Meeting- 4
Variables and Hypothesis
Variable

What is variable?

- **symbols** which given numbers or values
- **characteristic**, number or quantity that increases or decreases over time or takes different values in different situations
- (usually) **letters** or **other symbols** that represent unknown numbers or values
- **factors** or **conditions** that can change during the course of an experiment.
A variable may also be called a data item.

Age, sex, business income and expenses, country of birth, capital expenditure, class grades, eye colour and vehicle type are examples of variables.

It is called a variable because the value may vary between data units in a population, and may change in value over time.
Types of Variables

- Variable
  - Numeric
    - Continuous
    - Discrete
  - Categorical
    - Ordinal
    - Nominal
Types of Variables

1. **Numeric variables:**
   - have values that describe a measurable quantity as a number, like 'how many' or 'how much'.
   - are quantitative variables.
   - Numeric variables may be further described as either:
     a) **Continuous variable** → any value obtained by measuring.
        Examples: height, time, age, temperature
     b) **Discrete variable** → any value obtained by counting.
        Examples: number of cars, number of children, etc.

   - The data collected for a numeric variable are quantitative data.
2. Categorical variables:
   - have values that describe a 'quality' or 'characteristic' of a data unit, like 'what type' or 'which category'
   - are qualitative variables and tend to be represented by a non-numeric value.
   - Categorical variables may be further described as:
     a) **Ordinal variable** → any value can be ordered or ranked.
        Examples: GPA, clothing size (S,M,L), attitudes
     b) **Nominal variable** → any value is not able to be organized in a logical sequence.
        Examples: sex, business type, religion, brand, etc.
   • The data collected for a categorical variable are **qualitative** data.
Quantitative and qualitative data can be gathered from the same data unit depending on whether the variable of interest is numerical or categorical. For example:

<table>
<thead>
<tr>
<th>Data unit</th>
<th>Numeric variable</th>
<th>Quantitative data</th>
<th>Categorical variable</th>
<th>Qualitative data</th>
</tr>
</thead>
<tbody>
<tr>
<td>A person</td>
<td>&quot;How many children do you have?&quot;</td>
<td>4 children</td>
<td>&quot;In which country were your children born?&quot;</td>
<td>Australia</td>
</tr>
<tr>
<td></td>
<td>&quot;How much do you earn?&quot;</td>
<td>$60,000 p.a.</td>
<td>&quot;What is your occupation?&quot;</td>
<td>Photographer</td>
</tr>
<tr>
<td></td>
<td>&quot;How many hours do you work?&quot;</td>
<td>38 hours per week</td>
<td>&quot;Do you work full-time or part-time?&quot;</td>
<td>Full-time</td>
</tr>
<tr>
<td>A house</td>
<td>&quot;How many square metres is the house?&quot;</td>
<td>200 square metres</td>
<td>&quot;In which city or town is the Brisbane house located?&quot;</td>
<td>Brisbane</td>
</tr>
<tr>
<td>A business</td>
<td>&quot;How many workers are currently employed?&quot;</td>
<td>264 employees</td>
<td>&quot;What is the industry of the business?&quot;</td>
<td>Retail</td>
</tr>
<tr>
<td>A farm</td>
<td>&quot;How many milk cows are located on the farm?&quot;</td>
<td>36 cows</td>
<td>&quot;What is the main activity of the farm?&quot;</td>
<td>Dairy</td>
</tr>
</tbody>
</table>
Other Types of Variables

1. Independent Variable (IV):
   - variables that may impact the dependent variable
   - variable that the scientist manipulates

2. Dependent Variable (DV):
   - its values depend upon another variable - the independent variable

3. Intervening/intermediate Variable (ITV):
   - A variable that explains a relation or provides a causal link between other variables
Other Types of Variables

Example:

- The four day working work (IV) will lead to higher productivity (DV) by enhancing the job satisfaction (ITV).

- A promotion campaign (IV) will increase savings activity (DV), by increasing the motivation to save (ITV).
Other Types of Variables

- Four day Working Work
- Productivity
  - Independent Variable
  - Dependent Variable
- Job Satisfaction
  - Intervening/Intermediate Variable
  - Dependent Variable
The word **hypothesis** has the same meaning in statistics as it does in everyday use. What does this word mean? Some possibilities are:

- an idea
- an assumption
- a guess
- a theory

**Definition:**

- *Hypothesis is a tentative statement formulated for empirical testing.*
1. Descriptive hypothesis:
   - *Statement that expresses about the existence, big, form and distribution of a variable*

   • Example:
     - *The unemployment rate in the city is more than 6 percent.*
Types of Hypothesis

2. **Correlation hypothesis:**
   - The statements that describe any relationship between two variables that related to any particular case

   **Example:**
   - The young technicians are less productive than the older
3. Causal hypothesis:
   – Any changes in one variable will influence other variable

• Example:
  – An increasing in family income tends to raise saving rate

independent variable                    dependent variable
In statistics, a **hypothesis** is an idea, an assumption or theory about the characteristics of one or more populations.

**Example:**
1. The unemployment rate in the country is more than or equal 6 percent.
2. The mean of student’s age is 22 years old.
Hypothesis Formulation

• A hypothesis often follows a basic format of "If {this happens} then {this will happen}.

• One way to structure your hypothesis is to describe what will happen to the dependent variables if you make changes to the independent variables.

• The basic format might be: "If {these changes are made to a certain independent variable}, then we will observe {a change in a specific dependent variable}"
Hypothesis Formulation

A few examples:

1. "Students who eat breakfast will perform better on a math exam than students who do not eat breakfast."

2. "Students who experience test anxiety prior to an English exam will get higher scores than students who do not experience test anxiety."

3. "Motorists who talk on the phone while driving will be more likely to make errors on a driving course than those who do not talk on the phone."
Hypothesis Formulation

Descriptive hypothesis ➔ Research Question

The unemployment rate in the city is more than 6 percent.

What is the unemployment rate in the city?

Eighty percent share holder in PT ABC agree with increasing in devidend.

Do the share holder in PT ABC prefer an increasing devidend?
## Hypothesis Formulation

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Example Hypothesis</th>
<th>Possible Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Why is this place risky?</td>
<td>• It is due to a large number of targets being available.</td>
<td>• Count the targets at the location and calculate the crime rate. Compare this rate to rates for the surrounding area. If the risky place has a higher rate, the hypothesis is false, if it is about the same or lower, then the hypothesis is true.</td>
</tr>
<tr>
<td>2. Why are there more car thefts in the problem area than in nearby areas?</td>
<td>• Residents park their cars on the street, unlike residents of other areas.</td>
<td>• If the problem area has similar or lower on-street parking rates than the others, reject the hypothesis. If higher, accept it.</td>
</tr>
<tr>
<td>3. Why did the theft of copper piping from new construction suddenly increase?</td>
<td>• A new owner bought a nearby scrap metal dealership.</td>
<td>• Compare the thefts of piping before and after the change in owners. If the theft rate is the same before and after, or the trend in thefts was already going up before the change, then the hypothesis is probably false. If otherwise, the hypothesis appears reasonable.</td>
</tr>
</tbody>
</table>
Hypothesis for Testing

• **Null hypothesis** ($H_0$):
  – assumption that no difference exists between the sample statistic and the population parameter

• **Alternative hypothesis** ($H_1/H_A$):
  – the logical opposite of the null hypothesis
  – that a difference exists between the sample statistic and the population parameter to which it is compared.
Example:

1. \( H_0 \): The unemployment rate in the city is more than or equal 6 percent.
   \( H_1 \): The unemployment rate in the city is less than 6 percent.

2. \( H_0 \): The mean of student’s age is 22 years old.
   \( H_1 \): The mean of student’s age is not 22 years old.
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
What is a Good Hypothesis?

- A good hypothesis should fulfill three conditions:
  - Must fit with its purpose
  - Must be testable
  - Must be simple, need less assumption
### Group Practice - Hypothesis Formulation

Please formulate the hypothesis and its possible test for every hypothesis based on the given research question.

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Hypothesis (Make at least 3 hypotheses for each)</th>
<th>Possible Test (for each hypothesis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Why did the product sales decrease after Christmas?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>2. Why did the hotel occupation rate decline?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3. Why did the defect product increase?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4. Why did the production cost increase?</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>