

# What is Statistics?

## 🎯 Lesson Objectives

By the end of this lesson, you should be able to:

- Understand what statistics is
- Understand why statistics is important
- Differentiate between descriptive and inferential statistics
- Understand the concepts of population and sample
- Recognize how statistics is used in business and research



## 🧠 Why Do We Need Statistics?

As many of you know, I own **MyNetworkTraining.com**, where I provide networking and IT training to students around the world.

Imagine the following situation:

Last month:

- 1,000 active students

This month:

- 850 active students

This means I lost:

$$1,000 - 850 = 150 \text{ students}$$

Immediately, several questions come to mind:

- Why did students leave?
- Were students unhappy with the courses?
- Did a competitor offer lower prices?
- Did fewer students enroll because of summer holidays?
- Was it simply random variation?

Without data, I can only guess.

Statistics helps me stop guessing and start making decisions based on facts.

## 🔍 What is Statistics?

Statistics is the science of:

- Collecting data
- Organizing data
- Analyzing data
- Interpreting data
- Making decisions based on data

Think of statistics as a process:

Raw Data



Analysis



Information



Decision

The purpose of statistics is to transform raw numbers into useful information.

## **Statistics Is Everywhere**

Statistics is used in almost every field.

### **Business**

Examples:

- Customer retention
- Marketing campaigns
- Sales analysis
- Employee performance
- Product popularity

### **Education**

Examples:

- Student grades
- Course ratings
- Exam results
- Student satisfaction

### **Healthcare**

Examples:

- Recovery rates
- Treatment effectiveness
- Disease studies

### **Government**

Examples:

- Population growth
- Employment statistics
- Inflation rates
- Economic indicators

## **Example from MyNetworkTraining**

Suppose I want to know which training category is the most popular on my platform.

I collect the following information:

### **Category   Students**

MikroTik   500

UniFi   300

Cisco   150

OPNsense 50

### **What Can I Learn?**

- MikroTik is the most popular category
- UniFi is the second most popular category
- Cisco is the third most popular category
- OPNsense currently has the fewest students

Instead of reviewing thousands of enrollments manually, statistics helps me summarize the information quickly.

## **Two Main Types of Statistics**

Statistics is divided into two major branches:

**1** **Descriptive Statistics**

**2** **Inferential Statistics**

Everything we learn in this course belongs to one of these two categories.

### **Descriptive Statistics**

#### **Purpose**

Descriptive statistics helps us summarize and describe data.

#### **Main Question**

What happened?

#### **Examples**

- Mean
- Median
- Mode
- Range
- Standard Deviation

#### **Example**

Suppose one of my courses receives the following ratings:

5, 4, 5, 5, 3, 4, 5

Looking at these numbers individually does not tell me much.

Descriptive statistics helps me summarize the ratings and understand the overall performance of the course.

### **Inferential Statistics**

#### **Purpose**

Inferential statistics helps us make conclusions about a larger population.

#### **Main Question**

What can I conclude?

#### **Examples**

- Chi-Square
- Spearman's  $r_s$
- Paired t-Test
- Independent t-Test
- ANOVA

Inferential statistics helps determine whether observed results are likely real or simply caused by random chance.

### **Population and Sample**

Before learning inferential statistics, we must understand two important concepts.

#### **Population**

A population is the entire group we want to study.

Example:

All MyNetworkTraining students  
(5,000 students)

## Sample

A sample is a smaller group selected from the population.

Example:

300 students selected for a survey

A sample is used because it is often impossible to collect data from every individual in the population.

## The Most Important Diagram in Statistics

Population

(5,000 Students)



Sample

(300 Students)



Statistical Test



Conclusion

This process is the foundation of inferential statistics.

Almost every statistical test we learn later follows this exact model.

## Example: Student Satisfaction Survey

Suppose I send a survey to:

300 students

Results:

240 students satisfied

60 students not satisfied

This means:

$240 \div 300 = 80\%$

80% Satisfaction

The next question becomes:

**Can I use these 300 students to make conclusions about all 5,000 students?**

This is exactly why inferential statistics exists.

## Why Do We Need Statistical Tests?

Suppose I launch two marketing campaigns.

**Campaign A**

100 New Students

**Campaign B**

110 New Students

Question:

Is Campaign B really better?

Or:

Did random chance create the difference?  
Statistics helps answer this question objectively.

## What Will We Learn Later?

### Chi-Square

Used to determine whether two categories are related.

Example:

- Gender
- Preferred Course Category

### Spearman's $r_s$

Used to determine whether two rankings are related.

Example:

- Student Age Ranking
- Student Satisfaction Ranking

### Paired t-Test

Used to compare:

Before Training

vs

After Training

Example:

Did students improve after completing a course?

### Independent t-Test

Used to compare:

Group A

vs

Group B

Example:

Did students who followed Method A perform better than students who followed Method B?

## Why Learning Statistics Is Important

Statistics helps us:

- Understand data
- Make better business decisions
- Support research findings
- Analyze survey results
- Evaluate training effectiveness
- Complete dissertations and assignments
- Reduce decision-making based on assumptions

In today's world, organizations rely heavily on data.

Statistics provides the tools needed to transform that data into knowledge.

## Lesson Summary

Statistics is the science of collecting, organizing, analyzing, interpreting, and presenting data.

There are two main branches of statistics:

### **Descriptive Statistics**

Answers:

What happened?

Examples:

- Mean
- Median
- Mode
- Range

### **Inferential Statistics**

Answers:

What can I conclude?

Examples:

- Chi-Square
- Spearman's  $r_s$
- t-Tests

### **Key Takeaway**

If you remember only one thing from this lesson, remember this:

Descriptive Statistics



Describe Data

Inferential Statistics



Draw Conclusions

Everything we study in this course will be built on this foundation.

# Mean, Median, and Mode

## Lesson Objectives

By the end of this lesson, you should be able to:

- Understand Mean, Median, and Mode
- Calculate Mean, Median, and Mode
- Understand when each measure should be used
- Identify situations where Mean can be misleading
- Interpret data using Mean, Median, and Mode

## Why Do We Need Mean, Median, and Mode?

Suppose I launch a new MikroTik course on MyNetworkTraining.

After a few weeks, students provide the following ratings:

5, 4, 5, 5, 3

Looking at these ratings individually does not immediately tell me how well the course is performing.

I need a way to summarize the ratings.

This is where:

- Mean
- Median
- Mode

become useful.

These are called:

### **Measures of Central Tendency**

They help us identify the "center" of the data.

## What is Central Tendency?

Central Tendency tells us:

### **Where is the center of the data?**

Think of it as finding a value that best represents the entire dataset.

Example:

5, 4, 5, 5, 3

Instead of looking at all five numbers, we can summarize them using a

single representative value.

## Mean (Average)

### Definition

The Mean is the average value.

### Formula

Mean = Sum of Values

-----

Number of Values

### Example

Course Ratings:

5, 4, 5, 5, 3

Step 1:

$5 + 4 + 5 + 5 + 3 = 22$

Step 2:

$22 \div 5 = 4.4$

**Mean = 4.4**

### Visual Representation

Course ratings

Example ratings for a new course.

01223345

The Mean tells us that the average course rating is:

2.5 / 5

## Advantages of Mean

- Easy to calculate
- Uses all values in the dataset
- Most commonly used measure

## **Problem with Mean**

Mean can be heavily affected by unusual values.

These unusual values are called:

### **Outliers**

## **Outlier Example**

Suppose I sell the following number of courses per day:

10, 10, 10, 10, 100

### **Mean**

$$10 + 10 + 10 + 10 + 100 = 140$$

$$140 \div 5 = 28$$

$$\text{Mean} = 28$$

### **Is 28 a Good Representation?**

No.

Most days I sell:

10 courses

Only one day had:

100 courses

The Mean was distorted by the outlier.

## **Median**

### **Definition**

The Median is the middle value after sorting the data.

### **Example**

Course Ratings:

5, 4, 5, 5, 3

Step 1:

Sort the data:

3, 4, 5, 5, 5

Step 2:

Find the middle value:

3, 4, 5, 5, 5

↑

**Median = 5**

### **Why Median Is Important**

Median is much less affected by outliers.

### **Outlier Example Again**

Daily Sales:

10, 10, 10, 10, 100

Sorted:

10, 10, 10, 10, 100

Middle value:

10

**Median = 10**

This represents the data much better than Mean = 28.

### **When Should I Use Median?**

Use Median when:

- Outliers exist
- Income data is analyzed
- House prices are analyzed
- Data is heavily skewed

### **Mode**

#### **Definition**

Mode is the value that occurs most frequently.

## Example

Course Ratings:

5, 4, 5, 5, 3

Frequency:

### Rating Frequency

3      1

4      1

5      3

**Mode = 5**

Because it appears most often.

## Types of Mode

### Unimodal

One mode.

Example:

2, 3, 3, 4, 5

Mode:

3

### Bimodal

Two modes.

Example:

2, 2, 3, 3, 4

Modes:

2 and 3

### Multimodal

More than two modes.

Example:

1,1,2,2,3,3

Modes:

1,2,3

## Comparing Mean, Median, and Mode

Suppose a course receives:

5, 5, 5, 4, 1

**Mean**

$$20 \div 5 = 4$$

**Median**

1,4,5,5,5

Median = 5

**Mode**

5

Results:

**Measure Value**

Mean 4

Median 5

Mode 5

Notice that each measure tells a slightly different story.

## Excel Functions

**Mean**

=AVERAGE(A1:A5)

**Median**

=MEDIAN(A1:A5)

## Mode

=MODE.SNGL(A1:A5)

## Real MyNetworkTraining Example

Suppose a course receives:

5, 5, 5, 5, 4, 5, 5, 4, 5, 3

### What Can I Learn?

Mean:

Provides the average satisfaction.

Median:

Provides the middle rating.

Mode:

Provides the most common rating.

Together, these measures give me a much clearer picture of course quality.

## When to Use Each Measure

### Measure Best Use

Mean     General average

Median   Data with outliers

Mode     Most common value

## Lesson Summary

### Mean

Average Value

Uses all values.

Sensitive to outliers.

### Median

Middle Value  
Resistant to outliers.

## **Mode**

Most Frequent Value  
Shows the most common observation.

## **★ Key Takeaway**

Mean  
↓  
Average

Median  
↓  
Middle Value

Mode  
↓  
Most Frequent Value

These three measures form the foundation of descriptive statistics and are among the most frequently used statistical tools in business and research.

# Range and Standard Deviation

## Lesson Objectives

By the end of this lesson, you should be able to:

- Understand what Range is
- Calculate Range
- Understand what Standard Deviation is
- Understand data variability
- Interpret small and large Standard Deviations
- Use Excel to calculate Range and Standard Deviation

## Why Do We Need Range and Standard Deviation?

In Lesson 2, we learned:

- Mean
- Median
- Mode

These measures help us find the center of the data.

However, they do not tell us how spread out the data is.

Let's look at the following example.

### Group A Exam Scores

78, 79, 80, 81, 82

Mean:

80

### Group B Exam Scores

40, 60, 80, 100, 120

Mean:

80

## Important Observation

Both groups have:

Mean = 80

Does this mean the two groups performed similarly?

No.

Group A scores are very close together.

Group B scores are spread out over a much wider range.

We need a way to measure that spread.

This is why Range and Standard Deviation are important.



### What is Range?

#### Definition

Range measures the difference between:

Highest Value

-

Lowest Value



#### Range Formula

Range = Maximum Value - Minimum Value

#### Example 1

Exam Scores:

78, 79, 80, 81, 82

Maximum Score:

82

Minimum Score:

78

Range:

$82 - 78 = 4$

**Range = 4**

#### Example 2

Exam Scores:

40, 60, 80, 100, 120

Maximum Score:

120

Minimum Score:

40

Range:

$120 - 40 = 80$

**Range = 80**

### **Comparing the Two Groups**

#### **Group A**

Range:

4

Students achieved very similar scores.

#### **Group B**

Range:

80

Students achieved very different scores.

### **Advantages of Range**

- Very easy to calculate
- Easy to understand
- Gives a quick indication of variability

### **Disadvantages of Range**

Range only uses:

- Highest value
- Lowest value

It ignores all other values.

### **Example**

Dataset A

10, 10, 10, 10, 20

Range:

$$20 - 10 = 10$$

Dataset B

10, 12, 14, 16, 20

Range:

$$20 - 10 = 10$$

Both datasets have the same range.

However, the data is clearly different.  
Range alone is not enough.

## What is Standard Deviation?

### Definition

Standard Deviation measures how far the values are from the Mean.

In simple terms:

How spread out is the data?

### Easy Interpretation

#### Small Standard Deviation

Values are close to the Mean.

#### Large Standard Deviation

Values are far from the Mean.

### Example 1

Group A Scores

78, 79, 80, 81, 82

Mean:

80

The scores are very close to the Mean.

#### Small Standard Deviation

### Example 2

Group B Scores

40, 60, 80, 100, 120

Mean:

80

The scores are far from the Mean.

#### Large Standard Deviation

### Visual Comparison

#### Small Standard Deviation

78 79 80 81 82

↑

Mean

Most values stay close to the average.

### Large Standard Deviation

40   60   80   100   120  
          ↑  
          Mean

Values are spread out across a much wider range.

### MyNetworkTraining Example

Suppose I teach the same course to two groups.

#### Group A

Exam Results:

78,79,80,81,82

Mean:

80

Interpretation:

Students performed consistently.

#### Group B

Exam Results:

40,60,80,100,120

Mean:

80

Interpretation:

Student performance varies significantly.

Some students performed very poorly.

Some students performed exceptionally well.

### Why Businesses Care About Standard Deviation

Standard Deviation helps answer questions such as:

- Are student results consistent?
- Are customer satisfaction scores stable?

- Are monthly sales predictable?
- Is employee performance consistent?

Businesses often prefer consistency because it makes planning easier.

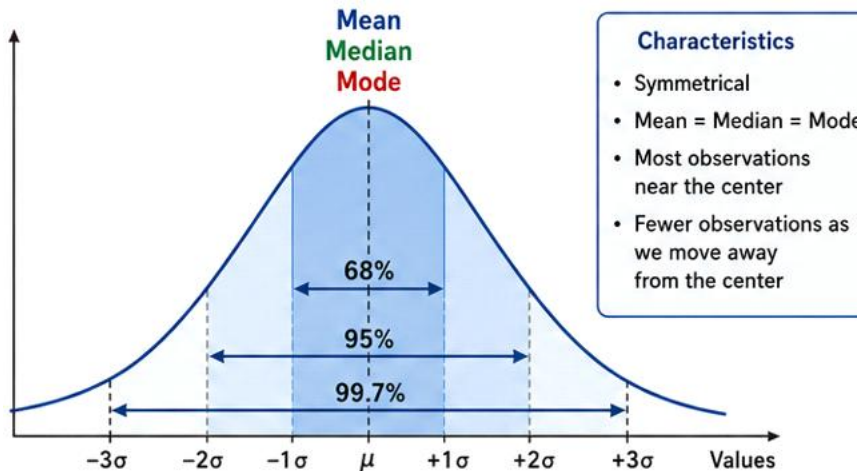
## Standard Deviation and the Bell Curve

Many datasets follow a pattern called:

### Normal Distribution

Also known as:

### Bell Curve



Most values are located near the Mean.

The further we move from the Mean, the fewer observations we find.

We will study the Bell Curve in detail in Lesson 4.

## Excel Exercise

### Dataset

#### Student Score

1	75
2	80
3	78
4	82
5	85
6	79
7	81
8	77

9            83

10           80

### **Calculate Range**

Maximum Score

=MAX(B2:B11)

Result:

85

Minimum Score

=MIN(B2:B11)

Result:

75

Range

85 - 75 = 10

### **Calculate Standard Deviation**

Formula

=STDEV.S(B2:B11)

Result:

Approximately

2.944

### **Interpretation**

A Standard Deviation of:

2.944

means that most student scores are relatively close to the Mean.

The students performed consistently.

### **When to Use Each Measure**

**Measure**

**Purpose**

Mean

Average value

Median

Middle value

Mode

Most common value

Range Overall spread  
Standard Deviation Variability around the Mean

## Lesson Summary

### Range

Maximum Value

-

Minimum Value

Measures the overall spread of the data.

### Standard Deviation

Measures:

How far values are from the Mean

### Small Standard Deviation

Consistent Data

### Large Standard Deviation

Highly Variable Data

## ★ Key Takeaway

Range



Measures Spread

Standard Deviation



Measures Variability Around The Mean

Two datasets can have exactly the same Mean but completely different Standard Deviations.

This is why Standard Deviation is one of the most important concepts in statistics.

# Normal Distribution, Skewness, and Kurtosis

## 🎯 Lesson Objectives

By the end of this lesson, you should be able to:

- Understand what a Normal Distribution is
- Recognize a Bell Curve
- Understand Positive Skewness
- Understand Negative Skewness
- Understand Kurtosis
- Interpret distribution shapes in business and research

## 🧠 Why Do We Need Distribution Shapes?

In Lesson 2, we learned:

- Mean
- Median
- Mode

In Lesson 3, we learned:

- Range
- Standard Deviation

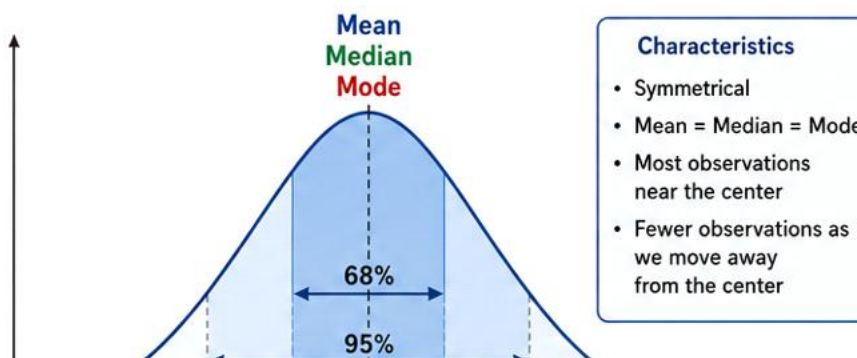
Now we want to understand something different:

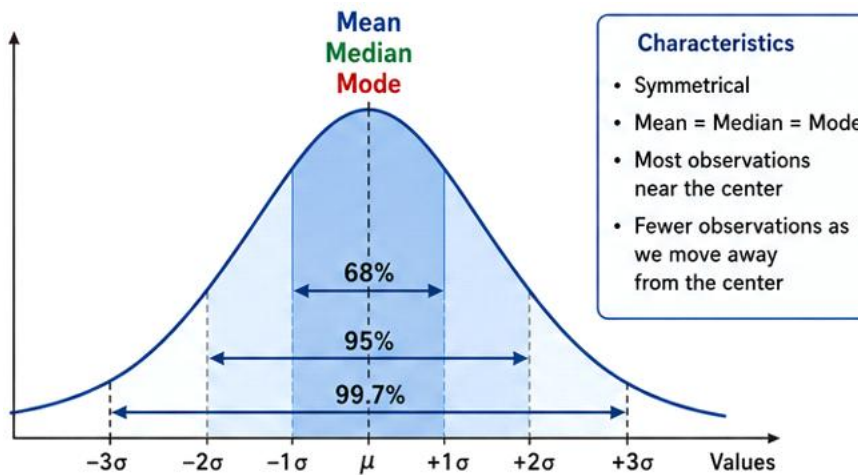
### **What shape does the data have?**

The shape of the data tells us a lot about how observations are distributed.

## 🔔 What is a Normal Distribution?

A Normal Distribution is the most common distribution in statistics. It is often called: Bell Curve





## Characteristics of a Normal Distribution

### Symmetrical

The left side is a mirror image of the right side.

### Mean = Median = Mode

All three measures are located at the center.

Mean = Median = Mode

### Most Values Are Near The Center

Values become less common as we move away from the Mean.

## MyNetworkTraining Example

Suppose 1,000 students complete an exam.

Most students score around:

70 - 80

A few students score:

90+

A few students score:

50-

The result often looks like a Bell Curve.

Most students are average.

Very few students are extremely weak or extremely strong.

## Why Is Normal Distribution Important?

Many statistical tests assume that data follows a Normal Distribution.

Examples:

- t-tests
- ANOVA
- Regression

Understanding the Bell Curve helps us understand many future topics.

## What is Skewness?

Skewness tells us:

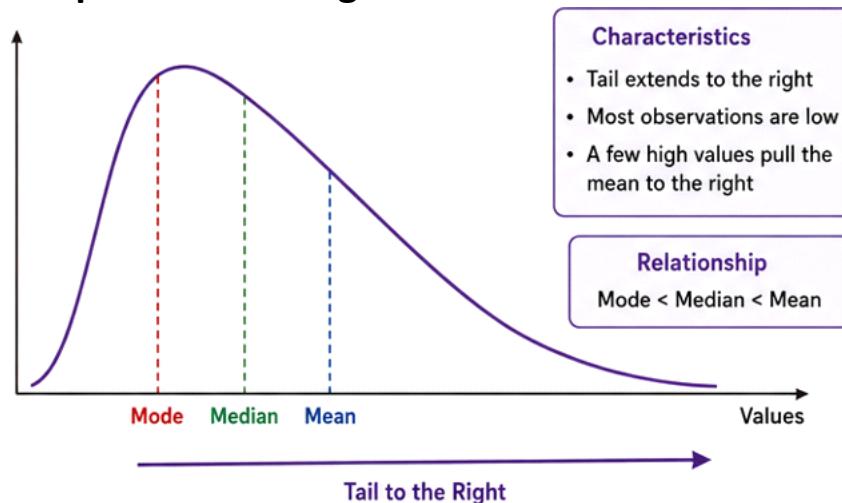
Is the data leaning to one side?

A distribution is skewed when it is not perfectly symmetrical.

## Positive Skewness (Right Skew)

Positive Skew means:

**Tail points to the right**



### Characteristics

- Most values are low
- A few very high values exist
- Tail extends to the right

## Example: Course Completion Time

Suppose students complete a course in:

2, 3, 3, 4, 4, 5, 20

Most students finish quickly.

One student takes much longer.

This creates:

**Positive Skewness**

## Real Business Example

Customer support response times:

Most tickets:

1-2 hours

A few difficult tickets:

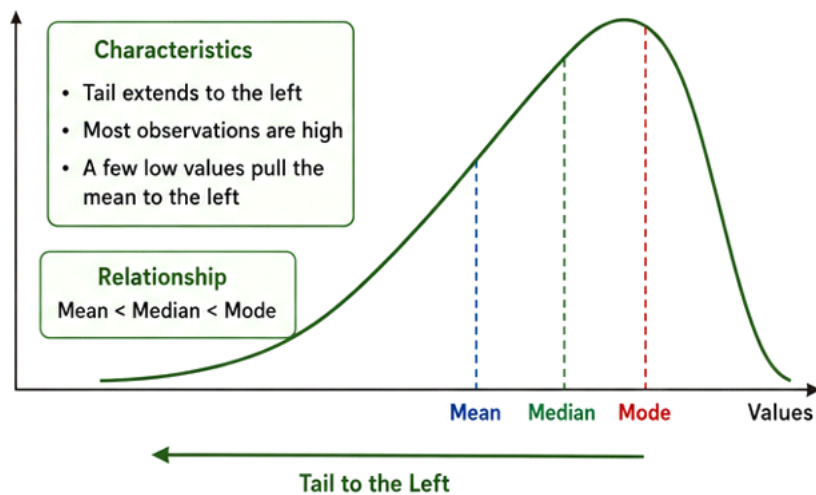
10-20 hours

This often creates a positively skewed distribution.

## Negative Skewness (Left Skew)

Negative Skew means:

**Tail points to the left**



## Characteristics

- Most values are high
- A few very low values exist
- Tail extends to the left

## Example: Course Ratings

Suppose a course receives ratings:

5,5,5,5,5,4,5,5,2

Most students love the course.

One student gives a very low rating.

This creates:

**Negative Skewness**

## Real MyNetworkTraining Example

Suppose one of my MikroTik courses receives:  
95% positive ratings.  
Only a few students provide poor ratings.  
The rating distribution becomes negatively skewed.

## 🎯 Easy Memory Trick

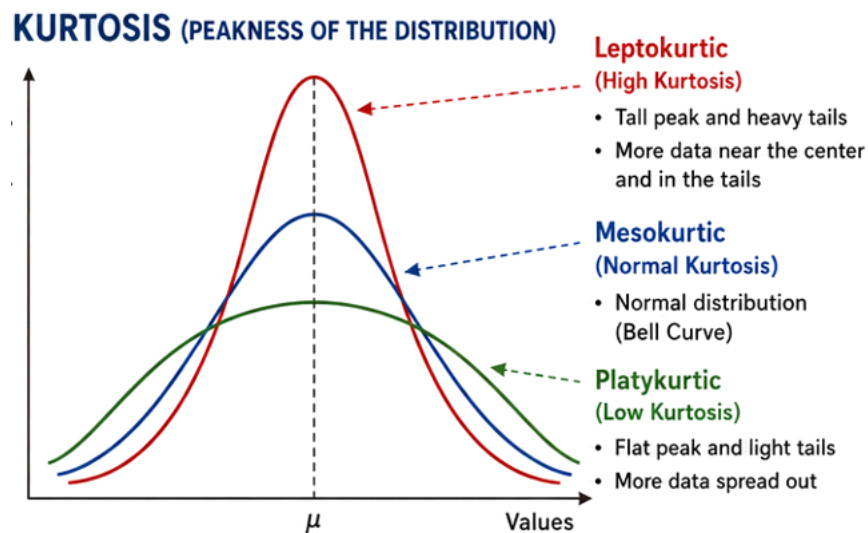
Positive Skew  
Tail Right

Negative Skew  
Tail Left

Always look at the tail.  
The tail determines the direction of skewness.

## △ What is Kurtosis?

Kurtosis measures:  
How peaked or flat a distribution is



**Normal Kurtosis**  
Normal peak.

**High Kurtosis**  
Very sharp peak.  
Many observations are concentrated near the center.

## Low Kurtosis

Flatter distribution.

Values are more spread out.

## Should We Worry About Kurtosis?

For business students:

**Not too much.**

In practice:

- Excel calculates it automatically
- SPSS calculates it automatically

The most important thing is to understand:

High Kurtosis

↓

Sharper Peak

Low Kurtosis

↓

Flatter Distribution

## Excel Exercise

Suppose we have the following exam scores:

### Student Score

1	65
2	70
3	75
4	80
5	85
6	75
7	78
8	82
9	76
10	79

## Calculate Skewness

Formula

=SKEW(B2:B11)

## Calculate Kurtosis

Formula

=KURT(B2:B11)

## Interpretation

Skewness Value	Interpretation
Between <b>-0.5 and +0.5</b>	Approximately symmetric (very close to normal)
Between <b>-1 and -0.5</b> or <b>+0.5 and +1</b>	Moderately skewed
Less than <b>-1</b> or greater than <b>+1</b>	Highly skewed

## Skewness Near 0

Approximately Normal Distribution

Kurtosis Value	Interpretation
Between <b>-1 and +1</b>	Approximately normal
Between <b>-2 and -1</b> or <b>+1 and +2</b>	Moderately different from normal
Less than <b>-2</b> or greater than <b>+2</b>	Substantially different from normal

## Normal Kurtosis

Normal peak.

## Lesson Summary

## **Normal Distribution**

Bell Curve

Most common distribution in statistics.

## **Positive Skewness**

Tail Right

Few large values pull the distribution to the right.

## **Negative Skewness**

Tail Left

Few small values pull the distribution to the left.

## **Kurtosis**

Measures:

How peaked or flat a distribution is

### **★ Key Takeaway**

Normal Distribution



Balanced

Positive Skew



Tail Right

Negative Skew



Tail Left

Kurtosis



Peak Height

Understanding the shape of data helps us choose the correct statistical tests and interpret results more accurately.



# Descriptive vs Inferential Statistics (In Depth)

## Lesson Objectives

By the end of this lesson, you should be able to:

- Differentiate between Descriptive and Inferential Statistics
- Understand when each type is used
- Identify real-world business applications
- Understand why inferential statistics is needed
- Recognize which statistical methods belong to each category
- Apply both concepts to research and business scenarios

## Why Is This Lesson Important?

In the previous lessons, we learned:

- Mean
- Median
- Mode
- Range
- Standard Deviation
- Normal Distribution
- Skewness
- Kurtosis

Now we need to understand:

**What are we actually doing with all these calculations?**

The answer is:

We are either:

**Describing Data**

or

**Drawing Conclusions**

These are the two major branches of statistics.

## The Two Branches of Statistics

Statistics can be divided into:

### **1** Descriptive Statistics

## **Inferential Statistics**

Everything we learn in this course belongs to one of these categories.

### **What is Descriptive Statistics?**

#### **Definition**

Descriptive Statistics summarizes and describes data.

It helps answer:

What happened?

### **Purpose of Descriptive Statistics**

Descriptive Statistics helps us:

- Organize data
- Summarize data
- Understand patterns
- Present information clearly

It does not make predictions.

It does not draw conclusions about a larger population.

### **MyNetworkTraining Example**

Suppose I review the ratings of one of my MikroTik courses.

Ratings:

5, 4, 5, 5, 3, 4, 5, 4, 5, 5

Using Descriptive Statistics I can calculate:

- Mean
- Median
- Mode
- Range
- Standard Deviation

These values help me understand the course ratings.

They describe the data I already have.

### **Common Descriptive Statistics**

#### **Measures of Central Tendency**

- Mean
- Median
- Mode

## Measures of Dispersion

- Range
- Standard Deviation

## Distribution Shape

- Normal Distribution
- Skewness
- Kurtosis

## What is Inferential Statistics?

### Definition

Inferential Statistics uses sample data to make conclusions about a larger population.

It helps answer:

What can I conclude?

## Why Do We Need Inferential Statistics?

Imagine MyNetworkTraining has:

5,000 students

Can I survey all 5,000 students?

Usually not.

It would take too much time and effort.

Instead, I select:

300 students

and collect information from them.

## Population vs Sample

### Population

Entire group of interest.

Example:

All 5,000 students

### Sample

Smaller group selected from the population.

Example:

300 surveyed students

## The Inferential Statistics Process

Population



Sample



Statistical Analysis



Conclusion

This process is the foundation of inferential statistics.

### Real Example

Suppose I survey:

300 students

Results:

240 satisfied

60 not satisfied

Satisfaction rate:

$$240 \div 300 = 80\%$$

Now I ask:

**Can I conclude that most of the 5,000 students are satisfied?**

This is an inferential statistics question.

### Descriptive vs Inferential Statistics

#### Descriptive Statistics

Focuses on:

Data You Have

#### Inferential Statistics

Focuses on:

Conclusions About Data You Don't Have

#### Example Comparison

Suppose I collect ratings from 300 students.

## Descriptive Question

What is the average rating?

Answer:

Mean = 4.5

## Inferential Question

Can I conclude that all students are satisfied?

Now I am moving from description to inference.



## Visual Comparison

DESCRIPTIVE STATISTICS

Data



Summary



Information

INFERENCE STATISTICS

Sample



Statistical Test



Population Conclusion



## Examples of Descriptive Statistics

**Technique**

**Purpose**

Mean

Average

Median

Middle Value

Mode

Most Common Value

Range

Spread

Standard Deviation

Variability

Skewness                      Distribution Shape

Kurtosis                      Peak Shape

### **Examples of Inferential Statistics**

<b>Technique</b>	<b>Purpose</b>
Chi-Square	Relationship between categories
Spearman's rs	Relationship between rankings
Paired t-Test	Before vs After comparison
Independent t-Test	Compare two groups
ANOVA	Compare multiple groups

### **Common Student Mistake**

Many students calculate:

Mean = 4.5

and immediately conclude:

All students are satisfied.

This is incorrect.

A Mean describes the sample.

It does not automatically describe the entire population.

Inferential Statistics is required before making broader conclusions.

### **Excel Exercise**

Suppose we have the following ratings:

#### **Student Rating**

1            5

2            4

3            5

4            5

5            3

6            4

7            5

8            5

9        4  
10       5

### **Descriptive Analysis**

Calculate Mean:

=AVERAGE(B2:B11)

Result:

4.5

This describes the ratings.

### **Inferential Question**

Can I conclude that all MyNetworkTraining students would rate the course similarly?

This question requires inferential statistics and cannot be answered using the Mean alone.

### **Why Researchers Use Inferential Statistics**

Researchers often cannot study:

- Every customer
- Every employee
- Every student
- Every citizen

Instead they:

1. Select a sample
2. Analyze the sample
3. Draw conclusions about the population

This is the foundation of modern research.

### **Lesson Summary**

#### **Descriptive Statistics**

Answers:

What happened?

Examples:

- Mean
- Median
- Mode

- Range
- Standard Deviation

Describes the data you already have.

## **Inferential Statistics**

Answers:

What can I conclude?

Examples:

- Chi-Square
- Spearman's  $r_s$
- t-Tests

Uses sample data to make conclusions about a population.

### **★ Key Takeaway**

Descriptive Statistics



Describe Data

Inferential Statistics



Draw Conclusions

Descriptive Statistics helps us understand the data we have.  
Inferential Statistics helps us make decisions about the larger population using sample data.

# Population and Sample (In Depth)

## Lesson Objectives

By the end of this lesson, you should be able to:

- Understand the difference between Population and Sample
- Understand why sampling is necessary
- Recognize the advantages of sampling
- Understand sampling error
- Understand representative samples
- Apply population and sample concepts in business and research

## Why Is This Lesson Important?

In Lesson 5, we learned that:

### **Inferential Statistics**

uses a sample to make conclusions about a larger population.

Before we can use statistical tests such as:

- Chi-Square
- Spearman's  $r_s$
- t-Tests

we must clearly understand:

Population

and

Sample

These are two of the most important concepts in statistics.

## What is a Population?

### **Definition**

A Population is the complete group that we want to study.

It contains all individuals, objects, or observations that are relevant to our research question.

## MyNetworkTraining Example

Suppose I want to know whether students are satisfied with my training platform.

I currently have:  
5,000 students  
The population is:  
All 5,000 students

## **Additional Examples of Populations**

### **University Research**

Population:  
All MBA students at a university

### **Business Research**

Population:  
All company customers

### **Employee Survey**

Population:  
All employees in the organization

## **What is a Sample?**

### **Definition**

A Sample is a smaller group selected from the population.  
The sample is used to represent the larger population.

## **MyNetworkTraining Example**

Population:  
5,000 students  
Sample:  
300 students

I collect information from only 300 students and use their responses to understand the entire population.

## **Population vs Sample**

### **Population**

Entire group  
Usually very large

### **Sample**

Smaller group  
Usually smaller

Difficult to study completely	Easier to study
More expensive	Less expensive
Takes more time	Faster to collect data

## **Why Don't We Study the Entire Population?**

In an ideal world, we would collect data from everyone. However, this is often impossible.

### **Example**

Suppose a company has:

250,000 customers

Surveying every customer would require:

- Significant time
- Significant cost
- Large resources

Instead, researchers select a sample.

## **Advantages of Sampling**

### **Saves Time**

Instead of surveying:

5,000 students

I may survey:

300 students

### **Saves Money**

Fewer surveys.

Less administration.

Lower research costs.

### **Easier Data Analysis**

Smaller datasets are easier to manage and analyze.

### **Faster Decision-Making**

Results can be obtained much sooner.

## **Important Requirement**

A sample must represent the population.

This is called:

### **Representativeness**

#### **Good Sample Example**

Population:

5,000 students

Composition:

- 70% Male
- 30% Female

Sample:

300 students

Composition:

- 70% Male
- 30% Female

This sample closely represents the population.

#### **Bad Sample Example**

Population:

5,000 students

Composition:

- 70% Male
- 30% Female

Sample:

300 students

Composition:

- 100% Male
- 0% Female

This sample does not represent the population properly.

The results may be misleading.

## **What is Sampling Error?**

Sampling Error is the difference between:

## Population Value and Sample Value

### **Example**

Suppose:

Actual population satisfaction:  
82%

Sample satisfaction:  
80%

Difference:  
2%

This difference is called Sampling Error.

### **Important Point**

Sampling Error is normal.

Almost every sample differs slightly from the population.

The goal is to make that difference as small as possible.

### **Larger Samples Usually Produce Better Results**

Imagine two surveys.

#### **Survey A**

Sample:  
20 students

#### **Survey B**

Sample:  
500 students

Which one is likely to represent the population better?

Usually:

500 students

because larger samples tend to reduce sampling error.

### **Population and Sample Diagram**

Population  
(5,000 Students)



Sample  
(300 Students)



Statistical Analysis



Conclusion

This is the foundation of inferential statistics.

### **Real Research Example**

Suppose I want to know:

Are students satisfied with my courses?

Population:

5,000 students

Sample:

300 students

Survey Result:

80% satisfied

Question:

Can I use this result to make conclusions about all students?

This is exactly what inferential statistics attempts to do.

### **Types of Sampling (Introduction)**

There are several ways to select a sample.

For now, we will introduce the two most common methods.

### **Random Sampling**

Every member of the population has an equal chance of being selected.

Example:

Randomly selecting 300 students from a list of 5,000 students.

### **Convenience Sampling**

Participants are selected because they are easy to reach.

Example:

Surveying only students currently online.

### **Which Method Is Better?**

Generally:

Random Sampling

because it reduces bias.

### **Common Student Mistake**

Students often assume:

Large Sample

=

Perfect Sample

This is not always true.

A large sample can still be biased if it is not representative of the population.

### **Excel Exercise**

Suppose we survey 20 students.

Results:

<b>Response</b>	<b>Count</b>
-----------------	--------------

Satisfied	16
-----------	----

Not Satisfied	4
---------------	---

### **Calculate Satisfaction Rate**

Formula:

=16/20

Result:

80%

## Discussion Question

Can we conclude that all MyNetworkTraining students are satisfied?

Answer:

Not yet.

We must consider:

- Sample size
- Sampling method
- Sampling error
- Representativeness

## Why Population and Sample Matter

Without understanding Population and Sample, we cannot correctly perform:

- Chi-Square Tests
- Spearman's  $r_s$
- t-Tests
- Research Projects
- Business Surveys

These concepts form the foundation of inferential statistics.

## Lesson Summary

### Population

Entire Group

Examples:

- All students
- All customers
- All employees

### Sample

Smaller Group Selected From Population

Used to represent the population.

### Sampling Error

Difference Between Population and Sample

### Representative Sample

A sample that accurately reflects the population.

★ **Key Takeaway**

Population



Sample



Statistical Analysis



Conclusion

Inferential Statistics relies on samples to help us understand larger populations.

The better the sample, the more reliable our conclusions will be.

# Variables and Measurement Levels

## Lesson Objectives

By the end of this lesson, you should be able to:

- Understand what a variable is
- Identify different types of variables
- Understand the four levels of measurement
- Differentiate between Nominal, Ordinal, Interval, and Ratio data
- Understand why measurement levels are important in statistics
- Choose appropriate statistical methods based on data type

## Why Is This Lesson Important?

Before we can perform statistical tests, we must understand:

What type of data do I have?

This is one of the most important questions in statistics.

Different types of data require different statistical techniques.

For example:

- Chi-Square works with certain types of data.
- Spearman's  $r_s$  works with other types.
- t-tests require different types of data.

If we do not understand our data, we may choose the wrong statistical test.

## What is a Variable?

### Definition

A Variable is any characteristic that can take different values.

In simple terms:

Something that can change

## MyNetworkTraining Examples

### Student Age

18, 25, 32, 45

Age varies from student to student.

Therefore:

Age = Variable

### **Course Rating**

1, 2, 3, 4, 5

Different students give different ratings.

Therefore:

Course Rating = Variable

### **Gender**

Male

Female

Different students belong to different categories.

Therefore:

Gender = Variable

### **Types of Variables**

There are two broad categories:

#### **Qualitative Variables**

Also called:

Categorical Variables

Examples:

- Gender
- Course Category
- Country
- Department

#### **Quantitative Variables**

Also called:

Numerical Variables

Examples:

- Age
- Salary
- Exam Score
- Revenue

### **Levels of Measurement**

Statistics classifies data into four levels:

1. Nominal
2. Ordinal
3. Interval
4. Ratio

Understanding these levels is essential because they determine which statistical tests can be used.

## **1** Nominal Level

### **Definition**

Nominal data consists of categories.

The categories have:

No order

### **Examples**

#### **Gender**

- Male
- Female

#### **Course Category**

- MikroTik
- Cisco
- UniFi
- OPNsense

#### **Country**

- Netherlands
- Germany
- France

### **Important Rule**

With Nominal data:

- ✗ Cannot rank
- ✗ Cannot average
- ✗ Cannot calculate meaningful differences

## Example

Male = 1

Female = 2

The numbers are labels only.

Female is not "greater" than Male.

## Key Point

Nominal Data = Categories Without Order

## 2 Ordinal Level

### Definition

Ordinal data consists of categories that have an order.

However:

Differences between categories are not equal

## Examples

### Customer Satisfaction

- Poor
- Fair
- Good
- Very Good
- Excellent

### Education Level

- High School
- Bachelor
- Master
- Doctorate

## Example

Student Satisfaction:

### Student Rating

- |   |           |
|---|-----------|
| A | Poor      |
| B | Good      |
| C | Excellent |

We know:

Excellent > Good > Poor

But we do not know the exact distance between them.

### Key Point

Ordinal Data = Categories With Order

### **3** Interval Level

#### Definition

Interval data consists of numbers where:

- Equal intervals exist
- Zero does not mean "nothing"

#### Example

##### Temperature in Celsius

10°C

20°C

30°C

Difference:

$$20 - 10 = 10$$

Difference:

$$30 - 20 = 10$$

Equal intervals exist.

#### Why Is It Not Ratio?

Because:

0°C

does not mean:

No Temperature

Therefore it is Interval data.

#### Additional Examples

- Calendar years
- IQ scores

### Key Point

Interval Data = Equal Intervals But No True Zero

## Ratio Level

### Definition

Ratio data has:

- Equal intervals
- A true zero

### Examples

#### Age

0 years

means:

No age

#### Revenue

€0

means:

No revenue

#### Number of Students

0 students

means:

No students

#### Exam Score

0 points

means:

No points earned

### Why Ratio Is Powerful

With Ratio data we can say:

40 is twice 20

because a true zero exists.

## Summary of Measurement Levels

Level	Ordered?	Equal Intervals?	True Zero?
-------	----------	------------------	------------

Nominal	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ordinal	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Interval	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Ratio	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

### MyNetworkTraining Examples

Variable	Measurement Level
Gender	Nominal
Course Category	Nominal
Course Satisfaction	Ordinal
Student Age	Ratio
Number of Students	Ratio
Monthly Revenue	Ratio
Completion Percentage	Ratio

### Why Measurement Levels Matter

Different data types require different statistical tests.

#### Example

##### Nominal Data

Gender  
Course Category  
Often analyzed using:  
Chi-Square

##### Ordinal Data

Satisfaction Ranking  
Often analyzed using:  
Spearman's  $r_s$

##### Ratio Data

Exam Scores  
Often analyzed using:

## t-Test

### **Common Student Mistake**

Students often assume:

All numbers are the same

This is incorrect.

Example:

Student Age = Ratio

and

Course Rating = Ordinal

Both contain numbers.

However, they represent different measurement levels.

### **Excel Exercise**

Create the following table:

<b>Student</b>	<b>Gender</b>	<b>Satisfaction</b>	<b>Age</b>
1	Male	Good	25
2	Female	Excellent	30
3	Male	Fair	22
4	Female	Very Good	28

### **Task**

Identify the measurement level of:

#### **Gender**

Answer:

Nominal

#### **Satisfaction**

Answer:

Ordinal

#### **Age**

Answer:

Ratio

## Lesson Summary

### **Nominal**

Categories

No Order

Examples:

- Gender
- Country
- Course Category

### **Ordinal**

Categories

With Order

Examples:

- Satisfaction Levels
- Education Levels

### **Interval**

Equal Intervals

No True Zero

Examples:

- Temperature
- IQ

### **Ratio**

Equal Intervals

True Zero

Examples:

- Age
- Revenue
- Number of Students
- Exam Scores

## Key Takeaway

Nominal



Categories

Ordinal



Ranked Categories

Interval



Equal Intervals

Ratio



Equal Intervals + True Zero

Understanding measurement levels is essential because they determine which statistical methods and tests can be used later in research and data analysis.

# Introduction to Statistical Testing

## Lesson Objectives

By the end of this lesson, you should be able to:

- Understand what a statistical test is
- Understand why statistical tests are used
- Understand how researchers move from a sample to a conclusion
- Recognize the most common statistical tests
- Understand when different statistical tests are used
- Prepare for advanced statistics courses

## Why Do We Need Statistical Tests?

In previous lessons, we learned:

- Mean
- Median
- Mode
- Range
- Standard Deviation
- Skewness
- Kurtosis
- Population and Sample
- Measurement Levels

These concepts help us understand data.

But eventually we want to answer questions such as:

**Is there a relationship between Gender and Course Preference?**

**Did students improve after training?**

**Are two groups different from each other?**

**Is there a relationship between Age and Satisfaction?**

To answer these questions, we use:

**Statistical Tests**

## What is a Statistical Test?

### Definition

A Statistical Test is a method used to analyze data and determine whether a conclusion can be made about a population.

In simple terms:

A statistical test helps us make decisions using data.

### **MyNetworkTraining Example**

Suppose I teach a MikroTik course.

Students take an exam:

#### **Before Training**

Average Score:

60

#### **After Training**

Average Score:

80

Question:

Did the course really improve student performance?

or

Was the difference caused by chance?

A statistical test helps answer this question.

### **The Statistical Testing Process**

Most statistical tests follow the same process.

Population



Sample



Collect Data



Apply Statistical Test



Draw Conclusion

This is the foundation of inferential statistics.

### **What Questions Can Statistical Tests Answer?**

Statistical tests help answer questions such as:

#### **Relationship Questions**

Example:

Is Gender related to Course Preference?

### **Difference Questions**

Example:

Did exam scores improve after training?

### **Comparison Questions**

Example:

Did Group A perform better than Group B?

### **Choosing the Correct Statistical Test**

One of the most important skills in statistics is selecting the correct test.

Different tests answer different questions.

### **Test 1: Chi-Square**

#### **Purpose**

Used to determine whether two categorical variables are related.

#### **Example**

Suppose I survey students.

Variables:

- Gender
- Preferred Course Category

Question:

Is Gender related to Preferred Course Category?

This is a job for:

#### **Chi-Square Test**

#### **Data Example**

##### **Student Gender Preferred Course**

1	Male	MikroTik
2	Female	Cisco
3	Male	MikroTik
4	Female	UniFi

Both variables are categorical.

Therefore:

Chi-Square

is appropriate.

## **Test 2: Spearman's Rank Correlation (Spearman's $r_s$ )**

### **Purpose**

Used to determine whether two rankings are related.

### **Example**

Suppose I want to know:

As student satisfaction increases,  
does course completion also increase?

### **Data Example**

#### **Student Satisfaction Rank Completion Rank**

A	1	2
B	2	3
C	3	4
D	4	5

Question:

Are the rankings related?

Use:

**Spearman's  $r_s$**

## **Test 3: Paired Sample t-Test**

### **Purpose**

Used when the same people are measured twice.

### **Example**

Students take an exam:

**Before Training**

and

**After Training**

Question:

Did students improve?

Use:

### **Paired Sample t-Test**

#### **Example Dataset**

**Student Before After**

1	60	80
2	65	85
3	55	75

The same students appear in both measurements.

### **Test 4: Independent Sample t-Test**

#### **Purpose**

Used to compare two separate groups.

#### **Example**

Group A:

Classroom Training

Group B:

Online Training

Question:

Did one group perform better?

Use:

### **Independent Sample t-Test**

#### **Example Dataset**

**Group A Group B**

80	70
85	72
82	75

Two different groups.

Therefore:

Independent Sample t-Test

## Quick Test Selection Guide

Research Question	Statistical Test
Relationship between categories	Chi-Square
Relationship between rankings	Spearman's rs
Before vs After comparison	Paired t-Test
Compare two groups	Independent t-Test

## Why Selecting the Correct Test Matters

Suppose I want to compare:

Exam Scores

between:

Group A

Group B

Using Chi-Square would be incorrect.

Using an Independent t-Test would be correct.

Choosing the wrong test can produce invalid conclusions.

## Easy Decision Tree

Do I have categories?



Yes



Chi-Square

-----  
Do I have rankings?



Yes



Spearman's rs

-----  
Do I have the same people measured twice?



Yes



Paired t-Test

-----

Do I have two separate groups?



Yes



Independent t-Test

## Real Business Examples

### **Marketing**

Question:

Is Gender related to Product Preference?

Use:

Chi-Square

### **Customer Satisfaction**

Question:

Does Satisfaction increase as Service Quality increases?

Use:

Spearman's rs

### **Employee Training**

Question:

Did employees improve after training?

Use:

Paired t-Test

### **Product Comparison**

Question:

Did Product A outperform Product B?

Use:

Independent t-Test



## Excel Exercise

Read each scenario and identify the correct statistical test.

### Scenario 1

Relationship between:

- Gender
- Preferred Course Category

Answer:

Chi-Square

### Scenario 2

Relationship between:

- Satisfaction Ranking
- Completion Ranking

Answer:

Spearman's rs

### Scenario 3

Exam Scores:

- Before Training
- After Training

Answer:

Paired Sample t-Test

### Scenario 4

Compare:

- Group A
- Group B

Answer:

Independent Sample t-Test



## What Comes Next?

Now that you understand:

- Descriptive Statistics
- Inferential Statistics
- Population and Sample

- Measurement Levels
- Statistical Testing

you are ready to study individual statistical tests in more detail.

Examples:

- Chi-Square Test
- Spearman's Rank Correlation
- Paired Sample t-Test
- Independent Sample t-Test

## Lesson Summary

### Statistical Tests

Help us:

Make Decisions Using Data

### Common Tests

#### Chi-Square

Relationship between categories.

#### Spearman's $r_s$

Relationship between rankings.

#### Paired Sample t-Test

Before vs After comparison.

#### Independent Sample t-Test

Comparison of two separate groups.

## ★ Key Takeaway

Research Question



Choose Correct Test



Analyze Data



Draw Conclusion

Understanding how to choose the correct statistical test is one of the most important skills in business research and data analysis.