

EXERCISE 5

Preparing your data (continued)

Opening SPSS

For this Exercise, you should have available the corrected merged data set that you corrected in the course of the previous Exercise and saved as *Questionnaire Data (corrected)*. Open SPSS with this data set in the usual way.

Describing interval data

The next part of the Exercise is to explore the *Age* variable by tabulating a range of statistics and drawing a histogram with a superimposed normal curve.

Choose **Analyze**→**Descriptive Statistics**→**Frequencies...** to see the **Frequencies** dialog box. Transfer *Age* to the *Variables* box.

Click **Statistics...** to see the **Frequencies: Statistics** dialog box. Choose Quartiles, Mean, Median, Std. Deviation, Range, Minimum and Maximum. Click **Continue**.

Back in the **Frequencies** dialog box, turn off **Display Frequency Tables**.

Now choose **Charts**→**Histograms** and click the **With Normal Curve** option. Click **Continue** to return to the **Frequencies** dialog box.

- **Edit the histogram to make it suitable for black-and-white printing**
- **Print the histogram.**

Manipulation of the data set – transforming variables

It is sometimes useful to change the data set in some way. For instance, in the current data set, some people have entered their weight in stones and pounds, others in kilograms. Likewise, some people may have entered their height in feet and inches (or just inches), others in metres. In order to have useful data on weight and height, you must use the same units of measurement. In this Exercise we shall adopt metric units (kilograms and metres). This will necessitate converting any other measurements into metric measurements. Use the **Compute** procedure (Section 4.4.2) by selecting **Transform**→**Compute Variable...** to see the **Compute Variable** dialog box. In the **Target Variable** box, type the name of the variable (*Kilos*) which contains the kilograms data. In the **Numeric Expression** box, enter the conversion function: entering **(Stones*14 + Pounds) * 0.453** to convert pounds to kilograms. (One pound is 0.453 kilograms and stones convert to pounds by multiplying by 14.) Note that, in computing, the symbol * is used for multiplication. Do not click **OK** yet!

There remains one further problem: what about those cases whose weight is already in *Kilos* and do not have any values in the *Stones* and/or *Pounds* variables? If the formula above were to be immediately applied, these people would end up with no values in the *Kilos* column.

To convert only the cases with stones and/or pounds measurements, you must select the **If** box in the **Compute Variable** dialog box, then the **Include if case satisfies condition** box

and enter the following expression **stones >= 0** which tells the program to calculate the kilograms if the entry in *Stones* is 0 or greater than 0. Select **Continue** and then **OK**.

You will receive a message which asks **Change existing variable?**. Select the **OK** option. Now the program will calculate all the missing *Kilos* data and enter them in the data set. Check that it has done this. Save the file using the **Save As** option, giving the amended file a new name (e.g. **Metric Data**). This ensures that you still have a copy of the old file, in case you have made any mistakes in calculation and you wish to retrieve the old data at some future time. (It is often regarded as a safer procedure to recode data into a new variable since it allows one to check that the correct recoding procedure has been requested. This has not been done here because we need to preserve the values in *kilos* already present for some cases.)

Now do a similar conversion for the height data, converting feet and inches to metres. To do this you will need to know that there are 12 inches in a foot and 1 inch is 0.0254 metres. Work out a conversion factor with this in mind. Remember to change the condition to **Feet >=0**. When you have converted the height data, save the file again, this time by simply clicking **Save** rather than **Save As** since you have already nominated **Metric Data** as the new file.

Describing interval data – means of cases categorised by a grouping variable

You can obtain a table of means for one variable at different categories (or combinations) of another variable (or variables). Use the **Means** procedure described in Section 4.3.2 to obtain a two-way table of means for *Metres* by *Sex*. Then use the same procedure to obtain a three-way table of *Kilos* by *Sex* by *Faculty*. (Look carefully at Section 4.3.2 to see how to layer the variables, using **Next**, to produce the three-way table.)

- **Print the output of this exercise.**

Finishing the session

Close down SPSS and any other windows before logging out of the computer.