

Deccan Education Society’s

NAVINCHANDRA MEHTA INSTITUTE OF TECHNOLOGY AND DEVELOPMENT

**NAAC Accredited “B++”**

**TITLE OF THE PROJECT**

**Time Series Project**

**Stock Market case-study**

**SUBMITTED BY**

**Aditya Arun Deshpande  
Roll no : C21027**

**Under the guidance of**

**Dr.Sulakshana Vispute**

**[ ACADEMIC YEAR ]**

**2022-23**

 Submitted to University of Mumbai

in partial fulfillment of the requirements for qualifying

MASTER OF COMPUTER APPLICATION   
 Examination

**Deccan Education Society’s**

NAVINCHANDRA MEHTA INSTITUTE OF TECHNOLOGY AND DEVELOPMENT

PROJECT CERTIFICATE

This is to certify that the Project done at Udemy by Mr. Aditya Arun Deshpande

(Seat No.C21027) in partial fulfillment for MCA Degree Examination has been

found satisfactory. This report had not been submitted for any other examination and does not

form part of any other course undergone by the candidate.

Internal Guide Director

EXAMINED BY

EXTERNAL EXAMINER ………………………………

DATE:

College Stamp



# Acknowledgement

|  |
| --- |
| I ,**Aditya Arun Deshpande** student of **Navinchandra Mehta Institute of Technology** **and Development** would like to thank the college and Incharge Director **Dr. Anita** **Bobade** and **Dr. Rasika Mallya**, **HOD,** for giving us best in class knowledge for the enrolled course.  Secondly, I would like to thank my Project Guide, **Dr. Sulakshana Vispute**, for her immense support and continuous encouragement to me. Without her support, it would have been difficult for me to complete the project.  I would also like to thank the entire college, the faculties, the librarians and the non-teaching staff for their support and cooperation and helping students to achieve their goals. |

**List of Figures, if any:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Sr. No.** | **Figure Number** | **Title** | **Page Number** |
| **1** | **1.1** | **System Overview(flowchart)** | **11** |
| **2** | **1.2** | **Module Hierarchy/ Work Breakdown Structure** | **13** |
| **3** | **1.3** | **Data Preprocessing specification** | **14** |
| **4** | **1.4** | **System Flow Chart/ Activity Diagram** | **16** |
| **5** | **1.5** | **Chart of different stocks ( Closing price )** | **18** |
| **6** | **1.6** | **daily returns in % of apple company stock** | **18** |
| **7** | **1.7** | **Heatmap of co-relation** | **19** |

**EXECUTIVE SUMMARY OF THE PROJECT**

# Stock market is always a mysterious place for Indian investors. Most of the Indian investors are traditional and conservative when it comes to investment options. Stock market is never an option for most. Reason for this is we Indians do not have in hand experience in stock market. Stock market is considered to be the riskiest and dangerous place to invest, a common misconception. This project deals with technical analysis of selected Blue-chip companies which are highly performing in NSE. Technical analysis along with fundamental analysis is the key to attain great profit from any investments in stock exchange. This project tries to give a brief information about stock market and a in depth view of technical analysis tools like Bollinger bands and Relative Strength Index. Both the tools are well used alone among investors. This project tries to analyze the results when both Bollinger and RSI is combined. This project gives a researcher’s version of conclusion to all the analysis that have been done during the research period. Analysis shows how an investor can trade with strong stocks and how to predict market direction. This can help Ventura securities to advise and educate their clients to use combined technical tools to have a safer trading section. Stock market is the most potent way of earning high profit than any other investments, risk involved with trading in stock market is also very high if traded blindly. Deal it with great research and analysis, stock market could bring fortune to an investor. INDEX

**Data Analytics**

|  |  |  |  |
| --- | --- | --- | --- |
| **Chapter** |  | **Contents** | **Page No.** |
| **I** |  | **Introduction** | 8 |
|  | 1.1 | Company Profile |  |
|  | 1.2 | Legacy System |  |
|  | 1.3 | Background (Domain Knowledge required) |  |
|  |  |  |  |
| **II** |  | **Proposed System** | 9 |
|  | 2.1 | About the Existing System, if any (optional) |  |
|  | 2.2 | Scope of Work |  |
|  | 2.3 | Operating environment |  |
|  | 2.4 | Proposed System | 10 |
|  | 2.5 | System Overview(flowchart) | 11 |
|  | 2.6 | Limitations of the System (legacy system) | 11-12 |
|  |  |  |  |
| **III** |  | **Analysis and Design** | 13 |
|  | 4.1 | Module Hierarchy/ Work Breakdown Structure |  |
|  | 4.2 | Data Preprocessing specification | 14 |
|  | 4.3 | System Flow Chart/ Activity Diagram | 16 |
|  | 4.4 | Algorithm’s flowchart | 16 |
|  | 4.5 | Algorithm specification(why to use) |  |
|  |  |  |  |
| **IV** |  | **Data Dictionary** | 17 |
|  |  |  |  |
| **V** |  | **Implementation and Evaluation(joins,stored procedures,functions) / Testing of the System**  **(Software Customization, User Interface (screen shot) and User Manual, Testing.)** | 20-44 |
|  |  |  |  |
| **VI** |  | **Conclusion** | 45 |
|  |  |  |  |
| **VII** |  | **Limitations and Future Enhancements** | 46 |
|  |  |  |  |
| **VIII** |  | **References** | 47 |
|  |  |  |  |

**I. Introduction**

**1.1 Company Profile**

With the permission of my mentor, I am doing a home project on role of ‘Time Series project’ and project title is “Stock market case study”.

**1.2 Legacy System**

BSE, the first ever stock exchange in Asia established in 1875 and the first in the country to be granted permanent recognition under the Securities Contract Regulation Act, 1956, has had an interesting rise to prominence over the past 143 years.  
  
The journey of BSE is as eventful and interesting as the history of India's securities market. Following are some of the key milestones and achivements.

**1.3 Background (Domain Knowledge required)**

1. **Python Programming**: A good grasp of Python programming language fundamentals, including data types, variables, loops, conditional statements, functions, and file handling, is essential.

2. **Data Visualization**: data visualization libraries such as Matplotlib, Seaborn, or Plotly enables effective exploration and presentation of textual data insights.

**II. Proposed System**

* 1. **Scope of Work**

1. Demand and Supply:  
   Demand and Supply of shares of a company is a major reason price change in stocks. When Demand Increase and Supply is less, price rises. and vice versa.
2. Corporate results: This will be regarding to the profits or progress of the company over a span of time say 3 months.
3. Popularity: Main Strength in hands of share buyer. Popularity of a company can effect on buyers. Like if any good news of a company, may result in rise of stock price. And a bad news may break dreams.
4. The stock value depends on other factors as well, but we are taking into consideration only these main factors.

**2.3 Operating environment**

**1. Data Collection**: Nowadays stocks data is available on different dataset websites like kaggle.

So, we have companies stock data dataset like Date of stock , Open price of that stock on that day, Close price of that stock on that day , High value of that day, Low value of that day, Name of that stock etc.

**2. Data Preprocessing**: Clean and preprocess the collected text data to make it suitable for analysis. This may involve tasks such as removing irrelevant information, converting text to numbers, removing null values , Checking null values, handling raw data .

**3. Exploratory Data Analysis (EDA)**: Perform exploratory analysis on the collected data to gain insights into the operating environment. This can include analyzing data frequencies, chart analysis, and visualizing trends or patterns in the data. Using Jupyter Notebook's data visualization libraries such as Matplotlib, Seaborn, or Plotly to create informative charts and graphs.

1. **Understanding use case :** Understanding the data what type of data is all about.
2. **Run ETL Pipeline :** In ETL, E means Extract the data means we need to extract the data , T means Transform the data , L means load the data. These all use for preparing the data.

**6. Reporting and Visualization**: Present your findings in a clear and concise manner using Jupyter Notebook. Utilize markdown cells to provide explanations, insights, and visualizations to support your analysis. Incorporate relevant charts, graphs, or word clouds to enhance the visual representation of your results.

**2.4 Proposed System**

**1. Setting up your environment**: Install Jupyter Notebook and the required Python libraries such as pandas, numpy, matplotlib, and seaborn for data analysis and visualization.

**2. Acquiring the Stock data**: obtaining The data using the kaggle website, which allows to access information of different datasets and allow to download it .

**3. Fetching YouTube data using the API**: Writing Python code to connect to the stock Data data such as stock price,high,low,open,close,volume,date,name of stock etc. information. Storing the data in a structured format we used a pandas liabrary for DataFrame.

**4. Perform exploratory data analysis (EDA)**: Using pandas, numpy, matplotlib, and seaborn to perform exploratory data analysis on the Stock data. Exploring various aspects such as stock price, high, low, open, close, volume, date, name of stock.

**5. Visualize the data**: Utilize matplotlib and seaborn to create visualizations such as bar plots, line charts, scatter plots, histograms, word clouds, and sentiment analysis plots. These visualizations will help to gain insights into the data and communicate findings effectively.

**2.5 System Overview(flowchart)**

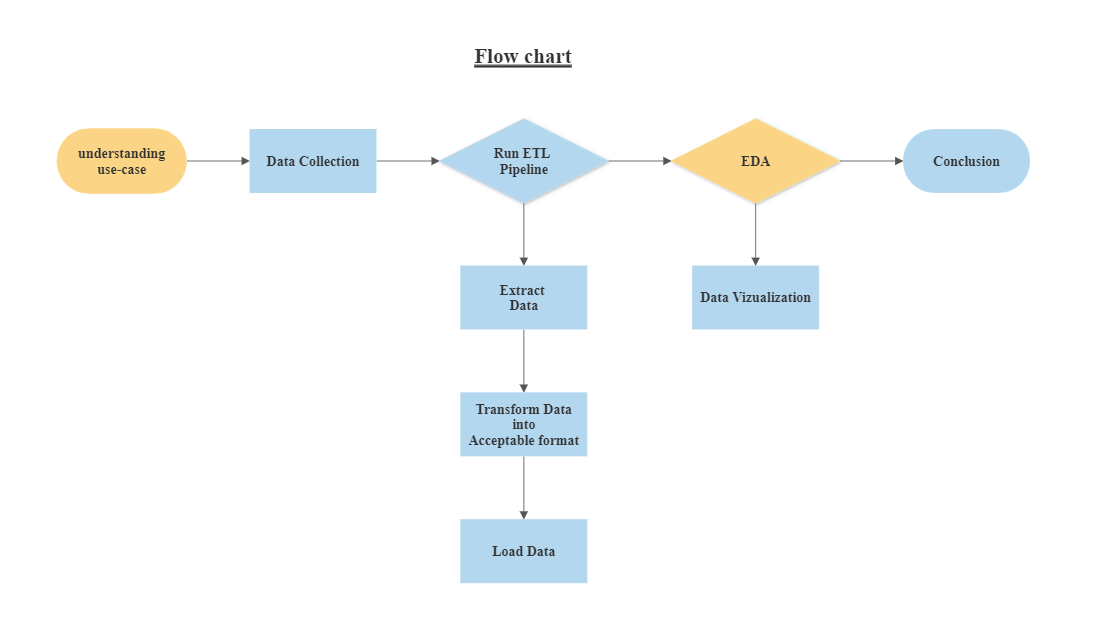


Fig – 1.1

**2.6 Limitations of the System (legacy system)**

**1. Outdated Technology**: Legacy systems often use outdated technology and programming languages that may not be compatible with modern analysis tools and frameworks. This can make it challenging to integrate and extract data from the legacy system effectively.

**2. Limited Scalability**: Legacy systems are typically not designed to handle large-scale data processing. As stock market generates a massive amount of data, legacy systems may struggle to scale and process the data efficiently. This can result in slower analysis and processing times.

**4. Data Format Incompatibility**: Legacy systems often use proprietary data formats or databases that are not easily compatible with modern analysis tools and platforms. This can require additional effort to convert and transform the data into a suitable format for analysis.

**5. Security Vulnerabilities**: Older legacy systems may have security vulnerabilities that make them more susceptible to cyber threats and attacks. This can be a significant concern when working with sensitive YouTube data that needs to be protected.

**6. Limited Analytical Capabilities**: Legacy systems may lack advanced analytical capabilities and algorithms required for sophisticated stock market analysis. This can restrict the types of analysis you can perform and limit the insights you can derive from the data.

**7. Lack of Support and Maintenance**: As legacy systems are often outdated and no longer actively supported by vendors, it can be challenging to find support or receive updates. This can lead to difficulties in troubleshooting issues and fixing bugs during the analysis project.

**8. Integration Challenges**: Integrating a legacy system with modern analysis tools, data pipelines, or third-party services can be complex and time-consuming. Incompatibilities in protocols, data exchange formats, or connectivity can hinder smooth integration, leading to delays and additional development effort.

**9. Case-sensitive :** Python is a case−sensitive programming language. This means that it considers uppercase and lowercase letters differently. As a result, in Python, we cannot use two terms with the same characters but different cases interchangeably.

**III. Analysis and Design**

4.1 Module Hierarchy/ Work Breakdown Structure

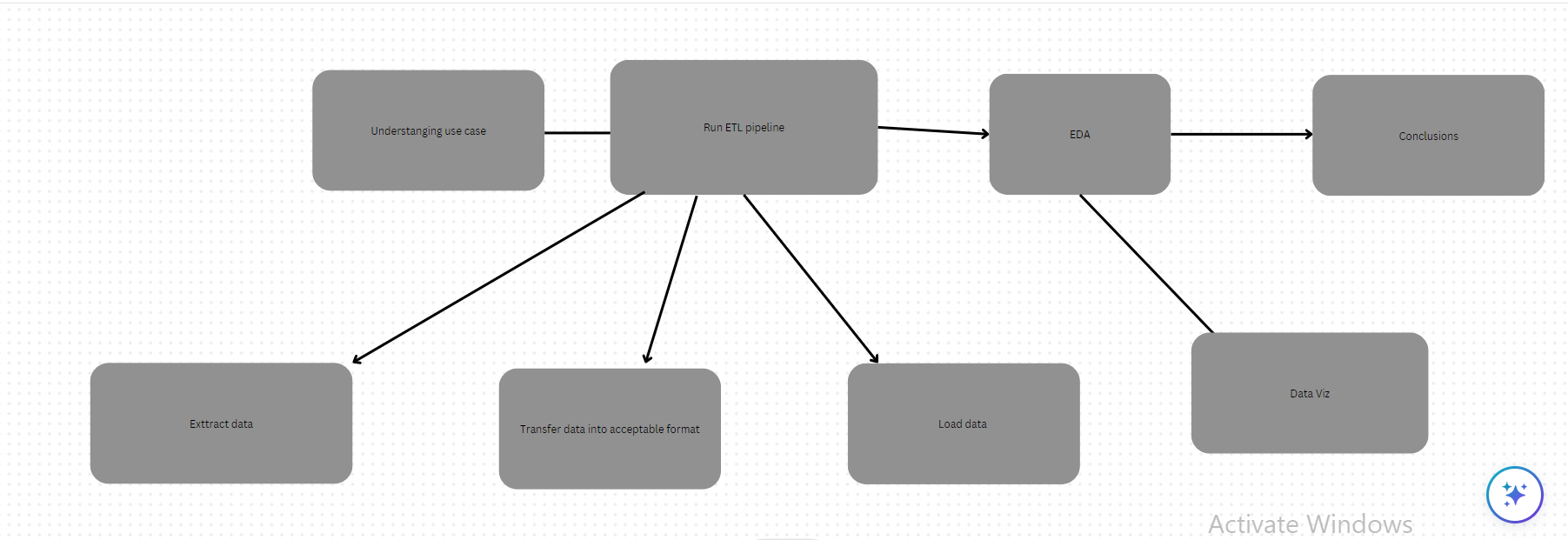
****

Fig – 1.2

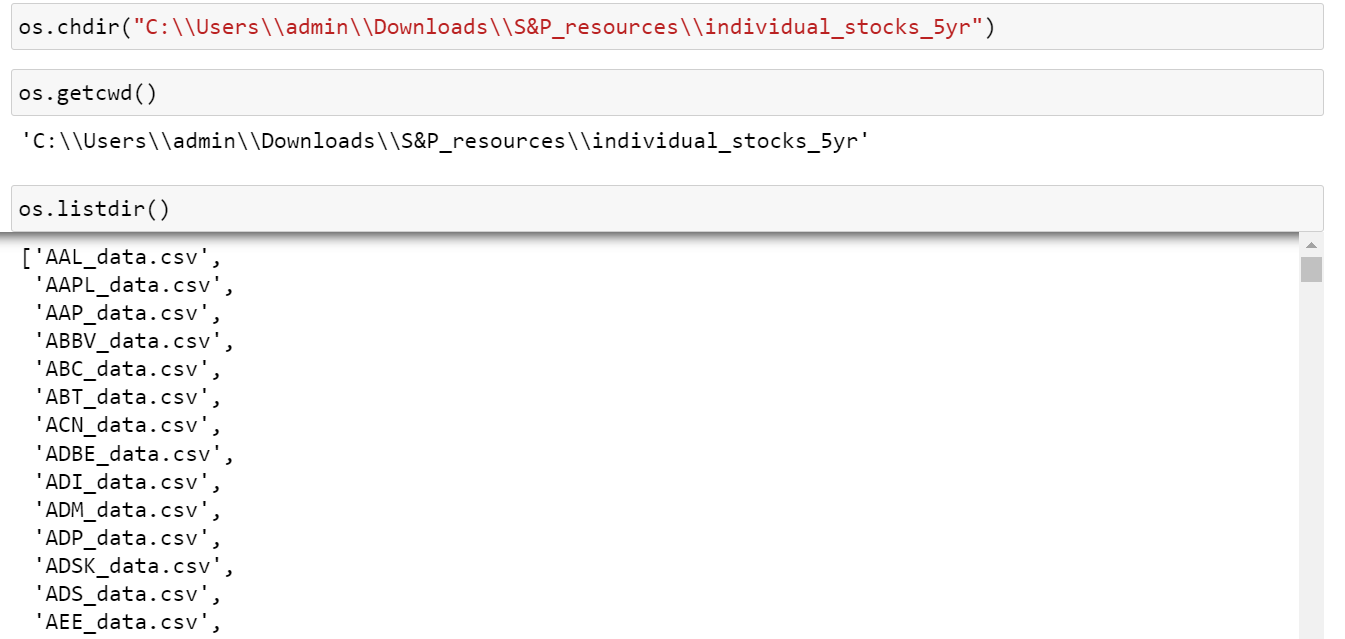
**4.2 Data Preprocessing specification**



Fig – 1.3

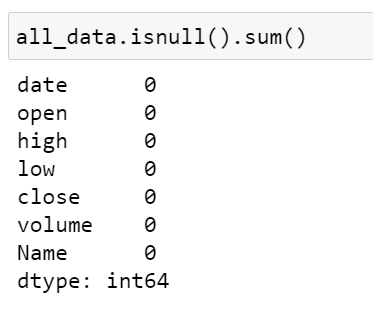
**Data Cleaning :**

After some simple manipulations and loading of the csv data into pandas DataFrame, we have the following dataset where open, high, low and close represent prices on each date and volume the total number of shares traded.



**Missing values** : missing values are not present which we confirmed by running the following command:



****

**Outliers :**   The key point to understand here is that our dataset now includes prices but prices are not the metric . will attempt to forecast because they are measured in absolute terms and therefore harder to compare across time and other assets.

**4.3 System Flow Chart/ Activity Diagram**

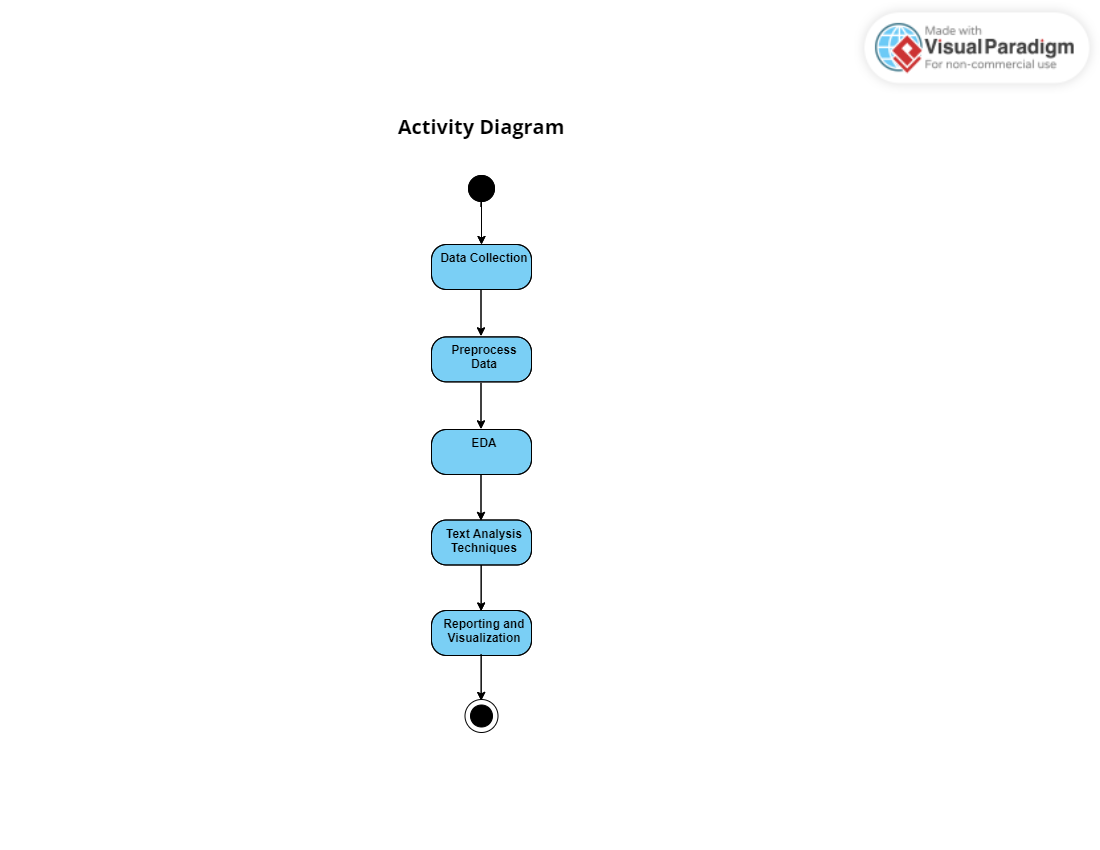
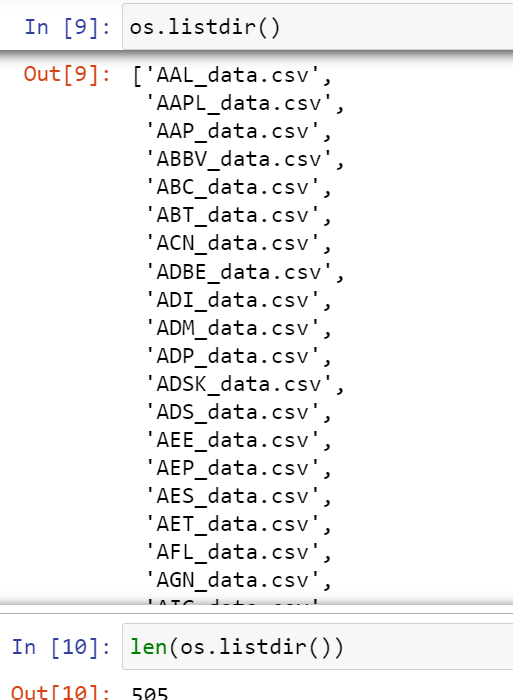


Fig – 1.4

**IV. Data Dictionary**

A data dictionary for a Stock market analysis project would contain information about the specific data elements and Dattributes related to the Stock text data being analyzed.

1. Date : Date of the stock which generated the data on a particular date.
2. Open : Open price of the particular stock which is in the numeric form.
3. High : High value of particular date of that stock.
4. Low : Low value of particular date of that stock.
5. Close : Close value of particular date of that stock.
6. Volume : volume of particular date of that stock on that particular day generated.
7. Name : Name of the stocks like Apple , Microsoft , Google , Amazon etc..



**# Chart of different stocks ( Closing price )**

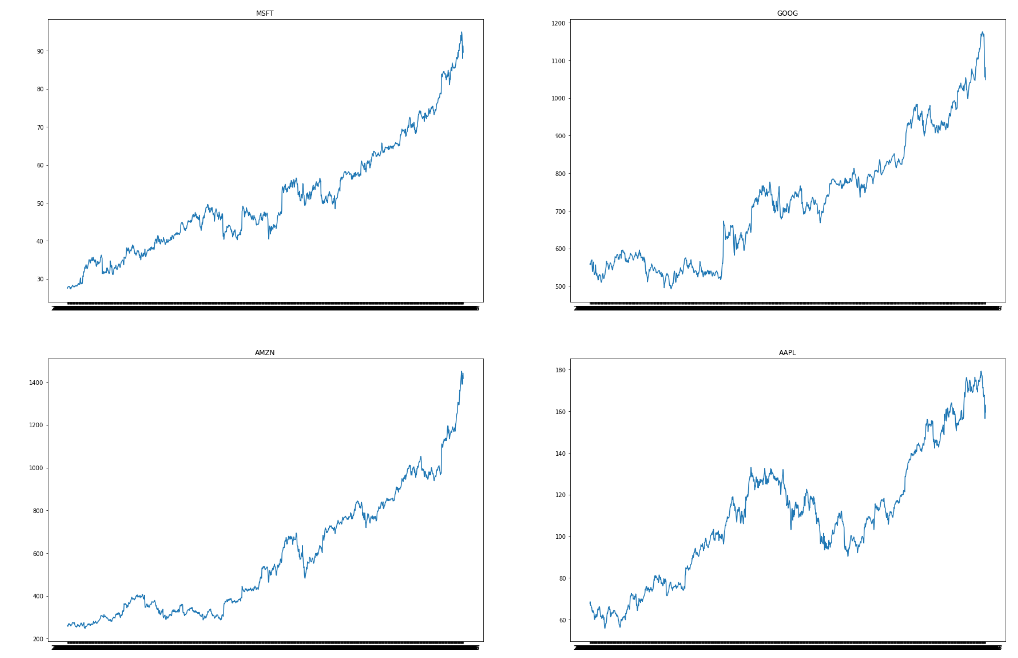


Fig – 1.5

**# daily returns in % of apple company stock**

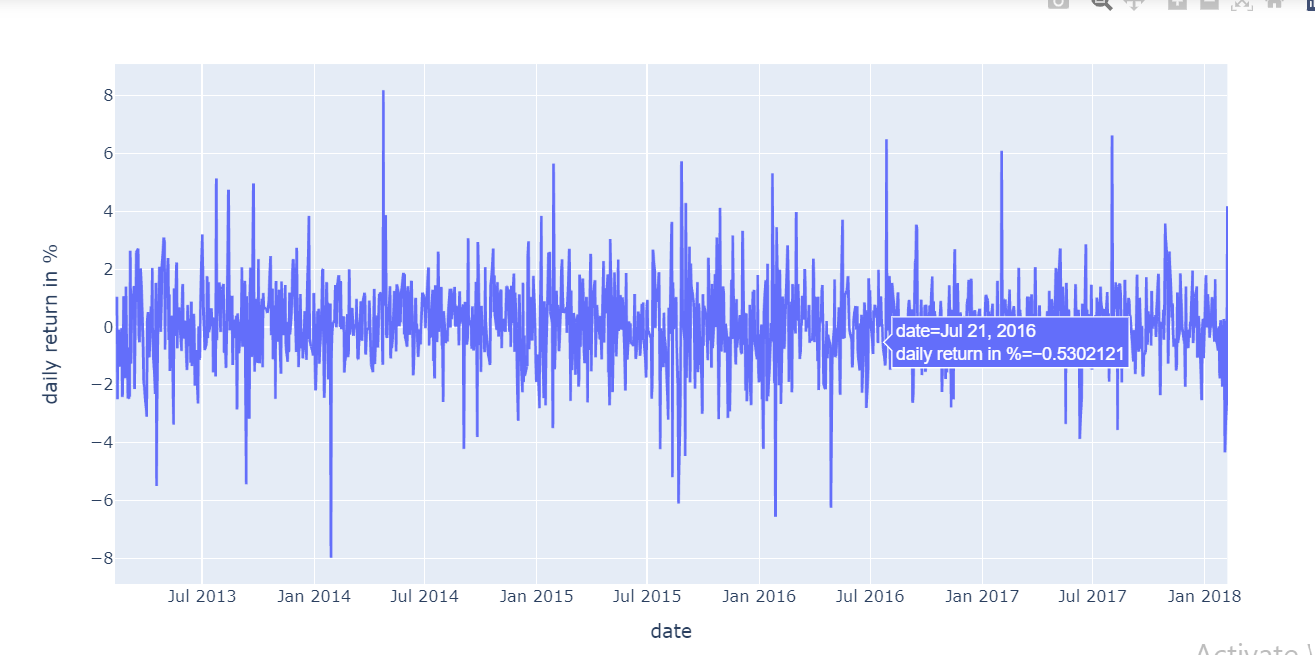
****

Fig – 1.6

**# Heatmap of co-relation**

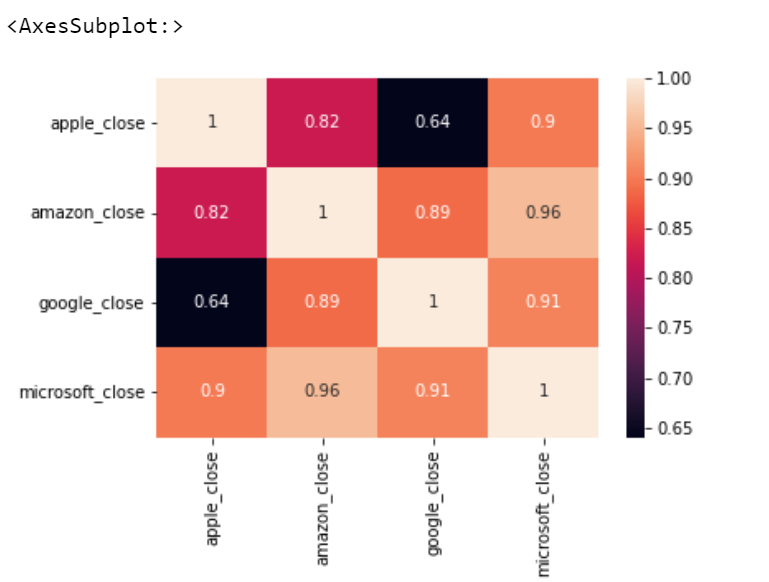
****

Fig – 1.7

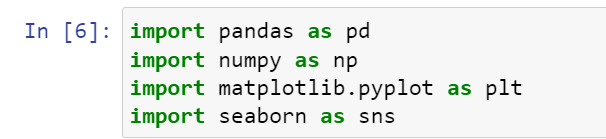
**V- Implementation and Evaluation(joins,stored procedures,functions) / Testing of the System**

**(Software Customization, User Interface (screen shot) and User Manual, Testing.)**

* **Importing csv files into the python tool :**

****

* **Importing libraries :**



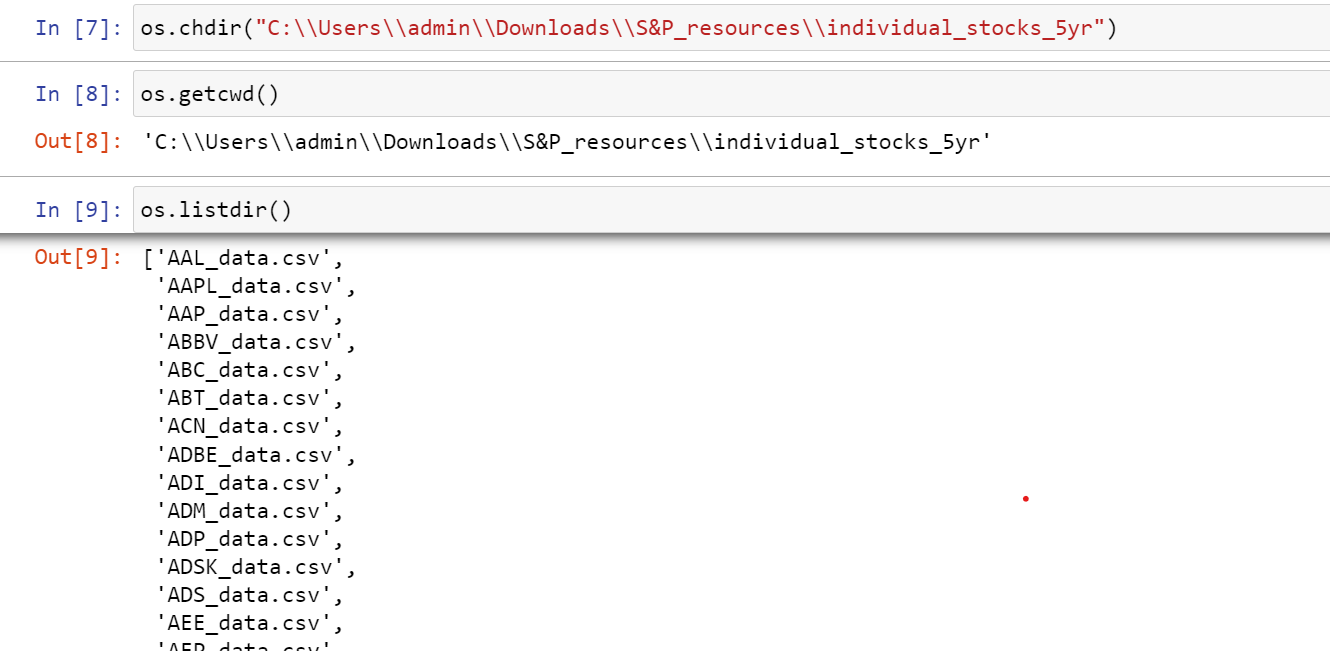
**Pandas :** Pandas is a Python library used for working with data sets. It has functions for analyzing, cleaning, exploring, and manipulating data.The name "Pandas" has a reference to both "Panel Data", and "Python Data Analysis" and was created by Wes McKinney in 2008.

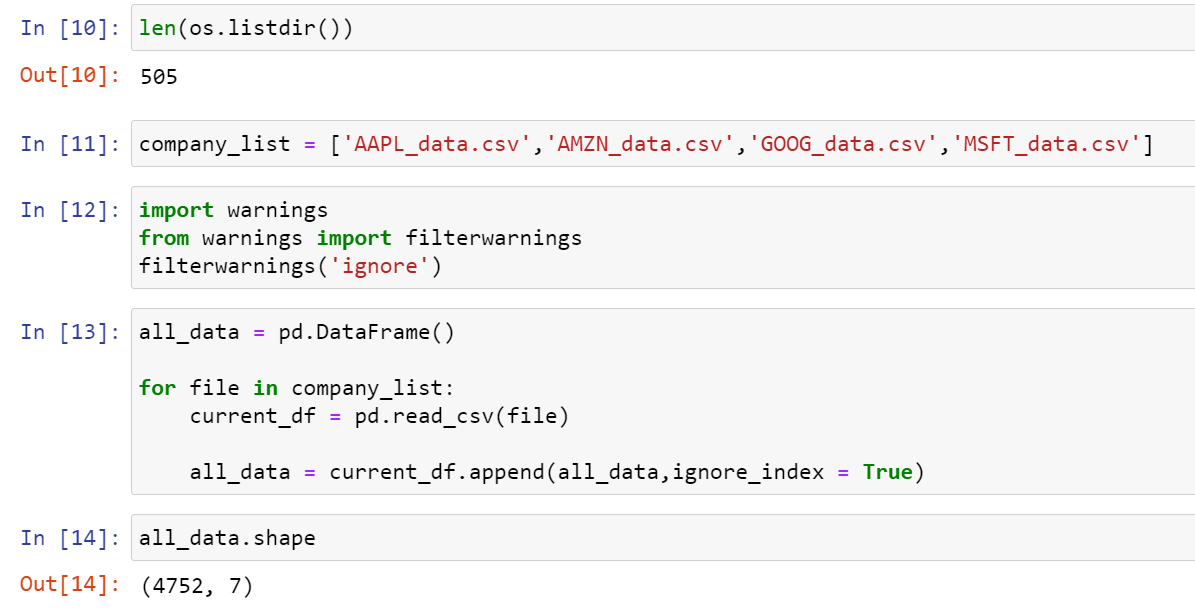
**Numpy :** NumPy is a Python library used for working with arrays. It also has functions for working in domain of linear algebra, fourier transform, and matrice. NumPy was created in 2005 by Travis Oliphant. It is an open source project and you can use it freely. NumPy stands for Numerical Python.

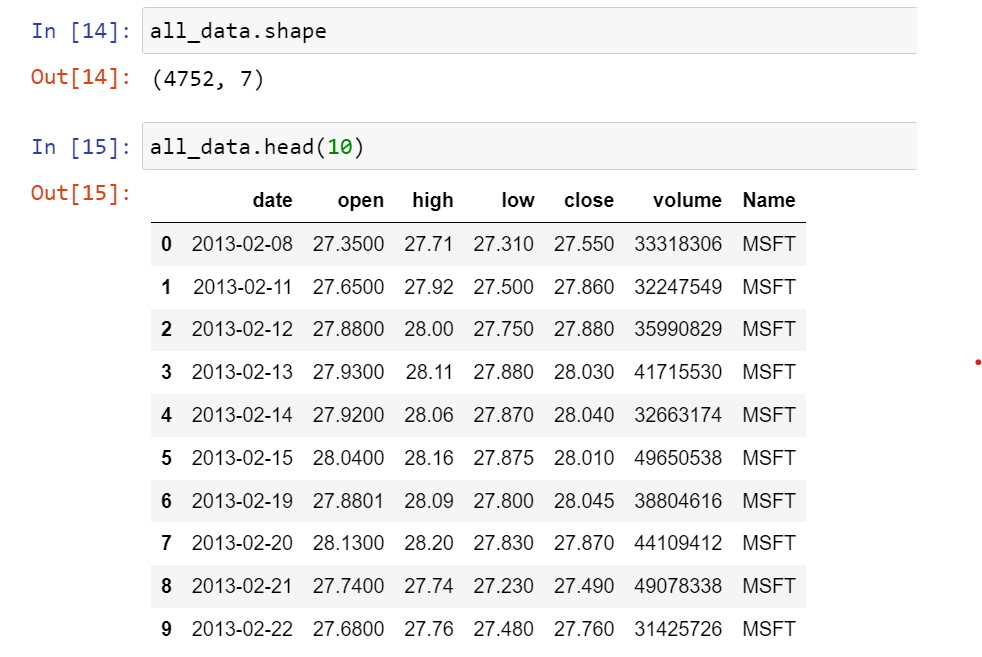
**Matplotlib :** Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. Matplotlib makes easy things easy and hard things possible.

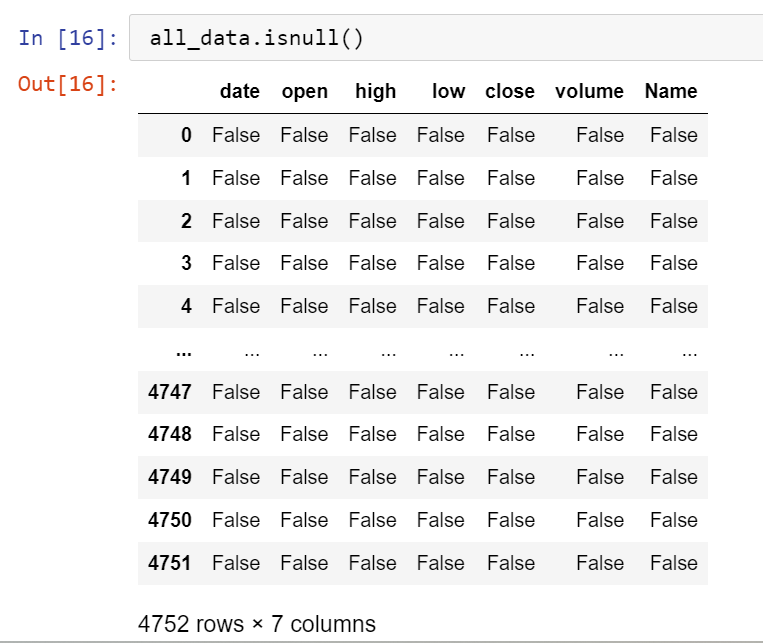
* Create [publication quality plots](https://ieeexplore.ieee.org/document/4160265/citations?tabFilter=papers).
* Make [interactive figures](https://mybinder.org/v2/gh/matplotlib/mpl-brochure-binder/main?labpath=MatplotlibExample.ipynb) that can zoom, pan, update.
* Customize [visual style](https://matplotlib.org/stable/gallery/style_sheets/style_sheets_reference.html) and [layout](https://matplotlib.org/stable/tutorials/provisional/mosaic.html).
* Export to [many file formats](https://matplotlib.org/stable/api/figure_api.html#matplotlib.figure.Figure.savefig).
* Embed in [JupyterLab and Graphical User Interfaces](https://matplotlib.org/stable/gallery/#embedding-matplotlib-in-graphical-user-interfaces).
* Use a rich array of [third-party packages](https://matplotlib.org/mpl-third-party/) built on Matplotlib.

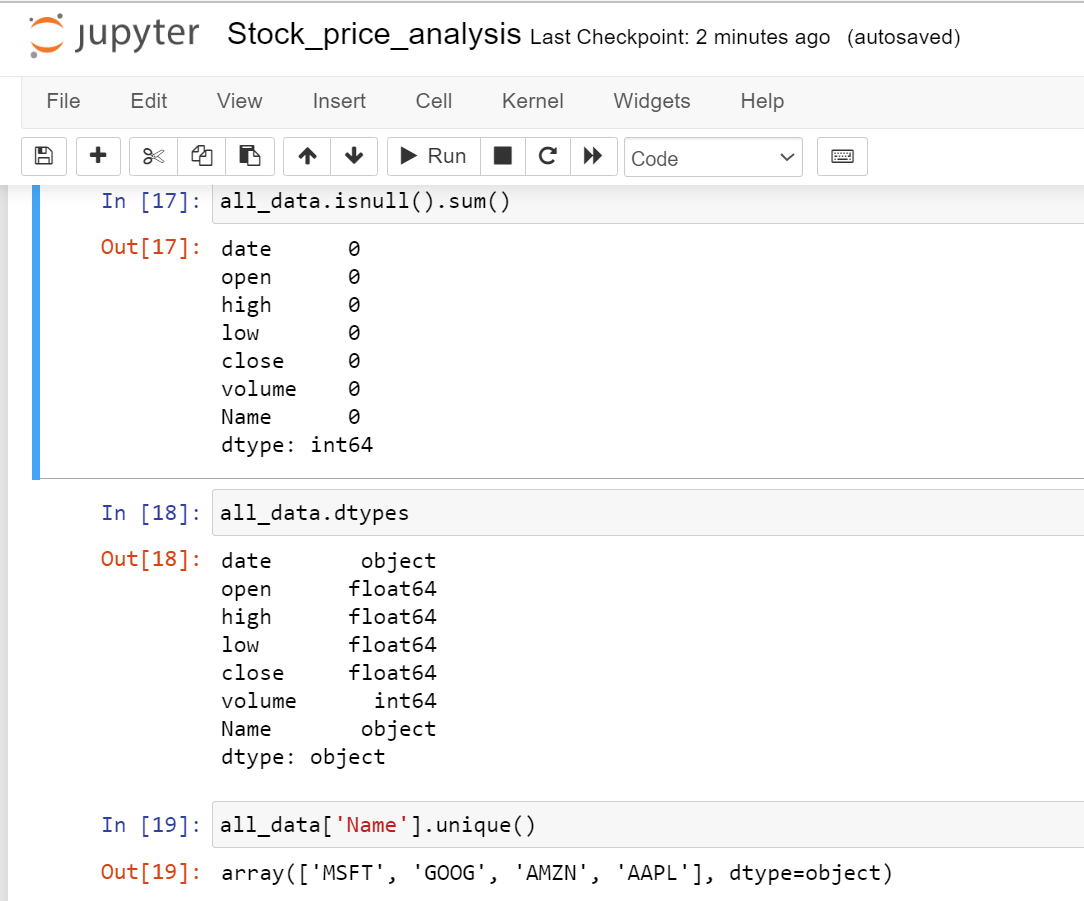
**Seaborn :** Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

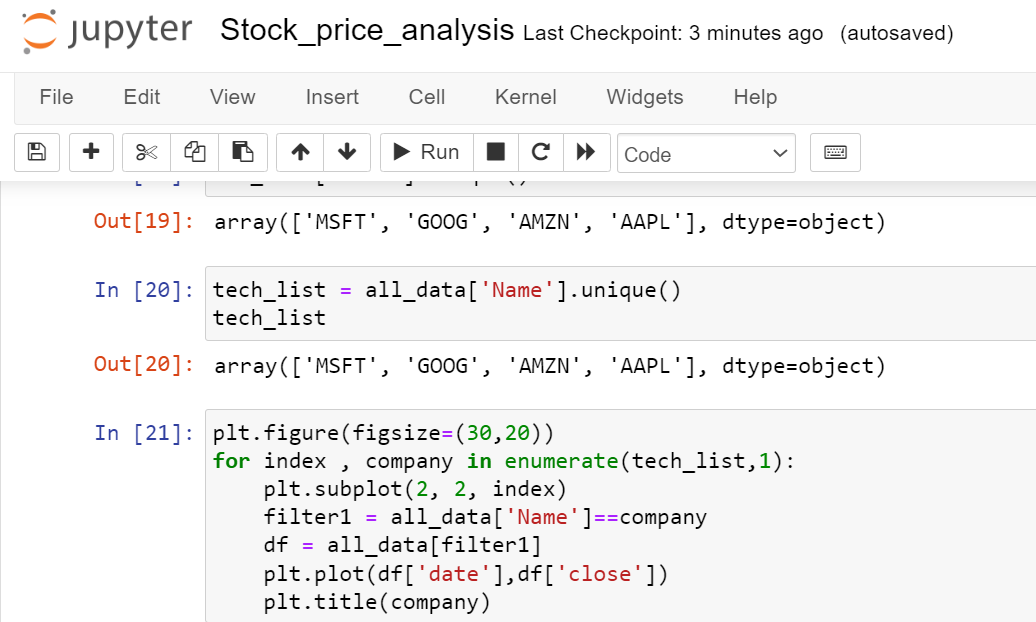
****

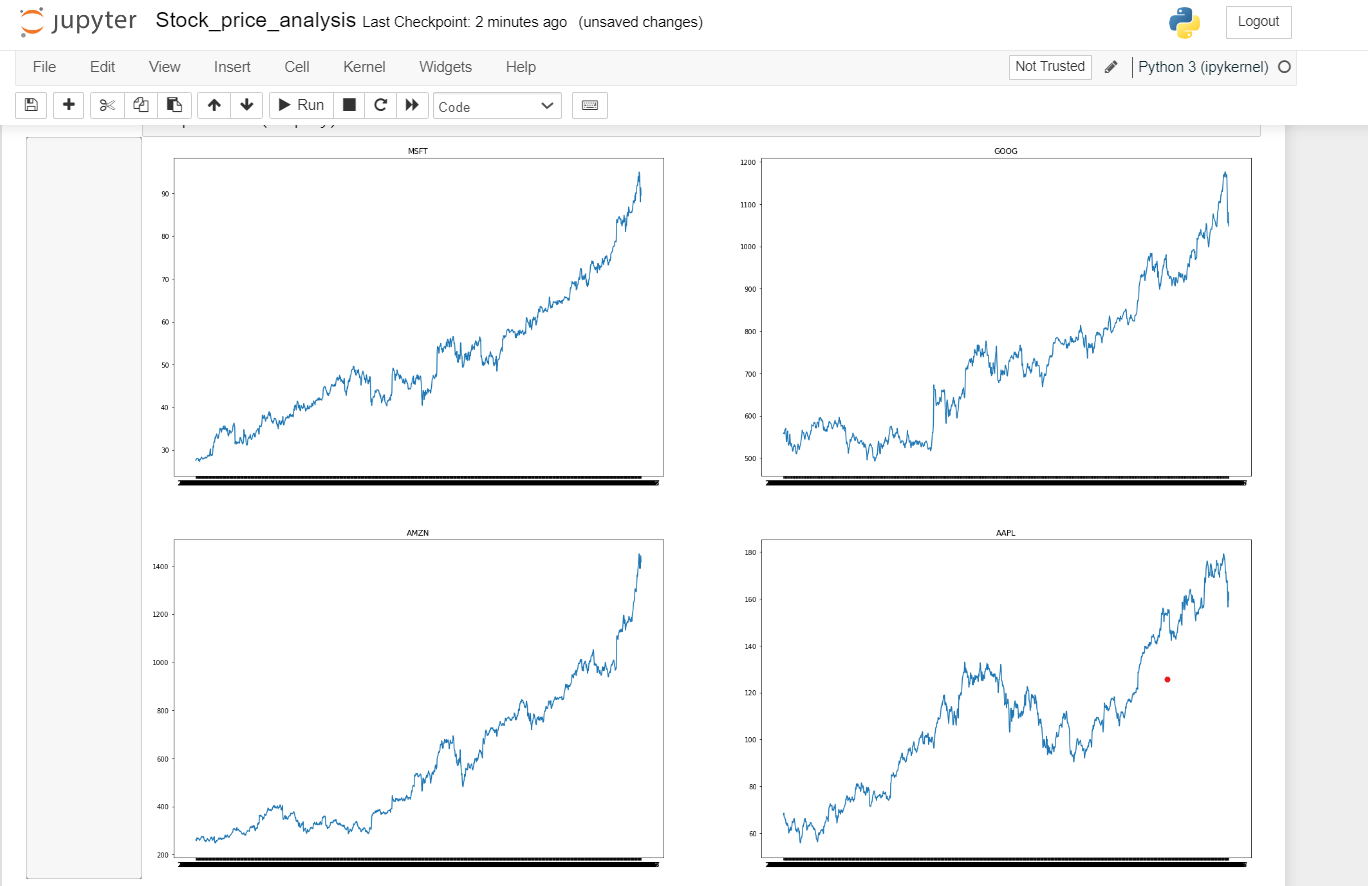


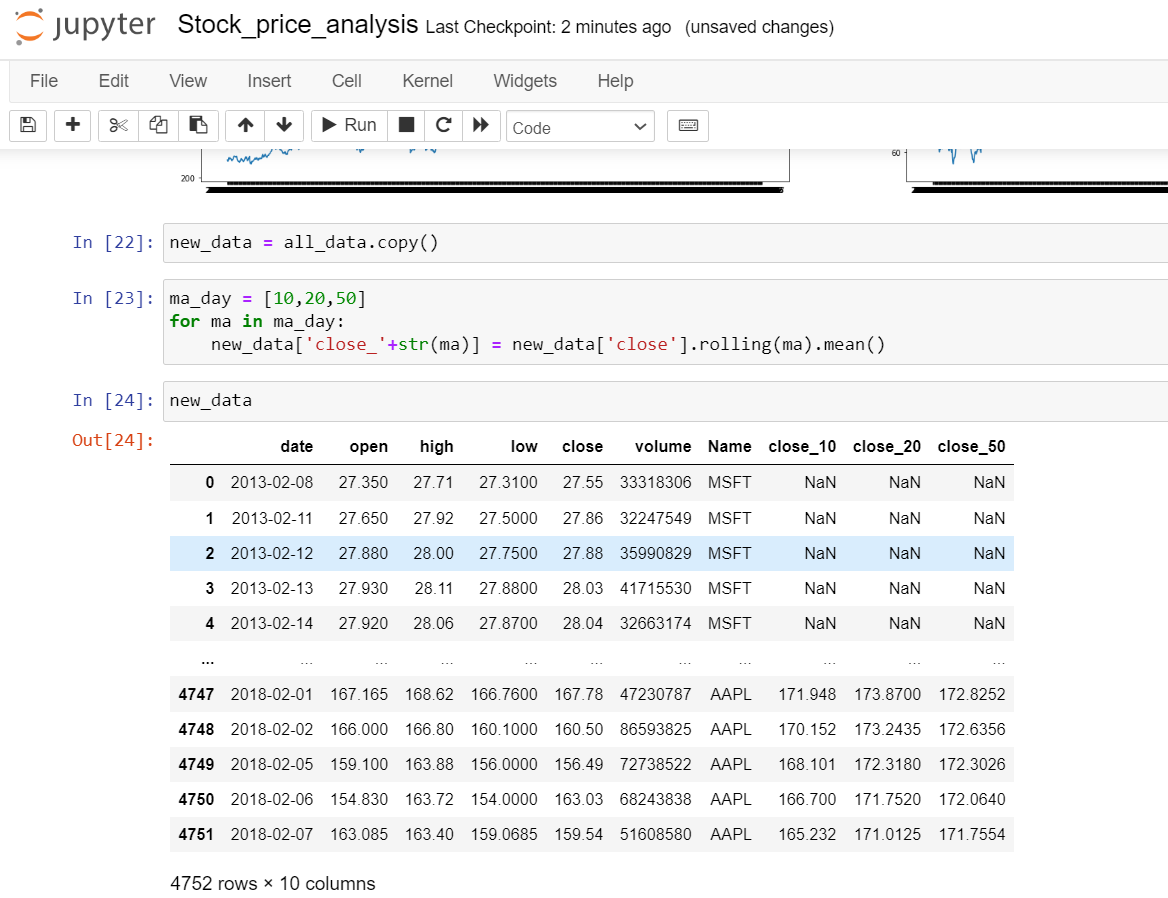


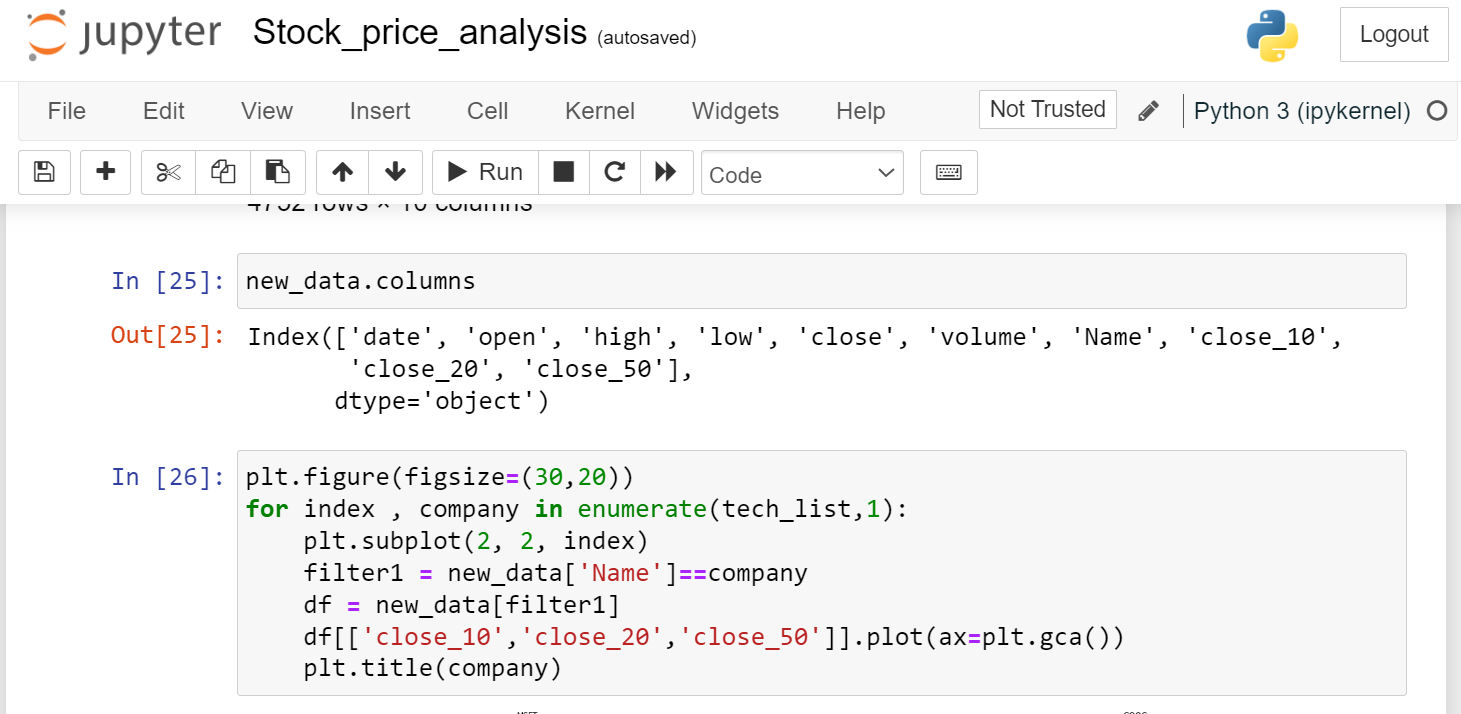


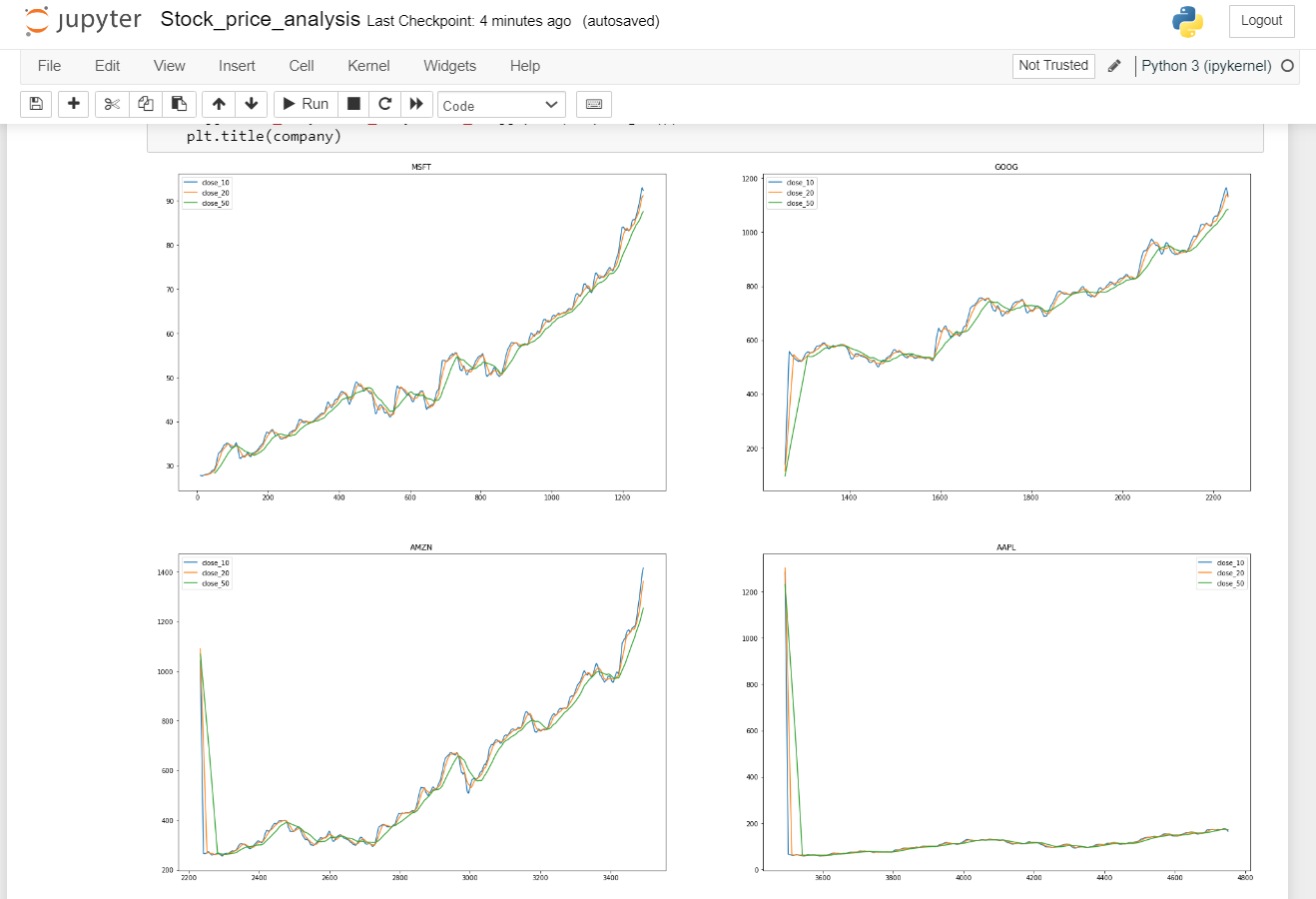


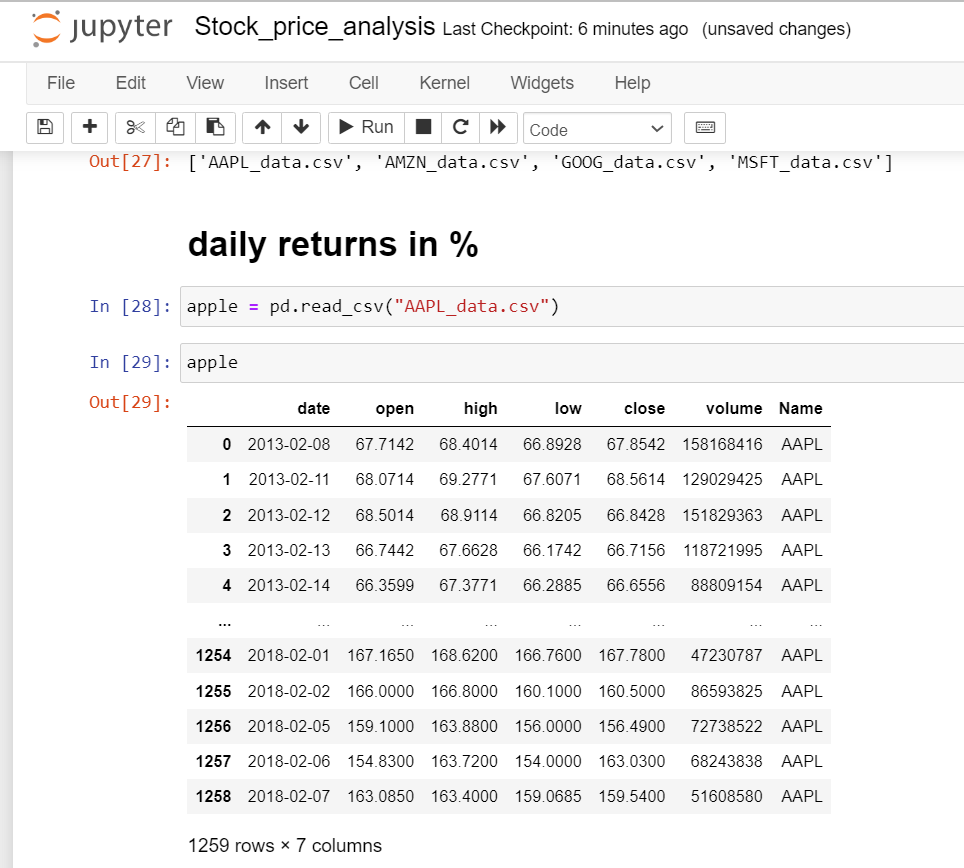


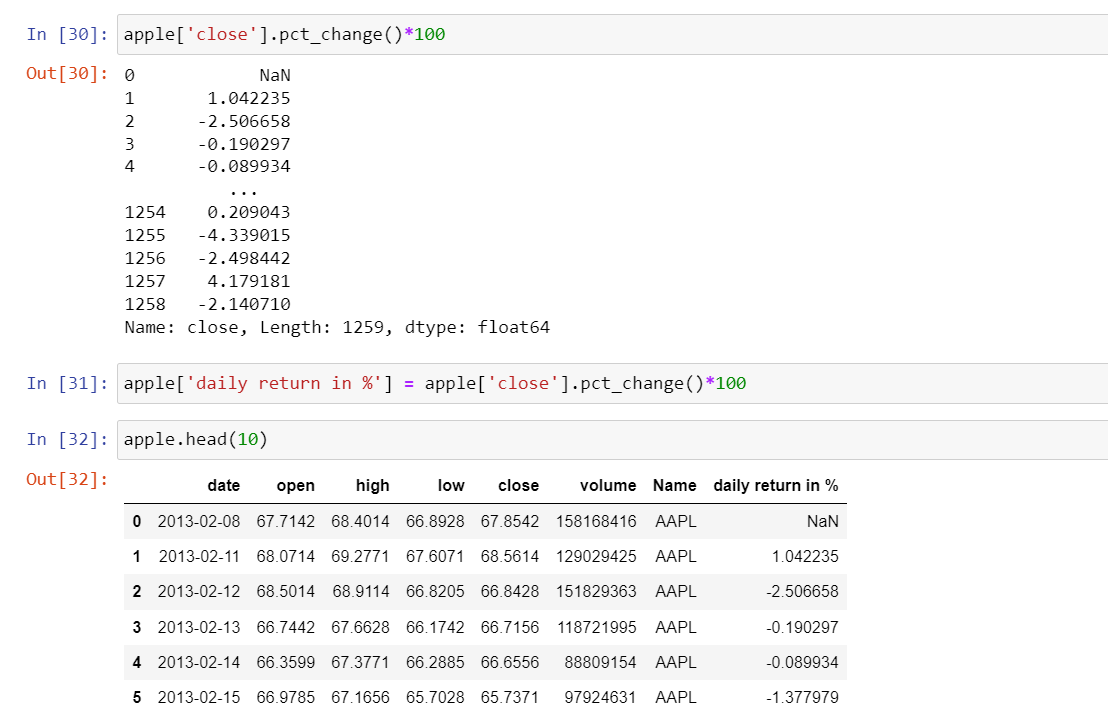






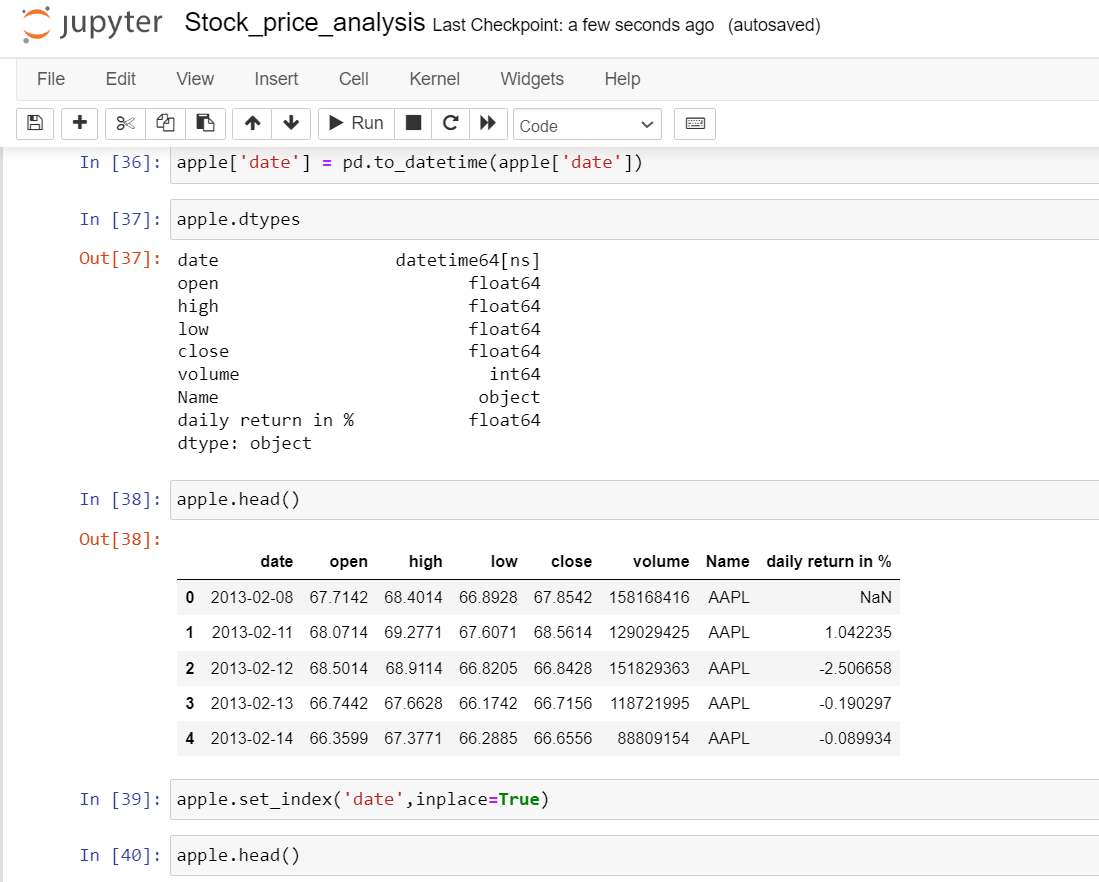


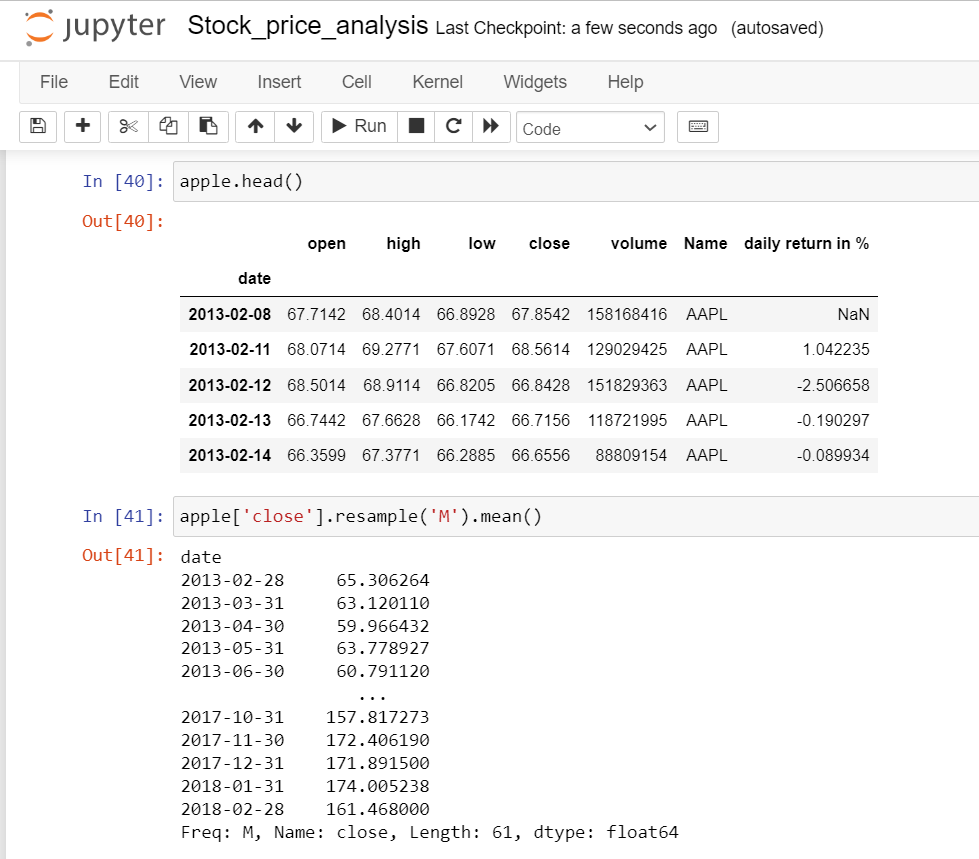


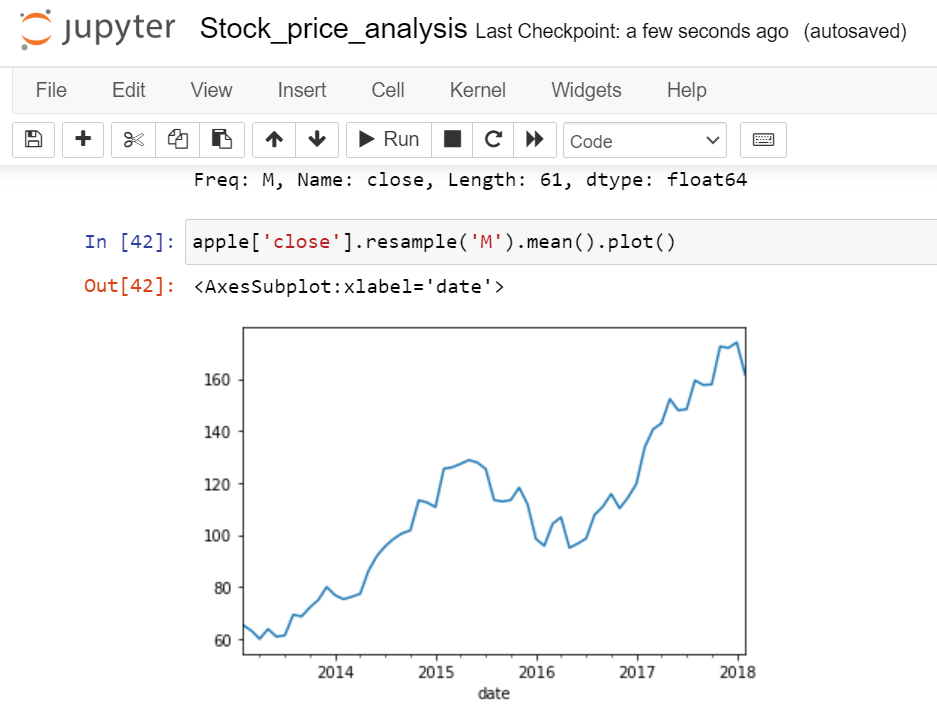


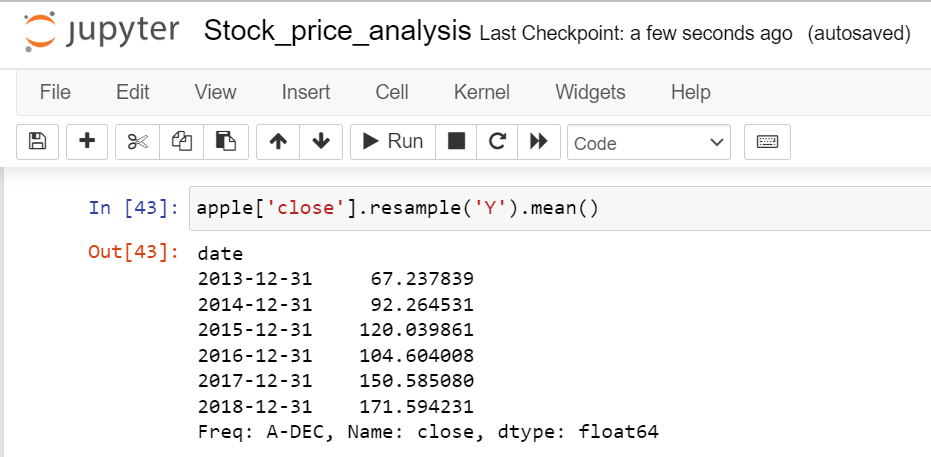


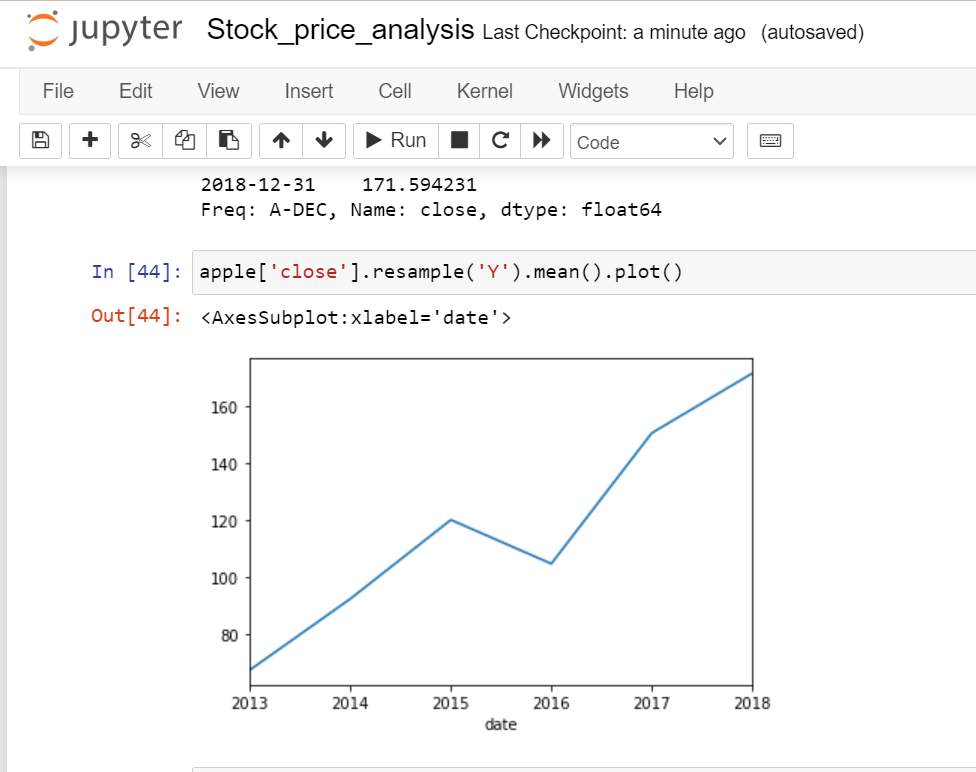


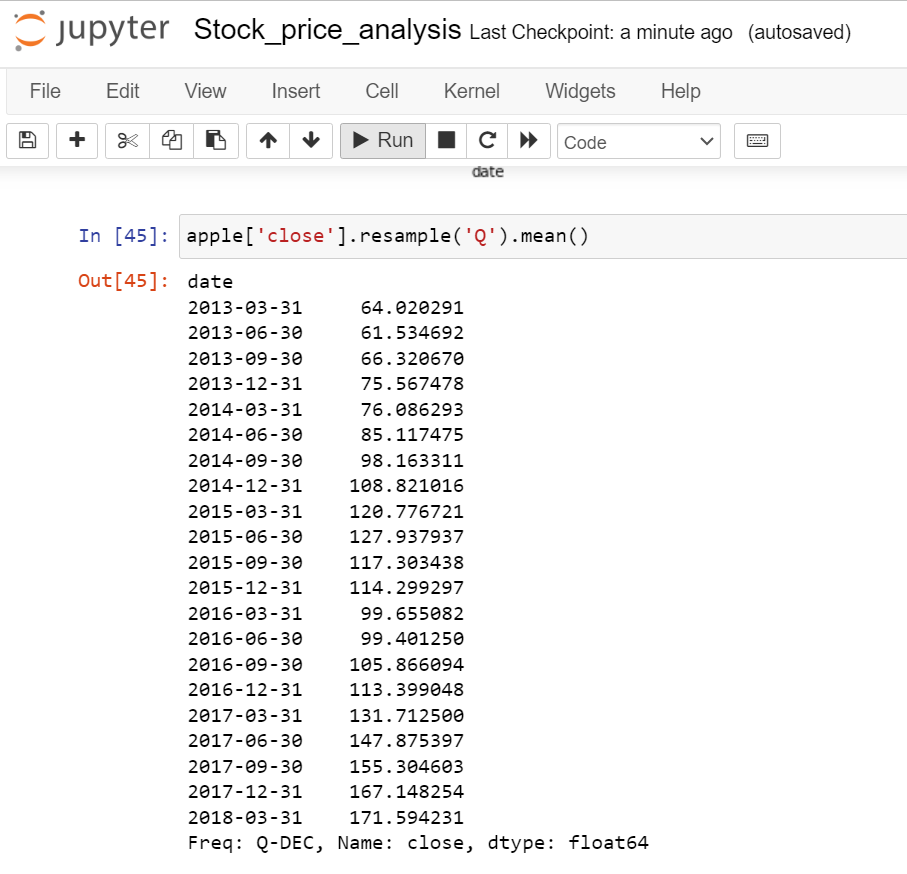


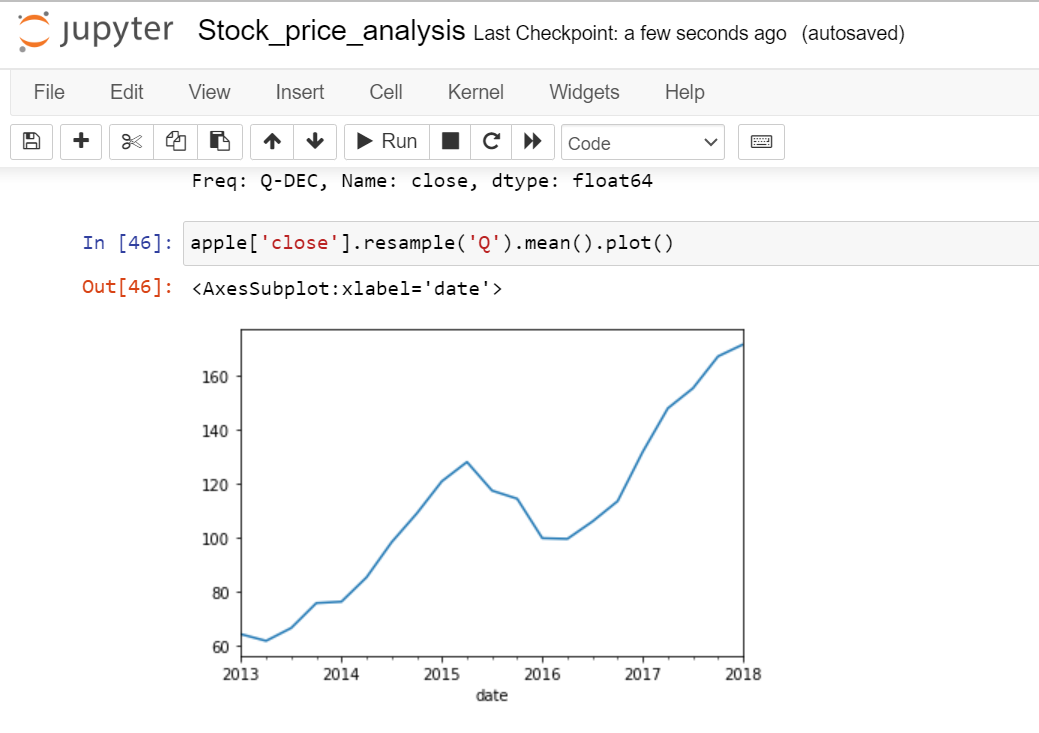


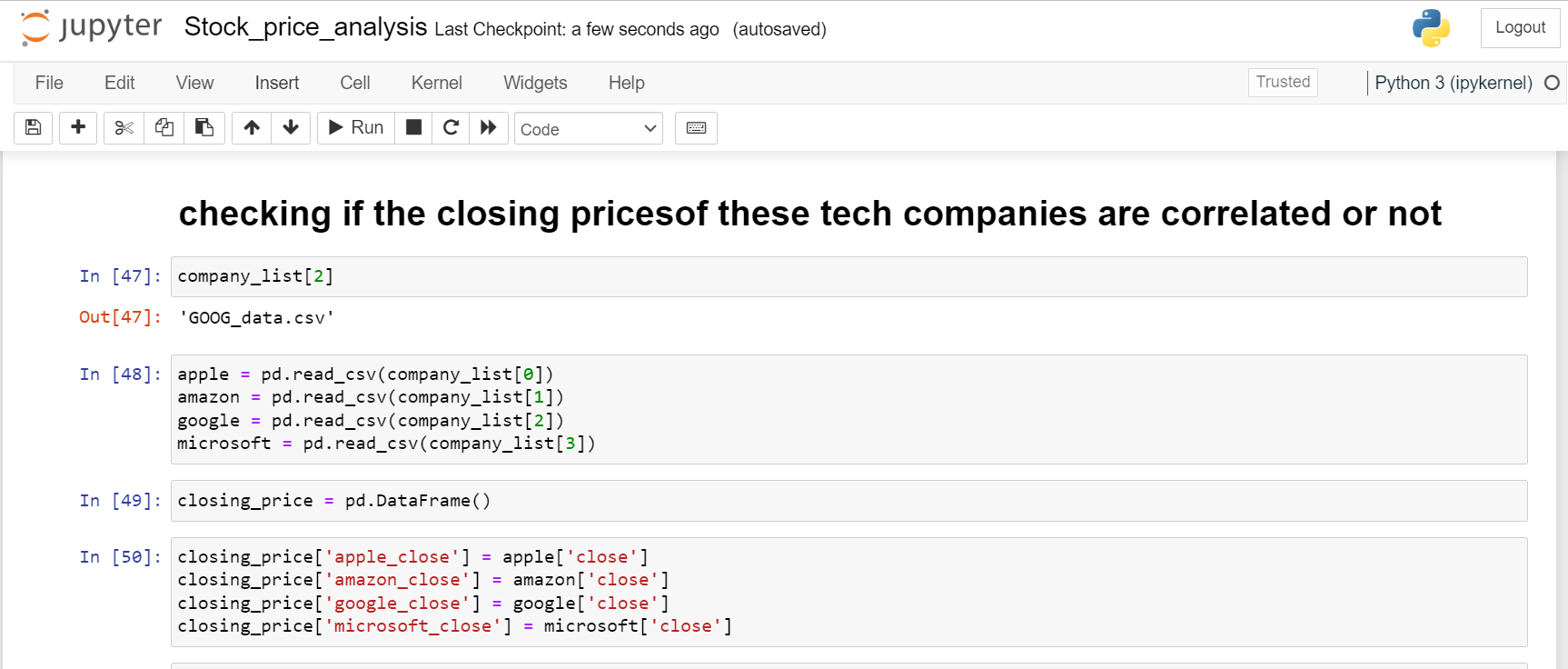


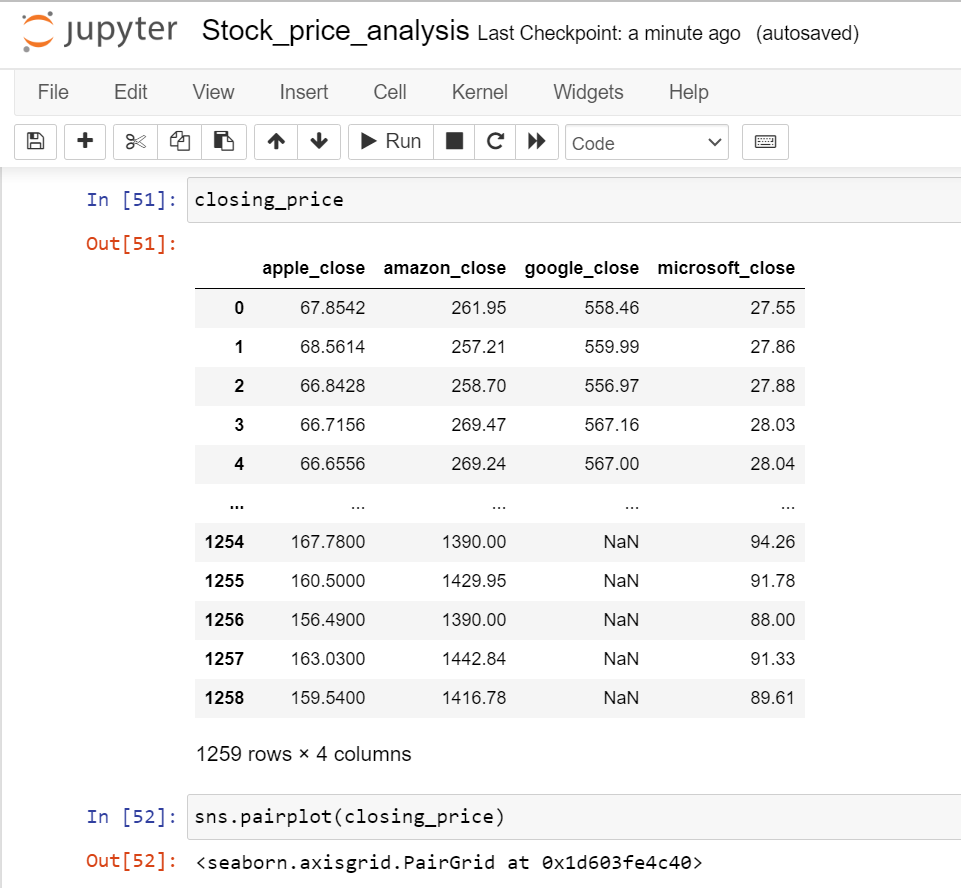


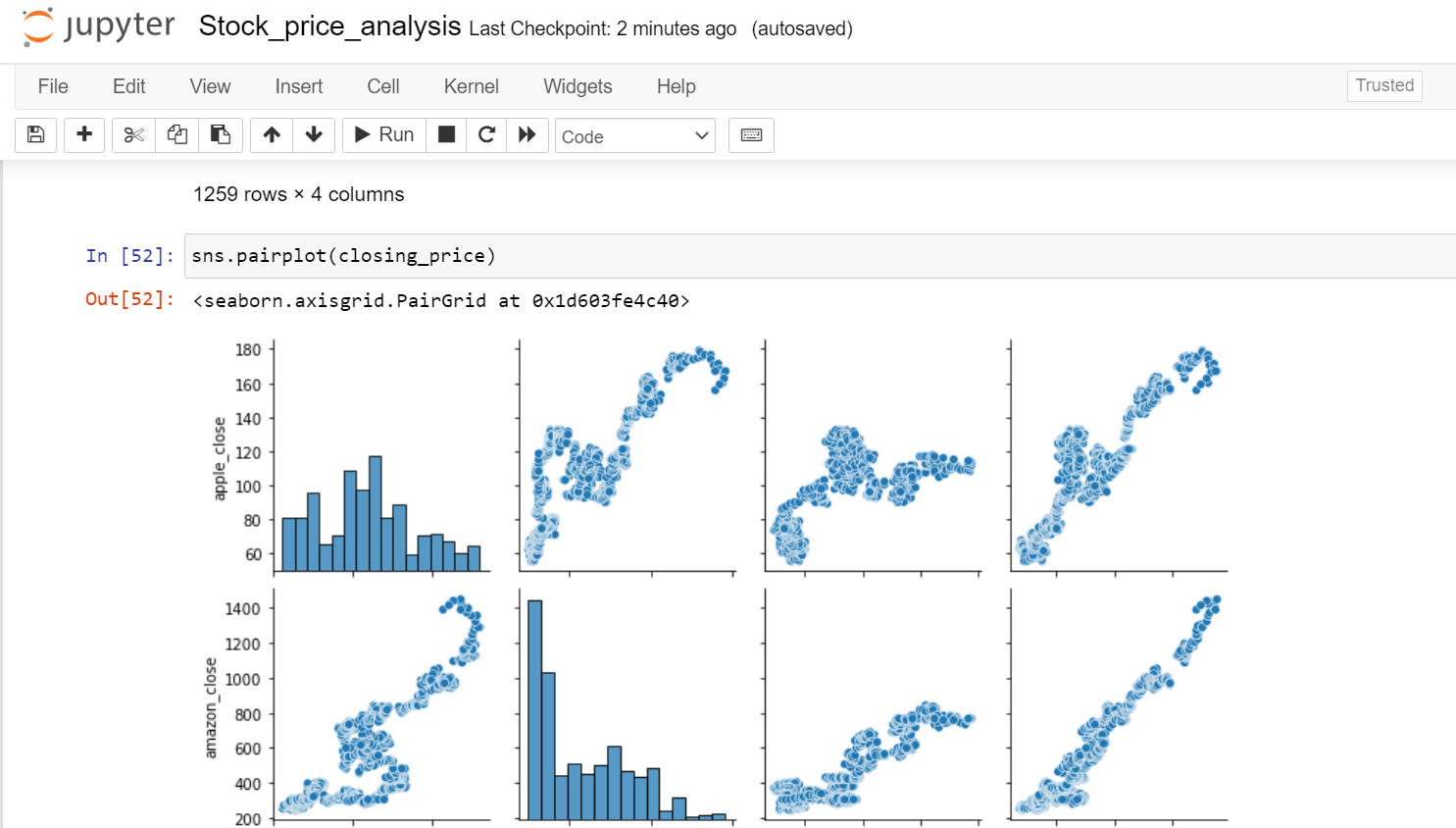


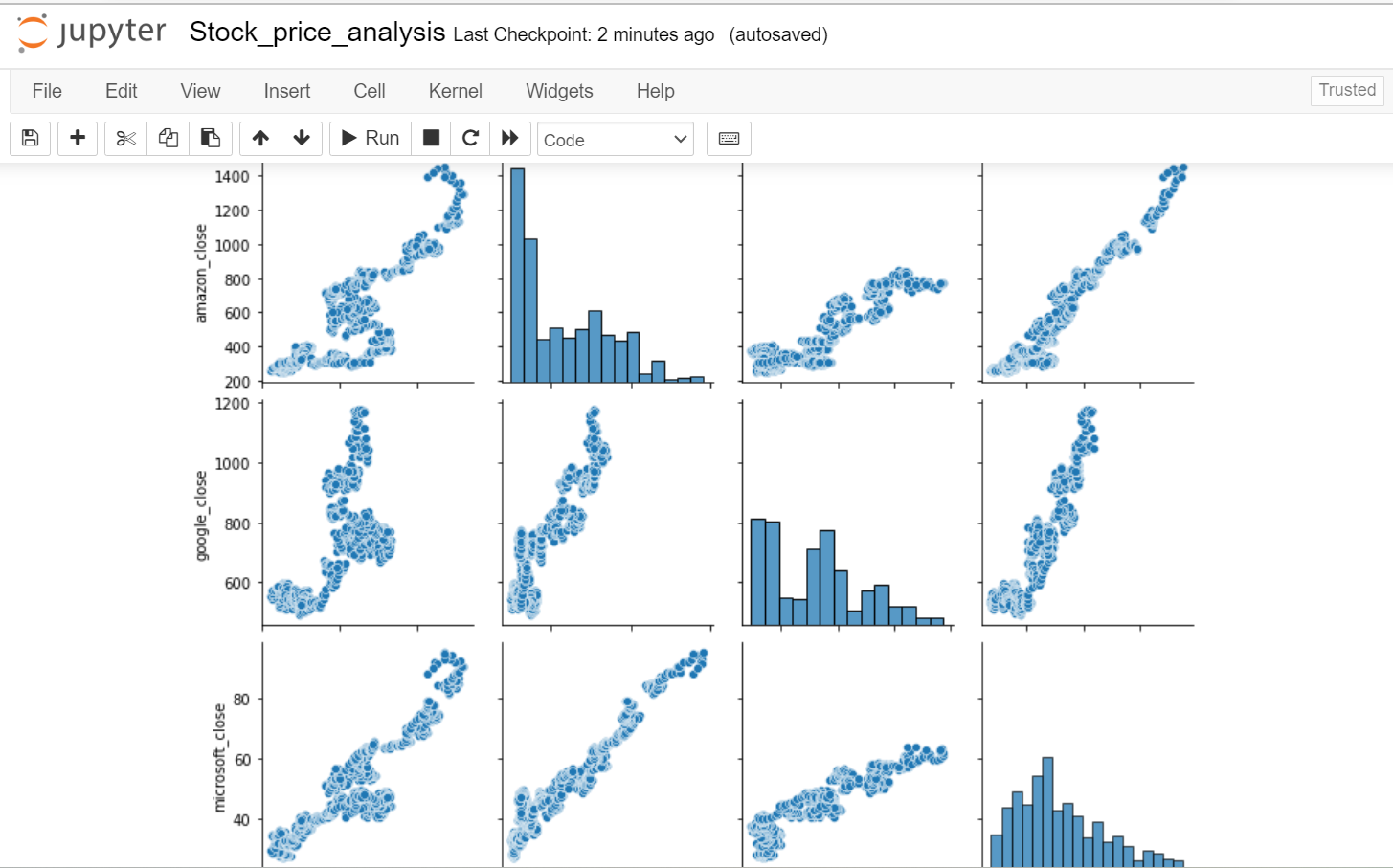






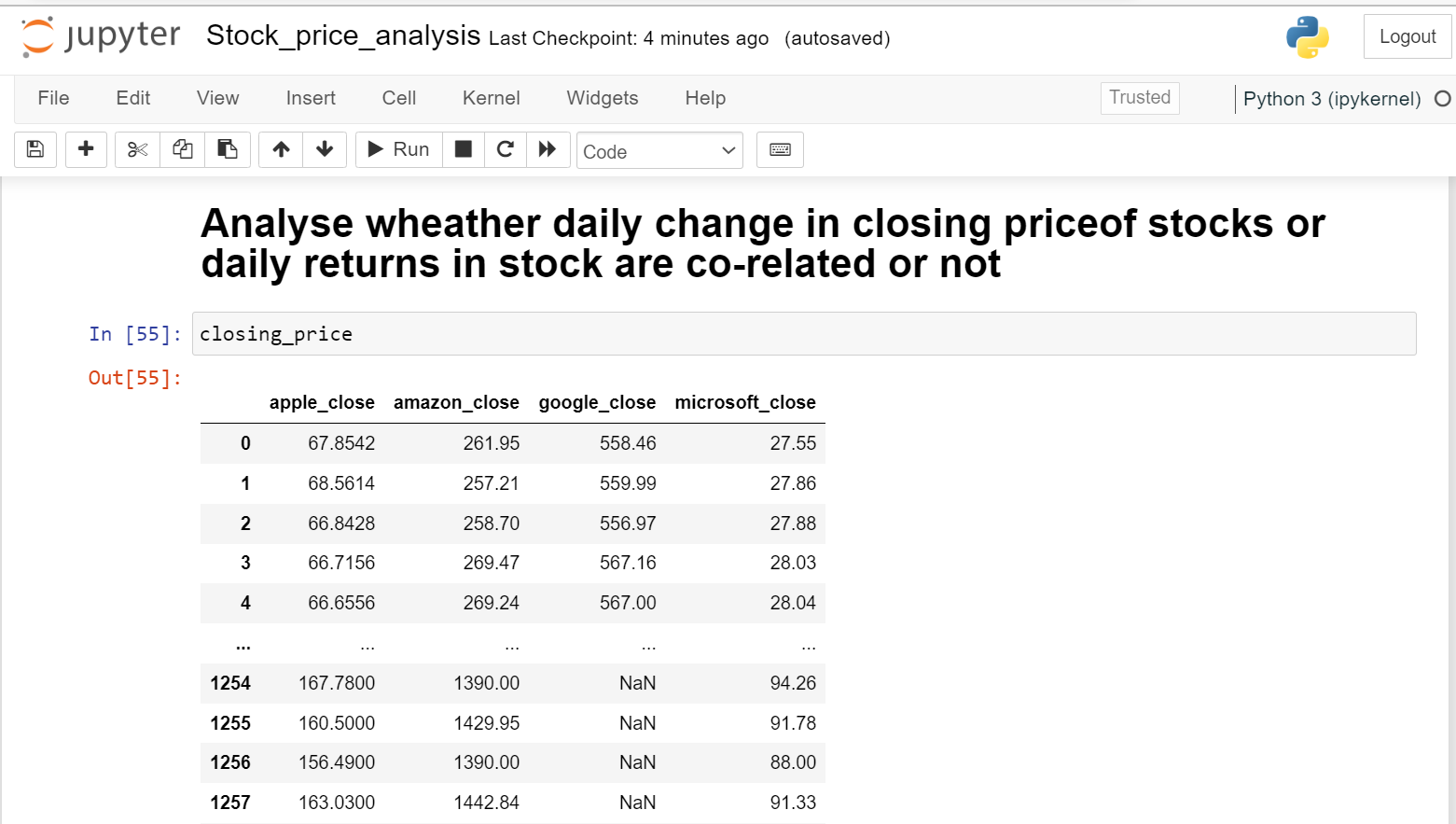


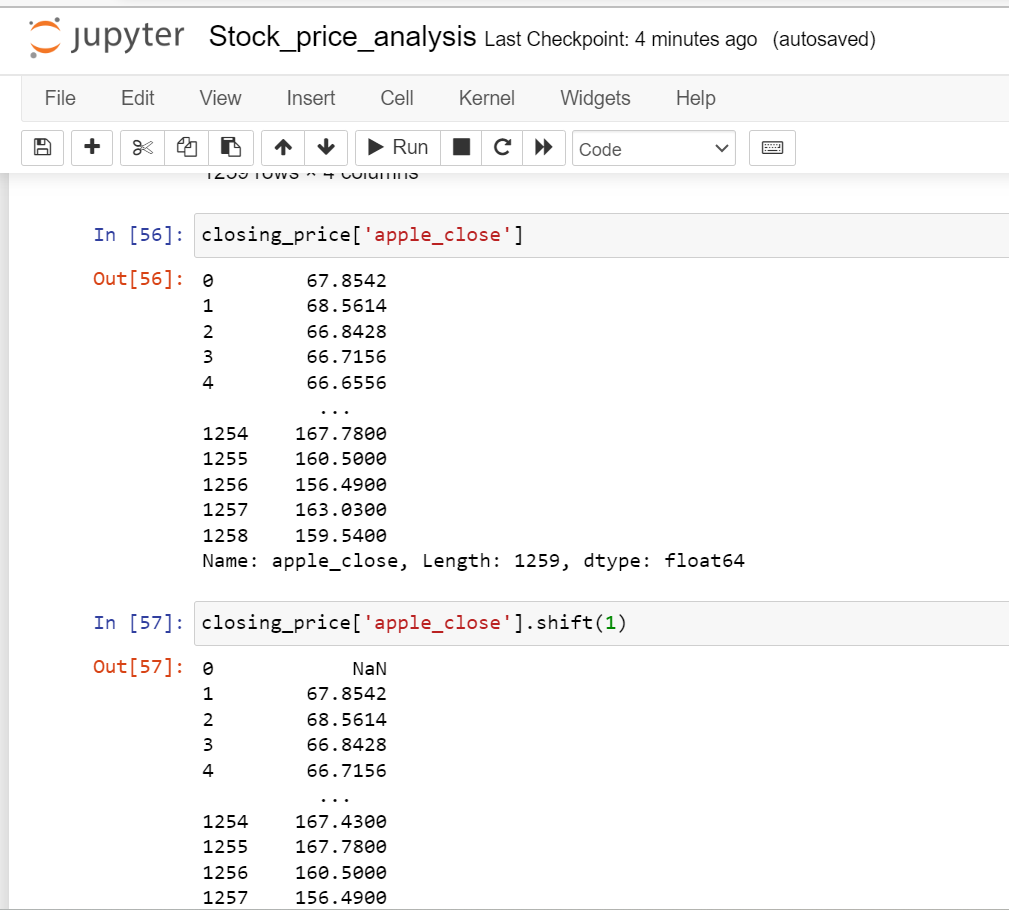


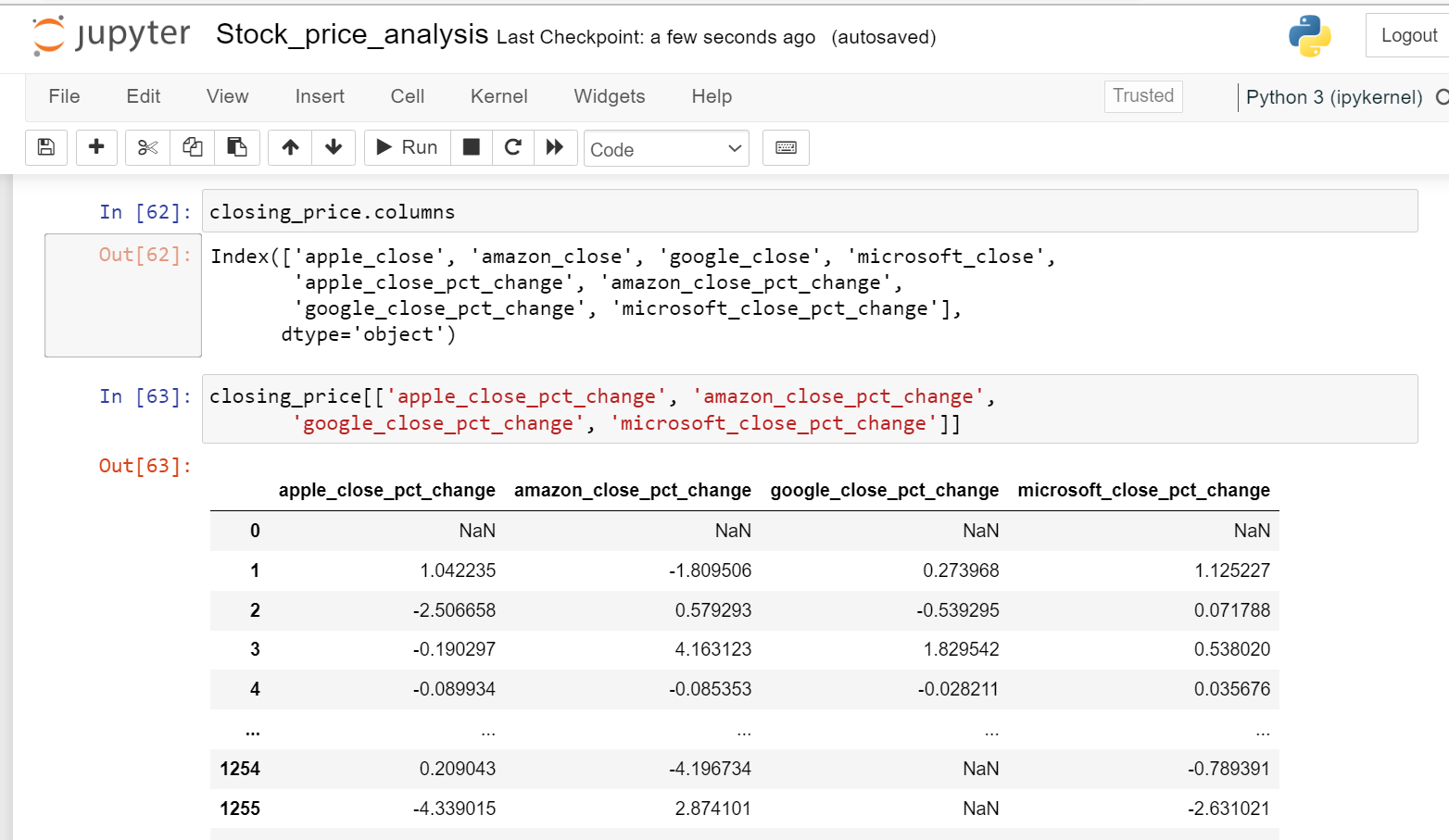


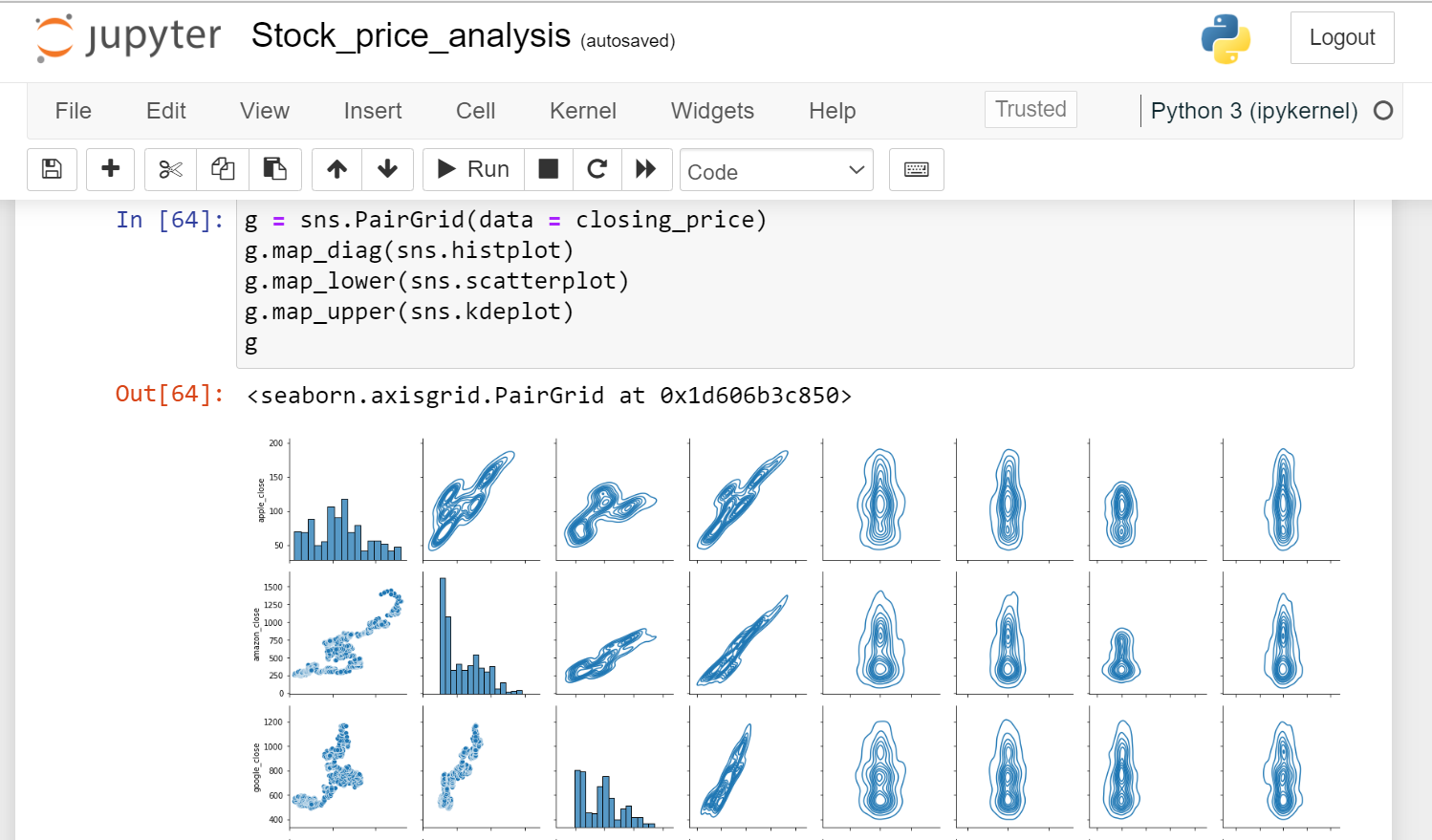


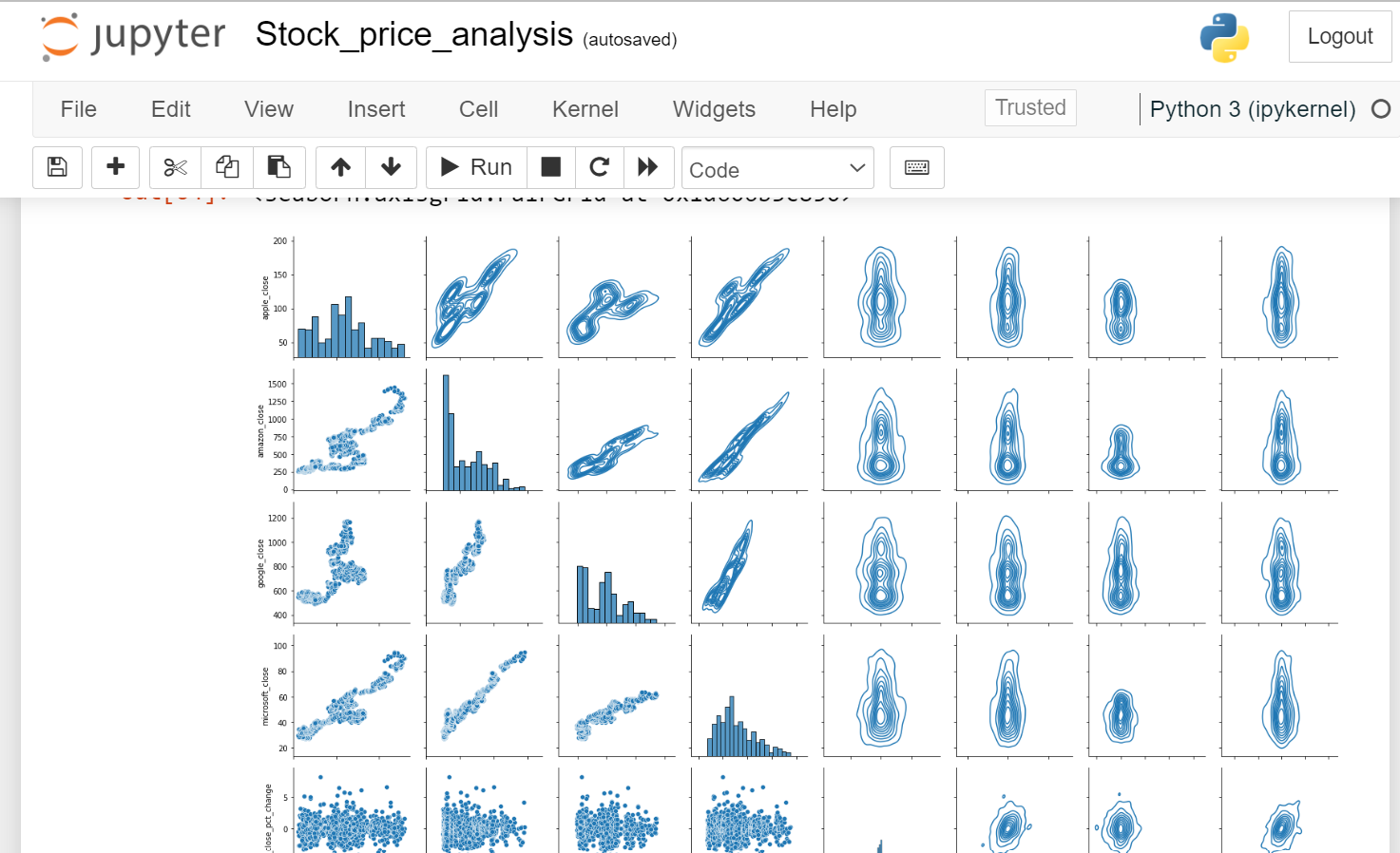


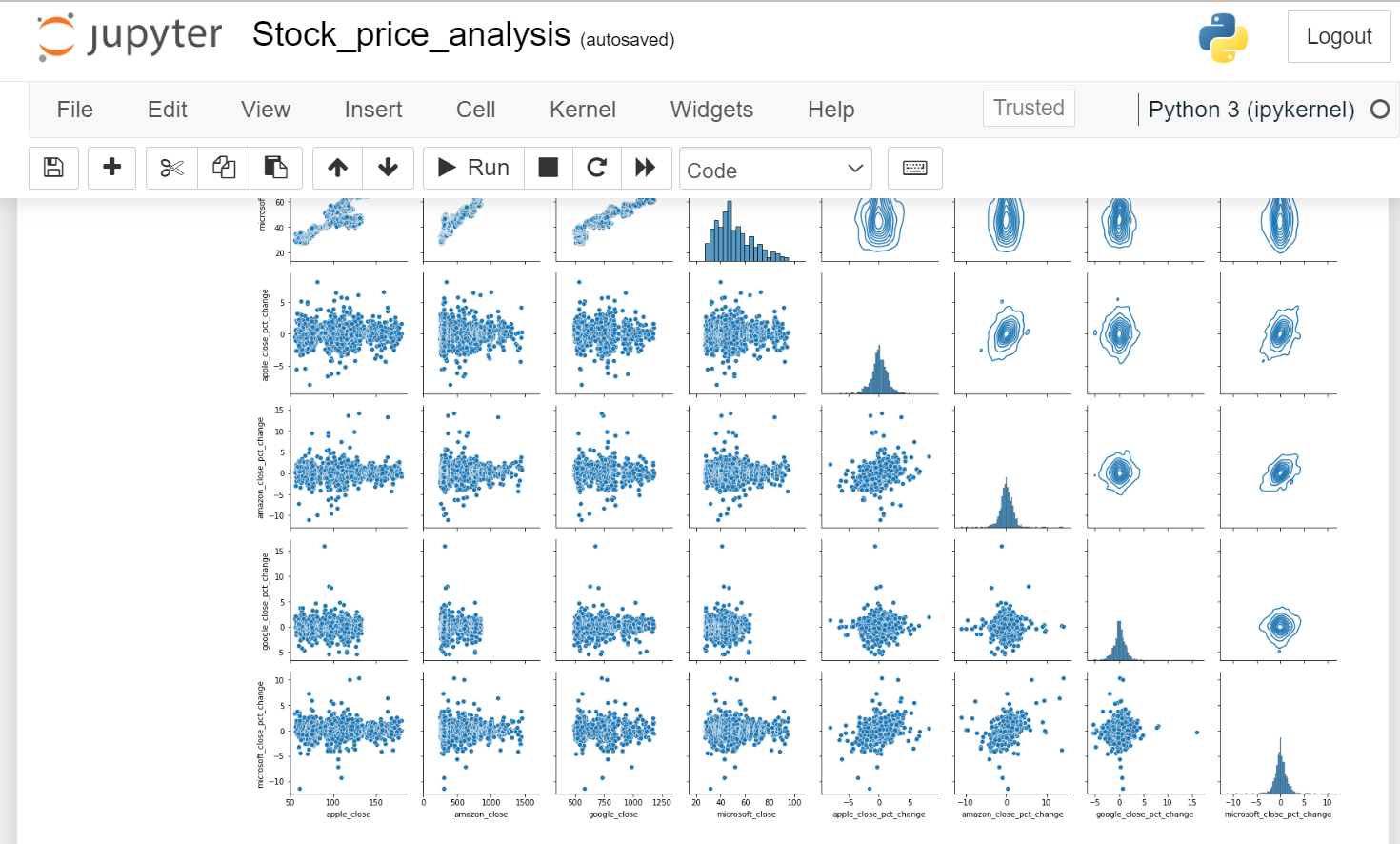


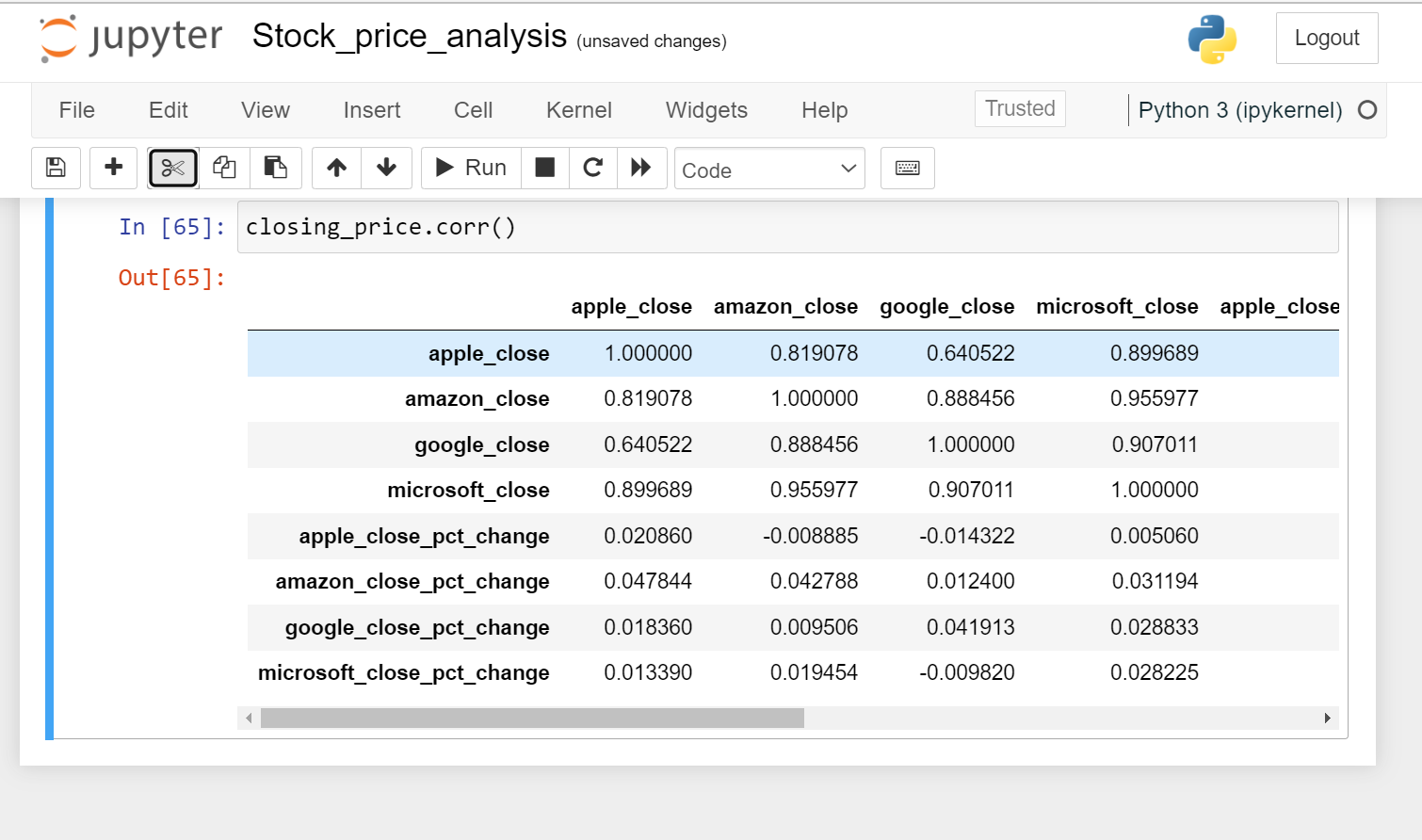












**VII. Conclusion**

The stock market is like a big auction where anyone interested can buy and sell stocks, or shares, which are small ownership slices of publicly traded companies. Stocks are bought and sold on a stock exchange. Individual investors can’t go straight to the stock exchange, so they work with a broker to buy and sell. A stock index tracks the performance of a group of stocks that are representative of the market overall. Investors can watch stock indexes to track past and current performance. Companies sell stocks and investors buy stocks with the goal of making money. A business might want to raise money to create a new product, and an individual might be using stocks to make money for retirement. Many people invest in mutual funds instead of buying individual stocks. Mutual funds let you pool your money with other investors to buy more and varied stocks across different companies. Not all stocks are created equal. Different types of stocks may have different growth rates, expected returns, dividend payments, and risk levels. Stock prices go up and down based on what investors are willing to pay. Things like a company’s profits, company news, and how investors think a company will do in the future can all impact a stock’s price. The economy can affect stock prices, how many shares are bought and sold, and the types of investments people make. On the flip side, the stock market’s performance can impact whether people think the economy is doing well or poorly. In a bull market, stock prices rise for an extended period of time. When stock prices drop and keep falling, it’s called a bear market. Price returns, total returns, and relative stock performance are all useful tools for keeping track of how well an individual stock does over time and deciding when to buy or sell shares.

**VII. Limitations and Future Enhancements**

Limitations is when huge amount of data is loading at that time compilation timing increasing rapidly.

Huge amount of data is generating day by day.

Data saving issue.

Need more drives or space for saving the generated data.

Time limit while trading.

Complication things while predicting.

In Future Enhancement we can add extra trading time in a india.

We can add extra cloud or drives for saving the data because huge amount of data is generated while trading (Mb,Gb,Tb,etc..)

Literacy about trading.

**VIII. References**

* <https://www.quickenloans.com/blog/stock-market-101-stock-market-work>
* <https://www.moneyunder30.com/how-to-read-a-stock-chart>
* <https://www.investor.gov/introduction-investing/investing-basics/investment-products/stocks>
* <https://www.avatrade.com/cfd-trading/stocks/stock-market>
* <https://www.udemy.com/course/data-analytics-projects-python/>
* <https://www.nism.ac.in/research-paper-articles/>
* <https://www.financialexpress.com/opinion/disclosures-and-disconnects/3110776/>
* <https://www.google.com/search?q=us+stock+market+today&rlz=1C1CHBD_enIN957IN957&oq=US+stock+market+&gs_lcrp=EgZjaHJvbWUqBwgCEAAYgAQyBggAEEUYOTINCAEQABiDARixAxiABDIHCAIQABiABDIHCAMQABiABDIHCAQQABiABDIHCAUQABiABDIHCAYQABiABDIHCAcQABiABDINCAgQABiDARixAxiABDIHCAkQABiABNIBDDc5MTk0MjJqMWoxNagCALACAA&sourceid=chrome&ie=UTF-8>
* <https://www.moneycontrol.com/us-markets/>
* Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose1, GiridharMaji, Narayan C. Debnath, Soumya Sen
* S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman - Stock price predictionusing LSTM, RNN and CNN-sliding window model - 2017.
* Murtaza Roondiwala, Harshal Patel, Shraddha Varma, “Predicting Stock Prices Using LSTM” in Undergraduate Engineering Students, Department of Information Technology, Mumbai University, 2015.
* Xiongwen Pang, Yanqiang Zhou, Pan Wang, Weiwei Lin, “An innovative neural network approach for stock market prediction”, 2018
* Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan Department of Computer Science and Engineering National Institute of Technology, Hamirpur – 177005, INDIA - Stock Market Prediction Using Machine Learning.