

Hypothesis = Assumption

Drinking Sugary drinks daily
Leads to being Overweight

Null Hypothesis

- **there is no statistical significance difference between the two variables in the hypothesis.**
- Example : There is no statistically significant relationship between the type of water I feed the flowers and growth of the flowers.
- often represented by H_0 (H-zero)

Alternate Hypothesis

- **A statistical hypothesis used in hypothesis testing, which states that there is a significant difference between the set of variables.**
- denoted by H_1 (H-one)

	H_0 is actually:	
	True	False
Reject H_0	Type I error	Correct
Accept H_0	Correct	Type II error

Classification of Data

- Process of arranging data into homogeneous (similar) groups according to their common characteristics.
- heterogeneous data is divided into separate homogeneous classes

Example : Separating data on the basis of Public, Internal, Confidential

Professor. Secrist

- “Classification is the process of arranging data into sequences according to their common characteristics or separating them into different related parts.”

Type of Classification of Data

Descriptive and Quantitative Classification

Descriptive characteristics are those that can be described in words and are expressible in qualitative terms.

- For instance, literacy, gender, caste and religion are descriptive characteristics.
- Descriptive or qualitative classification is termed classification according to attributes.

Type of Classification of Data

Descriptive and Quantitative Classification

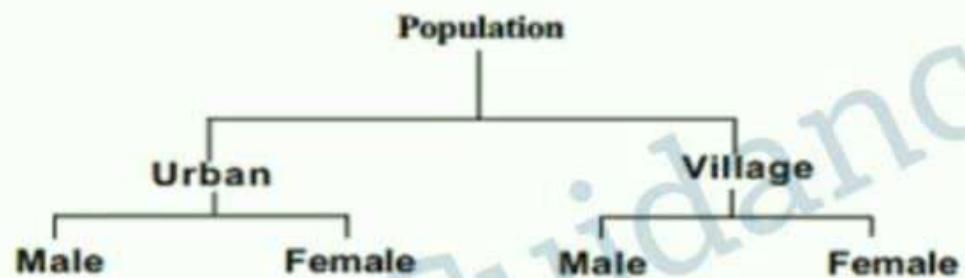
Descriptive characteristics are those that can be described in words and are expressible in qualitative terms.

- For instance, literacy, gender, caste and religion are descriptive characteristics.
- Descriptive or qualitative classification is termed classification according to attributes.

Numerical characteristics are quantitative in nature.

- Height, weight, age, income and expenditure are numerically expressible characteristics.
- Numerical or quantitative classification of data in certain class intervals

Type of Classification of Data



States	Production of food grains (in '000 tons)
Tamil Nadu	4500
Karnataka	4200
Andhra Pradesh	3600

Type of Classification of Data

- Simple and Manifold Classification

In **Simple classification**, only one attribute is studied.

- the data is classified into two separate classes under a single attribute.
- For Example, data collected on literacy in the country can be classified into two distinct classes: literate and illiterate.

Analysing and classifying collected data under several attributes in different classes is called **manifold classification**.

- For example, if each of the two classes, literate and illiterate, is divided into males and females, then there would be four classes.

- For example, if each of the two classes, literate and illiterate, is divided into males and females, then there would be four classes.
- Classification of data into a number of classes and classes within classes is known as manifold classification.