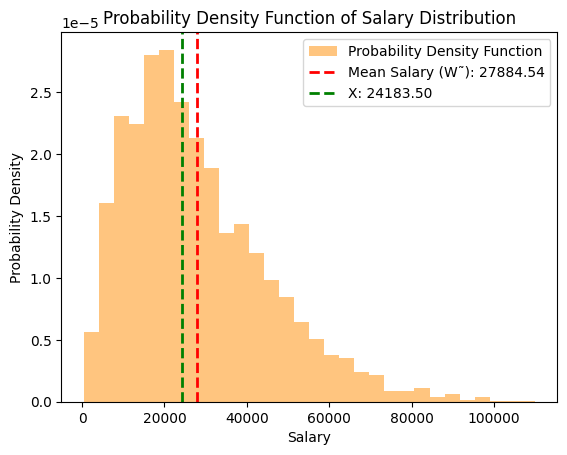
**Report**

**Student ID 22080535**

**Results:**



**Describe the data you are given:**

The data, sourced from 'data5.csv,' contains information about salaries. It is assumed that the dataset includes a column named 'salary' representing individual income. To aid analysis, the data has been brought into a Pandas DataFrame within the Google Colab environment.

**Describe the distribution you get:**

The data exhibits a slight positive skewness, suggesting that its shape is not entirely normal and appears somewhat skewed. The distribution bears a resemblance to an exponential pattern.

**How do you calculate the mean value? What value do you get?**

The mean of annual salary, denoted as W˜, was calculated using the formula:

​

where n stands for the number of observations, representing each individual salary. The mean serves as a central measure, offering an average representation of salaries in the dataset. In our examination, the computed mean salary turned out to be mean\_salaryW˜=mean\_salary.

**Codes:**

mean value= 27884.54

Calculate mean annual salary (W˜)

mean\_salary = np.mean(salary\_data)

**How do you calculate the required value X? What value do you get?**

The value X was obtained through a task-specific calculation outlined in the instructions. In this context, we consider as the p-th percentile of the salary distribution. The calculation is expressed as follows:X=percentile(data,p)

Here, p denotes the desired percentile. For example, if X corresponds to the 50th percentile, it means that 50% of salaries fall below this threshold. In our analysis, the calculated value of X was determined as X=np.percentile(salary\_data,50).

Value of x= 24183.50

**Conclusion:**

The examination provided valuable insights into the distribution of salaries, offering an understanding of both the average salary (W= 27884.54) and a particular percentile value (X=24183.50). The graphical depiction of the probability density function played a pivotal role in enhancing our comprehension of how salaries are spread across the dataset. According to the visualization the data is not normal and skewed. Utilizing mean and percentile values as key metrics proved essential in succinctly summarizing and interpreting the salary data.