TABLE OF CONTENTS

|  |  |  |
| --- | --- | --- |
| S.NO | NAME OF THE CONTENTS | PAGE.NO |
| 1. | ABSTRACT | 1 |
| 2. | INTRODUCTION | 2 |
| 3. | CODE | 5 |
| 4. | INPUT AND OUTPUT | 6 |
| 5. | CONCLUSION | 9 |
| 6. | REFERENCES | 9 |

STOCK MARKET ANALYSIS

USING DATA FRAMES

**1.Abstarct:**

In this document we are going to discuss about the data frames and its operations. And analyse the stock market data using data frames. The analysis framework encompasses data retrieval, preprocessing, exploratory data analysis (EDA), and the implementation of various statistical and machine learning models for predictive analysis. Leveraging pandas' robust data manipulation capabilities, the study aims to provide insights into historical stock market data and forecast future trends. The methodology encompasses techniques such as data cleaning, feature engineering, visualization, and model evaluation to uncover patterns and relationships within the data. Through practical examples and case studies, this research demonstrates the efficacy of pandas in conducting thorough stock market analysis, facilitating informed decision-making for investors and financial analysts.

Keywords: Stock market analysis, pandas, Python, data retrieval, data preprocessing, exploratory data analysis (EDA), statistical models, machine learning models, predictive analysis, data cleaning, feature engineering, data visualization, model evaluation

**2.Introduction:**

Data Frames are used to collect and store a large amount of data in the form of a table or rows and columns. It is like a 2-dimensional data structure or 2-dimensional array. We can create data using excel sheets, CSV files, lists of lists, lists of tuples, dictionaries, etc.

There are so many predefined functions, methods, and attributes in data frames. Let us see some of those now. Attributes

|  |  |
| --- | --- |
| at | Access a single value for a row/column label pair. |
| attrs | Dictionary of global attributes of this dataset. |
| axes | Return a list representing the axes of the DataFrame. |
| columns | The column labels of the Data Frame. |
| dtypes | Return the dtypes in the Data Frame. |
| empty | Indicator whether Series/Data Frame is empty. |
| flags | Get the properties associated with this pandas object. |
| iat | Access a single value for a row/column pair by integer position. |
| iloc | Purely integer-location based indexing for selection by position. |
| index | The index (row labels) of the Data Frame. |
| loc | Access a group of rows and columns by label(s) or a boolean array. |
| ndim | Return an int representing the number of axes / array dimensions. |
| shape | Return a tuple representing the dimensionality of the Data Frame. |
| size | Return an int representing the number of elements in this object. |
| style | Returns a Styler object. |

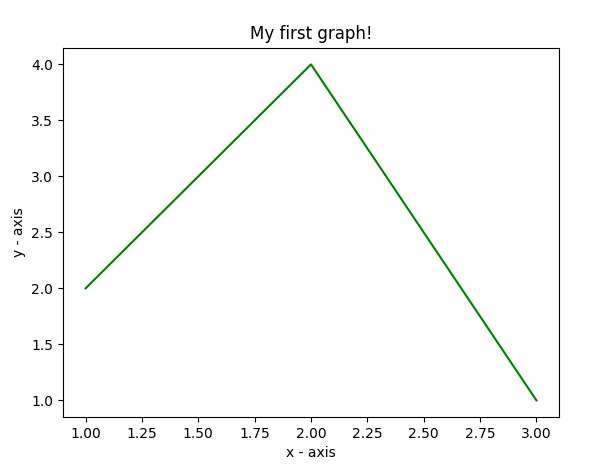
values Return a Numpy representation of the Data Frame.

Methods in data frames:

1. The describe () method returns description of the data in the Data Frame.If the Data Frame contains numerical data, the description contains this information for each column:etc Data visualization plays a significant role in the representation of both small and large data sets count - The number of not-empty values.

|  |  |  |
| --- | --- | --- |
| mean - The | average (mean) | value. |
| std - The | standard | deviation. |
| min - the | minimum | value. |
| 25% - The | 25% | percentile\*. |
| 50% - The  75% - The 75% percentile\*. | 50% | percentile\*. |

Data Visualization is the presentation of data in pictorial format. The process of finding trends and correlations in our data by representing it pictorially is Data Visualization. To perform data visualization in python, we can use various python data visualization modules such as Matplotlib, Seaborn, Plotly etc. Data visualization plays a significant role in the representation of both small and large data sets. Example:



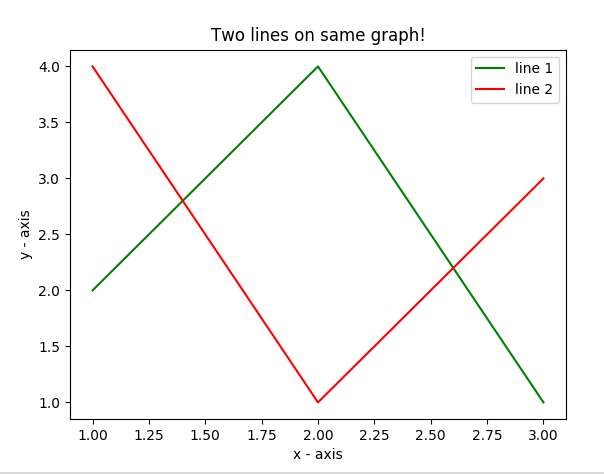
Define the x-axis and corresponding y-axis values as lists.

Plot them on canvas using. plot () function.

Give a name to x-axis and y-axis using. xlabel() and .ylabel() functions.

Give a title to your plot using. title () function.

Finally, to view your plot, we use. show () function.

 Here, we plot two lines on the same graph. We differentiate between them by giving them a name(label) which is passed as an argument of the. plot () function.

The small rectangular box giving information about the type of line and its color is called a legend. We can add a legend to our plot using. legend () function.

There are four types of financial markets. They are

1. Currency markets
2. Money markets
3. Derivative markets
4. Capital markets

Capital markets are used to sell equities(stocks), debt securities. These are also called as stock markets. It is a place where different financial instruments are traded between different entities. We are going to analyse the stock market data using stacks. A stock market or equality market is a market for trading company stocks also known as shares and derivatives at an agreed price. Shares represent a fraction of ownership in the company or business. The common feature of all these is equity participation. Different classes of shares have different voting rights. Ownership of the shares is documented by a legal document that specifies the number of shares owned by the shareholder, and other details of shares, such as par value or the class of the shares.

A shareholder is a person or a company that legally owns one or more shares of the company. Shareholders are granted privileges depending on the class of stock, including the right to vote on matters such as elections to board of directors, the right to share in distribution of the company’s income, the right to purchase new shares issued by the company, the right to company’s assets during a liquidation of the company.

**3.Code:**

import pandas as pd import matplotlib.pyplot as plt

AMZN=pd.read\_csv("D:\python\AMZN.csv")

AAPL=pd.read\_csv("D:\python\AAPL.csv")

GOOGL=pd.read\_csv("D:\python\GOOGL.csv")

AMZN.describe()

AAPL.describe()

GOOGL.describe()

AAPL.set\_index('Date',inplace=True)

AMZN.set\_index('Date',inplace=True)

GOOGL.set\_index('Date',inplace=True) AAPL["Close"].plot() plt.title("APPLE STOCKS") plt.show()

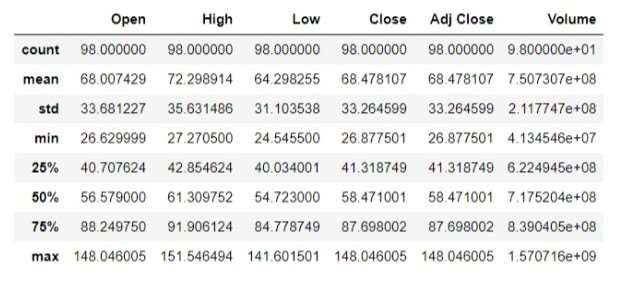
AMZN["Close"].plot() plt.title("AMAZON STOCKS") plt.show()

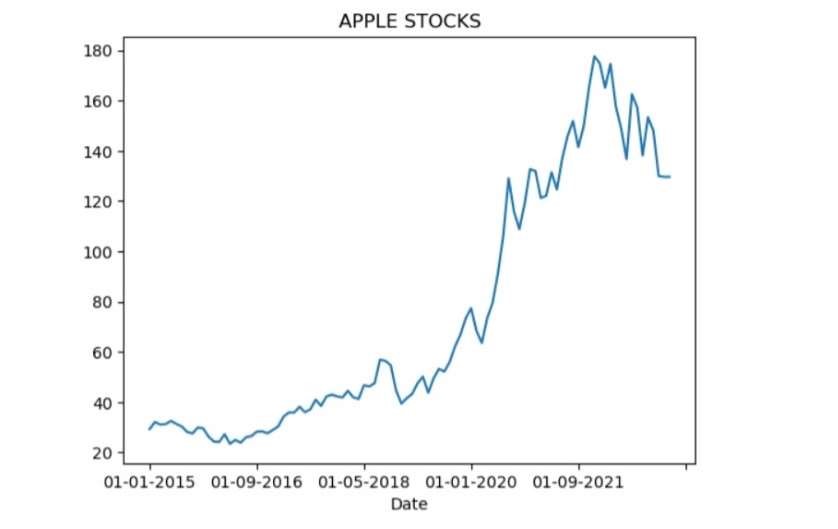
GOOGL["Close"].plot() plt.title("GOOGLE STOCKS") plt.show()

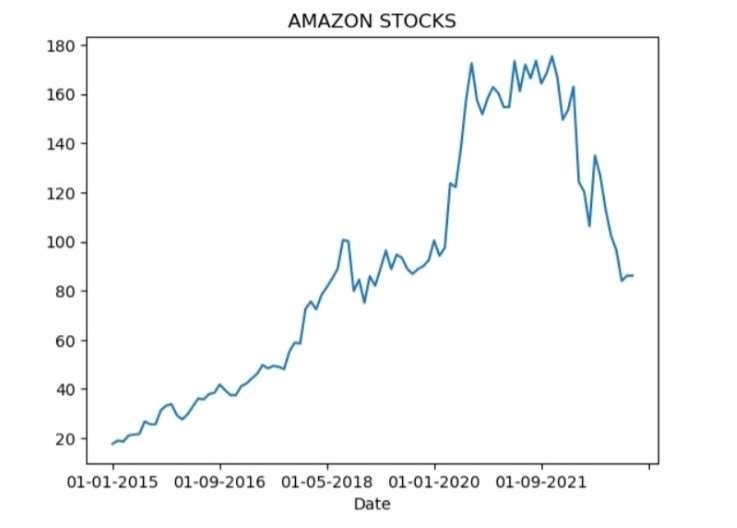
AAPL["Close"].plot(label="APPLE Close")

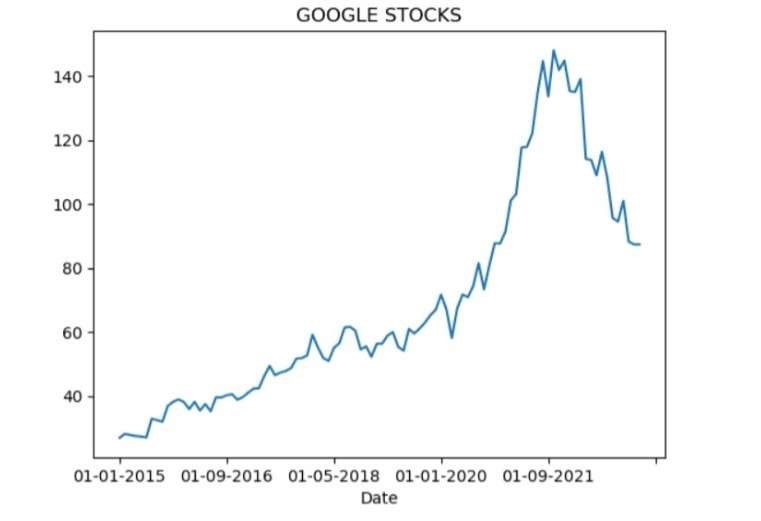
AMZN["Close"].plot(label="AMAZON Close") GOOGL["Close"].plot(label="GOOGLE Close") plt.title("MNC STOCKS CLOSING") plt.ylabel("STOCKS") plt.show()

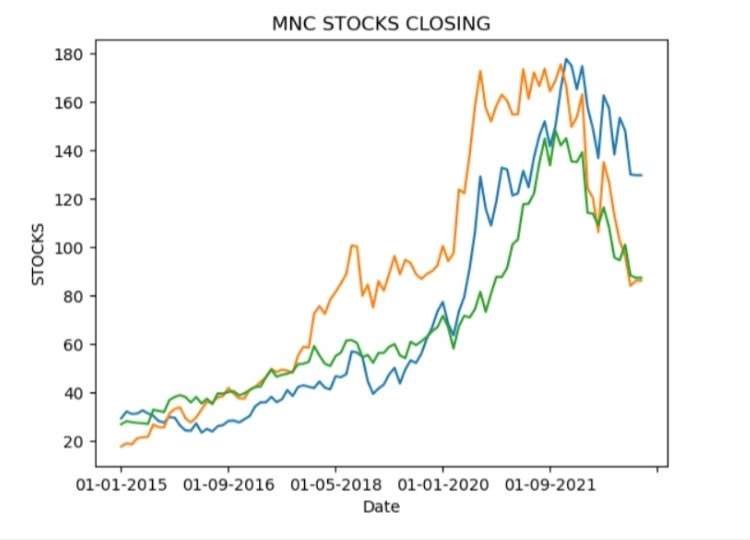
**4.Input and Output:**











**5.Conclusion:**

A Data frame is a two-dimensional data structure, i.e., data is aligned in a tabular fashion in rows and columns. Features of data frames are Potentially columns are of different types, Size – Mutable, Labelled axes (rows and columns), Can Perform Arithmetic operations on rows and columns. Data frame is a Pandas object. To create a dataframe, we need to import pandas. Data frame can be created using data frame () function. Data visualization deals with visual representation of data. It graphically plots data and is an effective way to communicate inferences from data. matplotlib and Seaborn are python libraries that are used for data visualization. They have inbuilt modules for plotting different graphs.

**6. References:**

1. https://www.kaggle.com
2. https://www.jovian.com
3. https://www.scribd.com
4. https://github.com