FACULTY OF BUSINESS, HUMANITIES & HOSPITALITY

MASTER IN BUSINESS ADMINISTRATION

BANKING AND FINANCE

BM4406
RESEARCH METHODOLOGY

SELF INSTRUCTIONAL MATERIALS

ACADEMIC YEAR 2021

Table of Contents

Topic 1	Developing a Research Topic	7
	1.1 Narrowing the topic	7
	1.2 Four Terms used in Research	13
	1.3 Theories and testing of Variables	14
	References	
Topic 2	Conducting a Literature Review	17
	2.1 The Purpose of Critical Review	17
	2.2 The content of the critical review	18
	2.3 Literature Sources	21
	2.5 Planning the literature search	21
	2.5 Conducting the literature search using a variety of approaches	22
	References	
Topic 3	Research Methodology/Research Approach	23
	3.1 The Research Process 'Onion'	23
	3.2 The Honeycomb of Research Methodology	24
	References	
Topic 4	Theoretical Framework and Hypothesis Development	30
	4.1 Theoretical Framework in Inductive Research (an aspect of research design)	30
	4.1.1 The variables	31
	4.1.2 A concept model (a schematic diagram)	34
	4.1.3 Theory	34
	4.1.4 Hypothesis Development	36
	4.2 Qualitative Definition and Purpose of Hypotheses	42
	References	
Topic 5	Negotiating Access and Research Ethics	44
	5.1 Access to collect data	44
	5.2 Strategies to gain access	45
	5.3 Research Ethics	45
	5.4 Ethical issues during data collection stage	47
	References	
Topic 6	Selecting Samples (Sampling)	50
	6.1 Representativeness of samples	50
	6.2 Sampling Frame	51
	6.3 Deciding on a Suitable Sample Size	51
	6.4 Sampling Designs (Techniques)	53
	References	
Topic 7	Using Secondary Data	59
	Secondary Data	59

	Measurement bias	61
	Costs and Benefits	61
	References	
		64
Topic 8	Collecting Primary Data through Observation	64
	8.1 Observation	64
	8.2 Researcher Roles in Participant Observation	65
	8.3 Factors that may determine the choice of participant observer role	66
	8.4 Data collection and analysis in participant observation	67
	References	
Topic 9	Collecting Primary Data using Interviews	73
	9.1 Interview	73
	9.2 Situations favoring Qualitative Research Interviews	75
	References	
Topic 10	Collecting Primary Data using Questionnaires	79
	10.1 Questionnaire	79
	10.2 Types of Questionnaire	80
	10.3 Designing the Questionnaire	83
	References	
Topic 11	Analysing Quantitative Data	93
	11.1 Quantitative data	93
	11.2 Preparing data for analysis	94
	11.3 Exploring and Presenting Data	99
	11.4 Describing Data using Statistics	106
	11.5 Examining relationships, differences and trends using statistics	109
	References	
Topic 12	Analysing Qualitative Data	115
	12.1 Qualitative data	115
	12.2 Steps in Qualitative Analysis	116
	12.3 Quantifying Qualitative Data: Content Analysis	118
	References	
Topic 13	Writing and Presenting Project	123
	13.1 Writing	123
	13.2 Suggested Structure for your Project Report	124
	References	

Welcome

Welcome to the study of Research Methodology (BM4406). This module is really interesting because it provides you with guidelines and advice to carry out a research in an area of interest to you or to find an answer to solve a problem confronting you. This module involves many areas of study and knowledge from the initial identification of a research problem to the final research report. This implies that the understanding of the sequence of flow of activities is central to inquiry. It introduces two aspects of research study: the quantitative and qualitative e approaches. It also addresses the concepts of theoretical framework, research designs and other procedures used in the conduct of a research. The study of this module will benefit you when you are conducting your project paper at later stage of your MBA programme.

Description of the Course

Research, generally speaking, is all about generating answers to research questions in order to advance knowledge. Fundamentally, the research questions are derived from a particular research topic or an area of interest. It, then begin by finding out who has studied the research problem you plan to examine to ensure that you do not conduct a research study that has already been done but to enable you to build on the existing knowledge and improving on the subject matter. After having settled down with the appropriate research topic, you will move on to identify through research review and other references to decide on the research approach whether it is going to be qualitative or quantitative or a mixed type. Next you will be in a position to decide on the research strategy and establish your research design. The research design then leads you to the methods to collect the data to find answers to the research questions and to achieve your research objectives.

Aim of the Course

This module aims to provide you the knowledge and understanding about operations management and its importance to an organisation. Operations management is an inclusive concept serving to encompass production of products and services and the non-core production related process such as purchasing, physical distribution, after sales service, literally involving the whole organisation. It has become apparent that operations management can make or break a business. Its strategic importance lies in reducing the cost of production and the development of capabilities to attain competitive advantage and in its absence, other functions in the organisation become redundant. Today 30% of the people are employed in the manufacturing sector of the economy and the remaining 70% in the service sector.

Course Learning Outcome

After the successful completion of this module, students should be able to:

- to explain the need to have an appropriate research topic that is specific, researchable and can be carried out within the available time and resources.
- to identify the two philosophical approaches to conduct the research in terms of qualitative or quantitative or a mixed method approach.
- to demonstrate the importance of the need to have a research design for the collection of data.
- to critically evaluate the data and to reach the approach conclusion of the research.

Support and Contact Details

Name:	:	Dr.Chandramalar Munusami
Email	:	chandramalar@nilai.edu.my
Tel	:	068502338 (208)

Assessment

Assessment Methods and Types:

Coursework: 60%	
i) Individual preparation and presentation of a research topic of your interest in 350 words.	10%
ii) Individual evaluation and presentation of five concepts applied in the module: (a) qualitative and quantitative research, (b) theoretical framework, (c) literature review, (d) research design and (e) research strategy	20%
iii) Group assignment: Write a research proposal - area or subject of interest in 2500 words.	30%
Final Examination :	40%
Total :	100%

Formative Assessment

1. Developing a Research Topic
   Knowledge check activity

2. Conducting a Literature Review
   Knowledge check activity

3. Research Methodology/Research Approach
   Knowledge check activity

4. Theoretical Framework & Hypothesis Development
   Knowledge check activity

5. Negotiating Access & Ethics
   Knowledge check activity

6. Sampling
   Knowledge check activity

7. Using Secondary Data
   Knowledge check activity

8. Collecting Primary Data - Observation
   Knowledge check activity

9. Collecting Primary Data - Interview
   Knowledge check activity

10. Collecting Primary Data - Questionnaire
    Knowledge check activity

11. Analysing Quantitative Data
    Knowledge check activity

12. Analysing Qualitative Data
    Knowledge check activity

Summative Assessment

1. Writing and Presenting Project

Text

Main reference supporting the course:

1. Jonathan Wilson (2014), *Essentials of Business Research, A Guide to Doing your Research project*, 2nd Edition, Singapore: SAGE.
2. Uma Sekaran & Roger Bougie (2016*), Research Methods for Business*, 7th Edition, UK: WILEY.
3. Mark Saunders, Philip Lewis & Adrian Thornhill (2014), *Research Methods for Business Students*, 6th Edition. UK: Prentice Hall.
4. Pamelo S. Schindler (2019). *Business Research Methods*, 13th Edition, McGraw Hill, New York.

Additional references:

1. John W. Creswell (2008), *Educational Research, Planning, Conducting and Evaluating Quantitative and Qualitative Research*, 3rd Edition, Singapore: Pearson.
   (For Research Designs: Chapters 11-18)
2. L.R. Gay, G.E. Mills & P. Airasian (2006). *Educational Research, Competencies for Analysis and Applications*, 8th edition, Pearson, New Jersey.

Topic 1: Developing a Research Topic

Learning Outcomes

By the end of this topic, you will be able to:

1. Able to state the research topic in quantitative term or qualitative term.
2. Refining research ideas and converting them into research project.
3. Develop and formulate research questions and the review of literature.
4. Understanding concepts: purpose statement, hypotheses and research objectives
5. Understanding of theories and testing of variables in quantitative research and in qualitative research.

Introduction

Before you can start your research you have to have some idea of what you want to do. This is in fact the most difficult and important part of the research project. You have first to be clear of what you are interested to research, if not; it would be difficult to plan how to go about researching the topic. This formulating and clarifying process is time consuming. This an imperative step and if done properly you are more likely to achieve a successful project.

The first step in selecting a research topic is to identify a general subject area that is related to your area of expertise and is of particular interest to you. Your idea does not have to be original or unique, and may be similar to existing studies. Still, your final choice of topic is likely to come from your own idea, rather than someone else's.

There are possibly four sources from where you would come up with research topics: theories, personal experiences, previous studies that can be replicated (repeated) and library sources.

1.1 Narrowing the topic

For quantitative or qualitative research, the topic area must be narrowed to a more specific, researchable one. This is to overcome any difficulty to carry out the research or to interpret the findings.

For quantitative research, the time to narrow the topic should be from the beginning of the research process. Whereas for qualitative research it is done at a later time, this is because the qualitative researcher has to start with a general topic area in mind, and then make the necessary observation of the research setting (the place/environment to conduct the research study) and the participants over a certain time before he can formulate a narrow topic.

For ideas to narrow the research topic, the researcher has to talk to his supervisor or to specialists in the area of interest, from literature review and other sources of information.

Upon successful formulation and clarification of the research topic, it is possible to choose the most appropriate research strategy, research design and data collection and analysis techniques. During this time, you will be coming out with research ideas and refining on them. Concurrently you will be looking at research questions and objectives for writing up the research proposal for the research project.

Attributes of a good Research Topic

1. The topic must fit the specifications and meet the standards set by the examining institution or body.
2. The topic is really fascinating or of interest.
3. The topic contains issues that have a clear link to theory.
4. It requires the necessary research skill to undertake the topic.
5. It can be achieved within the available time.
6. It can be achieved within the financial resources that are likely to be available.
7. It is able to gain access to data required for the topic.
8. It is able to state the research questions and objectives clearly.
9. It is related to the idea that has been given.
10. The findings for this research topic are likely to be symmetrical i.e. of similar value whatever the outcome.
11. The research topic meets the career goals.

Stating the Research Topic

Stating a quantitative statement:
For a quantitative study, a well written topic statement generally describes the variables of interest, the specific relationship between those variables and ideally the important characteristics of the participants (a gifted student, teenage-mothers or with learning disability).
An example of a topic statement: "***The topic to be investigated in this study is the effect of "positive reinforcement" on the "quality of 10th graders' English compositions.***"
The variables to be examined are "positive reinforcement" and "quality of English compositions"

Stating a qualitative statement
Qualitative statements are stated in more general language than quantitative ones. This is because qualitative research needs to spend time in the research context for the focus of the study to emerge. Qualitative topic statements will eventually narrow as the researcher leans more about the context, its inhabitants and these more precise statements will appear in the report.
An example of a qualitative statement: "The purpose of this study is to describe the nature of children's engagement with mathematics. The intention is to gather details about children's ways of entering into and sustaining their involvement with mathematics."
Another example:
"This qualitative study examines how members of an organisation identify, evaluate and respond to organisational change. The study examines what events members of an organisation identify as significant change events and whether different events are seen as significant by subgroups in the organisation."
Another example:

"The purpose of this research is to study the social integration of children with disabilities in a general education third-grade class."

Placement and Nature of the Topic Statement in a Study
The statement of the topic is the *first component* of the introductory sections of both the research plan and the completed research report and it gives direction to the remaining aspects of both the plan and report. The topic statement is accompanied by a presentation of:

- *the background* of the topic (scope of study). It provides information for readers to understand the nature of the topic;
- *statement of the problem*. Explain the problem in the proposed research topic you expected to solve.
- a *justification* of the study. It explains how investigation of the research topic might contribute to the theory or practice;
- the *significance* of the topic. It will help to develop a tentative hypothesis or a prediction of research findings. This tentative hypothesis may change over time as a result of the review of the literature. The hypothesis gives direction to the literature search and helps the research narrow its scope to include only relevant topics; and
- *limitations* of the research study. They are potential weaknesses or problems with the study identified by the researcher. The weaknesses are listed one by one. They often may be related to inadequate measures of variables, loss or lack of participants, small sample sizes, errors in measurements and other factors related to data collection and analysis. These limitations may provide useful warnings to other researchers who may want to conduct a similar or replication study.

How to go about generating and refining Research Ideas

Many techniques are used to generate research ideas. The more frequently used techniques for generating and refining research ideas are rational thinking and creative thinking.

Rational thinking	Creative thinking
• Examining your own strengths and interests. • Looking at past project titles • Discussion • Searching the literature	• Keeping a notebook of ideas. • Exploring personal preferences using past projects i.e. from past reports. • Relevance trees • Brainstorming

Other ways to refining Research Ideas

1. Delphi technique - A group of people who are involved or interested in the research idea will examine the research idea to come out with a more specific research idea (Robson, 2002).
2. Preliminary study - a review of some of the literature/an informal discussion with people who have personal experience of and knowledge about your research idea.
3. Integrating ideas - it involves classifying each research idea into its area, then its field and finally the precise aspect in which you are interested.

Turning Research Ideas into Research Project

1. Writing the research questions

This is important to define research questions clearly at the beginning of the research process, as it will help to draw clear conclusions. Research questions cannot be too easy or too difficult. Help may be needed from other people e.g. supervisor.

2. Writing research objectives

The research questions become the base to write the research objectives. Furthermore research objectives require more vigorous thinking. (Note: objective provides the direction and purpose.)

Phrasing research questions as research objectives	
Research questions	Research objectives
1. Why have organisations introduced team briefing?	1. To identify organisations' objective for team briefing scheme.
2. How can the effectiveness of team briefing schemes be measured?	2. To establish suitable effectiveness criteria for team briefing scheme.
3. How team briefing been effective?	3. To describe the extent to which the effectiveness criteria for team briefing have been met.
4. How can the effectiveness of team briefing be explained?	4a. To determine the factors associated with the effectiveness criteria for team briefing being met. b. To estimate whether some of those factors are more influential than other factors.
5. Can the explanation be generalised?	5. To develop an explanatory theory that associates certain factors with the effectiveness of team briefing scheme.

How to go about developing a research question

1. What do you want to find out?
2. Why do you want to research this topic?
3. Why does this research need to be carried out? Is there a gap in the literature or does it contribute to existing theory and/or management practice?
4. What data/information already exists in other similar studies?
   (Adopted from Wilkinson, 2000:16)

Why research questions are important for conducting a research project

1. Help to set boundaries when conducting literature review & identifying the key literature.
2. Help propose a suitable methodology.
3. Help produce a refined set of results.
4. Help to allow easier analysis.
5. Help to draw together a reasonable set of conclusions and make reference to previous research.

How to formulate a reasonable research question

1. it must not be too easy;
2. it should allow for suitable analysis;
3. it should allow the generation of new insight; and
4. it is to avoid common areas of research.

Review of Related Literature

1. The major purpose of reviewing the literature is to determine what has already been done that relates to your topic. It tells you what needs to be done.
2. Previous studies can provide the rationale for your research hypothesis and indications of what needs to be done can help you justify the significance of your study.
3. Another important purpose of reviewing the literature is to discover research strategies and specific data collection approaches that have or have not been productive in investigations of topics similar to yours. This information helps you to avoid other researchers' mistakes and profit from their experience. It may help you to reformulate your research topic.
4. Becoming familiar with previous research also facilitates interpretation of your research findings/result.

Guidelines how far you need to do your review of the literature:

- Avoid the temptation to include everything you find in your literature review. Bigger does not mean better.
- When investigating a heavily researched area, review only those works that are directly related to your specific problem.
- When investigating a new or little-researched problem area, review any study related in some meaningful way to your problem.

Qualitative Research and the Review of Related Literature

- For qualitative researcher the literature review is important but may serve a slightly different purpose. It should not be used to be the direction of the research, because it may defeat the inductive purpose of a qualitative research.
- However it is argued that the review of related literature is important early in the qualitative research process because it serves the following functions:
    1. The literature review brings out the underlying assumptions (propositions) behind the research questions that are central to the research proposal.
    2. It enables the researcher to acquire the knowledge about the proposed research and to defend his research proposition.
    3. It enables the researcher to identify gaps in the body of literature and encourage the researcher to direct his research study into the area where gaps exist.
    4. It helps the researcher refine the research questions and the development of the hypotheses that provide possible directions the researcher may follow.

Identifying Key Words

These key words can help you in your literature search. Sometime you may have to look for alternative words. In looking for the initial key words you can find them at initial sources and probably find additional key words from succeeding sources. After identifying the keywords you are ready to look appropriate sources.

Evaluating Your Sources

Once you have a source in hand, you need to evaluate it. The first thing you have to do is to determine if it really applies to your research topic. If it does then evaluate its quality of the information – for example does it come from a scholarly journal or a proper magazine; is the information someone's personal opinion or the result of a research study? Look at the date of publication and where the

source was found. If the research area is a continuing development then look up for the latest references.
Next identify and verify whether the information is objective and impartial. Does the author's evidence support the interpretations made? Does the information add to the information you have already gathered about your topic? If it does then it is useful and worth paying attention to.
Special care must be taken to evaluating World Wide Web sources.
Conducting effective library and Internet searches will yield an abundance of useful information about your topic. You should collect information that is both up-to-date and comprehensive.

Abstracting from references
After you have identified the primary references related to your topic, you are ready to read the article. You will create abstracts by reviewing, summarizing and classifying your references. You do that for 2 reasons:

1. You have made a summary although it may not be a perfect one.
2. You have read the article.

You should begin the abstracting process by reading the latest references because the most recent research is likely to have profited from previous research. Also recent references may cite preceding studies you may not have identified.

For each reference complete the following steps:

1. If the article has an abstract or a summary, read it to determine the article's relevancy to your problem.
2. Skim the entire article, making mental notes of the main points of the study.
3. Maintain the complete bibliographic reference for the work including the reference to the library classification.
4. Classify and code the article according to some system - a Photostat copy of the article.
5. Summarise the essential points.
6. Indicate any thought that comes to your mind.
7. Indicate any statements that are direct quotations or personal reactions.

Analysing, Organising, and Reporting the Literature
A literature requires a technical form of writing that is unlike most of the writing we do. In technical writing facts must be documented and opinions substantiated. The guidelines for technical writing:

- Document facts and substantiate opinions. Cite reference to support your facts & opinions.
- Define terms clearly and be consistent in your use of them.
- Direct your writing to a particular audience.
- Follow an accepted manual of style e.g. the chapter headings, tables constructed, footnotes and bibliographies and their current editions etc.
- Limit big words and avoid jargon.
- Start each major section with a brief overview of the section E.g. In this section, three main issues are examined. The first is …….″
- End each major section with a summary of the main ideas.

A review of Literature:

- Definition, Purpose and Scope
- Qualitative Research and the Review of Literature
- Identifying Keywords
- Identifying your sources

- Evaluating your sources
- Abstracting
- Analysing, Organising and Reporting the Literature.

1.2 Four Terms used in Research

1. The Purpose Statement

It is a statement that advances the overall direction or focus for the study. This is usually stated in one or two sentences. It is used both in quantitative and qualitative research and is typically found in the 'statement of the problem'. It often appears as the last sentence of an introduction. The statement is stated with the phrase "The purpose of this study is ……."
For example in quantitative research:

The purpose of this study is to examine the relationship between use of Internet communication between teachers and parents in a Midwestern school district and student achievement on tests in high school social studies.

An example in qualitative research would be written as follows:

The purpose of this study is to explore parent stories regarding Internet communications with teachers about their students in one Midwestern school district.

2. Research Questions

Research questions are questions in quantitative or qualitative research that narrow the purpose statement to specific questions that researchers seek to answer. Researchers develop them before identifying the methods of the study (i.e. the type of data to be collected, analysed and interpreted in a study). It is done at the theoretical framework stage. There are few research questions being made to fully explore the research topic. Research questions in quantitative research differ from qualitative research because of different elements required in each research.
In quantitative research, the questions relate attributes or characteristics (*variables*) of individuals or organisations. In qualitative research, the questions include the central concept (i.e. the central phenomenon) being explored.
The research questions are typically at the end of the introduction of the "statement of the problem" section or immediately following the review of the literature. An example of a research question in quantitative research:

Do parent-teacher Internet communications affect student performance in the classroom?

In qualitative research, the research question is as follows:

What types of Internet experiences do parents have with teachers about the performance of the parents' children?

3. Hypotheses

Hypotheses are statements in quantitative research in which the researcher makes a prediction about the outcome of a relationship among the variables. They serve to narrow the purpose statement to specific predictions. These predictions are based on past research and literature where investigators have found certain results and can now offer predictions as to what other investigators will find when they repeat the study with new people or at new sites.

These hypotheses are stated typically at the end of the introduction or immediately after the review of the literature or in a separate section titled "Hypotheses". Usually researchers advance several hypotheses such as three or four. For example:

Students in high schools in the school district in which parents and teachers communicate through the Internet will have higher grades than students whose parents and teachers do not communicate through the Internet.

4. Research Objectives

A research objective is a statement of intent used in quantitative research that specifies goals that the researcher plans to achieve in a study. Researchers often subdivide objectives into major and minor objectives. They appear frequently in survey or questionnaire studies or in evaluation research in which researchers have clearly identified objectives. Like hypotheses and research questions, objectives are found at the end of the "statement of the problem" section, after the literature review, or in a separate section of the study. Objectives can be identified by the phrases such as "The objectives in this study are ……..."

For example: the following represent objectives for a study:

1. To describe the frequency of Internet communication between parents and teachers regarding the parents' children in high school social studies classes.
2. To describe the types of Internet communication between parents and teachers.
3. To relate (a) frequency and (b) types of communication to student achievement in the class as measured by performance on tests.

Note: Research questions, hypotheses and objectives are means to narrow and focus purpose statements.

Why are these statements and questions important?

They provide the direction for the conduct of the research study. They direct the identification of the most appropriate research strategy and research design for collecting data. They also provide key components for understanding the results of a project. Good research links the purpose statement and questions to major results.

1.3 Theories and testing of Variables

In quantitative research we seek to test whether the independent variable influences the outcome of dependent variable. We make this test because we have found past research that suggests that this relationship exists. The researchers may even go further to establish a theory that predicts an impact of the independent variable on the dependent variable. They seek to test this theory. But because the theory is deals with humans in unpredictable situations, we say that the independent variable "probably causes" the dependent variable. The ideas of probable causation is that researchers attempt to establish a likely cause-and -effect relationship between variables rather than prove the relationship.

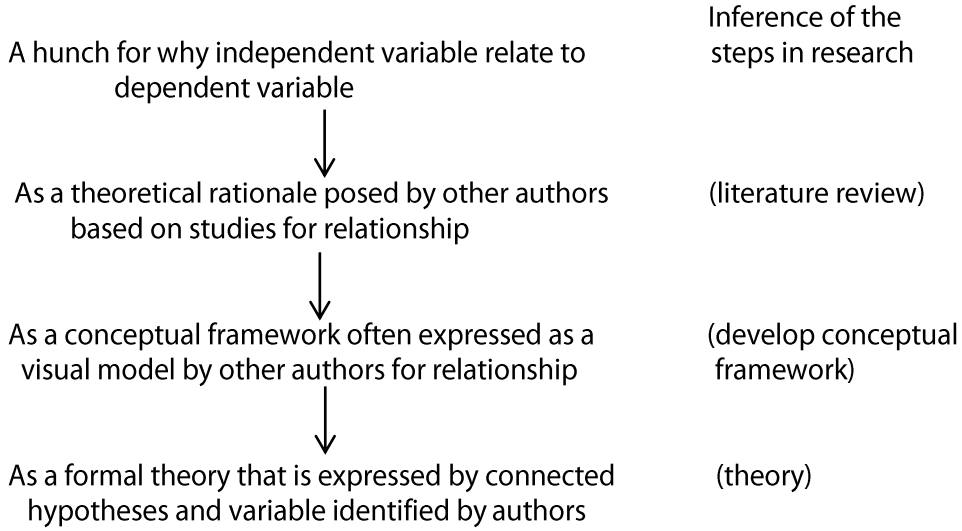
Therefore, *a theory in quantitative research explains and predicts the probable relationship between independent and dependent variables*.

For example, researchers test the relationship between peer groups' influence and adolescents. This relationship is tested over and over, with different groups and in other settings. Repeatedly the relationship of a positive effect holds true. Then some comes along, calls this relationship a theory and assigns a name to it. This theory is born, reported in the literature and test by other researchers.

Therefore, a theory may be thought of as a bridge that connects the independent and dependent variables. Theories are no more than broad explanations for what we would expect to find when we relate variables.
In quantitative research, investigators locate a theory in the literature, examine the predicted relationship, among variables in the theory, and then test the relationships with new participants or at new sites. To test the theory, researchers write purpose statements, research questions and hypotheses that advance the predicted relationship.

For example, a theory on leadership might predict that when managers use consensus-building decision styles, their subordinates feel more supported. Using a theory to test is better than basing variables on your own personal hunches that are subject to challenge by others.

The process of theory development:



Writing the Research Proposal

Writing a research proposal is a crucial part of the research process. It can help to get research funding or for approval by the academic research committee.

The purposes of the research proposal are:

1. for organising your ideas.
2. for convincing your audience.
3. for contracting with your 'client' (supervisor).

Content of your research proposal:

1. Title
2. Background - Tell why you feel the research that you are planning is worth the effort. This may be expressed in the form of a problem that needs solving or something you find exciting and has aroused your curiosity.
   This is also the section where you will demonstrate your knowledge of the relevant literature (Only an overview of the key literature sources from which you intend to draw.)
3. Research questions and objectives

4. Method
5. Time scale
6. Resources - consider the resources needed - finance, data access & equipment.
7. References

SELF CHECK 1

1. Why is it important to spend time formulating and clarifying your research topic?
2. You have decided to search the literature to 'try and come up with some research ideas in the area of Operations Management'. How will you go about this?
3. A colleague of yours wishes to generate a research idea in the area of accounting. He has examined his own strengths and interests on the basis of his assignment and has read some review articles, but has failed to find an idea about which he is excited. He comes and asks you for advice. Suggest two techniques that your colleague could use, and justify your choices.
4. You are interested in doing some research on the interface between business organisations and schools. Write three research questions that may be appropriate.
5. What may be the theory underpinning the decision by organisations sponsoring schools?
6. How would you demonstrate the influence of relevant theory in your research proposal?

References

- Wilson: Chapters 1 and 2 and Saunders: Chapter 2.
- Note: Lecture 1: Developing a Research Topic

Topic 2: Conducting a Literature Review

Learning Outcomes

By the end of this topic, you will be able to:

1. To acquire a good understanding and insight of relevant previous research and the trends that have emerged in relation to your chosen research topic.
2. To lead you to the appropriate direction in your research area in term of your research philosophy defined in your chosen research area.
3. To highlight the work most relevant to your research.
4. To highlight those issues where your research will provide fresh insights.
5. To guide you into subsequent sections of your research project report.

Introduction

It is really impossible to do a research without engaging in literature review related to your intended area of study. You may end up not achieving anything useful in your research study.

A review of the literature will enable you to know the current state of your intended research area. This will prevent you from duplicating what had been done and aiding you to go on to another area of your intended research study. In this way your research findings can contribute to the literature.

The literature review can provide you with useful ideas about developing the appropriate research topic, research questions and the research objectives, and hypotheses especially for deductive approach and the research model. All these ideas and guides eventually enable you to come out with the appropriate research design to enable you to go about collecting your data and analysing the data before you can be in a position to write up your research report.

Literature review is an imperative step in the research process and it must be done *critically* and with great expectations. This is because the success of your research study depended on it.

According to Jankowicz (2000:169) "knowledge does not exist in a vacuum and your work only has value in relation to other people's. Your work and your findings will be significant only to the extent that they're the same as or different from, other people's work and findings."

2.1 The Purpose of Critical Review

It will form the foundation on which the research is built. ***Its main purpose is to help you to develop a good understanding and insight into relevant previous research and the trends that have emerged.***

There are two ways by which the literature is used:

1. Deductive approach - it is the process by which we arrive at a reasoned conclusion by logically generalizing from a known fact. E.g. we know that all high performers are highly proficient in their jobs. If John is a high performer, we then conclude that he is highly proficient in doing his job. In the same way reviewing the literature can help the researcher to identify theories or conceptual framework.

2. Inductive approach - it is a process where we observe certain phenomena and on this basis arrive at conclusion. In other words, in induction we logically establish a general proposition based on observed facts. E.g. we see that the production processes are the prime features of factories or manufacturing plants. We therefore conclude that factories exist for production purposes. In the same way, we can review the literature and from the data acquired from the literature review, can start to develop theories from them.

Both of these ways could be used in a research.
However if you have a predetermined knowledge of the subject area, then it is an advantage because you do not have to review all of the literature before collecting the data. Then the analysis of the acquired data can help you to come out with new findings and theories that have not been thought of before.

Benefits of literature review:

1) It helps you to distinguish whether inductive or deductive approach to be used.
2) It helps to refine further research questions and objectives.
3) It can highlight research possibilities that have been overlooked.
4) It can lead to discover explicit recommendations for further research.
5) It helps to avoid repeating work already done.
6) It can sample current opinions in newspapers, professional and trade journals to gain insights into the aspects of the research questions and objectives.
7) It leads to discover and provide an insight into research approaches, strategies and techniques that may be appropriate to the research questions and objectives.
8) It provides the direction and scope of your research study/project.

2.2 The content of the critical review

Critical means to make judgement by providing a detailed and justified analysis of and commentary on the merits and faults of the key literature within the chosen area.
To make critical judgement, Dees (200) suggested that you should have:

- ✓ Refer to work by recognised experts in your chosen area.
- ✓ Consider and discuss work that supports and work that opposes your ideas.
- ✓ Make reasoned judgements regarding the value of others' work to your research.
- ✓ Support your arguments with valid evidence in a logical manner.
- ✓ Distinguish clearly between fact and opinion.

It is necessary to discuss the work that has already been done in your area of research and to reference that work. It involves drawing out the key points and trends and presents them in a logical way. The background knowledge of the research questions and objectives and established boundaries of the research area need to be reported. Any area of contradiction or problem with the literature it must be justified.

Therefore, the writing of the critical review should have:

1) The key academic theories within the research area.
2) To show that your knowledge of your chosen area is up to date.
3) To show how your research relates to previous published research.
4) To assess the strengths and weaknesses of previous work, including omissions or bias and take these into account in your arguments.
5) To justify your arguments by referencing previous research.
6) Through clear referencing, to enable those reading your project report to find the original work you cite.

The Structure of the Critical Review

The literature review is basically a description and critical analysis of what other authors have written. You also need to focus on the research questions and objectives. How far has the literature gone in answering your research questions? The precise structure of the critical review is usually your choice. It can be a single chapter, a series of chapters or throughout the project report as you tackle various issues.

The common mistake with critical literature reviews is that they become uncritical listings of previous research, often being little more than annotated bibliographies.

Generally a critical review has the following steps:

1. Start at a more general level before narrowing down to your specific research question(s) and objectives.
2. Provide a brief overview of key ideas.
3. Summarise, compare and contrast the work of the key writers;
4. Narrow down to highlight the work most relevant to your research;
5. Provide a detailed account of the finding of this work;
6. Highlight those issues where your research will provide fresh insights;
7. Lead the reader into subsequent sections of your project report, which explore these issues.

Whatever way is used to structure the literature review, it is imperative that you have read, understood and evaluated the items you have located and able to link together the different ideas you find in the literature to form a coherent and cohesive argument. It must relate to the research questions and objectives.

A check list for evaluating the literature review

1. Does your review start at a more general level before narrowing down?
2. Does the literature covered relate clearly to your research question and objectives?
3. Have you covered the key theories of recognised experts in the area?
4. Have you covered the key literature or at least a representative sample?
5. Are those issues highlighted where your research will provide fresh insights?
6. Is the literature you have included up to date?
7. Have you been objective in your discussion and assessment of other people's work?
8. Have you included references that are counter to your own opinion?
9. Have you distinguished clearly between facts and opinions?

10. Have you made reasoned judgements about the value of others work to your own?
11. Have you justified clearly your own ideas?
12. Is your argument coherent and cohesive - do the ideas link together?
13. Does your review lead the reader into subsequent sections of your project report?

An example of a critical review of the literature

Traditional measures of service quality (e.g. Parasuraman et al.); focus on measurement of the gap between service users' perceptions and expectations across a series of dimensions that characterise the service. Notwithstanding shortcomings of conceptualising service quality in this manner, recognised in the SERVQUAL, debates (e.g. Carmen, 1990; Cronin and Taylor, 193; Van Dyke et al., 1997), the use of such a disconfirmation approach is widely reported in the literature (e.g. Robinson, 1999).

The number and nature of constructs, which represent the service encounter, are a function of a service relationship in a particular industry or situation. Each of these relationships differs and is, in reality unique. Gummesson (1996) identifies a series of general qualities characterising relationships such as collaboration, dependency, trust, power, longevity, frequency, closeness, content, as well as personal and social properties. In so doing, he emphasises the breadth of properties that may be deemed relevant by the parties involved in a particular service relationship. However, it is unlikely that all of these properties are of similar relevance to every relationship. Consequently it has been argued that a series of generic dimensions against which to measure service quality is inappropriate (Carmen, 1990). In a review of quality, as one of the primary outcome measures of service relationships, Rosen and Supernant (1998) support this view. They conclude that global measures of service quality (such as SERVQUAL) may not provide the details necessary to assess the strengths and weaknesses of a relationship. In particular, they may fail to take account of the uniqueness and the realities of specific relationship and how they are interpreted and expressed by the parties involved.

Research has also highlighted that interdependencies between organisations are established and maintained through the encounters and interactions of individuals within each organisation (de Burca, 1995). The measurement of the quality of such encounters therefore needs to reflect the perspectives of all these individuals. Rosen and Supernant (1998) suggest that traditional measures fail to reflect fully the dyadic nature of service encounters as they general assess the quality construct from only one partner's point of view. They call for the evaluation of service relationships to accommodate this by including the perspectives of both parties. Although they suggest that this may result in the need to reconcile different views, they also highlight the need for awareness and understanding of the views of all parties involved in a service encounter. We would content that these processes could result in both parties involved in the service questioning the relevance of the norms against which they evaluate the encounters. This, we believe, support our contention that approaches which have the ability to capture a diversity of service users' and providers' experiences of such concepts, are likely to be of more value. Furthermore, we have argued that where measures focus only on specific transactions, they may fail to take account of the ongoing nature of service relationships that are based upon repeated encounters (William et al., 1999).

Rosen and Supernant (1998) support Smith's (1995) arguments when they point to the shortcoming o the global nature of the quality construed as a diagnostic tool for remedial action. This implies that the assessment of the relationship's quality should lead to action to enhance benefits obtained by both parties from it. Data collected to assess quality should therefore be useful. In this context, usefulness can be viewed from two key perspectives.

Source: An article published by Mark & Christine in the Journal of European Industrial Training (Saunders and Williams, 2000:220-221).

Identify:
1) The overall structure of starting at a more general level before narrowing down;
2) The provision of a brief overview of the key ideas.
3) Narrowing down to highlight that work which is most relevant to the research reported in the paper.
4) Providing more detail about the findings of that work which is most relevant.

2.3 Literature Sources

Literature sources are classified into 3 categories:

1. Primary sources: reports, theses, emails, conference reports, company reports, some government publications (white papers & planning documents) & unpublished manuscript sources (letters, memos & committee minutes). These are also referred to as the grey literatures they are difficult to trace/locate.
2. Secondary sources: newspapers, books, journals (periodicals, serials and magazines), internet and some government publications. They are easier to locate than the primary sources.
3. Tertiary sources: indexes, abstracts, catalogues, encyclopaedias, dictionaries, bibliographies & citation indexes.

2.4 Planning the literature search

Literature search is time consuming. It is then necessary to plan this search. It can start by:

1. Having clearly defined research question(s) and objectives.
2. Defining the parameters of the research.
   From the research questions and objectives you will have a good idea of which subject matter is going to be relevant. However you need to be clear (Bell, 1999):
   - The language of publication (e.g. English)
   - The subject area (e.g. accountancy)
   - Business sector (e.g. manufacturing)
   - Geographical area (e.g. Asia)
   - Publication period (e.g. the last 10 years)
   - Literature type (e.g. refereed journals and books)
3. Generating key words and search terms.
   Undertake to read key authors and recent review articles in your research area in order to help you to define your subject matter and to suggest appropriate key words. They may contain references to other work that is pertinent to your research question(s) and objective(s).
   Key words or search terms are important as they help you to find the relevant literature as they are the basic terms that describe your research question(s) and objectives. They help you to search the tertiary literature. Key words include the authors' surnames can be found by:
   - Discussion with colleagues, project tutor and librarians.
   - Initial reading
   - Dictionaries, thesauruses, encyclopaedias and handbooks.
   - Brainstorming
   - Relevance trees.

4. Discussing the ideas as widely as possible.

2.5 Conducting the literature search using a variety of approaches

These will include:

- Searching using tertiary sources and the Internet.
- Following up references in articles you have read.
- Scanning and browsing secondary literature in your library.

Once obtained, the literature must be evaluated for its relevance to your research questions and objectives. This must include a consideration of each item's currency. Each item must be read and noted. Bibliographic details, a brief description of the content and appropriate supplementary information should also be recorded.

SELF CHECK 2

1. What do you need to have in order to plan your research?
2. What are the possible sources for a literature review to be conducted?
3. In what ways a literature review can help you in your research project?
4. In what ways a literature review can be used?
5. What is a critical literature review and its benefits?
6. How would you plan your literature review?
7. How would a critical literature review benefit a researcher?

References

- Wilson: chapter 3, Saunder: chapter 3 and Uma: chapter 4.
- Note: Lecture 2: **Conducting a Literature Review**

Topic 3: Research Methodology/Research Approach

Learning Outcomes

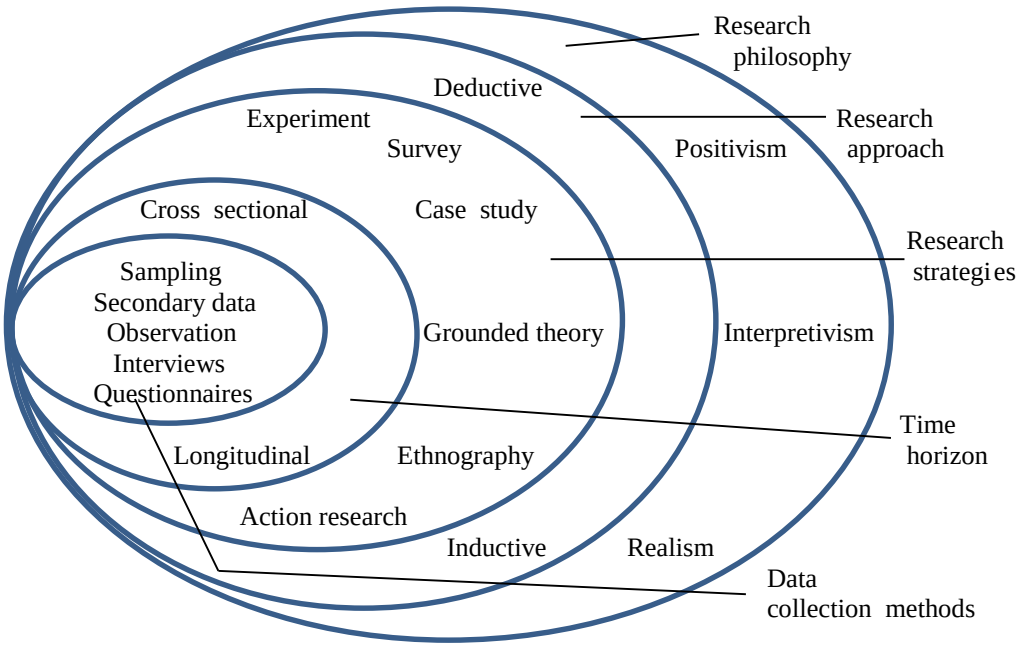
By the end of this topic, you will be able to:

1. To acquire an understanding of the importance of research philosophy in research method.
2. To distinguish the importance of inductive and deductive research approach and their implications in the conduct of the research.
3. To have the ability to distinguish the research strategies for the conduct of a research and their tactics.
4. To develop the appropriate research design to achieve the objectives of the research.
5. To apply the necessary ethical practices in the conduct of the research.

Introduction

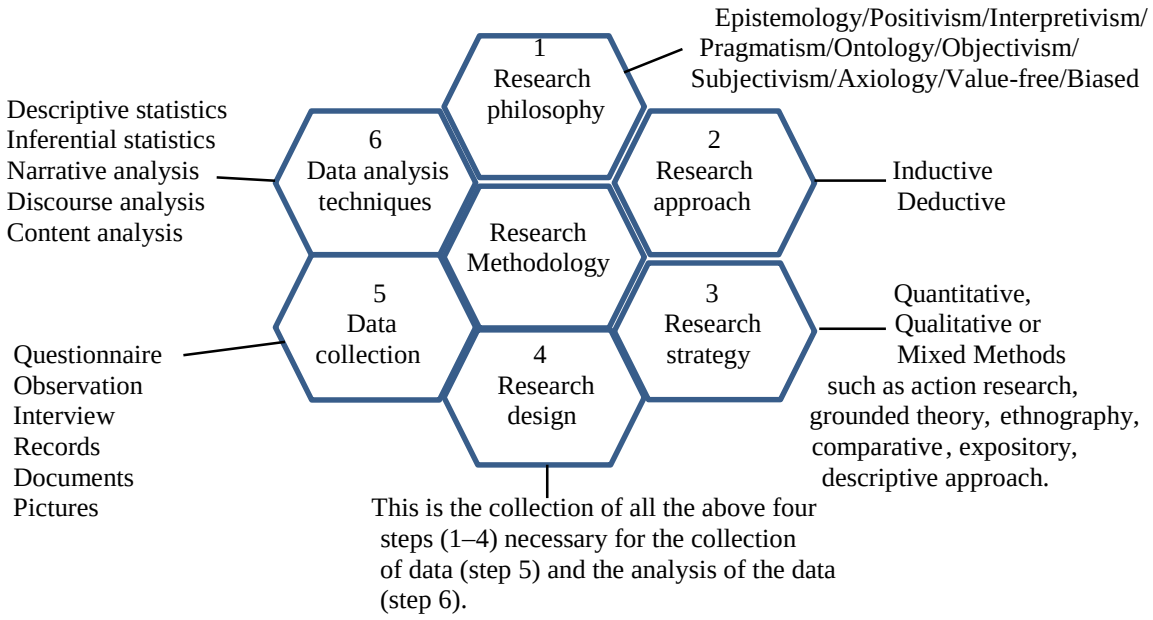
This chapter examine two approaches related to research methodology:
1. The Research Process 'Onion' (Saunders et al)
2. The Honeycomb of Research Methodology (Jonathan Wilson)

3.1 The Research Process 'Onion'



The above **onion research process** shows the different aspects as layers of an onion. The first represents the philosophy of research, then the research approach, research strategies, and time horizon before coming to data collection methods. All these **five** aspects constitute the **research design.**

3.2 The Honeycomb of Research Methodology



The honeycomb approach indicates **six** aspects of the research methodology with the inclusion of the research design compared to the 'onion process' approach. The research design is not included in the research process 'Onion'. In the research process 'Onion' the research design is constituted to be from the five aspects starting from the research philosophy to the data collection methods (see diagram).

A research design is a framework that is developed as a guide to the collection, analysis and interpretation of data using quantitative or qualitative strategy. There are many research designs and each of them has its unique characteristics to guide you in your research process.
As with regard to the research strategy in aspect 3 and as a part of a research design, it refers to the quantitative approach, qualitative approach or mixed approach that is applied to conduct a research study. The research strategy is referred to as the method that is developed by researchers for the collection of data and analysis of data such as a case study, grounded theory, experiment, action research, ethnography, and comparative study, expository study and descriptive study.

There is a relationship between research philosophy and research approach and research strategy. This implies that the choice of the research philosophy will likely to determine the research approach and research strategy. E.g. if you choose positivism as your research philosophy, then your research approach has to be deductive and your research strategy has to be quantitative. On the other hand, if you choose interpretivism, your research approach has to be inductive and your research strategy has to be qualitative as well.

Similarly, your choice of research philosophy also involves your choice of the research design. It leads you to the type of evidence required and how it is to be collected and interpreted. Furthermore, it helps you to identify and adapt the research design option according to the constraints of your research topic.

Review the 5 aspects in the Research Process Onion:

1. The Research Philosophy (principle or concept)

(*Philosophy originates from the Greek word "philosophia". "Philo" means 'loving/lover of' and "sophia" means 'wisdom'. Therefore philosophy means "love of wisdom". It is concerned with the search for meaning.*

Research philosophy is related to the systematic inquiry into the nature and meaning of reality, existence, truth, knowledge, reasoning and value.)

There are three philosophical views about the research process that dominate the literature:

1. Positivism (deductive/quantitative approach)
2. Interpretivism (inductive/qualitative approach)
3. Realism (mixed approach)

These are three different views about how knowledge is developed and judged as being acceptable. They are different but not mutually exclusive. All of them have an important part to play in business and management research. In other words there are three ways by which data or information is gathered to provide answers to the research study.

Positivism (Quantitative approach)

This approach reflects what a scientist does. He sets an objective and conducts an experiment and from the data collected, he determines the outcome. This method is highly structured and can be repeated as the environment is controlled. The data can be statistically analysed to determine whether the objective is achieved. This is a scientific quantitative approach.

Interpretivism (Quality approach)

This is a qualitative approach to study the activities of a selected group of people or a selected few persons in a natural environment and then attempt to draw meanings of the events and activities of these people. The role of the observer is to make sense of and understand their motives, actions and intentions in a way that is meaningful for these research participants.

Realism (Mixed approach)

This is based on the belief that a reality exists that is independent of human thoughts and beliefs. That there are unforeseeable forces existing in the external environment to affect the interpretations and behaviours of the people but somehow the people do not notice them.

Realism recognises the importance of understanding people's socially constructed interpretations and meanings within the context of seeking to understand broader social forces, structures or processes that influence the nature of people's view and behaviours.

Business and management research is not done just using one approach but a combination of positivism and interpretivism reflecting the stance of realism.

2. Choosing a research approach

This is related to whether the research should use the deductive approach or the inductive approach. In the deductive approach a theory/ hypothesis is developed and then design a *research strategy* to test the hypothesis. This is related to positivism.

On the other hand, the inductive approach involves collecting the data and then develops a theory from the data analysis. This is related to the philosophy of interpretivism (qualitative approach).

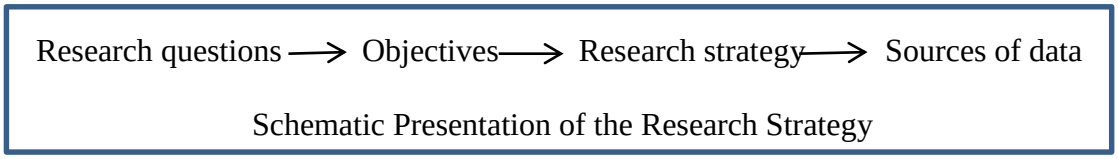
Deductive approach process	Inductive approach process
1. Come out with a hypothesis about the relationship between two or more vents from the theory. 2. Express the hypothesis in operational terms i.e. how the variables are to be measured. 3. Test the hypothesis i.e. do an experiment. 4. Examine the outcome to confirm the theory. 5. If necessary modify the theory.	1. Collect the data e.g. by observation or interview. 2. Analyse the data collected. 3. The result of the analysis leads to the formation of a theory. 4. It involves understanding the social aspects of human behaviour.

Major differences between deductive and inductive approaches to research	
Deduction emphasises (quantitative)	Inductive emphasises (qualitative)
1. Scientific principles 2. Moving from theory to data 3. The need to explain causal relationships between variables 4. The collection of quantitative data 5. The application of controls to ensure validity of data 6. The operationalization of concepts to ensure clarity of definition 7. A highly structured approach 8. Researcher independence of what is being researched 9. The necessity to select samples of sufficient size in order to generalise conclusions.	1. Gaining an understanding of the meanings humans attach to events. 2. A close understanding of the research context 3. The collection of qualitative data 4. A more flexible structure to permit changes of research emphasis as the research progresses 5. A realisation that the researcher is part of the research process 6. Less concern with the need to generalise

3. Research Strategy

It is *a general plan of how you will go about answering the research question.* It contains clear objectives, derived from the research question, specify the sources from which you intend to collect the data and indicate the constraints you will face such as time, financial resources, location, and access to data and ethical issues. It refers to quantitative approach or qualitative approach or even a mixed approach.

Research topic Quantitative/Qualitative/Mixed



Schematic Presentation of the Research Strategy

Differences between strategy and tactics: Strategy is concerned with the overall approach you adopt; tactics is about the finer detail of how the data is collected and the analysis methods used.

The research strategy options available are:

- Experiment - a systematic approach to gather data and interpretation of the data.
- Survey - a deductive approach. It allows the collection of a large amount of data from a sizable population. Often obtained by using a questionnaire, the data are standardised and compared. The importance of this strategy is the designing of the questionnaire. Other methods to collect data are observation and interview.
- Case study - an in-depth investigation of one unit (e.g. individual, group, institution, organisation, programme or document). It involves the understanding and translating the dynamics of the selected unit. It is a problem-solving technique of a culture-sharing group. It is basically to understand what is going on or the things a particular group of people do in a particular environment. Multiple methods of primary data collection such as observation, interviews and questionnaires are used to improve the qualitative research and concurrently to enable the research findings to be as authentic as possible.
- Grounded theory - data collection starts without the formation of an initial theoretical framework. Theory is developed from data generated by a series of observations. These data lead to the generation of predictions that are then tested in further observations which may confirm, or otherwise, the predictions. Constant reference to the data to develop and test theory leads Hussey and Hussey (197) to call grounded theory an inductive/deductive approach, theory being grounded in such continual reference to the data.
- Ethnography - an inductive approach; came from the field of anthropology. The purpose is to study the cultural patterns and perspectives of participants in their natural setting, a form of qualitative research.
- Action research - any systematic inquiry conducted by teachers, principals, school counsellors, or other stakeholders in the teaching-learning environment, to gather information about the ways in which their particular schools operate, the teachers teach, and the students learn.
- Cross-sectional and longitudinal studies

  Cross-sectional studies - involve the study of a particular phenomenon at a particular time. E.g. the study of the IT skills of managers on one organisation at a given time e.g. over a period of days or weeks or months.

  Longitudinal studies - where the researcher studies people or phenomena at more than one point in time in order to answer the research question. E.g. the research might want to study employees' behaviour before and after a change in the top management, to study the effect of the change. Here, because data are gathered at two different points in time, the study is not cross-sectional or one- shot kind, but is carried longitudinally across a period of time.
- Exploratory, descriptive and explanatory studies.

  Exploratory studies - this is usually done when there is not much information available about the situation at hand. Data are gathered through interviews, questionnaire or even observation or a search of the literature.

  Descriptive studies - this is to ascertain and to describe the characteristics of the variables of interest in a situation. E.g. a study of a class in terms of the percentage of members who are in their senior and junior years, sex composition, age groupings, number of semesters until graduation and number of business courses taken. The goal of a descriptive study is to offer a profile or to describe relevant aspects of the phenomena of interest to the researcher from an individual, organisational, industry-oriented or other perspective. Such information may be vital before even considering

certain corrective steps such as to bring about a change in the management. Presenting data in a meaningful form help to:
(1) understand the characteristics of a group in a given situation;
(2) think systematically about aspects in a given situation;
(3) offer ideas for further problem and research; and
(4) help make certain simple decisions.

Explanatory studies - it is a study of a situation or a problem to establish the relationship between variables e.g. is there a correlation between the two variables.

4. Research design

The research design is a detailed plan or a blueprint of the steps to be followed to achieve the objectives conceived by the researcher in terms of answering the research questions formulated by him in any area or discipline.

The research design is actually a collection of steps involved before starting to carry out the collection of data. It involves the followings:

- *Identifying the research topic*
- *Develop research question(s) and objectives.*
- *Selecting an appropriate research strategy most suitable to the research questions and objectives.*
- *Looking at constraints under which the research is being conducted e.g. is the pursuit of longitudinal research precluded.*
- *Does it involve combining different research methods (different techniques used for collecting data – Questionnaire, interview, observation & documents.*
- *The threats to credibility such as reliability and validity contained in your research design.*
- *Moral ethics in collecting data.*

The research design option that is chosen will provide the above steps for the collection of data.

It has the following properties:

- *Is time-based, procedural plan for every research activity.*
- *Is always focused on the research question.*
- *Guides selection of sources of information (cases).*
- *Provides the framework for specifying the relationships among the study's variables.*

Mixed Methods

This is a combination of quantitative and qualitative methods and primary and secondary data are used.

Credibility of research findings

This refers to how good are the data collected i.e. the trustworthiness of the research. This depends on how accurately the instrument developed to measure a particular concept. It is necessary to ensure that the instrument that is use in the research measures the variables accurately. To ensure that the measures developed are reasonably good, we use the reliability and validity tests.

Reliability - it indicates the extent to which the measure is without bias (error free) and hence offers consistent measurement across time and across the various items in the instrument. It therefore indicates the stability and consistency with which the instrument measures the concept and helps to assess the 'goodness' of a measure. This is more towards consistency.

Validity - It is the degree to which a test measures what it is intended to measure; a test is valid for a particular purpose for a particular group. In qualitative research, it is the degree to which qualitative data accurately gauge what the researcher is trying to measure. There are many threats to validity according to Robson (2002) such as history, testing, instrumentation, mortality, maturation and ambiguity about causal direction.
Generalizability (external validity) - to what extent the findings of the research can be applied to the population. In a case study it is not possible to apply to other situations because circumstances and conditions such as the settings and participants are different.

The ethics of *research design*

Moral ethics should exist when gaining access to the place for research to be carried out, to the participants, and the collection of the data. In fact moral ethics applies in all the stages of the research process starting from designing to the research strategy through the implementation stage and the evaluation and control stage. All these stages if are not done ethically have serious implications on the validity of the research.

SELF CHECK 3

1. Briefly explain the steps you would take to decide a research project.
2. What are the ethical issues and the constraints you would face as well as the reliability and validity of your research design?
3. What are the factors that determine a researcher to go for quantitative approach or qualitative research?

References

- Saunders: chapter 4, Wilson: chapter 5 and Uma: chapter 6
- Note: Lecture 3: **Research Methodology/Research Approach**

Topic 4: Theoretical Framework and Hypothesis Development

Learning Outcomes

By the end of this topic, you will be able to:

1. Explain in deductive research, the relationships of the variables are subjected to hypothetical testing.
2. Explain the association of the variables is derived from the theory.
3. Identify the differences between a deductive research framework from an inductive research framework.

Introduction

The key areas of learning in this topic are:

- The Research Process 'Onion' of Saunders et al and The Honeycomb research methodology of Johanthan Wilson.
- Understand the steps involved in a research approach.
- The choice of a research approach
- The concept of research strategies in relation to the research questions and their options
- The concept of research design as distinct from data collection

4.1 Theoretical Framework in Inductive Research (an aspect of research design)

A *theoretical framework* is a presentation of how certain variables or concepts) are related to each other and the explanation of how they are associated to each other. Draw *a model* to show their relationships and from which a theory is drawn. (In fact, the model and the theory in the research area can be derived from a review of the literature related to the research area.)

The process of building a theoretical framework involves:
1. Introducing the variables in the model.
2. Develop the conceptual model that indicates the relationships between the variables.
3. Coming up with a theory that explains the relationships between the variables in the model.
4. Hypotheses are drawn from these relationships of the variables

The hypothesized relationships are then tested through appropriate statistical analyses. This approach is known as *deductive* quantitative approach i.e. not for inductive (qualitative) approach.

In summary:
The process is as follows:

Theoretical framework ⟶ variables ⟶ a model ⟶ theory ⟶ hypotheses

4.1.1 The variables

What are variables?
A variable is anything that can take on differing or varying values. This value can be different at different times for the same object or persons or at the same time for different objects or persons. Examples of variables are productive units, absenteeism and motivation.

There are ***4 types of variables***:

a) The *dependent variable* (also known as criterion variable).
b) The *independent variable* (also known as the predictor variable).
c) The *moderating variable*. (A variable that has an effect on the dependent variable and independent variable relationship.)
d) The *mediating variable*. (A variable that stands between the dependent variable and independent variable.)

Each variable can be a discrete entity (e.g. male or female) or a continuous entity (e.g. the age of an individual).

a) Dependable variable
This type of variable is dependent on other variables to influence its outcome or change. It is possible to have more than one dependent variable in a study such as between quality and volume of output; low-cost production and customer satisfaction; and quality and cost.
In a research study the researcher looks for the factors that influence the dependent variables such as how they change.

b) Independent variable
It is the one that influences the dependent variable in either a positive or negative way. When the independent variable is present, the dependent variable is also there. The idea is to find whether the change of the dependent variable is influenced by the independent variable.

However, to establish that a change in the independent variable causes a change in the dependent variable, there are four conditions to be met:

1. A change in the dependent variable should be associated with a change in the independent variable.
2. The independent variable (the cause) must precede the dependent variable. The cause must take place before the effect occurs.
3. No other factor should be a possible cause of the change in the dependent variable.
4. A logical explanation (a theory) is needed and it must explain why the independent variable affects the dependent variable.

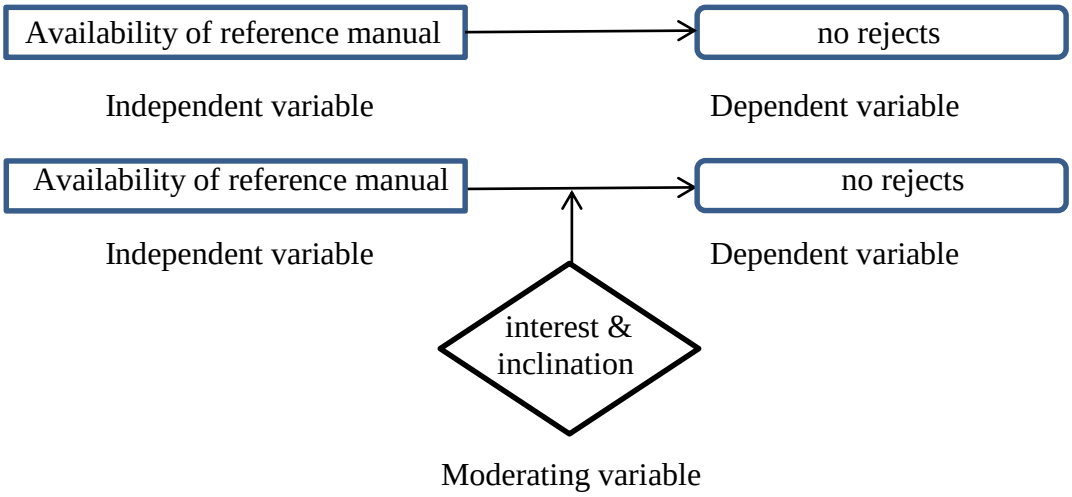
c) Moderating variable

- The moderating variable has a strong contingent effect on the independent variable-dependent variable relationship. It is a new variable constructed by the researcher to

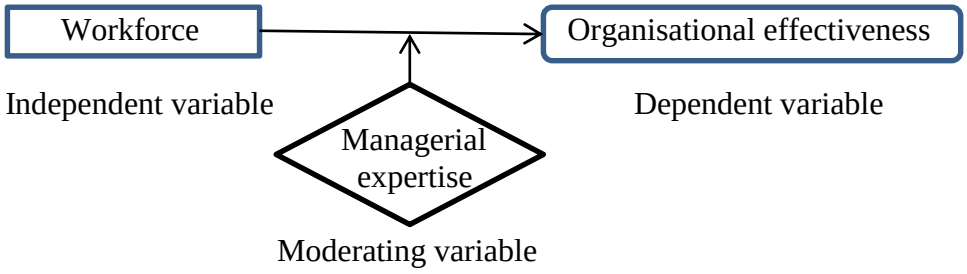
interact with the independent variable to bring about an *interaction effect*. This moderating variable is a special form of independent variable.

- It means that the presence of a third variable (the moderating variable) modifies the original relationship between the independent and dependent variables.
- The moderating variable has an interaction with the independent variable in explaining the variance.
- The moderating variable is an independent variable.

E.g. It has been found that there is a relationship between the availability of the reference manuals that manufacturing employees have access to and the product rejects. If the workers follow the procedures as laid down in the manuals, they are able to manufacture products that are flawless. This relationship holds true generally for all workers. However, it is also dependent on the inclination/attitude of the employees to look into the manuals whenever there is a new procedure to be adopted. In other words, only those who have the interest and urge to refer to the manual every time a new process is adopted will produce flawless products. Others who do not consult the manual will continue to produce defective products. This influence of the attitudes of the workers on the relationship between the independent and the dependent variables is known as the moderating variable and is illustrated in the figure below.



Another example: A company has a diversified workforce. The effectiveness of the organisation is dependent on the ability of the manager to capitalise on this diversified workforce. The diversified workforce is the independent variable and the effectiveness of the organisation is the dependent variable. The degree of effectiveness of the organisation is influenced by the ability of the manager to use the diversity workforce. The relationships among the variables can be shown as follows:



The distinction between an independent variable and a moderating variable

- Normally an independent variable directly influences the dependent variable and the effect is shown in figure (a).
- However a moderating variable tends to modify the influence of the independent variable on the dependent variable. This can be seen in figure (b), where the expertise of the manager can change the influence of the workforce on the effectiveness of the organisation.

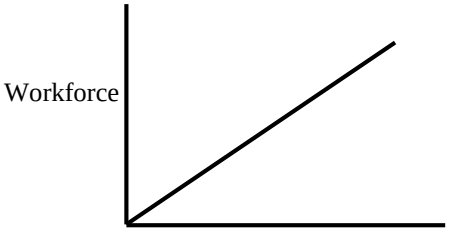


Figure (a)

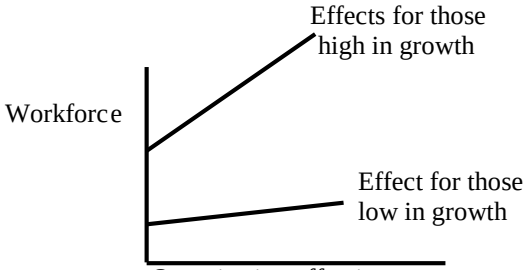


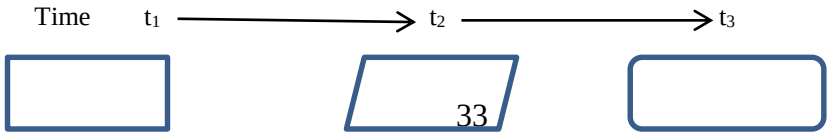
Figure (b)

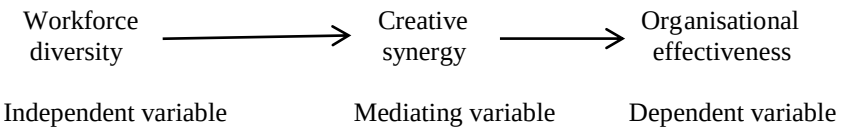
Note the steepness of the top line and the relative flatness of the bottom line in Figure (b).

d) Mediating variable (Intervening variable)

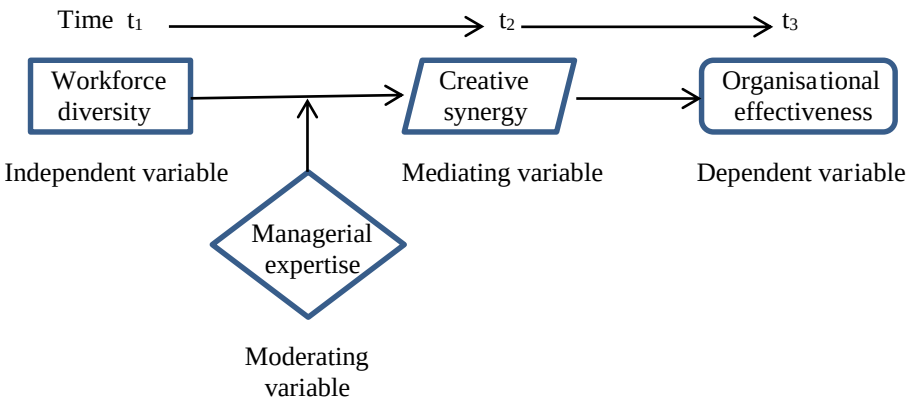
- Mediating/Intervening variables are different from dependent variables or any of the types of independent variables.
- An intervening variable is an attribute or characteristic that stands between the independent variable and dependent variable and exercises an influence on the dependent variable and the independent variable.
- Intervening variables transmit (or mediate) the effects of the independent variable on the dependent variable.
- It is a variable that surfaces between the time the independent variables start operating to influence the dependent variable and the time their impact is felt on it (i.e. later).
- There is a temporal quality or time dimension to the mediating variable. In other words, bringing a mediating variable into play helps you to model a process.
- The mediating variable surfaces as a function of the independent variable(s) operating in any situation, and helps to conceptualize and explain the influence of the independent variable(s) on the dependent variable.

E.g. in the previous example, the workforce as an independent variable influences the dependent variable (organisational effectiveness), the mediating variable that surfaces as a function of the diversity in the workforce is the "creative synergy" that comes from the differences in the culture, experience and skills of the different workers in the organisation. This creative synergy surfaces at time , t_2 and then helps organisation to attain organisational effectiveness subsequently in t_3. This development is illustrated in the figure below.





What can happen to the above situation if a moderating variable (managerial expertise) is included? The model is changed or the relationships of the four variables will be affected and this can be seen in the figure below.

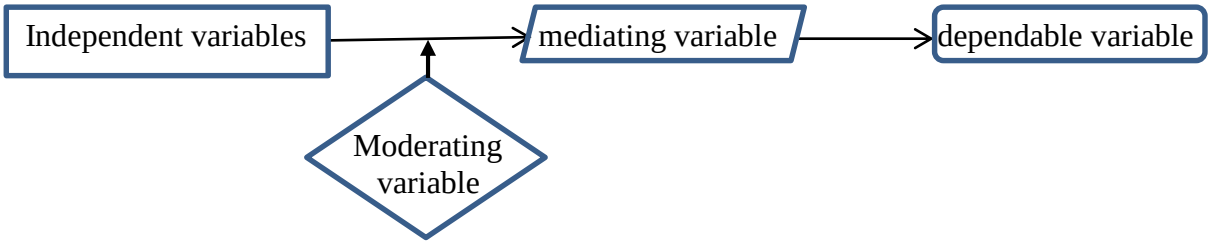


It shows that the managerial expertise moderates the relationship between workforce diversity and creative synergy. In other words, the creative synergy will not result from the problem-solving skills of the diverse workers unless the manager is capable to harnessing that synergy by creatively coordinating the problem-solving skills of the different skills of the workers. If the manager does not have the expertise to perform this role, then no matter how many different problem-solving skills the diverse workforce might have, synergy will not surface. Instead of functioning effectively, the organisation might just remain static, or even become worse.

4.1.2 A concept model (a schematic diagram)

It is presented to describe your ideas about how the concepts (variables) are related to each other.

For example:



4.1.3 Theory

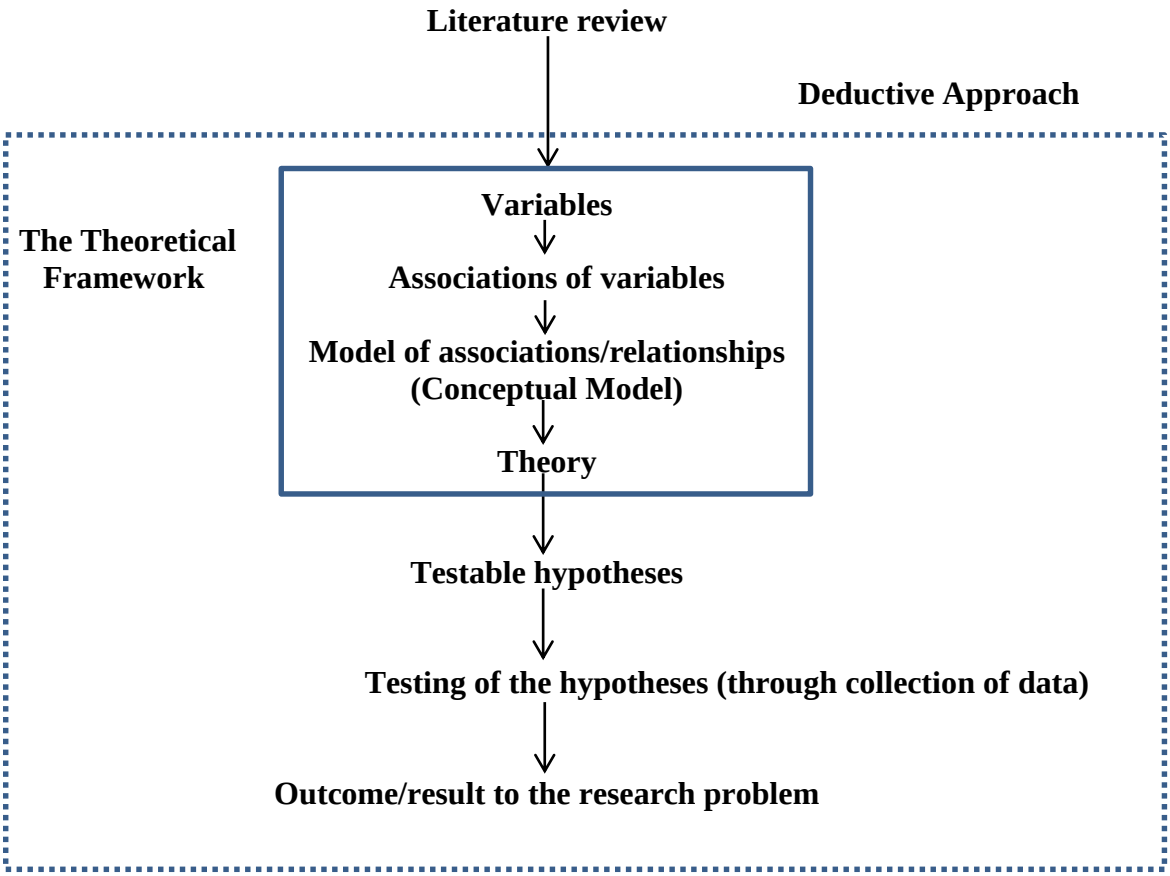
A theory or a clear explanation for the relationships in your model is the last component of the theoretical framework. From the theoretical framework, testable hypotheses are developed to examine whether the theory formulated is valid or not.

How Theory is generated

The literature review is an imperative activity because the literature review is the foundation for the development of the theoretical framework. From the theoretical framework the entire deductive research project is based (a quantitative approach). It is a logically developed, described and elaborated network of associations among the variables deemed relevant to the problem situation and identified through such process as interviews, observations and literature review (research methods/data collection instruments). Other factors like intuition and experience also guide the development of the theoretical framework. ***A theory is derived from the conclusion of the relationships of the variables in the model.***

Note that it is through the literature review that the variables are identified by previous research findings. It is through the connections of the variables that the theoretical model is evolved. The theoretical framework represents and elaborates the nature and direction of the relationships. Just as the literature review sets the stage fo*r a good theoretical framework, this in turn provides for the development of the testable hypotheses.*

Brief summary in schematic form:



4.1.4 Hypothesis Development

Quantitative Definition and Purpose of Hypotheses

When researchers conduct studies they want to know if a specific variable they are studying will have an effect on a specific population. Since it is unpractical to test the entire population, a sample is taken from the population to test the population. Then a *statistical hypothesis* is made on the assumption of the population from the sample taken. This hypothesis testing allows the researcher to use sample data and to infer the results from the sample to the population. The purpose of the hypothesis is to decide whether the results of the study indicate a real relationship between variables, or if the results simply show the random fluctuation that would be the result of chance.

What is a hypothesis?
Hypothesis is an educated guess that a researcher makes based on information available to him. This information can be obtained from researcher's own experience or from the literature review. The hypothesis so developed will be tested using an appropriate statistical procedure to decide it can be accepted or rejected. The researcher does not set out to prove the hypothesis but rather collect data that either support or do not support the hypothesis.

All quantitative research studies require having hypotheses. However the possible exception in quantitative research studies is the descriptive study whose purpose is to answer certain specific questions.

Hypotheses are derived from theories or from knowledge gained while reviewing the related literature. (See theoretical framework.) Therefore it is essential to conduct a literature review in order to develop appropriate hypotheses for the related topic.

Generally *a quantitative researcher formulates a hypothesis before conducting the study because the nature of the study is determined by the hypothesis. Every aspect of the research is affected by the hypothesis, including participants, measuring instruments, design, procedures, data analysis and conclusions.*

It is imperative to understand that although all hypotheses are based on theory or previous knowledge, they aim to extend the knowledge. There are certain criteria that are applied to determine the value of a given hypothesis.

Criteria for Hypothesis
In quantitative research a good hypothesis has the following characteristics:

1. It is based on sound reasoning that is consistent with theory or previous research.
2. It provides a reasonable explanation for the predicted outcome.
3. It clearly states the expected relationship between defined variables.
4. It is testable within a reasonable time frame.

Types of Hypotheses
1. In terms of how they are derived such as *Inductive* hypothesis or *Deductive* hypothesis.
2. In terms of expected relationship or difference between two variables such as

non-directional hypothesis or *directional* hypothesis.
3. In terms of significant or different relationship such as the null hypothesis.

Stating a Hypothesis
A good hypothesis is stated clearly and concisely, expresses the relationship between two variables and defines those variables in measurable terms.

Writing Quantitative Hypotheses
Guidelines:
1. State the variables in this order: independent (first position), dependent (second position) and control (third position).
2. If comparing groups in the hypothesis, explicitly state the groups, if variables are related, specify the relationship among the variables.
3. Make a prediction about changes you expect in your groups, such as less or more favourable or no changes (e.g. no difference). You will then test this prediction using statistical procedures.
4. You may state information about the participants and the site of the study, but this information may not be necessary if it repeats information stated in your purpose statement.

Formats regarding Statements of Hypotheses
1. If-the statements
Example:
If *young women are more frequently exposed to images of thin models in advertisements,* ***then*** *they will be more likely to express dissatisfaction with their body weight.*

2. Directional and nondirectional hypotheses
Examples:
Directional hypotheses:
The ***greater*** *the stress experienced in the job, the* ***lower*** *the job satisfaction of employees.*

Women are ***more*** *motivated* ***than*** *men.*

Nondirectional hypotheses:
There is a relation between arousal-seeking tendency and consumer preferences for complex designs.

There is a difference between the work ethic values of American and Asian employees.

3. Null and Alternate Hypotheses
There are two types of hypotheses: the null and the alternative to the null. They are needed in a research study.

Type of hypothesis	Null Hypothesis	Alternative Hypothesis
Purpose	To test in the general population that there is no	The hypothesis that may be true if the null is rejected; it suggests a

	change, no relationship, no difference	change, a relationship or a difference.
Specific language	There is no difference (or relationship) between.	Magnitude statements such as higher, lower, more positive, more favourable.
How researcher test the Hypothesis	A test of the hypothesis	A test of the hypothesis.

Null Hypotheses

A null hypothesis (labelled as H_0) is set up to be rejected in order to support an alternative hypothesis (labelled as H_A). The null hypothesis is presumed true until statistical evidence, in the form of hypothesis testing, indicates otherwise. However, it is stated typically in terms of there being no (significant) relationship between two variables or no (significant) difference between two groups. What is this all about? Well, in setting up the null hypothesis, we are stating that there is no difference between what we might find in the population characteristics and the sample we are studying. Since we do not know the true state of affairs in the population, all we can do is to draw inferences based on what we find in our sample. What we imply through the null hypothesis is that any differences found between two sample groups or any relationships found between two variables based on our sample are simply due to random sampling fluctuations and not due to any "true" differences between the two population groups or relationships between two variables. The null hypothesis is thus formulated so that it can be tested for possible rejection. If we rejection the null hypothesis then all permissible alternate hypotheses relating to the particular relationship tested could be supported. It is the theory that allows us to have faith in the alternate hypothesis that is generated in the particular research investigation. *This is one more reason why the theoretical framework should be grounded on sound, defendable logic to start with. Otherwise, other researchers are likely to refute and postulate other defensible explanations through different alternate hypotheses.*

The null hypothesis in respect of group differences stated in the example "Women are more motivated than men" would be:

$$H_0: \mu_m = \mu_w$$

or

$$H_0: \mu_m - \mu_w = 0$$

where H_0 represents the Null Hypothesis, μ_m is the mean motivational level of men and μ_w is the mean motivational level of the women.

The alternate for the above example would statistically be set as follows:

$$H_A: \mu_m < \mu_w$$

which is the same as

$$H_A: \mu_w > \mu_m$$

where H_A represents the alternate hypothesis and μ_m and μ_w are the mean motivation levels of men and women respectively.

For the *nondirectional hypothesis* of mean group differences in work ethic values in the example "There is a difference between ethic values of American and Asian employees."

$$H_0: \mu_{am} = \mu_{as}$$

or

$$H_0: \mu_{am} - \mu_{as} = 0$$

where H_0 represents the null hypothesis, μ_{am} is the mean of work ethic value of Americans and μ_{as} is the mean work ethic value of Asians.

The alternate hypothesis for the above example would statistically be set as:

$$H_A: \mu_{am} \neq \mu_{as}$$

where H_A represents the alternate hypothesis and μ_{am} is the mean work ethic of the Americans and μ_{as} is the mean work ethic of Asians.

The null hypothesis for the relationship between the two variables in the example "The greater the stress experienced in the job, the lower the job satisfaction of employees," would be $H_{0:}$ There is no relationship between stress experienced on the job and the job satisfaction of employees. This would be statistically expressed by:

$$H_{0:}\ p = 0$$

where p represents the correlation between stress and job satisfaction, which in this case is equal to 0 (i.e. no correlation).

The alternate hypothesis for the above null, which has been expressed directionally, can be statistically expressed as:

$$H_{0:}\ p < 0 \text{ (the correlation is negative)}$$

For the example "There is a relationship between age and job satisfaction," which has been stated nondirectionally, the null hypothesis would be statistically expressed as:

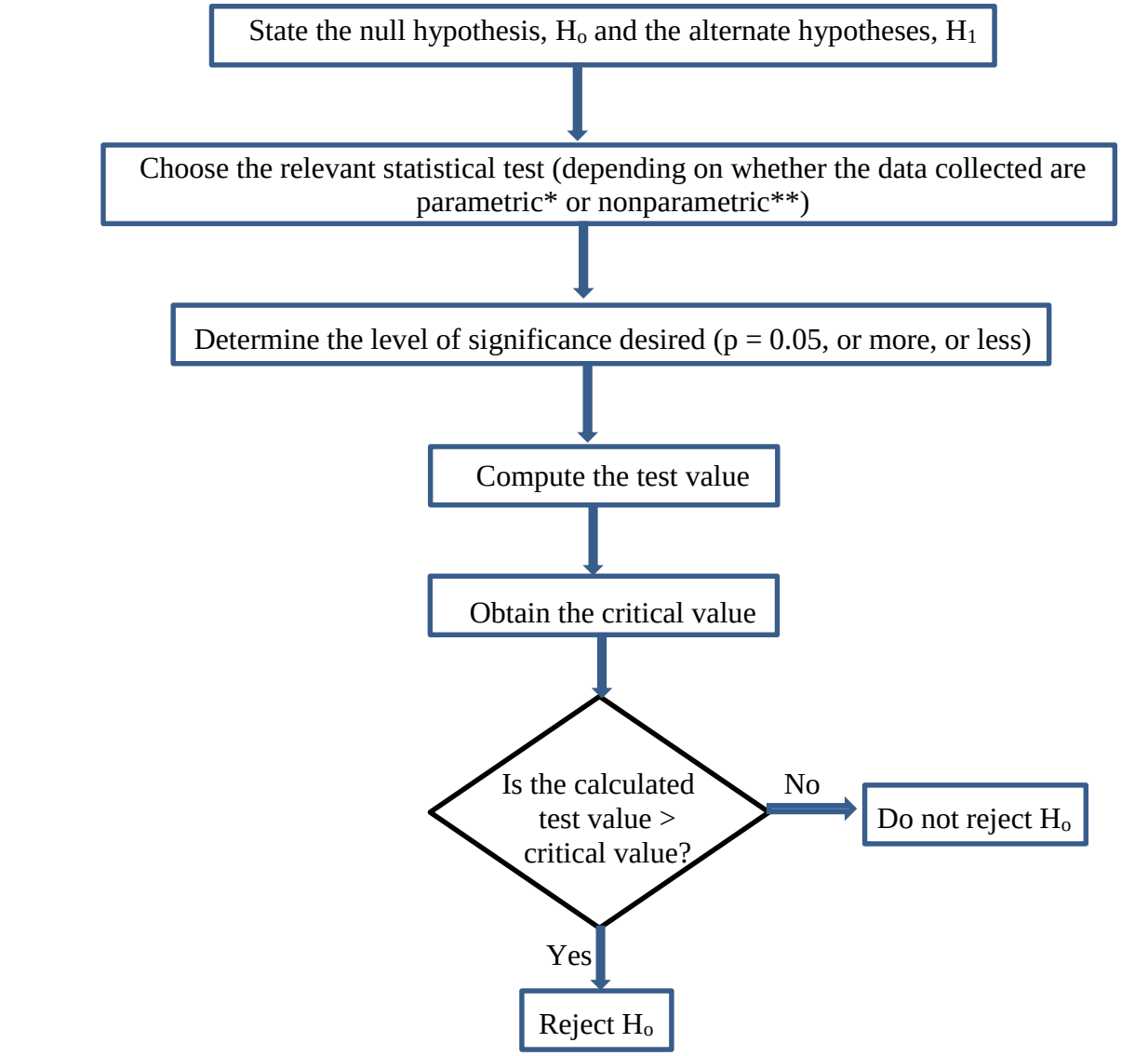
$$H_{0:}\ p = 0$$

whereas the alternate hypothesis would be expressed as

$$H_{0:}\ p \neq 0$$

Having formulated the null and alternate hypotheses, the appropriate statistical tests (t-tests, F-tests) can then be applied, which indicate whether or not support has been found for the alternate hypothesis – i.e. that there is a significant difference between groups or that there is a significant relationship between variables as hypothesized.

The steps to follow in testing hypotheses:



* Parametric test – data follow a normal distribution and using means and standard Deviations.
** Nonparametric test – not possible to have normal distribution and sample is of small size.

Both are derived from *inferential statistics* where the data is used to infer to a wider population.

1. State the null and alternative hypothesis
   The researcher starts the analysis usually by stating the null hypothesis (H_o) and the alternative hypothesis (H_1). Before stating the hypotheses, it is always advisable to state the research problem in terms of a question that identifies the population (s) of interest to the researcher, the parameter(s) of the variable under investigation, and the hypothesised value of the parameter(s).

This helps the researcher to clearly state the null and alternative hypotheses for the research problem under investigation.

For example, A lecturer has observed that the female students are generally more attentive in class than the male students. An obvious question may strike the mind: Is the performance of the female students better than the performance of male students? If this question is set for the purpose of investigation, the null and alternative hypotheses can be defined as:

H_o : There is no significant difference in the average marks scored by male and female students (H_o: $u_1 = u_2$)

H_1: The average marks scored by female students is significantly higher than the average marks scored by male students (H_1: $u_1 > u_2$).

2. Choose the relevant statistical test

   Here the researcher has to decide what test will be used based on the values of the parameters that will be tested. This is also where the researcher will decide whether to use a parametric or a non-parametric test. A clear statement of the hypotheses helps the researcher in deciding the types of variable(s) that should be used in sample data collection which further defines statistical test to be used. For example, in the above stated hypothesis, the parametric test will be most suitable as the variable (mars) is a ration scaled variable.

3. Select the relevant level of significance

   The researcher has to fix the criteria for the decision which means defining the alpha level and critical region. The critical region is composed of extreme scores that are very unlikely to be obtained if the null hypothesis is true. The result has to be beyond these scores in order to be statistically significant.

   The alpha level is a probability value. It defines the critical region and tells us the probability that a result beyond the critical region occurred by chance. Normally the levels of significance used by most of the researchers are 0.01 and 0.05. A level of significance ($\alpha = 0.05$) implies that there is a 5% chance that a result in the critical region occurred by chance.

4. Compute the calculated value

   Once the data is collected, the researcher's job is to calculate the test value of the relevant statistical test.

5. Obtain the critical test value based on the test in step 2

   After calculating the value based on the selected test, we have to look up the critical values for each test based on the particular distribution of the test (i.e. normal, X^2, t, *F*). This critical value will help us in the decision whether the hypothesis is supported or not supported.

6. Draw the conclusion

   If the calculated value is larger than the critical value then we reject the null hypothesis and conclude that the alternative hypothesis is supported.

Type I and Type II Errors

There is always a risk that the inference a researcher draws about a population may be incorrect. There are two types of errors associated with hypothesis testing. We may reject H_o when H_o is true and we may accept H_o when H_o is not true. The former is known as Type 1 error and the latter as

Type II error. The Type I error is denoted by alpha (α), known as α error or the level of significance of test. Type II error is denoted by beta (β), known as β error. These two types of errors can be represented in the tabular form:

	Decision	
	Accept H_o	Reject H_o
H_o (True)	Correct decision	Type I error (α error)
H_o (False)	Type II error (β error)	Correct Decision

The probability of Type I error is usually determined in advance and is understood as the level of significance of testing the hypothesis. If Type I error is fixed at 5%, it implies that there are about 5 chances out of 100 that we will reject H_o when H_o is true. The Type I error can be controlled by fixing it at lower level. For example, instead of fixing it at 5%, if we fix it at 1%, then the probability of committing Type I error reduces from 0.05 to 0.01.
But with a fixed same size, n, when we try to reduce Type I error, the probability of committing Type II error increases.

4.2 Qualitative Definition and Purpose of Hypotheses

Qualitative researchers do not state formal hypotheses before conducting a study. They seek to understand the nature of their participants and contexts before stating a research focus or hypothesis. They may develop hypotheses to guide them for their proposed research but they are not tested. The qualitative researchers gather data and then build patterns and associations in the participants' natural environment without having prior hunches of what to study or observe.

Identifying patterns and associations in the natural setting help the researchers to discover ideas and questions that lead to new hunches or hypotheses. Therefore, in qualitative research, the strength lies in generating hypotheses not testing hypotheses.

Having identified a guiding hypothesis, the qualitative researcher may "operationalize" the hypothesis through the development of research questions that provide the researcher with a focus for data collection. The qualitative research questions may encompass on topics such the participants' understanding of meanings and social life in a particular context.

SELF CHECK 4

1. Make up three different situations in which motivation to work would be an independent variable, a mediating variable and a moderating variable.
2. Failure to follow accounting principles causes immense confusion, which in turn creates a number of problems for the organisation. Those with vast experience in bookkeeping, however, are able to avert the problems by taking timely corrective action. List and label the variables in this situation, explain the relationships among the variables and illustrate these by means of diagrams.
3. A store manager observes that the morale of employees in her supermarket is low. She thinks that if their working conditions are improved, pay scales raised, and the vacation benefits made more attractive, the morale will be boosted. She doubts, however, if an increase in pay scales would raise the morale of all employees. Her conjecture is that those who have supplemental incomes will just not be "turned on" by higher pay, and only those without side income will be happy with increased pay, with a resultant boost to morale. List and label the variables in this situation. Explain the relationships among the variables and illustrate them by means of diagrams. What might be the problem statement or problems definitions for the situation?
4. Create a diagram to illustrate the relationships between the relevant variables in question 3 above, and develop five different hypotheses.

References/ Key Readings

- John W. Creswell (2008) pages 120-143.
- Um Sekaran, chapter 5

Note: Lecture 4: **Theoretical Framework and Hypothesis Development**

Topic 5: Negotiating Access and Research Ethics

Learning Outcomes

By the end of this topic, you will be able to:

1. Explain the importance of gaining accessibility to the collection of data from the sources.
2. Identify the strategies to gain access to collect data.
3. Examine the ethical issues in the research process.

Introduction

This chapter will explain about the negotiating access and research ethics issues. The ability to collect data depends on gaining access/entry to the source. This may face certain difficulties:

1. Organisations or individuals may not be prepared to engage in additional, voluntary activities because of time and resource constrain.
2. The request for access and cooperation may fail to interest the person who receives it and this could be due to a number of reasons:
   - a lack of perceived value in relation to the work of the organisation or individual.
   - the nature of the topic because it is potential sensitive or concerns the confidentiality of the information that would be required.
   - perceptions of your credibility and doubts about your competence.
3. The organisation may find itself in a difficult situation owing to external events totally unrelated to any perceptions about the nature of the request or the person making it.

5.1 Access to collect data

There are different types and levels of access and they provide different problems to gain access to the collection of the required data. E.g. physical access to an organisation, access to intended participants, continuing access in order to carry out further parts of the research or to be able to repeat the collection of data in another part of the organisation; cognitive access in order to get sufficiently close to find out valid and reliable data.

The feasibility to gain access to the data for your research is an important determinant of what you choose to research and how you undertake the research. In the event you gain an entry to an organisation, there are other levels of access to be considered in order to realise your research strategy. It is necessary to also gain informal acceptance by the participants who form your research sample. You need to conduct yourself in an unbiased way and ethically in getting your data as reliably as possible. In other words, you must be able to reach the position where you can reveal the reality of what is occurring in relation to your research question and objectives. It is important that you are able to gain the confidence of the people in the organisation about your credibility and competence. to gain the answers to your question and objectives. If your research

requires you to gain access on more than one occasion and at different times these issues must be formally relate to the management of the organisation in order avoid problems in the future.

5.2 Strategies to gain access

Summary of strategies to gain access

- Allowing yourself sufficient time to gain permission to gain access.
- Using existing contacts and developing new ones.
- Providing a clear account of purpose and type of access required.
- Overcoming organisational concerns about the granting of access.
- Identifying possible benefits to the organisation in granting you access.
- Using suitable language.
- Facilitating ease of reply when requesting access e.g. pro forma (page 126).
- Developing your access on an incremental basis.
- Establishing your credibility with intended participants.

5.3 Research Ethics

Ethical concerns extend to many areas in the research: planning the research, gaining access to organisation and to individuals, collect, analyse and report the data. Ethics actually refers to the appropriateness or acceptability of your behaviour in relation to the rights of those who become the subject of your work. Therefore, you need to consider the ethical issues throughout the period of your research and to remain sensitive to the impact of your work on those whom you approach to help, those who provide access and cooperation, and those affected by your results.

The conduct of your research may be in the form of a code of ethics where a statement of principles and procedures for the conduct of your research is provided. This is because in a research you will be in contact with other people for your information collection. E.g. in the face-to-face interview places the researcher in a position of some power because he is able to formulate questions including probing ones, which may cause levels of discomfort.

Ethical issues that affect the research process generally

- privacy of possible and actual participants;
- voluntary nature of participation and the right to withdraw partially or completely from the process;
- consent and possible deception of participants;
- maintenance of the confidentiality of data provided by individuals or identifiable participants and their anonymity;
- reactions of participants to the way in which you seek to collect data;
- effects on participants of the way in which you use, analyse and report your data;
- behaviour and objectivity of the researcher.

Implications of respecting privacy in business and management research

Respecting privacy in business and management research means the right:

- not to participate;

- not to be harassed or offered inducement to participate or to extend the scope of participation beyond that freely given;
- not to be contacted at unreasonable times or at home (where the scope of the research is related to an organisational setting)
- of participants to determine, within reason, when they will participate in the data collection process;
- of participants to expect the researcher to abide by the extent of the consent given and not to find that the researcher wishes to widen the scope of the research without first seeking and obtaining permission;
- of participants not to be subject to any attempt to prolong the duration of an interview or observation beyond that previously agreed unless the participant freely proposes this as an option;
- of the participants not to answer any question, or set of questions, or provide any related data where requested;
- of participants not to be subjected to questions that create stress or discomfort;
- to expect agreed anonymity and confidentiality to be observed strictly both in relation to discussions with other research or organisational participants and during the reporting of findings (including from those who gain subsequent access to data).

Ethical issues during the design and initial access stages

- Not to pressure anyone to grant access.
- Not to interfere the privacy of the intended participant e.g. by telephoning him at unacceptable time or confronting intended participant.
- Where personal data about individuals are obtained without their consent, you have to treat the data in the strictest confidence and not to abuse it in any way.
- When a person has consented to participate in the data collection, it does not mean that the person has also consented to the way in which the data are to be used.
- Ensure that informed consent is obtained. Three situations are shown with regard to the nature of consent:

<table>
<tr><th>Lack of consent</th><th>Implied consent</th><th>Informed consent</th></tr>
<tr><td>Participant lacks knowledge</td><td>Participant does not full understand her/his rights</td><td rowspan="2">Participant consent given freely on full information about participation rights and use of data.</td></tr>
<tr><td>Researcher uses deception to collect data</td><td>Researcher implies consent about use of data from fact of access or return of questionnaire</td></tr>
</table>

There are many ways to establish informed consent and an example in the form of a checklist is shown below:

Checklist of requirements for informed consent
About the nature of the research ✓ What is its purpose? ✓ Who is or will be undertaking it?

✓ Is it being funded or sponsored -- if so, by whom and why? ✓ Who is being asked to participate - i.e. broad details about the sampling frame, sample determination and size? ✓ How far has the research project progressed?
About the requirements of taking part ✓ What type of data will be required from those who agree to take part? ✓ How will these data be collected e.g. interview, observation or questionnaire? ✓ How much time will be required, and on how many occasions? ✓ What are the target dates to undertake the research and for participation?
About the implications of taking part and participants' rights ✓ Recognition that participation is voluntary. ✓ Recognition that participants have the right to decline to answer a question or set of questions, or to be observed in particular circumstances. ✓ Recognition that the participants have control over the right to record any of their responses where use of a tape recorder is contemplated. ✓ Recognition that participants may withdraw at any time. ✓ What are the consequences of participating - possible risks, depending on the nature of the approach and purpose, and expected benefits? ✓ What assurances will be provided about participant anonymity and data confidentiality?
About the use of the data collected and the way in which it will be reported ✓ Who will have access to the data collected? ✓ How will the results of the research project be disseminated? ✓ How will assurances about anonymity and confidentiality be observed at this stage? ✓ What will happen to the data collected after the project is completed? ✓ Where data are to be preserved, what safeguards will be 'built in' to safeguard the future anonymity and confidentiality of participants?

5.4 Ethical issues during data collection stage

General issues irrespective of any method used to collect data:

- The privacy rights of participants.
- The maintenance of objectivity in the collection of data is to be accurately and fully and avoiding any subjectivity in what is recorded. It will affect the reliability and validity of the research instrument.
- Confidentiality and anonymity have to be shown.
- Seeking explanations in qualitative methods must be done within the appropriate and acceptable parameters.

Issues related to particular method of collecting data

- The use of observation method must be within the boundaries of what is permissible to observe e.g. not to intrude into the private life of the participants or their private telephone calls and to take note of the reaction of the participant.
- In the face-to-face interviews avoid overzealous questioning and pressing the participant to respond.

- The questions to be asked need to be considered and avoid any way demeaning to your participant.
- Seeking to conduct an interview requires to arrange a time suitable to the participant or to prolong a discussion.
- If you have observed something you are not allowed to do so, then you have to inform those affected about what has occurred and why. This is known as debriefing.

Ethical issues to the analysis and reporting stages

Objectivity must be maintained during the analysis of the data collected. This means to avoid selecting the data to report or misrepresenting its statistical accuracy. The ethical issues of confidentiality and anonymity must be maintained during the reporting stage. Permission must be gained from the person concerned to revive any confident information.

The use of the research data by others is another ethical issue. It is viewed that the identity of the participant s should not be revealed even though the research conclusions are used to make decisions.

The introduction of the data protection legislation has led to the aspect of research assuming a greater importance and to a need for researchers to comply carefully with a set of legal requirements to protect the privacy and interests of their data subjects.

Summary checklist to anticipate and deal with ethical issues
1. Attempt to recognise potential ethical issues that will affect your proposed research. 2. Utilise your university's code on research ethics to guide the design and conduct of your research. 3. Anticipate ethical issues at the design stage of your research and discuss how you wil seek to control these in your research proposal. 4. Seek informed consent through the use of openness and honesty, rather than using deception. 5. Do not exaggerate the likely benefits of your research for participating organisations or individuals. 6. respect others' rights to privacy at all stages of your research project. 7. Maintain objectivity and quality in relation to the processes you use to collect fata. 8. Recognise that the nature of a qualitatively based approach to research wil mean that there is greater scope for ethical issues to arise and seek to avoid the particular problems related to interviews and observations. 9. Avoid referring to data gained from a particular participant when talking to others, where this would allow the individual to be identified with potentially harmful consequences to that person. 10. Covert research should be considered only where activity is likely to be significant issue or where access id denied (and a covert presence is practical). However, other ethical aspects of your research should still be respected when using this approach. 11. Maintain your objectivity during the stages of analysing and reporting your research. 12. Maintain the assurance that you gave to participating organisations with regard to confidentiality of the data obtained and their organisational anonymity. 13. Consider the implications of using the internet and email carefully in relation to the maintenance of confidentiality and anonymity of your research participants and their

data, before using this means to collect any data. Avoid using this technology to share any data with other participants.

14. Project individual participants by taking great care to ensure their anonymity in relation to anything that you refer to in your research project report, dissertation or thesis.
15. Consider how the collective interests of your research participants may be adversely affected by the nature of the data that you are proposing to collect, and alter the nature of your research question and objectives where this possibility is likely. Alternatively declare this possibility to those who you wish to participate in your proposed research.
16. Consider how you will use secondary data in order to protect the identities of those who contributed to its collection or who are named within it.
17. Unless necessary, base your research on genuinely anonymised data. Where it is necessary to process personal data, comply with all of the data protection legal requirements carefully.

SELF CHECK 5

1. How can you differentiate between types of access and why is it important to do this?
2. Which strategies to help to gain access are likely to apply to the following scenarios?
    a. an 'external' researcher seeking direct access to managers who will be the research participants;
    b. an 'external' researcher seeking access through an organisational gatekeeper/broker to her/his intended participants;
    c. an 'internal' researcher planning to undertake a research project within her/his employing organisation?
3. What are the principal ethical issues you will need to consider irrespective of the particular research methods that you use?
4. What problems might you encounter in attempting to protect the interests of participating organisations and individuals despite the assurances that you provide?

References/Key Readings

- Saunders: chapter 5 and Wilson: chapter 4
- Note: Lecture 5: **Negotiating Access and Research Ethics**

Topic 6: Selecting Samples (Sampling)

Learning Outcomes

By the end of this topic, you will be able to:

1. Deciding on a suitable sample size with 95% level of certainly.
2. Choose the right sample for a research investigation as it is imperative to ensure its representation of the population to validate the research project.
3. Explain the techniques in sampling (sampling designs) – probability sampling and nonprobability sampling.

Introduction

This chapter will explain about sampling and the process of selecting a sufficient number of participants from a population so that by studying the sample and understanding the characteristics of the sample subjects it would be possible to generalise the characteristics of the population.

The reasons for using a sample rather than collecting data from the entire population:

1. It would be practically impossible to collect data several hundreds or thousands of people in a population. Even if it were possible, it would be prohibitive in terms of cost, time and other human resources.
2. Studying a sample rather than the entire population is more likely to lead to more reliable results, mostly because fatigue is reduced, resulting in fewer errors in collecting data.

6.1 Representativeness of samples

The need for choosing the right sample for a research investigation is imperative. This is because a sample will rarely be the exact replica of the population from which it is drawn. However, it is possible to choose a sample in such a way that it is representative of the population as there is a small probability that the sample values might fall outside the population parameters.

The characteristics (height, weight) in the population are generally normally distributed. These attributes are such that most people will be clustered around the mean and there will only be a small number at the extremes that are very tall or very short, very heavy or very light. If we are able to estimate the population characteristics from the characteristics represented in a sample with reasonable accuracy, the sample then has to be chosen such that the distribution of the characteristics of participants follows the same type of normal distribution in the sample as it does in the population. The sampling distribution of the sample mean is then normally distributed. As the sample size ‘n’ increases the means of the random samples taken from practically any

population approach a normal distribution with mean 'U' and the standard of deviation ' '. If we take a sufficiently large number of samples and choose them with care, we will have a sampling distribution of the means that has normality.

In scientific research, the need of a sample to be representative of the population is important for generalizability. However in certain research, this generalizability of the sample may not be that important. This is seen in the case study situation.

(*Parameter refersto the characteristics of a population. On the other hand, statistics refers to the characteristics of a sample and it is used to estimate the value of the population.)*

6.2 Sampling Frame

Sampling frame is basically a complete list of all the cases/members in the population. E.g. in a Lion Club the complete sampling frame is the complete list of all its members. It is important to have the complete list of all members in order to avoid biasness and to make your sample a representative of the population.

Checklist for selecting a sampling frame

- ✓ Are cases listed in the sampling frame relevant to your research topic, for example are they current?
- ✓ Does the sampling frame include all cases in other words is it complete?
- ✓ Does the sampling frame exclude irrelevant cases, in other words is it precise?
- ✓ (For purchased list/s) Can you establish and control precisely how the sample will be selected?

6.3 Deciding on a Suitable Sample Size

A sample can be a representative of a population i.e. it has generalizability, if it is appropriate chosen. A large sample size reduces error. Sampling is like a compromise between the accuracy of the findings and the amount of time and money invested. Such a compromise is governed by:

1. The confidence you need to have in your data i.e. the level of certainty that the characteristics of the data collected will represent the characteristics of the population.
2. The margin of error that you can tolerate i.e. the accuracy you require for any estimate made from your samples.
3. The types of analyses you are going to make - each type of analysis may require a minimum threshold number.
4. The size of the total population from which your sample is being drawn.

As a rule of thunb, the Economist's (197) advised that for any statistical analysis, a minimum number of **30** is required for a sample size.
As for researchers they normally work to a **95%** level of certainty i.e. if your sample is selected 100 times at least 95 of these samples would be certain to represent the characteristics of the population.

Margin of Error
Margin of error describes the precision (accuracy) of the estimates of the population. A larger sample size normally reduces the margin of error.

The table below provides a rough guide to the different sample sizes required for different sizes of population at 95% level of certainty.

	Margin of Error			
Population	5%	3%	2%	1%
50	44	48	49	50
100	79	91	96	99
150	108	132	141	148
200	132	168	185	198
500	217	340	414	475
1000	278	516	706	906
10000	370	964	1936	4899
100000	383	1056	2345	8762
1000000	384	1066	2395	9511
10000000	384	1067	2400	9595

The Importance of High Response Rate

A sample that is selected must be a perfect representative sample of the population. This means that every member in the sampe is able to make the same response. If a sample is made up of 100 members in the small service business, it means the 100 members must be really involved in the small service business and they could make the necessary responses. This implies a high response rate because the sample provides a true representative of the population. It will then reduce the biasness because of the absence of non-responses and the cost of the survey.

Non-response can be due to 4 reasons:

1. Refused to respond.
2. Ineligibility to respond (not meeting the requirements).
3. Inability to locate respondent.
4. Respondent located but unable to make contact.

Calculation of response rate:

$$\text{Total response rate} = \frac{\text{Total number of responses}}{\text{Total number in sample - ineligible}}$$

Calculation of active response rate

$$\text{Active response rate} = \frac{\text{Total number of responses}}{\text{Total number in sample - (ineligible + unreachable)}}$$

Calculate Actual Sample Size from Estimated Response Rate

The sample size must be big enough to provide the necessay confidence in you data i.e. as a representative of the population.

First estimate the response rate from the sample and then calculate to increase the size of the sample using the formula:

$$n^2 = \frac{n \times 100}{re\%}$$

n^2 = actual sample size
n = minimum sample size
re% = estimated response rate in percentage

Example:
John wanted to make a survey of customers and calculated that an adjusted minimum sample size of 429 was required. He estimated the response rate would be 30%. What is the actual sample size John should have?
Applying the formula:

$$n^2 = \frac{n \times 100}{re\%}$$

$$= \frac{439 \times 100}{30}$$

$$= 1463$$

John's actual sample size should be 1463.

It is important to know the response rate when selecting the sampe size. E.g. using post survey, it is found that only 30% response rate is achievable and for interview only 50%.

6.4 Sampling Designs (Techniques)

There two types of sampling designs:
1. Probability sampling - the persons in the population have the known chance or probability of being selected. This design is used when the representativeness of the sample is of importance in the interests of widdr generalizability.
2. Nonprobability sampling - not every person in the population has the equal chance of being selected. This design is not for generalizability but critical for other factors.

Each of these two designs has different sampling strategies.

Probability Sampling
Types:
1. Simple random
2. Systematic
3. Stratefied random

4. Cluster

1. Simple random - Every person in the population has a 'known-and-equal' chance of being selected to make up a sample. E.g. there are 1000 people in the population and we need a sample of 100 persons. We can pieces of paper, each bearing the name of one of the persons into a hat and we can draw 100 of them from the hat. Each of those names has a 100/1000 chance of being drawn. The probability of any one of them being chosen as a subject is 1 and that everyone has the same or equal probability of being chosen. The sampling design is known as ***simple random sampling***. It has the least bias and offers the most generalizability. However, it is cumbersome and expensive and updating of the listing of the population is not possible. As a consequence complex probability sampling designs are used such as:

1. Systematic sampling
2. Stratified random samplinng
3. cluster sampling
4. Area sampling
5. Double sampling

2. Systematic sampling - This involves drawing every nth element in the population starting with a randomly chosen element between 1 and n.
E.g. There are 260 houses and we want to have 35 houses for a sample. We can draw every 7^{th} house starting from a random number from 1 to 7. Let us say that the random number is 7, then houses numbered 7, 14, 21, 28 and so on would be sampled until the 35 houses are selected.
However this systematic sampling design may suffer from systematic biasness. E.g. every 7^{th} house is a corner house.
For market surveys, consumer attitude surveys and the like, the systematicsampling design is often used.

3. Stratified random sampling - there are subgroups of elements existing within the population. E.g. the HR manager wants to train everyone employee in the organisation. Within the population of the organisation there different levels of people such as top management managers, middle-level managers and lower-level managers, first live supervisors, computer analysts, clerical workers and so on. The training requirements for different levels will be different. Data will need to be collected from each subgroup level in the population to assess the needs at each subgroup. In such a situation, stratified random sampling is used.
This stratified random sampling is used to study buying habits of customers on the basis of life stages, income levels or geographical areas. Stratification ensures homogeneity within each subgroup or stratum.

4. Cluster sampling - intact grops, not individuals are randomly selected. All the members of selected groups hhave similar characteristics. E.g. instead of randomly selecting all fifthgraders in a large school district, you can randomly select fifthgrade classrooms and use all the students in each classroom.
Cluster sampling is more convenient when the population is very large or spread out over a wide geographical area.

It may be the only possible method of selecting a sample when the researcher is unable to obtain a list of all members of the population.

Any location within which we find an intact group of similar characteristics is a cluster. Examples of clusters are classrooms, schools, city blocks, hospitals and department stores.

Benefits of cluster sampling:

1. Involves less time and expense
2. Generally it is more convenient than other techniques.
3. It is easier to get permission to work with all members (students) in several classrooms than to work with a few students in many classrooms or to survey all the people in a limited number of city blocks than a few people in many city blocks.

Drawbacks of Cluster Sampling

1. The chances are greater of selecting a sample that is not representative of the population.
2. The smaller the sample size, the more likely that the sample selected may not represent the population.
3. Cluster samples offer more heterogeneity within groups and more homogeneity among groups.
4. Cluster sampling lends itself to greater biases and is the least generalizable of all the probability sampling designs (techniques) because most naturally occurring clusters in the organisational context do not contain heterogeneous elements. In other words, the conditions of intracluster heterogeneity and inter-homogeneity are often not met. For this reason cluster sampling is not well accepted in organisational research.

Single Stage and MultiStage Cluster Sampling

Cluster sampling can also be done in several stages and is then called ‘multistage cluster sampling. E.g. A national survey is to be done to determine the monthly bank deposits. Cluster sampling would first be used to select the urban, semiurban and rural geographical locations for study.

At the next stage particular areas in each of these locations would be chosen.

At the third stage banks in each area would be chosen

Nonprobability Sampling

In nonprobability sampling designs,the members in the population do not have the known and equal chance of being selected as sample subjects. This implies that the findings from the study of the sample cannot be confidently generalized to the population. This is the reason for using the nonprobability sampling designs. However, some nonprobability sampling designs are more dependable than others and could offer some important leads to potentially useful information with regard to the population.

Types:

1. Quota sampling
2. Purposive sampling
3. Snowball sampling
4. Self-selection sampling

5. Convenience sampling

1. Quota Sampling

It is a form of proportionate stratified sampling in which a predetermined proportion of people are sampled from differentgroups, but on a convenience basis.

E.g. It is assumed that the work attitudes of blue-colar workers are quite different from those of white-collar workers. If there are 60% blue-colar workers and 40% white-collar workers in the organisation and that a total of 30 people are to be interviewed to find the answer to the research question. Then a quota of 18 blue-collar workers and 12 white-collar workers will form the sample because these numbers represent 60 and 40 percent of the sample size.

It is obvious that the sample will not be totally representative of the population and therefore the generalizability of the findings will be restricted.

However it is a convenience that the type of sampling technique can provide in terms of effort, costs and time. This technique becomes necessary when a subset of the population is underrepresented in the organisation e.g. minority groups, foremen and so on. Its good point is that it ensures that all the subgroups in the population are adequately represented in the sample. It also indicates that *quota samples are basically stratified samples from which subjects are selected nonrandomly.*

As the workplace or the society becomes more heterogeneous because of the changing demographics, quota sampling can be expected to be used more frequently in the future. E.g. quota sampling can be used to have some idea of the buying behaviour of various ethnic groups, for getting an understanding as to how employees from different nationalities perceive the organisational culture and so on.

Although quota sampling is not as generalizable as stratified random sampling, it does offer some information based on which further investigation can be done. This means that at the first stage of research the nonprobability design of quota sampling can be used and once some information has been attained, a probability design will follow. The opposite approach is also possible, by starting the research with probability sampling design and from the information received a new area for researh is indicated and a nonprobability sampling design might be used to explore further findings.

2. Purposive sampling

This approach is to obtain from specific target groups. Here, the sampling is confined to specific types of people who can provide the desired iinformation, either because they are the only ones who possess it or conform to some criteria set by the researcher. The two major types of purposive sampling are judgement sampling and quota sampling.

In judgement sampling, the choice of subjects are those who are in the best position to provide the information required. E.g. if a research wants to find out what it takes for women mangers to make it to the top, the only people who can provide firsthand information are the women who have risen to the positions of presidents, vice presidents and important top-level executives in work organisations. Therefore this sampling design is used when a limited number or category of people have the information that is sought. Obviously judgement sampling may curtail the generalizability of the findings. It is, however, the only viable sampling method for obtaining the type of information that is required from very specific pockets of people who alone possess the needed facts and can give the information sought. It requires the researcher to locate them.

3. Snowball sampling

This is normally used when it is difficult to identify members of the desired populationl e.g. people who are workingwhile claimingunemployment benefit. You need to:

1. Make contact with one or two persons in the population.
2. Ask these persons to identify further people.
3. Ask these new persons to identify further new people and so on.
4. Stop wheneither no new ones are givenor the sampe is as large as is manageable.

The main problem in the sampling method is making initial contact. It is used when there are difficulty to identify the people for the sample.

4. Self-selection sampling

This occurs when you allow an individual to identify his/her desire to take part in the research. You therefore:

1. Publicise your need for cases, either by advertising through appropriate media or by asking them to take part.
2. Collect data from those who respond.

Cases that self-select often do so because of their feelings or opinionsabout the research question(s) or stated objectives.

E.g. Sila's research was concerned with teleworking. She had decided to administer hor questionnaire using the Internet. She published her research on a range of bulletin boards and through the teleworkers' association asking for volunteers to fill in a questionnaire. Those who responded were sent a short questionnaire by email.

5. Convenience sampling

Convenience or haphazard sampling involves selectinghaphazardly those cases that are easiest toobtain for your sample, such as the person interviewed at random in a shopping centre for a television programme. The sample selection process is continued until your required sample size has been reached. This technique may be widely used but it is prone to bias and influneces that are beyond your control, as the individuals only appear in the sample because of the ease of obtaining them. Often the sample is intended to represent the total population, for example managers taking an MBA course as a surrogate for all managers! In such instances the choice of sample is likely to have biased the sample, meaniing that subsequent generalisations are likely to be at best flawed. These problems are less important where there is little variation in the population, and such samples often serve as pilots to studies using more structured samples.

SELF CHECK 6

1. Identify a suitable sampling frame for each of the following research questions:
   a. How do company directors of manufacturing firms of over 500 employees think a specified piece of legislation will affect their companies?
   b. Which factors are important in accountants' decisions regarding working in mainland China?
   c. How do employees at MAY Bank in Nilai think the proposed introduction of compulsory Saturday working will affect their working lives?
2. Identify the relevant population for the following research foci, and suggest the appropriate sampling design to investigate the issues, explaining why they are appropriate. Wherever necessary, identify the population frame as well.
   a. A gum manufacturing firm would like to know the types of guns possessed by various age groups in Kuala Lumpur.
   b. Hospital administrator wants to find out if the single parents working in the hospital have a higher rate of absenteeism than parents who are not single.
   c. A researcher would like to assess the extent of pilferage in the materials storage warehouses of manufacturing firms in Selangor.
   d. The director of human resources wants to investigate the relationship between drug abuse and dysfunctional behaviour of blue-collar workers in a particular plant.
3. a. Explain why cluster sampling is a probability sampling design.
   b. What are the advantages and disadvantages of cluster sampling?
   c. Describe a situation where you would consider the use of cluster sampling.
4. Use of a sample of 5000 is not necessarily better than using one of 500. How would you react to this statement?

   Nonprobability sampling designs ought to be preferred to probability sampling designs in some cases. Explain with examples.

References

- **Wilson: chapter 8 and Saunders: chapter 6**
- **Note: Lecture 6: Selecting Samples (Sampling)**

Topic 7: Using Secondary Data

Learning Outcomes

By the end of this topic, you will be able to:

1. Justify the use of the secondary data must meet the needs of the research questions and objectives.
2. Justify the credibility of such data to be applied in the research must enhance the quality of the research.
3. Discuss the secondary data provide the information to match the population or sample in the research.

Introduction

The keys areas of learning in this chapter are:

- These secondary data are in existence and do not required to be collected by the researcher.
- Their advantages – easily available, save money and time and are important when primary data could be obtained.
- The use of secondary data must be subjected to review of their suitability, quality and representativeness.
- The reliability and validity of the secondary data must be related to the source and the methods used to collect them.
- Other consideration of their usage are measurement bias and the costs and benefits.

7.1 Secondary Data

They are data that already exist and do not have to be collected by the researcher. Examples of secondary data: statistical bulletins, government publications, information published or unpublished and available from either within or outside the organisation, data available from previous research, case studies and library records, on-line data, web-sites and the internet.

Advantages of Secondary Data

1. They are available.
2. Save time and money.
3. Essential in instances when data simply cannot be obtained using primary data collection procedures.

Disadvantages of Secondary Data

1. May not match your need.
2. Access can be costly or difficult.
3. Aggregations and definitions may be unsuitable.
4. No real control over data quality.
5. Initial purpose may affect how data are presented.

Evaluating secondary data sources
You need to be sure:
1. The secondary data will enable you to answer your research question(s) and to meet your objectives.
2. The benefits associated with their use will be greater than the costs.
3. You will be allowed access to the data.

Overall suitability
Pay attention to measurement validity and coverage including unmeasured variables.

Measurement validity - the secondary data provide you with the information that you need to answer your research question(s) or meet your objectives. Sometimes the measures used may not match those that you need then the secondary data are invalid.

Coverage and unmeasured variable - ensure that the secondary data cover the population about which you need data, for the time period you need, and contain data variables that will enable you to answer your research question(s) and to meet your objectives.
Coverage will concern two issues:
1. Ensuring that unwanted data are or can be excluded.
2. Ensuring that sufficient data remain for analysis to be undertaken once unwanted data have been excluded.

Precise suitability
This refers to the secondary data being reliable and valid and no biasness.
Reliability - is assessing the reputation/authority of the source.
Validity - refers to the consistency and accuracy of the data, the method by which the data are collected and the precision needed by the original collector of the data. This could relate to the sampling technique used, the sampling error associated with it and the response rates.
The reliability and validity of secondary data are related to the method by which the data are collected and the source.
The source from which the secondary data re obtained can be assessed by its authority or reputation. Well known organisations are noted for their reliability and trustworthiness because their existence is dependent on their credibility.
The reliability and validity of secondary data are dependent on a detailed assessment of the method(s) used to collect the data. The following aspects should be examined:
1. Who were responsible for collecting or recording the information and the context in which the data were collected?

2. What are the possible error/biases?
3. The process by which the data were selected and collected.
4. The sampling design used to extract the data and the associated sampling error and response rates. E.g. survey gives high response rate and is more reliable than low response rate. For dairies, transcripts of interviews or meetings are unlikely to have formal methodology of how the data were collected and therefore reliability would be difficult to assess. Letters and memos are personal point of view and should only be considered as writer's perceptions and view but not an objective account of reality. In other words, reliability and validity of secondary data will be easier to assess, if the methodology used to collect the data is explained, the sampling technique used and the response rates are explained.
5. Where data are compiled as in a report, pay particular attention to how these data are analysed and how the results are reported. If percentages are given without the totals being given, examine the data very carefully.
6. Similarly where quotations of figures or data are given without supporting evidence of references or authorities again you should be very careful. These data may not be reliable.

7.2 Measurement bias

This can occur for two reasons:
1. Deliberate or intentional distortion of data.
2. Changes in the way data are collected.

In organisations, distortion of records occurred deliberately by managers and employees. However, measurement bias resulting from deliberate distortion is difficult to detect. This can be detected by triangulate the findings i.e. cross-checking verification, with other independent data sources. Where two or more conclusions are similar you can be confident that the data are not distorted.
Changes in which the data are collected can also cause distortion of data. There is no consistency in the collection of the data. Once the method is changed to collect the data, the bias also changes. You have to discover the way data are recorded has changed.

7.3 Costs and Benefits

According to Kervin (1999) the final criterion for assessing secondary data is a comparison of the costs of acquiring them with the benefits they will bring.
Costs include time and financial resources that are needed to obtain the data. Some data can be obtained free while others may need to be paid to get them. Data from consultant organisations may be costly to obtain.
Benefits may be assessed in terms of the extent to which data will enable you to answer your research question(s) and meet your objectives. The benefits can be assessed from the overall and precise suitability. This assessment is summarised in the checklist below.

Checklist to evaluate secondary data sources
Overall suitability ✓ Does the data set contain the information you require to answer your research question(s) and meet your objectives? ✓ Do the measures used match those you require? ✓ Is the data set a proxy for the data you really need? ✓ Does the data set cover the population that is the subject of your research? ✓ Can data about the population that is the subject of your research be separated from unwanted data? ✓ Are the data sufficiently up to date? ✓ Are data available for all the variables you require to answer your research question(s) and meet your objectives?
Precise suitability ✓ How reliable is the data set you are thinking of using? ✓ How credible is the data source? ✓ Is it clear what the source of the data is? ✓ Is the source of the data likely to be reliable? ✓ Do the data have an associated copyright statement? ✓ Do associated published documents exist? ✓ Is the methodology clearly described? ✓ If sampling was used what was the procedure and what were the associated sampling errors and response rates? ✓ Who were responsible for collecting or recording the data? ✓ (For surveys) Is a copy of the questionnaire or interview checklist included? ✓ (For compiled data) Are you clear how the data were analysed and compiled? ✓ Are the data likely to contain measurement bias? ✓ What was the original purpose for which the data were collected? ✓ Who was the target audience and what was their relationship to the data collector or compiler (were there any vested interests)? ✓ Have there been any documented changes in the way the data are measured or recorded including definition changes? ✓ How consistent are the data obtained from this source when compared with data from other source? ✓ Are you happy that the data have been recorded accurately?
Costs and benefits ✓ What are the financial and time costs of obtaining these data? ✓ Have the data already been entered into a computer? ✓ Do the overall benefits of using these secondary data sources outweigh the associated costs?

SELF CHECK 7

1. Give three examples of different situations where you might use secondary data as part of your research.
2. Explain the need to evaluate the use of secondary data.
3. What factors should be considered to ensure the reliability and validity of the secondary data?
4. What are the possible ways by which data can be distorted in organisations to create biasness?
5. How can the distortion of records in organisations be detected?

References

- **Wilson: chapter 7 and Saunders: chapter 7**
- **Note: Lecture 7 Using Secondary Data**

Topic 8: Collecting Primary Data through Observation

Learning Outcomes

By the end of this topic, you will be able to:

- Explain the observation is used to collect data about people's behaviour or actions and it does not involve asking question or communicating with others.
- Explain the data gathered in the inductive approach are descriptive in nature and a narrative account is evolved which eventually leads to the development of a theory (the central theme/phenomenon).

Introduction

The key areas of learning in this chapter are:

- One way to gather data about what people do or act is to watch them in a systematic manner.
- Essentially observation is the systematic approach of observing, describing, analysing and interpreting people's behaviour and there is no questioning or communicating with them.
- Collection of data by observation can be done in two possible ways:
  - participation observation in qualitative research approach.
  - structured observation in quantitative research approach
- There are factors that determine the application of observation as a research strategy to gather data.
- Data collected from observation are those that could not be obtained from interview.
- In qualitative research, data collected from observation are basically descriptive in nature and become the basis for development a framework of theory to understand and explain the happenings in a research environment.
- In a structured observation as in a quantitative research, data are collected based on quantitative aspect such as frequency of occurrence.

8.1 Observation

If your research question(s) and objectives are about what people do, an obvious way in which to discover this is to watch them do it. This is essentially what observation involves: the systematic observation, recording, description, analysis and interpretation of people's behaviour. There is no questioning or communicating with people.

There are two types of observation:

1. Participant observation which is a qualitative approach of discovering the meanings that people attach to their actions. This involves the researcher participating in the activities of the people.
2. Structured observation which is a quantitative approach and is more concerned with the frequency of those actions.

Participant Observation (a qualitative approach)
The researcher attempts to participate fully in the lives and activities of subjects and becomes a member of that group, organisation or community. This enables the researcher to share their experiences by not merely observing what is happening but also feeling it. It is not possible in many instances to understand the behaviour of people, and participant observation then serves the purpose to discover those delicate nuances of meaning. (Nuance = a very slight difference in meaning.) This implies that participant observation can help a researcher to get a more meaningful understanding of the human interactions in a group, organisation or community. It is attempting to find the meaning behind an interaction between or among the people.

Situations in which participant observation has been used
Whyte (1995) lived among a poor American-Italian community and he used participant observation to understand the behavioural pattern of this 'street corner society'.
Roy (1952) worked in a machine shop for 10 months as an employee. He wanted to understand how and why his 'fellow workers' operated the piecework bonus system.
Rosen (1991) worked as a participant observer in a Philadelphia advertising agency. He wanted to understand how organisations used social drama to create and sustain power relationships and social structures.

8.2 Researcher Roles in Participant Observation

There are four roles identified by Gill and Johnson (1997):

1. Complete participant
2. Complete observer

} The identity of the researcher is not revealed.

3. Observer as participant
4. Participant as observer

} The identity of the researcher is revealed.

The first two of these roles, the complete participant and complete observer, the researcher's identity is not revealed. This is to overcome the problem of the research subjects putting up a different behaviour or not really want to cooperate with you in your research.
The second two, observer as participant and participant as observer, entail the researcher revealing his purpose to those with whom he is mixing in the research setting.

Complete Participant Role

The researcher attempts to become a member of the group in which he is performing research. His identity is not revealed to enable him to discover the true meaning of the group member behaviour in order to obtain answers to the research questions and to attain the research objectives.

There may a question of ethics. The researcher might be accused of spying on people who have become his friends as well as colleagues. They may learn to trust him with information that they would not share were they to know his true purpose. The researcher has to avoid such a situation and not to lose sight of his research purpose.

Complete Observer Role

Here the researcher does not take part in the activities of the group. E.g. you may be a customer in a supermarket, but you act as a complete observer to study the behaviour of other customers in the supermarket.

Observer as Participant Role

You may act as an observer in a team building course you attended. Your identity as a researcher is clear to all concerned.

Participant as Observer Role

The identity of the researcher is made known to the members concerned in the research setting. This approach is useful in a situation for assessing the experience of the participants such as a fieldwork relationship. The revelation is necessary to gain the trust of the members. E.g.

Your research project is concerned with you wishing to understand what the experience of a one-day assessment centre is like for the graduates who attend these as prospective employees.

You decide that there is no better way of doing this than 'getting in on the action' and being a guinea pig. You negotiate access with the company managers who are running the assessment centre. You also explain your research to the graduates who are there 'for real'. You become involved in all the activities and speak to as many of your fellow graduates as possible in order to discover their feelings about the experience. At the end of the day you are utterly exhausted!

8.3 Factors that may determine the choice of participant observer role

1. The purpose of your research.
2. The time you have to devote to your research.
3. The degree to which you feel suited to participant observation.
   Not everyone is suited to this type of research. It relies very much on building relationships with others and a certain amount of personal flexibility is required. As a participant observer you have to be 'all things to all people'. Your own personality must be suppressed to a greater extent. This is not something with which you may feel comfortable.
4. Organisational access.

5. Ethical considerations

8.4 Data collection and analysis in participant observation

The data generated by participant observation may be:

1. Primary observations - statements where you would note what happened or what was said at the time. Keeping a diary is a good way of doing this.
2. Secondary observations - are statements by observers of what happened or was said. These are observers' interpretations.
3. Experiential data - data from your perceptions and feelings as you experience the process you are researching. Keeping a diary of these perceptions and feelings proves a valuable source of data when the time comes to write up your research.
4. Collect data on factors material to the research setting for example, roles played by key participants and how they may have changed, organisational structures and communication patterns.

Data collection

The data collected are those that could not be obtained at the interview. Questioning and observation can come together. First you can ask the informant to clarify the situations you have observed and second to yourself to clarify the situation and the accounts given of the situations which you have observed earlier.

The data in qualitative research is basically descriptive in nature classified as 'descriptive observation' and 'narrative account'.

In descriptive observation you may concentrate on observing the physical setting, the key participants and their activities, particular events and their sequence and the attendant processes and emotions involved. This description may be the basis for your writing of a narrative account. You have to go on and develop a framework of theory that will help you to understand and to explain to others what is going on in the research setting you are studying.

How you record your data is important. You must be 'open' to more possibilities so that you are able to make more notes at the time the events is being observed or reported. One golden rule is: recording must take place on the same day as the fieldwork in order that you do not forget valuable data.

Data analysis

In participant observation research your data collection and analysis activity may be part of the same process. This means you will be carrying out analysis and collection of data simultaneously. From the analysis you may come up with 'promising lines of enquiry' that you might want to follow up in your continued observation. At the end of the analysis of the data collected you should be able to come out with a theory to help you to understand 'what is going on'. This approach is known as analytic induction.

Threats to reliability and validity

Participant observation is very high on ecological validity because it involves studying social phenomena in their natural environment. It is also subjected to threats of validity because of

the situations in which the research is done being specific circumstances and conditions. As a consequence, there is little generalizability.
The threat of reliability could come from the observer bias. It is difficult to avoid such bias but should try to minimise it. This is because we have our own perceptions to colour our interpretation of what we believe to be true.
One possible way is to use triangulation to improve the reliability of the conclusions reached. Information for triangulation can come from the data collected from interviews and questionnaires from the participants.

Advantages and Disadvantages of participant observation

Advantages

1. It is good at explaining 'what is going on' in particular social situation.
2. It heightens the researcher's awareness of significant social processes.
3. It is particularly useful for researchers working within their own organisations.
4. Some participant observation affords the opportunity for the researcher to experience 'for real' the emotions of those who are being researched.
5. Virtually all data collected are useful.

Disadvantages

1. It can be very time consuming.
2. It can pose difficult ethical dilemmas for the researcher.
3. There can be high levels of role conflict for the researcher (e.g. colleagues versus researcher).
4. The closeness of the researcher to the situation being observed can lead to significant observer bias.
5. The participant observer role is a very demanding one, to which not all researchers will be suited.
6. Access to organisations may be difficult.
7. Data recording is often very difficult for the researcher.

Structured Observation (quantitative approach)

Differences between Unstructured Observation and Structured Observation	
Unstructured Observation	Structured Observation
1. Does not have a predetermined structure. 2. Answers to the questions and objectives come from observing the activities of the participants. 3. Researchers can get attached to the research setting. 4. Researchers' findings depend on the responsiveness of the participants. 5. Results of the findings may not be	1. Systematic approach to observation and with a predetermined structure. 2. Data collection is based on quantitative aspect such as frequency of occurrence.

generalised.	

Situations in which structured observation may be used

It is like a 'time and motion' study such as in the production areas or the study of how managers go about planning, leading and controlling activities in a certain period of time.

Advantages and Disadvantages

Advantages

1. It can be used by anyone after suitable training in the use of the measuring instrument. Therefore you can delegate this extremely time-consuming task. In addition, structured observation may be carried out simultaneously in different locations. This would present the opportunity of comparison between locations.
2. It should yield highly reliable results by virtue of its replicability.
3. Structured observation is capable of more than simply observing the frequency of events. It is also possible to record the relationship between events. For example, is the visit to the retail chemist's counter to present a prescription preceded by an examination of merchandise unrelated to the prescription transaction?
4. The method allows the collection of data at the time they occur in their natural setting. therefore there is no need to depend on 'second-hand' accounts of phenomena from respondents who put their own interpretation on events.
5. Structured observation secures information that most participants would ignore because to them it was too mundane or irrelevant.

Disadvantages of structured observation

1. The observer must be in the research setting when the phenomena under study are taking place.
2. Research results are limited to overt action or surface indicators form which the observer must make inferences.
3. Data are slow and expensive to collect.

Data collection and analysis in structured observation

Before you start the structured observation, you have to decide to use an 'off the shelf' coding schedule (like a standard format) or develop your own coding schedule.

E.g. an 'off-the-shelf' coding schedule: Recording sheet for observing behaviour in groups

Name of group:
Nature of activity:
Date: Name of observer:
Initial arrangement of group:

C D
B E
A F

	Name of group members (or reference letters)					
	A	B	C	D	E	F
Taking initiative - e.g. attempted leadership, seeking suggestions, offering direction						
Brainstorming - e.g. offering ideas or suggestion, however valid.						
Offering positive ideas - e.g. making helpful suggestions, attempting to problem-solving						
Drawing in others - e.g. encouraging contributions, seeking ideas and opinions						
Being responsive to others - e.g. giving encouragement and support, building on ideas						
Harmonising - e.g. acting as peacemaker, calming things down, compromising						
Challenging - e.g. seeking justification, showing disagreement in a constructive way						
Being obstructive - e.g. criticising, putting others down, blocking contribution						
Clarifying/summarising - e.g. linking ideas, checking progress, clarifying objectives or proposals						
Performing group roles - e.g. spokesperson, recorder, time-keeper, humourist						
Other comments						

Guidelines for developing your own coding schedule

Attribute	Comments
Focused	Do not observe and record all that is going on. Concern yourself only with what is strictly relevant
Unambiguous	Therefore requiring the absolute minimum of observer interpretation
Non-context dependent	The observer's job is more difficult if the coding of behaviours is dependent o the context in which the behaviour occurs. It may

	be essential for your research question(s) and objectives to record contextual data, but this should be kept to a minimum
Explicitly defined	Provide examples for the observer (even if this is you) of behaviours that fall into each category and those that do not
Exhaustible	Ensure that it is always possible to make a coding for those behaviours you wish to observe
Mutually exclusive	Ensure that there is no overlap between behaviour categories
Easy to record	The observer must be able to tick the correct box quickly without having to memorise appropriate categories

Data analysis

How complex is your analysis of the data will depend on your research question(s) and objectives. If you used for example the off-the -shelf schedule as shown above to record you observations, you can establish the amount of interactions by category in order to relate the results to the output of the meeting.

Alternatively you may use the above schedule to see what patterns emerge. It may be that the amount of interactions varies by the nature of the group or its activity or that seating position is associated with the amount of contributions.

Threats to validity and reliability

The main threats are ones of reliability such as:

1. Subject error - it may cause the data to be unreliable. This is dependent on who has been chosen for the research study.
2. Time error - this may be due to the fact that the time at which you conduct the observation does not provide data that are not typical of the total time period in which you are interested. A different output result occurred.
3. Observer effect - this is the result of a change in the behaviour of the participant you were observing especially when the subject becomes aware of being observed.

SELF CHECK 8

1. What are the two types of observation you can use to obtain primary data?
2. For qualitative research what type of observation would be more appropriate and why is it so?
3. What types of role a researcher can play in unstructured observation to gather the research data?
4. Explain why generalizability of the findings in unstructured observation is difficult to apply to the population.
5. What are the differences between unstructured observation and structured observation?
6. What are the possible threats to the reliability and validity faced by participant observation?

References

- Wilson: chapter 6 and Saunders: chapter 8
- Note: Lecture 8: **Collecting Primary Data Through Observation**

Topic 9: Collecting Primary Data using Interviews

Learning Outcomes

By the end of this topic, you will be able to:

- Justify the usefulness of interviews is to fill the gaps of missing data or to obtain further or deeper information about the behaviour of participants.
- Discuss the different purposes for the use of interviews are applied in qualitative research and in quantitative research.

Introduction

The key areas of learning in this chapter are:

- Interviews can help to obtain valid and reliable data relevant to the research questions and objectives.
- Interviews can be done in three possible ways: structured, semi-structured and unstructured.
- Interviews are used in exploratory research, descriptive research and explanatory research.
- Interviews are most frequently used in qualitative research in order to obtain in-depth information about the behaviour of the participations in a society or community.
- Data obtained through interviews in inductive research may not have generalizability or external validity, however, it has internal validity.
- In deductive research, interviews help to obtain data necessary to validate findings from the use of questionnaires.

9.1 Interview

What is an interview?

It is a purposeful discussion between two or more people.
The use of interviews can help you to gather valid and reliable data that are relevant to your research question(s) and objectives. If you have not established a research question(s) and objectives, an interview can help you to achieve this.

Types of Interview

1. Structured interviews - where questionnaires are used and they are based on a predetermined and standardised or identical set of questions. You read out the question and record the response on a standardised schedule.
2. Semi-structured interviews - the researcher has a list of themes and questions to be covered during the interview, although these may vary from interview to interview. This

means you may decide to omit some questions in particular interviews depending on the circumstances of the situation.

3. Unstructured interviews - are informal interviews. You may use these interviews to explore the depth of a general area in which you are interested. These are referred to as ***in-depth interviews***. There is no predetermined list of questions to work through in this situation. It is subject to what you would like to explore in the area related to your research study. The interviewee is allowed the opportunity to talk freely about events, behaviour and beliefs in relation to the topic area. This type of interview is sometimes called ***non-directive*** or ***informal interview***. This is different from ***respondent interview*** where the interviewer directs the interview and the interviewee responds to the questions of the researcher.

Interviews can be conducted on a face-to-face with the participant or with more than one person or a group. It could also be done by telephone.

Links to the purpose of research and research strategy

Each type of interview has a different purpose. Each type is used to gather data which will then be the subject of quantitative analysis.

In qualitative research, interviews are used to conduct discussions not only to reveal and understand the 'what' and the 'how' but also to place more emphasis on exploring the 'why'. Research can be classified into: exploratory, descriptive and explanatory. Interviews are specifically used in each of these three research categories:

1. In an exploratory study, in-depth interviews can be very helpful to 'find out what is happening and to seek new insights'. Semi-structured interviews may also be used in relation to an exploratory study.
2. In descriptive studies, structured interviews can be used as a means to identify general patterns.
3. In an explanatory study, semi-structured interviews can be used in order to understand the relationships between variables, such as those revealed from a descriptive study. Structured interviews may also be used in relation to an explanatory study.

These are summarised as follows:

	Exploratory	Descriptive	Explanatory
Structured		**	*
Semi-structured	*		**
In depth	**		

* = less frequent ** = more frequent

In a quantitative research more than one type of interview may be used moving from structured to unstructured (in-depth) interviews. Interviews can be linked up with the questionnaires they can act as a means to validate findings from the use of questionnaires. The key point to consider is the need for consistency between the research question and objectives, the strategy to be employed and the methods of data collection used - their fitness for purpose.

9.2 Situations favoring Qualitative Research Interviews

1. The nature of the approach to research.
   Where it is necessary to understand the reasons for the decisions that the research participants taken or to understand the reasons for their attitudes and opinions, it will be necessary to conduct a qualitative interview. Semi-structured or in-depth interviews will be used in order to get the interviewees to explain, or build on, their responses. Such interviews may lead to opportunities to new areas not previously considered but which are significant for better understanding and to address the research questions and objectives. The result is that you are able to collect a rich and detailed set of data.

2. The significance of establishing personal contact.
   Participants are more agreeable to interview than to complete a questionnaire. An interview provides the opportunity to reflect on events without needing to write anything down. It also provides the opportunity for interviewees to receive feedback and personal assurance about the way in which information will be used. Interviews provide great response rate compared to questionnaires. Furthermore interviewers have better control over who answers the questions.

3. The nature of the data collection questions.
   An interview will be the most advantageous to attempt to obtain data in the following circumstances:
   - where there are a large number of questions to be answered;
   - where the questions are either complex or open-ended;
   - where the order and logic of questioning may need to be varied.

   Semi-interview or in-depth interview will be most appropriate for the latter two types of situations.

4. Length of time required and completeness of the process.
   It is a wise move to discuss with the participants of the time required to complete the interview. Once they accepted the time requirement, they will agree to be interviewed. It is also appropriate to arrange a time suitable with the participants for the interview to take place.
   The aim of an interviewer is to obtain answers to all the questions asked in the interview. However you have to allow the right of the participant not to answer any of the questions and consideration to modify the question to get a response. This must take into consideration the ethical aspect as well.

Data quality issues and how to overcome them

There are a number of issues associated with qualitative research interviews:

1. Reliability
2. Forms of bias

3. Validity and generalizability

Reliability - whether alternative researchers would reveal similar information. It is also related to the issues of bias. The interviewer bias - where the comments, tone or non-verbal behaviour of the interviewer create bias in the way that interviewees respond to the questions being asked. Another aspect of bias is the interviewee or response bias - this may due to the perceptions the interviewee has on the interviewer. The interviewee chooses not to reveal the real thing. However there can be other factors that can cause this interviewee bias such as the sensitive information, the sample chosen and timing factor.

Validity may not be such an issue. Validity refers to the extent to which the researcher gains access to the participants' knowledge and experience, and is able to infer a meaning that the participant intended from the language that was used by this person. The high level of validity that is possible in relation to qualitative interviews is made clear by the following quotation:

> The main reason for the potential superiority of qualitative approaches for obtaining information is that the flexible and responsive interaction which is possible between interviewer and respondent(s) allows meanings to be probed, topics to be covered from a variety of angles and questions made clear to respondents.
>
> (Sykes, 1991:8, cited by Healey and Rawlinson, 1994:132)

However qualitative research using semi-structured or in-depth interviews will not be able to be used to make generalisations about the entire population where this is based on a small and unrepresentative member of cases. This will be the situation in a case study approach (Yin, 1994).

Overcoming data quality issues

Reliability - the findings derived from using non-standardized research methods are not necessarily intended to be repeatable since they reflect reality at the time they were collected, in a situation which may be subject to change. Besides the circumstances to be explored are complex and dynamic. The situation demands flexibility to explore the complexity of the research question(s) and objectives. Therefore any attempt to ensure that qualitative non-standardised research could be replicated by other researchers would not be realistic or feasible without understanding the strength of this type of research.

How to overcome bias in qualitative interviews

1. The interviewer must be well prepared for the interview. This means that he must have the level of knowledge of his research topic in order to establish his credibility in the view of his research participant. This knowledge may be acquired through a review of the literature.
2. Credibility may also be promoted through the supply of relevant information to the participants before the interview.
3. The appearance of the researcher may affect the perception of the interviewee. Dressing must be acceptable to the setting within which the interview is to occur.
4. Where the interviewee has not met you before, the first few minutes of conversation may

have a significant impact on the outcome of the interview - again related to the issue of credibility and the level of the interviewee's confidence.

5. When conducted appropriately, your approach to questioning should reduce the scope for bias during the interview and increase the reliability of the information obtained. Questions should avoid too many theoretical concepts or jargon since your understanding of such terms may vary from that of the interviewees.
6. Appropriate behaviour by the researcher should reduce the scope for bias during the interview. Comments or non-verbal behaviour such as gestures, which indicate any bias in your thinking should be avoided.
7. You must demonstrate paying attention and listening skills to the participants to avoid bias.
8. Able to summarise your understanding of the information provided by the interviewee. It is then repeated to the interviewee to ensure correctness as well as to avoid biasness.
9. A full record of the interview should be compiled as soon as possible to avoid any bias.
10. You must be able to recognise cultural differences and to overcome issues of bias.

The Researcher's interviewing competence

It involves in areas like:

- opening the interview
- using appropriate language
- questioning - open questions; probing questions (why, how & what); specific and closed questions (a yes/no answer type).
- listening
- testing and summarising understanding
- behavioural cues
- recording data

Recording information

The need to create a full record of the interview soon after its occurrence was identified as a means to control bias and to produce reliable data for analysis. It is an easy job. It demand is more than just to take notes of what has been said but to be able to comprehend the meaning of what is being said as well. It needs time to complete the notes and to be with the time the information is being revealed by the interviewee. One option is to tape-record the interview. You may meet an interviewee who prefers you to take down the notes. If the interview is being recorded, you still have to make a transcript of it as soon as possible and to complete the record of the interview you have made earlier.

Group interview

The researcher acts as a facilitator in a group interview. You will explain the theme of the discussion to the group, make the members relax and then initiate the discussion. .A question is posted to the group. The group members are allowed to discuss and present their information freely. You have then to manage the group carefully. Group interaction may lead to a highly productive discussion as interviewees respond to the question and may

reveal data that provide you with important insights. You must be careful not to allow a few to dominate the discussion and ensure that everyone has a chance to present his idea on the question.

SELF CHECK 9

1. What type of interview would you use in each of the following situations?
   a. a market research project?
   b. a research project seeking to understand whether trade union attitudes have
      changed?
   c. following the analysis of a questionnaire?
2. What are the advantages of using semi-structured and in-depth interviews?
3. During a presentation of your proposal to undertake a research project, which will be qualitatively based using semi-structured and in-depth interviews, you feel that you have dealt well with the relationship between the purpose of the research and the proposed methodology when one of the panel leans forward and asks you to discuss the trustworthiness and usefulness of your work for other researchers. This is clearly a challenge to see whether you can defend such a qualitative approach. How do you respond?

References

- **Wilson: chapter 9**
- **Note: Lecture 9: 9. Collecting primary data using semi-structured and in-depth Interviews**

Topic 10: Collecting Primary Data using Questionnaires

Learning Outcomes

By the end of this topic, you will be able to:

- Discuss the efficient way to obtain precise data from a large sample.
- Justify usage is dependent on the purpose of the research such as descriptive, explanatory or exploratory.
- Discuss the design of the questionnaire is imperative as it will influence the response rate, reliability and validity of the answers from participants.

Introduction

The key areas of learning in this topic are:

- Questionnaires are used to obtain precise data that are required to answer the research question as it will not be possible to go back to the same people with another questionnaire.
- Questionnaires with standardised questions are best for descriptive and explanatory research but for exploratory or analytical research the questions need to be different from those for descriptive or explanatory research.
- In exploratory or analytical research, the questions need to obtain data that can explain the cause and effect relationships.
- The same questionnaire is provided to all the participants and this is an efficient way to collect responses from a large sample.
- The design of the questionnaire is important as it will affect the response rate and the reliability and validity of the data collected.

10.1 Questionnaire

It is a general term to include all techniques of data collection in which each person is asked to respond to the same set of questions in a predetermined order (deVaus, 2002).
It is used in many research areas. All the people are provided the same questionnaire and it provides an efficient way to collect responses from a large sample. However many authors argued that it is more difficult to produce a good questionnaire than what has been expected. The questionnaire must ensure that it will collect the precise data that are required to answer the research questions and to achieve the objectives. This is important because it is difficult to go back to the same people with another questionnaire.
The design of the questionnaire will affect the response rate and the reliability and validity of the data collected. The response rate, reliability and validity can be maximised by:

- careful design of individual questions

- clear layout of the questionnaire form
- lucid explanation of the purpose of the questionnaire
- pilot testing
- carefully planned and executed administration.

When to use questionnaire?

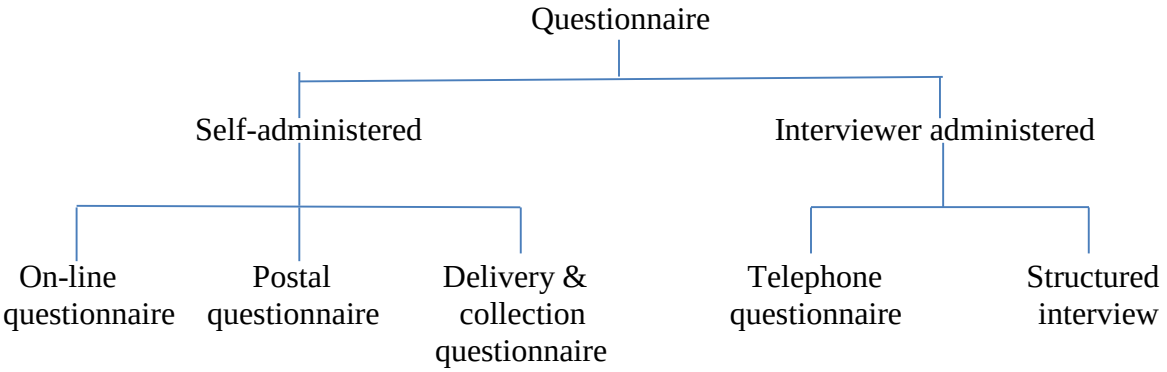
Questionnaires are not good for exploratory or other research that requires large numbers of open-ended questions.

They work best with standardised questions that all respondents will interpret the same way. They are therefore used for descriptive and explanatory research. Descriptive research requires the respondents to identify and describe variables such as attitudes and opinion and different occurrences in the workplace.

On the other hand, exploratory or analytical research requires you to examine the factors and want you to explain their relationships especially about the cause and effect relationships. As a consequence the questionnaire has to be designed differently from that of the descriptive and explanatory types.

10.2 Types of Questionnaire

1. Self-administered questionnaires which are to be completed by the respondents as seen in on-line questionnaire, postal questionnaire and delivered and collected questionnaire.
2. Interviewer administered questionnaires where the interviewer records the answers from the respondents, especially through telephone or as being known as *telephone questionnaire*. Another category is known as *structured interviews* (also known as *interview schedules*) where the interviewers physically meet respondents and ask the questions face-to-face.



The Choice of Questionnaire

Many factors will influence your choice of the questionnaire:

- characteristics of the respondents from whom you wish to collect data
- importance of reaching a particular person as respondent
- importance of respondents' answers not being contaminated or distorted
- size of sample you require for your analysis, taking into account the likely response rate
- types of question you need to ask to collect your data
- number of questions you need to ask to collect your data

These factors may not apply equally to your choice of questionnaire and for some research questions or objectives may not apply at all. The type of questionnaire you choose will dictate who your respondent should be. For example using the delivery and collection questionnaire,

you expect to have a particular respondent to answer the questionnaire and this is a way for you to ascertain that when you get back the answered questionnaire that particular person has answered it. This is to improve the reliability and validity of the data collected. If e-mail or postal questionnaire is used there is no certainty that the person to whom you intended would answer the questionnaire. On the contrary, if interviewer administered questionnaire is used you can ensure the right respondent answers it and this can improve the reliability and validity of your data.
Furthermore any contamination of the respondents' answers will reduce your data's reliability. This can occur when the respondents are not sure of the answer because they do not have the knowledge or experience and you get what is known as ***uninformed response***. If could be the outcome of uninformed response if the respondent finds the answer from other persons.
The type of questionnaire you used and the size of the questionnaire will influence the response rate. These are some examples of how the choice of questionnaire can affect the response rate and the reliability of the data gathered.

All data collected by questionnaires will be analysed by computer. Some software packages allow you to design your questionnaire and to enter and analyse the data.

Deciding what data need to be collected
Research design requirements (requirements to design the questions)
It is important that the questions you want to ask in the questionnaire must be clear and precise so that you can get the appropriate answers from the respondents and consider how you are going to analyse data gathered. Next you must know how to design the questionnaire to meet these requirements so that you are able to answer your research questions and meet your objectives.
For business research, the data collected from using questionnaires will be used for descriptive or explanatory purpose. If the questions are for describing the characteristics of a population, then they are normally administered to a sample. The sample however must be a close representative of the population so that generalizability is possible.
You can do a literature review or discuss with your colleagues, project tutor or interested parties to get some idea to designing the research questions.
For international or cross-cultural research, it is important to have an understanding of the countries or cultures when designing the research questions. As with explanatory research the data are required to test a theory or theories. This means you have to define the theory/theories to test relationships between variables prior to designing the questionnaire. To be able to do this you need to review the literature first, discussed your idea widely and conceptualised your own research clearly prior to designing the questionnaire. In particular you need to be clear about which relationships are likely to exist between variables.

Defining theories in term of relationships between variables
The relationships between variables are normally tested statistically from the data gathered by **the questionnaire.**

Types of variables

Dillman (2000) distinguishes between three types of data variable that can be collected through questionnaires:

- opinion variables - record how respondents feel about something or what they think or believe is true or false. It is about their experience.
- behaviour variables
- attribute variables

} record what respondent do and are.

When recording what the respondents do you are recording their behaviour. On the other hand, respondents' opinions refer to their experience.
Behavioural variables contain data on what people did in the past, do now or will do in the future.
Attributes are best thought of as things a respondent possesses, but not what the respondent does. They are the characteristics of the respondent and include age, gender, marital status, education, occupation and income.

Ensuring that essential data are collected

Data collected must enable the research question(s) are answered and objectives achieved. There is no universal method to do this but there are guidelines to enable this to happen:

1. Decide whether the main outcome of your research is descriptive or explanatory.
2. Subdivide each research question or objective into more specific investigative questions about which you need to gather data.
3. Repeat the second stage if you feel that the investigative questions are not sufficiently precise.
4. Identify the variables about which you will need to collect data to answer each investigative question.
5. Establish how to measure the data for each variable.

Investigative questions are questions that you need to answer in order to address satisfactorily each research question and to meet each objective. They are generated to meet the research question(s) and objectives. Some investigative questions may be subdivided into more detailed investigative questions. You need to be clear whether you are interested in the respondents' opinions, behaviours or attributes. It is found that literature review, discussions with interested parties and pilot studies to be of help here.
Then you need to identify the variables about which you need to collect data to answer each investigative question and how they are to be measured. You can find suggested possibilities through literature review.

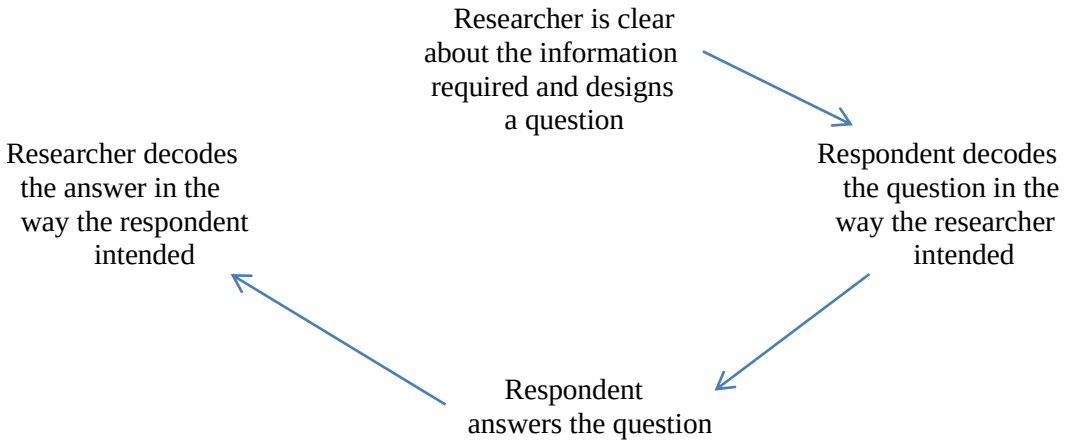
An example of Data requirements table
Sarah was asked to discover staff attitudes to the possible introduction of a no smoking policy of her workplace. Discussion with senior management and colleagues and reading relevant literature helped her to firm up her objective and investigative questions. A selection of these is included in the extract from her table of data requirements.

Research question/objective: To establish employees' attitude to the possible introduction of a no-smoking policy at their workplace.			
Type of research: Predominantly descriptive, although wish to examine differences between employees.			
Investigative questions	Variable(s) required	Detail in which data measured	Check included in questionnaire ✓
Do employees feel that they should be able to smoke in their office if they want to as a right? (opinion)	Opinion of employees to smoke in their office as a right.	Feel… should be allowed, should not be allowed, no strong feelings	
Do employees feel that the employer should provide a smoking room for smokers if smoking in office is banned?(opinion)	Opinion of employees to the provision of a smoking room for smokers	Feel … very strongly that it should, quietly strongly it should, no strong opinions, quiet strongly it should not, very strongly it should not	
Would employees accept a smoking ban at work if the major of people agreed to it? (behaviour)	Likely behaviour of employees regarding the acceptance of a ban	Would ….accept with no pre-condition, accept if a smoking room is provided, not accept without additional conditions (specify conditions), would not accept whatever the conditions	
Do employees opinions differ depending on ▪ age? (attribute) ▪ whether or not a smoker? (behaviour)	(Opinion of employees outlined above) ▪ Age of employee ▪ Smoker	(Included above) ▪ To nearest 5 years band (youngest 16 years, oldest 65) ▪ Non-smokers, smokes but not in office, smokes in office	
How representative are the responses? (attributes)	Age of employee Gender of employee Job	(Included above) Male, female Senior management, management, supervisory, other	

10.3 Designing the Questionnaire

The validity and reliability of the data you collect and the response rate you achieve depend, to a large extent, on the design of your questions, the structure of your questions, and the rigour of your pilot testing. A valid question will enable accurate data to be collected, and one that is reliable will mean that these data are collected consistently. According to Foddy (194:17) 'the question must be understood by the respondent in the way intended by the researcher and the answer given by the respondent must be understood by the researcher in the way intended by the

respondent'. This means that there are 4 stages that must occur if the question is to be valid and reliable as shown in the diagram below.



Designing Individual Questions

The data you need to collect should dictate the kind of question to be designed. There three possible ways to do this:

1. adopt questions used in other questionnaires
2. adapt questions used in other questionnaires
3. develop your own questions

Type and wording of individual questions
Clear wording of questions using terms that are likely to be familiar to and understood by respondents can improve the validity of the questionnaire.
Most types of questionnaire include a combination of open and closed questions. Open questions (also known as open-ended questions) allow respondents to give answers in their own way. However, closed questions (referred to as closed-ended questions or forced-choice questions) provide a number of alternative answers from which the respondent is instructed to choose. Closed-ended questions are quicker and easier to answer. Six types of closed-ended questions are identified:

1. list, where the respondent is offered a list of items, any of which may be selected.
2. category, where only one response can be selected from a given set of categories.
3. ranking, where the respondent is asked to place something in order.
4. scale or rating, in which a rating device is used to record responses.
5. quantity, to which the response is a number giving the amount.
6. grid, where responses to two or more questions can be recorded using the same matrix.

Example of an open question

Please list up to three things you lie about your job:

1. ……
2. ……
3. ……

Note: Open questions are time consuming when coding them at the time of analysis and it is advisable to keep their use to a minimum.

Example of a list question

Please tick ✓ the box in the provided column for services you provided as a home care assistant for this client in the past month.

If you have not provided a particular service, please leave the box blank.

service	provided
cleaning rooms	☐
shopping	☐
bed making	☐
laundry	☐
other	☐
(please describe) …….....	

Note: It must not assume that any unmarked response is a negative response. A non-response can mean uncertainty or an item does not apply.

Example of a category question

How often do you visit this shopping centre?

Interviewer: listen to the respondent's answer and tick ✓ as appropriate.

☐ first visit	2 or more times a week ☐
☐ once a week	less than once a week to fortnightly ☐
☐ less than fortnightly to once a month	less often ☐

Note: Responses should be arranged in a logical order for ease to locate the response category. The categories should be mutually exclusive (should not overlap) and should cover all possible responses. The layout of the questionnaire should make it clear which boxes refer to which response category by placing them close to the appropriate text.

Example of a ranking question

Please number each of the factors listed below in order of importance to you in your choice of a new car. Number the most important 1, the next 2 and so on. If a factor has no importance at all, please leave blank.

factor	importance
acceleration	[ ]
boot size	[ ]
depreciation	[ ]
safety features	[ ]
fuel economy	[ ]
price	[ ]
driving enjoyment	[ ]

other 〔 〕
…… (⇐ please describe)

Note: For ranking questions, ensure the instructions are clear and will be understood by the respondent. Generally, respondents find that ranking more than 7 or 8 items takes too much effort, so you should keep your list to this length or shorter. Respondents can rank accurately only when they can see or remember all items.

Example of a rating or scale question

For the following statement please tick ✓ the box that matches your view most closely.

	agree	tend to agree	tend to disagree	disagree
I feel that employees' views have influenced the decisions taken by management.	☐	☐	☐	☐

Note: You can expand this form of question further to record finer shades of opinion of the respondents.

Conversely for rating the respondent's attitude a 10-point numeric rating scale can be used. It is to reflect the feeling of the respondent to an issue as shown below.

For the following statement please circle ◯ the number that matches your view most closely.

This concert was ….. Poor value for money 1 2 3 4 5 6 7 8 9 10 Good value for money

Another variation is the semantic differential rating scale. The respondent is asked to rate a single object or idea on a series of bipolar rating scales. Each bipolar scale is described by a pair of opposite adjectives designed to capture respondent's attitudes towards service. For these rating scales, you should vary the position of positive and negative adjectives from left to right to reduce the tendency to read only the adjective on the left.

For example:

On each of the lines below place a x to show how you feel about the service you received at our restaurant.

Fast — Slow

Unfriendly — Friendly

Value for money — Over priced

Note: Instead of developing your own scales, it often makes sense to use or adapt existing scales developed by others. These can be found by referring to articles that use scales or refer to literature.

Example of a quantity question
The responses to a quantity question is a number, which gives the amount of a characteristic. For this reason, a quantity question is used to collect behaviour o attribute data. A common quantity question, which collects attribute data, is:

What is your year of birth | 1 | 9 | | | |

Example of a grid
It is a kind of a matrix that records two or more similar questions at the same time. Questions can be listed on the left-hand side of the page and the answer to each question is recorded in the cell where the row and column met. Using the grid may save space but respondents may have difficulties comprehending these designs and they are a barrier to response.
Note: The wording of each question must be carefully considered to ensure that the responses are valid i.e. measure what you think they do. Your questions have to be within the context for which they are written rather than in abstract.

A checklist is used to ensure that the question wording is within the context as shown below:

1. Does your question collect data at the right level of detail to answer your investigative question as specified in your data requirements table?
2. Will respondents have the necessary knowledge to answer your question?
3. Does your question talk down to respondents? It should not!
4. Are the words used in your question familiar, and will all respondents understand them in the same way? In particular you should use simple words and avoid jargon, abbreviations and colloquialisms.
5. Are there any words that sound similar and might be confused with those used in your question?
6. Are there any words in your question that might cause offence? These might result in biased response or a lower response rate.
7. Can your question be shortened?
8. Are you asking more than one question at the same time?
9. Does your question include a negative or a double negative? Question that include the word ‘not’ is sometimes difficult to understand.
10. Is your question unambiguous?
11. Does your question imply that a certain answer is correct? If it does, the question is biased and need to be reworded.
12. Does your question prevent certain answers from being given?
13. Is your question likely to embarrass the respondent?
14. Have you incorporated advice appropriate for your type of questionnaire (such as the maximum number of categories) outlined in the earlier discussion of question types?
15. Are answers to closed questions written so that at least one will apply to every respondent?
16. Are the instructions on how to record each answer clear?

Translating questions into other languages

It has to be done with care. This is extremely important if you are translating questions and associated instructions into another language.

According to Usunier (1998) translating questionnaire attention should pay attention to:

1. Lexical meaning - the precise meaning of individual words.
2. Idiomatic meaning - the meanings of a group of words that are natural to a native speaker and not deductible from those of the individual words (e.g. theEnglish expression for informal communication, 'grapevine', has a similar idiomatic meaning as the French expression telephone arabe, meaning literally 'arab telephone')
3. Grammar and syntax - the correct se of language, including the ordering of words and phrases to create well-formed sentences (e.g. in Japanese the ordering is quite different from English or Dutch, as verbs are at the end of sentences).
4. Experiental meaning - the equivalence of meanings of words and sentences for people in their everyday experiences (e.g. terms that are familiar in the source questionnaire's context such as 'dual career household' may be unfamiliar in the target questionnaire's context).

Usunier also provided a number of techniques for translating your source questionnaire. These, along with their advantages and disadvantages , are summarised below. The source questionnaire is the questionnaire that is to be translated and the target questionnaire is the translated questionnaire. Make sure the source questionnaire and the translated questionnaire are attached to the appendix.

	Direct translation	Back translation	Parallel translation	Mixed techniques
Approach	Source questionnaire to target questionnaire	Source questionnaire to target questionnaire to source questionnaire; comparison of two new source questionnaires; creation of final version.	Source questionnaire to target questionnaire by two or more independent translations; comparison of two target questionnaires; creation of final version.	Back translation undertaken by two or more independent translators; comparison of two new source questionnaires; creation of final version
Advantages	Easy to implement, relatively inexpensive	Likely to discover most problems	Leads to good wording of target questionnaire	Ensure best match between source and target questionnaires
Disadvantages	Can lead to many discrepancies (including those relating to meaning) between source and target questionnaire	Requires two translators, one a native speaker of the source language, the other a native speaker of the target language	Cannot ensure that lexical, idiomatic and experiential meanings are kept in target questionnaire.	Costly, requires two or more independent translators, implies that the source questionnaire can also be changed

Designing the survey form

The order and flow of questions - should be logical to your respondents and interviewer. rather than to follow the order of your data requirements. To assist the flow of the survey it may be

necessary to include filter questions i.e. if certain questions do not apply to the respondent, he/she can skip them and continue on to respond to other questions in the subsequent questions.

Example: Are you currently registered as unemployed? Yes ☐ 1
No ☐ 2

If 'no' go to question 25

Checklist for question order

1. Are questions at the beginning of your questionnaire more straightforward and ones the respondent will enjoy answering? Questions about attributes and behaviours are usually more straightforward to answer than those collecting data on opinions.
2. Are questions at the beginning of your questionnaire obviously relevant to the stated purpose of your questionnaire? E.g. question requesting contextual information may appear irrelevant.
3. Are questions and topics that are more complex placed towards the middle of your questionnaire? By this stage most respondents should be completing the survey with confidence but should not yet be bored or tired.
4. Are personnel and sensitive questions towards the end of your questionnaire, and is their purpose clearly explained? On being asked these a respondent may refuse to answer; however, if they are at the end of an interviewer-administered questionnaire you will still have the rest of the data!
5. Are filter questions and routeing instructions easy to follow so that there is a clear route through the questionnaire?
6. (For interviewer-administered questionnaires) Are instructions to the interviewer easy to follow?
7. Have you re-examined the wording of each question and ensured it is consistent with the position in the questionnaire as well as with the data you require?

Layout of the questionnaire

Layout is important for both self-administered and interviewer-administered questionnaires. Interviewer-administered questionnaires should be designed to make reading questions and filling in responses easy.

The layout for self-administered questionnaire should, in addition, be attractive to encourage the respondent to fill it in and to return it, while not appearing too long. A cramped questionnaire is not likely to be more acceptable to respondents.

Therefore is a wide spread view that long questionnaires should be avoided as this reduces response rate. On the contrary, a short questionnaire may suggest that your research is insignificant and hence not worth bothering with. In general, it has been found that a length between 4 and 8 A4 pages has been acceptable for self-administered questionnaires in organisations. For structured questionnaires can vary from few minutes in the street to 2 hours in comfortable environment. The advice from deVaus' (2002) is:

- Do not make the questionnaire longer than is really necessary to meet your research questions and objectives.
- Do not be too obsessed with the length of your questionnaire.

One way to reduce apparent length that will not reduce legibility is to record answers to questions in a table with the questions in one column and answers in the corresponding column.

A checklist for questionnaire layout

1. (For self-administered questionnaires) Do questions appear squashed on the page? This will put the respondent off reading it and reduce the response rate. Unfortunately a thick questionnaire is equally off- putting!
2. (For self-administered questionnaires) Is the questionnaire going to be printed on good quality paper? This will imply that the survey is important.
3. (For self-administered questionnaires) Is the questionnaire going to be printed on warm-pastel-coloured paper? Warm pastel shades such as yellow and pink generate more responses than coal colours such as green or blue. White is good neutral colour but bright or fluorescent colours should be avoided.
4. (For structured interviews) Will the questions and instructions be printed on one side of the paper only? You will find it difficult to read the questions on back pages if you are using a questionnaire attached to a clipboard.
5. Is your questionnaire easy to read? Questionnaires should be typed in 12 point or 10 point using a plain font. Excessively long and excessively short lines reduce legibility. Similarly, respondents find CAPITALS, *italics* and shaded backgrounds more difficult to read. However, if used consistently, they can make completing the questionnaire easier.
6. Have you ensured that the use of shading, colour, font sizes, spacing and the formatting of questions is consistent throughout the questionnaire?
7. Is your questionnaire laid out in a format that respondents are accustomed to reading? Research has shown that many people skim-read questionnaires (Dillman, 2000). Instructions that can be read one line at a time from left to right moving down the page are therefore more likely to be followed correctly.

Explaining the purpose of the questionnaire

The cover letter - explains the purpose of the survey. This is the first part of the questionnaire that a respondent will look at. Research by Dillman (2000) and others has shown that the message contained in a self-administered questionnaire's covering letter will affect the response rate.

Structure of a covering letter

Type of paper	Good quality, official letterhead, including address, telephone number and email address (if possible)
Maximum length of letter	One side (12 point font size if possible)
Date	In full e.g. 15 August 2015
Recipient's name	Title, forename, surname (absence suggest impersonality)
Recipient's address	In full (absent suggests impersonality)
Salutation	Use recipient's title and name (if possible)
First set of messages	What research is about, lwhy it is useful
Second set of messages	Why recipient's response is important, how long it will take to complete
Third set of messages	Promises of confidentiality and anonymity
Fourth set of messages	How results will be used, token reward or charity donation for participation (if any)
Final set of messages	Whom to contact if have any queries, who to return completed questionnaire to and date by which should be returned.

Closing remarks	Thank recipient for their help.
Signature	Yours, by hand, in blue
Name and title	Yours including forename and surname
Postscript	Express thanks or other appropriate message (optional, but postscript is often the most visible aspect of letter).
Source: Dillman, (2000)	

Introducing the questionnaire

Dillman (2000) argues that to achieve a high response rate, you should explain clearly and concisely the questionnaire to the respondents on the first page of the questionnaire.
You may include a clear unbiased title to convey the topic of the questionnaire; substitile which conveys the research nature of the topic and neutral graphic illustration or logo to add interest.

Closing the questionnaire

At the end of your questionnaire you need to explain clearly what you want the respondents to do with their completed questionnaires. It is usual to start this section by thanking her or him for completing the questionnaire, and by providing a contact name and telephone number for any queries she or he may have. You should give the details of the date by which you would like the questionnaire returned and how and where to return it.

Pilot Testing and Assessing Validity

The questionnaire that is developed must be pilot tested before using it in the research project. The purpose of the pilot test is to refine the questionnaire so that respondents wil have no problems in answering the questions. and there will be no problem in recording the data. In addition, it will enable you to obtain some assessment of the questions' validity and the likely reliability of the data that will be collected.
Pilot test is done by asking an expert or a group of experts to comment on the representativeness and suitability of your questions and to receive suggestions for amendments and improvements. This trial run of the questionnaire is important and necessary as it enables the knowing of how the questionnaire will succeed. How many people you require to do the pilot test will depend on the size of the research project, the financial resources and the research questions and objectives.

Testing for Reliability

Reliability refers to the consistency of responses to the questions in the questionnaire. This can be achieved by doing a test re-test run i.e. the same questionnaire is given to the same respondents twice but at different times but within a reasonable time break.
Another approach is called the alternative form. The questions on another questionnaire are arranged differently from the first questionnaire. The responses from the two questionnaires are then compared to look for consistency.
Another way is to use statistical method to look of correlations of the responses to find internal consistency e.g. Cronbach's alph.

Administering the questionnaire

Once the questionnaire is designed, pilot tested and amended and the sample selected, l the questionnaire is used to collect data. The stage of administering the questionnaire has arrived. Next thing is to gain access to the sample and attempt to gain maximum response rate.
The collection of the data is dependent on the type of questionnaire designed:

- on line questionnaire
- postal questionnaire
- delivery and collection questionnaire
- telephone questionnaire
- structured interviews

SELF CHECK 10

1. Under what circumstances would you choose to use a delivery and collection questionnaire rather than a postal questionnaire? Give reasons for your answer.
2. Develop a checklist of things to consider in questionnaire construction.
3. Describe how the proper sequence of questions may improve a questionnaire.
4. What is the importance of pre-testing of questionnaire?

References

- Saunders: chapter 10
- Note: Lecture 19: **Collecting primary data using Questionnaires**

Topic 11: Analysing Quantitative Data

Learning Outcomes

By the end of this topic, you will be able to:

- Identifying/classifying the type of quantitative data – categorical or quantifiable
- Exploring and presenting data in terms of tables, diagrams and graphs
- Identifying/exploring the variables and presenting them. in forms of charts, histograms, graphs
- Comparing the variables in terms of specific values and interdependence
- Examining relationships, differences and trends using statistics in terms of significant testing, chi-square testing for groups of data, correlation or regression.

Introduction

The key areas of learning in this topic are:

- Qualitative data are gathered based on the positivist research philosophy and they are numerical in nature.
- These data need to be analysed and interpreted to become useful.
- Analysis of quantitative data involves the preparation of the data for analysis, summarizing and presenting the data using tables and graphs, describing the data using suitable statistical methods and examining the relationships and trends between variables.

11.1 Quantitative data

Quantitative data are gathered based on the positivist research philosophy and they are numerical in nature. For those who have adopted the interpretivist philosophy may also use quantitative data in combination with qualitative data. For the data to be useful they need to be analysed and interpreted. It needs quantitative analysis to assist this process. This can range from creating simple tables or diagrams that show the frequency of occurrence through establishing statistical relationships between variable to complex statistical modelling. Computer has been used to analyse the data.

The analysis of quantitative data involves a process consisting of the following stages:

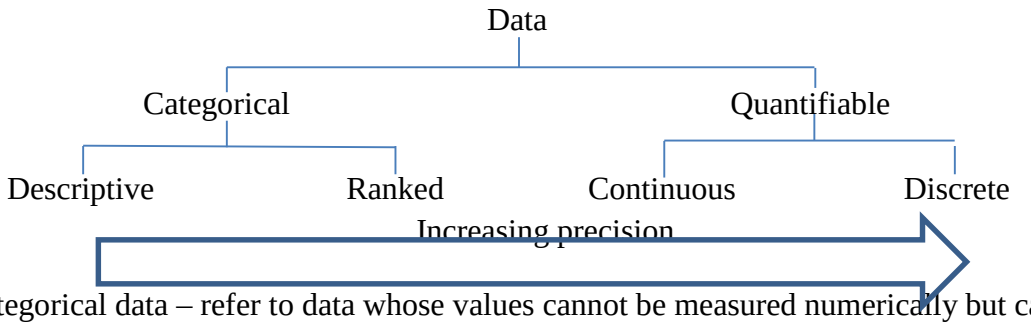
1. preparing your data for analysis;
2. summarizing and presenting your data using tables and graphs;
3. describing your data using suitable statistical methods; and
4. examining relationships and trends between variables.

11.2 Preparing data for analysis

Considerations must be given before obtaining the data for quantitative analysis:

- type of data (level of numerical measurement);
- format in which your data will be input to the analysis software;
- impact of data coding on subsequent analyses (for different data types);
- need to weight cases;
- methods you intend to use to check data for errors.

- Data types - Two distinct groups as shown in the diagram below:



i. Categorical data – refer to data whose values cannot be measured numerically but can be either classified into sets according to characteristics or placed in rank order.
The categorical data are subdivided into (a) descriptive (nominal) data and (b) rank (ordinal) data.
   (a) *Descriptive (nominal) data* cannot be measured numerically or to rank it. However, it is possible to count them and to establish which category has the most and whether they are spread evenly between categories.
   (b) *Ranked (or ordinal) data* are more precise as it is possible to know the position of each case within the data set.

ii. Quantifiable data – refer to those values that can be actually measured numerically as quantities and they are subdivided into two groups:
   (a) Continuous data – refer to those data whose values can be theoretically take any value i.e. can be measured accurately e.g. temperature, delivery distance and length of service and these are continuous data.
   (b) Discrete data – can be measured precisely. Each case takes a finite number of values from a scale that measures changes in discrete units. E.g. the size of a pair of shoes (a non-integer) or a finite number such as the number of mobile phones produced or customers served.

Note: The more precise the level of measurement the greater the range of analytical techniques can be used. E.g. data that have been collected and coded using precise numerical measurements can also be regrouped to a less precise level where they can also be analysed. E.g. a student's score in a test could be recorded as the actual mark (discrete data) or as the position in their class (ranked data). By contrast, less precise data cannot be made more precise. The implication is if you are not sure of the level of precise you require, then it is better for you to collect data at the highest level possible and to regroup them if necessary.

Data Layout

It is possible to use computer-aid programmes to assist in the collection of data such as computer-aided personal interviewing (CAPI), computer-aided telephone interviewing (CATI) and on-line questionnaire can automatically enter and save data to a computer file at the time of collection. Certain survey design and analysis software can help to bring about analysis and integration of the data. It is then possible to specify a data layout compatible with the analysis software. Other data collection methods require the entry of the data and then to proceed to analyse them. In this situation the data layout needs to meet the requirements of the analysis software.

Virtually all analysis software will accept the data if they are entered in table format. This is called a *data matrix*.

A simple data matrix

	id	age	gender	service	employed
Case 1	1	27	1	2	1
Case 2	2	19	2	1	2
Case 3	3	24	2	3	1

Within the data matrix, each column usually represents a single *variable* for the data obtained. Each matrix row contains the variables for an individual case. E.g. if the data have been obtained using a survey, each row will contain the data from one survey form.

On the other hand, for longitudinal data such as a company's share price over time, each row (case) might be a different time period.

In the above table, the first variable (id) is the *survey form identifier*. This means it is possible to link data for each case in the matrix to the survey form when error checking.

The second variable (age) contains quantifiable data, the age of each respondent (case) at the time of the survey.

Subsequent variables contain the remaining data: the third (gender) records this descriptive data using code 1 for male and 2 for female; the fourth (service) records each case's length of service to the nearest year with their most recent employer.

The final variable (employed) records whether each case is (code 1) or is not (code 2) currently in employment. Codes can have different meaning for different variables.

If you intend to enter data into a spreadsheet, the first variable is in column A, the second in column B and so on. Each cell in the first row (1) should contain a short variable name to enable you to identify each variable. Subsequent rows (2 onwards) will each contain the data for one case. Statistical analysis software follows the same logic, although the variable names are usually displayed 'above' the first row, as in the above table.

A worked example

Data Input

As part of a market research interview survey you need to discover which of four products (tomato, ketchup, brown sauce, soy sauce, vinegar) have been purchased within the last month by consumers. You therefore need to collect four data items from each respondent:

- Tomato ketchup purchased within the last month? Yes/No
- Brown sauce purchased within the last month? Yes/No

- Soy sauce purchased within last month? Yes/No
- Vinegar purchased within last month? Yes/No

Each of these data items is a separate variable. However, the data are collected using one question:

1. Which of the following items have you purchased within the last month?

Item	purchased	not purchased	not sure
Tomato ketchup	☐ 1	☐ 2	☐ 3
Brown sauce	☐ 1	☐ 2	☐ 3
Soy sauce	☐ 1	☐ 2	☐ 3
Vinegar	☐ 1	☐ 2	☐ 3

The data collected from each respondent will form separate variables in the data matrix using numerical codes (1=purchased, 2= not purchased, 3= not sure). This is known as multiple-dichotomy coding:

	tomato	brown	soy	vinegar
respondent	1	1	1	2

Question 2 (below) could theoretically have millions of possible responses for each of the 'things'. The number that each respondent mentions way also vary,. Our experience suggests that virtually all respondent will select five or less. Space therefore has to be left to code up to five responses after data have been collected.

2. List up to five things you like about your current job
   ……
   ……...
   ……..
   ……..
   ……
   ……

For office use only

☐☐☐☐
☐☐☐☐
☐☐☐☐
☐☐☐☐

The multi-dichotomy method uses a separate variable for each different answer. For Question 2 a separate variable could be used for each 'thing'' listed: e.g. salary, location, colleagues, hours, holidays, car and so on. Subsequently it is possible to code each variable as 'listed' or 'not listed' for each case. The alternative, the multiple-response method, uses the same number of variables as the maximum number of different responses from any one case. For question 2 these might be named 'liked 1', 'like2, 'liked3', 'liked4', and 'liked5'. Each of these variables would use the same codes and could include any of the responses as a category. Statistical analysis software often contains special multiple-response procedures to analyse such data.

Coding

All data types should, with few exceptions, be recorded using numerical codes. This enables the data to be quickly entered with fewer errors. It also makes subsequent analyses, in particular

those that require re-coding of data to create new variables, more straightforward. Unfortunately, analyses of limited meaning are also easier, such as calculating a mean (average) gender from codes 1 and 2! A common exception to using a numerical code for categorical data is where a postcode is used as the code for a geographical reference. If a spreadsheet is used, a list of codes for each variable needs to be kept. Statistical analysis software can store these so that each code is automatically labelled.

Coding quantifiable data

Actual numbers are often used as codes for quantifiable data. Once the data have been entered as a matrix it is possible to use analysis software to group or combine data to form additional variables with less detailed categories. The process is referred to as ***re-coding***.

E.g. A person's salary could be coded to the nearest dollar and entered into the matrix as 23453 (discrete data). Later, re-coding could be used to place it in a group of similar salaries, from $20,000 to $24,999 (categorical data).

Coding categorical data

Codes are also applied to categorical data.

Existing coding schemes can be used for many variables. These include industrial classification, occupation, and socioeconomic classification as using them would:

- Save time;
- Are normally well tested;
- Allow comparisons of the results with other (often larger) surveys.

These codes should include on the data collection form as pre-set codes (e.g. questionnaires) unless it is decided to use another code. The important point is that it is possible to compare the data with those already collected e.g. data collected in the earlier questionnaires.

Coding at data collection occurs when there is a limited range of well-established categories into which the data can be placed. These are included on the data collection form and the person filling in the form selects the correct category.

Coding after data collection is necessary when it is unclear of the likely responses or there are a large number of possible responses, it is better to wait until data from the first 50 to 100 cases are available and then develop the coding scheme. This is called codebook. While the data collection methods are being designed, it is essential to be clear about the intended analyses and in particular ***the level of precision required*** and ***the coding schemes used by surveys with which comparisons are to be made.***

To create the codebook for each variable it is necessary:

1. to examine the data and establish broad groupings;
2. to subdivide the broad groupings into increasingly specific subgroups dependent on the intended analyses;
3. to allocate codes to all categories at the most precise level of detail required;
4. to note the actual responses that are allocated to each category and produce a codebook;
5. to ensure that those categories that may need to be aggregated together are given adjacent codes to facilitate re-coding.

Coding missing data

Each variable for each case in the data set should have a code, even if no data has been collected. A missing data code is used to indicate why data are missing. deVaus (2002) identifies four main reasons for missing data:

- The data were not required form the respondent, perhaps because of a skip generated by a filter question in a survey.
- The respondent refused to answer the question (a non-response).
- The respondent did not know the answer or did not have an opinion. Sometimes this is treated as implying an answer; on other occasions it is treated as missing data.
- The respondent may have missed a question by mistake, or the respondent's answer may be unclear.
- leaving part of a question in a survey blank implies an answer; in such cases the data are not classified as missing.

Weighing cases

Most data a researcher uses will be a sample.

In stratified random sampling the researcher may have used a different sampling fraction for each stratum (a group of persons performing a certain function such as the management group i.e. the managers, middle managers or technical supervisors). Or the researcher may have obtained a different response rate for each of the strata. To obtain an accurate overall picture it is necessary to take into account these differences in response rates between strata. A common method of achieving this is to use cases from those strata that have lower proportions of responses to represent more than one case in your analysis. Most statistical analysis software allows a person to do this by weighing cases. To weight the cases the person:

1. calculate the percentage of the population responding for each stratum;
2. calculate which stratum had the highest percentage of the population responding;
3. calculate the weight for each stratum using the formula

$$\text{Weight} = \frac{\text{highest proportion of population responding for any stratum}}{\text{proportion of population responding in stratum for which calculating weight}}$$

4. apply the appropriate weight to each case.

An example: Weighting cases

To select your sample for a survey you have used stratified random sampling. The percentage of each stratum's population that responded is given below:

Upper stratum: 90%
Lower stratum: 65%

To account for the differences in the response rates between strata you decide to weight the cases prior to analysis.

The weight for the upper stratum is: 90/90 = 1

This means that each case in the upper stratum will count as 1 case in your analysis.

The weight for the lower stratum is: 90/65 = 1.38

This means that each case in the lower stratum will count for 1.38 cases in your analysis.

You enter these as a separate variable in your data set and use the statistical analysis software to apply the weights.

Checking for errors

Some errors will occur no matter how carefully you code and subsequently enter data . The main methods to check data for errors are as follows:

- Look for illegitimate codes. In any coding scheme, only certain numbers are allocated. Other numbers are therefore errors. Common errors are the inclusion of betters O and o instead of zero, letters l or 1 instead of 1, and number 7 instead of 1.
- Look for illogical relationships. E.g. if a person is coded to the professional socioeconomic group and their social class is unskilled manual an error has occurred.
- Check that rules in filter questions are followed. Certain responses to filter questions mean that other variables should be coded as missing values. If this has not happened there has been an error.

For each possible error there is a need to discover whether it occurred at coding or data entry and then correct it.

Data checking is very time consuming and so is often not undertaken. For not doing it is a very dangerous thing and can result in incorrect results from which false conclusions are drawn.

11.3 Exploring and Presenting Data

Once the data have been entered and checked for errors, the analysis can begin.

Exploratory data analysis approach is useful in these initial stages. It emphasises the use of diagrams to explore and understand the data. Always keep in mind the research questions and objectives when exploring the data. Exploratory data analysis approach looks for relationships in the data. Each diagram or table must be clearly labelled and structured.

A summary checklist of the points to remember when designing a diagram or table:

Checklist for diagrams and tables

For both diagrams and tables:

- ✓ Does it have a brief but clear and descriptive title?
- ✓ Are the units of measurement used clearly stated?
- ✓ Are the sources of data used clearly stated?
- ✓ Are there notes to explain abbreviations and unusual terminology?
- ✓ Does it state the size of the sample on which the values in the table are based?

For diagrams

- ✓ Does it have clear axis labels?
- ✓ Are bars and their component in the same logical sequence?
- ✓ Is more dense shading used for smaller areas?
- ✓ Is a key or legend included (where necessary)?

For tables

- ✓ Does it have clear column and row headings?
- ✓ Are columns and rows in a logical sequence?

It is best to begin exploratory analysis by looking at individual variables and their components. The key aspects to consider will be guided by the research questions and objectives. and to include:

- specific values;
- highest and lowest values;
- trends over time;
- proportions
- distributions.

After completing the exploratory analysis of the data, it moves to begin comparing and looking for relationships between variables, considering in addition:

- conjunctions (the point where values for two or more variables intersect).
- totals
- Interdependence and relationships.

These are summarised in a table below.

<table>
<tr><td></td><td colspan="2">Categorical</td><td colspan="2">Quantifiable</td></tr>
<tr><td></td><td>Descriptive</td><td>Ranked</td><td>Continuous</td><td>Discrete</td></tr>
<tr><td>To show one variable so that any specific value can be read easily</td><td colspan="4">Table/frequency distribution (data often grouped)</td></tr>
<tr><td>To show the frequency of occurrence of categories or values for one variable so that highest and lowest are clear</td><td colspan="2">Bar chart (data may need grouping)</td><td>Histogram or frequency polygon (data must be grouped)</td><td>Bar chart of pictogram (data may need grouping</td></tr>
<tr><td>To show the trend for a variable</td><td></td><td>Line graph or bar chart</td><td>Line graph or histogram</td><td>Line graph or bar chart</td></tr>
<tr><td>To show the proportion of occurrences of categories or values for one variable</td><td colspan="2">Pie chart or bar chart (data may need grouping)</td><td>Histogram or pie chart (data must be grouped</td><td>Pie or bar chart (data may need grouping</td></tr>
<tr><td>To show the distribution of values for one variable</td><td colspan="2"></td><td>Frequency polygon, histogram (data must be grouped) or box plot</td><td>Frequency polygon, bar chart (data may need grouping) or box plot</td></tr>
<tr><td>To show the interdependence between two or more variables so that any specific value can be read easily</td><td colspan="4">Contingency table/cross-tabulation (data often grouped)</td></tr>
<tr><td>To compare the frequency of occurrence of categories or values for two or more variables so that highest and lowest are clear</td><td colspan="4">Multiple bar chart (continuous data must be grouped, other data may need grouping)</td></tr>
<tr><td>To compare the trends for two or more variables so that conjunctions are clear</td><td></td><td colspan="3">Multiple line graph or multiple bar chart</td></tr>
<tr><td>To compare the proportions of occurrences of categories or values for two or more variables.</td><td colspan="4">Comparative pie charts or percentage component bar chart (continuous data must be grouped, other fata may need grouping)</td></tr>
</table>

To compare the distribution of values for two or more variables		Multiple box plot
To compare the frequency of occurrence of categories or values for two or more variables so that totals are clear.	Stacked bar chart (continuous data must be grouped, other data may need grouping)	
To compare the proportions and totals of occurrences of categories or values for two or more variables	Comparative proportional pie charts (continuous data must be grouped, other data may need grouping)	
To show the relationship between cases for two variables		Scatter graph/scatter plot

Exploring and presenting individual variables

To show specific values

Using a table (frequency distribution) is the simplest way of summarising data for individual variables so that specific values can be read.

For descriptive data, the table summarises the number of cases (frequency) in each category.

For variables where a large number of categories (or values for quantifiable data), there is a need to group the data into categories that reflect the research question(s) and objectives.

To show highest and lowest values

As a table, there is no visual significance to highest or lowest values unless emphasized by some other forms. Diagrams are used to provide visual clues. For categorical and discrete data, bar charts and pictograms are used. However, bar charts provide a more accurate representation. Bar charts are sometimes known as column charts

For continuous data, a histogram is used. In a histogram the area of each bar represents the frequency of occurrence and the continuous nature of the data is emphasised by the absence of gaps between the bars.

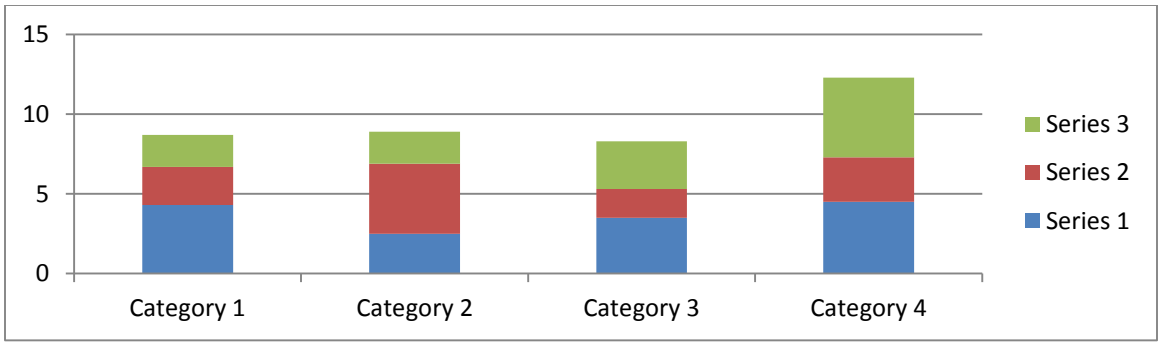
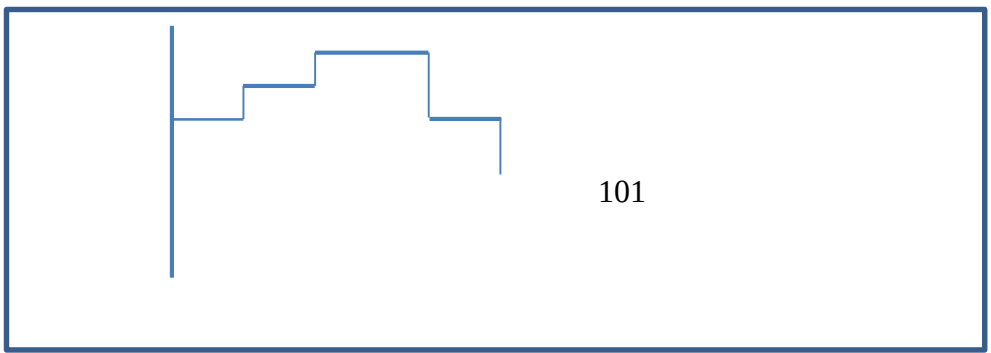


Figure 1 An example of a stacked bar/column chart



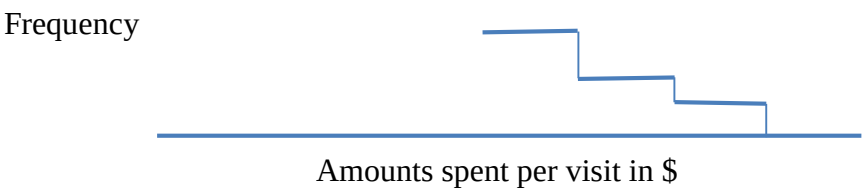


Figure 2: An example of a histogram

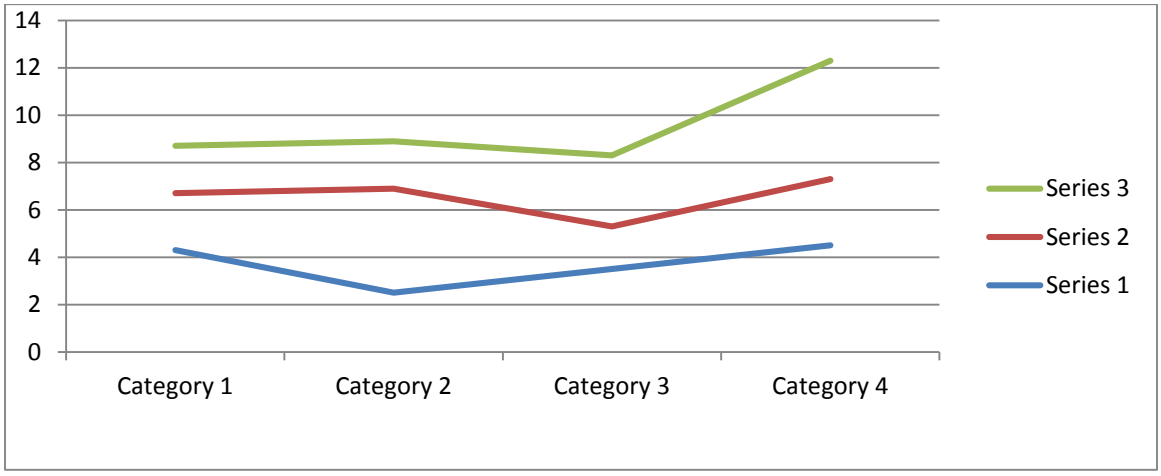


Figure 3: An example of frequency polygons/line graph

To show the trend

Trends can only be presented for variables containing quantifiable (and occasionally ranked) longitudinal data. The most suitable diagram for exploring the trend is a line graph. In which the data values for each time period are joined with a lone to represent the trend.(See above figure the line graph. Bar charts can also be used to show trends between discrete time periods and histograms for continuous time periods.

To show proportions

Research has shown that the most frequently used diagram to emphasise the proportion or share occurrences is the pie chart. A pie chart is divided into proportional segments according to the share each has of the total value.

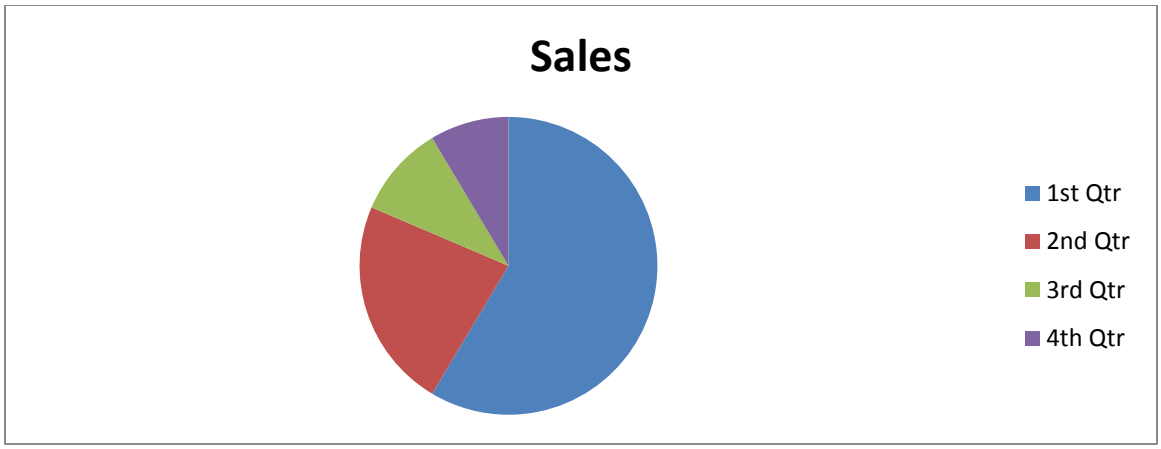


Figure 4: An example of a Pie Chart

To show the distribution of value

This can be done by plotting either a frequency polygon or a histogram for continuous data or a frequency polygon or bar chart for discrete data. If the diagram shows a bunching to the left and a long tail to the right the data is positively skewed. If the converse is true, the data are negatively skewed.

If the data are equally distributed either side of the highest frequency then they are symmetrically distributed. A special form of the symmetric distribution, in which the data can be plotted as a bell-shaped curve, is known as the normal distribution.

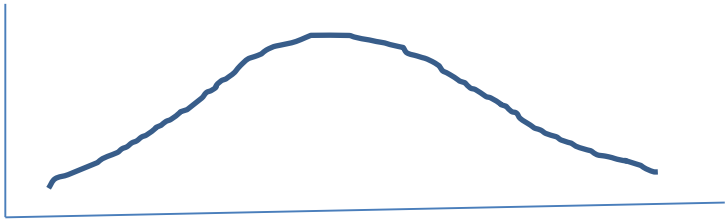
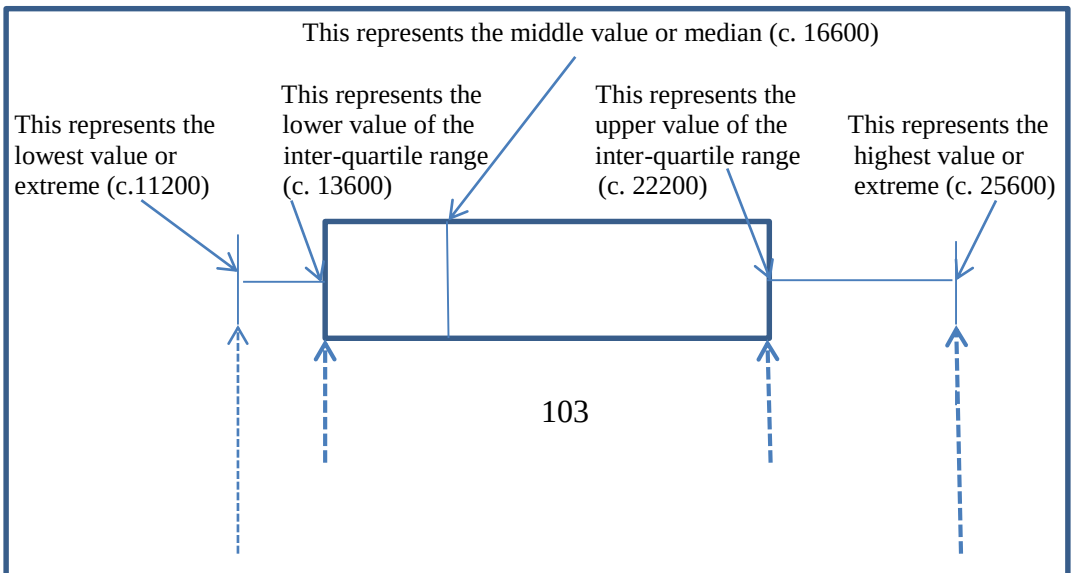
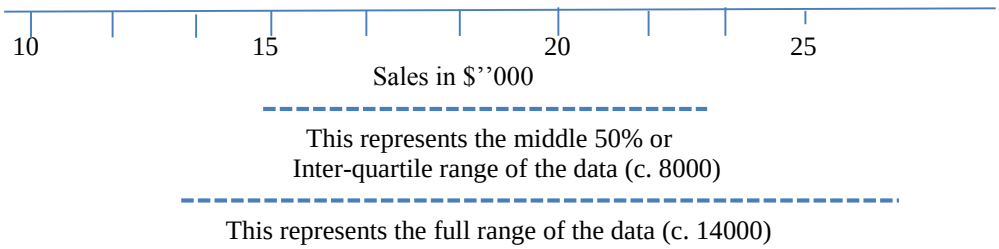


Figure 5: Example of a bell-shaped curve

The box plot is a more advanced statistical analysis. This diagram provides a pictorial representation of the distribution of the data for a variable. The plot shows where the middle value or median is, how this relates to the middle 50% of the data or inter-quartile range, and highest and lowest values or extremes.





Annotated sketch of box plot

Figure 6: An example of a box ploy

Comparing variables

To show specific values and interdependence

The best method is a table known as the contingency table or cross-tabulation as shown in the example below. Such a table enables the examination of interdependence between the variables. For variables where there are likely to be a large number of categories (or values for quantifiable data), there is a need to group the data to prevent the table from becoming too large.

Table 1: Contingency table: number of insurance claims by gender, 2002

Number of claims	Male	Female	Total
0	10032	13478	23500
1	2156	1430	3586
2	120	25	145
3	13	4	17
Total	12321	14937	27258

To compare highest and lowest values

The best approach is to use the multiple bar chart (also known as the compound bar chart). See example given below.

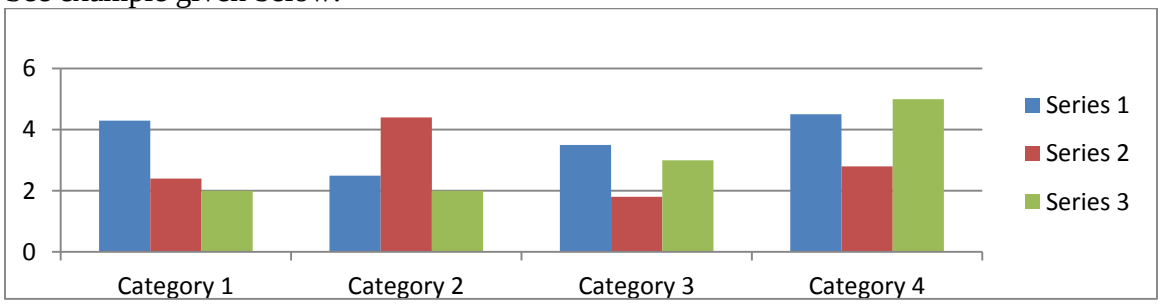


Figure 7: An example of a multiple bar chart

To compare proportions

Comparison of proportions between variables uses either a percentage compound bar chart or two or more pie charts.

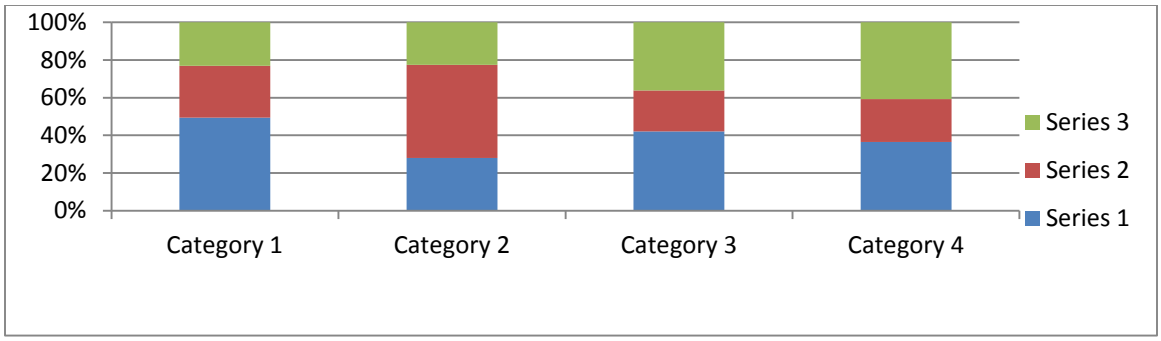


Figure 8: An example of a percentage compound bar chart

To compare trends and conjunctions

The most suitable diagram to compare trends for two or more quantifiable (or occasionally ranked) variable is a multiple line graph where one line represents each variable or use multiple bar charts in which bars for the same time period are placed adjacent.

Looking for conjunctions in the trends - that is, where values for two or more variables intersect - this is where the lines on a multiple line graph cross.

E.g. Breakeven analysis chart.

To compare totals

Comparison of totals between variables uses a variation of the bar chart. A stacked bar chart can be used for all data types provide that continuous data and data where there are more than six possible values or categories are grouped.

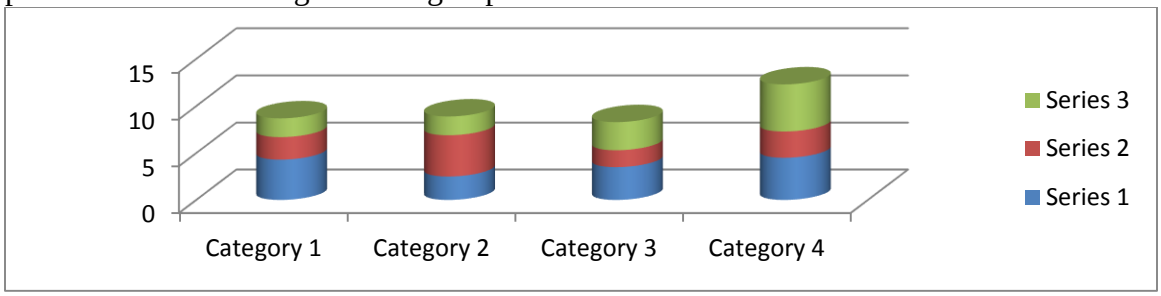


Figure 9: An example of a stacked bar chart

To compare proportions and totals

To compare both proportions of each category or value and the totals for two or more variables it is best to use *comparative proportional pie charts* for all types of data.

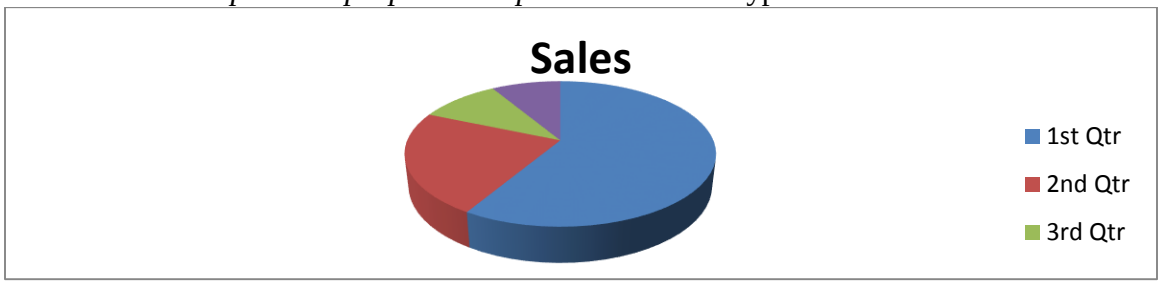


Figure 10: an example of a comparative proportional pie chart

To compare the distribution of values

Often it is used to compare the distribution of values for two or more variables. Plotting multiple frequency polygons or bar charts will enable the comparison of distribution for up to three or four variables (see figures 3 and 7). An alternative is to use a diagram of multiple box plots (see figure 6).

To show the relationship between cases for variables
To explore possible relationships between ranked and quantifiable data variables, it is done by plotting one variable against another. This is called scatter graph or scatter plot.
Convention dictates that the dependent variable i.e. the variable that changes in response to changes in the other (independent) variable - on the vertical axis.

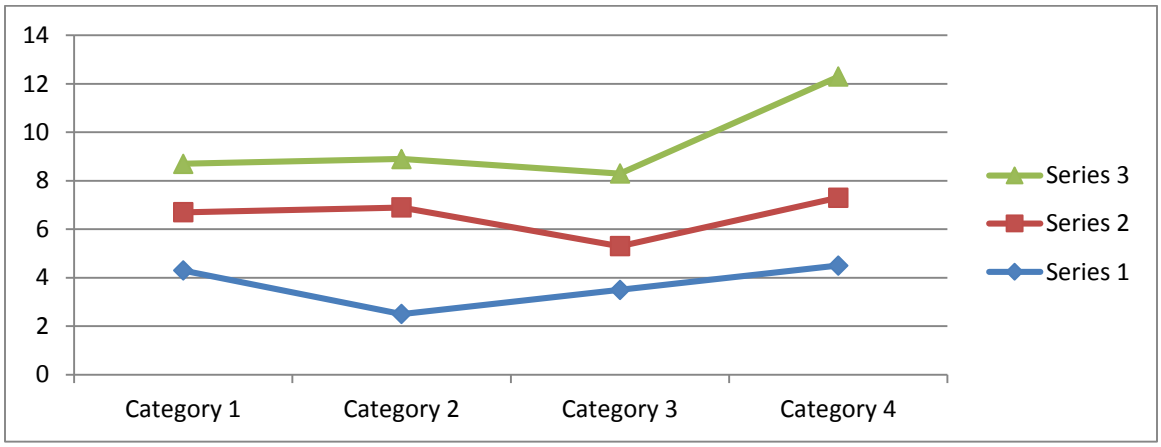


Figure 11: An example of scatter plots.

The strength of the relationship is indicated by the closeness of the points to an imaginary line. If, as the values for one variable increase, so do those for the other, then a positive relationship occurs. If, as the values for one variable decreased, those for the other variable increased, then a negative relationship is indicated.
Furthermore the strength of this relationship can be assessed statistically using techniques such as correlation or regression.

11.4 Describing Data using Statistics

In the exploratory data analysis approach the emphasis is on the use of diagrams in order to understand the data collected.
On the other hand, descriptive statistics enable enables a person to describe and compare variables numerically. The research question(s) and objectives should guide the choice of statistics. Statistics focus on two aspects to describe a variable:
1) the central tendency
2) the dispersion.
These are summarised in table 2.

Table 2: Descriptive statistics by data type: a summary

To calculate a measure of:	Categorical		Quantifiable	
	Descriptive	Ranked	Continuous	Discrete

Central tendency that ……	… represents the value that occurs most frequently	Mode	
	… represents the middle value		Median
	… includes all data values (average)		Mean
Dispersion that ……	… states the difference between the highest and lowest values		Range (data need not be normally distributed but must be placed in rank order)
	… states the difference within the middle 50% of values		Inter-quartile range (data need not be normally distributed but must be placed in rank order)
	… states the difference within another fraction of the values		Deciles or percentiles (data need not be normally distributed but must be placed in rank order)
	… describes the extent to which data values differ from the mean		Variance, or more usually, the standard deviation (data should be normally distributed)
	… compares the extent to which data values differ from the mean between variables		Coefficient of variation (data should be normally distributed)

Describing the central tendency

There are three ways of measuring the central tendency:

- Value that occurs most frequently (mode)
- Middle value or mid-point after the data have been ranked (median)
- Value, often known as the average, that includes all data values in its calculation (mean)

The mode is the value that occurs most frequently. For descriptive data, the mode is the only measure of central tendency that can be sensibly interpreted. E.g. the most common colour of motor cars sold last year was red or that the most common makes were Ford and Vauxhall.

Median value is obtained by ranking all the values in ascending order and finding the mid-point (or 50th percentile) in the distribution. For variables that have an even number of data values the median will occur halfway between the two middle data values. The median has the advantage that it is not affected by extreme values in the distribution.

Mean or average, is the arithmetic mean. It is defined for a set of n observations as the sum of all the values divided by n. The mean \ is indispensable in statistics and easy to work with mathematically, but it is not ideal for all situations. If there are values a long way apart from the others, then the median may be preferred.

Describing the dispersion

It is important to describe how the data values are dispersed around the central tendency. This is possible only for quantitative data. Two of the most frequently used ways of describing the dispersion are the:

1) difference within the middle 50% of values (inter-quartile range);
2) extent to which values differ from the mean (standard deviation).

Although these measures of dispersion are suitable only for quantitative data most statistical analysis software will also calculate them for categorical data if numerical codes are used.

To state the difference between values

It is possible to calculate the difference between the lowest and the highest values i.e. the range. However in statistics this is rarely used. A most frequently used statistics, is the inter-quartile range. The range is divided into 4 equal sections called quartiles. The lower quartile is the value below which is a quarter of the data values will fall, the upper quartile is the value above which a quarter of the data values will fall. The remaining half of the data values will fall between the lower and upper quartiles. The difference between the upper and lower quartiles is the inter-quartile range. The concern is the middle 50% of data values and ignores extreme values.

To describe and compare the extent by which values differ from the mean

It is important to look at the extreme to which the data values for a variable are spread around their means, as this is what is needed to know in order to assess its usefulness as a typical value for the distribution. If the data values are very close to the mean, then the mean is more typical than if they vary widely. To describe the extent of spread of quantifiable data it is to use the ***standard of deviation***. (If your data are a sample, this is calculated using a slightly different formula than if your data are a population, although if your sample is larger than about ***30 cases*** there is little difference in the two situations (Morris, 1999).

If there is a need to compare the relative spread of data between distributions of different magnitudes (e.g. one may be measured in hundreds of tonnes, the other in billions of tonnes). To make a meaningful comparison then a common way of doing this is:

1. to divide the standard deviation by the mean;
2. then to multiply the answer by 100.

This result in statistic called the ***coefficient of variation***. The values of this statistics can then be compared. The distribution with the largest coefficient of variation has the largest relative spread of data.

A worked example

A bank collects data on the total value of transactions at each of its main and sub-branches. The mean value of total transactions at the main branches is 5 times as high as that for the sub-branches. This makes it difficult to compare the relative spread in total value of transactions between the two types of branches. Calculating the coefficients of variation reveals that there is relatively more variation in the total value of transactions at the main branches than the sub-branches:

Branch type	Mean total transaction value	Standard deviation	Coefficient of variation

Main	$6 000 000	$1 417 000	23.62
Sub	$1 200 000	$217 000	18.08

This is because the coefficient of variation for the main braches is larger (23.62) than the coefficient for the sub-branches (18.08).

11.5 Examining relationships, differences and trends using statistics

Question: How does a variable relate to another variable? In statistic the test is known as ***statistical significance***. This is testing the likelihood of the relationship occurring by change alone, if there was no difference in the population from which the sample was drawn. The way in which this significance is tested can be thought of as answering one from a series of questions, dependent on the data type:

- Is the association statistically significant?
- Are the differences statistically significant?
- What is the strength of the relationship, and is it statistically significant?
- Are the predicted values statistically significance?

These are summarised in table 3 below along with statistics used to help examine trends.

Table 3: Statistics to examine relationships, differences and trends by data type: a summary

	Categorical		Quantifiable	
	Descriptive	Ranked	Continuous	Discrete
To test whether two variables are associated	Chi square (data may need grouping)		Chi square if variables grouped into discrete classes	
To test whether two groups (categories) are different	Kolmogorov-Smirnov (data may need grouping)		Independent t-test or paired t test (often used to test for changes over time)	
To test whether three or more groups (categories) are different			Analysis of variance (ANOVA)	
To assess the strength of relationship between two variables		Spearman's rank correlation coefficient	Pearson's product moment correlation coefficient (PMCC)	
To assess the strength of a relationship between one dependent and one or more independent variables			Regression coefficient	
To predict the value of a dependent variable from one or more independent variables			Regression coefficient	
To compare relative changes over time			Index number	
To determine the trend over time of a series of data			Times series: moving averages Regression equation	

Testing for significant relationships and differences

Testing the probability of a relationship between variables occurring by chance alone if there really was no difference in the population from which that sample was drawn is known as ***significant testing***.

If the probability of the test statistic or one more extreme having occurred by chance alone is very low (usually p = 0.05 or lower), then it is a statistically significant relationship. If the probability of obtaining the test statistic or one more extreme by chance alone is higher than 0.05, then it can conclude that the relationship is not statistically significant. There may still be a relationship between the variables under such circumstances, but you cannot make the conclusion with any certainty.
The statistical significance of the relationship indicated by the test statistic is determined in part by the sample size. One consequence of this is that it is very difficult to obtain a significant test statistic with a small sample. Conversely, because the impact of sample size declines rapidly for samples over about 30 cases, if an extremely large sample is chosen it is relatively easy to obtain a significant test statistic for a relationship that is in reality not that obvious.

Type I and Type II errors
Errors can occur when making inferences from samples. Statisticians refer to these as Type I and Type II errors. An error made by wrongly coming to a decision that something is true when in reality it is not, is known as a ***Type I error***. This could be coming to a conclusion that the two variables are related when in actual fact they are not.
A ***Type II error*** involves the opposite occurring i.e. a conclusion that something is not true, when in reality it is. For example, you conclude that the two variables are not related when they are.
Researchers consider Type I error to be more serious than Type II error. It is more important to minimise Type I error than Type II error.

To test whether two variables are associated
Often descriptive or quantifiable data will be summarised as a two-way contingence (see Table 1). The ***chi square test*** enables the determination of how likely that the two variables are associated. It is based on a comparison of the observed values in the table with what might be expected if the two distributions were entirely independent. It is in fact, assessing the likelihood of the data in the table occurring by chance alone by comparing it with what would be expected if the two variables were independent of each other.
The test relies on:

- the categories used in the contingency table being mutually exclusive, so that each observation falls into only one category or class interval;
- no more than 20% of the cells in the table having expected values of less than 5. For contingency tables of two rows and two columns no expected values of less than 10 are preferable.

If the latter assumption is not met, the accepted solution is to combine rows and columns.
The chi square test calculates the probability that the data in your table could occur by change alone.

To test whether two groups are different
Categorical data
Sometimes it is necessary to see whether the distribution of an observed set of values for each category of a variable differs from a specified distribution, e.g. whether your sample differs from the population from which it was selected. The Kolmogrov-Smirnov test enables you to establish this. It is based on a comparison of the cumulative proportions in the same categories

for the specified population. Therefore you are testing the likelihood of the distribution of your observed data differing from that of the specified population by chance alone.

Quantifiable data

If a quantifiable variable can be divided into two distinct groups using a descriptive variable you can assess the likelihood of these groups being different using an ***independent groups t-test.*** This compares the difference in the means of the two groups using a measure of the spread of the scores. If the likelihood of any difference between these two groups occurring by chance alone is low this will be represented by a large *t* statistic with a probability less than 0.05. This is termed statistically significant.

To test whether three or more groups are different

If a quantifiable variable is divided into three or more distinct groups using a descriptive variable, you can assess the likelihood of these groups being different occurring by chance alone by using one-way analysis or one-way ANOVA.

ANOVA analyses the variation within and between groups of data by comparing means. The *F*-ratio or *F* statistic represents these differences. If the likelihood of any difference between groups occurring by chance alone is low, this will be represented by a large *F* ratio with a probability of less than 0.05. This is termed statistically significant.

Hays (1994) lists the following assumptions that need to be met before using one way ANOVA:

- Each data value is independent and does not relate to any of the other data values. This means that you should not use one-way ANOVA where data values are related in some way, such as the same person being tested repeatedly.
- The data for each group are normally distributed. This assumption is not particularly important provided that the number of cases in each group is large.
- The data for each group have the same variance (standard deviation squared). However, provided that the number of cases in the largest group is not more than 1.5 times that of the smallest group, this appears to have very little effect on the test results.

Assessing the strength of relationship

As part of your exploratory data analysis you will already have plotted the relationship between cases for two ranked or quantifiable variables using a scatter graph. Such relationships might include those between weekly sales of a new product and those of a similar established product, or age of employees and their length of service with the company. These examples emphasise the fact that your data can contain two sorts of relationship:

- those where a change in one variable is accompanied by a change in another variable but it is not clear which variable caused the other to change, a correlation;
- those where a change in one or more (independent) variables causes a change in another (dependent) variable, a *cause-and-effect relationship*.

To assess the strength of relationship between pairs of variables

A correlation coefficient enables you to quantify the strength of the relationship between two ranked or quantifiable variables. This coefficient can take on any value between -1 to +1. A value of +1 represents a perfect positive correlation. This means the two variables are precisely related and that, as values of one variable increase, values of the other variable will increase. By contrast a value of -1 represents a perfect negative correlation. This means that the two variables are precisely related, however, as the values of one variable increase those of the other decrease.

Correlation coefficients between +1 and -1 represent weaker positive and negative correlations, a value of 0 meaning the variables are perfectly independent. Within business research it is extremely unusual to obtain perfect correlation.
If your variables contain quantifiable data you should use **Pearson's product moment correlation coefficient (PMCC)** to assess the strength of relationship. However, if one of your variables contains rank data you will need to rank the other variable and use **Spearman's rank correlation coefficient**. Although this uses a different formula to calculate the correlation coefficient it is interpreted in the same way.

To assess the strength of a cause-and-effect relationship between variables
In contrast to the correlation coefficient the ***regression coefficient (also known as coefficient of determination)*** enables you to assess the strength of relationship between a quantifiable dependent variable and one or more quantifiable independent variables. For a dependent variable and one (or perhaps two) independent variables you will have probably already plotted this relationship on a scatter graph. If you have more than two independent variables this is unlikely as it is very difficult to represent four or more scatter graph axes visually.
The regression coefficient (represented by r^2) can take on any value between 0 and +1. It measures the proportion of the variation in a dependent variable (amount of sales) that can be explained statistically by the independent variable (marketing expenditure) or variables (marketing expenditure, number of sales, staff etc). This means that if all the variation in amount of sales can be explained by the marketing expenditure and the number of sales staff the regression coefficient will be 1. If 50% of the variation can be explained the regression coefficient be will be 0.5, and if none of the variation can be explained the coefficient will be 0. Within our research we have rarely obtained a regression coefficient above 0.8.
The process of calculating a regression coefficient and regression equation using one independent variable is normally termed ***regression analysis***. Calculating the regression coefficient and regression equation using two or more independent variables is termed ***multiple regression analysis***.

To predict the value of a variable from one or more other variables
Regression analysis can also be used to predict the values of a dependent variable given the values of one or more independent variables by calculating a regression equation. You may wish to predict the amount of sales for a specified marketing expenditure and number of sales staff. This can be presented in a regression equation:

$$\text{Amount of sales} = a + (b_1 \times \text{marketing expenditure}) + (b_2 \times \text{number of sales staff}$$

Using regression analysis you would calculate the values of a, b_1 and b_2 from data you had already collected on amount of sales, marketing expenditure and number of sales staff. A specified marketing expenditure and number of sales staff could then be substituted into the regression equation to predict the amount of sales that would be generated.

If your equation is a perfect predictor then the regression coefficient will be 1. If the equation can predict only 50% of the variation then the regression coefficient will be 0.5, and if the equation predicts none of the variation the coefficient will be 0.

Examining trends

When examining longitudinal data the first thing is to draw a line graph to obtain a visual represent of the trend. Subsequent to this, statistical analysis can be undertaken. Two of the more common uses of such analyses are:

- to compare trends for variables measured in different units or of different magnitudes;
- to determine the long-term trend and forecast future values for a variable.

To compare trends

To answer some research question (s) and to meet some objectives you may need to compare trends between two or more variables measured in different unit or at different magnitudes. To compare changes in prices of fuel oil and coal over time is difficult as the prices are recorded for different units (litres and tonnes). One way of overcoming this is to use index numbers and compare the relative changes in prices rather than actual figures. Index numbers are also widely used in business publication and by organisations.

Although each indices can involve quite complex calculations they all compare change over time against a base period. The base period is normally given the value of 100 (or 1000 in the case of many share indices), and change is calculated relative to this. Thus a value greater than 100 would represent an increase relative to the base period, and a value less than 100 a decrease.

To calculate simple index numbers for each case of a longitudinal variable you use the following formulae:

$$\text{Index number for case} = \frac{\text{data value for case}}{\text{data value for base period}} \times 100$$

Thus if a company's sales were 125 000 units in 1999 (base period) and 150 000 units in 2000 the index number for 1999 would be 100 and for 2000 it would be 120.

To determine the trend and forecasting

The trend can be estimated by drawing a freehand line through the data on a line graph. However, these data are often subject to variations such as seasonal variations and so this method is not very accurate. A straight forward way of overcoming this is to calculate a moving average for the time series of data values. Calculating a ***moving average*** involves replacing each value in the time series with the mean of that value and those values directly preceding and following it. This smooths out the variation in the data so that you can see the trend more clearly. The calculation of a moving average is relatively straightforward using either a spreadsheet or statistical analysis software.

Once the trend has been established it is possible to forecast future values by continuing the trend forward for time periods for which data have not been collected. This involves calculating the long-term i.e. the amount by which values are changing each time period after variations have been smoothed out. Once again this is relatively straightforward to calculate using analysis software. Forecasting can also be undertaken using other statistical methods including regression.

SELF CHECK 11

1. What is the difference between categorical data and quantifiable data?
2. What is meant by significant testing?
3. In what way sample size can influence significant testing?
4. What are Type I error and Type II error? Which Type error is more serious?
5. When would you use chi-square test?
6. Explain hypothesis testing and give examples.
7. What is the purpose of independent group t-test? Give example of statistical method used.
8. What statistical approaches would you use to assess the strength of relationships between variables?
9. How would you assess the strength of cause and effect relationships among variables?
10. Identify the usefulness of trend analysis.

References

- Wilson: chapter 9 & Saunders: chapter 11

Note: Lecture 11: **Analysing Quantitative Data**

Topic 12: Analysing Qualitative Data

Learning Outcomes

By the end of this topic, you will be able to explain:

1. Qualitative data are obtained to find solution to a problem when it could be done using deductive approach.
2. Qualitative data are deep-seated information that can only be acquired through thorough and time-consuming approaches; from the participants at site and in small sample size, sometime only one participant.
3. It also required unbiased interpretation and understanding of the data that are ethically attained.
4. Most of the data is in the form of verbatim.
5. The resultant outcome of the qualitative data analysis is the development of a theory.

Introduction

The key areas of learning in this topic are:

- It is distinguished from quantitative data analysis, as any kind of analysis that produces findings or concepts and hypotheses that are not arrived by statistical.
- Data from qualitative approach can be derived from different research strategies such as grounded theory, case study, ethnography, or descriptive approach.
- Data are not gathered from hypotheses or theories but from the sites and participants in the selected environment.
- Data are analysed to reach a central phenomenon and to the development of a theory.
- There is no strong emphasis towards statistical approach but there is a need for triangulation for internal validity and reliability.

12.1 Qualitative data

Glaser (1992) defined qualitative analysis as 'any kind of analysis that produces findings or concepts and hypotheses, as in grounded theory, that are not arrived at by statistical methods'.

Grounded theory is a research design, a method that has been used to reach a theory or conclusion through the process of observation of behavioural changes of a participant (e.g. a drug addict) over a period of time. It is very much an inductive approach. The research started without any preconceived ideas of what theories might be developed from the data. This grounded theory was developed by two sociologists – Barney Glaser and Anselm Strauss (1967). It stressed the fact that the theory is based on actual field data from participants but not from discovering the concepts of variables and hypotheses (as seen in quantitative approach).

Qualitative data analysis is very much exploratory in nature. The likelihood is that you have gathered a huge amount of data, then you may face with the problem with what to include and what to discard from your data. In order to avoid this kind of situation, you need to have a clear 'plan' of the steps that need to be addressed to analyse your data in an effective manner. Unfortunately there is no definitive series of steps applicable to qualitative data analysis. This implies that there are many approaches in qualitative research e.g. in case study/multiple case study, grounded theory, ethnography and narrative research.

One key ingredient is you need to transcribe the qualitative data derived from an interview or a field note taken from observational research into a computer. Essentially this is the first stage in the analytical process and requires a great deal of time and patience before moving on to the next step in your analysis.

Before you move on to do your data analysis, you need to have a clear idea as to your research approach. Have you decided to adopt an inductive or deductive approach to analysing your qualitative findings? If you decide to use the inductive analysis then you have to develop categories of your data/findings that emerge from field notes, documents and interviews. These are categories that arise after the gathering of data but not from those imposed before the collection of data.

If you decide to follow a deductive approach in your qualitative analysis, then you will start with a predetermined set of categories. Then try to see relationships in them to reach your conclusion which is the theory you are attempting to reach in your research study.

12.2 Steps in Qualitative Analysis

1. Transcribing your data

This is immensely a time consuming process. You have to make sure that you transcribe exactly your respondents' answers. This must be in the ***verbatim*** answers from your respondents. Be careful not to make any changes to the wording because it may lose the clarity of the answer. After having fully transcribed your data you are ready to begin to organise your data.

Next, if you have carried out observational research, then the first stage in your qualitative analysis is likely to be markedly different from that of data arising from interviews. In the first place, there is no verbal data, but observational data. Your analysis of the observational data will be trying to reducing the data until you are familiar with them. Each phase of data analysis requires data reduction to bring meaning and insight to the words and acts of the participants in the study (Marshall and Rossman, 1995:113).

Patton (1990: 374-375) makes a distinction between data collection methods (i.e. observations and interviews) when considering qualitative analysis. Analytical considerations when analysing data from observation can be summarised as follows:

- Critical/major elements: Present the data by critical incidents or major events – not necessary in the order of occurrence but in the order of importance.
- Various settings: Describe various places, sites, settings or locations (doing case studies of each) before doing cross-setting pattern analysis.

- People: If individuals or groups are the primary unit of analysis then case studies of people or groups may be the focus for case studies.
- Processes: The data may be organized to describe important processes. e.g. control, recruitment, decision-making, socialization, communication.
- Issues: The observations may be pulled together to illuminate key issues, often the equivalent of the primary evaluation questions, such as how participants changed.

Whether analysing observational or interview data, essentially this first step is all about organizing your data into a manageable form and keep into a software package to facilitate analysis later on. (Note: data could be obtained through questionnaires. The more research methods/instruments are used helps to improve the quality of the data as seen in qualitative research triangulation plays an important role to improve the credibility of the data.)
It is imperative that before you start analyse your data, you must become 'familiar' with your data. That means you have to really read your data for the next step of the analytical process.

2. Reading and Generating Categories, Themes and Patterns

Reading through and understanding your data is an imperative aspect of qualitative analysis in order to enable you to find the key parts/areas and identify common patterns or themes. This aspect will lead you to the next important development of the qualitative analysis process i.e. the coding.

3. Approaches to coding your data

A code is a key word, theme or category within your transcript or notes. The idea of coding is to rearrange the data into categories in order to develop theoretical concepts (Strauss, 1987:29). Coding enables you to identify categories and subcategories.
There are two types of coding of data – emergent coding (inductive) and a prior coding (deductive). Emergent coding implies that your categories are developed from your data, whereas a prior coding implies that your categories are determined before your analysis. The important development of coding is to breakdown your raw data into something more manageable and meaningful. There are, however, three types of coding:

- open coding – labelling and categorizing your data.
- axial coding - basing on relating categories with subcategories.
- selective coding – basing on identifying a core category that represents the main theme of the research i.e. the focal category to have been generated from the data.

The purpose of coding procedures can be summarised as follows:

- build rather than test theory;
- provide researchers with analytic tools for handling masses of raw data;
- help analysts to consider alternative meanings of phenomenon;
- be systematic and creative simultaneously; and
- identify, develop and relate the concepts that are the building blocks of theory. (Strauss and Corbin, 1990: 13)

4. Interpreting your Findings

We have seen the stages in the qualitative analysis process – transcribing the collected data, finding of themes and patterns and coding of data. This aspect is the engagement in interpretation of the findings. What does it involve?

A major part of the interpretation of your data is looking for connections between categories that you have identified within each transcript. The types of questions that you might ask yourself concerning each category include:

- Is there a relationship between categories?
- How important are these relationships?
- Is this consistent with previous research?
- Why are there differences or similarities between categories?

These are questions that are challenging and rewarding to your research study. Able to explain especially an insight aspect of your findings can be extremely important and useful in your research topic.

12.3 Quantifying Qualitative Data: Content Analysis

If you decide to quantify your qualitative data, the approach you are likely to adopt is undertaking frequency counts. Simply counting the number of times that a word, phrase or theme is raised by a research participant, is an example of quantifying qualitative data. Another method is content analysis.

'Content analysis is a way of systematically converting text to numerical variables for quantitative data analysis' (Collis and Hussey, 2003:250). Content analysis usually involves the following steps:

1. Identify the unit of analysis – recording unit, sentence or paragraph.
2. Choose categories that are relevant to the issues being studied. They must be reliable, so that if someone else repeated the analysis they would find the same information (increased reliability).
3. Once you have chosen your categories, read through the material, and apply these codes to units of text.
4. Tabulate the material. Present the categories and list the assertions under them.

Content analysis can be applied to examine patterns in your data. It is used as word frequency counts. The advantage of content analysis is the ability to explore what is said (e.g. context) and not said (e.g. form) in successive stages of the interview (Merton and Kendall, 1949:541).

A major potential limitation is that frequency counts may include words that are used out of context or worlds that have multiple meanings (e.g. an institution can refer to a university, or a private or public sector organisation). A respondent's level of English may also mean that certain words and phrases are used more frequently as a result of 'limited' English language capability.

Advantages of content analysis:

- It is relatively straightforward.
- It is an unobtrusive method of data collection;
- It can provide interesting insights into how language is used to convey meaning.

- Finally it is useful for looking at frequencies of words and their change infrequncy over tome. It can be used for analysing historical trends, e.g. mention of the Internet in marketing magazines over the last ten years.
-

Which Qualitative Approach Should I use to analyse my Data?

There is no one best approach to qualitative data analysis. The approach largely depends on the nature of your research topic and your research objectives. For example, if your main purpose is to compare your findings to those of an earlier study, then this may well dictate how you analyse your raw data. You are likely to engage in deductive approach to your analysis. Your analysis may not necessarily need to be a straightforward choice between inductive and deductive methods. When coding your data, you may wish to have a predetermined set of codes (a deductive approach) but also to develop new codes and subsequent categories as you proceed in your analysis (an inductive approach). In short, the approach that you adopt often comes down to personal preference. Nevertheless, some approaches to qualitative analysis, such as grounded theory, are more challenging than others. As we have established, grounded theory is particularly challenging because of the time involved, and it generally commands an excellent grasp of the literature.

An Overview of Qualitative Research

There is no clear distinction between qualitative research from quantitative research. However three distinct differences between the two have been highlighted:

Quantitative data	Qualitative data
1. Based on meanings derived from numbers	1. Based on meanings expressed through words
2. Collection results in numerical and standardised data	2. Collection results in non-standardised data requiring classification into categories
3. Analysis conducted through the use of diagrams and statistics	3. Analysis conducted through the use of conceptualisation

The nature of qualitative data has implications for both its collection and its analysis. To be able to capture the richness and fullness associated with qualitative data they cannot be collected in a standardised way, like that of quantitative data. During analysis, the non-standardised and complex nature of the data collected need to be classified into categories before they can be meaningfully analysed. It may be possible to make use of diagrams and statistics at this stage, such as the frequency of occurrence of certain categories of data and the way to analyse the qualitative data collected is done through the creation of a conceptual framework (research model). This is usually formulated before or during the data collection.
The analysis of qualitative data involves a demanding process and should not be seen as an 'easy option'.

Marshall and Rossman (1999) have advised that data analysis should be considered:
(1) at the time the researcher is formulating a proposal to undertake qualitative research and
(2) the act of analysing qualitative data should occur at the same time as the data are collected and continue to do so after the data are collected.

Tesch (1990) has grouped the many strategies to analyse qualitative data into 4 main categories:

1. understanding the characteristics of language;
2. discovering regularities;
3. comprehending the meaning of text or action;
4. reflection.

These categories indicate a number of broad ways of differentiating approaches to qualitative analysis. The first two categories are associated with analytic strategies that require greater structure and set procedures to follow, in comparison with the second two categories. In fact the first two categories are associated with deductive approach where data categories and codes to analyse data are derived from theory and predetermined analytical framework. On the other hand, the second two categories are associated with inductive approach without predetermined categories and codes to direct the analysis of the data collected. The distinction between the two approaches in qualitative analysis of data is shown diagrammatical as follows:

These different approaches to qualitative analysis indicate different analytic strategies are used to analyse the qualitative data. However there are common features in those approaches that are highly or fairly highly structured and procedural. E.g. one common feature is categorising of data collected into meaningful parts. This rearranging and analysing of data systematically and rigorously is actually transforming the nature of the data collected in order to allow the researcher to:

1. comphrehend and manage them;
2. integrate related data drawn from different transcript and notes;
3. identify key themes or patterns from them for further exploration;
4. develop and/or test hypotheses based on these apparent patterns or relationships;
5. draw and verify conclusions.

The general activities involved in qualitative analysis are:

- categorisation
- 'unitising' data;
- recognising relationships and developing the categories you are using to facilitate this;
- developing and testing hypotheses to reach conclusions.

- Categorisation

  The first activity is the classification of the data into meaningful categories. These categories are in fact the codes or labels to rearrange the data. They provide an emergent structure that is relevant to the research project to organise and analyse the data further. The identification of these categories can be guided by the research question and objectives of the research project. The categories must form a well-structured, analytical framework for the analysis to be done. Dey (1993;96-97) states that 'categories must have two aspects, an internal aspect

– they must be meaningful in relation to the data – and an external aspect - they must be meaningful in relation to the other categories'. As the analysis develops, a more hierarchical approach to the categorisation of the data appears.

- 'Unitising' data
  The next activity of the analytical process is to put the relevant data into the appropriate category or categories. This is referred to as forming 'units' of data. A unit of data can be a number of words, a sentence, a number of sentences, a complete paragraph, or some other chunk of textual data that fits the category. This is actually a selective process which has the effect of reducing and rearranging the data gathered into a more manageable and comprehensive form. One way to bring about the reduction and rearrangement of the data is to use analytical techniques such as matrices, charts, graphs and networks. Using these analytical techniques may enable the researcher to recognise emergent patterns in the data gathered that will provide the researcher with an indication about how to further the data collection.
- Recognising relationships and developing categories
  Generating categories and reorganising the data according to the categories is actually engaging in analysing the data gathered. This analysis will continue as the researcher look for themes and patterns or relationships in the rearranged data. This may even lead the researcher to alter the categories or relationships as he searches for meaning in his data set. He may decide to subdivide or integrate categories as ways to refining or focusing. In qualitative analysis the idea is to come out with a more hierarchical structure by categorising and coding of the data gathered to reach an explanation for the research question and objectives that form the focus of the research study.
- Developing and testing hypotheses
  As the researcher is seeking to come out with patterns within the data gathered and recognising relationships between the categories he will be able to develop hypotheses in order to test them.
  According to Silverman, a hypothesis is a 'testable proposition'. The appearance of an apparent relationship or connection between categories will need to be tested if the researcher wants to reach a conclusion that there is actual relationship between the categories.

  It is important to test the hypothesis that emerges inductively from the data. Looking for new explanations and seeking to explain why negative cases occur, so that the researcher can move towards the development of valid and well-grounded conclusions.

The Interactive Nature of the Process

The course of events outlined above demonstrate that data collection, data analysis and the development and verification of relationships and conclusions are very much an interrelated and interactive set of processes. Analysis occurs during the collection of data as well as after it. This analysis helps to shape the direction of data collection, especially where the inductive, grounded approach is being followed.

Furthermore the interactive nature of data collection and analysis allow you to recognise important themes, patterns and relationship are present in the cases where the data are collected. It can also guide you to adjust future data collection .

The concurrent process of data collection and data analysis has implications for the way in which you will need to manage your time and organise you data and related documentation. It will be necessary arrange interviews or observation with enough time between them to allow yourself sufficient time to write up or type a transcript, or set of notes and to analyse this before proceeding to your next data collection session.

SELF CHECK 12

1. Identify 10 ways by which qualitative research can be distinguished from quantitative research.
2. What are the possible disadvantages of qualitative data analysis?
3. How long will it take me to conduct my qualitative analysis?
4. What are the steps I am likely to go through when analysing my qualitative data?

References

- Wilson: chapter 10 and Saunders: chapter 12
- Note: Lecture 12: **Analysing Qualitative Data**

Topic 13: Writing and Presenting Project

Learning Outcomes

By the end of this topic, you will be able to explain:

- The content of the report must have the abstract, introduction, literature review, method (research design and methodology), findings, conclusions (discussion, conclusion and recommendations), references and appendices.
- The topic must be clear with few words to tell a clear story, to the data and to reach the conclusion.
- The writing must possess higher levels of intellectual capability such as analysis, evaluation and creativity.
- Presentation of the research report required strategies too. It must have clear aims and objectives, and the right approach.

Introduction

The key areas of learning this topic are:

- Writing out of a research project is very important and there is no stopping as and when as desired, it can only stop after the complex section is completed to prevent being lost over the complex issue when writing starts again.
- Writing is also the demand for thinking.
- Writing out a research project required certain sequential order in the content.
- It is necessary to be reminded of the importance of having a clear topic that has the fewest number of words that can tell a clear story of what is intended, help to get the necessary data and reaching the conclusion.
- Presentation of the report requires certain strategies such as stating of clear aims and objectives of the research study and the method of delivery and as far as possible, get the audience involved.

13.1 Writing

Writing is a powerful way of learning (Griffiths, 1993). According to Phillips and Pugh (2000) that writing is the only time when we really think. If writing is synonymous with thinking, then writing is something you should do throughout the whole research process.
This means that writing your project report should be done as each part of the research process is completed and to write as a continuous process throughout the research.

You need to create time for writing your research project. It is helpful to have blocks of time where writing can take place on successive days. This is to ensure a degree of continuity of ideas into your research.

Furthermore writing is actually a creative process and that it should be done when your mind is at its freshest. You should find a place conducive to do your writing where there is little interruption in order to enhance your productivity.

You should set your own realistic targets for the period of writing e.g. how many words you would write for each period of time you are writing. This is the height of self-discipline.

You need to use a word processor for writing your research report. In writing a research report involves much changes and rewriting as more data are collected and new ideas are evolved. There is much updating and deleting of text to improve the research report. At the end of the day you have to express your ideas with brevity, clarity and accuracy.

As a guiding principle, it is essential to construct a plan as a guide, before you start writing. This plan is like a flow chart with the sequential steps or stages in an orderly and coordinated manner and more importantly it provides the direction and scope on what information to gather to achieve the purpose of the research study.

Take the 'clustering' method of Veroff (2001) where it is indicating the following stages:
1. Write the main topic in the middle of a sheet of paper.
2. Jot down the other ideas that occur to you at other points on the page.
3. As the page begins to fill, relationships between the ideas suggest themselves and lines between the ideas may be drawn.
4. This allows you to group the ideas into discrete but related 'chunks', which enables you to devise an outline structure for a section or chapter.

Many writers prefer to finish their writing session while they are in the middle of a section to which they will look forward to returning. You must not leave a complex section half completed as it will be difficult to pick up the threads later on.

It is advisable to get your friends to read your work, even it is in a draft form. Ask your frind to be constructively critical. Your friend must be prepared to tell you about things in the text that are not easy to understand – to point out omissions, spelling, punctuation and grammatical errors. Overall your friend must tell you whether the piece of writing makes sense and achieves its purpose. If your project report does not communicate to the reader in the way it should you will get it back for revision work in the long run. It is much better to try and to ensure that this does not happen.

13.2 Suggested Structure for your Project Report

Robson (2002) has provided a general structure for a project report:
1. Abstract
2. Introduction

3. Literature review
4. Method
5. Results
6. Conclusions
7. References
8. Appendices

This suggested structure is good for the deductive approach. It assumes that the literature was reviewed to establish the current state of knowledge on the topic and this informed the method adopted. Reporting the findings in a factual manner gives rise to a detailed consideration of what these findings mean to the specific piece of research that has been conducted and to the current state of knowledge on the topic.

However if your research is essentially inductive, you may structure your report in a different way. You may tell your story (i.e. to explain your conclusions) in the early part of the report. This may include a clear explanation of how this story relates to the existing literature on the topic. This could be followed by a detailed explanation of how you arrived at these conclusions.

Coming back to the general structure, each of the sections can be subdivided into one or more relevant chapters depending on the topic and the way in which you want to present your particular ***storyline***. This is a vital point. Your structure should have a clear flow.

The Abstract

This is probably the most important part of your report because it may be the only part that some will read. It is a short summary of the complete content of the project report. This abstract prepares the readers for what is to come. It should contain four short paragraphs with the answers to the following questions:

1. What were my research questions, and why were these important?
2. How did I go about answering the research questions?
3. What did I find out in response to my research questions?
4. What conclusions do I draw regarding my research questions?

Smith (1991) lists five principles for the writing of a good abstract:

1. It should be short. Try to keep it to a maximum of two sides of A4. Some universities stipulate a maximum length, often 300 - 500 words.
2. It must be self-contained. It must summarise the complete content of your report.
3. It must satisfy your reader's needs. Your reader must be told about the problem, or central issue, that the research addressed and the method adopted to pursue the issue. It must also contain a brief statement of the main results and conclusion.
4. It must convey the same emphasis as the report, with the consequence that the reader should get an accurate impression of the report's contents from the abstract.
5. It should be objective, precise and easy to read. The project report contents page should give you the outline structure for the abstract. Summarising each section should give you an accurate resume of the report. The abstract is not the place for elaborating any of your main themes. Be objective. You will need to write several drafts before you eliminate every word

that is not absolutely necessary. The purpose is to convey the content of your report in as clear and brief a way as possible.

Writing a good abstract is difficult. The obvious thing to do is to write it after you have finished the report. It is suggested that you draft it at the start of your writing so that you have got your storyline abundantly clear in your mind. You can then amend the draft when you have finished the report so that it conforms to the five principles above.

The Introductory Chapter

The introduction should give the reader a clear idea about

- the central issue of concern in your research and why you thought that this was worth studying.
- It should include a full statement of your research question(s) and objectives.
- If your research is based on an organisation, it is a good idea to include in this chapter some details of the organisation, such as its history, size, product and services. Note: this may be a general background to the more specific detail later on in the chapter on Method where you discuss about the research setting in conducting your research study.
- Come out with a 'route map' to guide the reader through the rest of the report. For example the research process.
- This chapter is usually a fairly brief but it is a vital chapter.

The literature review

This is a necessary activity you must do to commerce your research study.
The main purposes of the literature review are to set your study within its wider context and to show the reader how your study supplements the work that has already been done on your topic. The literature review may inform directly any specific hypothesis that your research was designed to test. These hypotheses may also suggest a particular research approach, strategy and data collection methods. However if you are working inductively (i.e. from data to theory) your literature review may serve the purpose of illuminating and enriching your conclusions.
The title of your literature review chapter should reflect the content of the chapter. It may draw on one of the main themes in the review. It is recommended not to call it simply 'literature review'. It may be that your literature is reviewed in more than one chapter. This would be the case, where you were using more than one body of literature in your research.
However if your literature review is only in one chapter then it is common to just say, literature review.

The Method Chapter (The Design and Methodology Chapter)

This chapter gives the reader sufficient information to make an estimate of the reliability and validity of your methods. The points to be included in this chapter are listed as follows:

1. Setting
    - ✓ What was the research setting?
    - ✓ Why did you choose that particular setting?
    - ✓ What ethical issues were raised by the study and how were these addressed?
2. Participants
    - ✓ How many?

- ✓ How were they selected?
- ✓ What were their characteristics?
- ✓ How were refusals/non-returns handled?

3. Materials
   - ✓ What tests/scales/interview or observation schedules/questionnaires were used?
   - ✓ How were purpose-made instruments developed?
   - ✓ How were the resulting data analysed?
4. Procedures
   - ✓ What were the characteristics of the interviewers and observers, and how were they trained?
   - ✓ How valid and reliable do you think the procedures were?
   - ✓ What instructions were given to participants?
   - ✓ How many interviews/observations/questionnaires were there; how long did they last; where did they take place?
   - ✓ When was the research carried out?

Source: Developed from Robson (2002)

The Results Chapter

This is the most straight forward chapter. It is reports the facts that your research discovered. This is where you will include tables and diagrams to illustrate your findings (do not put these in the appendices).

It also contains verbatim quotes from interviews or sections of narrative account that illustrate periods of unstructured observation. This is in fact a powerful way to convey the richness of your data. It is the qualitative equivalent of tables and graphs, A short verbatim quote can convey with penetrating simplicity a particularly difficult concept that you are trying to explain. Do not be afraid to capture precisely what the interviewee said. It can become a moment of most revelation and amusement.

Two important points to bear in mind when writing this chapter:

1. Stress that the purpose is to present facts (i.e. factual information) so do not offer opinions on the facts.
2. Structure your findings in a clear, logical and easily understood manner. Examine your objectives of your research study and let these dictate the order in which you present your findings.

The Conclusions Chapter

Logically there should be a conclusion for each finding. This suggests that the conclusion chapter should be as long as the findings chapter. This is because findings do not present reflective thoughts such as opinions.

The conclusions demonstrate whether you have answered the research question and show the degree of insight that you exhibit in reaching your conclusions. This part of the research is considered to be most difficult. In fact, this is the second major opportunity in the research process to demonstrate real originality of thought (the first time being in the stage where you choose the research topic). In the conclusions you are making judgements rather than reporting facts. This is where your maturity of understanding can shine through. The key questions to ask

of each of the findings are: 'So what?' and importantly, "To what extent have I answered my research question(s) and met my research objective(s)?' Therefore this chapter actually requires you to reflective critically of your findings and you should give time to think and reflect to reach your conclusions.
You may find that the clearest way to present your conclusions is to follow a similar structure to the one used in your findings section (i.e. method section). If that structure reflects the research objectives then it should make certain that your conclusions would address the research question(s). Drawing up a matrix or a 'mind map' may help you in structuring your findings and conclusions.

An example of using a matrix in the planning of the content for the results and conclusions chapters:

Research question	Result (what factual information did I discover in relation to the specific research question?)	Conclusion (what judgement can I make about the results in relation to the specific research question?)
What are the operational differences between different shifts in the production plant?	Cases of indiscipline in the last six months have been twice as frequent on the night shift as on the day shift.	The night shift indiscipline problem may be due to the reluctance of operators to work on that shift.

In the conclusions chapter, you may include 'discussion' as a section or place it in another chapter. Here you can ask such questions as: "What does this mean?" 'What are the implications for organisations?' 'What are the implications for the current state of knowledge of the topic?' 'How does it add to the literature?' 'What are the implications for future research?

Note: In this conclusion chapter no new material should be included but only in the discussion you may do so, so long it is linked to the point you are making about your conclusions.

Answering the research question(s), meeting the objectives and, if appropriate, supporting or otherwise the research hypotheses is the main purpose of the conclusions chapter. This is where you will consider the findings presented in the previous chapter. You should also return to your literature review and ask yourself 'What do my conclusions add to the understanding of the topic displayed in the literature?

Implications – you may include practical implications to your findings in this conclusions chapter i.e. what your research implies for any future research.

Limitations – what are the limitations you face in your research study? This is probably like a confession of your weaknesses, but as a mature reflection on the degree to which your findings and conclusions can be said to be the truth.

Recommendations – this can be in your conclusions chapter or in a separate chapter., especially if you are doing an exploratory research.

References

Use the Harvard system, It is advisable to start your references at the beginning of the writing process and add to it as you go along. This is to avoid confusion with regard to the references you have made if you were to do it at the later stage.

Appendices
In general, appendices should be kept to the minimum. This is because the points you are making should be in the body of the text. If the material is 'interesting to know' rather than 'essential to know', then it can be in the appendices. However your appendices should include a blank copy of your questionnaire, interview or observation schedule.. Where these have been conducted in a language different from that in which you write your submitted project report you will need to submit both this version and the translation.

Length of the project report
You will probably have guidelines on the amount of words your project report should contain. Do stick to these.

Organising the Project Report's Content

- Choosing a title
  This is the part of the project report that students spend the least time. It is indeed a big mistake. According to Day (1995:1-5), a good title is one that has 'the fewest possible words that adequately describe the content of the paper'.

- Tell a clear story
  Be prepared to be asked what is your main storyline? It is the writing of the abstract (summary). It helps you to think clearly about the storyline because you have to summarise it in so few words. The format for developing the storyline is as follows:
  1st step: What is my research question?
  2nd step: What is the answer to my research question?
  3rd step: What evidence led me to that answer?

- Helping the reader to get all the information out
  Dividing your work
  This is done by dividing the work in such a way that it is easy for the readers to find their way round it and for them always to be clear where they are, where they have come from and where they are going. You can use a matrix approach to help in developing the content in the chapter. You will see that each column of the matrix represents the broad content of a chapter. Each division may have a subdivision.

An example of using a matrix in the planning of the content for the results and conclusions chapters:

Research question	Result (what factual information did I discover in relation to the specific research question?)	Conclusion (what judgement can I make about the results in relation to the specific research question?)

What are the operational differences between different shifts in the production plant?	Cases of indiscipline in the last six months have been twice as frequent on the night shift as on the day shift.	The night shift indiscipline problem may be due to the reluctance of operators to work on that shift.

Previewing and summarising chapters

A further way in which you can signpost your work is to 'top and tail' each chapter. This is to include a few words at the beginning of the chapter that provide a description of how the chapter is to contribute to answering the research question, the methods used in this part of the study, and the points that are covered.

At the end of each chapter it is useful if the reader has a brief summary of the content of the chapter. This may see, like repetition. However, it helps the reader on her or his journey through your report and ensures that you, the writer, are on the correct road.

Tables and graphics

Your reader will find your project report more accessible and easier to read if you present some of your data and ideas in tables and grapics. It is not only numerical data that can be presented in tables. You can also present ideas that can be easily compared.

A final note of caution: to avoid confusing your reader, do make sure that wherever possible you have introduced the table or graphic before it appears in the text.

- One report or two?
  It depends who are the audiences. If for academic audience then the report has to be more complete and detail. If for the organisation's managers then the report will be less intense like little about literature review and the development of the theory but to include the recommendations.

- Clarity and simplicity
  You should write a report aiming to provide the reader with a report that she or he cannot put down until 2.00 a.m. or later for fear of spoiling the flow.
  Use simple sentences as far as possible. So you need to avoid jargon and careful not to use large numbers of quotations form the literature.

- Check your spelling and grammar
  You need to have your friends to check for spelling and grammar mistakes.

  The ten commandments of good writing:
  Day (1995:160) provides a useful checklist for those of us who make the sort of grammatical errors that threaten the credibility of our writing.

  Common grammatical errors

	Often we write	The correct way is
1	Each pronoun should agree with their antecedent	Each pronoun should agree with *its* antecedent.
2	Just between you and I, case is important	Just between you and *me*, case is important.
3	A preposition is a poor word to end a sentence with	A preposition is a poor word *with which* to end a sentence.
4	Do not use no double negatives	Do not use double negatives

5	Verbs have to agree with their subject	verbs *have* to agree with their subject
6	Remember to never split an infinitive	Remember *never to split* an infinitive.
7	When dangling, do not use participles.	Do not use dangling participles.
8	Avoid clichés like the plague.	To avoid clichés like the plague!
9	Do not write a run-on sentence it is difficult when you got to punctuate it so it makes sense when the reader reads what you wrote.	Do not write a run-on sentence. It is difficult to punctuate it so that it makes sense to the reader.
10	About sentence fragments	What about sentence fragments? (!)
11	The data is included in this section.	The data are included in this section.

- Person, tense and gender
  The style of writing the research report is dictated by the method you used and in the choice of personal pronoun. The observer is independent of what is being observed and as such an *impersonal style* is more appropriate. On the other hand the researcher may be the participant observer who participates in the daily life of people under study. The researcher is an intrinsic part of the research process, and the used of the first person seems more logical here. However, style is more important. The use of the term 'the author' sounds too impersonal and the excessive use of 'I' or 'we' may raise questions about your ability to stand outside your data and to be objective.

  Day (1995:160) identifies rules for the correct use of tense. He suggests that 'you' should normally use the present tense when referring to previously published work e.g. Day identifies and you should use the past tense when referring to your present results e.g. I found that …..

  Day (1995) and Becker (1986) champion the use of the active voice (I found that) because it is clearer, shorter and unambiguous.

  Care must be given to the use of gender-based terms. This is to avoid biasness and does not cause offence.

  Preserving anonymity is important and the use of pseudonyms is necessary.

- The need for continual revision
  Project reports are read by many people and they will judge your work. For this reason you need to polish your work with successive drafts until you are happy that you can do no better.

  Checklist for evaluating your first draft
  - ✓ Is there a clear structure?
  - ✓ Is there a clear storyline?
  - ✓ Does your abstract reflect accurately the whole content of the report?
  - ✓ Does your introduction state clearly the research question(s) and objectives?
  - ✓ Does your literature review inform the later content of the report?
  - ✓ Are your methods clearly explained?
  - ✓ Have you made a clear distinction between findings and conclusions in the two relevant chapters?
  - ✓ Have you checked all your references and presented these in the required manner?
  - ✓ Is there any text material that should be in the appendices or vice versa?

- ✓ Does your title reflect accurately your content?
- ✓ Have you divided up your text throughout with suitable headings?
- ✓ Does each chapter have a preview and a summary?
- ✓ Are you happy that your writing is clear, simple and direct?
- ✓ Have you eliminated all jargon?
- ✓ Have you eliminated all unnecessary quotations?
- ✓ Have you checked spelling and grammar?
- ✓ Have you checked for assumptions about gender?
- ✓ Is your report in a format that will be acceptable to the assessing body?

After having gone through the checklist you may decide to make the necessary amendments or corrections. Keep asking yourself 'How easy can I make the reader's task?' After each successive draft do leave a space of time for your thoughts to mature. It is amazing how something you wrote a few days before will now make no sense to you. On the other hand, you may be impressed with what you have written earlier due to its clarity and simplicity and the insight of some passages.
After completing the second draft you may want to give it to your friends to read. Ask your reader to use the checklist above , to which you can add specific points that you feel are important (e.g. are my arguments well reasoned?).

Meeting the assessment criteria
Your work will be assessed against the criteria that apply to your research programme. The Bloom's taxonomy of educational objectives can be used to illustrate your level that your project report should meet. At the lower level of Bloom's taxonomy exists knowledge and comprehension of the topic covered. At the intermediate levels they should contain evidence of application and analysis. Application refers to the ability to apply certain principles and rules in particular situations. Your method section should be the principal vehicle for demonstrating application. Analysis may be illustrated by your ability to break down your data and to clarify the nature of the component parts and the relationship between them. Whatever your assessment criteria, it is certain that you will be expected to demonstrate your ability at these lower and intermediate levels.
The higher levels are synthesis and evaluation. Synthesis is the ability to arrange and assemble various elements so as to make a new statement or plan or conclusion.
Evaluation refers to the ability to judge materials or methods in terms of internal accuracy and consistency or by comparison with external criteria. You have the opportunity to show your ability in the literature review and in the awareness of the limitations of your own research. It is necessary that each of the levels of educational objectives should be demonstrated in your project report.

Oral presentation of the report

- Planning and preparing
  All presentations should have clear aims and objectives.

  The aim is to give the audience an overview of your report in such a way that it will capture their interest (e.g. an abstract). How much detail to go into it will be determined largely by the time at your disposal. Keep it clear and simple.

Your objectives are more specific. These should be phrased in terms of what it is you want your audience members to do after your presentation. It is a good idea to share the objectives with the audience members so that they know about the journey on which they are being taken.

An example of objectives for presentation:

- Describe the purpose of the project
- Explain the context in which the project research was set.
- Identify the research strategy adopted and the reasons for its choice.
- List the main findings, conclusions and recommendations flowing from the research.

The final point is the approach to deliver your presentation. A good idea is to involve the audience members rather than simply tell them what it is you want them to know. Asking them to ask questions throughout the presentation is a good way of ensuring that the talk is not all in one direction. You have to be careful too, to ensure that you do not let questions and answers run away with time or else you will lose your control of time. Be careful not to turn your presentation into a teaching session.

- Using visual aids
  Rawlins (1993:37) says 'I hear and forget, I see and I remember'. The use of visual aids will do more than enhance the understanding of your audience. It will help you to look better prepared and therefore more professional.
  A simple set of overhead slides will perform the same function as a set of notes, in that it will ensure that you do not forget key points, and will help you to keep your presentation on track. Since you have already know the material so well that a key point noted on the overhead will be enough to trigger your thought process and focus the attention of the audience. Key points also ensure that you do not have to read the script but to keep the attention of the audience. Using the white board is actually to supplement your pre-prepared slides. It provides a means to explain further of your points.

- Making the presentation
  If you have prepared your presentation and practised it, you will have the confidence although you may feel nervous. You can try your presentation to a friend to ensure that it flows smoothly.
  You can use Parry's 'three-step' rule:
  1. First tell them what I am going to say;
  2. then I say to them;
  3. then I tell them what I've said'.

  Audience like to know where they are going, they like to know how they are progressing on the journey, and they like to know when they have arrived.
  Finally some practical points that will help:
  - Think about whether you would prefer to sit or stand at the presentation. The former may be better to foster debate, the latter is likely to give you a sense of control. Which you

choose may depend upon the circumstances of the presentation, including the approach you wish to adopt, the room layout, the equipment you are using and your preferred style.

- Consider how you will deal with difficult questions. Rehearse these and your answers in your mind so that you can deal with them confidently during the presentation.
- Avoid jargon.
- Check the room before the presentation to ensure you have everything you need, you are happy and familiar with the layout, and all your equipment is working.

SELF CHECK 13

1. What is the most important intrinsic element required when you are writing your research report?
2. Why is the content of the research report need to follow a certain pattern?
3. What is the importance of having a clear topic for the report?
4. What would be your strategy to present your report?

References

- **Wilson: chapter 11 and Saunders: chapter 13**
- **Note: Lecture 13: Writing and Presenting your Project Report**