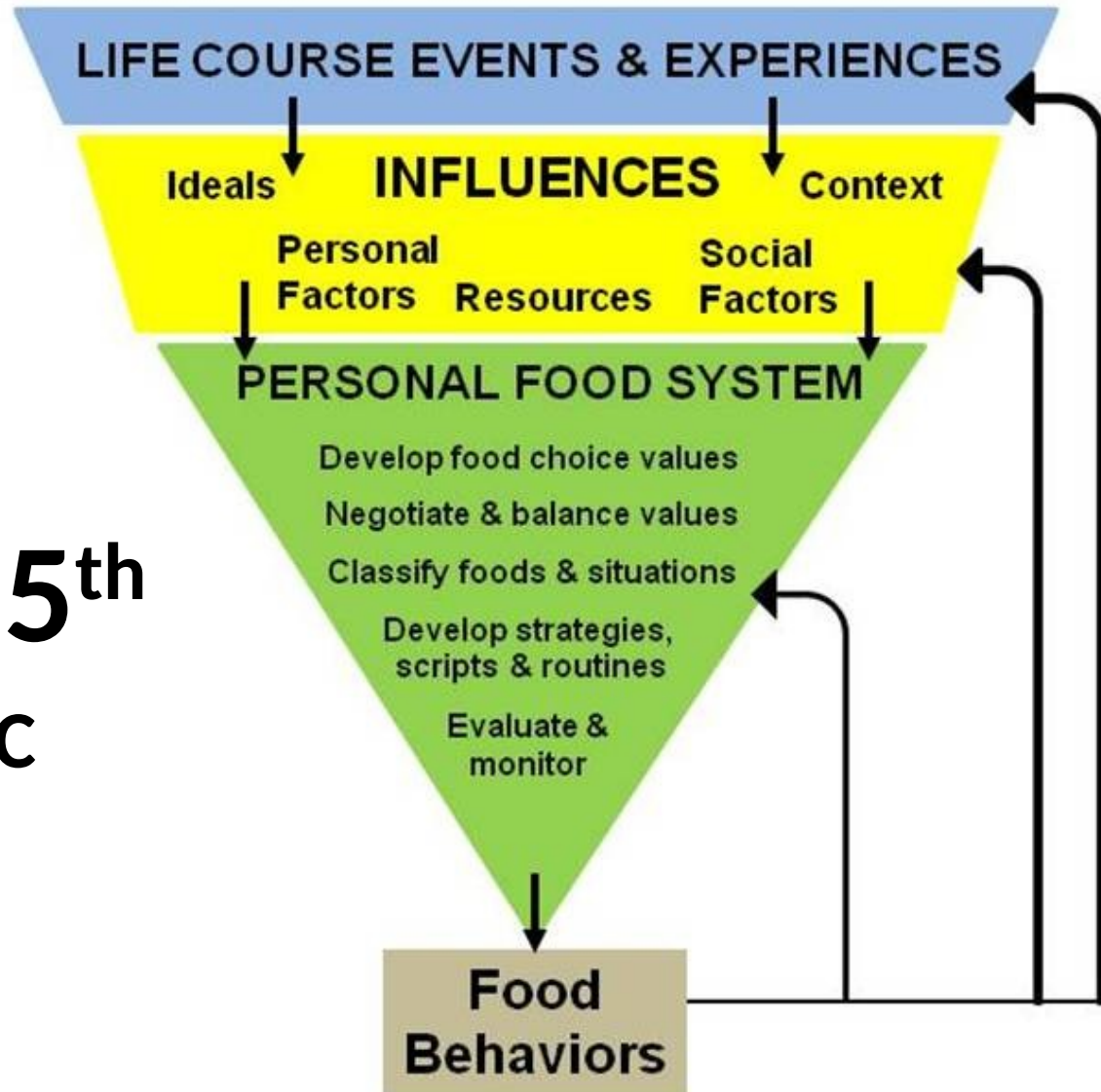


RESEARCH & DEVELOPMENT TECHNIQUE



Year: 5th

2008 E.C

CHAPTER 1

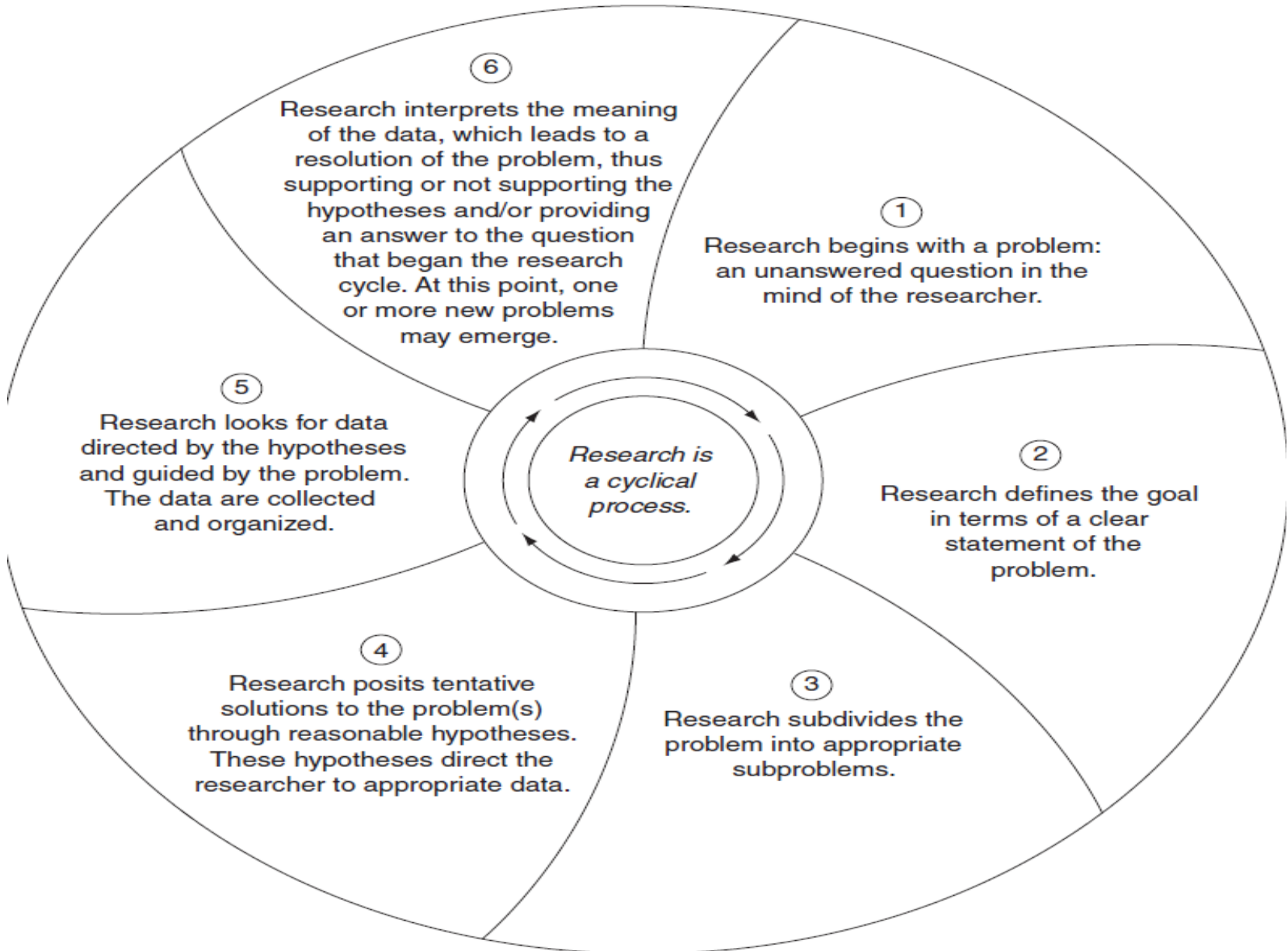
DEFINITION AND CONCEPT OF RESEARCH

- The word research is derived from the Latin word meaning to know.
- **Research is:**
 - a systematic process of
 - ✓ Collecting,
 - ✓ analyzing, and
 - ✓ Interpreting information (data) in order to increase our understanding of a phenomenon about which we are interested or concerned.

- The term 'research' refers to the systematic method consisting of
 - Enunciating the problem,
 - Formulating a hypothesis,
 - Collecting the facts or data,
 - analyzing the facts and
 - Reaching certain conclusions either in the form of
 - ✓ solutions(s) towards the concerned problem or
 - ✓ in certain generalizations for some theoretical formulation.

- Although research projects vary in complexity and duration, research typically has eight distinct c/x:
 - ✓ originates with a question or problem.
 - ✓ requires clear articulation of a goal.
 - ✓ requires a specific plan for proceeding.
 - ✓ usually divides the principal problem into more manageable sub-problems.
 - ✓ is guided by the specific research problem, question, or hypothesis.
 - ✓ accepts certain critical assumptions.
 - ✓ requires the collection and interpretation of data in an attempt to resolve the problem that initiated the research.
 - ✓ Research is, by its nature, cyclical or, more exactly, helical.

THE RESEARCH PROCESS IS CYCLICAL



OBJECTIVES OF RESEARCH

- The purpose of research is to
 - ✓ discover answers to questions through the application of scientific procedures.
 - ✓ find out the truth which is hidden and which has not been discovered as yet

- Though each research study has its own specific purpose, we may think of research objectives as falling into a number of following broad groupings:
- To gain familiarity with a phenomenon or to achieve new insights into it (studies with this object in view are termed as *exploratory* or *formulative* research studies);
- To portray accurately the characteristics of a particular individual, situation or a group (studies with this object in view are known as *descriptive* research studies);
- To determine the frequency with which something occurs or with which it is associated with something else (studies with this object in view are known as *diagnostic* research studies);
- To test a hypothesis of a causal relationship between variables (such studies are known as *hypothesis-testing* research studies).

Motivation of Research

- What makes people to undertake research? This is a question of fundamental importance. The possible motives for doing research may be either 1 or more of the following
 - ✓ Desire to get a research degree along with its consequential benefits;
 - ✓ Desire to face the challenge in solving the unsolved problems,
 - ✓ Desire to get intellectual joy of doing some creative work;
 - ✓ Desire to be of service to society;
 - ✓ Desire to get respectability.

- However, this is not an exhaustive list of factors motivating people to undertake research studies. Many more factors such as
 - ✓ directives of government,
 - ✓ employment conditions,
 - ✓ curiosity about new things,
 - ✓ desire to understand causal relationships,
 - ✓ social thinking and awakening, and the like may as well motivate people to perform research operations

Chapter 2

TYPES OF RESEARCH

- Research can be classified in terms of:
 - ✓ goal of research,
 - ✓ specific objectives of research,
 - ✓ approaches of research,
 - ✓ designs,
 - ✓ the type of data used in research, and
 - ✓ Fields of study.

Classification of Research based on the Goal of Research

1. fundamental research

- mainly concerned with generalisations and with the formulation of a theory.
- “Gathering knowledge for knowledge’s sake is termed ‘pure’ or ‘basic’ research.”
- Research concerning some natural phenomenon or relating to pure mathematics are examples of fundamental research
- Similarly, research studies, concerning human behaviour carried on with a view to make generalisations about human behaviour, are also examples of fundamental research

- The major **aims** of basic research include:
 - ✓ Obtaining and using empirical data to formulate, expand, or evaluate theory; and
 - ✓ Discovery of knowledge solely for the sake of knowledge.
- Hence, basic research may take any of the following forms:
 - **Discovery**: where a totally new idea or explanation emerges from empirical research which may revolutionize thinking on that particular topic.
 - **Invention** : where a new technique or method is created.
 - **Reflection** : where an existing theory, technique or group of ideas is re-examined possibly in a different organizational or social context.

2. Applied research

- aims at finding a solution for an immediate problem facing a society or an industrial/business organisation
- research aimed at certain conclusions (say, a solution) facing a concrete social or business problem is an example of applied research.
- Research to identify social, economic or political trends that may affect a particular institution or the copy research or the marketing research or evaluation research are examples of applied research

- Thus, the central aim of applied research is to
 - ✓ discover a solution for some pressing practical problem,
- whereas basic research is directed towards
 - ✓ finding information that has a broad base of applications and thus, adds to the already existing organized body of scientific knowledge.

- The purpose of applied research is
 - ✓ about testing theories, often generated by pure science, and applying them to real situations,
- Applied scientific research can be about
 - finding out the answer to a specific problem, such as
 - ✓ ‘Is global warming avoidable?’ or
 - ✓ ‘Does a new type of medicine really help the patients?’

- The primary purpose for applied research is discovering, interpreting, and the development of methods and systems for solving practical problems on a wide variety of real life situations of our world and the universe.
- Applied research:
 - ✓ Is conducted in relation to actual problems and under the conditions in which they are found in practice;
 - ✓ Employs methodology that is not as rigorous as that of basic research;
 - ✓ Yields findings that can be evaluated in terms of local applicability and not in terms of universal validity.

Classification of Research based on the Specific Objectives of Research

- According to the specific objectives research can be classified as:
 - ✓ descriptive,
 - ✓ explanatory, or
 - ✓ exploratory research

- ***Descriptive research***

- ✓ includes surveys and fact-finding enquiries of different kinds.
- ✓ The major purpose is description of the state of affairs as it exists at present.
- ✓ In social science and business research we quite often use the term *Ex post facto research* for descriptive research studies.
- ✓ The main characteristic of this method is that the researcher has no control over the variables; he can only report what has happened or what is happening.

- Most *ex post facto research* projects are used for descriptive studies in which the researcher seeks to measure such items as, for example,
 - ✓ frequency of shopping,
 - ✓ preferences of people, or
 - ✓ similar data.

Descriptive research

- sets out to describe and to interpret **what is**.
- It looks at individuals, groups, institutions, methods and materials in order to describe, compare, contrast, classify, analyze and interpret the entities and the events that constitute the various fields of inquiry.
- It aims to describe the state of affairs as it exists.
- The methods that come under descriptive research are:
 - ✓ Surveys
 - ✓ Correlation studies
 - ✓ Observation studies
 - ✓ Case studies

- The methods of research utilized in descriptive research are
 - ✓ survey methods of all kinds, including comparative and correlational methods.
- In *analytical research*, on the other hand, the researcher has to use facts or information already available, and analyze these to make a critical evaluation of the material.

Explanatory research

- aims at establishing the cause and effect relationship between variables.
- The researcher uses the facts or information already available to analyze and make a critical evaluation of the data/information.
- There are two types of explanatory research:
 1. experimental
 2. Ex post facto research

Experimental research

- Experimental research involves comparing two groups on one outcome measure to test some hypothesis regarding causation.
- The key element in true experimental research is scientific control and the ability to rule out alternative explanations

Ex post facto research

- is a method of teasing out possible antecedents of events that have happened and cannot, therefore, be engineered or manipulated by the investigator.
- means **after the fact** or **retrospectively** and refers to those studies which investigate possible cause-and-effect relationships by observing an existing condition or state of affairs and searching back in time for plausible causal factors.

Exploratory research

- Less formal, sometimes even unstructured and focuses on gaining background information and helps to better understand and clarify a problem.
- It can be used to develop hypotheses and to develop questions to be answered.
- is conducted when there are few or no earlier studies to which references can be made for information.
- is a type of research conducted because a problem has not been clearly defined.
- Its purpose is to gain background information and better understand and clarify a problem.

- Exploratory research helps to:
 - ✓ determine the best research design,
 - ✓ develop hypotheses
 - ✓ develop questions to be answered,
 - ✓ understand how to measure a variable,
 - ✓ determine data collection method, and
 - ✓ Determine selection of subjects.

- It can be said that exploratory research is generally a precursor to a more formal study.
- It helps save time, and resources.
- If a researcher is starting a new project, she/he probably should start with exploration.
- Exploratory research often relies on secondary research such as
 - ❖ reviewing available literature and/or data, or
 - ❖ qualitative approaches such as
 - informal discussions with consumers, employees, management or competitors, and
 - more formal approaches through
 - ✓ in-depth interviews, focus groups, projective methods, case studies or pilot studies.

Classification of Research based on Approaches of Research

- Research can be classified as
 - ✓ qualitative research and
 - ✓ quantitative research when the issue at hand is the approaches to be employed in conducting research.
- As mentioned above, this structure is mainly for educational purpose

- Quantitative methods are identified with the so-called “hard science” disciplines, whereas qualitative methods, with the social sciences.
- **Qualitative Research**
 - ✓ involves studies that do not attempt to quantify their results through statistical summary or analysis.
 - ✓ seeks to describe various aspects about behavior and other factors studied in the social sciences and humanities.
 - ✓ In qualitative research data are often in the form of descriptions, not numbers

- **Quantitative research**

- ✓ Is based on the measurement of quantity or amount.
- ✓ Involves the generation of data in quantitative form which can be subjected to rigorous quantitative analysis in a formal and rigid fashion.
- ✓ It is applicable to phenomena that can be expressed in terms of quantity.
- Qualitative approach can be further sub-classified into
- ***Experimental***: characterised by much greater control over the research environment and in this case some variables are manipulated to observe their effect on other variables.

- ***Inferential***: to form a data base from which to infer characteristics or relationships of population.
- This usually means survey research where a sample of population is studied (questioned or observed) to determine its characteristics, and it is then inferred that the population has the same characteristics.
- ***simulation approaches*** : involves the construction of an artificial environment within which relevant information and data can be generated.
- This permits an observation of the dynamic behaviour of a system (or its sub-system) under controlled conditions.

Qualitative research,

- Is concerned with subjective assessment of attitudes, opinions and behaviour
- Generally, the techniques of focus group interviews, projective techniques and depth interviews are used
- Is concerned with qualitative phenomenon, i.e., phenomena relating to or involving quality or kind.
- For instance, when we are interested in investigating the reasons for human behaviour (i.e., why people think or do certain things), we quite often talk of 'Motivation Research', an important type of qualitative research.
- This type of research aims at discovering the underlying motives and desires, using in depth interviews for the purpose.

- Other techniques of such research are word association tests, sentence completion tests, story completion tests and similar other projective techniques.
- Attitude or opinion research i.e., research designed to find out how people feel or what they think about a particular subject or institution is also qualitative research.
- Qualitative research is especially important in the **behavioural sciences** where the aim is to discover the underlying motives of human behaviour.

Characteristics of Qualitative and Quantitative Research

Characteristics	Qualitative Research	Quantitative Research
Typical Data Collection Methods	Participant observation, semi structured interviews, introspection.	Laboratory observations, questionnaire, schedule or structured interviews.
Formulation of Questions and Answers	Open loosely specified questions and possible answers. Questions and answers are exchanged in two way communication between researcher and research participant.	Closed questions and answer categories to be prepared in advance.
Selection of Respondents	Information maximization guides the selection of respondent. Every respondent may be unique (key person).	Representativeness as proportion of population N. Sample selection, sample size according to assumptions about distribution in population N. Respondents should be directly comparable.
Timing of Analysis	Parallel with data collection	After data collection
Application of Standard Methods of Analysis	Are rarely used. Methods of analysis are formulated during the data collection process.	Standard statistical methods are frequently used
Typical forms of Analysis	Critical analysis and interpretation of source materials. Selection, systematizing and summarizing interview transcripts and observations.	Cross tabulations, correlation analysis and tests of significance on numerical data
The Role of Theories in the Analysis	Existing theories are typically used only as point of departure for the analysis. Theories are further developed by forming new concepts and relations. The contents of the new concepts are studied and illustrated. Practical application of theory is illustrated by cases.	A-priori deducted theories are operationalised and tested on data. The process of analysis is basically deductive

Quantitative Research

- is the systematic and scientific investigation of quantitative properties and phenomena and their relationships.
- The objective of quantitative research is to develop and employ mathematical models, theories and hypotheses pertaining to natural phenomena.
- It usually starts with a theory or a general statement proposing a general relationship between variables.

Table 1. Comparison of quantitative and qualitative research approach

Table 1. Comparison of quantitative and qualitative research approaches		
	Quantitative	Qualitative
General framework	Seek to confirm hypotheses about phenomena	Seek to explore phenomena
	Instruments use more rigid style of eliciting and categorizing responses to questions	Instruments use more flexible, iterative style of eliciting and categorizing responses to questions
	Use highly structured methods such as questionnaires, surveys, and structured observation	Use semi-structured methods such as in-depth interviews, focus groups, and participant observation
Analytical objectives	To quantify variation	To describe variation
	To predict causal relationships	To describe and explain relationships
	To describe characteristics of a population	To describe individual experiences
		To describe group norms
Question format	Closed-ended	Open-ended
Data format	Numerical (obtained by assigning numerical values to responses)	Textual (obtained from audiotapes, videotapes, and field notes)
Flexibility in study design	Study design is stable from beginning to end	Some aspects of the study are flexible (for example, the addition, exclusion, or wording of particular interview questions)
	Participant responses do not influence or determine how and which questions researchers ask next	Participant responses affect how and which questions researchers ask next
	Study design is subject to statistical assumptions and conditions	Study design is iterative, that is, data collection and research questions are adjusted according to what is learned

Classification of Research based on Designs

- Once the researcher has determined the specific question to be answered and has operationalized the variables and research question into a clear, measurable hypothesis, the next task is to consider a suitable research design.
- Although there are endless ways of classifying research designs, they usually fall into one of three general categories:
 - ✓ experimental,
 - ✓ quasi-experimental, and
 - ✓ non-experimental

Classification or Research by Type of Data

- Depending of the type of data generated and used research can be classified as
 - ✓ Primary research (also called field research)
 - involves the collection of data that does not already exist
 - ✓ Secondary research (also known as desk research).
 - involves the summary, collation and/or synthesis of existing data rather than generating primary data, where data are collected from, for example, research subjects or experiments.

Classification of Research by Fields of Study

- Research can also be classified based on fields of study. Therefore, there are:
 - ✓ natural science research,
 - ✓ social science research,
 - ✓ educational research,
 - ✓ behavioral science research,
 - ✓ Health science research, *etc.*

Other types of research

- The objective of exploratory research is the development of hypotheses rather than their testing, whereas
- formalized research studies are those with substantial structure and with specific hypotheses to be tested.
- *Historical research* is that which utilizes historical sources like documents, remains, etc. to study events or ideas of the past, including the philosophy of persons and groups at any remote point of time. Research can also be classified as *conclusion-oriented* and decision-oriented.

- While doing conclusion oriented research, a researcher is free to pick up a problem, redesign the enquiry as he proceeds and is prepared to conceptualize as he wishes. Decision-oriented research is always for the need of a decision maker and the researcher in this case is not free to embark upon research according to his own inclination. Operations research is an example of decision oriented research since it is a scientific method of providing executive departments with a quantitative basis for decisions regarding operations under their control.

CHAPTER 3

PROCESS IN RESEARCH PROPOSAL DEVELOPMENT

What is the Research Proposal?

- **A research proposal is:**
 - ✓ A document written by a researcher that provides a detailed description of the proposed program.
 - ✓ The detailed plan of study.
 - ✓ It is a document which sets out your ideas in an easily accessible way.

- Research proposals are written for various reasons, such as:
 - ✓ Budget request for the research they describe,
 - ✓ Certification requirements for research
 - ✓ As a task in tertiary education (e.g., before performing research for a dissertation), or
 - ✓ As a condition for employment at a research institution.

- A well-thought out and well-written proposal can be judged according to three main criteria.
 - ✓ Is it adequate to answer the research question(s), and achieve the study objective?
 - ✓ Is it feasible in the particular set-up for the study?
 - ✓ Does it provide enough detail that can allow another investigator to do the study and arrive at comparable results?
- Before an attempt is made to start with a research project, a research proposal should be compiled

Identifying a research topic

- Criteria for selecting a research topic:
 - ✓ Relevance/Significance
 - ✓ Avoidance of duplication
 - ✓ Urgency of data needed (timeliness)
 - ✓ Feasibility of study
 - ✓ Applicability of results
 - ✓ Interest to the researcher
 - ✓ Ethical acceptability

Components of a Research Proposal

- The basic components of a research proposal are the same in many fields. However, how they are phrased and staged may vary by discipline.
- ❖ Components of a research proposal are:
 - **1. Title page**
 - The title of your research proposal should state your topic exactly in the smallest possible number of words. Put:
 - ✓ your name,
 - ✓ the name of your department/faculty/college,
 - ✓ the name of your advisor(s) and
 - ✓ date of delivery under the title

2. Summary/Abstract

- ✓ The abstract is a one page brief summary of the thesis proposal.
- ✓ Do not put information in the abstract that is not in the main text of your research proposal.
- ✓ Do not put references, figures, or tables in the abstract
- A well-prepared summary enables the reader to
 - Identify the basic content of a document quickly and accurately,
 - Determine its relevance to their interests, and
 - Decide whether they need to read the document in it's entirely

3. Introduction/Background

- Its purpose is to establish a framework for the research, so that readers can understand how it is related to other research.
- The introduction also should address the following points:
 - ✓ Sufficient background information to allow the reader to understand the context and significance of the question you are trying to address.
 - ✓ Proper acknowledgement of the previous work on which you are building.
 - ✓ Sufficient references such that a reader could achieve a sophisticated understanding of the context and significance of the question.

- ✓ The introduction should be focused on the research question(s).
- ✓ All cited work should be directly relevant to the goals of the research.
- ✓ Explain the scope of your work, what will and will not be included.

4. Statement of the problem

- The problem statement describes the context for the study and it also identifies the general analysis approach.
- Effective problem statements answer the question
 - ✓ *“Why does this research need to be conducted.”*
- A well-articulated statement of the problem establishes
 - ✓ the foundation for everything to follow in the proposal and
 - ✓ will render less problematic most of the conceptual, rhetorical and methodological obstacles typically encountered during the process of proposal development.

5. Literature review

- A literature review is a description of the literature relevant to a particular field or topic.
- It gives an overview of
 - ✓ what has been said,
 - ✓ who the key writers are,
 - ✓ what are the prevailing theories and hypotheses,
 - ✓ what questions are being asked, and
 - ✓ what methods and methodologies are appropriate and useful.

6. Hypotheses /Questions

- In research studies, the term hypothesis implies a derivation, within a hypothetic-deductive theoretical system, of a particular assertion or prediction.
- The hypothesis is subject to test, i.e., to confirmation or rejection on empirical grounds.
- The term question implies an interrogative statement that can be answered by data, which is logically related to the same conceptual framework,
- but which does not necessarily stem from that framework through logical deduction

- **Questions** are most often used in qualitative inquiry, although their use in quantitative inquiry is becoming more prominent.
- **Hypotheses** are relevant to theoretical research and are typically used only in quantitative inquiry.
- A research question poses a relationship between two or more variables but phrases the relationship as a question; a hypothesis represents a declarative statement of the relations between two or more variables.

- **7. Conceptual framework**
- A conceptual framework is described as a set of broad ideas and principles taken from relevant fields of enquiry and used to structure a subsequent presentation.

8. Objective/Aim of the study

- The objectives of a research project summarise
 - ✓ what is to be achieved by the study.
- Objectives should be
 - ✓ simple (not complex),
 - ✓ specific (not vague),
 - ✓ stated in advance (not after the research is done),
and
 - ✓ stated using “action verbs” that are specific enough
to be measured

- Commonly, research objectives are classified into
 - ✓ **general objectives and specific objectives.**
- The general and specific objectives are logically **connected to each other** and the specific objectives are commonly considered as **smaller portions of** the general objectives.
- It is important to ascertain that the general objective is closely related to the statement of the problem.
- Properly formulated objectives will facilitate the development of your research methodology and will help to orient the
 - ✓ collection, analysis, interpretation and utilization of data.

9. Research methods, materials and procedures

- is really **the heart** of the research proposal
- decide exactly how you are going to achieve your stated objectives: *i.e.*,
 - ✓ what new data you need in order to shed light on the problem you have selected and how you are going to collect and process this data.
- Indicate the methodological steps you will take
 - ✓ to answer every question,
 - ✓ to test every hypothesis illustrated in the Questions /Hypotheses section or
 - ✓ address the objectives you set

Important components of the materials and methods section are:

Study design

- The type of research design chosen depends on:
 - ✓ The type of problem;
 - ✓ The knowledge already available about the problem; &
 - ✓ The resources available for the study.

- Generally, there are two main categories of research design:
 - ✓ observational study, and
 - ✓ experimental or intervention study.
- In the observational study,
 - ✓ the researchers **stand apart from events** taking place in the study. They simply observe and record.
- In the experimental or intervention study,
 - ✓ the researchers **introduce an intervention** and observe the events which take place in the study

Sampling

- **Sampling for quantitative studies**
- Sampling is the process of selecting a number of study units from a defined study population.
- Often research focuses on a large population that, for practical reasons, it is only possible to include some of its members in the investigation.
- You then have to draw a sample from the total population.
- The way you define your study population and your study unit depends on the problem you want to investigate and on the objectives of the study.

- The key reason for being concerned with sampling is that of *validity*—the extent to which the interpretations of the results of the study follow from the study itself and the extent to which results may be generalized to other situations with other people or situation.
- All probability samples involve the idea of *random sampling* at some stage.
- Probability sampling requires that a listing of all study units exists or can be compiled.
- This listing is called the **sampling frame**.

Examples of probability sampling

- **Simple random sampling:**
 - ✓ The guiding principle behind this technique is that each element must have an equal and nonzero chance of being selected.
- **Systematic sampling:**
 - ✓ The systematic random sampling technique begins with selecting one element at random in the sampling frame as the starting point; however, from this point onward, the rest of the sample is selected systematically by applying a predetermined interval.

Stratified sampling:

- ✓ begins with the identification of some variable, which may be related indirectly to the research question and could act as a confounder (such as geography, age, income, ethnicity, or gender).
- ✓ This variable is then used to divide the sampling frame into mutually exclusive *strata* or subgroups.

- **Cluster sampling:**

- ✓ It may be difficult or impossible to take a simple random sample of the units of the study population at random, because a complete sampling frame does not exist.
- ✓ Logistical difficulties may also discourage random sampling techniques (e.g., interviewing people who are scattered over a large area may be too time-consuming).

- **Multistage sampling:**

- ✓ Multistage cluster sampling is used when an appropriate sampling frame does not exist or cannot be obtained.
- ✓ Multistage cluster sampling uses a collection of pre-existing units or clusters to “stand in” for a sampling frame.

Sampling size

- **Sample size in quantitative studies**
 - ✓ Having decided **how** to select the sample, you have to determine the **sample size**.
 - ✓ The research proposal should provide information and justification about sample size.
 - ✓ It is not necessarily true that the bigger the sample, the better the study.
 - ✓ The larger the sample size, the higher the power of the study.
 - ✓ Large sample sizes are needed to detect small differences.

- **Sample size in qualitative studies**

- ✓ There are no fixed rules for sample size in qualitative research.
- ✓ The size of the sample depends on what you try to find out, and from what different informants or perspectives you try to find that out.
- ✓ You can start with two or four Focus Group Discussions (FGDs) depending on the complexity of the research objectives.

10. Work plan

- Work plan is a schedule, chart or graph that summarizes the different components of a research proposal and how they will be implemented in a coherent way within a specific time-span.
- It may include:
 - ✓ The tasks to be performed;
 - ✓ When and where the tasks will be performed;
 - ✓ Who will perform the tasks and the time each person will spend on them;
 - ✓ It describes the plan of assessing the ongoing progress toward achieving the research objectives;
 - ✓ The plan specifies how each project activity is to be measured in terms of completion, the time line for its completion;
 - ✓ A good work time plan enables both the investigators and the advisors to monitor project progress and provide timely feedback for research modification or adjustments

11. Budget and funding

- This is to secure funds from a funding organization to cover the cost of conducting a research project.
Typically, a proposal budget reflects direct and indirect costs.
- **Direct costs:**
- **Personnel:** Salaries and wages of all participants of the study. Principal investigator; supervisor; data collector; drivers; guards; data entry clerks, data analysis, report writing, etc
- **Consumable supplies:** stationeries, computers, chemicals, and educational materials
- **Equipments:**
- **Travel:** cost of project-related travel

- **Communications:** postage, telephone, telegram, fax, e-mail charges
- **Publication:** the cost incurred of preparing and publishing the results of the research. It includes: technical reports, manuscripts, illustrations, graphics, photography, slides, and overheads
- **Indirect costs:**
- Those costs incurred in support and management of the proposed activities that can not be readily determined by direct measurement. Examples include;
 - ✓ Overhead costs for institutions or associations
 - ✓ General administrative cost
 - ✓ Operational and maintenance
 - ✓ Depreciation and use allowance

12. References

- You must give references to all the information that you obtain from books, journals, and other sources.
- References may be made in the main text using
 - ✓ index numbers in brackets (Vancouver style) or
 - ✓ authors name (Harvard style).
- **For a journal paper give:**
 - ✓ the names of the authors,
 - ✓ the year of publication,
 - ✓ the title of the paper,
 - ✓ the title of the journal,
 - ✓ the volume number of the journal,
 - ✓ the first and last page numbers of the paper

- **For a book give:**
 - ✓ the author,
 - ✓ the year of publication,
 - ✓ the title, and the edition number if there is one,
 - ✓ the name of the publisher,
 - ✓ the page numbers for your reference.

- **For an internet reference give:**
 - ✓ the author of the web page,
 - ✓ the title of the item on the web page,
 - ✓ the date the item was posted on the web page
 - ✓ the date the item was accessed from the web page
 - ✓ the complete and exact URL.

- **13. Appendices/Annexes**
- Include any additional information you think might be helpful to a proposal reviewer.
- For example, include:
 - ✓ Questionnaire & other collection forms
 - ✓ Dummy tables
 - ✓ Biographical data on the principal investigator
 - ✓ The consent form (if any)

Flowchart: Steps in the development of a research proposal

Questions you must ask	Steps you will take	Important elements of each step
What is the problem and why should it be studied?	Selection, analysis and statement of the research problem	<ul style="list-style-type: none"> - problem identification - prioritising problems - analysis - justification
What information is available?	Literature review	<ul style="list-style-type: none"> - literature and other available information
Why do we want to carry out the research? What do we hope to achieve?	Formulation of research objectives	<ul style="list-style-type: none"> - general and specific objectives - hypotheses
What additional data do we need to meet our research objectives? How are we going to collect this information?	Research methodology	<ul style="list-style-type: none"> - variables - types of study - data collection techniques - sampling - plan for data collection - plan for data processing and analysis - ethical considerations - pre-test or pilot study
Who will do what, and when?	Work plan	<ul style="list-style-type: none"> - human resources - timetable
What resources do we need to carry out the study? What resources do we have?	Budget	<ul style="list-style-type: none"> - material support and equipment - money
How will the project be administered? How will utilisation of results be ensured?	Plan for project administration and utilisation of results	<ul style="list-style-type: none"> - administration - monitoring - identification of potential users
How will we present our proposal to relevant authorities, community and the funding agencies?	Proposal summary	<ul style="list-style-type: none"> - briefing sessions and lobbying

- NB: Development of a research process is a cyclical process.
- The double-headed arrows indicate that the process is never linear.
- *Adapted from the International Development Research Center: Designing and Conducting Health Systems Research Projects, VOLUME 1*

CHAPTER 4

PROCESS IN CONDUCTING RESEARCH

- **The research processes**
- There are a number of stages in the research process, although their number and description tends to vary between authors.
- One simplified view would see these as five stages which include
 - ✓ conceptualization,
 - ✓ contextualization,
 - ✓ data collection and / or generation,
 - ✓ data analysis and
 - ✓ reporting conclusions.

- **Conceptualization:**
 - ✓ Defining the 'problem';
 - ✓ establishing the research questions;
 - ✓ identifying the aims;
 - ✓ specifying the testable hypotheses;
 - ✓ deciding on the research approach;
 - ✓ identifying the most appropriate way of undertaking the data collection and/or
 - ✓ generation.
- This stage is the most difficult one for novice researchers.

- **Contextualization:**

- ✓ Putting the research in the context of similar research that has been done in the past.
- ✓ If there is similar research was done elsewhere, looking carefully the methodology used and the way data collected and the limitation of the work may help.

- **Data Analysis:**

- ✓ Method to collect data and the following data analysis technique should match; occasionally problem may emerge as one start to do data analysis

- **Reporting Conclusions:**

- ✓ Writing up and / or further dissemination.

- **Data Collection and /or generation:**

- ✓ Applying the chosen method(s).

Research Methods – data collection and analysis

- You need to operationalise the research problem by choosing the most suitable research method, or technique, for your specific study.
- There are a number of different research methods available and one should be selected which is most likely to meet the objective of the research and gather the correct type of information
- Research methods should be seen as complementing and supporting each other.
- In a more generalized way, research is often broken down into two different approaches:

- **Quantitative research:**

- ✓ research involving numerical or statistical data.
- ✓ Emphasis is on the **quantifiable observations** of the research i.e. numbers involved.
- ✓ This type of research is mainly objective.

- **Qualitative research:**

- ✓ an approach to gather non-numerical data and related ways of analysis where emphasis is on the **qualitative** results.
- ✓ Words and observations are used to express the reality where 'getting close to the data' and an 'in-depth' approach are key concerns.
- ✓ This type of research is mainly subjective.

- Below follows a list of individual research methods (observation studies, questionnaires, interviews, focus groups, document analysis, experimental research, and mathematical modeling) incorporating a brief description, advantages and disadvantages, ways of collecting data, general hints for carrying out the research and some examples of their use.

2.1 Observation studies

- Observation refers to the process of observing and recording events or situations.
- The technique is particularly useful for discovering how individuals or groups of people or animals behave, act or react.
- There are two main types of observation :
- **Participant observation**
 - ✓ is usually limited to studies of human subjects.
- **In non-participant observation**
 - ✓ the researchers simply observe the activities without taking part themselves.
 - ✓ Beside the study of human subjects, non-participant observation can also be used to study animal behavior.

2.2. Questionnaires

- A questionnaire is a type of survey where respondents write answers to questions posed by the researcher on a question form.
- Questionnaires are extremely flexible and can be used to gather information on almost any topic involving large or small numbers of people.
- The commonest type of questionnaire involves closed choice or fixed questions where the respondent is required to answer by choosing an option from a number of given answers, usually by ticking a box or circling an answer.
- These types of questionnaires only gather straight forward, uncomplicated information, and only simple questions can be asked.
- The open-ended questionnaire differs in that it allows the respondent to formulate and record their answers in their own words.
- These are more qualitative and can produce detailed answers to complex problems.

Example:

- **Closed choice question:**
- *People go to a bar for different reasons; for which of the following four reasons do you most go?*
 - a. *to meet friends*
 - b. *to drink alcohol*
 - c. *to watch sport*
 - d. *to play pool or darts*
- *What is the primary reason you applied to the BDU undergraduate program?*
 - a. *to expand my knowledge base*
 - b. *to earn a better salary*
 - c. *to get a promotion*
 - d. *had no specific reason, but did because my friends were applying*

Open ended question:

- *People go to the bar for different reasons; for what reason do you most go?*
.....
- *List the top five reasons why you applied to the BDU undergraduate program?*
.....
- There are advantages and disadvantages associated with each type of method.
- Open ended questions give a greater insight and understanding of the topic researched but may be difficult to classify and quantify and must be carefully interpreted.

- Fixed choice questions are
 - ✓ easy to classify and quantify,
 - ✓ require less time,
 - ✓ effort and ingenuity to answer but do not allow the respondents to qualify, develop or clarify their answers.

2.3 Interviews

- Interviews are limited to cases where the subjects of study are humans.
- Interviews are a type of survey where questions are delivered in a face-to-face encounter by an interviewer.
- The interview is like a conversation and has the purpose of obtaining information relevant to a particular research topic.

- There are advantages and disadvantages associated with each type of method
- Structured interviews
 - ✓ maximize reliability and are easier to classify and quantify. By contrast
- unstructured interviews
 - ✓ can give a greater insight and more in-depth understanding of the topic researched, but need more expertise to control and more time for analysis.

2.4 Focus Groups

- The focus group is a type of interview that involves carefully selected individuals who usually do not know each other.
- They generally consist of 7-10 members alongside the researcher.
- These individuals are selected as they hold particular characteristics which the researcher believes are necessary to the topic of focus.
- A group discussion is held in a permissive environment in order to extract opinions and share ideas and perceptions through group interaction.

- It is not necessary to reach a consensus.
- Focus groups are extremely useful in providing qualitative data which gives an insight into attitudes and perceptions difficult to obtain using other procedures.
- The researcher acts as a moderator and listener posing predetermined open ended questions which the respondents answer in any way they choose.

2.5 Document Analysis

- This refers to the process of using any kind of document, films, television programs and photographs as well as written sources, such as books, papers and letters, for analysis in relation to a particular research question.
- It can be used as the singular method of research or as a supplementary form of inquiry.

- Document analysis, also referred to as content analysis, differs from the majority of research methods in two major ways.
 - It is an indirect form of research; it is something that has been produced, so the investigator is not generating original data.
 - It is an 'unobtrusive', or 'non-reactive' method.
 - This refers to the fact that the document will not be affected in any way by your research; it cannot react as a human can.

- **2.6 Mathematical Modeling**

- ✓ Mathematical modeling can be used to analyze relationships between different variables and to predict possible outcomes, or causal effects.

Advantages Disadvantages	Advantages Disadvantages
Can extend powers of deductive reasoning.	Does not explain why variables are linked to particular outcomes - can not explain why particular variables are important.
Attempts to be objective - maths is 'neutral'.	Model produced is limited to one situation and therefore may not apply to others.
Is an aid to causal explanation and can therefore help calculate the effects of actions.	Inability to distinguish causal from accidental relations.
	Could be built on preconceptions.

2.7 The experimental method

- This method involves setting up an experiment in order to test a particular theory or hypothesis.
- In its simplest terms experimentation is concerned with seeing what changes occur if something new is tried out and with the effects of these changes on something else.
- It is a method particularly associated with the physical and life sciences although the approach is also used in social sciences such as psychology, health care and education.
- There are two different types of experiment, the laboratory experiment and the field experiment.

- **1. Laboratory experiments:** the researcher will conduct a small-scale study where subjects can be:
 - ✓ manipulated,
 - ✓ observed and
 - ✓ tested in a highly controlled environment.
- In these types of studies, the data obtained is often subjected to statistical analysis
- **2. A field experiment:** is an experiment that takes place outside the laboratory. This leads to a decrease in researcher control, which may hide the effects of changes made, but the results gained can be still be generalized to the real world.

- Experimental Research is often used where:
 - ✓ There is time priority in a causal relationship (cause precedes effect),
 - ✓ There is consistency in a causal relationship (a cause will always lead to the same effect), and
 - ✓ The magnitude of the correlation is great

Advantages and disadvantages of experimental method are given below.

Advantages	Disadvantages
Ideas can be tested in a controlled way.	Where human subjects are involved it is generally viewed as unethical.
Ideal for investigating causal relationships	Results may be different in the real world to those discovered in a controlled environment.
Can generalize effects.	The influence of all variables can never be eliminated; many different circumstances potentially function as variables that can affect the outcome.
Scientifically validated findings give greater value to research.	Large amount of preparation is required
	Humans may respond to expectations of the experiment not to the experiment itself

Steps in Research Process:

- Formulating the Research Problem
- Extensive Literature Review
- Developing the objectives
- Preparing the Research Design including Sample Design
- Collecting the Data
- Analysis of Data
- Generalisation and Interpretation
- Preparation of the Report or Presentation of Results-
Formal write ups of conclusions reached.

CHAPTER 5

PRINCIPLES OF EXPERIMENTAL DESIGN AND ANALYSIS

5.1 Fundamental concepts in experimental design

- Experiments are **performed** by **investigators** in virtually all fields of inquiry, usually
 - ✓ to **discover** something about a particular process or system
- An experiment is:
 - ✓ a test or series of tests in which purposeful **changes are made** to the **input** variables of a process or system so that we may observe and identify the reasons for changes **that may be observed in the output** response

- *Experimentation plays an important role in*
 - ✓ *new product development*
 - ✓ *process improvement*
 - ✓ *process optimization*
 - ✓ *Ingredient optimization etc.*
- The **objective** in many cases may be
 - ✓ to **develop a robust process**, i.e., a process affected minimally by **external sources of variability**

❖ Example

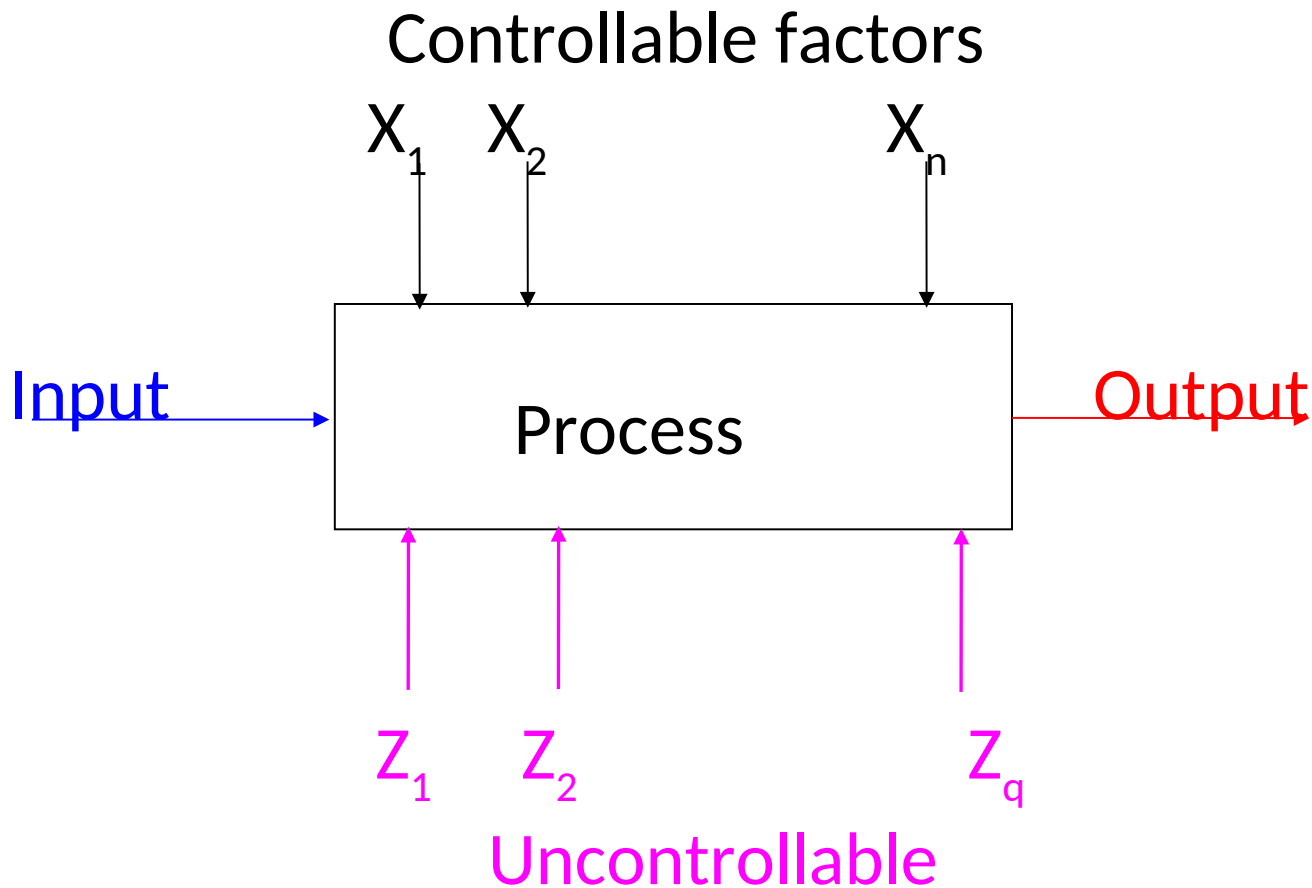
- ✓ The effect of barrel temperature on the rehydration kinetics of corn-based breakfast cereal
 - Barrel temperatures considered
 - ✓ 1. **110°C** 2. **130°C**, 3. **150°C**
 - **Objective:** to identify which barrel temperature results best rehydration (hydration) rate

- While considering this simple experiment, a number of important questions could be raised
 - ✓ *Are these three T° s the only ones* of potential interest?
 - ✓ *Are there any other factors that might affect hydration rate* that should be *investigated* or *controlled* in this experiment?
 - ✓ *How many samples should be tested at each temperature level?*
 - ✓ *How should the test samples be assigned to the barrel temperatures*, and in what order should the data be collected?
 - ✓ *What method of data analysis should be used?*
- All of these questions, and perhaps many others, will have to be answered satisfactorily *before the experiment is performed*

- In any experiment, *the results and conclusions* that can be drawn **depend** to a large extent on the manner in which the *data were collected*.
- Suppose that the investigator failed to realize
 - ✓ the initial moisture content of the samples extruded *using different barrel temperature* could possibly differ.

- When the mean hydration rate is **compared**, the investigator is **unable** to say:
 - ✓ how much of **the observed difference** is the result of the barrel temperature and
 - ✓ how much is the result of inherent differences between the initial moisture contents.
- Thus, the method of **data collection** has adversely **affected the conclusions** that can be drawn from the experiment.
- **Hence: The effect of barrel temperature and initial moisture content is confounded**

- In general, experiments are used to study
 - ✓ the performance of **processes** and **systems**.
- One can usually visualize the process as a combination of
 - ✓ machines, methods, people, and other resources ***that transform some*** input (often a material) into an output that has one or more **observable responses**.
- ✓ The process or system can be represented by the model shown below.



- Some of the process variables $X_1, X_2, \dots, X_n$ are controllable, whereas other variables $Z_1, Z_2, \dots, Z_q$ are uncontrollable (although they may be controllable for purposes of a test).

- The objectives of the experiment may include the following:
 - ✓ Determining which **variables are most influential** on the response Y
 - ✓ Determining where to set the influential X 's so that Y is almost always near the **desired nominal value**
 - ✓ Determining where to set the influential X 's so that "**variability in Y is small**"
 - ✓ Determining where to set the influential X 's 'so that the 'effects of uncontrollable variables $Z_1, Z_2, \dots Z_q$ are minimized
- Experiments often involve several factors.
- Usually the objective is to determine the influence that these factors have on the output response.

- The general approach to planning and conduction the experiment is called the **strategy of the experimentation** such as
 - ✓ *Best-guess approach*
 - ✓ *One-factor at a time approach*
 - ✓ *Factorial experiments*

APPLICATION OF EXPERIMENTAL DESIGN

- Experimental design methods have found broad application in many disciplines.
- Experimentation is viewed as **part of the scientific process** and as **one of the ways we learn about how systems or processes work**.
- Generally, one learn through a series of activities in which we make conjecture (assumptions) about a process, perform experiments to generate data from the process, and then use the information from the experiment to establish new conjectures, which lead to new experiments, and so on.

- **Experimental design** is a critically important **tool** in the
 - ✓ engineering, science and technology world for improving the ***performance of various processes***.
- It also has extensive application **in the development of new processes**.
- The application of experimental design techniques early in process development can result in:
 - ✓ 1. Improved process **yields**
 - ✓ 2. Reduced **variability** and closer conformance to nominal or target requirements
 - ✓ 3. Reduced development **time**
 - ✓ 4. Reduced overall **costs**

BASIC PRINCIPLES OF EXPERIMENTAL DESIGN

- If an experiment is to be performed most efficiently, a scientific approach to planning the experiment must be employed.
- Statistical design of experiments refers to the process of planning the experiment so that appropriate data that can be analyzed by statistical methods will be collected, resulting in valid and objective conclusions.
- The statistical approach to experimental design is necessary if one wishes to draw meaningful conclusions from the data.

- When the problem involves data that are subject to experimental errors, statistical methodology is the only **objective** approach to analysis.
- Thus, there are two aspects to any **experimental problem**:
 - ✓ *the design of the experiment*
 - ✓ *the statistical analysis of the data.*
- *These two aspects are closely related because the **method of analysis depends directly on the design employed***

- The three basic principles of experimental design are:
 - ✓ Replication
 - ✓ Randomization
 - ✓ Blocking

I Replication

- Replication has two important properties.
- **1. It allows the experimenter to obtain an estimate of the experimental error.**
 - ✓ This estimate of error becomes a basic unit of measurement for determining whether observed differences in the data are really *statistically significant*.
- 2. If the sample mean (e.g., $\bar{y}$) is used to estimate the effect of a factor in the experiment, *replication permits the experimenter to obtain a more precise estimate of this effect*.
- There is an important distinction between **replication** and **repeated measurements**

II Randomization

- Randomization is the **cornerstone** underlying the use of statistical methods in experimental design.
- Randomization refers to
 - ✓ 1. The **allocation (assignment)** of the experimental material
 - ✓ 2. The **order** in which the individual runs or trials of the experiment are to be performed
- Statistical methods require that the observations (or errors) be independently distributed random variables. Randomization usually, makes this assumption valid.
- By properly randomizing the experiment, **one also assists in "averaging out" the effects of extraneous factors that may be present.**

- Suppose that the specimens in the above experiment are slightly different in feed moisture content and that the rehydration rate may be affected by **feed moisture content**.
- If all the specimens subjected to the 110°C barrel temperature have higher feed moisture content than those subjected to 130 and 150°C barrel temperature, we may be **introducing systematic bias** into the experimental results.
- This bias handicaps one of the barrel temperatures and consequently invalidates the results.
- *Randomly assigning the specimens to the barrel temperature alleviates this problem*

III Blocking

- Blocking is a design technique used to improve the precision with which comparisons among the factors of interest are made.
- *Often blocking is used to reduce or eliminate the variability transmitted from nuisance factors*; that is, factors that may influence the experimental response but in which we are not directly interested.
- ❖ For example, an experiment in a food process may require two batches of raw material to make all the required runs.
- However, there could be differences between the batches due to supplier-to-supplier variability.
- ✓ If we are not specifically interested in this effect, we would think of the batches of raw material as a nuisance factor.

- Generally, a block is a set of **relatively homogeneous** experimental conditions.
- In the food process example, each batch of raw material would form a block, because
 - ✓ the variability within a batch would be expected to be smaller than the variability between batches.

GUIDELINES FOR DESIGNING EXPERIMENTS

- To use the statistical approach in designing and analyzing an experiment, it is necessary to have a clear idea, in advance, of exactly
 - ✓ what is to be studied
 - ✓ how the data are to be collected and
 - ✓ at least a qualitative understanding of how these data are to be analyzed.

- **An outline of the recommended procedure is:**
 - ✓ 1. *Recognition of & statement of the problem*
 - ✓ 2. *Choice of factors levels and range*
 - ✓ 3. *Selection of the response variable*
 - ✓ 4. *Choice of experimental design*
 - ✓ 5. *Performing the experiment*
 - ✓ 6. *Statistical analysis of the data*
 - ✓ 7. *Conclusions and recommendations*

1 Recognition of and statement of the problem

- This may seem to be a rather obvious point, but in practice it is often not simple to realize that a problem requiring experimentation exists, nor is it simple to develop a clear and generally accepted statement of this problem.
- It is necessary to develop all ideas about the objectives of the experiment.
- Usually, it is important to solicit input from all concerned parties.
- For this reason a team approach to designing experiments is recommended

- It is usually helpful to prepare a list of specific problems or questions that are to be addressed by the experiment.
- A clear statement of the problem often contributes substantially
 - ✓ to better understanding of the phenomenon being studied and the final solution of the problem.
 - ✓ It is also important to keep the **overall objective** in mind; for example,
 - is this a new process or system -in which case the initial objective is likely to be **characterization** or factor screening-or
 - is it a mature or reasonably well-understood system that has been previously **characterized-in** which case the objective may be **optimization**?

- There are many possible **objectives** of an experiment, including
 - ✓ **confirmation** (is the system performing the same way now that it did in the past?),
 - ✓ **discovery** (what happens if we explore new materials, variables, operating conditions, etc.), and
 - ✓ **stability** (under what conditions do the response variables of interest seriously' degrade?).
- Obviously, the **specific questions** to be addressed in the experiment relate directly to the **overall objectives**

2 Choice of factors, levels and range

- When considering the factors that may influence the **performance of a process or system**, the experimenter usually discovers that these factors **can be classified** as either
 - ✓ **Potential Design Factors** or
 - ✓ **Nuisance Factors**

1 POTENTIAL DESIGN FACTORS

- The potential design factors are those factors that the experimenter may wish to vary in the experiment.
- Often we find that there are a lot of potential design factors, and some further classification of potential factor is helpful.
- Some useful classifications are
 - ✓ design factors,
 - ✓ held-constant factors, and
 - ✓ allowed-to-vary factors.

- **The design factors:** are the factors **actually selected** for study in the experiment.
- **Held-constant factors:** are variables that may, exert some effect on the response, but for the purposes of the present experiment these factors **are not of interest**, so they will be held at a specific level.
- **Allowed-to-vary factors:** the **experimental units** or the "**materials**" to which the design factors are applied are **usually non-homogeneous**, yet we often ignore this unit-to-unit variability and rely on **randomization** to **balance out** any material or experimental unit effect.
- It is often assumed that the effects of **held-constant factors** and **allowed-to-vary factors** is relatively small

2 NUISANCE FACTORS

- These factors **may have large effects** that must be accounted for, yet **we may not be** interested in them in the context of the present experiment.
- Nuisance factors are often classified as:
 - ✓ controllable
 - ✓ uncontrollable and
 - ✓ noise factors.

- **A controllable nuisance factor:** is one whose levels may be set by the experimenter.
 - ❖ For example,
 - ✓ the experimenter can select different batches of raw material or
 - ✓ different days of the week when conducting the experiment.
- **The blocking principle** discussed in the previous section, is often useful in dealing with controllable nuisance factors

Uncontrollable nuisance

- If a nuisance factor is **uncontrollable** in the experiment, **but it can be measured**, an analysis procedure called the **analysis of covariance** can often be used to compensate for its effect.
 - ❖ *For example,*
 - ✓ *The relative humidity in the process environment may affect process performance, and if the humidity cannot be controlled, **it probably can be measured and treated as a covariate.***

noise factor

- When a factor **that varies naturally** and uncontrollably in the process can be controlled for purposes of an experiment, we often call it a **noise factor**.
- In such situations, our objective is usually to find the **settings of the controllable design factors** that **minimize** the variability transmitted from the noise factors.
- This is sometimes called
 - ✓ *a process robustness study or*
 - ✓ *a robust design problem*

- Levels and ranges of the experiment

3. Selection of the response variable.

- In selecting the response variable, the experimenter should be **certain that this variable really provides useful information** about the process under study.
- **Most often, the average or standard deviation (or both) of the measured characteristic will be the response variable.**
- *It is critically important to identify issues related to **defining the responses of interest and how they are to be measured** before conducting the experiment*

4. *Choice of experimental design*

- If the **pre-experimental planning** activities are done correctly, this step is relatively easy.
- **Choice of an experimental design involves:**
 - ✓ The consideration of **sample size** (number of replicates)
 - ✓ The selection of a **suitable run order** for the experimental trials, and
 - ✓ The determination of whether or not **blocking** or other **randomization restrictions** are involved.

- In selecting the experimental design it is important to keep the experimental objectives in mind as shown below (**two possible scenarios**):
- 1. In many engineering, science and technology experiments we may already know at the outset that **some of the factor levels will result in different values** for the response:
 - ✓ Consequently, *we are interested in identifying which factors cause this difference and in estimating the magnitude of the response change.*

- 2 In other situations, we may be more interested in verifying uniformity.
- ❖ For example, two production methods A and B may be compared, A being the standard and B being a more cost-effective alternative.
 - ✓ *Hence the experimenter will then be interested in demonstrating whether there is or there is no difference in the **quality of products** produced using the two methods.*

5. *Performing the experiment*

- When running the experiment, it is vital to monitor the process carefully to ensure that everything is being done according to **plan**.
- **Errors** in experimental procedure at this stage will usually **destroy experimental validity**.
- Up-front planning is crucial to success.
- *It is easy to underestimate the logistical and planning aspects of running a designed experiment in a complex research and development environment*

- Prior to conducting the experiment a few trial runs or pilot runs are often helpful (preliminary study or experiment).
- The trial or pilot runs (preliminary study or experiment) provide:
 - ✓ **Information** about consistency of **experimental material**
 - ✓ a check on the **measurement** system
 - ✓ a rough idea of experimental **error**
 - ✓ a chance to practice the overall experimental technique.
- *This also provides an opportunity to revisit the decisions made in steps 1-4, if necessary.*

6. Statistical analysis of the data:

- Statistical methods should be **used to analyze the data** so that results and conclusions are **objective** rather than **judgmental** in nature.
- If the experiment has been designed correctly and if it has been performed according to the design, the statistical methods required **are not complicated**.
- Remember that **statistical methods** cannot prove that a factor (or factors) has a particular effect.
- ✓ They only **provide guidelines** as to the reliability and validity of results.

- Properly applied statistical methods do not allow anything to be proved experimentally, but they do allow us to
 - ✓ measure the likely error in a conclusion or
 - ✓ to attach a level of confidence to a statement.
- The primary advantage of statistical methods is that
 - ✓ they add *objectivity* to the decision-making process.
- Statistical techniques coupled with
 - ✓ good engineering or process knowledge and
 - ✓ common sense **will usually lead to sound conclusions.**

7. Conclusions and recommendations

- Once the data have been analyzed, the experimenter must draw *practical conclusions* about the results and **recommend a course of action**.
- Graphical methods are often useful in this stage, particularly in presenting the results to others.
- **Follow-up runs and confirmation testing** should also be performed to validate the conclusions from the experiment.

- Throughout this entire process, it is important keep in mind that **experimentation is an important part of the learning process** where the experimenter:
 - ✓ tentatively formulate hypotheses about a system
 - ✓ perform experiments to investigate these hypotheses, and
 - ✓ on the basis of the results formulate new hypotheses, and so on.
- Hence experimentation is iterative.
- It is usually a major mistake to design a single, large, comprehensive experiment at the start of a study.

- **A successful experiment requires:**
 - ✓ **knowledge** of the important factors
 - ✓ the **ranges** over which these factors should be varied
 - ✓ the appropriate number of **levels** to use, and
 - ✓ the proper **units of measurement** for these variables.
- Generally, we do not perfectly know the answers to these questions, but we learn about them as we go along

- As an experimental program progresses
 - ✓ one often drops some input variables and add others
 - ✓ change the region of exploration for some factors, or
 - ✓ add new response variables.
- Consequently, it would be wise to usually experiment *sequentially*, and as a general rule, no more than about 25% of the available resources should be invested in the first experiment.
- This will ensure that sufficient resources are available to perform confirmation runs and ultimately accomplish the final objective of the experiment



THANK YOU

CHAPTER 6

SINGLE AND MULTI-FACTOR (FACTORIAL) EXPERIMENTS

Reading assignment

1. Normal distribution, z-test, t-distribution and t-test
2. Hypothesis testing, confidence interval

- Based on **nature of treatments** under investigation and **nature of comparison** required among them, experiments are classified as
 - ✓ *comparative* and *factorial* experiments.
- **Comparative experiments** –
 - ✓ There is only **one factor** whose effect is to be studied.
 - ✓ It is also called (One-at a time experiment)
- **Example:**
 - ✓ Comparison of drying methods;
 - ✓ The effect of chemical preservative dose,
 - ✓ Comparison of the performance of three machines etc.

- ***Factorial experiment :-***

- ✓ Combination of two or more levels of more than one factor are the treatments.

Example:

- ✓ The effect of blanching time and frying temperature on the crispiness of potato chips
- ✓ The effect of germination time and temperature on the quality of malt
- ✓ The effect of type of ingredients, blending ratio and processing method on the quality of extrudate

Frying temperature	Blanching time	
	time 1	time 2
T1	Comb 11	Comb 12
T2	Comb 21	Comb 22

6.1 Analysis of single factor experiments

- **6.2.1 Completely randomized design (CRD) and analysis**
- **One-way analysis of variance**
- The analysis of variance often referred to as **ANOVA** or AOV
 - ✓ is probably the most frequently used method for data analysis **for comparing means.**
- The method is used **to test for significant** differences in means of a variable across groups of observations.

Suppose we have t-treatments and n-number of replications,
the data summary looks like:

Treatment 1	Treatment 2	-	-	-	-	-	Treatment t
X_{11}	X_{21}	-	-	-	-	-	X_{t1}
X_{12}	X_{22}	-	-	-	-	-	X_{t2}
X_{13}	X_{23}	-	-	-	-	-	X_{t3}
-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-
X_{1n}	X_{2n}	-	-	-	-	-	X_{tn}
$\bar{X}_1$	$\bar{X}_2$	-	-	-	-	-	$\bar{X}_t$

$$\bar{X} = \frac{\sum_i^t X_i}{t} \text{ grand mean}$$

- Model for the data

- **1 Mean model**

$$X_{ij} = \mu_i + \varepsilon_{ij} ; i = 1, 2, \dots, t \text{ and } j = 1, 2, \dots, n$$

- **2 Effects model**

$$\mu_i = \mu + \tau_i$$

- The effects model will then be:

$$X_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

- where X_{ij} is the observed data, μ_i is the population mean, τ_i is the effect of the treatment i and ε_{ij} is the error term (the random error).

- Both models called **one-way analysis of variance**
 - ✓ **Null hypothesis (H_0):** $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \dots = \mu_t$
 - ✓ **Alternative hypothesis (H_A):** $\mu_i \neq \mu_j$ (at least for one pair)
- The analysis of variance of CRD is carried out by **partitioning** the total variation into components:
 - ✓ **treatment variation** and
 - ✓ **residual variation** \
 - (decomposition of the sum of squares)

The total variation (total sum of squares) is calculated as

$$SS_T = \sum_{i=1}^t \sum_{j=1}^n (X_{ij} - \bar{X})^2$$

$$\sum_{i=1}^t \sum_{j=1}^n (X_{ij} - \bar{X})^2 = \sum_{i=1}^t \sum_{j=1}^n ((X_{ij} - \bar{X}_i) + (\bar{X}_i - \bar{X}))^2$$

$$\sum_{i=1}^t \sum_{j=1}^n (X_{ij} - \bar{X})^2 = n \sum_{i=1}^t (\bar{X}_i - \bar{X})^2 + \sum_{i=1}^t \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2 + 2 \sum_{i=1}^t \sum_{j=1}^n (X_{ij} - \bar{X}_i) (\bar{X}_i - \bar{X})$$

$$\text{but } \sum_{i=1}^t \sum_{j=1}^n (X_{ij} - \bar{X}_i) = 0 \text{ (check)}$$

Therefore

$$SS_T = \sum_{i=1}^t \sum_{j=1}^n (X_{ij} - \bar{X})^2 = n \sum_{i=1}^t (\bar{X}_i - \bar{X})^2 + \sum_{i=1}^t \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2$$

$$SS_T = SS_{Tr} + SS_{Er}$$

$$\frac{\sum_{i=1}^t \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2}{N - t}$$

= pooled variance (common variance or the variance is equal)

$$\frac{n \sum_{i=1}^t (\bar{X}_i - \bar{X})^2}{t - 1}$$

= Variation of the treatment means from the grand mean (variance)

- But if the treatment means are equal, the variation of the treatment means from the grand mean to estimate the **common variance**.
- Therefore the **two variances are independent measures of the variability of the data** and are expected to be close to each other
- *If the variance among the treatments exceeds the variation within the groups, the treatment means are said to be statistically different*

Working formula to calculate sum of squares and mean squares

$$SS_T = SS_{Tr} + SS_{Er}$$

$$SS_T = SS_{Tr} + SS_{Er}$$

$$\text{Total } SS = \sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}^2 - C$$

$$\text{Where : } C = \frac{\left(\sum_{i=1}^t \sum_{j=1}^{n_i} x_{ij} \right)^2}{N}; \quad C = \text{correction term}$$

$$\textit{Treatment SS} = \sum_{i=1}^t \frac{\left(\sum_{j=1}^{n_i} x_{ij} \right)^2}{n_i} - C; \text{ where } \sum_{j=1}^{n_i} x_{ij} = \text{the sum of the } n_i \text{ data from group } i,$$

$$\textit{Error SS} = \sum_{i=1}^t \sum_{j=1}^{n_i} x_{ij}^2 - \sum_{i=1}^t \frac{\left(\sum_{j=1}^{n_i} x_{ij} \right)^2}{n_i}$$

$$SS_{ER} = SS_T - SS_{Tr}$$

$$treatment\ MS = \frac{treatment\ SS}{treatment\ DF} \quad error\ MS = \frac{error\ SS}{error\ DF} \quad total\ MS = \frac{total\ SS}{total\ DF}$$

$$F = \frac{treatment\ MS}{error\ MS}$$

The critical value for this test (obtained from F-table) is $F_{\alpha, (1), (t-1), (N-t)}$

(1), (t-1), (N-t)

t-1 = Df of nominator

N-t = Df of denominator

t = factor levels

N = number of observations

Decision:

- If the calculated F is at least as large as the critical value (obtained from F -table) then we reject the null hypothesis (H_0)
- **Note that the conclusions are reached based on only calculation of MS values.**
- The ANOVA does not identify which mean(s) is (are) **different** but at least one mean is different.
- The purpose of ANOVA is simply **to provide evidence** concerning the presence or absence of treatment effects.

Summary of ANOVA table

Source of variation	Degree of freedom	Sum of Squares (SS)	Mean squares (MS)	Variance ratio (F - value)
Total	N-1	Total SS	----	$F = \frac{\text{treatment MS}}{\text{error MS}}$
Among groups	t-1	treatment SS	$\text{treatment MS} = \frac{\text{treatment SS}}{\text{treatment DF}}$	
Within groups (error)	N-t (N = nt)	Error SS	$\text{error MS} = \frac{\text{error SS}}{\text{error DF}}$	

Example 1

- Four brands of dried meat were evaluated by eight judges for their firmness using the score sheet below. The anchor words were agreed upon by the judges during preliminary sessions. Numerical values were given to the ratings by measuring the distance of the judges' marks from the left end of the line in units of 0.1 0.25 cm. Carryout analysis of variance and draw conclusion whether there is a significant difference between the products.

QUESTIONNAIRE FOR DESCRIPTIVE ANALYSIS WITH SCALING

Please evaluate the firmness and chewiness of these samples of dried meat.

1 Firmness - make vertical lines on the horizontal line to indicate your rating of the firmness

of each sample. Label each vertical line with the code number of the sample it represents.

Please taste the samples in the following order

572 681 437 249

Very soft

very firm

Judges	Firmness scores				
	Brands				Total
	A	B	C	D	
1	1.9	2.1	3.4	1.1	8.5
2	4.1	2.4	5.1	3	14.6
3	4.6	3.2	5.3	1.8	14.9
4	2.3	5.3	5.3	1.3	14.2
5	1.5	4.7	5.3	0.6	12.1
6	2	3.3	4.7	1.8	11.8
7	2	3.5	5.2	2.7	13.4
8	2.1	4.8	5.4	1.7	14.0
Total	20.5	29.3	39.7	14.0	103.5
Mean	2.6	3.7	5.0	1.8	

Analysis

$$C = \frac{(\sum X)^2}{N} = \frac{103.5^2}{8 \times 4} = 334.6$$

$$SS_T = \sum X^2 - C = 1.9^2 + \dots + 1.7^2 - C = 72.59$$

Treatment (product) sum of squares

$$SS_{Tr} = \sum \frac{(T_i^2)}{b} - C = \frac{20.5^2 + 29.3^2 + 39.7^2 + 14}{8} - C = 46.6$$

Error sum of squares

$$\text{Error sum of squares} = SS_T - SS_{Tr} = 25.99$$

ANOVA Table

Source	SS	df	MS	F_{cal}	F_{crit}
Treat	46.60	3	15.53	16.73	2.95
Error	25.99	28	0.93		
Total	72.59	31			

Conclusion: There is significant difference

MEAN SEPARATION (MULTIPLE COMPARISON)

- In the example, we have calculated the F-value and reported that **there is a significant difference** between treatments variability.
- But ANOVA **doesn't indicate** which means are actually different.
- The process of finding which of the means from the array of means are different is known as **multiple comparisons (mean separation)**.
- One of the methods of multiple comparison (mean separation) is the **least significant difference (LSD) Test**
- **LSD** should be used only when the F test is significant .
- **LSD** test is derived from t-test and it is computed as
 - ✓ a product of standard error of the difference of the two means &
 - ✓ t at error degree of freedom and 5% significance level

$$LSD = SE_{\bar{X}_A - \bar{X}_B} \times t_{tab} @ Err_{df} \times \alpha \% SL$$

The derivation is computed as

$$t = \frac{d_i - \mu}{SE_{\bar{d}}} = \frac{LSD}{SE_{\bar{d}}} = \frac{LSD}{\sqrt{2 S_p^2 / n}}$$

The smallest difference which gives rise to a significant difference is the value, which gives the t-tabulated value at the given degree of freedom

$$\text{Hence } LSD = t_{tab} \cdot \sqrt{\frac{2MS_E}{n}}$$

$$LSD = t_{tab} \cdot \sqrt{\frac{2S_p^2}{n}} = 2.048 \times \sqrt{\frac{2 \times 0.93}{8}} = 0.988$$

	1	
Means	8	
2.6a	a	3.7b 5.0c

(means followed by different letters are significantly different)

6.3.2 Randomized Complete Block Design (RCBD) and analysis

- **RCBD**
 - ✓ Is one of the most widely used experimental design in Agri
 - ✓ Is especially suited for field experiments where the number of treatments are small
- *The objective* of using RCBD is
 - ✓ to keep the variability among experimental units within a block as small as possible and
 - ✓ to maximize differences among blocks.
- This is because the greater is the sum of squares attributable to known causes and the smaller is the error sum of squares.

- **Suppose** a study is conducted to study the effect of barrel temperature (T) on bulk density of an extruded product
- If three different brands of extruders are used which are expected to have different extrusion performance, the extruders could be used as blocks to control variation due to extruders
- *Definition* - When equal numbers of **measurements** are made on each treatment in each block and if the order of tests **within a block is randomized**, the experiment is called ***Randomized Complete Block Experiment***

Extruder type (Blocks)		
A	B	C
T_1	T_3	T_2
T_2	T_1	T_3
T_3	T_2	T_1

T_1, T_2 and T_3 are the temperature levels used

Analysis of a RCBD

The statistical model is

$$X_{ij} = \mu + \alpha_i + \beta_j + \xi_{ij}$$

where μ is the population mean, α_i is treatment effect, β_j block effect and ξ_{ij} is random error

TREATMENTS	BLOCKS				Treatment totals
	I	II	III	IV	
Trt-1	X_{11}	X_{12}	X_{13}	X_{14}	-----
Trt-2	X_{21}				-----
Trt-3	X_{31}				-----
Trt-4	X_{41}			X_{44}	-----
Block totals	-----	-----	-----	-----	Total sum

The analysis of variance is carried out the same way as CRD analysis **by partitioning** the total variation into components, but with one more sum of squares (one more source of variation).

Total variation	=	Variation between blocks	+	Variation between treatments	+	Error variation
------------------------	----------	---------------------------------	----------	-------------------------------------	----------	------------------------

Suppose there are t -treatments and b - blocks, the computation of the sum of squares will be:

$$SS_T = \sum X_{ij}^2 - \frac{\left(\sum X_{ij}\right)^2}{N} \quad \text{at } (bt - 1) \text{ df}$$

$$SS_{Tr} = \sum_{i=1}^t \frac{\left(\sum_{j=1}^b X_j\right)^2}{n_b} - CF \quad \text{at } (t - 1) \text{ df}$$

$$SS_{block} = \sum_{j=1}^b \frac{\left(\sum_{i=1}^t X_i\right)^2}{n_t} - CF \quad \text{at } (b-1) \text{ df}$$

$$SS_{Er} = SS_T - SS_{Tr} - SS_{block} \quad \text{where } df_{Er} = df_T - df_{Tr} - df_{block}$$

ANOVA TABLE FOR RCBD

Source of Variation	SS	Df	MS	F_{calc}	F_{tab}
Blocks	SS_B	$b-1$	$MS_B = \frac{SS_B}{b-1}$	$\frac{MS_B}{MS_{Er}}$	
Treatment	SS_{Tr}	$t-1$	$MS_{Tr} = \frac{SS_{Tr}}{t-1}$	$\frac{MS_{Tr}}{MS_{Er}}$	
Error	SS_{Er}	$(b-1)(t-1)$	$MS_{Er} = \frac{SS_{Er}}{(b-1)(t-1)}$		
Total	SS_T	$bt-1$			

Example 2

- The effect of barrel temperature on the bulk density of extruded breakfast cereal was studied. Three barrel temperatures (T_1 , T_2 , and T_3) were used. To shorten the experimental period, four extruders of different brand were used. Because of the expected difference in the performance of the extruders, the **experimenter** wanted to control their effect. The bulk density (kg/m^3) of the products are shown in the table below.
- Carry out the ANOVA and draw your conclusion???

Treatment	Extruders				Treatment Total
	I	II	III	IV	
Tr-1	330	288	295	313	1226
Tr-2	372	340	343	341	1396
Tr-3	359	337	373	302	1371
Block Total	1061	965	1011	956	3993

Computing the Components of the Variations

$$\text{Correction factor CF} = \frac{(\sum X)^2}{bn} = \frac{3993^2}{3 \times 4} = 1328671$$

Sum of squares for totals

$$SS_T = \sum X^2 - CF = 330^2 + \dots + 302^2 - CF = 8864$$

Sum of squares for treatments

$$SS_{Tr} = \sum \frac{(T_i^2)}{b} - CF = \frac{1226^2 + 1396^2 + 1371^2}{4} - CF = 4212$$

Block sum of squares

$$SS_B = \sum \frac{B_i^2}{t} - CF = \frac{1061^2 + 965^2 + 1011^2 - 956^2}{3} - CF = 2330$$

$$\text{Error sum of squares} = SS_T - SS_{Tr} - SS_B = 8864 - 4212 - 2330 = 2322$$

Source	SS	df	MS	VR (F_{calc})	F_{tab}	
					5%	1%
Blocks	2330	3	776.7	2.01	4.76	9.78
Treatments	4212	2	210.6	5.44*	5.14	10.92
Error	2322	6	387			
Total	8864	11				

The following information is inferred from the ANOVA table:

1.No significant difference in the variability among the four blocks. F_{cal} is less than F_{tab} and hence the experiment could be analyzed as CRD

2.Treatment means – bulk density

Tr-1	306.5
Tr-2	349.0
Tr-3	343.0

The standard error for comparing the **difference of two means**:-

$$SE_{\bar{X}_A - \bar{X}_B} = \sqrt{2 \times \frac{S_p^2}{n}} = \sqrt{2 \times \frac{387}{4}} = 13.9$$

To compare which of the means are significantly different, compute the LSD and compare after arranging the arrays of means

$$LSD = SE_{\bar{d}} \times t \quad \text{at Error degree of freedom} = 13.9 \times 2.45 = 34$$

Arranging the means in the order

Tr-1	Tr-3	Tr-2
306.5	343	349

Treatment 1 shows a significantly different bulk density compared to the others. But there was **no significant difference between treatments 2 and 3**

REPEATED MEASURES EXPERIMENTAL DESIGN VS RCBD

- **Repeated experimental designs** are commonly used in experiments involving **people** (e.g. sensory evaluation of foods)
- Because of differences in background, training, experience etc., differences in the responses of different people for same products **may be large**.
- Unless these variabilities are controlled they become part of the experimental error
- It would inflate the error sum of squares making it difficult to detect real differences between treatments
- **Remedy:** each of the treatments should be used **on each person (subject)**
- Such a design is called ***Repeated Measures Design*** (Treatment-by-panelist design)

The **treatment by panelists design** in sensory evaluation is preferred because it:

- (1) provides a more effective use of subjects, but also more importantly,
- (2) it results in a separate panelist term that can be partitioned from the error
- **Data Structure**

Treatment	Subjects				Treatment totals
	1	2		N	
1	Y_{11}	Y_{12}		Y_{1n}	
2	Y_{21}	Y_{22}		Y_{2n}	
.					
.					
t	Y_{t1}			Y_{tn}	
Subject totals	----	----		---	Totals

The **linear model** for repeated measures design

$$Y_{ij} = \mu + T_i + P_j + \varepsilon_{ij}$$

where T_i the effect of the i^{th} treatment, P_j a parameter associated with the j^{th} subject

$$SS_T = SS_{Tr} + SS_{subject} + SS_E$$

$$df_T = tn - 1; \quad df_{Tr} = t - 1; \quad df_{sub} = n - 1; \quad df_E = df_T - df_{Tr} - df_{sub}$$

Example 3

An experiment was conducted to investigate the effect of changes in ingredients on a specific sensory attribute of a product. Accordingly three products (Product A, B and C are evaluated and the scores obtained are shown in the table below.

Drive a conclusion whether the products are significantly different or the differences stem from chance?

Brands (A, B and C)					
Judges	A	Judges	B	Judges	C
1	6.6	11	8.0	21	9.0
2	6.0	12	7.0	22	8.0
3	7.0	13	10.0	23	12.0
4	5.0	14	5.0	24	5.0
5	6.0	15	5.0	25	7.0
6	5.0	16	6.0	26	9.0
7	7.0	17	7.0	27	8.0
8	4.0	18	6.0	28	8.0
9	7.0	19	6.0	29	5.0
10	1.0	20	2.0	30	3.0
Total	54.0	62.0		74.0	
Mean	5.4	6.2		7.4	

Analysis

The analysis starts from computation of sum of squares

Total sum of squares

$$\text{Total SS } (SS_T) = \sum_{i=1}^t \sum_{j=1}^{n_i} x_{ij}^2 - C$$

$$\text{Where: } C = \frac{\left(\sum_{i=1}^t \sum_{j=1}^n x_{ij} \right)^2}{N}; \quad C = \text{correction term}$$

$$C = \frac{(\sum X)^2}{N} = \frac{190^2}{10 \times 3} = 1203.33$$

$$SS_T = \sum X^2 - CF = 6.6^2 + \dots + 3.0^2 - C = 148.47$$

Treatment (product) sum of squares

$$SS_{Tr} = \sum \frac{(T_i^2)}{b} - C = \frac{54^2 + 62^2 + 74^2}{10} - C = 20.27$$

Error sum of squares

$$\text{Error sum of squares} = SS_T - SS_{Tr} = 148.47 - 20.27 = 128.40$$

ANOVA Table

Source of var	SS	df	MS	F_{cal}	F_{crit}
Treatment	20.27	2	10.13	2.13	3.35
Error	128.40	27	4.76		
Total	148.67	29			

Conclusion: No significant difference

Example 4

In example 3 if only 10 judges were used to evaluate the samples. The experimental design will be considered as treatment by panelist design where the effect of panelists is considered and the residual sum of squares will be portioned into error and panelists. Carry out ANOVA and draw conclusions.

Judges	Brands			Total	Mean
	A	B	C		
1	6.0	8.0	9.0	23.0	7.7
2	6.0	7.0	8.0	21.0	7.0
3	7.0	10.0	12.0	29.0	9.7
4	5.0	5.0	5.0	15.0	5.0
5	6.0	5.0	7.0	18.0	6.0
6	5.0	6.0	9.0	20.0	6.7
7	7.0	7.0	8.0	22.0	7.3
8	4.0	6.0	8.0	18.0	6.0
9	7.0	6.0	5.0	18.0	6.0
10	1.0	2.0	3.0	6.0	2.0
Total	54.0	62.0	74.0	190.0	
Mean	5.4	6.2	7.4		

Analysis

$$\text{Total } SS (SS_T) = \sum_{i=1}^t \sum_{j=1}^n x_{ij}^2 - C$$

$$\text{Where: } C = \frac{\left(\sum_{i=1}^t \sum_{j=1}^n x_{ij} \right)^2}{N}; \quad C = \text{correction term}$$

$$C = \frac{(\sum X)^2}{N} = \frac{190^2}{10 \times 3} = 1203.33$$

$$SS_T = \sum X^2 - CF = 6.6^2 + \dots + 3.0^2 - C = 148.47$$

Treatment (product sum of squares)

$$SS_{Tr} = \sum \frac{(T_i^2)}{b} - C = \frac{54^2 + 62^2 + 74^2}{10} - C = 20.27$$

Panelists sum of squares

$$SS_{panelists} = \sum \frac{(P_j^2)}{n} - C = \frac{23^2 + 21^2 + \dots + 6^2}{3} - C = 106.00$$

$$\text{Error sum of squares} = SS_T - SS_{Tr} - SS_{panelists} = 148.47 - 20.27 - 106 = 22.4$$

ANOVA table

Source of var.	SS	Df	MS	F _{calc}	F _{crit}
Treatment	20.27	2	10.13	8.14*	3.55
panelists	106.00	9	11.78	9.46*	2.46
Error	22.40	18	1.24		
Total	148.47	29			

* sig. at p<0.05

1. Conclusion????? Compare the result with example 1 !!!

2. Carry out mean separation and find out which means are different

Missing Observation in Randomized Complete Block Design

- Unlike CRD, RCBD does not deal simply when the result for some of units are missing..
- The absence of one or more observations changes the form of the design since the balanced nature of the relationship between treatments and blocks is destroyed.
- Although there are various statistical methods of dealing with this problem, the commonly used practice is to approximate the missing value and carry on the ANOVA as normal.
- The simplest situation is when one observation is missing.
- Say the observation which receives treatment one in block one is missing, X_{11} .
- Assume also that the RCBD has b blocks and t treatments.

- Then to calculate the estimated value for this unit we require the total of non missing observation in block 1, , the missing observation for treatments, , and grand total of all recorded observations, .

-

- The estimated value for the missing total is then

$$X_{11} = \frac{bB_1 + tT_1 - G_1}{(b - 1)(t - 1)}$$

Where X_{11} is the estimate of the missing observation,

Using this estimated value we complete the ANOVA, except that because our estimate is derived from the remaining values, the degree of freedom for both total variation and error variation is one less than usual.

In calculating the standard errors of treatments means we must obviously take account of the fact that the mean for the treatment with a missing unit is based only (b-1) actual observations. This leads to a revised standard error of a treatment means, when the treatment has a missing unit.

$$SE_{\bar{x}} = \sqrt{\left[\frac{S_p^2}{b} \left(1 + \frac{t}{(b-1)(t-1)} \right) \right]}$$

and the corresponding standard error of the difference between this treatment mean and say any other treatment mean

$$SE_{\bar{d}} = \sqrt{\frac{S_p^2}{b} \left(2 + \frac{t}{(b-1)(t-1)} \right)}$$

The standard errors are, of course, based on the degree of freedom of

$$S_p^2, \quad \text{i.e., } (b-1)(t-1)-1.$$

The standard errors of treatment means and of differences between treatment means not involving the missing unit are in the usual form, i.e.

$$\sqrt{S_p^2/b} \text{ and } \sqrt{2 S_p^2/b}$$

Example.5

Suppose that in the extruded breakfast cereals experiment (example 2), the record of the bulk density for second treatment in the third block has been missed, the record then looks:

Treatmen ts	Mechanical extractors				
	I	II	III	IV	Total
Tr-1	330	288	295	313	1226
Tr-2	372	340	-	341	[1053]
Tr-3	359	337	373	302	1371
Total	1061	965	[668]	956	3650

B'_3 - Block total for recorded count in block 3 is 668

T'_2 - Treatment total for recorded yield for treatment 2 is 1053

G' - Overall total of records = 3650

The estimate of the missing observation is, therefore,

$$X_{23} = \frac{4 \times 668 + 3 \times 1053 - 3650}{3 \times 2} = 363.5$$

The revised treatment, block and grand totals including this estimate are:

T_{r-2} - 1416.5, B_3 (III) - 1031.5, and the G - 4013.5.

And the analysis of variance can now be calculated to give the following:

Source of variation	Sum of squares	Degree of freedom	Mean Square	Variance Ration	
				F_{cal}	F_{tab}
Blocks	2609	3	869.7	2.06	
Treatment	4948	2	2474	5.86	
Error	2112	5	422.4		
Total	9669	10			

Treatment means

$$T_{r-1} = 306$$

$$T_{r-2} = 354$$

$$T_{r-2} = 343$$

$$\text{Standard error for comparing } T_1 \text{ and } T_3 = \sqrt{\frac{2 \times S_p^2}{4}} = 14.5$$

$$\text{Standard error for comparing } T_2 \text{ with } T_1 \text{ or } T_3 = \sqrt{\frac{S_p^2}{4} \left(2 + \frac{3}{3 \times 2} \right)} = 16.2$$

Both standard errors being based on 5 degree of freedom

6.4 ANALYSIS OF FACTORIAL EXPERIMENTS

- Researchers often wish to study the *combined effects of two or more factors* on a response variable (quality of food products) and the *interaction* of the factors. For example:
 - ✓ the combined effects and the interaction of two or more sweeteners on the degree of sweetness
 - ✓ the combined effect of panelists and ingredients and panelist-ingredient interaction etc.
- A simple sample problem
- two sweeteners, sucrose and high-fructose corn syrup (HFCS), being blended in a food (say, a breakfast cereal).

- Two levels of each sweetener, in a *factorial design*. This means that every level of one factor is combined with every level of the other factor (factors are crossed).
- One would like to know whether:
- the levels of sucrose had any effect => Main effect sucrose
- the levels of HCFS had any effect => Main effect HCFS
- the two sweeteners in combination produced any result that would not be predicted from the average response to each sweetener (if there is interaction between the two factors)

Data set for a two-factor experiment (the effect of panelists not considered)

Factor B HFCS	Factor A: levels of sucrose	
	Level 1(2%)	Level 2 (4%)
Level 1 (2%)	A_1B_1	A_2B_1
Level 2 (6%)	A_1B_2	A_2B_2

$$a = 2, b=2, n = 4$$

The concept of interaction

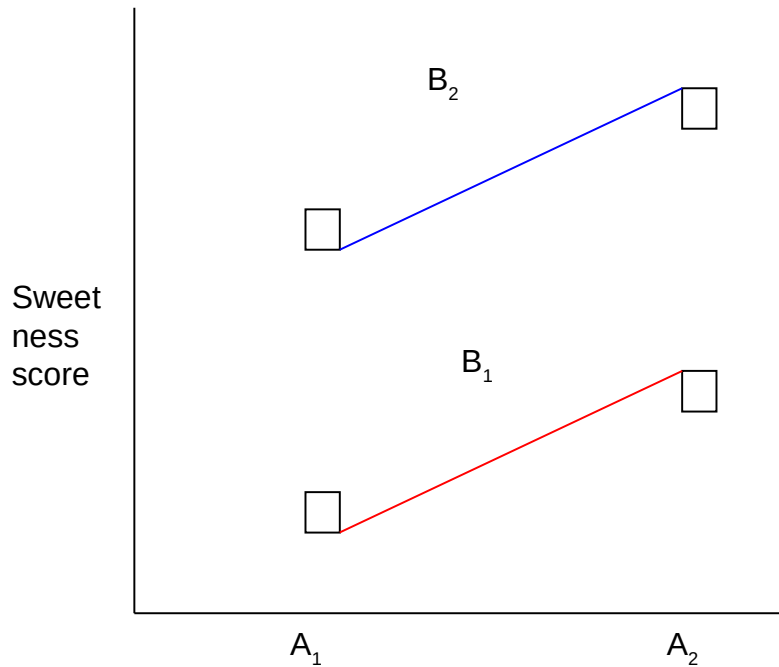
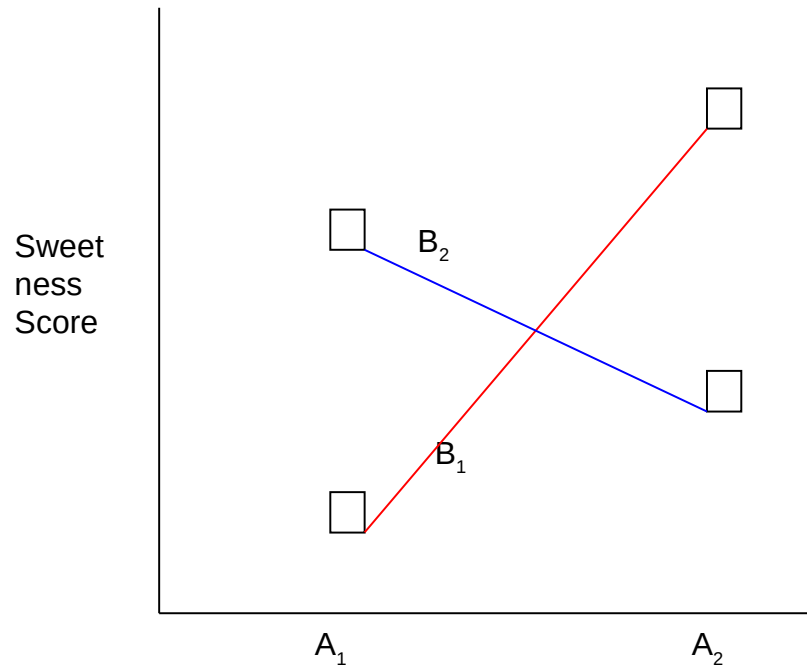
- Suppose that in the above experiment we got the following information the numbers indicating the sweetness score

B_2	15	A_2B_2
B_1	10	12
	A_1	A_2

If the average observation on A_2B_2 is 17, then the effects of the two factors are said to be additive (**not interacting**). i.e., the difference between the average observations at (A_2) and (A_1) is the same for both values of B and vice versa.

But if the average observation at A_2B_2 is any thing but 17, then the two factors are said to **interact**.

The following are two possible combinations:



The results of the factorial experiment can be combined to estimate the effects of the individual factors also called *the main effects* and *the interaction between treatments*.

Data set for a two-factor experiment

Factor B	Factor A (different ingredient composition)			
	Level 1	Level 2	..	Level a
Level 1	A_1B_1	A_2B_1	..	A_aB_1
Level 2	A_1B_2	A_2B_2	..	A_aB_2
.	.	.	..	.
.	.	.	..	.
Level -b	A_1B_n	A_2B_b	..	A_aB_b

Note: Each cell value is average of three values in the cell (i.e .if $n = 3$),

$$A_1B_1 = \frac{X_{111} + X_{112} + X_{113}}{3}; \quad A_aB_b = \frac{X_{a,b,1} + X_{a,b,2} + X_{a,b,3}}{3}$$

Factorial experiments have the following two advantages over “one at a time” (studying the effect of one factor at a time) experiments:

1. Factorial experiments provide an opportunity for studying not only the **individual effects** of each factor but also their **interaction**.

❖ *The most serious defect of one at a time experiments is that they are unable to detect interactions between factors, whereas complete factorial experiments can*

2. Factorial experiments have the advantage of **economizing** on **experimental resources**.

❖ *When experiments are conducted factor by factor much more resources are required for the same precision than when factorial experiments are conducted factor by factor*

Disadvantages of Factorial experiments

- 1.the large and prohibitive number of treatment combinations that may be required to study many factors at several levels.
- 2.the decrease in efficiency as a result of the increased size of replicate which is associated with large number of treatment combinations,

TWO-FACTOR ANALYSIS OF VARIANCE

- The linear statistical model assuming for a two-factor experiment looks like
- The ANOVA will test hypotheses from a general linear model. This model states that any score in the data set is determined by a number of factors:

$$X_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

where

X_{ijk} is observed data

μ is the population mean

α_i is the effect of i^{th} level of factor A

β_j is the effect of j^{th} level of factor B

$(\alpha\beta)_{ij}$ the interaction between A_i and B_j

ε_{ijk} is the random error

k represents the replicate of the combination of level i of A and level j of B

An observation = grand mean + factor 1 effect + factor 2 effect + interaction effect + error

Computation of sum of squares

1. Correction factor,

$$CF = \frac{\left(\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n X_{ijk} \right)^2}{abn} = \frac{(\sum X)^2}{N}$$

2. Total sum of squares

$$SS_T = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n X_{ijk}^2 - CF = \sum X^2 - CF$$

3. Treatment sum of squares (cells sum of squares)

$$SS_{Tr} = \sum_{i=1}^a \sum_{j=1}^b \frac{\left(\sum_{k=1}^n X_{ijk} \right)^2}{n} - CF = \frac{T_{11}^2 + T_{12}^2 + \dots + T_{ab}^2}{n} - CF$$

4. Error variation obtained by subtraction

Error sum of squares = Total sum of squares – Treatment squares

5. Treatment variation can be separated into components

5.1 Main effect of factor A

$$SS_A = \sum_{i=1}^a \frac{\left(\sum_{j=1}^b \sum_{k=1}^n X_{ijk} \right)^2}{b \times n} - CF = \frac{T_{A1}^2 + T_{A2}^2 + \dots + T_{Aa}^2}{nb} - CF$$

5.2. Main effects of factor B,

$$SS_B = \sum_{j=1}^b \frac{\left(\sum_{i=1}^a \sum_{k=1}^n X_{ijk} \right)^2}{a \times n} - CF = \frac{T_{B1}^2 + T_{B2}^2 + \dots + T_{Bb}^2}{na} - CF$$

5.3. Interaction sum of squares,

$$SS_{A \times B} = SS_{Tr} - SS_A - SS_B$$

ANOVA Table

Sources	SS	Df	MS	(F _{cal})	F _{crit}
Treatments	SS _{Tr}	ab-1			
<i>Effects of A</i>	SS _A	a-1	MS _A =SS _A /df _A	MS _A / MS _E	
<i>Effects of B</i>	SS _B	b-1	MS _B =SS _B /df _B	MS _B / MS _E	
<i>Interaction A×B</i>	SS _{AxB}	(a-1)(b-1)	MS _{AxB} =SS _{AxB} /df _{AxB}	MS _{AxB} / MS _E	
Error	SS _T -SS _{Tr}		MS _E =SSE/df _E		
Total	SS _T	abn-1			

Example 7 (two- factor analysis CRD)

The effect of ingredient composition and processing method on the acceptability of a product was studied using eight trained judges for each ingredient composition. The acceptability scores are presented below.

Carryout ANOVA and draw your conclusion on whether the composition and the processing methods affect the acceptability of the products??

Replications	Ingredient 1			Ingredient 2		
	fried	baked	roasted	fried	baked	roasted
1	8	6	5	7	4	2
2	7	7	6	6	5	3
3	7	6	5	6	4	2
4	8	7	5	7	5	2
5	6	6	6	5	4	3
6	7	6	4	6	4	1
7	7	7	5	6	5	2
8	8	7	5	7	5	2
Sum	58	52	41	50	36	17
Mean	7.25	6.5	5.125	6.25	4.5	2.13

Analysis

$$CF = \frac{\left(\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n X_{ijk} \right)^2}{abn} = \frac{(\sum X)^2}{N} = \frac{254^2}{48} = 1344.083$$

$$SS_T = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n X_{ijk}^2 - CF = 1496 - 1344.08 = 151.92$$

$$SS_{Tr} = \sum_{i=1}^a \sum_{j=1}^b \frac{\left(\sum_{k=1}^n X_{ijk} \right)^2}{n} - CF = \frac{58^2 + 52^2 + 41^2 + 50^2 + 36^2 + 17^2}{8} - CF = 135.17$$

$$SS_E = SS_T - SS_{cells} = 151.92 - 135.17 = 16.75$$

Portioning Treatment (Cells) sum of squares

	Sums -Ingredient		Sums - Process method		
	Level 1	level2	Level 1	level2	Level 3
sum	151	103	108	88	58
n (obsrv.)	24	24	16	16	16

Main factor Ingredient- sum of squares

$$SS_A = \frac{151^2 + 103^2}{24} - CF = 48$$

Main factor Process methods- sum of squares

$$SS_B = \frac{108^2 + 88^2 + 58^2}{16} - CF = 79.17$$

Interaction sum of squares

$$SS_{A \times B} = SS_{cells} - SS_A - SS_B = 135.17 - 48 - 79.17 = 8.00$$

Source	SS	df	MS	F _{calc}	F _{crit}
Treat(Cell)	135.17	5	27.03	67.79	2.44
Main F1	48	1	48.00	120.36	4.07
Main F2	79.17	2	39.59	99.26	3.22
F1xF2	8.00	2	4.00	10.03	3.22
Error	16.75	42	0.40		
Total	151.92	47			

P>0.0

Conclusion???

There is significant effect



THANK YOU