

Tutorial 1: Getting Up and Running with SPSS

Introduction

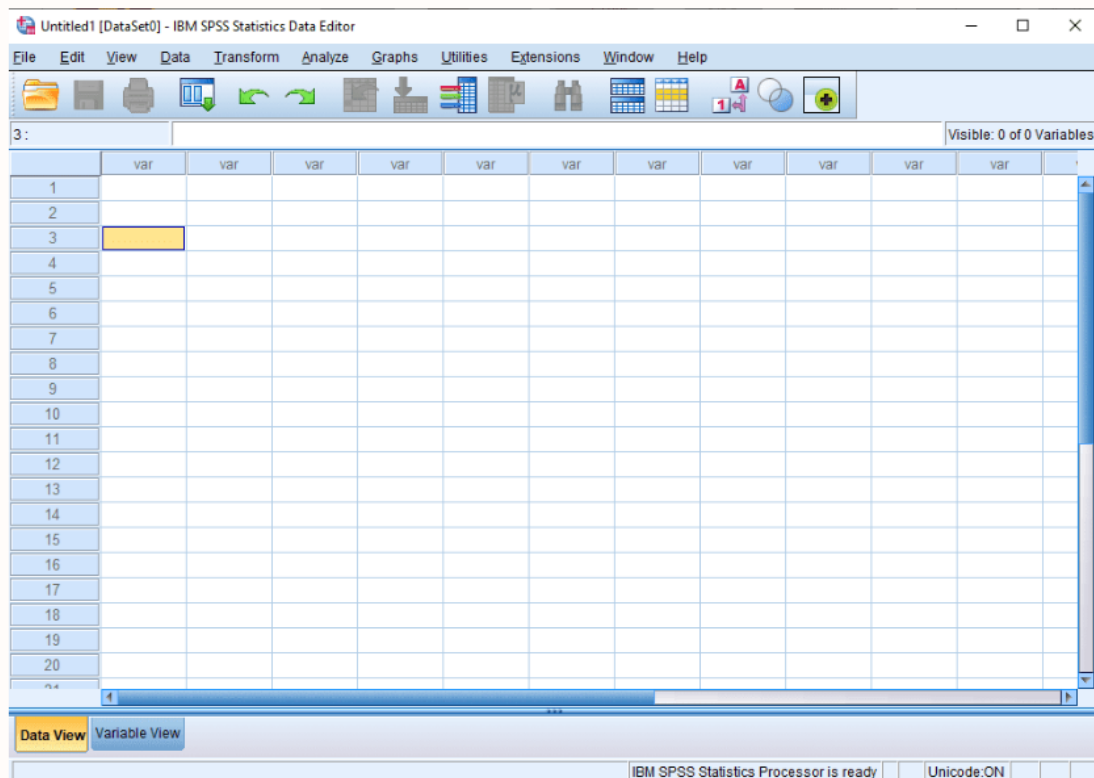
Welcome to this beginner's tutorial on getting started with SPSS! SPSS, which stands for Statistical Package for the Social Sciences, is a powerful statistical analysis software widely used in psychology, sociology, education, business, and other fields.

If you are new to SPSS, you may feel overwhelmed by its many features and options. However, this tutorial will guide you step-by-step through the basics of using SPSS. We will learn how to import a data file, open it in SPSS, clean and sort it, create variables, enter data, and develop basic charts.

By the end of this tutorial, you will have a good understanding of how to navigate SPSS and perform some of the most commonly used statistical procedures when getting a dataset ready. So, let's get started and explore the world of SPSS together.

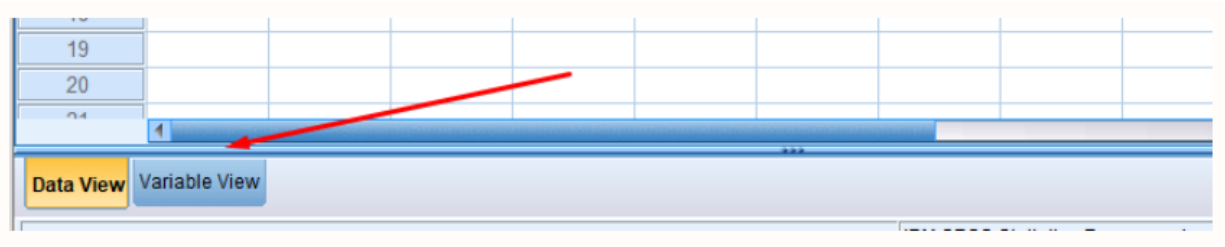
The Main Interface

When you open SPSS, you will see the main interface window or **Data Editor** window, as shown below.



SPSS Data editor window

An important feature of the data editor window is that it presents two views: Data View and Variable View.



The Data View is where you can edit the data you loaded into SPSS.

The Variable View is where you define the properties of the variables in your data set.

These views are best understood after seeing an example of what can be done in each view with an actual data set.

Consider the data set below. Ten fish were caught in Lake Jennings, measured, and released back into the water. The data is shown in the Excel spreadsheet below.

	A	B	C	D	E	F	G	H	I	J
1	ID	GENDER	LENGTH	WEIGHT						
2	1	1	50	122						
3	2	2	57	120						
4	3	2	41	116						
5	4	1	54	135						
6	5	2	54	125						
7	6	1	38	121						
8	7	1	63	109						
9	8	2	48	116						
10	9	1	59	103						
11	10	2	46	137						
12										
13										

There are several ways to import this file into SPSS. We will now see the simplest method, that of direct manual transference.

Loading the Data into SPSS: Method 1

In Excel, we highlight all the data *except for the first row, which contains the headings (variable names)*.

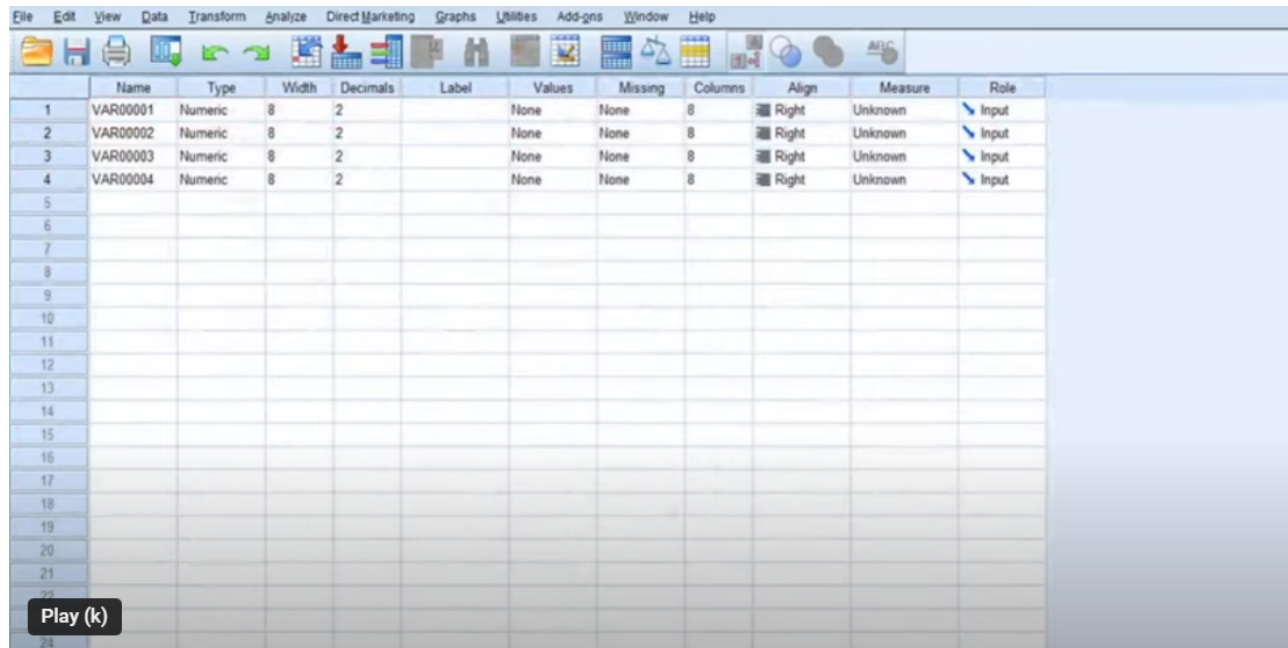
	A	B	C	D
1	ID	GENDER	LENGTH	WEIGHT
2	1	1	50	122
3	2	2	57	120
4	3	2	41	116
5	4	1	54	135
6	5	2	54	125
7	6	1	38	121
8	7	1	63	109
9	8	2	48	116
10	9	1	59	103
11	10	2	46	137

Now, we go over to SPSS, click on an empty cell, say A1, and simply paste the copied data.

	VAR00001	VAR00002	VAR00003	VAR00004	var	var	var	var
1	1.00	1.00	50.00	122.00				
2	2.00	2.00	57.00	120.00				
3	3.00	2.00	41.00	116.00				
4	4.00	1.00	54.00	135.00				
5	5.00	2.00	54.00	125.00				
6	6.00	1.00	38.00	121.00				
7	7.00	1.00	63.00	109.00				
8	8.00	2.00	48.00	116.00				
9	9.00	1.00	59.00	103.00				
10	10.00	2.00	46.00	137.00				

Assigning Variable Properties

When you enter data directly and manually, as we did, you will see that SPSS has generated default names for our variables. In this case, VAR00001, ..., VAR00004. In order to give actual names and enter more properties for the variables, we go to Variable View. There, you will see that SPSS has also assigned default properties to each variable. This is the case when we load the dataset using the direct Copy and Paste method, as we did.



Before we begin to make changes in the variable view, we must familiarize ourselves with some conditions for naming variables in SPSS.

Condition 1: The names given to the variables must not have more than 64 characters.

Condition 2: There should be no spaces between words in the variable name, for example, **Section Number**. To separate the two words, use the underscore: **Section_Number**.

Condition 3: The names must be unique. Two variables cannot have the same name.

Condition 4: Names cannot start with a number or any special character, for example. 3, &, *, #.

Now, we will see how to change or alter properties in Variable View by examining each column heading, beginning with Column 1: Name.

NAME

Here, we can name each variable. However, these names are only for identification purposes and do not appear anywhere in the analysis.

To change the first variable name from VAR00001 to ID, we click on each variable name. This will highlight the name, which we will change to "ID." Change the other names:

VAR00002 = Gender, VAR00003 = Length and VAR00004 = Weight.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ID	Numeric	8	2		None	None	8	Right	Unknown	Input
2	Gender	Numeric	8	2		None	None	8	Right	Unknown	Input
3	Length	Numeric	8	2		None	None	8	Right	Unknown	Input
4	Weight	Numeric	8	2		None	None	8	Right	Unknown	Input
5											
6											

TYPE

This is where we declare the variable type. All variables say “Numeric” by default. The only non-numeric variable is Gender, which is coded as numeric, with 1 = Male and 2 = Female. We, therefore, leave the default “numeric” values as they are.

	Name	Type	Width	Decimals
1	ID	Numeric	8	2
2	Gender	Numeric	8	2
3	Length	Numeric	8	2
4	Weight	Numeric	8	2
5				
6				

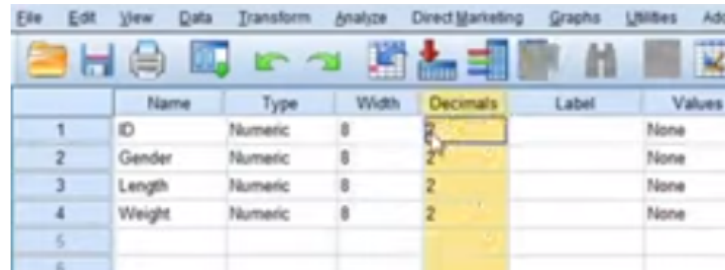
WIDTH

The third column is Width. The width is the number of characters (numbers or letters) in the variable name, which is eight (8). The default width is 8. However, we can increase or decrease this value by clicking in the cell and using the down or up arrows that appear to change the values.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ID	Numeric	8	2		None	None	8	Right	Unknown	Input
2	Gender	Numeric	8	2		None	None	8	Right	Unknown	Input
3	Length	Numeric	8	2		None	None	8	Right	Unknown	Input
4	Weight	Numeric	8	2		None	None	8	Right	Unknown	Input
5											
6											

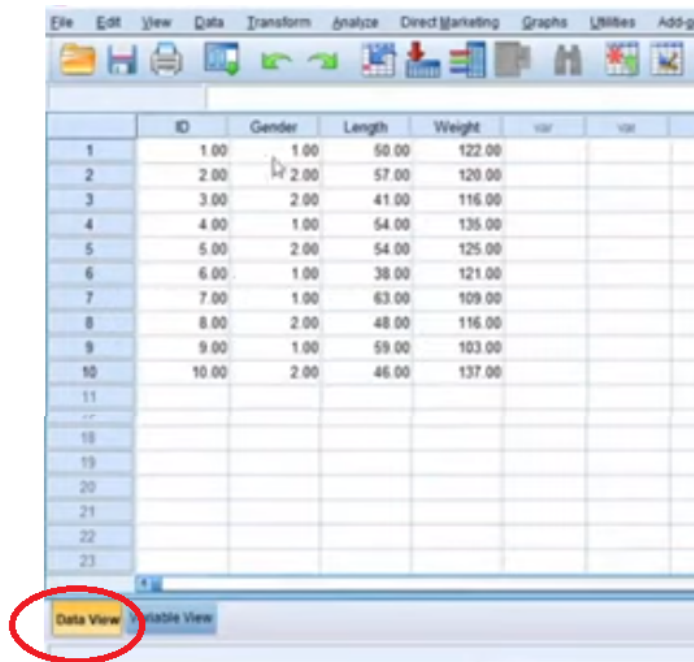
DECIMALS

This column gives the number of decimal places for variables that are numeric. The variables are given two decimal places by default.



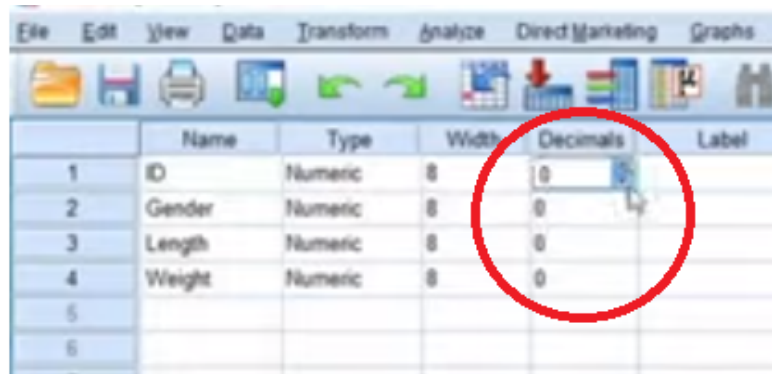
	Name	Type	Width	Decimals	Label	Values
1	ID	Numeric	8	2		None
2	Gender	Numeric	8	2		None
3	Length	Numeric	8	2		None
4	Weight	Numeric	8	2		None
5						
6						

If we go to Data View, we will see that all values are expressed in two decimal places.



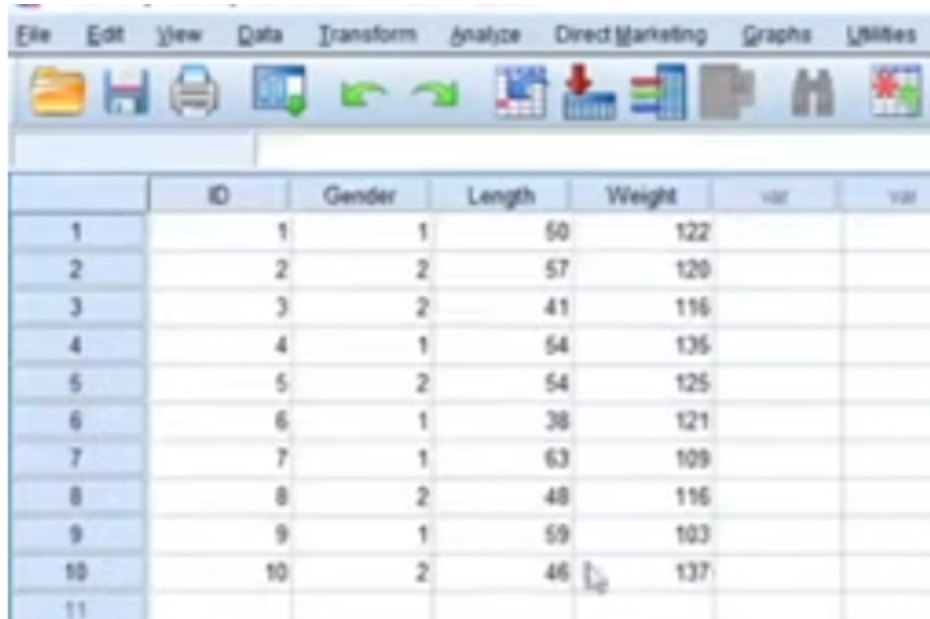
	ID	Gender	Length	Weight	var	var
1	1.00	1.00	50.00	122.00		
2	2.00	2.00	57.00	120.00		
3	3.00	2.00	41.00	116.00		
4	4.00	1.00	54.00	135.00		
5	5.00	2.00	54.00	125.00		
6	6.00	1.00	38.00	121.00		
7	7.00	1.00	63.00	109.00		
8	8.00	2.00	48.00	116.00		
9	9.00	1.00	59.00	103.00		
10	10.00	2.00	46.00	137.00		
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						

Now, I don't want the values expressed with decimal places, so I return to Variable View, click on each cell in the Decimal column, and change to zero.



	Name	Type	Width	Decimals	Label
1	ID	Numeric	8	0	
2	Gender	Numeric	8	0	
3	Length	Numeric	8	0	
4	Weight	Numeric	8	0	
5					
6					

In Data View, we could see the results of this change, with all numbers being expressed as integers.

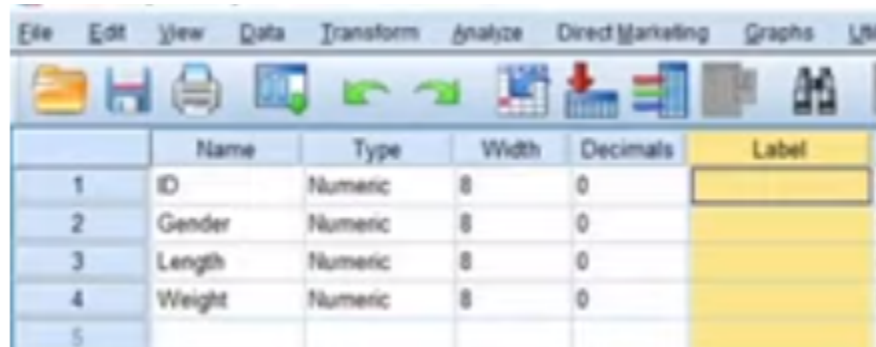


	ID	Gender	Length	Weight	var	var
1	1	1	50	122		
2	2	2	57	120		
3	3	2	41	116		
4	4	1	54	135		
5	5	2	54	125		
6	6	1	38	121		
7	7	1	63	109		
8	8	2	48	116		
9	9	1	59	103		
10	10	2	46	137		
11						

Let us return to Variable View to continue entering the properties of the variables.

LABEL

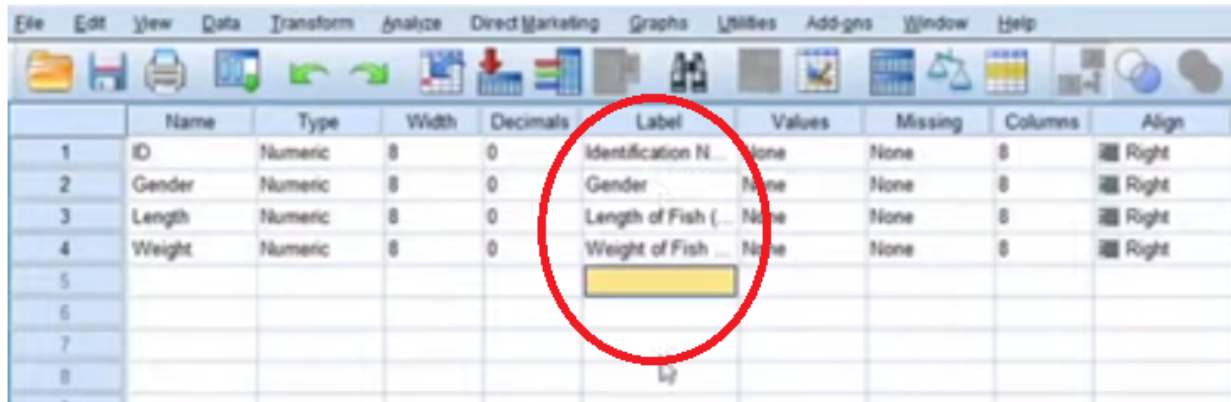
This is where you enter a description of the variable. This column is one of the most important because what you enter here will show up in all analyses of the data.



	Name	Type	Width	Decimals	Label
1	ID	Numeric	8	0	
2	Gender	Numeric	8	0	
3	Length	Numeric	8	0	
4	Weight	Numeric	8	0	
5					

We change the values as follows:

For the first variable, we click in the cell and write “Identification Number” (there can be spaces in the Label). In the second variable, we write “Gender.” We write “Length of fish (cm)” for the third variable and “Weight of fish (kg)” for the fourth. This is shown below.



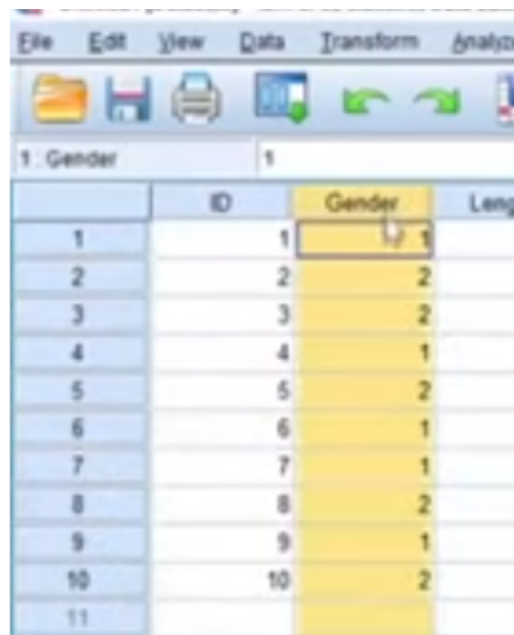
The screenshot shows the SPSS Variable View window. A red circle highlights the 'Label' column. The table below represents the data shown in the window.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align
1	ID	Numeric	8	0	Identification N...	None	None	8	Right
2	Gender	Numeric	8	0	Gender	None	None	8	Right
3	Length	Numeric	8	0	Length of Fish (...)	None	None	8	Right
4	Weight	Numeric	8	0	Weight of Fish ...	None	None	8	Right
5									
6									
7									
8									

We can widen the column to see the entries better.

VALUES

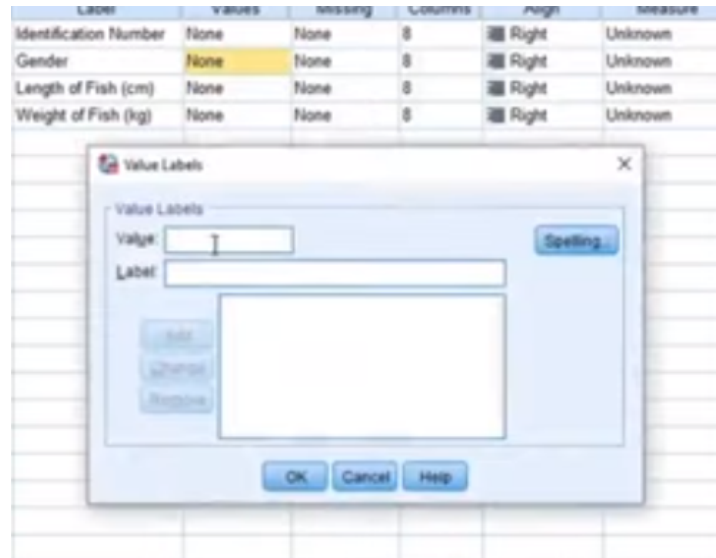
The Values column is where you assign values to categorical and ordinal variables. The only categorical variable we have is Gender. So, we go to Data View to see how many groups we have in the Gender Variable.



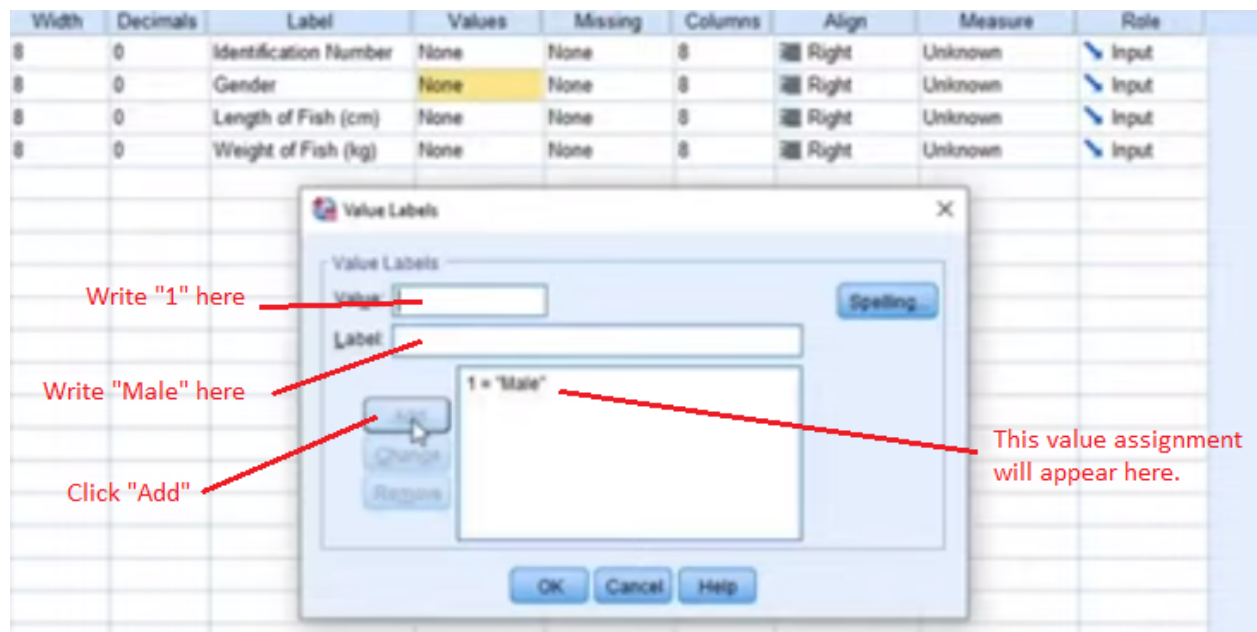
The screenshot shows the SPSS Data View window for the 'Gender' variable. The table below represents the data shown in the window.

	ID	Gender	Length
1	1	1	
2	2	2	
3	3	2	
4	4	1	
5	5	2	
6	6	1	
7	7	1	
8	8	2	
9	9	1	
10	10	2	
11			

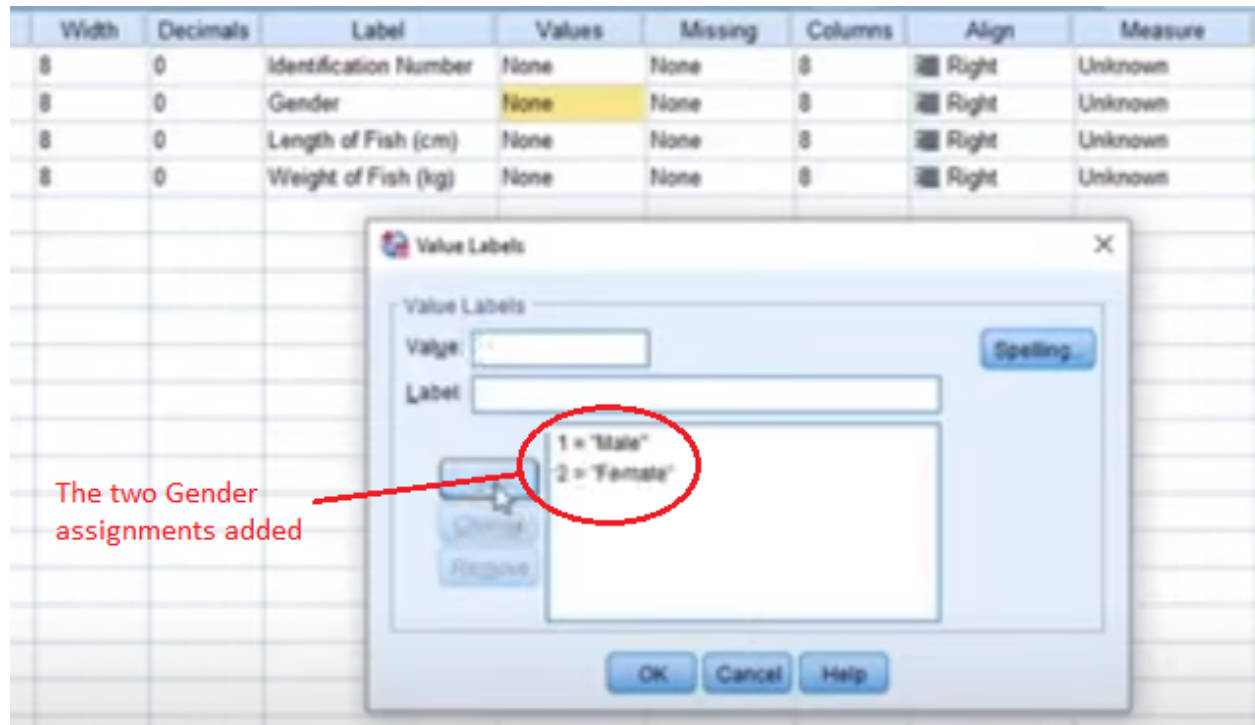
We see there are only two groups: 1 and 2. We go back to Variable View and make the changes. We click in the cell of the Gender variable; click on the blue arrow, and a dialog box will appear.



To make the assignments, proceed as follows.



Make the second assignment by writing "2" in the Value box and "Female" in the Label box.



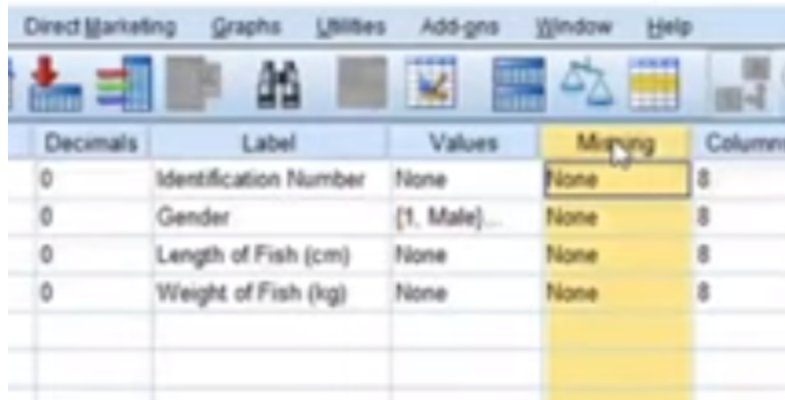
Click OK. Now, let us go to Data View to see the results of our assignments.

The screenshot shows the SPSS Data View with the 'Gender' variable assigned to the values 'Male' and 'Female'. The table has columns: ID, Gender, Length, and Weight. The 'Gender' column is highlighted in yellow.

ID	Gender	Length	Weight
1	Male	50	122
2	Female	57	120
3	Female	41	116
4	Male	54	135
5	Female	54	125
6	Male	38	121
7	Male	63	109
8	Female	48	116
9	Male	59	103
10	Female	46	137
11			

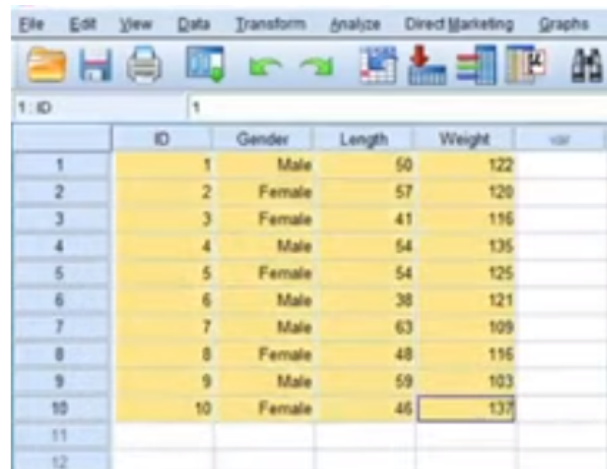
MISSING

This is where we indicate whether there are missing values in any of the variables.



Decimals	Label	Values	Missing	Columns
0	Identification Number	None	None	8
0	Gender	{1, Male}...	None	8
0	Length of Fish (cm)	None	None	8
0	Weight of Fish (kg)	None	None	8

There are no missing values, as we can see from the Data View.



	ID	Gender	Length	Weight	var
1	1	Male	60	122	
2	2	Female	57	120	
3	3	Female	41	116	
4	4	Male	54	135	
5	5	Female	54	125	
6	6	Male	38	121	
7	7	Male	63	109	
8	8	Female	48	116	
9	9	Male	59	103	
10	10	Female	46	137	
11					
12					

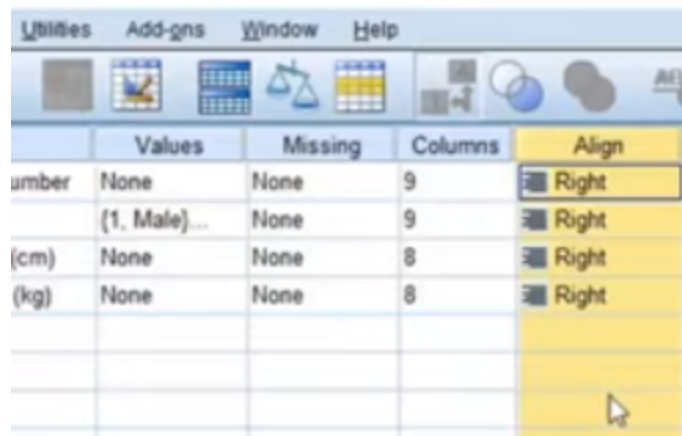
We, therefore, move on to the next column.

COLUMNS

This is where you can increase or decrease the width of each variable column in Data View by entering a value in “Column” of Variable View. The application of the property is self-explanatory, so we move on to the next column.

ALIGN

This is where you tell SPSS how you want your data arranged in each of the variable columns in Data View. Default is “Right.”

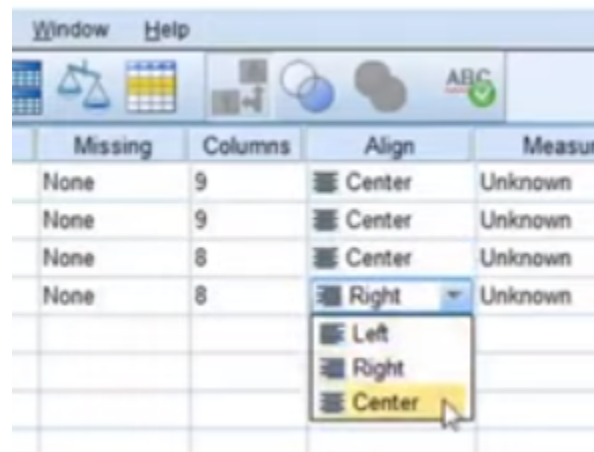


This means that all data is right-aligned in Data View.

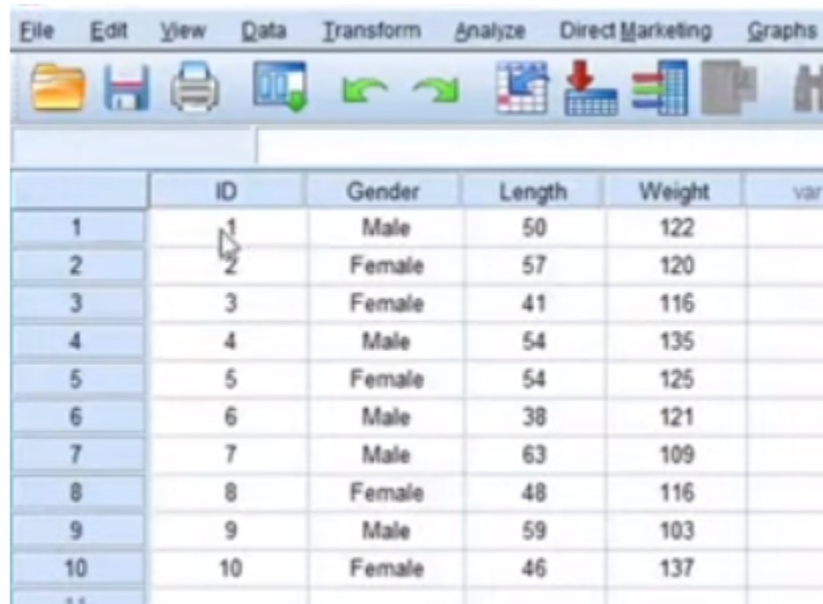
This screenshot shows the SPSS Data Editor window with a table containing 10 rows of data. The columns are labeled 'ID', 'GENDER', 'LENGTH', 'WEIGHT', and 'VIB'. The data is as follows:

	ID	GENDER	LENGTH	WEIGHT	VIB
1	1	1	50	122	
2	2	2	57	120	
3	3	2	41	116	
4	4	1	54	135	
5	5	2	54	125	
6	6	1	38	121	
7	7	1	63	109	
8	8	2	48	116	
9	9	1	59	103	
10	10	2	46	137	

Now, let us change the alignment of the data in all variables to "Center." To do this, we simply click on the right edge of the cell of each variable and choose "Center" from the drop-down list that appears.



Now, let us go to Data View to see the results of these changes.

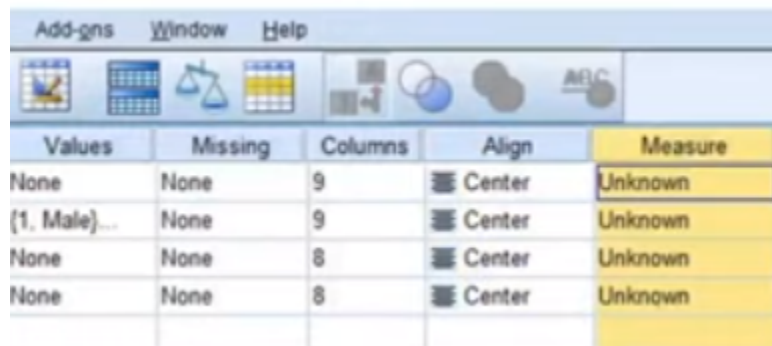


	ID	Gender	Length	Weight	var
1	1	Male	50	122	
2	2	Female	57	120	
3	3	Female	41	116	
4	4	Male	54	135	
5	5	Female	54	125	
6	6	Male	38	121	
7	7	Male	63	109	
8	8	Female	48	116	
9	9	Male	59	103	
10	10	Female	46	137	

We see that the data are all center-aligned.

MEASURE

This is where you declare the level of measurement of each variable, an important property. The default is “Unknown.”

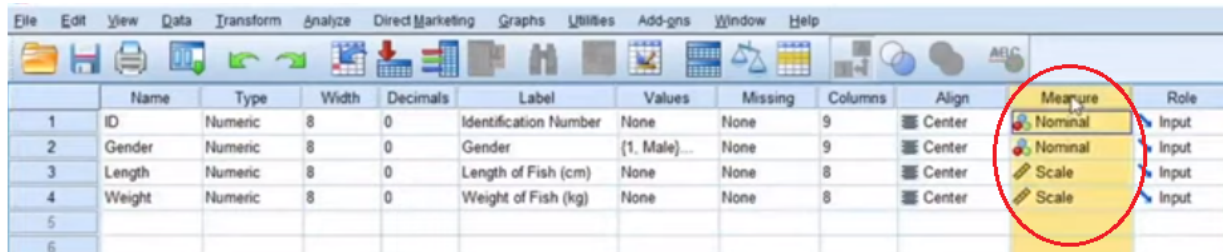


Values	Missing	Columns	Align	Measure
None	None	9	Center	Unknown
(1, Male)...	None	9	Center	Unknown
None	None	8	Center	Unknown
None	None	8	Center	Unknown

In SPSS, there are three options for Measure: Scale (numeric), Nominal (categorical unordered, e.g., Blue, Red, etc.) and Ordinal (categorical ordered, e.g., Large, Medium, Small). Therefore, we need to change the measure for each variable from the default of “Unknown” to its appropriate measure. The first variable, “ID,” is Nominal because it is a categorical variable that does not represent any ordering or ranking.

The variable “Gender” is also Nominal because it is a categorical variable that does not represent any ordering or ranking: (Male, Female).

The variables “Length” and “Width” are both numeric. Therefore, we choose “Scale.”



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ID	Numeric	8	0	Identification Number	None	None	9	Center	Nominal	Input
2	Gender	Numeric	8	0	Gender	{1, Male}...	None	9	Center	Nominal	Input
3	Length	Numeric	8	0	Length of Fish (cm)	None	None	8	Center	Scale	Input
4	Weight	Numeric	8	0	Weight of Fish (kg)	None	None	8	Center	Scale	Input
5											
6											

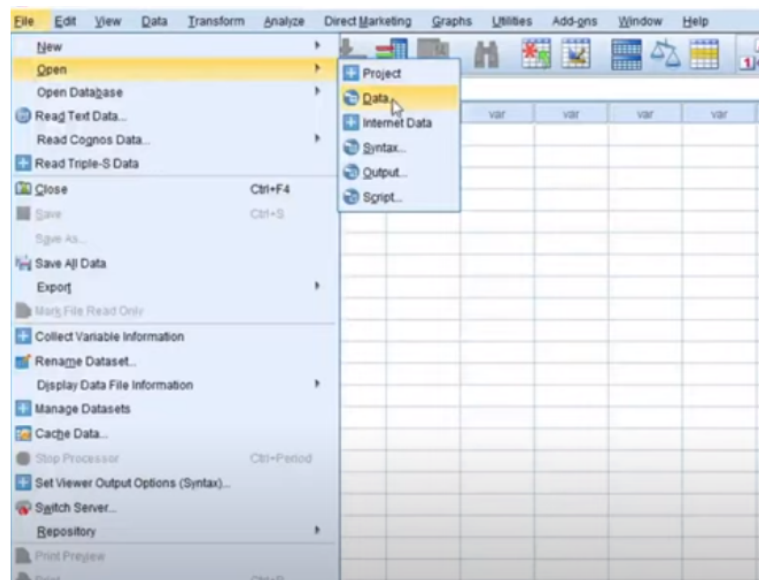
ROLE

We will leave this property as the default “Input” for all variables.

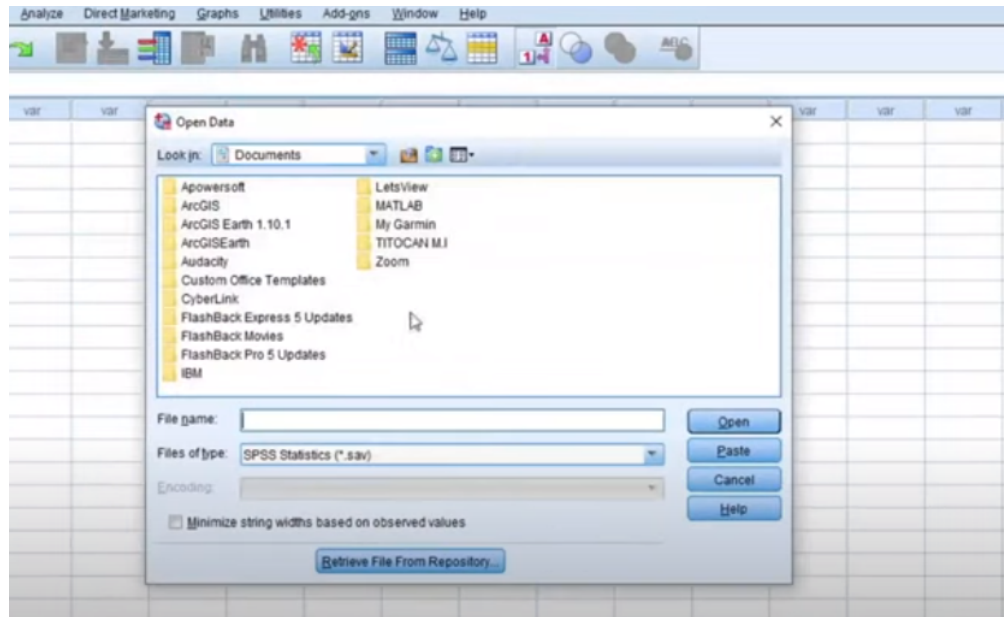
We have now declared and defined all properties and can go to Data View to do our analysis. However, let us now see how to load the data using a second method that results in less work for us.

Loading the Data into SPSS: Method 2

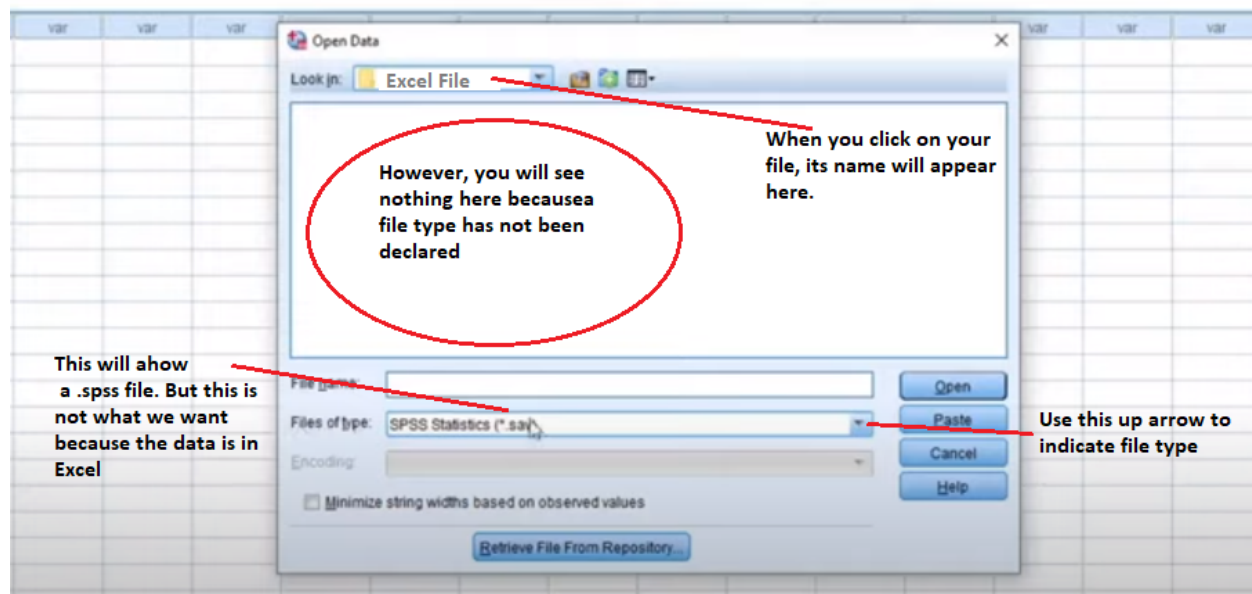
In the SPSS main interface, click on the File menu > Open > Data.



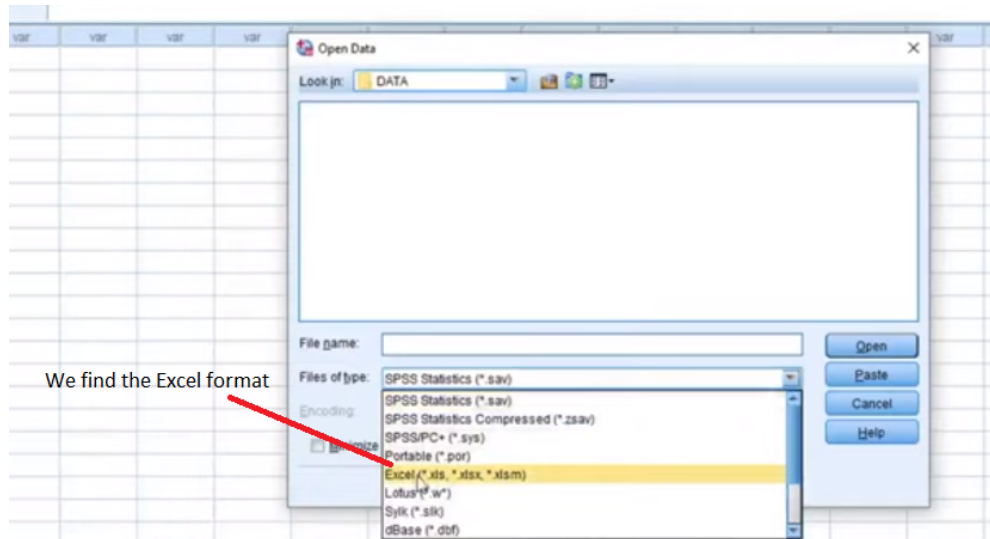
The following dialog box will appear.



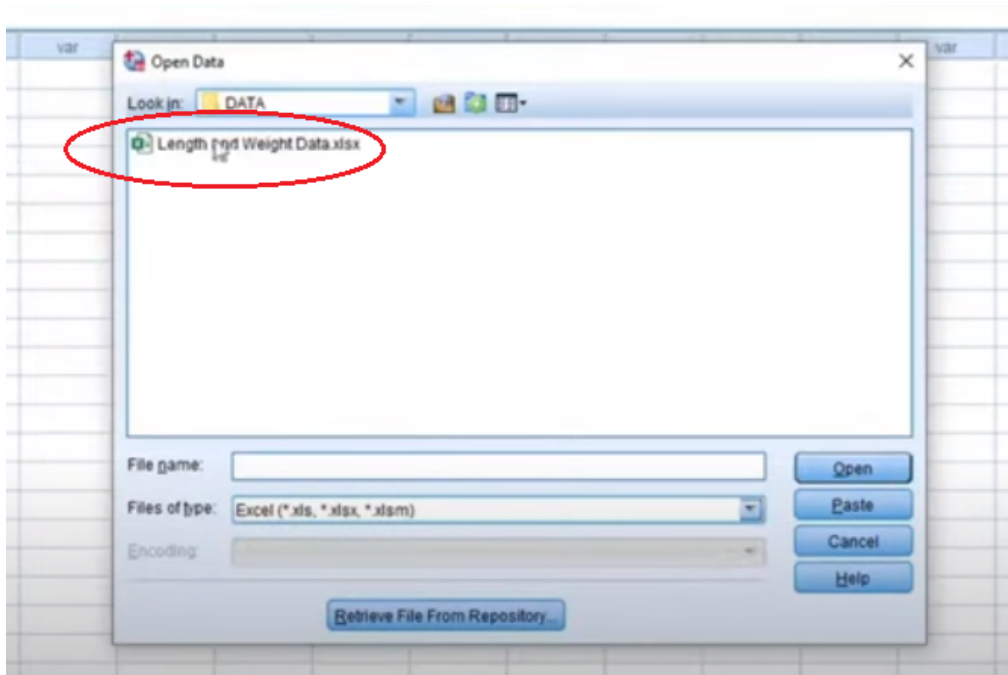
Here, you navigate to your Excel file.



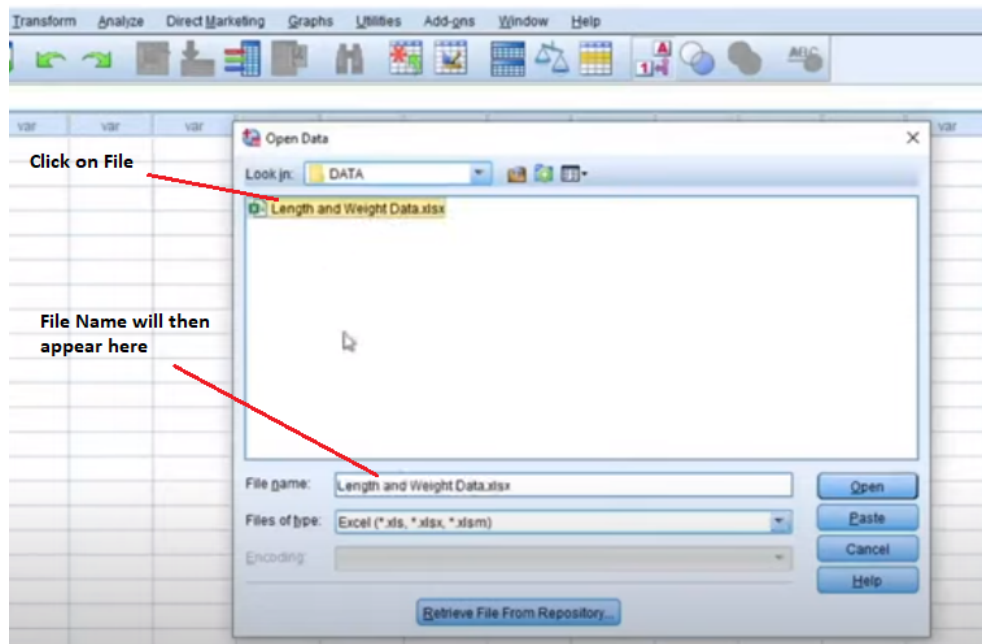
We use the blue arrow to the right of the File Type box to find the “Excel” format.



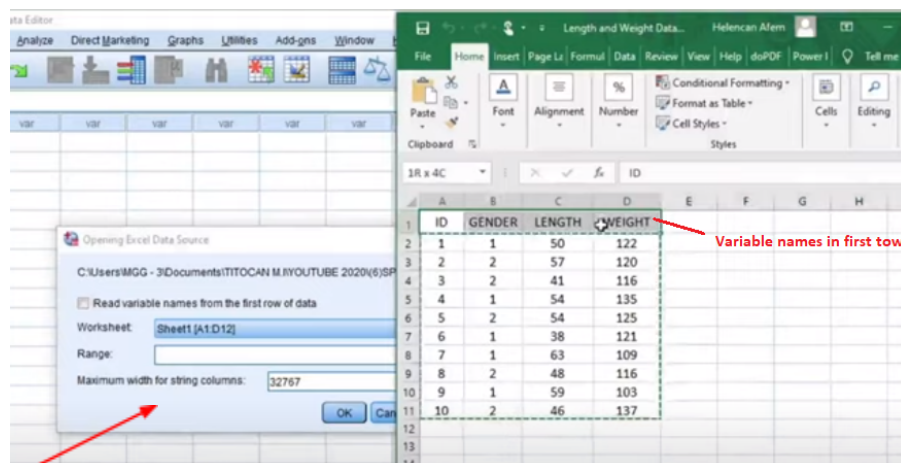
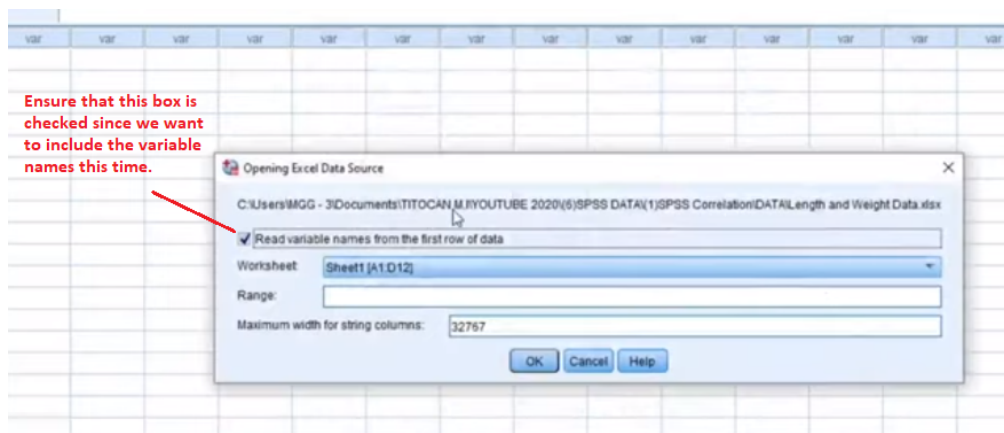
And immediately, our file appears in the open space.



Now, we do as the figure shows.

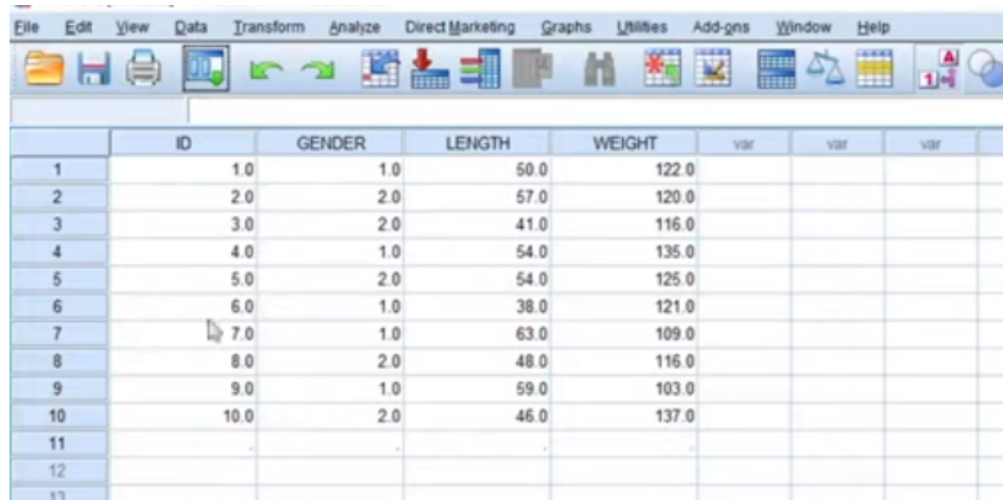


Click "Open," and the following dialog box appears.



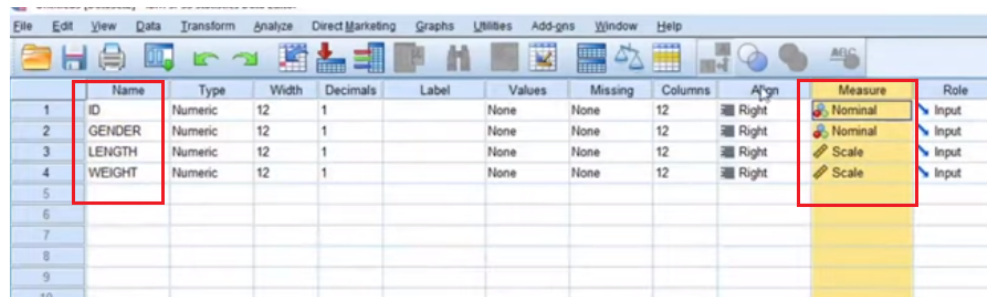
Now, click “OK” in the blue dialog box.

The dataset will be loaded into SPSS with the correct variable names and not the default. This is shown in the Data View below.



	ID	GENDER	LENGTH	WEIGHT	var	var	var
1	1.0	1.0	50.0	122.0			
2	2.0	2.0	57.0	120.0			
3	3.0	2.0	41.0	116.0			
4	4.0	1.0	54.0	135.0			
5	5.0	2.0	54.0	125.0			
6	6.0	1.0	38.0	121.0			
7	7.0	1.0	63.0	109.0			
8	8.0	2.0	48.0	116.0			
9	9.0	1.0	59.0	103.0			
10	10.0	2.0	46.0	137.0			
11							
12							
13							

Now, let us go into Variable View to examine our data set.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	ID	Numeric	12	1		None	None	12	Right	Nominal	Input
2	GENDER	Numeric	12	1		None	None	12	Right	Nominal	Input
3	LENGTH	Numeric	12	1		None	None	12	Right	Scale	Input
4	WEIGHT	Numeric	12	1		None	None	12	Right	Scale	Input
5											
6											
7											
8											
9											
10											

As we can see, the columns for “Name” and “Measure” have already been filled with the correct properties. We, therefore, would now fill in the properties for the remaining columns as we did in Method 1.