

UNIT-I

INTRODUCTION TO SPSS

What is SPSS?

SPSS (Statistical Package for the Social Sciences) is a statistical software used for data analysis, visualization, and decision-making. It is widely used in research, business, healthcare, and social sciences for analyzing both small and large datasets efficiently.

Features of SPSS

- **Data Management:** Allows users to enter, edit, and manage datasets.
- **Statistical Analysis:** Performs both descriptive and inferential statistical analysis.
- **Graphical Representation:** Generates charts and graphs for data visualization.
- **Automation:** Uses syntax and scripts to automate repetitive tasks.
- **Compatibility:** Supports data import/export from Excel, CSV, databases, and other formats.

Getting Started with SPSS

When **SPSS** is opened, it has several key components:

1. **Data View** – Displays the dataset in a spreadsheet format where each row represents a case (observation) and each column represents a variable.
2. **Variable View** – Defines variable properties such as name, type, measurement scale (nominal, ordinal, scale), and labels.
3. **Output Window** – Displays results of statistical tests, tables, and graphs.

UNIT-I

4. **Syntax Editor** – Allows users to write and execute SPSS commands for advanced analysis.

Entering Data in SPSS

- Each row represents an observation (case).
- Each column represents a variable (question/feature).
- Data can be entered manually or imported from Excel, CSV, or databases.

Example: Dataset of Student Exam Scores

ID	Gender	Age	Exam Score
1	Male	22	78
2	Female	21	85
3	Male	23	90

Descriptive Statistics in SPSS

Descriptive statistics summarize and describe the characteristics of a dataset using:

1. **Frequency Distributions**
2. **Measures of Central Tendency** (Mean, Median, Mode)
3. **Measures of Dispersion** (Variance, Standard Deviation, Range, Interquartile Range)

UNIT-I**Frequency Distribution in SPSS**

A frequency distribution shows how often each value appears in a dataset. It is mainly used for categorical variables such as Gender, Education Level, etc.

Steps to Generate a Frequency Table in SPSS

1. Open SPSS and load the dataset.
2. Click on **Analyze** → **Descriptive Statistics** → **Frequencies**.
3. Select the variable(s) (e.g., Gender, Exam Score).
4. Click on **Statistics** and check **Mean, Median, Mode, Standard Deviation** if needed.
5. Click **Charts** → Select **Bar Chart** or **Pie Chart** for visualization.
6. Click **OK** → The results appear in the Output Window.

Example Output (Frequency Table for Gender):

Gender	Frequency	Percentage
Male	20	40%
Female	30	60%
Total	50	100%

UNIT-I

Example Chart: Bar Chart of Gender Distribution

▣ A bar chart will display the frequency of Male vs. Female responses.

Measures of Central Tendency in SPSS

Central tendency refers to the middle or average value of a dataset.

Mean (Arithmetic Average)

➤ **Definition:**

- The mean is the **sum of all values divided by the number of values.**

➤ **Formula:**
$$\text{Mean} = \frac{\sum X}{N}$$

where **X** = sum of values, **N** = total number of values.

SPSS Steps to Calculate Mean:

1. Click **Analyze** → **Descriptive Statistics** → **Descriptives.**
2. Select **Exam Score** variable.
3. Click **Options**, select **Mean**, then **OK.**

Median (Middle Value)

➤ **Definition:**

The median is the middle number when data is arranged in order.

- If **odd** number of values: Median = middle value.
- If **even** number of values: Median = average of two middle values.

UNIT-I

SPSS Steps to Calculate Median:

1. Click **Analyze** → **Descriptive Statistics** → **Frequencies**.
2. Select **Exam Score** variable.
3. Click **Statistics**, check **Median**, then **OK**.

Mode (Most Frequent Value)

- **Definition:** The mode is the value that appears most frequently.

SPSS Steps to Calculate Mode:

1. Click **Analyze** → **Descriptive Statistics** → **Frequencies**.
2. Select **Exam Score** variable.
3. Click **Statistics**, check **Mode**, then **OK**.

Measures of Dispersion (Variability) in SPSS

Dispersion shows how spread out the data is around the central value.

Range

➤ **Formula:** $Range = Max - Min$

➤ **SPSS Steps:**

- ❖ Click **Analyze** → **Descriptive Statistics** → **Frequencies**.
- ❖ Select **Minimum and Maximum** in **Statistics**.

UNIT-I**Variance****Definition:**

Measures how much data points deviate from the mean.

Formula:
$$\text{Variance} = \frac{\sum(X - \bar{X})^2}{N-1}$$

SPSS Steps:

Click **Analyze** → **Descriptive Statistics** → **Descriptives**.

Select **Variance** under Statistics.

Standard Deviation (SD)

Definition: Shows the average deviation from the mean.

Formula:
$$SD = \sqrt{\text{Variance}}$$

SPSS Steps:

❖ Click **Analyze** → **Descriptive Statistics** → **Descriptives**.

❖ Select **Standard Deviation** under Statistics.

Interquartile Range (IQR)**Definition:**

Measures spread between the 25th percentile (Q1) and 75th percentile (Q3).

UNIT-I

Formula: $IQR = Q_3 - Q_1$

SPSS Steps:

- ❖ Click **Analyze** → **Descriptive Statistics** → **Explore**.
- ❖ Select **IQR** under Statistics.

Data Visualization – Charts in SPSS

Bar Charts

- ❖ Used for **categorical data** (e.g., Gender distribution).
- ❖ Steps: Analyze → Descriptive Statistics → Frequencies → Charts → Bar Chart → OK.

Pie Charts

- ❖ Shows proportions in percentage.
- ❖ Steps: Analyze → Descriptive Statistics → Frequencies → Charts → Pie Chart → OK.

Histograms

- ❖ Shows the **distribution of numerical data**.
- ❖ Steps: Graphs → Legacy Dialogs → Histogram → Select Variable → OK.