

Using survey data

UK Data Service





Author: Jen Buckley and Sarah King-Hele
Updated: August 2015
Version: 1

Acknowledgement/Citation

These pages are based on the following workbook, funded by the Economic and Social Research Council (ESRC).

Williamson, Lee, Mark Brown, Jo Wathan, Vanessa Higgins (2013) Secondary Analysis for Social Scientists; Analysing the fear of crime using the British Crime Survey. Updated version by Sarah King-Hele. Centre for Census and Survey Research

We are happy for our materials to be used and copied but request that users should:

- link to our original materials instead of re-mounting our materials on your website
- cite this as an original source as follows:

Buckley, Jen and Sarah King-Hele (2015). *Using survey data*. UK Data Service, University of Essex and University of Manchester.

Contents

1.	Introduction	4
2.	Before you start	5
2.1.	Research topic and questions	5
2.2.	Survey data and secondary analysis	6
2.3.	Concepts and measurement	7
2.4.	Change over time	9
2.5.	Worksheets	10
3.	Find data	11
3.1.	Survey microdata	11
3.2.	UK Data Service	13
3.3.	Other ways to find data	15
3.4.	Evaluating data	16
3.5.	Tables and reports	18
3.6.	Worksheets	19
4.	Get started with survey data	20
4.1.	Registration and access conditions	20
4.2.	Download	21
4.3.	Statistics packages	22
4.4.	Survey weights	23
4.5.	Worksheets	25
5.	Data analysis	26
5.1.	Types of variables	26
5.2.	Variable distributions	28
5.3.	Missing data	30
5.4.	Analysing more than one variable	31
5.5.	Manipulate data	33
5.6.	Worksheets	35
6.	Report your results	36
6.1.	Describe your data	36
6.2.	Good tables and graphs	37
6.3.	Discussing results	39
6.4.	Limitations and future work	41
6.5.	Citing the data	42
6.6.	Tips for writing a research report	43
7.	References and resources	44



1. Introduction

When starting a research project, you may assume that you need to collect data. However, there are many good quality datasets available for you to use.

Indeed, the UK is particularly rich in survey data. The government regularly uses representative surveys to inform policy and academics and research organisations conduct many large scale social surveys.

Data from large social surveys is routinely archived and made available for others to (re)use in secondary data analysis.

Aims of this guide

This guide aims to help new researchers utilise some of the extensive survey data available. In particular, the guide is designed to support those starting small research projects, especially students doing dissertations.

The guide includes material to read, worksheets for getting started and questions to think about and answer.

An example: Understanding fear of crime

In the guide, key topics are related to real data and research questions using an example research project. The example project examines what determines fear of crime and uses the Crime Survey for England and Wales 2013-2014.

The details of other research projects will vary from the example; however, the main ideas and skills can be easily applied to other research projects.

Structure of the guide

The different sections move through the stages of research, including:

- developing research questions and designs
- finding and accessing relevant survey data
- getting started with data analysis
- reporting research

The supporting worksheets are available to download from the online version of this guide (www.ukdataservice.ac.uk/teaching-resources/using-survey-data).

Many of the worksheets use the fear of crime example and include analysis of the Crime Survey for England and Wales 2013-2014 in SPSS.

Warning: these pages are not comprehensive

As with all research, you need to read, think and develop relevant skills to complete a research project. The guide includes suggestions for further reading; however, researchers need to read literature related to the research topic, which is up to you and your imagination!

2. Before you start

Knowing what you want your research to achieve, will make it easier to do. In this section, you will find information about:

- developing research questions
- surveys and using survey data
- concepts and measurement

The worksheet [Thinking before doing](#) can also help you consider the issues in relation to your own research project.

2.1. Research topic and questions

When starting, you might know the general research area and topics that interest you but still need to develop research questions. Research questions may be very broad to begin with, but they can be refined and developed as you progress.

Read around your topic

In these early stages, reading around your topic is vital. The relevant literature can give you important insights into your topic including information about what is already known, commonly used research methods and the kinds of data available for researching the topic. For example, what do we learn about fear of crime as a research topic from the extract below?

Surveys have repeatedly shown that crime has surpassed unemployment and health as an issue of major public concern (Jacobs and Worcester, 1991). This has led some criminologists to conclude that fear of crime poses almost as large a threat to society as crime itself (Clemente and Kleinmen, 1977). Both official statistics and research show that there is an uneven distribution in criminal victimisation and in fear of crime (Barclay et al., 1995; Mirrlees-Black et al., 1996). Some individuals and some geographical areas are more vulnerable to crime than others and certain groups fear crime more (Pantazis and Gordon 1999:198).

Research questions to research design

When you have a clear idea of your research questions, selecting a suitable research design will be easier. However, you may find initial research questions need further thought as they are not easily researched in practice.

In the next section, we consider the types of research that can be examined using large scale survey data.

Further reading

Bryman, A. (2008). *Social Research Methods*. Oxford: Oxford University Press (Ch. 1 'Social Research Strategies, Ch. 2 'Research designs' and Ch. 3 'Planning a research project and formulating research questions' Ch. 4 'Getting started: reviewing the literature')



2.2. Survey data and secondary analysis

Secondary survey data offers a rich resource for social researchers. Data collections typically come from well-designed surveys of large representative samples and cover a variety of themes and topics.

There are too many survey datasets available to list them all. As a glimpse, here are some key surveys available from the UK Data Service:

- **British Social Attitudes (BSA)** – Over 30 years of data on the attitudes of the British public towards a wide range of social issues
- **Crime Survey for England Wales (CSEW)** – The most widely used survey data for researching crime rates and the effects of crime in the UK
- **Health Survey for England (HSE)** – An annual survey carried out since 1991 with a mixture of questionnaires, physical measurements and blood samples
- **Labour Force Survey (LFS)** – A survey of the employment circumstances of the UK population used to provide official measures of employment and unemployment

You can find more information about available surveys by exploring the [Key Data](#) and [theme](#) sections on our website (ukdataservice.ac.uk). Also, further sections of this guide cover how to search the UK Data Service for suitable data.

How can we use survey data?

Survey data can be used for varied research projects and questions. For example, in relation to the theme 'fear of crime', the *Crime Survey for England and Wales* could be used to examine:

- the prevalence of fear of crime in the population
- how fear of crime varies across groups within the population
- what determines the fear of crime such as 'victimisation'
- how fear of crime affects behaviour

Advantages of using existing social surveys

Secondary analysis of existing survey data has many advantages. Perhaps most of all, survey data from the UK Data Service is often more suitable and better quality than what you could collect yourself. For instance, large social surveys tend to be:

- designed by experts using tested methods of data collection
- provide large representative samples that can be used to describe a population
- repeated, which allows you to look at change over time

Further reading

Bryman, A. (2008). *Social Research Methods*. Oxford: Oxford University Press (*Ch. 13 'Secondary Analysis and official statistics'*)

2.3. Concepts and measurement

When doing social research, you need to define key terms and concepts. For example, what do we mean by 'fear of crime'? This stage will require some careful thinking.

The literature can help show you how others understand and define key terms and concepts. For example, can you identify the issues raised in the following discussion of the term 'fear of crime'?

"The term 'fear of crime' has gained almost universal use. Some believe it is overused and misused (Ferraro and La Grange 1987). Pantazis (2000) points to the confusion that exists in the literature on the precise meaning of fear of crime and how it should be measured. Nevertheless, it is widely accepted as a short-hand term to refer to a whole range of attitudes feelings, reactions and emotions that [...] people have towards crime and victimisation. It can include worry and anxiety (Hough 1995, cited by Pantazis 2000); terror, panic and unease (Bannister and Fyfe 2001) (Burnett 2006, p.127-128)

Concepts and variables

Having a clear idea of key terms and concepts is very important for working out how they can be measured in your research. For the secondary analysis of existing quantitative data, you need to:

- relate concepts to variables in a dataset
- think critically about what variables measure

Think critically about the survey questions

In a survey dataset, variables typically relate to survey questions. You therefore need to assess how well the underlying questions measure your concepts of interest.

- For example, 'How safe do you feel walking alone in this area after dark?' This question is used to measure 'fear of crime' as part of the Crime Survey for England Wales (CSEW). Do you think this survey question will measure fear of crime well? Can you identify any limitations?

The literature can give insights into the strengths and weaknesses of common ways of measuring key concepts. For instance, in relation to the question above, a report from the British Crime Survey (as the CSEW was previously called) notes:

certain characteristics of the question are notable. It presents a more or less specific situation or stimulus – walking alone after dark. Assessments of risk are implicit, since feelings of safety on neighbourhood streets at night presumably reflect beliefs about the likelihood of victimisation. The question asks specifically about night-time safety primarily because other studies have found that few people feel unsafe walking in their neighbourhood during the day (Maxfield 1984, p.4).

The report also identifies limitations of the question as a measure of fear of crime:

it may conceal the true extent of fear among men: whereas women may feel little reluctance to admit to feeling uneasy on the streets at night, men may be wary of expressing similar fears. Secondly, the measure obviously refers only to street crime, which is relatively rare, spatially concentrated, only one of a number of crime threats people fear (Maxfield 1984, p.4)



Choices and compromise

During your project, you may need to make some difficult choices about which measures to use.

Many key social science concepts are complex and difficult to measure. For instance, social class is a much debated and difficult to define concept and unsurprisingly, there are many approaches to measuring social class. For example,

- one approach is to ask respondents directly about their social class (producing a 'subjective measure' of social class)
- another (common) approach is to use information about occupation and employment status to classify individuals (an occupational classification approach)

Standardised measures

Researchers invest lots of time and resources in developing appropriate measures. As a result, you will often find similar and standardised approaches to measuring certain concepts.

For example, in UK surveys the National Statistics Socio-economic Classification (NS-SEC) is common. The NS-SEC uses information about occupation and employment status to create groups such as "Managerial and professional occupations", "Intermediate occupations" and "Routine and manual occupations".

Like all measures, standardised approaches have strengths and weaknesses. However, you should be able to find lots of material examining the validity for standardised measures such as the NS-SEC.

- as another example, the General Health Questionnaire (GHQ-12), which is used in the Health Survey for England, is a common way of measuring mental health

Derived variables

Measures of complex concepts such as NS-SEC and GHQ-12 are typically in the forms of derived variables. For derived variables, raw data collected from survey respondents is changed in some way, most often information from several survey questions is combined.

Combining information from several variables can improve measurement. For example, rather than asking about household income directly, surveys often use multiple questions asking about different forms of income. This information can then be combined to create a (more accurate) measure of overall household income.

Further reading

De Vaus, D. (2013) *Surveys in Social Research London: 6th ed.*, London: Routledge (Ch. 4 'Developing indicators for concepts' and Ch7 'Constructing questionnaires')

Bryman, A. (2008). *Social Research Methods*. Oxford: Oxford University Press (Ch. 6 'The nature of quantitative research')

2.4. Change over time

Many research questions refer to change over time, either at a collective or individual level. For instance,

- has society become more fearful of crime?
- how does growing up in a high crime area affect someone's fear of crime?

Survey data with some element of time can help address these types of question.

Survey data with a time element

Different types of survey data include some form of time element. For instance,

- when cross-sectional surveys, which collect data from a sample in a single point of time, are repeated with new samples we get repeated cross-sectional data. Comparing data from the different samples then allows analysis of trends
- longitudinal studies collect data from the same sample over time. For example, the survey Understanding Society interviews the same households each year to build a picture of how their lives are changing over time
- a further way of bringing in an element of time is to ask respondents about their past. For instance, to measure differences in 'upbringing', surveys might ask a respondent about their childhood

Data in these different forms can be used for different purposes. For instance,

- to address whether society has become more fearful of crime, we can use repeated cross-sectional data to examine the prevalence of the fear of crime at different points in time
- to examine how individual life experiences affect fear of crime, we would ideally use longitudinal data, or alternatively, cross-sectional data with measures of past life experiences

The challenge of analysing change over time

Questions about how societies, groups and individuals change over time are often some of the most interesting.

However, introducing time to your analyses can bring some added challenges. For example, analysis of longitudinal data requires the use of advanced statistical methods.

You therefore need to make sure you have the relevant skills and resources to draw upon.

For further information, see our guide ['Analysing changing over time'](#)



2.5. Worksheets

Thinking before doing

This worksheet can help you get started with your research project. In the different sections, you will think about:

- your research questions
- the population and units of analysis
- the key concepts in your research

Thinking through these key steps will help you identify what data you need for your research project.

The supporting worksheets are available to download from the online version of this guide (www.ukdataservice.ac.uk/teaching-resources/using-survey-data).

3. Find data

With a clear idea of your research questions and concepts, you can start searching for suitable data. In this section, you will find information about:

- how to find data for your research project
- evaluating if a data collection suits your needs

There are also two supporting worksheets. [*Finding the right survey data for your research project*](#) will help you search for data for your own research project and [*Finding and Using published tables and reports*](#) shows how to find and access readily available statistics on key social science topics.

3.1. Survey microdata

Data from surveys is available as microdata, which means the data relates to individual cases. In social surveys, the cases are most often individual people who responded to a survey. However, cases could be other units such as households.

Microdata in SPSS

The screenshot below shows some data from the Crime Survey for England Wales, 2013-2014 in SPSS (a computer software package for analysing quantitative data).

Figure 1 Survey microdata in SPSS (Crime Survey for England Wales, 2013-2014)

	rowlabel	serial	screen	year	samptype	split	subsplit	hselec	adults_s	n
1	135230110	523011	0	13	1	3	5	1	1	
2	135230170	523017	0	13	1	1	1	1	1	
3	135230190	523019	0	13	1	3	5	1	2	
4	135230210	523021	0	13	1	1	2	1	1	
5	135231010	523101	0	13	1	1	1	1	1	
6	135231020	523102	0	13	1	2	3	1	3	
7	135231210	523121	0	13	1	1	2	1	4	
8	135232230	523223	0	13	1	3	6	1	2	
9	135233140	523314	0	13	1	2	4	1	2	
10	135233210	523321	0	13	1	1	2	1	1	
11	135234310	523431	0	13	1	3	6	1	1	
12	135235130	523513	0	13	1	1	2	1	1	
13	135236070	523607	0	13	1	3	6	1	1	
14	135236110	523611	0	13	1	3	5	1	1	
15	135236190	523619	0	13	1	3	5	1	1	
16	135237100	523710	0	13	1	2	3	1	4	
17	135237250	523725	0	13	1	1	1	1	1	
18	135238140	523814	0	13	1	2	4	1	1	

This image illustrates some of the main features of survey microdata:

- each row contains a case (i.e. a person)
- each column contains values for the different variables
- variables are characteristics that vary from case to case



In survey datasets, variables typically relate to survey questions with each respondent's responses stored as values.

For example, in the picture above:

- the shaded cell contains a value for the case in Row 3 on the variable 'adult_s'
- 'Adults_s' indicates the number of adults (aged 16+) in household
- the cell, therefore, indicates that for the case in Row 3 lives in a household with 2 adults

When responses to a question are in the form of categories (such as 'Agree' or 'Disagree'), each response category is assigned a specific value

Why use microdata?

It may not at first seem obvious how we can use microdata to answer our questions. However, using the detailed information about each individual case means we can, for example,

- make tables and graphs to our own specification
- examine the associations between variables using statistical methods
- examine specific sub-groups
- manipulate the data to regroup or combine information

Further reading

Bryman, A. (2008). *Social Research Methods*. Oxford: Oxford University Press. (Ch. 7 Sampling, Ch. 13 'Secondary Analysis and official statistics' and Ch. 14 'Quantitative data').

3.2. UK Data Service

The UK Data Service is the principal source of secondary data in the UK. We provide access to and support in using thousands of archived data collections.

To allow you to find data, data collections all have

- a unique study number
- a catalogue entry

Catalogue entries

Catalogue entries provide detailed information about a data collection along with links to further documentation and access to the data. See, for example, the catalogue page for the [Crime Survey for England and Wales 2013-2014](#).

Figure 2 UK Data Service catalogue entry (Crime Survey for England and Wales 2013-2014)

The screenshot displays the UK Data Service website interface. At the top, there is a navigation bar with links: 'About us', 'Get data', 'Use data', 'Manage data', 'Deposit data', and 'News and Events'. A search bar is located on the right. The main header area includes the 'UK Data Service' logo and the breadcrumb 'Discover > Catalogue'. The left sidebar contains a 'Discover' link and a 'Variable and question bank' section with a 'QualiBank' link. The main content area is titled 'Catalogue' and features a 'SHARE' button. Below this, it states 'UK Data Service data catalogue record for:' followed by the title 'Crime Survey for England and Wales, 2013-2014'. There are links for 'Documentation', 'Related Studies', and 'Publications', along with a 'Download/Order' button and a 'DDI XML' link. The 'TITLE DETAILS' section lists various identifiers and roles: SN: 7619, Title: Crime Survey for England and Wales, 2013-2014, Alternative title: CSEW, Persistent identifier: 10.5255/UKDA-SN-7619-2, Series: Crime Survey for England and Wales [Crime Survey for England and Wales, 1982-], Depositor: Office for National Statistics, Principal investigator(s): Office for National Statistics, Data collector(s): TNS BMRB, and Sponsor(s): Home Office, Ministry of Justice, Office for National Statistics. The 'CITATION' section provides the citation text: 'Office for National Statistics, *Crime Survey for England and Wales, 2013-2014* [computer file], 2nd Edition. Colchester, Essex: UK Data Archive [distributor], May 2015. SN: 7619, <http://dx.doi.org/10.5255/UKDA-SN-7619-2>'. A note below the citation says 'Select the text above to add data citation in your outputs.' The 'SUBJECT CATEGORIES' section lists 'Crime and law enforcement - Law, crime and legal systems'.



Series pages

Surveys that are repeated regularly also have Series pages, which includes summary information about the series as a whole and links to the individual catalogue pages. See for example, the Series page for the [Crime Survey for England and Wales](#).

Figure 3 UK Data Service Series page (Crime Survey for England and Wales)

UK Data Service

Site Search FAQ Help Contact

About us Get data Use data Manage data Deposit data News and Events

Discover > Series

Series

UK Data Service series record for:

Crime Survey for England and Wales

[Abstract](#) | [Access](#) | [Get started](#) | [FAQ](#) | [Related](#) | [Links](#) | [Search](#)

SERIES ABSTRACT

The Crime Survey for England and Wales (CSEW) (formerly the British Crime Survey, provides an important source of information about levels of crime, public attitudes to crime and other related issues. The results play an important role in informing government policy. The CSEW measures the amount of crime in England and Wales by asking people about crimes they have experienced in the last year. This includes crimes not reported to the police, so it is an important alternative to police records.

DATA ACCESS

GN 33174 CRIME SURVEY FOR ENGLAND AND WALES, 1982-	+
GN 33408 CRIME SURVEY FOR ENGLAND AND WALES, 1996-: SPECIAL LICENCE ACCESS	+
GN 33465 CRIME SURVEY FOR ENGLAND AND WALES, 2011-: SECURE ACCESS	+

GETTING STARTED

How can I find out about the datasets – variables, population, sample size etc.?

[Read more...](#)

What format are the data in and where can I download them from?

Ways to find data

Through the [UK Data Service website](#) (ukdataservice.ac.uk), there are several ways to find data.

- [Key data](#) includes details of the most popular data collections, by data type. Survey datasets are listed under UK Surveys, Cross-national Surveys and Longitudinal Studies
- [Theme pages](#) list data collections by themes, which currently including crime, health and housing. The theme pages also include advice, research examples and links to relevant resources
- you can search the '[Discover](#)' search engine using key words and topics. However, since the UK Data Service catalogue contains lots of data, try refining your search (e.g. by date) and sorting results
- there is also [the Variable and Question Bank](#), a tool that allows you to search a range of UK surveys for specific questions and variables

Further help

For help searching for data, you can watch a short [video](#) or see our written guide with illustrated instructions '[Ways to get data from the UK Data Service](#)'. Links can be found on the Get Data section of our website (<http://ukdataservice.ac.uk/get-data>).

The worksheet 'Finding the right data for your research project' can also guide your search.

3.3. Other ways to find data

There are many ways of finding data beyond the UK Data Service. If you are struggling in your search, here are a few useful suggestions:

- data from the UK Census data is available through the [UK Data Service Census](#) support pages
- for data in other countries visit the websites of [CESSDA](#) and [IFDO](#).
- try data collectors. National statistics organisations (such as the [UK Office for National Statistics](#)) and major social survey organisations such as [Natcen](#) give details of data they collect on their websites
- remember your literature search. If a major data source is available it is likely that someone will already have used it. A good knowledge of the literature will help you to identify the data sources that are available



3.4. Evaluating data

Once you have identified potentially useful datasets, you need to evaluate their suitability more closely.

The catalogue entry

You can find most of the necessary information from the catalogue entry. The catalogue entry also gives you access to any further documentation for the survey such as the survey questionnaire and any user guides.

Suggested steps for evaluating a dataset

1) Record the full name and study number so you can easily return to the data.

2) Read the first few sections of the catalogue entry to get a general overview.

- TITLE DETAILS will tell you who commissioned and collected the data
- is the year the data relates to clear in the title?
- is the data collection part of a series?
- the TITLE and SUBJECT CATEGORIES help indicate the general topics covered
- the ABSTRACT provides more detailed information about the data collection

3) Establish the details of data collection.

The COVERAGE, UNIVERSE, METHODOLOGY section includes detailed information about the data and how it was collected.

- 'Dates of field work' - If unclear before, check the data is from the right time period
- Country and Universe – Does the sample relate to the right population (for example, is the whole or the UK or only England, are any significant groups excluded?)
- does the data have the right units of analysis (e.g. individuals or households)?
- Number of units tells you the sample size. There are no set rules about what sample size will be sufficient, a larger sample will generally give more precise results and you need a larger sample when examining small population sub-groups (such as ethnic minorities)
- Sample design – If you are looking for data to make inferences about a wider population, you will typically need data collected through random sampling.
- Time element – Is there any time dimension to the survey? Is this suitable for your needs?
- what data collection methods are used? You should consider how methods of data collection have different strengths and weaknesses and how these relate to your project
- are you looking at survey data? The UK Data Services provides access to varied forms of data including data from smaller scale 'qualitative studies'. Most data collections can be fruitfully used in secondary data analysis, but using different methods. One way to identify if you have survey data is to check under 'Kind of data' for the terms 'Numeric data' and 'Individual (micro) level'

- 4) Examine the survey documentation for more detailed information about the survey questions and variables.

A list of further documents comes under 'Documentation'. The format of documents varies across surveys. Usually, there will be a copy of the questionnaire and a user guide. Sometimes the questionnaire is part of the user guide.

To explore the survey documentation (which usually comes in PDF files), you can:

- scroll through the document looking at questions
- examine the contents page
- search for key words (using 'Control F' to open the search box)

When reviewing survey questions and variables, consider:

- which survey questions relate to your concepts and if there are variables for all concepts
- how the questions are asked and if this might affect the results? For example, are they included in a face to face interview or in a self-completion module? Are all questions asked to everyone?
- the question wording. Is there anything ambiguous about the wording? Are the response categories suitable?

In addition to the questionnaire you may find information about variables and measurement in the user guide, especially details of derived variables.

The documentation will include other useful information such as details of the fieldwork procedures and survey response rates.



3.5. Tables and reports

The agencies and government departments that carry out surveys often publish tables and reports with lots of relevant information and statistics.

For example:

- the Office for National Statistics (ONS) regularly publishes [reports and bulletins on its website](#) with findings from the Crime Survey for England and Wales.
- reports on the Health Survey for England from the Health and Social Care Information Centre, [see this example from 2013](#).
- each year NatCen publishes a report from the British Social Attitudes survey such as the [BSA 31 Report from 2014](#).

Why examine published results rather than just use the data?

This guide focuses on using microdata to do your own analyses of survey data; however, published tables and reports may also be useful for your research project.

For example, published tables and reports can:

- offer useful insights for developing the rationale for your own research
- help you to understand what data surveys provide
- give answers to some of your research questions

How to find published tables and reports

Our website includes a section with links to sources of [statistics](#).

You can find statistics through national statistics websites such as the [ONS](#).

Published reports and tables are often available on the websites of the organisations who conduct and commission surveys. To find out who commissioned and conducted a specific survey look at the survey catalogue page or other documentation.

You can also use a general web search (e.g. through Google).

Your library may also be able to direct you to good sources of statistics.

The worksheet [Finding and Using published tables and reports](#) uses an example of crime statistics to show how to find published tables and reports and to manage the data in Excel.

3.6. Worksheets

Finding the right survey data for your research project

This worksheet will help you search for data for your own research project. Following the worksheet, you will:

- try different ways of finding data through the UK Data Service
- assess the suitability of a dataset

Finding and Using published tables and reports

This worksheet shows you how to find and access readily available statistics on key social science topics. Following the worksheet, you will:

- find links to statistics through the UK Data Service
- explore statistics from the Office for National Statistics (ONS)
- download and manage tables and graphs in Excel

The worksheet uses the theme of crime as an example, with the task of finding data on rates of crime.

The supporting worksheets are available to download from the online version of this guide (www.ukdataservice.ac.uk/teaching-resources/using-survey-data).



4. Get started with survey data

If you register with the UK Data Service, you will be able to download data collections from the catalogue. You can then analyse data using statistics software packages such as SPSS, Stata or R. In this section, you will find information about:

- registering with the UK Data Service
- downloading data
- statistics software

There are also two supporting worksheets: [Register and download data](#) and [Get started with SPSS](#).

4.1. Registration and access conditions

To access most data collections from the UK Data Service, you need to:

- register with the UK Data Service
- agree the conditions of use

Licence agreements are made individually. Therefore, you cannot pass data to someone who has not agreed to the conditions of use. In other words, you must not share data with friends. For a friend to have a dataset, they need to register and download the data themselves.

Why do I need to register and agree to the conditions?

Registration and agreeing to the access conditions are part of your ethical responsibility in secondary analysis.

The conditions of use require you to act responsibly and ethically with the data. In particular, you promise to respect the anonymity of the survey respondents.

Registering the use of data helps the organisation providing data to understand who and how it is used. Information about how data is used helps organisation to ensure data continues to be collected and made available.

How to register

Full instructions of how to register are on our [registration](#) pages.

4.2. Download

Registered users can download datasets direct from the UK Data Service web site www.ukdataservice.ac.uk.

Survey microdata is generally available in SPSS, STATA and tab-delimited formats.

Data can be downloaded from the catalogue page. There will be a link on the top right of the catalogue entry.

Figure 4 Download data through the dataset's catalogue page



To download, simply click the link and follow the instructions.

The Worksheet [Register and download UK Data Service data](#) will guide you through registering and downloading data. The worksheet uses the Crime Survey for England Wales but the same process will apply to all survey datasets.



4.3. Statistics packages

Survey microdata are usually analysed using a statistics software package such as SPSS, Stata or R.

These specialised packages make it easy for you to:

- view data
- recode and compute variables
- weight data to make it representative of the population
- create tables and graphs and do a wide range of statistical analyses
- store commands and results

What statistics packages are there?

SPSS, Stata and R are commonly used statistics packages.

In our worksheets, we use SPSS as many people find it the easiest to use. Other packages look different but the analyses and key issues are the same.

Further reading and resources

For guides to basic analyses and data manipulation in SPSS, Stata and R, see:

- [What is SPSS 20 for Windows?](#)
- [What is Stata?](#)
- [The R guide to UK Data Service key UK surveys](#)

The Worksheet 'Getting started with SPSS' shows you some of the key features of SPSS using the Crime Survey for England and Wales, 2013-2014.

4.4. Survey weights

The terms survey weights, weighting variables and weighted data may be unfamiliar or confusing. You might be tempted to avoid thinking about survey weights. However, you should be aware that your results are likely to be incorrect if you do.

Luckily, with some background information and a little work, using survey weights is relatively straightforward.

What do weights do?

Survey data comes from a sample of a population. Weights help make the sample data representative of the population.

The main way of getting a representative sample is to select cases at random. However, a randomly selected sample might not be representative because:

- some people are less likely than others to agree to participate in surveys (often termed non-response bias)
- common way of selecting a sample can make some cases more or less likely to be selected (referred to as unequal selection probabilities)

Survey producers create weights in response to these problems. The weights adjust a sample to make it more representative of the population it was designed to reflect.

How do weights work?

Weights appear in survey datasets as a variable, which assigns a value to each case to indicate how much 'weight' it should have during data analysis.

A weighting variable can make several adjustments to the data; for example, it can simultaneously adjust for non-response and unequal selection probabilities.

The weights in your dataset

Most large survey datasets include at least one weighting variable. You need to use the weight for your results to reflect the population accurately (unless you take other steps to avoid biased results).

Procedures for applying weights are relatively simple but vary across statistics software packages. Typically, you will specify the weight before you start analyses.

Which weight?

Survey datasets often contain more than one weight. Different weights may relate to

- different samples (such as a 'core' and 'ethnic boost' sample)
- sample units (such as individuals or households).

The details of any weights will be included in the survey documentation. It is important to read the documentation carefully to find out which weight to use for your analysis.



If unsure, you can contact the UK Data Service helpdesk (ukdataservice.ac.uk/help/get-in-touch).

Grossing weights

Weights can also adjust a sample to make it look the same size as the population. These are called grossing weights.

Grossing weights are useful when describing the prevalence of social phenomena in the population (such as incidences of crime). They are used in some official surveys including the Crime Survey for England and Wales and the Labour Force Survey.

When using grossing weights, results from data analysis will look like they come from a sample of millions, rather than a few thousand; thus, results can appear more precise than they might be in reality. As a result, some researchers prefer to rescale the weights to stop the artificial inflating of the sample size.

The 'Data Analysis Workbook' that accompanies this guide includes more on using grossing weights.

Further reading and resources

The following UK Data Service guides focus on weights in more detail

- [An introduction to using weights in social surveys \(Video\)](#)
- [What is weighting?](#)
- The guides to using [Stata](#) and [SPSS](#) also include sections on weighting.

To access any of our guides, see <http://www.ukdataservice.ac.uk/use-data/guides/>.

4.5. Worksheets

The following worksheets will help you get started with using survey data.

For an example research project, the worksheets use the Crime Survey for England and Wales 2013-2014. However, the same process applies to other survey datasets.

[Registering and downloading data from the UK Data Service](#)

This worksheet shows you how to get data from the UK Data Service. Following the worksheet, you will:

- *register with the UK Data Service*
- *download data in SPSS format*

[Getting started with SPSS](#)

This worksheet shows you some of the key features of SPSS. Following the worksheet, you will:

- explore the Data Editor using the Data View and Variable View
- see how to conduct an analysis using the drop-down menus
- see how to view and save your results in the output Viewer

The supporting worksheets are available to download from the online version of this guide (www.ukdataservice.ac.uk/teaching-resources/using-survey-data).



5. Data analysis

When planning and doing data analysis, there are several steps to consider. In this section, you will find information to help you:

- understand and explore key variables
- examine relationships between variables
- manipulate data to aid and improve analysis

There is a supporting [Data Analysis Workbook](#) with examples relating to the fear of crime and using the Crime Survey for England and Wales, 2013-2014.

5.1. Types of variables

Variables measure characteristics that vary among cases. In survey data, examples can include a respondent's sex, age and responses to questions such as how safe they feeling walking alone after dark.

Measurement scales

Variables use measurement scales; for instance, a variable for 'sex' typically categorises individuals into two groups ('male' and 'female') and 'age' measures discrete numerical values across a wide range such as 16, 17, ...96, 97...

Variables are classified into types according to their measurement scale. The distinctions are important because different statistical methods apply to each type.

- for example, when discussing the demographic composition of a group, does it make more sense to discuss the mean in relation to age or sex?

Interval and ratio variables

Ratio and interval variables consist of numerical values that indicate differences in magnitude. They are therefore thought of as 'quantitative' measures. Common examples in survey data include age and income.

The distance between levels ('intervals') is equivalent across the range of values. For example, a 50 year old is ten years older than a 40 year old and a 40 year old is ten years older than a 30 year old.

Ratio and interval variables often have a wide range of possible values. For this reason, they are often termed *continuous*. (However, true 'continuous' variables have an infinite range of values, unlike age where values are a finite list of numbers such as 16, 17, ...96, 97, ...).

If using SPSS, you will also see interval and ratio variables labelled as 'Scale' measures.

There is a difference between ratio and interval variables; in addition to the above characteristics, ratio variables have a fixed zero point. However, the difference is not

very important when dealing with data from social surveys and the terms tend to be used interchangeably.

Nominal variables

Nominal variables differentiate groups sharing some common attribute such as

- sex (male or female)
- ethnic group (such as White, Black, Asian)

These forms of measurement are often called *categorical* variables. Values signify which cases belong to which categories; for example 1=Male and 2=Female. In contrast to quantitative variables:

- categories do not vary in terms of the level or magnitude
- the actual numerical value is not meaningful

Ordinal

Ordinal variables have categories that can be ranked in order. For example, the question 'How safe do you feel walking alone in this area after dark?' has four response options: 1. Very Safe, 2. Fairly Safe, 3. A bit unsafe, and 4. Very unsafe.

Ordinal measures are common in survey data due to the use of closed questions and rating scales. For example, ordinal variables are created by questions that measure:

- approval with options such as 'poor, fair, good, excellent'
- agreement with options such as 'strongly agree, agree, disagree, strongly disagree'
- frequency with options such as 'frequently, occasionally, rarely, never'

Ordinal variables possess both categories like nominal variables and quantitative features (i.e. levels have a greater or smaller magnitude level). As a result, you have to think more about which techniques to use with ordinal variables. Sometimes it makes most sense to use techniques for analysing nominal variables and sometimes techniques for quantitative variables can be useful. There are also specific techniques for analysing nominal variables.

The next section considers some of the ways to describe the distribution of key variables, but first can you identify what type of variables you have?

Further reading and resources

Agresti, A. and B. Finlay (2009) *Statistical Methods for Social Sciences* (fourth edition). New Jersey: Pearson International Edition (Section 2.1 Variables and their measurement p, 12-14).

Bryman, A. (2008). *Social Research Methods*. Oxford: Oxford University Press. (Section on 'Types of Variables, pp.321-322).

Dr Nic from the Blog *Learn and teach Statistics and Operations* has a useful post on the issues with analysing ordinal data [Oh Ordinal data, what do we do with you? 8 July 2013](#)



5.2. Variable distributions

Examining variable distributions is one of the first tasks in data analysis.

A distribution refers to the way cases are distributed across the possible values of a variable; for example, for sex we consider the proportion of males to females.

Variable distributions can be described in different ways. For example, we can use

- tables of counts and percentages
- descriptive statistics such as means
- charts and graphs

The best way to examine a distribution depends upon the type of variable.

Categorical variables

Key ways to explore categorical variables are:

- frequency tables (with percentages)
- graphs such as bar charts or pie charts

Frequency tables list all possible values of a variable with the number of observations as each value and can also include percentages (or proportions). The values and percentages tell you how many people belong in each category.

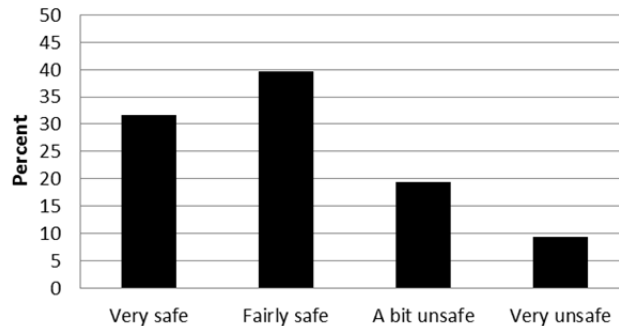
Table 1: How safe do you feel walking along after dark?

	<i>Frequency</i>	<i>Percent</i>
Very safe	2658	31.7
Fairly safe	3320	39.6
A bit unsafe	1632	19.4
Very unsafe	781	9.3
Total (n=)	8390	100

Source: Crime Survey for England and Wales, 2013-2014

Bar charts and pie charts show the same information visually.

Figure 1 How safe do you feel walking along after dark?



Source: Crime Survey for England and Wales, 2013-2014

Quantitative variables

Frequency tables are generally less useful for quantitative variables: the number of values means tables are difficult to absorb, interpret and even fit on a page!

One option is to group values before making a table; for example, we could group age using 10 year intervals (20-29, 30-31 etc).

A histogram shows information in a similar way. Values are grouped into 'bins' with bars to show the distribution of responses across the values.

However, quantitative variables are often best described by summary statistics such

- the minimum and maximum values (e.g. the youngest and oldest respondent)
- the mean and/or median (mean is the average and the median is the central value)
- the standard deviation

Further reading

- Bryman, A. (2008). *Social Research Methods*. Oxford: Oxford University Press (Ch. Quantitative data analysis and Ch.15 Getting started in SPSS)
- De Vaus, D. (2013) *Surveys in Social Research London: 6th ed.*, London: Routledge (Ch. 12 'Overview of analysis' and Ch. 13 'Univariate analysis')



5.3. Missing data

When examining variables you will probably find that data is missing for some cases. Evaluating the nature of missing data is an important part of good survey research.

Reasons for missing data

There are two main forms of missing data. These are where respondents

- were asked a question and either replied 'don't know' or 'refused' to answer
- were not asked a question

'Don't know' or 'refused' missing data

It is common to find a small amount of missing data because respondents refused to answer or responded 'don't know'.

Cases responding don't know or refusing are typically treated as having missing data and excluded from analyses.

If excluded, it is important to consider how those answering a question may differ from those not answering as the results from that question may be slightly biased.

It can also make sense to include 'don't know' as a distinct response category (rather than exclude these cases as missing). But, to do this successfully, there needs to be a fairly high number of don't know responses.

A high number of cases responding 'don't know' or refusing could indicate that there is a problem with the survey question. You may want to review the question wording and, if problematic, consider if there any alternative variables.

'Not applicable' missing data

In surveys, some participants may not be asked certain questions for two main reasons:

1. Sometimes it makes sense not to ask all respondents a particular question. For example, it does not make sense to ask those who unemployed about their current job. As a result, questionnaires are often structured so that if a certain answer is given, subsequent questions are skipped.
2. Surveys can include modules of questions asked to only a sub-sample of respondents.

For example, some surveys divide a sample into sub-samples and use slightly different versions of the questionnaire for each group. Varying the questions across groups extends the range of topics that can be covered without lengthening the questionnaire.

If only a small sub- sample get asked a question, the number of missing cases will be high. The missing cases will however be random, which means bias is not a concern. The main problem with this form of missing data is that the smaller sample can be a limit on further analysis.

5.4. Analysing more than one variable

The aim of data analysis is generally to examine relationships between variables. Analyses of more than one variable can be described as

- 'bivariate' when looking at two variables
- 'multivariate' when looking at two or more variables

With 'bivariate' and 'multivariate' analyses, typically:

- one variable is considered to be an outcome variable (also called the dependent or response variable)
- and the other variables are explanatory variables (also called independent variables)

Bivariate and multivariate analyses can be done using a variety of statistical techniques; the choice of technique depends upon the measurement of the different variables.

Techniques for examining two variables

When examining the relationship between two variables, you may use:

- crosstabulations (also called a two-way frequency or contingency table) to look at the relationship between two categorical variables
- correlations and scatterplots to examine two quantitative variables
- comparisons of means to examine how a quantitative variable varies across the levels of a categorical variable

Statistical tests

Statistical tests are used to establish if relationships found in the sample data can be considered to apply in the population (i.e. are 'statistically significant'). There are many different tests that can be used, and the most appropriate depends on the types of variable and aims of the analysis. Most statistics textbooks will include explanations for the different tests.

More complex analyses

When analysing survey data, we often want to examine the relationships between three or more variables. For example, sometimes we might find a relationship between two variables but suspect the relationship observed is due to a 'third' variable.

Techniques for analysing relationships between more than two variables include

- adding a further layer to a crosstabulation (a three-way table)
- regression analysis (including multiple regression or logistic regression)
- factor analysis



Learning statistical methods

The 'Data Analysis workbook' accompanying this guide includes analysis of using crosstabulations (along with a chi-square test) to examine relationships between variables.

However, to learn about statistical methods, you will need to consult a good book, web-resource and ideally course materials. See below for some suggestions.

Further reading and resources

Introductory chapters on statistics and quantitative data analysis:

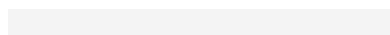
Bryman, A. (2008). *Social Research Methods*. Oxford: Oxford University Press (Ch. 14 'Quantitative data analysis – section on 'Bivariate analysis' and Ch. 15 'Using SPSS')

De Vaus, D. (2013) *Surveys in Social Research London: 6th ed.*, London: Routledge (Ch. 14-17)

More comprehensive textbooks introducing statistics and data analysis

Agresti, A. and B. Finlay (2009) *Statistical Methods for Social Sciences* (fourth edition). New Jersey: Pearson International Edition

Field, A. (2013) *Discovering Statistics using IBM SPSS Statistics: Fourth Edition*. London: Sage



5.5. Manipulate data

You can manipulate survey data to aid and enhance your analyses. Data manipulation can be very beneficial but needs to be done carefully.

Here are some ways you can manipulate data.

Filtering

Filtering is a way of temporarily selecting cases for analysis. For example, you might want to only examine women or people of working age. A filter would identify the cases of interest and can be used and removed as required.

Recoding variables

Variables can be recoded in many ways.

A common use of recoding is to group levels of a categorical variable. For example, you could recode an ethnic group variable with categories 'white', 'black', 'Asian', 'other' to create a simpler variable with categories 'white' and 'other'.

Another form of recording is to group scale measures such to create categorical variables. For example, recoding an age variable into a variable with age groups ('18-34 years', '35-64 years' and '65+').

Recoding variables is useful when:

- there are many response categories and results are difficult to present and interpret
- the sample is spread too thinly over many cells
- you want to use a crosstabulation to examine how an interval variable (such as age) relates to a categorical variable

Important! Be careful about deleting and overwriting original variables. When recoding a variable it is sensible to create a new variable in case you make a mistake or change your mind.

Computing variables

Variables can also be computed using information from one or more variables. For example, you could compute a variable

- 'average number of units of alcohol per day' by dividing values of a variable 'number of units of alcohol per week' by 7
- 'total weekly pay' by adding values for 'weekly pay excluding overtime' and 'weekly overtime pay'

Further reading and resources

The 'Data Analysis workbook' accompanying this guide includes examples of how to recode and compute variables in SPSS.

The following UK Data Service guides also contain information about data manipulation:



- [What is SPSS 20 for Windows?](#)
- [What is Stata?](#)
- [The R guide to UK Data Service key UK surveys](#)

As you progress with survey data, you may need to use other forms of data manipulation. For example, it is sometimes necessary to merge two datasets or create household level variables from individual data. See, the above guides for more information.

5.6. Worksheets

Data Analysis Workbook

Using the theme of 'fear of crime' and the Crime Survey for England Wales 2013-2014, this workbook shows you how to use SPSS to explore data and conduct some analyses around fear of crime. Following the workbook, you will:

- use weighting variables
- examine missing values
- explore categorical variables using frequency tables and charts
- produce descriptive statistics for a continuous variable
- analyse the relationship between categorical variables using crosstabulations
- recode and compute variables

The supporting worksheets are available to download from the online version of this guide (www.ukdataservice.ac.uk/teaching-resources/using-survey-data).



6. Report your results

Communicating your results well is crucial to your research project. In this section, you will find advice on writing about your survey research, including

- describing and citing data
- using tables and graphs
- discussing results and their limitations

6.1. Describe your data

Before presenting analyses, you need to tell your reader about the data. Your reader needs to know the full name of the dataset and year the data is from.

You should also include details of the:

- population (e.g. England, Scotland, Wales or the UK)
- sampling methods
- sample size
- response rate for the survey
- survey mode (telephone interview, online survey, face-to-face survey)

Why this dataset?

Explain the reasons for using this particular dataset. Thinking back to your earlier decisions, why is this dataset the most suitable for your research? Does it:

- contain the best selection of variables
- best relate to the geography or population of interest
- provide the most recent data or data from the right time period

Equally, indicate any limitations of the data; for example, is the data now fairly old? Are any groups of interest excluded from the sample? Is it missing a useful measure?

Discuss the key variables

You need to give information about your key variables, including:

- question wording and possible responses
- why they were chosen (with any relevant references to other studies and relevant theory)
- any strengths and weakness (e.g. how might the question wording affect responses?)
- the distribution of responses using a frequency table or summary statistics
- information about missing data
- details of any recoding.

6.2. Good tables and graphs

You will need to present data using tables, charts and graphs.

It can be difficult to choose what information to include and how to present it. When deciding, remember that:

- you do not need to include everything
- focus on the results that answer your research question

Tips for presenting results

When presenting results, remember to

- number each output (e.g. Figure 1, Figure 2, Table 1, Table 2 etc)
- give each output an informative title (and include the key topic, population, year)
- include in the title, subtitle or as a note below relevant details of the data (such as source, year(s), sample size, weighting variable used and (if appropriate) question wording)
- make sure all text and numbers are large enough to read
- avoid including unnecessary information or precision such as lots of decimal places

Table, graph or chart?

There are no firm rules about whether to use a table or graph. However, in general:

- tables tend to be best for showing detailed figures
- charts and graphs can be good to show overall trends and patterns

It is, however, important to choose appropriate tables, graphs or charts for the type of variables you are using (consult a statistics text book if unsure).

Tables

Make sure there is not too much information (e.g. too many rows or columns) or precision (e.g. decimal places).

Ideally a table should fit on one page.

Tables of percentages can help show the relationship between categorical variables. However, make sure you include the correct percentages:

- for comparing groups, percentages for each group should add up to 100 percent
- as a rule of thumb, put the dependent variable in the row, the explanatory variable in the column and include COLUMN percentages.

Graphs and charts

The pattern in the data should be clearly visible in a chart or graph. To help, try to

- not include too many numbers



- make sure numbers on the axes are readable
- label axes clearly
- be careful with pie charts (with more than a couple of categories, they are difficult to interpret)
- use a clear choice of colours (e.g. does it work in black and white print?)

Re-format your output

You can edit output from statistics packages to improve its appearance.

This may be done within the statistics package or using software such as Microsoft Excel. For example, by copying and pasting a table into Microsoft Excel, you can use Excel to create charts and graphs. (Tip: you can find excellent online tutorials on using Excel.

6.3. Discussing results

You cannot just include tables and figures and expect reader to make sense of it. You need to provide the interpretation.

The content of the discussion will mostly depend upon your research question and results; however, the following points may help to make your discussion clear.

Keep your research questions in focus

You do not need to report every result. Include relevant background information but focus mainly on the results that answer your research question(s) and explain how the results relate to your research questions.

Think about the structure

A logical structure may be to start with the descriptive analysis (such as frequency tables of your key variables) and move on to more advanced results (such as those from multivariate analyses).

Use the structure of your writing to help link the results to your research questions, for example, consider including section titles to make the topic of each section clear.

Refer to figures in the text

A good way to highlight the key findings is to refer to specific results (e.g. "Table 1 shows nearly 60 percent of men agree with the statement compared to 51 percent of women"). But when referring to results,

- use the correct table and figure number in the text (e.g. write "Table 1 shows that" instead of "the table shows that")
- include just the key figures, rather than listing all values
- use equivalent values if comparing groups
- be careful, especially, for attitudinal scales not to focus only on extreme categories such as 'agree strongly'
- ignoring results that do not support your hypotheses

Be clear about how you got your results

Make sure it is clear in your discussion how you got the results. Say if,

- the data is weighted, you recoded any variables, there are missing values (and why)
- provide clear explanations of any procedures used; for example, if referring to p values and statistical significance, explain the test and the meaning of the results

Be precise in your language

Be precise in your use of language. For example, saying you have used data "between 2005 and 2010" implies you used several datasets from the period but it is not clear how many and from when. It might be better to say "data from each year between 2005 and 2010" or "data from 2005 and 2010".

Make sure you use any technical terms correctly.



Be concise

Keeping the discussion concise can help keep the focus on the key findings. Too much detail can make the bigger picture hard to see. For example, you only need to provide a detailed explanation of tests and procedures the first time they are used.

Tables or graphs that interfere with the flow of your report, but still provide useful background information, could be put in an appendix (however, use this sparingly).

Cautious interpretation

All research has limitations which can affect the insights we get from our results. Noting limitations is therefore an important part of the discussion. For example, if the sample is too small for meaningful analysis of a sub-group of interest (such as young adults), it is good to note this limitation when reporting results that are likely to be imprecise.

Association and causation

Discussions of the relationships between variables are another important place for caution.

For instance, you may, hypothesise that “Y causes X”. However, causation is very difficult to prove (especially using data from a single point in time), therefore, when discussing the results, it might be better to use phrases such as “Y is related to X” or that “X increases as Y increases”.

Further reading

Aldridge, A. and K. Levine (2001). *Surveying the Social World: Principles and Practice in survey research*. Open University Press. (Ch. 9 ‘Presenting your findings’)

Bryman, A. (2008). *Social Research Methods*. Oxford: Oxford University Press. (Ch 27 ‘Writing up Social Research’)

6.4. Limitations and future work

When finishing a research project, you often conclude that you could have done it better or that further research is needed. These kinds of reflections are an important part of the research process and should be discussed when writing about the project.

Identifying problems and limitations

Problems and limitations can take many forms. For example, it may be that

- there was a disadvantage with the data such as being out of date, an issue with the question wording or the sample was too small
- you could have done the analysis in a different way such as including another variable or using a more sophisticated statistical technique
- analysis of survey data can only partly address the question and other methods (such as in-depth interviews) might be needed

Discussing problems and limitation

Comments on the problems and limitations can be included in your introduction to the data, discussion of specific results or concluding discussion. Wherever the discussion fits best, you should aim to

- acknowledge the problems and limitations
- indicate any ways you have tried to mitigate problems; for instance, did you recode a variable because the sample was spread too thinly over the categories
- explain how they may shape interpretation of the results (for example, do we need to be cautious in our interpretation or does this data only allow a partial answer)
- consider how further research could help

Highlighting limitations shows critical awareness

Highlighting the limitations will not undermine your project; instead, it can show that you have thought about the research critically and understood the research process.

These discussions can be especially important for a dissertation where you need to show critical thinking.



6.5. Citing the data

It is important to cite or reference the data you use in your research project report: (as well anything else taken from another source).

Good citation and referencing of data

- makes it clear how data was created (and where your results come from)
- gives credit to the data producers or authors of the tables, graphs or research you have used
- helps data producers evidence the use and value of large scale survey data

How do I cite data?

You can find the citation for each of the datasets on its catalogue page in the CITATION section. For example, the citation for the Crime Survey for England and Wales 2013-2014 is:

CITATION 

The citation for this study is:

Office for National Statistics, *Crime Survey for England and Wales, 2013-2014* [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor], May 2015. SN: 7619, <http://dx.doi.org/10.5255/UKDA-SN-7619-2>

See the [data citation page](#) for further information.

6.6. Tips for writing a research report

Before you start

View research project reports for examples of structure and language. If for a university course, can you see reports submitted in previous years?

Be clear about word count, structure and deadline

Structure

All research project reports whether quantitative or qualitative, should include:

- the aim(s) of the report
- motivation for the research project – previous research and how the report adds to existing knowledge
- where data and information has come from
- what was done and why
- results
- conclusions and further research

When to start writing

Start writing early. You do not need to finish your analyses before starting to write. You can edit early drafts later on. Write as you go while you still remember the details

When you have finished

Ask friends and family to read your report (or parts of it). They should not need to be experts in your topic to understand the key points or to spot mistakes.

Put the final draft of your report aside for a day or two and then re-read it before submission.



7. References and resources

References in text

Burnett, A. (2006). Reassuring older people in relation to fear of crime. In Wahidin, A. & Cain, M, eds. *Ageing, Crime and Society*. Tavistock: Willan Publishing,

Maxfield, M.G. (1984). *Fear of Crime in England and Wales*. Home Office Research Study No.78. London: HMSO

Pantazis, C. and Gordon, D. (1999). "Are Crime and Fear of Crime More Likely to be Experienced by the 'Poor'?" in Dorling, D. and Simpson, S. (eds) *Statistics in Society: the Arithmetic of Politics*. London: Arnold, pp. 198-211.

Other references/resources

Textbooks

Agresti, A. and B. Finlay (2009) *Statistical Methods for Social Sciences* (fourth edition). New Jersey: Pearson International Edition

Aldridge, A. and K. Levine (2001). *Surveying the Social World: Principles and Practice in survey research*. Open University Press.

Bryman, A. (2008). *Social Research Methods*, 3rd edition. Oxford: Oxford University Press.

Blaxter, L. C. Hughes and M. Tight, (2006) *How to Research*, 3rd edition, Buckingham: Open University Press.

David, M. and C. Sutton (2004) *Social Research: The Basics*, London: Sage

De Vaus, D. (2013) *Surveys in Social Research London: 6th ed.*, London: Routledge

Field, A. (2013) *Discovering Statistics using IBM SPSS Statistics: Fourth Edition*. London: Sage

Fink, A. (2009) *Conducting Research Literature Reviews*, 3rd edition, London: Sage.

Gilbert, N. ed (2008) *Researching Social Life*, 3rd edition, London: Sage.

Reardon, D. (2006) *Doing Your Undergraduate Project*, London: Sage.

Trochim, W. M. The Research Methods Knowledge Base:
<http://www.socialresearchmethods.net/kb/index.php>

Yates, Simeon J. (2003) *Doing Social Science Research*. London: Sage/Open University Press

Other Guides

- [Analysing changing over time](#)
- [An introduction to using weights in social surveys \(Video\)](#)
- [The R guide to UK Data Service key UK surveys](#)

- [Ways to get data from the UK Data Service](#)
- [What is weighting?](#)
- [What is SPSS 20 for Windows?](#)
- [What is Stata?](#)

To access any of our guides, see <http://www.ukdataservice.ac.uk/use-data/guides/>.

1 August 2015

T +44 (0) 1206 872143
E help@ukdataservice.ac.uk
W ukdataservice.ac.uk

The UK Data Service delivers
quality social and economic
data resources for researchers,
teachers and policymakers.

© Copyright 2015
University of Essex and
University of Manchester

UK Data Service

