Student Name:

Student ID:

7PAM2000 Applied Data Science 1

Assignment 1: Visualisation

Table of Contents

[Dataset Source 2](#_Toc118913432)

[Line Chart 2](#_Toc118913433)

[Histogram 3](#_Toc118913434)

[Pie Chart 4](#_Toc118913435)

[References 5](#_Toc118913436)

# Dataset Source

https://storage.googleapis.com/dft-statistics/road-traffic/downloads/data-gov-uk/region\_traffic\_by\_road\_type.csv%22

# Line Chart

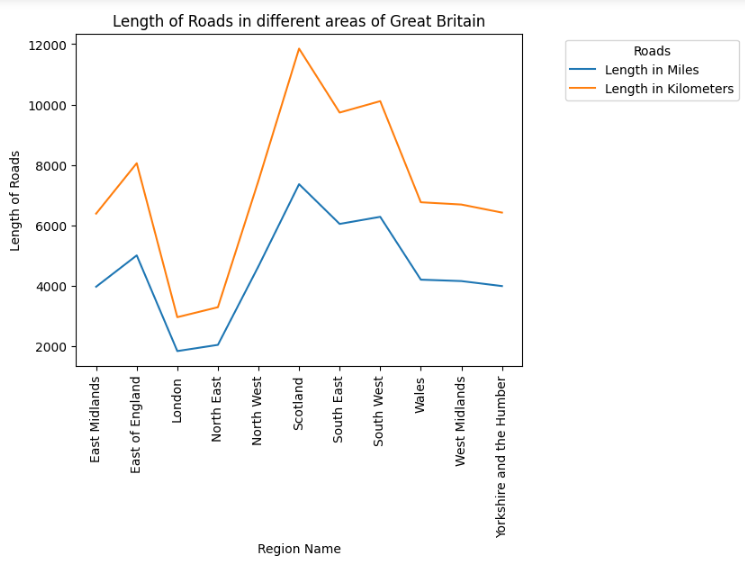
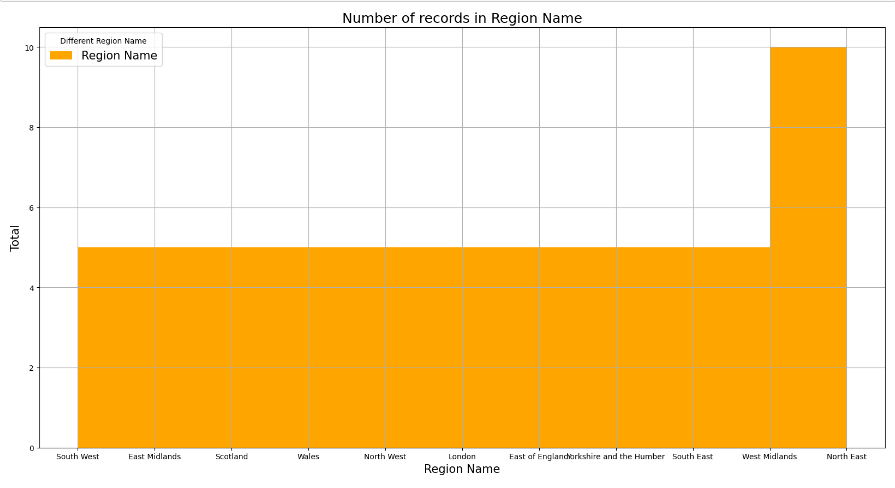


Figure 1: Line plot between length of roads and region name

Matplotlib can generate line charts right out of the box for you to use in your projects. Alterations to a line chart may be made in a variety of ways, including the addition of additional lines, alterations to the line's colour or type, and many more. There is a charting package available for the Python programming language that is named Matplotlib. Line charts are only one of the many kinds of charts that can be created with it.

In this line plot, the attributes that are being taken into consideration are the Length of Roads and the name of the Region. The plot displays the Length of Roads in each of the many Regions, displaying the length in both kilometres and miles. The name of the area is specified along the axis to the left of the graph's x-coordinate, and the various types of length are shown in the graph's legend using the colours. The correct labels have been placed along the x axis and the y axis of the graph, and it also has a title (Raoniar, 2022). The length in kilometres is shown in blue, while the length in miles is shown in blue. Both lengths are provided in parentheses. It is plain to see that length expressed in kilometres is greater than length expressed in kilometres.

# Histogram



It is possible to get a visual representation of the probability distribution of one-dimensional numerical data by using something called a histogram. Instead of displaying the values, histograms provide the data in the form of a frequency distribution. To do this, the whole value range is segmented into several bins, which may be thought of as intervals varying in width. After that, it determines the total number of values that fit into each category and displays the findings in a manner that is easy to comprehend (Singh, 2022).

The histogram that was just shown uses the property "Region Name" as its basis. It determines the total number of observations that may be found in each geographical area. The name of the area is written along the axis of the graph, and the number of observations is written along the y axis of the graph. It is clear that the total number of observations in each of the areas is the same, with the exception of the Northeast, which has the largest total number of observations. The graph has been labelled accurately, and the caption has been included. Additionally, the graph has been given the colour orange.

# Pie Chart

Chart, pie chart

Description automatically generated

The pie charts that can be found in Matplotlib may be constructed with the help of the pie() function. It's very possible that you'll like browsing through the Matplotlib gallery. Python programmers have access to a wide variety of plotting and charting options because of the matplotlib package. One of the charts that it is capable of making is a pie chart, but it can also produce many other types of charts. One data series at a time is all that can be shown in a pie chart. The amount of information included in one data series is represented by the size of a wedge in a pie chart. The size of the wedge is proportional to the total quantity of data. A pie chart presents the values of its individual data points as a proportion of the total area of the chart.

In the pie chart that was just shown, the attributes "Road Category" and "Motor Vehicles" were the ones that were selected. The primary information that can be gleaned from the graph is the number of vehicles that fall into each type of road. The distinct colours, as displayed in the legend along with their respective names, are used to designate each group. It is possible to draw the conclusion that the TN category has the greatest number of operational motor vehicles, whilst the PA category contains the smallest number of operational motor vehicles. The caption for the graph is presented appropriately inside the graph itself, complete with a variety of colours and values, and both axes are labelled with pertinent information. Every one of the functionalities that were included in the graph as a result of the matplotlib package.

# References

Raoniar, R. (2022) *Introduction to line plot - matplotlib, Pandas and Seaborn Visualization Guide (part 1)*. One Zero Blog. [Online] [Accessed on November 9, 2022]https://onezero.blog/introduction-to-line-plot-matplotlib-pandas-and-seaborn-visualization-guide-part-1/.

Singh, P. (2022) *Tutorial – Matplotlib Histogram – Naukri Learning*. Naukri.com. [Online] [Accessed on November 9, 2022]https://www.naukri.com/learning/articles/matplotlib-histogram/.