



Chapter 3: Managing Your Data

Discussion Points

1. A sample of 2000 will result in a huge volume of data to manage. If everyone responded to your questionnaire, and there were 20 coded variables in each record, this would result in 40,000 cells of data to enter accurately into your database. However, the reality is that few surveys will result in a 100% response rate. All the same, the sheer amount of data you will collect will present you with a considerable amount of work to ensure it is all entered accurately into your database. Once the data has been entered, it must be checked for accuracy. This whole task is very time consuming and you must build this into your survey timetable. There are several points to consider:
 - How accurate is the person entering the data?
 - Are you using database software that has an easy to use data entry screen?
 - Will the software allow for the data to be checked easily?
 - How much time will all this take?
 - If this survey is being paid for by someone else, has the cost of data entry been accounted for?
2. The question of computer software is linked to the first discussion point. I have a preference for using different software applications for different stages of the process. For example, I use Microsoft Access for data entry and checking. This program allows me to configure each field to accept only certain codes. This helps to promote accuracy because an incorrectly keyed entry is more likely to generate a rejection warning. This isn't foolproof, but it does have a considerable effect on promoting accuracy. Most database packages will have this facility. For data analysis itself, I use SPSS because it was designed for the task, but a powerful spreadsheet package such as Microsoft Excel is very capable of carrying out data analysis. Excel's graphics are also very good so the results of your analysis can be displayed very effectively. If you are using different packages for different tasks you need to ensure that data can be transferred between them. Current commercial software packages are generally able to handle each other's data. This may be a direct ability, or an indirect one in which the data needs to be 'exported' into a format that both are able to recognise. For example, Access will export data tables in Excel format, and SPSS is able to read Excel files.

What I have written so far looks at some general points to consider, and here is a checklist of the main considerations:

- Can your choice handle the data entry and data analysis parts of the survey?
- If you are using different applications for each process, can you transfer the data between applications with ease and in a way that will not result in errors (i.e. sometimes, data transfer results in the loss of some formatting that will lead to the data being misread)?
- Can your applications display the results of your analysis in a suitable format, or will you have to type it out again? This is not terminal, but it is extra work.