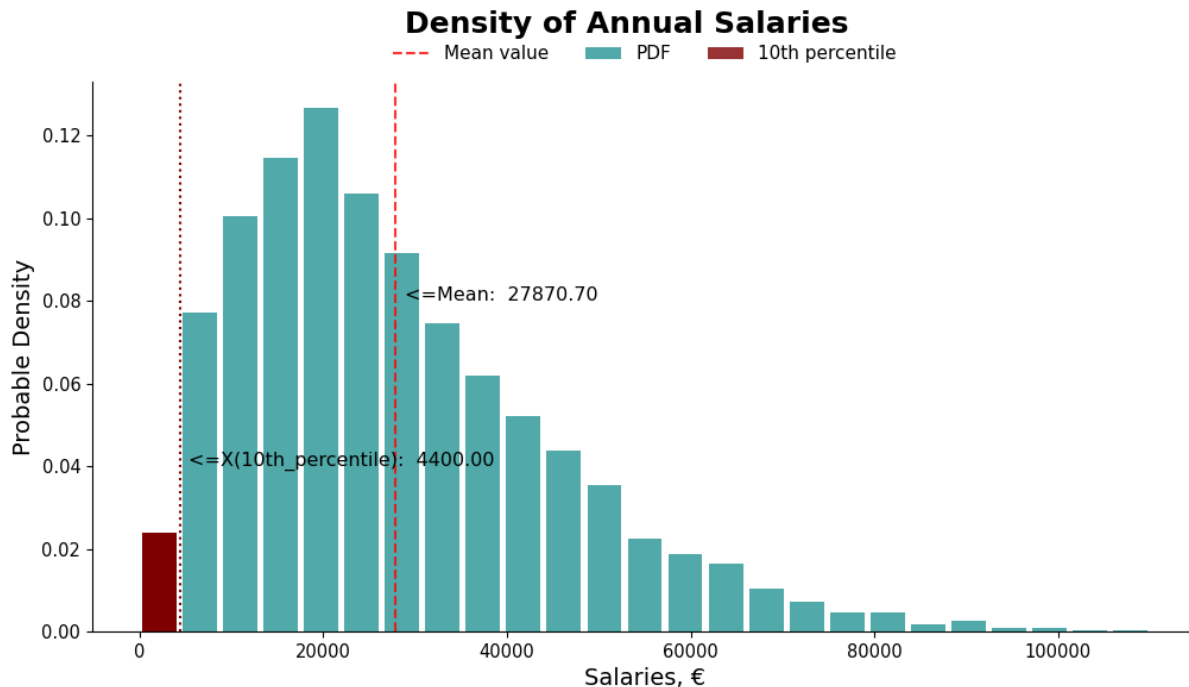


## Distribution of Annual salaries €

The dataset provided for the report included annual salaries in euros within a span of **4000** no zero values ranging from **€426 - €109,837**. The data also portrays a high dispersion between values with a standard deviation value calculated as **€17,185**. Also, the values are observed to be capped according to percentiles as **25%** values within **€15,185**, **50%** values at **€24,183** and **75%** values at **€34,237**.



The distribution of data highlighted the highest density of values to be in the **€18,000 to €22,000** range with a density value above **0.12**, demonstrating most individuals to be earning the lower-middle annual salary. The massive decline from **0.10** from **0.02** between **€25,000 to €60,000** portrays number of individuals decreasing in the upper-middle income range and finally a density **below 0.02** is observed in high incomes, ranging from **€70,000 to €110,000**.

The **mean value** was acquired by the summed product of center points of each distribution bar and their density acquired by the normalized distribution, with a resulting value of **€27,870**.

The **X value** was acquired by using index values of cumulative distributions below **10%** and using those indices to elicit and highlight values that fall below the 10<sup>th</sup> percentile value of **€4,400**. The python functions employed for this operation were:

- **Np.cumsum()**: To acquire a cumulative distribution of values within ranges.
- **Np.argmin()**: To acquire indexes that were below the density level of 0.10
- **Np.abs()**: To convert the acquired values in to absolute values for removing negatives.