MANONMANIAM SUNDARANAR UNIVERSITY

TIRUNELVELI

Directorate of Distance and Continuing Education

M.A. Criminology & Police Science

(Effective from the Academic Year 2016-2017 onwards)

FUNDAMENTALS OF RESEARCH METHODS AND STATISTICAL APPLICATIONS

UNIT-I: RESEARCH: NATURE AND DEFINITION

Introduction

Introduction to research methodology provides students with a comprehensive overview of a broad range of research paradigms and methodologies, with their ontological and epistemological underpinnings, as well as associated methods and techniques, in order to inform the design of methodologically sound research proposals and to develop their interdisciplinary methodological literacy as future researchers. On successful completion of this subject, anyone will be able to: demonstrate an advanced understanding of a broad range of research paradigms and methodologies, including their ontological and epistemological foundations; critically reflect on a range of research paradigms and methodologies, their relationship with disciplines and bodies of literature, and their relevance to specific research problems and research methods and techniques; critically evaluate a range of studies that employ very different research paradigms and methodologies.Research is the systematic process of collecting and analysing information (data) in order to increase our understanding of the phenomenon with which we are concerned or interested. Research involves three main stages:

- Planning
- Data collection and
- Analysis.

The Research Process

It is research involving social scientific methods, theories and concepts, which can enhance our understanding of the social processes and problems encountered by individuals and groups in society. It is conducted by sociologists, psychologists, criminologist, economists, political scientists and anthropologists. It is not just common sense, based on facts without theory, using personal life experience or perpetuating media myths. It basically

- Originates with a question or problem.
- Requires a clear articulation of a goal.
- Follows a specific plan of procedure.
- Usually divides the principal problems into more manageable sub-problems (hypotheses), which guide the research.
- Accepts certain critical assumptions.
- Requires collection and interpretation of data to answer original research question.

Research: Nature, definition & purposes

- It is re-search involving social scientific methods, theories and concepts, which can enhance our understanding of the social processes and problems encountered by individuals and groups in society.
- It is conducted by sociologists, psychologists, economists, political scientists and anthropologists.
- It involves the systematic collection of methods to produce knowledge.
- It is objective.
- It can tell you things you do not expect.
- It consists of theory and observation.
- Sometimes called 'soft sciences' because their subject matter (humans) are fluid and hard to measure precisely.
- It is an empirical research – i.e. facts are assumed to exist prior to the theories that explain them.

It is **not** just common sense, based on facts without theory, using personal life experience or perpetuating media myths.

Concept: A concept is generally accepted collection of meanings or characteristics associated with certain events, objects, conditions, situations, and behaviours.

The importance of concept to research is to

- Device measurement concepts by which to test the hypothetical statements
- Gather data using these concepts.

Inductive VS Deductive- Reasoning

Direction of reasoning (deductive/inductive)

- Level of social reality that it is explaining (macro/ meso /micro)
- Whether it is formal (general) or substantive (specific).

Deduction (deductive)

In the process of deduction, you begin with some statements, called 'premises', that are assumed to be true, you then determine what else would have to be true if the premises are true. For example, you can begin by assuming that God exists, and is good, and then determine what would logically follow from such an assumption. You can begin by assuming that if you think, then you must exist, and work from there. In mathematics, you can also start will a premise and begin to prove other equations or other premises. With deduction you can provide absolute proof of your conclusions, given that your premises are correct. The premises themselves, however, remain unproven and un-provable, they must be accepted on face value, or by faith, or for the purpose of exploration.

Examples of deductive logic

All men are mortal. Joe is a man.
Therefore Joe is mortal.

If the first two statements are true, then the conclusion must be true.

Bachelors are unmarried men. Bill is unmarried.
Therefore, Bill is a bachelor.

Induction (inductive)

In the process of induction, you begin with some data, and then determine what general conclusion(s) can logically be derived from those data. In other words, you determine what theory or theories could explain the data. For example, you note that the probability of becoming schizophrenic is greatly increased if at least one parent is schizophrenic, and from that you conclude that schizophrenia may be inherited. That is certainly a reasonable hypothesis given the data. However, induction does not prove that the theory is correct. There are often alternative theories that are also supported by the data. For example, the behavior of the schizophrenic parent may cause the child to be schizophrenic, not the genes. What is important in induction is that the theory does indeed offer a logical explanation of the data. To conclude that the parents have no effect on the schizophrenia of the children is not supportable given the data, and would not be a logical conclusion.

Examples of inductive logic

This cat is black. That cat is black A third cat is black. Therefore all cats are black.
This marble from the bag is black. That marble from the bag is black. A third marble from the bag is black. Therefore all the marbles in the bag black.
Two-thirds of my latino neighbors are illegal immigrants. Therefore, two-thirds of latino immigrants come illegally.
Most universities and colleges in Utah ban alcohol from campus.
That most universities and colleges in the U.S. ban alcohol from campus.

Deduction and induction by themselves are inadequate to make a compelling argument. While deduction gives absolute proof, it never makes contact with the real world; there is no place for observation or experimentation, no way to test the validity of the premises. And, while induction is driven by observation, it never approaches actual proof of a theory. Therefore an effective paper will include both types of logic.

Quantitative vs Qualitative research

Qualitative research gathers information that is not in numerical form. For example, diary accounts, open-ended questionnaires, unstructured interviews and unstructured observations. Qualitative data is typically descriptive data and as such is harder to analyze than quantitative data.

Qualitative Research

Qualitative research gathers information that is not in numerical form. For example, diary accounts, open-ended questionnaires, unstructured interviews and unstructured observations. Qualitative data is typically descriptive data and as such is harder to analyze than quantitative data. Qualitative research is useful for studies at the individual level, and to find out, in depth, the ways in which people think or feel (e.g. case studies).

Analysis of qualitative data is difficult and requires an accurate description of participant responses, for example, sorting responses to open questions and interviews into broad themes. Quotations from diaries or interviews might be used to illustrate points of analysis. Expert knowledge of an area is necessary to try to interpret qualitative data and great care must be taken when doing so, for example, if looking for symptoms of mental illness.

An interest in qualitative data came about as the result of the dissatisfaction of some psychologists (e.g. Carl Rogers) with the scientific study of psychologists such as the behaviorists (e.g. Skinner). Since psychologists study people, the traditional approach to science is not seen as an appropriate way of carrying out research, since it fails to capture the totality of human experience and the essence of what it is to be human.

Exploring the experience of participants is known as a phenomenological approach (re: Humanism). It is argued that to focus on isolated pieces of behavior, as is most often the case in studies interested in collecting quantitative data, is rather superficial, and ignores the social context within which behavior takes place. Given that psychological research is something which happens in a social context, the objectivity of the researcher, central to traditional methods, is seen as essentially false within psychology. As people studying people, researchers necessarily have attitudes and values which they bring to their research. It is therefore more honest that researchers' attitudes and values should be acknowledged, and form part of the context of research.

A good example of a qualitative research method would be unstructured and group interviews which generate qualitative data through the use of open questions. This allows the respondent to talk in some depth, choosing their own words. This helps the researcher develop a real sense of a person's understanding of a situation. However, it can be time consuming to conduct the unstructured interview and analyze the qualitative data.

Quantitative Research

Quantitative research gathers data in numerical form which can be put into categories, or in rank order, or measured in units of measurement. This type of data can be used to construct graphs and tables of raw data. Experiments typically yield quantitative data, as they are concerned with measuring things. However, other research methods, such as observations and questionnaires can produce both quantitative and qualitative information.

For example, a rating scale or closed questions on a questionnaire would generate quantitative data as these produce either numerical data or data that can be put into categories (e.g. "yes",

"no" answers). Whereas open-ended questions would generate qualitative information as they are a descriptive response.

Experimental methods limit the possible ways in which a research participant can react to and express appropriate social behavior. Findings are therefore likely to be context-bound and simply a reflection of the assumptions which the researcher brings to the investigation.

Criminological Research: Meaning, objective and scope

Analysis of crimes and criminal behavior needs scientific basis. Following scientific methodology in gathering facts about crimes and criminal behavior and consequently analyzing them assures objectivity and impartiality of those involved in solving crimes. This review course will refresh the criminology students who will take the board examination on the basic principles and methods of conducting research, technical writing, and basic statistics which he or she can apply in the practice of his or her profession.

Objectives

- The Identify and apply the concepts of criminological research.
- Determine the types and methods of research.
- Know approaches in analyzing and interpreting crime statistics.

Nature and Scope of Criminology Research

Meaning and Nature of Research

The word "research" is composed of two syllables, re and search. Dictionary define the former syllable as a prefix meaning again, anew, or over again, and the latter as a verb meaning to examine closely and carefully.

- There are two basic complementary research approaches - quantitative and qualitative.

- There are two main goals of social (criminological) research – pure (to develop theory and expand the knowledge base) and applied (to develop solutions for problems and relevant application for criminological practice).

- There are three possible reasons for conducting criminological research – exploration (conducted when there is a little prior knowledge); description (yield to additional information only when some knowledge has been obtained) and; explanation (when substantial knowledge is available, it attempts to explain the facts already gathered).

- Research is simply a systematic, controlled, empirical and critical investigation or refined technique of thinking, employing specialized tools, instruments, and procedures in order to obtain a more adequate solution of a problem than would possible under ordinary means.

- Research process starts with (a) Identifying the problem (SMART), (b) Formulation of hypothesis, (c) collects data or facts, (d) analyzes these critically, and (e) reaches decisions based on actual evidence.
- Research involves original work (literature, studies, and readings) instead of a mere exercise of opinion.
- Research evolves from a genuine desire to know (probe) rather than a desire to prove something.

Ethical Considerations in Research

- Veracity/Accurate Analysis and Reporting (obligation to tell the truth, not to lie or deceive others)
- Privacy (obligation to maintain the state or condition of limited access to a person)
- Anonymity and Confidentiality (obligation not to divulge information discovered without the permission of the subject)
- Fidelity (obligation to remain faithful to one's commitments, which includes keeping promises and maintaining confidentiality)
- Informed consent (seeking permission to the person/guardian)
- No Harm (obligation not to inflict harm/endanger either physical or psychological or socially)
- Voluntary Participation
- Avoiding Deception (reveal real purpose of the research)

Research Methods

Methods in Criminological Research

- Descriptive method (to describe systematically a situation or area of interest factually and accurately)
- Historical method (to reconstruct the past objectively and accurately, often in relation to the tenability of a hypothesis)
- Case and Field method (to study intensively the background, current status, and environmental interactions of a given social unit)
- Co-relational method (to investigate the extent to which variations in one factor correlate with variations in one or more other factors based on correlation coefficient)
- Causal-comparative or "Ex post facto" method (to investigate possible cause-and- effect relationships by observing some existing consequences and looking back through the data for plausible causal factors)
- Experimental method (to investigate possible cause-and-effect relationship between two or more treatment conditions and comparing the results to a control group(s) not receiving the treatment; "What will happen")

Types of Criminological Research

- Action Research (to develop new skills or new approaches and to solve problems with direct application to the workplace or other applied setting)

- Survey (descriptive) Research (to know of interest "what is"; typically employs questionnaires and interviews to determine attitudes, opinions, preferences, and perceptions of interest to the researcher)
- Close-ended Questionnaire (pre-categorized by the researcher's words)
- Open-ended Questionnaire (in respondent's words)
- Observational Research (collecting direct information about human behavior)
- Historical Research (investigating documents and other sources that contains facts that existed in the past; "What was")
- Evaluation Research (to study processes and procedures for the improvement of a system)

Types of Criminological Research According to Purpose

1. Exploration (to develop an initial, rough understanding of a phenomenon)
   - Methods: literature reviews, interviews, case studies, key informants

2. Description (precise measurement and reporting of the characteristics of the population or phenomenon)
   - Methods: census, surveys, qualitative studies

3. Explanation (why "Is x the case?" or "Is x the relationship?")
   - Methods: experimental

➢ **Variables are the conditions or characteristics that the researcher manipulates, controls, or observes. (Independent Variable, Dependent Variable, Moderator Variable)**

➢ **Hypothesis ("wise guess") Null hypothesis; alternative hypothesis (operational hypothesis)**

Sources of information

- Related Literature (books, magazine)
- Related Reading (legal documents, memos)
- Related Studies (journals, thesis, dissertation)
- Key informants
- Artifacts
- Other material evidences

UNIT-II: STEPS IN RESEARCH

Sources of Research Problems

In criminology, as in any other science, theory plays an important role as a basis for formulating research questions and later understanding the larger implications of one's research results. Another motivation for research is one's personal interests. There are other motivational sources for research that we will explore in this chapter, including helping to answer questions illuminated by earlier research. We use the Minneapolis experiment and the SARP replication research to illustrate the three main research strategies: deductive, inductive, and descriptive research.

In all three, theory and data are inextricably linked. The chapter ends with scientific and ethical guidelines that should be adhered to no matter what research strategy is used, and shows how the Minneapolis experiment followed these guidelines. By the chapter's end, you should be ready to formulate a criminological research question, design a general strategy for answering this question, and critique previous studies that have addressed this question.

- Major Reasons for Doing Literature Reviews
- The Search Process (Step 1 ~ 9)
- Critical Analysis of Literature Reviews
- Publication Bias
- Variable Quality in the Primary Research Studies
- Inclusion-Exclusion Decisions
- With the post-positivist paradigm, the researcher who plans to conduct quasi-experimental research needs to be able to develop a hypothesis.
- With an interpretive/constructivist orientation, the researcher should have a good understanding of previous research.

In this transformative paradigm, the researcher should consult with persons who have experienced oppression and seek out literature that represents their viewpoints.

Step 1: Identify Research Topic

Two pieces of advice should guide researchers.

- They should be flexible in their conceptualization of the research problem being investigated.
- They should begin with a broad idea and be prepared to narrow it down as they progress through the research.

Source of research topics

- A research topic can emerge from a wide variety of source.
- the researcher's interests, knowledge of social conditions, etc
- Any of these is appropriate as a source to help identify the primary research topic.

- For researchers interested in conducting a comprehensive review of literature, they must study topics that appear in the literature (Cooper, 1989).
- For sponsored research, the researcher needs to clarify with the funding agency what the research problem.
- Scholars working in the transformative paradigm have been instrumental in stimulating research on a variety of topics.

Step 2: Review Secondary Sources to Get an Overview

- Review of Research in Education: Each volume contains a series in diver topics such as violence in the schools, welfare reform and education, etc
- Yearbook of the National Society for the Study of Education: Recent topics include inter-professional partnerships that facilitate the integration of services to enhance both teaching and learning.
- The Annual Review of Criminology: contains literature reviews on topics of interest in Criminology, Criminal Justice and allied area.
- Research in Race and Ethnic Relations: is published annually to address race relations and minority and ethnic group research.
- Other handbooks have been published on specific topics.

Primary and Secondary sources in criminology

For some research projects you may be required to use primary sources. How can you identify these?

Primary Sources

A primary source provides direct or firsthand evidence about an event, object, person, or work of art. Primary sources include historical and legal documents, eyewitness accounts, and results of experiments, statistical data, pieces of creative writing, audio and video recordings, speeches, and art objects. Interviews, surveys, fieldwork, and Internet communications via email, blogs, list serves, and newsgroups are also primary sources.

In the natural and social sciences, primary sources are often empirical studies—research where an experiment was performed or a direct observation was made. The results of empirical studies are typically found in scholarly articles or papers delivered at conferences.

Primary Sources Primary sources are un-interpreted, original, or new materials—e.g. an activist gave a speech, a scientist conducted original research, a student drew original conclusions from others' works, an artist created a piece of artwork, or your grandmother wrote an autobiography.

Primary sources are first-hand and not interpreted by anyone else, they offer a personal point of view, and are created by a witnesses of, or participants in, an event (except in cases of historical research written after the fact). Researchers also create primary sources.

Questions to ask when determining if something is a Primary Source

Did the author conduct original research on the topic?

- Is the information the result of a survey?
- Is the information un-interpreted data or statistics?
- Is the source an original document or a creative work?
- Did the information come from personal experience?

Why Use Primary Sources?

Sources that present new research, original conclusions based on the research of others, or an author's original perspective are more helpful and effective for your needs. They allow you to interpret the information rather than relying on the interpretations of others. This is why your instructors may require you to seek out original research for your assignments.

Note: Keep in mind that because primary sources reflect the true meanings and ideas put forth by authors, the information itself may not be completely objective, well-reasoned, or accurate.

Examples
- Scholarly journal article that reports new
- Research and findings
- Newspaper/magazine articles written soon after the event/fact Court records
- Translation/excerpt of an original document
- Art or music
- Autobiographies
- Manuscripts
- Correspondence, letters, Speeches
- Interviews
- Data from a research study
- Websites

Secondary Sources

Secondary sources describe, discuss, interpret, comment upon, analyze, evaluate, summarize, and process primary sources. Secondary source materials can be articles in newspapers or popular magazines, book or movie reviews, or articles found in scholarly journals that discuss or evaluate someone else's original research.

Secondary Sources Secondary sources are information sources that interpret, include, describe, or draw conclusions based on works written by others. Secondary sources are used by authors to present evidence, back up arguments and statements, or help represent an opinion by using and citing multiple sources. Secondary sources are often referred to as being “one step removed” from the actual occurrence or fact. Questions to Ask When Determining If Something Is a Secondary Source: Did the author consult multiple sources to create this work?

- Is this information an interpretation or paraphrasing of another author's work?
- Did the information come from second-hand reporting?
- Is the source a textbook, review, or commentary?
- Does the source include quotations or images?
- Why Use Secondary Sources? Secondary sources are best for uncovering background or historical information about a topic and broadening your understanding of a topic by exposing you to others' perspectives, interpretations, and conclusions. However, it is better to critique an original information source (primary source) if you plan to reference it in your work.

Examples
- Most books (including textbooks)
- Documentary movies
- Art, book, movie, and theater reviews
- Analysis of a clinical trial
- Newspaper/magazine articles written as historical, opinionated, or reflective accounts
- Commentaries
- Biographies
- Dictionaries, encyclopedias
- Websites (also primary)
- A research paper written by you
- Literature reviews and meta-analyses

Note: Many times literature reviews and meta-analyses make up part of a peer-reviewed journal article. If the article includes new data or draws new conclusions, remember that overall it is a primary source.

Tertiary Sources

Tertiary sources consist of information which is a distillation and collection of primary and secondary sources - they provide overviews of topics by compiling and synthesizing information gathered from other resources. Why Use Tertiary Sources? Tertiary sources are convenient and easy-to-use; they are great resources to use as introductions to a new topic.

Examples
- Bibliographies
- Dictionaries, encyclopedias (also secondary)
- Handbooks
- Fact books
- Guide books
- Indexes, abstracts, bibliographies used to locate primary and secondary sources
- Manuals
- Almanacs
- Textbooks (also secondary)

What are Variables?

A variable is an observable and measurable element (or attribute) of an event. **Variables** are concepts that have been operationalized. A variable, then, is any entity that can take on different values. OK, so what does that mean? Anything that can vary can be considered a variable. For instance, age can be considered a variable because age can take different values for different people or for the same person at different times. Similarly, country can be considered a variable because a person's country can be assigned a value. Theoretically, variables can be of a qualitative nature. For example, qualitative distinctions could be made regarding a person's age (old or young). The variable gender consists of two text values: male and female.

But, we can, if it is useful, assign quantitative values instead of the text values, but we don't have to assign numbers in order for something to be a variable. It's also important to realize that variables aren't only things that we measure in the traditional sense. For instance, in much social research and in program evaluation, we consider the treatment or program to be made up of one or more variables (i.e., the 'cause' can be considered a variable), hence even the program can be considered a variable.

Independent variables

An Independent Variable is "a variable that stands alone and isn't changed by the other variables you are trying to measure. For example, someone's age might be an independent variable. Other factors (such as what they eat, how much they go to school, how much television they watch) aren't going to change a person's age. In fact, when you are looking for some kind of relationship between variables you are trying to see if the independent variable causes some kind of change in the other variables, or dependent variables." (Independent variable) causes a change in (Dependent Variable) and it isn't possible that (Dependent Variable) could cause a change in (Independent Variable).

For example

(Time Spent Studying) causes a change in (Test Score) and it isn't possible that (Test Score) could cause a change in (Time Spent Studying).

We see that "Time Spent Studying" must be the independent variable and "Test Score" must be the dependent variable because the sentence doesn't make sense the other way around.

What are Independent and Dependent Variables?

Question: What's a variable?
Answer: A variable is an object, event, idea, feeling, time period, or any other type of category you are trying to measure. There are two types of variables-independent and dependent.

Question: What's an independent variable?
Answer: An independent variable is exactly what it sounds like. It is a variable that stands alone and isn't changed by the other variables you are trying to measure. For example, someone's age

might be an independent variable. Other factors (such as what they eat, how much they go to school, how much television they watch) aren't going to change a person's age. In fact, when you are looking for some kind of relationship between variables you are trying to see if the independent variable causes some kind of change in the other variables, or dependent variables.

Question: What's a dependent variable?
Answer: Just like an independent variable, a dependent variable is exactly what it sounds like. It is something that depends on other factors. For example, a test score could be a dependent variable because it could change depending on several factors such as how much you studied, how much sleep you got the night before you took the test, or even how hungry you were when you took it. Usually when you are looking for a relationship between two things you are trying to find out what makes the dependent variable change the way it does.

Many people have trouble remembering which the independent variable is and which the dependent variable is! An easy way to remember is to insert the names of the two variables you are using in this sentence in the way that makes the most sense. Then you can figure out which is the independent variable and which is the dependent variable:
(Independent variable) causes a change in (Dependent Variable) and it isn't possible that (Dependent Variable) could cause a change in (Independent Variable).

Dependent Variables

A dependent variable is what you measure in the experiment and what is affected during the experiment. The dependent variable responds to the independent variable. It is called dependent because it "depends" on the independent variable. In a scientific experiment, you cannot have a dependent variable without an independent variable.

Example: You are interested in how stress affects heart rate in humans. Your independent variable would be the stress and the dependent variable would be the heart rate. You can directly manipulate stress levels in your human subjects and measure how those stress levels change heart rate.

Important distinction having to do with the term variable is the distinction between an independent and dependent variable. This distinction is particularly relevant when you are investigating cause-effect relationships. We must learn this distinction.

In all fairness, it's as "easy" as the signs for arrivals and departures at airports?

-- Do I go to arrivals because I'm arriving at the airport?
or
-- Does the person I'm picking up go to arrivals because they're arriving on the plane?!

The **dependent variable** (outcome) is the variable one is attempting to predict. By convention is represented by the letter **Y**. Common dependent variables in criminal justice are concepts such as crime and recidivism. The **independent variable** (predictor) is the variable that causes, determines, or precedes in time the dependant variable and is usually denoted by the letter **X**. An

independent variable in one study could become a dependent variable in another. For example, a study of the impact of poverty (X) upon crime (Y) [poverty-crime] finds poverty as the independent variable, whereas a study that looks at race (X) as a predictor of poverty (Y) [race-poverty] finds poverty as a dependent variable. As a rule of thumb, the treatment variable is always an independent variable, as are demographic variables, such as age, sex, and race. The dependent variable usually is the behaviour/attitudes.

Table 2.1 Independent and Dependent variables: synonyms

Independent Variable	Dependent Variable
Predictor	Criterion
Presumed Cause	Presumed Effects
Stimulus	Response
Predicated from	Predicated to
Antecedent	Consequence
Manipulated	Measured outcome

Main Steps in Social Research

Research involves the natural human actions of asking questions and finding answers. The beginning state of a research project is an engaging and relevant question and the end state is an answer to this question. The heart of research is validation. When a researcher validates an answer, they bend over backwards to perform every reasonable check to ensure that their answer is an appropriate representation of reality. The context of research is social. New knowledge created during research is linked to knowledge that was created by previous researchers. New knowledge is only accepted after it has been presented to and critically reviewed by a community of researchers. Social science research, the topic of this page, involves research questions that involve interactions between people.

Relevance

The research process is a time effective way to discover best practices for elevating learning and professional growth of others. At present, most teachers advance their practices by a trial and error process. In general, a trial and error process is time consuming, ineffective, and leads to anecdotal findings that cannot be transferred to others. In contrast, the research process helps a teacher focus on that which is most important, it fosters review and application of known knowledge, and it leads to results that extend the body of knowledge.

Following are the main steps in social research

- Selection of Research Problem.
- Extensive Literature Survey.
- Making Hypothesis.
- Preparing the Research Design.
- Sampling.
- Data collection.
- Data Analysis.
- Hypothesis Testing.

Different Purposes of Social Research

Social research is research conducted by social scientists following a systematic plan. Social research methodologies can be classified along a quantitative/qualitative dimension.

Exploratory

- Goal is to generate many ideas.
- Develop tentative theories and conjectures.
- Become familiar with the basic facts, people and concerns involved.
- Formulate questions and refine issues for future research.
- Used when little is written on an issue.
- It is the initial research.
- Usually qualitative research.

Descriptive research

- Presents a profile of a group or describes a process, mechanism or relationship or presents basic background information or a context.
- Used very often in applied research.
- E.g.: General Household survey – describes demographic characteristics, economic factors and social trends.
- Can be used to monitor changes in family structure and household composition.
- Can also be used to gain an insight into the changing social and economic circumstances of population groups.
- Often survey research.

Analytical (or explanatory)

- Goes beyond simple description to model empirically the social phenomena under investigation.
- It involves theory testing or elaboration of a theory.
- Used mostly in basic research.

Evaluation

- Characterised by the focus on collecting data to ascertain the effects of some form of planned change.
- Used in applied research to evaluate a policy initiative or social programme to determine if it is working.
- Can be small or large scale, e.g.: effectiveness of a crime prevention programme in a local housing estate.

Case Study

A case study is a detailed analysis of a single event, group, or person for the purpose of understanding how a particular context gives rise to this event, group, or person.

Ethnography

Ethnography is an in-depth study of a culture for the purpose of understanding that culture and its inner workings.

Grounded Theory Research

In grounded theory research, a researcher uses the inductive reasoning process to develop a theory that explains observed behaviors or processes. Grounded theory is more of an approach to qualitative research than a specific method.

Action Research

Action research is either research initiated to solve an immediate problem or a reflective process of progressive problem solving led by individuals working with others in teams or as part of a "community of practice" to improve the way they address issues and solve problems.

Table No. 2.2 The list of research methods is not exhaustive

Statistical–quantitative methods	Qualitative methods	Mixed methods
• Cluster analysis • Correlation and association • Multivariate statistics • Regression analysis • Social network analysis • Social sequence analysis • Surveys and questionnaire • Structural equation modeling • Survey research • Quantitative marketing research	• Analytic induction • Case study • Ethnography • Life history • Morphological analysis • Most significant change technique • Participant observation • Textual analysis • Unstructured interview	• Archival research • Content analysis • Longitudinal study • Focus group • Historical method • Semi-structured interview • Structured interview • Triangulation (social science)

Formulation of research problem

A clear statement of objectives will help you develop effective **research**. It will help the decision makers evaluate your project. It's critical that you have manageable objectives. (Two or three clear goals will help to keep your **research** project focused and relevant).

1. Specify the Research Objectives
A clear statement of objectives will help you develop effective research. It will help the decision makers evaluate your project. It's critical that you have manageable objectives. (Two or three clear goals will help to keep your research project focused and relevant.)

2. Review the Environment or Context of the Research Problem
As a marketing researcher, you must work closely with your team. This will help you determine whether the findings of your project will produce enough information to be worth the cost. In order to do this, you have to identify the environmental variables that will affect the research project.

3. Explore the Nature of the Problem

Research problems range from simple to complex, depending on the number of variables and the nature of their relationship. If you understand the nature of the problem as a researcher, you will be able to better develop a solution for the problem.

To help you understand all dimensions, you might want to consider focus groups of consumers, sales people, managers, or professionals to provide what is sometimes much needed insight.

4. Define the Variable Relationships

Marketing plans often focus on creating a sequence of behaviors that occur over time, as in the adoption of a new package design, or the introduction of a new product.
Such programs create a commitment to follow some behavioral pattern in the future.

Studying such a process involves

- Determining which variables affect the solution to the problem.
- Determining the degree to which each variable can be controlled.
- Determining the functional relationships between the variables and which variables are critical to the solution of the problem.

During the **problem formulation** stage, you will want to generate and consider as many courses of action and variable relationships as possible.

5. The Consequences of Alternative Courses of Action

There are always consequences to any course of action. Anticipating and communicating the possible outcomes of various courses of action is a primary responsibility in the research process.

What is Research Design?

- A research design provides the framework for the collection and analysis of data.
- A choice of research design reflects decisions about the priority being given to a range of dimensions of the research process.

- Involves research method.
  - Research method is simply a technique for collecting data. It can involve a specific instrument such as a self-completion questionnaire or a structured interview etc.

Tools of Research

- The library and its resources
- The computer and its software
- Techniques of measurement
- Statistics
- Facility with language
- Tools are not research methods – e.g. library research and statistical research are meaningless terms.
- Tools help your research methods.

Research Proposal (More formal than Research Design)

- **Title**
- **Statement of research question**
  - **Remember to stress why the problem is important!**
- **Background/information**
- **Aims and objectives of the study**
- **Methods**
- **Timetable**
- **Data analysis**
- **Limitations of the study**
- **Ethical issues**
- **In Funding applications, add**
  - **Resources/Budget**
  - **Dissemination**

Selection of Problem

When selecting a research problem for your study, there are a few factors which you need to consider. These factors will ensure that your research process is more manageable and you will remain motivated. Below given are the factors to consider in selecting a research problem.

Considerations in Selecting Research Problem

The most important criterion in selecting a research problem

1. Interest research problem

. The whole research process is normally time. Consuming and a lot of hard work are needed. If you choose a topic which does not greatly interest you, it would become difficult to keep up the motivation to write. The whole research process is normally time; consuming and a lot of hard work is needed. If you choose a topic which does not greatly interest you, it would become difficult to keep up the motivation to write. Before selecting a research problem, you need.

2. Expertise

Before selecting a research problem,

- You need to ensure that you met certain level of expertise in the area you are proposing.
- Make use of the facts you learned during the study and of course your research supervisors will lend a hand as well.

*** Remember, you need to do most of the work yourself.

3. Data availability

If your research title needs collection of information (journal, reports, proceedings) before finalising the title, you need to make sure you have these materials available and in the relevant format.

4. Relevance

Always choose a topic that suits your interest and profession. Ensure that your study adds to the existing body of knowledge. Of course, this will help you to sustain interest throughout the research period.

5. Ethics

In formulating the research problem, you should consider some ethical issues as well. Sometimes, during the research period, the study population might be adversly affected by some questions. In ICT, some scenarios might occur especially research related information security, which might concern certain authorities. Therefore, it is always good for you to identify ethics related issues during the research problem formulation itself.

Review of Literature

Major Reasons for doing Literature Reviews

- The purpose of the literature review is to provide the reader with an overall framework.

- Literature review serves to explain the topic of the research and to build a rational for the problem that is studied.

- Researchers use the literature review to identify a rationale for their own study.

Some of specific rationales might emerge from your literature review

1. You may find a lack of consistency in reported result

e.g. Born (1993) chose to study site based management and shared decision making because the outcomes of previous research were unclear.

2. You may have uncovered a flaw in previous research based on its design, data collection instruments, sampling, or interpretation

e.g. Lips (1993) notes the gender-sensitive nature of tests used to support differences between males and females in mathematics skills.

3. Research may have been conducted on a different population.

e.g. Sullivan, Vernon, and Scanlan (1987) note that incidence data on sexual abuse were available for the general population but not for deaf children.

4. You may document an ongoing educational or psychological problem and propose studying the effect of an innovative intervention to try to correct that problem.

e.g. B.T. Anderson (1993) noted the ongoing problem of underrepresentation of minority women in scientific careers.

5. Uncertainty about the interpretation of previous studies' findings may justify further research.

e.g. Eagly and Carli (1981) reported that the sex if the research influenced the size of differences associated with specific psychological characteristics.

6. The process for conducting this type of literature review varies, depending on your purpose.

e.g. B. T. Anderson (1993) reviewed literature for the purpose of proposing a model intervention program to increase the representation of minority women in scientific careers.

7. When a literature review is conducted to provide a comprehensive understanding of what is known about a topic, the process is much longer.

e.g. Scarr and Eisenberg (1993) reviewed almost 200 references in their review of child care research.

Sample Collection

Sampling is the process of selecting observations (a sample) to provide an adequate description and inferences of the population. What you want to talk about what you actually observe in the data Population Sampling Frame Sampling Process Inference Sample. Sampling is the process of selecting units (e.g., people, organizations) from a population of interest so that by studying the sample we may fairly generalize our results back to the population from which they were chosen.

What is the purpose of taking a sample?

To draw conclusions about populations from samples, we must use inferential statistics which enables us to determine a population`s characteristics by directly observing only a portion (or **sample**) of the population. We obtain a sample rather than a complete enumeration (a census) of the population for many reasons. A sample is "a smaller (but hopefully representative) collection of units from a population used to determine truths about that population"

Why sample?

- Resources (time, money) and workload
- Gives results with known accuracy that can be calculated mathematically

The sampling frame is the list from which the potential respondents are drawn

- Registrar's office
- Class rosters
- Must assess sampling frame errors

Steps in the Sampling Process

- Specify Sampling Method: - The method by which the sampling units are to be selected is described.
- Determine sample Size: - The number of elements of the population to the sample is chosen.
- Specify Sampling Plan: - The operational procedures for selection of the sampling units are selected.
- Select the sample: - The office and field work necessary for the selections of the sample are carried out.

Data Analysis

Data Analysis is the process of systematically applying statistical and/or logical techniques to describe and illustrate, condense and recap, and evaluate data. An essential component of ensuring data integrity is the accurate and appropriate analysis of research findings.

Once you have selected the topic of the research and have gone through the process of literature survey, established your own focus of research, selected the research paradigm and methodology,

prepared your own research plan and have collected the data; the next step is analysis of the data collected, before finally writing the research report.

Data analysis is an ongoing activity, which not only answers your question but also gives you the directions for future data collection. Data analysis procedures (DAP) help you to arrive at the data analysis. The uses of such procedures put your research project in perspective and assist you in testing the hypotheses with which you have started your research. Hence with the use of DAP, you can

- convert data into information and knowledge, and
- explore the relationship between variables.

Understanding of the data analysis procedures will help you to

- appreciate the meaning of the scientific method, hypotheses testing and statistical significance in relation to research questions

- realise the importance of good research design when investigating research questions

- have knowledge of a range of inferential statistics and their applicability and limitations in the context of your research

- be able to devise, implement and report accurately a small quantitative research project

- be capable of identifying the data analysis procedures relevant to your research project

- show an understanding of the strengths and limitations of the selected quantitative and/or qualitative research project

- demonstrate the ability to use word processing, project planning and statistical computer packages in the context of a quantitative research project and report

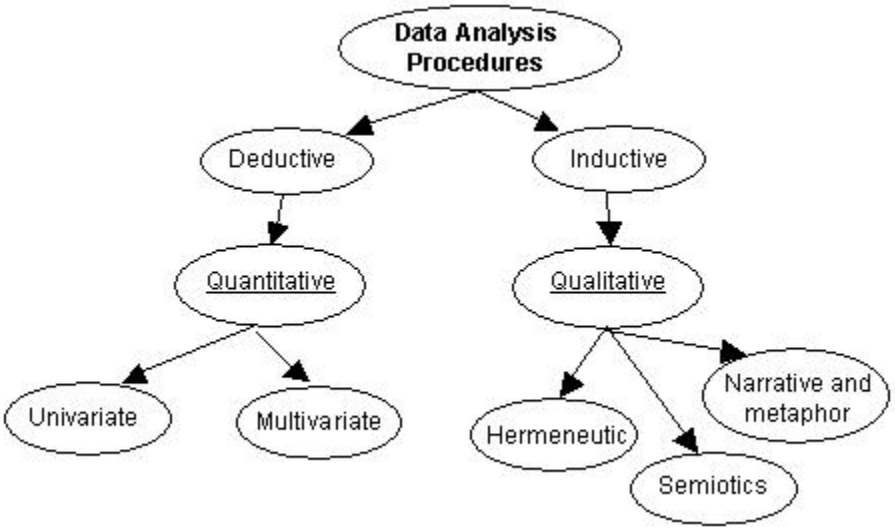
- be adept of working effectively alone or with others to solve a research question/ problem quantitatively.

The literature survey which you carried out guides you through the various data analysis methods that have been used in similar studies. Depending upon your research paradigm and methodology and the type of data collection, this also assists you in data analysis. Hence once you are aware of the fact that which particular procedure is relevant to your research project, you get the answers to

- What kinds of data analysis tools are identified for similar research investigations? and
- What data analysis procedures should you use for your purpose?

There are numerous ways under which data analysis procedures are broadly defined. Still there certain variable that just considered very important in data analysis. The following diagram makes it evident.

Figure 2.1 Flow Chart of Data Analysis



There are, in fact, a number of software packages available that facilitate data analysis. These include statistical packages like SPSS, SAS, and Microsoft Excel etc. Similarly tools like spreadsheets and word processing software are multipurpose and very useful for data analysis. The following links are useful for getting to know more about data analysis procedures and packages.

Report writing

Parts of a Research Paper (Thesis)

A. Preliminary Pages

- Cover page
- Approval Sheet
- Abstract
- Table of Contents
- List of Tables

Chapter 1 Introduction

- Background of the Study (includes significance of the study)
- Conceptual framework
- The Problem and hypotheses)

Chapter 2 Review of Literature

Chapter 3 Method and Procedures

- Research design
- Population (includes scope and delimitation of the study)
- Data-gathering procedures
- Data gathering tools (includes the description of the research instruments, Validity and Reliability of the instruments)
- Statistical tools

Chapter 4 Interpretation and Analysis of Findings

- Presentation of data
- Analysis and Interpretation
- Drawing implications out of the research findings
- Corroboration from related sources of information

Chapter 5 Conclusions and Recommendations

B. Appendices (References, forms/tools. Related articles published by the researcher / if required Curriculum Vitae)

C. Reference

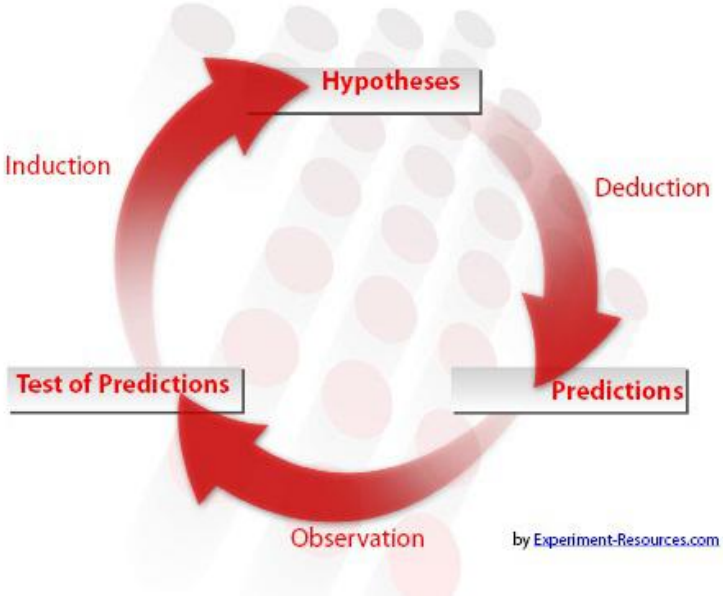
- APA format makes use of parenthetical citation (old format use latin citations – ibid; op. cit ; or loccit and endnotes or footnotes)

UNIT-III: HYPOTHESIS AND SAMPLING

Hypothesis: Definition, types and sources

A research hypothesis is the statement created by researchers when they speculate upon the outcome of a research or experiment. Every true experimental design must have this statement at the core of its structure, as the ultimate aim of any experiment. The hypothesis is generated via a number of means, but is usually the result of a process of inductive reasoning where observations lead to the formation of a theory. Scientists then use a large battery of deductive methods to arrive at a hypothesis that is testable, falsifiable and realistic.

Figure 3.1 Process of Framing Hypothesis



Proposition: - Proposition is a statement about the observable phenomenon (CONCEPTS) that may be judged as true or false.

Hypothesis: - When a proposition is formulated for empirical testing, we call it a hypothesis.

- Declarative statement about the relationship between two or more variables
- Hypothesis is of tentative and conjectural nature.
- Hypothesis has also been described as statements in which we assign variables to cases.

Types of Hypothesis

Descriptive hypothesis: - Descriptive hypothesis states the existence, size, form, or distribution of some variable.

Relational Hypothesis: - these are statements that describe a relation between two variables with respect to some case.

Correlation Hypothesis: - these states that the variables occur together in some specified manner without implying that one causes the other

Explanatory (Causal) Hypothesis: - there is an implication that the existence of or a change in one variable causes or leads to change in the other variable.

Sources of Hypothesis

This write-up will throw a light on the four important sources of hypothesis in social research, i.e, (1) General Culture in which a Science Develops, (2) Scientific Theory, (3) Analogies, and (4) Consequences of Personal, Idiosyncratic Experience as the Sources of Hypothesis.

1. General Culture in which a Science Develops

A cultural pattern influences the thinking process of the people and the hypothesis may be formulated to test one or more of these ideas. Cultural values serve to direct research interests. The function of culture has been responsible for developing today's science to a great dimension. In the words of Goode and Hatt, "to say that the hypotheses are the product of the cultural values does not make them scientifically less important than others, but it does at least indicate that attention has been called to them by the culture itself.

For example in the Western society race is thought to be an important determinant of human behaviour. Such a proposition can be used to formulate a hypothesis. We may also cite metaphysical bias and metaphysical ideas of Indian culture to have been responsible for the formulation of certain types of hypotheses. It implies that cultural elements of common cultural pattern may form a source of the formulation of hypotheses.

2. Scientific Theory

A major source of hypothesis is theory. A theory binds a large body of facts by positing a consistent and lawful relationship among a set of general concepts representing those facts. Further generalizations are formed on the basis of the knowledge of theory. Corollaries are drawn from the theories.

These generalizations or corollaries constitute a part of hypothesis. Since theories deal with abstractions which cannot be directly observed and can only remain in the thought process, a scientific hypothesis which is concerned with observable facts and observable relationship between facts can only be used for the purpose of selecting some of the facts as concrete instances of the concepts and for making a tentative statement about the existence of a relation among the selected facts with the purpose of subjecting the relation to an empirical test."

A hypothesis emerges as a deduction from theory. Hence, hypotheses become "working instruments of theory" Every worthwhile theory provides for the formulation of additional hypothesis. "The hypothesis is the backbone of all scientific theory construction; without it, confirmation or rejection of theories would be impossible."

The hypotheses when tested are "either proved or disproved and in turn constitute further tests of the original theory." Thus the hypothetical type of verbal proposition forms the link between the empirical propositions or facts and the theories. The validity of a theory can be examined only by means of scientific predictions or experimental hypothesis.

3. Analogies

Observation of a similarity between two phenomena may be a source of formation of a hypothesis aimed at testing similarity in any other respect. Julian Huxley has pointed out that "casual observation in nature or in the framework of another science may be a fertile source of hypothesis. The success of a system in one discipline can be used in other discipline also. The theory of ecology is based on the observation of certain plants in certain geographical conditions. As such, it remains in the domain of Botany. On the basis of that the hypothesis of human ecology could be conceived.

Hypothesis of social physics is also based on analogy. "When the hypothesis was born out by social observation, the same term was taken into sociology. It has become an important idea in sociological theory". Although analogy is not always considered, at the time of formulation of hypothesis; it is generally satisfactory when it has some structural analogies to other well established theories. For the systematic simplicity of our knowledge, the analogy of a hypothesis becomes inversely helpful. Formulation of an analogous hypothesis is construed as an achievement because by doing so its interpretation is made easy.

4. Consequences of Personal, Idiosyncratic Experience as the Sources of Hypothesis

Not only culture, scientific theory and analogies provide the sources of hypothesis, but also the way in which the individual reacts to each of these is also a factor in the statement of hypotheses. Certain facts are present, but every one of us is not able to observe them and formulate a hypothesis.

Referring to Fleming's discovery of penicillin, Backrach has maintained that such discovery is possible only when the scientist is prepared to be impressed by the 'unusual'. An unusual event struck Fleming when he noted that the dish containing bacteria had a green mould and the bacteria were dead. Usually he would have washed the dish and have attempted once again to culture the bacteria.

But normally, he was moved to bring the live bacteria in close contact with the green mould, resulting in the discovery of penicillin. The example of Sir Issac Newton, the discoverer of the theory of Gravitation, is another glaring example of this type of 'personal experience'. Although prior to Newton's observation, several persons had witnessed the falling of the apple, he was the right man to formulate the theory of gravitation on the basis of this phenomenon.

Thus emergence of a hypothesis is a creative manner. To quote Mc Guigan, "to formulate a useful and valuable hypothesis, a scientist needs first sufficient experience in that area, and second the quality of the genius." In the field of social sciences, an illustration of individual perspective may be visualized in Veblen's work. Thorstein Veblen's own community

background was replete with negative experiences concerning the functioning of economy and he was a 'marginal man', capable of looking at the capitalist system objectively.

Thus, he could be able to attack the fundamental concepts and postulates of classical economics and in real terms Veblen could experience differently to bear upon the economic world, resulting in the making of a penetrating analysis of our society. Such an excellent contribution of Veblen has, no doubt, influenced social science since those days.

Research Design: Meaning and types

It can simply be called as "A framework or blueprint for conducting the research project". It specifies the details of the procedures necessary for obtaining the information needed to structure and / or solve marketing research problems. A research design provides the framework for the collection and analysis of data. A choice of research design reflects decisions about the priority being given to a range of dimensions of the research process.

This involves

- Defining the problem/research question
- Review of related literature
- Planning the research
- What methodology will you use?
- What data do you want to use/produce?
- How feasible is your research approach?
- Ethical consideration

Steps in Research Design

- Define the information needed.
- Design the exploratory descriptive, and / or causal phases of the research
- Specify the measurement and scaling procedures.
- Construct and pretest a questionnaire (interviewing form) or an appropriate form for data collection.
- Specify the sampling process and sample size.
- Develop a plan of data analysis.

Reliability and validity

Issues of research reliability and validity need to be addressed in methodology chapter in a concise manner.

Reliability

Reliability refers to the extent to which the same answers can be obtained using the same instruments more than one time. In simple terms, if your research is associated with high levels

of reliability, then other researchers need to be able to generate the same results, using the same research methods under similar conditions. It is noted that "reliability problems crop up in many forms. Reliability is a concern every time a single observer is the source of data, because we have no certain guard against the impact of that observer's subjectivity". According to Wilson (2010) reliability issues are most of the time closely associated with subjectivity and once a researcher adopts a subjective approach towards the study, then the level of reliability of the work is going to be compromised.

Validity

Validity of research can be explained as an extent at which requirements of scientific research method have been followed during the process of generating research findings. Oliver (2010) considers validity to be a compulsory requirement for all types of studies. There are different forms of research validity and main ones are specified by Cohen et al (2007) as content validity, criterion-related validity, construct validity, internal validity, external validity, concurrent validity and face validity. Measures to ensure validity of a research include, but not limited to the following points:

- Appropriate time scale for the study has to be selected;
- Appropriate methodology has to be chosen, taking into account the characteristics of the study;
- The most suitable sample method for the study has to be selected;
- The respondents must not be pressured in any ways to select specific choices among the answer sets.

It is important to understand that although threats to research reliability and validity can never be totally eliminated, however researchers need to strive to minimize this threat as much as possible.

Sampling: Non Probability and Probability types

Sampling methods are classified as either probability or non-probability. In probability samples, each member of the population has a known non-zero probability of being selected. Probability methods include random sampling, systematic sampling, and stratified sampling.

Probability Samples

A probability sampling method is any method of sampling that utilizes some form of random selection. In order to have a random selection method, you must set up some process or procedure that assures that the different units in your population have equal probabilities of being chosen. The following sampling methods are examples of probability sampling

- Simple Random Sampling (SRS)
- Stratified Sampling.
- Cluster Sampling.
- Systematic Sampling.
- Multistage Sampling (in which some of the methods above are combined)

What is a Simple Random Sample?

A simple random sample is often mentioned in elementary statistics classes, but it's actually one of the least used techniques. In theory, it's easy to understand. However, in practice it's tough to perform.

Technically, a simple random sample is a set of n objects in a population of N objects where all possible samples are equally likely to happen. Here's a basic example of how to get a simple random sample: put 100 numbered bingo balls into a bowl (this is the population N). Select 10 balls from the bowl without looking (this is your sample n). Note that it's important not to look as you could (unknowingly) bias the sample. While the "lottery bowl" method can work fine for smaller populations, in reality you'll be dealing with much larger populations.

How to Perform Simple Random Sampling: Example

A larger population might be "All people who have had strokes in the United States." That list of participants would be extremely hard to obtain. Where would you get such a list in the first place? You could contact individual hospitals (of which there are thousands and thousands…) and ask for a list of patients (would they even supply you with that information? If you could somehow obtain this list then you will end up with a list of 800,000 people which you then have to put into a "bowl" of some sort and choose random people for your sample. This type of situation is the type of real-life situation you'll come across and is what makes getting a simple random sample so hard to undertake.

Sample question: Outline the steps for obtaining a simple random sample for outcomes of strokes in GH hospitals.

Step 1: Make a list of all the trauma hospitals in the GH (there are several hundred: the CDC keeps a list).

Step 2: Assign a sequential number to each trauma center (1,2,3…n). This is your sampling frame (the list from which you draw your simple random sample).

Step 3: Figure out what your sample size is going to be. See: (Sample size) (how to find one).

Step 4: Use a random number generator to select the sample, using your sampling frame (population size) from Step 2 and your sample size from Step 3. For example, if your sample size is 50 and your population is 500, generate 50 random numbers between 1 and 500.

Warning: If you compromise (say, by not including ALL trauma centers in your sampling frame), it could open your results to bias.

Simple Random Sample vs. Random Sample

A simple random sample is similar to a random sample. The difference between the two is that with a simple random sample, each object in the population has an equal chance of being chosen. With random sampling, each object does not necessarily have an equal chance of being chosen.

Unequal probability sampling isn't usually addressed in basic statistics courses, but if you're interested in an example of when it might be used, read this article.

Stratified sampling is a method of sampling from a population in statistics. When sub-populations vary considerably, it is advantageous to sample each subpopulation (stratum) independently. Stratification is the process of grouping members of the population into relatively homogeneous subgroups before sampling.

Systematic sampling relies on arranging the target population according to some ordering scheme and then selecting elements at regular intervals through that ordered list.

- Systematic sampling involves a random start and then proceeds with the selection of every *k*th element from then onwards. In this case, *k*=(population size/sample size).
- It is important that the starting point is not automatically the first in the list, but is instead randomly chosen from within the first to the *k*th element in the list.
- A simple example would be to select every 10th name from the telephone directory (an 'every 10th' sample, also referred to as 'sampling with a skip of 10').

Cluster sampling is an example of 'two-stage sampling'

- First stage a sample of areas is chosen;
- Second stage a sample of respondents *within* those areas is selected.
- Population divided into clusters of homogeneous units, usually based on geographical contiguity.
- Sampling units are groups rather than individuals.
- A sample of such clusters is then selected.
- All units from the selected clusters are studied.

Multi-stage Sampling

- Complex form of cluster sampling in which two or more levels of units are embedded one in the other.
- First stage, random number of districts chosen in all states.
- Followed by random number of talukas, villages. Then third stage units will be houses.
- All ultimate units (houses, for instance) selected at last step are surveyed.
- This technique is essentially the process of taking random samples of preceding random samples.
- Not as effective as true random sampling, but probably solves more of the problems inherent to random sampling.
- An effective strategy because it banks on multiple randomizations. As such, extremely useful.
- Multistage sampling used frequently when a complete list of all members of the population not exists and is inappropriate.
- Moreover, by avoiding the use of all sample units in all selected clusters, multistage sampling avoids the large, and perhaps unnecessary, costs associated with traditional cluster sampling.

Multi-Phase Sampling

- Part of the information collected from whole sample & part from subsample.
- In Tb survey MT in all cases – Phase I
- X –Ray chest in MT +ve cases – Phase II
- Sputum examination in X – Ray +ve cases - Phase III
- Survey by such procedure is less costly, less laborious & more purposeful

Non-Probability Samples

Any sampling method where some elements of population have *no* chance of selection (these are sometimes referred to as 'out of coverage'/'undercovered'), or where the probability of selection can't be accurately determined. It involves the selection of elements based on assumptions regarding the population of interest, which forms the criteria for selection. Hence, because the selection of elements is nonrandom, non-probability sampling not allows the estimation of sampling errors..

Example: We visit every household in a given street, and interview the first person to answer the door. In any household with more than one occupant, this is a non-probability sample, because some people are more likely to answer the door (e.g. an unemployed person who spends most of their time at home is more likely to answer than an employed housemate who might be at work when the interviewer calls) and it's not practical to calculate these probabilities.

- Convenience sample
- Purposive sample
- Quota
- Accidental

Judgmental or Purposive Sampling

The researcher chooses the sample based on who they think would be appropriate for the study. This is used primarily when there is a limited number of people that have expertise in the area being researched

Convenience Sampling

- Sometimes known as grab or opportunity sampling or accidental or haphazard sampling.
- A type of non-probability sampling which involves the sample being drawn from that part of the population which is close to hand. That is, readily available and convenient.
- The researcher using such a sample cannot scientifically make generalizations about the total population from this sample because it would not be representative enough.
- For example, if the interviewer was to conduct a survey at a shopping center early in the morning on a given day, the people that he/she could interview would be limited to those given there at that given time, which would not represent the views of other members of society in such an area, if the survey was to be conducted at different times of day and several times per week.
- This type of sampling is most useful for pilot testing.

- In social science research, snowball sampling is a similar technique, where existing study subjects are used to recruit more subjects into the sample.

Purposive Sample

A purposive sample is a non-probability sample that is selected based on characteristics of a population and the objective of the study. Purposive sampling is also known as judgmental, selective, or subjective sampling.

This type of sampling can be very useful in situations when you need to reach a targeted sample quickly, and where sampling for proportionality is not the main concern. There are seven types of purposive samples, each appropriate to a different research objective.

Types of Purposive Samples

A maximum variation/heterogeneous purposive sample is one which is selected to provide a diverse range of cases relevant to a particular phenomenon or event. The purpose of this kind of sample design is to provide as much insight as possible into the event or phenomenon under examination. For example, when conducting a street poll about an issue, a researcher would want to ensure that he or she speaks with as many different kinds of people as possible in order to construct a robust view of the issue from the public's perspective.

A homogeneous purposive sample is one that is selected for having a shared characteristic or set of characteristics. For example, a team of researchers wanted to understand what the significance of white skin--whiteness--means to white people, so they asked white people about this. This is a homogenous sample created on the basis of race.

Typical case sampling is a type of purposive sampling useful when a researcher wants to study a phenomenon or trend as it relates to what are considered "typical" or "average" members of the effected population. If a researcher wants to study how a type of educational curriculum affects the average student, then he or she chooses to focus on average members of a student population.
Conversely, extreme/deviant case sampling is used when a researcher wants to study the outliers that diverge from the norm as regards a particular phenomenon, issue, or trend. By studying the deviant cases, researchers can often gain a better understanding of the more regular patterns of behavior.

If a researcher wanted to understand the relationship between study habits and high academic achievement, he or she should purposively sample students considered high achievers.
Critical case sampling is a type of purposive sampling in which just one case is chosen for study because the researcher expects that studying it will reveal insights that can be applied to other like cases. When sociologist C.J. Pascoe wanted to study sexuality and gender identity develop among high school students, she selected what was considered to be an average high school in terms of population and family income, so that her findings from this case could be more generally applicable.

With total population sampling a researcher chooses to examine the entire population that has one or more shared characteristics. This kind of purposive sampling technique is commonly used to generate reviews of events or experiences, which is to say, it is common to studies of particular groups within larger populations.

Expert sampling is a form of purposive sampling used when research requires one to capture knowledge rooted in a particular form of expertise. It is common to use this form of purposive sampling technique in the early stages of a research process, when the researcher is seeking to become better informed about the topic at hand before embarking on a study. Doing this kind of early-stage expert-based research can shape research questions and research design in important ways.

Quota Sampling

- The population is first segmented into mutually exclusive sub-groups, just as in stratified sampling.
- Then judgment used to select subjects or units from each segment based on a specified proportion.
- For example, an interviewer may be told to sample 200 females and 300 males between the age of 45 and 60.
- It is this second step which makes the technique one of non-probability sampling.
- In quota sampling the selection of the sample is non-random.
- For example interviewers might be tempted to interview those who look most helpful. The problem is that these samples may be biased because not everyone gets a chance of selection. This random element is its greatest weakness and quota versus probability has been a matter of controversy for many years

Panel Sampling

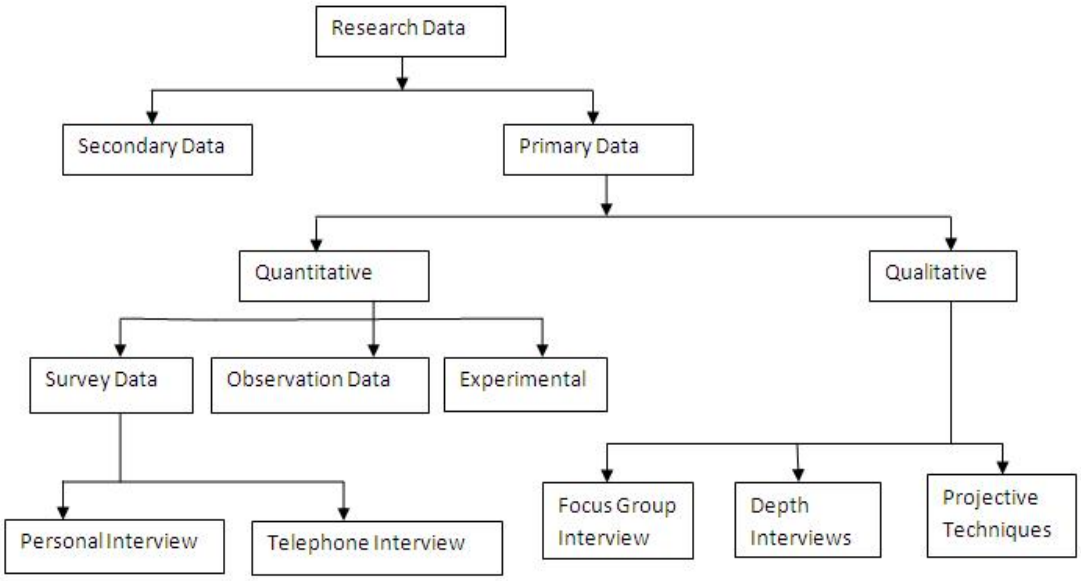
- Method of first selecting a group of participants through a random sampling method and then asking that group for the same information again several times over a period of time.
- Therefore, each participant is given same survey or interview at two or more time points; each period of data collection called a "wave".
- This sampling methodology often chosen for large scale or nation-wide studies in order to gauge changes in the population with regard to any number of variables from chronic illness to job stress to weekly food expenditures.
- Panel sampling can also be used to inform researchers about within-person health changes due to age or help explain changes in continuous dependent variables such as spousal interaction.
- There have been several proposed methods of analyzing panel sample data, including growth curves.

In addition, non-response effects may turn *any* probability design into a non-probability design if the characteristics of non-response are not well understood, since non-response effectively modifies each element's probability of being sampled.

Methods of Data Collection

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate outcomes. Data collection possibilities are wide and varied with any one method of collection not inherently better than any other. Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, and evaluate outcomes. The data collection component of research is common to all fields of study including physical and social sciences, humanities, business, etc. While methods vary by discipline, the emphasis on ensuring accurate and honest collection remains the same.

Table 3.1 Means of Data Collection



The importance of ensuring accurate and appropriate data collection

Regardless of the field of study or preference for defining data (quantitative, qualitative), accurate data collection is essential to maintaining the integrity of research. Both the selection of appropriate data collection instruments (existing, modified, or newly developed) and clearly delineated instructions for their correct use reduce the likelihood of errors occurring.

Consequences from improperly collected data include

- inability to answer research questions accurately
- inability to repeat and validate the study
- distorted findings resulting in wasted resources
- misleading other researchers to pursue fruitless avenues of investigation

- compromising decisions for public policy
- causing harm to human participants and animal subjects

While the degree of impact from faulty data collection may vary by **discipline** and the nature of investigation, there is the potential to cause disproportionate harm when these research results are used to support public policy recommendations.

- Data Collection Strategies
- Characteristics of Good Measures
- Quantitative and Qualitative Data
- Tools for Collecting Data

The Data Collection Process

All methods of collection require rigorous and systematic design and execution that includes

- thorough planning
- well considered development
- effective piloting
- weighed modification
- deliberate implementation and execution
- appropriate management and analysis

Secondary Data

Secondary data – data someone else has collected

– This is what you were looking for in your assignment.

Examples

- County health departments
- Vital Statistics – birth, death certificates
- Hospital, clinic, school nurse records
- Private and foundation databases
- City and county governments
- Surveillance data from state government programs
- Federal agency statistics - Census, NIH, etc.

Primary Data

Data – data you collect

- Surveys
- Focus groups
- Questionnaires
- Personal interviews
- Experiments and observational study

Surveys

- Surveying involves gathering information from individuals using a questionnaire
- Surveys can
  - reach a large number of respondents
  - generate standardized, quantifiable, empirical data - as well as some qualitative data and offer confidentiality / anonymity
- Designing survey instruments capable of generating credible data, however, can be difficult

Survey Types

Surveys can be
- descriptive or explanatory
- involve entire populations or samples of populations
- capture a moment or map trends
- can be administered in a number of ways

Focus Group

- The first step will involve deciding on the purpose of the focus group in relation to how and why other methods are being used to inform your evaluation questions.
- Preparation of a question guide will help you to keep the discussion reasonably focused. Prompt cards with statements asking students to rate the extent to which they agree or disagree with them on a Likert scale might be a useful way of getting some measure of the diversity of perspectives, as a precursor to more open discussion of the underlying explanation for these ratings.
- You will probably have to rely on volunteers for the focus group who will have appropriate experience to express opinions, perspectives, attitudes, emotions etc about these experiences that will be informative to the evaluation. A group of around 7 is ideal, and it is recommended that there should be no more than 12. Any more than this and the discussion is likely to become unmanageable. The risk from having too few students is that their perspective will not fairly reflect the range of perspectives in the population as a whole.
- If the discussion is not scheduled in class time, a suitable venue will need to be booked and participants informed of arrangements.
- It can be difficult to facilitate the discussion and note-take at the same time. If you do not have access to suitable or reliable equipment to record the discussion, or it may be too sensitive to record, you may want to think about enlisting assistance to help facilitate or take notes.
- Although it is usual to approach the discussion with a draft guide of questions, a flexible and adaptable approach will help you follow-up unanticipated but relevant issues that emerge. It can be a difficult balance to both keep the discussion reasonably focused but also allow sufficient freedom for different issues and perspectives to emerge.
- Analysis of the data will depend on qualitative approaches which help to identify key themes in student opinions, attitudes, feelings etc about their learning experience.

Observation, Questionnaires and Personal interviews

Observation

Observational research (or field research) is a type of correlational (i.e., non-experimental) research in which a researcher observes ongoing behavior. There are a variety of types of observational research, each of which has both strengths and weaknesses.

It is typically divided into naturalistic (or "nonparticipant") observation, and participant **observation**. Cases studies and archival research are special types of observational research. Naturalistic (or nonparticipant) observation has no intervention by a researcher.

Questionnaire

A **questionnaire** is a **research** instrument consisting of a series of questions and other prompts for the purpose of gathering information from respondents. Although they are often designed for statistical analysis of the responses, this is not always the case.

Types of questionnaire

1. Structured questionnaire
a) Have definite and concrete questions.
b) Is prepared well in advance.
c) Initiates a formal inquiry.
d) Supplements and checks the data, previously accumulated.
e) Used in studies of the economics and the social problems, studies of the administrative policies and changes etc.

2. Unstructured questionnaire

a) Used at the time of the interview.
b) Acts as the guide for the interviewer.
c) Is very flexible in working.
d) Used in studies related to the group of families or those relating to the personal experiences, beliefs etc.

A questionnaire can also be divided as the follows depending on the nature of the questions therein

i. Open ended questionnaire
a) Respondent is free to express his views and the ideas.
b) Used in making intensive studies of the limited number of the cases.
c) Merely an issue is raised by such a questionnaire.
d) Do not provide any structure for the respondent's reply.
e) The questions and their orders are pre – determined in the nature.

ii. Close ended questionnaire
a) Responses are limited to the stated alternatives.
b) One of the alternatives is simply YES or NO.
c) Respondent cannot express his own judgment.

iii. Mixed questionnaire
a) Questions are both close and open ended.
b) Used in field of social research.

iv. Pictorial questionnaire
a) Used very rarely.
b) Pictures are used to promote the interest in answering the questions.
c) Used in studies related to the social attitudes and the pre – judices in the children.

Interview

An **interview** in qualitative **research** is a conversation where questions are asked to elicit information. The interviewer is usually a professional or paid researcher, sometimes trained, who poses questions to the interviewee, in an alternating series of usually brief questions and answers.

Types of Interview

- Personal Interview
- Telephone Interview
- Focus Group Interview
- Depth Interview
- Projective Techniques

Structured Approach
- All data collected in the same way
- Especially important for multi-site and cluster evaluations so you can compare
- Important when you need to make comparisons with alternate interventions

Use Structured Approach When
- need to address extent questions
- have a large sample or population
- know what needs to be measured
- need to show results numerically
- need to make comparisons across different sites or interventions

Semi-structured Approach
- Systematic and follow general procedures but data are not collected in exactly the same way every time
- More open and fluid
- Does not follow a rigid script

- may ask for more detail
- people can tell what they want in their own way

Use Semi-structured Approach when

- conducting exploratory work
- seeking understanding, themes, and/or issues
- need narratives or stories
- want in-depth, rich, "backstage" information
- seek to understand results of data that are unexpected

Characteristics of Good Measures

- Is the measure relevant?
- Is the measure credible?
- Is the measure valid?
- Is the measure reliable?

Participatory Methods

Involve groups or communities heavily in data collection

Examples:

- community meetings
- mapping
- transect walks

Community Meetings

One of the most common participatory methods

- Must be well organized
  - agree on purpose
  - establish ground rules
    - who will speak
    - time allotted for speakers
    - format for questions and answers

Mapping

- Drawing or using existing maps
- Useful tool to involve stakeholders
  - increases understanding of the community
  - generates discussions, verifies secondary sources of information, perceived changes

Types of Mapping

– natural resources, social, health, individual or civic assets, wealth, land use, demographics

Transect Walks

- Evaluator walks around community observing people, surroundings, and resources
- Need good observation skills
- Walk a transect line through a map of a community — line should go through all zones of the community

Observation

See what is happening
- traffic patterns
- land use patterns
- layout of city and rural areas
- quality of housing
- condition of roads
- conditions of buildings
- who goes to a health clinic

Observation is Helpful when

- need direct information
- trying to understand ongoing behavior
- there is physical evidence, products, or outputs than can be observed
- need to provide alternative when other data collection is infeasible or inappropriate

Degree of Structure of Observations

- Structured: determine, before the observation, precisely what will be observed before the observation
- Unstructured: select the method depending upon the situation with no pre-conceived ideas or a plan on what to observe
- Semi-structured: a general idea of what to observe but no specific plan

Ways to Record Information from Observations

- Observation guide
  - printed form with space to record
- Recording sheet or checklist
  - Yes/no options; tallies, rating scales
- Field notes
  - least structured, recorded in narrative, descriptive style

Mail / Phone / Internet Surveys

- Literacy issues
- Consider accessibility
  - reliability of postal service
  - turn-around time
- Consider bias
  - What population segment has telephone access? Internet access?

Interviews

- Often semi-structured
- Used to explore complex issues in depth
- Forgiving of mistakes: unclear questions can be clarified during the interview and changed for subsequent interviews
- Can provide evaluators with an intuitive sense of the situation

Diaries and Self-Reported Checklists

- Use when you want to capture information about events in people's daily lives
- Participants capture experiences in real-time not later in a questionnaire
- Used to supplement other data collection

Delphi Technique

- Enables experts to engage remotely in a dialogue and reach consensus, often about priorities
- Experts asked specific questions; often rank choices
- Responses go to a central source, are summarized and fed back to the experts without attribution
- Experts can agree or argue with others' comments
- Process may be iterative

Pilot study

Pilot study is a small scale preliminary study conducted in order to evaluate feasibility, time, cost, adverse events, and affect size (Statistical variability) in an attempt to predict an appropriate sample size and improve upon the study design prior to performance of a full scale research project.

Objectives

- Pilot study is a small experiment designed to test logistics
- Gather information prior to a large study
- Improve the actual study's quality and efficiency

- Reveal deficiencies in the design of a proposed experiment or procedure and these can then be addressed before time
- A good research strategy requires careful planning and a pilot study will often be a part of this strategy

Advantages

- Carried out before large scale quantitative research in an attempt to avoid time and money being wasted on an inadequately designed project.

- It is a potentially valuable insight and should anything be missing in the pilot study, it can be added to the full scale (and more expensive) experiment to improve the chances of a clear outcome.

- Pilot experiments are used to sell a product and provide quantitative proof that the system has potential to succeed on a full scale basis.

- In social science, pilot studies can be referred to as small scale studies that will help identify design issues before the main research is done.

- It permits preliminary testing of the hypothesis that leads to testing more precise hypothesis in the main study. It may lead to changing some hypothesis, dropping some or developing new hypothesis.

- It often provides the researcher with ideas, approaches, and clues you may not have foreseen before conducting the pilot study.

- It permits a thorough check of the planned statistical and analytical procedures, giving you a chance to evaluate their usefulness for the data. You may then be able to make needed alterations in the data collecting methods, and therefore, analyze data in the main study more efficiently.

- It can greatly reduce the number of unanticipated problems because you have an opportunity to redesign parts of your study to overcome difficulties that pilot study reveals.

- In the pilot study, the researcher may try out a number of alternative measures and then select those that produce the clearest results for the main study.

Primary Scales of Measurement

SCALE	BASIC CHARACTERISTICS	COMMON EXAMPLES	MARKETING EXAMPLES	PERMISSIBLE STATISTICS	
				DESCRIPTIVE	INFERENTIAL
Nominal	Numbers identify and classify objects	Social Security numbers, numbering of football players	Brand numbers, store types, sex classification	Percentages, mode	Chi-square, binomial test
Ordinal	Numbers indicate the relative positions of the objects but not the magnitude of differences between them	Quality rankings, rankings or teams in a tournament	Preference rankings, market position, social class	Percentile median	Rank order correlation, Friedman ANOVA
Interval	Difference between objects can be compared; zero point is arbitrary	Temperature (Fahrenheit Centigrade)	Attitudes, opinions, index numbers	Range, mean, standard deviation	Product-moment correlations, t-tests, ANOVA, regression, factor analysis
Ratio	Zero point is fixed; ratios of scale values can be computed	Length, weight	Age, income, costs, sales, market shares	Geometric mean, harmonic mean	Coefficient of variation

- Nominal (observations are categorized or sorted based on defined properties; each category is distinct, mutually exclusive, and exhaustive. Ex. Gender, Religious Affiliation, College major, hair color, birthplace, nationality, tribe)

- Ordinal (scores or observations are ranked in order without distance between individuals. Ex. Age group when ranked, Socio-Economic-Status, Level of Conflict)

- Interval (with equal intervals between numbers where there is no absolute absence of the attribute because zero is assigned and represents an arbitrary point. Ex. Temperature, IQ score)

- Ratio (this is in contrast to interval where there exist an absolute absence of the attribute or rational zero. Ex. Age, height, weight, length of time)

What are the Methods and Techniques used in Criminology?

While scientific methods are basically alike for all sciences, scientific techniques differ, for techniques are the particular ways in which scientific methods are applied to a particular problem. Each science, therefore, develops a series of techniques which fit the body of material studies. What are the techniques used in criminological research? The criminologists generally use

survey method, case study method, and statistical method in studying criminal behaviour. Occasionally, experimental method is also used.

The survey method collects facts by putting questions to a large number of persons under scientific controls. The three tools often used in this technique are questionnaire, schedule and interview guide. A questionnaire is filled out by the informant personally, while a schedule is filled out by a trained investigator.

Questions in both are pre-structured. An interview guide consists of only points on which questions are put to the informants; it does not have structured questions. These tools have their pitfalls. The informants may not understand a question; they may pick up an answer even though they may not have any firm opinion on the matter; they may give an 'acceptable' answer rather than the real one; or they may be swayed by the way the question is worded.

Even though these tools may have a margin of error, yet they are very useful, for they are more reliable than guesswork. While non-participant observation may be used in criminology, the use of participant observation is almost impractical. The participant observer seeks insight by wishing to take part in whatever is being studied, and, this is not feasible in criminological studies.

Case Study Method

The case study method is a method of studying social phenomena through an intensive and in-depth analysis of an individual case. The case may be a person (a juvenile delinquent), a group (youth criminals), an institution (Borstal School), an event (riot in prison), a situation (collective violence), an organisation (the police), or any other unit of social life. This method provides an opportunity for a thorough analysis of many specific details that are often overlooked in other methods.

The investigator delves into the social, medical, psychological and sometimes psychiatric and mental background of the individual. Attention is given to attitudes and perceptions as well as to behaviour, following the theory of W.I. Thomas that "if individuals define situations as real, they are real in their consequences". Information is gathered from families, school records, neighbours, peers, work-groups, and other sources; numerous interviews with the individual are also made.

This approach rests on the assumption that the case being studied is typical of cases of a certain type so that, through intensive analysis, generalizations may be made which will be applicable to other cases of the same type. The greatest value of the case study is in the suggestion of hypotheses which can then be tested by other methods. Much of our reliable knowledge about juvenile delinquency, for instance, has developed through the testing of hypotheses which were suggested by early case studies of delinquents (Thomas, 1923; Shaw, 1931).

Similarly, much of our present knowledge about female criminals in India stems from hypotheses suggested by Ram Ahuja (1969) and others. These hypotheses are often not tested by the case study method but by other methods.

Advantages of Case Study

There are a number of advantages in using case studies. First, the examination of the data is most often conducted within the context of its use (Yin, 1984), that is, within the situation in which the activity takes place. A case study might be interested, for example, in the process by which a subject comprehends an authentic text. To explore the strategies the reader uses, the researcher must observe the subject within her environment, such as reading in classroom or reading for leisure. This would contrast with experiment, for instance, which deliberately isolates a phenomenon from its context, focusing on a limited number of variables (Zaidah, 2003). Second, variations in terms of intrinsic, instrumental and collective approaches to case studies allow for both quantitative and qualitative analyses of the data. Some longitudinal studies of individual subjects, for instance, rely on qualitative data from journal writings which give descriptive accounts of behaviour. On the other hand, there are also a number of case studies which seek evidence from both numerical and categorical responses of individual subjects.

While cautions researchers not to confuse case studies with qualitative research, he also notes that "case studies can be based … entirely on quantitative evidence". Third, the detailed qualitative accounts often produced in case studies not only help to explore or describe the data in real-life environment, but also help to explain the complexities of real life situations which may not be captured through experimental or survey research.

A case study of reading strategies used by an individual subject, for instance, can give access to not only the numerical information concerning the strategies used, but also the reasons for strategy use, and how the strategies are used in relation to other strategies. As reading behaviours involve complex cognitive processes, each reading strategy cannot be examined in isolation but rather in relation to other strategies.

How to Design and Conduct a Case Study?

The advantage of the case study research design is that you can focus on specific and interesting cases. This may be an attempt to test a theory with a typical case or it can be a specific topic that is of interest. Research should be thorough and note taking should be meticulous and systematic.

The first foundation of the case study is the subject and relevance. In a case study, you are deliberately trying to isolate a small study group, one individual case or one particular population. For example, statistical analysis may have shown that birthrates in African countries are increasing.

A case study on one or two specific countries becomes a powerful and focused tool for determining the social and economic pressures driving this.
In the design of a case study, it is important to plan and design how you are going to address the study and make sure that all collected data is relevant. Unlike a scientific report, there is no strict set of rules so the most important part is making sure that the study is focused and concise; otherwise you will end up having to wade through a lot of irrelevant information.

It is best if you make yourself a short list of 4 or 5 bullet points that you are going to try and address during the study. If you make sure that all research refers back to these then you will not be far wrong. With a case study, even more than a questionnaire or survey, it is important to be passive in your research. You are much more of an observer than an experimenter and you must remember that, even in a multi-subject case, each case must be treated individually and then cross case conclusions can be drawn.

How to Analyze the Results?

Analyzing results for a case study tends to be more opinion based than statistical methods. The usual idea is to try and collate your data into a manageable form and construct a narrative around it. Use examples in your narrative whilst keeping things concise and interesting. It is useful to show some numerical data but remember that you are only trying to judge trends and not analyze every last piece of data. Constantly refer back to your bullet points so that you do not lose focus.
It is always a good idea to assume that a person reading your research may not possess a lot of knowledge of the subject so try to write accordingly. In addition, unlike a scientific study which deals with facts, a case study is based on opinion and is very much designed to provoke reasoned debate. There really is no right or wrong answer in a case study.

Statistical Methods

The statistical method enables us to reduce a complex mass of data to simple units of measurement. For example, the researcher can calculate an average (a mean, a median, or a mode) and point out the central tendency of the whole group of items and can also give another figure (standard deviation) to measure the dispersion of items around the central tendency.

For example, a criminologist can collect data pertaining to the period for which criminals remain under trial and are detained in prisons without any work, or the period for which petty offenders are imprisoned, pointing out the lack of utility of such imprisonment in the reformation of criminals. By means of statistical device, known as coefficient of correlation, a criminologist can also compare the figures of one group with those of other groups in order, to reveal potential relationships.

In statistical technique, quite frequently the correlation technique is used in criminology. Investigators pick up a variable considered to be relevant to the study of crime or delinquency and measure the correlation between that variable and crime. The correlation coefficient measures the degree to which variables occur together. Perfect correlation is 1.00, and the direction may be positive or negative. For example, age and vision are usually negatively correlated, i.e., as age increases, vision decreases.

In criminology, suppose in one study, we find a correlation coefficient of -0.51 between median years of education and juvenile delinquency, and a correlation coefficient of +0.70 between the rate of delinquency and the percentage of low-caste persons, it would mean that as the rate of delinquency increases, the median years of education decrease.
But this relationship is not as strong as the relationship between class and delinquency. The strong correlation found by Kewalramani (Child Abuse, 1991) in his study of sexual abuse of

children was negative to correlation of -0.56 between poverty and rate of sexual abuse. That is, as the percentage of children who lived in well-off homes increased, the rate of sexual abuse of children decreased.

The criticism against the statistical technique is that it can give us bases upon which to predict delinquency/crime but it tells us nothing about causation. For example, the correlation between the social contacts a prisoner is able to maintain and his adjustment in prison is high. We can predict that one will occur when the other occurs; but we would certainly not argue that maladjustment in prison is caused by lack of social contacts.

More likely, in many cases of correlation, factors other than those correlated may be the 'cause'. Could we say that because more rapes occur in homes than in streets, sleeping in the streets is safer for women than sleeping at home? This type of correlation ship will be ludicrous. Other factors have to be considered in the analysis of statistics. A correlation only directs inquiry; it does not show a causal relationship. A sociologist has wisely said: "Where there is causation, there is also correlation, but where there is correlation, and there may be no corresponding causation." In short, statistical analyses are used by criminologists only to measure the frequency with which certain factors such as age, income, education, intelligence, and others occur with regard to delinquency and crime.

Experimental method

The experimental method, often used by physical scientists, is not as adaptable to use in the social sciences because control is more difficult to achieve. The closest the criminologist has come to the use of this method is through controlled observations. We may take an example: Suppose we attempt to measure the differences between short-term and long-term methods of detention of male juvenile delinquents. Knowing the failure of traditional methods used in the reformation of juveniles in Children Homes, a new type of treatment facility is introduced in a selected Home that would treat boys for a short period, say, up to three months. The new programme could be a guided group interaction combined with work and recreation programmes, all aimed at changing the self concept of the juvenile. Its results then would be compared with those of the longer treatment period at the same Home where the sentences ran from twelve months to eighteen months.

Thus, in an experimental method, ideally an experiment begins with two or more equivalent groups and an experimental variable is introduced only in the 'experimental group'. The 'control group' does not experience the experimental variable. The investigator measures the phenomenon under study before and after the introduction of the experimental variable, thus getting a measure of the change presumably caused by the variable.

In using the experimental method in criminology, the investigator encounters the problem of selecting equivalent groups which proves to be an insurmountable problem. Thus, the difficulty of equalizing the experimental and control groups, the difficulty in getting sufficient data, the difficulty of isolating all variables that are not being tested are only some of the problems which make the use of the experimental design in delinquency and crime research doubtful.

All the above-mentioned methods have thus relative merits as well as limitations. As such, no one method can be described as superior to the other. Each is important for its own purpose. The merit of the survey method lies in the fact that it makes the study of criminal behaviour in its natural setting (family, prison, police station, court, correctional reformatory, etc.) eminently practicable; but the great weakness of this method is that the convicted and the incarcerated criminals often do not give accurate information. They only say what they think a researcher wants to hear, either because they just wish to get rid of the researcher, or because they think that the researcher might help them with the official agency. However, the greatest advantage of this method is that researchers get out of their armchairs and move into the field and study criminals in their own environment.

Unobtrusive Measures

Unobtrusive research (or unobtrusive measures) is a method of data collection used primarily in the social sciences. The term "unobtrusive measures" was first coined by Webb, Campbell, Schwartz, & Sechrest in a 1966 book titled Unobtrusive Measures: nonreactive research in the social sciences. Unobtrusive measures are measures that don't require the researcher to intrude in the research context. Direct and participant observation requires that the researcher be physically present. This can lead the respondents to alter their behavior in order to look good in the eyes of the researcher. A questionnaire is an interruption in the natural stream of behavior. Respondents can get tired of filling out a survey or resentful of the questions asked.

Unobtrusive measurement presumably reduces the biases that result from the intrusion of the researcher or measurement instrument. However, unobtrusive measures reduce the degree the researcher has control over the type of data collected. For some constructs there may simply not be any available unobtrusive measures.

Three types of unobtrusive measurement are discussed here

Indirect Measures

An indirect measure is an unobtrusive measure that occurs naturally in a research context. The researcher is able to collect the data without introducing any formal measurement procedure. The types of indirect measures that may be available are limited only by the researcher's imagination and inventiveness. For instance, let's say you would like to measure the popularity of various exhibits in a museum. It may be possible to set up some type of mechanical measurement system that is invisible to the museum patrons.

In one study, the system was simple. The museum installed new floor tiles in front of each exhibit they wanted a measurement on and, after a period of time, measured the wear-and-tear of the tiles as an indirect measure of patron traffic and interest. We might be able to improve on this approach considerably using electronic measures. We could, for instance, construct an electrical device that senses movement in front of an exhibit. Or we could place hidden cameras and code patron interest based on videotaped evidence.

One of my favorite indirect measures occurred in a study of radio station listening preferences. Rather than conducting an obtrusive survey or interview about favorite radio stations, the researchers went to local auto dealers and garages and checked all cars that were being serviced to see what station the radio was currently tuned to. In a similar manner, if you want to know magazine preferences, you might rummage through the trash of your sample or even stage a door-to-door magazine recycling effort.

These examples illustrate one of the most important points about indirect measures - you have to be very careful about the ethics of this type of measurement. In an indirect measure you are, by definition, collecting information without the respondent's knowledge. In doing so, you may be violating their right to privacy and you are certainly not using informed consent. Of course, some types of information may be public and therefore not involve an invasion of privacy.

There may be times when an indirect measure is appropriate, readily available and ethical. Just as with all measurement, however, you should be sure to attempt to estimate the reliability and validity of the measures. For instance, collecting radio station preferences at two different time periods and correlating the results might be useful for assessing test-retest reliability. Or, you can include the indirect measure along with other direct measures of the same construct (perhaps in a pilot study) to help establish construct validity.

Content Analysis

Content analysis is the analysis of text documents. The analysis can be quantitative, qualitative or both. Typically, the major purpose of content analysis is to identify patterns in text. Content analysis is an extremely broad area of research. It includes: Thematic analysis of text.

The identification of themes or major ideas in a document or set of documents. The documents can be any kind of text including field notes, newspaper articles, technical papers or organizational memos.

- Indexing

There are a wide variety of automated methods for rapidly indexing text documents. For instance, Key Words in Context (KWIC) analysis is a computer analysis of text data. A computer program scans the text and indexes all key words. A key word is any term in the text that is not included in an exception dictionary. Typically you would set up an exception dictionary that includes all non-essential words like "is", "and", and "of". All key words are alphabetized and are listed with the text that precedes and follows it so the researcher can see the word in the context in which it occurred in the text. In an analysis of interview text, for instance, one could easily identify all uses of the term "abuse" and the context in which they were used.

Quantitative descriptive analysis

Here the purpose is to describe features of the text quantitatively. For instance, you might want to find out which words or phrases were used most frequently in the text. Again, this type of analysis is most often done directly with computer programs.

Content analysis has several problems you should keep in mind. First, you are limited to the types of information available in text form. If you are studying the way a news story is being handled by the news media, you probably would have a ready population of news stories from which you could sample. However, if you are interested in studying people's views on capital punishment, you are less likely to find an archive of text documents that would be appropriate. Second, you have to be especially careful with sampling in order to avoid bias. For instance, a study of current research on methods of treatment for cancer might use the published literature as the population. This would leave out both the writing on cancer that did not get published for one reason or another as well as the most recent work that has not yet been published. Finally, you have to be careful about interpreting results of automated content analyses. A computer program cannot determine what someone meant by a term or phrase. It is relatively easy in a large analysis to misinterpret a result because you did not take into account the subtleties of meaning.
However, content analysis has the advantage of being unobtrusive and, depending on whether automated methods exist, can be a relatively rapid method for analyzing large amounts of text.

Secondary Analysis of Data

Secondary analysis, like content analysis, makes use of already existing sources of data. However, secondary analysis typically refers to the re-analysis of quantitative data rather than text. In our modern world there is an unbelievable mass of data that is routinely collected by governments, businesses, schools, and other organizations. Much of this information is stored in electronic databases that can be accessed and analyzed. In addition, many research projects store their raw data in electronic form in computer archives so that others can also analyze the data. Among the data available for secondary analysis is:

- census bureau data
- crime records
- standardized testing data
- economic data
- consumer data

Secondary analysis often involves combining information from multiple databases to examine research questions. For example, you might join crime data with census information to assess patterns in criminal behavior by geographic location and group.

Secondary analysis has several advantages. First, it is efficient. It makes use of data that were already collected by someone else. It is the research equivalent of recycling. Second, it often allows you to extend the scope of your study considerably. In many small research projects it is impossible to consider taking a national sample because of the costs involved. Many archived databases are already national in scope and, by using them; you can leverage a relatively small budget into a much broader study than if you collected the data yourself.

However, secondary analysis is not without difficulties. Frequently it is no trivial matter to access and link data from large complex databases. Often the researcher has to make assumptions about what data to combine and which variables are appropriately aggregated into

indexes. Perhaps more importantly, when you use data collected by others you often don't know what problems occurred in the original data collection. Large, well-financed national studies are usually documented quite thoroughly, but even detailed documentation of procedures is often no substitute for direct experience collecting data.

One of the most important and least utilized purposes of secondary analysis is to replicate prior research findings. In any original data analysis there is the potential for errors. In addition, each data analyst tends to approach the analysis from their own perspective using analytic tools they are familiar with. In most research the data are analyzed only once by the original research team. It seems an awful waste. Data that might have taken months or years to collect is only examined once in a relatively brief way and from one analyst's perspective. In social research we generally do a terrible job of documenting and archiving the data from individual studies and making these available in electronic form for others to re-analyze. And, we tend to give little professional credit to studies that are re-analyses. Nevertheless, in the hard sciences the tradition of replicability of results is a critical one and we in the applied social sciences could benefit by directing more of our efforts to secondary analysis of existing data.

Victimization Surveys

The very aim of the victimization survey is to unearth the dark/ hidden/ unreported figure in crime. The dark (or hidden) figure of crime is a term employed by criminologists and sociologists to describe the amount of unreported or undiscovered crime. The only way to find out actual happening of crime and victimization is to organize/conduct a victimization survey. Victimisation surveys have gradually become accepted as a major innovation in the assessment of certain crime-related issues. Much has been written on the shortcomings of the victimisation survey approach, and in criticism of a mechanical application of the victimisation surveys as a substitute for other measures of „crime“. Nevertheless, a tendency is observable that victimisation surveys begin to form part of the basic crime information systems widely applied in some countries, with other countries likely to follow suit in the near future. It is true that the approach is ridden with many inherent flaws, just like any other method of measurement.

Interests of knowledge

When we speak of knowledge-based criminal policy, reference is not made to just any kind of arbitrary „knowledge“. Instead, this discourse makes reference to the overall social costs of crime and crime control, and to the fair distribution of such costs. The first problem is to assess the situation in an unbiased manner. The second problem, then, is to see what can be done about it. A further issue is what is today often addressed in terms of the „what works“ paradigm, i.e. issues that refer to crime prevention and crime reduction, and the prevention or reduction of re-offending – in brief: what are the most justifiable and effective ways of spending resources on reducing the social harms caused by crime and crime control?

Within this frame of reference, the interests of knowledge served by victimisation surveys include objectives such as:

1) to learn about unrecorded crime (victimisation surveys unveil large amounts of unrecorded events that may be crimes). In practical terms, this means estimates of the overall prevalence and incidence of, surveyable " victimisation experiences, and estimates of unrecorded crime.

2) to measure psychological harm and other consequences, and material damage and other costs caused by victimisation. Also broader cost issues may be approached.

3) to measure repeat, serial, and multiple victimisation, victim careers, accumulation of victimisation risks, vulnerable population groups.

4) to compare survey findings with police data. Victimisation surveys allow for insights into how recorded crimes are selected from all possible events that share certain characteristics, including the reporting behaviour of the population. Here, questions on reporting/not reporting crimes to the police and experiences related to reporting are asked.

5) to measure satisfaction with police performance both generally and in each concrete case.

6) to assess popular confidence in the criminal justice system, including police, and punishment attitudes and expectations as to authority performance.

7) to give voice to the victims of crimes and their needs for support. This approach demonstrates that there is a large volume of events or experiences that may be crimes and opens the possibility of reassessing the relative importance of given types of crime-related events, in particular those that are typically not recorded in other standard sources.

8) to find information that is relevant in terms of the rights of crime victims and for constructing indicators related to such rights and needs to enforce them.

9) to learn about public opinion related to crime and crime control: knowledge-based and informed criminal policy should be aware of public opinion about these matters regardless of whether there is agreement or disagreement on what the central crime problems are or on how to deal with crime. The survey is also helpful for finding out what people understand by „crime".

10) to learn about public fear or concern about crime (knowledge-based criminal policy must be aware of and address popular concerns related to crime and crime control). This is addressed by questions on fear and concern, and the deterioration of quality of life caused by crime.

11) to learn what people have done about victimisation (such as preventive measures adopted by the general public or corporate bodies subject to crime risks). Survey questions on precautionary and avoidance behaviour, and the use of protection measures serve this end.

12) to make international comparisons of rates and trends (national or local crime issues are often mistaken as unique and in need of extreme measures, while international comparisons may reveal that the situation is not unusual; also, if the victimisation survey would provide evidence supporting a contrary conclusion, this would be equally important). International comparisons will bring crime assessment into comparative perspective.

13) to measure trends nationally (shortcomings in the standard administrative crime recording systems, such as their inability to account for variations in reporting behaviour, may cause erroneous conclusions concerning trends of certain types of events that are reflected in victimisation surveys).

14) to make regional comparisons within one country (variations in standard administrative crime recording systems and of reporting behaviour may even hamper comparisons across areas within the same jurisdiction).

15) to assess the outcome of crime prevention programmes (more recently, this aspect has gained much support, as local crime prevention projects have become more popular). The victimisation survey has also some unique features that are distinct from criminal justice-related administrative data:

16) the victimisation survey is flexible, it can use standard and changing modules at need.

17) the victimisation survey is able to combine events that are recorded and attended to by different agencies (health care, social services, police, nongovernment organisations).

18) the victimisation survey is able to combine events other than crime with the victimisation experiences (such as the physical safety approach that combines crime victimisation with accidents), and personal characteristics (lifestyle, intoxication, risk-taking behaviour).

This list is not likely to be comprehensive. Many more relevant interests of knowledge could probably be served by victimisation surveys. This list does, however, already demonstrate that the victimisation survey approach could provide better (albeit, of course, probably not the full) answers to a wide scope of relevant questions than what can be drawn from standard administrative crime data that are traditionally relied upon, such as police-recorded crimes, arrest statistics, or statistics on sentenced persons or prisoners. In particular, the focus is shifted from the offender to the victim and the consequences of crimes.

UNIT-IV: DATA ANALYSIS

Types of data: qualitative and quantitative

There are two general types of data. Quantitative data is information about quantities; that is, information that can be measured and written down with numbers. Some examples of quantitative data are your height, your shoe size, and the length of your fingernails. Qualitative research gathers information that is not in numerical form. For example, diary accounts, open-ended questionnaires, unstructured interviews and unstructured observations. Qualitative data is typically descriptive data and as such is harder to analyze than quantitative data. (For more details please refer Data Collection)

Analysis and interpretation of data

Data Analysis

By the time you get to the analysis of your data, most of the really difficult work has been done. It's much more difficult to: define the research problem; develop and implement a sampling plan; conceptualize, operationalize and test your measures; and develop a design structure. If you have done this work well, the analysis of the data is usually a fairly straightforward affair.

In most social research the data analysis involves three major steps, done in roughly this order:

- Cleaning and organizing the data for analysis (Data Preparation)
- Describing the data (Descriptive Statistics)
- Testing Hypotheses and Models (Inferential Statistics)

Data Preparation involves checking or logging the data in; checking the data for accuracy; entering the data into the computer; transforming the data; and developing and documenting a database structure that integrates the various measures.

Descriptive Statistics are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures. Together with simple graphics analysis, they form the basis of virtually every quantitative analysis of data. With descriptive statistics you are simply describing what is, what the data shows.

Inferential Statistics investigate questions, models and hypotheses. In many cases, the conclusions from inferential statistics extend beyond the immediate data alone. For instance, we use inferential statistics to try to infer from the sample data what the population thinks. Or, we use inferential statistics to make judgments of the probability that an observed difference between groups is a dependable one or one that might have happened by chance in this study. Thus, we use inferential statistics to make inferences from our data to more general conditions; we use descriptive statistics simply to describe what's going on in our data.

In most research studies, the analysis section follows these three phases of analysis. Descriptions of how the data were prepared tend to be brief and to focus on only the more unique aspects to your study, such as specific data transformations that are performed. The descriptive statistics that you actually look at can be voluminous. In most write-ups, these are carefully selected and organized into summary tables and graphs that only show the most relevant or important information.

Usually, the researcher links each of the inferential analyses to specific research questions or hypotheses that were raised in the introduction, or notes any models that were tested that emerged as part of the analysis. In most analysis write-ups it's especially critical to not "miss the forest for the trees." If you present too much detail, the reader may not be able to follow the central line of the results. Often extensive analysis details are appropriately relegated to appendices, reserving only the most critical analysis summaries for the body of the report itself.

Data Interpretation

Data analysis and interpretation is the process of assigning meaning to the collected information and determining the conclusions, significance, and implications of the findings. The steps involved in data analysis are a function of the type of information collected; however, returning to the purpose of the assessment and the assessment questions will provide a structure for the organization of the data and a focus for the analysis.

Numerical vs. Narrative – Quantitative vs. Qualititative

The analysis of numerical (quantitative) data is represented in mathematical terms. The most common statistical terms include:

- Mean – The mean score represents a numerical average for a set of responses. For a data set, the terms arithmetic mean, mathematical expectation, and sometimes average are used synonymously to refer to a central value of a discrete set of numbers: specifically, the sum of the values divided by the number of values. If the data set were based on a series of observations obtained by sampling from a statistical population, the arithmetic mean is termed the sample mean to distinguish it from the population mean.
- Standard deviation – The standard deviation represents the distribution of the responses around the mean. It indicates the degree of consistency among the responses. The standard deviation, in conjunction with the mean, provides a better understanding of the data. For example, if the mean is 3.3 with a standard deviation (StD) of 0.4, then two-thirds of the responses lie between 2.9 (3.3 – 0.4) and 3.7 (3.3 + 0.4).
- Frequency distribution – Frequency distribution indicates the frequency of each response. For example, if respondents answer a question using an agree/disagree scale, the percentage of respondents who selected each response on the scale would be indicated. The frequency distribution provides additional information beyond the mean, since it allows for examining the level of consensus among the data.
- Higher levels of statistical analysis (e.g., t-test, factor analysis, regression, ANOVA) can be conducted on the data, but these are not frequently used in most program/project assessments.

Some Data Interpretation and Analysis Tips

- Consider **the data from various perspectives**. Whatever your project may be or whatever data you have collected from your business it's always best to ask what that data means for various actors or participants.
- Think beyond the data but do not stray too far from the data. Be mindful that you are not making too much of your data or too little. Make the link between the data and your interpretations clear. Base your interpretations in your research.
- Make visible the assumptions and beliefs, or mental models, that influence your interpretation. We each carry images, assumptions, and stories in our minds about ourselves, others, the organizations we work in, etc. As a composite, they represent our view of our world. Because these models are generally unarticulated, i.e., below our level of our awareness, if left unexamined, these assumptions and beliefs can lead to incorrect interpretations. Reflect on your own thinking and reasoning. Individually and/or collectively list your assumptions about the inquiry focus.
- Take care not to disregard outlying data or data that seems to be the exception.
- Data that is surprising, contradictory or puzzling can lead to useful insights (insites.org)

Data processing

In social science, coding is an analytical process in which data, in both quantitative form (such as questionnaires results) or qualitative (such as interview transcripts) is categorized to facilitate analysis. Coding means the data transformation. Substantial changes have taken place also in the field of data processing, particularly over the past few years, making it difficult to lay down generally applicable quality standards for the individual steps. The pluralisation in the methods of data collection already noted in the previous chapter has led to an increasing differentiation of procedural techniques in the field of data processing too. For instance, for the purposes of coding and data editing computer-aided interviews must be treated differently to conventional personal interviews conducted with paper and pencil.

The corporate structures within market and social research agencies have also changed substantially in the field of data processing. There is a marked trend towards decentralised data processing. Central mainframe computers are only rarely used or only for specific purposes. Often the project manager himself is responsible for the final stages of data processing – i.e. tabulation and diagrammatic representation of the research results and other forms of reporting. This places higher demands on the intra-corporate procedures for data processing.

Manual coding is becoming less and less common and is often confined to the coding of open-ended questions. Increasingly, questionnaires are read automatically using optical scanners. In the case of computer-aided interviews, a data record is immediately created and stored on an electronic data carrier. The use of multivariate methods of analysis based on more and more powerful software packages is constantly increasing. The multitude of presentation techniques available leads to different forms of presentation of the research findings.

Developing standards that take into account all the possibilities available here would go far beyond the scope of a catalogue of standards for quality assurance in market and social research.

Nevertheless, there are certain basic rules for the field of data processing too, which at least apply to most of the procedures commonly used today.

A. Devising Coding Instructions

i. Closed Questions

With closed questions, coding is generally already defined before fieldwork begins. If the interviews are conducted with the aid of computers, the coding is part of the programming software. This aspect of programmes, in particular, should be tested accordingly before fieldwork commences.

For coding, a distinction must be made between the two answer categories “don’t know” and “no answer”. The default category “other” must not lead to any distortions in the distribution of responses.

ii. Open-ended Questions

With open-ended questions, the answers occurring in a sufficiently large number of questionnaires should first be listed. The necessary number of questionnaires is determined by the proportion of respondents expected to respond to the particular question. The code frame should then be developed on the basis of this list. The responsibility for the content of the categories in a code frame lies with the project manager in charge. If in the course of coding it becomes necessary to modify or change a code frame, the questionnaires that have already been coded must be recoded.

B. Conducting and Checking Coding

i. Briefing Coding Staff

Coding staff must be briefed by the project manager or some other person who is thoroughly acquainted with the design and objectives of the research project.

ii. Checking Coding

The person in charge of briefing the coding staff should check at least five percent of the coding.

iii. Dealing with Mistakes

Erroneous coding must be corrected. If the number of errors made accumulates, all or some of the coding staff must be briefed once again, and the questionnaires affected must be recoded.

C. Conducting and Checking Data Entry

i. Manual Data Entry

Suitable checks must be conducted to try to ensure that the data transferred to an automatically readable data carrier agree as closely as possible with the data contained in the questionnaires.

For this purpose, at least ten percent of the questionnaires entered should be checked and any errors found should be corrected. When checking manual data entry, all operators should be included. When an above average number of data are found to be incorrect, all the questionnaires whose data has been entered by the data entry operator in question must be checked.

ii. Automated Data Entry

When entering data automatically, those questionnaires which cannot be read easily or at all for technical reasons must be picked out and entered manually. In addition it must be ensured that when data is entered by automatic means, manifestly illogical formal details are recognised as such. In such cases, the data must be corrected manually. Automatically entered data must be spot-checked for correctness, whereby particular emphasis should be placed on the correct entry of numerical values, since these are difficult or impossible to check by means of plausibility tests.

D. Editing Raw Data

Even when data are coded and entered with great care, it cannot be ruled out that the data may still contain errors. Here it is necessary to distinguish between formal, logical errors (e.g. filter errors or multiple responses when only a single response is permitted) and inconsistent or implausible results.

i. Formal, Logical Errors

As a rule, formal logical errors may be eliminated using a suitable editing programme. The project manager must be informed about the programme and the type of data correction performed by it. It is his responsibility to instruct any changes to be made to the editing programme in accordance with contextual aspects.

ii. Inconsistent or Implausible Results

When inconsistent or implausible results occur, the project manager should first be informed. He will decide how to proceed with these cases. To sort out the problem it is usually advisable to refer back to the original data, i.e. the individual questionnaires.

iii. Documentation

Both the raw data and the edited data should be stored. Changes that were made by the editing programme employed should be documented in detail.

E. Weighting Data

i. Procedures Used

Quota samples are in most cases samples of individuals. If the required quota characteristics and their combinations have been obeyed, there is no need for any correction by means of weighting the results – apart from a possible weighting dictated by the sample design itself.

Random samples, too, are mostly intended to represent a total population of persons or households. Since there is no central register of individuals in Germany, and sampling at the Residents' Registration Offices is extremely complicated, most random samples are drawn in several stages. First of all, the households are chosen, followed by the selection of the individuals. This means that people belonging to the underlying population have different probabilities of being selected, depending on the size of the household in which they live.

A multi-stage sampling method makes it necessary to perform an appropriate weighting of the results. In the first stage, the design-related bias is corrected and the sample of households is converted into a sample of individuals. After this, in a second stage, the structural bias of the sample caused particularly by non-responses is corrected. Only after the second stage of the weighting procedure is a representative sample of individuals available that can be used for projections.

When samples of households are required, the first stage in the weighting procedure is, of course, unnecessary. In these cases the correction of any structural bias in the second stage of weighting is done using the known distribution of a number of household characteristics. In order to determine the necessary weighting factors, it is normally possible – for both samples of individuals and for samples of households – to resort to the data from official statistics or to structural data from other generally recognised research projects.

ii. Documentation

The structural data of the sample should be documented in both their weighted and un-weighted forms. In particular it must be apparent from this how the weighting affects the structure of the sample. The weighting procedure used must be described adequately, including mention of the weighting variables.

F. Presentation and Analysis of Data

Research results can be passed on to clients in very different forms – depending on what arrangements have been agreed upon. The following typical forms may be distinguished:

- The client receives an edited and usually also weighted data record on an automatically readable data carrier. In addition he receives a method report containing the essential project-related details of data collection and data processing.
- The client receives an uncommented table of results and a method report.
- The client receives a report in which the results of the research project are interpreted, including diagrams and tables, as well as a method report.
- The client receives an interpretative report which also contains recommendations based on the results of the research project, as well as a method report.
- The results of the research project are presented to the client in person and the resulting recommendations are discussed with him.

These various typical forms of supplying the results of a research project to clients can be combined with each other. In all cases care should be taken that the client and the research

agency conducting the research project agree about the form of analysis and presentation of the results already at the time of commissioning a research project.

Survey method

Survey research is one of the most important areas of measurement in applied social research. The broad area of survey research encompasses any measurement procedures that involve asking questions of respondents. A "survey" can be anything forms a short paper-and-pencil feedback form to an intensive one-on-one in-depth interview.

We'll begin by looking at the different types of surveys that are possible. These are roughly divided into two broad areas: Questionnaires and Interviews. Next, we'll look at how you select the survey method that is best for your situation. Once you've selected the survey method, you have to construct the survey itself. Here, we will be address a number of issues including: the different types of questions; decisions about question content; decisions about question wording; decisions about response format; and, question placement and sequence in your instrument. We turn next to some of the special issues involved in administering a personal interview. Finally, we'll consider some of the advantages and disadvantages of survey methods.

Report writing

So now that you've completed the research project, what do you do? I know you won't want to hear this, but your work is still far from done. In fact, this final stage -- writing up your research -- may be one of the most difficult. Developing a good, effective and concise report is an art form in itself. And, in many research projects you will need to write multiple reports that present the results at different levels of detail for different audiences. There are several general considerations to keep in mind when generating a report:

The Audience

Who is going to read the report? Reports will differ considerably depending on whether the audience will want or require technical detail, whether they are looking for a summary of results, or whether they are about to examine your research in a Ph.D. exam.

The Story

I believe that every research project has at least one major "story" in it. Sometimes the story centers on a specific research finding. Sometimes it is based on a methodological problem or challenge. When you write your report, you should attempt to tell the "story" to your reader. Even in very formal journal articles where you will be required to be concise and detailed at the same time, a good "storyline" can help make an otherwise very dull report interesting to the reader. The hardest part of telling the story in your research is finding the story in the first place. Usually when you come to writing up your research you have been steeped in the details for weeks or months (and sometimes even for years). You've been worrying about sampling response, struggling with operationalizing your measures, dealing with the details of design, and wrestling with the data analysis. You're a bit like the ostrich that has its head in the sand. To find

the story in your research, you have to pull your head out of the sand and look at the big picture. You have to try to view your research from your audience's perspective. You may have to let go of some of the details that you obsessed so much about and leave them out of the write up or bury them in technical appendices or tables.

Formatting Considerations

Are you writing a research report that you will submit for publication in a journal? If so, you should be aware that every journal requires articles that you follow specific formatting guidelines. Thinking of writing a book. Again, every publisher will require specific formatting. Writing a term paper? Most faculties will require that you follow specific guidelines. Doing your thesis or dissertation? Every university I know of has very strict policies about formatting and style. There are legendary stories that circulate among graduate students about the dissertation that was rejected because the page margins were a quarter inch off or the figures weren't labeled correctly.

To illustrate what a set of research report specifications might include, I present in this section general guidelines for the formatting of a research write-up for a class term paper. These guidelines are very similar to the types of specifications you might be required to follow for a journal article. However, you need to check the specific formatting guidelines for the report you are writing -- the ones presented here are likely to differ in some ways from any other guidelines that may be required in other contexts.

I've also included a sample research paper write-up that illustrates these guidelines. This sample paper is for a "make-believe" research project. But it illustrates how a final research report might look using the guidelines given here.

Researcher Fraud

Research fraud is a form of misconduct involving fabrication, falsification, plagiarism, or other practices that seriously deviate from those that are commonly accepted within the academic Page 3 Tulane University Faculty Handbook Page 70 of 166 **research** communities for proposing, conducting or reporting research. Behavior that would be considered scientific could occur at all points in a research protocol. You could encounter different types of scientific misconduct at different stages, right from the origination of the research study itself to the publication of the results.

Listed below are the top 10 transgressions that peer reviewers and journal editors look for, incorporating content from both the *World Association of Medical Editors* (WAME), and the *US Office of Research Integrity* :

1. *Misappropriation of Ideas* – taking the intellectual property of others, perhaps as a result of reviewing someone else's article or manuscript, or grant application and proceeding with the idea as your own.
2. *Plagiarism* – utilizing someone else's words, published work, research processes, or results without giving appropriate credit via full citation.

3. *Self-plagiarism* – recycling or re-using your own work with appropriate disclosure and/or citation.
4. *Impropriety of Authorship* – claiming undeserved authorship on your own behalf, excluding material contributors from co-authorship, including non-contributors as authors, or submitting multi-author papers to journals without the consensus of all named authors.
5. *Failure to Comply with Legislative and Regulatory Requirements* – willful violations of rules concerning the safe use of chemicals, care of human and animal test subjects, inappropriate use of investigative drugs or equipment, and inappropriate use of research funds.
6. *Violation of Generally Accepted Research Practices* – this can include the proposal of the research study, manipulation of experiments to generate preferred results, deceptive statistical or analytical practices to generate preferred results, or improper reporting of results to present a misleading outcome.
7. *Falsification of Data* – rather than manipulate the experiments or the data to generate preferred results, this transgression simply fabricates the data entirely.
8. *Failure to Support Validation of Your Research* – by refusing to supply complete datasets or research material needed to facilitate validation of your results through a replication study.
9. *Failure to Respond to Known Cases of Unsuccessful Validation Attempts* – published research that is found to be flawed should be retracted from the journal that published it.
10. *Inappropriate Behavior in Relation to Suspected Misconduct* – failure to cooperate with any claims of misconduct made against you, failure to report known or suspected misconduct, destruction of any evidence related to any claim of misconduct, retaliation against any persons involved in a claim of misconduct, knowingly making false claims of misconduct.

Plagiarism

Plagiarism is presenting someone else's work or ideas as your own, with or without their consent, by incorporating it into your work without full acknowledgement. All published and unpublished material, whether in manuscript, printed or electronic form, is covered under this definition. Plagiarism has become an increasingly serious problem in the University. It is aggravated by the easy access to and the ease of cutting and pasting from a wide range of materials available on the internet. Plagiarism is regarded as a very serious offence in the academic world. It constitutes academic theft - the offender has 'stolen' the work of others and presented the stolen work as if it were his or her own. It goes to the integrity and honesty of a person. It stifles creativity and originality, and defeats the purpose of education.

In this University, plagiarism is a disciplinary offence. Any student who commits the offence is liable to disciplinary action. It is disappointing to witness an increase of plagiarism in the University at both undergraduate and graduate levels in recent years. Time and again students who appeared before the University Disciplinary Committee alleged that they did not know what constituted plagiarism. This booklet attempts to give you some guidance on what constitutes plagiarism, why it is wrong, and how to avoid it.

What constitutes plagiarism?

For instance in a University, based on the Regulation 6 of the University's Regulations Governing Students' Academic Conduct Concerning Assessment provides:

"A candidate shall not engage in plagiarism nor employ nor seek to employ any other unfair means at an examination or in any other form of assessment. Plagiarism is defined as direct copying of textual material or willful use of other people's data and ideas, and presenting them as one's own without acknowledgement, whether or not such materials, data and ideas have been published."

Put it simply, plagiarism is copying the work of another person without proper acknowledgement. There are two parts in the definition: *copying* and the *absence of proper acknowledgement*. As a result, it gives an impression to an ordinary reader that the work is the original work of the author when in fact it was copied from some others' work. The idea underlying plagiarism is very simple: if you appropriate the work of another person, you should give proper recognition to that person.

Plagiarism covers "any other form of assessment". It covers theses, dissertations, take-home examinations, assignments, projects, and other forms of coursework. It applies to both undergraduate and graduate students.

** The University would like to acknowledge the contribution of Professor Johannes Chan of the Faculty of Law for preparing the first edition of this booklet in August 2002.*

Copying does not necessarily mean copying word for word. Closely paraphrasing or substantial copying with minor modifications (such as changing grammar, adding a few words or reversing active/passive voices) is still copying for this purpose. It is not so much the form of the copying that is important, but the substance of what is copied. It does not matter what the nature of the source is. It may be a book, an article, a dissertation, a Government report, a table from the internet, a memorandum, or simply an assignment of another student or even teaching material distributed to you. The source may also be graphics, computer programmes, photographs, video and audio recordings or other non-textual material. It does not matter whether the source has been published or not.

The most common form of plagiarism

The most common form of plagiarism is copying from the published works of writers or the essays of other students without any acknowledgement. The source may be published in traditional text or on the internet.

Confidentiality in Criminal Justice Research

The need for confidentiality arises in relationships where one party is vulnerable because of the trust reposed in the other and includes relationships where one party provides information to another because of the latter's commitment to confidentiality. For this reason, a confidential

relationship is a fiduciary relationship par excellence. The researcher-participant relationship is unique among relationships in which confidentiality may be considered integral to the functioning of the relationship. If research participants face a risk because of their participation in our research, it is usually because we have walked into their lives and exposed them to it. Generally, research participation is voluntary, and conducted without undue coercion or promise of reward. While some research participants are paid for their time, such as students in psychology experiments or criminal offenders, the payment is usually nominal.

Although participants may derive satisfaction at someone lending them an ear and perhaps even hope to change the world by revealing private information to a researcher, research participants are generally motivated by their desire to help create knowledge for the greater good. The primary purpose of research ethics is to ensure research participants are not harmed by their involvement in research.

With respect to risks produced by third party intrusion, our commitment to confidentiality and anonymity provide the key foundations for this protection. When research can be conducted in a way that maintains research participant anonymity, the threat of violating confidentiality because of some unwanted third party intrusion is minimal. Clearly, whenever data can be gathered anonymously, they should be.

However, in many types of research, anonymity is not an option. Researchers who want to link particular individuals across different databases or track certain participants over time must have a means of identifying research participants. When working with large databases, a variety of techniques can be used to sever identifying information and destroy it, or store it separately, preferably in another country.

Unless they are involved in longitudinal studies or tracking individuals across databases, quantitative researchers generally do not experience the problem of recording incriminating information, and they rarely carry it in their heads. But the more qualitative and inductive the research becomes, the more difficult it is to use technical devices to protect research participants.

Ethnographers and other researchers working with a limited number of key informants or on case histories, where participants are chosen on the basis of their reputation or institutional affiliation and position, can delete identifying information from interview transcripts and field notes, but they cannot delete it from their memories. Of course, confidentiality is not essential across the entire spectrum of research. With relatively innocuous topics confidentiality may be of little concern to participants.

Also some research participants prefer to be named, as is their right. However, when research participants divulge personal information that could harm their reputation, self-esteem and/or well being, and especially when secrecy is cherished in the social world under study, the researcher-participant relationship is predicated on trust. Trust cannot be built on a promise of confidentiality that, depending on whether law enforcement authorities or interested third parties want the information, may or may not be kept.

Potential Conflicts between Research Ethics and Law

There are at least four areas of potential conflict between the law and the ethical requirement of confidentiality:

1. When researchers learn about certain crimes and are statutorily obliged to report them (i.e., mandatory reporting laws that do not exempt researchers; these vary by jurisdiction, but may include, for example, elder abuse, child abuse, and/or spousal assault).

2. When researchers learn about potential future crime and may be held liable for harm to third parties they could have prevented.

3. When non-governmental third parties subpoena researchers to testify about issues arising in high stakes litigation.

4. When prosecutors, grand juries, congressional committees and various public bodies subpoena researchers to testify about crimes and/or other offences research participants may have revealed to the researcher. Coroners also can subpoena researchers who they think might have information relevant to an inquest.

In general, the first two are situations where the researcher's violation of confidentiality would be a matter of their own initiative, independent of compulsion. In contrast, subpoenas create the threat of compelled revelation after the fact. Depending on the kind of research they are conducting, criminologists could find themselves confronting dilemmas in any of these four areas. However, historically, in the U.S. at least, it is the third and fourth categories — both involving the possibility of subpoena and orders for disclosure — that have represented the greatest threat to researchers and their participants.

UNIT-V: BASIC STATISTICS

Statistics-Meaning

Statistics is a branch of mathematics dealing with the collection, analysis, interpretation, presentation, and organization of data. In applying statistics to, e.g., a scientific, industrial, or social problem, it is conventional to begin with a statistical population or a statistical model process to be studied. Statistics is a branch of mathematics dealing with the collection, analysis, interpretation, presentation, organization of data and interpretation of numerical facts, for drawing inferences on the basis of their quantifiable likelihood (probability).

In applying statistics to, e.g., a scientific, industrial, or social problem, it is conventional to begin with a statistical population or a statistical model process to be studied. Statistics can Interpret aggregated of data too large to be intelligible by ordinary observation because such data (unlike individual quantities) tend to behave in regular, predictable manner. It is subdivided in to descriptive statistics and inferential statistics.

The study of statistics involves the extensive use of numbers and information meant to describe a situation or derive a logical inference or conclusion from analyzing both numerical figures and non-numerical facts. Statistics broadly refers to a range of techniques and procedures for analyzing, interpreting, displaying and making decisions based on data. Statistics may either be descriptive or inferential. The explosion of available data made possible by technology advances has resulted in a great need to make sense of all data collection, thus statistical analysis has never been more important.

Today, data mining is being undertaken across many fields and organizations and statistical analysis becomes indispensable in order to convert all the data into information that will be useful in a variety of ways. In addition, information systems of organizations rely on good statistics so that users are able to rely on the information they generate. The use of statistics and statistical analysis on crime and criminal justice is critical in order for governments to be effective in implementing anti-crime programs and protect the well-being of the population and assess the social impact of public expenditures and policies.

In an increasingly globalized and interconnected world, in the age of cybercrimes even, the collection of reliable and comprehensive criminal justice statistics cannot be understated. Countries collect huge amount of data pertaining to crimes, criminal profiles and related socio-economic, political and geographic profile of different communities and the country as a whole in order to sustain its respective criminal justice system.

In many cases, the governments undertake extensive research using quantitative or qualitative techniques or both in order to generate the data. During the research process, statistical analysis is required to transform all these data into useful information for effective criminal justice decision-making as well as policy research.

Significance and Scope of Criminal Justice Statistics

Statistical data in criminal justice are gathered to answer questions and often data is collected during the crime investigation process. In defining the scope of criminal justice statistics, it is important to know who the past and present users are and for what purpose the data will be used. The questions may also deal with who submits the data, who receives the data, what data is submitted and in what form and at what interval. In considering the criminal justice statistics system by itself, there is also a need to consider what critical policy issues should be included in the formulation of any program to improve it. The same UN document discloses that an analysis of the requirements for a good justice statistics program might look into the following concerns as a starting point:

(a) Incidence of crime (seriousness, trends, structure, others)
(b) Profile of offenders and their characteristics
(c) Workload of the system (crimes, arrests, dispositions, offenders under supervision)
(d) Offenders and cases moving through the system
(e) Recidivism
(f) Characteristics of victims
(g) Resources expended (human and fiscal)
(h) Crime correlates like economic, demographic and other data
(i) Social and economic cost of crime
(j) Citizens attitudes towards and concerns about crime and criminal justice

The consideration of underlying questions with regard to these starting points may very well lead to an exhaustive description of the scope and concern of criminal justice statistics and most importantly its analysis. Data must be transformed into information in order to be truly useful and practical.

In this material, we discuss the influence of statistical analysis on criminal justice research and the role of research in the delivery of the services under the criminal justice system. Statistical Analysis in Criminal Justice Research The use of statistics and statistical analysis on crime and criminal justice is critical in order for governments to be effective in implementing anti-crime programs and protect the well-being of the population and assess the social impact of public expenditures and policies.

In an increasingly globalized and interconnected world, in the age of cybercrimes even, the collection of reliable and comprehensive criminal justice statistics cannot be understated. Countries collect huge amount of data pertaining to crimes, criminal profiles and related socioeconomic, political and geographic profile of different communities and the country as a whole in order to sustain its respective criminal justice system.

In many cases, the governments undertake extensive research using quantitative or qualitative techniques or both in order to generate the data. During the research process, statistical analysis is required to transform all these data into useful information for effective criminal justice decision-making as well as policy research.

Classification of Tabulation

Tabulation of Data

The process of placing classified data into tabular form is known as tabulation. A table is a symmetric arrangement of statistical data in rows and columns. A process of condensing data and presenting it in a compact form, by putting data into statistical table, is called tabulation. Classification of data is done after data collection process is completed. Conversely, in tabulation the data is arranged in rows and columns, in a systematic way.

The process of placing classified data into tabular form is known as tabulation. A table is a symmetric arrangement of statistical data in rows and columns. Rows are horizontal arrangements whereas columns are vertical arrangements. It may be simple, double or complex depending upon the type of classification.

Classification of Tabulation

(1) Simple Tabulation or One-way Tabulation: When the data are tabulated to one characteristic, it is said to be simple tabulation or one-way tabulation.
For Example: Tabulation of data on population of world classified by one characteristic like Religion is example of simple tabulation.

(2) Double Tabulation or Two-way Tabulation: When the data are tabulated according to two characteristics at a time. It is said to be double tabulation or two-way tabulation.

For Example: Tabulation of data on population of world classified by two characteristics like Religion and Sex is example of double tabulation.

(3) Complex Tabulation: When the data are tabulated according to many characteristics, it is said to be complex tabulation.

Diagrammatic and Graphic Representation of Data

Diagrammatic presentation of data

Generally a statistical data is first classified and then tabulated**,** that is the data is presented in the form of a table. A table consists of numbers which may not always be interesting. They can be confusing especially when they are large and complex. So it may not be easily understood by layman. Hence a tabulated data can be represented through diagrams and graphs, which is more interesting and appealing. These are more attractive and easily understood, when compared to the tables.

Diagrams

Diagram is a pictorial representation of the data. A diagram is a visual form of presentation of statistical data. Diagram refers to various types of devices such as bars, circles, maps, pictorials, cartograms etc. These devices can take many attractive forms. Strictly speaking, these are not graphic devices. Diagrams do not add any new meaning to the statistical facts, but they exhibit the results more clearly. The use of diagrams is becoming more and more popular in the present time. Diagrams, occupy an important place, because (merits).

1. They are attractive and impressive: Diagrams are attractive and create interest in the mind of the readers. They are more appealing to the eye; even a layman can understand them very easily. Diagrams have greater attention than mere figures.
2. They save time and labour: Diagrams save mush time and labour to understand it and enables one to draw meaningful inferences from it.
3. They have universal applicability: Diagrammatic presentation of statistical data is followed universally. It is greatly used in almost all walks of life as a good guide in economics, business, social institutions, administration and other fields.
4. They make data simple: Diagrams can be remembered easily, as they render comparison in an easy and possible way. They render the whole data readily intelligible. For example, the study of profit pattern of two firms with the help of figures may not be clear, but when the figures are put in the medium of a diagram the trend can be very clear.
5. They make comparison easy: Diagrams render comparison between two or more sets of data. In absolute figures comparison may not be clear, but diagrammatic presentation makes it earlier and simpler.
6. They provide more information: A diagram will reveal more information than the data in a table. Cold figures can speak in clear tones, if translated into diagrammatic language.

Limitations of a diagram

The presentation of a diagram without a careful study will be misleading. In brief, the following are the deficiencies or restricted uses:

1. Diagrams cannot be analyzed further.
2. Diagrams show only approximate values.
3. The uses of certain diagrams are limited to the experts (example: multi-dimensional ones).
4. It exposes only limited facts. All details cannot be presented diagrammatically.
5. To draw a table is easy but construction of a diagram is not so easy.
6. It is a supplement to the tabular presentation but not an alternative to it.
7. Minute readings cannot be made. Small differences in large measurements cannot be studied.
8. If there is a wide gap between two different measurements, the diagram will not give a meaningful look. For example, 10 and 900 cannot be shown in a diagram, whatever the scale be adopted.
9. Diagrams are drawn when comparisons needed otherwise, they are of little use.
10. Diagrams drawn on false base are illusory.

Rules for making a diagram

Diagrammatic presentation of a statistical table is simple and effective as photographic memory will last long in the mind than any other form. The construction of a diagram is an art, which can be acquired through practice. However, the following guide line will help in making them more effective

1. Heading: every diagram must have suitable title. The title, in bold letters, conveys the main facts depicted by the diagram. If needed, sub-headings can also be given. It must be brief, self-explanatory and clear.
2. Size: The size of the diagram should neither be too big nor too small. It must match with the size of the paper. It should be in the middle of the paper.
3. Length and Breadth: An appropriate proportion should be maintained between length and breadth. Lutz has suggested that proportions of length and breadth should be 2:1, 1:4 or 4:1. If it is so, the diagram looks attractive. Care should be taken to ensure that the diagram does not look ugly.
4. Drawing: Since impression is needed it should be drawn neatly and accurately with the help of drawing instruments. Each diagram should also be numbered for ready reference.
5. A proper Scale: A proper Scale must be chosen for the diagram to look attractive and create a visual impact on the reader. It must suit the space available. Accuracy should not be sacrificed to attractiveness.
6. Selection of appropriate diagram: The most important point is the selection of proper diagram to present a set of figures. All types of diagrams are not suitable for all types of data. A wrong selection of the diagram will distort the true characteristics of the phenomenon to be presented and might lead to very wrong and misleading interpretations.
7. Right method: The important point, which must be borne in mind at all times, is that the pictorial presentation, chosen for any situation, must depict the true relationship and point out the proper conclusion. Use of an inappropriate chart may distort the facts and mislead the reader.
8. Index: When many items are shown in a diagram, through different colors, dotting, crossings etc an index must be given for identifying and understanding the diagrams.
9. Sources: If the data presented have been acquired from some external source, the fact should be indicated at the bottom of the diagram.
10. Simplicity: Diagram should be very simple. It must be so simple that even a lay man who does not have the knowledge of mathematical or statistical background, can understand the diagram. If the data is very more diagrams can be used to represent the data. Too much of information presented in a diagram will be confusing. Therefore, it is suggested to draw several simple diagrams which are more effective than a complex one.

Types of diagram

There are various diagrammatic devices by which statistical data can be presented. The following are some of the common types of diagrams.

- One dimensional diagram (line and bar)
- Two dimensional diagram (Rectangle, square, circles etc)
- Three dimensional diagram (cube, sphere, cylinder etc)

- Pictogram
- Cartogram

One dimensional diagram

In one dimensional diagram, the length of the lines or bars is considered and the width of the bars is not taken into consideration. The term 'bar' means a thick wide line; the following are the main types:

A. Line diagram: This is the simplest of all the diagrams. On the basis of size of the figures, heights of bars or lines are drawn. The distance between lines is kept uniform. It makes comparison easy. This diagram is not attractive and hence it is less important.

B. Simple Bar diagram: A simple bar diagram is used represent a single variable. A single variable may mean 'population of various countries', 'production of sugar in different states of our country'. 'Number of deaths in some localities' 'number of employees in various branches of a bank' and so on for various years. However, a simple bar diagram can represent only one category of data.

C. Multiple Bar diagram: Multiple bar diagrams are used to denote more than one phenomenon, for example, import and export trend. Multiple bars are useful for direct comparison between two values. The bars are drawn side by side. In order to distinguish the bars, different colours, shades, may be used and a key index to be given to understand the different bars. These diagrams are easy to understand.

D. Sub-divided bar diagram: The bar is subdivided into various parts in proportion to the values given in the data and may be drawn on absolute figures or percentages. Each component occupies a part of the bar proportional to its share in the total. Here also to differentiate different components from one another, different shades or colours may be used.

E. Percentage sub-divided bar diagram: To make comparison on a relative basis various components are expressed as percentage to the total. Percentages are cumulated to divide the bars. Here bars are all of equal heights; each segment shows the percentage to total.

F. Other bar diagrams

a. Deviation bars: Deviation bar diagram is used to depict the net deviations in different values, that are surplus or deficit, profit or loss, net import or export etc which has either positive or negative values. Positive values are shown above the base line and negative value below the baseline.

b. Broken bars: In certain cases we may come across data which contain very wide variations in values- very small or very large. In order to provide adequate and reasonable shape to the smaller bars, the larger bars may be broken at the top. The value of each bar is written at the top of the bar.

2. Two-dimensional diagram

Pie-chart: A pie-chart is used to represent the total into the breakup of various components. For example, the pie chart may represent the budget of a family, for a month and the various sections may represent portions of the budget allotted to food, rent, clothing, and education so on. The pie-chart is so called because the entire graph looks like a pie and the components resemble slices cut from the pie. Pie charts are also called as sector graphs or angular diagrams. They can be constructed for either a single set of data or for two or more sets of data.

Differences between tables and diagrams

1. A table consists of accurate numbers, whereas a diagram gives only an approximate idea.
2. A table contains more information when compared to a diagram.
3. A Table requires careful observation whereas a diagram is easy to follow and interpret.
4, Diagrams are readily understood and they can be retained in memory for a long time, whereas it is difficult and impractical to memorize the numbers present in tables.
5. Diagrams and graphs are attractive and they give good impression when compared to tables.
6. Diagrams and graphs are easy for comparison, when compared to tables.

Graphic Representation

Graphic presentation of statistical data gives a pictorial effect. The collected data will generally be complex. It will be very difficult to understand the importance of collected data. Yet the classification and tabulation will reduce the complexity, still they are not easily understood by the common people.

If the mass of data are depicted graphically, they become easy to be understood and grasped. Statisticians have since long discovered the importance of graphic presentation. It enables us to present the data in a simple, clear and effective manner.

The wandering of a line is more powerful in its effect on the mind than a tabulated statement; it shows what is happening and what is likely to take place just as quickly what is happening and what is likely to take place just as quickly as the eye is capable of working". Graphic presentation of numerical data is becoming popular because of various merits. The advantages of graphic presentation are as follows:

1. It is attractive and impressive.
2. It simplifies complexity of data.
3. It provides easy comparison of two or more phenomena.
4. It needs no special knowledge of mathematics to understand a graph.
5. It provides the basis to locate the statistical measures, like median, mode, quartiles etc.
6. Apart from simplicity, it saves the time and energy of the statistician as well as the observer.
7. Graphic method is probably the simplest method of presenting statistical data.
8. It shows any trend that may be present and the direction in which the trend may change.

Procedure for the construction of a graph

Graphs are prepared on a graph paper in which two lines are drawn, which intersect each other at right angles. These lines are called axis; the point of intersection is 'o' (point of origin). The horizontal line is called y axis.

Generally an independent variable is represented along 'x axis' and the dependent variable along 'Y-axis'. For each axis, a convenient scale, representing the units of the variable is chosen in such a manner that it accommodates the entire data that is given in the table. The scales of x and y need not be the same.

After fixing the origin and scale, the given data is plotted by marking points or dots, corresponding to various x and y values. Then, these points are joined by straight lines to get the graph.

General Rules:

While graphing statistical data, the following guidelines may be observed:
1. Every graph must have a title, indicating the facts presented by the graph.
2. It is necessary to plot the independent variable on the horizontal axis and dependent variables on the vertical axis.
3. Problem arises regarding the choice of a suitable scale. The choice must accommodate the whole idea.
4. The principle of drawing graph is that the vertical scale must start from zero. If the fluctuations are quite small compared to the size of variables, there is no need of showing the entire scale from the origin.
5. For showing proportional relative changes in the magnitude, the ratio or logarithmetic scale should be used.
6. The graph must not be over-crowded with curves.
7. If more than one variable is plotted on the same graph. It is necessary to distinguish them.
8. Index should be given to show the scales and the meaning of different curves.
9. All lettering must be horizontal.
10. It should be remembered that for every value of independent variable there is a corresponding value of the dependent variable. It is these matched values that are to be plotted.
11. Source of information should be mentioned as foot note.

Measures of Central Tendency: Mean, Median, Mode,

We're now ready to talk about descriptive statistics used to characterize and summarize quantitative (interval-level and ratio-level) variables. These descriptive statistics fall into three groups.

- measures of **central tendency** (e.g., mean or average)
- measures of **spread**, **dispersion**, or **variation** (e.g., range)
- measures of distribution **shape** (e.g., if the distribution is symmetrical)

The Mean

The **mean** (also called the **arithmetic mean**) is the same as the **average**.

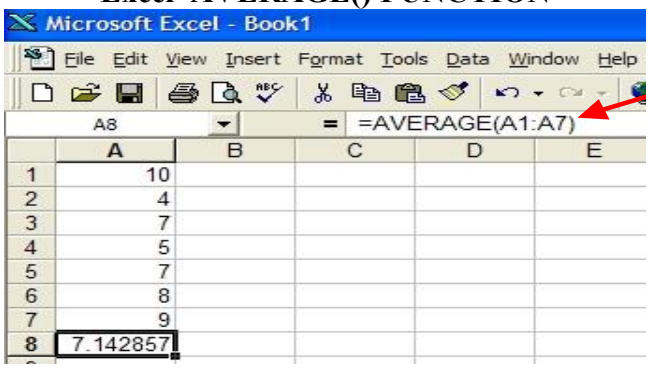
$\mu = \frac{\sum x}{N}$ (for a population) $\qquad \bar{x} = \frac{\sum x}{n}$ (for a sample)

Ex: The data represent the number of textbooks purchased by a sample of seven students:

10 4 7 5 7 8 9

$$\bar{x} = \frac{10+4+7+5+7+8+9}{7} = \frac{50}{7} = 7.14$$

Excel AVERAGE() FUNCTION



The mean is affected by outliers (unusual extreme values). Even one extreme value can throw the mean off.

Example: 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 100
Mean = 110/2 = 55. This value isn't a very good summary description of the dataset, since most values were exactly 1.

The Median

The *median* is a measure of central tendency more resistant to the effects of extreme values. The median is the value that occupies the *middle position* of data when data are put in rank order by magnitude.

Let n be the number of cases in your data.
If n is ***odd***, the median is the middle number of the data values sorted by magnitude. It occupies the $\left(\frac{n+1}{2}\right)^{th}$ position.

If n is ***even***, the median is the average of the middle two numbers of the data sorted by magnitude. It is the average of the numbers in the $\left(\frac{n}{2}\right)^{th}$ and $\left(\frac{n+2}{2}\right)^{th}$ positions.

Example (odd number of values):

1 3 **4** 8 10

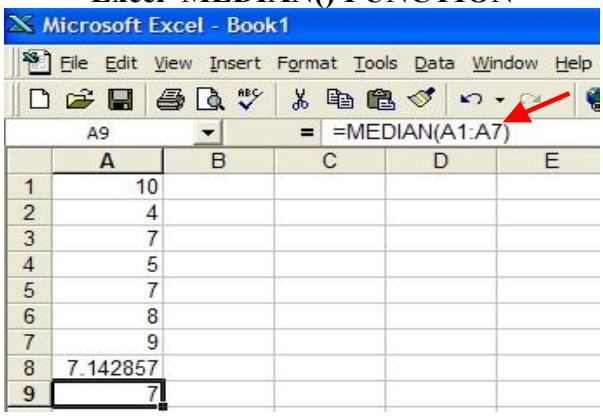
The middle value is 4 (two values are higher, and two lower. This is the median.

Example (even number of values):

2 3 4 **4 5** 8 9 9

The two middle values are 4 and 5. The median is the average of these two values, or 4.5.

Excel MEDIAN() FUNCTION

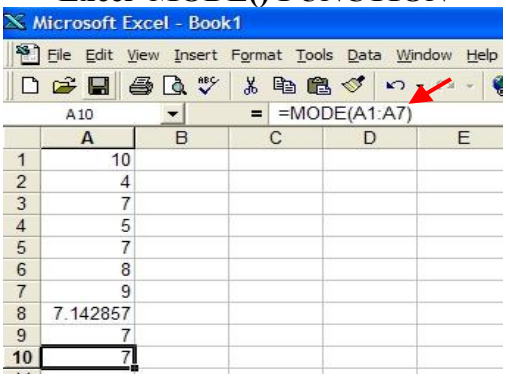


The Mode

The *mode*, by definition, is the *most frequently occurring value* in a series.

- There can be more than one modes
- There can be no mode

Excel MODE() FUNCTION



Measures of dispersion

Range

Range = Maximum - Minimum

Ex Books: 4 5 7 7 8 9 10 Range = 10 - 4 = 6

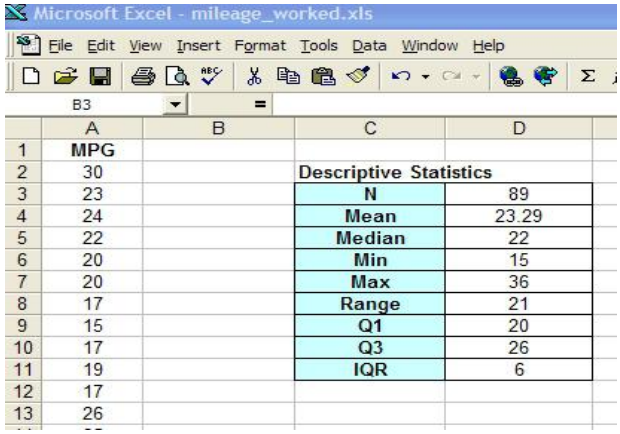
Interquartile Range (IQR)

$IQR = Q_3 - Q_1$

Ex The sample of books: $Q_1 = 5$, $Q_3 = 9$, IQR = 9 - 5 = 4

Example
Make an Excel spreadsheet to compute the mean, median, range, Q1, Q3, and IQR.

- Name the spreadsheet **Mystats**.
- Make it look like this



- Data are placed in Column A.
- Use these formulas:

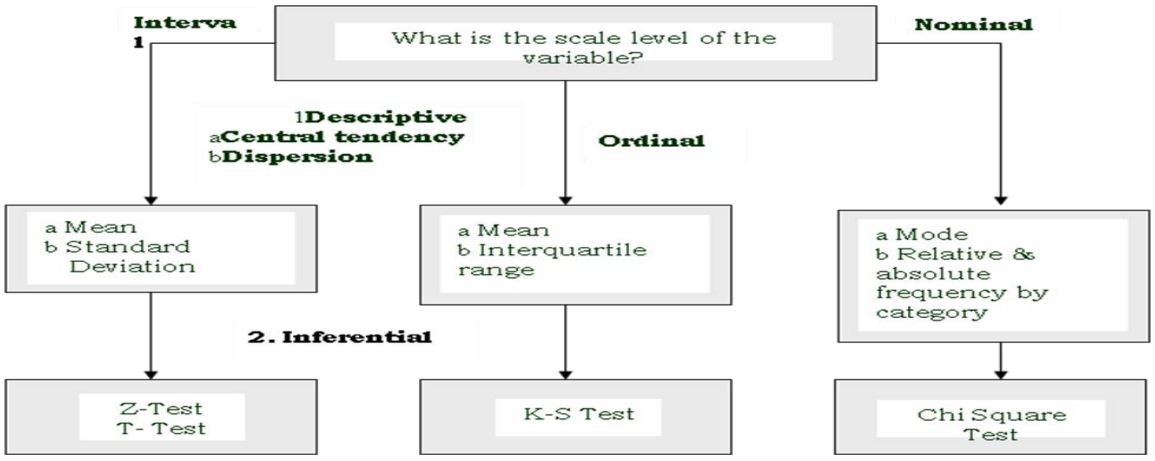
Statistic	Formula
N	=COUNT(a:a)
Mean	=AVERAGE(a:a)
Median	=MEDIAN(a:a)
Min	=MIN(a:a)
Max	=MAX(a:a)
Range	=MAX(a:a)-MIN(a:a)
Q1	=QUARTILE(a:a, 1)
Q3	=QUARTILE(a:a, 3)
IQR	=QUARTILE(a:a,3)-QUARTILE(a:a,1)

Note that a:a is shorthand for the range consisting of all numerical cells in Column A.

- Format the mean to **two decimal places**.
- Use the spreadsheet to calculate these statistics for the **female** blood cholesterol data on the class web-page.
- Take a snapshot, or cut-and-paste the results into Word, print, and turn in.

Univariate analysis

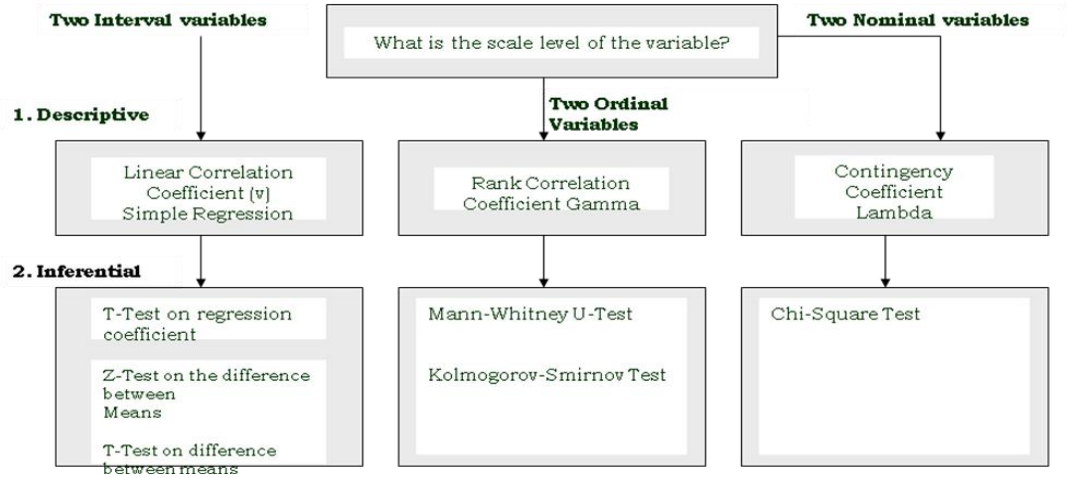
Statistical techniques appropriate for analyzing data when there is a single measurement of each element in the sample or, if there are several or, if there are several measurements on each element, each variable is analysed in isolation.



Bivariate Analysis

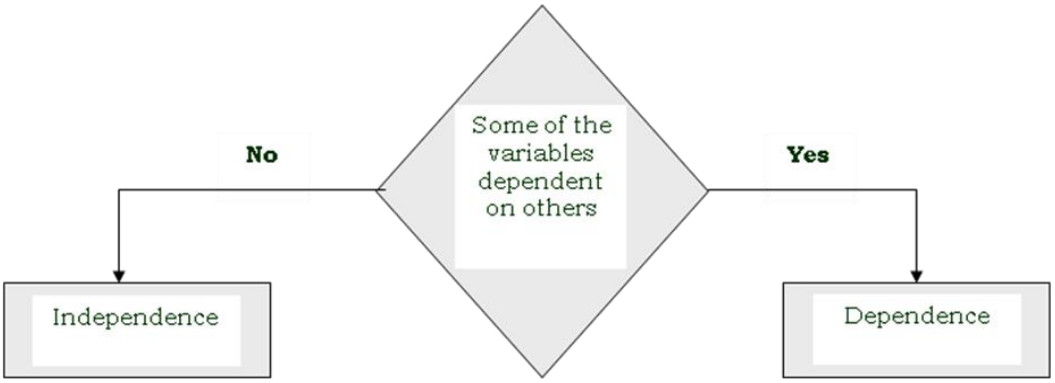
Bivariate analysis is the simultaneous **analysis** of two variables (attributes). It explores the concept of relationship between two variables, whether there exist an association and the strength of this association, or whether there are differences between two variables and the significance of these differences. The Bivariate platform shows the relationship between two continuous variables. It is the *continuous by continuous* personality of the Fit Y by X platform. The word bivariate simply means involving two variables instead of one (univariate) or many (multivariate).

The Bivariate analysis results appear in a scatterplot. Each point on the plot represents the X and Y scores for a single subject; in other words, each point represents two variables. Using the scatterplot, you can see at a glance the degree and pattern of the relationship between the two variables. You can interactively add other types of fits, such as simple linear regression, polynomial regression, and so on.



Multivariate Analysis

Multivariate Data Analysis refers to any statistical technique used to analyze data that arises from more than one variable. This essentially models reality where each situation, product, or decision involves more than a single variable. Multivariate techniques: Statistical techniques suitable for analysing data when there are two or more measurements on each element and the variables are analysed simultaneously. Multivariate techniques are concerned with the simultaneous relationships among two or more phenomena.



1Factor Analysis
2Cluster Analysis
3multidimensional Analysis

Percentage Analysis

Data collected are edited and coded by using the tally bars. This helps in converting the gathered data into a tabulated grouped data. Percentage Analysis is applied to create a contingency table from the frequency distribution and represent the collected data for better understanding.

What is a percentage?

A percentage is a part of a whole. It can take on values between 0 (none of the whole) and 100 (the entire whole). The whole is called the base. The base must ALWAYS be reported whenever a percentage is determined.

Example: There are 20 students in a classroom, 12 of whom are males and 8 of whom are females. The percentage of males is 12 "out of" 20, or 60%. The percentage of females is 8 "out of" 20, or 40%. (20 is the base.)

To how many decimal places a percentage should be reported?

One place to the right of the decimal point is usually sufficient, and you should almost never report more than two. For example, 2 out of 3 are 66 2/3 %, which rounds to 66.67% or 66.7%.

[To refresh your memory, you round down if the fractional part of a mixed number is less than 1/2 or if the next digit is 0, 1, 2, 3, or 4; you round up if the fractional part is greater than or equal to 1/2 or if the next digit is 5, 6, 7, 8, or 9.] Computer programs can report numbers to ten or more decimal places, but that doesn't mean that you have to. I believe that people who report percentages to several decimal places are trying to impress the reader (consciously or unconsciously).

Lang and Secic (2006) provide the following rather rigid rule:

"When the sample size is greater than 100, report percentages to no more than one decimal place. When sample size is less than 100, report percentages in whole numbers. When sample size is less than, say, 20, consider reporting the actual numbers rather than percentages."

[Their rule is just as appropriate for full populations as it is for samples. And they don't say it, perhaps because it is obvious, but if the size of the group is equal to 100, be it sample or population, the percentages are the same as the numerators themselves, with a % sign tacked on.]

How does a percentage differ from a fraction and a proportion?

Fractions and proportions are also parts of wholes, but both take on values between 0 (none of the whole) and 1 (the entire whole), rather than between 0 and 100. To convert from a fraction or a proportion to a percentage you multiply by 100 and add a % sign. To convert from a percentage to a proportion you delete the % sign and divide by 100. That can in turn be converted to a fraction. For example, 1/4 multiplied by 100 is 25%. .25 multiplied by 100 is also 25%. 25% divided by 100 is .25, which can be expressed as a fraction in a variety of ways, such as 25/100 or, in "lowest terms", 1/4. (See the excellent On-Line Math Learning Center website for examples of how to convert from any of these part/whole statistics to any of the others.) But, surprisingly (to me, anyhow), people tend to react differently to statements given in percentage terms vs. fractional terms, even when the statements are mathematically equivalent.

One well-known author (Gerd Gigerenzer) prefers fractions to both percentages and proportions. In his book (Gigerenzer, 2002) and in a subsequent article he co-authored with several colleagues (Gigerenzer, et al., 2007), he advocates an approach that he calls the method of "natural frequencies" for dealing with percentages. For example, instead of saying something like "10% of smokers get lung cancer", he would say "100 out of every 1000 smokers get lung cancer" [He actually uses breast cancer to illustrate his method]. Heynen (2009) agrees. But more about that in Chapter 3, in conjunction with positive diagnoses of diseases.

Is there any difference between a percentage and a percent?

The two terms are often used interchangeably (as I do in this book), but "percentage" is sometimes regarded as the more general term and "percent" as the more specific term. The AMA Manual of Style, the BioMedical Editor website, the Grammar Girl website, and Milo Schield have more to say regarding that distinction. The Grammar Girl (Mignon Fogarty) also explains whether percentage takes a singular or plural verb, whether to use words or numbers

before the % sign, whether to have a leading 0 before a decimal number that can't be greater than 1, and all sorts of other interesting things.

Do percentages have to add to 100?

A resounding YES, if the percentages are all taken on the same base for the same variable, if only one "response" is permitted, and if there are no missing data. For a group of people consisting of both males and females, the % male plus the % female must be equal to 100, as indicated in the above example (60+40=100). If the variable consists of more than two categories (a two-categoried variable is called a dichotomy), the total might not add to 100 because of rounding. As a hypothetical example, consider what might happen if the variable is something like Religious Affiliation and you have percentages reported to the nearest tenth for a group of 153 people of 17 different religions. If those percentages add exactly to 100 I would be terribly surprised.

Several years ago, Mosteller, Youtz, and Zahn (1967) determined that the probability (see Chapter 3) of rounded percentages adding exactly to 100 is perfect for two categories, approximately 3/4 for three categories, approximately 2/3 for four categories, and approximately $\sqrt{6/c\pi}$ for $c \geq 5$, where c is the number of categories and π is the well-known ratio of the circumference of a circle to its diameter (= approximately 3.14). Amazing!
[For an interesting follow-up article, see Diaconis & Freedman (1979). Warning: It has some pretty heavy mathematics!]

Here's a real-data example of the percentages of the various possible blood types for the U.S.:

O Positive	38.4%	
A Positive	32.3%	
B Positive		9.4%
O Negative		7.7%
A Negative		6.5%
AB Positive		3.2%
B Negative		1.7%
AB Negative	.7%	

[Source: American Red Cross website]

Those add to 99.9%. The probability that they would add exactly to 100%, by the Mosteller, et al. formula, is approximately .52.

Can't a percentage be greater than 100?

I said above that percentages can only take on values between 0 and 100. There is nothing less than none of a whole, and there is nothing greater than all of a whole. But occasionally [too often, in my opinion, but Milo Schield disagrees with me] you will see a statistic such as "Her salary went up by 200%" or "John is 300% taller than Mary". Those examples refer to a comparison in terms of a percentage, not an actual percentage. I will have a great deal to say about such comparisons in the next chapter and in Chapter 5.

Why are percentages ubiquitous?

People in general, and researchers in particular, have always been interested in the % of things that are of a particular type, and they always will be. What % of voters voted for Barack Obama in the most recent presidential election? What % of smokers get lung cancer? What % of the questions on a test do I have to answer correctly in order to pass?

An exceptionally readable source about opinion polling is the article in the Public Opinion Quarterly by Wilks (1940a), which was written just before the entrance of the U.S. into World War II, a time when opinions regarding that war were diverse and passionate. I highly recommend that article to those of you who want to know how opinion polls SHOULD work. S.S. Wilks was an exceptional statistician.

What is a Rate?

A rate is a special kind of percentage, and is most often referred to in economics, demography, and epidemiology. An interest rate of 10%, for example, means that for every dollar there is a corresponding $1.10 that needs to be taken into consideration (whether it is to your advantage or to your disadvantage). Some rates, like some percentages, can be greater than 100.

There is something called "The Rule of 72" regarding interest rates. If you want to determine how many years it would take for your money to double if it were invested at a particular interest rate, compounded annually, divide the interest rate into 72 and you'll have a close approximation.

To take a somewhat optimistic example, if the rate is 18% it would take four years (72 divided by 18 is 4) to double your money. [You would actually have "only" 1.93877 times as much after four years, but that's close enough to 2 for government work! Those of you who already know something about compound interest might want to check that.]

Birth rates and death rates are of particular concern in the analysis of population growth or decline. In order to avoid small numbers, they are usually reported "per thousand" rather than "per hundred" (which is what a simple percent is). For example, if in the year 2010 there were to be six million births in the United States "out of" a population of 300 million, the ("crude") birth rate would be 6/300, or 2%, or 20 per thousand. If there were to be three million deaths in that same year, the (also "crude") death rate would be 3/300, or 1%, or 10 per thousand.

One of the most interesting rates is the "response rate" for surveys. It is the percentage of people who agree to participate in a survey. For some surveys, especially those that deal with sensitive matters such as religious beliefs and sexual behavior, the response rate is discouragingly low (and often not even reported), so that the results must be taken with more than the usual grain of salt.

Some rates are phrased in even different terms, e.g., parts per 100,000 or parts per million (the latter often used to express the concentration of a particular pollutant).

What kinds of calculations can be made with percentages?

The most common kinds of calculations involve subtraction and division. If you have two percentages, e.g., the percentage of smokers who get lung cancer and the percentage of non-smokers who get lung cancer, you might want to subtract one from the other or you might want to divide one by the other. Which is it better to do? That matter has been debated for years. If 10% of smokers get lung cancer and 2% of non-smokers get lung cancer (the two percentages are actually lower than that for the U.S.), the difference is 8% and the ratio is 5-to-1 (or 1-to-5, if you invert that ratio). I will have much more to say about differences between percentages and ratios of percentages in subsequent chapters.

Percentages can also be added and multiplied, although such calculations are less common than the subtraction or division of percentages. I've already said that percentages must add to 100, whenever they're taken on the same base for the same variable. And sometimes we're interested in "the percentage of a percentage", in which case two percentages are multiplied. For example, if 10% of smokers get lung cancer and 60% of them (the smokers who get lung cancer) are men, the percentage of smokers who get cancer and are male is 60% of 10%, or 6%. (By subtraction, the other 4% are female.)

You also have to be careful about averaging percentages. If 10% of smokers get lung cancer and 2% of non-smokers get lung cancer, you can't just "split the difference" between those two numbers to get the % of people in general who get lung cancer by adding them together and dividing by two (to obtain 6%). The number of non-smokers far exceeds the number of smokers (at least in 2009), so the percentages have to be weighted before averaging. Without knowing how many smokers and non-smokers there are, all you know is that the average lung cancer % is somewhere between 2% and 10%, but closer to the 2%. [Do you follow that?]

What is Inverse percentaging?

You're reading the report of a study in which there is some missing data (see the following chapter), with one of the percentages based upon an n of 153 and another based upon an n of 147. [153 is one of my favorite numbers. Do you know why? I'll tell you at the end of this book.] You are particularly interested in a variable for which the percentage is given as 69.8, but the author didn't explicitly provide the n for that percentage (much less the numerator that got divided by that n). Can you find out what n is, without writing to the author?

The answer is a qualified yes, if you're good at "inverse percentaging". There are two ways of going about it. The first is by brute force. You take out your trusty calculator and try several combinations of numerators with denominators of 153 and 147 and see which, if any, of them yield 69.8% (rounded to the nearest tenth of a percent). OR, you can use a book of tables, e.g., the book by Stone (1958), and see what kinds of percentages you get for what kinds of n's.

Stone's book provides percentages for all parts from 1 to n of n's from 1 to 399. You turn to the page for an n of 153 and find that 107 is 69.9% of 153. (That is the closest % to 69.8.) You then turn to the page for 147 and find that 102 is 69.4% of 147 and 103 is 70.1% of 147. What is your best guess for the n and for the numerator that you care about? Since the 69.9% for 107 out of 153 is very close to the reported 69.8% (perhaps the author rounded incorrectly or it was a typo?),

since the 69.4% for 102 out of 147 is not nearly as close, and the 70.1% is also not as close (and is an unlikely typo), your best guess is 107 out of 153. But you of course could be wrong.

Chi-square Test

Chi-square test (χ^2): a statistical procedure used to analyze categorical data

We will explore two different types of χ^2 tests:

1. One categorical variable: Goodness-of-fit test
2. Two categorical variables: Contingency table analysis

1. **Parametric tests**: Use at least one numeric rating, so scores can be placed in a frequency distribution that usually has a normal shape
   - Correlation/Regression: Continuous variables only
   - t-test/ANOVA: One categorical variable, one continuous variable

2. **Non-parametric tests**: Do not use numerical values, scores cannot be placed on a frequency distribution (also sometimes used for numeric variables that have non-normal distributions)
   - Chi-square (χ^2): Categorical variables only

Pronunciation note: Chi is pronounced "khi" (from Greece), not "chai" (India)

B. Two Types of χ^2

- χ^2 Test for Goodness of Fit
  - Involves a single categorical variable only

- χ^2 Test for Independence
  - Involves 2(+) categorical variables

C. χ^2 Test for Goodness of Fit

- Generally just involves one categorical variable
- Null hypothesis specifies the proportion of the population in each category
- Determines how well sample data conform to proportions set forth by the null hypothesis
- Test statistic (χ^2) examines whether the proportions in the sample reliably differ from the null hypothesis

Examples:

- Use logic or past research to guide the null hypothesis

H_0 for Gender:

Female	50%

Male	50%

H_0 for Negative Childhood Emotion:

Shame	25%
Anger	25%
Anxiety	25%
Sadness	25%

- Often an equal percentage (proportion) is chosen for each group
- Alternatively, null hypothesis could be based on known proportions in a larger population

Examples:

H_0 for Ethnicity:

White	75%
non-White	25%

H_0 for Vegetarianism:

No	97%
Yes	3%

H_0 for Religious Affiliation:

Christianity	84%
Non-religious / Don't care	10%
Agnosticism	2%
Atheism	1%
Other	3%

- χ^2 used to examine whether the proportions in a sample reliably differ from those hypothesized
  Like F…
  χ^2 ranges from 0 to ∞
  Is small when the null hypothesis is likely true.
  Is large when the null hypothesis is rejected.

H_0 for Gender:

Female	50%
Male	50%

$N = 279$
Proportion Female = 50% = .50
Hypothesized frequency Female = .50 * 279 = 139.5
Proportion Male = 50% = .50
Hypothesized frequency Male = .50 * 279 = 139.5

Chi-Square Test- Frequencies

15. Gender

	Observed N	Expected N	Residual
Female	213	139.5	73.5
Male	66	139.5	-73.5
Total	279		

Test Statistics

	15. Gender
Chi-Square[a]	77.452
df	1
Asymp. Sig.	.000

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 139.5.

- The Observed N indicates the actual frequency in each group
- The Expected N indicates the frequency that is hypothesized, based on the null hypothesis.
- The Residual just indicates the Observed frequency minus the Expected frequency
- The Test Statistics box indicates that the χ^2 value is 44.17, the degrees of freedom (df) are 1, and $p < .001$.

The chi-square test for goodness of fit was significant, $\chi^2(1, N = 279) = 77.45$, $p < .001$. The sample included an unexpectedly high number of females.

T-Test

The t-test calculates the t statistic, degrees of freedom, and P value of the specified data. These results are displayed in the t-test report which automatically appears after the t-test is performed. The other results displayed in the report are enabled and disabled in the Options for t -test dialog.

Result Explanations

In addition to the numerical results, expanded explanations of the results may also appear. To turn off this explanatory text, choose the Statistics menu Report Options... command and uncheck the Explain Test Results option.

The number of decimal places displayed is also set in the Report Options dialog. For more information on setting report options, see SETTING REPORT OPTIONS.

Normality Test

Normality test results show whether the data passed or failed the test of the assumption that the samples were drawn from normal populations and the P value calculated by the test. All parametric tests require normally distributed source populations.

This result is set in the Options for t-test dialog.
Equal Variance Test
Equal Variance test results display whether or not the data passed or failed the test of the assumption that the samples were drawn from populations with the same variance and the P value calculated by the test. Equal variance of the source population is assumed for all parametric tests.

Summary Table

SigmaStat can generate a summary table listing the sizes N for the two samples, number of missing values, means, standard deviations, and the standard error of the means (SEM). This result is displayed unless you disable the Summary Table option in the Options for t-test dialog.

- N (Size) The number of non-missing observations for that column or group.

- Missing The number of missing values for that column or group.

- Mean The average value for the column. If the observations are normally distributed the mean is the center of the distribution.

- **Standard Deviation**: A measure of variability. If the observations are normally distributed, about two-thirds will fall within one standard deviation above or below the mean, and about 95% of the observations will fall within two standard deviations above or below the mean.

- **Standard Error of the Mean**: A measure of the approximation with which the mean computed from the sample approximates the true population mean.

t Statistic

The t-test statistic is the ratio

T= difference between means of groups / standard error of the difference between means

The standard error of the difference is a measure of the precision with which this difference can be estimated.

You can conclude from "large" absolute values of t that the samples were drawn from different populations. A large t indicates that the difference between the treatment group means is larger than what would be expected from sampling variability alone (i.e., that the differences between the two groups are statistically significant). A small t (near 0) indicates that there is no significant difference between the samples.

Degrees of Freedom Degrees of freedom represents the sample sizes, which affect the ability of the t-test to detect differences in the means. As degrees of freedom (sample sizes) increase, the ability to detect a difference with a smaller t increases.

P Value

The P value is the probability of being wrong in concluding that there is a true difference in the two groups (i.e., the probability of falsely rejecting the null hypothesis, or committing a Type I error, based on t). The smaller the P value, the greater the probability that the samples are drawn from different populations. Traditionally, you can conclude there is a significant difference when $P < 0.05$.

Confidence Interval for the Difference of the Means

If the confidence interval does not include zero, you can conclude that there is a significant difference between the proportions with the level of confidence specified. This can also be described as $P < a$ (alpha), where a is the acceptable probability of incorrectly concluding that there is a difference.

The level of confidence is adjusted in the Options for t-test dialog; this is typically 100(1-a), or 95%. Larger values of confidence result in wider intervals and smaller values in smaller intervals. For a further explanation of a, see Power below. This result is set Options for t-test dialog.

Power

The power, or sensitivity, of a t-test is the probability that the test will detect a difference between the groups if there really is a difference. The closer the power is to 1, the more sensitive the test.

t-test power is affected by the sample size of both groups, the chance of erroneously reporting a difference, a (alpha), the difference of the means, and the standard deviation.

This result is set in the Options for t-test dialog.

Alpha (a)

Alpha (a) is the acceptable probability of incorrectly concluding that there is a difference. An error is also called a Type I error (a Type I error is when you reject the hypothesis of no effect when this hypothesis is true).

The value is set in the Options for t-test dialog; a value of $a = 0.05$ indicates that a one in twenty chance of error is acceptable, or that you are willing to conclude there is a significant difference when $P < 0.05$.

Smaller values of a result in stricter requirements before concluding there is a significant difference, but a greater possibility of concluding there is no difference when one exists (a Type II error). Larger values of a make it easier to conclude that there is a difference but also increase the risk of reporting a false positive (a Type I error).

RECOMMENDED READINGS

1. Black, Hubert M., (1972) *Social Statistics*, New York: McGraw-Hill Book Co.

2. Denscombe, M., 2007. *The good research guide for small-scale social research projects*. 3rd ed. Maidenhead, UK: Open University Press.
3. Goode W. and P. Hatt (1952) *Methods in Social Research*, NY.
4. Hagan F.E., (1993). *Research Methods in Criminal Justice and Criminology*, Prentice Hall Englewood Cliffs, NJ.
5. Hays, William L., (1973) *Statistics for Social Sciences*, New York: Holt, Rinehart and Winston.
6. Hirchi, T., and Selivin, H.C., (1967). *Delinquency Research: An appraisal of Analytical methods*, New York: Free Press.
7. Kerlinger, Fred N., (1983). *Foundations in Behavioral research*, Delhi: Surjeet Publications, 1983.
8. Krishnaswami O.R. (1993) *Methodology of Research in Social Sciences*, Himalaya Publishing House, Bombay.
9. Nie, Normal H., et al. (1975) *Statistical Packages for the Social Sciences*, New York, McGraw Hill.
10. Wilkinson and Bhandarkar (1993). *Methodology and Techniques in Social Research,* Himalaya Publishing House, Bombay.

OTHER RELATED WEBSITES

guides.library.utoronto.ca/c.php?g=251500&p=1674315
http://www.acrl.org/ala/mgrps/divs/acrl/about/sections/is/iscommittees/webpages/research/researchscholarship_.pdf
http://www.ala.org/ala/mgrps/divs/acrl/about/sections/is/projpubs/researchagendalibrary.cfm>.
https://archive.org/...fundamentalOfResearchMethodologyAndStatistics/1.fundamental
https://arxiv.org/pdf/physics/0601009
https://books.google.co.in/books?isbn=1446206602
https://islmblogblog.files.wordpress.com/2016/05/rme-edu-helpline-blogspot-com.pdf
web.ftvs.cuni.cz/hendl/metodologie/introduction-to-research-methods.pdf
www.ala.org/acrl/sites/ala.org.acrl/files/.../is/iswebsite/.../bibliographyresearch.doc
www.davidkelly.ie/research/...resources/complete-methodology-bibliography/
www.emeraldgrouppublishing.com/research/guides/methods/
www.statcan.gc.ca/pub/12-587-x/12-587-x2003001-eng.pdf
www.pondiuni.edu.in/sites/default/files/Research%20Methodology.pdf
www.springer.com › Home › Psychology › Cognitive Psychology
https://books.google.co.in/books?isbn=1410606724
https://books.google.co.in/books?isbn=1118626117
http://ieeexplore.ieee.org/abstract/document/1709988/
https://books.google.co.in/books?hl=en&lr=&id=FOTzBwAAQBAJ&oi=fnd&pg=PA1&dq=FUNDAMENTALS+OF+RESEARCH+METHODS+AND+STATISTICAL+APPLICATIONS+websites+references&ots=5Sp5S84nmV&sig=bDMojkwVMJl4RtTzLwziurt03t4#v=onepage&q