UNIT — II: Research Problem Formulation & Design (09 Hours)
Problem identification and research gap analysis, Hypothesis: Types, formulation, testing, Research design: Exploratory, experimental, simulation-based studies, Data sources: Primary vs. Secondary.

Research Problem Formulation & Design

---

1. Problem Identification and Research Gap Analysis

1.1 What is a Research Problem?

- It is a **statement about an area of concern** that requires a solution, improvement, or deeper understanding.
- Example:
    - General: "Why is employee turnover increasing in IT companies?"
    - Specific Research Problem: "What are the key factors influencing voluntary turnover among mid-level software engineers in Indian IT companies?"

1.2 Steps in Identifying a Research Problem

1. **Observation of real-life issues**
    - Example: Farmers facing low crop yield despite using fertilizers → Research problem: "Investigating the impact of soil quality on fertilizer effectiveness."
2. **Literature review**
    - Example: Existing studies show customer loyalty in retail, but little research in *online grocery* → Gap: "Customer loyalty in online grocery apps."
3. **Discussions with experts**
    - Example: Doctors report rising antibiotic resistance → Problem: "Effectiveness of alternative herbal medicines."
4. **Personal experiences/curiosity**
    - Example: A student interested in AI notices bias in Chatbots → Problem: "Analyzing fairness in AI-powered language models."
5. **Policy/organizational needs**
    - Example: Government wants to promote electric vehicles (EVs), but adoption is slow → Problem: "Barriers to EV adoption in urban India."

.3 Research Gap Analysis

- A research gap = **what is missing in existing knowledge.**
- Example of gaps:
    - **Knowledge gap:** No study on mental health of gig workers in India.
    - **Methodological gap:** Previous studies used only surveys; new study could use **in-depth interviews**.
    - **Theoretical gap:** No clear theory explaining why some startups succeed while others fail.
    - **Practical gap:** Many studies on renewable energy adoption, but none focus on *rural households*.

How to identify gaps?

- Review 20–30 papers on your topic.
- Make a table: *what they studied, methods used, limitations mentioned*.
- Example:
    - 10 studies on "Digital payment adoption" → All focused on *urban* customers.
    - Gap: "Digital payment adoption in *rural* areas."

---

2. Hypothesis: Types, Formulation, Testing

2.1 What is a Hypothesis?

- A **proposed explanation** made on the basis of limited evidence, tested using data.

2.2 Types with Examples

1. **Null hypothesis (H_0):**
    - "There is no significant difference in exam performance between students who study 2 hours and 5 hours daily."
2. **Alternative hypothesis (H_1):**
    - "Students who study 5 hours daily perform significantly better than those who study 2 hours."
3. **Directional hypothesis:**
    - "Increased social media usage decreases students' academic performance."
4. **Non-directional hypothesis:**
    - "There is a relationship between social media usage and academic performance." (but not saying positive or negative)

5. **Simple vs. Complex:**
    - Simple: “Advertising affects sales.”
    - Complex: “Advertising and product quality together affect sales and customer loyalty.”
6. **Associative vs. Causal:**
    - Associative: “Stress is related to sleep quality.”
    - Causal: “Stress causes poor sleep quality.”

2.3 Formulating a Hypothesis

- From **theories** (Maslow’s hierarchy, Technology Acceptance Model, etc.).
- From **previous studies**.
- From **observations**.
- Example:
    - Observation: Many people abandon shopping carts online.
    - Hypothesis: “High delivery charges increase shopping cart abandonment in e-commerce.”

2.4 Hypothesis Testing Example

- Study: Testing whether a new fertilizer increases crop yield.
- H_o: “New fertilizer has no effect on crop yield.”
- H_1: “New fertilizer increases crop yield.”
- Method: Conduct experiment on 100 farms, half use old fertilizer, half use new.
- If **$p < 0.05$** → reject H_o → conclude fertilizer works.

---

3. Research Design

3.1 Exploratory Research

- Used when problem is not well defined.
- Example: A company notices declining product sales but doesn't know why.
- Methods: Interviews with customers, focus groups, analyzing reviews.

3.2 Descriptive Research

- Used to **describe characteristics**.
- Example: "What percentage of people in Bangalore use UPI payments daily?"
- Method: Conduct survey → report statistics.

3.3 Experimental Research

- Used to **test cause-effect**.
- Example: Does a new teaching method improve student performance?
    - Experimental group → taught with new method.
    - Control group → taught with traditional method.
    - Compare exam results.

3.4 Simulation-Based Research

- Used when real-world testing is costly or impossible.
- Example:
    - Simulating **traffic flow** to test smart traffic lights.
    - Using **computer models** to predict climate change impact.

---

4. Data Sources

4.1 Primary Data Examples

- **Surveys:** Asking 500 college students about online learning satisfaction.
- **Experiments:** Testing effectiveness of a new medicine.
- **Interviews:** Talking to CEOs about digital transformation challenges.

4.2 Secondary Data Examples

- **Government reports:** Census, World Bank data.
- **Company data:** Annual reports of Tata Motors.
- **Research articles:** IEEE, Elsevier, Springer journals.
- **Databases:** IMF, UNDP, NSSO data.

4.3 Example Comparison

- Research Question: "What factors influence electric vehicle (EV) adoption?"
    - **Primary data:** Survey of 1,000 people about their EV preferences.
    - **Secondary data:** Government EV sales statistics from 2015–2024.

---

5. Summary

- A **research problem** comes from observation + literature + needs.
- A **gap** shows what is missing in knowledge.
- A **hypothesis** is a testable prediction about variables.
- **Research design** gives a structured plan (exploratory, descriptive, experimental, simulation).
- **Data sources** may be primary (first-hand) or secondary (existing).

Mini Case Studies for Research Problem Formulation & Design

---

1. Exploratory Research Case Study

Scenario:

An e-commerce company notices that **many customers abandon their shopping carts** before completing the purchase. Management wants to understand *why* this is happening, but they don't have clear ideas.

Tasks for Students:

1. Identify possible **research problem statement.**
    - Example: "Why do customers abandon online shopping carts in e-commerce?"
2. Suggest **exploratory research methods** (interviews, focus groups, analyzing customer reviews, website analytics).
3. Prepare a set of **5–6 open-ended interview questions.**
    - Example: "What makes you hesitate before completing an online purchase?"
4. Discuss how findings could lead to **hypothesis formulation.**

✅ *Learning outcome:* Students understand how to explore vague problems and turn them into specific, testable research questions.

---

2. Descriptive Research Case Study

Scenario:

A government education department wants to know the **extent of online learning adoption** among university students during and after COVID-19.

Tasks for Students:

1. Formulate a **research problem statement.**
    - Example: "What percentage of students in India are satisfied with online learning?"
2. Identify **variables**:
    - Independent: Type of university, student location (urban/rural)
    - Dependent: Satisfaction level, hours spent online
3. Design a **survey questionnaire** with 5–10 questions.
    - Example: "On average, how many hours per week do you attend online classes?"
4. Suggest methods for **data collection and analysis** (e.g., frequency tables, charts, descriptive statistics).

✅ *Learning outcome:* Students learn how to describe patterns and behaviors using survey data.

3. Experimental Research Case Study

Scenario:

A university wants to test whether a **new teaching method (flipped classroom)** improves student performance compared to traditional lectures.

Tasks for Students:

1. Define **hypotheses**:
    - H_0: No difference in exam scores between flipped classroom and traditional lectures.
    - H_1: Flipped classroom improves exam scores.
2. Identify **control and experimental groups**:
    - Control group → Traditional lecture
    - Experimental group → Flipped classroom method
3. Decide **sample size** (e.g., 100 students: 50 control, 50 experimental).
4. Select appropriate **statistical test** (e.g., independent t-test to compare mean exam scores).
5. Discuss possible **validity threats** (bias, motivation differences, teacher effect).

✅ *Learning outcome:* Students learn how to test cause–effect relationships with experiments.

---

4. Simulation-Based Research Case Study

Scenario:

A city government wants to test different **traffic signal systems** to reduce congestion, but real-life experiments would be costly and disruptive.

Tasks for Students:

1. Define **research problem statement.**
    - Example: “Can traffic congestion in city X be reduced using an adaptive traffic signal system?”
2. Build a **simulation model** (using software like AnyLogic, MATLAB, or even simple Excel modeling).
    - Inputs: Number of vehicles per hour, signal timing, road width.
    - Output: Average waiting time, congestion level.
3. Run **two scenarios:**
    - Fixed-time signals
    - Adaptive signals (adjust timing based on traffic flow)
4. Compare results (average waiting time reduced by 20%).

5. Suggest **policy recommendations**.

✅ *Learning outcome:* Students see how computer models can test real-world solutions when experiments are not feasible.

---

5. Mixed Example for Data Sources (Primary vs. Secondary)

Scenario:

You want to research **factors influencing electric vehicle (EV) adoption** in India.

Tasks for Students:

1. Identify **primary data** sources:
    - Surveys of 500 car buyers
    - Interviews with EV owners
2. Identify **secondary data** sources:
    - Government EV adoption statistics (Ministry of Road Transport)
    - Reports from Tata Motors, Ola Electric, or IEA
3. Discuss **advantages and limitations** of each.
4. Propose a **research design** combining both.

✅ *Learning outcome:* Students understand how to use both primary and secondary data effectively.