Hilla University
College of Medical and Health Technology
Anthesis department
Computer Application
Third stage / First course

Lec. No. 1



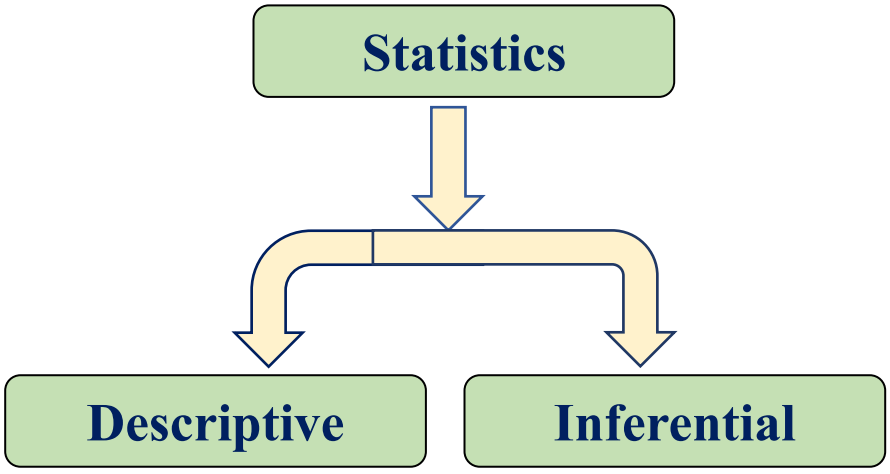
Assist. Lec. Zainab Hussein Ali
M.Sc. Information technology

Introduction to Statistics

What is Statistics?

Statistics is the science of collecting, analyzing, interpreting, presenting, and organizing data, it helps us make informed decisions based on data.

Types of Statistics:



- Descriptive Statistics: Summarizes and describes the features of a dataset (e.g., mean, median, mode, standard deviation).
- Inferential Statistics: Draws conclusions and makes predictions about a population based on a sample (e.g., hypothesis testing, confidence intervals).

Importance of Statistics

- ❖ Data Analysis: Enables effective data interpretation and decision-making.
- ❖ Research: Essential in scientific research for validating hypotheses and understanding trends.
- ❖ Real-world Applications: Used in various fields such as business, healthcare, economics, social sciences, and more.

*There are several programs and software tools commonly used in statistics for data analysis, visualization, and modeling, the most popular ones:

(SPSS, Excel, SAS, MATLAB)

What is SPSS?

SPSS (Statistical Package for the Social Sciences) is a powerful software tool widely used for statistical analysis in social science research, market research, healthcare, and various other fields. Developed in the late 1960s, SPSS has evolved into a comprehensive tool for data management, statistical analysis, and graphical representation.

Key Features

- User-Friendly Interface: SPSS offers a straightforward graphical interface that allows users to perform analyses without extensive programming knowledge.
- Data Management: It supports various data formats, including spreadsheets and databases, making it easy to import, manipulate, and manage data.
- Statistical Analysis: SPSS provides a wide range of statistical tests and procedures, from basic descriptive statistics to advanced inferential techniques.
- Visualization: Users can create various types of charts and graphs to visualize data effectively, aiding in the interpretation of results.
- Output Viewer: The software includes an output viewer for displaying results in a clear and organized manner.

Common Applications

- ❖ Social Sciences: Analyzing survey data, conducting experiments, and performing demographic studies.
- ❖ Healthcare: Evaluating clinical trial data and patient surveys.
- ❖ Market Research: Analyzing consumer behavior and preferences.
- ❖ Education: Conducting research studies and evaluating educational programs.

Basic Functions

- Descriptive Statistics: Calculate measures like mean, median, mode, standard deviation, and frequency distributions.
- Inferential Statistics: Perform hypothesis testing, correlation, regression analysis, and ANOVA.
- Data Transformation: Create new variables, recode existing ones, and handle missing data.
- Graphical Analysis: Generate bar charts, histograms, boxplots, and scatterplots to visualize data distributions and relationships.

Running SPSS

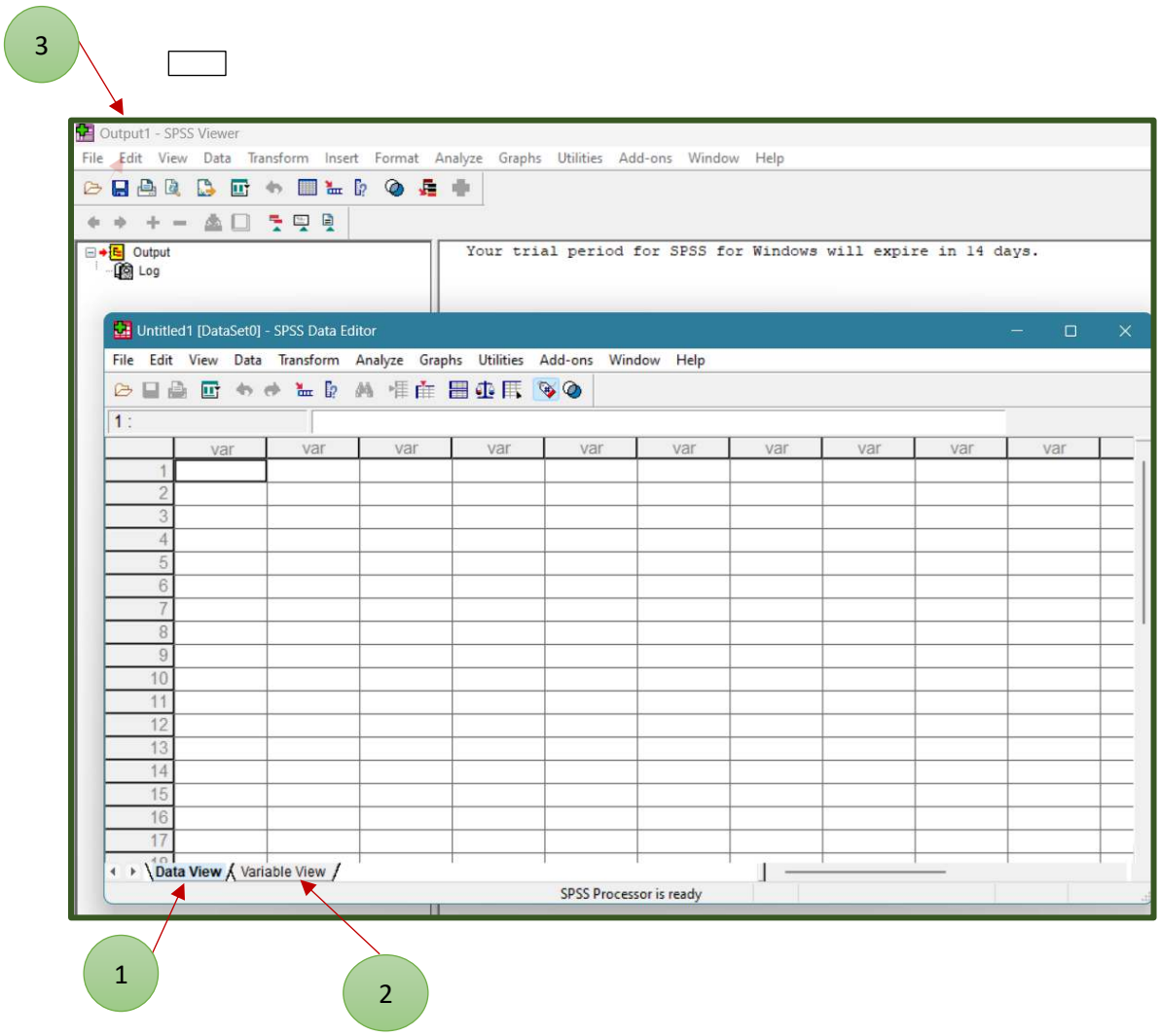
Depending on how the computer you are working on is structured, you can open SPSS in one of two ways.

1. If there is an SPSS shortcut like this on the desktop, simply put the cursor on it and double click the left mouse button.
2. Click the left mouse button on the Start menu, then put your cursor on Programs or All Programs and left click the mouse.

Layout of SPSS

The SPSS interface consists of several key components, each serving a specific function:

1. Data View: This is where you enter and view your data in a spreadsheet format. Each row represents a case (or participant), and each column represents a variable.
2. Variable View: This tab allows you to define the properties of each variable
3. Output Viewer: This window displays the results of your analyses, including tables, charts, and statistical outputs. You can edit and export results from here.



Variable View

It provides a grid-like layout where each row represents a single variable, and columns contain details about each variable.

Key Columns in Variable View:

1. **Name**:

   This is where you assign a name to each variable. Variable names should start with a letter, contain no spaces, and generally be concise (e.g., age, income, gender).

2. **Type:**

   Defines the data type for the variable. The common types are:

   - Numeric: For numbers, both integers and decimals.
   - String: For text variables (e.g., names).
   - Date: For date formats.



3. **Width:**

   Specifies the maximum number of characters for the variable (relevant for string variables) or the number of digits for numeric variables.

4. **Decimals:**

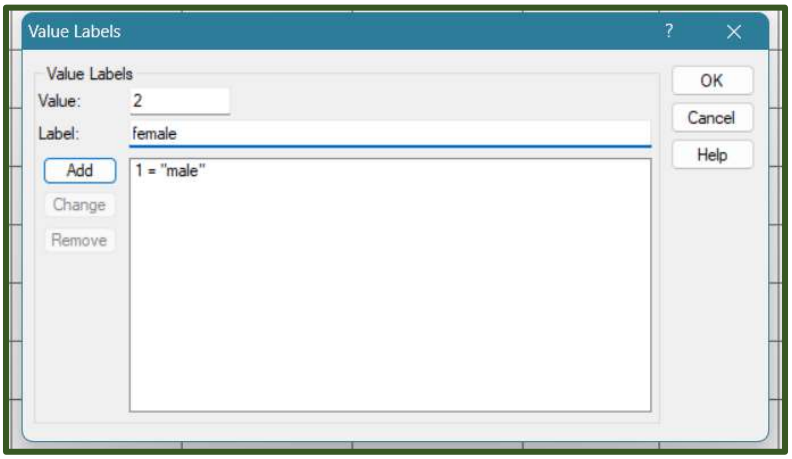
   Determines the number of decimal places for numeric variables. This is often set to 0 for whole numbers or 2 for currency.

5. **Label**

   Provides a descriptive label for the variable. Unlike the **Name** column, the **Label** can be longer and more descriptive (e.g., "Participant's Age in Years").

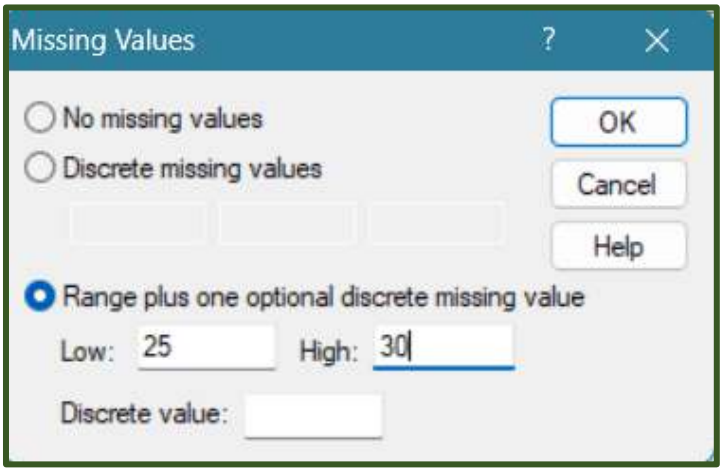
6. Values:

For categorical variables (e.g., gender or survey responses), you can define the values and their labels. For example, you can assign “1 = Male” and “2 = Female.” Click the cell in the **Values** column to open a dialog box for defining these labels.



7. Missing:

Specifies the values to be treated as missing in your data (e.g., “99” for a question left unanswered). This helps ensure missing data isn’t included in analyses accidentally.



8. **Columns:**

Determines the display width of the variable in the Data View. This setting doesn't affect the data itself, just how much space it takes up on screen.

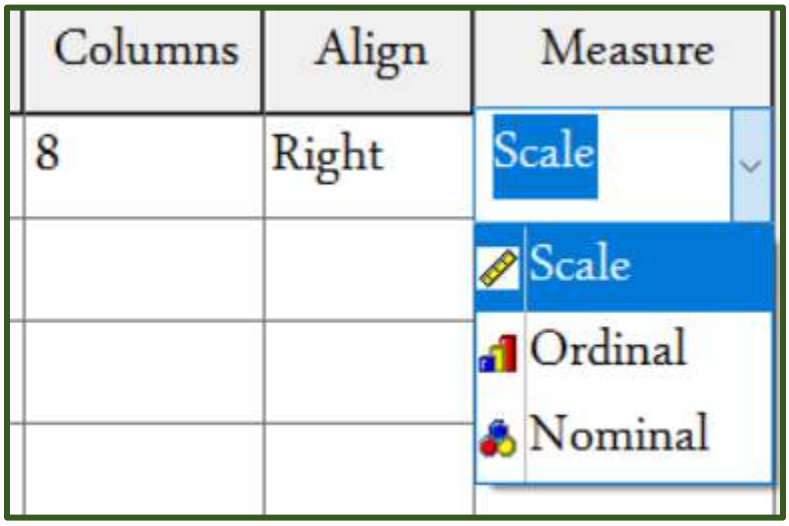
9. **Align:**

Specifies the alignment of the data in Data View (left, center, or right). This is mostly for visual organization.

10. **Measure**:

Defines the level of measurement for the variable:

- **Nominal**: Categories without a specific order (e.g., gender, eye color).
- **Ordinal**: Categories with a specific order (e.g., rating scales).
- **Scale**: Continuous data, often numeric (e.g., age, height).

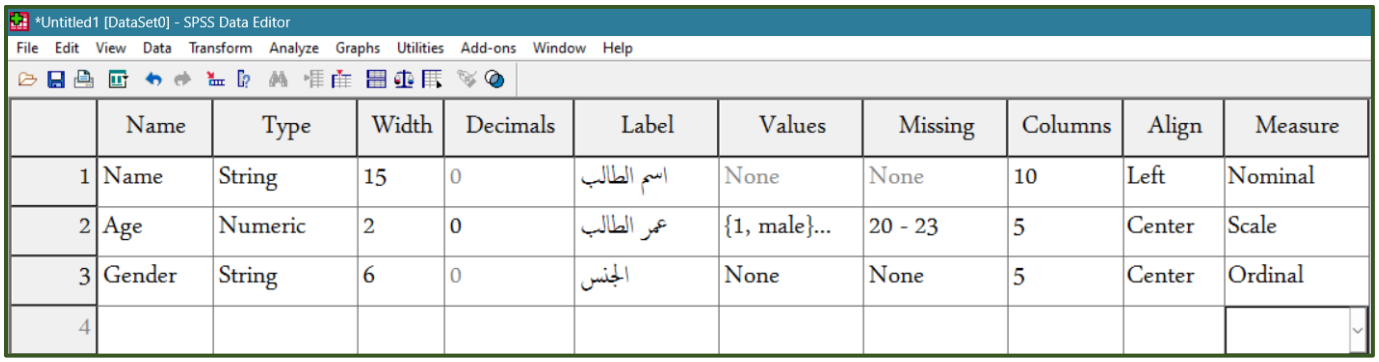


Example of Setting Up Variables in Variable View:

Name as "Name" (String, no decimals, Nominal).

Age as "Age" (Numeric, 0 decimals, Scale).

Gender with values (1 = Male, 2 = Female) as Ordinal.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Name	String	15	0	اسم الطالب	None	None	10	Left	Nominal
2	Age	Numeric	2	0	عمر الطالب	{1, male}...	20 - 23	5	Center	Scale
3	Gender	String	6	0	الجنس	None	None	5	Center	Ordinal
4										

Data View

It displays data in a spreadsheet-like grid where each row represents a single case (e.g., a survey response, participant, or observation), and each column represents a variable (e.g., age, gender, income).

Layout and Structure of Data View

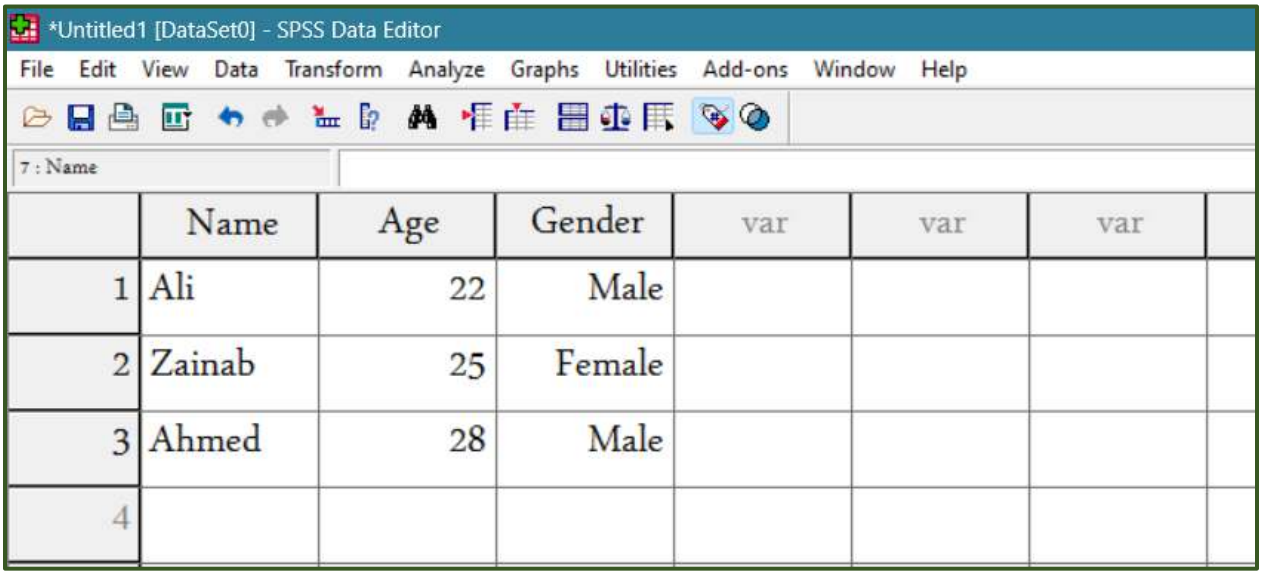
- Rows as Cases
- Columns as Variables
- Data Entry and Editing

Example of Setting Up Data in Data View:

Name as "Ali"

Age as "22"

Gender (1 = Male) as Ordinal.



	Name	Age	Gender	var	var	var	
1	Ali	22	Male				
2	Zainab	25	Female				
3	Ahmed	28	Male				
4							