Australian Domestic & Family Violence Clearinghouse

READING STATISTICS: A GUIDE FOR PROFESSIONALS WORKING WITH DOMESTIC AND FAMILY VIOLENCE

ROCHELLE BRAAF
Senior Researcher, ADFV Clearinghouse
CHRISTINE EASTMAN
Senior Researcher, Social Policy Research Centre, UNSW
ISOBELLE BARRETT MEYERING
Research Assistant, ADFV Clearinghouse
KAREN WILCOX
Senior Researcher, ADFV Clearinghouse

KEY POINTS

- Statistical literacy is crucial for professionals working in the field of domestic and family violence, or for those who engage with domestic and family violence literature in their work.
- Practitioners and professionals with an understanding of statistics can determine what data is appropriate for their needs, assess data quality and limitations, critique data misuse, draw on statistical findings to inform their practice and include statistical analyses in their own research study design.
- Statistical data also provide valuable information to guide policy and program development, and inform evaluations of response effectiveness. For example, we can use statistical information to set benchmarks for rates of violence or intervention outcomes.
- In using and interpreting statistical data, we need to remain mindful that such information can be influenced by a range of factors, including the social and political context in which domestic violence occurs (such as typically low reporting and disclosure rates).
- Some of the most common statistical measures and terms used in quantitative research and explained in this Research and Practice brief include *samples, variables, measures of central tendency, significance and regression analysis*.

WHY STATISTICS MATTER

Increasingly, governments, services and judicial decision makers have been asked to consider 'the evidence base' in relation to their work with men, women or children to address domestic and family violence. Whether it is as evaluators of program effectiveness, monitors of participant progress in an intervention, or as careful assessors of new social science research or new theories, professionals require a sufficient understanding of statistics in order to critically consider new evidence, new theories and the value of new initiatives. Various types of questions confront domestic violence, human service, policy and legal system professionals that all require engagement with statistics if they are to be evidence-informed:

- *How common is domestic violence?*
- *Are protection order application rates increasing or decreasing?*
- *Does men's unemployment contribute to risk for perpetration of violence?*
- *Do respectful relationship programs prevent future violence?*
- *How do we make sense of findings that children who spend low levels of time with a non-resident parent are also more likely to be hostile towards that parent? Does this mean that hostility can be decreased by increasing parenting time?*

These types of questions can all be answered by reference to measurable data and to statistical analyses of that data. In becoming more comfortable with statistics, professionals are also better able to assess research they may encounter, and the ideas and presumptions that underpin this research. They will, therefore, be better able to engage with the arguments, assumptions and controversies that emerge. The credibility of new theories and explanations of social issues all rely, to an extent, on statistics.

Statistics matter in the everyday work of service providers, as well. Service evaluations, information about program outcomes and research studies all use ***quantitative*** research, or research that uses measurable information. Practitioners with an understanding of statistics can determine what data is appropriate for their needs and then better explore what is available. They can assess data quality and limitations, and interpret studies by other researchers to inform their practice. They can include statistical analyses in their own research study design or evaluations, and draw justifiable conclusions. Practitioners who can accurately and effectively report on data they obtain are better able to address their practice or program goals.

Statistical data also provide valuable information to guide the development of policies and programs. We can use statistical information to set benchmarks for intervention outcomes. We can monitor trends and patterns in domestic and family violence data, such as changes in rates of reporting to police. Monitoring of outcomes can assist us to target service delivery to identified needs and allocate resources effectively.

Additionally, we can use statistical information to inform evaluations of program or response effectiveness. For instance, we can compare client outcomes for different support interventions to see which ones are more likely to result in successful outcomes, such as increased feelings of safety, or reduced victimisation. This in turn can lead to improved quality of service reviews and evaluations, by better understanding the results of the information gathering exercises, such as outcome monitoring, client feedback surveys, and client demographics.

In this Research and Practice Brief, we list some of the most common statistical measures and terms used in quantitative research. Our purpose is to explain what these measures and terms tell us and to outline their uses, as well as some of the common pitfalls. Unless otherwise indicated, the examples given are fictional and do not derive from real data. We do not provide many formulas but the **Helpful Resources** will direct readers to a number of excellent texts and references that do provide that information.

How to use this guide

You can use this guide in a number of ways:

- If you are looking for a general overview of statistical analyses, you will benefit from reading the entire paper, in conjunction with the **Helpful Resources**.
- If you are seeking an explanation about a specific term or measure, you can look up the relevant section. Some of the topics covered include interpreting tables, samples, variables, measures of central tendency, significance and regression analysis.
- If you regularly read reports or research studies that make use of quantitative data, you may want to keep this paper as a reference guide.

WHAT ARE STATISTICAL ANALYSES?

We apply statistical analyses to ***quantitative***, or numerical, data, concerned with measurement. For example, statistical analyses can tell us *how many* clients there were, *how much* participants valued a program, *how often* clients sought help, and *what types or groups* of people accessed a service.

Statistical analyses also look for ***patterns*** and ***relationships*** in data. For example, they tell us if ***rates*** of help seeking are increasing or decreasing over time. They can tell us whether a factor that can be different or can change, known as a ***variable,*** changes, as another variable increases or decreases (for example, whether certain factors, such as the consumption of substances or alcohol, increase the likelihood that a perpetrator will use violence). They can also tell us how ***typical*** a value is within a certain data set (for example, whether most clients attending a program are likely to be within a particular age range or from a certain cultural group).

Unlike ***qualitative*** data, which tends to be presented as text only, statistical data can be presented in a number of different ways – in tables, graphs, charts, figures, maps or as text. Later in this paper we offer detailed advice on interpreting tables **(see p.7).** There are many sources of statistical data in Australia that relate to domestic and family violence. These include statistics collected by the Australian Bureau of Statistics (ABS), Australian Institute of Health and Welfare, Australian Institute of Criminology, Australian Institute of Family Studies and individual state and territory police, courts, hospitals and other services. Most sources will be found in the *Directory of family and domestic Violence statistics* (2011); a comprehensive guide to statistical data sources relating to domestic and family violence across the health, welfare, family and community services, and crime and justice sectors.

STATISTICAL TERMS AND MEASURES

The list below covers some of the key terms in statistical analysis that may be encountered by those seeking to understand domestic and family violence prevalence, patterns, interventions and research. They have been grouped into two sections; the first referring to terms and measures used in the design and planning stage of research or data collection, and the second, arising from the presentation and analysis of data. Understanding these terms, and applying them to evaluation, reporting or critical thinking about theories assists professionals to thoughtfully consider whether an idea or theory is credible (such as the idea or theory of 'parental alienation'), whether a program is valuable (such as a criminal justice system diversion program), and whether an intervention is effective (such as an adolescent violence intervention).

UNDERSTANDING TERMS USED IN RESEARCH PLANNING & DESIGN

Hypothesis

A ***hypothesis*** is a proposed explanation for a research problem that can be tested using quantitative analysis. If a research study is conducting a statistical test, it will typically have at least two hypotheses: a ***research*** hypothesis and at least one either ***null*** or ***alternative*** hypothesis.

A ***hypothesis*** *is a statement that describes a research problem that quantitative analysis can be used to answer*

Research hypothesis or research statement

A ***research statement***, or ***research hypothesis***, should clearly define the question that the research will examine. It describes the overall problem that the research is setting out to examine. It often takes the form of a general, but clear, statement or a broad question that is then narrowed down into a number of specific, well-defined questions that are then tested using statistical analysis. A hypothesis is a preliminary explanation of the relationship between elements being studied.

For example, a research hypothesis might be: 'Advertising of help lines on television increases help-seeking rates among victims'. Put another way, the hypothesis might take the form of a question: 'Does advertising of help lines on television increase help-seeking rates among victims?'

Terminology differs slightly depending on the field of study, or the research context. For example, it is common in psychology to specifically state the *hypothesis*, which is then tested, whereas in social research there may be a stated research *objective* or a range of research questions that the research sets out to explore.

Null and alternative hypothesis

The ***null*** and ***alternative hypotheses*** are used to turn the research hypothesis into a statistical test, as it is easier to disprove a claim than to prove it. They are often used interchangeably with the research hypothesis. They are typically seen in scientific research and reported in statistical results. In the example above, the null hypothesis would be 'Advertising of help lines on television does not increase help-seeking rates among victims'. Conversely, the alternative hypothesis states that there *is* a relationship, and the statistical tests developed would address proving this to be so.

Variables

Variables refer to factors being measured, or counted, in statistical analyses. They may be demographics like gender and age, numbers of clients attending a service, satisfaction levels of participants, or activities such as rates of reporting or help-seeking. The variables in a study can be established by the hypothesis. For example, in the example above, the presence or absence of help line advertising on television, and help-seeking rates, are the variables. There are four types of variables: ***nominal, ordinal, interval and ratio.***

Variables *are the factors being measured*

Nominal, or categorical, variables

Also known as categorical variables, nominal variables have no numerical value but instead are names of categories. For example:

- yes/no
- male/female
- working full-time/ working part-time/ unemployed
- married/de-facto/separated/widowed/single

Ordinal variables

Ordinal variables contain categories that are in *order*; they may be assigned a numerical value and may be ranked from high to low, or most to least. However, they do not have a standard distance between each of the values and, therefore, may be interpreted differently between respondents. The most common examples of ordinal measurements are ***Likert-type*** scales:

1. Very effective	2. Effective	3. Neutral	4. Somewhat ineffective	5. Very ineffective
1. Very happy	2. Somewhat happy	3. Neither happy nor unhappy	4. Somewhat unhappy	5. Very unhappy

However, caution needs to be used with ordinal variables, as participants may define 'very effective' and 'very happy' differently.

Interval variables

As with ordinal variables, ***interval variables*** are indicators of numerical values but the distance between each of the values is the same. Common examples of interval variables are age and height. Test scores are another example.

Ratio-level variables

Ratio-level variables are interval measurements that have an ***absolute zero*** starting point. This means the numbers cannot start below zero, which allows the intervals between variables to be assessed as a fraction, or ratio. For example, the total number of protection order breaches investigated per police command region might range between 0 and 20 a week. These are ratio variables: it is possible to compare regions. If there are 4 breaches investigated per week in region A and 15 breaches investigated per week in region B, then there are 15/4 or 3.75 more breaches investigated in region B.

Relationships between variables

When talking about variables in terms of their relationships to each other, we refer to ***independent*** and ***dependent*** variables.

Independent variables *are the variables that are assumed to cause change*

Independent, or explanatory, variables

Where a research hypothesis states that one variable is associated with another variable, the ***independent variable*** is the one assumed to be causing the influence. For this reason, it is also sometimes referred to as the ***explanatory variable***, in that it explains the change in another variable. An independent variable is changed to find out the effect on the dependent variable. For example, we may be testing the influence of police training (independent variable) on arrest rates.

Dependent variables *are those that are influenced, or changed, by other variables*

Dependent, or response, variables

A ***dependent variable*** is affected by an independent variable. It is the factor that is measured to see whether it is changed when changes are made to the independent variable. Because it responds to the independent variable, it is often called the ***response variable***. In the example given above, the arrest rate is the dependent variable.

Sometimes other factors can influence a dependent variable or interfere with the relationship between variables. These are known as ***confounding factors***, and are sometimes referred to as ***confounding variables***. When we plan research, or when we interpret results, it is important to consider any other factors that may influence the relationship between variables. For example, in the example given above, it may be unclear whether police training is influencing the arrest rate or whether changes in arrests are due to other factors, such as a pro-arrest policy, increased commission of offences or police rostering changes.

Sample

A ***sample*** refers to a subset of a population that will be tested or investigated to identify certain variables and patterns and relationships. For example, an agency might sample 100 people of 1000 participants in its healthy relationships program, to ask about their satisfaction with the program. A ***census*** differs from a sample because it aims to capture everyone or everything under study. The Australian Census of Population and Housing, conducted every five years by the Australian Bureau of Statistics, is the most well known local example. This captures information on population numbers as well as demographics, such as age, gender, ethnicity, religion, employment, income, education and location.

The sample of a research study impacts on the conclusions that can be drawn about any data or results from the study. In particular, in quantitative research, by looking at the sample we can decide whether the results may be ***generalisable*** to the larger population. For example, if the sample is comprised of English speaking, non-Indigenous, middle class women, then any conclusions drawn from the study are limited to English speaking, non-Indigenous, middle class women. They cannot be generalised to the population of Australian women.

There are different ways to sample a group. We discuss some of the main ones below.

Probability sampling techniques

Researchers often want a sample that represents the population that the researcher wants to make a statement about. For example, a statewide program evaluator may wish to investigate outcomes for program participants. In order to do this, the evaluator needs to provide a sample that represents the total population of participants across the state. This allows the results to be generalised to program participants as a whole. A ***probability sample*** is a method of sampling that provides a random selection of participants; each participant in the study has a chance of being selected, the chance is not zero, and the chance can be calculated. Probability samples are often used in large scale studies. As such, service providers are more likely to encounter these in research they are reading about, rather than in their own service evaluations or client feedback studies.

There a several types of probability samples, each with differing uses, depending on the size of the population being studied, the ease of access to information, such as lists of names, and the variations within the population sub groups. Most commonly, professionals working in the domestic and family violence sector will come across ***simple random*** sample and ***stratified*** samples.

Simple random sample

A ***simple random sample*** is chosen by randomly selecting the required number of participants from the sampling frame, or group. Each respondent has an equal chance of being selected. Random allocation of numbering to a population, through a spreadsheet program, can be used to provide simple random sampling. Alternatively, a systematic sample can be used, which takes a respondent at a fixed interval, say, every tenth enquirer to an enquiry service.

A ***probability sample*** *is one in which each participant has a known chance of being selected*

Stratified, or quota, sample

If it is important that Indigenous people are proportionally represented in the sample and Indigenous people make up 25% of the population, then a sample needs to be chosen that comprises 75% non-Indigenous and 25% Indigenous people. We can do this by randomly selecting subjects from within specified sub groups of the population, to make sure that the sample proportionally represents the subgroups of the population. This is known as a ***stratified sample***, or a ***quota sample***. It takes account of one or more characteristics of the population to ensure that these attributes are represented in the sample. Each respondent's chance of being selected is in proportion to their characteristic of interest.

Non-probability sampling techniques

The other main type of sampling used in quantitative analysis is called ***non-probability sampling***. Realistically, non-probability sampling is common in service and program reporting. This means that the results are not always *generalisable* to the population as a whole.

Accidental, haphazard or convenience sampling

Several common methods of non-probability sampling are ***accidental, haphazard or convenience sampling***, in which participants are selected (or self-selected) to be part of the study on the basis of ease of access. This could include interviewing people on the street (in no systematic way). In a clinical setting, it might involve selecting clients involved in a program or practice who are available for interviews in the morning, or sending an email with an online survey link to every member on a mailing list. In a research context, the sample may consist of those who volunteered to take part. Importantly, all of these types of samples provide no evidence that they are representative of the populations under study. So a survey of women visiting a website about protection orders which asks whether they've ever experienced domestic violence does not provide results about the rates of domestic violence amongst women generally.

Because a ***non-probability sample*** *does not involve random selection of participants there is no evidence that participants are representative of the wider population*

Purposive sampling

Purposive sampling requires sampling with a purpose to recruit one or more specific pre-defined groups. For example, we could be researching the experiences of young women who have experienced dating violence, but it is unlikely that the researcher will have access to data (a list of people who have experienced dating violence) from which to take a random sample. Instead, the young women may be recruited through a youth counselling service. The researcher may then verify that the respondent does in fact meet the criteria for being in the sample, in this case, having experienced dating violence. Purposive sampling can be very useful for situations where you need to reach a targeted sample quickly, where the group is 'hidden' or hard to reach through random sampling, and where sampling for proportionality is not the primary concern.

Over-sampling

Over-sampling is a process for including proportionally more people from some groups in a sample than are represented in the whole population. This is done to provide information about groups who tend to be poorly represented in research, in order to ensure that more statistically accurate results are gained for that group. For example, a study of domestic violence and sexuality may over-sample transgender, queer and intersex people to generate more data about this group.

Sample sizes

Typically, the larger the ***sample size***, the more likely it is that the sample reflects the general population. Some statistical tests also require samples of a certain size to be useful and/or accurate.

Sometimes it is not appropriate to do a statistical test, given the size of the data we are working with or type of data we have. For example, a small convenience sample of clients would not be an appropriate data set to perform statistical tests on but a simple random sample of 100 clients using probability sampling would be appropriate for use with some tests.

Error in statistics

There are generally two types of error reported in statistical testing: ***sampling*** and ***non-sampling errors***. These both relate to the ability of the sample being tested to accurately represent the true characteristics of the population, and therefore how valid and accurate any results may be.

Errors *impact on the validity of data*

Sampling error

Sampling error is a term that relates to the difference between the observed characteristics of a sample, versus the actual *true* characteristic in the population. Sources of sampling error reflect sample size and the representativeness of a sample. Small samples, non-probability samples and differences in the composition of the sample compared to the population (***sampling bias***) lead to sampling error. A census has no sampling error because it records every member of the research population. By contrast, differences between the statistics of a sample and those of its corresponding population increase as a sample gets smaller. A probability sample of 500 taken from a population of 10 000 will be more representative and the results more accurate than a sample of 20 taken from the same population.

Non-sampling error

The second type of error reported in statistical testing is ***non-sampling error*** or error that does not relate to sampling. This cannot be measured. Examples include: ***non-response error***, which can occur when respondents skip questions in a survey or interview (a common example is questions on sexual activity or drug use); and ***response error***, which can be caused by ambiguous or misleading survey questions, or questions which respondents do not, or may not wish to reply to truthfully. Examples of this may be questions about perpetration of domestic and family violence asked of a sample of men enquiring about family mediation. It is important that non-sampling error is controlled as much as possible to maximise the reliability of the results. Ensuring that good research methods are used, such as good survey design, helps to minimise non-sampling error. Similarly, pretesting of interview questions ensures that they are not ambiguous or misleading.

UNDERSTANDING DATA AND RESULTS

The results obtained from a study depend upon the research hypothesis (what is being asked), the quality of the methodology, and how well the study is undertaken, such as the *research methods*, the *quality of sampling*, and how accurately data is *compiled*.

Reading tables

Often results are presented in the form of a table. ***Tables*** typically summarise data across one or more categories. For example, tables could provide information about counts of incidents, clients or sessions, or measures such as income. Often they present the information from different categories of variables across the rows and columns of the table, to indicate the distribution of those responses or counts. This is known as ***cross-tabulation***.

Cross tabulations *compare different independent variables to find out the distribution of responses within categories*

Cross-tabulation

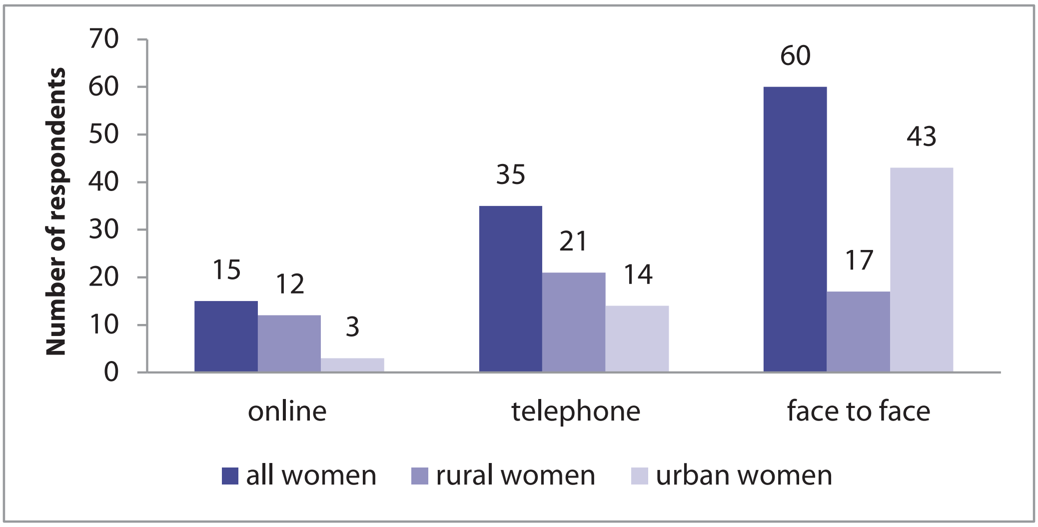
Cross-tabulation is a simple form of analysis which shows how often values occur for each categorical variable under study. For example, if we asked a group of 110 women about how they sought counselling in a survey which also included demographic information, we could cross-tabulate responses about the counselling mode used, with responses about where women live (see Table 1).

Table 1: Cross-tabulation of counselling mode and women's geographical location

	All women	Rural women	Urban women	Total
Online	15	12	3	30
Telephone	35	21	14	70
Face to face	60	17	43	120
TOTAL	**110**	**50**	**60**	**220**

Comparisons of frequency may also be presented in a bar graph (see Figure 1).

Figure 1: Comparison of counselling mode and women's geographical location



Common guidelines for interpreting a table are outlined below.

Interpreting tables: An example from the Personal Safety Survey

Table 2 below provides an example of a typical table produced by the Australian Bureau of Statistics from the Personal Safety Survey (ABS 2006):

- Look at the title to identify the table content. In Table 2, the title indicates that the table relates to respondents' experience of stalking and their relationship to the stalker. Note that the table only includes figures for *respondents* who experienced stalking, not the total population who experienced stalking.
- Identify whether there are subtables present and how they relate to each other. The first subtable relates to a person's experience during the 12 months prior to the survey, whereas the second subtable at the bottom relates to experiences during their lifetime. The larger period of time reported in the bottom table explains the much larger reported numbers.
- Look at the column (vertical) and row (horizontal) labels to identify variables in the table and how they relate to each other. In this instance, the columns indicate the gender of the respondent and the rows indicate the relationship of the perpetrator to the respondent. The column headers indicate that within each gender, the number of respondents in thousands ('000) is on the left and the percentage of respondents (%) is on the right.
- In this table, note how two independent variables are cross tabulated, i.e. victim gender and relationship of victim to stalker.
- Establish whether percentages are given in rows or in columns, and what this tells us. In the example below, the column percentages add up to 100, which indicates that the table is showing a distribution of perpetrator types by gender.
- In the table below, we can easily compare the distribution of perpetrator types across gender. For example, a higher percentage of females than males who were stalked in the last 12 months reported that the perpetrator was a previous partner (14.1% compared to 5.2%). We can also examine the distribution of stalker types within each gender. For example, strangers represented the largest percentage of stalker types reported by females, with 39.6% of females responding they had experienced a stranger stalking episode during the 12 months prior to the survey.
- Read any additional information and footnotes under the table. These usually identify any limitations, add comments or explanations. In Table 2, we notice a number of footnotes that specify definitions and add caveats to the data.
- Determine if any conclusions can be drawn. In this case we might suggest that the likelihood of being stalked by a previous partner is greater for women than men. Looking at the totals, we can conclude that since the age of 15 (the reporting period), almost four times as many women (295 300) as men (75 500) have been stalked by a previous partner.

Table 2: Experience of stalking [a] and relationship to perpetrator in most recent incident [b]

During the last 12 months	Males		Females		Persons	
	'000	%	'000	%	'000	%
Stranger	43.0	38.9	77.4	39.6	120.4	39.3
Boyfriend, girlfriend or date	*9.2	*8.3	16.0	8.2	25.3	8.3
Previous partner	**5.8	**5.2	27.6	14.1	33.4	10.9
Family or friends	48.3	43.7	65.0	33.3	113.4	37.0
Other known persons [c]	*7.6	*6.9	20.3	10.4	27.9	9.1
Total persons who experienced stalking [d]	**110.7**	**100.0**	**195.4**	**100.0**	**306.1**	**100.0**
In their lifetime [e] [f]						
Stranger	205.3	30.1	496.8	33.7	702.1	32.6
Boyfriend, girlfriend or date	84.9	12.4	173.6	11.8	258.5	12.0
Previous partner	75.5	11.1	295.3	20.1	370.8	17.2
Family or friends	266.0	39.0	421.5	28.6	687.4	31.9
Other known persons [c]	69.5	10.2	123.3	8.4	192.8	8.9
Total persons who experienced stalking [d]	**681.7**	**100.0**	**1472.3**	**100.0**	**2153.9**	**100.0**

* estimate has a relative standard error of 25% to 50% and should be used with caution
** estimate has a relative standard error greater than 50% and is considered too unreliable for general use
(a) Stalking involves various activities such as loitering which the respondent believes was intended to harm or frighten.
(b) Includes male and female perpetrators.
(c) Includes acquaintance, neighbour, counsellor, ex-boyfriend or ex-girlfriend.
(d) Components may not add to the total as a person may have experienced stalking by more than one perpetrator.
(e) Includes incidents that occurred more than 20 years ago.
(f) Questions asked about stalking refer to a person's experience over their whole lifetime

Adapted from the Personal Safety Survey (2006, p. 26)

Common errors

Common errors made when interpreting tables include:

- adding together categories that are not the same or which may be overlapping; for example, adding figures for those who experienced violence by a previous partner with those that experienced violence by a current partner – some may be the same people and so would be double counted if the numbers were simply added together.
- confusing incidents, or occasions, with individuals; for example, confusing the number of domestic violence incidents with domestic violence victims – victims may experience many incidents and so the number of actual victims may be far fewer than the number of incidents.

Using tables in your reports

Tables should always be introduced in the text with an explanation of what the table shows, and why its evidence is important. Poorly presented tables can be confusing and easy to misread, so when producing a table, take time with the formatting. See *Snooks & Co 2002*, for a summary of tabulation styles and standards.

In reporting statistics, no value judgements should be made that may cause a reader to misinterpret the findings.

Analysis of variables

When we analyse variables, we can describe them individually, or by looking at the relationships between variables.

Descriptive, or univariate, analysis

We may wish to summarise and discuss a single variable, such as client gender or age. ***Descriptive analysis***, also known as ***univariate analysis***, is a way of summarising and presenting data such as categories, interval data, measurements, counts. There are several ways of presenting data summaries of single variables. These include measures of ***central tendency,*** analysis of ***distribution*** and ***measures of dispersion***.

Central tendency

The ***central tendency*** of the data refers to what is most common or typical in a data set, and includes measures of ***mean, mode*** and ***median***. These are helpful measures which allow us to summarise data. For example, a service may say that it saw a mean, or average, of 12 clients in its older women's program per week. Central tendency measures also provide a common reference point for comparing data. For example, we can compare the average age of clients attending two programs to inform marketing and program content development.

Mean

The ***mean***, or ***average***, is perhaps the best known and most useful of the central tendency measures. It tells us the *typical value* of a variable within a data set, such as the age of clients. We calculate the mean by taking the sum of all the values in a data set and dividing that sum by the total number of values. In the example below, we can calculate the typical age of clients attending a children's program by adding together the ages of each client and then dividing that sum by the total number of clients in the program.

*The **mean** or average tells us the typical value of variables within a data set*

Ages of children attending a program

3, 3, 4, 4, 5, 5, 5, 5, 6, 6, 7, 7 = **60** (sum of all ages)

60 ÷ 12 (sum divided by the number of clients) = **5**

The average age of participants is **5** years.

One point to note is that the mean can be distorted if the values in the data set are skewed, particularly if the data set is small. For example, in the data set below, the age of the last client distorts the mean age of the group.

Ages of young people attending a service

12, 12, 13, 13, 13, 14, 18 = 108 (sum of ages)

108 ÷ 7 (sum divided by the number of clients) = **15**

The mean age of clients is calculated as **15** years, although almost all the clients are under that age. The older client lifted the average age of the service.

While the mean can be calculated for ***interval-level*** or ***ratio-level*** variables, it cannot be used for ***nominal*** or ***ordinal*** variables (e.g. as in the case of rankings or categories such as gender).

Mode

The ***mode*** tells us the *distribution* of values in a data set that occur most frequently. For example, if we had a data set of the number of times women attending a post-separation program experienced domestic violence incidents before reporting to the police, we may be able to tell the most frequent or common number of prior incidents. In the following data set, the mode is 7, because that is the number that occurs most frequently.

*The **mode** tells us which values in a data set occur most frequently*

Numbers of prior incidents experienced by women before reporting domestic violence to the police

3, 3,4,5,5, 6, 6, **7, 7, 7, 7**, 9, 9, 10, 11, 12,15,17

A data set might be referred to as ***bi-modal*** if there are two values that occur equally frequently. In the following data set, referring to attendance at counselling, the two modes are 5 and 14.

Number of counselling sessions attended per client

2, 3, **5, 5, 5, 5, 5,** 6, 7, 8, 9, 9, 10, 12, 12, **14, 14, 14, 14, 14**

We can see in the example above that the mode may not be very helpful in describing the centre of the distribution, since 5 and 14 appear at the low and high end of the distribution.

*The **median** is the middle number of a group of numbers, exactly dividing a data set into two equal halves*

Median

The ***median*** tells us the *middle number* of a group of numbers. That is, it divides the group into two equal halves so we can identify the exact middle point. In the following array, the median is 7 because there are as many values or cases below the number 7 as there are above.

Number of contacts with other services per client per year

1, 1, 2, 4, 4, 5, 6, 6, **7**, 8, 8, 9, 9, 10, 11, 11, 13

If there is an even number of values in an array, as in the example below, the median may be taken as the higher number (7), the lower number (6) or the average of the central numbers (6.5).

Number of counselling sessions per client

1, 1, 2, 3, 3, 4, 5, **6, 7**, 8, 10, 11, 13, 15, 16, 28

We can see in the above example that the last client has had many more counselling sessions than the other clients. This represents an ***outlier*** in the array. The median is much less affected by outliers than is the mean.

Distribution

Distribution is a type of univariate analysis. It refers to how frequently values, numbers or counts of a variable are spread across the range being considered. Distributions can be presented in table format or in graphs.

Where the variables are categories, such as gender, the information is often presented as proportions, or percentages, calculated from the number of counts in each category.

For example, Table 3, below shows the distribution of clients accessing a program over a year, by gender.

Table 3: Program x Client Data, 2012

	Number	Percentage
Women	98	78.4
Men	27	21.6
Total	125	100.0

Where the information being presented refers to intervals, such as measurements, we may wish to show how variables are spread across a data set. The spread, or distribution, of data can be plotted on a graph. Many statistical tests refer to a normal ***distribution***, where data are spread in an even way around a central value.

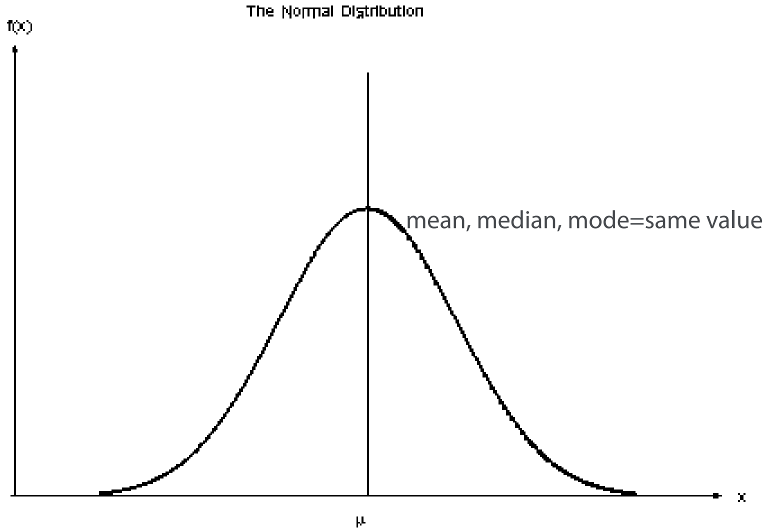
Normal distribution

Often the spread of values will be a ***normal distribution***, which is a distribution of values where the median, mode and mean are all the same value and there is an even spread of values on each side, with 50% below the mean and 50% above the mean.

Many interval and ratio level variables (for example, test scores or ages within a population) have a normal distribution, with most values clustering around the mean, few outliers and increasingly lower frequencies as values move further away from the mean. The normal distribution, when graphed, forms a symmetrical, bell shaped curve (as in Figure 2).

*A **normal distribution** is a representation of values for a variable that form a symmetrical bell shaped curve*

Figure 2: Normal distribution



There are many other distributions that data can take, however the normal distribution is the most important because many statistical tests assume the data are *normally distributed*. Data suitable for a test that assumes normality in distribution include normalised test scores, age, height and weight. Data that exhibit extreme differences from normality will not be suited for many of the most common statistical tests.

Measures of dispersion

While the central tendency variables are useful, they do not tell us whether the values cluster around a particular point or how widely they vary from the average or median. We can measure the spread or dispersion of data by examining the ***standard deviation, range*** and ***interquartile range***.

*The **standard deviation** tells us how closely values in a data set cluster around the mean or average*

Standard deviation

We can further analyse how typical a value is (that is, the degree to which other values or results are clustered around the mean) by using the ***standard deviation***. The standard deviation measures dispersion of values relative to the mean (thus, it is more useful where the mean is a good measure of central tendency). As a measure, standard deviation shows how far from the mean a value is. A standard deviation is also represented as a sigma value (δ). It is often represented graphically, as in Figure 3.

Figure 3: Standard deviation

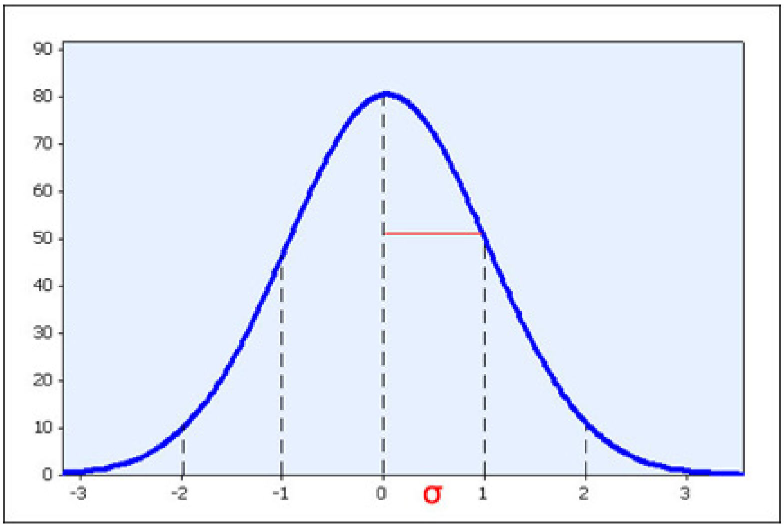


Image: Usable Stats http://www.usablestats.com/images/sd_main.jpg

The standard deviation is calculated by:

1. subtracting the mean from each data point and squaring the result of each
2. finding the average of the squares (that is, adding the squares and dividing by the total number of values) to give the variance
3. taking the square root of the variance to give the standard deviation

For example, for the following range of ages (6, 6, 7, 7, 8, 8) where **7** is the mean:

$(6 - 7)^2 = 1$

$(6 - 7)^2 = 1$

$(7 - 7)^2 = 0$

$(7 - 7)^2 = 0$

$(8 - 7)^2 = 1$

$(8 - 7)^2 = 1$

$\frac{(1 + 1 + 0 + 0 + 1 + 1)}{7}$ = **0.57** (variance)

We take the square root of the variance to give the standard deviation:

$\sqrt{0.57}$ (variance) = **0.76** (standard deviation)

The standard deviation can be very useful for comparing measurements in variation of the same variable, such as satisfaction rates of two groups of clients. We might find that while Group A and Group B both had an average satisfaction score of 70, Group A actually showed much greater variation in satisfaction than Group B, with some clients scoring very low and some very high.

The standard deviation is usually reported as the statistic either plus or minus the standard deviation. For example, research might determine that the average age of perpetrators is 19 +/- 2 years. In tables it is usually reported in a separate column or in brackets following the statistic.

Range

The ***range*** refers to the difference between the minimum and maximum values within a data set, and is calculated simply by subtracting the smallest observed value from the largest observed value.

The ***range*** *represents the spread of data, calculated using the highest and lowest values*

In the example below, we can calculate the range of ages of clients attending a perpetrator intervention program.

Ages of clients attending a perpetrator intervention program

19, 20, 24, 28, 30, 31, 35, 36, 37, 38, 42, **46.**

46-19 =27. The ***range*** is **27.**

Importantly, the range is vulnerable to the influence of outliers. For example, if the last value in the data set above had been 66 instead of 46, the range would be 47 (66 – 19 = 47). This is a much greater range than the one given above, even though only one value has changed in the data set. A range of 47 would imply a greater diversity of ages for attendees of the program than actually exists.

Quartiles and the interquartile range

A useful method of describing the spread of data is by computing ***quartiles*** and the ***interquartile range*** (IQR). The data is first divided into four quartiles which represent four equal groups of observations. The first quartile (Q1) represents 25% of the observations. The second quartile is the median which represents the halfway point of the data – 50% of the data lie on either side of this value. The third quartile (Q3) represents 75% of the observations.

The IQR is calculated by taking the difference between the third quartile and the first quartile. Because the third and first quartiles represent 75% and 25% of the observations respectively, the resulting statistic indicates the range in which half of the observations lie. A larger IQR indicates that the data is widely spread.

The following example shows the calculation of the IQR:

Age of people calling a counselling service:

12	13	14	14	20	22	26	27	31	31
		Q1 = 14		Q2 (Median) = 21			Q3 = 27		

The ***quartile*** *divides data into four equal groups*

The median is the middle point of the data. Here it is 21 (the halfway point between 20 and 22). There are five observations on the left hand side of the median and the middle point of those is 14 which represents Q1. The third quartile (Q3) is the middle point between the median and the largest observation.

We would then compute the IQR as being 27-14 = 13. We would report this by stating that half of the participants were aged between 14 and 27.

In addition to quartiles, researchers often divide data into ***deciles***, which represent 10 even groups of observations.

Relationships between variables

Often we would like to examine whether there are relationships between variables, such as whether one variable coexists with, or effects, another. For example, we may wish to consider whether social marketing campaigns affect community attitudes which condone violence against women. In order to investigate these questions, we must use ***bivariate*** or ***multivariate*** analyses. Various statistical tests are of value in these analyses.

Bivariate and multivariate analysis

***Bivariate and multivariate analysis** examines the relationship between two or more variables*

Unlike univariate analysis, ***bivariate*** analysis is more explanatory in nature. It examines the relationships between two or more variables; that is, the different ways in which they may be related to each other.

For example, a bivariate analysis might examine the relationship between men's unemployment and their perpetration of violence to see if the former is correlated with the latter and by how much.

Multivariate analysis examines relationships between three or more variables simultaneously.

For example, multivariate analysis could be used to identify relationships between parenting time, domestic violence exposure concerns and children's wellbeing.

Correlation and causation

In examining relationships between variables, as in bivariate and multivariate analysis, we are looking to see if one variable influences another. For example, do one or more variables increase, as another decreases? Or is the reverse true? In the first instance, we may try to determine whether the variables are correlated and then try to determine whether one causes the other.

Correlation

*A **correlation** is the measure of a relationship between two variables*

A ***correlation*** is the measure of a relationship between two variables. It means that the two variables are linked, either positively (as one increases, so does another), or negatively (as one increases, the other decreases). The strength of a correlation is often expressed using a ***correlation coefficient***, with a value between -1 and 1. This is expressed using the symbol '*r*'. If *r* is either 1 or -1, then there is a strong correlation (positive or negative) between the variables. If there is no correlation, then *r* is 0. Correlation may also be plotted on a scatter graph. Correlation analyses can determine:

- the strength and direction of the relationship between variables
- how much of the variation in one variable can be attributed to the variation of another variable
- the probability that the relationship between variables is due to a sampling error.

For example, we may look at community anti-violence advertising and reported incidents of violence. Correlation analyses may find a positive correlation between these two variables; that is, as one variable increases, the other does as well.

Causation

*A **causal** relationship is where the presence or absence of an independent variable determines the presence or absence of a dependent variable*

When it can be established that one variable causes an increase, decrease or other change in another variable, we examine whether there is a ***causal relationship***. A causal relationship is where an independent variable determines the presence or absence of change of a dependent variable. Effectively, when a particular cause is present, we anticipate a certain outcome or effect. For example, living with domestic violence may be found to be a cause of depression in children.

Be careful – correlation does not equal causation!

An association between two variables does not always mean that variable A causes variable B. It may be that B causes A or that there are other factors influencing the dependant variable or both variables.

To illustrate, researchers may observe that the number of calls to police to report domestic violence is greater on average for non-migrants compared with migrant populations. This could be interpreted as less domestic violence experienced by migrant populations or, put another way, a correlation between ethnicity and rates of violence. However, lower rates of reporting may be due to other factors associated with migrant communities, such as language barriers, lack of information about rights to safety, lack of knowledge of services available or legal processes. These may all be ***confounding*** factors.

Similarly, community anti-violence campaigning may be introduced as a means of violence 'prevention', and assumptions made that the campaigns will lead to reduced rates of violence. However, data may demonstrate that the campaigns are associated with an *increase* in reports of violence and demands for services, because victims become more aware of the issue. Whilst the two variables may be correlated, the campaign cannot be viewed as causing increased domestic violence.

In order to show causation, relationships between variables need to be reasoned out using formal theories and, in some cases, statistical tests involving quantitative research designs (such as controlled trials or experiments)

Significance

Findings are often discussed in terms of whether or not they are ***statistically significant***. A test of significance can tell us whether there is an association between two or more variables, and whether that association is likely to be real rather than due to chance.

Statistical significance

A ***statistically significant*** finding means that the relationship between variables is likely to be real and not due to chance. It does not mean that the finding is important, meaningful or necessarily strong.

A statistically ***significant*** *finding means that the relationship between variables is likely to be real and not due to chance*

Importantly, a finding that is significant does not rule out a possibility that it was caused by a design flaw in the research such as a sampling error. It is important to take context and the 'real life' interpretation of the finding into consideration when interpreting statistical significance.

Effect size

In addition to being concerned if a relationship between variables is a real one, we should also be concerned about the size of the effect of one variable on another. For example, if our intervention only reduces domestic violence perpetration by .05%, we may ask, 'Is it worth the money?' Reporting the ***effect*** is a standard way of reporting the magnitude and, therefore, the practical value of the results. Effect size needs to be interpreted in conjunction with the statistical significance to measure the practical significance of the results.

Tests of significance

There are different tests for statistical significance, including two of the most common, the chi-square (χ^2) and the Student's t-test.

Chi-square test of association

Chi-square tests *are used to test associations between categorical variables*

A ***chi-square*** is used to determine whether there is a relationship between two categorical variables. It is used to test variables like gender, job type, program attendance or whether people had a successful outcome. A chi-square examines the frequencies (counts) of each category against the frequency of a distribution observed by chance, to determine whether the relationship between the variables is due to chance or likely to be a real effect in the population.

An example of the use of a chi-square test would be to establish whether there is a relationship between the region in which a woman lives (urban, rural or remote) and whether she has reported the most recent domestic violence episode to police, or whether any apparent relationship is chance.

A ***t-test*** *tells us if there is a significant difference between the means of two groups, or between a group's mean and its population mean*

Student's t-test

A ***t-test*** tells us if there is a significant difference between the means (or averages) of two groups or between a group's mean and its population mean. It is typically used with interval data such as income, age, height and weight.

Examples of three types of t-test are:

- to test whether the average income of women aged 20-40 who have left a violent relationship is different to the overall population income for women aged 20-40
- to test whether there are differences between the average rates of violence experienced by one group of clients versus another group of clients. – this is called an independent means t-test, which means that the two groups are comprised of different clients
- to test whether the rate of violence reported by partners changed, following contact with a men's behaviour change service – this is considered to be a dependent means t-test, which means that the two groups are the same people who are measured twice

A ***p-value*** *is the probability that a relationship between variables is produced by a sampling error*

p-value

Tests of significance usually generate a ***p-value***, which is an indicator of the probability that the test statistic has been produced by chance. The smaller the p-value, the more confident we can be that a relationship between the variables exists.

Generally, anything less than a p-value of 0.05 is considered acceptable, meaning that there is 5% error, or we would expect to be 95% sure that the test statistic is genuine and has not been produced by chance.

Regression

Once we have established that there is a relationship between variables, we can use ***regression analysis*** to make predictions about the value of a dependent variable by looking at one or more independent variables. There are different types of regression analysis, including ***linear regression***.

Simple linear regression *is used to plot a relationship between an independent and dependent variable*

Linear regression

Linear regression is used to plot a relationship between an independent and one or more dependent variables. It is an extension of a ***correlation analysis*** that accounts for the correlation between multiple variables at once. The accuracy of the predictions we can make about a dependent variable depends on the strength of the correlation; if the correlation is weak, we can predict, but not confidently.

A ***simple regression*** is where an outcome variable is predicted from a single independent variable. A multiple regression predicts the outcome variable from a number of independent variables.

For example, we could plot the relationship between the number of training sessions (independent variable) and test results (dependent variable).

A regression works by computing a *line of best fit* between the observations that minimises the distance between each observed point and the line. The vertical distance between the data point and the line is referred to as *residual error*.

The simple regression equation is $Y = a + bX$

The dependent variable is represented by the letter 'Y'. It represents the outcome that we are predicting.

The gradient of the line is called the ***coefficient***, which is represented by the letter 'b'. It indicates the amount of change in the dependent variable estimated to occur for each incremental increase of the dependent variable(s).

The point at which the line meets the y-axis (zero) is called the ***intercept*** which represents the value of Y (the dependent variable) when the x-axis is zero. This is represented by the letter 'a'.

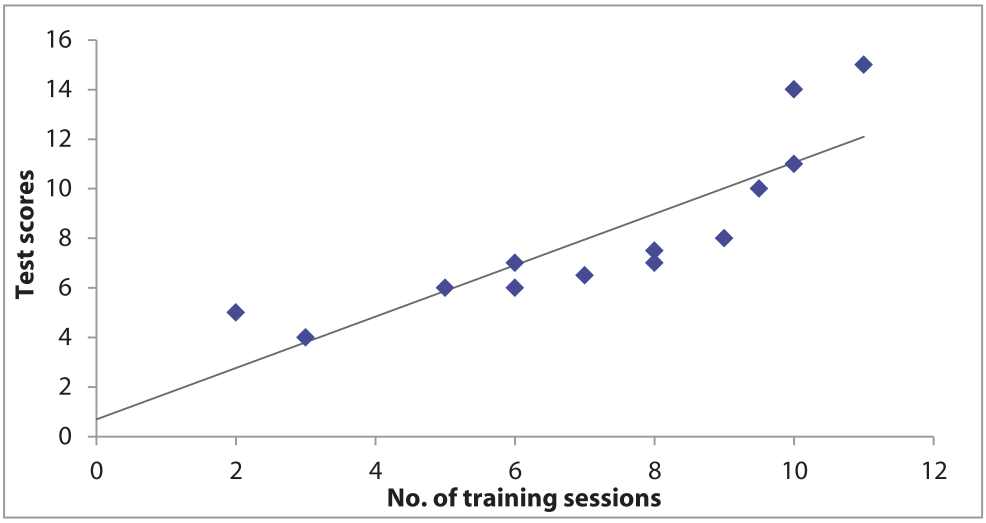
More detail on this equation and how to calculate it can be found in Weinbach and Grinnell (2004) and other statistical texts.

The example below in Figure 4 shows a scattergram of points where the y-axis (the outcome variable) is the test score and the x-axis (dependent variable) is the number of training sessions. The straight line running through the points represents the line of best fit (the coefficient) and the point at which the line meets the y-axis is the intercept.

Expressed as an equation, the results of this regression are Y = 0.7 + 1.0x. The intercept is 0.7 (the average test score when no training sessions were attended). For every additional training session attended, the test score is expected to increase by an average of 1 point.

Another way to examine the results of a regression is to interpret the equation for a value of X. In this case, a person who has attended 10 training sessions would on average be expected to have a test score of Y = 0.7 + (1*10) = 0.7 +10 = 10.7.

Figure 4: Linear regression of training session attended and test results



LIMITATIONS OF STATISTICS

In using and interpreting statistical data, we need to remain mindful that such information can be influenced and limited by a range of factors. An earlier Clearinghouse paper documented limits on statistical data that relate to the social and political context in which violence occurs (Marcus & Braaf 2007). To reiterate some of those points:

- There are typically low reporting and disclosure rates for domestic and family violence, resulting in incidence figures not reflecting the actual number of victims.
- Cultural, language and other barriers for some groups of victims can prevent or inhibit their disclosures of violence and skew statistical profiles; availability of statistical data is especially limited for Indigenous women, women from culturally and linguistically diverse backgrounds, women with disability and women in same-sex relationships.
- Limited access to police, courts and support services in some regions affects a comprehensive collection of statistical data.
- Government or agency policies can impede the collection and publication of accurate statistics.

In addition to these social and political factors, a range of methodological issues can also influence data collection, research participation and response rates:

- The establishing of an appropriate research hypothesis is often a key, but overlooked threshold step. Can the data be found to address the hypothesis? For example, if a criminal justice strategy purports to reduce domestic and family violence, will it be possible to design research to demonstrate that it is effective in this regard? Can we even hypothesise that the strategy does, or doesn't, reduce domestic and family violence?
- The mode of data collection can influence study participation. For example, respondents to an online survey may be younger than people participating in a mail survey.

- Response rates can be influenced by whether participation in the research is compulsory or voluntary.
- Definition of key terms will influence the scope of responses and scope of the sample. For example, how *domestic violence*, *perpetrator* and *victim* are conceived and operationalised may vary between different organisations collecting data, and between jurisdictions. These categories may also be perceived differently by respondents.
- The level of training provided to survey administrators may influence the depth or quality of information gathered.
- The number of survey questions, definitions or sample sizes and groups differ from one study to another. These variations will affect how easily their findings can be compared. For example, comparisons of the Australian Bureau of Statistics' *Personal Safety Survey* over time may be influenced by changes made to questions.
- Finally, as discussed, probability/statistical testing should not be performed on convenience sampling or on very small samples.

In assessing statistical information, the Australian Bureau of Statistics (2010) reminds us to:

- **Be sceptical –** what has really been said and what has been left out?
- **Be aware –** ignoring definitions or comparing statistics inappropriately can lead to misinterpretations.
- **Be critical –** analyse and interpret information and data to draw your own conclusions.

CONCLUSION

Statistics can generate valuable measures of domestic and family violence-related data, and identify patterns and relationships between factors that are invaluable to informing decision making, policy and practice. Statistical data inform our understanding of the scale of issues, help us identify how, or whether, different factors influence each other, and provide an understanding of how situations are changing over time. The use of statistical data, together with qualitative research studies, provides a fuller picture of the nature of domestic and family violence.

Statistical data are also the foundation for reporting the results of interventions, so they are necessary for evaluations of service effectiveness, and to guide planners of strategies aimed at addressing social problems, such as domestic and family violence. Given the plethora of explanations and theories about domestic and family violence perpetration, statistics also help professionals and service providers make sense of literature and evidence in their field of expertise.

In doing this, we must keep in mind that statistical analyses are only as good as the data we put into them; where the data are limited or inappropriate for the purpose, there is not much value to be drawn from the statistics that are then derived.

For those who would like to deepen their knowledge of statistical tools, we refer you to the **Helpful Resources**, listed on the following page.

ACKNOWLEDGEMENTS

The authors are grateful to a number of people who read and provided comments on this paper: Dr Lucy Browning, Research Assistant in the Centre for Gender-Related Violence Studies, University of New South Wales; Dr Matthew McDonald, Tutor in the School of Social Sciences, University of New South Wales; Dr George Matheson, Lecturer in the School of Social Sciences, Media and Communication, University of Wollongong; and Dr Michael Wearing, Senior Lecturer in the School of Social Sciences, University of New South Wales.

REFERENCES

Australian Bureau of Statistics (ABS) 2006, *Personal safety survey*, Australia, reissue, cat. no. 4906.0, Commonwealth of Australia, Canberra

Australian Bureau of Statistics (ABS) 2010, 'Why statistics matter', Commonwealth of Australia, Canberra, viewed 17 April 2013, <http://www.abs.gov.au/websitedbs/a3121120.nsf/home/Why+understanding+statistics+matters>

Australian Bureau of Statistics (ABS) 2011, *Directory of family and domestic violence statistics*, cat. no. 4533.0, Commonwealth of Australia, Canberra

Marcus G & Braaf R 2007, *Domestic and family violence studies, surveys and statistics: pointers to policy and practice*, Stakeholder Paper 1, Australian Domestic and Family Violence Clearinghouse, Sydney

National Statistical Service 2010, *A guide for using statistics for evidence based policy*, cat no.1500.0, Australian Bureau of Statistics, Canberra

Snooks & Co 2002, *Style manual for authors, editors and printers*, 6th edn, John Wiley & Sons, Milton, Qld.

Weinbach R & Grinnell R 2004, *Statistics for social workers*, Pearson Education, Boston

HELPFUL RESOURCES

General statistics

The Australia Bureau of Statistics (ABS) website features a wide range of resources on understanding statistics.

www.abs.gov.au/websitedbs/a3121120.nsf/home/Understanding%20statistics

Statistical sources

The ABS' *Directory of family and domestic violence statistics* (2011) is a definitive guide to statistical data sources relating to domestic and family violence across the health, welfare, family and community services, and crime and justice sectors in Australia.

www.abs.gov.au/ausstats/abs@.nsf/Lookup/4533.0main+features12011

Using statistics in domestic violence research

Publications that discuss the use of statistics in domestic violence research include:

- Ellsberg M & Heise L 2005, *Researching violence against women: a practical guide for researchers and activists*, World Health Organization Program for Appropriate Technology in Health, Geneva

  www.path.org/projects/researching_violence_practical_guide.php

- Marcus G & Braaf R 2007, *Domestic and family violence studies, surveys and statistics: pointers to policy and practice*, Stakeholder Paper 1, Australian Domestic and Family Violence Clearinghouse, Sydney

  www.adfvc.unsw.edu.au/PDF%20files/Stakeholderpaper_1.pdf

- National Sexual Violence Resource Center 2012, *Reading, understanding and evaluating research: glossary of research terms*, National Sexual Violence Resource Center, Enola, PA

  www.nsvrc.org/sites/default/files/Publications_NSVRC_Factsheet_Research-Terms-Glossary.pdf

Did you find this publication useful?
If so, let us know by email, facebook or twitter.

Publication information

ISSN: 1839-8766

Australian Domestic & Family Violence Clearinghouse
The University of New South Wales
Sydney NSW 2052

phone: +61 2 9385 2990
fax: +61 2 9385 2993
freecall: 1800 753 382

email: clearinghouse@unsw.edu.au
web: adfvc.unsw.edu.au
facebook: facebook.com/ADFVC
twitter: @ADFVC_Info

The Australian Domestic and Family Violence Clearinghouse is a project of the Centre for Gender-Related Violence Studies, based in the University of New South Wales, School of Social Sciences. The Clearinghouse is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs.

UNSW AUSTRALIA

Australian Government