**Quick Tips
22**

Quick Tips

Make Certain Your Electronic Data Are Accurate

When you enter your data into Excel, Access or another software program, make certain that what is entered accurately reflects the data you collected and the coding decisions that you and your support staff made.

The dangers of not double-checking your data files can include:

- Getting results that do not make sense
- Getting results that you cannot support with the data you collected
- Wasting time and energy on analyzing information that may lead you to draw inaccurate conclusions

To make certain you have entered electronic data accurately, use one of the following methods or combine methods. Your goals are to make certain that the information is accurate as entered from the original questionnaires or interviews and treated consistently across all questionnaires or interviews.

- **At a minimum, scan the data for values that do not belong.**

  For example, Does a "6" appear where the maximum value available is a "5"? If there are values that lie outside the realm of what is possible given a particular range of responses, return to the original questionnaire. What did the respondent say or write? Does this "impossible" value seem to appear several times in the data? If so, check several of the original questionnaires. Does an "impossible" value seem to arise often in relation to a specific individual doing the data entry? If so, check with that individual to see if he or she made a judgment call on a response category, such as adding a category that did not exist on the questionnaire.

- **If it is appropriate to the level of accuracy you desire, double-check each entry.**

  One easy way to speed this up is to have one person read the values aloud and another person check the screen as each number is read. You could choose to examine the answers for each of the questionnaires or a smaller group of them. If a smaller group is examined, say five or ten percent of the total, then you also accept a certain level of error in the finished product. If more than one person has done data entry and they have not used consistent decision rules, you may have to fix any mistakes.

- **Be aware of trouble spots in data files, which can include areas where some respondents are asked to skip specific questions and did not do so.**

  When a respondent answers a question that exempts him or her from one or more questions that follow, the questionnaire is using what is called a "skip pattern."

  It is important to look for inconsistencies, such as respondents who ignored the intended skip pattern and answered a question that did not apply to them. Decide how to handle the data that results from respondents ignoring directions on skip patterns. Be consistent across all the questionnaires. Usually you will want to consider their inappropriate answers as "missing" data.

Watch for places where respondents do not provide an answer, also known as "missing" data. Decide how you will code "missing" data and code it consistently for each variable. In Excel, it makes sense to leave missing values as blanks. In SPSS, it makes sense to assign a number to the missing value and later define that value in the program.

- **Be familiar with the conventions of the data entry program you are using.**

  Some may set a default to define the information as "string" data (i.e. words, text, language, etc.). Unless you tell the program the information is a set of numbers, it will not treat it as numeric. This causes a problem when you ask a mathematical function or when the data is input into a statistics package. The computer will not be able to do the analysis. Be sure to define the values as numeric if that meets the needs of the software you are using. Each software has its own conventions so read the manual and use the "help" features.

- **Keep track of the decisions you make by writing them down.**

  Over time, you could forget the decisions you made and why. Your written record is a reminder and helps you answer the questions of others.

Further reading:

Rose, D., & Sullivan, O. (1993). *Introducing Data Analysis for Social Scientists.* Buckingham, UK: Open University Press.