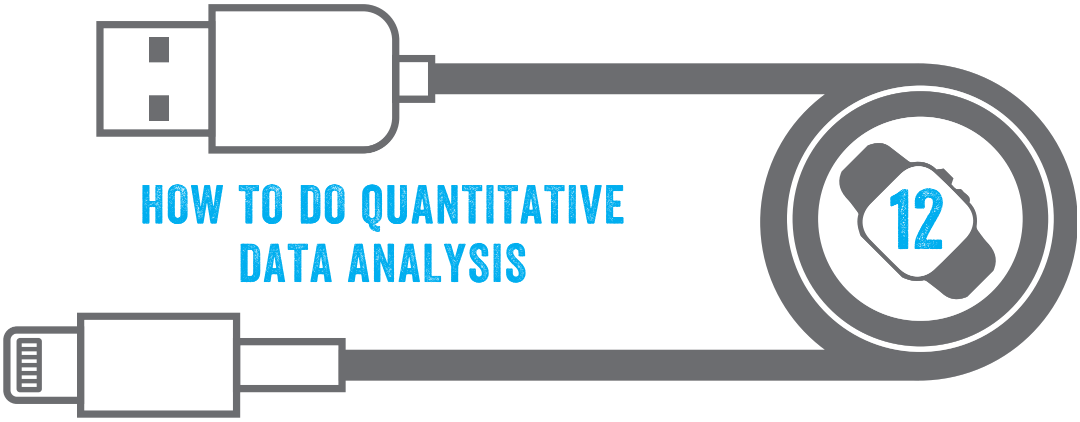


HOW TO DO QUANTITATIVE DATA ANALYSIS

THIS CHAPTER COVERS THE FOLLOWING TOPICS

- How to use content analysis to create quantitative data out of qualitative content.
- How to use quantitative methods to describe patterns and make comparisons in your data.
- How to use the chi-square statistic to evaluate the association between two categorical variables.
- How to use Pearson's correlation, and its non-parametric alternative, to investigate relationships between two interval variables.
- How to use the *t*-test, and its non-parametric alternatives, to compare means.
- How to decide which advanced statistical options may be worth pursuing to answer your research question.

12.1 INTRODUCTION

It is not surprising that quantitative methods in social research have gained a strong position in many countries, given that generalization and prediction are highly prized by governments, businesses and other outcome-oriented institutions. Counting and comparing are the primary activities in quantitative analysis. Quantitative analysis tends to prioritize objectivity and impartiality, generally taking the view that truths exist independently of human opinions about them and can be discovered through empirical observation and/or measurement. Quantitative analysis follows standardized, transparent approaches that can be replicated by others. This chapter guides you through the process of developing your quantitative analysis step by step, from selecting an appropriate type of analysis to implementing it with *statistical software*.

While primarily focused on detecting overall patterns rather than individual variation, statistics can be used to uncover patterns within and across different groups. That is, after finding the overall patterns for the population, the next question is what kind of patterns there might be below that population level. You can achieve this by running analyses comparing people with different demographic characteristics (e.g., did men and women respond differently? Did people's responses vary with age?). You can also separate the sample based on the outcome data collected: for example, you can separate your data into deciles (dividing into 10 groups) to see what patterns apply to each 10 per cent step within your sample (e.g., bottom 10%, next 10%, etc.).

You can answer some research questions by adding up the number of individuals with a certain characteristic or by calculating how often something is said or done. Answering this kind of research question requires what is known as *univariate* (one *variable*) statistics. Using such one-variable statistics, you can:

- Find the frequency, that is, a simple count.
- Identify the proportion of times something occurs, for example, using a percentage or ratio.
- Calculate a measure of *central tendency*, such as a *mean* or *median*.
- Calculate the level of variability in scores, for example by using *standard deviation*.

Other research questions require a comparison between two or more categories of participant or points in time. For example, you might want to know whether there are any differences in your data on the basis of age, gender, ethnicity, income, psychological orientations, etc. You can answer this kind of research question by using:

- Content analysis to quantify qualitative data.
- Crosstabular analysis to evaluate whether two variables are related.
- *Correlation* analysis to evaluate the strength of relationships between variables.
- Comparing means to assess whether two groups are significantly different in some way (*t-test*).

This chapter builds on guidance we have provided in previous chapters about quantitative sampling and knowledge claims (Chapter 5) and survey design (Chapter 7). Here, we begin by describing the main principles of quantitative analysis, then offer guidance on when and how to use the different types of analysis listed above. While this chapter addresses some key statistical concepts, there are no formulae or other mathematics. This is a brief introduction to the most essential practical information you need to know about the statistical tests you're most likely to need. For more detail and advice on a greater number of statistical tests, see the suggestions for further reading at the end of this chapter.

12.2 RELY ON THE PRINCIPLES AND CONCEPTS OF QUANTITATIVE ANALYSIS

Quantitative data analysis involves both finding an appropriate explanation and testing that explanation. There are some essential concepts for you to understand in order to make good decisions as you develop your quantitative data analysis.

First, you should know the distinction between cases and variables. Social scientists seek to identify cases – people, organizations, events – about which they will assemble evidence. Any phenomenon for which change can be observed from case to case (e.g., employees' awareness of institutional aims or political attitudes) can be a *variable*. Variables may be observed, surveyed and otherwise measured. Once captured, they can be subjected to statistical analysis.

It is a well-established principle in statistical theory that the normal distribution of values in a population follows a set pattern. This distribution of data is most commonly recognized from its bell shape. This bell-shaped distribution has the following characteristics:

- The mean is located in the centre of the distribution.
- The greater the distance from the mean, the lower the frequency of occurrence.

As will be discussed later, this normal distribution can be evaluated statistically.

12.2.1 DEVELOP AND TEST YOUR MODELS AND HYPOTHESES

Quantitative analysis is based on the idea that you can establish generalizable knowledge about a social phenomenon through deductive reasoning. That is, you first develop a *hypothesis* (a statement about how a particular phenomenon works or how variables are related) based on theory or prior research, etc. (or initial inductive observations). Once you have a hypothesis, you can then identify what you would expect to find from particular *dependent variables*, if the hypothesis is correct. You then collect 'observed data' to test its 'fit' with expected outcomes defined by the hypothesis. Evidence supporting the model increases faith in its accuracy, while evidence that contradicts the model detracts from belief in its accuracy. For example, if we had a hypothesis that 'money is the root of all evil', we could test this hypothesis by analysing in a survey whether people who have more money are more likely to report committing evil acts. If we found a statistical relationship between self-reported earnings and propensity to commit evil acts (however we choose to define those), this would support a broader model linking money and evil.

Statistical analysis uses a rather peculiar and long-winded logical structure. Instead of directly evaluating the accuracy of a hypothesis (e.g., men will earn larger salaries than women due to advantages gained from gender inequality), statistical analyses operate by testing the opposite proposition (e.g., there will be no difference between men and women's salaries). If this opposite view (or *null hypothesis*) is shown to be unlikely to be true by the statistical analysis, then it is rejected and the *alternative hypothesis* is considered to be the likely truth.

In addition, there is a crucial distinction in quantitative analysis between the variables that you think might be causing or predicting an outcome (*independent variable*) and the variables that you think of as outcomes or effects (dependent variable). For example, if you were researching whether the saying 'money is the root of all evil' is accurate, money would be the independent variable and evil would be the dependent variable.

12.2.2 UNDERSTAND THE TYPES OF VARIABLES

Your data can take the form of different types of variables. You must be able to identify these different types of variable in order to use the statistical analysis that corresponds with the variable type(s) of your data. There are three different types of quantitative data, although they are sometimes referred to with different names.

- ***Categorical (nominal) variables***. A categorical or nominal variable is one that has two or more categories, but there is no intrinsic ordering to the categories. For example, you can't calculate an average sexual orientation, gender, ethnicity or hair colour for people: people with blond, red or black hair are simply in different categories. Other examples of categorical variables include gender (male/female), place of birth (native-born/immigrant) and marital status (married/single).
- ***Ordinal variables.*** The difference between an ordinal and a categorical variable is that there is a clear, natural ordering within ordinal variables. For example, the categories of social class and education level are clearly ordered, such as from higher to lower. However, even though the categories within each variable can be ordered, the spacing between each category may not be the same. For example, the size of the step from 'working class' to 'middle class' is not the same (in terms of income) as the step from 'middle class' to 'upper class'. Likewise, the step from bachelor's degree to master's is not the same as the step from master's to doctorate. This makes education level an ordinal variable.

- ***Interval (continuous or scale) variables***. An interval variable is similar to an ordinal variable, except that the intervals between the values of the interval variable are equally spaced. For example, annual income measured in British pounds or US dollars would be interval data because the step from $1 to $2 is the same as the step from $125 to $126, and so on: it is always the same interval (i.e., space) between each value within the variable. The same goes for age as measured in years.

Variable type matters greatly because particular statistical analyses are intended for use with only certain types of variables. You can't compute an average level of hair colour in a room because the statistical calculation of the arithmetic average requires an interval variable.

12.3 USE CONTENT ANALYSIS TO CONVERT QUALITATIVE INTO QUANTITATIVE DATA

Content analysis is a systematic method of converting qualitative content (such as words or images), for example from a survey or existing data, into quantitative data. The method provides you with the tools to ensure this quantification process is done in a reliable way. It uses a process of categorization called content analysis *coding* (see also Chapter 11 for discussion on coding). This involves categorizing raw data using a limited number of options (cf. Hayes & Krippendorff, 2007). Your coding must lead to numerical data in order for you to analyse it statistically (even if you start with categories, you must assign numbers to those categories).

A key fork in the road for your content analysis is whether you need to analyse your content for manifest or latent meaning.

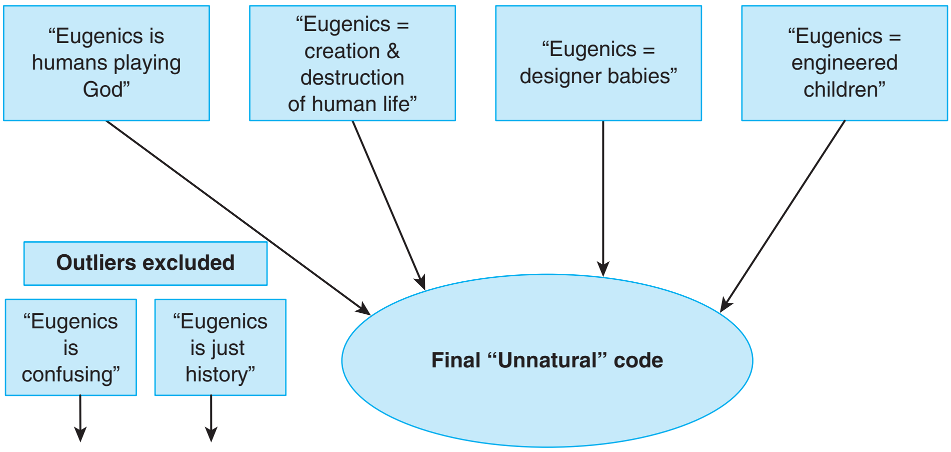
- **Manifest content**. This is the objective, surface-level or concrete content. Examples of this kind of research include counting the number of times the word 'homosexual' appears on a news website, the number of pictures of men versus women in a science textbook or the number of times a set of wildlife conservation-related words are used in open-ended survey responses.
- **Latent content**. This is defined by a focus on underlying or implicit meanings. For example, you might consider how approvingly or disapprovingly homosexual behaviour is mentioned in a newspaper, or whether women are performing traditionally masculine or feminine tasks when they are pictured in a textbook, or whether there is evidence that learning has occurred when comparing a survey response from an individual at the beginning of the year and the end.

While identifying manifest content is relatively straightforward, latent meaning is much more difficult to reliably code. It is possible to use automated text analysis tools to identify manifest content. However, using human coders can be particularly valuable when analysing latent content as (with training) they will be able to identify implicit meaning that may evade an automated analysis tool.

12.3.1 DEVELOP CATEGORIES THROUGH OPEN CODING

While you can get pre-made categories for your analysis from theory or prior research, one of the ways to develop your categories for the content analysis is an initial inductive analysis (from data) known as open coding (see Glaser & Strauss, 1967). Often this will involve identifying more specific categories that can be collapsed into an overarching one. For example, in Figure 12.1 four similar codes on eugenics were collapsed into a final all-encompassing code called 'Unnatural', while two outlier (infrequently mentioned) codes were not included, as they would have belonged to different overarching themes and did not appear in a cluster with other similar codes.

Figure 12.1 Example of final code development from open coding for "Unnatural"



There are different ways of showing the content analysis categories you've developed. However, you should usually prepare a table that includes category names, definitions and examples (see Table 12.1). This kind of table can help maintain clarity for all involved in the study, including the reader, by showing the key steps involved in going from specific quotations to the coding category.

12.3.2 EVALUATE INTER-CODER RELIABILITY

Inter-coder reliability (also called inter-rater agreement) refers to the extent to which independent analysts (or 'coders') evaluating the same content characteristics will reach the same conclusion. A high level of agreement is taken as evidence that the content analysis has identified characteristics that are objectively evident in the texts being analysed. As Neuendorf (2002, p. 141) argues, 'given that a goal of content analysis is to identify and record relatively objective (or at least intersubjective) characteristics of messages, reliability is paramount. Without the establishment of reliability, content analysis measures are useless.' The practice of testing for inter-coder reliability serves two major purposes. Firstly, it is a

Table 12.1 Example of codes for open-ended survey responses to pre-festival question: 'What comes to mind when you think of the Silicon Valley Technology Festival?'

Code	Code Description	Code Examples
Lack of awareness	Comments by people who have never heard of or never before visited the Silicon Valley Technology Festival expressing uncertainty about the event	• 'I don't know, first time attending' [Master's, Male, age 33, White] • 'Wasn't aware of it before this week' [PhD, Male, age 35, White] • 'My first time hearing about the festival so opinions will be formed from here henceforth' [No formal qualifications (still in education), Male, age 23, White]
Silicon Valley	References to the city which plays host to the Silicon Valley Science Festival	• 'A big event in an industrial city' [Foundation Diploma or equivalent, Female, age 20, White]
Educational	Reference is made to having the chance to learn new things and being educational both for children and adults. Providing youth with opportunities for the future	• 'That we will as adults be learning too! :)' [Degree, Female, age 36, Asian/Asian American] • 'Learning for kids, broadening horizons, the "wow" factor that science has. I want my daughter to be marvelled by science' [Master's, Male, age 44, White] • 'Public education' [Degree, Male, age 62, White]
Interactive Engagement	Focus is on the interactive nature of the Technology Festival and 'learning through doing'. Participants comment on taking part in activities and experiments	• 'Experiment interaction (no reading)' [Degree, Female, age 37, White] • 'Hands on approach, fun and interactive, inspires our children to ask questions and find out more' [Master's, Male, age 45, White] • 'Innovative science, interesting presentations, a University of Stanford academic pouring liquid nitrogen onto my kids shoes (it was safe and I knew the academic, the kids loved it)' [High school, Male, age 38, White]
Technological Advancement	Participants comment on the future of technology, science and new inventions	• 'Technology for the future' [Foundation Diploma or equivalent, Female, age 34, White] • 'Scientific progress and spirit of future innovation' [Master's, Female, age 41, White]

quality assurance mechanism that can signal when there are problems with your content analysis design: 'High levels of disagreement among judges suggest weaknesses in research methods, including the possibility of poor operational definitions, categories, and judge training' (Kolbe & Burnett, 1991, p. 248). Secondly, inter-coder reliability establishes the value of your analysis for the reader: high inter-coder agreement means that you've produced results that are worthy of being believed by readers.

The first step in testing for inter-coder reliability is to have the members of your coding team independently code the same (randomly selected) subset of your sample texts. The rule of thumb is that you want to randomly select at least 10 per cent of your articles for overlapping coding. Once you have a set of items that have been coded by at least two analysts, you can statistically test for inter-coder reliability. There are many different statistics options available to you for evaluating the level of consistency between coders' scores, each with different strengths and limitations. We recommend using Krippendorf's alpha (or Kalpha), because it can handle almost any relevant research situation, and because it is the most valid measure of inter-coder reliability (Krippendorff, 2004). Another reason for choosing Kalpha is that there is a simple SPSS macro you can use for calculating this statistic, with easy-to-understand instructions for downloading, installing and running the macro. To access this information, visit http://www.afhayes.com. See also Hayes and Krippendorff (2007).

REAL WORLD EXAMPLE

Content-analysing news coverage of eugenics

Dimitri conducted a study of representations of eugenics in different kinds of news magazines and newspapers. News coverage of eugenics across three distinct genres of US-based news publications was subjected to content analysis. This content analysis addressed the research question: what concepts and rhetorical themes dominate mainstream, pro-science, and Christian fundamentalist news coverage of the human cloning debate? He conducted a preliminary qualitative analysis of the data, which pointed to the importance of 'unnatural' rhetoric, resulting in a reorientation towards coding for unnatural-related content during the quantitative content analysis. His study identified four thematic coding categories (see Figure 12.1), which all connected to the issue of the perceived unnaturalness of eugenics. Here is how the inter-coder reliability procedure and results for this study could be explained:

> Inter-coder reliability was calculated based on a random sample of 900 cases in three distinct genres of US-based news publications. The second coded random sample showed an excellent level of inter-coder reliability with Kalpha = .87.

This example shows how a content analysis study can establish its reliability.

It is generally advisable to test for reliability near the beginning of the coding process, during the coding process and again at the end of the coding process (MacQueen, McLellan, Kay, & Milstein, 1998). Begin by having the full coding team analyse about 15–30 randomly selected articles (no more than 5 per cent of your total sample), and then run your reliability statistics. Then have your team meet together (if practicable) and examine where there were disagreements. If there was a particular code with poor reliability, it may indicate that you need to rephrase, or rewrite, the coding procedures. Once you feel comfortable that your coding team members are all on the same page, have everyone continue analysing the rest of the articles. As the overlapping coding procedure continues, and as you begin to get results, make sure that you are regularly rerunning your inter-coder reliability statistics to ensure everyone is still maintain consistency.

The question of what constitutes a 'good' level of inter-coder reliability often depends upon the nature of your research. Generally, it is more acceptable for a complicated or new coding procedure to have somewhat lower inter-coder reliability than a simple or well-established procedure. Regardless of the circumstances, a Kalpha coefficient above 0.9 should be seen as very good. Unless there is previous research establishing that the coding mechanism employed should be higher, anything above a 0.8 is considered good. For more exploratory research or complex coding procedures a score of 0.7 may be acceptable. Anything below a 0.6 should be considered unreliable. A score of 0 indicates that the level of inter-coder agreement was not better than a completely random distribution. A negative score indicates that there was systematic disagreement on how to apply particular codes (Krippendorff, 2008).

REAL WORLD EXAMPLE

Content-analysing children's zoo visit drawings

Eric conducted a study at London Zoo to evaluate the educational effects of visiting the zoo on schoolchildren (Jensen, 2014). The main measure of educational impact was an open-ended survey question that asked children to draw their favourite wildlife habitat and all the plants and animals that live there (putting labels on everything). Children did this the day before the zoo visit and the day after the zoo visit. The pre- and post-visit drawings for the same children were matched up and then analysed using the following simple content analysis system:

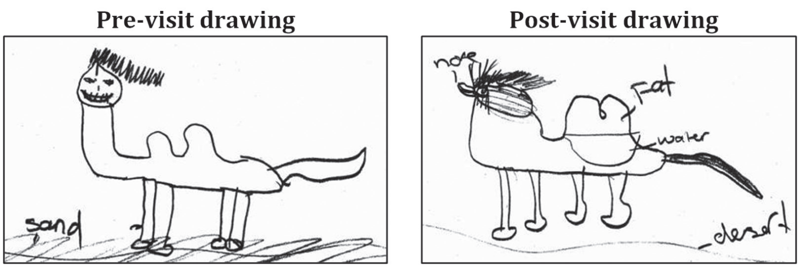
1 = negative change in accuracy of representation (animals/habitat)
2 = no change in accuracy
3 = positive change

Learning was defined for the purposes of this measurement as including:

- Level of elaboration in terms of accurately showing the physical characteristics of the animal
- Level of conceptual sophistication in terms of use of scientific concepts relating to animals or habitats
- Accuracy of depiction and labelling of animal's habitat

For example, the two drawings in Figure 12.2 are from the same child, and were coded as 'positive change'. This was because of the specific criteria set out for the content analysis: the nose had become more accurately represented, there was evidence of increased conceptual sophistication as the child shifted from the term 'sand' to the more abstract scientific concept of 'desert' (a habitat category). Also, the child showed an understanding of the purposes served by the camel's humps.

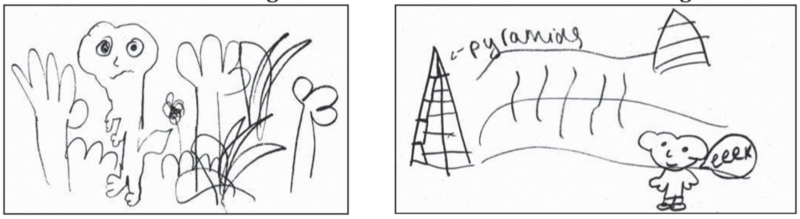
Figure 12.2 Pre- and post-visit drawings by the same child around a visit to London Zoo



It is worth highlighting that the criteria for this particular study are not concerned with other forms of inaccuracy in the drawings, such as representations of animals in human-like terms (known as 'anthropomorphism'). So the drawings in Figure 12.3 were coded as showing 'positive change',

Figure 12.3 Pre- and post-visit drawings around a visit to London Zoo showing positive change

Pre-visit drawing **Post-visit drawing**



despite the fact that the meerkat in the example is given shoes in the post-visit drawing and the desert is represented using pyramids.

This is how content analysis works. You set criteria that are made clear to both analysts and to readers, and then these are applied strictly and consistently across the entire sample. At the end of this process, Eric had systematically converted the open-ended visual data into statistical data that he was able to analyse using some of the techniques discussed in this chapter.

12.3.3 ACCOUNT FOR THE STRENGTHS AND WEAKNESSES OF CONTENT ANALYSIS

The strengths of content analysis include that you can create quantitative data that can be used in statistical analyses, easily repeat portions of the analysis if necessary and apply the same analysis to data collected over different points in time. It is a way of reliably measuring the meaning communicated in content in a way that is in principle inter-subjective (i.e., its meaning is shared across individuals). Moreover, the level of inter-subjectivity achieved is known (through inter-coder reliability). Finally, it offers a clear, well-established method, enabling transparency and procedural clarity for both the researcher and readers/users of the research.

Meanwhile, the main weakness of content analysis is that you can develop inaccurate understandings of the meaning communicated in content because you are working with static information (i.e., you can't normally ask follow-up questions to clarify). This can lead to attributing meanings to a text that are not actually there. This issue is particularly challenging when it comes to analysing latent content, which inevitably involves subjective interpretation (although this can be mitigated through clear documentation of categories and inter-coder reliability measures). Overall, however, some form of content analysis is likely to be required to convert qualitative data into quantitative data that can be analysed statistically.

12.4 USE SPSS TO CONDUCT YOUR QUANTITATIVE DATA ANALYSIS

SPSS is a software package that performs a wide variety of statistical tests. Both data management and analysis can be conducted with it. You can use it to produce graphs and perform statistical analyses ranging from calculating simple percentages to very sophisticated analyses well beyond the scope of this chapter. There are other statistics software providers, including the free open-source software R (http://www.r-project.org), which is widely accepted as a good (if less user-friendly) alternative to SPSS. However, we focus our guidance in this chapter on SPSS because it is the most widely used across the social sciences. This section provides you with a basic introduction to SPSS to help you get comfortable with its main features.

File types associated with SPSS include:

- **Data:** filename.sav
- **Output:** filename.spo
- **Commands:** filename.sps

SPSS can open these files as well as Excel, CSV and other file formats.

The first SPSS program window is the Data Editor, which has two views: Data View and Variable View. The Data View (Figure 12.4) is where you enter your data, or where your data will appear if you are importing them into SPSS from another programme such as Excel. In this view, the columns are where the variables are indicated (not editable in this view) and the rows are where the individual cases are recorded (editable in this view).

The SPSS Variable View (Figure 12.5) is where you enter your list of variables and define their characteristics. You enter details into the Variable View by clicking on the cells in each row. You normally only need to adjust the following categories:

1. **Name**. Enter a variable name. This name has to be all one word, such as GenderCategory or HairColour.
2. **Decimals**. If your data only appear as whole numbers (e.g., age in years), then change this to '0'. This will apply to most of your variables. If it is important to have the data show decimal places (e.g., 1.22), then enter the number of decimal places you would like to appear here.

Figure 12.4 Data View in SPSS

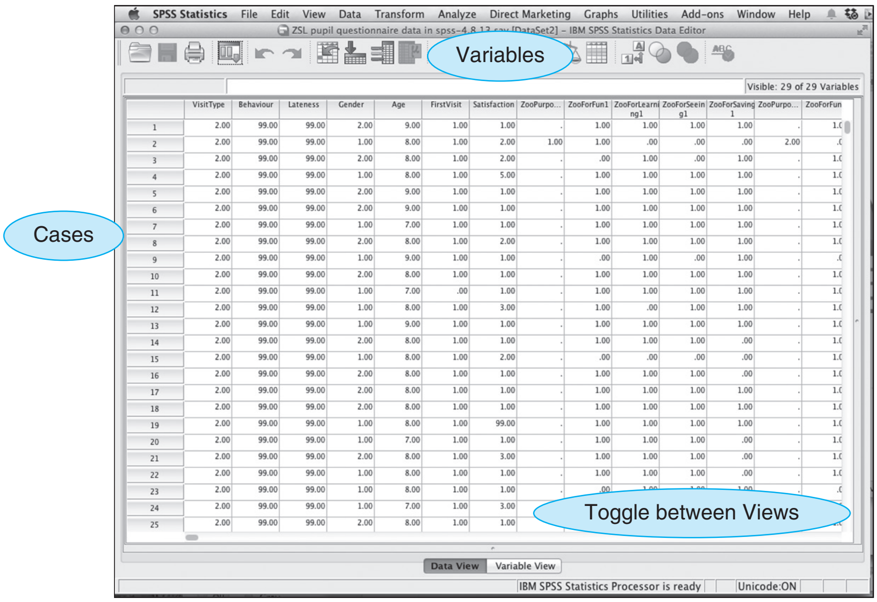
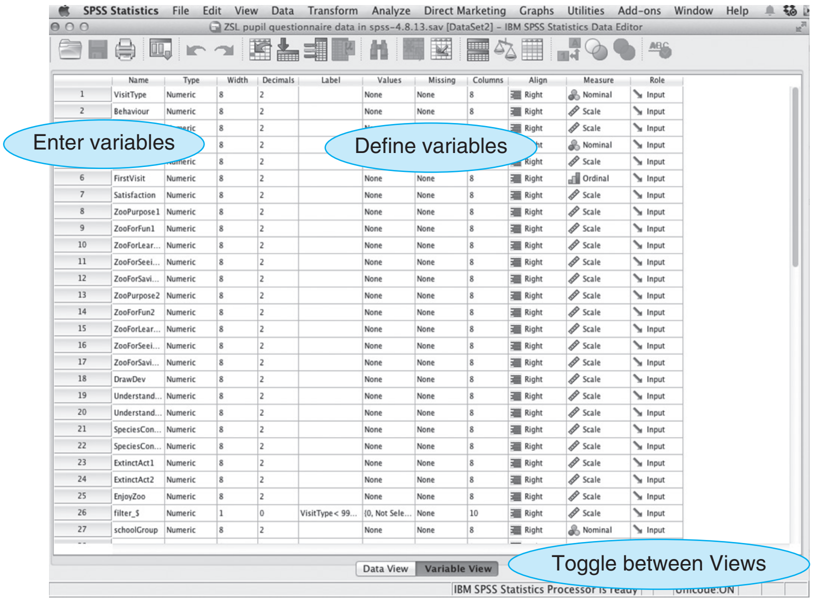


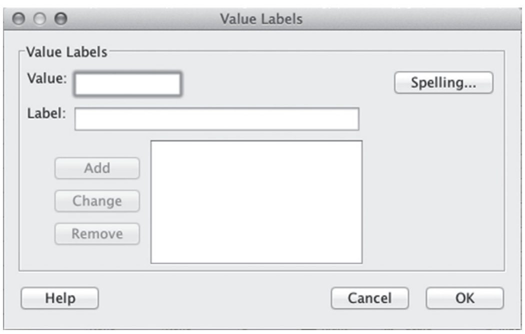
Figure 12.5 Variable View in SPSS



3 **Label**. This is where you can enter a variable name with spaces in it. Enter the label for the variable that you would like to have appear in any reports, as this label is carried through onto any results tables or charts produced through SPSS.

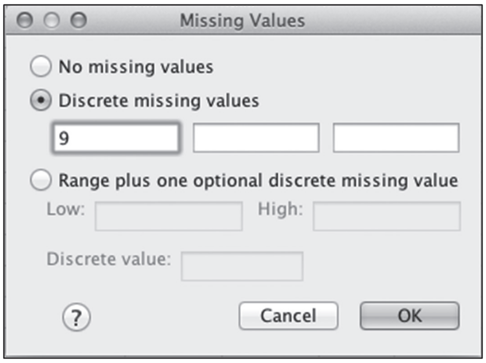
4 **Values**. For many variables, you need to assign an arbitrary number to its different categories to be able to use them in statistical analysis (see Figure 12.6). For example, you might enter the variable gender as follows: 1 = Female, 2 = Male, 3 = Other. You make these entries here in the Values cell. When you click in the Values cell, you will see an ellipsis (…). Click on this and a small window will come up that says 'Value Labels'. Here you put the value ('1' for Female in the example above), then the category name associated with that value goes below that in the Label field ('Female' in the example above). Then click on the Add button to save this label. Repeat this process until all of the categories for the variable have been entered. Finally, you can leave the Values cell in its default position of 'None' for variables that are created in a numerical form such as age in years or weekly income in American dollars or British pounds.

Figure 12.6 Value Labels window in SPSS



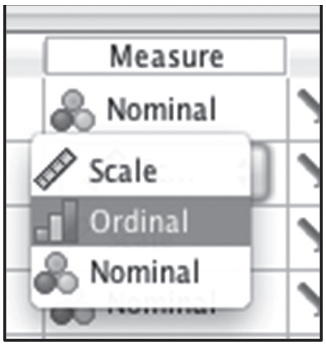
5 **Missing**. Here you define the values that should not be included in calculations. For example, if you had a 'don't know' response category on a survey question that was entered as '9' in SPSS, you could exclude this category from your analysis by clicking in the cell of the Missing column for the variable, choosing 'Discrete missing values' (see Figure 12.7) and then entering the value you wish to exclude (9 in this example). If you have more than three different codes that signify missing data (as in the case of having more than three reasons why the data is missing) you can use the 'Range plus one optional discrete missing value' option. Here, you enter a range of numbers where all the values should be considered missing (e.g., –99 to –10).

Figure 12.7 Missing Values window in SPSS



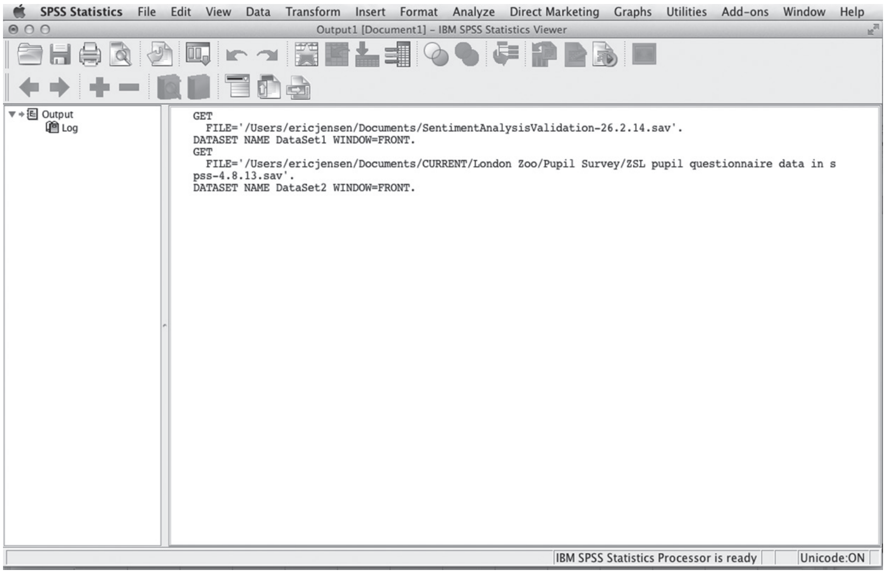
6 **Measure**. This is not essential to complete, but you can use it to note down the variable type as a reminder for yourself. SPSS uses the following terms for variable type: Nominal (categorical), Ordinal and Scale (interval) (see Figure 12.8).

Figure 12.8 Entering the Measure column



The second SPSS Program Window is the Output Viewer (Figure 12.9). This is where the results of your statistical tests appear once you run them in the Data Editor. This division between conducting analyses and viewing their results means that you have to know to go looking in the Output Viewer once you hit OK to run a statistical analysis.

Figure 12.9 Output Viewer in SPSS



12.5 CONDUCT DESCRIPTIVE STATISTICAL ANALYSIS

Descriptive statistics tend to be based on analysis of one variable at a time. Many of the statistics we hear about in the news and from government are descriptive statistics, such as the percentage of people holding a particular attitude, the fractions of the population in different ethnic categories and the median income in a population. Most descriptive statistical analyses you'll need to do will fit into three categories:

- Measuring the **prevalence of a characteristic** in the population based on data from a sample.
- Identifying the **general pattern** in a sample and population.
- Evaluating the **extent to which a characteristic varies** across a sample and population.

These three categories of descriptive statistical analysis will each be addressed in this section. Throughout this section, the focus will be on generalizing from sample to population. Therefore, it is important to keep in mind the distinction between the results for the sample (also known as the 'observed values') and the results for the population, which we estimate based on the sample data.

One important concept that relates to this distinction between sample and population results is the *confidence interval*. Imagine you want to know how many people hold a particular political viewpoint in a large city. As we discussed in an earlier chapter, if you systematically gather a sample from that population, you can generalize within a specified range and level of certainty. The results of such an analysis will include a confidence interval, that is, a range of values (based on your *sample*) within which the true value for the *population* is very likely to be (Cohen, Cohen, West, & Aiken, 2003, p. 15). For example, imagine you randomly selected a sample of people in the city of London to ask about whom they supported in the last mayoral election. If you found that 51% supported the Conservative Party candidate, that would not mean that the real percentage of Conservative support in the larger population of millions of Londoners was 51%. Your statistical analysis would give you a range, for example, plus or minus 5% within which the real value is likely to be (95% or better level of probability for statistical significance). This would mean that the confidence interval was from 46% to 56%. The real value could be anywhere in this range (with 95% probability or better); it could also be outside this range, but that would be much less likely (5% probability or less).

A related concept is *statistical power*, that is, the ability of a statistical test to detect a real pattern that exists in the broader population based on your sample data. This is traditionally defined as $1 - \beta$, that is, one minus 'the probability of failing to reject the null hypothesis when it is false' (Cohen et al., 2003, p. 51). This statistical power is affected by sample size (a larger sample size increases statistical power) the size of the real pattern in the population (a larger population effect is easier to detect through a sample), and where you set your α probability, meaning the probability of failing to accept the null hypothesis when it is true (Cohen et al., 2003, p. 52). For example, if you change your confidence interval from 95% to 99%, you decrease the statistical power.

12.5.1 DESCRIBE THE POPULATION

The simplest form of statistical analysis in social research is describing the distribution of characteristics or behaviours in a population using data you've gathered (i.e., your sample). Examples of this kind of analysis include the following:

- **Gender** – The number of women, number of men, proportion of women compared to men.
- **Religiosity** – How many people go to church once per year, per month, per week or never?
- **Age** – What's the oldest/youngest person you gathered data from? What's the average age of people in your sample?
- **Ethnicity** – How many people in your sample belong to each ethnic group?

These are important pieces of information, and they can be used to learn about the characteristics of the population.

CLARIFYING YOUR RESULTS

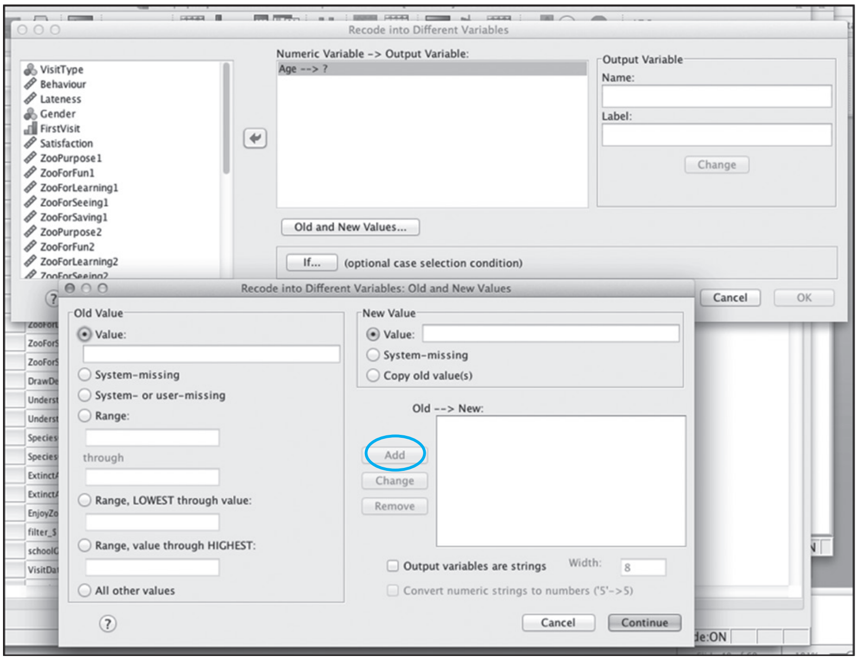
You may want to simplify and thus clarify the story your data are telling by merging response categories together at the analysis stage. For example, imagine you and 100 other people are asked the following question: 'On a scale of 1 (strongly disagree) to 5 (strongly agree), please indicate your response to the following statement: "I am enjoying learning about statistical analysis".' The responses would likely range from strongly disagreeing to strongly agreeing, but we might only need to have the overall picture of whether the learning experience is positive or negative. With that in mind, the categories of 'agree' and 'strongly agree' could be merged into one positive category, 'neutral' could be entered as a 'missing value', and 'strongly disagree' and 'disagree' could be merged into one negative category. The results remaining would then present a simplified view of the rate of positive versus negative responses to learning about statistical analysis.

The way you do this in SPSS is to go to the Transform menu and select Recode into Different Variables. This option creates a new transformed variable rather than making the changes in the original variable (see Figure 12.10). Then highlight the variable you want to adjust and move it over to the right to the Numeric Variable -> Output Variable box.

1. Fill in a Name and Label for the new variable.
2. Click on Old and New Values.
3. Specify the Old Value (e.g., 1 to 10, 11 to 20, etc.).
4. Specify a New Value (e.g., 1 [for 1–10], 2 [for 11–20], etc.).
5. Click on the Add button.
6. Repeat until all old and new values are specified (you will get errors if you leave any of the old values unaccounted for, so specify the translation even if it is going, for example, from 6 (old value) to 6 (new value).
7. Old values can be defined as single values, ranges or missing values.

Finally, if you do this recoding procedure, be sure to make a note describing how data categories were merged so that you can interpret your data without having to refer back to the Variable View in SPSS.

Figure 12.10 Merging categories in SPSS using Recode into Different Variables

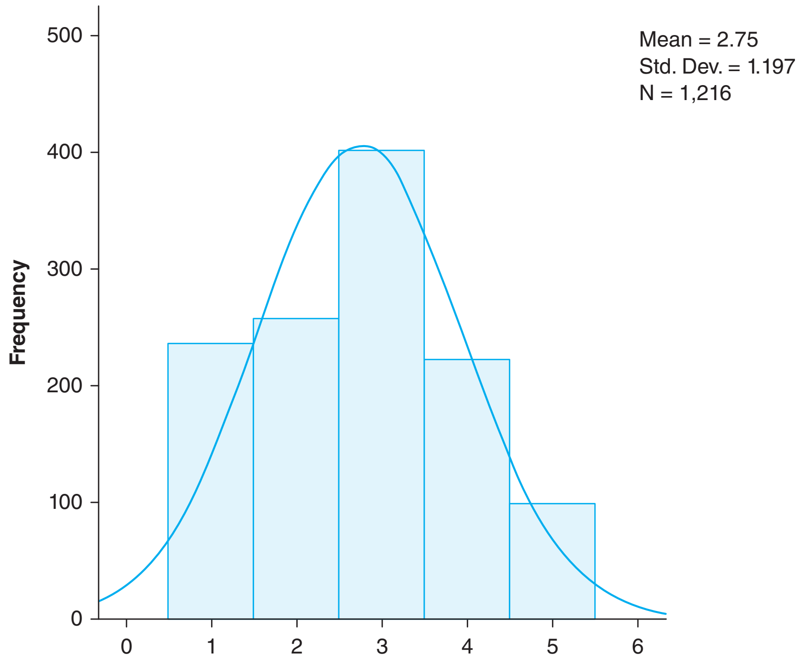


REPRESENTING UNIVARIATE DATA: USING SPSS TO VISUALIZE YOUR DATA

When presenting your data in charts, your goal should be to provide the reader with an accurate understanding of the data, while presenting it in a manageable form. There are three specific types of charts that are especially useful: the histogram, population pyramid and scatterplot.

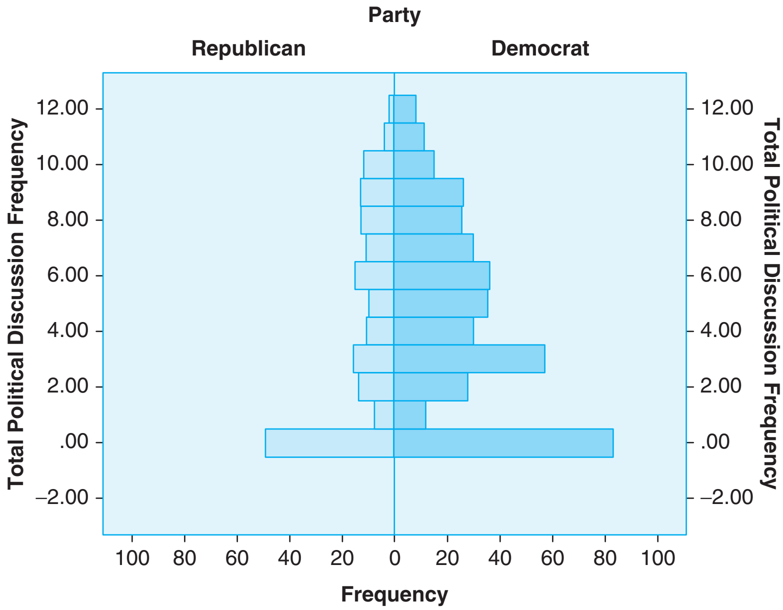
- **Histogram**. A histogram is a simple chart for showing the distribution of your data (see Figure 12.11). To build it, select Graphs from the SPSS menu, then Legacy Dialogs > Histogram. Select the variable on the left you wish to graph, and drag it over to Variable. Check the box labelled 'Display normal curve', and then click OK.

Figure 12.11 Example histogram



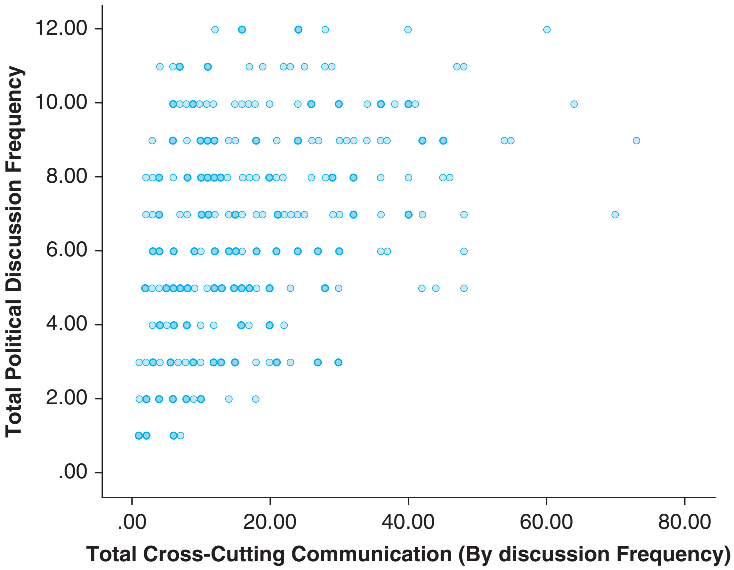
- **Population pyramid**. A population pyramid is similar to a histogram, in that it shows you the distribution of your data (see Figure 12.12). However, a population pyramid allows you to take this one step further, by comparing the distribution across two categories (e.g., Male/Female). To build it, select Graphs from the SPSS menu, then Legacy Dialogues > Population Pyramid. Move whichever variable you're interested in graphing to the box labelled 'Show Distribution over', and move the categorical variable you want to split the data by into the box labelled 'Split by'. If you would like to display a normal curve over your distributions, select 'Scale options', and select the box labelled 'Display normal curve'. Click OK, and then OK again.

Figure 12.12 Example population pyramid graph



- **Scatterplot**. Scatterplots help you to visualize the relationship between two variables, and are useful for testing important statistical assumptions (as will be discussed later). To build it, select Graphs from the SPSS menu, then Legacy Dialogs > Scatter Dot. Select Simple Scatter and then click Define. Move your outcome variable to the Y-Axis box, and your predictor variable to the X-Axis box. Click OK. You will then get a scatterplot (see Figure 12.13).

Figure 12.13 Example scatterplot from SPSS



12.5.2 LOOK FOR THE CENTRAL TENDENCY IN YOUR DATA

When you have interval data, identifying the typical pattern in the data can be very useful. This can be done by looking for the central tendency in the data. Selecting an appropriate measure of central tendency is straightforward once you understand some basic principles. The mean is probably the most commonly used measure of central tendency. To calculate it, you add up all of the values in your sample and divide by the number of cases. This can provide a general sense of the average pattern across the entire sample.

The median, on the other hand, is a measure of central tendency in which all of the values in the sample are spread out into a list from smallest to largest. The value that is in the exact middle of this list is the median. For variables that tend to have extreme values such as income (where the richest 1 per cent can stretch the income scale to extremes), the median is the more appropriate measure of central tendency because it's not distorted by such extremes. Indeed, the outcome measure can drastically affect the statistical picture that emerges, so it is important to use the one that is most appropriate. It is essential to compare like with like when making statistical comparisons.

REAL WORLD EXAMPLE

Gender inequality and statistical analysis

Gender inequality is a long-standing problem in Western societies, with women systematically paid less than men for doing the same or similar work. However, this gender pay gap has narrowed in recent years. In an effort to maintain pressure on the issue, however, campaigners have (mis)used the statistical evidence to make the problem seem worse than it is.

For example, the UK equal pay campaign group the Fawcett Society published a report describing the pay gap as follows:

> More than 40 years after the Equal Pay Act was enacted in 1970, women's hourly earnings continue to be significantly lower than men's with the gender pay gap (mean) for full-time employees in 2013 sitting at 15.7 per cent. This means that women effectively stop earning relative to men on the 4th November 2014 – this day is referred to as Equal Pay Day. … A woman working full-time now earns, on average, £5,000 less a year than a man. For all workers – both part-time and full-time – the gender pay gap stands at 19.1 per cent … . This means for every £1 earned by a man in the UK, a woman earns only 81 pence. (Fawcett Society, 2014, p. 2)

The use of the mean creates an inaccurate picture of central tendency for income because it is heavily affected by the richest people's income. In a footnote in the report, this fact is acknowledged and highlighted as the reason for using the mean rather than the median:

> A note on why we use *the mean measure*: Using median estimates mostly leads to lower estimates. This is because it neutralises the effect of having a small group of very highly paid male employees. … Using the mean measure is helpful precisely because it highlights that the economic elite in the UK is still predominantly male. (Fawcett Society, 2014, p. 4)

The problem with this explanation is that the claims presented in their report are about the central tendency, with references to an 'average' pay gap between men and women of £5,000 per year.

This presents a misleading picture of the problem, as the real source of the mean level of inequality is a relatively small number of men at the highest levels of the income spectrum. Men and women lower down the pay scale have much more similar earnings (albeit that there is still a small gender pay gap even at the lower income levels). Moreover, once gender differences in job type, job role, length of employment and number of hours worked per week are taken into account, the gender pay gap is very small in Western developed nations. This example highlights the importance of selecting appropriate statistical measures to gain an accurate understanding of the real sources of social problems.

12.5.3 LOOK AT HOW YOUR DATA SPREAD TO IDENTIFY INTERESTING PATTERNS

While the central tendency is an important part of the picture, it is also important to understand how your data are spread out. That is, you need to know how widely scattered your data are. Do they cluster tightly around your mean? If they don't, then that indicates your mean is a poor summary of your overall results. 'Standard deviation' is a measure used to summarize how spread out your data is: it is a measure of the average level of variability away from the mean in your data. If the standard deviation of your data is larger than the mean, that suggests your data are highly spread out. If it is much smaller than your mean, then the data are likely to be clustering together around the mean.

If you have data that are spread out, then you need to be able to identify the implications of that spread. One way of doing this is to split your data into what are called 'deciles', that is, break your data up into 10 parts (each part being one decile). Once you do this, you will sometimes see some interesting patterns, particularly at the extremes. (For more on this subject, see Williams & Monge, 2001, pp. 31–46.)

REAL WORLD EXAMPLE

The variability in drinking behaviour

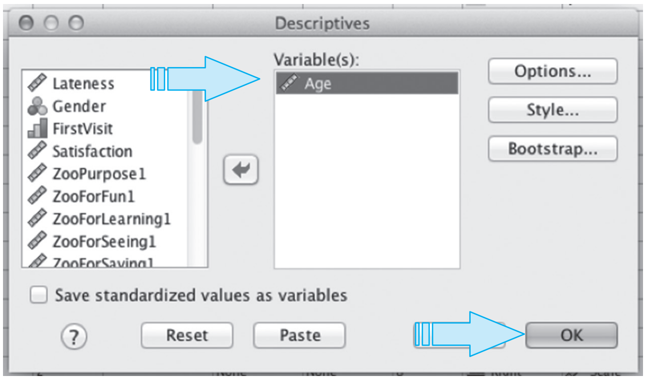
In Phillip Cook's book *Paying the Tab* (2007) about alcohol policy and consumption, he points out that while the central tendency for those who drink alcohol in the United States is a median of 3 drinks per week, this statistic masks a surprising level of variability. If you split all the adults in the United States into deciles, you find that the first three deciles (30%) don't drink any alcohol at all. The next three deciles have average consumption levels of 0.02 drinks per week, 0.14 drinks per week and 0.63 drinks per week, respectively. So 60 per cent of the US population is consuming less than 1 drink per week. The seventh and eighth deciles are closer to the median, with 2.17 and 6.25 drinks per week, respectively. But it is the top two deciles that account for the vast majority of all alcohol consumption in the USA, with averages of 15.28 and 73.85 drinks per week, respectively! 'That works out to a little more than four-and-a-half 750 ml bottles of Jack Daniels, 18 bottles of wine, or three 24-can cases of beer. In one week' (Ingraham, 2014). This means that 10 per cent of the population is consuming the majority of all the alcohol that people are drinking in the USA. This paints a very different picture of the situation than the measure of central tendency.

12.5.4 HOW TO USE THE DESCRIPTIVES FUNCTION TO CALCULATE A MEAN, STANDARD DEVIATION AND RANGE IN SPSS

To calculate the mean in SPSS, you can use the Descriptives function to create tables with summaries of values for variables (see Figure 12.14). To do this, take the following steps:

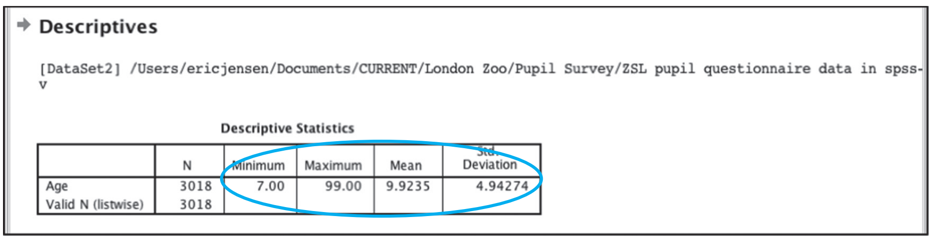
1. Go to the Analyze menu, select Descriptive Statistics, then Descriptives.
2. Highlight variables to create tables, click on the arrow (or drag and drop) to add to the Variable(s) list, then click OK.

Figure 12.14 Using the Descriptives function in SPSS to calculate the mean



This will give you an SPSS output like the one pictured in Figure 12.15, which shows you the mean and number of cases included in your analysis (N).

Figure 12.15 Using the Descriptives function to calculate the mean – Interpreting the SPSS output


Descriptives

[DataSet2] /Users/ericjensen/Documents/CURRENT/London Zoo/Pupil Survey/ZSL pupil questionnaire data in spss-v

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Age	3018	7.00	99.00	9.9235	4.94274
Valid N (listwise)	3018				

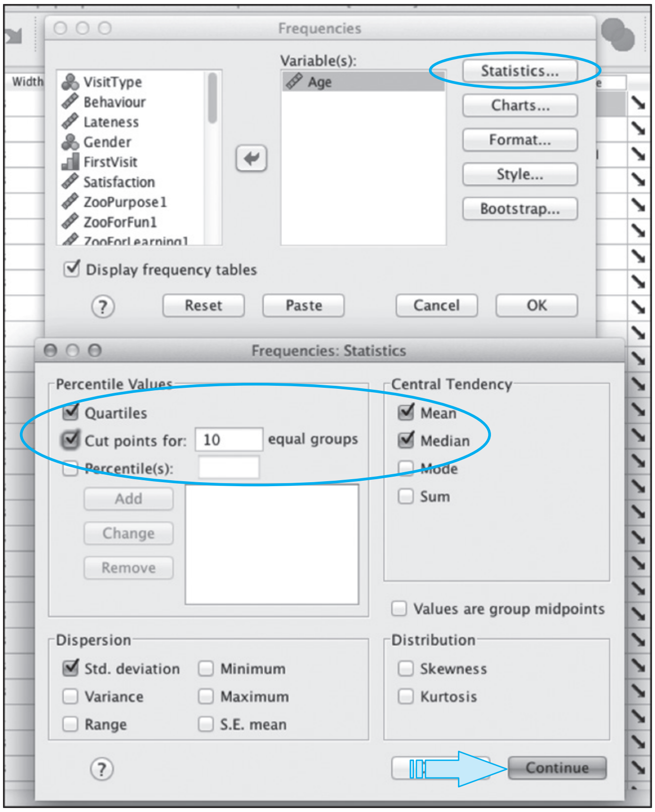
This output also provides the minimum and maximum (lowest and highest) values in the sample for the variable, as well as the standard deviation. Of course, these are only the default outputs. If you select Options prior to clicking OK, you will be given the ability to specify additional (although often unnecessary) statistics, including: variance, range, standard error of the mean, kurtosis, skewness and sum (i.e., the sum of all response values within the sample). You can also alter the order in which the variables are displayed in the output, by changing the option under Display Order'

12.5.5 HOW TO USE THE FREQUENCIES FUNCTION TO CALCULATE THE MEAN, MEDIAN, STANDARD DEVIATION AND QUARTILES IN SPSS

To calculate the median in SPSS, you can use the Frequencies function to create tables with counts of cases for each value of the variable (see Figure 12.16). To do this, take the following steps:

1. Go to the Analyze menu, select Descriptive Statistics, then Frequencies.
2. Highlight variables to create tables, click on the arrow to add to the Variable(s) list, then click OK.
3. Click on the Statistics button.
4. Tick the Mean and Median boxes. You can select 'Std. deviation' and Quartiles if you would also like that information.
5. Click Continue twice.

Figure 12.16 Using the Frequencies function in SPSS to calculate the mean, median, standard deviation and quartiles



The output for the Frequencies function provides two tables. The first (labelled 'Statistics') provides the mean, median, standard deviation, and percentile values (see Figure 12.17). Percentiles (including the cut points for 10 equal points) are evaluated by

Figure 12.17 Frequencies output in SPSS: descriptive statistics for the Age data set

Statistics

Age

N	Valid	3010
	Missing	8
Mean		9.6867
Median		9.0000
Std. Deviation		1.82861
Percentiles	10	8.0000
	20	8.0000
	25	8.0000
	30	8.0000
	40	9.0000
	50	9.0000
	60	10.0000
	70	11.0000
	75	11.0000
	80	12.0000
	90	12.0000

looking at the number on the right column. In the example in Figure 12.17, a value of 10 is at the 60th percentile. This means that 60% of the sample have a score below that value (and 40% have a score above).

The other table contains data on the frequencies of each value in the data set (Figure 12.18). It provides information on how often the value occurred, and the percent of cases with that value. It also provides the valid percent (i.e., the percentage of cases with that value, after removing 'missing' values), and the cumulative percentage.

Figure 12.18 Frequencies output in SPSS: frequency of each value in the Age data set

Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	7.00	123	4.1	4.1	4.1
	8.00	855	28.3	28.4	32.5
	9.00	818	27.1	27.2	59.7
	10.00	260	8.6	8.6	68.3
	11.00	342	11.3	11.4	79.7
	12.00	338	11.2	11.2	90.9
	13.00	185	6.1	6.1	97.0
	14.00	66	2.2	2.2	99.2
	15.00	21	.7	.7	99.9
	16.00	2	.1	.1	100.0
	Total	3010	99.7	100.0	
Missing	99.00	8	.3		
Total		3018	100.0		

12.6 CONDUCT INFERENTIAL STATISTICAL ANALYSIS

The role of *inferential statistics* is to generalize to a larger population (e.g., a community or nation) from your sample (i.e., the smaller group selected from that population) with a quantifiable risk of error (Cohen et al., 2003, p. 41). That is, through inferential statistics you can find out whether a research result is likely to indicate a real effect in the population (i.e., it is statistically significant) or whether it is due only to chance variation in the selection of a sample from that larger population.

One of the key turning points in your statistical analysis will be a decision about whether to use parametric or non-parametric methods. Parametric statistics are generally used with probability (or near-probability) samples, while non-parametric statistics are used with non-probability samples (see Chapter 5). If you're using parametric statistics, then you'll need to know the key assumptions underpinning this type of statistical test. One of the most common assumptions for parametric statistics is 'normality'. While the precise nature of this assumption differs depending on the statistical test you are using, the important point is that you can run tests to confirm it. These tests will be discussed where they are needed under specific statistical analysis options later in this chapter.

Another common assumption required when using certain parametric statistics is that the levels of variability across all parts of the sample data are equivalent. This is called 'equality of variance' ('homogeneity of variance' or 'homoscedasticity'). When comparing groups, this assumption can be confirmed using an inferential test called 'Levene's test'. This test is built into the SPSS functions for the independent samples *t*-tests, and will be discussed in more detail in those sections later in the chapter.

As with the descriptive statistics, the correct statistical analysis will depend on the type of data that you have and the kind of analysis you want to conduct. To find the relevant statistical analyses covered in this chapter, consult Table 12.2.

Table 12.2 Matching data type, analysis purpose and statistical test

Research Question	Number and Type of Dependent Variable(s)	Number and Type of Independent Variable(s)	Test
Group differences	1 categorical or ordinal	1 categorical or ordinal	Chi-square
	1 interval (continuous or scale)	1 categorical (with only two categories, e.g., male/female)	*t*-test
Degree of relationship	1 interval (continuous or scale)	1 interval (continuous or scale)	Pearson correlation

Once you've identified the correct statistical test to use, you can begin to understand how it works. Let's start with the chi-square test.

12.7 USE CHI-SQUARE AND CRAMÉR'S *V* TO DETERMINE SYSTEMATIC RELATIONSHIPS

You may want to evaluate the statistical relationship between categorical variables such as gender, ethnicity and religious affiliation. When both variables are categorical, you can't produce a mean. Therefore, to conduct this analysis, you use what is called a contingency table, which shows the frequency with which cases fall into each combination of categories such as ‘man’ and ‘Christian’.

When you conduct statistical analysis of categorical data like this, what you are trying to work out is whether there is any systematic relationship between the different variables being analysed or whether cases are randomly distributed across the cells in the contingency table. You do this by comparing what you find in your sample data with what would be expected if there were no relationship between the two variables (this ‘no relationship’ scenario is the null hypothesis in this case).

For each cell in the table there is an expected frequency, which SPSS will calculate for you and then compare with the observed frequencies from your sample data to see how much difference there is. The expected frequency is what you would get in each cell if there was an exactly equal spread of the data across the different categories in the table. The statistical test that tells you whether the difference is large enough to reject the null hypothesis and conclude that the variables are related is called a *chi-square test*.

If the result of this test, the chi-square statistic, is large enough to be statistically significant, this means that the difference in our sample is large enough that we would see it less than 5% of the time *if there were no real difference between them in the population* (which is the null hypothesis). This would mean you could reject the null hypothesis, and say that there is in fact a relationship between the variables you are analysing.

As is always the case, there are some assumptions to keep in mind for the chi-square statistic.

- You can only use chi-square tests where at least 80% of the cells in the table have an expected value of at least 5 (and all should be 1 or greater). If this assumption is violated the test loses power and may not detect a genuine effect (Yates, Moore, & Starnes, 1999, p. 734).
- The categories must be discrete. That is, no case should fall into more than one category (e.g., a respondent cannot be both male *and* female at the same time).

Once you have verified these chi-square test assumptions, you can run your chi-square test. To do so, take the following steps:

1. From the Analyze menu, select Descriptive Statistics (see Figure 12.19).
2. Select Crosstabs.
3. Highlight the variables you want, clicking the arrow to add to the Row(s) and Column(s) variable lists, then click OK.
4. Select Statistics (see Figure 12.20). Tick Chi-square. Tick Phi and Cramer’s V. Click Continue, then click OK.
5. Select Cells, look in Percentages and tick Column. Click Continue, then click OK.

Your first SPSS test result output will be a crosstabular table (see Figure 12.21), which presents the percentages and raw numbers in each category of your analysis, as well as the row and column totals.

Figure 12.19 Running the chi-square analysis

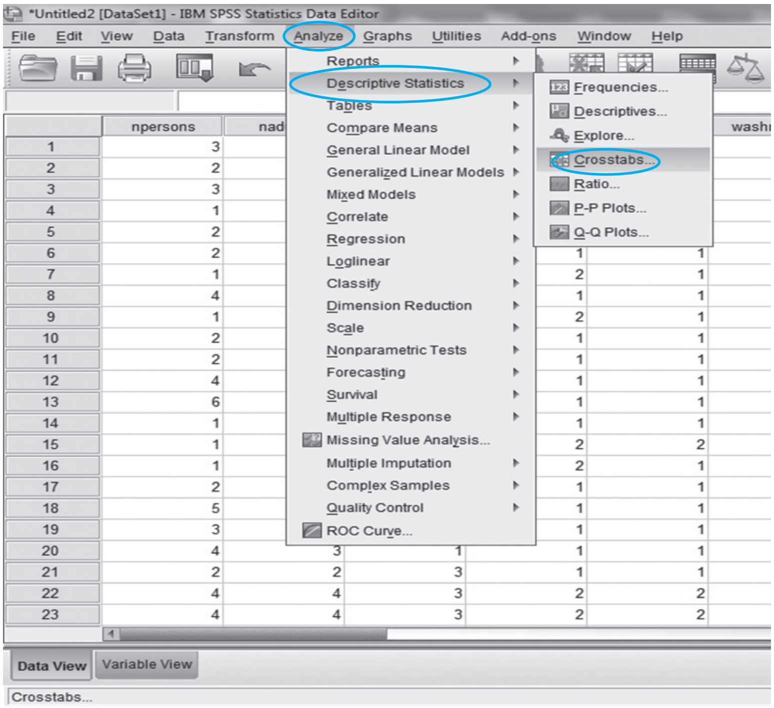


Figure 12.20 Selecting chi-square and Cramér's *V* analysis in SPSS

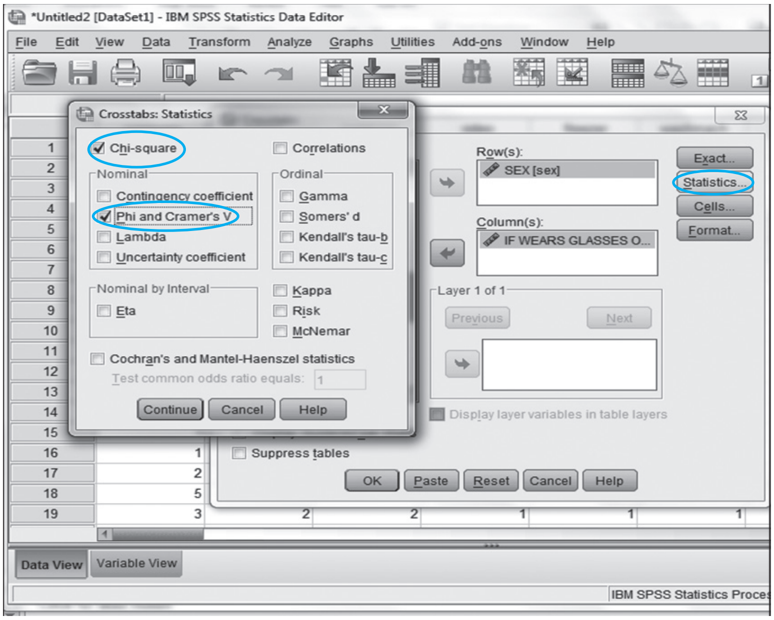


Figure 12.21 SPSS output – crosstabular table

HaveyouvisitedChesterZoobefore * Gender Crosstabulation

			Gender		Total
			Boy	Girl	
HaveyouvisitedChesterZoobefore	Yes	Count	133	117	250
		% within Gender	89.9%	98.3%	93.6%
	No	Count	15	2	17
		% within Gender	10.1%	1.7%	6.4%
Total		Count	148	119	267
		% within Gender	100.0%	100.0%	100.0%

The second part of the test result output in SPSS will be your chi-square results table (see Figure 12.22). The important details for you to extract from this table are as follows:

Figure 12.22 SPSS output – chi-square test result

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	165.785[a]	5	.000
Likelihood Ratio	164.320	5	.000
Linear-by-Linear Association	76.914	1	.000
N of Valid Cases	6303		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 165.69.

1 Look at the footnote at the bottom of the table. This tells you whether there are any cells in the crosstabular table with an expected count less than 5. If this number is 0, you are fine to continue. If not, then check to ensure you have at least 80 per cent of cells with expected counts of 5 or greater and all cells have values of at least 1. If so, then you can continue. If not, then you need to adjust your data (e.g., by merging low-frequency categories) to reduce the number with expected counts that are less than 5.

2 Look at the first row ('Pearson Chi-Square') under the column 'Asymp. Sig. (2-sided)'. This is your *p*-value. If it is less than 0.05, then the result is statistically significant. This means you can reject the null hypothesis and conclude that there is a relationship between the two categorical variables in this analysis.

3 Look at the first row ('Pearson Chi-Square') under the column 'Value'. This is your chi-square value, which is your evidence that the result is (or is not) statistically significant. You will need this information when your write up the results of your chi-square analysis.

At this stage, if you've determined that your chi-square result is statistically significant, you'll next need to determine the strength of the relationship between the two variables.

12.7.1 FIND THE EFFECT SIZE FOR YOUR STATISTICALLY SIGNIFICANT CHI-SQUARE RESULT

If you find a statistically significant relationship with your chi-square analysis, this means there is likely to be a real pattern in the population from which you sampled. However, to determine the size of the effect, you must run a second test called *Cramér's V*. This is a proportional reduction in error statistic, which indicates the strength of associations in contingency tables. It also enables us to compare different associations and decide which is stronger. Another way of describing proportional reduction in error is how much better your prediction of the outcome variable will be if you know something about the predictor variable. Proportional reduction in error statistics such as Cramér's *V* range from 0 to ±1. Roughly speaking, the following guidelines indicate the strength of the relationship between the variables (Cohen, 1988):

- 0.10 – Small effect size
- 0.30 – Medium effect size
- 0.50 – Large effect size

The example in Figure 12.23 shows a very weak effect (0.046 is very close to 0). This very weak relationship means that other variables are needed to provide a more complete explanation for the outcome variable being measured in this analysis.

Figure 12.23 Cramér's *V* results table from SPSS output

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.065	.013
	Cramer's V	.046	.013
N of Valid Cases		3018	

12.7.2 WRITE UP YOUR CHI-SQUARE AND CRAMÉR'S *V* RESULTS

The chi-square result tells you if there is a relationship between two variables (yes or no). When there is a relationship, Cramér's *V* is used to determine the strength of that relationship. When writing up your results, it is important to include information on both.

Begin your write up by restating the problem. For example: 'A chi-square test of independence was conducted comparing the previous attendance at Chester Zoo across boys and girls.' Follow this by telling the reader whether the result was significant, and by providing the relevant statistics: 'A significant interaction was found ($\chi^2(\#) = \#.\#\#$, $p = 0.\#\#\#$, Cramér's $V = 0.\#\#\#$).' The values provided are the degrees of freedom (df), followed by the chi-square test value, the *p*-value, and finally the value for Cramér's *V*. Conclude the write-up by restating the findings, and by providing the proportions: 'boys ($n = \#$) were more likely than girls ($n = \#$) to be attending the Zoo for the first time'.

As a final piece of advice, avoid using contingency tables with interval variables such as age in years, as most people would fall into different categories and so the tables would be enormous and unmanageable.

12.8 CONDUCT CORRELATION ANALYSIS TO MEASURE RELATIONSHIPS BETWEEN VARIABLES

Correlation analysis measures the relationship between two interval variables. This form of analysis tracks whether deviations from the mean 'covary' (i.e., vary together) in a systematic way.

Based on this covariance, Pearson's sample correlation coefficient (*r*) represents both the direction and the strength of the association between two numerical variables. A correlation can be either positive or negative. A positive correlation occurs when two variables change in the same direction together (going up together or down together). A negative correlation occurs when two variables change together but in opposite directions (as one goes down, the other goes up). You can tell how strong the correlation is just by looking at the *r*-value. The closer the *r*-value is to 1 (which represents a perfect positive correlation) or to –1 (which represents a perfect negative correlation), the stronger the correlation. When the *r*-value is closer to 0 (which represents no linear correlation at all), you know the correlation is weaker.

KEY TIPS

Avoid inflating the risk of Type I error with lots of correlation analyses

It may be tempting to just run your correlation analyses on as many different variable combinations as possible until you get one that is statistically significant. However, you should only run this test when you have a good reason to do so. Given that the statistical significance threshold is that there is only a one out of 20 chance of getting a significant result that is due to random variation between sample and population, rather than a real difference that exists in the population, if you run 20 correlation tests, the odds are that one of them will come up as a false positive, that is, it will show as statistically significant when there is no real correlation in the population (Leek & Storey, 2011).

12.8.1 TEST THE ASSUMPTIONS OF PEARSON'S CORRELATION TEST

As you prepare to use a Pearson's correlation test, you must take a moment to check that the test's assumptions are met by the data you're using. We've already covered the fact that the Pearson correlation test assumes you will be using interval data. However, there are other assumptions it shares in common with other parametric statistical tests:

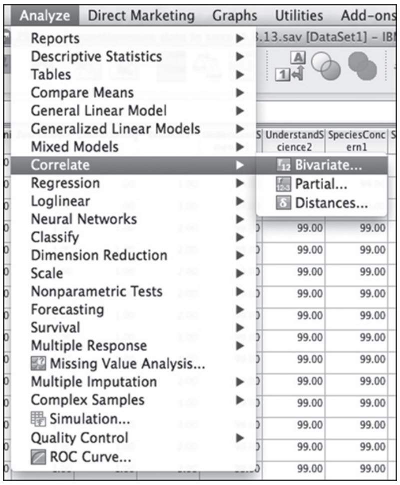
- **Normality**. The easiest way to determine if a variable is normally distributed is simply by looking at the histogram (as discussed earlier in this chapter). If the distribution of the variable appears to follow the shape of a bell curve, then it is usually safe to assume that the variable is normally distributed.
- **Linear relationship**. As with normality, the easiest way to determine whether two variables share a linear relationship (as opposed to a curvilinear relationship) is visually. Create a scatter plot (as discussed earlier in this chapter) placing the outcome variable on the *y*-axis and the predictor variable on the *x*-axis. You may assume a linear relationship as long as there is not a U-shaped curve.

- **No significant outliers**. The scatter plot is also a good way to look for significant outliers. If, when you look at the plot, there is one value that appears far away from the rest of the data, it is most likely an outlier. While there are mathematical ways to determine whether a case is an outlier, a simple scatter plot is sufficient for the purposes of this book.

12.8.2 RUNNING YOUR CORRELATION ANALYSIS AND INTERPRETING THE SPSS OUTPUT

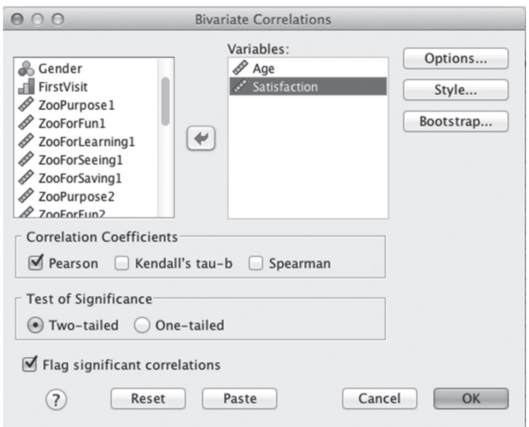
If the correlation assumptions have been upheld, you can now conduct your correlation analysis by selecting Analyze → Correlate → Bivariate (see Figure 12.24).

Figure 12.24 Selecting bivariate Pearson correlation in SPSS



In the Bivariate Correlations dialog box, you will need to move the variables you wish to analyse over into the Variables field (see Figure 12.25). Be sure to only use interval variables!

Figure 12.25 Bivariate Correlations dialog box in SPSS



When you click OK, what you get from SPSS is a correlation matrix (see Figure 12.26). This is a table that shows how each variable correlates against each of the other variables, including itself. In the table, Sig. is the *p*-value and *N* is the number of cases. So in this case, the *p*-value is 0.127 (not statistically significant, because this number is larger than 0.05), and we have $r = -0.028$ and $N = 3018$. As the result is not statistically significant, we would conclude that there is no relationship between age and satisfaction.

Figure 12.26 Correlation matrix (non-significant result)

Correlations

		Age	Satisfaction
Age	Pearson Correlation	1	-.028
	Sig. (2-tailed)		.127
	N	3018	3018
Satisfaction	Pearson Correlation	-.028	1
	Sig. (2-tailed)	.127	
	N	3018	3018

12.8.3 EVALUATE THE EFFECT SIZE FOR YOUR PEARSON CORRELATION RESULTS

For the Pearson correlation, you can use your *r*-value to determine the strength of the association between the two variables you are testing. The *r*-value without any adjustment gives an indication of effect size, as follows (based on Hopkins, 2013):

- 0.90 to 1.00 (almost perfect relationship)
- 0.70 to 0.90 (very strong relationship)
- 0.50 to 0.70 (strong relationship)
- 0.30 to 0.50 (moderate relationship)
- 0.10 to 0.30 (weak relationship)
- 0.00 to 0.10 (very weak relationship)

However, you can more precisely identify the effect size by squaring the *r*-value (i.e., multiplying it by itself). This is how you get r^2 (also called the coefficient of determination). The coefficient of determination tells you the percentage of the variation in one variable that is explained by variations in the other variable. Here's an example. If you have an *r* value of

0.3, then r^2 would be $0.3 \times 0.3 = 0.09$. If $r^2 = 0.09$, this means that 9% of the variation in one variable can be explained based on variation in the other variable.

WRITE UP YOUR PEARSON'S CORRELATION RESULTS

Using the key details from the SPSS results output, you can present your findings. If your result is significant, you could explain it as follows:

> A Pearson's correlation analysis showed that there is a statistically significant relationship between variable A and variable B ($r = 0.6$, $n = 98$, $p = 0.004$). The r^2-value is 0.36, which indicates that the correlation between variable A and variable B accounts for 36% of variance.

On the other hand, if your result is non-significant, you could simply explain it as follows:

> A Pearson's correlation analysis showed that there was no relationship between variable A and variable B ($r = 0.02$, $n = 98$, $p = 0.07$).

12.8.4 NON-PARAMETRIC ALTERNATIVE TO THE PEARSON CORRELATION

If you find that one or more of the key assumptions for parametric statistical tests has been violated, you will need to turn to the non-parametric alternative test known as *Spearman's rank-order* correlation. This test does not give you *r*-values and it is less sensitive to patterns in your data set than the Pearson correlation, which is why it is considered a second-best option only to be used when violated assumptions make it unavoidable.

The test statistic for the Spearman rank-order correlation is known as 'Spearman's rho' (or ρ). The process for calculating Spearman's rho in SPSS is almost identical to that for calculating Pearson's *r*. Set up the test just as you would set up the *Pearson correlation* test, by selecting Analyze > Correlate > Bivariate. Move the two variables you want to compare into the Variables box. Under Correlation Coefficients, uncheck the Pearson box, and instead check Spearman. When this is done, click OK.

The output for this test is almost identical to the output for Pearson's *r*, and can be interpreted in a similar way. As with the other form of correlation, the important information includes the significance level (*p*-value), the correlation coefficient (rho) and *N* (the number of cases).

WRITE UP YOUR SPEARMAN'S RHO RESULTS

Using the key details from the SPSS results output, you can present your findings. If your result is significant, you could write it as:

> A Spearman's rank-order correlation shows that Group A tended to rank higher in OUTCOME than Group B, r_s = 0.##, p = 0.###.

On the other hand, if your result is non-significant, you could simply write it as follows:

> A Spearman's rank-order correlation showed no difference in OUTCOME between Group A and Group B, r_s = 0.##, p = 0.###.

12.8.5 INTERPRETING CORRELATION

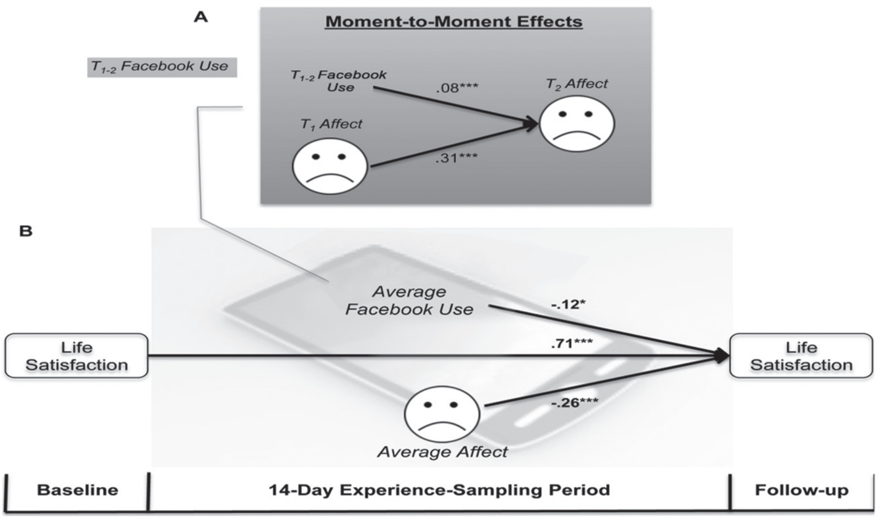
Correlation does not necessarily indicate causality, because there may be some other variable affecting the relation between the two variables under question. For example, a study finding a correlation between annual income levels and brain size might conclude that larger brains lead to higher incomes. However, both brain size and income could be affected by many other factors such as physical size of person, which may in fact be the real causal factor involved. In sum, the correlation coefficient says nothing about which variable is *causing* the other to change.

REAL WORLD EXAMPLE

Facebook and life satisfaction

Kross, Verduyn, Demiralp, Park, and Lee (2013) conducted a study of the relationship between Facebook use and moment-by-moment happiness and life satisfaction. They did this by text-messaging 82 people over a two-week period, sending them questions five times per day about how they felt at the time of receiving the text message. They then ran statistical analysis on the data they got back and found that Facebook use at one time point was correlated with negative shifts in mood and 'life satisfaction' the next time they contacted the participants. They also found that there was a general correlation between the amount of Facebook use over the two-week period and their life satisfaction levels (as Facebook use went up, life satisfaction levels went down). They concluded: "Facebook use predicts declines in the two components of subjective wellbeing: how people feel moment to moment and how satisfied they are with their lives" (Kross et al., 2013, p. 4). Thus, the authors of this study are implying that Facebook use causes declines in people's perceived well-being.

Figure 12.27 Facebook and life satisfaction



Source: Kross et al. (2013)

However, this claim of 'prediction' is actually only referring to correlation. This study only asked about a few things during the 'experience sampling' five times per day: measures of feelings, stated Facebook use and self-reported level of social contact through face-to-face or phone interactions. This means that anything else outside of these factors could have affected people's feelings, For example, perhaps a factor not accounted for in the analysis such as the time of day might have affected people's decision to use Facebook. Or a text message (not measured in the study) from a friend or loved one could have triggered negative feelings. With such a limited frame around the analysis, the causes of the outcomes measured in the study are effectively unknown. It is important to be very careful about drawing causal inferences in such cases.

12.9 USE *T*-TESTS TO COMPARE MEANS

If you have one interval outcome variable (e.g., income) and you want to compare means across two categories (e.g., men and women), then the *t*-test is where you should start. There are two different types of *t*-tests that will be discussed in this chapter:

- ***t*-tests with two independent samples** ('independent samples *t*-test'). You might use this statistical analysis to compare boys' and girls' results on a math test. The defining feature of this kind of statistical analysis is that the data you're comparing come from different groups of people.
- ***t*-test with two paired samples** (also called the 'dependent samples *t*-test'). This statistical test compares two means using data gathered from the same people measured at different times. For example, if you compared responses before and after a visit to the zoo for the same sample of children. That is, these data are from 'matched' samples.

When two samples of data are collected and the sample means are calculated, these means might differ by either a little or a lot. If the samples come from the same population, then we must start with the expectation that their means will be roughly equal (this is the null hypothesis). Although it is possible for their means to differ by chance alone, we would expect large differences between sample means to occur very infrequently. We compare the difference between the sample means that we collected to the difference between the sample means that we would expect to obtain if there were no effect (i.e., if the null hypothesis were true). If the difference between the samples we have collected is larger than what we would expect if the null hypothesis is true, then we can conclude one of two things:

1. **There is no effect.** Sample means in our population fluctuate and we have, purely by chance, gathered two samples that are not representative of the population from which they came.
2. **There is a statistically significant difference**. The two samples come from different populations but are typical of their respective parent population. In this scenario, the difference between samples represents a genuine difference between the samples (and so the null hypothesis is false).

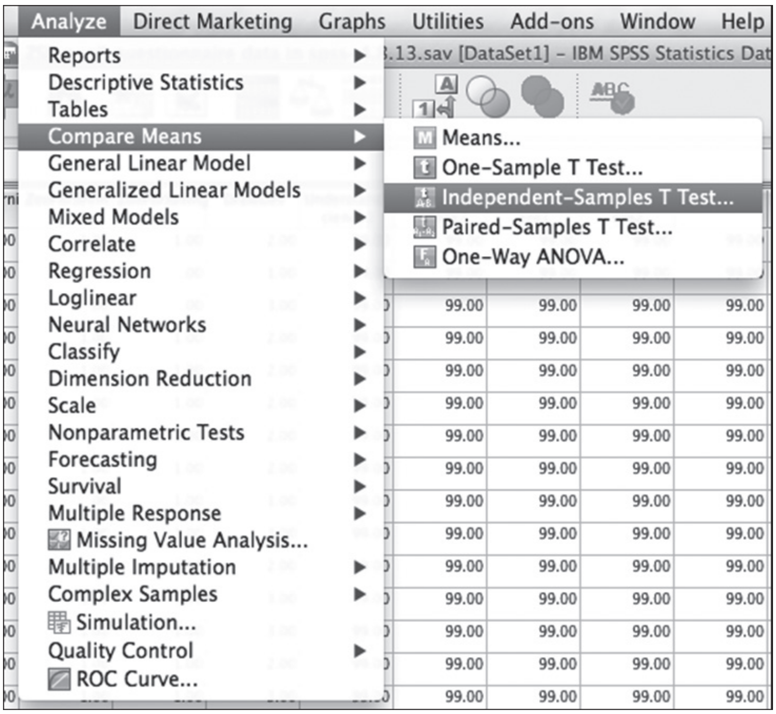
As the observed difference between the sample means gets larger, there can be more confidence that the second conclusion is correct (i.e., that the null hypothesis should be rejected). In sum, the *t*-value is effectively the difference between what you would expect if the null hypothesis (identical means in population) is true and the actual sample data you've collected. If the data you've collected indicates a large enough difference, you can reject the null hypothesis and say you have discovered a statistically significant difference between the two means. For the independent samples *t*-test, the main assumptions are *normality* and *equality of variances*. For the paired samples *t*-test, the main assumption is normality. Once you've decided the *t*-test is appropriate for your analysis, you need to test its assumptions. However, in this case, you can test the assumptions in SPSS as part of your *t*-test analysis.

To check whether your data have equality of variances, Levene's test is used; it tests the hypothesis that the variances in the groups are equal. If Levene's test is statistically non-significant ($p > 0.05$), then equality of variances is confirmed (i.e., all is well and you can carry on with your analysis as planned). If Levene's test is statistically significant ($p < 0.05$), this means that the variances are not equal (the assumption is violated). This indicates that you shouldn't be using the standard *t*-test. If you find that the *t*-test assumptions are not supported, then you should skip to the non-parametric alternatives (see Section 12.9.3).

12.9.1 RUN YOUR *T*-TEST

If the *t*-test assumptions have been supported, you can now run your *t*-test (see Figure 12.28). Select Analyze → Compare Means → *t*-test (choose between Independent-Samples T Test and Paired-Samples T Test). Each type of *t*-test will bring up a different dialog box.

Figure 12.28 Running a *t*-test in SPSS

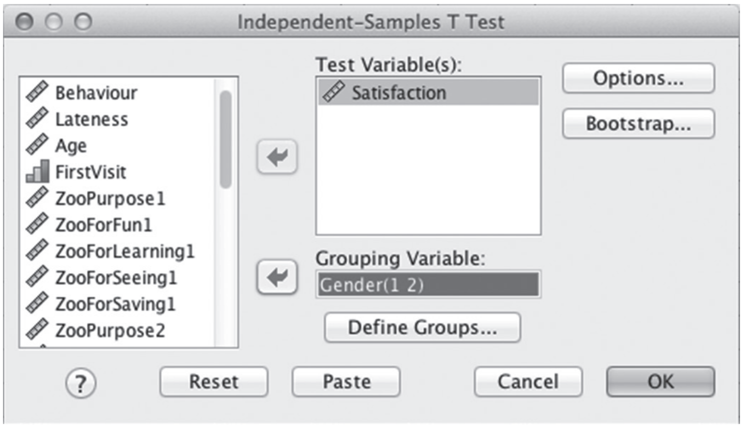


INDEPENDENT SAMPLES *T*-TEST

For the independent samples *t*-test, you need to take the following steps (see Figure 12.29):

- Choose an interval variable as your Test Variable.
- Choose a categorical variable as your Grouping Variable.
- Click on Define Groups and identify the numbers that have been assigned to each group for this variable (e.g., 1 and 2). Then click Continue, then OK to run the analysis.

Figure 12.29 Selecting variables for independent samples *t*-test in SPSS



You will receive a multi-part output from SPSS. The top part of the output (see Figure 12.30) provides information on the mean, standard deviation and standard error of the mean for the two variables you're analysing, as well as the sample sizes for each category you're looking at (N). You will need the sample sizes and standard error of the mean when you write up your results.

The Levene's test shown in Figure 12.31 has a statistically significant result (Sig. is less than 0.05). This means that the equality of variances assumption of this statistical test may have been violated. This in turn indicates that you shouldn't use the *t*-test, but will need to switch to a non-parametric alternative.

Figure 12.30 Top part of *t*-test SPSS output: Summarizing your data

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Satisfaction	1.00	1452	17.3485	35.81107	.93980
	2.00	1526	14.2192	32.56893	.83373

Figure 12.31 Part of *t*-test SPSS output: checking equality of variance assumptions – equality of variances assumption violated

		Levene's Test for Equality of Variances	
		F	Sig.
Satisfaction	Equal variances assumed	26.082	.000
	Equal variances not assumed		

If you have a non-significant Levene's test result (i.e., Sig. greater than 0.05), then you can continue with your *t*-test. Your independent samples *t*-test will have an SPSS output, which tells you whether the result is statistically significant (see Figure 12.32).

Figure 12.32 Example *t*-test results output in SPSS

Independent Samples Test

t-test for Equality of Means						
					95% Confidence Interval of the Difference	
t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
2.497	2976	.013	3.12928	1.25336	.67174	5.58683
2.491	2915.467	.013	3.12928	1.25631	.66593	5.59264

You can immediately see if a result is statistically significant by checking the column labelled 'Sig. (2-tailed)'. If this number is lower than 0.05, then your result is statistically significant. This means you would be able to reject the null hypothesis and conclude that there is a real difference between the two groups in the population from which your sample was drawn.

In order to write up your results, the important details you need to gather from this table are the sample sizes (N), standard errors of the means (SE), *t*-value and degrees of freedom (df), the Sig. value (p = .013 in the example above), as well as your measures of effect size. Here is an example of a write-up using this information:

> An independent two-samples *t*-test was conducted to compare the mean levels of satisfaction when zoo visits were self-guided or education officer-led. The results show that self-guided visits yielded significantly greater satisfaction levels (M = 1.821, SD = 0.988): $t(1592) = 3.944$, $p = 0.000$,* when compared with education officer-led visits (M = 1.659, SD = 0.913). These results indicate that education officer led visits lead to *lower* levels of satisfaction with the zoo.

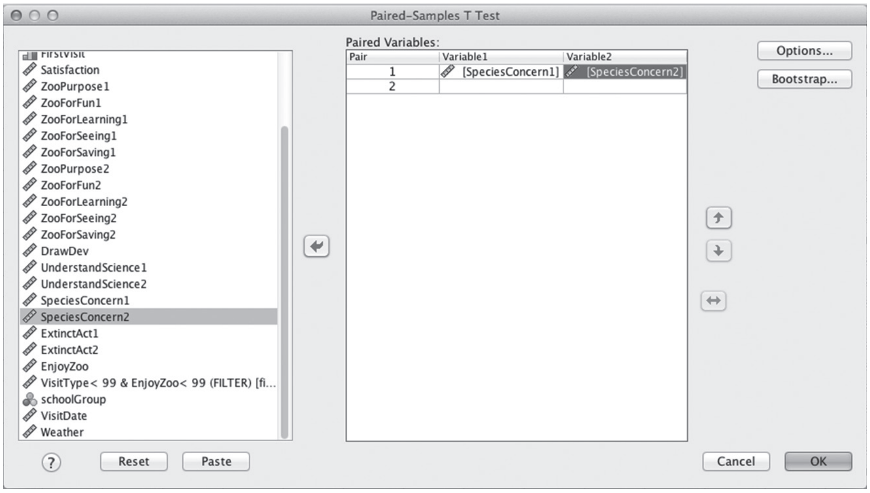
*(footnote) Levene statistic = 5.549, $p < 0.05$ (outcome suggests non-equal variances, therefore variance ratio calculated to confirm *t*-test assumption of equality of variances; this assumption was confirmed by Hartley's F_{max} = 1.170 (n = 2415).

While you can use different words when you write up your *t*-test, you should be sure to include the key statistical information provided here.

PAIRED SAMPLES *T*-TEST

For the paired samples *t*-test, you'll get a dialog box that has spaces for you to place your two paired variables (see Figure 12.33). Select the variables and then click 'OK'.

Figure 12.33 Selecting variables for paired samples *t*-test



Your paired-samples *t*-test results output includes a set of three tables. The important details for you to extract from the top table (Figure 12.34) are the mean, the sample size and the standard deviation, and from the bottom table, the *t*-value, degrees of freedom (df), and *p*-value (Sig.).

Figure 12.34 Paired samples *t*-test output in SPSS

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	SpeciesConcern1	88.4344	3018	30.36727	.55277
	SpeciesConcern2	88.0457	3018	30.85138	.56158

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	SpeciesConcern1 & SpeciesConcern2	3018	.868	.000

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	SpeciesConcern1 - SpeciesConcern2	.38867	15.71088	.28598	-.17207	.94941	1.359	3017	.174

Here is how you could write up a paired samples *t*-test result:

> A paired samples *t*-test compared pre- and post-visit data on children's concern for the conservation of animal species before and after a visit to London Zoo. There was no significant difference on this measure assessing pre-visit scores (M = 88.43, SD = 30.37) and post-visit (M = 88.05, SD = 30.85): $t(3017) = 1.359$, $p = 0.174$. These results indicate that the zoo visit had no effect on this outcome for children.

12.9.2 EVALUATE THE EFFECT SIZE OF YOUR *T*-TEST RESULT

If your result is statistically significant, you next need to determine the effect size. For *t*-tests, the appropriate effect size measure is known as Cohen's *d*, which is simply a standardized measure of the difference between two means. This value is not automatically calculated in SPSS, although the *t*-test output does provide all the necessary elements for calculating this effect size on your own. To avoid having to do these calculations yourself, we suggest using http://www.uccs.edu/~lbecker/ for calculating Cohen's *d*.

For determining the effect size for a comparison of two groups (whether between subjects or within subjects), all you need is the mean of each group (i.e., M_1 and M_2), and the standard deviation of each group (i.e., SD_1 and SD_2). These values are found in the first set of outputs for each *t*-test. Enter the information into the appropriate boxes on the website, and click on Compute.

12.9.3 HOW TO USE NON-PARAMETRIC ALTERNATIVES TO THE *T*-TEST

If the *t*-test assumptions are broken, your results may not be valid. Specifically, if the *t*-test assumptions are violated, it can result in a lack of statistical power, meaning that you may fail to reject the null-hypothesis when in fact you should have (that is, a false negative). It may also result in seemingly nonsensical results. This is especially true if you have small sample sizes and/or drastically unequal groups you're comparing. In these cases, you should use a non-parametric test instead.

NON-PARAMETRIC ALTERNATIVE TO INDEPENDENT SAMPLES *T*-TEST (MANN–WHITNEY)

An alternative to the independent samples *t*-test is the Mann–Whitney *U* test. The basic function of the Mann–Whitney *U*-test is similar to the *t*-test, except that you compare differences from within the *mean rank* rather than differences in the *mean*. The 'mean rank' refers to where each individual in the sample sits in reference to the other individuals. For example, the 10th highest score would be given a rank of 10, and the 11th highest a rank of 11. Then, for each group, the test is run on the mean of these ranks, instead of on the mean of the scores themselves. Of course, you don't need to rank each case by hand: This is done automatically by SPSS. To run a Mann–Whitney *U*-test, take the following steps:

3. Go to Analyze > Nonparametric Tests > Legacy Dialogs > 2 Independent Samples.
4. You'll now see a dialog box that is very similar to the one for the independent samples *t*-test. Move your outcome variable to the Test Variable List box, and the predictor variable to the Grouping Variable box.

5 Under the grouping variable box, click Define Groups.
6 Input the grouping values, just as you would for the *t*-test.
7 Click Continue.
8 Make sure the Mann-Whitney U box is checked, and no others. Click OK.

To determine whether the test is statistically significant, look at the value labelled "Asymp. Sig. (2-tailed)" (i.e., asymptotic significance). If this value is less than 0.05, then your test is statistically significant. You can report the results of the Mann–Whitney *U*-test exactly the same as you would with a *t*-test, just substituting *U* for *t* (i.e., *U* = ###, *p* = 0.###), and noting that there are no degrees of freedom" to report.

NON-PARAMETRIC ALTERNATIVE TO PAIRED-SAMPLES *T*-TEST (WILCOXON SIGNED RANK)

An alternative to the paired-samples *t*-test is called the Wilcoxon signed rank test. Similar to the Mann–Whitney *U*-test, the Wilcoxon signed rank test shares the same basic function as a *t*-test, although instead of comparing means, it compares rank means. To run the Wilcoxon signed rank test, take the following steps:

1 Go to Analyze > Nonparametric Tests > Legacy Dialogs > 2 Related Samples.
2 Transfer the variables you are interesting in analysing into the Test Pairs box, making sure that the time 1 scores go into Variable1 and the time 2 scores go into Variable 2.
3 Under Test Type, make sure that Wilcoxon is selected, and no others.
4 Click OK.

Next you will need to interpret the results. There should be two output tables, one labelled 'Ranks' and one labelled 'Test Statistics'. The first gives you an indication of how many individuals improved their rank from time 1 to time 2, while the second tells you whether the score is significant. In the Ranks table, 'Negative Ranks' are the number of individuals who scored lower at time 2, 'Positive Ranks' are the number of individuals who scored higher at time 2, and 'Ties' are individuals who scored the same at time 2. This information is interesting, although you generally don't need to report it.

The Test Statistics table gives you everything you need in order to report the significance of the test. The *p*-value is labelled 'Asymp. Sig. (2-tailed)'. If this value is less than 0.05, then your test is statistically significant.

You can report the results of the Wilcoxon signed rank test in exactly the same way as you would a *t*-test, substituting Z for *t* (i.e. Z = #.###, *p* = 0.###). As with the Mann–Whitney test, there are also no degrees of freedom to report.

12.10 LOOK TO THE STATISTICAL HORIZON BEYOND THIS CHAPTER

As can be seen in Figure 12.35, there are several other statistical tests you may need to use that are not covered in this chapter. In Table 12.3, we signpost some of the statistical analyses beyond the scope of this chapter so that you know where to go next if your statistics needs are not addressed here. If your situation calls for statistical tests not covered in this chapter, we recommend you consult Field (2014) for detailed guidance.

Figure 12.35 Map of statistical test options by number and type of variables

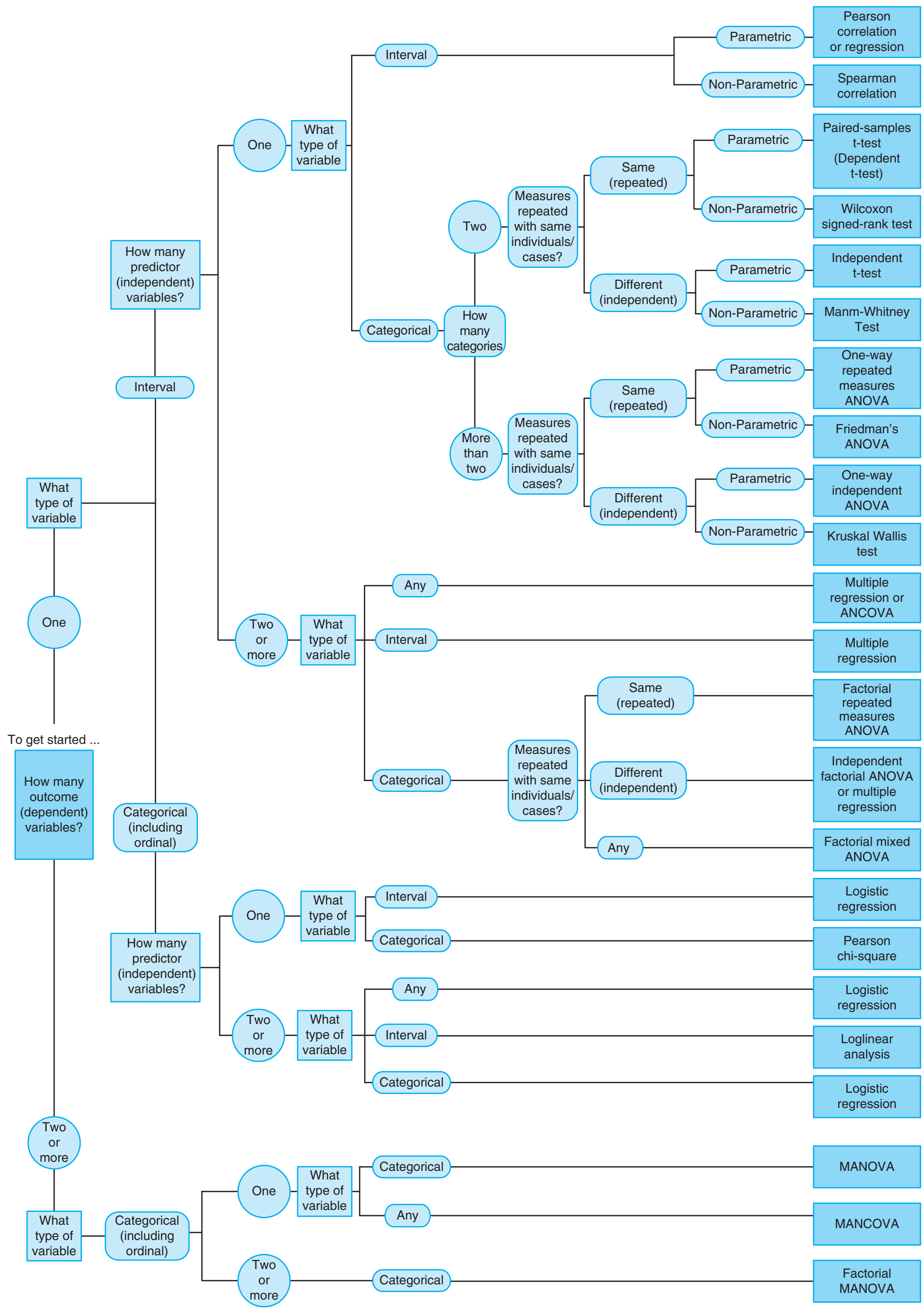


Table 12.3 Steps to the statistical horizon

Type of Research Question	Number and Type of Outcome Variable(s)	Number and Type of Predictor Variable(s)	Covariates	Test to Use
Group differences	1 Interval	1 Categorical	-	One-way ANOVA
		1 Categorical	1+	One-way ANCOVA
		2+ Categorical	-	Factorial ANOVA
			1+	Factorial ANCOVA
	2+ Interval	1 Categorical	-	One-way MANOVA
			1+	One-way MANCOVA
		2+ Categorical	-	Factorial MANOVA
			1+	Factorial MANCOVA
Degree of relationship	1 Interval	2+ Interval	-	Multiple Regression
	1+ Interval	2+ Interval	-	Path Analysis
Prediction of group membership	Nominal with 2 levels (e.g., Male/ Female)	2+ Categorical, Ordinal or Interval	-	Logistic Regression

12.11 CONCLUSION

Quantitative analysis aims to create efficient descriptions and summaries of patterns using standardized data sets and statistics. This chapter began with an introduction to content analysis as a method of converting qualitative data into a systematic numerical form that can be used in statistical analysis. The linchpin in the content analysis process is inter-coder reliability, which is a necessary (although not a sufficient) criterion for validity in the study. Without it, all results and conclusions in the research project may justifiably be doubted. We emphasized that all content analysis projects should be designed to include the assessment and reporting of this crucial quality assurance measure.

Whether created through content analysis or other methods, your quantitative data can be used to describe patterns in your data and make comparisons. They can also be used to test hypotheses using statistics to calculate whether a result is 'statistically significant', that is, whether it points to a real pattern in the population. For example, the chi-square statistic can tell you whether there is a 'significant' association between two categorical variables (or whether there exists an association that would be unlikely to be found by chance). Once a significant pattern is detected, you'll need to determine the strength of association using a different test (Cramér's *V* in the case of a significant chi-square result).

Statistical tests often have assumptions that must be met in order to use them. Parametric statistical tests tend to assume both normality and equality of variances. Correlation analysis also assumes linearity. If these assumptions are satisfied, you can use the Pearson correlation to investigate whether there is a relationship between two interval variables. However, it is important to limit the risk of a false positive result by only running correlation analyses when there is a good reason.

We then discussed options for comparing means between different categories of a predictor variable. If two samples come from populations with identical population means, then you would expect the difference between the sample means to be (close to) zero. Thus, the larger the difference between the two sample means, the more likely it is that the two population means are different. This is what the *t*-test evaluates.

Finally, we highlighted a range of other statistical options you can learn more about in other books or articles. This chapter only provides the bare essentials for you to be able to run statistical analyses in SPSS. It's likely you'll need to do some further reading if you have any unusual situations or analysis needs that are not covered here.

SUGGESTIONS FOR FURTHER READING

- Field, A. (2014). *Discovering statistics with IBM SPSS Statistics*. London: Sage. This book discusses the theoretical and practical aspects of a number of key statistical concepts and analyses. It strikes a very informal tone, but provides a good level of detail about how to conduct statistical tests using SPSS and related contextual information about these statistics.
- D'Agostino, R. B., Belanger, A., & D'Agostino, R. B. (1990). A suggestion for using powerful and informative tests of normality. *American Statistician*, *44*, 316–321. This article provides guidance about different options for assessing whether normality assumptions have been met in your statistical analyses.
- Kotrlik, J. W., Williams, H. A., & Jabor, M. K. (2011). Reporting and interpreting effect size in quantitative agricultural education research. *Journal of Agricultural Education*, *52*(1), 132–142. This article identifies different options for measuring effect size for a range of statistical tests.
- Cohen, J. (1992). Power primer. *Psychological Bulletin*, *112*(1): 155–159. This article is a brief, accessible and prescriptive look at the concept of statistical power. Perhaps the most notable thing about this article (and the reason it has been cited over 8500 times!) is the simple rules of thumb Cohen provides for interpreting effect size measures. This article is a quick read, and an incredibly useful citation.
- Williams, F., & Monge, P. (2001). *Reasoning with statistics: How to read quantiative research* (5th ed.). Boston: Wadsworth Cengage Learning. The world of quantitative statistical testing expands far beyond what we cover in this chapter. This book offers an excellent cursory look at many additional statistical methods which you may run into as you read social science research. Importantly, the book is not (primarily) designed to teach you *how to do* statistics, but instead on *how to read and understand* a large verity of advanced statistical analyses.
- Krippendorf, K. H. (2013). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage. Krippendorf is by many accounts the father of modern-day content analysis. This textbook covers everything you need to know in order to be an expert on doing and critiquing content analysis.

GLOSSARY

Central tendency – This is the typical pattern in your data, which can be assessed by calculating the mean or median.

Chi-square test – This test is used when you want to evaluate whether two categorical variables are related.

Coding – The process of systematically converting qualitative content into quantitative data by categorizing raw data using a limited number of standardized categories that are suitable for analysis.

Correlation – The extent to which two variables have a relationship dependent on each other. For example, there is a correlation between eating high quantities of fatty foods and gaining weight.

Cramér's V – This test is used as a follow-up after a statistically significant chi-square result to determine the size of the effect.

Dependent variable – The variable that is changed or altered in a study. For example, the amount of exercise, the voting preference or the amount eaten can all be changed in a study. The easiest way to remember the difference between dependent and independent variables is to insert the variables into the following sentence: (your independent variable) brings about a change in (your dependent variable) and it is not possible for (your dependent variable) to bring about a change in (your independent variable).

Descriptive statistics – Mathematical methods used to summarize and interpret data properties. They are distinct from inferential statistics in that they do not infer the properties of the population. For example, the percentage of goals scored during a football season is a descriptive statistic of individual player or team performance.

Hypothesis – The null hypothesis, usually denoted H_0, states that there is no association/relationship between the variables, while the alternative hypothesis, denoted H_1, states that there is an association/relationship between the variables.

Independent variable – The variable that stands alone and is not changed by any process in your analysis. For example, someone's age could be an independent variable and is not going to change no matter how much exercise they do, how they vote or what they eat.

Inferential statistics – Mathematical methods that are based on probability theory to deduce or infer the properties of a population larger than that of the sample tested. This is done, for example, by testing hypotheses and deriving estimates.

Mean – Result of dividing the sum of the values by the total number of cases to give an average of all values.

Median – The exact mid-point in your data.

Mode – The most frequently occurring attribute in your data.

Standard deviation – A standardized measure of the variability in the sample. It tells you how good a 'fit' there is between the full data set and the mean as a measure of central tendency.

Statistical software – Computer programs, such as SPSS and STATA, that specialize in analysing statistical data. These purpose-built programs make it far easier for you to carry out a host of basic analyses from descriptive statistics to more complex regressions.

t-test – Analysis of two population means used to understand whether or not the difference between the means of two populations is significant. For example, a *t*-test could help you determine if there is a significant difference in alcohol consumption between men and women.

Univariate analysis – Simplest form of statistical analysis, used to describe a single variable.

Variable type – There are three main types of variable: categorical/nominal, ordinal and interval. *Categorical variables* have no intrinsic ordering to them. Gender, ethnicity and hair colour categories (e.g., blonde, brunette, ginger) are examples of categorical variables: male and female are just categories, not an attribute you can have more or less of. *Ordinal variables* do have an intrinsic order. However, the steps between each value are not equal. For example, educational qualifications are ordered (some are higher or lower than others) but the steps between qualifications are not equal. The step from bachelor's degree to master's is not the same as from master's to doctorate. *Interval variables* have intrinsic ordering and the steps between each value are equal. For example, money (e.g., US dollars) is this kind of data. An interval variable is similar to an ordinal variable, except that the intervals between the values of the interval variable are equally spaced.

Glossary of Quantitative Analysis Symbols

Symbol	Symbol Name	Meaning / definition
μ	population mean	mean of population values
σ^2	variance	variance of population values
σ_X	*standard deviation*	standard deviation value of random variable *X*
$\bar{x}$	sample mean	average/arithmetic mean
s^2	sample variance	variance estimator
s	sample standard deviation	standard deviation estimator

REFERENCES

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). New York: Routledge.

Cook, P. J. (2007). *Paying the tab: The costs and benefits of alcohol control*. Princeton, NJ: Princeton University Press.

Fawcett Society (2014). The time to act is now: Fawcett's gender pay gap briefing. Retrieved 22 February 2015, from http://www.fawcettsociety.org.uk/wp-content/uploads/2014/11/Fawcett-Equal-Pay-Day-report-November-2014.pdf

Field, A. (2014). *Discovering statistics with IBM SPSS Statistics*. London: Sage.

Glaser, B. G., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine de Gruyter.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*, 77–89.

Hopkins, W. G. (2013). A scale of magnitudes for effect statistics. Retrieved 24 February 2015, from http://www.sportsci.org/resource/stats/effectmag.html

Ingraham, C. (2014, 25 September). Think you drink a lot? This chart will tell you. http://www.washingtonpost.com/blogs/wonkblog/wp/2014/09/25/think-you-drink-a-lot-this-chart-will-tell-you/

Jensen, E. (2014). Evaluating children's conservation biology learning at the zoo. *Conservation Biology*, *28*(4), 1004–1011.

Kolbe, R. H., & Burnett, M. S. (1991). Content analysis research: An examination of applications with directives for improving research reliability and objectivity. *Journal of Consumer Research*, *18*, 243–250.

Krippendorff, K. (2004). Reliability in content analysis. *Human Communication Research*, *30*, 411–433.

Krippendorff, K. (2008). Systematic and random disagreement and the reliability of nominal data. *Communication Methods and Measures*, *2*(4), 323–338.

Kross, E., Verduyn, P., Demiralp, E., Park, J., & Lee, D. S. (2013). Facebook use predicts declines in subjective well-being in young adults. *PLoS ONE*, *8*(8), 1–6.

Leek, J. T., & Storey, J. D. (2011). The joint criterion for multiple hypothesis tests. *Statistical Applications in Genetics and Molecular Biology*, *10*(1), article 28.

MacQueen, K. M., McLellan, E., Kay, K., & Milstein, B. (1998). Codebook development for team-based qualitative analysis. *Cultural Anthropology Methods*, *10*(2), 31–36.

Neuendorff, K. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.

Williams, F., & Monge, P. (2001). *Reasoning with statistics: How to read quantiative research* (5th ed.). Boston: Wadsworth Cengage Learning.

Yates, D., Moore, D., & Starnes, D. (1999). *The practice of statistics*. New York: W. H. Freeman.