

Survey data is the question answers, such as 'yes' or 'no' or perhaps a number, where a person has ranked a question on a scale. The survey data output will depend on the way in which the survey was constructed, it will be shaped by the survey questions asked, the format of the survey itself and the method in which data was collected. For example, if the survey was completed by the respondent, in a written form, then you will have a collection of written documents which require analysis of the question answers. If the survey has been completed by a researcher, then a more sophisticated method of data collection may have occurred e.g. tallies and counts of responses. If using an internet or email survey, a computer programme may have collected the data in a format which can easily be analysed. Consideration of the output needs to be made at the outset of the process, and time considerations need to be given as to how this data will be collected and analysed. Before you can input your data in a computer program or application you will need to undertake a process of coding. This involves assigning a code (often numeric) to each possible answer in your survey. So if question 1 in your survey asked the gender of the respondent, you may seek to code the answer 'male' with the number 0, while you may seek to code the answer 'female' with the number 1. Establishing these 'codes' on the distributed questionnaire can help at data entry time, but obviously has the downside of putting numbers on the questionnaire that are of no relevance to the respondent and therefore could make the questionnaire look more confusing than it needs to. Internet based survey tools can distribute your survey via email and also collect your results, often allowing you to view your results as they are collected in real-time. You can download live graphs and charts of the responses, and often filter the responses and dig down to get individual responses. While this offers significant benefits there needs to be careful consideration of the pros and cons of email or internet surveys and whether this method of collection suits the population you are targeting.

Microsoft Excel is useful for data summary, presentation, and for other basic statistical analysis. The program provides a set of data analysis tools called the Analysis ToolPak which you can use to save steps when you develop complex statistical analyses. You provide the data and parameters for each analysis and the tool uses the appropriate statistical macro functions and then displays the results in an output table. Some tools generate charts in addition to output tables. The Analysis ToolPak is not loaded by default, instructions for installing it, along with guides on how to use it can be found on the Microsoft website. SPSS is among the most widely used program for statistical analysis in social science. This is a data analysis package for quantitative research. It is particularly useful for the analysis of survey data as it covers a broad range of statistical procedures. There are other packages available such as SAS, Stata or Minitab however all are expensive to purchase, especially if only to be used for a one off survey. It may be possible to work with an academic institution to utilise their statistical packages, and organisations such as the Cathie Marsh Centre for Census and Survey Research (CCSR) provide training on the use of these packages.

Before you can input your data in a computer program or application you will need to undertake a process of coding. This involves assigning a code (often numeric) to each possible answer in your survey. So if question 1 in your survey asked the gender of the respondent, you may seek to code the answer 'male' with the number 0, while you may seek to code the answer 'female' with the number 1. Establishing these 'codes' on the distributed questionnaire can help at data entry time, but obviously

has the downside of putting numbers on the questionnaire that are of no relevance to the respondent and therefore could make the questionnaire look more confusing than it needs to.

A variable is any measured characteristic or attribute that differs for different subjects. Quantitative variables are measured on an ordinal, interval, or ratio scale, whereas qualitative variables are measured on a nominal scale (note in SPSS the Interval and Ratio levels are grouped together and called scale). There are a range of variables that need to be understood, dependent/independent, controlled/continuous/discrete in the application of statistical tests. The independent variable answers the question “What do I change?”, the dependent variable answers the question “What do I observe?” and the controlled variable answers the question “What do I keep the same?”. A variable which can have any numerical value is called a continuous variable (e.g. time). A variable which can only have whole numbers (integers) is called a discrete variable (e.g. the number of people in a group). It is important to understand the variable you have for analysis of data in statistical packages such as SPSS.