REPORT BASED ON ANNUAL SALARIES DISTRIBUTION

Name: AKSA SIBI ANNAM

Student ID: 22070385

1.Data Description:

The 'data5-1.csv' file provided the dataset for this study, which shows yearly salary. NumPy, Matplotlib, and SciPy are a few of the Python libraries that were used to analyse the annual salary distribution and extract significant statistical measures. For visual purposes, a histogram with thirty bins was made, providing information about the distribution properties. In addition, the data's underlying probability density function was modelled using a normal distribution.

2.Distribution Overview:

After producing a 30-bin histogram, the yearly salary distribution was examined visually. A normal distribution model was used to better represent the data, and important parameters were calculated for additional analysis, including the mean (μ) and standard deviation (σ).

3.Calculation of Mean (W~):

The mean annual salary (W~) is calculated using the numpy package's np.mean() function. The formula is as follows:

$$w\sim = \frac{1}{n}\sum_{i=1}^{n} xi$$

Where 'n' is the number of data points (salaries) and 'xi' represents individual salary values.

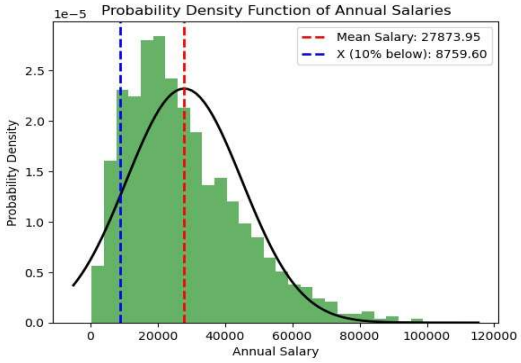
Calculation of Required Value X (10th Percentile):

The value at the 10th percentile (X) is calculated using np.percentile(data, 10). This statistic represents the wage that 10% of the data falls below. The formula for calculating percentiles is as follows:

Value of the data point at the p-th percentile, denoted by X=Percentile(p)

In this example, p=10.

Graphical Representation:



Results:

The graph produced by the Python method shows the computed mean salary (W) of 27873.95 and standard deviation (X) of 8759.60 for the provided dataset.

Conclusion:

This research provides insight into the distribution of annual revenues by displaying significant statistical metrics such as the mean and the value at a certain percentile. The normal distribution fit and graphical representation of the dataset enhance our understanding of its features and facilitate well-informed pay decision-making.