Name: **Hamza**

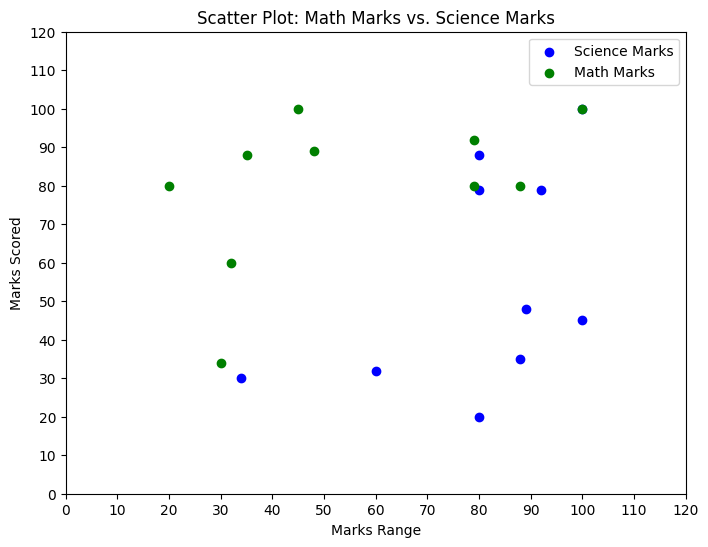
Assignment 1: Visualization

Git link: https://github.com/HamzaQureshi12/DataScience\_Assignment1.git

Data Source: Five Big companies Dataset, Programming Languages and two Subjects Dataset

1. **Scatter Plot - Math Marks vs. Science Marks**

I have created scatter plot to analyze the relationship between Science and Math marks. Each point on the plot represents a data entry with its values for the two variables. The x-axis represents marks scored in subjects and the y-axis represents range of marks.



**Data Description:**

This scatter plot visualizes the relationship between math and science marks. It is a suitable representation for comparing two continuous data sets.

**Benefits of Scatter Plot:**

This scatter plot can be used as a good example for a scatter plot that can visualize a relationship between two variables.

In our assignment, we can replace the sample data with a dataset that represents two related variables you want to compare.

But we should ensure that the x-axis and y-axis labels are meaningful based on our dataset, and provide a title explaining the context.

**Conclusions:**

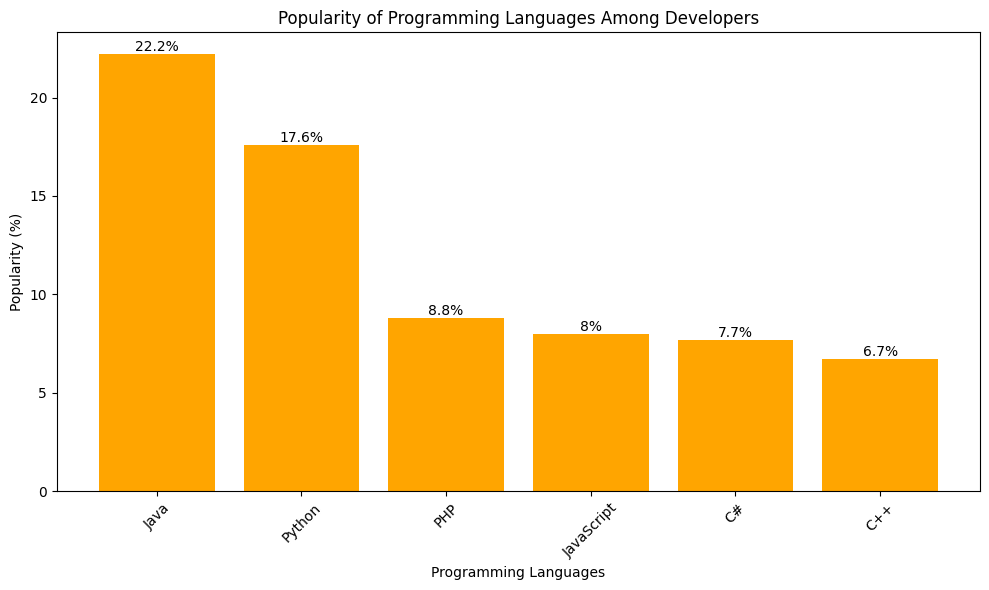
There appears to be a positive correlation between math and science marks.

Some students achieved high marks in both subjects, while others had lower marks in both.

A few outliers scored very high in one subject but much lower in the other.

2. Bar Chart - Programming Language Popularity

A bar chart was employed to compare Java, Python, PHP, JavaScript, C# and C++ programming languages based on their popularity among developers. Each bar represents a category, and the height of the bar corresponds to the value of the measurement.



**Data Description:**

This bar chart illustrates the popularity of various programming languages among developers. It is a suitable choice for comparing values in a categorical context.

**Benefits of Using Bar Chart:**

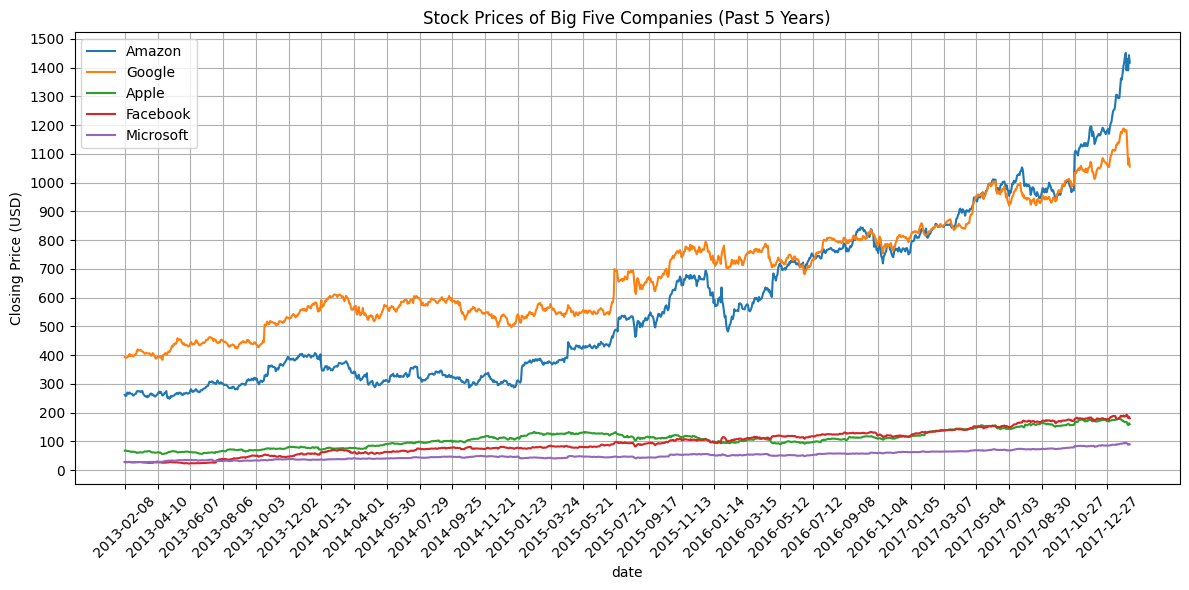
The bar chart I have created is suitable for comparing the popularity of programming languages, which is a good choice for categorical data.

We can use this code snippet in our assignment to visualize the popularity of something based on categories.

Replace the language names and popularity percentages with data from any dataset, and customize labels and titles accordingly.

3. Line Plot - Stock Prices of Big Five Companies

In this visualization, I have created a line plot to represent the trend of stock prices of five different companies over time. Multiple lines were used to show the trend for different companies. The x-axis represents date, and the y-axis represents the closing prices in USD. The legend indicates the different companies represented by the lines.



Data Description:

This line plot displays the historical stock prices of five major companies over the past five years. It is well-suited for visualizing time series data.

**Benefits of Using Line Plot:**

The line plot I have created can be used to visualize time series data, which aligns with our assignment requirements.

We can use this code to represent the stock prices of multiple companies over time.

Replace the company names, dates, and stock prices with data from any dataset and adjust labels and titles to match any other specific data.