UNIT 9

Data Analysis

Overview

Once the data collection phase has been completed, you are now ready to move to the data analysis stage in the research process. During data collection you will have large amounts of data including completed questionnaires, field notes, audio recordings, pictures, articles and much more, depending on the research approach that you adopt for your study. Data analysis begins from the moment you start thinking about how to sort through, organize and clean all that you have collected. Data analysis therefore is much more than statistical tests and measurements. It includes data preparation, cleaning, entering the data in various formats (also depending on whether you conduct quantitative or qualitative research), analysis and presentation of the findings. Unit 9 takes you through this process of data analysis.

Learning Objectives

By the end of this Unit you will be able to:

1. Identify inaccuracies found in the data prior to analysis.
2. Describe the data logging and entry processes.
3. Distinguish between descriptive statistics and inferential statistics.
4. Explain qualitative data analysis.
5. Describe the various types of data presentation (bar graphs, charts etc).

This Unit is divided into four Sessions as follows:

Session 9.1: Data Preparation, Cleaning and Entering

Session 9.2: Data Analysis – Descriptive Statistics

Session 9.3: Data Analysis – Inferential Statistics

Session 9.4: Qualitative Data Analysis

Readings & Resources

Required Readings

Dean, S., & Illowsky, B. (n.d.). Elementary statistics: Video Lecture. Hypothesis testing with two mean. Connexions. Retrieved from http://cnx.org/content/m17577/latest/

Dean, S. and B. Illowsky (2012). Collaborative statistics. Retrieved at: http://cnx.org/content/m16849/latest/?collection=col10522/latest

eSource - Behavioural and Social Science Research. Software and qualitative analysis. Retrieved at: http://www.esourceresearch.org/Default.aspx?TabId=537

Gilbert, N. (2006). *Researching social life*. London: Sage Publications.

Kumar, R. (1996). Research methodology. Melbourne: Longman.

Neuman, W.L. (2000). Results with one variable. In *Social Research Methods: Qualitative and Quantitative Approaches*. Boston: Allyn and Bacon, pp. 317 – 322.

Polkinghorne, D.E. (2005). Language and Meaning: Data collection in qualitative research. *Journal of Counseling Psychology*. Vol 52, No. 2, pp137-145. Robson, C. (2002). Real World Research. Blackwell Publishing.

Trochim, W. (2006). Research methods knowledge base. Web Centre for Social Research Methods. http://www.socialresearchmethods.net/kb/contents.php

Univariate analysis. Retrieved at: www.sagepub.com/upm-data/9913_040279ch03.pdf

You are also advised to locate and read: Additional papers relevant to the topics covered.

Session 9.1

Data Preparation, Cleaning and Entering the Data

Data Preparation

According to Trochim (2006), data preparation involves checking or logging the data in; checking the data for accuracy; entering the data into the computer; transforming the data; and developing and documenting a database structure that integrates the various measures.

Logging Data

In any research project, especially if you employ a mixed methods approach, you may have data coming from a number of different sources at different times:

- mail surveys returns
- coded interview data
- observational data
- audio recordings and transcripts of same

To deal with large amounts of data, you need to set up a procedure for logging the information and keeping track of it until you are ready to do a comprehensive data analysis. Different researchers differ in how they prefer to keep track of incoming data. In most cases, you will want to set up a database that enables you to assess at any time what data is already in and what is still outstanding. It is also critical that you retain the original data records for a reasonable period of time (returned surveys, audio files, field notes and so on). Most professional researchers will retain such records for at least 5-7 years (Trochim, 2006). For important or expensive studies, the original data might be stored in a data archive. The data analyst should always be able to trace a result from data analysis back to the original forms on which the data was collected. A database for logging incoming data is thus a critical component in good research record-keeping.

Checking the Data for Accuracy

As soon as data is received you should screen it for accuracy. In some circumstances, doing this right away will allow you to go back to the sample to clarify any problems or errors. There are several questions you should ask as part of this initial data screening:

- Are the responses legible/readable?
- Are all important questions answered?
- Are the responses complete?
- Is all relevant biographical and contextual information included (e.g., data, time, place, researcher)?

In most social research projects, quality of measurement is a major issue. Assuring that the data collection process does not contribute inaccuracies will help assure the overall quality of subsequent analyses. This is therefore why checking for accuracy and cleaning the data is an important part of the analysis.

Developing a Database Structure

The database structure is the manner in which you intend to store the data for the study so that it can be accessed in subsequent data analyses (Trochim, 2006). You might use the same structure you used for logging in the data or in large, complex studies you might have one structure for logging data and another for storing it. As mentioned above, there are generally two options for storing data on your computer: database programs (example data matrix in Excel spreadsheet) and statistical programs (example SPSS in a quantitative research or NVivo for a qualitative research study). As you move on to your applied research methods courses in years 2 and 3, you will learn more about SPSS and NVivo but for now you should just be aware of their value to data storage and analysis. Usually database programmes are the more complex of the two to learn and operate, but they allow the analyst greater flexibility in manipulating the data.

Entering the Data into the Computer

There are a wide variety of ways to enter the data into the computer for analysis. Probably the easiest is to just type the data in directly. In order to assure a high level of data accuracy, the analyst should use a procedure called double entry. In this procedure you enter the data once. Then, you use a special program that allows you to enter the data a second time and checks each second entry against the first. If there is a discrepancy, the program notifies the user and allows the user to determine the correct entry (Trochim, 2006). This double entry procedure significantly reduces entry errors. However, these double entry programs are not widely available and require some training. An alternative is to enter the data once and set up a procedure for checking the data for accuracy. For instance, you might spot check records on a random basis.

LEARNING ACTIVITY 9.1

A research study has conducted 3 focus group studies with youth on issues of youth and delinquency in their communities. In addition to a short questionnaire administered to each participant to gain biographical data for each, an hour-long discussion took place and was audio recorded for each focus group. If you were the researcher in charge of the overall studies, what steps would you take to organize, clean and prepare the data for analysis?

Session 9.2

Data Analysis: Descriptive Statistics

Descriptive Statistics

Descriptive statistics are used in quantitative data analysis to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures. Together with simple graphics analysis, they form the basis of virtually every quantitative analysis of data (Trochim, 2006). Descriptive statistics may be used for univariate analyses of data, that is, examining one variable at a time. Even if you plan to take your analysis further to explore the linkages, or relationships, between two or more of your variables you initially need to look very carefully at the distribution of each variable on its own.

Data collected using quantitative methods, for example a questionnaire administered for a survey, can be described and presented in many different formats. Suppose you are interested in buying a house in a particular area. You may have no clue about the house prices, so you might ask your real estate agent to give you a sample data set of prices. Looking at all the prices in the sample often is overwhelming. A better way might be to look at the median price and the variation of prices. The median and variation are just two ways that you will learn to describe data. Your agent might also provide you with a graph of the data. This area of statistics is called descriptive statistics (Dean and Illowsky, 2012).

Descriptive statistics are very important because if we simply presented our raw data it would be hard to visualize what the data was showing, especially if there was a lot of it. Descriptive statistics therefore enables us to present the data in a more meaningful way, which allows simpler interpretation of the data. For example, if we had the results of 100 pieces of students' coursework assignments, we may be interested in the overall performance of those students. We would also be interested in the distribution or spread of the marks. **Descriptive statistics** allow us to do this.

A **statistical graph** is a tool that helps you learn about the shape or distribution of a sample (Dean and Illowsky, 2012). The graph can be a more effective way of presenting data than a mass of numbers because we can see where data clusters and where there are only a few data values. Graphs are used in data analysis to show trends and to enable researchers to compare facts and figures quickly. You can also use graph data first to get a picture of the data. Then, more formal tools may be applied. Some of the types of graphs that are used to summarize and organize data

are the dot plot, the bar chart, the histogram, the stem- and-leaf plot, the frequency polygon (a type of broken line graph), pie charts, and the boxplot (see Dean and Illowsky for examples of these: http://cnx.org/content/m16849/ latest/?collection=col10522/latest)

Distribution and Dispersion of Data

In looking at distribution of the data, percentiles and averages can be calculated. As researcher, you may want to carry out these observations of a single variable, example the age distribution of your respondents. A percentile indicates the relative standing of a data value when data are sorted into numerical order, from smallest to largest. p% of data values are less than or equal to the pth percentile (Dean and Illowsky, 2012). For example, 15% of data values are less than or equal to the 15th percentile. Low percentiles always correspond to lower data values. High percentiles always correspond to higher data values.

A percentile may or may not correspond to a value judgment about whether it is good or bad. The interpretation of whether a certain percentile is good or bad depends on the context of the situation to which the data applies. In some situations, a low percentile would be considered good while in other contexts a high percentile might be considered good. In many situations, there is no value judgment that applies. Understanding how to properly interpret percentiles is important not only when describing data, but is also important in later chapters of this textbook when calculating probabilities.

It is important to note:

When writing the interpretation of a percentile in the context of the given data, the sentence should contain the following information:

- Information about the context of the situation being considered,
- The data value (value of the variable) that represents the percentile,
- The percent of individuals or items with data values below the percentile.
- Additionally, you may also choose to state the percent of individuals or items with data values above the percentile.

The centre of a data set is also a way of describing location. The two most widely used measures of the centre of the data are the mean (average) and the median. To calculate the mean weight of 50 people, add the 50 weights together and divide by 50. To find the median weight of the 50 people, order the data and find the number that splits the data into two equal parts. The median is generally a better measure of the centre when there are extreme values or outliers because it is not affected by the precise numerical values of the outliers. The mean is the most common measure of the centre. Another measure of the center is the mode. The mode is the most frequent value. If a data set has two values that occur the same number of times, then the set is bimodal. Statistical software can be used to easily calculate the mean, the median, and the mode for you.

Standard Deviation

An important characteristic of any set of data is the variation in the data. In some data sets, the data values are concentrated closely near the mean; in other data sets, the data values are more widely spread out from the mean. The most common measure of variation, or spread, is the **standard deviation**. The standard deviation is a number that measures how far data values are from their mean. The standard deviation provides a numerical measure of the overall amount of variation in a data set and can be used to determine whether a particular data value is close to or far from the mean. You should know as a rule that, the standard deviation is small when the data are all concentrated close to the mean, exhibiting little variation or spread. The standard deviation is larger when the data values are more spread out from the mean, exhibiting more variation.

Suppose that we are studying waiting times at the checkout line for customers at supermarket A and supermarket B and want to have a closer look at the variable "waiting time". Let us say the average wait time at both markets is 5 minutes. At market A, the standard deviation for the waiting time is 2 minutes; at market B the standard deviation for the waiting time is 4 minutes. Since market B has a higher standard deviation, we know that there is more variation in the waiting times at market B. Overall, wait times at market B are more spread out from the average; wait times at market A are more concentrated near the average. Again, statistical programmes can be used to easily calculate the standard deviation for you.

LEARNING ACTIVITY 9.2

Read more online graphs. Suppose you want to use a line graph to summarize data collected during your research project. What information can be gleaned about the data from the line graph? Discuss in the discussion forum.

Session 9.3

Data Analysis: Inferential Statistics

Inferential Statistics

While descriptive statistics simply summarize the data and describe what is or what the data shows, inferential statistics attempt to reach conclusions that extend beyond the immediate data alone (Trochim, 2006). Thus, while we use descriptive statistics simply to describe what's going on in our data or for one variable, we use inferential statistics to make predictions from our data to more general conditions or to explain relationships among multiple variables. For instance, we use inferential statistics to try to infer from the sample data ideas or conclusions about the overall population under study. Inferential statistics is related to multivariate analysis, in that it looks at the relationship among multiple variables where there are independent and dependent variables.

Often, you do not have access to the entire population you are interested in investigating, but only a limited number of data instead. For example, you might be interested in the exam marks of all students in your country. It is not feasible to measure all exam marks of all students in the country so you have to measure a smaller sample of students (e.g.,

100 students), which are used to represent the larger population of all students. Inferential statistical techniques would thus allow you to use this sample to make generalizations about the population from which the sample was drawn. For example, you might be able to predict future trends in students' marks or draw conclusions about differences in female and male marks. It is, therefore, important that the sample accurately represents the population. Inferential statistics are also based on the assumption that sampling naturally incurs sampling error and thus a sample is not expected to perfectly represent the population. The methods of inferential statistics are: the estimation of parameter(s) and testing of statistical hypotheses.

When you undertake quantitative research, you are automatically testing hypotheses and therefore you should always attempt to test them effectively. To do so, the following steps should be followed:

- Define the **research hypothesis** and set the parameters for the study.
- Set out the **null** and **alternative hypothesis** (or more than one hypothesis; in other words, a number of hypotheses).
- Explain how you are going to **operationalise** (that is, **measure** or **operationally** define) what you are studying and set out the variables to be studied.

- Set the **significance level**.
- Make a **one**- or **two-tailed prediction**.
- Determine whether the distribution that you are studying is **normal** (this has implications for the types of statistical tests that you can run on your data).
- **Select** an appropriate statistical test based on the variables you have defined and whether the distribution is normal or not.
- **Run** the statistical **tests** on your data and interpret the output.
- **Accept** or reject the **null hypothesis**.

Both descriptive and inferential statistics are used in quantitative data analysis. However, because the sample being studied under the qualitative approach is hardly ever representative of the wider population, then generalizations cannot be made to the population. It means therefore that inferential statistics do not apply to qualitative data analysis. You may, however, use some descriptive statistics in qualitative data analysis. For example, if you conduct in depth interviews with a sample of 30 inner city youth, you may use descriptive statistics (pie chart or bar graph) to display basic biographical data such as ages, gender, and social class of these 30 interviewees.

Since inferential statistics makes inferences from the sample to the wider population, it is important that the sample is representative of the population about which it is making predictions. In social research to address the issue of generalization, there are tests of significance. A Chi-square or T-test, for example, can tell us the probability that the results of our analysis on the sample are representative of the population that the sample represents. In other words, these tests of significance tell us the probability that the results of the analysis could have occurred by chance when there is no relationship at all between the variables we studied in the population we studied.

Other examples of inferential statistics include linear regression analyses, logistic regression analyses, ANOVA, correlation analyses, structural equation modeling, and survival analysis, to name a few. At this level of introductory social research you are not expected to carry out these various tests, but you should know the various tests or options available in inferential statistics and what each one achieves in order to plan your research design. The accompanying power point presentation to this unit will provide you more information on these various tests and what they achieve in inferential statistics.

LEARNING ACTIVITY 9.3

View the following videos:

- Dean, S., & Illowsky, B. (n.d.). Elementary statistics: Video Lecture. Hypothesis testing with two mean. Connexions. Retrieved from http://cnx.org/content/m17577/latest/
- Peridisco's Introductory Statistics - Chapter 8: Hypothesis testing available at: http://www.youtube.com/watch?v=HmMjS88eSVE&feature=related

Your tutor will divide your class into two subgroups. Each group will view one video and then in the discussion forum, each group will summarize and explain what they have learned from their video and respond to comments.

Session 9.4

Qualitative Data Analysis

What Is Qualitative Data?

Whereas quantitative data deals with numbers, qualitative data deals with meanings. Meanings are gleaned mainly through language, action and observation of action. Qualitative research methods study behaviours and experiences and so data in these methods are derived from an intensive exploration with a participant or group of participants. Such an exploration results in languaged data. The languaged data are not simply single words but interrelated words combined into sentences and sentences combined into discourses. The interconnections and complex relations of which discourse data are composed make it difficult to transform them into numbers for analysis (Polkinghorne, 2005).

You should note the richness and diversity of qualitative data. It can also encompass sounds, pictures, videos, music, songs, prose, poetry or text (from documents and content analysis). Text is by no means the only, nor is it always the most effective, means of communicating qualitative information. Qualitative data embraces a wide and rich spectrum of cultural and social artefacts. What do these different kinds of data have in common? They all convey meaningful information in a form other than numbers. Not only is that meaning revealed from a reading and examination of the text or languaged data but also during the interactions and interpretations between the researcher and the researched. Therefore, qualitative data analysis tends to be more subjective that quantitative data analysis and there is greater possibility of researcher bias during qualitative data analysis. The quote below aptly describes qualitative analysis:

> *"Qualitative analyses can be evocative, illuminating, masterful—and wrong. The story [the researcher relates], well told as it is, [may] not fit the data. Reasonable colleagues double-checking the case [may] come up with quite different findings. The interpretations of case informants [may] not match those of the researchers." (Miles and Huberman, 1994:247)*[1]

The purpose of data gathering in qualitative research is to provide evidence for the experience it is investigating. The evidence is in the form of accounts people have given of the experience.

---

[1]Source: eSource - Behavioural and Social Science Research. Software and qualitative analysis. Retrieved at: http://www.esourceresearch.org/Default.aspx?TabId=537

The researcher analyzes the evidence to produce a core description of the experience. The data serve as the ground on which the findings are based. In constructing the research report, the researcher draws excerpts from the data to illustrate the findings and to show the reader how the findings were derived from the evidential data (Polkinghorne, 2005). Most often the evidence takes the form of written texts. Written evidence is gathered from documents and data originally generated in oral form (e.g., through interviews) are transformed into written texts through transcription. However, the evidence itself is not the written texts on the paper but the meanings represented in these texts. It is not the printed words themselves that can be analyzed by counting how many times a particular word appears in the text. Rather, the evidence is the ideas and thoughts that have been expressed by the participants. In this sense, the textual evidence is indirect evidence.

Analyzing Qualitative Data

You will recall that quantitative research generates reliable, population based and generalizable data and is well suited to establishing cause-and-effect relationships while qualitative research generates rich and detailed data that contribute to in-depth understanding of the context. It is important to bear this difference in mind too when collecting and analyzing data for your research project. Therefore, while quantitative analysis will employ various statistical techniques for testing hypotheses and assessing relationships between and among variables, qualitative analysis involves a continual interplay between theory and analysis. In analyzing qualitative data, you will seek to discover common themes, patterns such as changes over time or possible causal links between variables.

Robson (2002) identifies four main approaches to qualitative data analysis. He explains them below:

- Quasi-statistical approaches – Use word or phrase frequencies and inter- correlations as key methods of determining the relative importance of terms and concepts – typified by content analysis.
- Template approaches – Key codes are determined either on an a priori basis or from an initial read of the data. These codes then serve as a template for data analysis – typified by matrix analysis
- Editing approaches – More interpretive and flexible than the above, with no (or few) a priori codes – typified by grounded theory approaches.
- Immersion approaches – Least structured and most interpretive, emphasizing researcher insight, intuition and creativity.

For now, it is important to have an elementary idea of the approaches above, but you will learn more about these as you move to more advance research methods. In all of the approaches above, the researcher applies subjective interpretations and exercises judgment in the coding of the data. Computers can also be a big help in qualitative analysis and various qualitative analysis software can be used to make more rigorous (and thus more scientifically sound) analyses that we otherwise could not or would not undertake. For example, Ethnograph and NUD*IST specifically analyze qualitative data.

Staying With Your Data

As stated above, in analysing your data it is important to continually make the link between your analysis and the theory. What this means is that the findings or results that emerge from the data must be discussed within the context of the theory/theories within which the study is framed. Sometimes students in using theory to give meaning to or explain their research findings, go beyond the data. This is a problem because students/ researchers start discussing what they have NOT found; they go beyond the boundaries of the data. As researchers you need to try to avoid this and learn the importance of 'staying with your data'.

Qualitative data are collected as descriptions, anecdotes, opinions, quotes, interpretations, etc., and are generally either not able to be reduced to numbers, or, are considered more valuable or informative if left as narratives. Qualitative data also includes very rich, descriptive, narrative and detailed information and can tell a lot about the issue or phenomenon being research. According to Johnson et al. (2010), qualitative research creates "mountains of words". No matter how large or small the project, the qualitative methodology depends primarily upon eliciting self-reports from subjects or observations made in the field that are transcribed into field notes. Even a small qualitative project easily generates thousands of words. Major ethnographic projects easily generate millions of words. This is a strength of qualitative data but this can also be a weakness in that it can be very difficult to organize, manage and analyse. Proper management of the data is therefore an important aspect of staying with your data and reducing misrepresentation of data or error in reporting.

Some ways in which you can ensure that you stay with your data include:

- Making detailed field notes during data collection. Field notes reflect what you observe during field work
- Recording exactly what you hear and see. This is where recording and transcribing interviews come in. Of course, you must get the participants permission before audio/video recording their interviews. During the write up of your analysis or discussion, always refer to the transcripts and field notes to make sure that you are not moving outside of what was said or observed

- Verifying what was said. If you are unsure about something, it is fine to go back to the participant and seek clarification. Most researchers after transcribing interviews, often send the transcripts to the participants to have them confirm/ verify that transcripts reflect what was said and intended
- Placing the information/data in context. Much of qualitative data and the interpretation of it is a subjective endeavour. To ensure that you are not reporting of giving meaning beyond the data it also helps to place the data within its context. For instance, if during the course of a focus group discussion one participant expressed very negative and hostile views about something, it would be important to discuss the reasons why or the context within which those views were expressed. It could be that the participant's views were expressed in response to provocation from another participant in the focus group.

Other factors that impact the analysis of qualitative data include proper sorting and organizing of the data itself. The next section addresses these stages in the analysis process.

Steps in Qualitative Data Analysis – Sorting and Coding

Just like quantitative data analysis, under qualitative methods large amounts of data – texts and otherwise – are collected during data collection. You must therefore sort, organize and code this data. You can begin by sorting the data by topic and this can be done by developing some sort of **coding scheme**: a set of tags or labels representing the conceptual categories into which to sort the data. These may be developed either before analysis takes place from the conceptual framework driving the study, during the analysis when you begin to identify issues in the data, or by some combination of the two. Next, segments of the data - often paragraphs or sentences - are marked with relevant codes (coded). This is the critical step in which the data are sorted into conceptual categories or chunks. In many cases, researchers write **memos** as they code, recording emerging ideas and early conclusions about both theory and methods. Memos are the researcher 's personal insights, interpretations or observations of the data. As insights accrue, it often becomes useful to search back through the data for places where specific words or phrases are used, and to locate related phenomena in the text, both in order to code these new chunks, and to check the validity of emerging conclusions.

Linking Codes to Texts

The next step in reducing the data is often to retrieve the chunks of text associated with particular codes, reading these passages to refine the understanding of that conceptual category. This is retrieving evidence from the texts to support or validate the codes. You would then be able to read all of the chunks (text passages) where the code has been applied in order to make sense of the code. Alternatively, you might want to identify all of the cases where a particular code applies.

In this way, the researcher begins to be able to write summaries of the main conceptual issues that appear in the data. These preliminary write-ups have the virtues of being much smaller in physical length than the original transcripts and of representing a move from the original concrete data to a more conceptual level or understanding. They also have the disadvantage of being one step removed from the original data, with the danger that some original meaning and context may be lost. This is because the researcher begins to apply his/her understanding to the data. It is therefore important for the researcher to have the ability to examine and re-examine the underlying data as analysis and write-up proceeds, in order to continually check interpretations against the data.

Displaying the Data

Finally, the researcher may enter these summaries into displays, for example text matrices or network diagrams, that aid in summarizing cases and themes, and identifying patterns and relationships in the data (MilesandHuberman,1994). In a text matrix or sometimes called a data matrix, the rows can be used to represent several codes or themes (example bottom up approach and knowledge transfer) while the columns represent some key stakeholder groups (example policy makers and researchers). The cells of the matrix would then be filled with summaries of each stakeholder group's views of each code or theme, allowing the researcher to look for patterns across the data. The network diagram, similarly, shows the researcher 's representation of the connections among the different codes. These displays are intended primarily as analytical tools for the researcher, rather than as illustrations for a reader.

From these summaries of the data the researcher draws conclusions. In order to verify these conclusions, the researcher can employ many of the same mechanisms of searching through the data (looking, for example, for disconfirming evidence) and constructing matrices (for example, to check a conclusion by triangulating from multiple sources or methods) to determine whether the conclusions reached are in fact supported by the data. Finally, a report is produced from the conclusions reached.

LEARNING ACTIVITY 9.4

Recall that:

The deductive approach - Reasoning from the general to the specific. In this approach, you begin by specifying a theory. From the theory, you generate hypotheses about what should happen in specific circumstances. The direction of reasoning is often thought of as "top down," from theory (the general) to data (the specific).

The inductive approach - Reasoning from the specific to the general. In this approach, you begin by examining concrete events or phenomena—your data. From the data move to build theory. The direction of reasoning is often thought of as "bottom up," from the data (the specific) to theory (the general).

Why is quantitative data analysis more suited to the deductive approach while qualitative data analysis is more suited to the inductive approach? Discuss with your peers in the discussion forum.

It provides a good discussion on the development of understanding of bias amongst a group of students on a Masters course in conflict resolution. In the discussion forum, have one of your peers summarize the main points in the article. The rest of the group can then provide feedback and their own additions to that discussion.

UNIT SUMMARY

This unit has attempted an elementary discussion on the data analysis processes under both quantitative and qualitative analysis. Since in this course you are not expected to carry out the actual research but rather to design your proposals for a planned study, we have tried to keep the data analysis processes simple for now. What is important at this stage is that you understand the differences between the two data analysis approaches, the various steps and advantages of these so that you can make informed decisions on deciding which approach to use in your study. You also need to understand the links among the various stages in the research process as you begin to plan and design your methodology, that is, how the research objectives are linked to the data collection methods which are then linked to the data analysis techniques employed. Therefore, if you are going to conduct a quantitative study which most often lends itself to a deductive approach, it means that you will begin with a theory and then a set of hypotheses which need to be tested. You would then conduct a survey, collect quantitative data through a questionnaire. That data would have to be cleaned and coded then entered into some statistical computer programme such as SPSS and various statistical tests run in order to determine relationship between and among variables and thus verify or reject the hypotheses that you began with. Confirmed hypotheses then lend credence to the initial theory that you began with and also contribute to your understanding of the relationships among your variables.

References

Belgrave, L.L. et al. (2002). Qualitative health research. Retrieved at: http://www.ssc.wisc.edu/gender/Resources/Qual-for-quants.pdf

Dean, S. and B. Illowsky (2012). Collaborative statistics. Retrieved at: http://cnx.org/content/m16849/latest/?collection=col10522/latest

eSource - Behavioural and Social Science Research. Software and qualitative analysis. Retrieved at: http://www.esourceresearch.org/Default.aspx?TabId=537

Gilbert, N. (2006). *Researching social life*. London: Sage Publications.

Johnson, B.D. et al. (2010). Structured qualitative research: Organizing 'Mountains of Words' for data analysis, both quantitative and qualitative. Retrieved at: http:// www.ncbi.nlm.nih.gov/pmc/articles/PMC2838205/

Kumar, R. (1996). *Research methodology*. Melbourne: Longman.

Neuman, W.L. (2000). Results with one variable. In *Social Research Methods: Qualitative and Quantitative Approaches*. Boston: Allyn and Bacon, pp. 317 – 322.

Polkinghorne, D.E. (2005). Language and Meaning: Data collection in qualitative research. *Journal of Counseling Psychology*. Vol 52, No. 2, pp137-145. Robson, C. (2002). Real world research. Blackwell Publishing.

Trochim, W. (2006). Research methods knowledge base. Web Centre for Social Research Methods. http://www.socialresearchmethods.net/kb/contents.php

Univariate analysis. Retrieved at: www.sagepub.com/upm-data/9913_040279ch03.pdf