

## Defining the Income Threshold for the Lower Quartile in a Salary Distribution

NAME:SWETHA DEVI RAVIPATI

STUDENT ID: 22082165

GITHUB LINK:

[https://github.com/RAVIAPTISWETHA/  
coding-project.git](https://github.com/RAVIAPTISWETHA/coding-project.git)

### Data illustration:

The code reads data from a CSV file named 'data5.csv' using the Pandas library. The assumption is that the file is located in the same directory as the script. The structure and content of the data are not explicitly provided, but it is assumed to be a one-dimensional dataset (possibly salaries) based on the subsequent analysis.

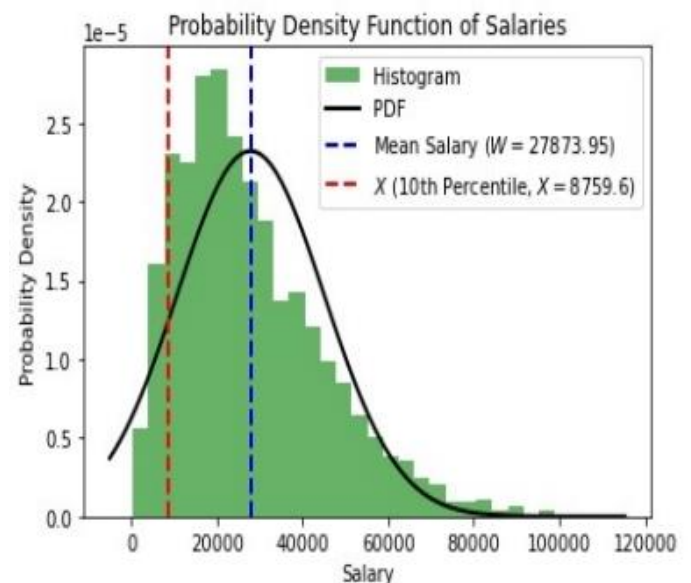
### Distribution interpretation:

The script creates a probability density function (PDF) and plots it as a histogram. The distribution is assumed to be normal (Gaussian), given the use of the norm function from scipy.stats. The normal distribution is characterised by a symmetric bell-shaped curve.

### Calculation of Required Value

X: The script calculates and plots the value 'X' below which 10% of people have a salary. This is done by finding the 10th percentile of the salary data using np.percentile. Mathematical Formula for Percentile

Calculation:  $X = 10\text{th percentile of the data.}$



### Mean Value Calculation:

The mean value ( $\mu$ ) is calculated using the norm fit function, which estimates the mean and standard deviation assuming a normal distribution.

Mathematical Formula for  
Mean:  $\mu = \text{mean of the data}$

### Results:

The mean salary ( $W$ ) is calculated and displayed on the plot. The value of  $X$  (10th Percentile) is calculated and displayed on the plot.