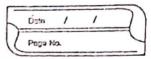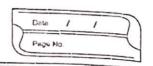Measures of Central Tendency

Arithmetic mean (Average)

The arithmetic mean is the most widely used measure of central tendency. It is calculated by adding all the values in a data set and then diving by the number of values.

Formula:

for population mean (μ): $\mu = \Sigma x / N$

for a sample mean (M): $M = \Sigma x / N$

Example from NPL:

The mean number of runs scored by 4 teams during

2024 was calculated using the following data:

Runs scored by 4 teams: 180, 160, 220, 140, 140, 170, 180,
188, 78, 88, 90, 92, 94, 199, 80

The sum of runs scored is $\sum x = 2090$

The number of teams is $N = 4$

∴ mean (x) is $= \sum X/N = 522.5$

While mean is the most commonly used it is sensitive to extreme values (outliers) which can skew the results.

Median

Median is another common measure of central tendency. It is defied as the middle value of a data set when it is ordered from smallest to largest

How to compute the Median

(1) for an odd Number of observations:

If the dataset has odd number of values, the median is the middle value when the data is arranged in order.
for example 3, 5, 7, 9, 11 . 7 is median Here.

(2) for an even Number of observations:

If the dataset has even number of values, the median is the average of the two middle values.
Example 2, 4, 7, 12 the median value is 4, 7 so the median
is $\frac{4+7}{2} = 5.5$

The median is the most appropriate measure of central tendency when dealing with skewed distribution or outliers because it is not affected by extreme values.

Mode

Mode is the measure of central tendency that refers to the most frequently occurring value in a dataset

Example from NPL 2020 "180" is mode as it is most frequently occuring value in a dataset

Median and Mean: Understanding Their Differences and Usous

Mean and Median, the key measures of central tendency. Both serve to represent the "center" of a data set, but they do so in different ways

When are the mean and median the same?

* Symmetric Distributions

When the data is symmetrical, the mean and median are the same. This is often the case in a bell shaped, normal distribution

Example: Imagine you're measuring the height of a group of people 1.6m, 1.7m, 1.8, 1.9 m and 2.0m

Mean : (1.6 + 1.7 + 1.8 . 1.9 + 2.0) / 5 = 1.8m

Median = 1.8m (The middle value)

Conclusion: Both the mean and median are the same because the distribution is symmetric.

Mean vs Median in Minimizing Deviations

(1) Minimizing Absolute Deviation

The median minimizes the sum of absolute Deviation. This means the sum of the distances from the median to each data point is the smallest

Consider these data points: 2, 3, 4, 9, 16

The median is 4

Absolute deviation from 4

$|2-4|$ $|3-4|$ $|4-4|$ $|9-4|$ $|16-4|$ = 20

∴ Sum of absolute deviation = 20

The mean is 6.8

Absolute deviations from 6.8

$|2-6.8|$ $|3-6.8|$ $|4-6.8|$ $|9-6.8|$ $|16-6.8|$ = 9.2

Sum of absolute deviations = 22.8

Conclusion The median minimizes the sum of absolute deviations (20) more than the mean (22.8)

(2) Minimizing Squared Deviations:

This mean minimizes the sum of squared deviations. This is important when you want to penalize larger differences more strongly

• Squared deviation from 4 (median):

$(2-4)^2$ $(3-4)^2$ $(4-4)^2$ $(9-4)^2$ $(16-4)^2$ = 174

∴ Sum of squared deviation = 174

squared deviation from 6.8 (mean):

$(2-6.8)^2$ $(3-6.8)^2$ $(4-6.8)^2$ $(9-6.8)^2$ $(16-6.8^2)^2$ = 134

Conclusion: The mean minimizes the sum of squared deviations (134.8) more than the median (114)

Balancing Point. The Mean
- The mean is the balance point on a distribution, meaning if you imagine the data points as weights on a number line, the mean is the point where the distribution balances out

Real world Analogy
Imagine you are on even planer organizing a party. You have the choice to arrange seating by

Mean : You balance all the people on a long table, and the exact middle of the table is the balance point where the "weight" of everyone is evenly distributed

Median : You find the middle person based on their seating number, ensuring half the guests are on either side of that person.

Keywords Takeways
- Symetric Data : For symetric distributions, mean and median are the same
- Minimizing Absolute Deviations : The median minimizes the sum of absolute deviation.
- Minimizing Squared Deviation : The mean minimizes the sum of squared deviations