# Introduction to SPSS

Today we begin our SPSS workshops. Not assuming prior knowledge of SPSS, we start with the basics. This first introductory workshop covers the following:

1. How to access SPSS
2. Navigating SPSS user interface (data and variable views)
3. How to enter data manually
4. How to save and open data files
5. How to create a new variable by recoding an existing variable

## How to access SPSS

SPSS is the primary statistical analytic tool for SSCI202. It is available in the computers of our labs, the library and digital lounge (C5C - 17 Wally’s Walk). Students can also access SPSS remotely through iLab (for details, see the next section).

If you want to install SPSS on your personal computer, you need to consider buying it. Several online stores offer an academic discount for students. It currently costs USD 35.95 for a 6-month license or USD 68.99 for a one-year license. If you buy it, choose “IBM SPSS Statistics Grad Pack 25.0 Base”.

* <https://studentdiscounts.com/?cm_mc_uid=95050930491614920117784&cm_mc_sid_50200000=1493219758&cm_mc_sid_52640000=1493125316>
* <http://www.academicsuperstore.com/products/IBM+SPSS/IBM+SPSS+Statistics>

## iLab

iLab is a Remote Desktop system that allows you to connect to university computers remotely. It gives you an access to SPSS over the internet. More information is available at <https://wiki.mq.edu.au/display/iLab/About>.

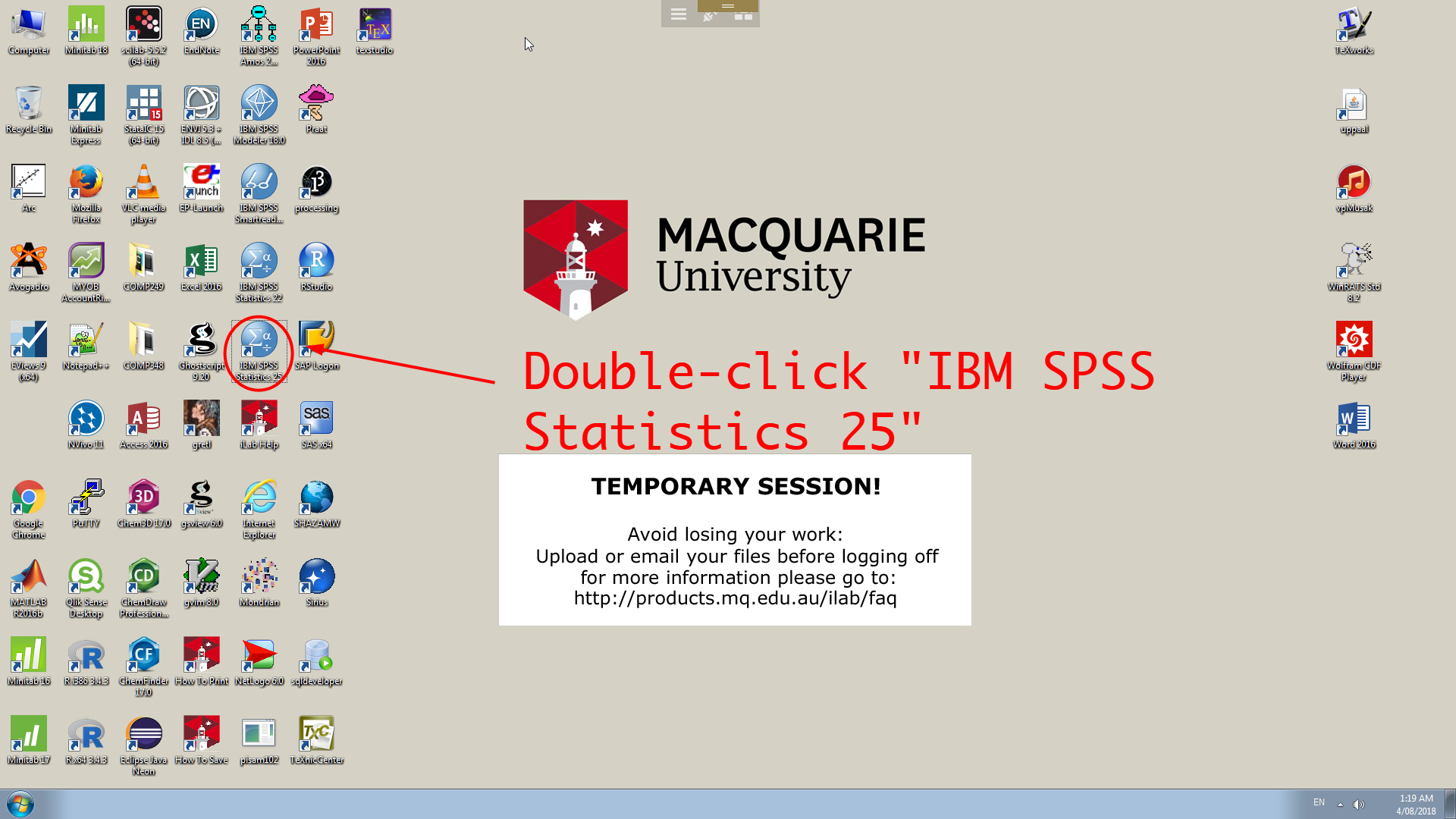
iLab is accessible from the Microsoft Windows, Mac OS X, Linux, iOS and Android operating systems. To get started with iLab, you need to take three steps.

* Download a client at <https://wiki.mq.edu.au/display/iLab/Clients>.
* Install it.
* Launch it.

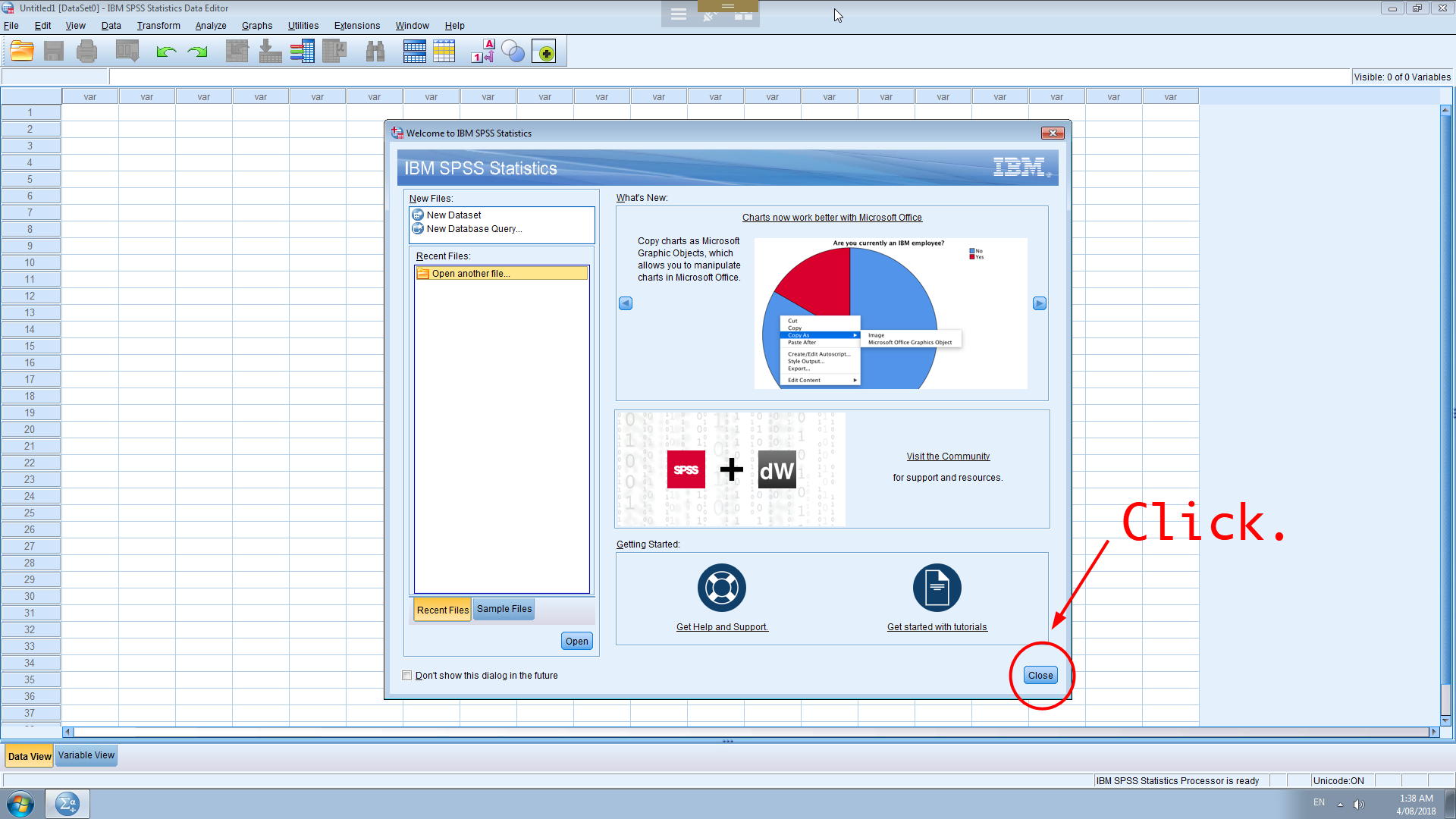
If you have an issue in installing and using iLab, please contact IT service desk (see <https://students.mq.edu.au/support/technology/service-desk>).

## Using SPSS via iLab

After logging into iLab or in the workshop computers, you will see the Windows desktop. **Double-click the icon of “IBM SPSS Statistics 25”**.



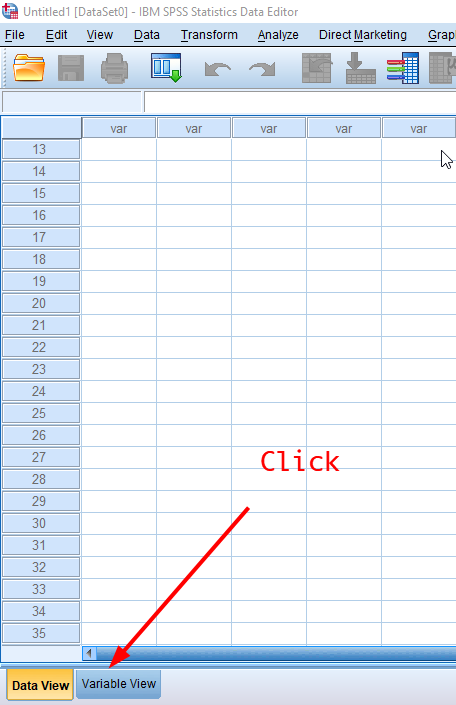
Then, it will launch SPSS. You will see your first SPSS window. **Just** **click the button of “Close”**. Then, you are ready for the next step.



Whenever you use SPSS via iLab, **I strongly recommend saving your work very frequently** (See p.14 for how to save your work). Because iLab provides just a temporary session, sometimes it abruptly ends without any notice. This means that you will lose your work unless you save it. And **you have to send your file to your mail account before logging off**. All personal files will be removed from iLab, and you can’t restore them.

## Navigating SPSS user interface (data and variable views)

Now let’s navigate how to manage data in SPSS. SPSS has two tabs at the bottom, Data View and Variable View. **Click Data View**.



Data can be manually entered into the Data Editor (as seen in the below figure), which may be useful for inputting small datasets or making minor edits to larger datasets.

**Now, we will manually enter a subsample of 30 respondents from Aussa (Australian Survey of Social Attitudes) dataset** using <table 1> It includes four variables: gender, age, political orientation and social class.

<Table 1> does not have any identification variable, that is an assigned number for each respondent to help researchers identify respondents, so we will need to create an ID variable following the common practice of data entry.

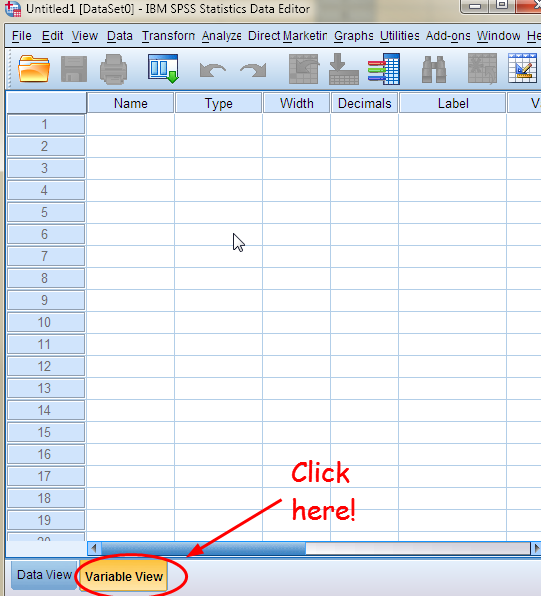
<Table 1> An Aussa Subsample of 30 Respondents

|  |  |  |  |
| --- | --- | --- | --- |
| **Gender** | **Age** | **Political Orientation** | **Social Class** |
| Male | 66 | Right | Middle class |
| Female | 72 | Right | Upper middle class |
| Female | 59 | Left | Middle class |
| Female | 20 | Left | Lower middle class |
| Female | 68 | Right | Upper middle class |
| Male | 76 | Right | Middle class |
| Male | 61 | Left | Upper middle class |
| Male | 90 | Right | Middle class |
| Female | 64 | Left | Lower middle class |
| Female | 39 | Left | Upper middle class |
| Male | 57 | Right | Middle class |
| Male | 47 | Left | Lower class |
| Female | 56 | Left | Middle class |
| Female | 51 | Left | Middle class |
| Male | 34 | Left | Working class |
| Male | 18 | Center | Middle class |
| Female | 18 | Left | Working class |
| Female | 30 | Left | Upper middle class |
| Female | 65 | Right | Middle class |
| Male | 35 | Right | Middle class |
| Female | 44 | Right | Upper class |
| Female | 40 | Right | Middle class |
| Male | 57 | Left | Upper middle class |
| Male | 40 | Left | Lower middle class |
| Female | 59 | Left | Middle class |
| Female | 82 | Right | Middle class |
| Female | 44 | Far right | Working class |
| Female | 30 | Left | Middle class |
| Male | 77 | Left | Working class |
| Female | 60 | Right | Lower middle class |

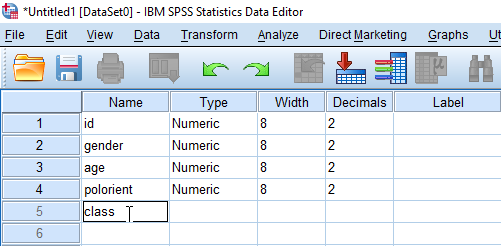
## Variable View Tab

**After completing data entry, now move on to the Variable View tab** at the bottom of the Data Editor window.

You need to define the variables that will be used. In our case, five variables are needed: *id, gender, age, political orientation* and *social class*.



* In the first row of the first column, type **id**.
* In the second row, type **gender**.
* In the third row, type **age**.
* In the fourth row, type **polorient** (which indicates *political orientation*).
* In the fifth row, type **class** (which indicates *social class*).



Next, **Click the Data View tab at the bottom** to continue entering the data.

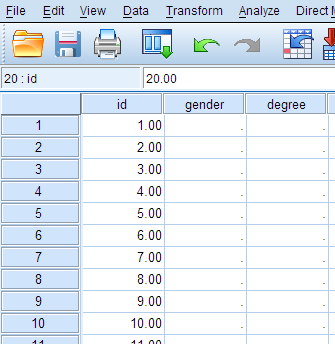
The names that you entered in Variable View are now the headings for the first five columns in Data View.

## How to Enter Data

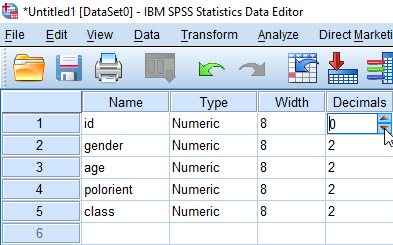
## Entering id variable (Nominal variable)

*id* variable is used to identify respondents. Since we have 30 respondents, let’s make id variable from 1 to 30.

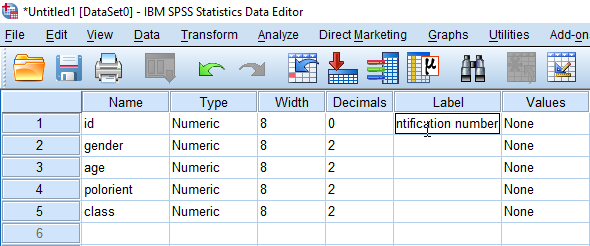
* First, **enter values for id in the first row, starting at the first column**. The value is 1. In the next row, the value is 2 and so on.



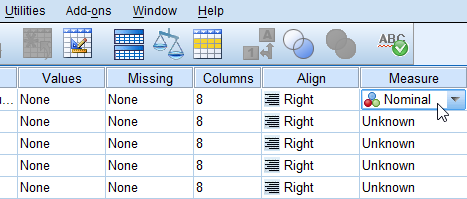
* After finishing to enter 30 for 30th respondent, we need to define characteristics of this variable. **Click the Variable View tab at the bottom** again.
* In the Variable View window, **click the first row of the Decimals column**. The default value is 2, which indicate two decimal places. We change this to 0 (no decimal) since *id* is a nominal variable.



* **Give a short description of this variable by typing “Identification number” in the Label column**.



* SPSS allows three levels of measurement: Nominal, Ordinal, and Scale. Level of measurement determines types of statistical methods or graphs we can use with the variable. So, assigning the appropriate level of measurement to a variable is critical. Since *id* is a nominal variable, we will define *id* as such. You just **click the Measure column for *id***. **and change the default value, “Unknown” to “Nominal”**.



* **Go back to the Data View tab**, and check how your id variable is changed (i.e. no decimal).

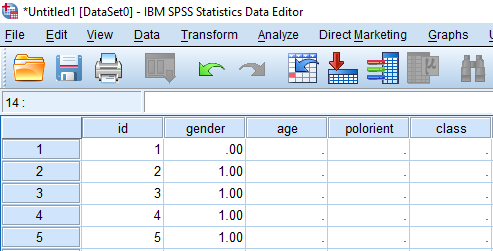
## Entering gender variable (Nominal variable)

*gender* variable has two categories: Male and Female. For convenience, we assign a numerical value to each category.

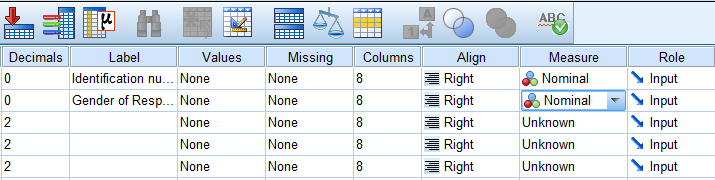
0 = Male

1 = Female1

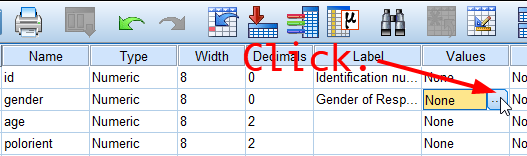
* In the Data Editor window, **enter the numerical values corresponding to each category under the column of *gender***.



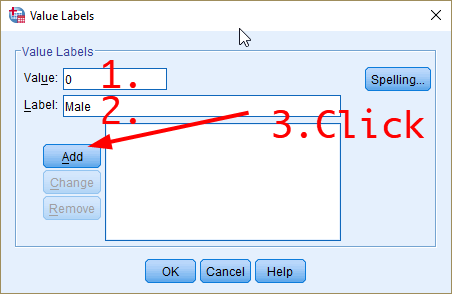
* **Set the variable characteristics of *gender* accordingly in the Variable View tab** again. It should have no decimal, proper label, and correct level of measurement.



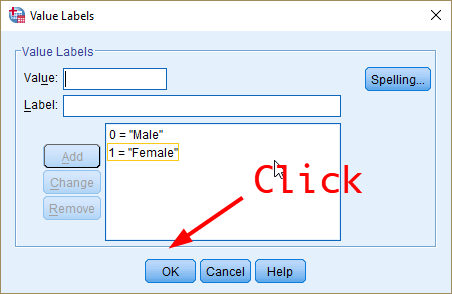
* Now, we will assign two categories (Male and Female) to each numerical value. First, **click the Values column of *gender***. Then, **click the blue shaded square in the cell as in the below figure**.



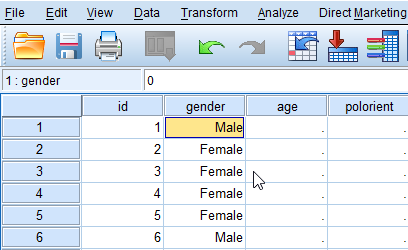
* You will see the box of Value Labels. In the box **1) type 0 in the Value**, **2) type “Male” in the Label**, and **3) Click Add**.

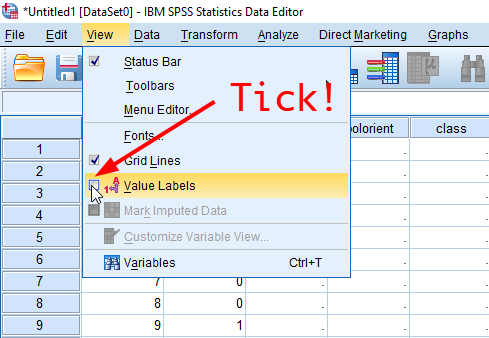


* In the same manner, **1) type 1 in the Value, 2) type “Female” in the Label, and 3) Click Add**. Then you will see the below dialogue box.



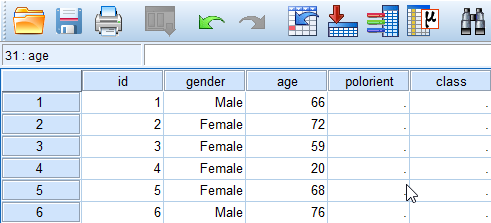
* Close the dialogue box by clicking OK.
* Let’s **come back to the Data View window**. Do you see your gender variable now have letters instead of numbers like below? If not, go to View menu at the top and then **tick Value Labels**.

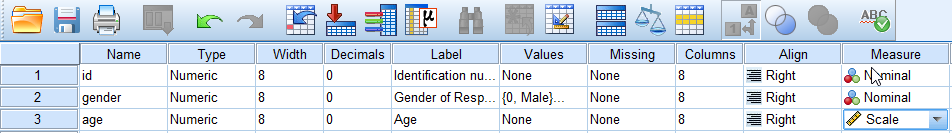




* Enabling the option of value Labels will set SPSS to display value labels instead of numerical values in the Data View window.

## Entering age variable (interval-ratio level variable)

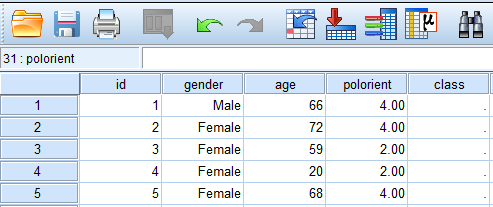
* In the Data View enter the values of age as in the <Table 1>.
* 
* In the Variable View, **set the Decimals to 0, type “Age” in the Label column and change the Measure to “Scale”**.



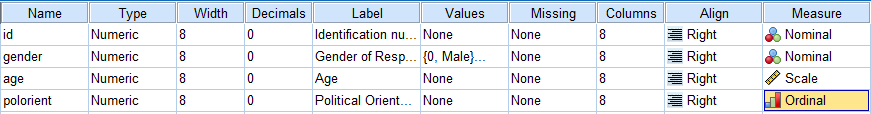
## Entering political orientation variable (ordinal variable)

*polorient* is an ordinal variable with five categories: 1 = Far left, 2 = Left, 3 = Central, 4 = Right, 5 = Far right.

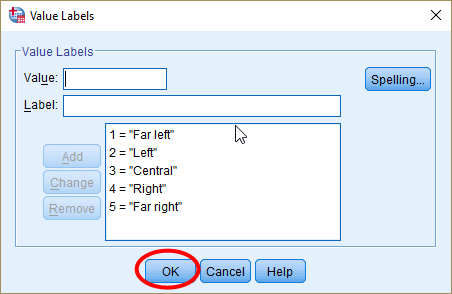
* In the Data View, **enter corresponding numerical values for each category**.



* In the Variable View window, **change the Decimals to 0, the Measure to Ordinal. And type “Political Orientation” in the Label**.



* Next, we will again assign the categories (value labels) to the numerical values.
* As in the gender variable case, click the blue shaded square in the Value column of *polorient*. Then type “1” in the Value and “Far left” in the Label. And click Add. In the same way, assign the other four categories. Then you will see the below dialogue box. And click OK.

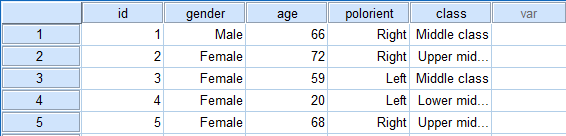


## Entering social class variable (ordinal variable)

*class* is is an ordinal-level variable with six categories: 1 = Lower class, 2 = Working class, 3 = Lower middle class, 4 = Middle class, 5 = Upper middle class, 6 = Upper class.

Enter class information by yourself. The procedure is very similar to what you have done for *polorient*.

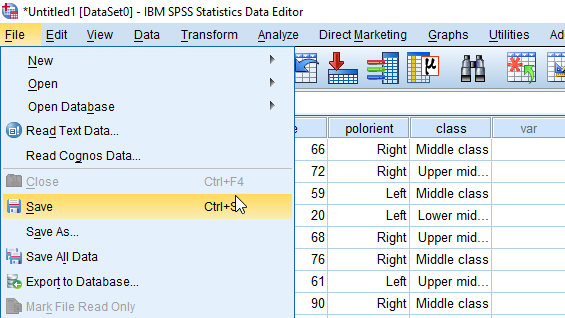
The final data should look like:



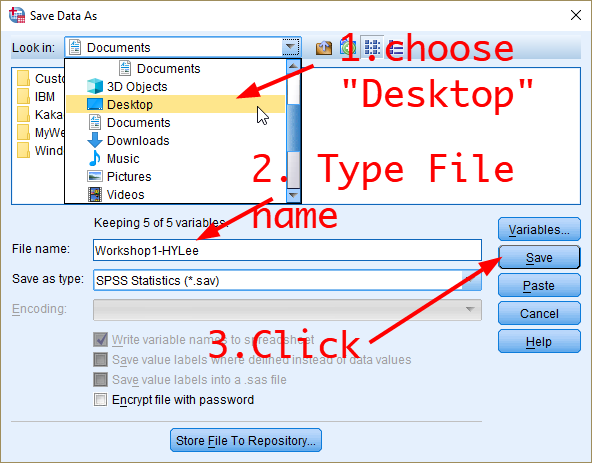
## How to Save and Open Data File

Now, you need to save your work so that you can use it in the future.

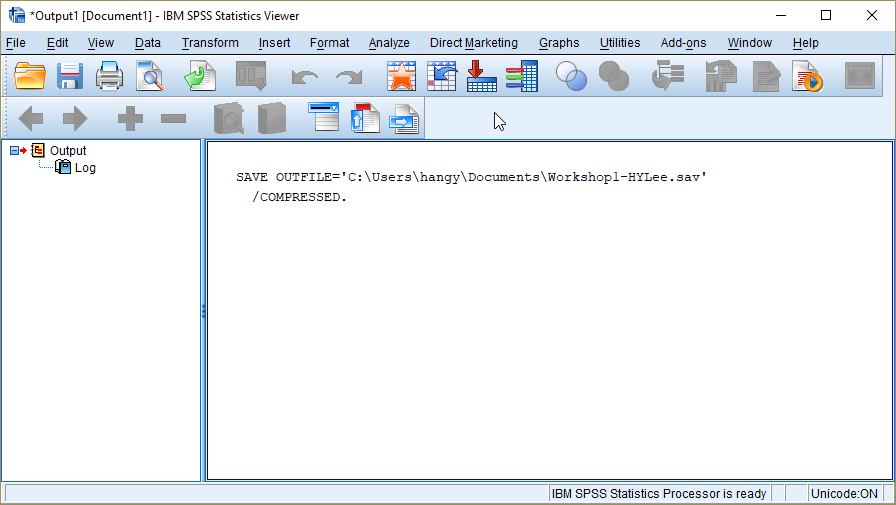
**Go to File > Save**. And you will see a dialogue box where you can specify the file location.



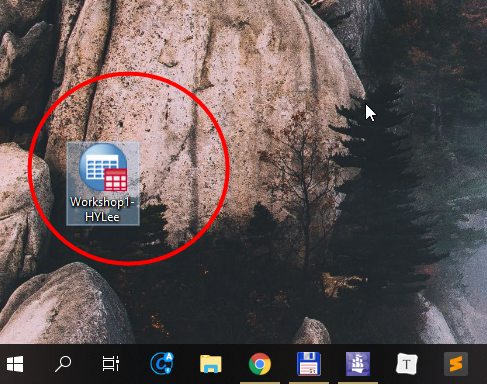
* In the popped-up box, **choose “Desktop” in the box of Look in** (it is the location where your file is saved) or wherever you want.
* **Type a file name in the box of File name**; I typed “Workshop1-HYLee”.
* Then, **click Save**.



Then, you will see the Output1 window confirming that SPSS has saved your file.



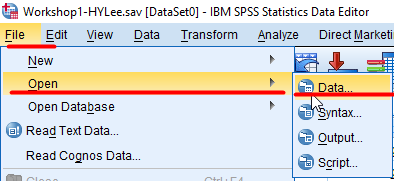
Check your windows Desktop. You will see your saved file.



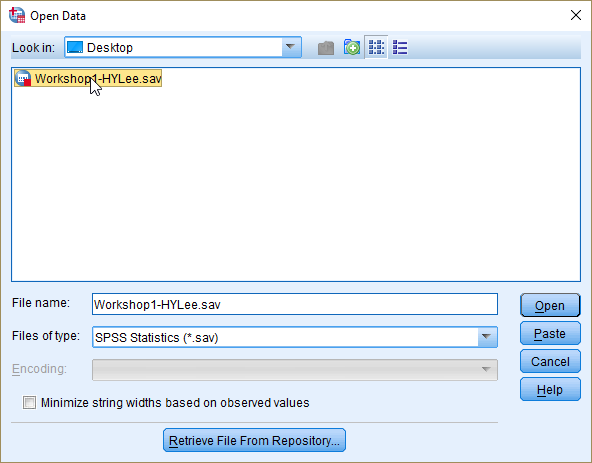
For those who use iLab, **send your work file as an attachment to your email address**. You will need to use this file in the next workshop 2. It is always recommended to save your file frequently, to send it to your email account before you log off, and to store it in your personal computer for the next use.

## Opening the data file

When you want to open SPSS files (\*.sav), go to File > Open > Data.



And choose the file you want to open. The extension of SPSS data files is SAV. This time, we will open the work file you have worked on so far.



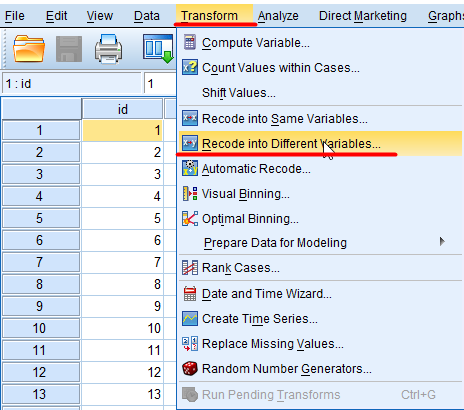
Then, you will see your data in SPSS.

## Generating a new variable of age group (Recoding variables)

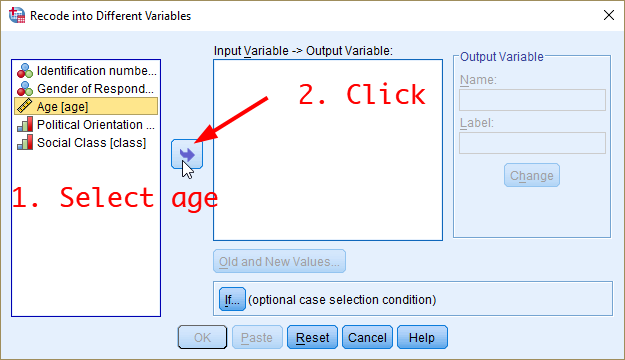
We often need to make a new variable from an existing variable. Suppose we would like to know how many respondents have an age between 10 and 19, 20 and 29, and so forth. The best way to explore the distribution of this grouped age is to use a variable which groups age in such a way. Thus, we will make a new age variable called *agegrp* using the original *age* variable.

The new age variable, *agegrp*, will have nine categories: Less than 20, 20—29, 30—39, 40—49, 50—59, 60—69, 70—79, 80—89 and more than 90. We will make this variable by using the original *age* variable.

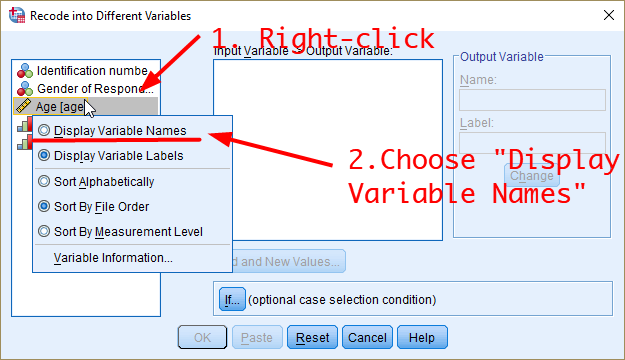
* **Go to Transform > Recode into Different Variables at the top menu**.



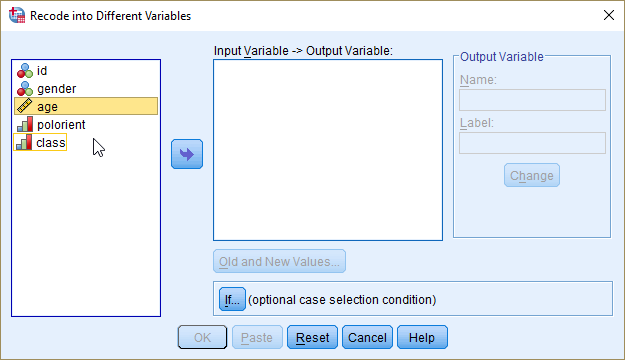
* In the popped-up dialogue box, **select age in the left variable list and move it to the right pane by clicking the right arrow**.



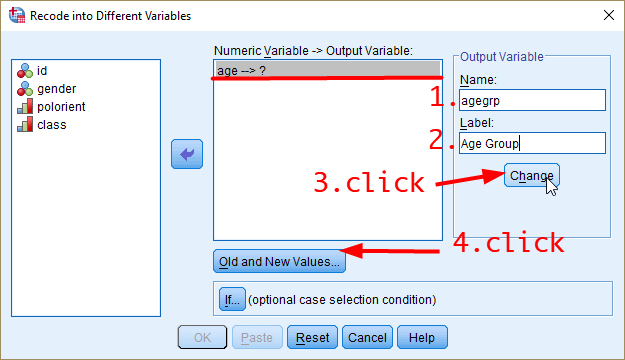
* The above box shows the variable label in the left box. If you want to see variable names rather than variable labels, right-click at any areas in the variable box and then choose “Display Variable Names” in the popped-up window.



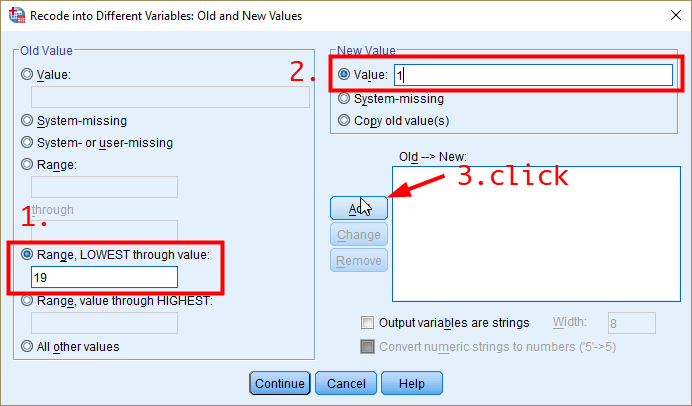
* Then, you will see a list of variable names in the left variable box.



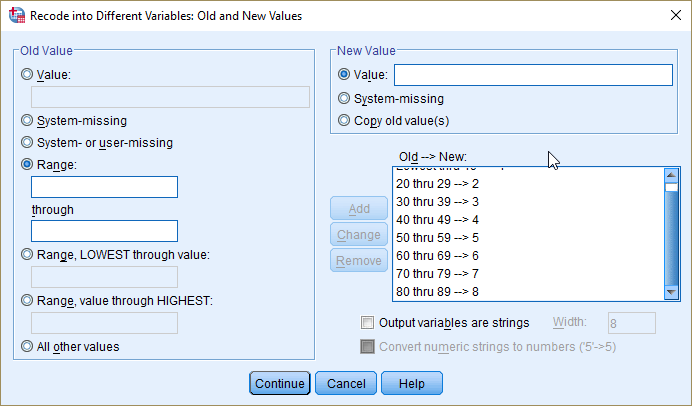
* In the middle white box, you will see “age → ?”. **In the right box titled Output Variable, type a new variable name (*agegrp*) and its label (Age Group)**. Then, **click Change.**



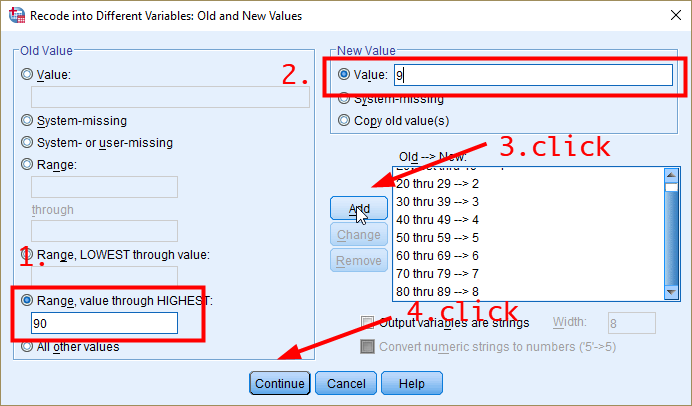
* “age → ?” will be changed into “age → agegrp”. Then, **click Old and New Values**.
* In the Old Value section of the next dialogue box, **select “Range, LOWEST through value”**: and **type 19**. Then, **type 1 in the “New Value” section**. **Click Add**.



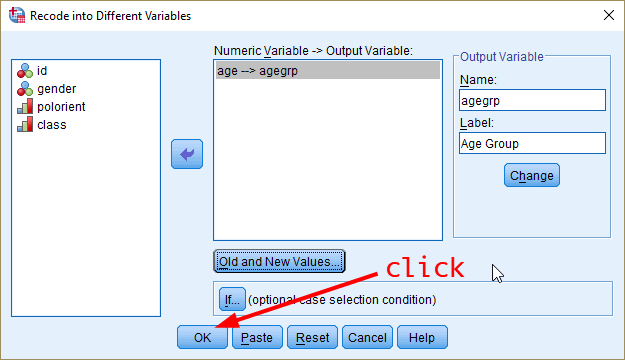
* Next, **select “Range” and type 20 through 29**. And **type 2** in the “New Value” section. **Click Add**.
* **Do the same procedure for other categories**. Then you will see the below dialogue box.



* Finally, **select “Range, value through HIGHEST”**: and **type 90**. Then, **type 9 in the “New Value” section**. **Click Add**.



* Then, **click Continue**. You will be back to the previous dialogue box. **Click OK**

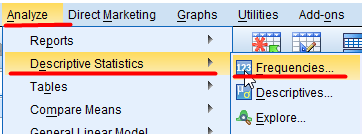


* In the Data View, you will see the newly generated variable, *agegrp*.
* In the Variable View, **assign categories to numerical values in the Values column** as you did for *gender*, *polorient* and *class*.
  + 1 = “Less than 20”
  + 2 = “20—29“
  + 3 = “30—39“
  + 4 = “40—49“
  + 5 = “50—59“
  + 6 =“60—69“
  + 7 = “70—79“
  + 8 = “80—89“
  + 9 = “More than 90”
* **Change the Measure to “Ordinal”**, because SPSS does not have the category of “Interval”. And **set Decimals to 0**.

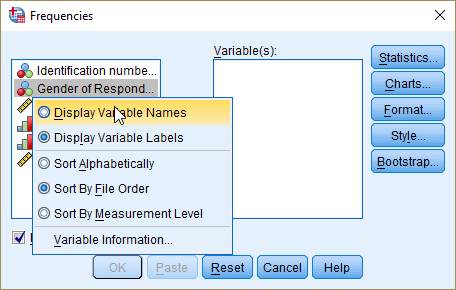
## How to Make a Frequency Table

Now, I will show how to generate a frequency table for a variable.

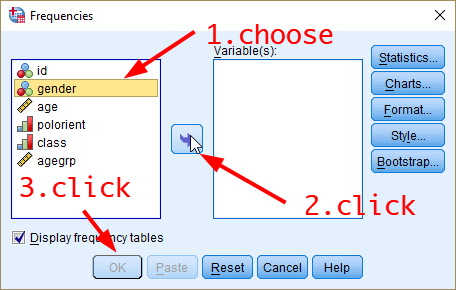
**Go to Analyze > Descriptive Statistics > Frequencies**.



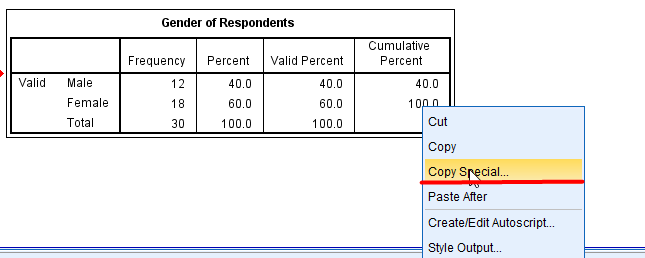
In the dialogue box, right-click at the left variable box. And choose “Display Variable Names”. Variable labels will be changed to variable names.



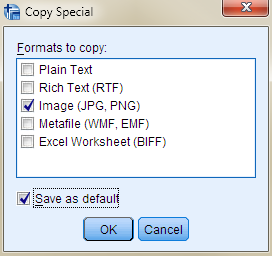
**Choose a variable you want to make a frequency table for**. And **click the arrow in the middle**. The variable will be moved to the right pane. Then, **click OK** in the below.



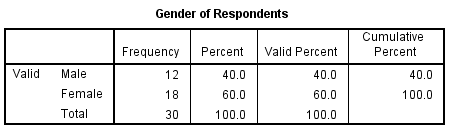
SPSS will open a new Output window where your results are displayed. In the output window, **choose the table that you want to copy and then right-click**. **Select “Copy Special”**.



In the new dialogue box, **choose “Image (JPG, PNG)”**. Then, **Click OK**.



Paste it in the MS Word. You can see the final table in your document. In this table, the second column, Frequency, shows the number of respondents who belong to each category. For instance, the number of males is 12, and that of females is 18. Other statistics (percent, valid percent, and cumulative percent) will be introduced in the Week 3 lecture.



## Workshop 1 Activities

You have so far constructed the dataset manually. Now, we need to check whether your dataset is constructed well without errors. Please answer the following three questions.

1. Make a frequency table of ***polorient***. How many respondents define themselves as Right?
2. Make a frequency table of ***class***. How many respondents define themselves as an Upper middle class?
3. Make a frequency table of ***agegrp***. Which age category has the highest frequency? And report the frequency of this age group.

***NOTE) External students should post their answers to these three questions on the iLearn. These activities will contribute to your workshop participation marks.***