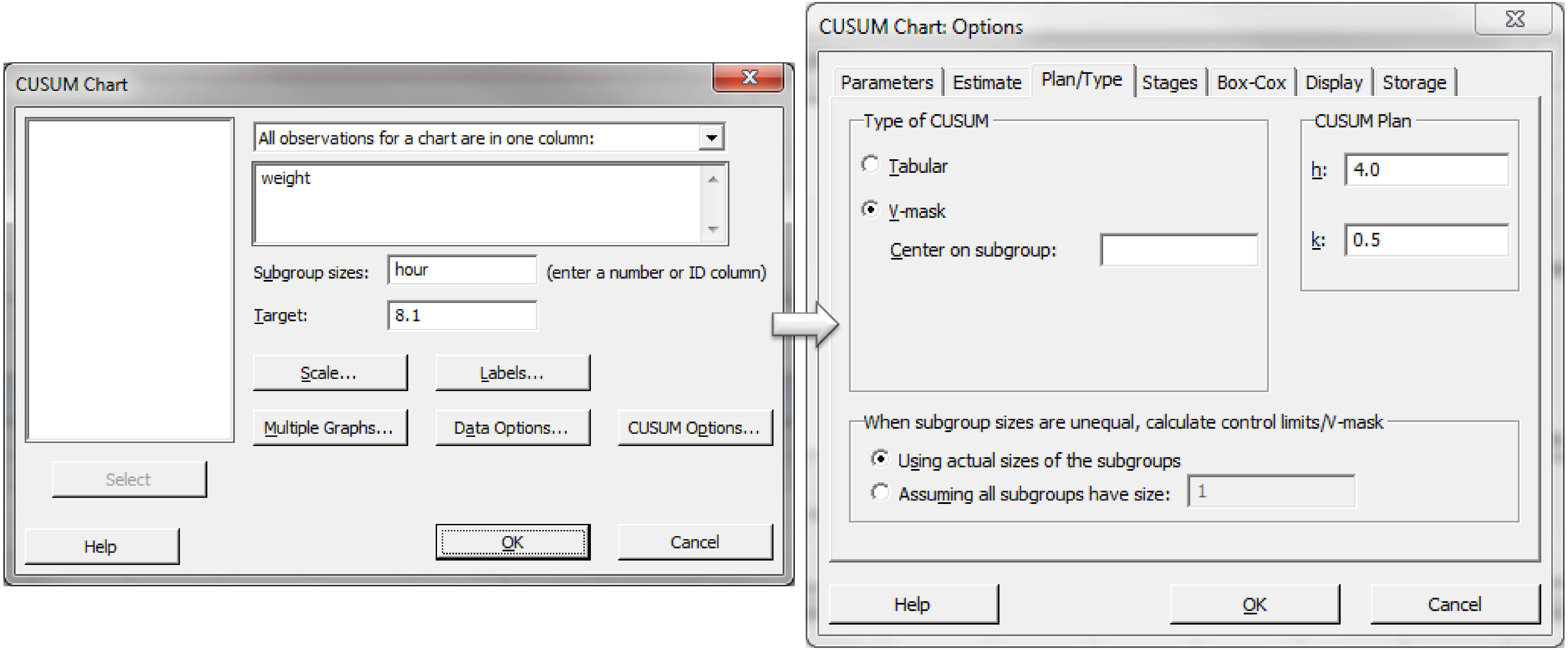
where $\bar{x}_i$ is the mean of the ith subgroup;
T is the process target;
k is the slope of the lower boundary of the V-mask.

Use Minitab to Plot a CumSum Chart

- Data File: “CumSum” tab in “Sample Data.xlsx”
- Steps in Minitab to plot CumSum charts

  1. Click Stat → Control Charts → Time Weighted Charts → CUSUM.
  2. A new window named “CUSUM Chart” appears.
  3. Select “weight” in the box below “All observations for a chart are in one column.”
  4. Select “hour” as the “Subgroup sizes.”
  5. Enter “8.1” as “Target.”
  6. Click the button “CUSUM Options” and another window named “CUSUM Chart – Options” pops up.
  7. Click the tab “Plan/Type.”
  8. Select the radio button “V-mask.”
  9. Click “OK” in the window “CUSUM Chart Options.”
  10. Click “OK.”
  11. The CUSUM chart appears in the newly-generated window.

Use Minitab to Plot a CumSum Chart



CUSUM Chart

All observations for a chart are in one column: weight

Subgroup sizes: hour (enter a number or ID column)

Target: 8.1

CUSUM Chart: Options – Plan/Type

Type of CUSUM

○ Tabular

● V-mask

Center on subgroup:

CUSUM Plan

h: 4.0

k: 0.5

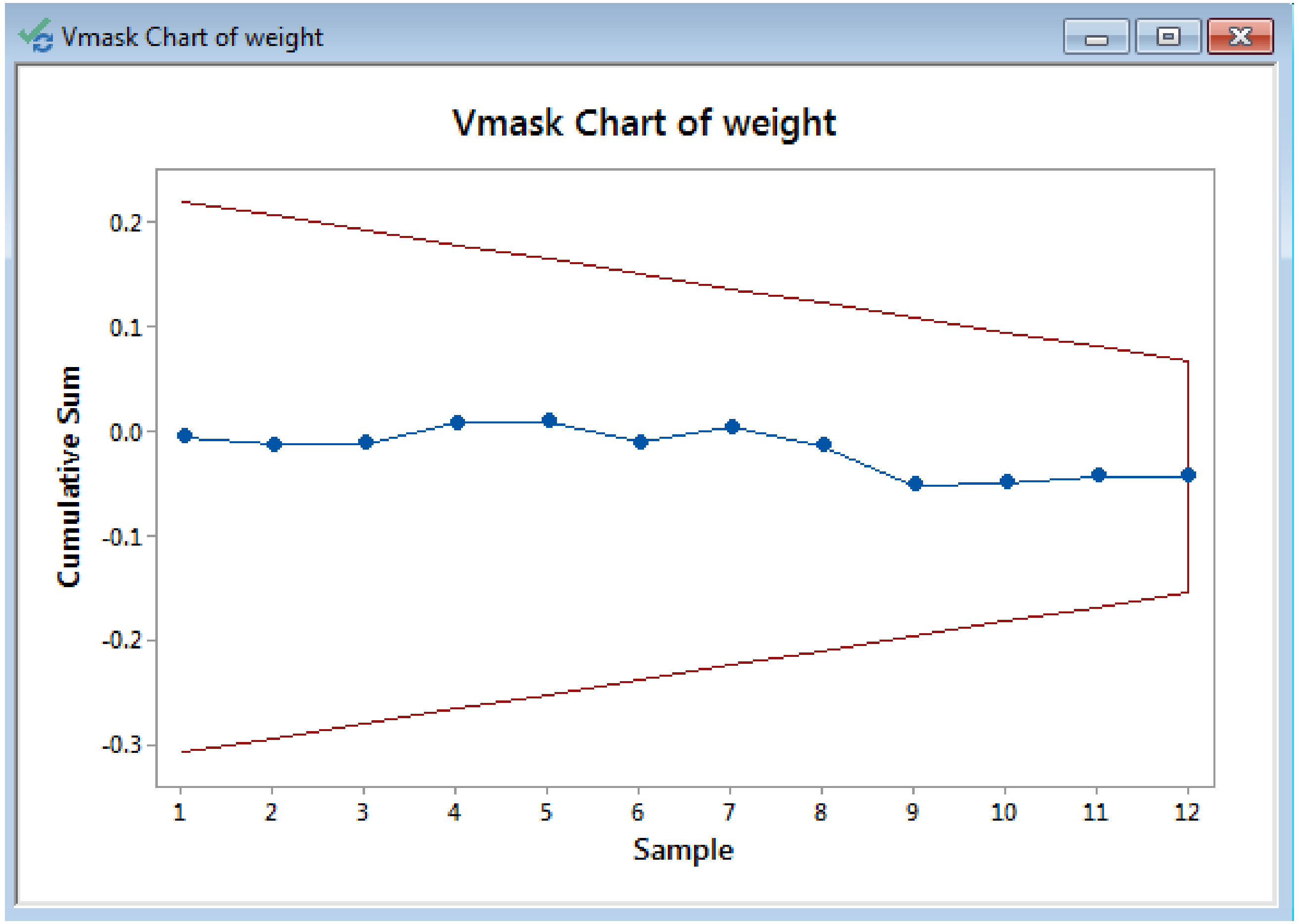
When subgroup sizes are unequal, calculate control limits/V-mask

● Using actual sizes of the subgroups

○ Assuming all subgroups have size: 1

Use Minitab to Plot a CumSum Chart

- CumSum Chart Interpretation:
  - The process is in control since all of the data points fall between the two arms of the V-mask.



5.2.9 EWMA Chart

EWMA Chart

- The **EWMA chart** (Exponentially-Weighted Moving Average Chart) is a control chart monitoring the exponentially-weighted average of previous and present subgroup means.
- The more recent data get more weight than more historical data.
- It detects the shift of the process mean from the process target over time.
- The underlying distribution of the EWMA chart is normal distribution.

EWMA Chart Equations

- EWMA Chart
  - Data Point: $z_i = \lambda \overline{x}_i + (1-\lambda) z_{i-1}$ where $0 < \lambda < 1$
  - Center Line: $\overline{X}$
  - Control Limits: $\overline{X} \pm k \cdot \frac{s}{\sqrt{n}} \sqrt{\left(\frac{\lambda}{2-\lambda}\right)\left[1-(1-\lambda)^{2i}\right]}$

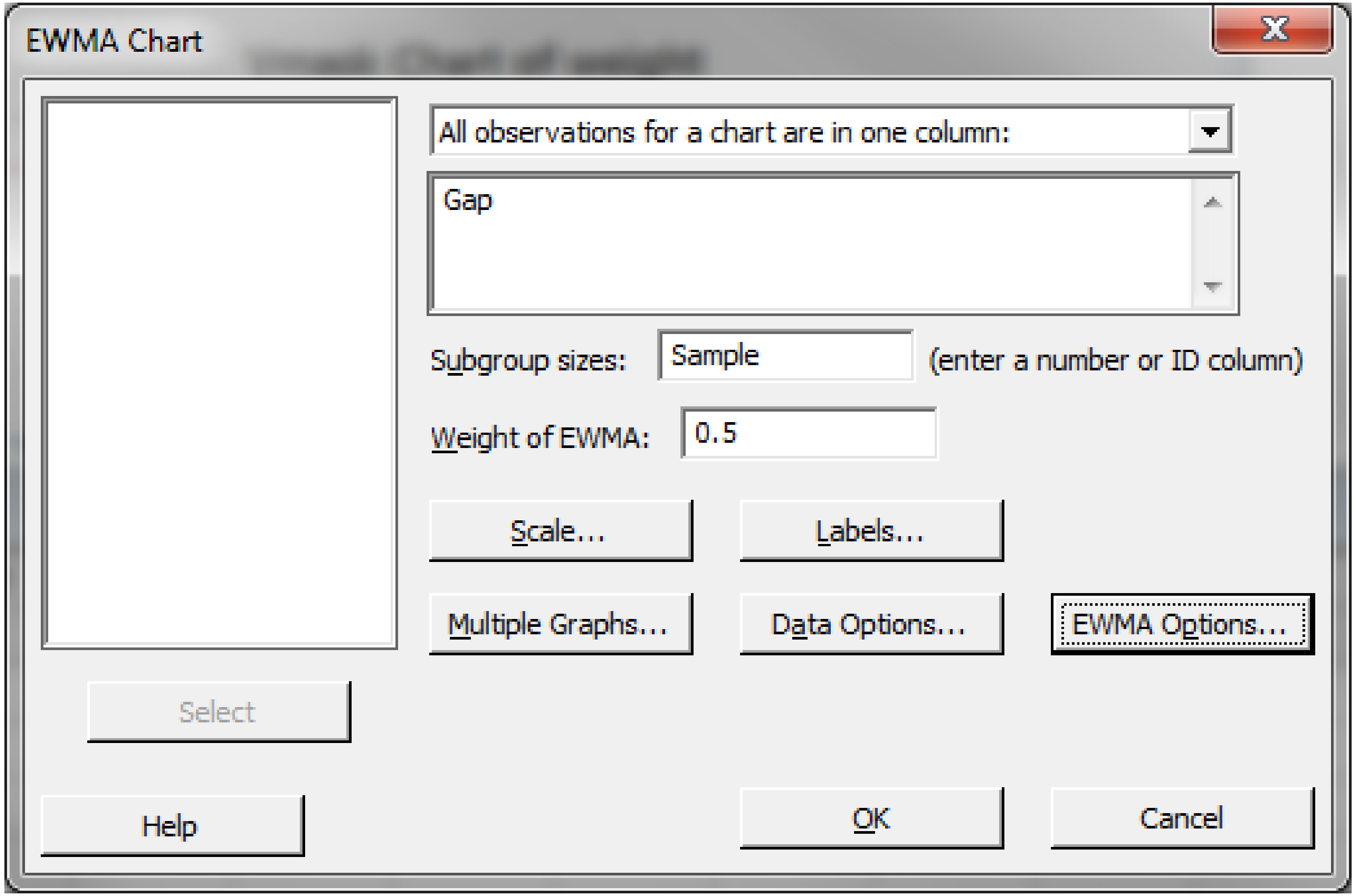
where $\overline{x}_i$ is the mean of the ith subgroup;
λ and *k* are user-defined parameters to calculate the EWMA data points and the control limits.

Use Minitab to Plot an EWMA Chart

- Data File: “EWMA” tab in “Sample Data.xlsx”
- Steps in Minitab to plot EWMA charts

  1. Click Stat → Control Charts → Time Weighted Charts → EWMA.
  2. A new window named “EWMA Chart” appears.
  3. Select “Gap” in the box below “All observations for a chart are in one column.”
  4. Select “Sample” as the “Subgroup sizes.”
  5. Enter “0.5” as “Weight of EWMA.”
  6. Click “OK.”
  7. The EWMA chart appears in the newly-generated window.

Use Minitab to Plot an EWMA Chart



All observations for a chart are in one column:

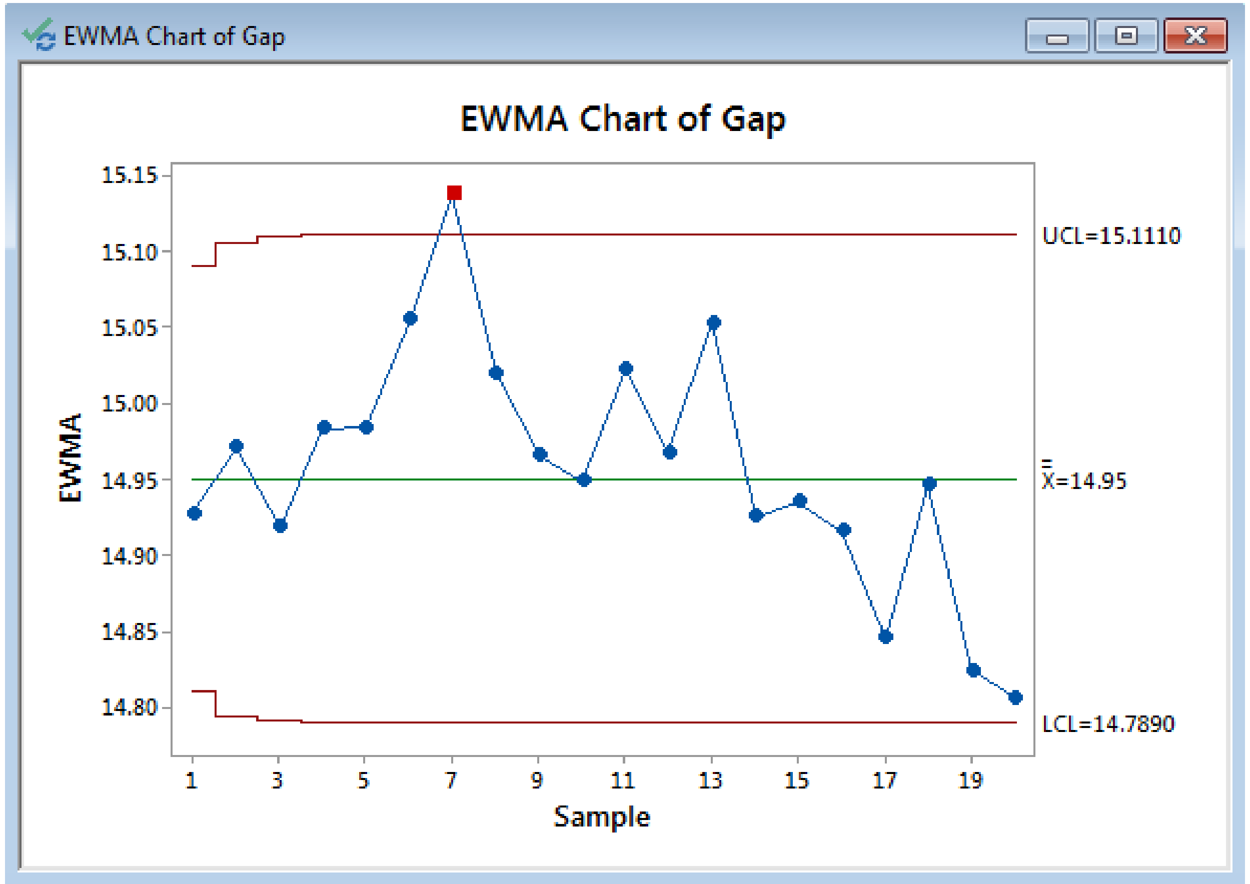
Gap

Subgroup sizes: Sample (enter a number or ID column)

Weight of EWMA: 0.5

Use Minitab to Plot an EWMA Chart

- **EWMA Chart Interpretation:**
  - One data point falls beyond the upper control limit and we conclude that the process is out of control.
  - Further investigation is needed to discover the root cause for the outlier.



5.2.10 Control Methods

Control Methods

- There are many control methods available to keep the process stable and minimize the variation.

- Most common control methods:
    - SPC (statistical process control)
    - 5S method
    - Kanban
    - Poka-Yoke (mistake proofing).

SPC

- **SPC (Statistical Process Control)** is a quantitative control method to monitor the stability of the process performance by identifying the special cause variation in the process.
- It uses control charts to detect the unanticipated changes in the process.
- Which control chart to use depends on:
  - Whether the data are continuous or discrete
  - How large the subgroup size is
  - Whether the subgroup size is constant
  - Whether we are interested in measuring defects or defectives
  - Whether we are interested in detecting the shifts in the process mean.

5S

- **5S** is a systematic approach of cleaning and organizing the workplace.
  - Seiri (sorting)
  - Seiton (straightening)
  - Seiso (shining)
  - Seiketsu (standardizing)
  - Shisuke (sustaining)

Kanban

- A **Kanban** system is a demand-driven system.
- The customer demand is the signal to trigger or pull the production.
- Products are made only to meet the immediate demand. When there is no demand, there is no production.
- It was designed to reduce the waste in inventory and increase the speed of responding to immediate demand.

Poka-Yoke

- **Poka-yoke** is a mechanism to eliminate the defects as early as possible in the process.

- Contact Method
  - Use of the shape, color, size, or any other physical attributes of the items

- Constant Number Method
  - Use of a fixed number to make sure a certain number of motions are completed

- Sequence Method
  - Use of a checklist to make sure all the prescribed process steps are followed in the right order.

5.2.11 Control Chart Anatomy

Control Chart Calculations Summary

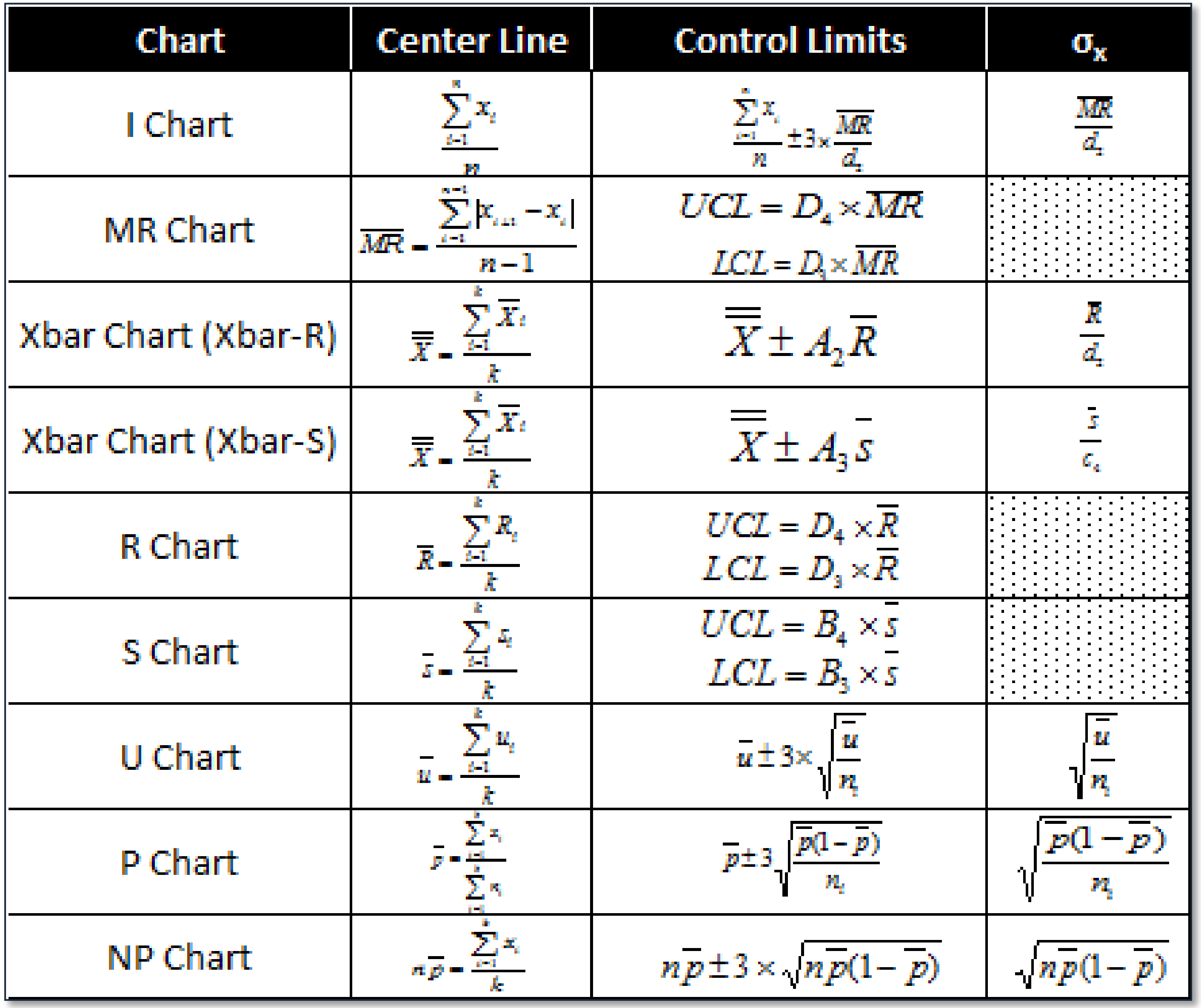
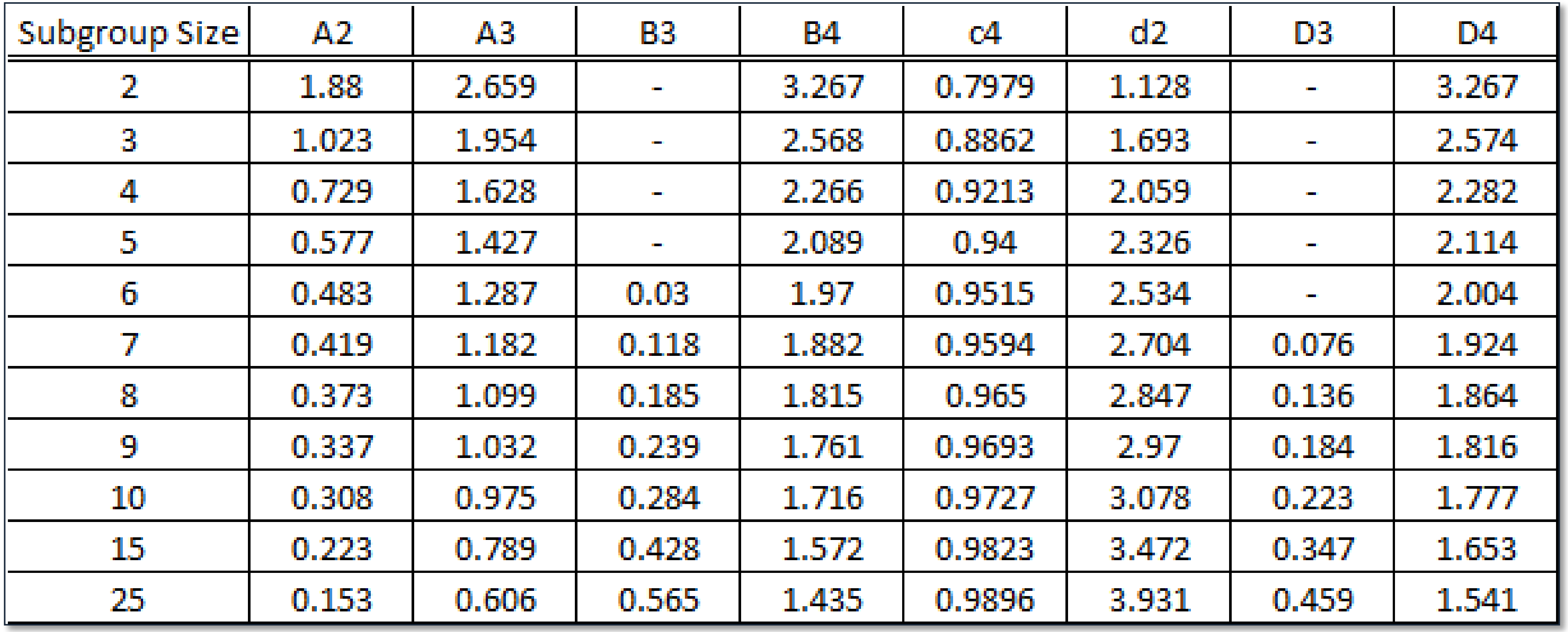


Chart	Center Line	Control Limits	σ_x
I Chart	$\frac{\sum_{i=1}^{n} x_i}{n}$	$\frac{\sum_{i=1}^{n} x_i}{n} \pm 3 \times \frac{\overline{MR}}{d_2}$	$\frac{\overline{MR}}{d_2}$
MR Chart	$\overline{MR} = \frac{\sum_{i=1}^{n-1} \lvert x_{i+1} - x_i \rvert}{n-1}$	$UCL = D_4 \times \overline{MR}$ $LCL = D_3 \times \overline{MR}$	▨
Xbar Chart (Xbar-R)	$\overline{\overline{X}} = \frac{\sum_{i=1}^{k} \overline{X}_i}{k}$	$\overline{\overline{X}} \pm A_2 \overline{R}$	$\frac{\overline{R}}{d_2}$
Xbar Chart (Xbar-S)	$\overline{\overline{X}} = \frac{\sum_{i=1}^{k} \overline{X}_i}{k}$	$\overline{\overline{X}} \pm A_3 \overline{s}$	$\frac{\overline{s}}{c_4}$
R Chart	$\overline{R} = \frac{\sum_{i=1}^{k} R_i}{k}$	$UCL = D_4 \times \overline{R}$ $LCL = D_3 \times \overline{R}$	▨
S Chart	$\overline{s} = \frac{\sum_{i=1}^{k} s_i}{k}$	$UCL = B_4 \times \overline{s}$ $LCL = B_3 \times \overline{s}$	▨
U Chart	$\overline{u} = \frac{\sum_{i=1}^{k} u_i}{k}$	$\overline{u} \pm 3 \times \sqrt{\frac{\overline{u}}{n_i}}$	$\sqrt{\frac{\overline{u}}{n_i}}$
P Chart	$\overline{p} = \frac{\sum_{i=1}^{k} x_i}{\sum_{i=1}^{k} n_i}$	$\overline{p} \pm 3 \sqrt{\frac{\overline{p}(1-\overline{p})}{n_i}}$	$\sqrt{\frac{\overline{p}(1-\overline{p})}{n_i}}$
NP Chart	$n\overline{p} = \frac{\sum_{i=1}^{k} x_i}{k}$	$n\overline{p} \pm 3 \times \sqrt{n\overline{p}(1-\overline{p})}$	$\sqrt{n\overline{p}(1-\overline{p})}$

Control Chart Constants



Subgroup Size	A2	A3	B3	B4	c4	d2	D3	D4
2	1.88	2.659	-	3.267	0.7979	1.128	-	3.267
3	1.023	1.954	-	2.568	0.8862	1.693	-	2.574
4	0.729	1.628	-	2.266	0.9213	2.059	-	2.282
5	0.577	1.427	-	2.089	0.94	2.326	-	2.114
6	0.483	1.287	0.03	1.97	0.9515	2.534	-	2.004
7	0.419	1.182	0.118	1.882	0.9594	2.704	0.076	1.924
8	0.373	1.099	0.185	1.815	0.965	2.847	0.136	1.864
9	0.337	1.032	0.239	1.761	0.9693	2.97	0.184	1.816
10	0.308	0.975	0.284	1.716	0.9727	3.078	0.223	1.777
15	0.223	0.789	0.428	1.572	0.9823	3.472	0.347	1.653
25	0.153	0.606	0.565	1.435	0.9896	3.931	0.459	1.541

Unnatural Patterns

- If there are *unnatural patterns* in the control chart of a process, we consider the process out of statistical control.

- Typical unnatural patterns in control charts:
  - Outliers
  - Trending
  - Cycling
  - Auto-correlative
  - Mixture.

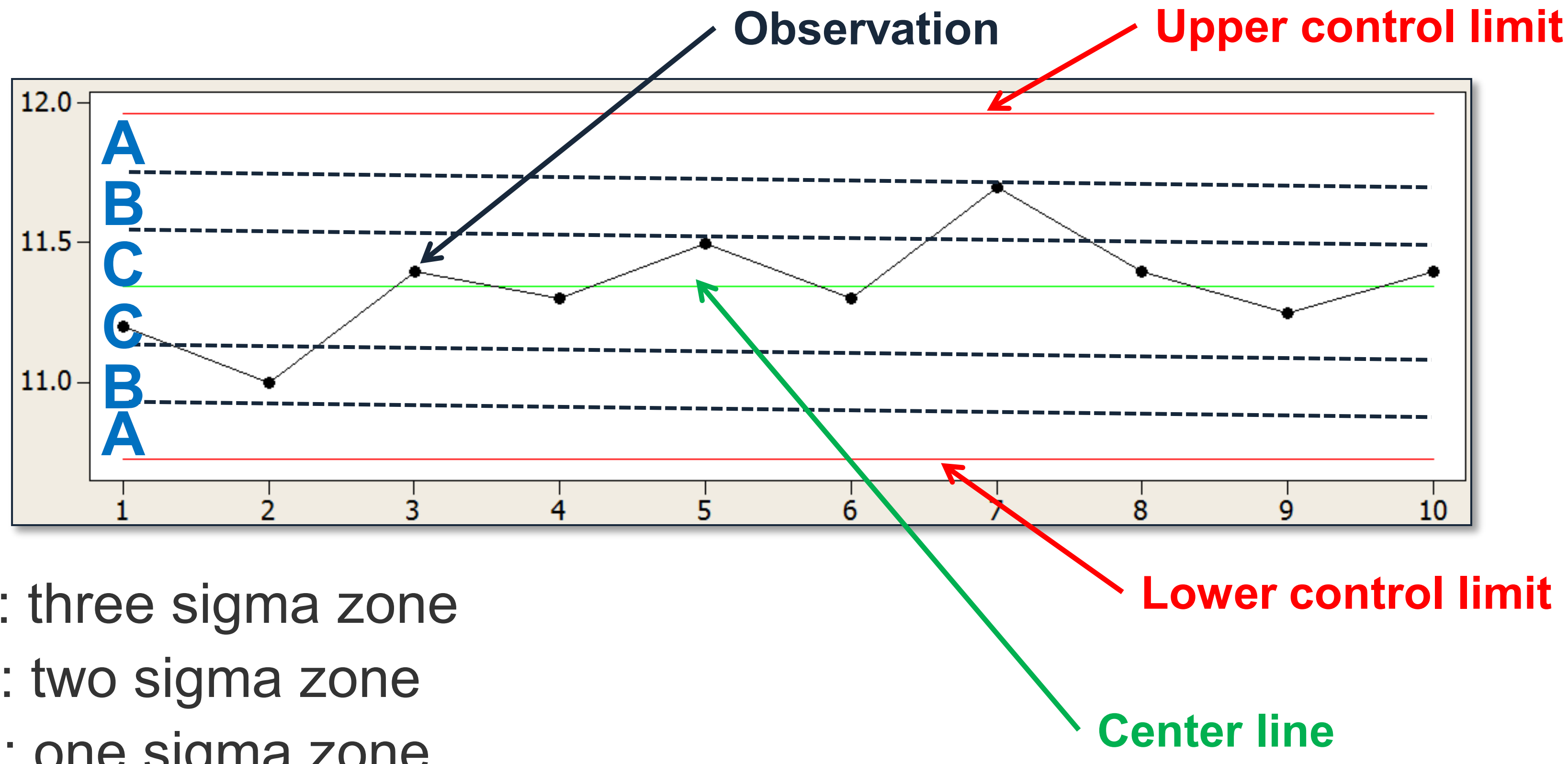
- A process is *in control* if all the data points on the control chart are randomly spread out within the control limits.

Western Electric Rules

- **Western Electric Rules** are the most popular decision rules to detect unnatural patterns in the control charts. They are a group of tests for special causes in a process.

- The area between the upper and lower control limits is separated into six subzones.
  - Zone A: between two and three standard deviations from the center line
  - Zone B: between one and two standard deviations from the center line
  - Zone C: within one standard deviation from the center line.

Western Electric Rules

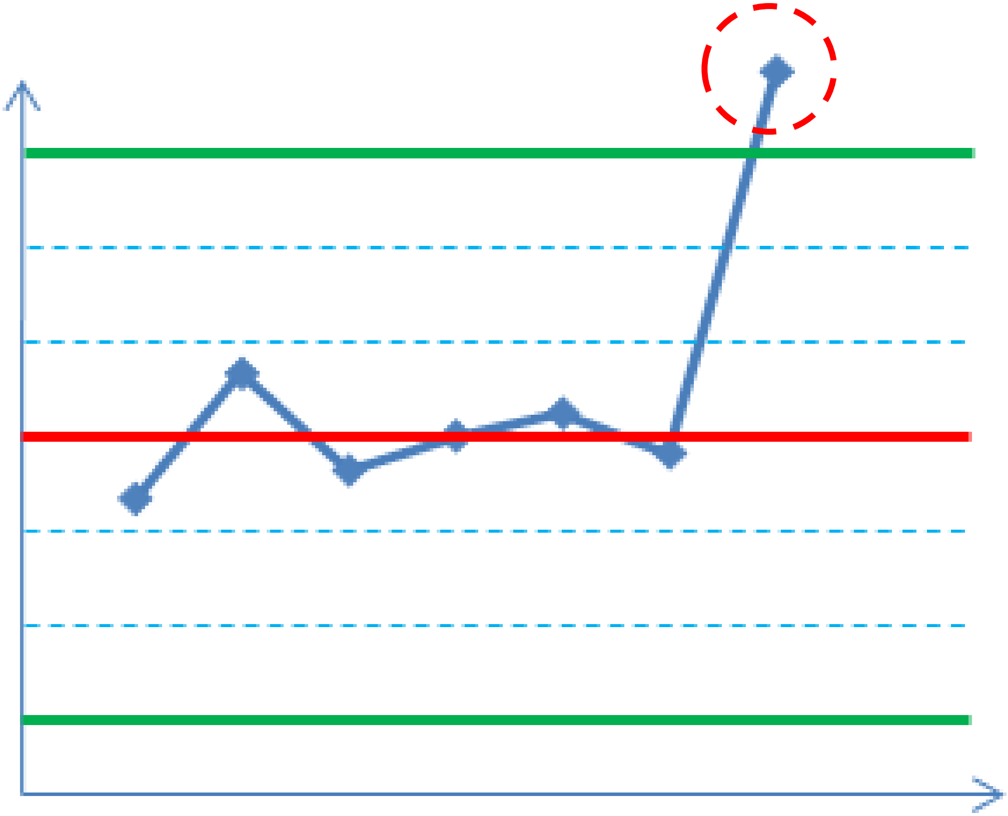


- Zone A: three sigma zone
- Zone B: two sigma zone
- Zone C: one sigma zone

- If a data point falls onto the dividing line of two consecutive zones, the point belongs to the outer zone.

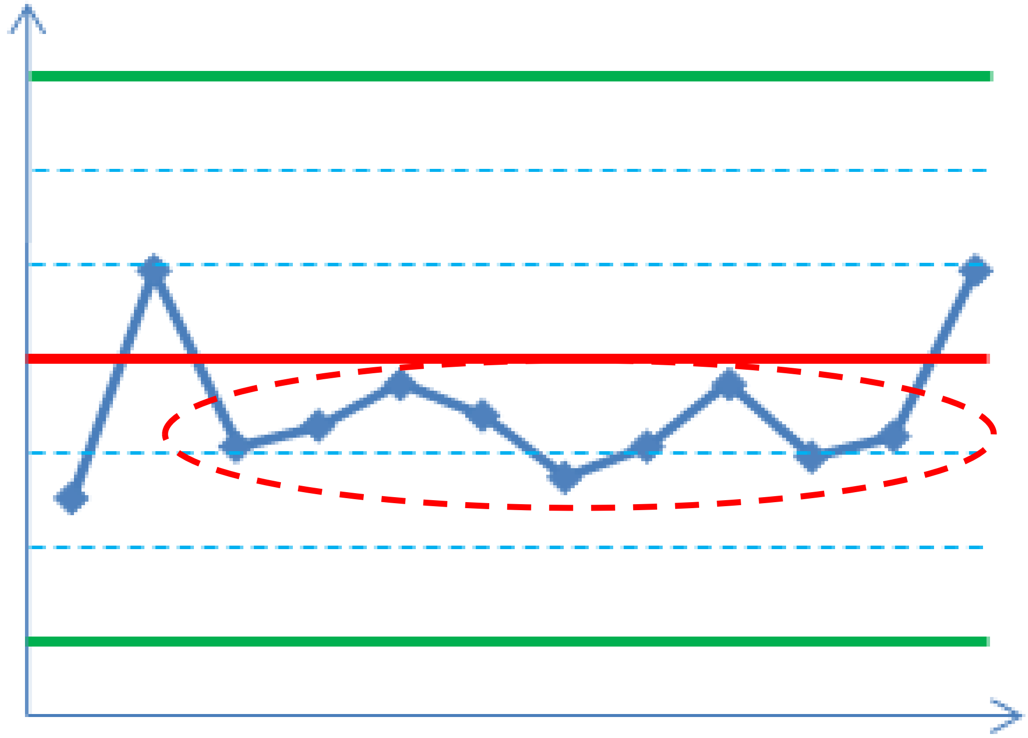
Western Electric Rules

- Test 1: 1 point more than 3 standard deviations from the center line (i.e., 1 point beyond zone A)



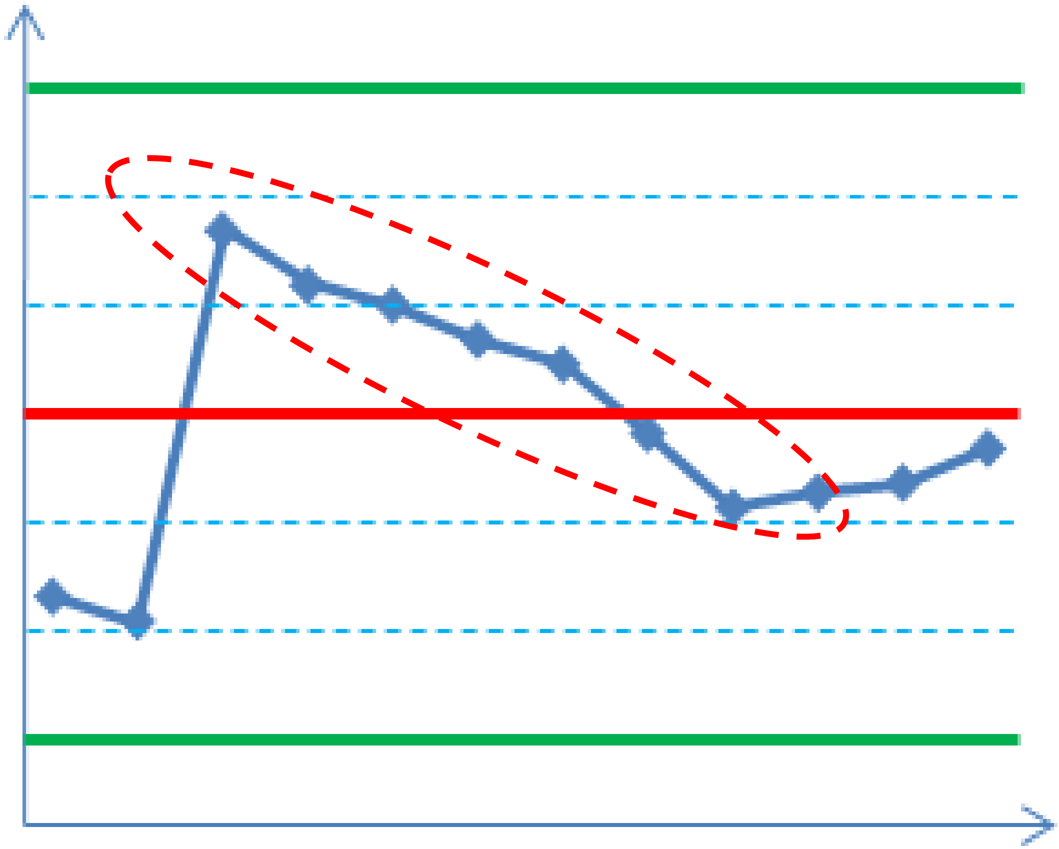
Western Electric Rules

- Test 2: 9 points in a row on the same side of the center line



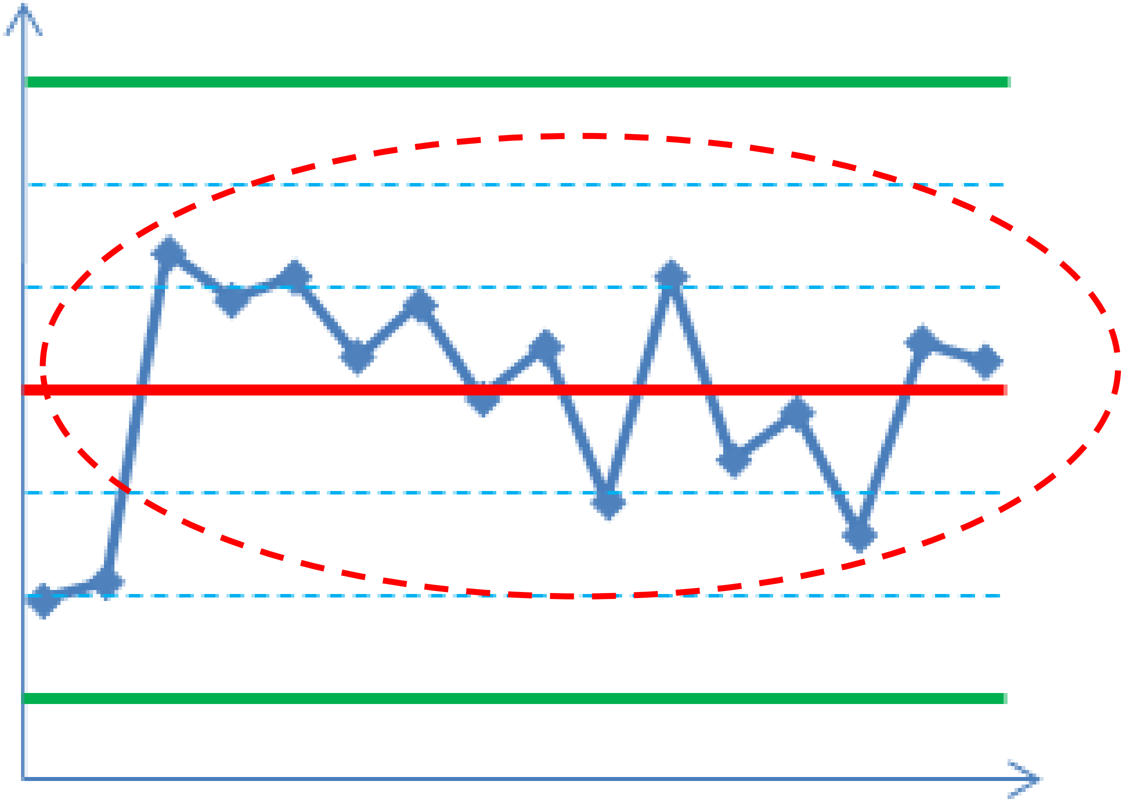
Western Electric Rules

- Test 3: 6 points in a row steadily increasing or steadily decreasing



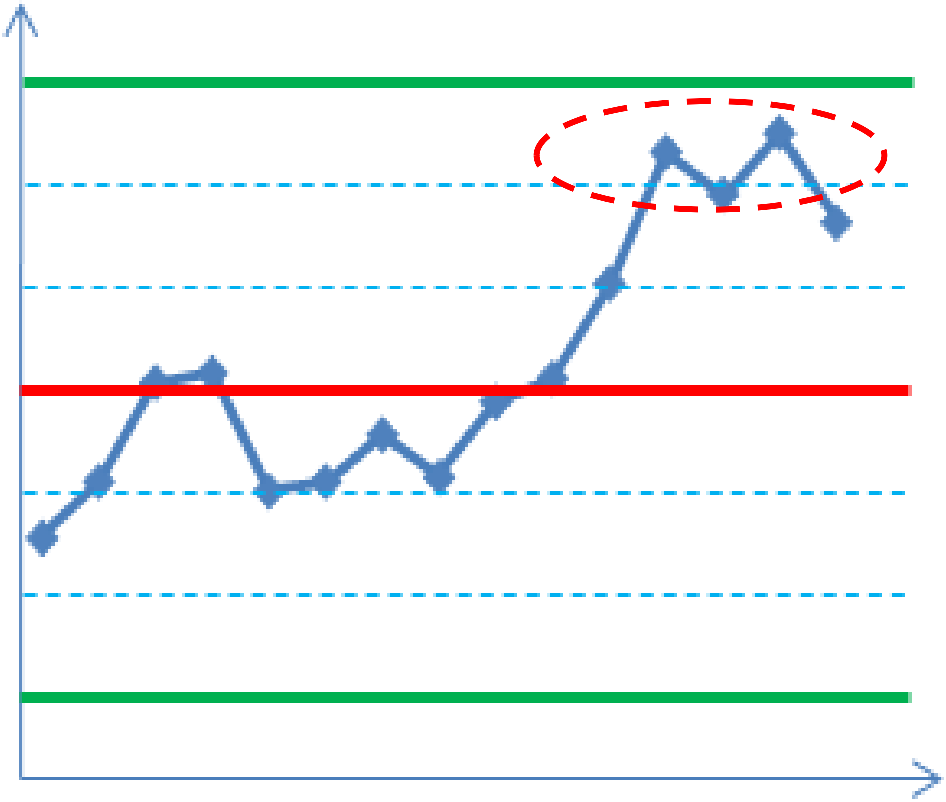
Western Electric Rules

- Test 4: 14 points in a row alternating up and down



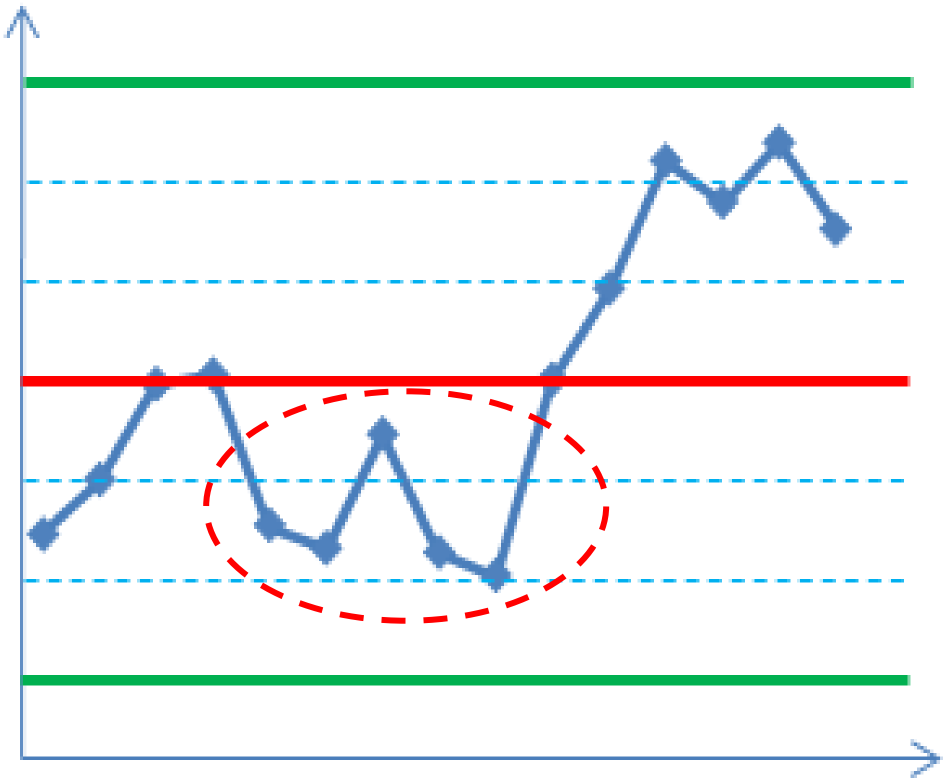
Western Electric Rules

- Test 5: 2 out of 3 points in a row at least 2 standard deviations from the center line (in zone A or beyond) on the same side of the center line



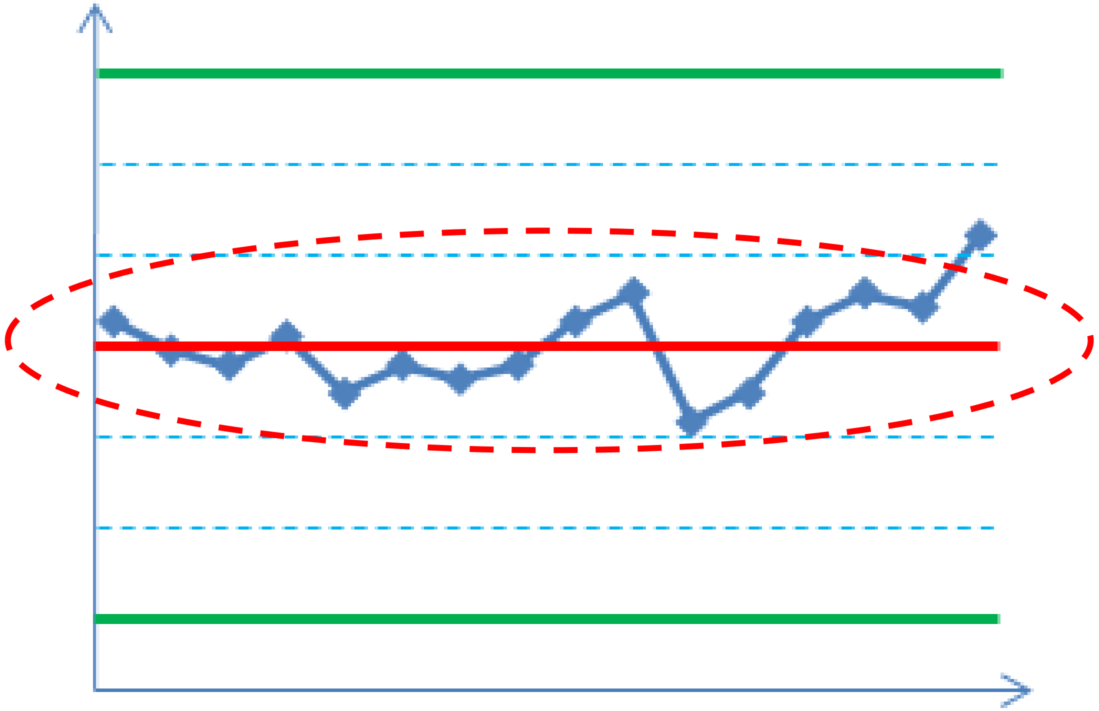
Western Electric Rules

- Test 6: 4 out of 5 points in a row at least 1 standard deviation from the center line (in zone B or beyond) on the same side of the center line



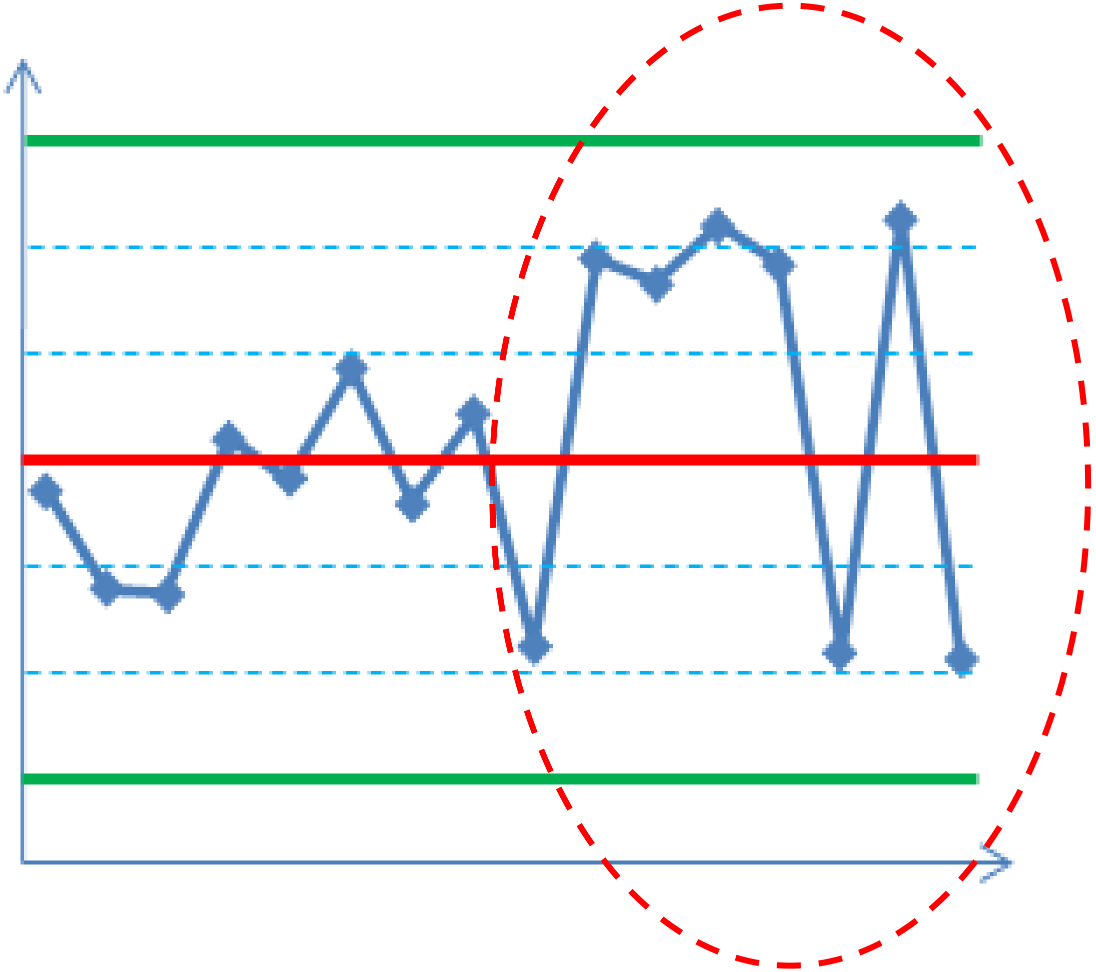
Western Electric Rules

- Test 7: 15 points in a row within 1 standard deviation from the center line (in zone C) on either side of the center line



Western Electric Rules

- Test 8: 8 points in a row beyond 1 standard deviation from the center line (beyond zone C) on either side of the center line



Next Steps

- If no data points fail any tests for special causes, the process is in *statistical control*.
- If any data point fails any tests for special causes, the process is *unstable* and we will need to investigate the observation thoroughly to discover and take actions on the special causes leading to the changes.
- Process stability is the prerequisite of process capability analysis.

5.2.12 Subgroups & Sampling

Subgroups

- **Rational subgrouping** is the basic sampling scheme in SPC (Statistical Process Control).
- When sampling, we randomly select a group (i.e., a subgroup) of items from the population of interest.
- The subgroup size is the count of samples in a subgroup. It can be constant or variable.
- Depending on the subgroup sizes, we select different control charts accordingly.

Impact of Variation

- The rational subgrouping strategy is designed to minimize the opportunity of having special cause variation within subgroups.

- If there is only random variation (background noise) within subgroups, all the special cause variation would be reflected between subgroups. It is easier to detect the out-of-control situation.

- Random variation is inherent and indelible in the process. We are more interested in identifying and taking actions on special cause variation.

Frequency of Sampling

- The **frequency of sampling** in SPC depends on whether we have sufficient data to signal the changes in a process with reasonable time and costs.
- The more frequently we sample, the higher the costs sampling may trigger.
- We need the subject matter experts' knowledge on the nature and characteristics of the process to make good decisions on sampling frequency.

5.3 Six Sigma Control Plans

Black Belt Training: Control Phase

5.1 Lean Controls
- 5.1.1 Control Methods for 5S
- 5.1.2 Kanban
- 5.1.3 Poka-Yoke (Mistake Proofing)

5.2 Statistical Process Control (SPC)
- 5.2.1 Data Collection for SPC
- 5.2.2 I-MR Chart
- 5.2.3 Xbar-R Chart
- 5.2.4 U Chart
- 5.2.5 P Chart
- 5.2.6 NP Chart
- 5.2.7 X-S chart
- 5.2.8 CumSum Chart
- 5.2.9 EWMA Chart
- 5.2.10 Control Methods
- 5.2.11 Control Chart Anatomy
- 5.2.12 Subgroups, Variation, Sampling

5.3 Six Sigma Control Plans
- 5.3.1 Cost Benefit Analysis
- 5.3.2 Elements of the Control Plan
- 5.3.3 Elements of the Response Plan

5.3.1 Cost Benefit Analysis

What is Cost-Benefit Analysis?

- The **cost-benefit analysis** is a systematic method to assess and compare the financial costs and benefits of multiple scenarios in order to make sound economic decisions.

- A cost-benefit analysis is recommended to be done at the beginning of the project based on estimations of the experts from the finance team in order to determine whether the project is financially feasible.

- It is recommended to update the cost-benefit analysis at each DMAIC phase of the project.

Why Cost-Benefit Analysis?

- In the Define phase of the project, the cost-benefit analysis helps us understand the financial feasibility of the project.

- In the middle phases of the project, updating and reviewing the cost-benefit analysis helps us compare potential solutions and make robust data-driven decisions.

- In the Control phase of the project, the cost-benefit analysis helps us track the project's profitability.

Return on Investment

- **Return on investment** (also called ROI, rate of return, or ROR) is the ratio of the net financial benefits (either gain or loss) of a project or investment to the financial costs.

$$ROI = \frac{TotalNetBenefits}{TotalCosts} \times 100\%$$

where

$$TotalNetBenefits = TotalBenefits - TotalCosts$$

Return on Investment (ROI)

- The return on investment is used to evaluate the financial feasibility and profitability of a project or investment.
    - If ROI < 0, the investment is not financially viable.
    - If ROI = 0, the investment has neither gain nor loss.
    - If ROI > 0, the investment has financial gains.
- The higher the ROI, the more profitable the project.

Net Present Value (NPV)

- The **net present value** (also called NPV, net present worth, or NPW) is the total present value of the cash flows calculated using a discount rate.

$$NPV = \frac{NetCashFlow_t}{(1+r)^t}$$

Where

$NetCashFlow_t$ is the net cash flow happening at time t;
r is the discount rate;
t is the time of the cash flow.

Cost Estimation

- Examples of costs triggered by the project:
  - Administration
  - Asset
  - Equipment
  - Material
  - Delivery
  - Real estate
  - Labor
  - Training
  - Consulting.

Benefits Estimation

- Examples of benefits generated by the project:
  - Direct revenue increase
  - Waste reduction
  - Operation cost reduction
  - Quality and productivity improvement
  - Market share increase
  - Cost avoidance
  - Customer satisfaction improvement
  - Associate satisfaction improvement.

Challenges in Cost and Benefit Estimation

- Different analysts might come up with different cost and benefit estimations due to their subjectivity in determining:
    - The discount rate
    - The time length of the project and its impact
    - Potential costs of the project
    - The tangible/intangible benefits of the project
    - The specific contribution of the project to the relevant financial gains/loss.

5.3.2 Elements of Control Plans

Control Plans

- The **control plans** ensure that the changes introduced by a Six Sigma project are sustained over time.

- Benefits of the Control phase:
    - Methodical roll-out of changes including standardization of processes and work procedures
    - Ensure compliance with changes through methods like auditing and corrective actions
    - Transfer solutions and learning across the enterprise
    - Plan and communicate standardized work procedures
    - Coordinate ongoing team and individual involvement
    - Standardize data collection and procedures
    - Measure process performance, stability, and capability
    - Plan actions that mitigate possible out-of-control conditions
    - Sustain changes over time.

What is a Control Plan?

- A **control plan** is a management planning tool to identify, describe, and monitor the process performance metrics in order to meet the customer specifications steadily.
- It proposes the plan of monitoring the stability and capability of inputs and outputs of critical process steps in the Control phase of a project.
- It covers the data collection plan of gathering the process performance measurements.
- Control plans are the most overlooked element of most projects. It is critical that a good solution be solidified with a great control plan!

Control Plan Elements

- Control Plan
  - The clear and concise summary document that details key process steps, CTQs metrics, measurements, and corrective actions.
- Standard Operating Procedures (SOPs)
  - Supporting documentation showing the “who does what, when, and how” in completing the tasks.
- Communication Plan
  - Document outlining messages to be delivered and the target audience.
- Training Plan
  - Document outlining the necessary training for employees to successfully perform new processes and procedures.
- Audit Checklists
  - Document that provides auditors with the audit questions they need to ask.
- Corrective Actions
  - Activities that need to be conducted when an audit fails.

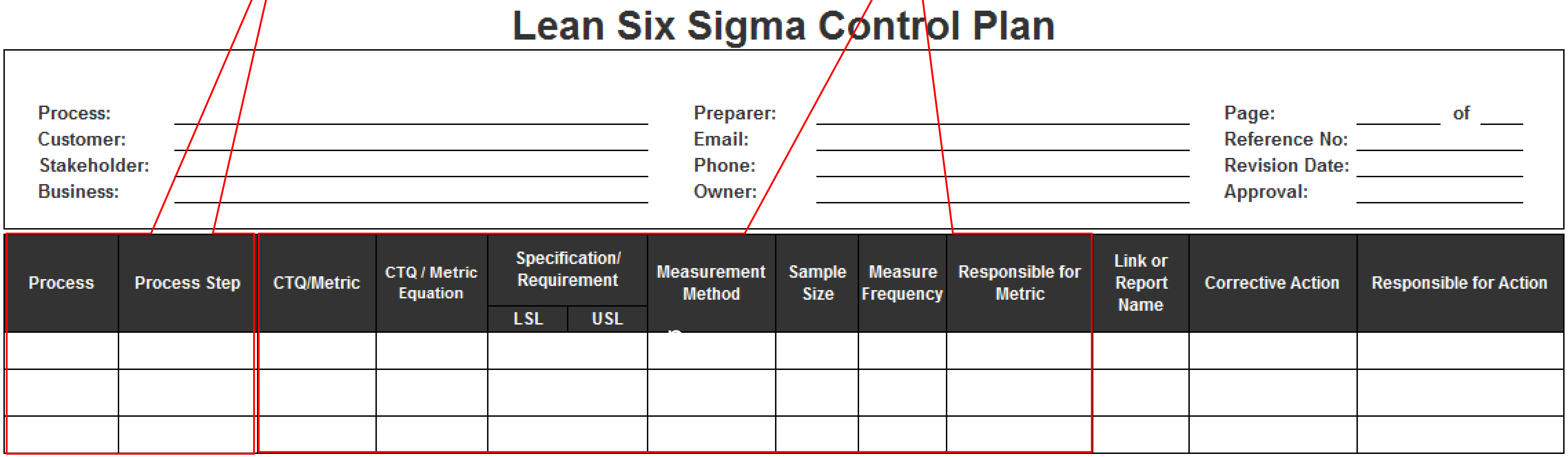
Control Plan

- The control plan identifies critical process steps that have significant impact on the products or services and the appropriate controls mechanisms.
- The control plan includes measurement systems that monitor and help manage key process step performance.
- Specified limits and targets of the performance metrics are clearly defined and communicated.
- Sampling plans to collect the measurements are declared:
  - How many samples are needed?
  - How often do we need to sample?
  - Where should we sample?

Control Plan

Key processes and process steps are identified

Critical information regarding key measurements is documented and clarified



Lean Six Sigma Control Plan

Process: ______
Customer: ______
Stakeholder: ______
Business: ______

Preparer: ______
Email: ______
Phone: ______
Owner: ______

Page: ______ of ______
Reference No: ______
Revision Date: ______
Approval: ______

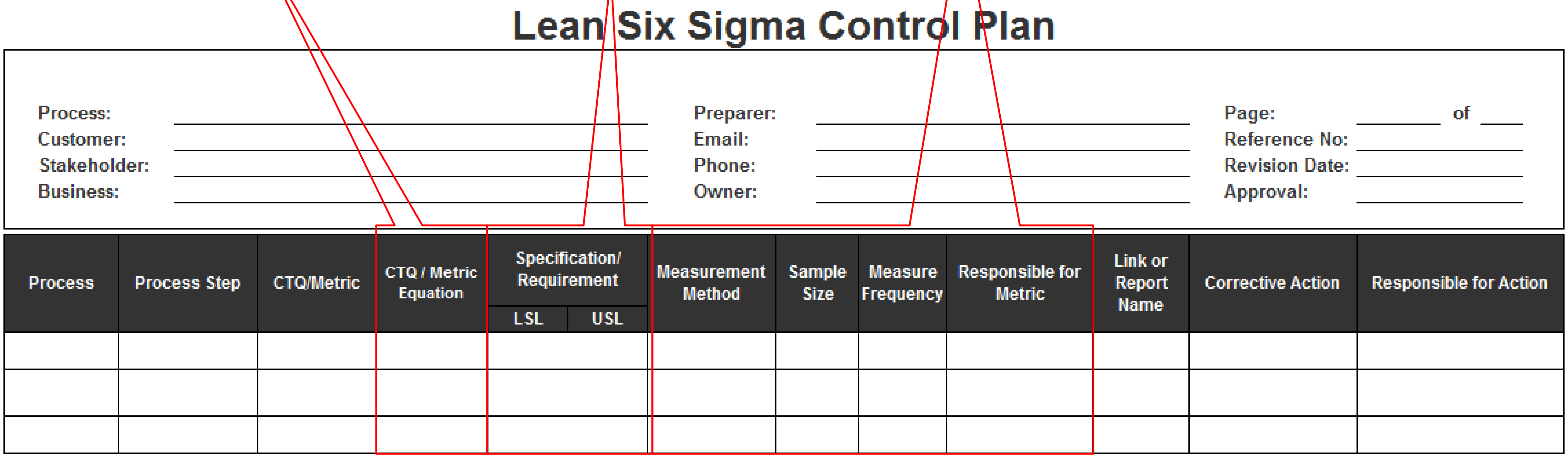
Process	Process Step	CTQ/Metric	CTQ / Metric Equation	Specification/ Requirement		Measurement Method	Sample Size	Measure Frequency	Responsible for Metric	Link or Report Name	Corrective Action	Responsible for Action
				LSL	USL							

Control Plan

Measurements are clearly defined with equations

Customer specifications are declared

Other key measurement information is documented: sample size, measurement frequency, people responsible for the measurement, etc.



Lean Six Sigma Control Plan

Process: ______________________
Customer: ______________________
Stakeholder: ______________________
Business: ______________________

Preparer: ______________________
Email: ______________________
Phone: ______________________
Owner: ______________________

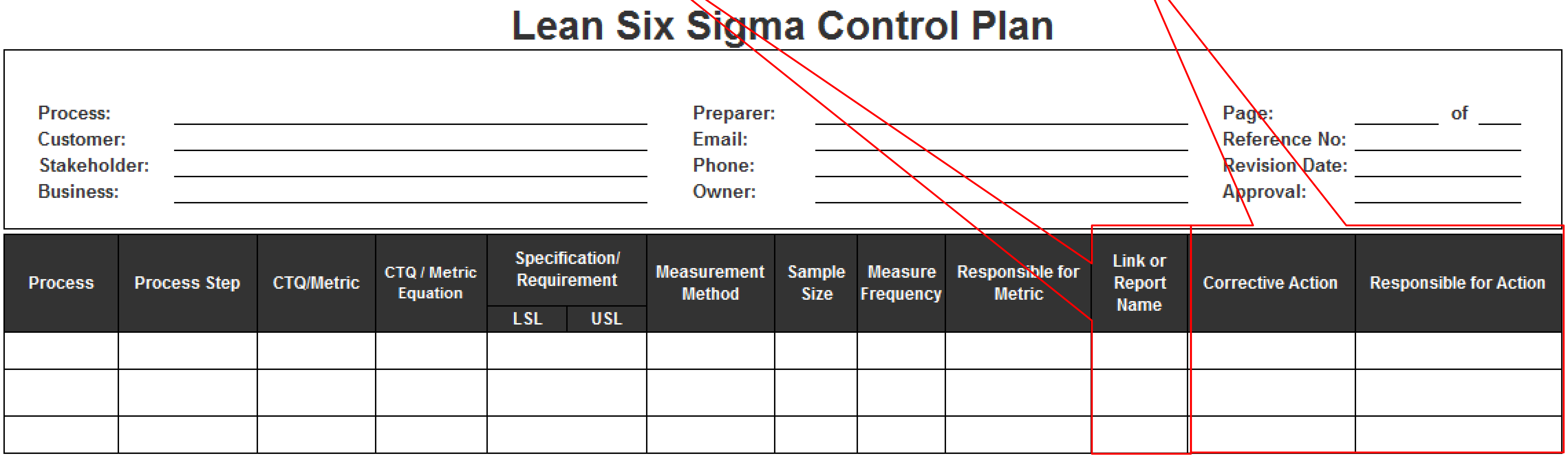
Page: ________ of ____
Reference No: ______________________
Revision Date: ______________________
Approval: ______________________

Process	Process Step	CTQ/Metric	CTQ / Metric Equation	Specification/ Requirement		Measurement Method	Sample Size	Measure Frequency	Responsible for Metric	Link or Report Name	Corrective Action	Responsible for Action
				LSL	USL							

Control Plan

Where will this measurement or report be found? Good control plans provide linking information or other report reference information.

Control plans identify the mitigating action or corrective actions required in the event the measurement falls out of spec or control. Responsible parties are also declared.



Lean Six Sigma Control Plan

Process: ____________________
Customer: ____________________
Stakeholder: ____________________
Business: ____________________

Preparer: ____________________
Email: ____________________
Phone: ____________________
Owner: ____________________

Page: ________ of ____
Reference No: ____________
Revision Date: ____________
Approval: ____________

Process	Process Step	CTQ/Metric	CTQ / Metric Equation	Specification/ Requirement		Measurement Method	Sample Size	Measure Frequency	Responsible for Metric	Link or Report Name	Corrective Action	Responsible for Action
				LSL	USL							

Standard Operating Procedures (SOPs)

- **Standard Operating Procedures** (SOPs) are documents that focus on process steps, activities, and specific tasks required to complete an operation.
- SOPs should not be much more than two to four pages.
- SOPs should be written to the user's level of required detail and information.
  - The level of detail is dependent on the position's required skills and training
- Good SOPs are auditable, easy to follow, and not difficult to find.
  - Auditable characteristics are: observable actions and countable frequencies. Results should be evident to a third party (*compliance to the SOP must be measurable*).

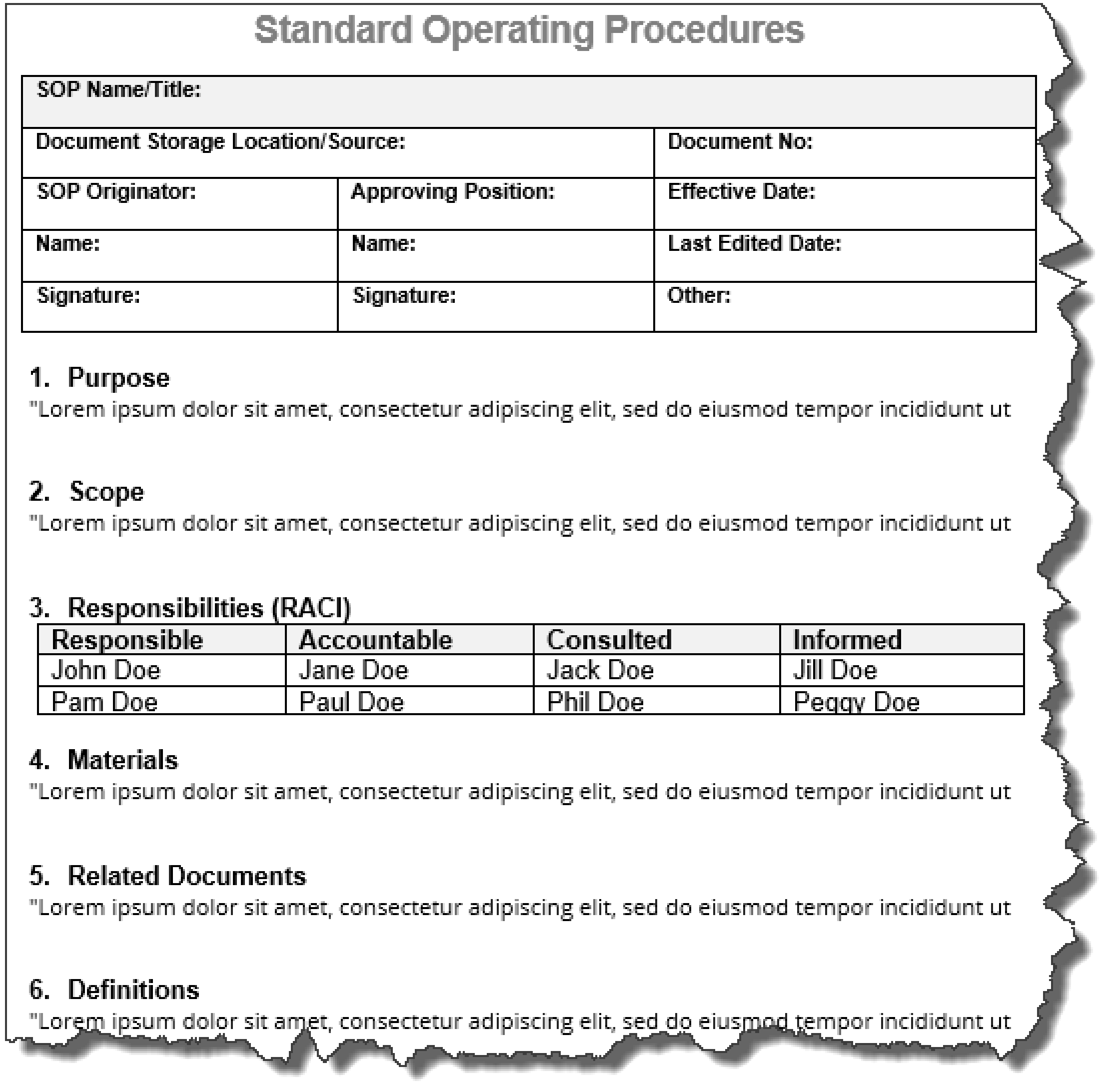
SOP Elements

- SOPs are intended to impart high value information in concise and well-documented manner.
- SOP Title and Version Number:
  - Provide a title and unique identification number with version information.
- Date:
  - List the original creation date; add all revision dates.
- Purpose:
  - State the reason for the SOP and what it intends to accomplish.
- Scope:
  - Identify all functions, jobs, positions, and/or processes governed or affected by the SOP.

SOP Elements

- Responsibilities:
  - Identify job functions and positions (not people) responsible for carrying out activities listed in the SOP.
- Materials:
  - List all material inputs: parts, files, data, information, instruments, etc.
- Process Map:
  - Show high level or level two to three process maps or other graphical representations of operating steps.
- Process Metrics:
  - Declare all process metrics and targets or specifications.
- Procedures:
  - List actual steps required to perform the function.
- References:
  - List any documents that support the SOP.

SOP Template



Standard Operating Procedures

SOP Name/Title:		
Document Storage Location/Source:		**Document No:**
SOP Originator:	**Approving Position:**	**Effective Date:**
Name:	**Name:**	**Last Edited Date:**
Signature:	**Signature:**	**Other:**

1. Purpose

"Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut

2. Scope

"Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut

3. Responsibilities (RACI)

Responsible	Accountable	Consulted	Informed
John Doe	Jane Doe	Jack Doe	Jill Doe
Pam Doe	Paul Doe	Phil Doe	Peggy Doe

4. Materials

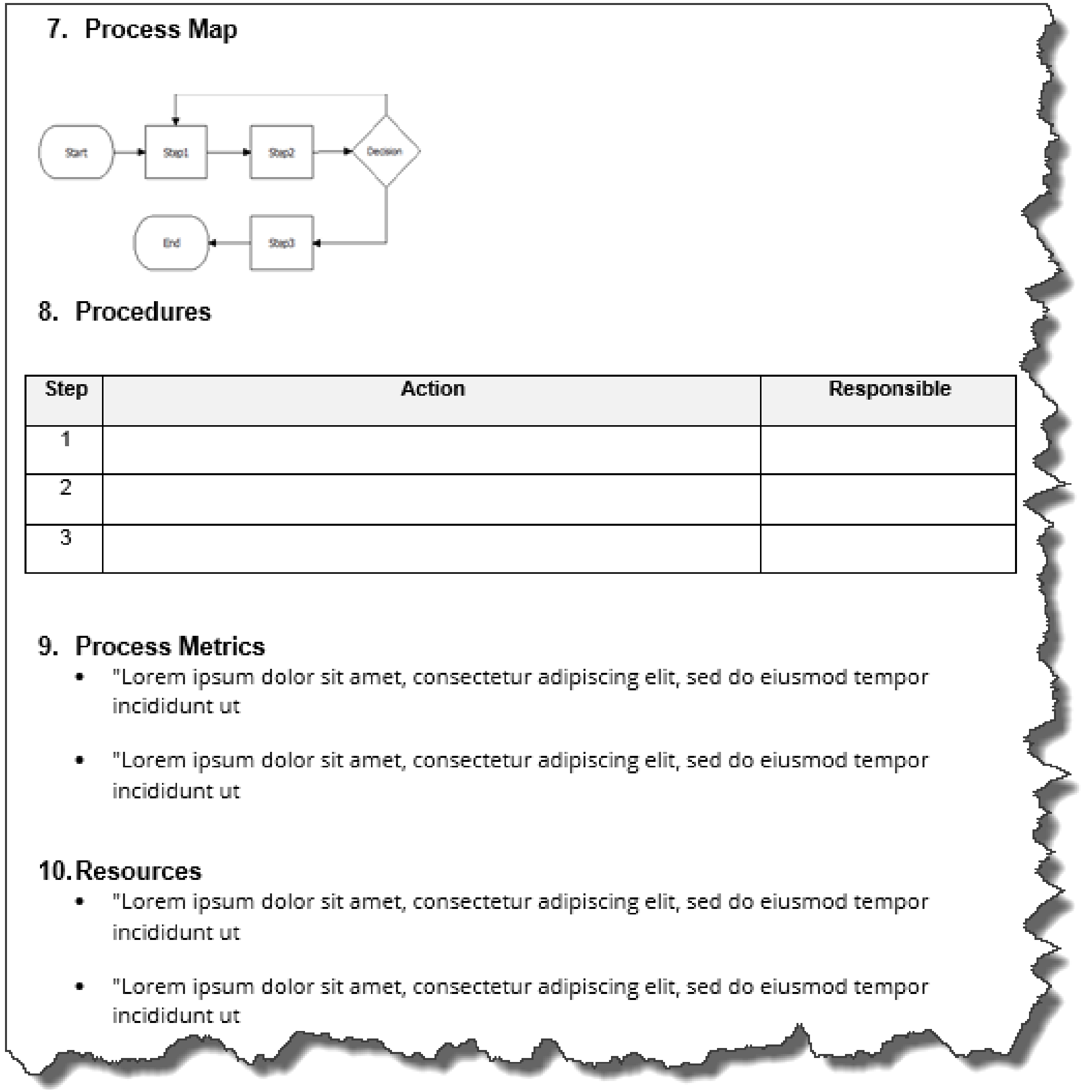
"Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut

5. Related Documents

"Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut

6. Definitions

"Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut



7. Process Map

8. Procedures

Step	Action	Responsible
1		
2		
3		

9. Process Metrics

- "Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut
- "Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut

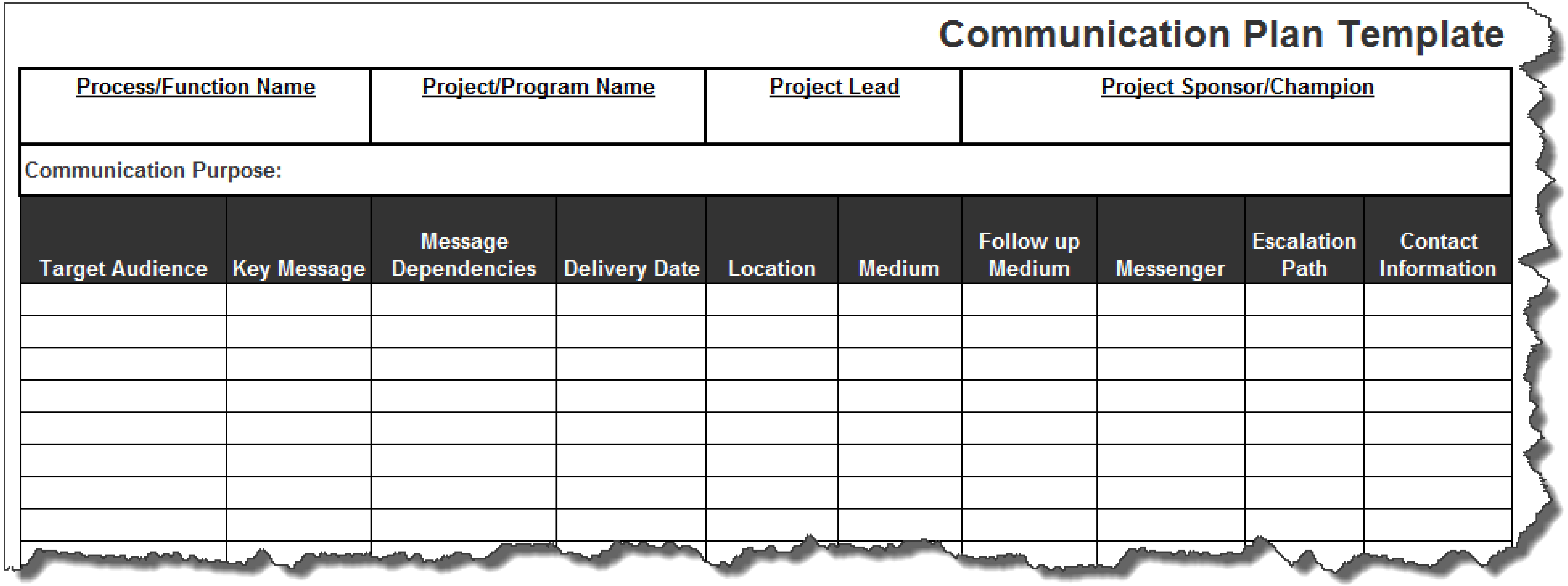
10. Resources

- "Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut
- "Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut

Communication Plans

- **Communication plans** are documents that focus on planning and preparing for the dissemination of information.
- Communication plans organize messages and ensure that the proper audiences receive the correct message at the right time.
- A good communication plan identifies:
  - Audience
  - Key points/message
  - Medium (how the message is to be delivered)
  - Delivery schedule
  - Messenger
  - Dependencies and escalation points
  - Follow-up messages and delivery mediums.
- Communication plans help develop and execute strategies for delivering changes to an organization.

Communication Plan Template



Communication Plan Template

Process/Function Name	**Project/Program Name**	**Project Lead**	**Project Sponsor/Champion**

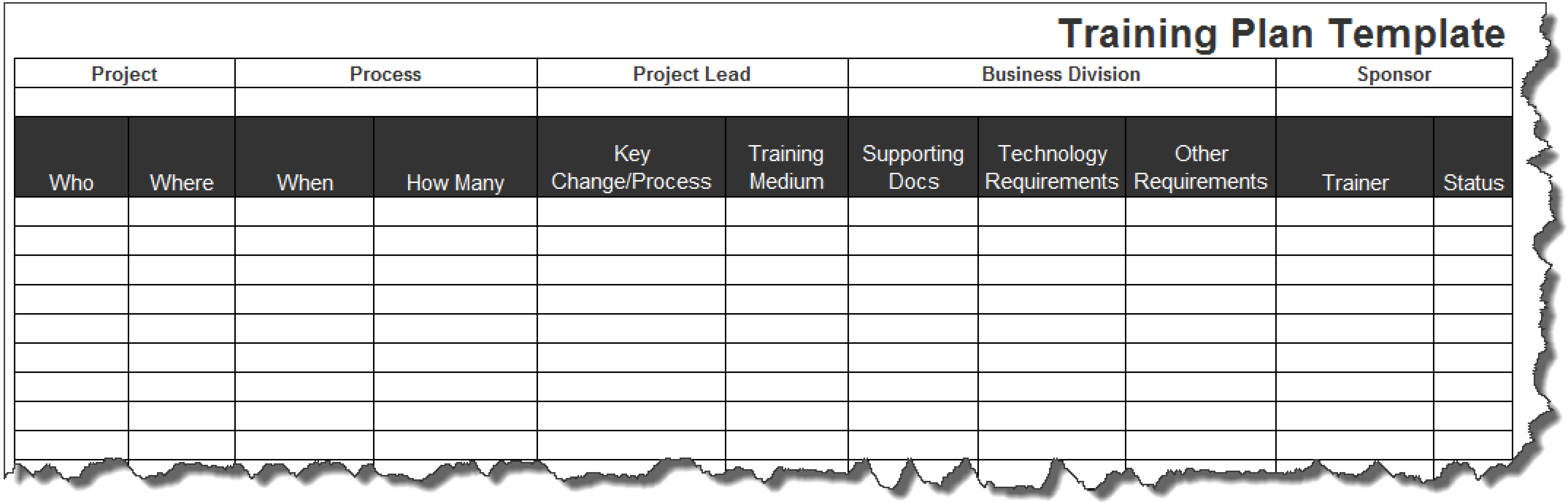
Communication Purpose:

Target Audience	Key Message	Message Dependencies	Delivery Date	Location	Medium	Follow up Medium	Messenger	Escalation Path	Contact Information

Training Plans

- **Training plans** are used to manage the delivery of training for new processes and procedures.
- Most GB or BB projects will require changes to processes and/or procedures that must be executed or followed by various employees.
- Training plans should incorporate all SOPs related to performing new or modified tasks.
- Training plans use and support existing SOPs and do not supersede them.
- Training plans should include logistics:
  - One-on-one or classroom
  - Instruction time
  - Location of training materials
  - Master training reference materials
  - Instructors and intended audience
  - Trainee names.

Training Plan Template



Training Plan Template

Project		Process		Project Lead		Business Division			Sponsor	
Who	Where	When	How Many	Key Change/Process	Training Medium	Supporting Docs	Technology Requirements	Other Requirements	Trainer	Status

Audits

- What is an **audit**?
  - ISO 9000 defines an Audit as “a systematic and independent examination to determine whether quality activities and related results comply with planned arrangements and whether these arrangements are implemented effectively and are suitable to achieve objectives.”

- Audits are used to ensure actions, processes, procedures, and other tasks are performed as expected.

Audit Guidelines

- Audits should be directed by managers, supervisors, and other accountable positions.

- An audit’s purpose must be well-defined and executed by independent unbiased personnel.

- Auditors must:
  - Be qualified to perform their tasks
  - Attend and successfully complete an internal auditing training session
  - Be able to identify whether or not activities are being followed according to the defined SOP
  - Base conclusions on facts and objective evidence
  - Use a well documented audit checklist.

- Audits should confirm compliance or declare non-compliance.

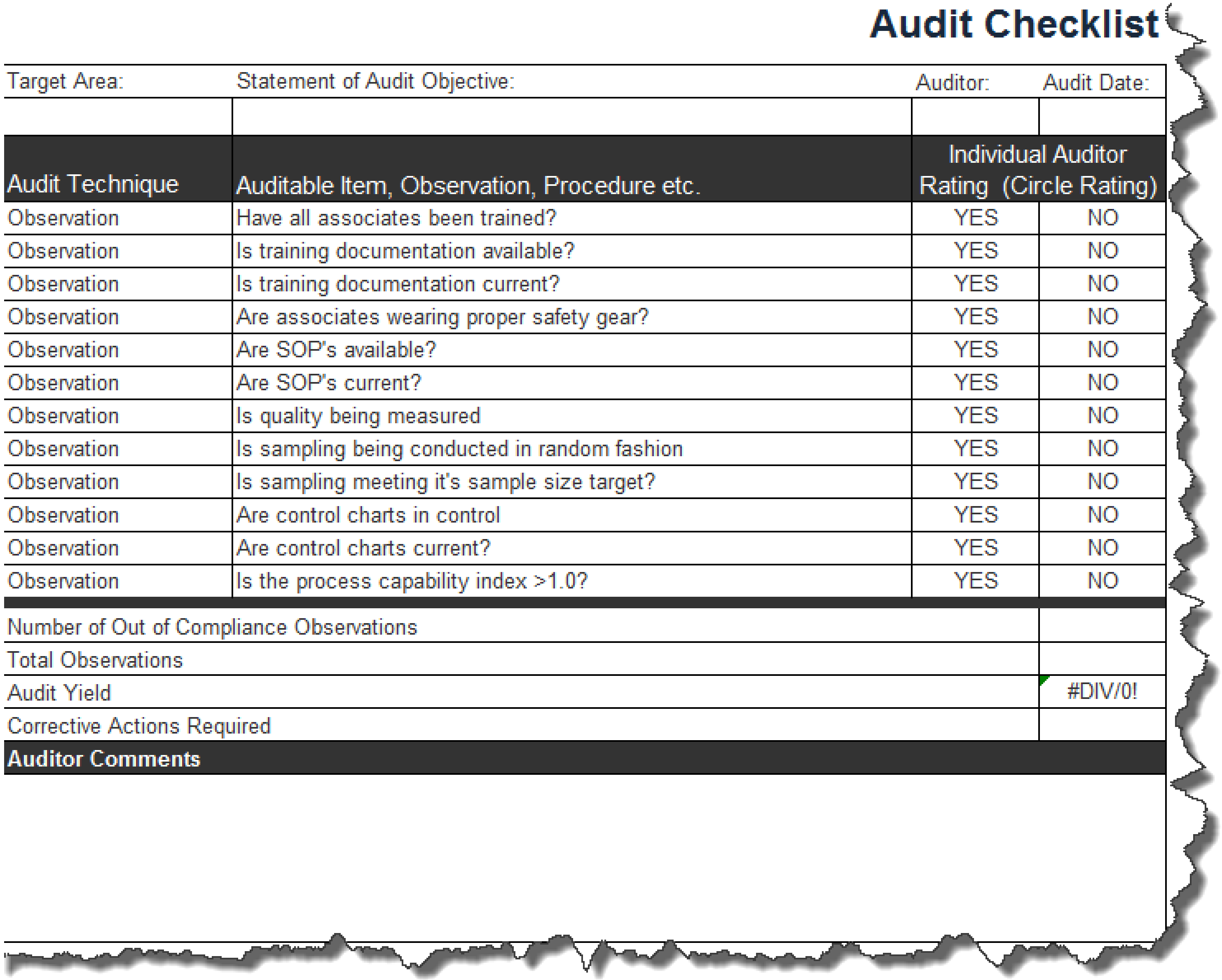
Audit Checklists

- Auditors should review the SOPs before preparing checklists or ensure that existing checklists properly reference SOPs.

- Audit checklists:
  - Serve as guides for identifying items to be examined
  - Are used in conjunction with understanding of the procedure
  - Ensure a well-defined audit scope
  - Identify needed facts during audits
  - Provide places to record gathered facts.

- Checklists should include:
  - A review of training records
  - A review of maintenance records
  - Questions or observations that focus on expected behaviors
  - Questions should be open-ended where possible
  - Definitive observations yes/no, true/false, present/absent, etc.

Audit Checklist Template



Audit Checklist

Target Area:	Statement of Audit Objective:	Auditor:	Audit Date:
Audit Technique	Auditable Item, Observation, Procedure etc.	Individual Auditor Rating (Circle Rating)	
Observation	Have all associates been trained?	YES	NO
Observation	Is training documentation available?	YES	NO
Observation	Is training documentation current?	YES	NO
Observation	Are associates wearing proper safety gear?	YES	NO
Observation	Are SOP's available?	YES	NO
Observation	Are SOP's current?	YES	NO
Observation	Is quality being measured	YES	NO
Observation	Is sampling being conducted in random fashion	YES	NO
Observation	Is sampling meeting it's sample size target?	YES	NO
Observation	Are control charts in control	YES	NO
Observation	Are control charts current?	YES	NO
Observation	Is the process capability index >1.0?	YES	NO
Number of Out of Compliance Observations			
Total Observations			
Audit Yield			#DIV/0!
Corrective Actions Required			
Auditor Comments			

5.3.3 Response Plan Elements

What is a Response Plan?

- A **response plan** should be a component of as many control plan elements as possible.
- Response plans are a management planning tool to describe corrective actions necessary in the event of out-of-control situations.
- There is never any guarantee that processes will always perform as designed. Therefore, it is wise to prepare for occasions when special causes are present.
- Response plans help us mitigate risks and, as already mentioned, should be part of several control plan elements.

Response Plan Elements

- Action triggers
  - When do we need to take actions to correct a problem or issue?
- Action recommendation
  - What activities are required in order to solve the problem in the process? The action recommended can be short-term (quick fix) or long-term (true process improvement).
- Action respondent
  - Who is responsible for taking actions?
- Action date
  - When did the actions happen?
- Action results
  - What actions have been taken?
  - When were actions taken?
  - What are the outcomes of the actions taken?

Response Plan Elements

Note the response plan element in this control plan template

Lean Six Sigma Control Plan

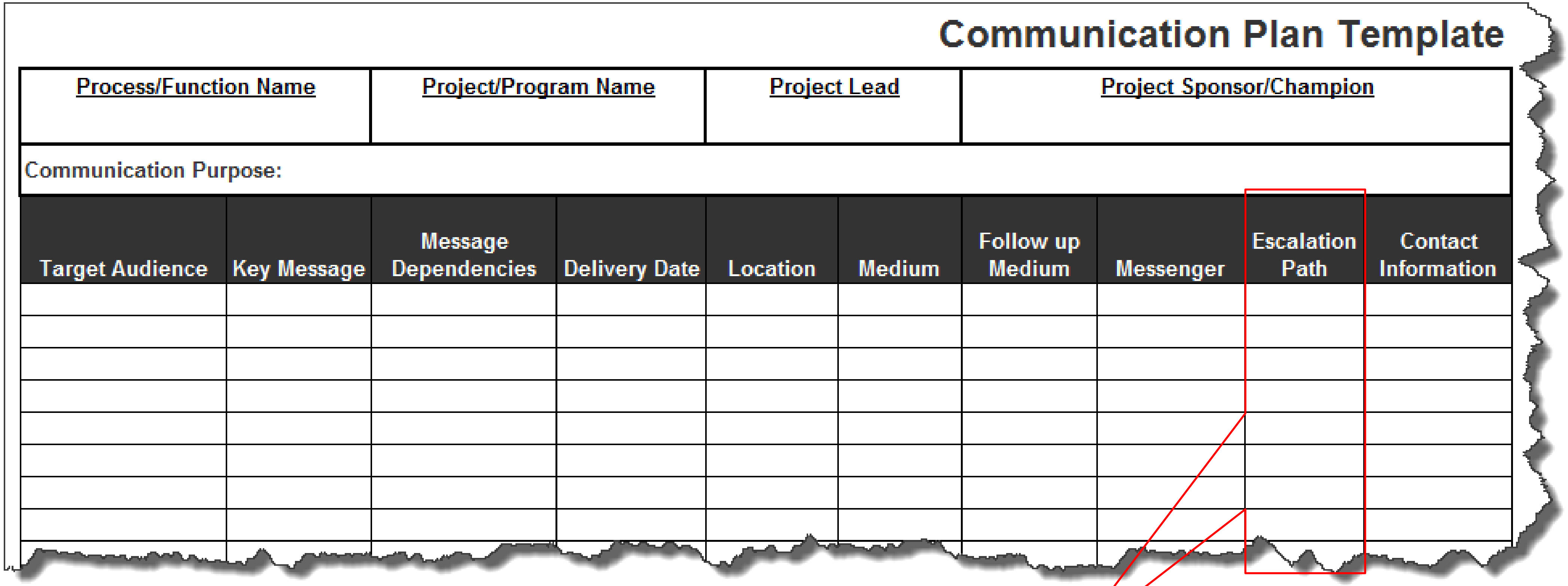
Process: ____________________
Customer: ____________________
Stakeholder: ____________________
Business: ____________________

Preparer: ____________________
Email: ____________________
Phone: ____________________
Owner: ____________________

Page: ________ of ____
Reference No: ____________________
Revision Date: ____________________
Approval: ____________________

Process	Process Step	CTQ/Metric	CTQ / Metric Equation	Specification/ Requirement		Measurement Method	Sample Size	Measure Frequency	Responsible for Metric	Link or Report Name	Corrective Action	Responsible for Action
				LSL	USL							

Response Plan Elements



Communication Plan Template

Process/Function Name	Project/Program Name	Project Lead	Project Sponsor/Champion

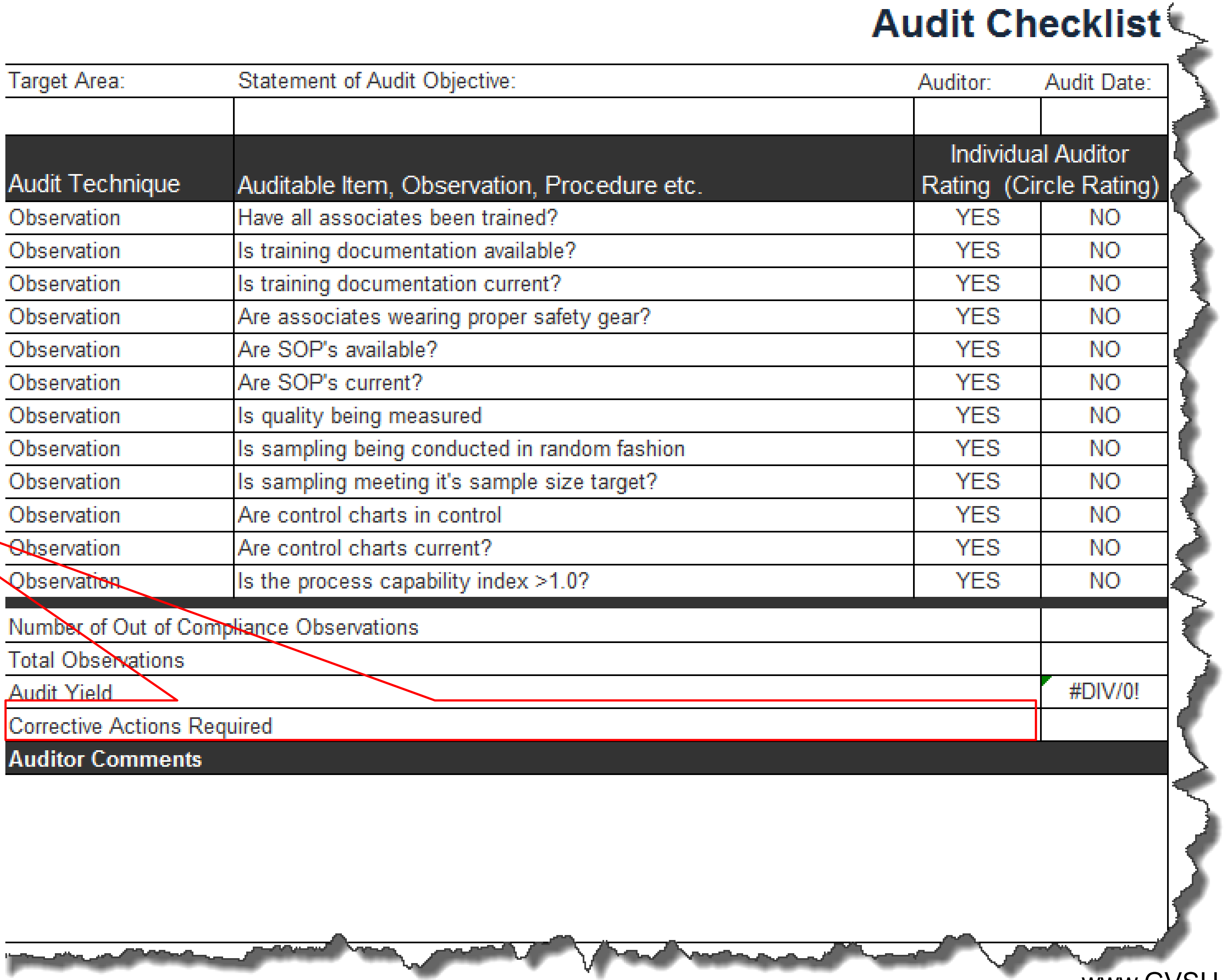
Communication Purpose:

Target Audience	Key Message	Message Dependencies	Delivery Date	Location	Medium	Follow up Medium	Messenger	Escalation Path	Contact Information

Note the response plan element in this communication plan template

Response Plan Elements

Note the response plan element in this audit checklist



Audit Checklist

Target Area:	Statement of Audit Objective:	Auditor:	Audit Date:
Audit Technique	**Auditable Item, Observation, Procedure etc.**	**Individual Auditor Rating (Circle Rating)**	
Observation	Have all associates been trained?	YES	NO
Observation	Is training documentation available?	YES	NO
Observation	Is training documentation current?	YES	NO
Observation	Are associates wearing proper safety gear?	YES	NO
Observation	Are SOP's available?	YES	NO
Observation	Are SOP's current?	YES	NO
Observation	Is quality being measured	YES	NO
Observation	Is sampling being conducted in random fashion	YES	NO
Observation	Is sampling meeting it's sample size target?	YES	NO
Observation	Are control charts in control	YES	NO
Observation	Are control charts current?	YES	NO
Observation	Is the process capability index >1.0?	YES	NO
Number of Out of Compliance Observations			
Total Observations			
Audit Yield			#DIV/0!
Corrective Actions Required			
Auditor Comments			

GRAND VALLEY STATE UNIVERSITY®
SCHOOL OF PUBLIC, NONPROFIT, AND
HEALTH ADMINISTRATION

Lean Six Sigma Black Belt Training

Featuring Examples from Minitab 18