

DAY 2

Saturday, May 4, 2024 10:32 AM

MATRICES AND DATA CLASSIFICATION:

1. NOMINAL SCALE: *Identification*

- When numbers assigned to objects serves as labels for identification or categorization
- Nominal scale variables are often used to label or identify different groups or characteristics
- Such numbers have no quantitative meaning
- Ex:
 - i. Gender: Categories include male, female, and non-binary.
 - ii. Marital Status: Categories include single, married, divorced, widowed.
 - iii. Eye Color: Categories include blue, brown, green, hazel.
 - iv. Blood Type: Categories include A, B, AB, O.
 - v. Types of cars (sedan, SUV, truck).

2. ORDINAL SCALE: *Rank*

- When assigned numbers to indicate relation b/n entities in terms of greater than , less than , but do not state how much greater than or less than
- For example, consider a survey asking respondents to rate their satisfaction with a product on a scale of "very dissatisfied," "dissatisfied," "neutral," "satisfied," and "very satisfied."
- In this ordinal scale, there is a clear order from least to most satisfaction, but the difference in satisfaction between "dissatisfied" and "neutral" may not be the same as between "neutral" and "satisfied."
- Permissible: Median , Quartile , Percentiles, Rank Order, etc.

3. INTERVAL SCALE: *Difference*

- When assigned number are such that difference in numbers is valid but no ratios allowed

- An interval scale is a level of measurement used in statistics where the distance or interval between adjacent points on the scale is equal, but there is no true zero point. In other words, the values on an interval scale have a meaningful order, and the difference between any two adjacent points is consistent, but there is no absolute starting point or true zero.
- A classic example of an interval scale is temperature measured in degrees Celsius or Fahrenheit. The difference between 20°C and 30°C is the same as the difference between 30°C and 40°C (i.e., 10 degrees), but a temperature of 0°C does not represent the complete absence of temperature; it's just a point on the scale.
- EX: Calendar Dates: Dates on the Gregorian calendar can be considered as an interval scale. The difference between January 1st and January 10th is the same as the difference between January 10th and January 20th, but there's no true zero date.

4. **RATIO SCALE:** *Ratio / Absolute Zero*

- When scale contains absolute zero
- All mathematical functions allowed here
- Ex: Height , Weight, Distance, etc.

SOURCES OF DATA:

PRIMARY DATA	SECONDARY DATA
Collected by Self	Collected by others(3rd party)
Reliable	Non-Reliable
Fresh data (recently collected)	Outdated possibility
As per our requirement data can be get	We can't collect as per our need
Time taking and Costly	Highly economic

CENTRAL TENDANIES:

1. **MEAN:**

- Mean is average

2. **MODE**

- Mode is the middle number of sorted data

3. MEDIAN

- Number with most occurrence is the median

MEASURES OF DISPERSION:

1. RANGE:

- Highest and Lowest Value

2. QUARTILE DEVIATION:

- The values which divide given data into 4 equal parts

Q1 Q2 Q3



- Q1 contains 25% of observations on its left
- Q2 contains 50% of observations on left
- Q3 contains 75% of observations on left

3. MEAN DEVIATION:

- Mean deviation, also known as mean absolute deviation (MAD), is a measure of the average absolute deviation of values from the mean of a dataset.
- It quantifies the average distance between each data point and the mean of the dataset.

$$\text{Mean Deviation} = \frac{\sum_{i=1}^n |x_i - \text{mean}|}{n}$$

4. STANDARD DEVIATION:

- Standard deviation is a measure of the dispersion or spread of a set of data points around the mean (average) of the dataset.
- It quantifies the amount of variation or deviation from the mean. A low standard deviation indicates that the data points tend to be close to the mean, while a high standard deviation indicates that the data points are spread out over a wider range of values.

$$\text{Standard Deviation}(\sigma) = \sqrt{\frac{\sum_{i=1}^n (x_i - \text{mean})^2}{n}}$$