



5. Introduction to SPSS

Dr. Dr. Albulene Grajcevci

Today's agenda

- Measurement scales
- Introduction to SPSS
- Coding variables in SPSS
- EXCEL to SPSS
- Missing values
- Activities
 - Recode variables
 - Compute new variable

Measurement levels

- Nominal
 - Only used to classify data, no value assigned
- Ordinal
 - Implies an ordered relationship
- Ratio (continuous)
 - Equal intervals, but including a 0
- Interval/Scale (continuous)
 - Equal intervals

Measurement levels

- Nominal
 - Gender, location, hair color
- Ordinal
 - Income, education level
- Ratio
 - Temperature, points in an exam
- Interval/Scale
 - agreement, preference

Activity 1

- Type of movie you prefer
- The quality of service you receive at a restaurant
- Amount of protein in your drink
- Calories in a cookie
- Your score in an exam
- Time it takes to finish an exam

SPSS: INTRO

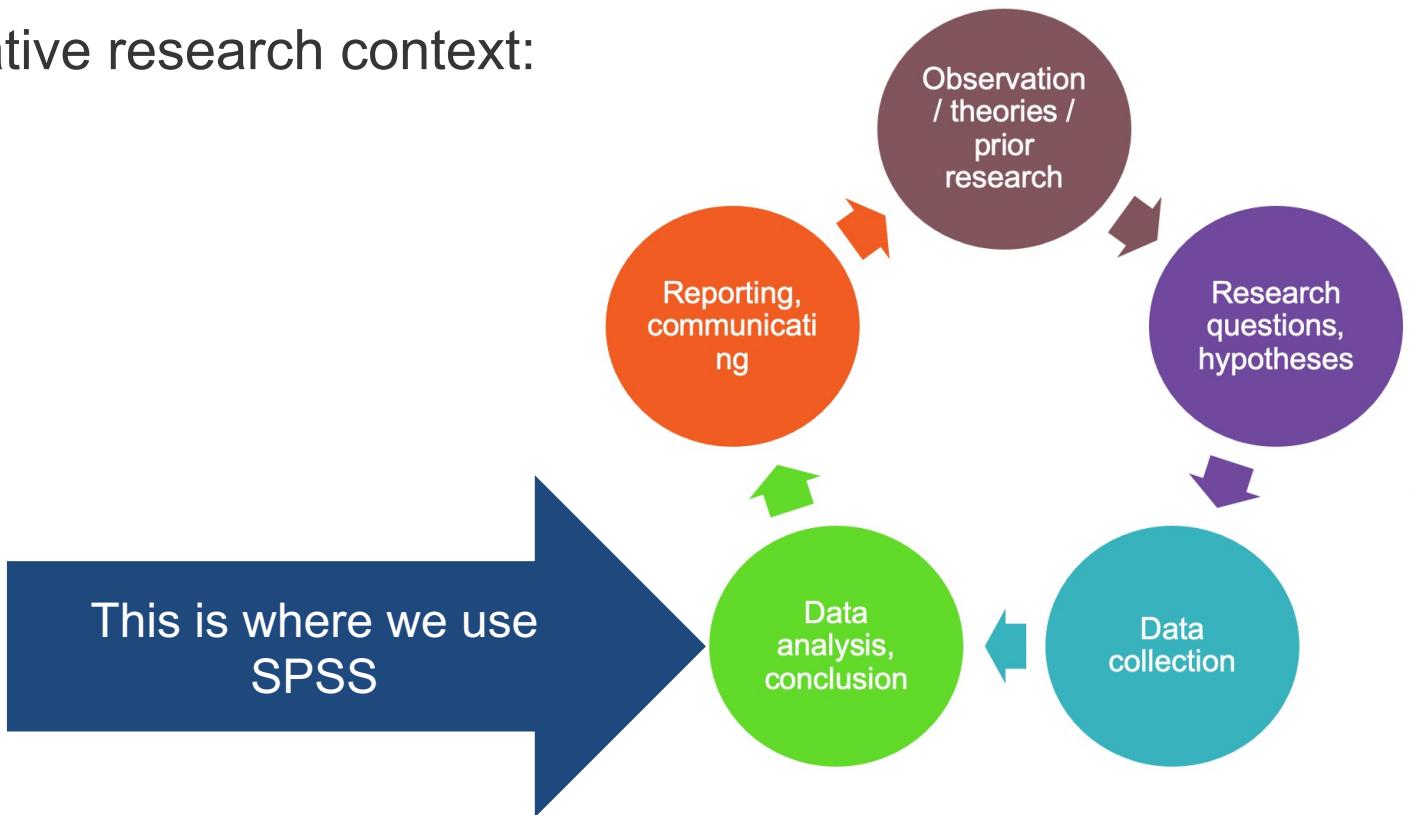
What is SPSS Statistics?

- ◊ Statistical Package for the Social Sciences
- ◊ Officially named "IBM SPSS Statistics"
- ◊ Current version: 27 (Windows / macOS / Linux)



What is SPSS Statistics used for?

- ◊ Manipulation and statistical analysis of survey data
- ◊ In a quantitative research context:



Adapted from the slides from Claudio Alves da Silva, 2020

SPSS: SPSS ENVIRONMENT

SPSS Environment

Each SPSS window is associated with specific tasks and types of SPSS files.

The screenshot shows the SPSS Data Editor window with the title bar "SPSS Statistics" and "SPSS-Workshop.sav [DataSet2] - IBM SPSS Statistics Data Editor". The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Extensions, Window, and Help. The toolbar contains various icons for file operations and data analysis. The main area displays a data grid with 16 rows and 12 columns. The columns are labeled: F1_Media, F2_Ind_Emo, Exp_Condition, Gender, Age, Res_pre, Res_post, Mot_01, Mot_02, and Mot_05. The first row is highlighted in blue. A status bar at the bottom indicates "IBM SPSS Statistics Processor is ready" and "Unicode:ON".

Data Editor
window:
Data View

Input and see all your collected
data

The screenshot shows the SPSS Data Editor window with the title bar "SPSS Statistics" and "SPSS-Workshop.sav [DataSet2] - IBM SPSS Statistics Data Editor". The menu bar and toolbar are identical to the Data View window. The main area displays a table of variable attributes for 19 variables. The columns are: Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, and M. The first few rows are: 1. F1_Media, Numeric, 8, 0, Factor 1: Learni..., {1, Tablet}..., None, 8, Right, Nc; 2. F2_Ind_Emo, Numeric, 8, 0, Factor 2: Induc..., {1, Joy}..., None, 8, Right, Nc; 3. Exp_Condit..., Numeric, 8, 0, Experimental C..., {1, Tablet+J...}, None, 8, Right, Nc; 4. Gender, Numeric, 8, 0, Gender, {1, Male}..., None, 8, Right, Nc; 5. Age, Numeric, 8, 0, Age, None, None, 8, Right, Sc. The status bar at the bottom indicates "IBM SPSS Statistics Processor is ready" and "Unicode:ON".

Data Editor
window:
Variable View

Overview of your variables and
attributes (name, scale level
(measure, labels, values, etc.)

SPSS Environment: Data Editor window

- ❖ Default window
- ❖ Displayed in spreadsheet format
- ❖ Contains 2 views: Data and Variable
- ❖ Data View: columns represent variables, rows represent cases
- ❖ Variable View: Shows information about variables present in the open data (but not the data itself)

SPSS Environment

Each SPSS window is associated with specific tasks and types of SPSS files.

The screenshot shows the SPSS Statistics Viewer window with the following content:

T-Test

Group Statistics

Gender	N	Mean	Std. Deviation	Std. Error Mean
Result Posttest	Male	10	47.2000	17.67484
	Female	6	80.1667	11.08903
				4.52708

Independent Samples Test

F	Sig.	t	df	t-test f			
				Sig. (2-tailed)	M Diff		
Result Posttest	Equal variances assumed	3.091	.101	-4.081	14	.001	-.32
	Equal variances not assumed			-4.583	13.908	<.001	-.32

Independent Samples Effect Sizes

Standardizer	Point Estimate	95% Confidence Interval Lower	Upper		
Result Posttest	Cohen's d	15.64434	-2.107	-3.357	-.810
	Hedges' correction	16.54981	-1.992	-3.174	-.766
	Glass's delta	11.08903	-2.973	-4.998	-.898

a. The denominator used in estimating the effect sizes.

Viewer
window

Executed syntax
Results from your analyses

The screenshot shows the SPSS Syntax Editor window with the following content:

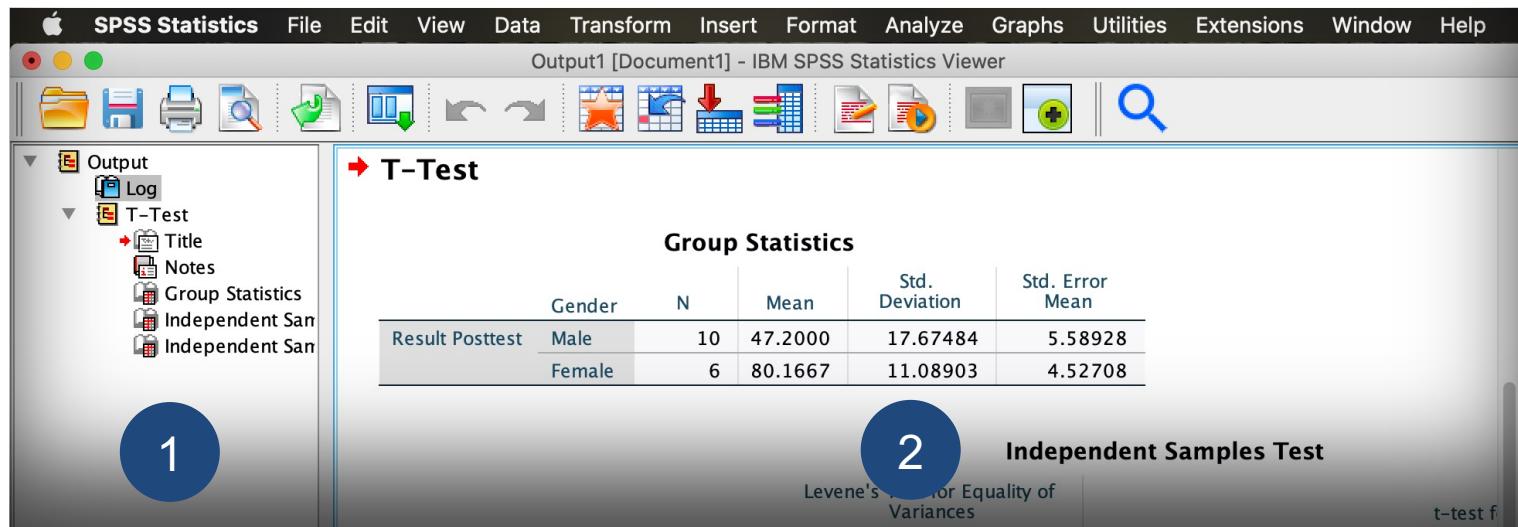
```
T-TEST GROUPS=Gender(1 2)
/MISSING=ANALYSIS
/VARIABLES=Res_post
/ES DISPLAY(TRUE)
/CRITERIA=C(.95).
```

Syntax Editor
window:
Variable View

Paste, edit, debug and run
syntax commands

SPSS Environment: (Output) Viewer window

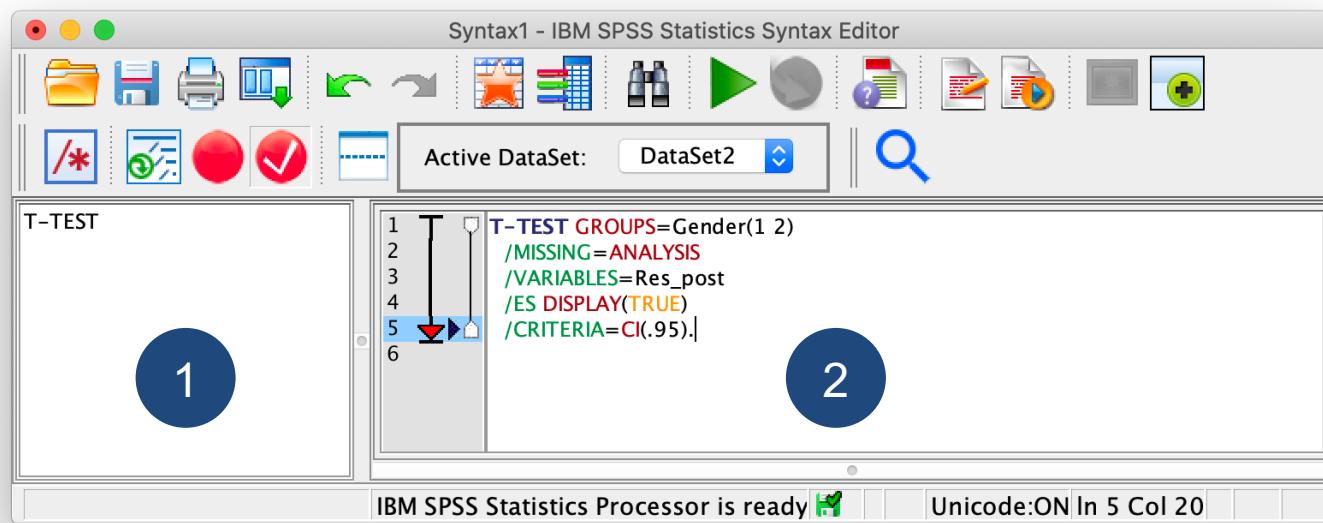
- ◆ Opens automatically
- ◆ Displays a log and output (syntax and results) of the actions taken as well as the result of statistical analysis
- ◆ Contains the outline of the content in the viewer (1) and the actual output (2)
- ◆ Contents can be modified and saved as a viewer file (*.spv) and other formats (e.g.: *.xlsx, *.doc, *.png)



Adapted from the slides from Claudio Alves da Silva, 2020

SPSS Environment: Syntax Editor window

- ◊ SPSS syntax is a programming language (alternative to menus)
- ◊ Users can write, debug and run
- ◊ Shows an outline of commands (1) and the syntax editor (2)
- ◊ Can be saved as an *.sps file
- ◊ To open a syntax window, go to: File > New > Syntax
- ◊ To run syntax code click the “Run Selection”  button in the Syntax Editor window



Adapted from the slides from Claudio Alves da Silva, 2020

SPSS Environment

SPSS affordances:

- ❖ It has a user interface → easier to learn
- ❖ Can be used for presenting results, e.g. result tables using APA formatting
- ❖ (Output) Viewer window: shows the syntax of every analysis/action. That syntax can be copied and then pasted in a Syntax Editor window
- ❖ Inclusion of add-ons (called macros) for more sophisticated analyses. Example PROCESS macro:
<https://www.processmacro.org/download.html>

SPSS Environment

SPSS disadvantages:

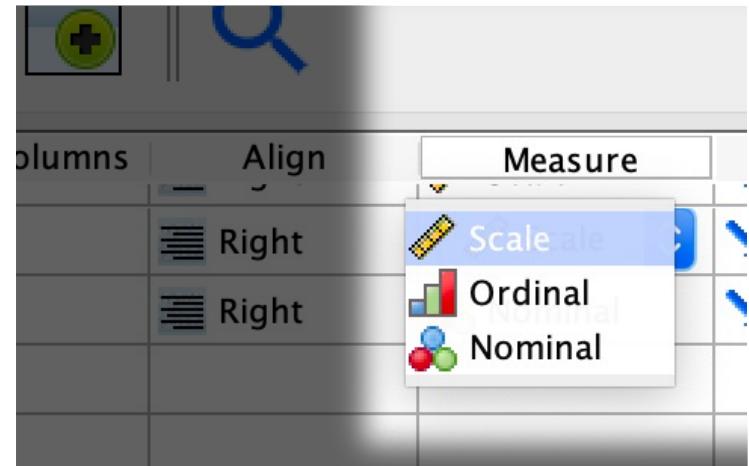
1. Not free in terms of freedom to use and modify the code (i.e. not open-source)
2. Not free in terms of money (i.e. you need pay/acquire a user license)
3. Related to 1, latest trends in statistics and plotting take a while to appear on SPSS

SPSS: SCALE LEVELS

Scale levels

Categories:

1. Nominal: for nominal data
2. Ordinal: for ordinal data
3. Scale: for interval and ratio data



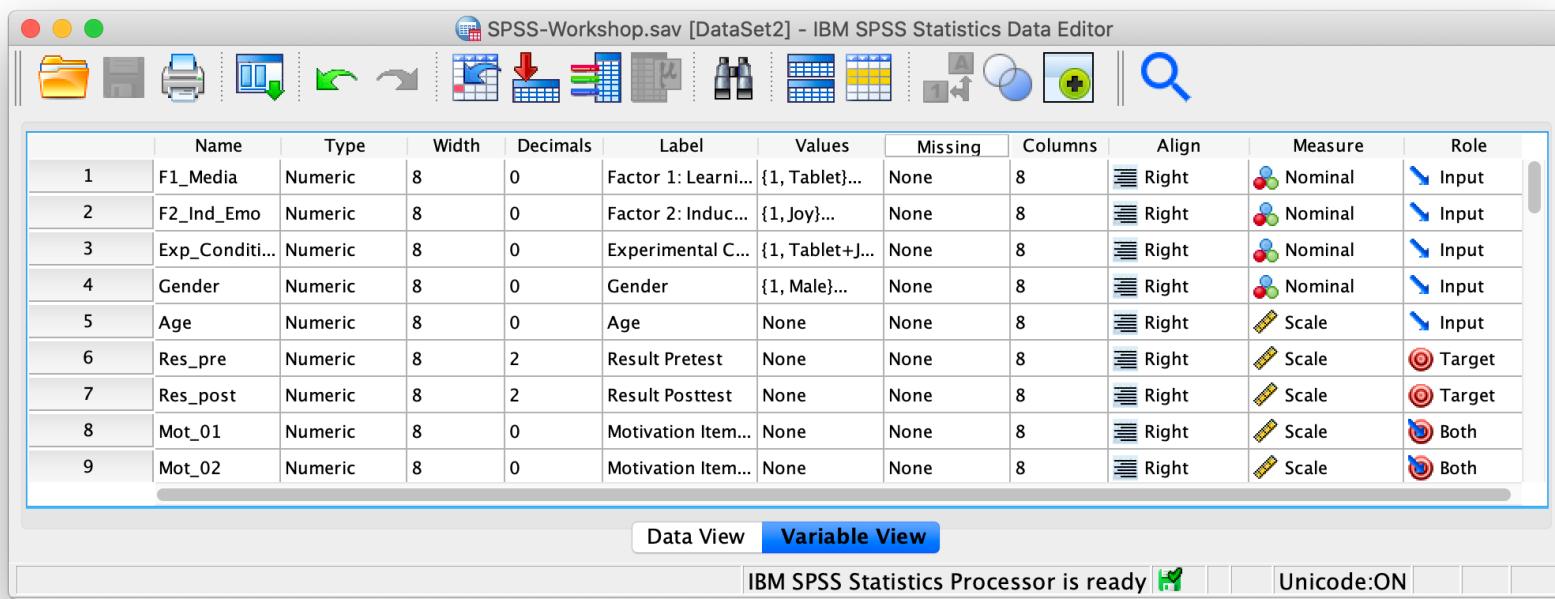
Important: Ensure to set your variables on the correct measure category!

SPSS: VARIABLES

Variables in Data and Variable view & Value labels

Variables

You can manage variables using the Variable View, once you enter the variable name, all the fields will be filled with default values that you can adjust if needed



The screenshot shows the IBM SPSS Statistics Data Editor window titled "SPSS-Workshop.sav [DataSet2] - IBM SPSS Statistics Data Editor". The main area displays a table of variables with the following columns: Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, Measure, and Role. The "Variable View" tab is selected at the bottom. The table data is as follows:

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	F1_Media	Numeric	8	0	Factor 1: Learn...	{1,Tablet}...	None	8	Right	Nominal	Input
2	F2_Ind_Emo	Numeric	8	0	Factor 2: Induc...	{1,Joy}...	None	8	Right	Nominal	Input
3	Exp_Conditi...	Numeric	8	0	Experimental C...	{1,Tablet+J...}	None	8	Right	Nominal	Input
4	Gender	Numeric	8	0	Gender	{1, Male}...	None	8	Right	Nominal	Input
5	Age	Numeric	8	0	Age	None	None	8	Right	Scale	Input
6	Res_pre	Numeric	8	2	Result Pretest	None	None	8	Right	Scale	Target
7	Res_post	Numeric	8	2	Result Posttest	None	None	8	Right	Scale	Target
8	Mot_01	Numeric	8	0	Motivation Item...	None	None	8	Right	Scale	Both
9	Mot_02	Numeric	8	0	Motivation Item...	None	None	8	Right	Scale	Both

At the bottom, the status bar shows "IBM SPSS Statistics Processor is ready" and "Unicode:ON".

Adapted from the slides from Claudio Alves da Silva, 2020

Variables

Best practices for naming variables in SPSS:

- ◊ Each variable name must be unique
- ◊ Keep variable names up to 64 characters long.
- ◊ No spaces, you may include alphanumeric characters, non-punctuation characters. Also periods and underscores within (not at the end).
- ◊ You can use upper and lower cases (Camel case - e.g. typeOfcar).
- ◊ SPSS has reserved keywords, you can't use them as a variable name (i.e.: ALL, AND, BY, NOT, OR, TO or WITH).
- ◊ You can add any type of character and quantity at the “Label”, but also keep it short to keep the output more “readable”.

Variables in Data View



The screenshot shows the SPSS Data Editor window with the title "SPSS-Workshop.sav [DataSet2] - IBM SPSS Statistics Data Editor". The toolbar at the top includes icons for file operations, data manipulation, and search. Below the toolbar is a table titled "Variable View" showing two variables:

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	F1_Media	Numeric	8	0	Factor 1: Learn...	{1,Tablet}...	None	8	Right	Nominal	Input
2	F2_Ind_Emo	Numeric	8	0	Factor 2: Induc...	{1_low}	None	8	Right	Nominal	Input

- ◊ **Name:** name of the variable
- ◊ **Type:** type of variable (e.g.: string, numeric, comma, date)
- ◊ **Width:** number of digits displayed (for numeric values) or the number of characters (for a string)
- ◊ **Decimals:** number of decimal digits to be shown
- ◊ **Columns:** the width of the actual column in the Data View (not the number of digits).
- ◊ **Align:** the alignment of content of the cell.
- ◊ **Measure:** level of measurement for the variable nominal, ordinal, or scale (interval or ratio). This setting affects everything from graphs to internal algorithms for statistical analysis.
- ◊ **Role:** how the variable will be used in your analysis (e.g.: input for IV or target for DV)

Variables in Variable View

Labels:

- ◊ Brief and descriptive display name for the variable
- ◊ Labels will be shown in the output

	Type	Width	Decimals	Label	Values	Miss
1	Numeric	8	0	Factor 1: Learning Medium	{1,Tablet}...	None
2	Numeric	8	0	Factor 2: Induced Emotion	{1,Joy}...	None
3	Numeric	8	0	Experimental Condition	{1,Tablet+J...}	None
4	Numeric	8	0	Gender	{1, Male}...	None
5	Numeric	8	0	Age	None	None
6	Numeric	8	2	Result Pretest	None	None
7	Numeric	8	2	Result Posttest	None	None

Value labels:

- ◊ For coded categorical variables (e.g. 1 = "Tablet"; 2 = "Book")
- ◊ Helps understand what each value represents
- ◊ The value labels can be displayed in the output instead of the original numerical codes
- ◊ To edit the values of a variable click on the button located on each cell from the "Values" column

	Decimals	Label	Values	Missing	Columns	Align
1	0	Factor 1: Learning Medium	Tablet}...	None	8	Right
2	0	Factor 2: Induced Emotion	{1,Joy}...	None	8	Right
3	0	Experimental Condition	{1,Tablet+J...}	None	8	Right
4	0	Gender	{1, Male}...	None	8	Right
5	0	Age	None	None	8	Right
6	2	Result Pretest	None	None	8	Right
7	2	Result Posttest	None	None	8	Right

Value Labels

Value:

Label:

1 = "Tablet"
2 = "Book"

Add Change Remove

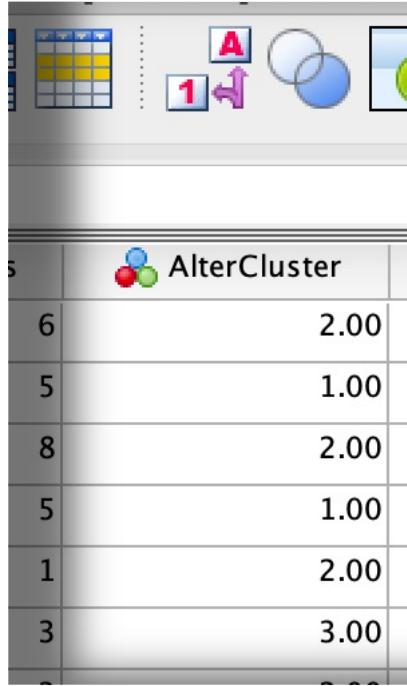
Help Cancel OK

Adapted from the slides from Claudio Alves da Silva, 2020

SPSS Environment: Variable value labels in Data View

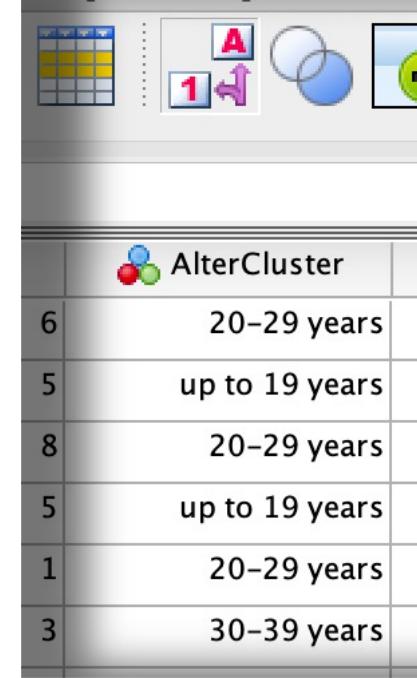
Labels:

- After defining values, in the **Data View** you can click on the “Value labels” button  to switch between values and their labels



The screenshot shows the SPSS Data View window. At the top, there are several icons: a grid, a red circle, a blue circle, a green circle, a red square with 'A', a blue square with '1', and a pink square with arrows. Below the toolbar is a header row with three columns: 'AlterCluster' and two other columns whose names are partially visible. The main area contains seven rows of data. The first column has values 5, 6, 5, 8, 5, 1, and 3. The second column has values 2.00, 1.00, 2.00, 1.00, 2.00, 3.00, and 2.00.

	AlterCluster	
5	2.00	
6	1.00	
5	2.00	
8	1.00	
5	2.00	
1	3.00	
3	2.00	



The screenshot shows the SPSS Data View window after switching to value labels. The same toolbar and header row are present. The data rows now show labels instead of values. The first column has values 6, 5, 8, 5, 1, and 3. The second column has labels "20-29 years", "up to 19 years", "20-29 years", "up to 19 years", "20-29 years", and "30-39 years".

	AlterCluster	
6	20-29 years	
5	up to 19 years	
8	20-29 years	
5	up to 19 years	
1	20-29 years	
3	30-39 years	

Adapted from the slides from Claudio Alves da Silva, 2020

SPSS: VARIABLES

Missing values & Define/Recode missing values

SPSS Environment: Missing values

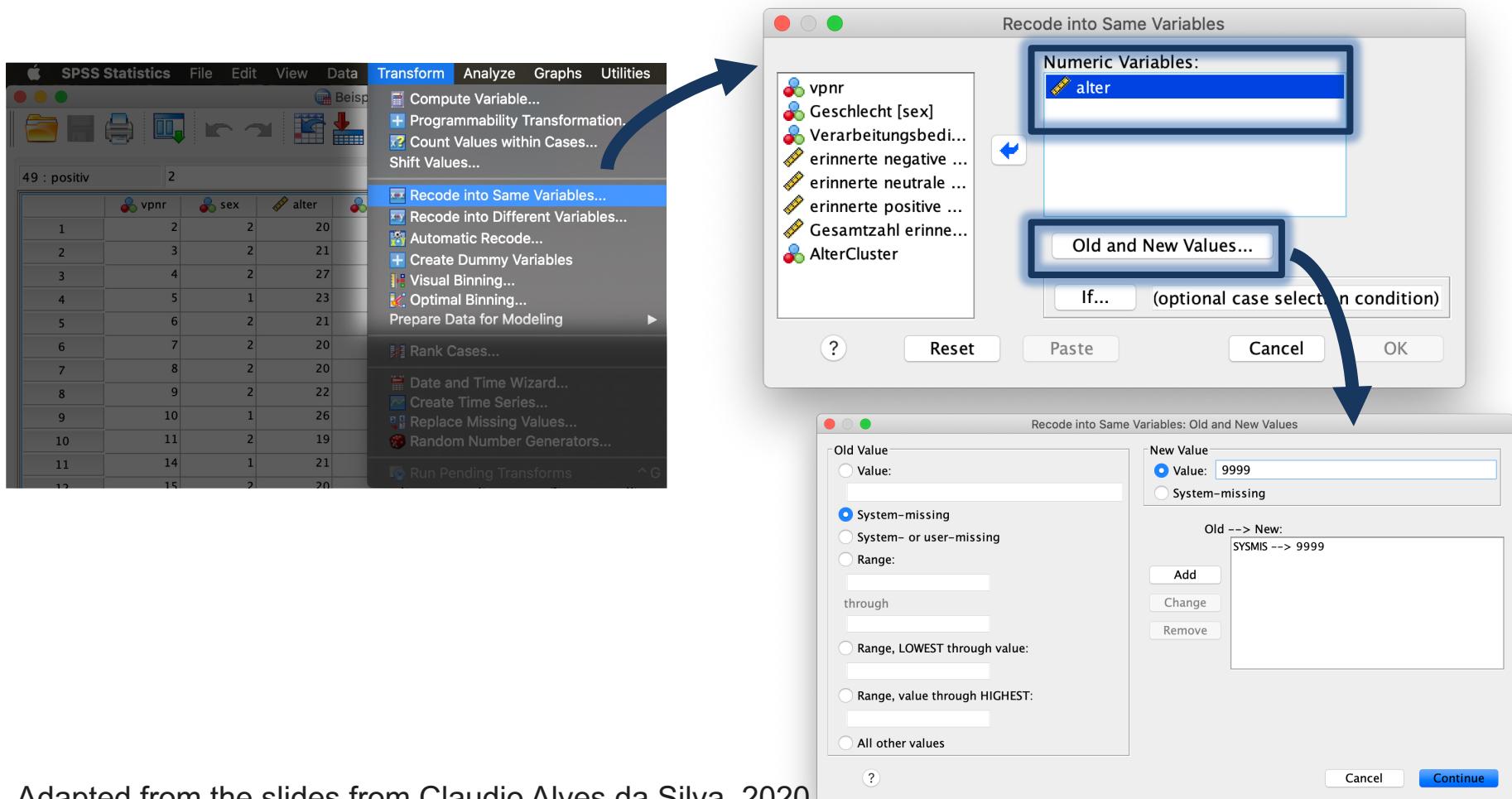
- ❖ When can it happen? When a participant doesn't answer a question, when an answer is illegible and can't be added to the dataset
- ❖ String variables when blank are not automatically recognized as missing values
- ❖ Numeric variables when missing are attributed a default value ":"

SPSS Environment: Define/Recode missing values

- ❖ We can recode a specific value to be treated as missing value data
- ❖ This new recoded missing value **must** be a value outside of the range of the real data ← Why you think that's the case?

SPSS Environment: How-to recode missing values

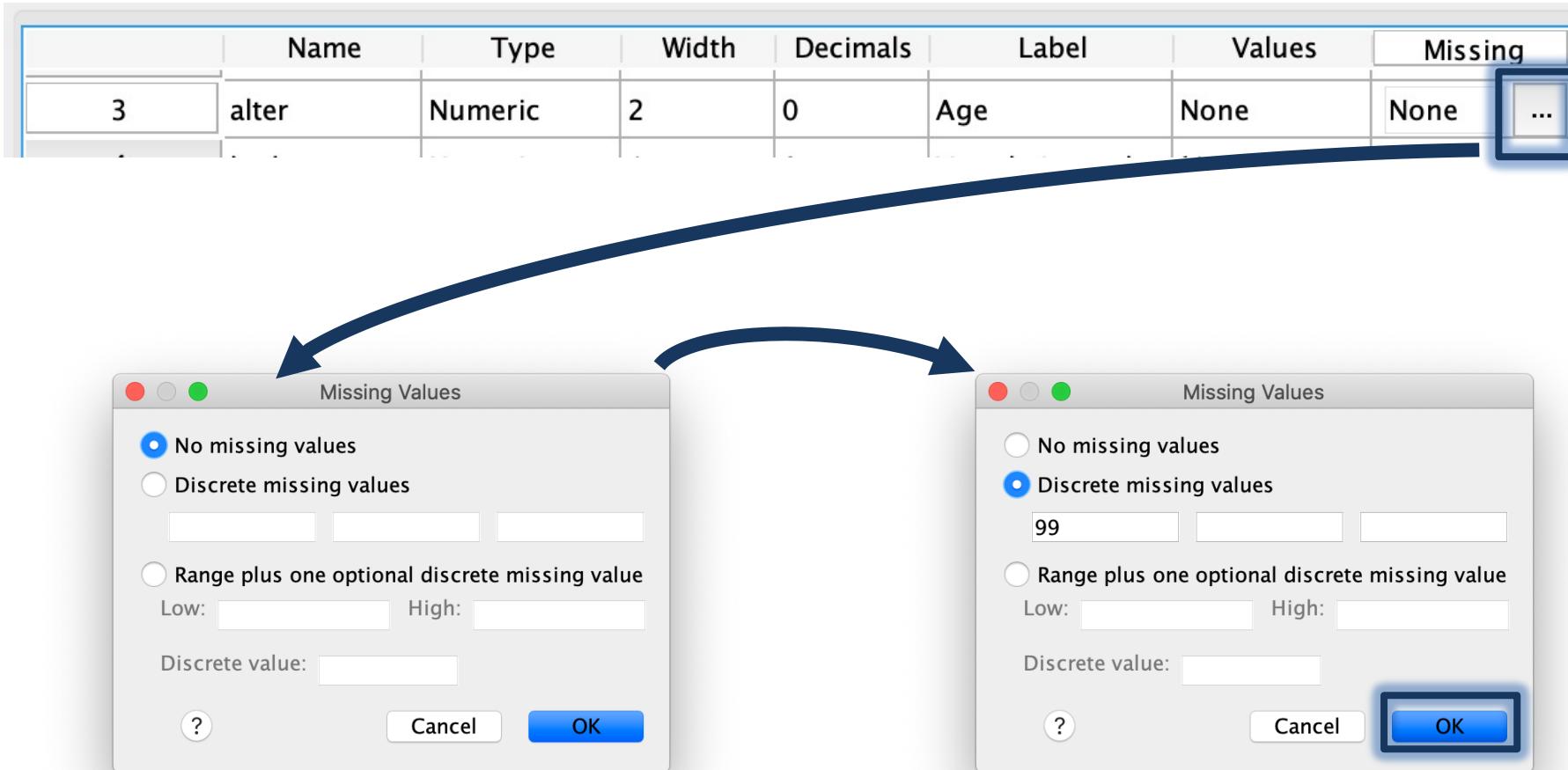
Recoding missing values in Data View



Adapted from the slides from Claudio Alves da Silva, 2020

SPSS Environment: How-to define missing values

Defining missing values in Variable View

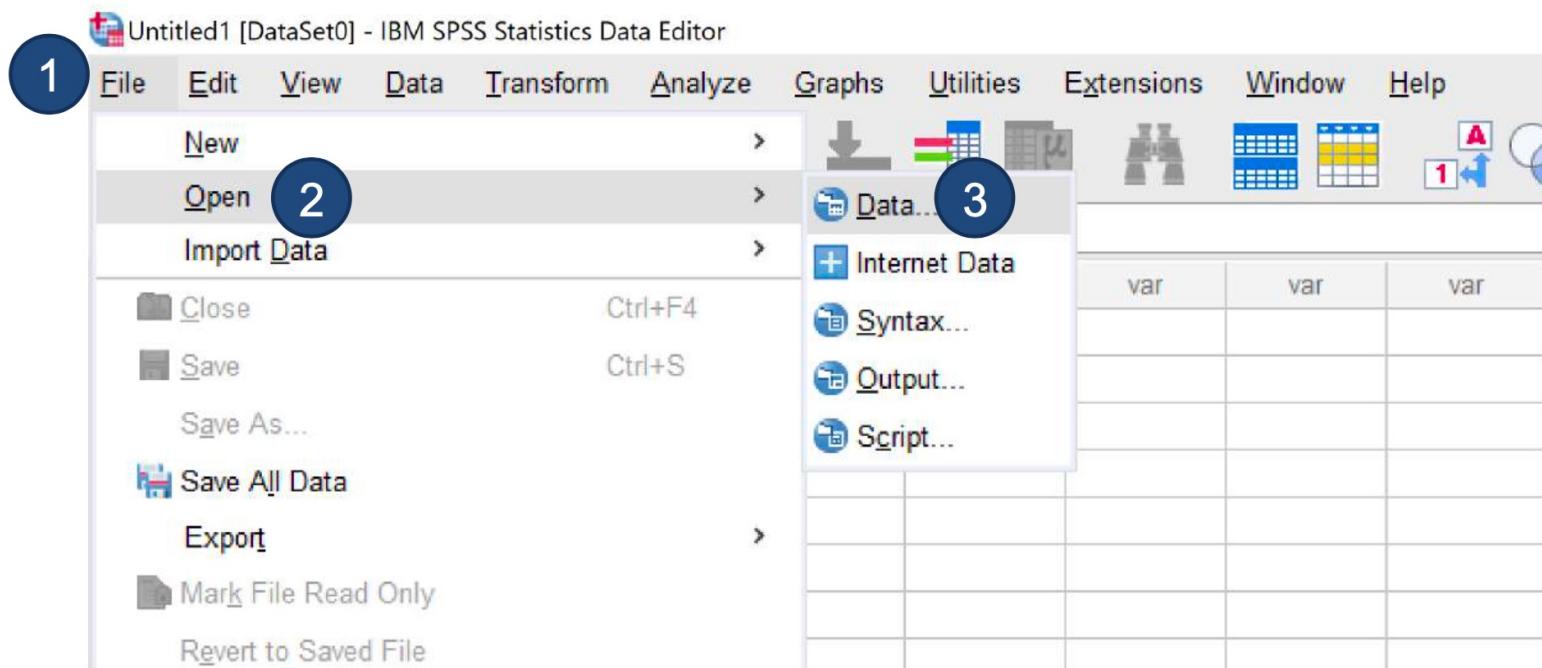


Adapted from the slides from Claudio Alves da Silva, 2020

SPSS: IMPORTING EXCEL DATA

Importing data from Excel to SPSS

- ❖ In the Data Editor window:

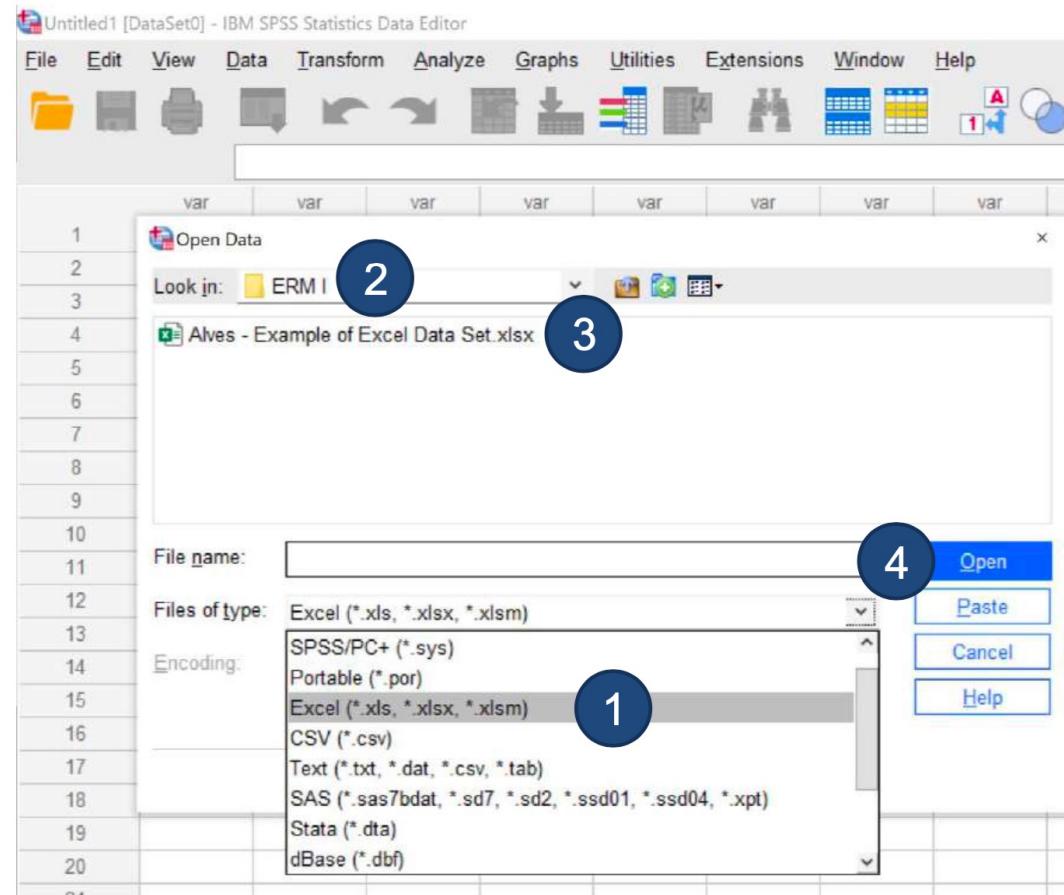


These instructions are validated for SPSS Version 25 to 27.

Importing data from Excel to SPSS

In the “Open Data” window, select the “Excel” option* (1) then navigate to the directory were the file is located (2), select the file (3) and click to open (4).

*If you do not specify the type of file that you wish to open, your file may not appear in the list of available files



Importing data from Excel to SPSS

In the “Read Excel File” window, choose the worksheet that contains your data* (1).

If you want, you can specify the range of rows / columns to import (2).

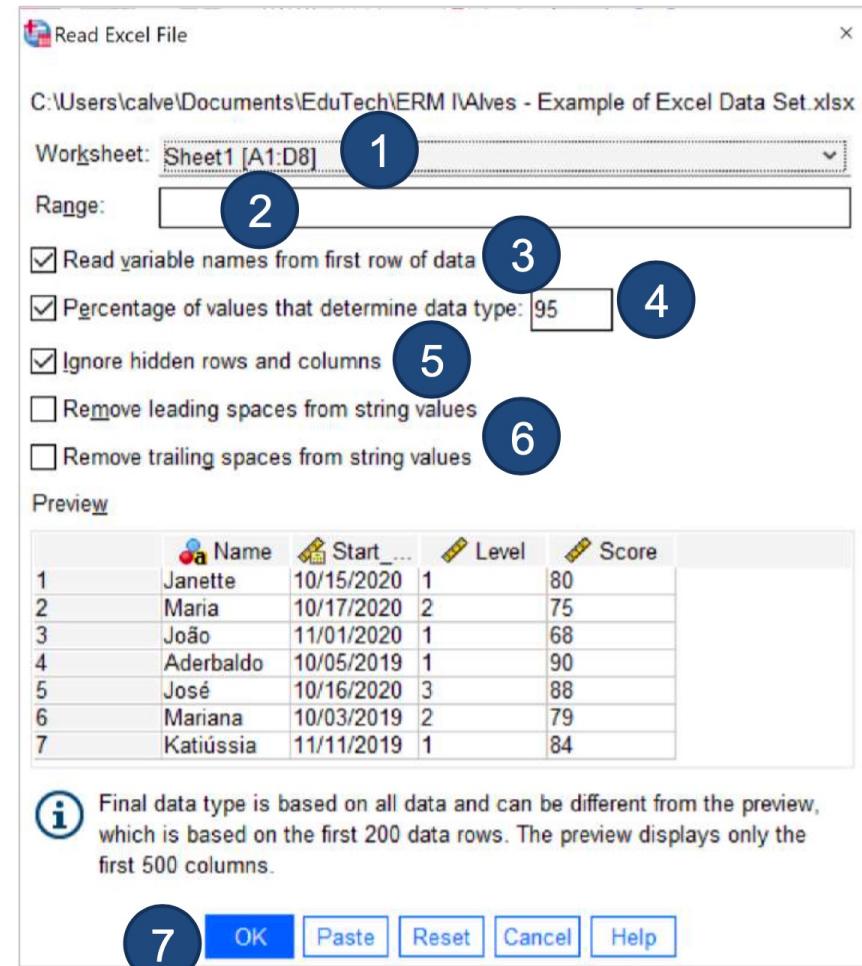
Select (3) if your variable names are in the first row of data, and (4) if you want SPSS to assign a data type to your variables**, you can also ignore hidden data (5).

The options to remove leading and trailing spaces from string values (6) removes any whitespace characters that appear at the beginning or the end of the string.

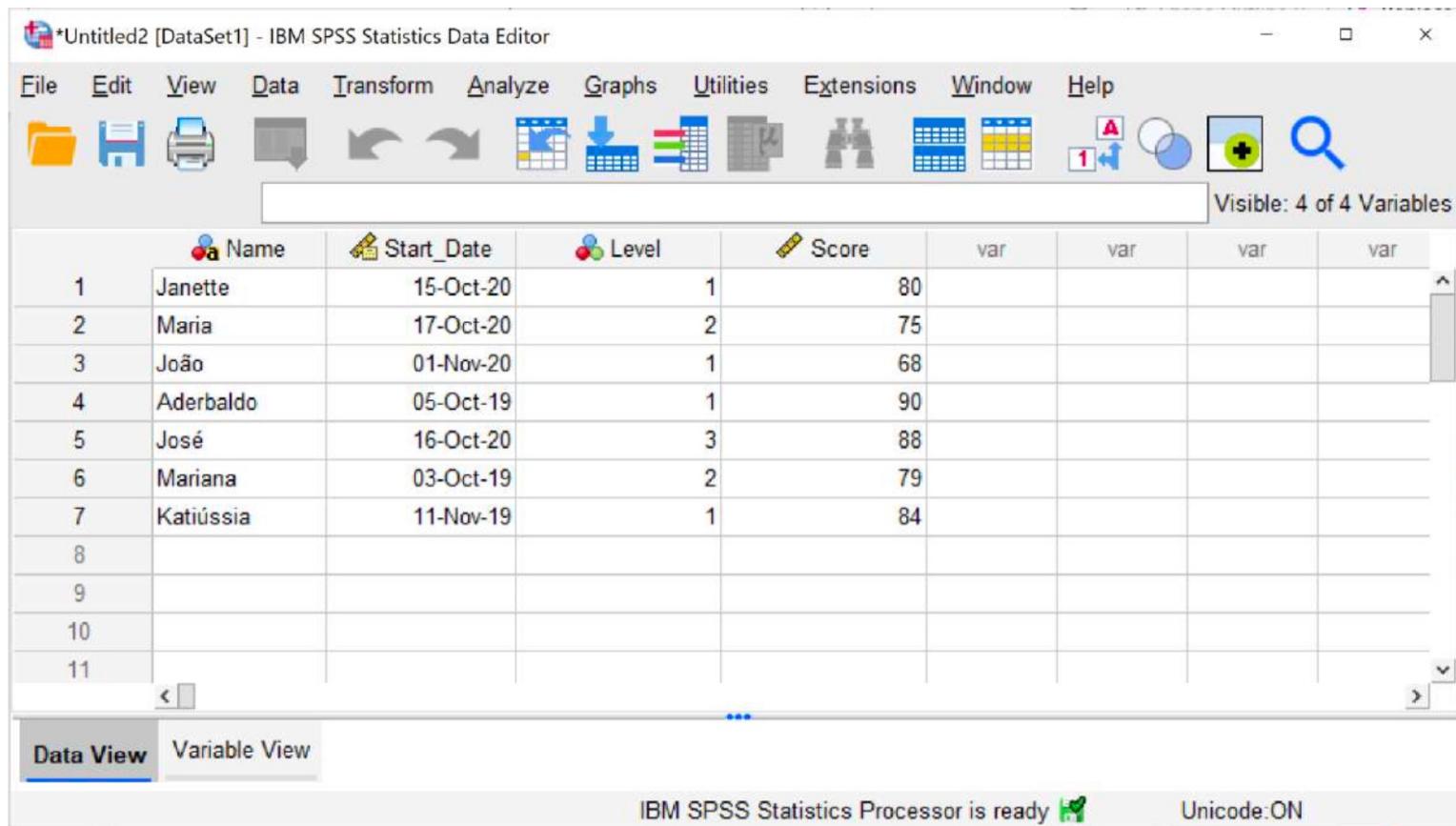
Click OK (7) and the data will appear in SPSS.

* SPSS will import one sheet at a time.

** Any number above 50.



Importing data from Excel to SPSS



The screenshot shows the IBM SPSS Statistics Data Editor window. The title bar reads '*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor'. The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Extensions, Window, and Help. The toolbar contains icons for file operations like Open, Save, Print, and various data analysis tools. A status bar at the bottom indicates 'Visible: 4 of 4 Variables' and 'IBM SPSS Statistics Processor is ready' with a green icon.

	Name	Start Date	Level	Score	var	var	var	var
1	Janette	15-Oct-20	1	80				
2	Maria	17-Oct-20	2	75				
3	João	01-Nov-20	1	68				
4	Aderbaldo	05-Oct-19	1	90				
5	José	16-Oct-20	3	88				
6	Mariana	03-Oct-19	2	79				
7	Katiússia	11-Nov-19	1	84				
8								
9								
10								
11								

Data View Variable View

Example of how data appears in the SPSS Data view window once the data have been successfully imported.

SELF-PACED ACTIVITIES

Activity 1 & 2

- Activity 1
 Recode variable
- Activity 2
 Compute variable