محاضرات في مادة تحليل قواعد المعطيات I التسويقية I

GHICHI ALI

السنة الدراسية 2017 - 2018

تحرير البيانات (Data editing)

- استحداث/ حساب متغير جديد • Creating/calculating a new variable

Suppose that the researcher would like to include an extra column in the example which indicates the 'body-mass index (BMI)'. The BMI is defined as the body weight in kilograms divided by the square of the height in metres. let use the table below:

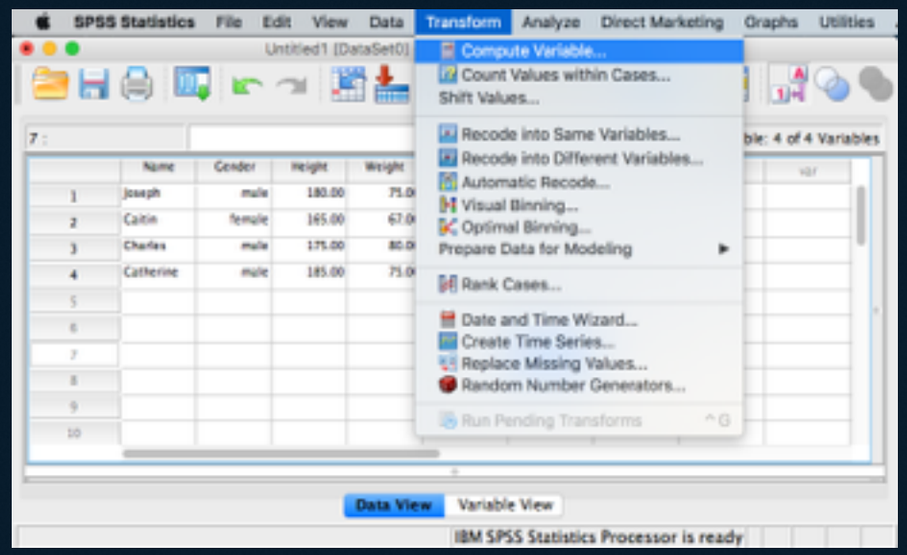
Table 1.1

Name	*Gender*	*Height (cm)*	*Weight (kg)*
Joseph	1	180	75
Caitlin	0	165	67
Charles	1	175	80
Catherine	0	170	70
Peter	1	185	75

The path to be followed to calculate an additional variable is Transform/Compute Variable

تحرير البيانات (Data editing)

- استحداث/ حساب متغير جديد
- Creating/calculating a new variable



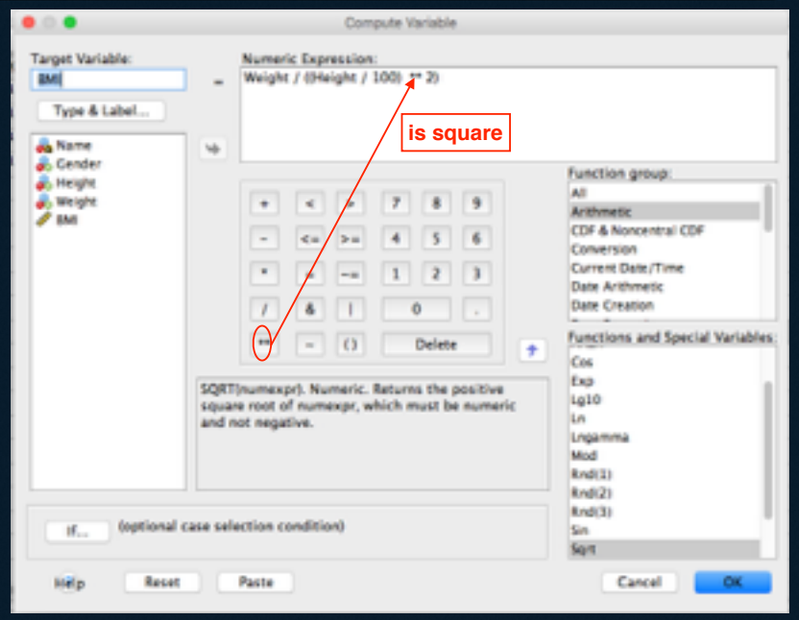
A dialogue window will be displayed

تحرير البيانات (Data editing)

- استحداث/ حساب متغير جديد
- Creating/calculating a new variable

1. type the name of the new variable (BMI in this case)
2. In the 'Numeric Expression' field, type the formula .
3. you may also select the variable names from the left box and click the button).
4. click OK

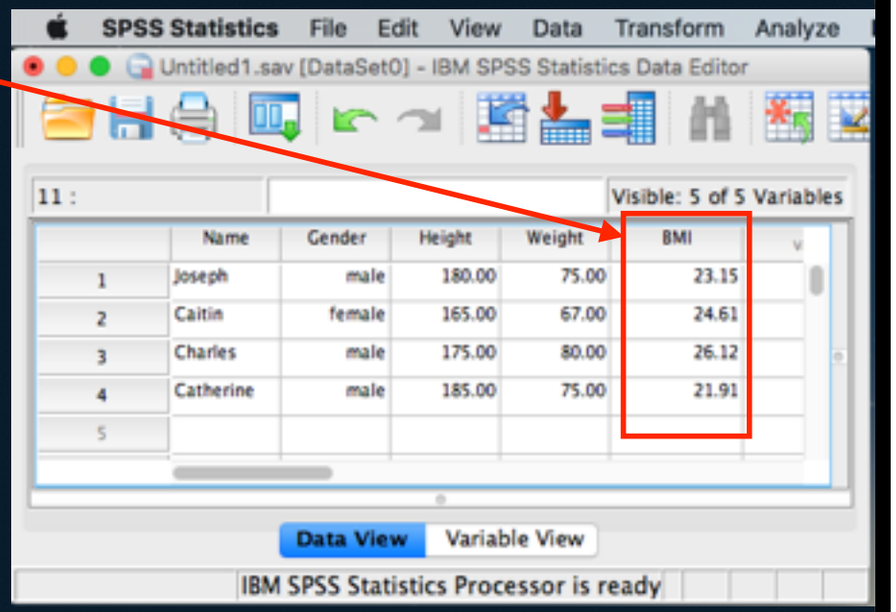
5. The new variable will now be shown in the 'Data View' screen



تحرير البيانات (Data editing)

- استحداث/ حساب متغير جديد • Creating/calculating a new variable

The new variable



تحرير البيانات (Data editing)

- البحث في مجموعة فرعية من الملاحظات • Research on a subset of observations

Selecting cases

- Sometimes a certain subanalysis requires that the analysis to be performed may only be done using a number of specific observations (cases).
- It is then possible to create separate files by deleting the non-relevant observations in the total data file each time, however this method is not efficient.
- There is a procedure in SPSS which may be used to temporarily turn off the observations which the user does not wish to include in the sub-study (thereby not deleting them permanently).

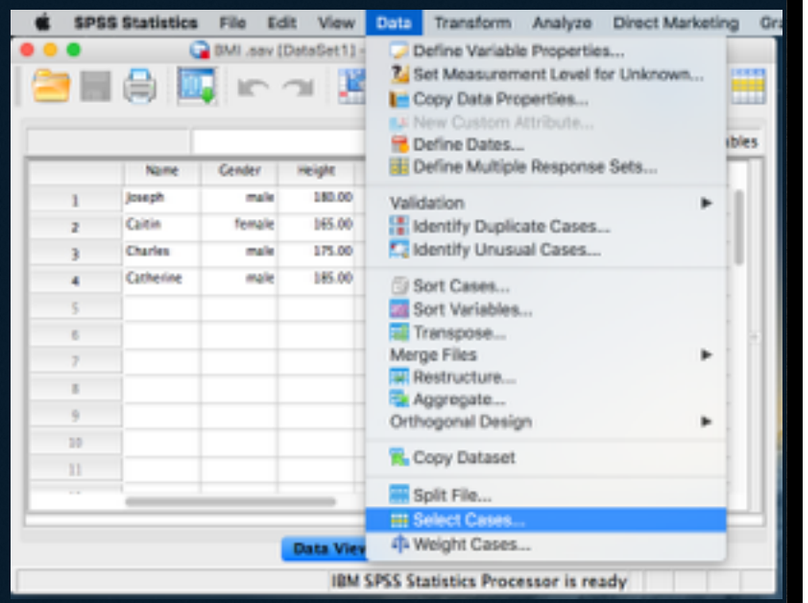
تحرير البيانات (Data editing)

- البحث في مجموعة فرعية من الملاحظات • Research on a subset of observations

Selecting cases

Suppose the researcher in the example would like to select only the male cases (e.g. for a subanalysis), but at the same time, does not wish to permanently delete the other observations (the females).

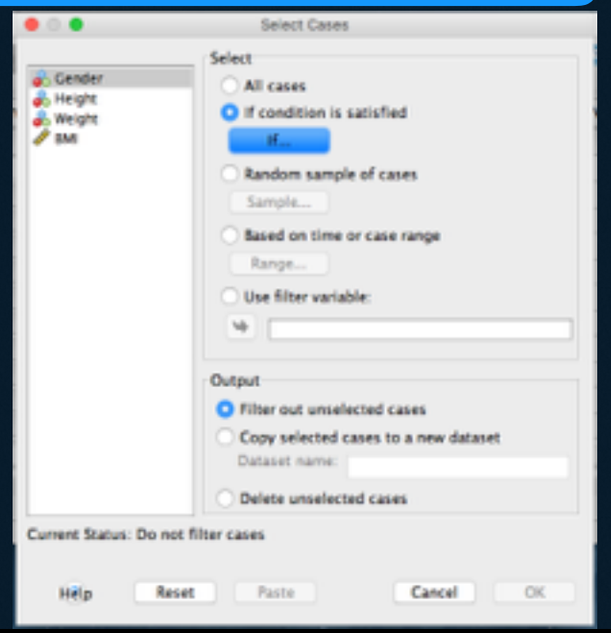
- The path that then must be followed is **Data/Select Cases**
- **Then it will appear a box see in the next slides**



تحرير البيانات (Data editing)

- البحث في مجموعة فرعية من الملاحظات • Research on a subset of observations

Selecting cases The default setting 'All Cases' must be changed by checking the option 'If condition is satisfied' and then clicking the 'If' button
then it will open a box, see in the next slide

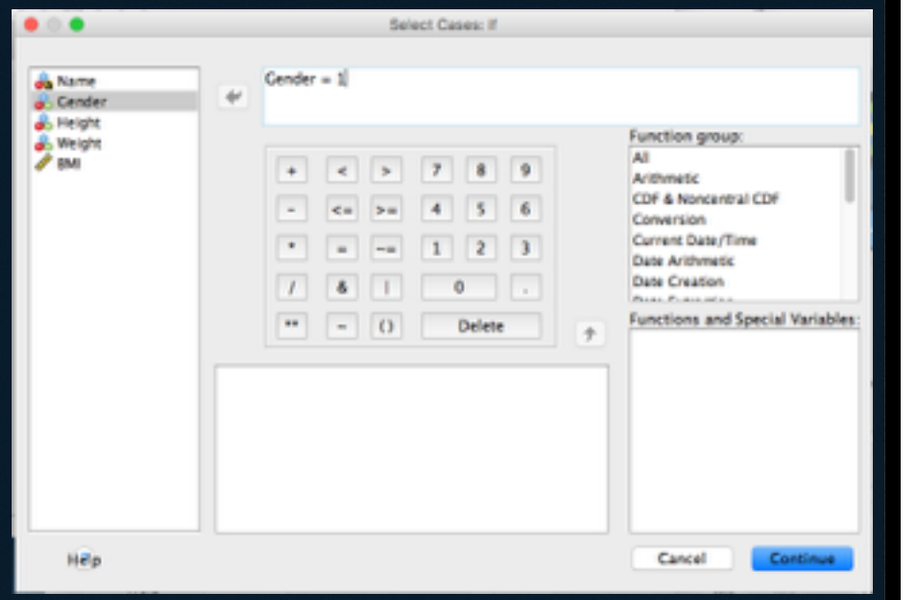


تحرير البيانات (Data editing)

- البحث في مجموعة فرعية من الملاحظات • Research on a subset of observations

Selecting cases

Click on ‘gender’ and make this equal to 1 (1 is the code for the male gender). Then click on ‘Continue’ and then on ‘OK’.

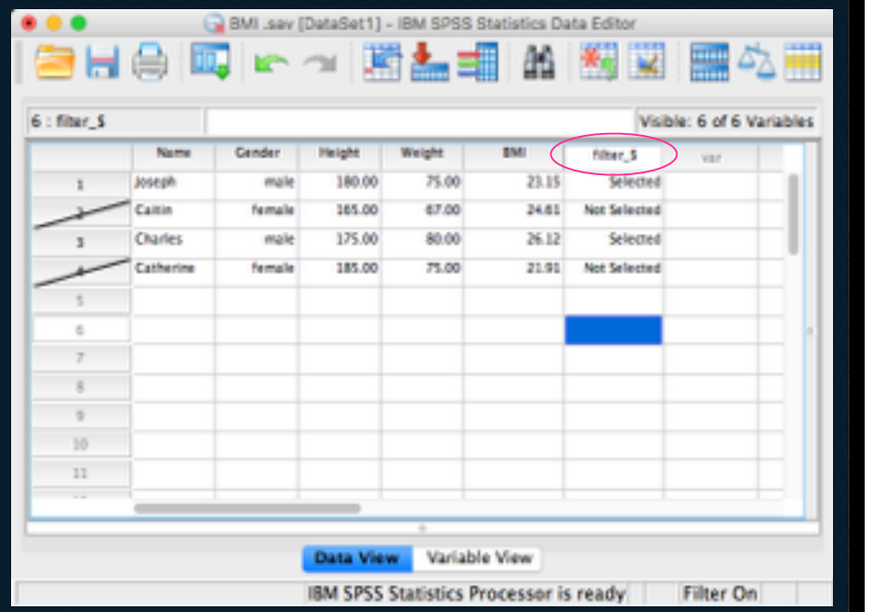


تحرير البيانات (Data editing)

- البحث في مجموعة فرعية من الملاحظات • Research on a subset of observations

Selecting cases

In the 'Data View' window, one will see a slanted line through certain respondent numbers indicating that the observations for the females have been turned off. An extra variable has also been created (filter_$) .



1. When the researcher would like to go back and work on all of the observations, he will once again follow the path Data/Select Cases and recheck the default setting 'All Cases'.

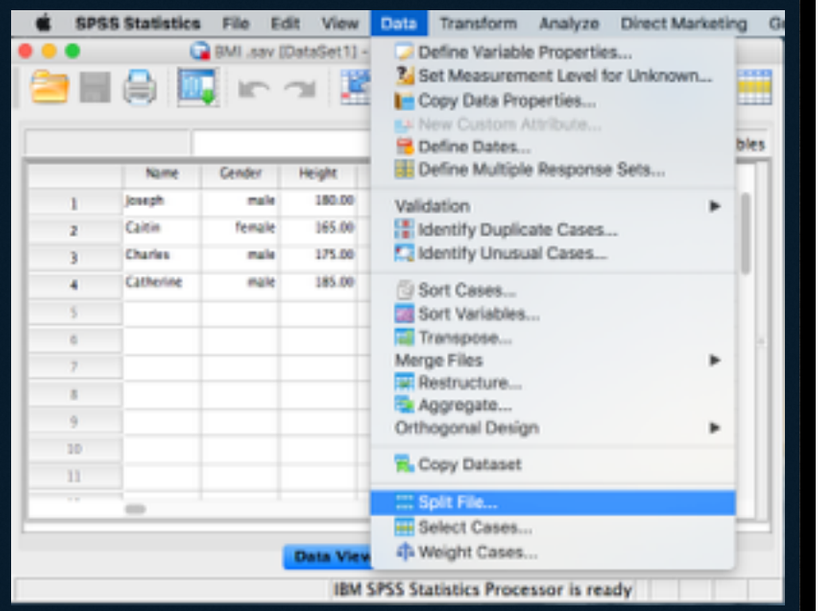
تحرير البيانات (Data editing)

- تقسيم قاعدة البيانات • Splitting the data file

Split file:

This means that when an analysis is performed, the user will obtain the results for the different groups for the variable for which the file has been split.

EXP: Suppose that the researcher wishes to perform separate analyses for the women as well as the men. The path is **Data/ Split file**

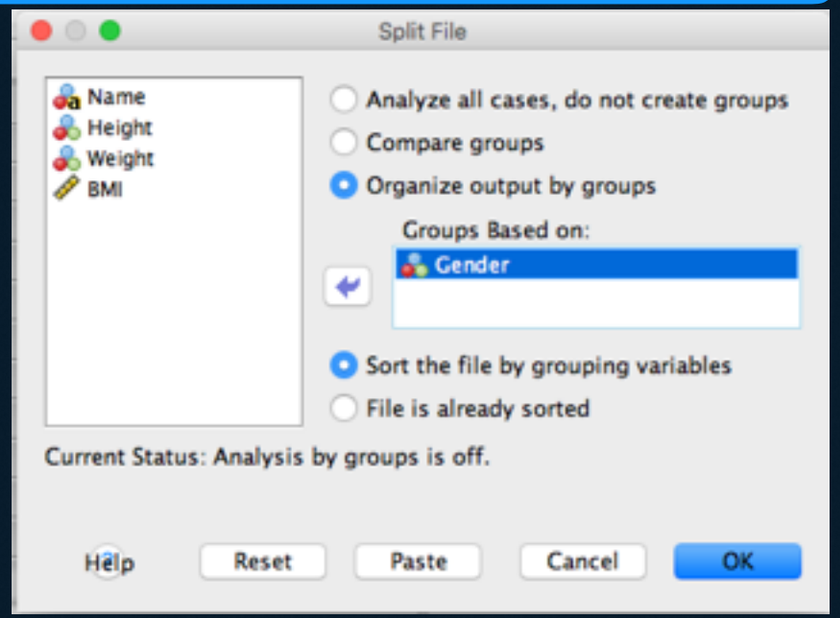


تحرير البيانات (Data editing)

- تقسيم قاعدة البيانات • Splitting the data file

Split file:

Change the default setting 'Analyze all cases, do not create groups' in 'Organize output by groups'. Next, move 'gender' to the 'Groups Based on:' subscreen. Then click on 'OK'.

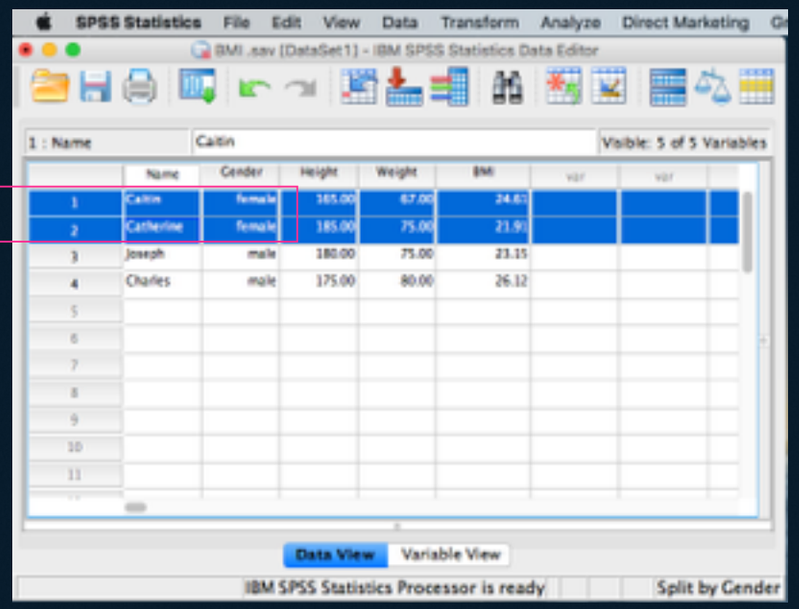


تحرير البيانات (Data editing)

- تقسيم قاعدة البيانات • Splitting the data file

Split file:

You can now see that the observations have been ranked by 'gender' in the Data View tab. the output for the indicated analysis will be grouped separately for men and women.



تحرير البيانات (Data editing)

- إعادة ترميز المتغيرات • Recoding variables

Recoding variables:

Sometimes you will want to transform a variable by grouping its categories or values together. For example, you may want to change a continuous variable into a categorical variable, or you may want to merge the categories of a nominal variable. In SPSS, this type of transform is called recoding.

In SPSS, there are three basic options for recoding variables:

1. Recode into Different Variables
2. Recode into Same Variables
3. DO IF syntax

1. In general, it is best to recode a variable into a different variable so that you never alter the original data and can easily access the original data if you need to make different changes later on.

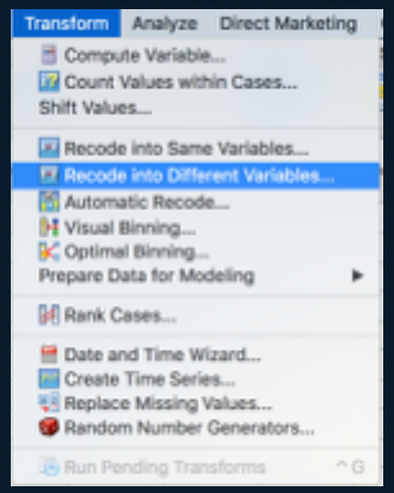
تحرير البيانات (Data editing)

- إعادة ترميز المتغيرات • Recoding variables

Recoding variables: Recoding into a different variable transforms an original variable into a new variable. That is, the changes do not overwrite the original variable; they are instead applied to a copy of the original variable under a new name.

To recode into different variables,
click Transform > Recode into Different Variables.

The Recode into Different Variables window will appear. See the example in the next slide.



تحرير البيانات (Data editing)

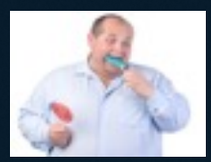
- إعادة ترميز المتغيرات • Recoding variables

Recode into Different Variables: Let's say that these five people must complete a questionnaire. we assume that this questionnaire consists of three questions (statements) in which their preferences regarding candy are being studied. The three statements must be evaluated on a 7-point scale, ranging from 'totally disagree (1)' to 'totally agree (7)'.

Question 1: When I watch television in the evening, I eat candy on a regular basis.
Question 2: If I'm hungry between meals, I will eat fruit more often than candy.
Question 3: I always like to add extra sugar to my dessert.

Their answers are shown in this table:

Name	Question 1	Question 2	Question 3
Joseph	7	3	6
Caitlin	2	5	3
Charles	5	1	5
Catherine	3	6	2
Peter	6	2	6



تحرير البيانات (Data editing)

- إعادة ترميز المتغيرات • Recoding variables

If the researcher wishes to perform an analysis of this data (e.g. calculate an 'average for candy preference'), he must first determine whether the questions were all scaled 'in the same direction'.

	Name	Gender	Height	Weight	BMI	Question1	Question2	Question3
1	Joseph	1.00	180.00	75.00	23.15	7.00	3.00	6.00
2	Caitin	.00	165.00	67.00	24.61	2.00	5.00	3.00
3	Charles	1.00	175.00	80.00	26.12	5.00	1.00	5.00
4	Catherine	.00	185.00	70.00	24.22	3.00	6.00	2.00
5	peter	1.00	185.00	75.00	21.91	6.00	2.00	6.00

- Take question 2 for example.
- Question 2: If I'm hungry between meals, I will eat fruit more often than candy.
- question 2 is not scaled in the same direction as questions 1 and 3 and for this reason needs to be recoded.

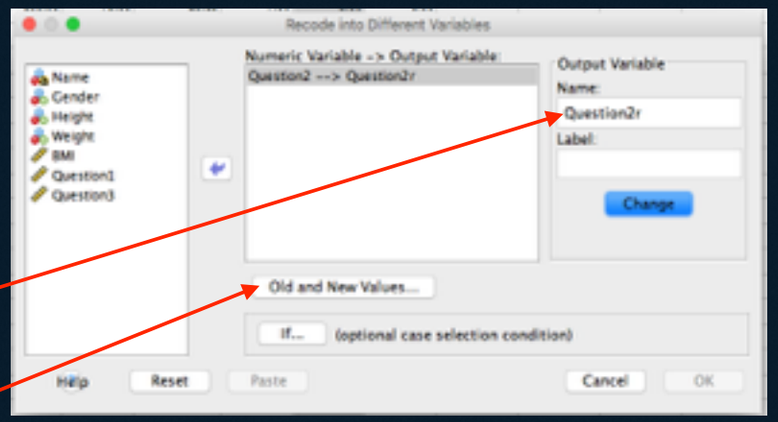
1. A high score indicates that these people are not so quick to reach for candy, while a high score for questions 1 and 3 indicates that there is a great preference for candy.

تحرير البيانات (Data editing)

- إعادة ترميز المتغيرات • Recoding variables

Go to Transform/Recode into Different Variables which will bring up this subscreen.

1. Click on 'question2' in the list of variables on the left and click
2. Under 'Output Variable' enter the name of the recoded variable (question2r) and click 'Change'
3. Click on the 'Old and New Values' button
4. you see a dialogue window
5. see in the next slide

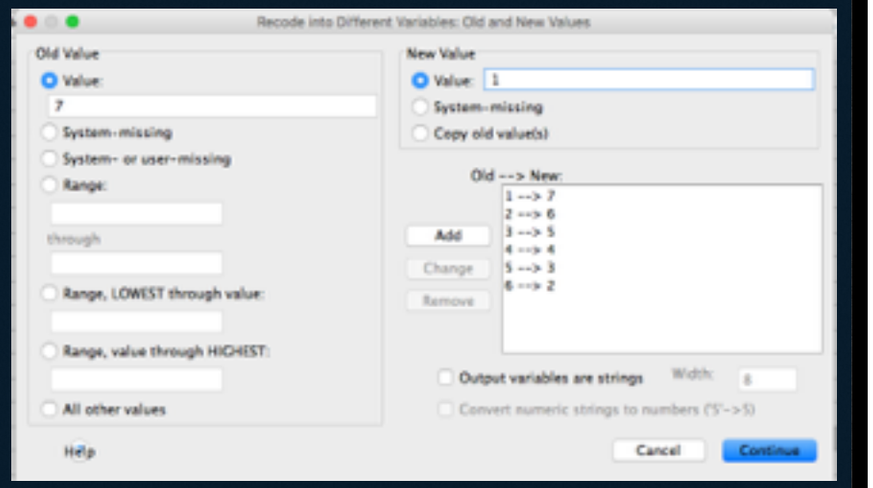


تحرير البيانات (Data editing)

- إعادة ترميز المتغيرات • Recoding variables

For each value to be recoded, the researcher must input the old and the new value.

1. For 'Old Value', fill in the value to be changed and under 'New Value', type the new value (1). Next, click on the 'Add' button
2. Click Continue
3. then Click Ok

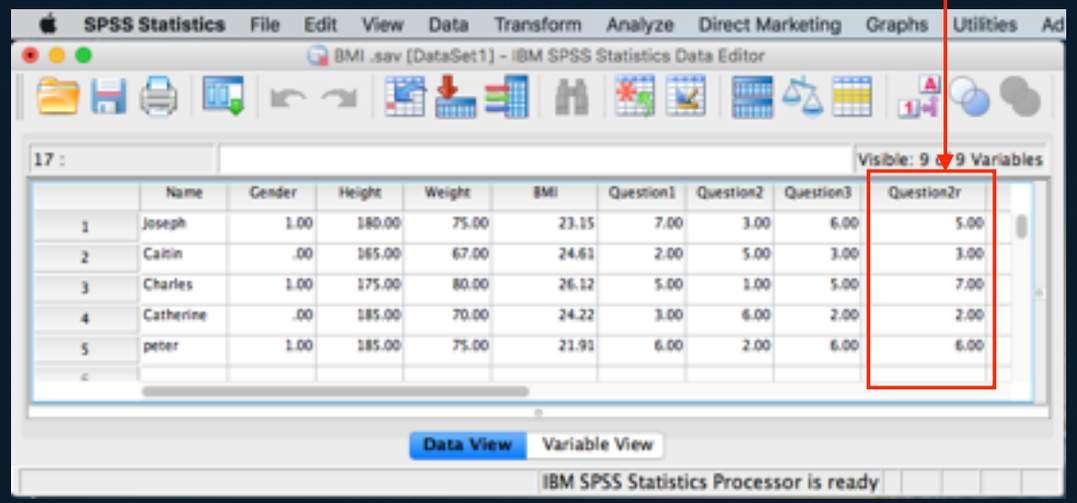


4. An extra variable with the recoded values has been created in the 'Data View' tab. See in the next slide

تحرير البيانات (Data editing)

- إعادة ترميز المتغيرات • Recoding variables

Now, a useful average of the variables 'question1', 'question2r' and 'question3' may be calculated if desired.



محاضرات في مادة تحليل قواعد المعطيات I التسويقية I