Salary Distribution And Statistics

Student Name : Hari Prasad Ghanapuram
Student Id: 22051910

This report details the analysis of a dataset containing annual salary information in Euros for a statistically representative sample of individuals in a European country. The goal is to create a probability density function, calculate the mean salary (W), determine a specific value X , and visualize the results on a histogram.

Dataset :
The dataset consists of a statistically representative sample of annual salaries. Each entry in the csv file represents one annual salary in Euros.

Distribution Analysis :
From the distribution we can see that the distribution of salaries is skewed to the right (positively skewed). The following formula is used to calculate pdf.

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}$$

Mean :
The mean salary (W) has been calculated as the average of all salaries in the dataset . The mean value that we got using dataset is 12579.48 Euros.

$$\frac{mean = \sum_{i=1}^{n} arr[i]}{n}$$

Value X calculation :
Calculated X as the 25th percentile of the salary distribution. This value, X, is such that 25% of the population has a salary above 7574.00 Euros.

X = percentile (data, 25)