

SPSS handout 2: Analysing a Subset of Cases

Richard Buxton and Rosie Cornish. 2007.

Sometimes we want to carry out a particular analysis for a subset of cases (e.g. a particular age group), rather than for the entire data set. Alternatively, we may want to carry out a particular analysis for a number of different subsets separately (e.g. carry out a separate analysis for males and females). This handout describes how to do these tasks in SPSS.

1 Select Cases

Using this option in SPSS you can carry out an analysis on a particular subset of data. The subset is defined by specifying a particular value (or range of values) for one of our variables. (Note that this option can also be used to delete particular observations / subsets of cases.)

- 1 Choose **Data...Select Cases**.
- 2 Under **Select** choose **If condition is satisfied**.
- 3 Click on **If**. In the blank box provided type in an expression specifying the cases you wish to select — e.g. Agegroup=1 — then click on **Continue**.
- 4 In the **Output** section, click **Filter out unselected cases**. SPSS will not alter the original data but will create an additional variable called 'Filter'. This will take the value 1 if a case is in the subset that you want to be analysed and will take the value zero otherwise. You can IGNORE this variable — it is simply a variable used by SPSS so that it knows which cases to include in any subsequent analysis.
- 5 Click **OK**.
- 6 Any analysis you do after this point will be done **ONLY** for the selected subset.

If you want to return to analysing the complete data set, you will need to turn the filtering off as follows:

- 1 Choose **Data...Select Cases**.
- 2 Now do ONE of the following: (1) Under **Select** choose **All cases** then click on **OK** OR (2) Click on **Reset** then **OK**.

2 Split File

This option can be used to carry out the same analysis on a number of different subsets of your data. Once you've split the file, any subsequent analysis will be carried out separately for each subset.

- 1 Choose **Data...Split File**.
- 2 Select either **Compare groups** or **Organise output by groups**. The first of these options puts corresponding pieces of output for the different subsets together, while the second option lists all the output for the first subset, then all the output for the second subset, and so on.
- 3 In the **Groups based on** box, select the variable you are using to define the subsets.
- 4 Check that your file is sorted by your grouping variable. If not, click on **Sort file by grouping variable**.
- 5 Click **OK**.
- 6 Any analysis you do after this point will be done for subsets separately.

If you want to return to analysing the complete data set, you will need to turn the split file off as follows:

- 1 Choose **Data...Split File**.
- 2 Now do ONE of the following: (1) Choose **Analyse all cases, do not create groups** then click on **OK** OR (2) Click on **Reset** then **OK**.