7PAM2000 APPLIED DATA SCIENCE 1

Assignment 1 : Visualization

By : Aladesuru Thomas

ID : 22086778

<https://github.com/taladesuru/UH-codeworks>

This report shows generation of three visualization graphics with the first part of the code generating two plots. A Bar Chart and an Histogram for the income by Deciles for Men and Women with 1st degree qualifications. While the 3rd plot is a Line Graph that shows multiple business sizes plotted against the cost for compliance.

* **Bar Chart**: **Income by Decile for Men and Women with 1st degree qualifications.**

This chart was used to compare the earnings distribution of people with 1st degree qualifications by gender in the entire labor force. deciles was used to display the variations in income and the Bar Chart was especially chosen because it is known to be a great method when comparing values.

The Men and women's incomes are represented by side-by-side bars for each decile so as to easily identify the disparities in income for both genders.



* **Histogram: Distribution of Men's Income**

This histogram visualizes the frequency distribution of men's income. The bins represent income ranges, and the height of each bar shows the frequency of incomes falling within that range. The method was selected because of its renowned usefulness when distribution of income is involved.



The plot shows that men have higher income compared to women in each decile. This data can be used to advice stakeholders and policy makers to step in and create policies that address this disparities

The second part of the Python code creates the third visualization method which is a Line plot/graph as directed in the assignment brief.

* **Line Graph:** **One-off Compliance Costs to Business Size based on HMRC estimate of number of businesses with at least one graduate in Paye As You Earn (for 2014 uprated by growth rates from BIS SME statistics, assuming a linear growth rate between 05-09 continues to 2014)**

In this plot, a line graph is displayed showcasing the size of employees of the firm on the x-axis and three lines representing the number of firms employing graduates, the average cost per hour of a personnel officer and the average hourly cost of an IT personnel.

Each point on the lines is marked with a circle.

The legend indicates which line corresponds to each category

The Blue colored line represents the number of enterprises employing graduates in the year 2014

The Orange colored line represents the mean hourly cost of personnel officer in the year 2014

The Green colored line represents the mean hourly cost of IT personnel/technician.

Overall, this visualization helps in understanding how the number of enterprises employing graduates and costs of personnel and IT technicians vary across different sizes of firms.

What can be inferred from this graph is the fact that the data distribution is better understood thanks to the visual depiction and the distribution of income, the distribution of men's income, and the relationship between the size of firms and the number of enterprises employing graduates, as well as the associated costs. These visualizations can aid in drawing valuable business insights and making informed decisions based on the data and ways to reduce cost by adopting other strategies such as outsourcing if the need requires it.