

# Business Research Methods

## 1. Foundation of Research

Dr. Asita Ghewari

# Research

Research can be defined as the search for knowledge , or as any systematic investigation, with an open mind to;

- ◉ establish novel facts,
  - ◉ solve new or existing problems,
  - ◉ prove new ideas,
  - ◉ develop new theories,
- usually using a scientific method

# Research

- Research is the process of gathering the information needed to answer certain questions and thereby helping in solving problems faced by an individual, firm, organisation or society.

# Business Research Defined

Business research is defined as the organised, systematic, data based, critical, objective, scientific enquiry or investigation for aid in making business decisions.



# Why Business research?

- ◉ Solve problems
- ◉ Decision making tool
- ◉ Competition
- ◉ Risk
- ◉ Investment....



Dr. Asita Ghewari



Dr. Asita Ghewari



Dr. Asita Ghewari



Dr. Asita Ghewari





**:: SUMMER SPECIAL! ::** Order \$65 or more and use coupon code "**COLDSTONE**" to get a FREE Cold Stone® gift box!



**CHOOSE A COUNTRY**

United States



**STORE LOCATOR**

enter zip code



more options

**SEARCH JELLYBELLY.COM**



OUR CANDY

OUR SHOP



**LOGIN  
MY ACCOUNT  
CHECKOUT**

Cart contains 0 item(s) | Total: \$0.00

**BEAN  
BOOZLED**  
DARE TO COMPARE



## BeanBoozled™ Jelly Beans

### Are you ready to be BeanBoozled™?

Jelly Belly's wildest collection dares you to compare some of our tastiest, most popular flavors with our craziest ones. But here's the catch - you won't know which ones are which!

The black Licorice bean looks exactly like the Skunk Spray bean! Sweet, luscious Caramel Corn might also be Moldy Cheese. You may think you're tasting our world-famous Buttered Popcorn bean, but what you'll be biting into could actually be Rotten Egg. The only way to find out what beans you're getting is to eat them!

The Jelly Belly BeanBoozled™ collection currently contains 10 colors of beans, and 20 flavors (10 "real" ones and 10...well...different ones). You can make a fun, and somewhat hair-raising, game of it.

Dr. Asita Ghewari



# EV SALES ZOOM PAST 1 MILLION IN 2023

NITIN KUMAR

Electric vehicle (EV) sales in India topped the 1 million mark in less than nine months in 2023, a milestone that took an entire year in 2022.

According to data from the Ministry of Road Transport and Highways' Vahan Dashboard, 1,037,011 EVs were registered with regional transport offices till September 19, accounting for 6.4 per cent of the total automobile sales in the country this year.

The sales surge can be attributed to increased individual purchases as well as business-to-business (B2B) acquisitions by EV fleet operators, industry sources said. Additionally, the Faster Adoption and Manufacturing of (Hybrid &) Electric Vehicles (FAME-II) initiative and the expansion of charging infrastructure have played a significant role in boosting EV adoption, they said.

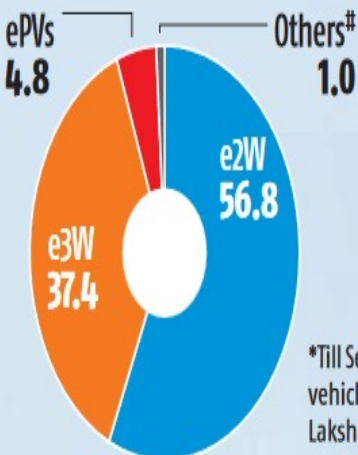
Turn to Page 3 ▶

## Record run

### Expanding EV presence

Year	Total vehicles sold	Total EVs sold	EV penetration (%)
2019	24,131,572	166,822	0.7
2020	18,613,324	124,643	0.7
2021	18,891,716	331,452	1.8
2022	21,556,313	1,024,778	4.8
2023*	16,081,652	1,037,011	6.4

### Segment-wise share in 2023 EV sales (in %)

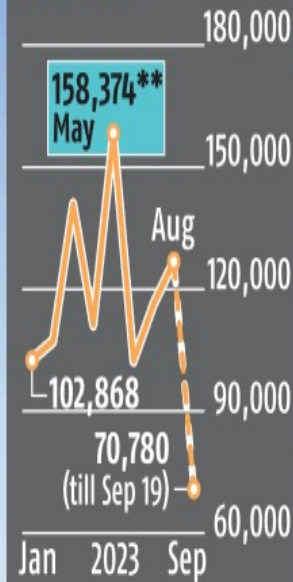


### Segment-wise penetration

Type of vehicles sold	Overall	EVs	2023* sales Share of total (%)
Two-wheeler	11,291,624	588,957	5.2
Three-wheeler	728,684	388,043	53.3
Passenger vehicle	2,491,950	49,640	2.0
Others#	1,569,394	10,371	0.7

\*Till Sept 19; #Goods vehicles, public service vehicles, and special category vehicles; Note: The numbers do not include data from Telangana and Lakshadweep Source: Vahan, Ministry of Road Transport and Highways

### Monthly EV sales



\*\*Highest ever monthly sales

Total  
1,037,011

Dr. Asita Ghewari



## EV sales...

Preetesh Singh, specialist CASE and alternate powertrains, NRI Consulting and Solutions, said the EV industry's growth also stemmed from the recent launch of high-quality vehicles, besides the presence of tax incentives, improved financing options, and a robust swapping facility. The two-wheeler segment accounted for 56 per cent of the EVs sold in 2023, followed by three-wheelers and passenger vehicles.

Throughout 2023, monthly EV sales have exceeded 100,000 units. The peak was reached in May, with a record sale of 158,374 vehicles.

While the impact of FAME-II subsidy cuts was most pronounced in June, resulting in a 35 per cent decline with sales dropping to 102,447 units, the industry began to recover in July. By August, with 126,741 units sold, sales had rebounded. The surge in August sales represented a 23 per cent increase from the June levels.

# Share of women in workforce: Going from bad to worse

ASHLI VARGHESE  
New Delhi, 19 September

Recent efforts to increase women's representation in politics come even as their participation in the economy retreats. The government on Tuesday introduced the Narishakti Vandan Adhiniyam to provide 33 per cent reservation to women in the Lok Sabha and state assemblies.

Women are less likely to work now than they were a few years ago, and continue to be less likely to own assets than men.

India has the lowest share of working women among the G20 countries. Its female labour force participation rate is 24 per cent, lower than the next-in-line

Saudi Arabia at 27.8 per cent. On the top is Australia (62.1 per cent), followed by China (61.1 per cent) and Canada (60.9 per cent). Other BRICS peers like

Brazil, South Africa, and Russia have over 50 per cent female labour force participation rate.

The female labour force participation rate is the proportion of the population aged 15 years and older that is working or seeking work.



## LAGGING BEHIND

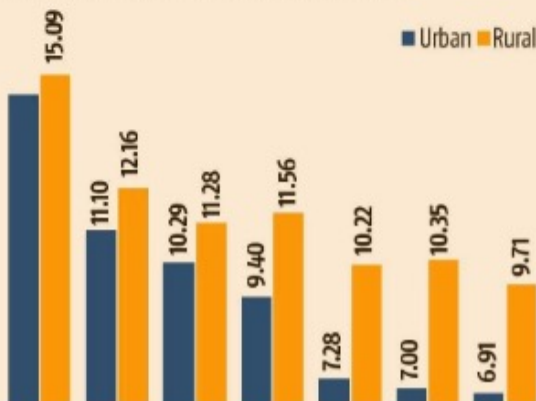
Lowest female labour force participation rate in G20  
Female labour force participation rate (in %)

Australia	62.1		Brazil	53.6	
China	61.1		Indonesia	52.7	
Canada	60.9		France	52.5	
UK	58.5		European Union	51.9	
US	56.5		Argentina	50.8	
Germany	56.4		South Africa	50.6	
Russia (2021)	55.3		Mexico	45.7	
Republic of Korea	54.6		Italy	41.0	
Japan	54.0		Türkiye	34.2	
			India	24.0	

Note: Data is as of 2022 for all the group of 20 (G20) countries, except Russia where it is for 2021. Female labour force participation rate is the proportion of the population ages 15 and older who are working or seeking work  
Source: World Bank

## WORSE IN URBAN AREAS

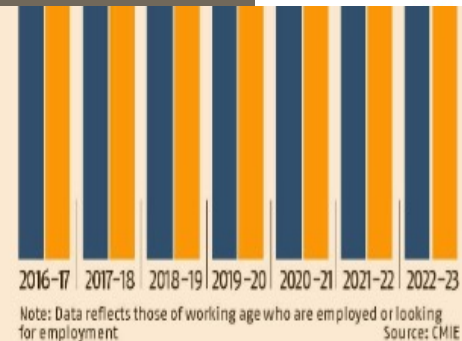
Female labour participation rate (in %)



The number of employed women has been largely in decline, falling by a third over the past seven years. In 2016-17, the labour pool was around 446 million, of which women constituted 68 million, making up just 15 per cent of total workforce. In FY23, of the 439 million total labour force, women were 45 million – around 10 per cent. There were around 22.7 million women workers in urban India in FY17; that number dedined to 12 million in FY23. The rural Indian female labour force dedined from 45 million to 33 million during the same period.

While female labour participation in both urban and rural India was around 15 per cent in FY17, their share in total workforce fell to 7 per cent for urban regions by FY23, and 9.7 per cent for rural areas. Women are also less likely to own assets like a house or land.

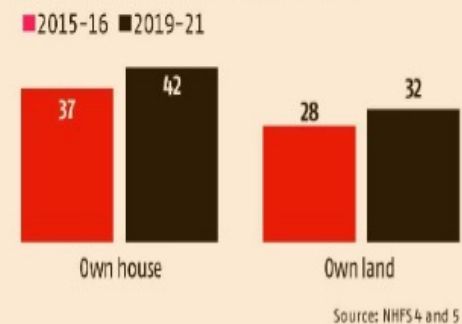
According to the National Family Health Survey (NFHS), while female land ownership increased from 37 per cent in 2015-16, it was still only 42 per cent in 2019-21. The house ownership has also increased from 28 per cent to 32 per cent.



Note: Data reflects those of working age who are employed or looking for employment  
Source: CMIE

## MAJORITY DON'T OWN ASSETS

Among women aged 15-49 (share, in %)



Source: NFHS 4 and 5

Dr. Asita Ghewari

# Business Research

- Research information is neither intuitive nor haphazardly gathered.
- Literally, research (re-search) -“search again”
- Business research must be objective
- Detached and impersonal rather than biased
- It facilitates the managerial decision process for all aspects of a business.

# Business Research Types

Basic research

Applied research



# Basic Research

- Attempts to expand the limits of knowledge.
- Not directly involved in the solution to a pragmatic problem.
- **Basic research** is experimental and theoretical work undertaken to acquire new knowledge without looking for long-term benefits other than the advancement of knowledge.

# Basic Research Example

- Factors involved in organization commitment
- Are members of highly cohesive work groups more satisfied than members of less cohesive work groups?
- Do consumers attract more by ambience?



# Applied Research

- Conducted when a decision must be made about a specific real-life problem
- **Applied research** is research undertaken to solve practical problems rather than to acquire knowledge for knowledge sake.

Applied research Example  
Kellogg's



Dr. Asita Ghewari

<b>Kind of Research</b>	<b>Key Characteristics</b>
Basic research	Focuses on generating fundamental knowledge
Applied research	Focuses on real-world questions and applications



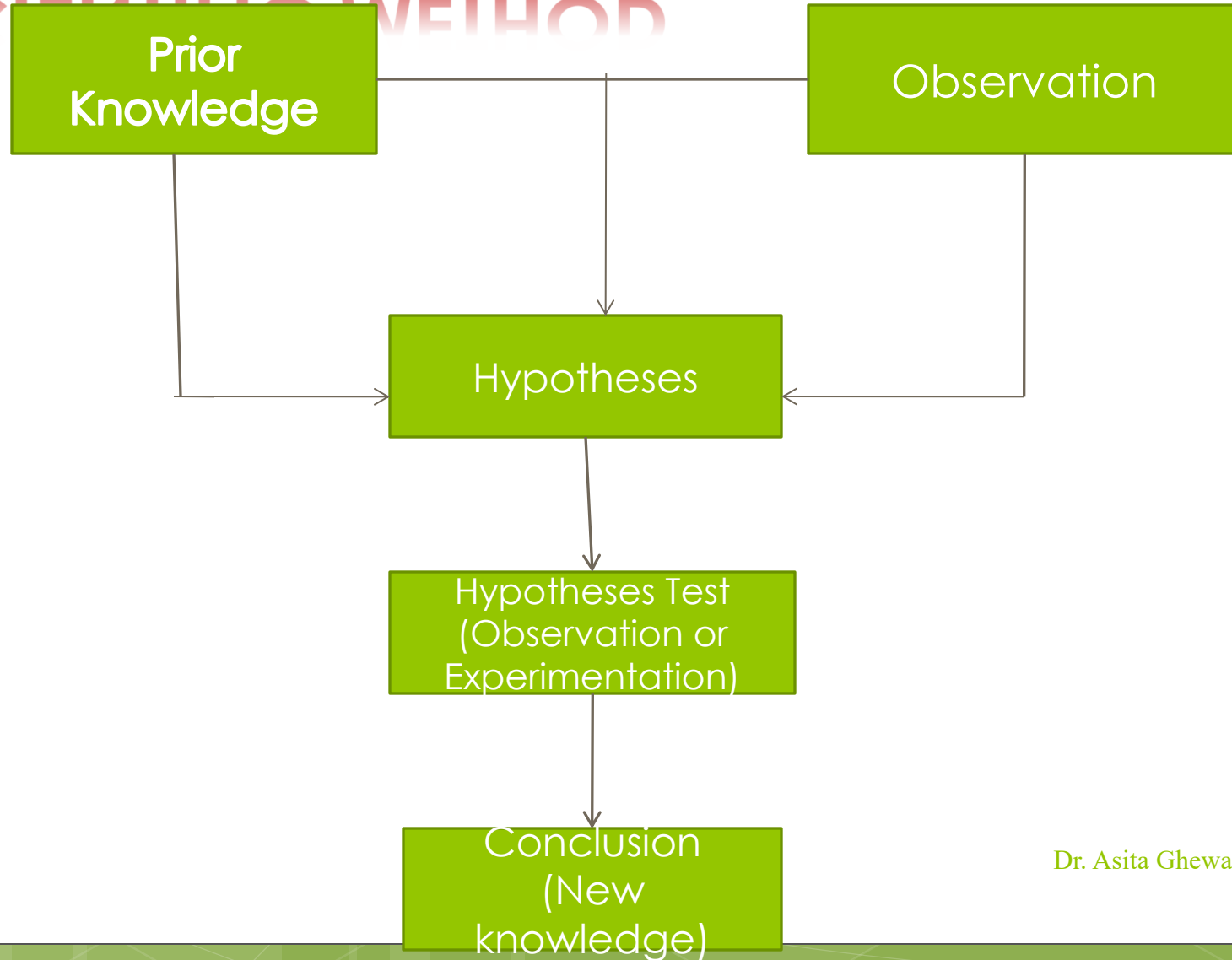
# Scientific Method

- The analysis and interpretation of empirical evidence (facts from observation or experimentation) to confirm or disprove prior conceptions.



Dr. Asita Ghewari

# SCIENTIFIC METHOD



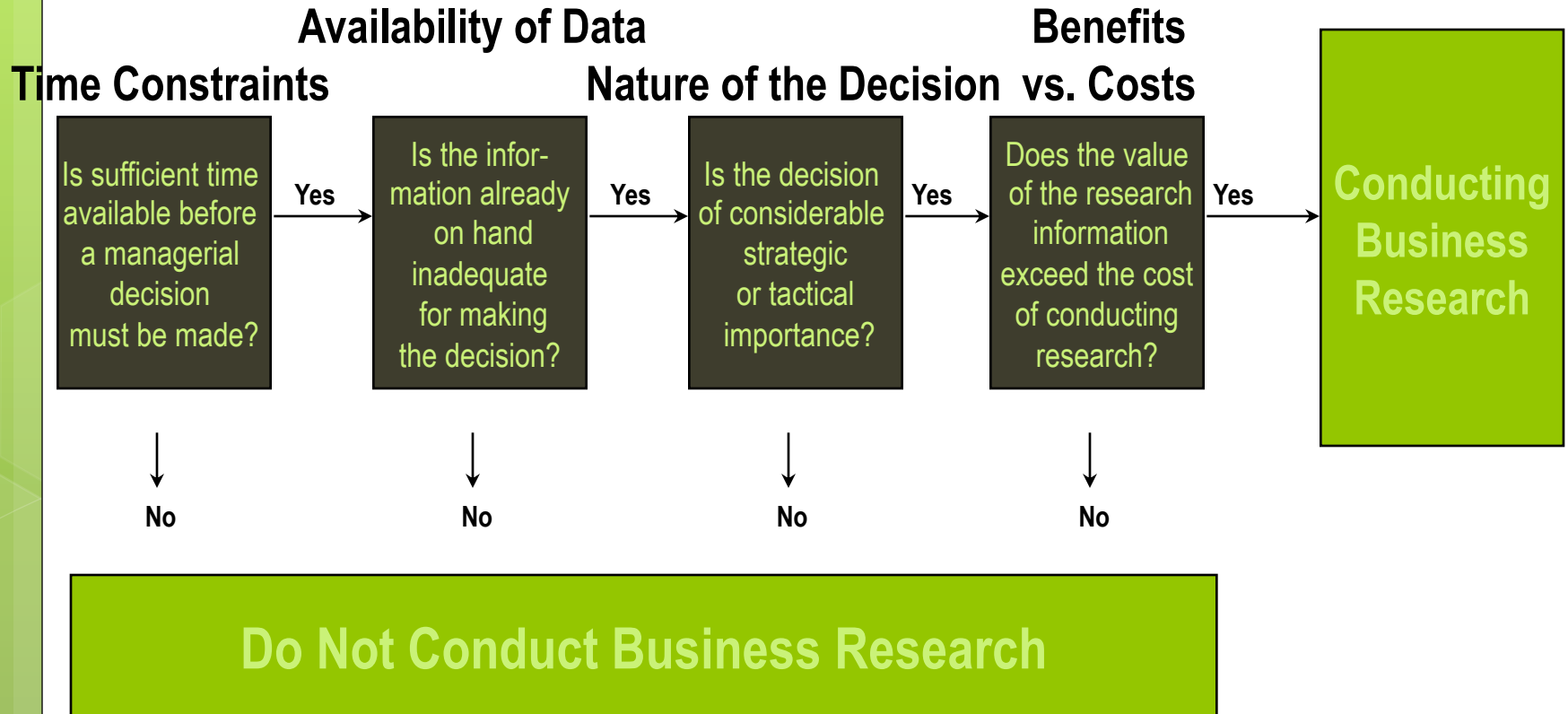
# Scientific Method

- Purposiveness
- Rigor
- Testability
- Replicability
- Objectivity
- Generalizability
- Parsimony
- Precision

# Determining When to Conduct Business Research

- Time constraints
- Availability of data
- Nature of the decision
- Benefits versus costs

# Determining When to Conduct Business Research



# Value versus Costs

- Potential Value of a Business Research Effort Should Exceed Its Estimated Costs

# Value and Costs of Undertaking Business Research

## VALUE

**Decreased  
Uncertainty**

**Higher Likelihood  
of Correct Decisions**

**Better Business  
performance**

**Higher Profits**

**Better Reputation**

## COSTS

**Research Costs**

**Delay in Making  
Business Decisions**

**Disclosure of  
Information to  
Rivals**

**Possibility of Error**

Dr. Asita Ghewari

# The Research Process: 11 Steps

- ◉ Step One: Establishing the Need for Research
- ◉ Step Two: Defining the Problem
- ◉ Step Three: Establishing Research Objectives
- ◉ Step Four: Literature Review
- ◉ Step Five: Determining Research Design
- ◉ Step Six: Identifying Information Types and Sources



# The Research Process: 11 Steps

- ◉ Step Seven : Decide the data collection Methods
- ◉ Step Eight: Determining Sample Plan and Size
- ◉ Step Nine: Collecting Data
- ◉ Step Ten: Analyzing Data
- ◉ Step Eleven: Preparing and Presenting the Final Report

Dr. Asita Ghewari

# Steps in the Research Process

## Steps

- 1 Establish Need
- 2 Define Problem
- 3 Establishing Research Objectives

**Figure out what to do in research**

- 4 Literature Review
- 5 Determine Research Design
- 6 Identify Information Sources
- 7 Decide Data Collection Method
- 8 Determine Sample Plan & Size

**Design the way to do the research**

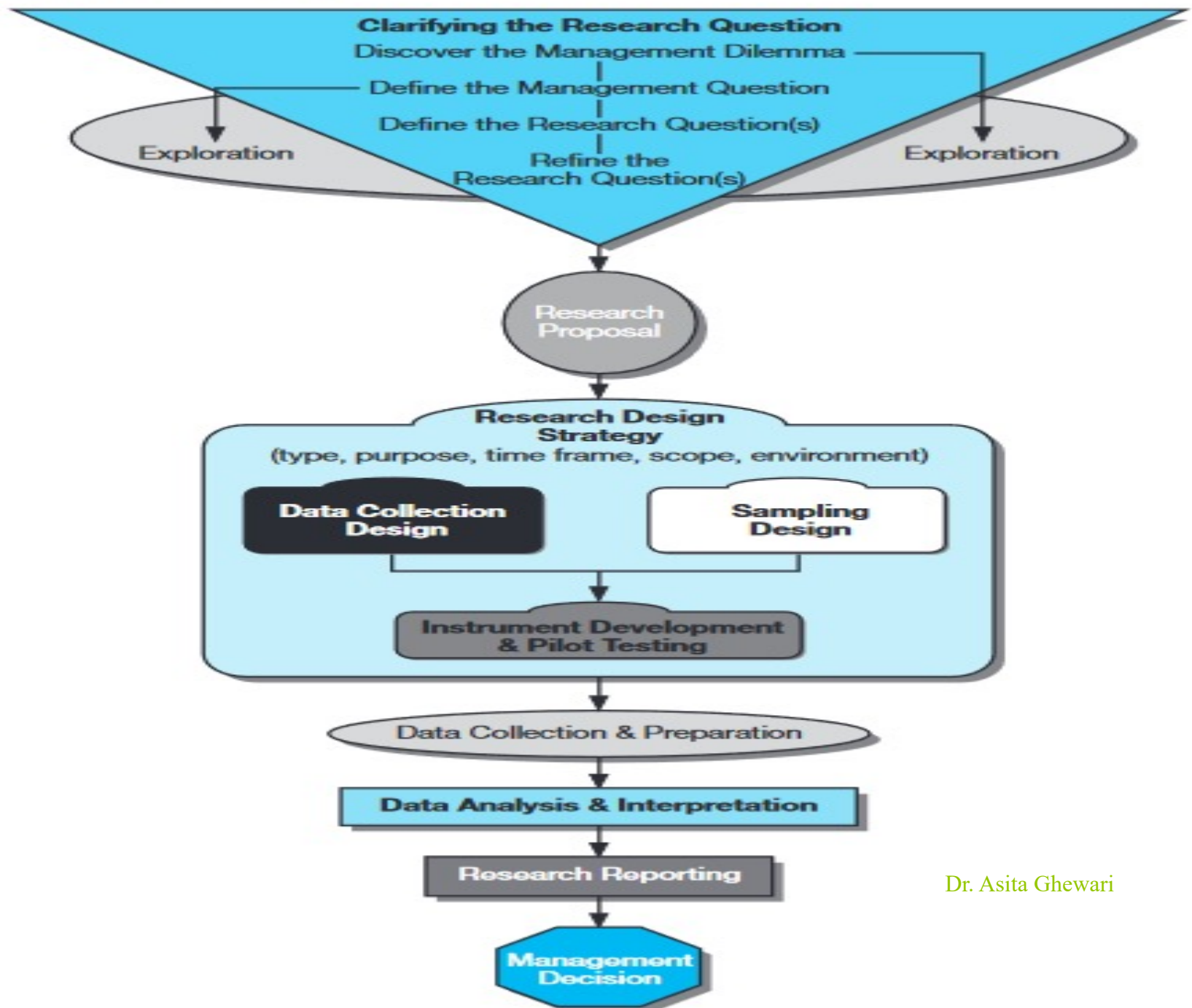
- 9 Collect Data

**Gather data from respondents**

- 10 Analyze Data
- 11 Write and Present Report

**Generate findings and interpret them**

Dr. Asita Ghewari



# Characteristics of Good Research

- ✖ Propose clearly defined
- ✖ Research process detailed
- ✖ Research design thoroughly planned
- ✖ Limitations frankly revealed
- ✖ High Ethical Standards Applied
- ✖ Adequate analysis for decision maker's needs
- ✖ Presented unambiguously
- ✖ Conclusions justified
- ✖ Researcher's experience reflected



# 81% intent on festival season e-shopping

**PEERZADA ABRAR**

Bengaluru, 25 September

About 81 per cent of consumers have indicated strong sentiment and intent to shop online during this festival season, according to a study by Nielsen Media India on behalf of Amazon India.

About 78 per cent trust online shopping and one in two intend to increase online spending, compared to the last season. Consumers expect a wide selection, unmatched value with competitive pricing, plus convenience of easy returns and exchange, which is driving the online shopping experience.

The study also revealed that for 68 per cent of consumers, Amazon.in is the go-to online shopping destination. About 75 per cent find the widest range and selection of products and brands

## MEESHO PROMISES 500K JOBS DURING UPCOMING SEASON

SoftBank-backed Meesho said it would enable nearly 500,000 seasonal job opportunities within its seller and logistics network to fulfill growing demand during the upcoming festival season. This is a 50 per cent increase compared to the seasonal jobs generated by the e-commerce firm last year.

Meesho aims to create roughly 200,000 job opportunities through its partnership with third-party logistics players such as Ecom Express, DTDC, Elastic Run, Loadshare, Delhivery, Shadowfax and Xpressbees. Over 60 per cent of the opportunities will be from Tier-III and -IV places. These roles will primarily encompass first-mile and delivery associates responsible for tasks such as delivery picking, sorting, loading, unloading and return inspections, Meesho said.

**PEERZADA ABRAR**

ence with options like faster delivery. More than 70 per cent wait for such events to purchase electronic items, with attractive offers (76 per cent) and new brand launches (75 per cent) among the core drivers. Over 75 per cent consumers are looking forward to attractive bank offers and no-cost equated monthly installments for shopping.

Three in four also wait for online festive shopping events to purchase large appliances, as they find access to new brands and product launches across categories like refrigerators, washing machines, geysers, and air conditioners.

Seventy-six per cent consumers wait for these dates to purchase smartphones — about 60 per cent are interested in the ₹10,000-20,000 price range, while two in three seek 5G features.

on the e-commerce platform.

The study said that 87 per cent of consumers in metros and 86 per cent in Tier-II cities (population of 10-40 lakh) prefer to shop

online during this festival period. Over 77 per cent of the respondents affirmed that festive day sales enhanced the convenience of their online shopping experi-

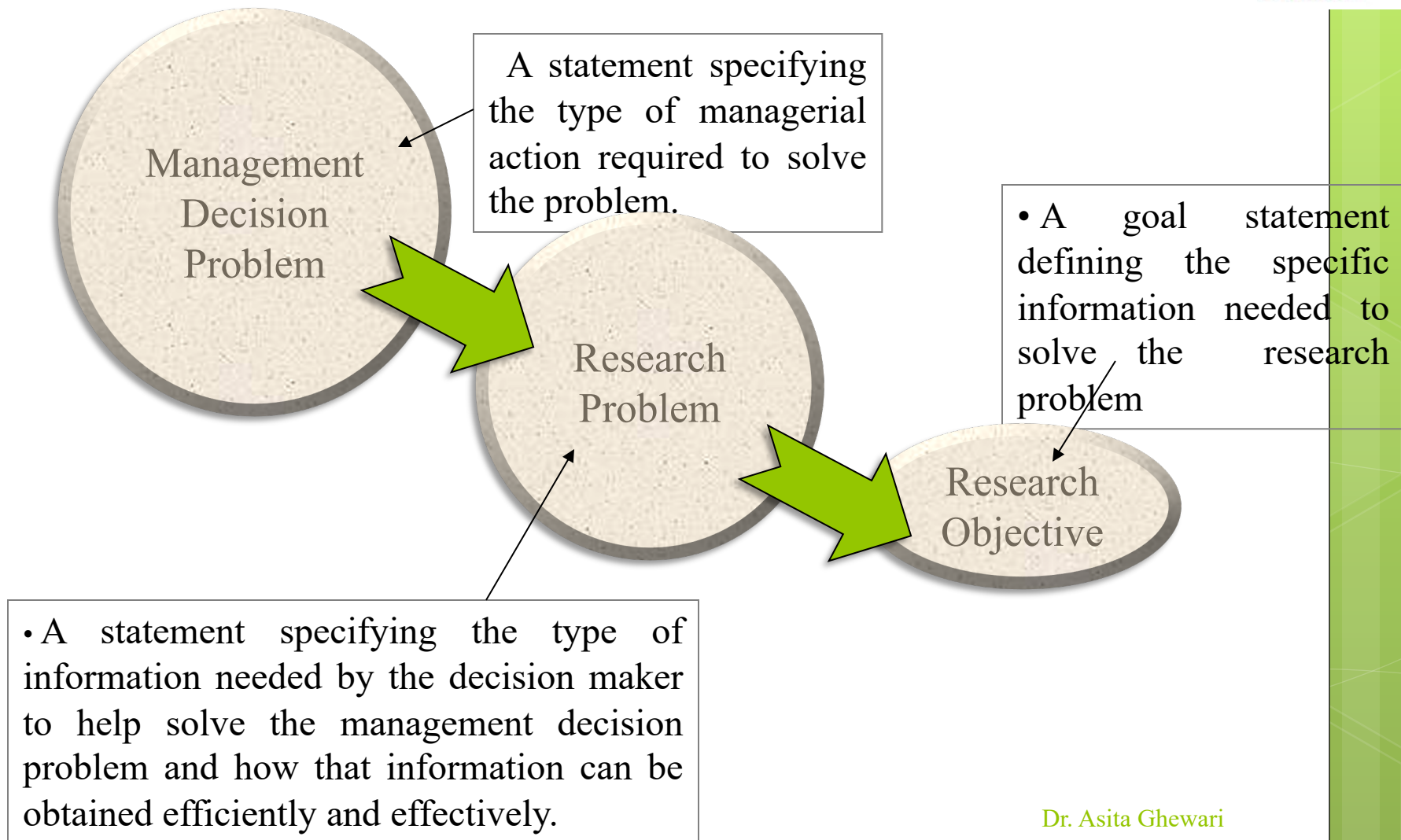
# Why define the Research Problem?

- Defining your destination before beginning a journey.
- It determines,
  - what you will do,
  - will it withstand scientific scrutiny,
  - how you will do it, and
  - what you may achieve!



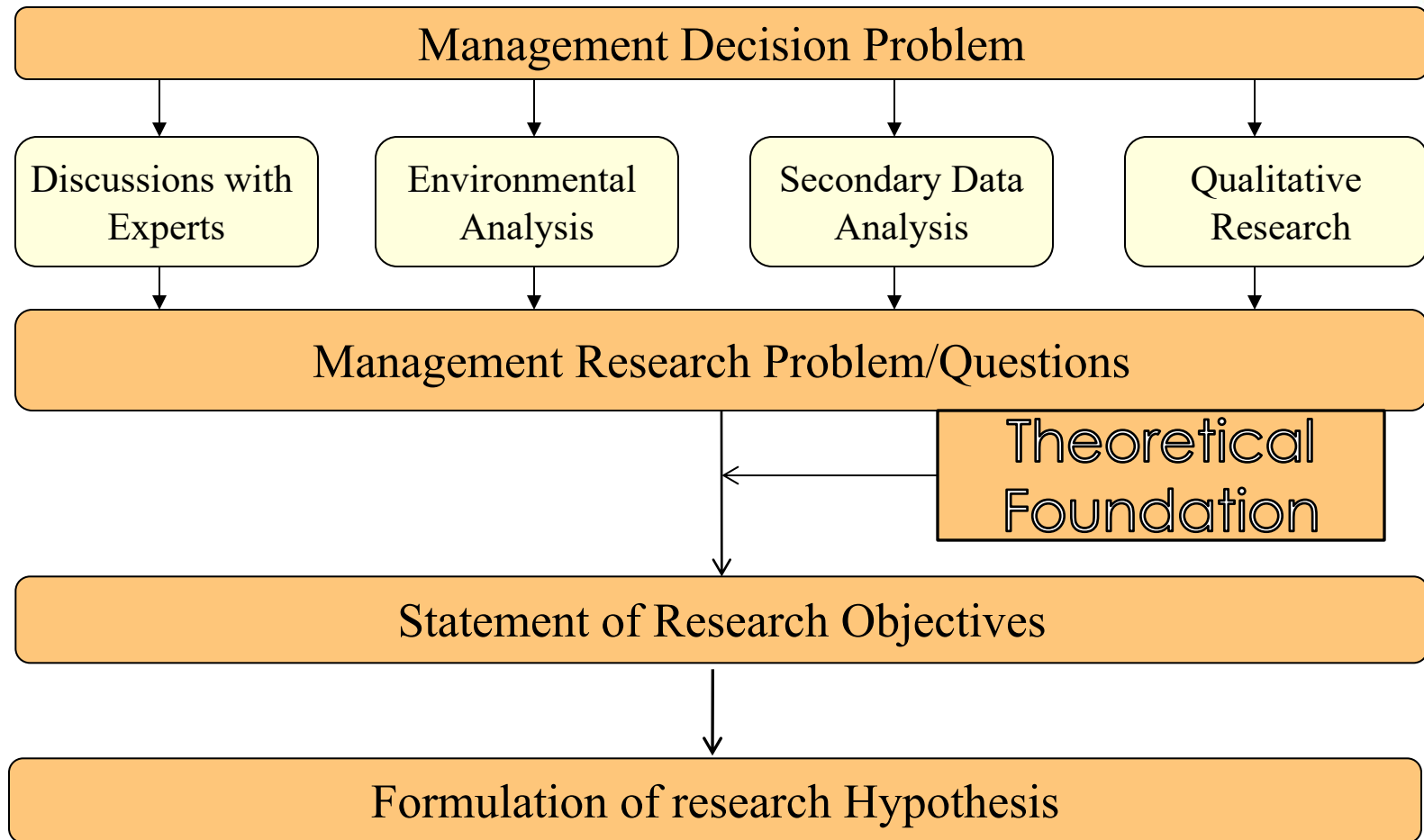
# From a Management Problem to a Research Problem

## *Making the Transition*



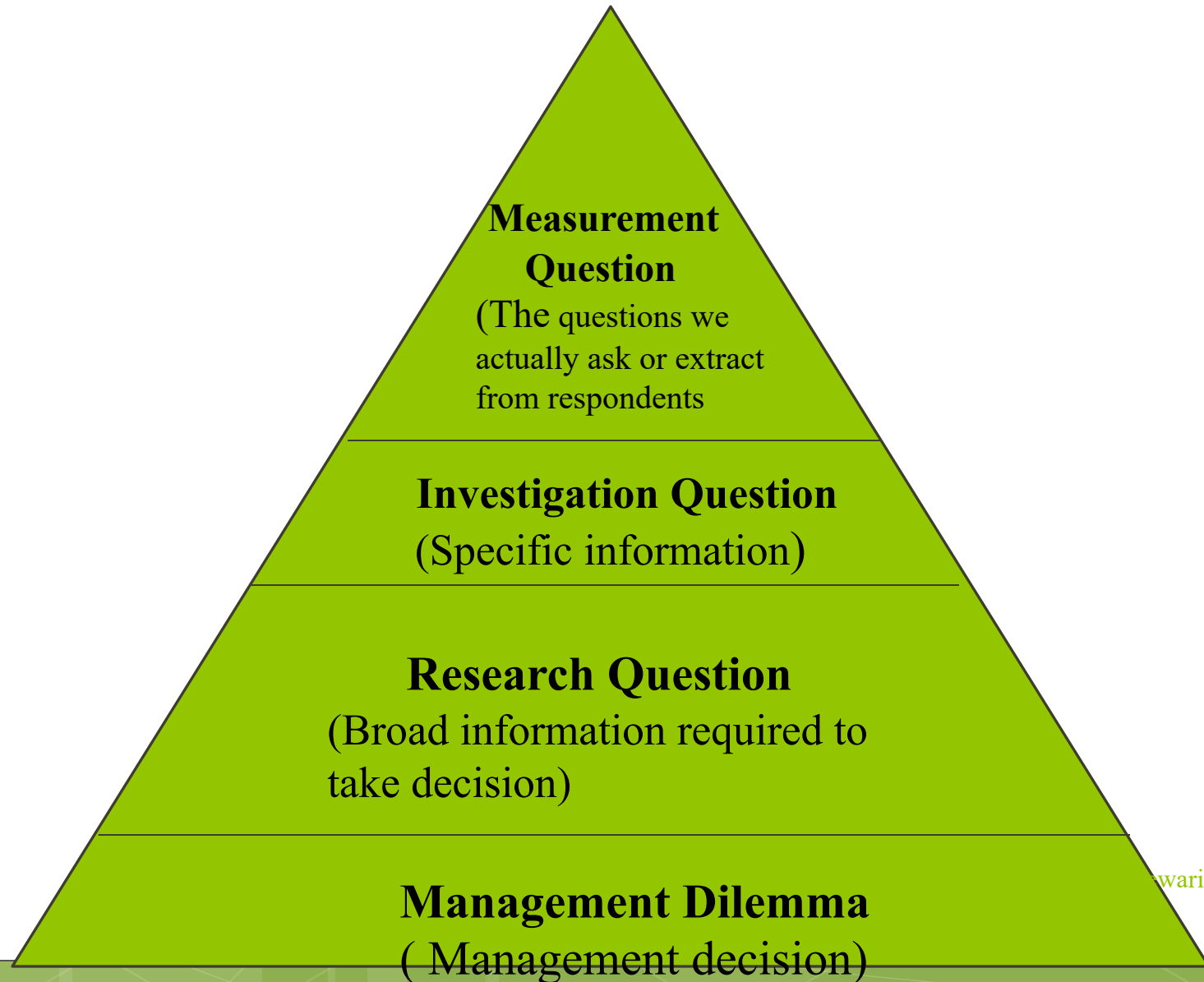
Dr. Asita Ghewari

# THE PROBLEM IDENTIFICATION PROCESS





# Scientific Enquiry





# Management Decision Problem Vs Research Problem

- Management Problem

Asks what the decision  
Maker needs to do

Action Oriented  
Focuses on symptoms

Research Problem

Asks what information  
is needed and how

Information oriented  
on causes



Dr. Asita Ghewari

# An Example

Situation: A coffee shop was concerned about low traffic and sales. Management did not know customers perceptions of the coffee shop nor did they have a feel for awareness of the coffee shop among residents.



Dr. Asita Ghewari

# Possible Management Problems & Research Questions

## Management's Problems/Questions:

- Why are store sales so low?
- How does the coffee shop increase customer traffic?
- What is the product mix that would best satisfy customers?
- What is the awareness of the coffee shop?
- What needs do our customers have that are currently not being met?

## **Redefined as Research Questions**

- What is the satisfaction among current coffee shop customers?
- What are perceptions of the coffee shop and its competitors?
- What is overall awareness of the coffee shop in the town?
- What is competitive position of coffee shop?

Management Problems	Research Problems
1. Allocation of advertising budget among different media type	
2. Whether a training program launched last year is to be continued this year?	
3. Should a new product be introduced?	
4. Should the advertising campaign be changed?	
5. Should the compensation package be changed to motivate employees?	
6. Develop package for new product.	
7. Increase store traffic.	

Dr. Asita Ghewari



# Putting It All Together



- Management Problem

- The U of L Placement office has noticed, while major companies make annual recruiting visits to campus for engineers, few local companies formally recruit business management majors through the placement office
- What should the Placement office do get local companies to hire business management majors ?

# THE RESEARCH PROPOSAL

- A written statement of the research design (strategy, blueprint) that often also includes
  - a statement explaining the purpose of the study (in the form of research objectives or deliverables)
  - a definition of the problem (often in the form of a decision statement)
  - the particular research methodology that will be employed
  - details of procedures that will be used during each stage of the research process, and
  - a schedule of costs and deadlines

Dr. Asita Ghewari

# Contents of the Research Proposal

1. Executive Summary
2. Problem Statement
3. Research Objective
4. Review of Literature
5. Importance or Benefits of the study
6. Research Design
  - Specific research method (or methods) to be used
  - Justification for choice of methods

# Contents of the Research Proposal - Continued



7. **Data Processing/Analysis**  
How the data will be processed  
What analysis techniques will be used for specific research objectives
8. **Presentation of Results**  
How results will be presented  
Written report (dummy table)  
Oral presentation
10. **Qualification of researchers**
11. **Time Schedule**  
Time schedule for completion of the study (if possible in the form of a Gantt Chart)
12. **Budget**  
Overall cost of the study (including breakdown of different cost components)
13. **Bibliography**

# Time Schedule

	TIME FROM START OF STUDY (WEEKS)								
RESEARCH ACTIVITY	1	2	3	4	5	6	7	8	9
<u>DISTRIBUTOR STUDY</u>									
1. Observation and store audits									
2. Depth interviews									
<u>CONSUMER STUDY</u>									
3. Focus Group Interviews									
4. Observation of consumer shopping									
5. Consumer Survey									
6. Data analysis and report writing									
7. Oral presentation									

Dr. Asita Ghewari

# Functions of a Research Proposal

## For Management (the client):

1. Evaluation
  - It helps the client to evaluate the ability of the research to accomplish its objectives
2. Choosing research suppliers
  - Client judges quality of different research suppliers from their proposals
3. Judgment standard
  - It is a standard for determining if the research was conducted as planned

# Functions of a Research Proposal



## For the Researcher:

1. Thinking ahead
  - + It forces the researcher to think through the different stages of the research process
2. Bidding document
  - + It serves as the researcher's bid to offer a specific service

## For Both Parties:

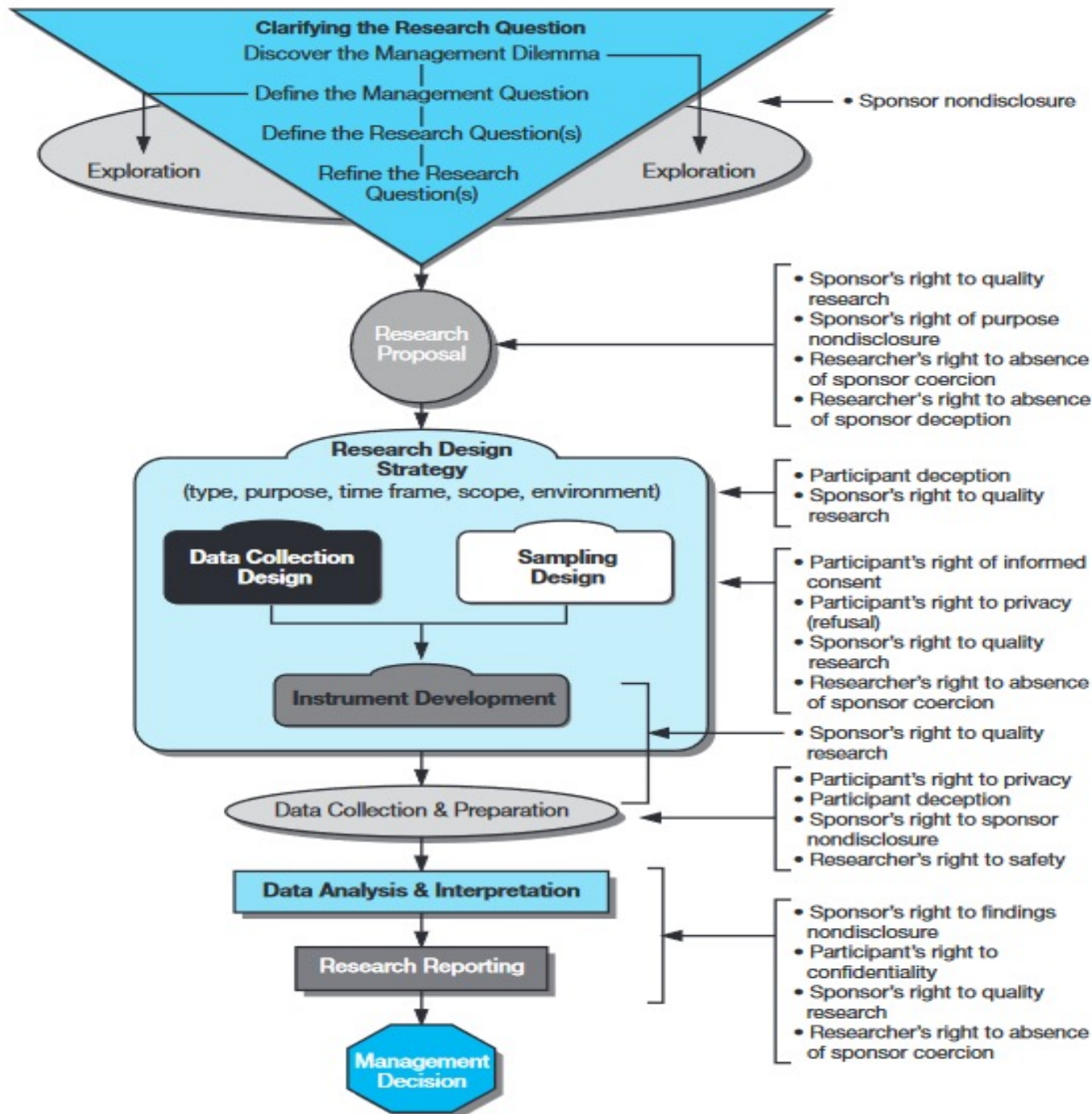
1. Communication
  - + It serves as a means of communication between the researcher and the client (decision maker)
2. Written Contract
  - + It serves as a written contract that can always be referred to

Dr. Asita Ghewari



# Ethical Principles

- ◉ Harm to Participants
- ◉ Lack of Informed Consent
- ◉ Invasion of Privacy
- ◉ Deception
- ◉ Reciprocity and trust
- ◉ Affiliation & conflict of interests



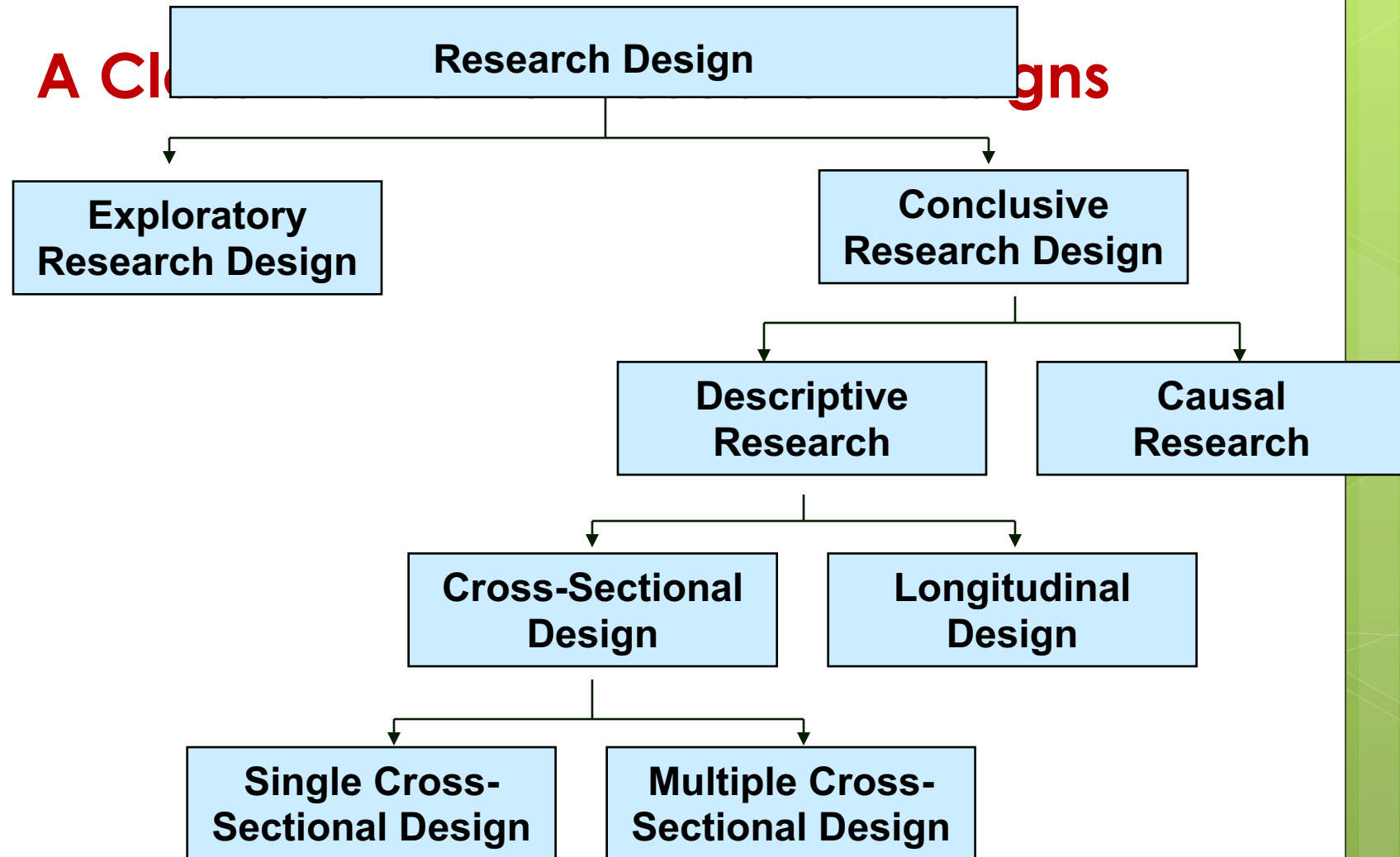
# RESEARCH DESIGN

## UNIT 2

---

# Research Design: Definition

- A **research design** is a framework or blueprint for conducting the research project. It details the procedures necessary for obtaining the information needed to structure or solve research problems.



# Exploratory & Conclusive Research Differences

## Exploratory

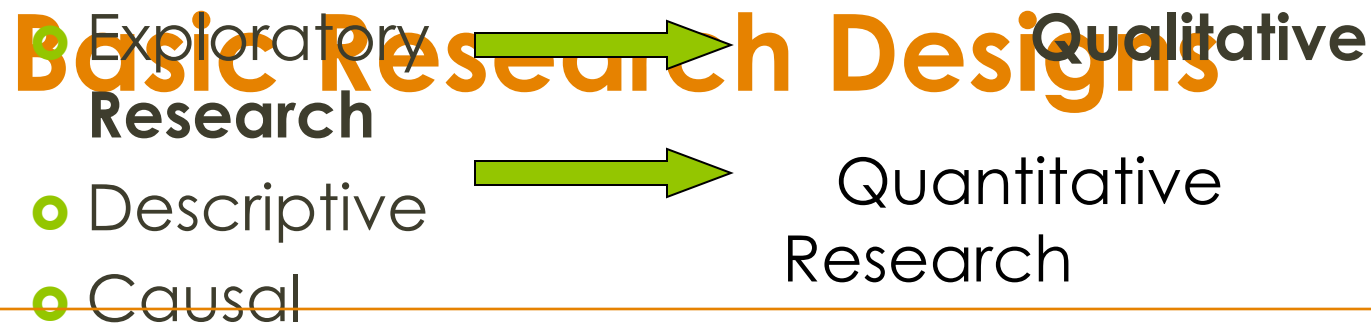
## Conclusive

<b>Objective:</b>	<b>To provide insights and understanding.</b>	<b>To test specific hypotheses and examine relationships.</b>
<b>Characteristics:</b>	<b>Information needed is defined only loosely. Research process is flexible and unstructured. Sample is small and non-representative. Analysis of primary data is qualitative.</b>	<b>Information needed is clearly defined. Research process is formal and structured. Sample is large and representative. Data analysis is quantitative.</b>
<b>Findings /Results:</b>	<b>Tentative.</b>	<b>Conclusive.</b>
<b>Outcome:</b>	<b>Generally followed by further exploratory or conclusive research.</b>	<b>Findings used as input into decision making.</b>



# A Comparison of Basic Research Designs

	<b>Exploratory</b>	<b>Descriptive</b>	<b>Causal</b>
<b>Objective:</b>	Discovery of ideas and insights	Describe characteristics or functions	Determine cause and effect relationships
<b>Characteristics:</b>	<p>Flexible, versatile</p> <p>Often the front end of total research design</p>	<p>Marked by the prior formulation of specific hypotheses</p> <p>Preplanned and structured design</p>	<p>Manipulation of one or more independent variables</p> <p>Control of other mediating variables</p>
<b>Methods:</b>	<p>Focus group</p> <p>In depth Interview</p> <p>Projective Technique</p>	<p>Secondary data</p> <p>Surveys</p> <p>Panels</p> <p>Observation and other data</p>	<p>Experiments</p>



### Types of Research

Degree of Problem Definition	Exploratory Research (Unaware of Problem)	Descriptive Research (Aware of Problem)	Causal Research (Problem Defined)
Possible Situation	<p>"Our sales are declining and we don't know why."</p> <p>"Would people be interested in our new product idea?"</p>	<p>"What kind of people buy our product? Who buys our competitor's product?"</p> <p>"What features do buyers prefer in our product?"</p>	<p>"Will buyers purchase more of our product in a new package?"</p> <p>"Which advertising campaign is more effective?"</p>

# Uses of Exploratory Research

- ◉ Formulate a problem or define a problem more precisely
- ◉ Identify alternative courses of action
- ◉ Develop hypotheses
- ◉ Gain insights for developing an approach to the problem
- ◉ Establish research priorities

# Use of Descriptive Research

- To describe the characteristics of certain groups.
- To estimate the percentage of units in a specified population exhibiting a certain behavior
- To determine the perceptions of respondent about characteristics
- To determine the degree to which variables are associated
- To make specific predictions

# Uses of Casual Research

- To understand which variables are the cause (independent variables) and which variables are the effect (dependent variables) of a phenomenon
- To determine the nature of the relationship between the causal variables and the effect to be predicted
- METHOD: Experiments

# RESEARCH DESIGN

- Descriptive research
  - A type of conclusive research that has its major objectives the description of something—usually characteristics or functions.
  - Descriptive research seeks to determine the answers to who, what, when, where and how questions.



- **Qualitative Research** involves finding out what people think, and how they feel - or at any rate, what they say they think and how they say they feel. This kind of information is subjective. It involves feelings and impressions, rather than numbers'

## **Qualitative**

**The aim of qualitative analysis is a complete detailed description.**

**The design emerges as the study unfolds**

**Researcher is the data gathering instrument.**

**Data is in the form of words (interviews), pictures (videos), or objects (artifacts).**

**Qualitative data is more rich, time consuming, and less able to be generalized.**

## **Quantitative**

**In quantitative research we classify features, count them, and construct statistical models in an attempt to explain what is observed.**

**All aspects of the study are carefully designed before data is collected.**

**Researcher uses tools (questionnaires or equipment) to collect data.**

**Data is in the form of numbers and statistics.**

**Quantitative data is more efficient, able to test hypotheses, but may miss contextual data.**

# Most frequent uses

- ◉ Understanding basic issues
  - ◉ why do people buy/use our product?
- ◉ Pretesting ideas or questions
  - ◉ do people want a product that cleans their refrigerator?
- ◉ Message testing
  - ◉ How do people like this ad?
- ◉ Recommended to capture the basic feel of a problem prior to conducting a more analytical study

# Strengths

- **Good for examining feelings and motivations**
- **Allows for complexity and depth of issues**
- **Provides insights**

# Weaknesses

- **Can't extrapolate to the whole population**
- **Volume of data**
- **Complexity of analysis**
- **Time-consuming nature of the clerical efforts**

require

# General approaches

- Individual interviews
  - Nonstructured
  - Structured
- Projective Techniques
- Group interviews
  - Structured or unstructured
  - Focus groups
- Observation

# Focus Group

*An interview conducted by a trained moderator in a non-structured and natural manner with a small group of respondents.*

# Characteristics of Focus Groups



Group size	8-12
Group composition	Homogeneous; respondents prescreened
Physical setting	Relaxed, informal atmosphere
Time duration	1 - 3 hours
Recording	Use of audio cassettes and videotapes
Moderator	Observational, interpersonal, and communication skills of the moderator



# Focus Group



## **2. DEPTH INTERVIEWS**

- One-on-one interviews between professional researcher and a respondent
- Try to uncover underlying motivations, prejudices and attitudes – lots of probing
- Laddering technique

# When to use depth interviews:

- ◉ Sensitive or confidential subject matter
- ◉ Situations where strong social norms exist
- ◉ Need intensive probing
- ◉ When respondent interaction unlikely to be helpful (e.g., with children)
- ◉ Have lots of money and time
- ◉ Need detailed responses

# **observation**

**. . . systematic witnessing  
and/or recording of behavioral  
patterns of objects, people, and  
events without directly  
communicating with them – can  
collect both qualitative and  
quantitative data.**

# Types of Observed Phenomena

- Physical actions
- Verbal behavior
- Expressive behavior
- Spatial relations and locations
- Temporal patterns
- Verbal and pictorial records

# Types of Observation Techniques



- ◉ Natural versus Contrived Observation
- ◉ Structured versus Unstructured Observation
- ◉ Disguised versus Nondisguised Observation
- ◉ Physical-trace evidence Observation
- ◉ Mechanical Observation

# Types of Mechanical Observation

- ◉ Eye-Tracking: Oculometer, Pupilometer
- ◉ Response Latency
- ◉ Voice Pitch Analysis
- ◉ People Meter
- ◉ Monitoring Web Site Traffic



# PROJECTIVE TECHNIQUES

- A collection of exploratory research techniques based on indirect questioning in which respondents are asked to project themselves into a particular person, object, or situation
- Often used where direct questioning is not likely to provide honest responses
  - Word association tests
  - Completion method
  - Construction Technique
  - Expressive Technique

# Continue...

- An unstructured, indirect form of questioning that encourages respondents to project their underlying motivations, beliefs, attitudes or feelings regarding the issues of concern.
- In projective techniques, respondents are asked to interpret the behavior of others.
- In interpreting the behavior of

# Causal Research Design

- ✖ Causal research designs that can identify cause-and-effect relationships between variables/constructs.
- ✖ Variable: the observable and measurable elements (or attributes) of an object or an event. Examples include demographics, attitudes, behaviors, etc.

# Experiment

- Scientific investigation in which an investigator manipulates and controls one or more independent variables and observes the degree to which the dependent variables change.

# Conditions for Causality

- **Concomitant variation** is the extent to which a cause,  $X$ , and an effect,  $Y$ , occur together or vary together in the way predicted by the hypothesis under consideration.
- The **time order of occurrence** condition states that the causing event must occur either before or simultaneously with the effect; it cannot occur afterwards.
- The **absence of other possible**

# Definitions and Concepts

- **Independent variables** are variables or alternatives that are manipulated and whose effects are measured and compared, e.g., price levels.
- **Test units** are individuals, organizations, or other entities whose response to the independent variables or treatments is being examined, e.g., consumers or stores.
- **Dependent variables** are the variables which measure the effect of the independent variables on the test units, e.g., sales, profits, and market shares.
- **Extraneous variables** are all variables other than the independent variables that affect the response of the test units, e.g., store size, store location, and competitive effort.
- Experimental group
- Control group

# Experiments – Validity

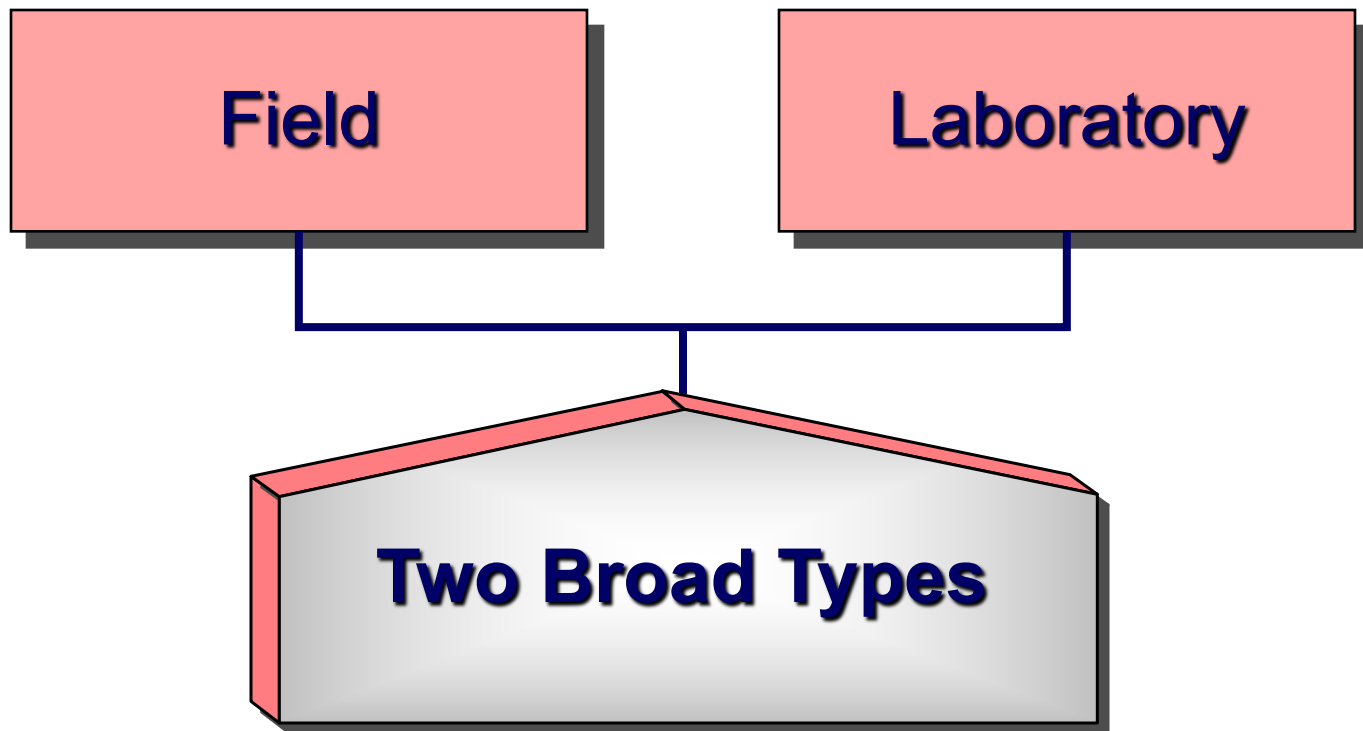
**Validity – the extent to which the conclusions drawn from the experiment are true.**

**Internal validity – the extent to which the research design accurately identifies causal relationships.**

**External validity – the extent to which a causal relationship is true for the defined target population.**



# Types of Experiments

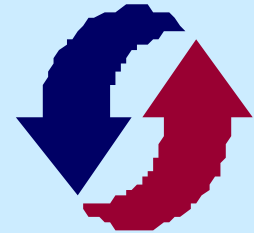


# Experiments

- Potential causes are “controlled” by using experimental designs and manipulation:
  - Manipulation – the causal variable is altered over different conditions.
    - Lab experiment – manipulation takes place in artificial setting.
      - \* maximizes control.
    - Field experiment – manipulation takes place in the relevant business context.
      - \* increases representativeness.

# Laboratory versus Field Experiments

Factor	Laboratory	Field
Environment	Artificial	Realistic
Control	High	Low
Reactive Error	High	Low
Demand Artifacts	High	Low
Internal Validity	High	Low
External Validity	Low	High
Time	Short	Long
Number of Units	Small	Large
Ease of Implementation	High	Low
Cost	Low	High



# Limitations of

## Experimentation

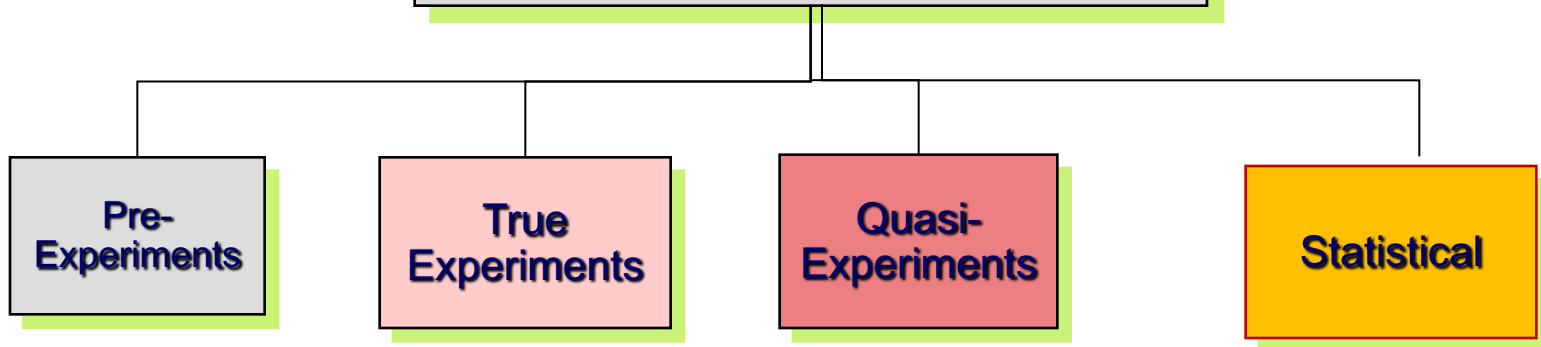
- Experiments can be time consuming, particularly if the researcher is interested in measuring the long-term effects.
- Experiments are often expensive. The requirements of experimental group, control group, and multiple measurements significantly add to the cost of research.
- Experiments can be difficult to administer. It may be impossible to control for the effects of the extraneous variables, particularly in a field environment.
- Competitors may deliberately contaminate the results of a field experiment.

# Controlling for extraneous (confounding) variables...

1. Randomization
  2. Matching
  3. Experimental design
  4. Statistical control
- analysis of covariance (ANCOVA)

# Experimental Research

## Types of Designs



# A Classification of Experimental Designs

- **Pre-experimental designs** do not employ randomization procedures to control for extraneous factors: the one-shot case study, the one-group pretest-posttest design, and the static-group.
- In **true experimental designs**, the researcher can randomly assign test units to experimental groups and treatments to experimental groups: the pretest-posttest control group design, the posttest only control



# A Classification of Experimental Designs

- ◉ **Quasi-experimental designs** result when the researcher is unable to achieve full manipulation of scheduling or allocation of treatments to test units but can still apply part of the apparatus of true experimentation: time series and multiple time series designs.
- ◉ **Statistical Designs** are CRD, RBD, LCD, factorial design

- **X** refers to the treatment or manipulation of the independent variable (*more than one X refers to a different level of treatment*).
- **O** refers to the observation or measurement of the dependent variable.
- Experimental designs vary widely in their power to control contamination of the relationship between the independent and dependent variables.
- Experiments can be categorized as pre-experiments, true experiments, and Quasi experiments based on the characteristic of control.

# Pre-experiment

Pre-experimental research designs are research designs that are characterized by a lack of random selection and assignment.

Types of Pre-experiments:

- ✖ After-Only Case Study
- ✖ One Group Pretest-Posttest Design
- ✖ Static Group Comparison

## After Only Case Study

X

O

- ◉ In this type of experimental design only one treatment (X) or manipulation is done on the independent variable.
- ◉ Then, the dependent variable is measured.

X

O

An example is a media campaign about a product's features without a prior measurement of consumer knowledge.

Results would reveal only how much target consumers know after the media campaign, but there is no way to judge the effectiveness of the campaign.

The lack of a pretest and control group makes this design inadequate for establishing causality.

# One Group Pretest-Posttest Design

$O_1$

X

$O_2$

This design meets the threats to internal validity better than the one-shot case study, but it is still a weak design.

For example, a researcher examining the effect of a commercial on brand liking would begin by taking a pre-test to determine current levels of brand liking among the participants.

$O_1$

X

$O_2$

The commercial would be shown.

Then a post-test would measure brand liking after the commercial.

A comparison between the post-test and the pre-test shows the change in liking.

However, any changes in liking are not necessarily due to the commercial.

The act of giving a pre-test could have influenced liking (testing effect).

# Static Group Comparison

Experimental Group:

X

O<sub>1</sub>

Control Group:

O<sub>2</sub>

This design provides for two groups, one of which receives the experimental stimulus while the other serves as a control.



For example, imagine that a new type of cheeseburger is being introduced, and an advertisement campaign is run.

After the ad airs, those who remember seeing it would be in the experimental group (X). Those who have no recall of the ad would be in the control group.

The intent of each group to purchase the cheeseburger would be measured.

The main weakness of this design is that there is no way to be certain that the two groups are equivalent or that the individuals are representative.

# True experiment



- ✖ A **true experiment** is a method of research in which there are two kinds of variables. The independent variable is manipulated by the experimenter, and the dependent variable is measured.
- ✖ The signifying characteristic of a true experiment is that it randomly allocates the subjects in order to neutralize the potential to ensure equivalence.
- ✖ There is also a control group for comparison.

Types of True experiments:

- ✖ Pretest-Posttest Control Group Design
- ✖ Posttest-Only Control Group Design
- ✖ Solomon four group design

# Pretest-Posttest Control Group Design

Experimental Group:	R	O <sub>1</sub>	X	O <sub>2</sub>
Control Group:	R	O <sub>3</sub>		O <sub>4</sub>

- ✖ The symbol R means that the true experimental designs use randomly assigned groups to ensure equivalence.
- ✖ The effect of the experimental is:  $E = (O_2 - O_1) - (O_4 - O_3)$ .
- ✖ This design deals with many of the threats to internal validity.
- ✖ External validity is threatened.

# Posttest-Only Control Group Design

Experimental Group:	R	X	O <sub>1</sub>
Control Group:	R		O <sub>2</sub>

- In this design, the pretest measurements are omitted.
- Pretests are well established in classical research design but are not really necessary **when it is possible to randomize.**
- The experimental effect is measured by the difference between O<sub>1</sub> and O<sub>2</sub>.

## Example for Posttest-Only Control Group Design

- Buick dealerships wish to determine the effectiveness of a special “test-drive” incentive.
- Buick dealerships nationwide are randomly assigned to either the control group or the experimental group.
- Those in the experimental group use a promotion to encourage test drives.
- The control group does not use any such promotions.
- The number of test drives throughout are measured and compared to determine if the promotion resulted in significantly more test

# Solomon Four Group

## Design 1

(EG):	[R]	$O_1$	$\rightarrow$	X	$\rightarrow$	$O_2$
(CG):	[R]	$O_3$		$\rightarrow$		$O_4$

## Design 2

(EG):	[R]			X	$\rightarrow$	$O_5$
(CG):	[R]					$O_6$

## Quasi-Experimental Designs: Time Series Design

$O_1$   $O_2$   $O_3$   $O_4$   $O_5$   $X$   $O_6$   $O_7$   $O_8$   $O_9$   
 $O_{10}$

- There is no randomization of test units to treatments.
- The timing of treatment presentation, as well as which test units are exposed to the treatment, may not be within the researcher's control.

# Group Time Series Design

R	O <sub>1</sub>	O <sub>2</sub>	O <sub>3</sub>	X	O <sub>4</sub>	O <sub>5</sub>	O <sub>6</sub>
R	O <sub>7</sub>	O <sub>8</sub>	O <sub>9</sub>		O <sub>10</sub>	O <sub>11</sub>	O <sub>12</sub>

A time series design introduces repeated observations before and after treatment and allows participants to act as their own controls.

The single treatment group design has before-after measurements as the only controls.

There is also a multiple design with two or more comparison groups as well as the repeated measurements in each treatment group.



# One-Shot Case Study

$X \quad O_1$

- A single group of test units is exposed to a treatment  $X$ .
- A single measurement on the dependent variable is taken ( $O_1$ ).
- There is no random assignment of test units.
- The one-shot case study is more appropriate for exploratory than for

# One-Group Pretest-Posttest Design

$O_1$      $X$      $O_2$

- A group of test units is measured twice.
- There is no control group.
- The treatment effect is computed as  $O_2 - O_1$ .
- The validity of this conclusion is questionable since extraneous variables are largely uncontrolled.

# Static Group Design

EG:                       $X$          $O_1$   
CG:                       $O_2$

- A two-group experimental design.
- The experimental group (EG) is exposed to the treatment, and the control group (CG) is not.
- Measurements on both groups are made only after the treatment.
- Test units are not assigned at random.

## True Experimental Designs: Pretest-Posttest Control Group Design

EG:	R	$O_1$	X	$O_2$
CG:	R	$O_3$		$O_4$

- Test units are randomly assigned to either the experimental or the control group.
- A pretreatment measure is taken on each group.
- The treatment effect (TE) is measured as:  $(O_2 - O_1) - (O_4 - O_3)$ .
- Selection bias is eliminated by randomization.
- The other extraneous effects are controlled as follows:

$$O_2 - O_1 = TE + H + MA + MT + IT + I + SR + MO$$

$$O_4 - O_3 = H + MA + MT + I + SR + MO$$

= EV (Extraneous Variables)

# Posttest-Only Control Group Design

EG : R X  $O_1$

CG : R  $O_2$

- The treatment effect is obtained by  
 $TE = O_1 - O_2$
- Except for pre-measurement, the implementation of this design is very similar to that of the pretest-posttest control group design.

# Limitations of

## Experimentation

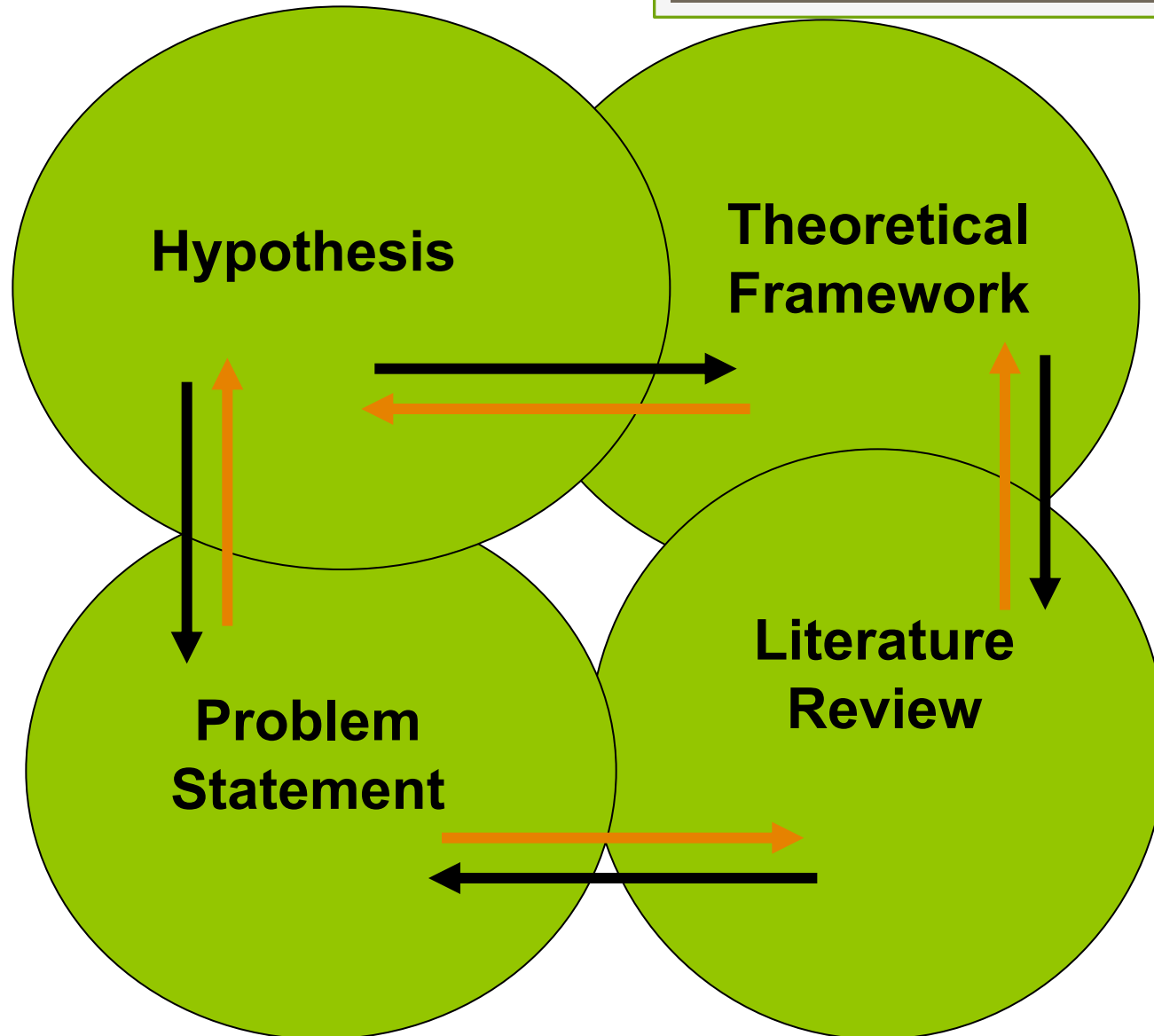
- Experiments can be time consuming, particularly if the researcher is interested in measuring the long-term effects.
- Experiments are often expensive. The requirements of experimental group, control group, and multiple measurements significantly add to the cost of research.
- Experiments can be difficult to administer. It may be impossible to control for the effects of the extraneous variables, particularly in a field environment.
- Competitors may deliberately contaminate the results of a field experiment.

# Propositions & Hypotheses

- Proposition
  - A statement about concepts that may be judged as TRUE or FALSE if it refers to observable phenomenon
  - Proposition formulated for empirical testing is Hypothesis
  - Example
    - Infosys employees have higher than average achievement motivation

- **Hypotheses:** intelligent hunches, guesses, or predictions that assist the researcher in seeking the solution or answer to the research question.
- A **theory's** validity is not directly examined. Instead, it is through the hypotheses that the merit of a theory can be evaluated





# The Role of the Hypothesis

- ◉ Guides the direction of the study
- ◉ Identifies facts that are relevant
- ◉ Suggests which form of research design is appropriate
- ◉ Provides a framework for organizing the conclusions that result

# Characteristics & Types of a Good Hypothesis

- A good hypothesis should fulfill conditions:
  - Must be adequate for its purpose
  - Must be testable
  - Must be state relationship if relation hypothesis
  - Logical, simple, consistent.
- Hypothesis types
  - Descriptive
  - Relational: Correlation & Causal

# Types of Hypothesis

- Descriptive

- Describes the existence, size, form or distribution of some variables
- Eighty percent of shareholders of HLL favour increasing the company's cash dividend

# Types of Hypothesis

- Relational
  - Statements that describe the relationship between two variables with respect to some case
  - Foreign (variable) refrigerators are perceived to be of better quality (variable) by Indian consumers (case)

# Types of Relational Hypothesis

- Two types: Correlation & Causal
- Correlation
  - Merely states that variables occur together without implying that one causes the other
    - People in Kerela give more importance to education than people in Punjab
    - In an office old employees are more responsive than young employees

# Types of Relational Hypothesis

- Causal (or Explanatory)
  - There is an implication that existence of (or a change in) one causes or leads to a change in the other
    - Causal variable is called Independent variable and the other Dependent variable
    - Advertisement causes higher sales
    - Increase in income leads to higher savings

# MEASUREMENT AND DATA

BRM Unit 3

Dr. Asita Ghewari



# MEASUREMENT

# Measurement and Scaling

**Measurement** means assigning numbers or other symbols to characteristics of objects according to certain pre specified rules.

- ▣ One-to-one correspondence between the numbers and the characteristics being measured.
- ▣ The rules for assigning numbers should be standardized and applied uniformly.
- ▣ Rules must not change over objects or time.

1. Selecting observable empirical event
2. Developing a set of mapping rules: a scheme for assigning numbers or symbols to represent aspects of the event being measured
3. Applying the mapping rules

# Measurement and Scaling

**Scaling** involves creating a continuum on which objects are located according to certain characteristics of the objects of interest according to some pre specified rules.

Consider an attitude scale from 1 to 100. Each respondent is assigned a number from 1 to 100, with 1 = Extremely Unfavorable, and 100 = Extremely Favorable. Measurement is the actual assignment of a number from 1 to 100 to each respondent. Scaling is the process of placing the respondents on a continuum with respect to their attitude toward department stores.

# LEVELS OF SCALE MEASUREMENT



- ❑ Numbers assigned in measurement can take on different levels of meaning depending on one of four mapping characteristics possessed by the numbers:
  1. **Classification** - The numbers are used only to group or sort responses. No order exists
  2. **Order** - The numbers are ordered. One number is greater than, less than, or equal to another
  3. **Distance** – The characteristics of distance means the absolute difference between the scale descriptors are known and may be expressed in units. The difference between any pair of numbers is greater than, less than, or equal to the difference between any other pair of numbers
  4. **Origin** - The number series has a unique origin indicated by the number zero
- ❑ The type of mapping characteristic assumed depends on the properties of the attribute being measured

Dr. Asita Ghewari

# The Four Levels of Scale Measurement

- Four levels of scale measurement result from this mapping
- 1. **Nominal Scale**: a scale in which the numbers or letters assigned to an object serve only as labels for identification or classification, e.g. Gender (Male=1, Female=2)
- 2. **Ordinal Scale**: a scale that arranges objects or alternatives according to their magnitude in an ordered relationship, e.g. Academic status (Sophomore=1, Freshman=2, Junior=3, etc)
- 3. **Interval Scale**: a scale that both arranges objects according to their magnitude, distinguishes this ordered arrangement in *units of equal intervals*, but does not have a natural zero representing absence of the given attribute, e.g. the temperature scale (40°C is not twice as hot as 20°C)
- 4. **Ratio Scale**: a scale that has absolute rather than relative quantities and an absolute (natural) zero where there is an absence of a given attribute, e.g. income, age.

# Nominal, Ordinal, Interval, and Ratio Scales Provide Different Information

Nominal



Ordinal

Show



Place



Win



Interval



20 Seconds



1.0 Second

Win



Ratio

Win



40 to 1  
long shot  
pays  
\$80

Dr. Asita Ghewari



Level	Examples	Numerical Operations	Descriptive Statistics
Nominal	Student ID number Yes – No Male – Female Buy – Did Not Buy East region Central region West region	Counting	<ul style="list-style-type: none"> <li>• Frequencies</li> <li>• Mode</li> </ul>
Ordinal	Student class rank Please rank your three favorite movies. Choose from the following: <ul style="list-style-type: none"> <li>• Dissatisfied</li> <li>• Satisfied</li> <li>• Very satisfied</li> <li>• Delighted</li> </ul> Indicate your level of education: <ul style="list-style-type: none"> <li>• Some high school</li> <li>• High school diploma</li> <li>• Some college</li> <li>• College degree</li> <li>• Graduate degree</li> </ul>	Counting Ordering	<ul style="list-style-type: none"> <li>• Frequencies</li> <li>• Mode</li> <li>• Median</li> <li>• Range</li> </ul>
Interval	Student grade point average (GPA) Temperature (Celsius and Fahrenheit) Points given on an essay question 100-point job performance rating provided by supervisor	Common arithmetic operations	<ul style="list-style-type: none"> <li>• Frequencies</li> <li>• Mode</li> <li>• Median</li> <li>• Range</li> <li>• Mean</li> <li>• Variance</li> <li>• Standard deviation</li> </ul>
Ratio	Amount spent on last purchase Salesperson sales volume Number of stores visited on a shopping trip Annual family income Time spent viewing a Web page	All arithmetic operations	<ul style="list-style-type: none"> <li>• Frequencies</li> <li>• Mode</li> <li>• Median</li> <li>• Range</li> <li>• Mean</li> <li>• Variance</li> </ul>

# Classification of Different Levels of Scale Measurement



Type of Scale	Data Characteristics	Numerical Operation	Descriptive Statistics	Examples
Nominal	Classification but no order, distance, or origin	Counting	Frequency in each category Percent in each category Mode	Gender (1=Male, 2=Female)
Ordinal	Classification and order but no distance or unique origin	Rank ordering	Median Range Percentile ranking	Academic status (1=Secondary, 2=Higher Secondary, 3=Under graduate, 4=Graduate)
Interval	Classification, order, and distance but no unique origin	Arithmetic operations that preserve order and magnitude	Mean Standard deviation Variance	Temperature in degrees Attitudinal scale
Ratio	Classification, order, distance and unique origin	Arithmetic operations on actual quantities	Geometric mean Coefficient of variation	Age in years Income in Saudi riyals

Dr. Asita Ghewari

Note: All statistics appropriate for lower-order scales (nominal being lowest) are appropriate for higher-order scales (ratio being the highest)



• Which of the soft drinks in the following list do u like



- Coke
- Mountain Dew
- Pepsi
- Seven- up
- Sprite

Rank the soft drinks according to how much u like each  
(most preferred=1, least preferred=5)

- Coke
- Mountain Dew
- Pepsi
- Seven- up
- Sprite

Please indicate how much u like each soft drink by checking the appropriate position on the scale

☐ Dislike a lot      Dislike    like    like a lot

- Coke
- Mountain Dew
- Pepsi
- Seven- up
- Sprite

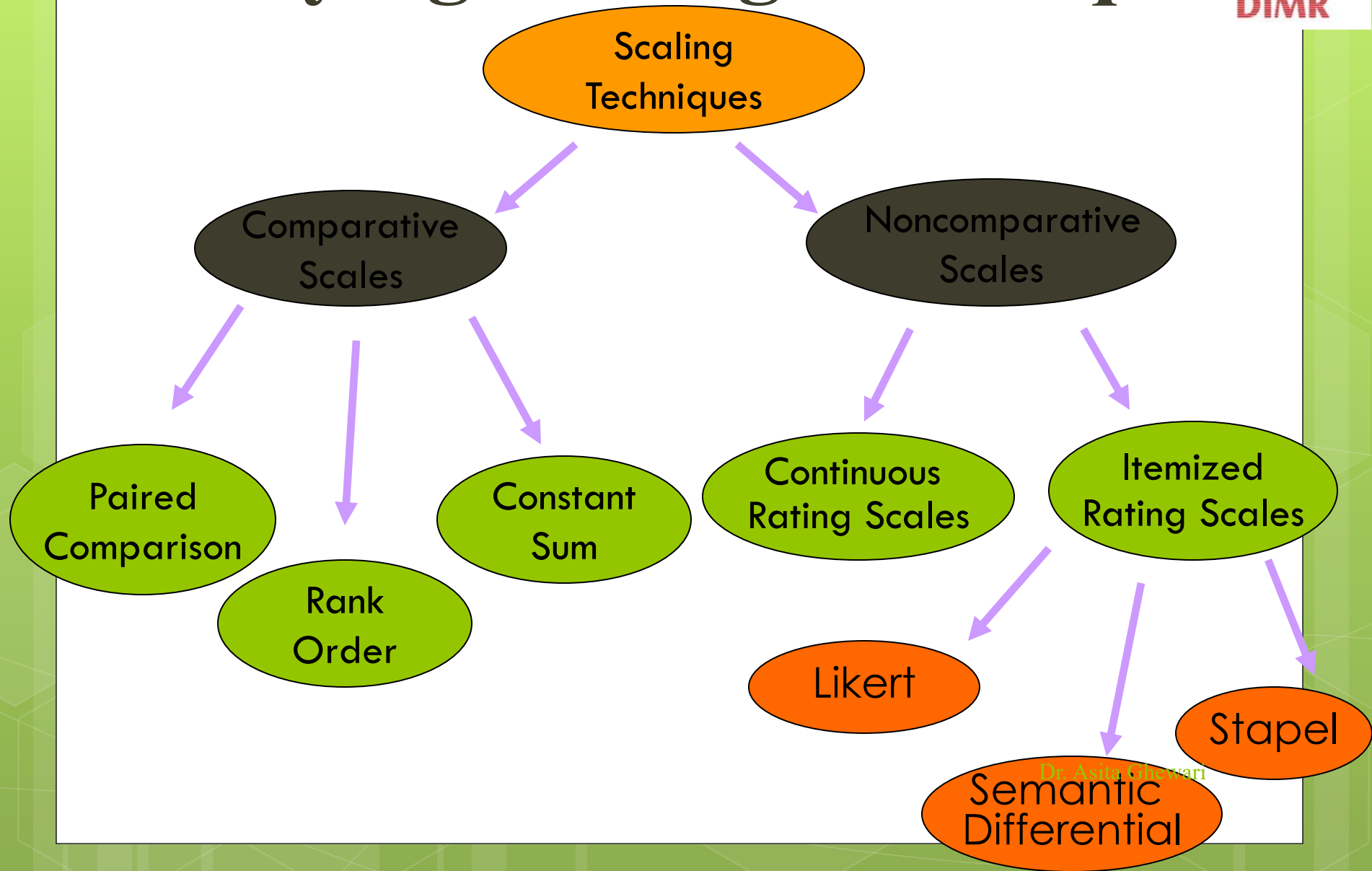
Please divide 100 points among these soft drink to represent how much u like

- Coke
- Mountain Dew
- Pepsi
- Seven- up
- Sprite

# A Comparison of Scaling Techniques

- **Comparative scales** involve the direct comparison of objects. Comparative scale data must be interpreted in relative terms and have only ordinal or rank order properties.
- In **Non comparative scales**, each object is scaled independently of the others in the set. The resulting data are generally assumed to be interval or ratio scaled.

# Classifying Scaling Techniques



# Comparative Scaling Techniques

## Paired Comparison Scaling

- A respondent is presented with two objects and asked to select one according to some criterion.
- The data obtained are ordinal in nature.
- Paired comparison scaling is the most widely used comparative scaling technique.

**Instructions:** We are going to present you with five pairs of shampoo brands. For each pair, please indicate which one of the two brands of shampoo you would prefer for personal use.

**Recording Form:**

	Dove	Pantine	Clinic plus	Head & houlders	Lorel
Dove		0	0	1	0
Pantine	1 <sup>a</sup>		0	1	0
Clinic plus	1	1		1	1
Head & Shoulders	0	0	0		0
Lorel	1	1	0	1	
Number of Times Preferred <sup>b</sup>	3	2	0	4	1

Dr. Asita Ghewari

# COMPARATIVE SCALES

## Paired Comparison

Please indicate which of the following airlines you prefer by circling your more preferred airline in each pair:

Air India

WestJet

Air Transat

Air India

Indigo

WestJet

WestJet

Air Transat

Air India

Indigo

Indigo

Air Transat

# Comparative Scaling Techniques

## Rank Order Scaling

- Respondents are presented with several objects simultaneously and asked to order or rank them according to some criterion.
- It is possible that the respondent may dislike the brand ranked 1 in an absolute sense.
- Furthermore, rank order scaling also results in ordinal data.

<u>Brand</u>	<u>Rank Order</u>	
1. <u>Vicco</u>		6. <u>Patanjali Dant Kanti</u>
		7. Close Up _____
2. Colgate _____		8. Pepsodent _____
		9. Babbol _____
3. Dabar _____		10. Sensodyne _____
4. Anchor _____		
5. Meswak _____		

# Comparative Scaling Techniques

## Constant Sum Scaling



- Respondents allocate a constant sum of units, such as 100 points to attributes of a product to reflect their importance.
- If an attribute is unimportant, the respondent assigns it zero points.
- If an attribute is twice as important as some other attribute, it receives twice as many points.
- The sum of all the points is 100. Hence, the name of the scale.



# Importance of Bathing Soap Attributes Using a Constant Sum Scale



## Instructions

On the next slide, there are eight attributes of bathing soaps. Please allocate 100 points among the attributes so that your allocation reflects the relative importance you attach to each attribute. The more points an attribute receives, the more important the attribute is. If an attribute is not at all important, assign it zero points. If an attribute is twice as important as some other attribute, it should receive twice as many points.

Dr. Asita Ghewari

# Importance of Bathing Soap Attributes Using a Constant Sum Scale



Form

## Average Responses of Three Segments

Attribute	Segment I	Segment II	Segment III
1. Mildness	8	2	4
2. Lather	2	4	17
3. Shrinkage	3	9	7
4. Price	53	17	9
5. Fragrance	9	0	19
6. Packaging	7	5	9
7. Moisturizing	5	3	20
8. Cleaning Power	13	60	15
Sum	100	100	100

Dr. Asita Ghewari

# **Constant Sum Scale**

**Please divide 100 points among the following characteristics so the division reflects the relative importance of each characteristic to you in the selection of a bank**

**Hours of service**

---

**Friendliness**

---

**Distance from home**

---

**Investment vehicles**

---

**Parking facilities**

---

# COMPARATIVE SCALES

Compared to Chevrolet, Ford is:

less  
innovative

about the  
same

more  
innovative

☐☐☐☐☐☐☐

1

2

3

4

5

6

7

# NON-COMPARATIVE SCALING TECHNIQUES



**ONE OF THE TWO TYPES OF SCALING TECHNIQUES IN WHICH EACH STIMULUS OBJECT IS SCALED INDEPENDENTLY OF THE OTHER OBJECT IN THE STIMULUS SET**

# Non comparative scaling techniques

- Respondents using a non comparative scale employ whatever rating standard seems appropriate to them.
- They do not compare the object being rated either to another object or to some specified standard.
- They evaluate only one object at a time, and for this reason non comparative scales are often referred to as monadic scales.
- Non comparative techniques consist of continuous and itemized rating scales.

# 1. Continuous rating scales

- In a continuous rating scale, also referred to as a graphic rating scale, respondents rate the objects by placing a mark at the appropriate position on a line that runs from one extreme of the criterion variable to the other.
- The form of the continuous scale may vary considerably.
- For example the line may be vertical or horizontal, Scale points in the form of numbers or brief descriptions, may be provided.

# Continuous rating scales

- Continuous rating scales are easy to construct.
- They can also be easily implemented on the internet.
- The cursor can be moved on the screen in a continuous fashion to select the exact position on the scale that best describes the respondent's evaluation.
- Moreover, the scale values can be automatically scored by the computer, thus increasing the speed and accuracy of processing the data.



# Continuous Rating Scale

Respondents rate the objects by placing a mark at the appropriate position on a line that runs from one extreme of the criterion variable to the other.

The form of the continuous scale may vary considerably.

How would you rate D mart as a department store?

Version 1

Probably the worst -----|----- Probably the best

Version 2

Probably the worst -----|----- Probably the best

0 10 20 30 40 50 60 70 80 90 100

Version 3

Very bad

Neither good  
nor bad

Very good

Probably the worst -----|----- Probably the best

0 10 20 30 40 50 60 70 80 90 100

Dr. Asita Ghewari

## 2. Itemized Rating Scales

- In an itemized rating scale, the respondents are provided with a scale that has a number of brief description associated with each category.
- The categories are ordered in terms of scale position, and the respondents are required to select the specified category that best describes the object being rated.

# Likert Scale

- Named after its developer, Rensis Likert, the Likert scale is a widely used rating scale that requires the respondents to indicate a degree of agreement or disagreement with each of a series of statements about the stimulus objects.
- Typically, each scale item has 5 response categories, ranging from “strongly disagree” to “strongly agree”.
- Illustration: A Likert scale for evaluating attitude towards D Mart in the context of the department store project.

**1. Strongly disagree 2. Disagree 3. Neither agree nor disagree 4. Agree 5. Strongly agree**



1. D mart sells high – quality merchandise. 1 2X 3 4 5
2. D mart has poor in – store service. 1 2X 3 4 5
3. I like to shop at D mart . 1 2 3X 4 5
4. D mart does not offer a good mix of different brands within a product category. 1 2 3 4X 5
5. The credit policies at D mart are terrible. 1 2 3 4X 5
6. D mart is where Indians shops. 1X 2 3 4 5
7. I do not like the advertising done by D mart . 1 2 3 4X 5
8. D mart sells a wide variety of merchandise. 1 2 3 4X 5
9. D mart charges fair prices. 1 2X 3 4 5

Dr. Asita Ghewari

# Likert scale

- The Likert scale has several advantages.
- It is easy to construct and administer.
- Respondents readily understand how to use the scale, making it suitable for mail, telephone or personal interview.
- The major disadvantage of the likert scale is that it takes longer to complete than other itemized scales because respondents have to read each statement.

# AGREEMENT

<ul style="list-style-type: none"> <li>•Strongly Agree</li> <li>•Agree</li> <li>•Undecided</li> <li>•Disagree</li> <li>•Strongly Disagree</li> </ul>	<ul style="list-style-type: none"> <li>•Agree Strongly</li> <li>•Agree Moderately</li> <li>•Agree Slightly</li> <li>•Disagree Slightly</li> <li>•Disagree Moderately</li> <li>•Disagree Strongly</li> </ul>	<ul style="list-style-type: none"> <li>•Agree</li> <li>•Disagree</li> </ul>	<ul style="list-style-type: none"> <li>•Agree</li> <li>•Undecided</li> <li>•Disagree</li> </ul>
<ul style="list-style-type: none"> <li>•Agree Very Strongly</li> <li>•Agree Strongly</li> <li>•Agree</li> <li>•Disagree</li> <li>•Disagree Strongly</li> <li>•Disagree Very Strongly</li> </ul>	<ul style="list-style-type: none"> <li>•Yes</li> <li>•No</li> </ul>	<ul style="list-style-type: none"> <li>•Completely Agree</li> <li>•Mostly Agree</li> <li>•Slightly Agree</li> <li>•Slightly Disagree</li> <li>•Mostly Disagree</li> <li>•Completely Disagree</li> </ul>	<ul style="list-style-type: none"> <li>•Disagree Strongly</li> <li>•Disagree</li> <li>•Tend to Disagree</li> <li>•Tend to Agree</li> <li>•Agree</li> <li>•Agree Strongly</li> </ul>

# FREQUENCY

<ul style="list-style-type: none"> <li>•Very Frequently</li> <li>•Frequently</li> <li>•Occasionally</li> <li>•Rarely</li> <li>•Very Rarely</li> <li>•Never</li> </ul>	<ul style="list-style-type: none"> <li>•Always</li> <li>•Very Frequently</li> <li>•Occasionally</li> <li>•Rarely</li> <li>•Very Rarely</li> <li>•Never</li> </ul>	<ul style="list-style-type: none"> <li>•Always</li> <li>•Usually</li> <li>•About Half the Time</li> <li>•Seldom</li> <li>•Never</li> </ul>	<ul style="list-style-type: none"> <li>•Almost Always</li> <li>•To a Considerable Degree</li> <li>•Occasionally</li> <li>•Seldom</li> </ul>
<ul style="list-style-type: none"> <li>•A Great Deal</li> <li>•Much</li> <li>•Somewhat</li> <li>•Little</li> <li>•Never</li> </ul>	<ul style="list-style-type: none"> <li>•Often</li> <li>•Sometimes</li> <li>•Seldom</li> <li>•Never</li> </ul>	<ul style="list-style-type: none"> <li>•Always</li> <li>•Very Often</li> <li>•Sometimes</li> <li>•Rarely</li> <li>•Never</li> </ul>	

IMPORTANCE



<ul style="list-style-type: none"><li>•Very Important</li><li>•Important</li><li>•Moderately Important</li><li>•Of Little Importance</li><li>•Unimportant</li></ul>	<ul style="list-style-type: none"><li>•Very Important</li><li>•Moderately Important</li><li>•Unimportant</li></ul>
---	--

QUALITY

<ul style="list-style-type: none"><li>•Very Good</li><li>•Good</li><li>•Barely Acceptable</li><li>•Poor</li><li>•Very Poor</li></ul>	<ul style="list-style-type: none"><li>•Extremely Poor</li><li>•Below Average</li><li>•Average</li><li>•Above Average</li><li>•Excellent</li></ul>	<ul style="list-style-type: none"><li>•Good</li><li>•Fair</li><li>•Poor</li></ul>
--	---	---

Dr. Asita Ghewari



# LIKELIHOOD

<ul style="list-style-type: none"> <li>•Like Me</li> <li>•Unlike Me</li> </ul>	<ul style="list-style-type: none"> <li>•To a Great Extent</li> <li>•Somewhat</li> <li>•Very Little</li> <li>•Not at All</li> </ul>	<ul style="list-style-type: none"> <li>•True</li> <li>•False</li> </ul>
<ul style="list-style-type: none"> <li>•Definitely</li> <li>•Very Probably</li> <li>•Probably</li> <li>•Possibly</li> <li>•Probably Not</li> <li>•Very Probably Not</li> </ul>	<ul style="list-style-type: none"> <li>•Almost Always True</li> <li>•Usually True</li> <li>•Often True</li> <li>•Occasionally True</li> <li>•Sometimes But Infrequently True</li> <li>•Usually Not True</li> <li>•Almost Never True</li> </ul>	<ul style="list-style-type: none"> <li>•True of Myself</li> <li>•Mostly True of Myself</li> <li>•About Halfway True of Myself</li> <li>•Slightly True Of Myself</li> <li>•Not at All True of Myself</li> </ul>

# Semantic Differential Scale

- The semantic differential is a 7 – point rating scale with end points associated with bipolar labels that have semantic meaning.
- In a typical application, respondents rate objects on a number of itemized, 7 – point rating scales bounded at each end by one of the two bipolar adjectives, such as “cold” and “warm”.
- Illustration: Evaluation of sears on five attributes.

# Example

## Instructions:

This part of the study measures what certain department stores mean to you by having you judge them on a series of descriptive scales bounded at each end by one of two bipolar adjectives. Please mark (X) the blank that best indicates how accurately one or the other adjective describes what the store means to you. Please be sure to mark every scale; do not omit any scale.

# Form

Rate you Manager :

Strong \_:\_:\_:\_X:\_:\_: Weak

Unreliable \_:\_:\_:\_X:\_: Reliable

Modern \_:\_:\_:\_:\_X: Old-fashioned

Fair \_:\_:\_:\_X:\_: Unfair

Cheerful \_X:\_:\_:\_:\_: Un cheerful

# Stapel Scale

- The stapel scale, named after its developer, Jan Stapel, is a unipolar rating scale with 10 categories numbered from -5 to +5, without a neutral point zero.
- This scale is usually presented vertically.
- Respondents are asked to indicate how accurately or inaccurately each term describes the object by selecting an appropriate numerical response category.
- The higher the number, the more accurately the term describes the object.
- In this example, sears is evaluated as not having high quality and having somewhat poor service.

# Stapel Scale

## Instructions:

- Please evaluate how accurately each word or phrase describes each of the department stores.
- Select a plus number for the phrases you think describe the store accurately.
- The more accurately you think the phrase describes the store, the larger the plus number you should choose.
- You should select a minus number for phrases you think do not describe it accurately.
- The less accurately you think the phrase describes the store, the larger the minus number you should choose.
- You can select any number, from +5 for phrases you think are very accurate, to -5 for phrases you think are very inaccurate.

Dr. Asita Ghewari

# Stapel Scales

- Modern versions of the Stapel scale place a **single adjective** as a substitute for the semantic differential when it is difficult to create pairs of bipolar adjectives.
- The advantage and disadvantages of a Stapel scale, as well as the results, are very similar to those for a semantic differential.
- However, the Stapel scale tends to be easier to conduct and administer.
- Measuring a Store's Image

Department  
Store Name

+5

+4

+3

+2

+1

-1

-2

-3

-4

-5

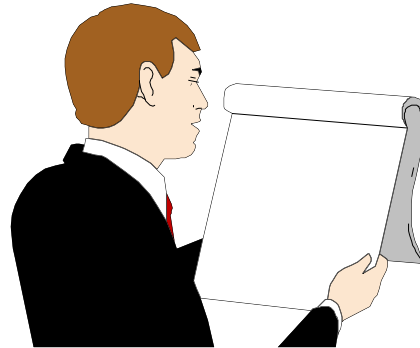
Wide Selection

# Some Basic Considerations When Selecting a Scale



**Selecting a Rating, Ranking,  
Sorting, or Purchase Intent Scale**

**Number of Categories**



**Odd or Even Number of Scale  
Categories**

**Forced Versus Non-forced  
Choice**

**Balanced Versus Non-  
balanced Alternatives**

Dr. Asita Ghewari



# Odd versus even

if neutral responses likely, use odd number

## Odd

Strongly Agree \_\_\_\_\_

Agree \_\_\_\_\_

Neutral \_\_\_\_\_

Disagree \_\_\_\_\_

Strongly disagree \_\_\_\_\_

## Even

Strongly Agree \_\_\_\_\_

Agree \_\_\_\_\_

Disagree \_\_\_\_\_

Strongly disagree \_\_\_\_\_

# Balanced and Unbalanced Scales



## Balanced Scale

*JOVAN MUSK FOR MEN IS*

**Extremely good**

**Very good**

**Good**

**Bad**

**Very bad**

**Extremely bad**

## Unbalanced Scale

*JOVAN MUSK FOR MEN IS*

**Extremely good**

**Very good**

**Somewhat Good**

**Good**

**Bad**

**Very bad**

# Forced vs. Unforced

## Forced

Extremely Reliable

\_\_\_\_\_

Very Reliable

\_\_\_\_\_

Somewhat Reliable

\_\_\_\_\_

Somewhat Unreliable

\_\_\_\_\_

Very Unreliable

\_\_\_\_\_

Extremely Unreliable

\_\_\_\_\_

## Unforced

Extremely Reliable

\_\_\_\_\_

Very Reliable

\_\_\_\_\_

Somewhat Reliable

\_\_\_\_\_

Somewhat Unreliable

\_\_\_\_\_

Very Unreliable

\_\_\_\_\_

Extremely Unreliable

\_\_\_\_\_

Don't know

\_\_\_\_\_

# Labeled vs. End Anchored

## Labeled

Excellent

\_\_\_\_\_

Very Good

\_\_\_\_\_

Fair

\_\_\_\_\_

Poor

\_\_\_\_\_

Very Poor

\_\_\_\_\_

## End Anchored

Excellent

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Poor

\_\_\_\_\_

# Number of Scale Points

## How good D Mart Departmental store

- ☐ Very bad
- ☐ Bad
- ☐ Neither good nor bad
- ☐ Good
- ☐ Very good

- ☐ Very bad
- ☐ Somewhat bad
- ☐ A little bad
- ☐ Neither good nor bad
- ☐ A little good
- ☐ Somewhat good
- ☐ Very good

# Characteristics of Good Measurement Scales



## 1. Reliability

- The degree to which a measure accurately captures an individual's true outcome without error; Accuracy
- synonymous with repetitive consistency

## 2. Validity

- The degree to which a measure faithfully represents the underlying concept; Fidelity

## 3. Sensitivity

- The ability to discriminate meaningful differences between attitudes. The more categories the more sensitive (not less reliable)

## 4. Generalizability

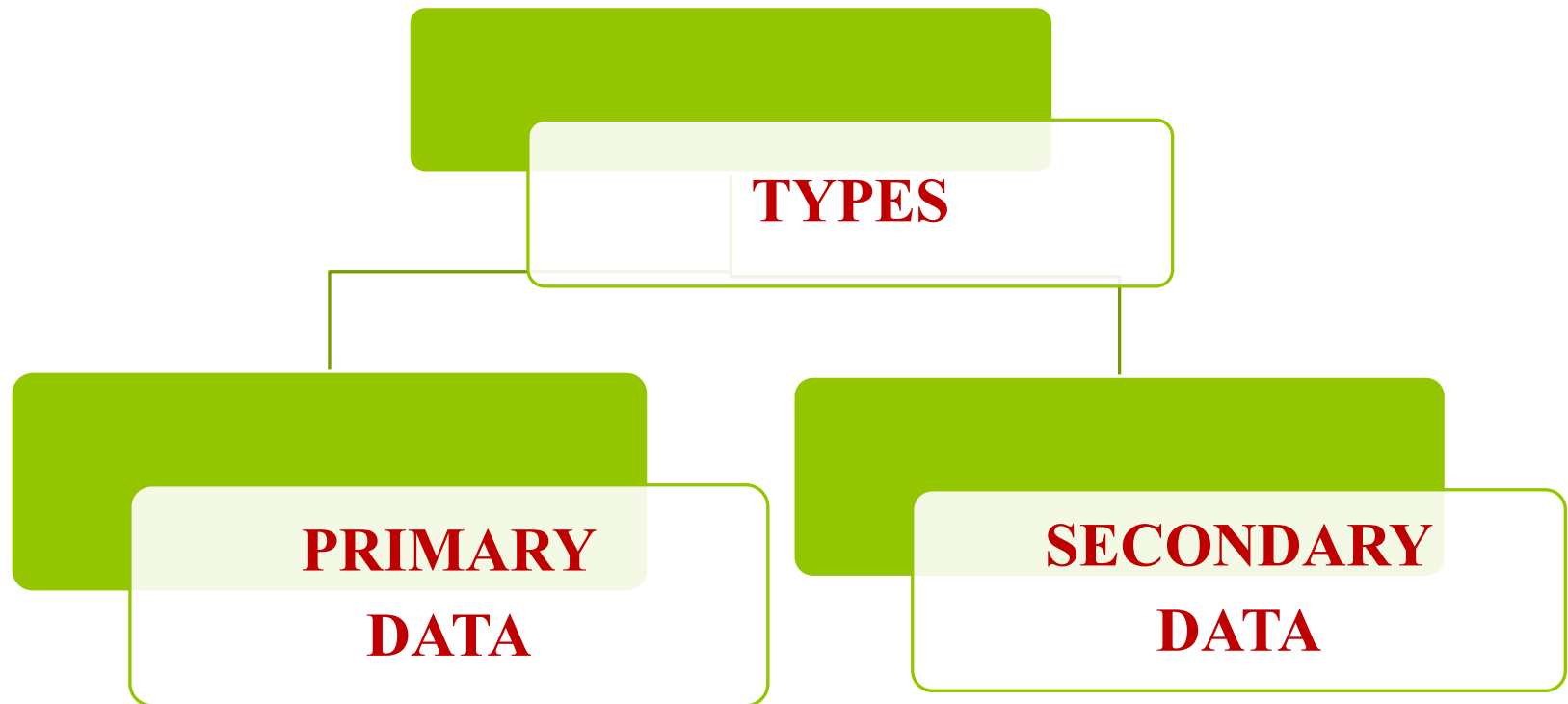
- How easy is scale to administer and interpret

Dr. Asita Ghewari

# Data

- The term data refers to any kind of information researchers obtain on the subjects, respondents or participants of the study. In research, data are collected and used to answer the research questions or objectives of the study.

# CLASSIFICATION OF DATA





# Types of Data

- **Primary data** are originated by a researcher for the specific purpose of addressing the problem at hand. The collection of primary data involves all six steps of the research process
- **Secondary data** are data which have already been collected for purposes other than the problem at hand. These data can be located quickly and inexpensively.

# A Comparison of Primary & Secondary Data

	Primary Data	Secondary Data
Collection purpose	For the problem at hand	For other problems
Collection process	Very involved	Rapid & easy
Collection cost	High	Relatively low
Collection time	Long	Short

Secondary Sources

Internal Corporate Information

Government Agencies

Trade and Industry Associations

Marketing Research Firms

Commercial Publications

News Media

# Internal Sources of Secondary Data



- Accounting system
- Routine company documents
- Salespeople's call reports
- Customer complaints
- Customer service records

# Criteria for Evaluating Secondary Data

- Specifications:
- Error:
- Currency:
- Objective(s):
- Nature:
- Dependability

# Criteria for Evaluating Secondary Data

- **Specifications:** Methodology Used to Collect the Data
- **Error:** Accuracy of the Data
- **Currency:** When the Data Were Collected
- **Objective(s):** The Purpose for Which the Data Were Collected
- **Nature:** The Content of the Data
- **Dependability:** Overall, How Dependable Are the Data

# Advantages and Disadvantages of Secondary Data



## Advantages

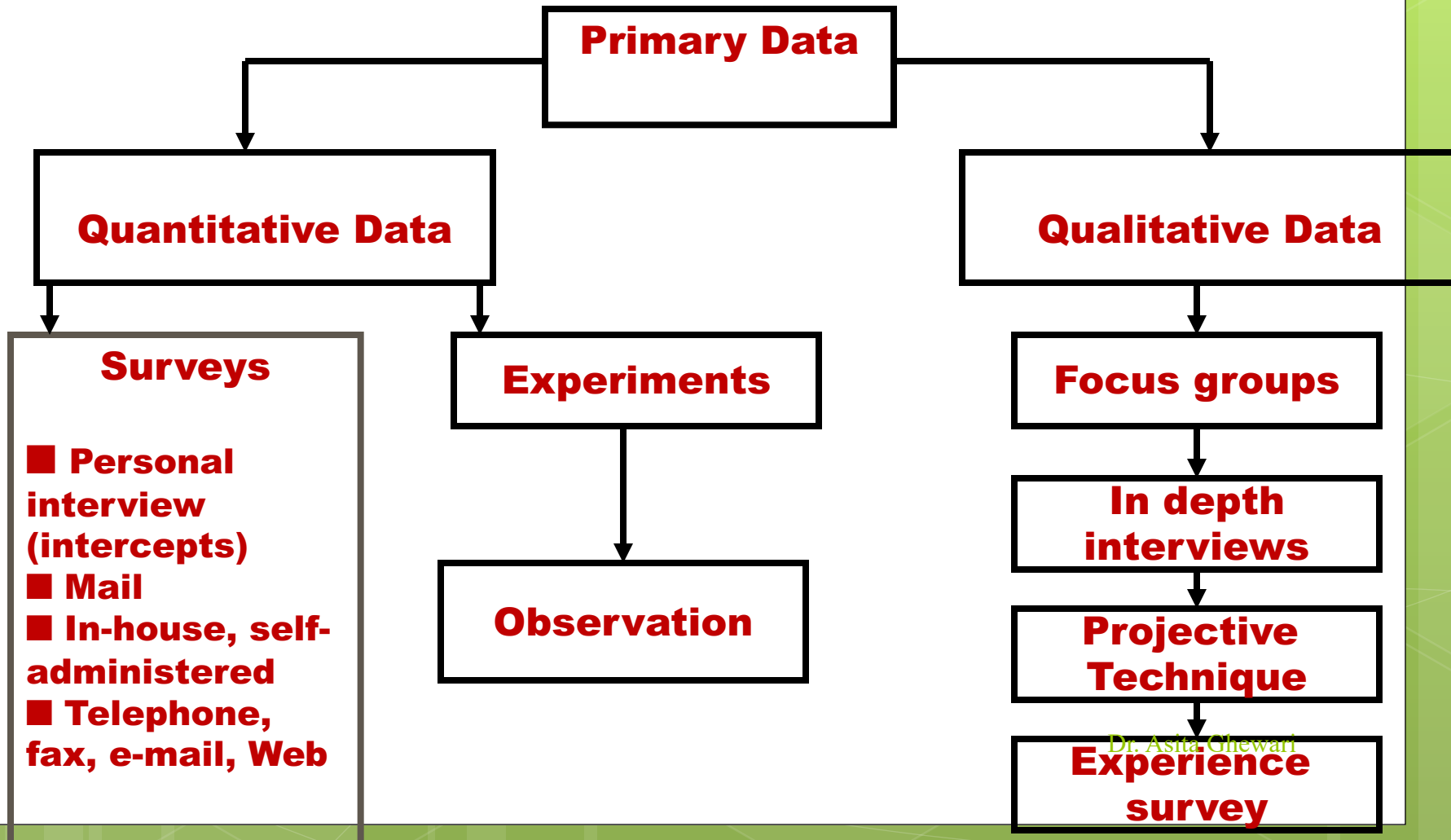
1. Easy availability – Secondary is easily available from a wide variety of sources
2. Faster to collect – Compared to primary data, secondary data can generally be collected faster → suitable for projects with time constraints
3. Cheaper to collect – Secondary data is generally cheaper to collect than primary data

## Disadvantages

1. Currency - Available information may be outdated and not suitable for our needs
2. Not specific to researcher's need
3. Incomplete information
4. Lack of Control – Researcher has no control over accuracy of the data → Need for cross-checks and cross verification

Dr. Asma Ghewari

# Primary Data Methods





# Advantages & Disadvantages of Primary Data



## Advantages

- ☐ Targeted Issues are addressed
- ☐ Data interpretation is better
- ☐ Efficient Spending for Information
- ☐ Decency of Data
- ☐ Proprietary Issues
- ☐ Addresses Specific Research Issues
- ☐ Greater Control

## Disadvantages

- ☐ High Cost
- ☐ Time Consuming
- ☐ Inaccurate Feed-backs
- ☐ More number of resources is required

Dr. Asita Ghewari

# Types of Surveys

## I. Personal Interview

- Door-To-Door Interviews
- Mall Intercepts
- Executive Interviews for Business Markets

## II. Telephone Interview

- Central Location Interviews
- Computer Assisted Telephone Interviewing (CATI)

## III Self-Administered or “Paper & Pencil” Survey

- Mail (One-Shot or Panel)
- Personally Administered
- Diskette / CD
- FAX

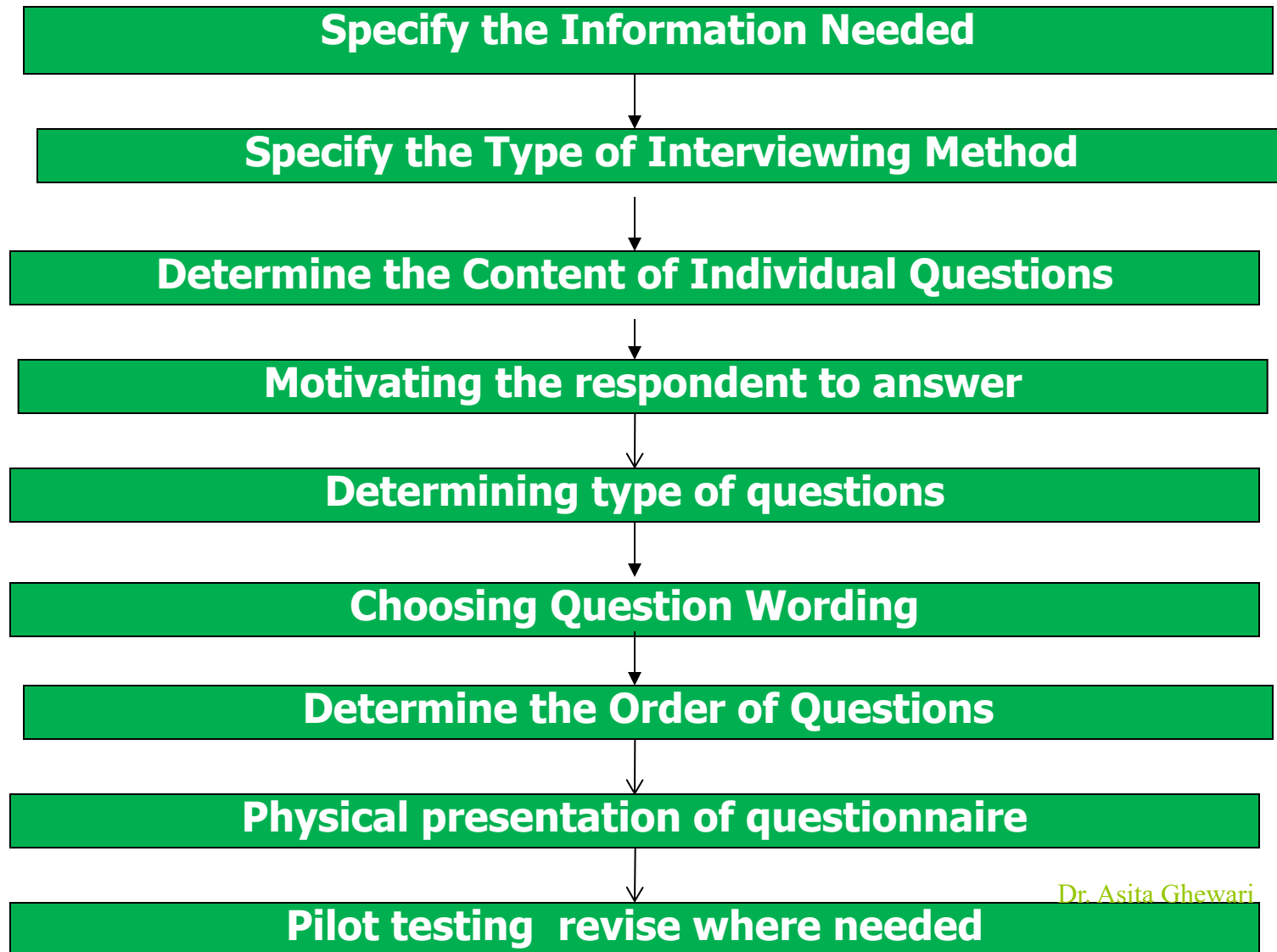
## IV. Electronic Survey

- E-Mail / Internet
- Computer-administered

## COMPARISON OF SURVEY TECHNIQUES

<b>Basis of Comparison</b>	<b>Mail, Internet, &amp; Fax Surveys</b>	<b>Telephone Surveys</b>	<b>Personal And Mail Intercept Surveys</b>
<b>Cost per completed survey</b>	Usually the least expensive, assuming adequate return rate	Moderately expensive, assuming reasonable completion rate	Most expensive because of interviewer's time and travel expenses
<b>Ability to probe and ask complex questions</b>	Little, since self-administered format must be short and simple	Some, since interviewer can probe and elaborate on questions	Much, since interviewer can show visuals, probe, establish rapport
<b>Opportunity for interviewer to bias results</b>	None, since form is completed without interviewer	Some, because of voice inflection of interviewer	Significant, because of voice and facial expressions of interviewer
<b>Anonymity given respondent</b>	Complete, since no signature is needed	Some, because of telephone contact	Dr. Asita Ghewari Little, because of face-to-face contact

# Questionnaire Construction



Dr. Asita Ghewari

## TYPICAL PROBLEMS IN WORDING QUESTIONS

### PROBLEM

### SAMPLE QUESTION

### EXPLANATION OF PROBLEM

Leading question

Why do you like Wendy's fresh meat hamburgers better than those of competitors?

Consumer is led to make statement favoring Wendy's hamburgers.

Ambiguous question

Do you eat at fast-food restaurants regularly?  
☐ Yes ☐ No

What is meant by word *regularly*—once a day, once a month, or what?

Unanswerable question

What was the occasion for eating your first hamburger?

Who can remember the answer? Does it matter?

Two questions in one

Do you eat Wendy's hamburgers and chili?  
☐ Yes ☐ No

How do you answer if you eat Wendy's hamburgers but not chili?

Nonexhaustive question

Where do you live?  
☐ At home  
☐ In dormitory

What do you check if you live in an apartment?

Nonmutually exclusive answers

What is your age?  
☐ Under 20 ☐ 20–40  
☐ 40 and over

Dr. Asita Ghewari  
What answer does a 40-year-old check?



# Sample questions from Wendy's survey



1. What things are most important to you when you decide to eat out and go to a fast-food restaurant?

---



---



---

2. Have you eaten at a fast-food restaurant in the past month?

☐ Yes ☐ No

3. If you answered yes to question 2, how often do you eat fast food?

☐ Once a week ☐ 2 to 3 times a month ☐ Once a month or less

4. How important is it to you that a fast-food restaurant satisfies you on the following characteristics? [Check the box that describes your feelings for each item listed]

CHARACTERISTIC	VERY IMPORTANT	SOMEWHAT IMPORTANT	IMPORTANT	UNIMPORTANT	SOMEWHAT UNIMPORTANT	VERY UNIMPORTANT
• Taste of food	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Cleanliness	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Price	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Variety of menu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5. For each of the characteristics listed below, check the space on the scale that describes how you feel about Wendy's. Mark an X on only **one** of the five spaces listed for each item listed.

CHARACTERISTIC		CHECK THE SPACE THAT DESCRIBES THE DEGREE TO WHICH WENDY'S IS . . .					
• Taste of food	Tasty	_____	_____	_____	_____	_____	Not tasty
• Cleanliness	Clean	_____	_____	_____	_____	_____	Dirty
• Price	Inexpensive	_____	_____	_____	_____	_____	Expensive
• Variety of menu	Broad	_____	_____	_____	_____	_____	Narrow

Dr. Asita Ghewari

# Sample questions from Wendy's survey



6. Check one box that describes your agreement or disagreement with each statement listed below:

STATEMENT	STRONGLY AGREE	AGREE	DON'T KNOW	DISAGREE	STRONGLY DISAGREE
• Adults like to take their families to fast-food restaurants	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Our children have a say in where the family chooses to eat	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7. How important are each of the following sources of information to you when selecting a fast-food restaurant to eat at? [Check one box for each source listed]

SOURCE OF INFORMATION	VERY IMPORTANT	SOMEWHAT IMPORTANT	NOT AT ALL IMPORTANT
• Television	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Newspapers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Radio	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Billboards	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Flyers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8. How often do you eat out at each of the following fast-food restaurants? [Check one box for each source listed]

RESTAURANT	ONCE A WEEK OR MORE	2 TO 3 TIMES A MONTH	ONCE A MONTH OR LESS
• Burger King	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• McDonald's	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
• Wendy's	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9. Please answer the following questions about you and your household. [Check only one for each item]

- What is your gender? ☐ Male ☐ Female
- What is your marital status? ☐ Single ☐ Married ☐ Other (widowed, divorced, etc.)
- How many children under age 18 live in your home? ☐ 0 ☐ 1 ☐ 2 ☐ 3 or more
- What is your age? ☐ Under 25 ☐ 25-44 ☐ 45 or older
- What is your total annual individual or household income?  
☐ <\$15,000 ☐ \$15,000-49,000 ☐ \$50,000 or more

Dr. Asita Ghewari

# Sampling

## Unit 4

Dr. Asita Ghewari



# Sample

- ◉ Subset of a larger population

# Sample

- Units (typically people) selected from the **population of interest** for purposes being able to draw conclusions about that population of interest.
- We use the “sample” statistics as a representation of the true “population” parameters (value).

# Sampling

- The process of obtaining information from a subset (sample) of a larger group (population)
- The results for the sample are then used to make estimates of the larger group
- Faster and cheaper than asking the entire population
- Two keys
  1. Selecting the right people
    - Have to be selected scientifically so that they are representative of the population
  2. Selecting the right number of the right people
    - To minimize sampling errors I.e. choosing the wrong people by chance

# Population

- Any Complete Group

# Census

- Investigation of all individual elements that make up a population

# Census Vs. Sample

- ◉ Sampling less expensive
- ◉ Sampling less time
- ◉ Greater accuracy of results
- ◉ Availability of population elements

# When Would One Conduct a Census

- ◉ Feasible (due to size of pop.)
- ◉ Necessary (due to differences in the population)

# Characteristics of Good Samples

- ◉ Representative
- ◉ Accessible
- ◉ Low cost



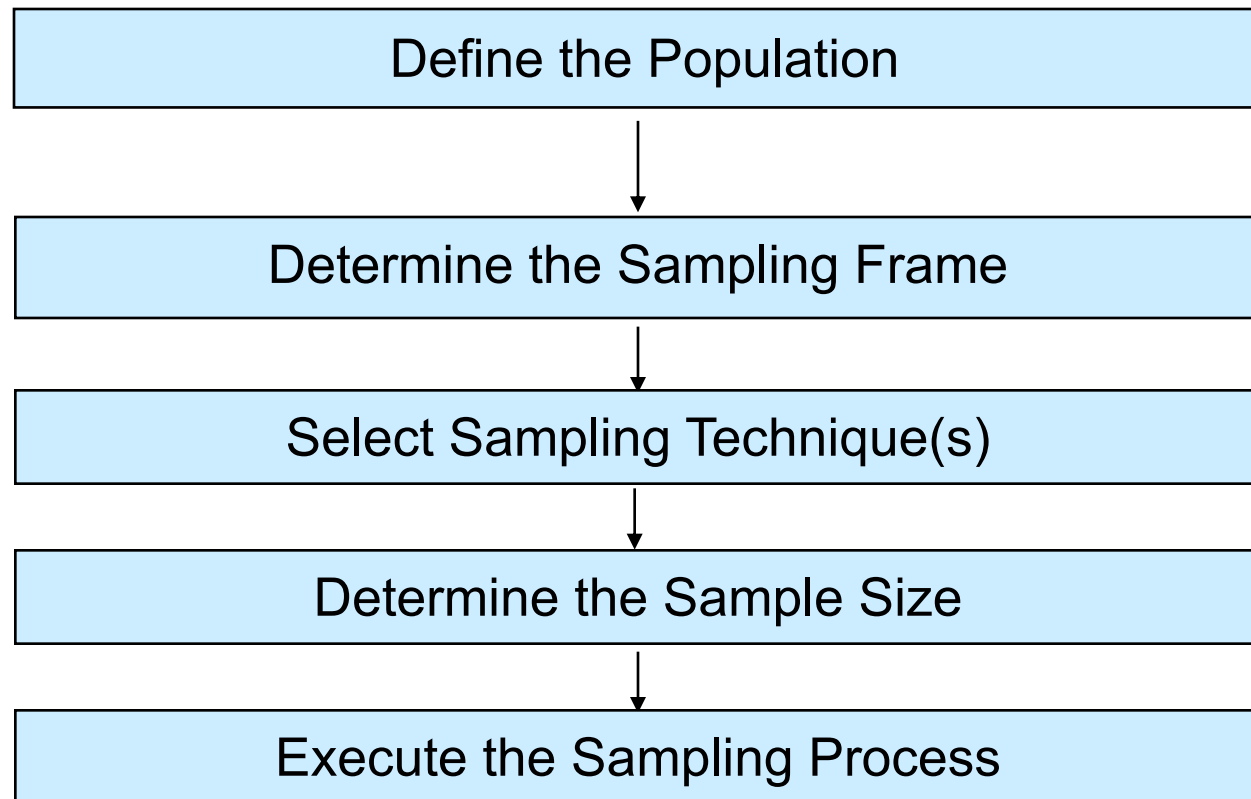
# Sample vs. Census

Type of Study

Conditions Favoring the Use of  
Sample                      Census

<b>1. Budget</b>	<b>Small</b>	<b>Large</b>
<b>2. Time available</b>	<b>Short</b>	<b>Long</b>
<b>3. Population size</b>	<b>Large</b>	<b>Small</b>
<b>4. Variance in the characteristic</b>	<b>Small</b>	<b>Large</b>
<b>5. Cost of Sampling errors</b>	Low	High
<b>6. Cost of non sampling error</b>	High	Low
<b>5. Nature of measurement</b>	<b>Destructive</b>	<b>Nondestructive</b>
<b>6. Attention to individual cases</b>	<b>Yes</b>	<b>No</b>

# The Sampling Design Process



# Define the Target Population

The target population is the collection of elements or objects that possess the information sought by the researcher and about which inferences are to be made. The target population should be defined in terms of elements, sampling units, extent, and time.

- An **element** is the object about which or from which the information is desired, e.g., the respondent.
- A **sampling unit** is an element, or a unit containing the element, that is available for selection at some stage of the sampling process.
- **Extent** refers to the geographical boundaries.
- **Time** is the time period under consideration.

# Sampling Frame

- The list of elements from which the sample may be drawn.
- Sometimes called working population
- **Problems**
  - omissions
  - ineligibles
  - duplications

# Two Major Categories of Sampling

- Probability Sampling
  - Known, nonzero probability for every element
- Nonprobability Sampling
  - Probability of selecting any particular member is unknown

# Determine the sample size...

- The size of the sample influences both the representativeness of the sample and the statistical analysis of the data
  - ... larger samples are more likely to detect a difference between different groups
  - ... smaller samples are more likely not to be representative

# Rules of thumb for determining the sample size...

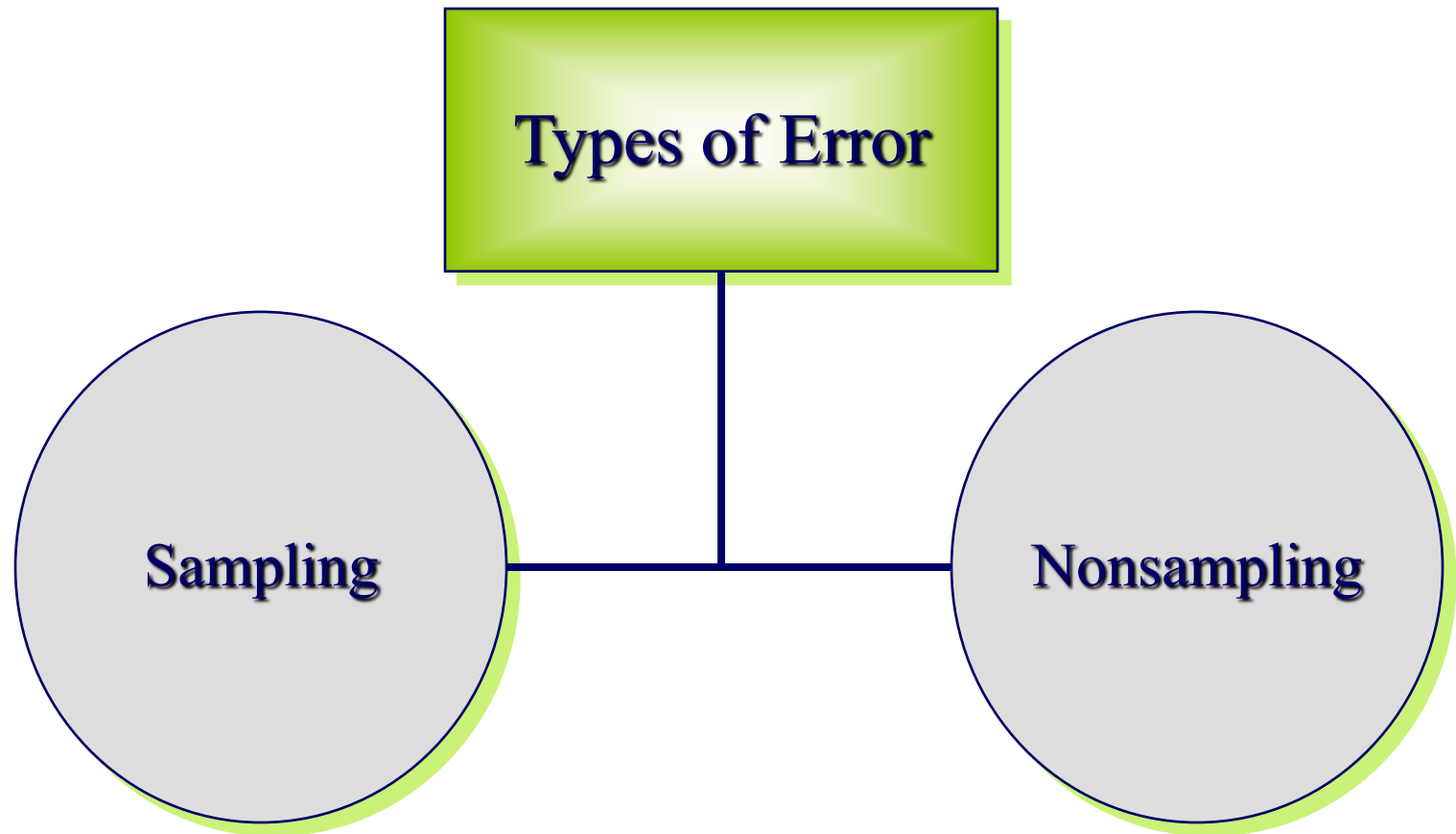
1. The larger the population size, the smaller the percentage of the population required to get a representative sample
2. For smaller samples ( $N < 100$ ), there is little point in sampling. Survey the entire population.
3. If the population size is around 500 (give or take 100), 50% should be sampled.
4. If the population size is around 1500, 20% should be sampled.
5. Beyond a certain point ( $N = 5000$ ), the population size is almost irrelevant and a sample size of 400 may be adequate.

# Sample Size

Important qualitative factors in determining the sample size

- the importance of the decision
- the nature of the research
- the nature of the analysis
- sample sizes used in similar studies
- resource constraints





# Sampling Error

- The difference between the sample results and the result of a census conducted using identical procedures
- Statistical fluctuation due to chance variations

# Minimizing Sampling Errors

- Increase the sample size
- Use a statistically efficient sampling plan
- Make the sample as representative of the population as possible

# Nonsampling Error

... bias that occurs in a research study regardless of whether a sample or census is used; e.g., bias caused by measurement errors, response errors, coding errors, etc.

# Types of non-sampling errors

Non-sampling errors can be of various types

- Coverage (or Frame) errors
- Non-response errors
- Measurement errors
- Data handling errors

Note that the first more often applies to sample surveys, while the last three apply to both surveys and censuses.

# Coverage (frame) errors

In surveys, the sample is selected from a list, i.e. a **sampling frame**, of all population members.

An inadequate frame leads to coverage errors. Often can have either

- **under-coverage** (missing elements), or
- **over-coverage** (duplicates)

Both lead to biased results.

# Non-response errors

Non-response errors are all errors arising from:

- ◉ **Unit non-response**, i.e. failure to obtain information from a pre-chosen sampling or population unit unit
- ◉ **Item non-response**, i.e. failure to get a response to a specific question or item in the data recording form.

# Measurement Errors

Measurement errors arise when the **recorded** response differs from the **true** value.

They can occur for a variety of reasons, e.g.

- by **respondent** (e.g. heads of households) giving an incorrect answer
- because of **instrument** or question error
- by **interviewer** error.

Further, errors may be greater for some sub-groups of the population, e.g. those less literate, or those unwilling to co-operate.



# Data handling errors

Data handling errors can occur from the stage of data collection up to the final stages of data analysis. Types of errors that can arise include:-

- errors in transmission of data from the field to the office
- errors in preparing the data in a suitable format for computerisation, e.g. during coding of qualitative answers
- errors in computerisation of the data
- errors during data analysis, e.g. imputation and weighting.

# Two Major Categories of Sampling

- Probability Sampling
  - Known, nonzero probability for every element
- Nonprobability Sampling
  - Probability of selecting any particular member is unknown

# Probability Sampling

- ◉ Simple Random Sample
- ◉ Systematic Sample
- ◉ Stratified Sample
- ◉ Cluster Sample
- ◉ Multistage Area Sample

# Simple Random Sampling

- A sampling procedure that ensures that each element in the population will have an equal chance of being included in the sample

## *advantages...*

- ...easy to conduct

- ...strategy requires minimum knowledge of the population to be sampled

## *disadvantages...*

- ...need names of all population members
- ...may over- represent or under- estimate sample members
- ...there is difficulty in reaching all selected in the sample

# Systematic Sampling

- A simple process
- Every  $n$ th name from the list will be drawn

# Stratified Sampling

- ◉ Probability sample
- ◉ Subsamples are drawn within different strata
- ◉ Each stratum is more or less equal on some characteristic
- ◉ Do not confuse with quota sample



# Cluster Sampling

- The purpose of cluster sampling is to sample economically while retaining the characteristics of a probability sample.
- The primary sampling unit is no longer the individual element in the population
- The primary sampling unit is a larger cluster of elements located in proximity to one another

# Nonprobability Sampling

- ◉ Convenience
- ◉ Judgment
- ◉ Quota
- ◉ Snowball

# Convenience Sampling

- Also called haphazard or accidental sampling
- The sampling procedure of obtaining the people or units that are most conveniently available

# Judgment Sampling

- ◉ Also called purposive sampling
- ◉ An experienced individual selects the sample based on his or her judgment about some appropriate characteristics required of the sample member

# Quota Sampling

- Ensures that the various subgroups in a population are represented on pertinent sample characteristics
- To the exact extent that the investigators desire
- It should not be confused with stratified sampling.

# Snowball Sampling

- A variety of procedures
- Initial respondents are selected by probability methods
- Additional respondents are obtained from information provided by the initial respondents

# What is the Appropriate Sample Design?

- ◉ Degree of Accuracy
- ◉ Resources
- ◉ Time
- ◉ Variability of the Population
- ◉ Statistical Consideration

# After the Sample Design is Selected

- ◉ Determine sample size
- ◉ Select actual sample units
- ◉ Conduct fieldwork



# Choosing Non probability vs. Probability Sampling

## Conditions Favoring the Use of

**Nonprobability sampling**

**Probability sampling**

**Nature of research**

**Exploratory**

**Conclusive**

**Relative magnitude of sampling and nonsampling errors**

**Nonsampling errors are larger**

**Sampling errors are larger**

**Variability in the population**

**Homogeneous (low)**

**Heterogeneous (high)**

**Statistical considerations**

**Unfavorable**

**Favorable**

**Operational considerations**

**Favorable**

**Unfavorable**

# **Data Analysis and Report Writing**

## **Unit 5**

Dr. Asita Ghewari

---

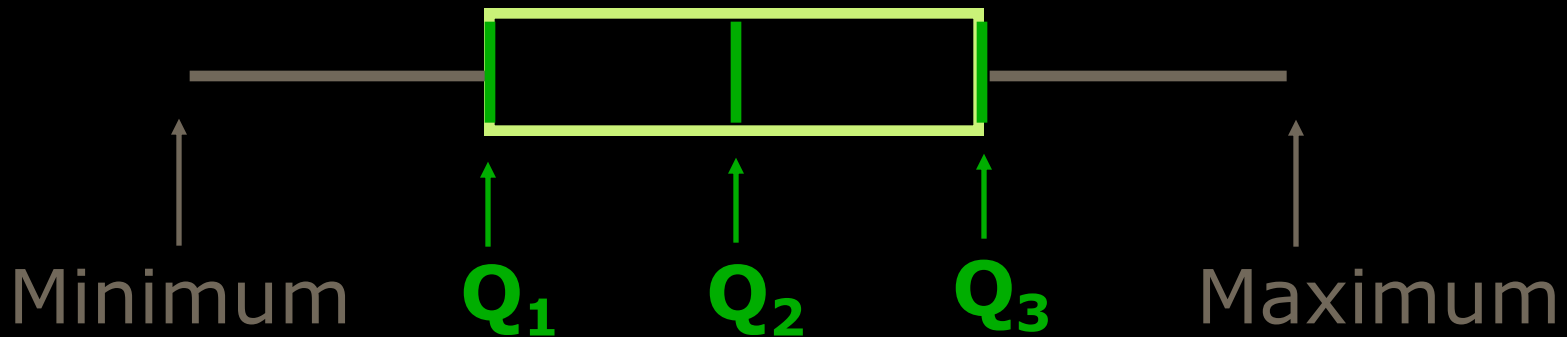
# Basic Analytical Tools

- Using
  - MS-EXCEL
  - MINITAb
  - SPSS
  - Jamovi

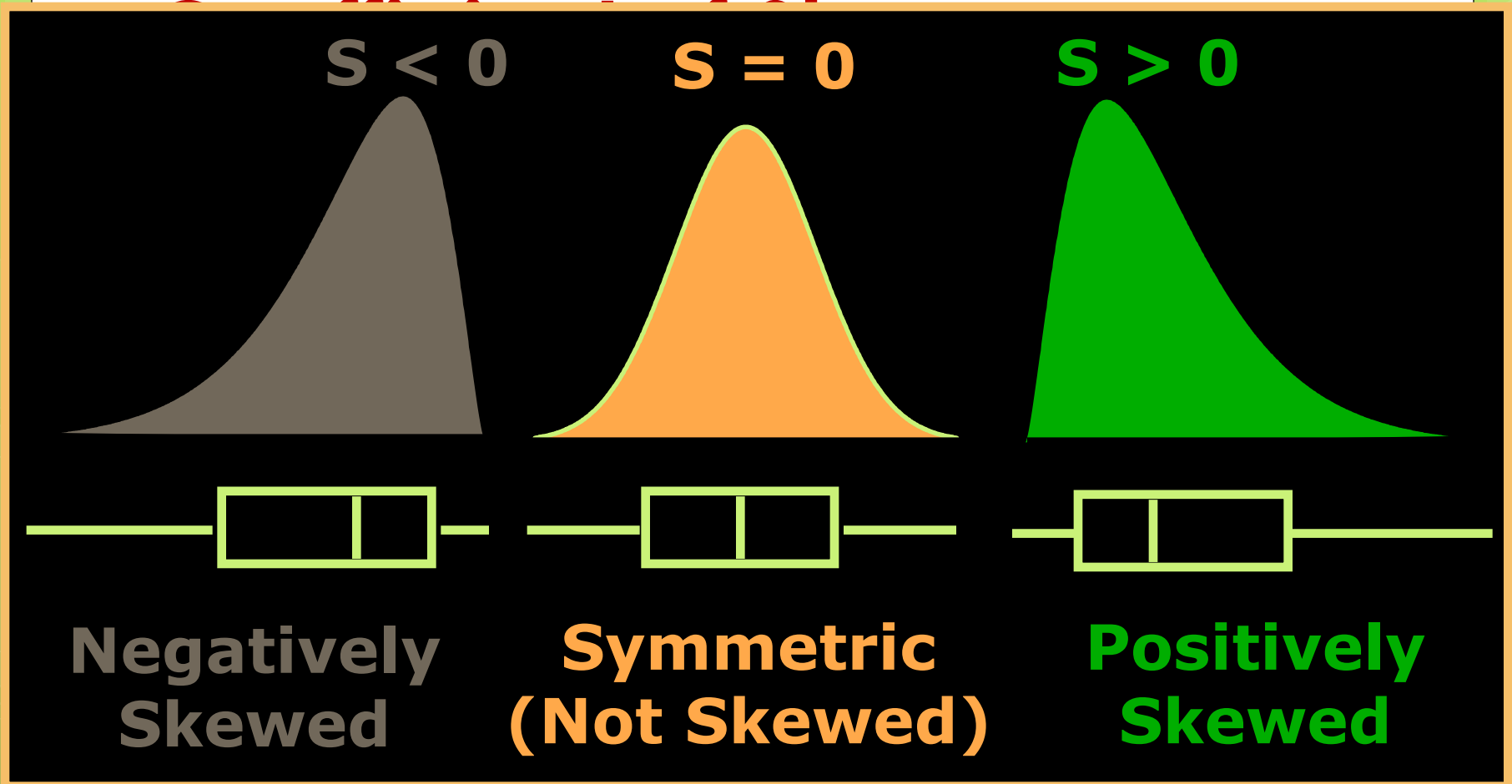
# Box and Whisker Plot

- Five specific values are used:
  - Median,  $Q_2$
  - First quartile,  $Q_1$
  - Third quartile,  $Q_3$
  - Minimum value in the data set
  - Maximum value in the data set

# Box and Whisker Plot



# Skewness: Box and Whisker Plots, and



- The box plot of a negatively skewed distribution has the median off-centre and generally to the right. The right-hand whisker will be short, while the left-hand whisker will be long, reflecting the gradual tailing off of data values to the left
- The box plot of a positively skewed distribution has the median off-centre and generally to the left. The left-hand whisker will be short, while the right-hand whisker will be long, reflecting the gradual tailing off of data values to the right.

# Stem-and-leaf:

Listed in Table is the number of 30-second radio advertising spots purchased by each of the 45 members of the Greater Buffalo Automobile Dealers Association last year. Organize the data into a stem-and-leaf display. Around what values do the number of advertising spots tend to cluster? What is the fewest number of spots purchased by a dealer? The largest number purchased?

**TABLE 4–1** Number of Advertising Spots Purchased by Members of the Greater Buffalo Automobile Dealers Association

96	93	88	117	127	95	113	96	108	94	148	156
139	142	94	107	125	155	155	103	112	127	117	120
112	135	132	111	125	104	106	139	134	119	97	89
118	136	125	143	120	103	113	124	138	Dr. Asita Ghewari		



# Stem and leaf: Example

Stem	Leaf
8	8 9
9	6 3 5 6 4 4 7
10	8 7 3 4 6 3
11	7 3 2 7 2 1 9 8 3
12	7 5 7 0 5 5 0 4
13	9 5 2 9 4 6 8
14	8 2 3
15	6 5 5

Dr. Asita Ghewari

# **DATA PROCESSING**

- **One of the important stages of the research process is data processing and analysis.**
- **Once data are collected the researcher turns his focus of attention on its processing.**

# DATA PROCESSING .....

- **Data processing refers to :**
  - **Editing,**
  - **Coding,**
  - **Recoding**
  - **Computing of the scores,**
  - **Preparation of master charts.**

# Editing of Data

- **Missing entries**
- **Illegible entries**
- **Discrepancies**

# Editing of .....

- **Contingency Responses :**

**Marital Status :**

**Married/Unmarried**

**Number of Children :**

**One/Two/Not Applicable**

# TYPES OF DATA

- ◉ NOMINAL
- ◉ ORDINAL
- ◉ INTERVAL/ RATIO

- **Nominal Scale - groups or classes**
  - ✓ Gender
- **Ordinal Scale - order matters**
  - ✓ Ranks (top ten videos)
- **Interval Scale - difference or distance matters – has arbitrary zero value.**
  - ✓ Temperatures ( $^{\circ}\text{F}$ ,  $^{\circ}\text{C}$ )
- **Ratio Scale - Ratio matters – has a natural zero value.**
  - ✓ Salaries

# UNIVARIATE DATA ANALYSIS

- **UNIVARIATE : One Variable**



# UNIVARIATE DATA ANALYSIS

## ☺ **Univariate Analysis**



### **Description**

**To describe profile and Independent Variables**

## ☺ **Descriptive Statistics**

**Percentages, ratio, proportion**

**Mean, Mode, Median, Standard Deviation**

# NOMINAL VARIABLE



## RELIGION OF RESPONDENTS

Religion	Distribution of Respondents	
	Frequencies	Percentages
Hindu	110	38.6
Muslim	107	37.5
Christian	68	23.9
Total	285	100.0

Dr. Asita Ghewari

# Ordinal Variable

## LEVEL OF EDUCATION

Level of Education	Distribution of Respondents	
	Frequencies	Percentages
Primary	83	29.1
Middle	136	47.7
High School	56	19.6
Graduate	10	3.6
Total	285	100.0

Dr. Asita Ghewari

# Summary Measures: Population

## Parameters      Sample Statistics

- Measures of Central Tendency

- ✓ Median
- ✓ Mode
- ✓ Mean

- Measures of Variability

- ✓ Range
- ✓ Interquartile range
- ✓ Variance
- ✓ Standard Deviation

- Other summary measures:

- ✓ Skewness
- ✓ Kurtosis

- Mean – an average computed by summing the values of several observations and dividing by the number of observations.
- Mode- an average representing the most frequently observed value or attribute.
- Median – an average representing the value of the “middle” case in a rank-ordered set of observations.

# DESCRIPTIVES.....



- You get Output showing the Descriptives Tables
- Descriptive Statistics

	N	Min	Max	Mean	SD
Age	240	18	58	30.34	4.45

Dr. Asita Ghewari



# Permissible Statistic

Scale	Definition	Permissible Statistics
Nominal	Categories, no value assessed. Gender, region, ethnic background.	Percentages, mode, binomial, chi square
Ordinal	Rank. Attitudes, preferences, social classes, age classes.	Percentile, median
Interval	Equality of differences. $Y=a+bx$ . Attitudes, opinions, index numbers.	Range, mean, standard deviation
Ratio	Proportion applies. $Y=cx$ . Zero means none. Age, cost, number purchased, sales.	Geometric mean

# Data Analysis consideration

<b><i>OK to compute....</i></b>	<b><i>Nominal</i></b>	<b><i>Ordinal</i></b>	<b><i>Interval</i></b>	<b><i>Ratio</i></b>
<b><i>frequency distribution</i></b>	<b><i>Yes</i></b>	<b><i>Yes</i></b>	<b><i>Yes</i></b>	<b><i>Yes</i></b>
<b><i>median and percentiles</i></b>	<b><i>No</i></b>	<b><i>Yes</i></b>	<b><i>Yes</i></b>	<b><i>Yes</i></b>
<b><i>sum or difference</i></b>	<b><i>No</i></b>	<b><i>No</i></b>	<b><i>Yes</i></b>	<b><i>Yes</i></b>
<b><i>mean, standard deviation, standard error of the mean</i></b>	<b><i>No</i></b>	<b><i>No</i></b>	<b><i>Yes</i></b>	<b><i>Yes</i></b>
<b><i>ratio, or coefficient of variation</i></b>	<b><i>No</i></b>	<b><i>No</i></b>	<b><i>No</i></b>	<b><i>Yes</i></b>



# BIVARIATE DATA

## APPLICATION OF INFERENCE STATISTICS

Dr. Asita Ghewari

**BIVARIATE ANALYSIS → ASSOCIATION BETWEEN TWO VARIABLES**

**😊 Bivariate Analysis → Inferences**

# BIVARIATE TABLE

## GENDER AND CUSTOMER SATISFACTION

Gender	Customer Satisfaction			TOTAL
	High	Moderate	Low	
Men	30	70	100	200
Women	65	90	45	200
Total	95	160	145	400

Dr. Asita Ghewari

# PERCENTAGING IN BIVARIATE TABLES

- **Row Percentage**

**Row Total (100) is base for calculating row percentage.**

- **Column Percentage**

**Column Total (100) is base for calculating column percentage.**

# ROW PERCENTAGE

**TABLE 4.8**  
**GENDER AND CUSTOMER SATISFACTION**

<b>Gender</b>	<b>Customer Satisfaction</b>			<b>TOTAL</b>
	<b>High</b>	<b>Medium</b>	<b>Low</b>	
<b>Men</b>	<b>30(15.0)</b>	<b>70 (35.0)</b>	<b>100 (50.0)</b>	<b>200 (100.0)</b>
<b>Women</b>	<b>65 (32.5)</b>	<b>90 (45.0)</b>	<b>45 (22.5)</b>	<b>200 (100.0)</b>
<b>Total</b>	<b>95</b>	<b>160</b>	<b>145</b>	<b>400</b>

Dr. Asita Ghewari

# COLUMN PERCENTAGE

**TABLE 4.8**  
**CUSTOMER SATISFACTION BY GENDER**

Gender	Customer Satisfaction			TOTAL
	High	Medium	Low	
Men	30 (31.6)	70 (43.8)	100 (69.0)	200
Women	65 (68.4)	90 (56.2)	45 (31.0)	200
Total	95 (100.0)	160 (100.0)	145 (100.0)	400

Dr. Asita Ghewari

# The Principle of Table Reading

- **Look for the title** : The title describes the information that is contained in the table. In Table 4.8 the title tells about differences in customer satisfaction between men and women respondents.
- **Determine in which direction** the percentages have been computed. In Table 4.8, the percentages have been computed across the rows.

# The Principle of Table Reading

- **Make comparisons:** Comparing the percentage differences in the table is a quick method for assessing the extent of a relationship between the variables. Comparisons are always made in a direction opposite to the one in which the percentages have been computed. If the percentages have been computed across the rows, as in the Table 4.8, then we compare percentages down the column.

Dr. Asita Ghewari



# INTERPRETATION

- **Gender divided into men and women.**
- **Each gender grouping is described in terms of “High”, “Moderate” and “Low” customer satisfaction.**
- **Men and women are compared in terms of customer satisfaction.**

# INTERPRETATION .....

- **Men with High Customer Satisfaction = 30 (15%)**
- **Women with High Customer Satisfaction = 65 (32.5%)**
- **Men with Low Customer Satisfaction = 100 (50%)**
- **Women with Low Customer Satisfaction = 45 (22.5%)**
- **Inference**



Dr. Asita Ghewari

**Women customers are more satisfied with services than men**

## LIMITATIONS OF INTERPRETATION BY PERCENTAGE DIFFERENCE

- Percentage difference, though, is very useful and simple way of assessing association has its own limitations.
- Assessment of association between variables by percentage difference just gives us an understanding of association likely to exist in sample.

## LIMITATIONS OF INTERPRETATION BY PERCENTAGE DIFFERENCE.....

- It does not tell us if the inference drawn will hold good to the population from which the sample is drawn.
- That means the inference may be just a chance.
- Therefore, to know that the inference is not by chance we need to compute some inferential statistical tests.

- **INFERENTIAL STATISTICAL TESTS**
  - **Tests of Association**
    - **Coefficient of correlation ( $r$ )**
    - **Chi-Square ( $\chi^2$ )**
    - **Coefficient of Contingency ( $C$ )**
  - **Tests Of Difference: t-tests**

# Coefficient of Correlation ( $r$ )

## Meaning

- **It ascertains whether a change in one variable is associated with change in another variable.**
- **Whether the higher income of parents is associated with high expectations of children in respect of salaries and the job they will subsequently take up.**

# Coefficient of correlation ( $r$ ) .....

- **Magnitude / Extent / Strength**
- The coefficient of correlation ( $r$ ) changes in value from positive one (+ 1.0) down through zero (0.0) to negative one or unity (- 1.0).
- **Coefficient of Correlation ranging from :**

<b>1 to 0.7</b>		<b>'high'</b>
<b>0.7 to 0.4</b>	→	<b>'substantial'</b>
<b>0.4 to 0.2</b>	→	<b>'low'</b>
<b>below 0.2</b>	→	<b>'negligible'.</b>

## AGE AND CUSTOMER SATISFACTION (SCORES)

S.No.	Age	Customer Satisfaction Scores
1	40	42
2	26	29
3	70	66
4	32	24
5	52	58
6	30	28
7	48	46
8	48	50
9	42	44
10	62	66

Dr. Asita Ghewari



# Chi-square Test

- The  $\chi^2$  (Greek letter  $\chi^2$  and pronounced as Ki-square) test provides us with a method to evaluate whether or not frequencies which have been empirically observed differ significantly from those which would be expected under a certain set of theoretical assumption.

# Bivariate Association between Gender and Customer Satisfaction

**TABLE 4.8**  
**CUSTOMER SATISFACTION BY GENDER**

Gender	Customer Satisfaction			TOTAL
	High	Medium	Low	
Men	30	70	100	200
Women	65	90	45	200
Total	95	160	145	400

Dr. Asita Ghewari

- ◉ **Null Hypothesis: There is no association between 'Gender' and 'Customer Satisfaction'**
- ◉ **Interpretation of Chi-square results**

The table values of chi-square are available at various probability levels.

We can find out from the table the values of chi-square at certain levels of significance.

Usually, the value of chi-square at 0.05 or .01 levels of significance from the given degrees of freedom is seen from the table and is compared with observed value of Chi-square.

From the sample table of chi square test results we have:

Dr. Asita Ghewari

- $\chi^2$  (calculated value) = 22.678
- Level of significance = .05
- Exact significance (p value) = 0.0000
- Degree of freedom (df) = 4
- $\chi^2$  (table value) at (P = .05 and df = 4) = 9.488

## Decision Rules

- If the observed value of  $\chi^2$  is more than the table value we reject the null hypothesis.

## Chi-square .....

- Chi square test results show that the calculated value is more than the table value /critical value , hence the null hypothesis “There is no association between ‘Gender’ and ‘Customer Satisfaction’ is rejected.
- That means research hypothesis is true and hence accepted.
- In other words we can say that the association between ‘Gender’ and ‘Customer Satisfaction’ is statistically significant.

Dr. Asita Ghewari

## Chi-square .....

- Alternatively, since SPSS output provides exact significance, we can modify the decision rule as : “If the exact significance level is more than **.05** we reject the null hypothesis”.
- The test results show that the exact significance is **.0000** which is less than **.05** hence we reject the null hypothesis and accept the research hypothesis.

Dr. Asita Ghewari

# Chi-square .....

- Since the association is statistically significant , now we can say that the inference drawn will hold good to the population from which the sample is drawn.
- Finally, we may infer that the association is due to some reasons and certainly it is not by chance.

## COEFFICIENT OF CONTINGENCY ( C ).....

- While the chi-square measures indicate that there is an association between two variables, they do not indicate the strength of the association.
- As such, if you are interested in explaining the strength of association you have to find out Coefficient of Contingency (C).



## COEFFICIENT OF CONTINGENCY ( C ).....

- This measure indicates the strength of the association between the row and column variables in a crosstable.
- The value of C has been found to be 0.232.
- We can say that the association is 'positive' and 'moderate'.

# Regression Analysis

- A statistical technique for characterising the pattern of a relationship between quantitative variables in terms of a linear equation and for summarising the strength of this relationship in terms of its deviation from that linear pattern.

# Regression Equation

$$Y = C + bX$$

**Y = Dependent Variable.**

**C = Constant**

**b = Regression Coefficient**

**X = Independent Variable**

# EXPENDITURE ON AD AND INCREASE IN NET PROFIT DURING LAST 10 YEARS



Year	Expenditure on AD (‘000)	Net Profit (‘00,000)
1999	26	29
2000	30	34
2001	32	38
2002	40	46
2003	42	47
2004	48	49
2005	48	50
2006	52	58
2007	62	66
2008	70	80

Dr. Asita Ghewari

# Regression Analysis

## - Out Put

### Model Summary

			Adjusted	Std. Error of
1	.988 <sup>a</sup>	.977	.974	2.46724

a.

Dependent Variable: EXPENDITURE ON LAB

# Regression Analysis -Out Put

Coefficients<sup>a</sup>

		Unstandardized		Standardized			
		B	Std. Error	Beta	Std. Error		
1	(Constant)	1.420	2.752		.516		.603
	EXPENDITURE ON AD	1.073	.059	.988	18.295		.000

a.

Dependent Variable: MONTHLY MORTGAGE INTEREST

# REGRESSION EQUATION

$$Y = C + b.X$$

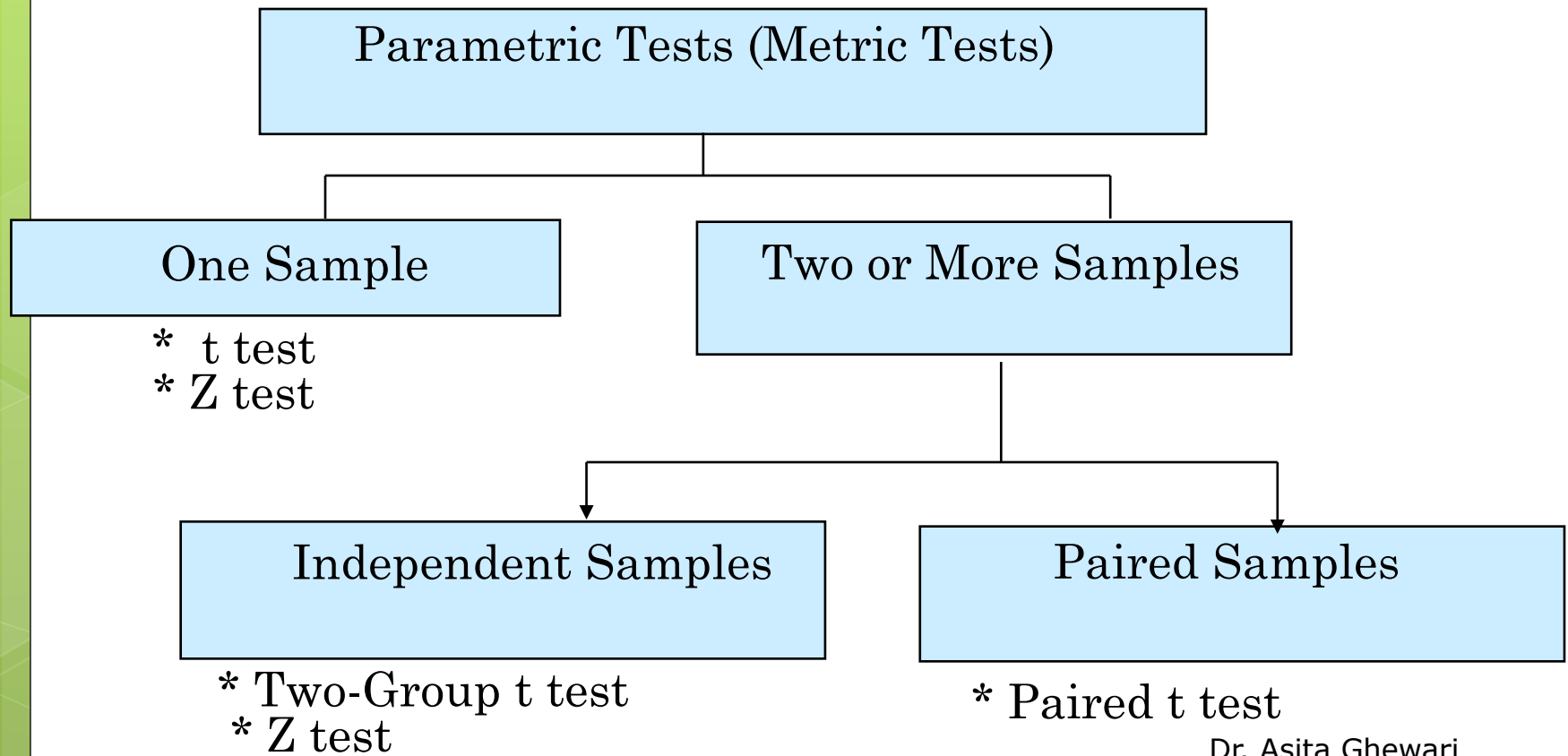
$$\text{Net Profit} = \text{Constant} + b.X$$

$$= 1.073 + .988 \times \text{Expenditure on AD}$$

**Multilinear Equation :**

$$Y = C + b_1 .X_1 + b_2 .X_2 + b_3 .X_3 + b_4 .X_4 + \dots$$

# A Classification of Hypothesis Testing Procedures for Examining Differences



Dr. Asita Ghewari



# One-Sample *t*-Test Example

Null	Ho: MEAN = 50 gm
Statistical test	<i>t</i> -test
Significance level	.05, n=100
Calculated value	1.786
Critical test value	1.66
Table value	Calculated value > critical value Reject H0

# T – Test for Independent Samples

- **T–test for independent samples is used for the scores (values) of two independent groups (two samples).**
- **That means there is no logical relationships between the scores (values) that have been obtained for one group when compared with other group.**

# The Paired t – test

- **The test is used when the researcher draws two random samples from the same population, introduces a treatment to one group while holding the other without the treatment, and then compares the groups to see whether there is a statistically significant difference between the groups.**

## Sale Proceeds of 8 Sales Executives before and after Sales Promotion Scheme

Sales Executive	Sales Before Promotion (in '000) (X)	Sales After Promotion (in '000) (Y)
1	9	12
2	9	10
3	15	15
4	12	14
5	8	14
6	10	11
7	9	10
8	9	8
n = 8		

Dr. Asita Ghewari

# PAIRED T-TEST OUT PUT

Null	$H_0: \mu_d = 0$
Alternative	$H_1: \mu_d > 0$
Statistical test	<i>t</i> -test dependent
Significance level	.05, $n=8$
Calculated value	1.650
Critical test value	1.895
Table value	Calculated value < critical value Accept $H_0$

Dr. Asita Ghewari

# Interpretation

The test results show that the exact significance is .069 which is greater than .05 hence we accept the null hypothesis and reject the alternative hypothesis and it can be inferred that there statistically no significant difference in sales before sales promotion and after sales promotion.

Dr. Asita Ghewari

# T– test for Independent Samples

- **The t – test for two independent samples examines the difference between their means to see how close or apart they are.**

## T – test for Independent Samples .....

- **The data in the next slide compares 10 days sales of two stores, say for example, Store 1 and Store 2.**



## SALE OF TWO STORES

Day	Sale of Store 1	Sale of Store 2
1	8	12
2	11	9
3	9	6
4	12	5
5	16	8
6	10	12
7	7	11
8	16	
	$\Sigma X = 100$	$\Sigma Y = 90$

## INDEPENDENT SAMPLES T-TEST OUTPUT

Null	$H_0: \mu_1 = \mu_2$
Alternative	$H_1: \mu_1 \neq \mu_2$
Statistical test	<i>t</i> -test independent
Significance level	.05, $n=14$
Calculated value	3.912
Critical test value	2.16
Table value	Calculated value > critical value Reject $H_0$

Dr. Asita Ghewari

## T – test for Independent Samples .....

### o Interpretation

The test results show that the exact significance is .809 which is MORE than .05 hence we reject the null hypothesis and accept the research hypothesis and it can be inferred that there is DIFFERENCE in sales of TWO stores.

# One-Way Analysis of Variance

- **The analysis of variance (some times obtained as ANOVA or F – Test) determines whether there is a statistically significant difference between more than two groups on the basis of their means.**

# One-Way Analysis of Variance Continues.....

- For example , let us consider 5 days sales of 3 departmental stores.
- The Sales Manager wishes to know if the sales of three departmental stores are significantly different.
- Sales data are shown in the next slide.

# 5 DAYS SALES OF 3 DEPARTMENTAL STORES



Day	Store A	Store B	Store C
1	14	38	34
2	20	42	40
3	22	24	30
4	18	12	48
5	30	40	58

Dr. Asita Ghewari

# One-Way Analysis of Variance

## SPSS Out Put

## One Way ANOVA OUTPUT

Null	$H_0: \mu_1 = \mu_2 = \mu_3$
Alternative	$H_1$ : At least two of departmental store sales are not same
Statistical test	ANOVA
Significance level	.05, n=12
Calculated value	29.02
Critical test value	3.89
Table value	Calculated value > critical value Reject $H_0$

Dr. Asita Ghawari



# One-Way Analysis of Variance

- Interpretation**

The test results show that the exact significance is **.0458** which is less than **.05** hence we reject the null hypothesis and accept the research hypothesis and it can be inferred that there is **DIFFERENCE** in 5 days sales of 3 stores.

**Alternatively, we can say that the difference in the 5 days sales of 3 stores is statistically significant.**

# Z-test

Test Name	Null Hypothesis	Example	Decision Rule
One sample Proportion Test	Ho: $P=0.1$ Population proportion is 10%	Proportion of defective in a batch 10%	Calculated Value of statistics > Table value , Reject $H_0$
Two sample Proportion Test	Ho : $P_1=P_2$ Two population has equal proportion	Proportion of defectives in two batches are same	Calculated Value of statistics > Table value , Reject $H_0$

Dr. Asita Ghewari

# Format of Research Report

- ◉ Executive Summary / Abstract
- ◉ Introduction
- ◉ Literature Review
- ◉ Objectives, Hypothesis
- ◉ Research Methodology
- ◉ Data Analysis
- ◉ Findings
- ◉ Suggestions / Further Research
- ◉ Conclusion
- ◉ References

Dr. Asita Ghewari