Fundamentals of Data Science

Mohammed Nihad Kaipalli
22081746

The graph presents an overview of salary distribution in a European country based on a dataset of 4000 observations. Key observations include:

- **Distribution Overview:** The histogram illustrates a diverse range of income levels within the population.
- **Probability Density Function (PDF):** Smooth curve indicating the likelihood of different salary ranges.
- **Mean Salary:** Calculated at 26847.65, providing a central reference point.
- **Population Fraction:** Approximately 0.12, falls between the mean and **1.25 times the mean salary.**

Statistical Insights:

Mode: 12251

Median: 23247.5

Standard Deviation: 17100.633591636855

Kurtosis: 1.289985827348517

Skewness: 1.071820842386697

Range: 117271

Maximum Salary: 117477

Minimum Salary: 206

Mean Value

Mean can be calculated in the following way.

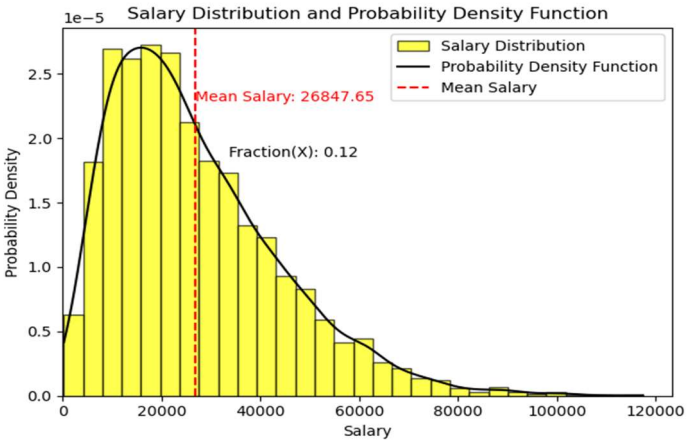
$$W^{\sim} = \frac{1}{n}\sum_{i=1}^{n} X\ i$$

- n is the total number of data points

- xi represents each individual salary

Value of X

The value of X represents the proportion of the population with salaries ranging between approximately the mean (W̃) and 1.25 times the mean. It is computed by dividing the number of salary data points falling within these bounds by the total count of salary observations. In this distribution, the calculated value of X stands at 0.12.



Conclusion:

In summary, the visualization and statistical analysis of the salary distribution highlight the diverse income landscape prevalent in the sampled European country. These insights can guide various stakeholders in making informed decisions regarding economic policies, social initiatives, and resource allocations.