

## Exercise 4 Data Handling

These exercises have been prepared for use in conjunction with Chapter 4 of the 4<sup>th</sup> edition of “SPSS for Psychologists” by Brace, Kemp and Snelgar (2009)

Load the data file **EX4Q1.sav**. This data file contains data collected using a simple questionnaire. Variables GK1 to GK20 code the responses to a series of general knowledge questions. For each question, a correct answer is scored 1 and an incorrect answer 0. Variables LIK1 to LIK10 code the response to a series of Likert-type items. For each of these items the participants rated their level of agreement with a statement using a seven point scale ranging from Strongly Disagree (1) to Strongly Agree (7).

1. Check the data file, looking at both the Variable view and the Data view. Check for errors in the variable settings and in the data file. Correct any errors you find.
2. The experimenter now realises that she has made an error – items GK2 and GK 8 have been scored incorrectly. Use the Recode command to reverse the scoring of these items.
3. One of the errors in this data file is that for each of the Variables GK1 to GK20 and LIK1 to LIK10 the Missing value has been set to 9. However, inspection of the data file will show that the researcher has actually used the value 99 to represent missing values. If you haven't already done so, correct this error. (Note, there are at least two different ways to do this).
4. Calculate the following sub-scales:
  - a.  $GKSub1 = GK1 + GK4$
  - b.  $GKSub2 = GK1 + GK2 + GK3 + GK5 + GK8 + GK11$
  - c.  $GKRatio = GKSub2 / GKSub1$
  - d.  $LIKSub1 = LIK1 + LIK2 + LIK4 + LIK6$
5. Look at the new variables you have computed. You will notice that for some participants the value of some of the variables has not been computed, and instead the cell contains a dot (“.”).
  - a. Why is this?
  - b. Compute new versions of these variables which overcome this problem. Name these new versions “NewGKSub1” etc. (TIP: Give careful thought to how you calculate the new value for LIKSub1).
  - c. Look at the new variable NewGKRatio. You probably still have some empty cells (containing a dot). Why is this?

6. Create a new variable that defines whether each participant is older or younger than the mean age.
7. Split the data set on the basis of this new variable and use the Descriptives command to produce descriptive statistics and charts of GKSub1 for each of the two age groups.
8. Undo the Split and then Select only the female participants. Now recalculate the statistics and charts for GKSub1.