



Assessment Submission Form

Student Name	Satish Mohan; Navin Mutha; Gautam Gupta (GROUP 10)
Student Number	18200575; 18202954; 18201797
Assessment Title	Group Assignment 1: STOCK MARKET ANALYSIS by the Team Data Whisperers
Module Code	MIS41110
Module Title	Programming for Analytics
Module Coordinator	Miguel Nicolau
Date Submitted	November 23 rd 2018
Office use only	
Date Received	

A SIGNED COPY OF THIS FORM MUST ACCOMPANY ALL SUBMISSIONS FOR ASSESSMENT.

STUDENTS SHOULD KEEP A COPY OF ALL WORK SUBMITTED.

Procedures for Submission and Late Submission

Ensure that you have checked the University's procedures for the submission of assessments.

Note: There are penalties for the late submission of assessments. For further information please see the *Assessment Guidelines* publication or the website at <http://www.ucd.ie/governance/resources/policypage-latesubmissionofcoursework/>

Plagiarism is the unacknowledged inclusion of another person's writings or ideas or works, in any formally presented work (including essays, examinations, projects, laboratory reports or presentations). The penalties associated with plagiarism are designed to impose sanctions that reflect the seriousness of the University's commitment to academic integrity. Ensure that you have read the University's *Briefing for Students on Academic Integrity and Plagiarism* and the UCD *Plagiarism Statement, Plagiarism Policy and Procedures*, (<http://www.ucd.ie/registrar/>)

Declaration of Authorship

I declare that all material in this assessment is my own work except where there is clear acknowledgement and appropriate reference to the work of others.

Signed: NAVIN MUTHA SATISH MOHAN GAUTAM GUPTA

Date: 23rd November 2018

Table of Contents

STOCK MARKET ANALYSIS.....	3
FLOW CHART.....	3
UNIFIED MODELING LANGUAGE.....	5
USER-MANUAL.....	6
ERROR HANDLING and EXCEPTION HANDLING	9
BIBLIOGRAPHY	10

STOCK MARKET ANALYSIS

In today's world most of the individuals spend lots of money in pursuing higher studies in pursuit of better jobs and placement. Unfortunately, they are reluctant to spend even 10% of the amount they have earned in investing. It is very important for an individual to understand investing as it helps securing the future over a longer period and requires patience. The purpose is to make the investor make a sound decision and evaluate the organizations performance based on certain financial parameters (Julia writer, 2017).

When it comes to investing in stocks, analyst should be capable enough of conducting a thorough technical analysis of the stock before investing. This includes simple to advanced analysis of numbers to provide a descriptive and predictive model based on price movements in the past for a stock.

A Stock Market Analysis model is built using Python Programming language to help investors analyse any stock which they want and make further financial decisions of buying, holding or selling stocks (Julia writer, 2017).

FLOW CHART

Before understanding any model, it is imperative to understand the flow of the code and identify how exactly it works. It helps interpreting the flow of the model along with different processes and decisions as required and describe it in a chronological manner.

The below chart is the basic graphic about how the initialisation and describing the main functionalities of the model.

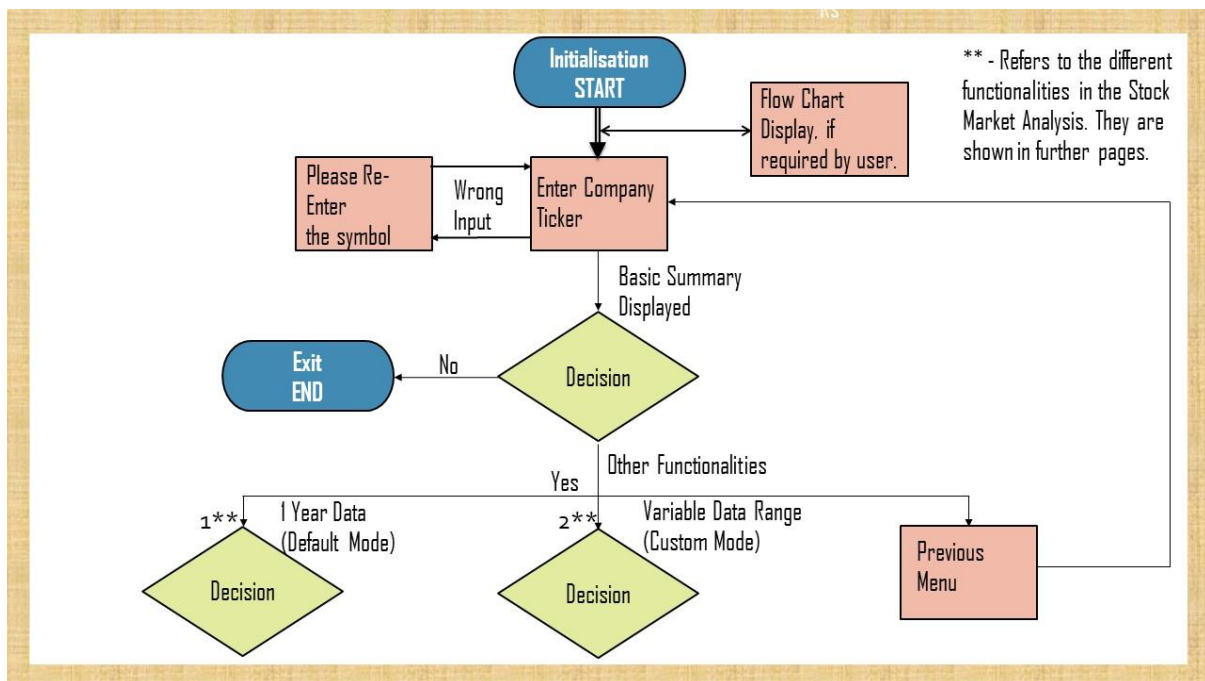


Figure 1: Flow Chart – Initialisation

The Below 2 figures now describes the Other Functionalities which have been incorporated in the model for analysis purposes.

1. Default Mode: 52 Week Range

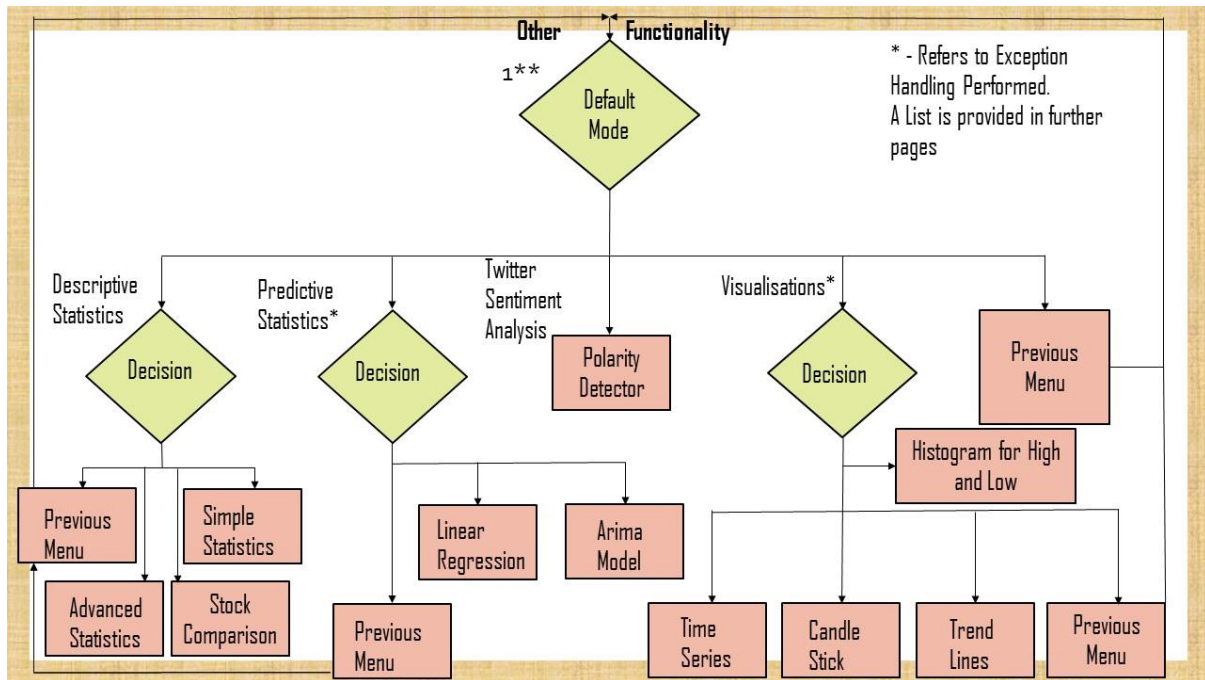


Figure 2: 52 Week Range

2. Custom Range: Variable Date Range

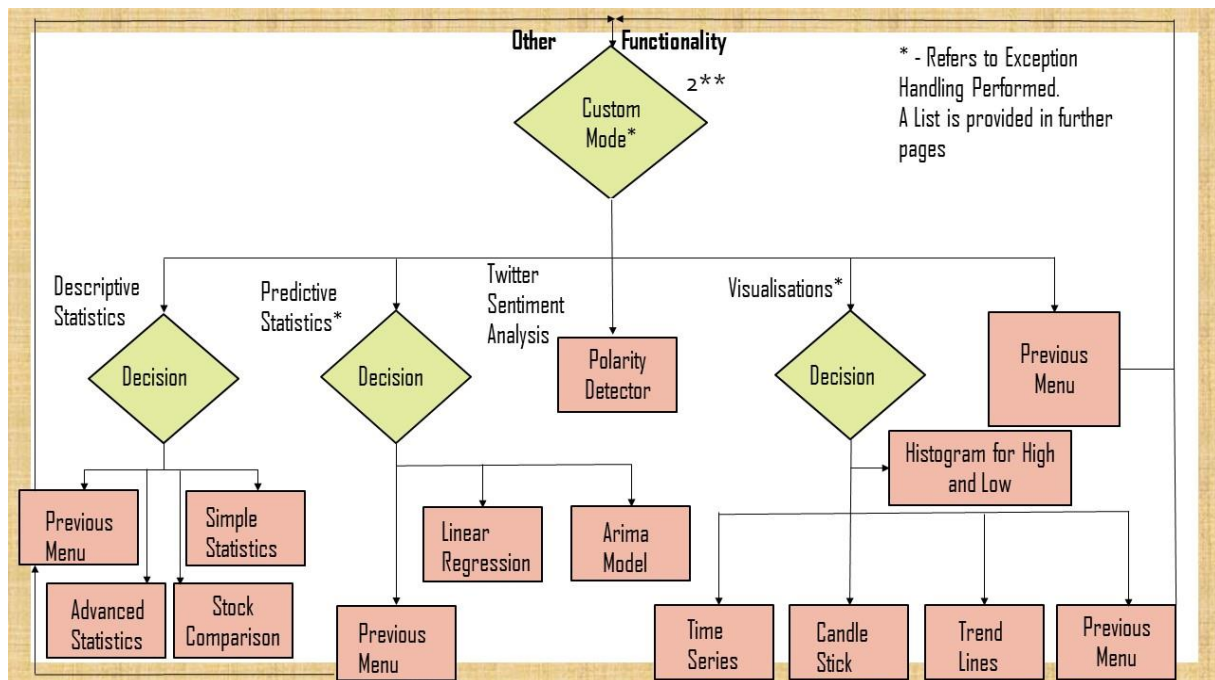


Figure 3: Variable Data Range

UNIFIED MODELING LANGUAGE

UML provides an overview of the flow control for interactions where similar information to sequence diagram, but it is focussed primarily on object relationships. The sequence of messages is shown through a numbering scheme. The following diagram corresponds to a communication diagram between the actor/user and the model of the code which isn't immediately visible but communicates set of messages between adjacent objects.

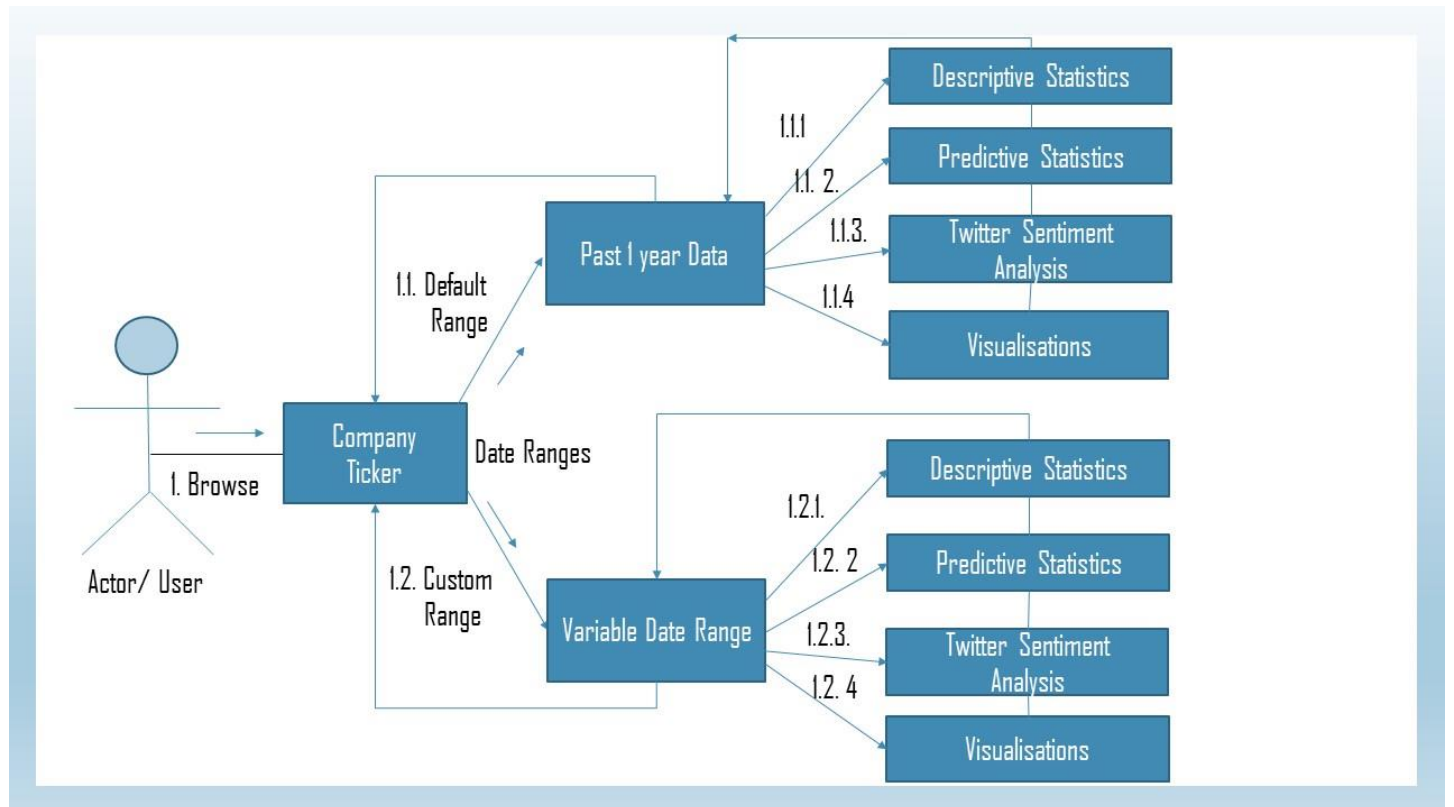


Figure 4: UML - Sequence Diagram

USER-MANUAL

A user-manual is a technical communication document intended to give assistance to people using the above mode. Below is a list of steps which will be helpful for the same.

1. There are 5 Python files; “home_page.py”, “descriptive_analytics.py”, “predictive_analytics”, “candlestick_timeseries”, “twitter”. Download the above-mentioned files for executing Stock Market analysis.
2. It is important to import certain Python Packages for successful execution of the code. Below is a list of those packages: -
 - a. pandas_datareader
 - b. pandas
 - c. numpy
 - d. datetime
 - e. calendar
 - f. pandas_market_calendars
 - g. descriptive_analytics
 - h. predictive_analytics
 - i. candlestick_timeseries
 - j. twitter
 - k. warnings
 - l. sys
 - m. PIL
 - n. image
 - o. scikit
 - p. matplotlib
 - q. statsmodels.tsa.stattools
 - r. statsmodels.tsa.arima_model
 - s. ARIMA
3. Execute the “home_page.py” file.
4. Enter the Company Ticker for which you want to do an analysis.
5. A general summary will be shown describing the organization’s Current price, Market Capitalisation, IPO year etc.
6. Another Menu option opens asking from 3 different options.
 - a. 52 Week Range: Showcasing Past 1 Year Data by Default
 - b. Custom Date Range: User can manually specify the Date for which details are required.
 - c. Go Back to Previous Menu
7. If option 6.a is chosen, it further throws certain Menu options which are required for analysis. They are: -
 - a. Stock Analysis for the {company}.
 - b. Stock Prediction for {company}.
 - c. Visualisations for {company}
 - d. Twitter Sentiment Analysis for {company}
 - e. Home Menu to change to go back to Step 6.
8. If option 7.a. is chosen – it describes the analytics associated with Descriptive Statistics.
 - a. Basic Statistics analysis for {company}.
 - b. Advanced Statistical Analysis for {company}.

- c. Comparison of Stocks between different companies.
 - d. Main Menu
9. 8.a. route selection (Basic Statistics analysis) will display Mean, Median, Range, 1st and 3rd Quartile, Standard Deviation, Coefficient of Variation for Various Menu options as listed below: -
- a. Open Price
 - b. Close Price
 - c. High Price
 - d. Low Price
 - e. Adjusted Close Price
 - f. Volume
 - g. Previous Menu

For a-f, all the details will be shown as mentioned above. Option g. selection will be going back to Step 8 for further analysis as required by the user.

10. 8.b. route selection (Advanced Statistical Analysis) will further display a menu option, which will plot charts and graphs for all provided options
- a. Simple Moving Average
 - b. Exponential Weighted Moving Average
 - c. Moving Average Convergence Divergence
 - d. Go Back to Previous Menu

For a-c, options additional details such as window size will be requested for further graph plotting. Option d. selection will be going back to Step 8 for further analysis as required by the user.

11. 8.c. route selection (Compare Stocks) helps in comparison between different stocks based on their Statistical and Graphical Views as required. Maximum Stocks which can be compared are 3. We can enter different stocks ticker symbols and can compare based on analysis as required. In comparison of stocks, Statistical will give basic descriptive mathematical differences such as Mean, Std deviation, Min and Max values, quartiles etc., whereas Graphical methods compares the Simple Moving Averages and Raw time Series options for the chosen stocks.
12. In the same route, there is a previous menu option which will redirect back to Step 8 for further analysis.
13. All the options are explored in Step 8. Clicking on 8.d., will traverse back to Step 7 for other options.
14. 7.b. is the Predictive Analysis model for the Stock Analysis. There are two options to be selected under the same. They are: -
- a. Prediction using Linear Regression
 - b. Prediction using ARIMA Model
 - c. Main Menu
15. If option 14.a, is chosen Linear Regression is done where different functionalities are incorporated such as prediction can be made for any past date as well based on the training data set chosen and displaying the Mean Square Value, R Square and RMS value as well. Graphical representation is also present for visualisation purposes.
16. If option 14.b, is chosen ARIMA Model for advanced predictive statistics is also used. It works on the concept of Auto Regressive Moving average with graphs provided for the same. It gives advanced variables such as coefficient of error, standard deviation, Z and P value as well.
17. If option 14.c is chosen it goes back to previous menu i.e. Step 7.

18. 7.c. is the Visualisation models for Date Range selected. There are 4 further options which are provided below for 7.c.
 - a. Time Series and Trend Lines
 - b. Candle Stick View
 - c. Histogram for High and Low
 - d. Previous Menu
19. 18.a, if chosen will ask for 2 other options which are: -
 - a. Time Series
 - b. Trend Lines
 - c. Previous Menu

Options a-b will display Time Series and Trend Lines for the Date range through a graphical representation. If option 'c' is chosen, it will go back to Step 18

20. If 18.b. is chosen, it will plot a Candlestick for the data based on their prices and display the values as well.
21. If 18.c. is chosen it will plot Histogram for Open and Low Prices. It will return to main menu after this display at Step 18 by choosing 18.d.
22. All the options have been explored for Step 7.a, 7.b. & 7. c.. Now if option 7.d. is chosen Twitter Sentiment Analysis comes into picture.
23. Twitter Analysis helps in finding the polarity of the company based on the sentiment analysis. An input needs to be provided to check the sentiment based on the limit (n) of tweets which user enters for the company. Based on this limit of tweets it will give either a positive or negative response for the recent trend which can help in decision making process of buying, selling or holding a stock.
24. Last option in Step 7 i.e. 7.e. will roll back to Step 6.
25. The other option in Step 6 in Custom Date Range where a user can input the desired date range i.e. a Start and an end Date.
26. For Custom Date Range the Steps for execution are the same from Steps 7 to 25 apart from the variable Date range provided. The performances and outputs are similar in nature.
27. Step 7.c will exit from the code altogether.
28. Exit from Code.

ERROR HANDLING and EXCEPTION HANDLING

This provides an extensive list in the model where we have applied and handled the errors and exceptions using different conditions under different functions.

- a. `def getCompanyDetails (Company Ticker):`
This function extracts the Ticker Symbol for all the companies listed in NASDAQ website and cross verifies with the ticker symbol entered by the user. If an unknown symbol is added, it will throw an error and return to back to the home screen for entering the correct symbol
- b. `def main () (Company Ticker):`
This function handles the ticker symbol if entered in lower case. It automatically converts the alphabets into Upper case before hitting the companies listed in NASDAQ.
- c. `"__main__":`
While executing the code for first time, if a user wants to understand the flow of the model, they can press Y and view that else exit from that for analysis.
- d. `def validate_date_format (date_string):`
The Date format needs to be YYYY-MM-DD format. If added in any other format it will throw and error and ask the user to enter in specified format.
- e. `def dateValidation:`
In the case for Custom Data Range, where we need to add the start date, it validates the date entered is a valid date
- f. `def endDateValidation:`
In the case for Custom Data Range, an end date needs to be inputted by the user. It validates whether the end date is valid format and not less than the start date. It can be greater than equal to the start date i.e. dateValidation.
- g. `def validate_date_range:`
It validates the Date Range specified is not entered as a future Date and not a valid case.
- h. `def validateWorkingDays:`
Need to check for the days when Start and End Date can either be both Saturday and Sunday. In that case a validation check is placed where it will throw an error in that case. Also, if a date is entered 6 months apart but both are weekends, in that case it will not throw an error. Only consecutive weekends will throw an error and ask the user to input the new dates.
- i. `def_simple_moving_avg:`
In descriptive statistics while displaying Simple Moving Average, integer check has been placed on the Windows size and a match window of 100 is taken into consideration for a longer date range i.e. beyond 250.
- j. `def_exponential_Weighted_MA:`
In descriptive statistics while displaying Exponential Weighted Moving Average, integer check has been placed on the Windows size and a match window of 100 is taken into consideration for a longer date range i.e. beyond 250.
- k. `def comparestocks:` This will compare the stocks for multiple tickers entered and validate the ticker symbol in each case.
- l. In all the cases where Menu options are provided, if a non-provided option is entered, it will throw an error showcasing as invalid option, please chose from the valid options.

BIBLIOGRAPHY

Julia writer, 2017. *Benzinga*. [Online]

Available at: <https://www.benzinga.com/general/17/06/9602734/why-is-stock-chart-technical-analysis-important>

[Accessed 13 June 2017].

Contribution of Each Individual Member (Equal Contribution as “The Data Whisperers”:

- 1. Navin Mutha: - Twitter Analysis, Visualization, Main Menu, Validation**
- 2. Satish Mohan: - Descriptive Analytics, Predictive Analytics, Validation**
- 3. Gautam Gupta: - Report Writing, UML Flow Chart, Code Integration and Testing, Validation**