

Self -study BMDS7010-02:

Getting to Know SPSS – Interface, Data Entry, Menus

Objectives:

- Understand the SPSS interface (Data View, Variable View).
- Manually enter a small, simple dataset.
- Define variables with appropriate names, types, and labels.
- Save an SPSS data file.

Relevance: This session focuses on entering patient demographics (age, gender, diagnosis code – simplified) and understanding data types (numeric, string, date) common in NHS datasets

Step-by-Step Instructions:

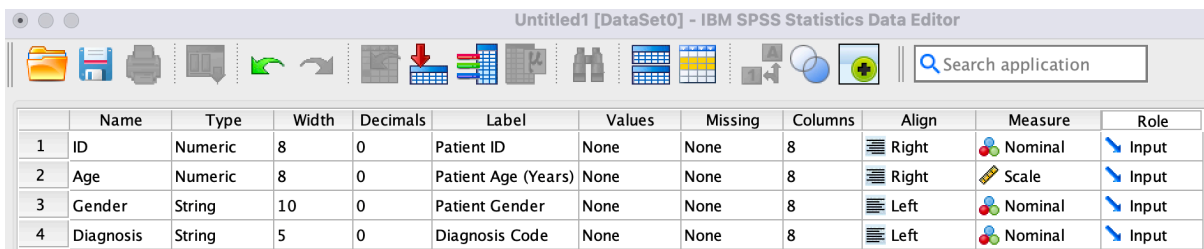
1. Launching SPSS and Exploring the Interface:

- When SPSS starts, you'll likely see a "Welcome" dialog or a blank spreadsheet. This is the **Data View**, where you'll see the actual data values.
- At the bottom left, find two tabs: **Data View** and **Variable View**. Click on **Variable View**. This is where you define the characteristics of each variable in your dataset.

2. Defining Variables (in Variable View):

- Create a few variables typical for patient demographics:
 - **ID:** A unique identifier for each patient.
 - **Age:** The patient's age in years.
 - **Gender:** The patient's gender.
 - **Diagnosis:** A simplified diagnosis code.
- For each variable, you'll need to specify several properties:
 - **Name:** A short, descriptive name (e.g., "ID", "Age", "Gender", "Diagnosis"). SPSS variable names have some restrictions (no spaces, start with a letter).
 - **Type:** This defines the kind of data.
 - ID: Choose "Numeric" (even if it contains leading zeros, we'll treat it as an identifier).
 - Age: Choose "Numeric".
 - Gender: Choose "String" (for text like "Male", "Female", "Other"). You could also use Numeric and assign value labels (1="Male", 2="Female"), but "String" is simpler for now.
 - Diagnosis: Choose "String" (e.g., "A1", "B2", "C3").
- - **Width:** The number of characters allowed for the variable. For ID, Age, make it reasonably large (e.g., 8). For Gender, make it large enough to accommodate the longest entry (e.g., 10). For Diagnosis Code, 5 might be enough.

- **Decimals:** For Age, set this to 0 (we want whole numbers). For ID, Gender, and Diagnosis (as strings), this setting doesn't apply. If decimals is set to = 2 for a numeric variable, values will display as 12.34 although the number may be stored as of 12.34567.
- **Label:** This is a longer, **more descriptive** name for the variable (e.g., "Patient ID", "Patient Age (Years)", "Patient Gender", "Diagnosis Code"). This makes your output easier to read.
- **Values:** We won't use this for now, but it's useful for coding categorical variables (e.g., 1 = "Male", 2 = "Female").
- **Missing:** We won't define missing values in this session.
- **Columns:** Affects the display width in Data View.
- **Align:** Affects the alignment of data in Data View (left, right, centre).
- **Measure:** This is important for SPSS to know how to treat the variable in analyses. Nominal variables represent categories or labels without any inherent order e.g. blood type A, B, AB, O). Scale variables represent numeric data where mathematical operations (addition, subtraction, averages, etc.) can be performed.
 - ID: Set to "Nominal".
 - Age: Set to "Scale" (for continuous, numerical data).
 - Gender: Set to "Nominal" (for categorical data).
 - Diagnosis: Set to "Nominal".
- **Role:** Usually leave this as "Input". Input – Used as an independent variable (predictor) in analysis. One can also have a “Target” variable used as a dependent variable (outcome) in analysis sometimes both.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ID	Numeric	8	0	Patient ID	None	None	8	Right	Nominal	Input
2	Age	Numeric	8	0	Patient Age (Years)	None	None	8	Right	Scale	Input
3	Gender	String	10	0	Patient Gender	None	None	8	Left	Nominal	Input
4	Diagnosis	String	5	0	Diagnosis Code	None	None	8	Left	Nominal	Input

3. Entering Data (in Data View):

- **Switch back to Data View** by clicking the tab at the bottom left.
- You'll now see the variable names you defined as column headers.
- Start entering data for a few hypothetical patients. For example:
 - Row 1: ID: 1001, Age: 45, Gender: Male, Diagnosis: A1
 - Row 2: ID: 1002, Age: 62, Gender: Female, Diagnosis: B2
 - Row 3: ID: 1003, Age: 38, Gender: Male, Diagnosis: A1
 - And so on. Create at least 5 rows of data.

4. Saving the Data File:

- Go to File > Save As.
- Choose a location to save your file (e.g., your Documents folder).
- Give the file a descriptive name (e.g., "PatientData.sav"). The ".sav" extension indicates an SPSS data file.
- Click "Save".
- The "Save As" dialog box.

To recap on the interface conventions when in Variable (defining) view:

- **Width:** These are suggestions; you can adjust them based on your expected data.
- **Decimals:** Not applicable (N/A) for string variables.
- **Values:** We are not using value labels in this initial session, though it is useful to know that option exists for the future. The Values column can be used to assign value labels to numeric variables. This is especially useful for categorical variables where numbers represent categories but do not have mathematical meaning. For example, a value of 1 means Male, a value of 2 means Female, a value of 3 means Other. It is also used in score banding for example, a Value of 1 means Very Dissatisfied, a Value of 2 means Dissatisfied etc.
- **Missing:** We are not defining missing values yet. However, understand that SPSS needs to know how to handle missing or special values in a dataset. Having this field can prevent errors in analysis by ensuring that missing data is properly accounted for.
- **Align:** This is a suggested setting and affects data view only. You just need to know it controls how data is aligned (justified) when displayed in Data View. The setting affects only the APPEARANCE of data in SPSS and does not impact calculations or analyses.
- **Measure :** This is an important setting. It lets SPSS know how to deal with a variable during analysis. What is in the Measure column defines a variable's level of measurement: **Nominal** (categorical with no order, e.g. gender), **Ordinal** (ordered categories, e.g. satisfaction levels), or **Scale** (numeric and measurable, e.g. age or income). Choosing the correct type ensures appropriate statistical analysis and prevents errors.

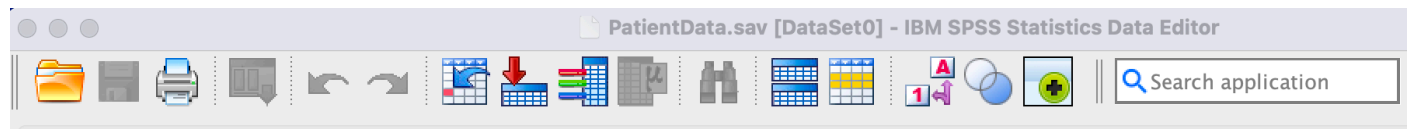
=== This concludes the sample data entry and saving task ===

Before finishing, and while SPSS is open in the Data View tab, try some of the menus across the interface:

This is only to make you aware that they are there!



Also notice the set of icons:



The best way of learning the main functions is step-by-step.

You can however find help on any menu item by doing one of the following:

Use the Help Menu – Click Help > Topics to access SPSS's built-in user guide.

Use the "Help" Button – Many dialog boxes have a Help button that opens relevant documentation for that feature.

Tooltip Hover – Hover over some menu items or icons to see brief descriptions in tooltips.

These options provide explanations, examples, and guidance on using SPSS features effectively.

Although not recommended at this point (it may overwhelm you) there is a YouTube tutorial that walks through each menu and icon – you could revisit this when you are more familiar with the interface.

<https://youtu.be/HInZEF20WuY?>