RegdNo. 322201297

Assignment No. 1

Course Code: ECAP790 Registration Number: 322201297

Instructions:

a. Attempt all questions given below in your own handwriting. Assignment in typed format will not be considered for evaluation.

b. The student has to complete the assignment in the allocated pages only. Any other page in case utilized shall not be considered.

Q1. Describe various measure of variability and how they are used in data analysis?

[10 Marks] [CO1, L1]

Measures of variability are statistical tools used to describe the spread or dispersion within a dataset. Here are the main measures of variability and their uses in data analysis.

1. Range: The range is the difference between the maximum and minimum values in a dataset. It is the simplest measure of variability but can be influenced by extreme values. For example, in a dataset {1, 3, 5, 7}, the range is 7 − 1 = 6. The range gives a quick sense of the data spread but does not provide information about the distribution of values within the dataset.

2. Variance: Variance measures how far a set of numbers are spread out from their average value. It is calculated by taking the average of the squared differences between each data point and the mean.

3. Standard deviation: It is the square root of the variance, denoted as σ (for population) or s (for sample). It provides a measure of spread in the same units as the data.

4. Quartiles and Quartile Deviation: Quartiles divide a ~~set~~ dataset into four equal parts. The first quartile (Q1) is the median of the lower half of the data, the second quartile (Q2) is the median, and the third quartile (Q3) is the median of the upper half. The quartile deviation also known as the semi-interquartile range.

5. Mean Absolute Deviation (MAD): MAD is the average of the absolute difference between each data point and the mean. It gives an idea of the average distance from the mean without squaring the difference, making it less sensitive to outliers than variance and standard deviation.

6. Coefficient of Variation (CV): The coefficient of variation is the ratio of the standard deviation of the means, expressed as a percentage. $CV = (\sigma/\mu) * 100\%$. It is useful for comparing the degree of variation between datasets with different units or widely different means.

Signature of the Student Tanya

Note:-

CO: is the Course Outcome as per your course syllabus.

L1-L6: Learning level objectives as per Revised Bloom Taxonomy (RBT).