Analysis Of European Salary Distribution

The dataset offered is a sample of yearly earnings in a European nation that is statistically representative. The annual wage in euros of everyone is represented by each entry in the dataset.

Probability Density Function:
We can see that probability density is highest around the mean salary, gradually decreasing as salaries deviate from the mean. The spread of salaries is relatively symmetric, suggesting a balanced distribution. Norm.pdf method from SciPy. Stats is used to calculate pdf. The formula is as follows.

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

Mean:
The mean salary (W) was calculated as approximately 39531.50 Euros, serving as a measure for the distribution. NumPy is used to calculate the mean. The formula is as follows.

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

X calculation:
The x value was calculating as approximately 82961.75. This implies that 5% of individuals have a salary above 9638.55. The calculation was performed using the NumPy library in python. The formula as follows.

X = percentile (data, 95)