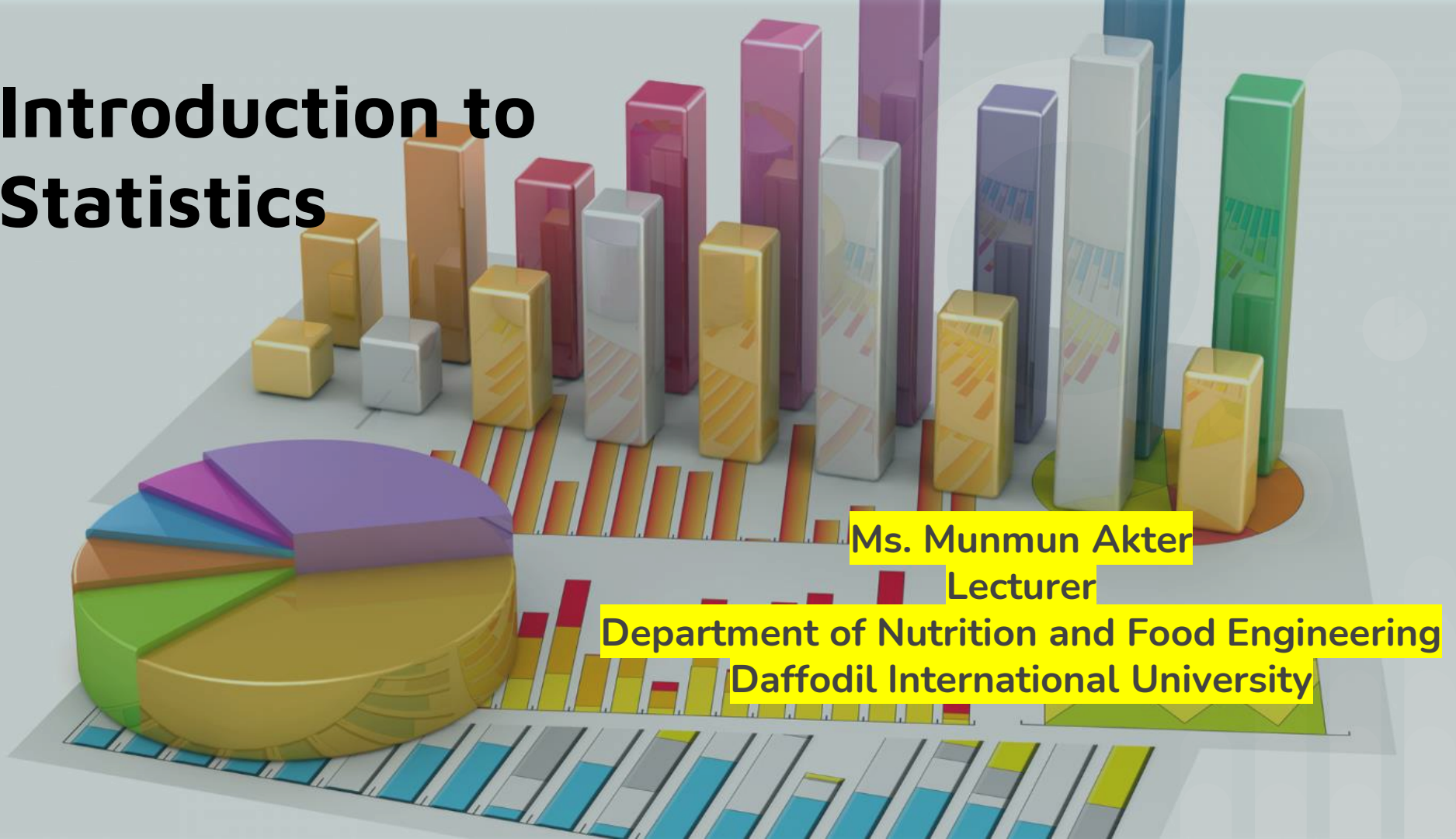
Introduction to Statistics



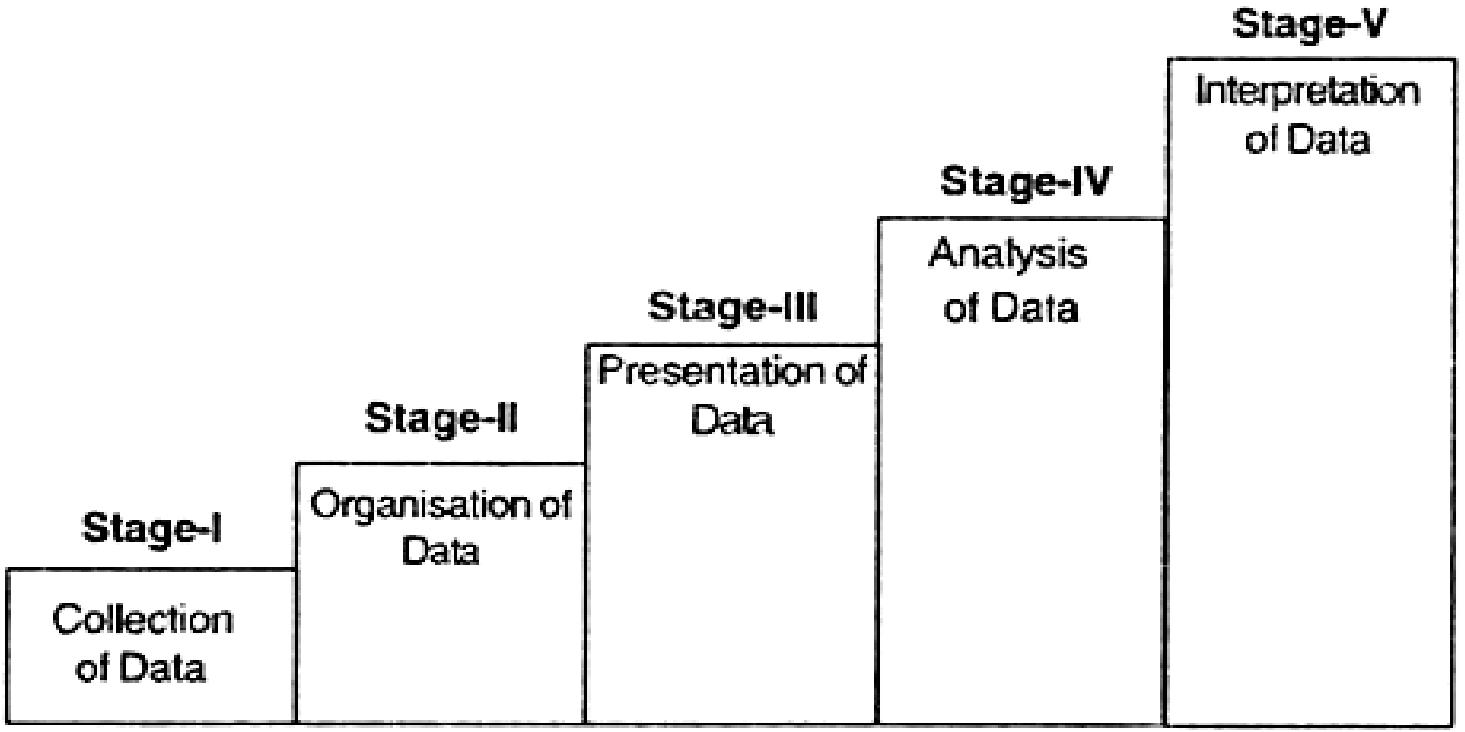
Ms. Munmun Akter
Lecturer
Department of Nutrition and Food Engineering
Daffodil International University

What is statistics

- Statistics as a discipline refers to statistical methodology techniques and procedures dealing with design of experiments, Collection, Classification, Summarization, Organization, and interpretation of information contained in a data set

- Statistics is science & art of dealing with variation in such a way as to obtain reliable results

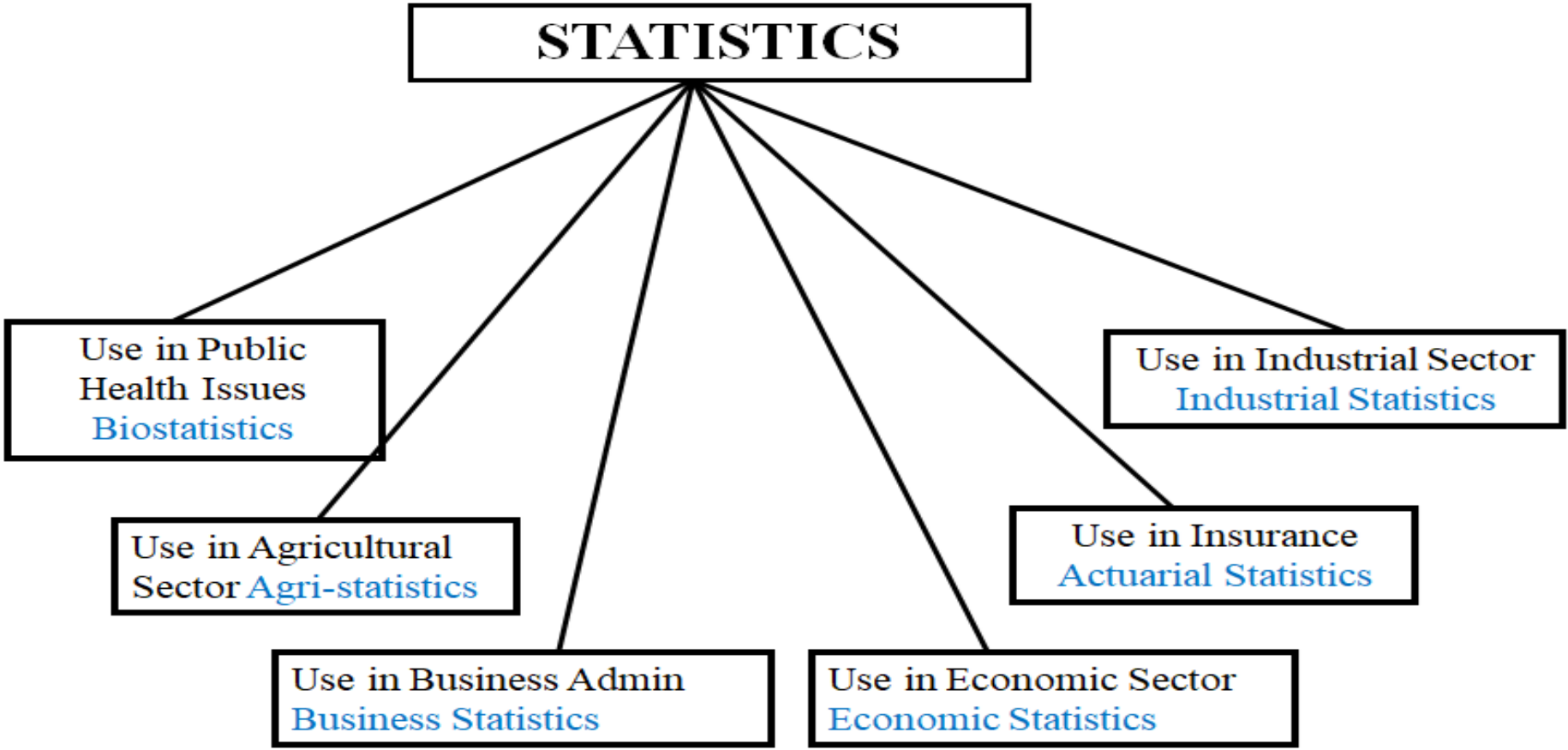
Steps of Statistics



Work of Statistician

- To guide the design of an experiment or survey prior to the data collection
- To analyze data using proper statistical procedures and techniques
- To present and interpret results to the researchers and other decision makers including the government and industries

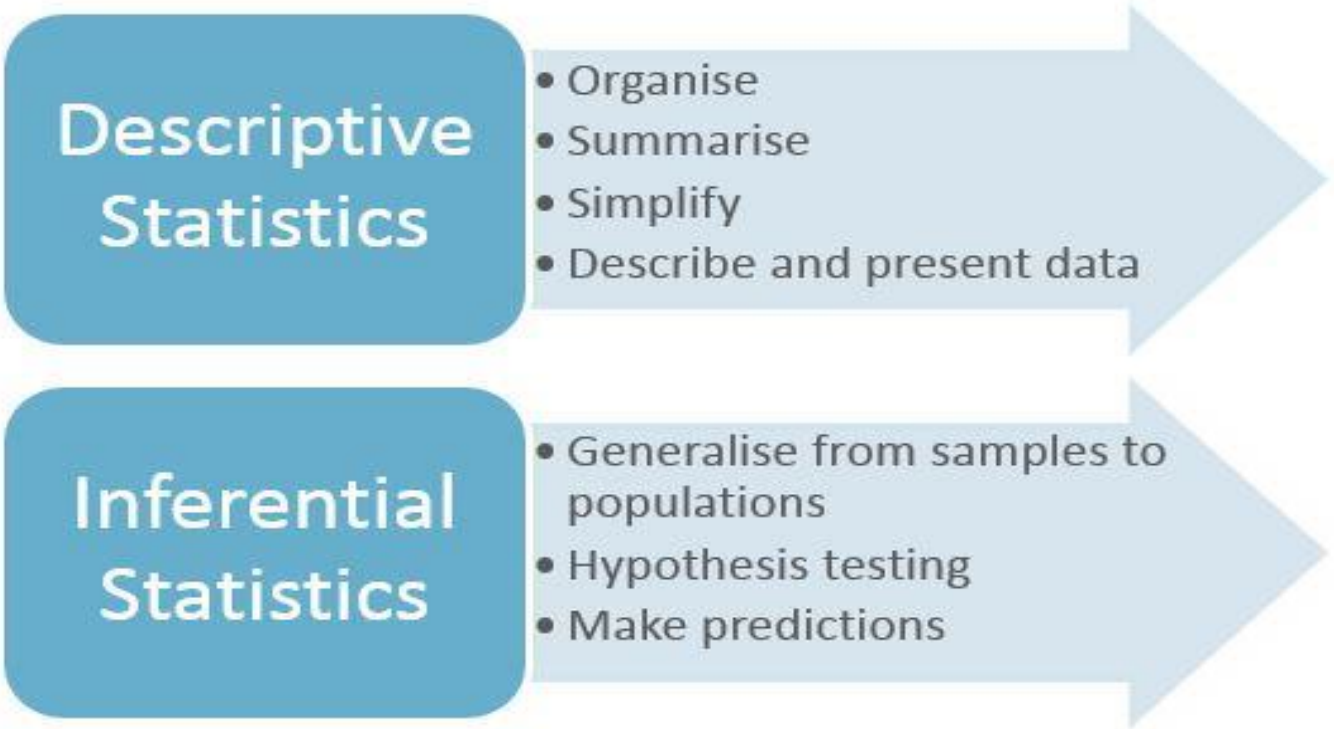
Flavor of Statistics



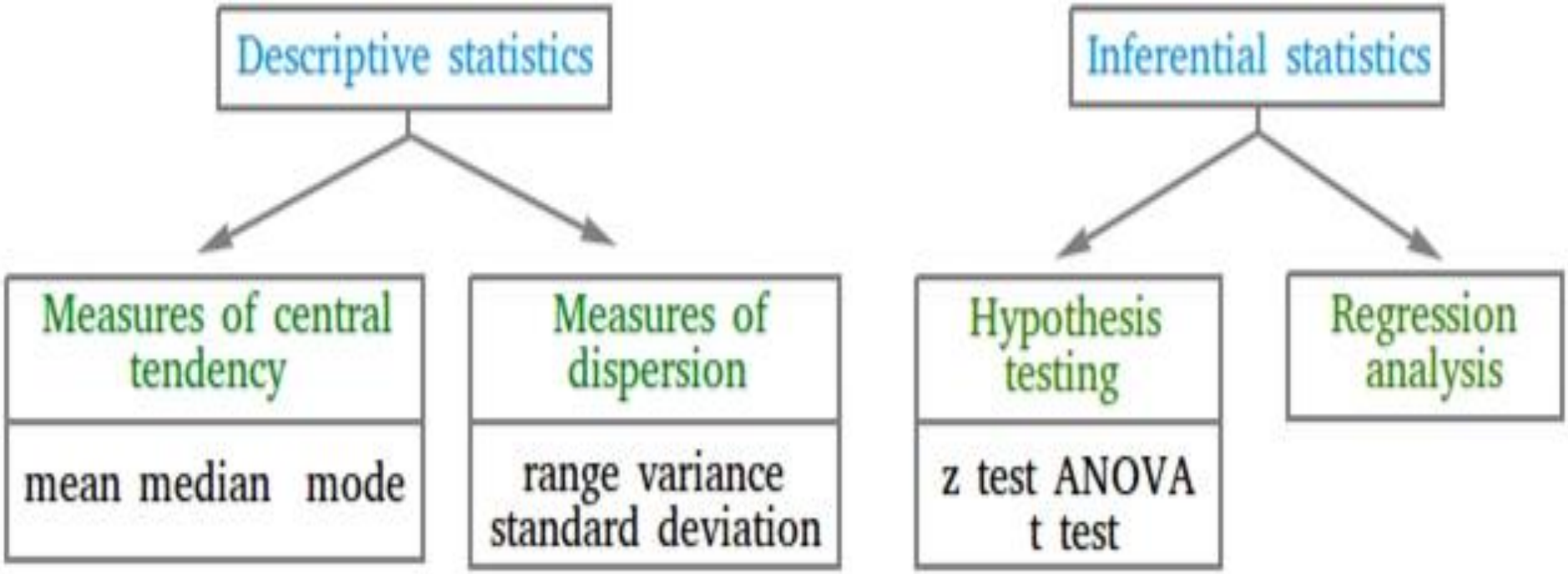
Biostatistics

- ★ When statistics is applied in Biology (including human biology, medicine, and public health), it is caled biostatistics
- ★ It is generally used to refer recorded data such as number of patient attending a hospital, no of malnourisjed children, percentage of dietary diversity in rural and urban population etc.
- ★ Francis Galton (1822-1911) hjas been called the father of Biostatistics.

Types of Statistics

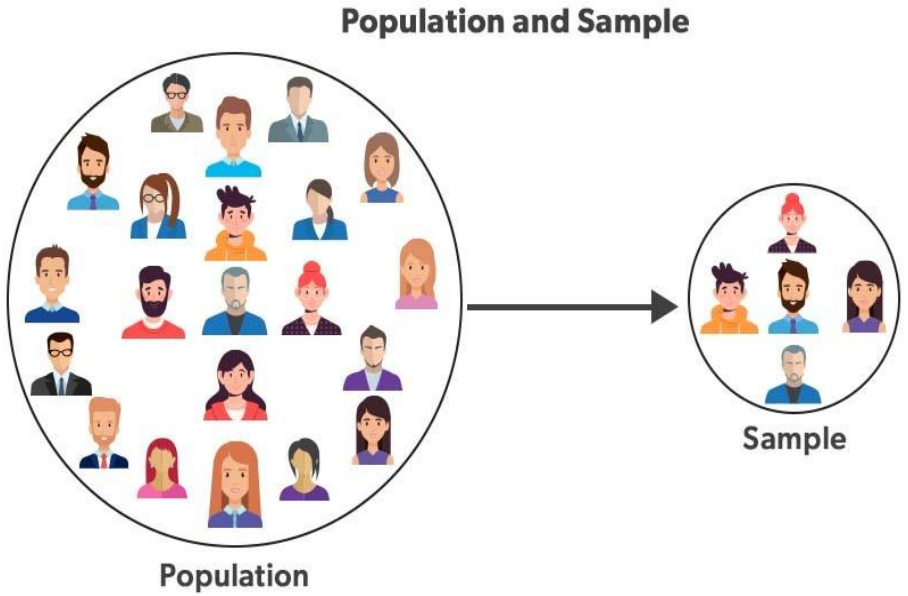


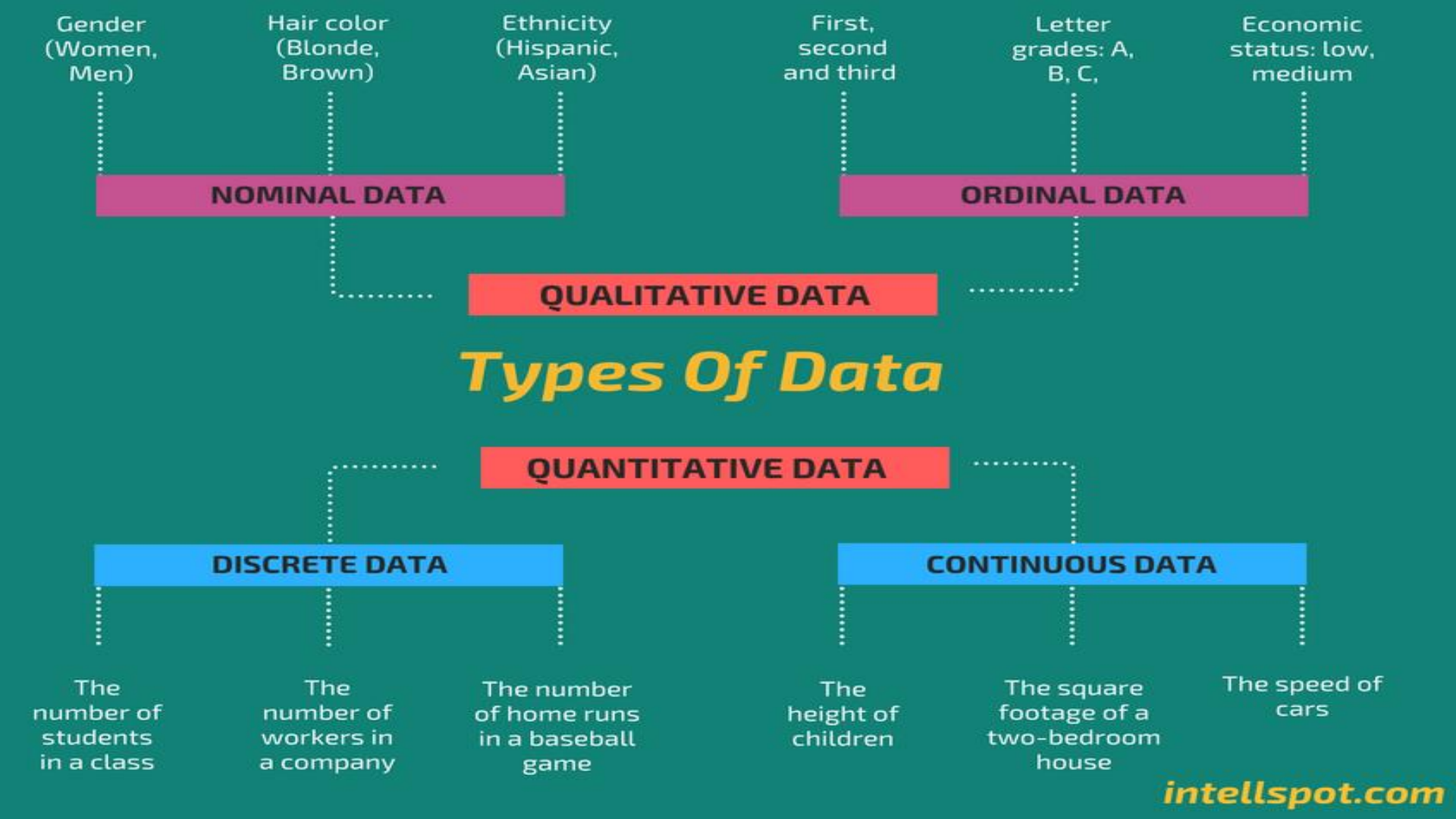
Measure of statistics



What is Population and Sample

- A population is a set of measurements of interest to the researcher
- A subset of the population is called sample A sample is usually selected such that it is representative of the population




Gender (Women, Men)
Hair color (Blonde, Brown)
Ethnicity (Hispanic, Asian)
First, second and third
Letter grades: A, B, C,
Economic status: low, medium
NOMINAL DATA
ORDINAL DATA
QUALITATIVE DATA
Types Of Data
QUANTITATIVE DATA
DISCRETE DATA
CONTINUOUS DATA
The number of students in a class
The number of workers in a company
The number of home runs in a baseball game
The height of children
The square footage of a two-bedroom house
The speed of cars
intellspot.com

Nominal Data: Nominal data represent categories or groups with no inherent order or ranking. Examples include gender, eye color, marital status, or car brands. Nominal data can be represented using labels or codes.

Ordinal data: Ordinal data also represent categories or groups, but they have a natural order or ranking. Examples include survey responses with Likert scales (e.g., strongly agree, agree, neutral, disagree, strongly disagree) or educational levels (e.g., high school, bachelor's, master's, PhD). The order of categories is important in ordinal data.

Discrete Data: It can only take specific, countable values. They are usually whole numbers or integers. Examples include the number of siblings, the number of cars in a household, or the count of occurrences of an event. Discrete data can further be categorized as nominal or ordinal data.

Continuous Data: It can take any numerical value within a certain range or interval. They are measured on a continuous scale. Examples include height, weight, temperature, and time. Continuous data can be further divided into interval and ratio variables.

Scale of Data/ Label of measurement

The scale of data refers to the level of measurement or the properties associated with the values that a variable can take.

1. **Nominal Scale:**
    - Nominal variables represent categories or groups with no inherent order or ranking.
    - Examples: Gender (male, female), eye color (blue, brown, green), marital status (single, married, divorced).
    - Properties: Categories are mutually exclusive and exhaustive, and there is no inherent numerical value or magnitude associated with them.
2. **Ordinal Scale:**
    - Ordinal variables represent categories or groups with a natural order or ranking.
    - Examples: Likert scale responses (e.g., strongly agree, agree, neutral, disagree, strongly disagree), educational levels (high school, bachelor's, master's, PhD).
    - Properties: Categories have a relative ranking or order, but the difference between categories may not be uniform or quantifiable.

3. Interval Scale:

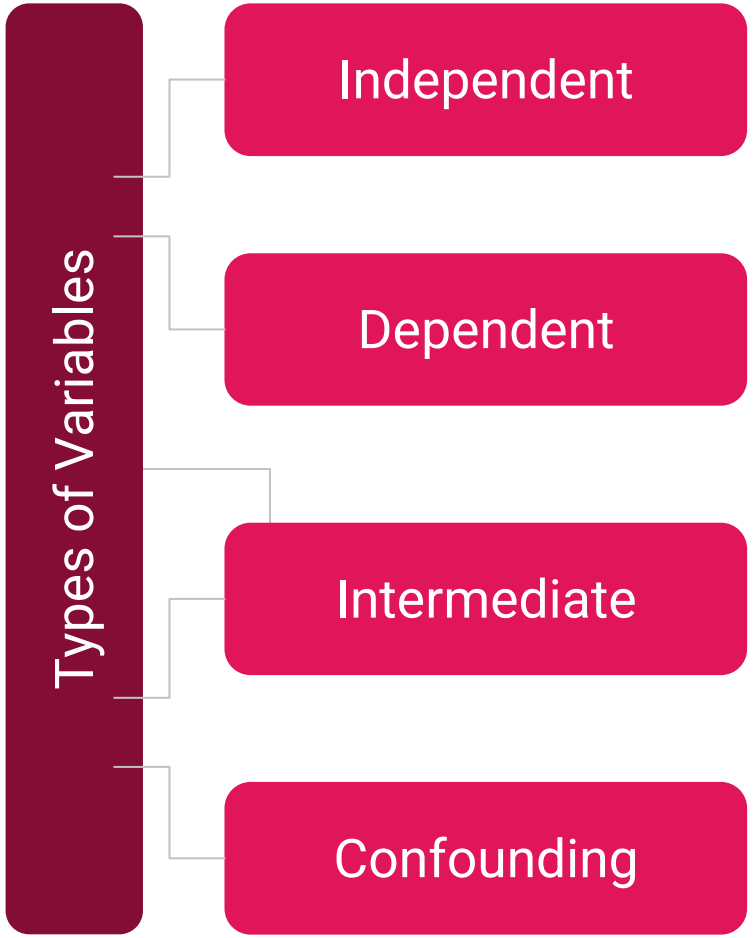
- Interval variables have equal intervals between values, and the zero point is arbitrary.
- Examples: Temperature measured in Celsius or Fahrenheit, calendar dates.
- Properties: Differences between values are meaningful, but ratios and proportions are not. Zero does not represent the absence of the attribute being measured.

4. Ratio Scale:

- Ratio variables have equal intervals between values and a true zero point.
- Examples: Height, weight, time, income.
- Properties: Ratios and proportions between values are meaningful. Zero represents the absence of the attribute being measured.

Variables

A variable is a characteristic or attribute that can vary or take different values. It is a measurable quantity or a characteristic of an object, individual, event, or phenomenon that can be observed or recorded. Variables are the building blocks of data and play a central role in statistical analysis.



Types of variables

- **Independent Variable:** The variable that is manipulated or controlled by the researcher. It is the presumed cause or predictor variable that is hypothesized to have an effect on the dependent variable.
- **Dependent Variable:** The variable that is measured or observed to determine the outcome or response. It is the variable that is expected to be influenced or affected by changes in the independent variable

Types of variables

- **Intermediate Variable:** A variable that is influenced by the independent variable and, in turn, influences the dependent variable. It acts as a mediating variable in explaining the relationship between the independent and dependent variables.
- **Confounding Variable:** An extraneous variable that is related to both the independent and dependent variables and can distort or confuse the relationship between them.

Common statistical symbols

Title	Symbol
Sample Mean	$\bar{x}$
Population mean	μ
Sample standard deviation	s
Population standard deviation	σ
Sample variance	s^2
Population variance	σ^2
Summation	Σ
Correlation coefficient	r
Coefficient of determination	r^2
Degree of freedom	df

Common statistical symbols

Title	Symbol
Chi-square value	$X2$
Sample proportion	p
Population proportion	$\prod$
Null hypothesis	Ho
Alternative hypothesis	H_1 or H_A
Sample Size	n
Type I error	a error
Type II error	b error
Power of the test	1- b