7PAM2000 Applied Data Science 1

Assignment 1: Visualisation 1

***Table of Contents***

[Code Overview of the UK Rainfall dataset 3](#_Toc128260643)

[First Visualisation: Line Plot showcasing rainfall in the UK over time 4](#_Toc128260644)

[Second Visualisation: The bar chart which showcases every rainfall in the UK by year 5](#_Toc128260645)

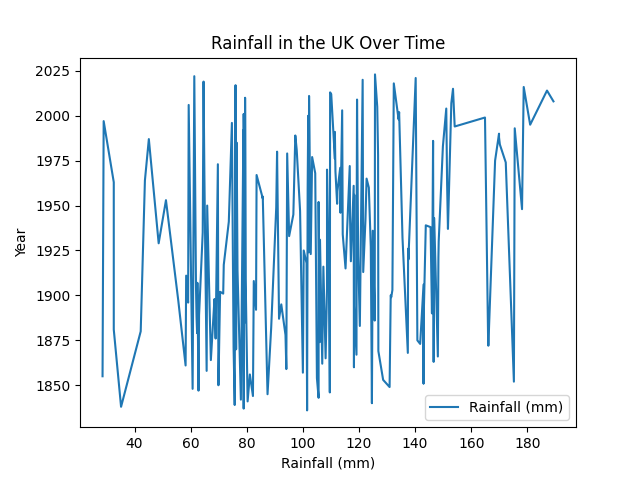
[Third Visualisation: Scatter Plot pinpointing rainfall in the UK by year 6](#_Toc128260646)

# Code Overview of the UK Rainfall dataset

The source code is a python script that utilizes libraries such as pandas and matplotlib to analyse the rainfall dataset for the United Kingdom. Initially the script imports all the necessary libraries, first the pandas library as pd for data manipulation and rainfall data analysis of the UK. After importing the libraries, the script then loads the data from the dataset URL into the pandas DataFrame object using the read\_csv function. The data is tab-separated with seven rows of header information, so the skiprows and header parameters are set accordingly. The resulting DataFrame is assigned to the variable df. Next, the script removes the last two columns of the DataFrame using iloc slicing to select only the first four columns. After this, the script renames the columns of the DataFrame to more descriptive labels using the columns attribute.

The code then creates three different plots of the data using the pandas plot() function. The first plot is a line plot of "Year" on the x-axis and "Rainfall (mm)" on the y-axis. This plot displays the trend of rainfall in the UK over time. The second plot is a bar chart of "Year" on the x-axis and "Rainfall (mm)" on the y-axis. This plot displays the amount of rainfall each year. The third plot is a scatter plot of "Year" on the x-axis and "Rainfall (mm)" on the y-axis. This plot displays the relationship between year and rainfall. Finally, each plot is labelled using title, xlabel, and ylabel methods of the plt object and displayed using the show() method.

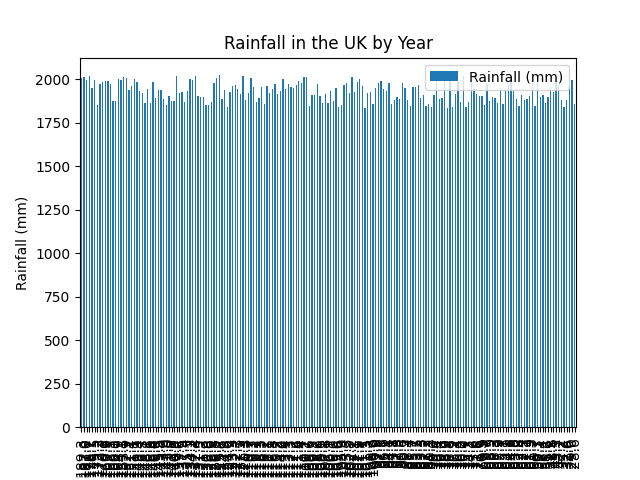
# First Visualisation: Line Plot showcasing rainfall in the UK over time



***Figure 1: Line Plot for rainfall in the UK***

The first visualisation justifies rainfall in the UK over time with the implementation of a line plot. A line plot is a graph that shows the trend or pattern of a set of data points over time. Its benefits include highlighting trends, identifying outliers, and showing changes in data over time. The figure interpretation suggests that the dataset has rainfall data for the years between 1850 – 2020. It can be visualised that the rainfall over the years lies between 40 mm and 180 mm. The average rainfall over the years is between 80 mm to 120 mm. The rainfall is highly influenced by the Atlantic Ocean and the North Atlantic drift due to which warm water arrives in the high northern latitudes and causes rainfall.

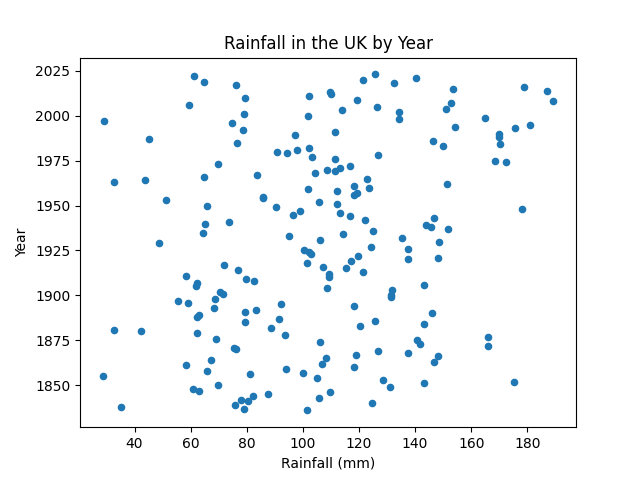
# Second Visualisation: The bar chart which showcases every rainfall in the UK by year



***Figure 2: Bar chart for rainfall in the UK***

The second visualisation generated by the python script provides a bar chart that provides every rainfall that occurred in each year from 1850 to 2020. A bar chart is a graph that compares different categories of data by using rectangular bars of different lengths. Its benefits include easy visualization of comparisons, highlighting differences, and identifying the highest or lowest values. It can be analysed that each bar represents the total rainfall (in mm) for the associated year. The rainfall in the UK varies widely from year to year. The year 2000 is recorded as the year with the highest rainfall of all other years whereas the year 1855 is recorded as the one with the lowest rainfall. The reason behind high rainfall within the UK is highly influenced by rainy seasons and in the regions with wet climates.

## Third Visualisation: Scatter Plot pinpointing rainfall in the UK by year



***Figure 3: Scatter Plot for rainfall in the UK***

The third visualisation from the python script depicts a scatter plot that identifies the rainfall in the UK over the years. A scatter plot is a graph that displays the relationship between two variables by using dots on a two-dimensional plane. Its benefits include showing patterns or trends, identifying potential correlations, and detecting outliers or unusual data points. It can be analysed that the figure indicates that the rainfall in the UK has adopted an increasing trend over the years. This also represents a weak relationship between the two variables (rainfall and years) as the scatter points do not follow the increasing trend in flow, for which, climate change might be a reason as it can directly affect rainfall in the UK.

<https://www.metoffice.gov.uk/research/climate/maps-and-data/data/index>