Experiment No.2

Title: Measuring Central Tendency and variability of the Data

Batch: **Roll No.:** **Experiment No.:2**

Aim: 1. Find out measures of central tendency for single and multi-attributes data using Statistical analysis
2. Find out measures of variability of data using statistical analysis

---

Resources needed: Any statistical analysis tools and any programming language

---

Theory:

Mathematically central tendency means measuring the center or distribution of location of values of a data set. It gives an idea of the average value of the data in the data set and also an indication of how widely the values are spread in the data set. That in turn helps in evaluating the chances of a new input fitting into the existing data set and hence probability of success.

There are three main measures of central tendency:

- Mean - It is the Average value of the data which is a division of sum of the values with the number of values.
- Median - It is the middle value in distribution when the values are arranged in ascending or descending order.
- Mode - It is the most commonly occurring value in a distribution.

Variability describes how far apart data points lie from each other and from the center of a distribution. Along with measures of central tendency, measures of variability give you descriptive statistics that summarize your data. Variability is also referred to as spread, scatter or dispersion. It is most commonly measured with the following:

- Range: the difference between the highest and lowest values
- Interquartile range: the range of the middle half of a distribution
- Standard deviation: average distance from the mean
- Variance: average of squared distances from the mean
- Coefficient of variation: ratio of the standard deviation to the mean.

Procedure / Approach /Algorithm / Activity Diagram:

1. Analyze the data to find out frequency, mean, mode and median, standard deviation, variance, interquartile of data?

2. Write a programming code with a procedure to compute mean mode and median, variance, standard deviation, interquartile of a given sample data without using readymade function?

3. Compute and analyze the group data to find out frequency, mean, mode and median, standard deviation, variance, interquartile of data?

Results: (Program printout with output / Document printout as per the format)

Questions:

1. What are the various applications of central tendency and variability of data?
2. What are the various applications of finding Central Tendency of Data?

Outcomes:

Conclusion: (Conclusion to be based on the objectives and outcomes achieved)

Grade: AA / AB / BB / BC / CC / CD /DD

Signature of faculty in-charge with date

References:

Books/ Journals/ Websites:

1. Han, Kamber, "Data Mining Concepts and Techniques", Morgan Kaufmann 3nd Edition

2. S.C. Gupta , V. K. Kapoor Fundamentals of mathematical statistics Sultan Chand and Sons 2014