Stock Analysis

23rd October 2019

KBC Trading™

Authored by:

Kaushal Kumar (19200110)
Bhagyashree Janyani (19200337)
Chaitanya Dave (19201077)

Contributions

Authors of this report Kaushal Kumar, Chaitanya Dave & Bhagyashree Janyani as a team are thankful to Dr. Miguel Nicolau for providing a holistic learning curve throughout the project. We certainly have learnt throughout the journey of this project.

Before going ahead, we would like to express what analytics means to us, something that we learned along the way to our project:

> Business Analytics leverages governance, culture, people and technology capabilities and a systematic thinking process that applies qualitative, quantitative, and statistical computational tools and methods to analyse data, gain insights, inform, and support impactful decision-making [O'Neill & Brabazon (2019)]
>]

The Project targets to help user to answer **When, How** and **Why** of the data of stock exchange via:

a) Descriptive and Current Data
b) Predictive Analytics

Contributors were not limited to any particular segment of the project and had a say in all aspects of it. The outcome of the project was analyzed and agreed mutually amidst all the members. Although there might be instances where a module was led by an individual, but nevertheless had pro-active participation from other members.
Our Contributions to this project would have been in vain if there wasn't any presence of "NASDAQ" and "quotes.wsj.com" for providing unparallel data experience from their repositories.

Table of contents:

Introduction

The KBC Trading™ during the longevity of the document will be referred as tool.

As stated in the contributions, the project targets to present Descriptive and Predictive analytics which in the scope of our project is based on Python.

The tool's purpose is to help Stock Analyst with their day-to-day activity with minimum use of any other technology. Be the one stop solution for what analysts seek. The tool presents its display in a command line where users select one option (out of many) when asked a question by the system. The tool asks necessary pre-defined questions which the users need to answer to go ahead.

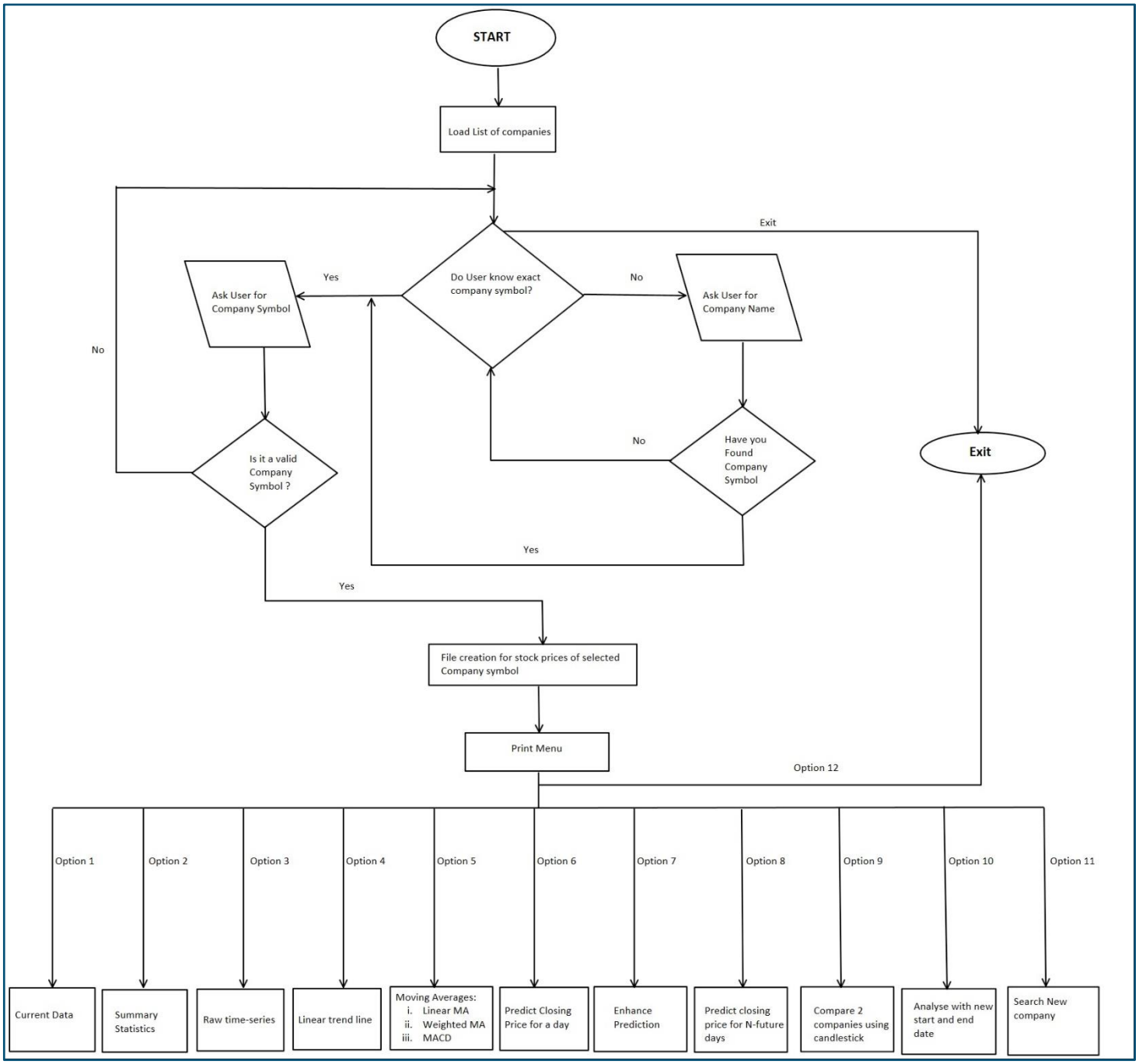
Users can see interactive graphs where they can download and save the graphs as per the user's convenience. User can select amongst analytical options available at their disposal. Tool uses linear, non-linear algorithms and voting mechanism (ensemble) to find the best result for prediction of the future values.

The tool pulls data from NASDAQ, "quotes.wsj.com" and alpha vantage API, these resources are open and reliable sources. In order to reduce the internet hits to these repositories, most of the data is stored locally on the user's machine acting as local database. All the files are stored on the default directory and the program also reads the same files from the same path. It is advised to the users to not change the file locations as it will lead to re-downloading of such data again and again.

The tool comes loaded with usage of some libraries, it is imperative that user check out the installation of those libraries either in their Anaconda PowerShell or Python environment. There are 2 ".py" files which constitute the tool. Both the files are to be placed at the same location and each file requires the presence of the other otherwise the tool will just not start.

The tool gathers current data, historical data and based on timeframe predicts a closing value for a particular date in the future and past. The tool also brings forth the important statistics and graphs on the historical data, along with showing comparison between two companies. The prediction model of the tool employs models from much known Sklearn(also known as Scikit) library of python and used interactive "Plotly" library to depict all the charts.

UML DIAGRAM


START
Load List of companies
Exit
Do User know exact company symbol?
Yes
No
Ask User for Company Symbol
Ask User for Company Name
No
Is it a valid Company Symbol ?
No
Have you Found Company Symbol
Exit
Yes
Yes
File creation for stock prices of selected Company symbol
Print Menu
Option 12
Option 1
Option 2
Option 3
Option 4
Option 5
Option 6
Option 7
Option 8
Option 9
Option 10
Option 11
Current Data
Summary Statistics
Raw time-series
Linear trend line
Moving Averages:
i. Linear MA
ii. Weighted MA
iii. MACD
Predict Closing Price for a day
Enhance Prediction
Predict closing price for N-future days
Compare 2 companies using candlestick
Analyse with new start and end date
Search New company

User Manual

Step 1: The tool can be executed by running the "**kbc_trading.py**" file from either

a) Atom PowerShell
b) Spyder (from Anaconda Suite)
c) any other python environment [out of scope of this document]

Following are the commands to download the libraries:
Commands in Anaconda Environment

1. **conda install scikit-learn**
2. **conda install -c plotly plotly**
3. **conda install -c conda-forge regex**
4. **conda install -c jmcmurray os**
5. **conda install -c conda-forge fbprophet**

or

Commands in Python Environment

1. **pip install -U scikit-learn**
2. **pip install plotly==4.3.0**
3. **pip install regex**
4. **pip install os-win**
5. **pip install fbprophet**

Pro Tip:
[All the files are stored in the user's default directory. Users are advised not to meddle with the files created by the system, unless users want to retrigger fresh stock of all the data which could be a burden to the tool. Tool downloads list of company from NASDAQ (stock exchange in New York, United States) the list is updated only when a new day has started in New York (UTC-4 time zone) because a new company is added at end of day or beginning of day. So, the tool runs once a day to download the list of company and even though the tool runs multiple times from the start, company names are not downloaded again and again.**]**

Step 2: Analyst/User is now requested to either search the company symbol (if the analyst is aware, type 'Y' in any case) or search for the company (if the analyst is not sure, type 'N' in any case) by typing a part of the company name. The later, finds a list of all the company names matching the entered text. If user wants to exit from searching company symbol, type 'E'.

```
Welcome HP to KBC home page !!

If you are asked a question,please check the options and then answer appropriately

Enjoy the Analysis

_____________________________________________________________________________
Data is loaded

Are you sure of company symbol you want to search? (Y/N/E(Exit)) : Y
```

Step 3: On refining the company search symbol, system asks the user to confirm the company symbol, so that other analytical tools can be made available. The chosen company symbol will be validated, if all holds good tool goes to the next step otherwise, tool continues to ask user to finalize the symbol. (User Inputs are case insensitive, although user needs to follow specific date format of YYYY-MM-DD).

[check the UML diagram]

```
Are you sure of company symbol you want to search? (Y/N/E(Exit)) : N

Enter the approx company name to be searched for : alpha
Showing Top 10 results for your search which gave  41  results
         Symbol          Name
         _________       _________
         AOSL            Alpha and Omega Semiconductor Limited
         GOOG            Alphabet Inc.
         GOOGL           Alphabet Inc.
         SMCP            AlphaMark Actively Managed Small Cap ETF
         ATEC            Alphatec Holdings, Inc.
         FPA             First Trust Asia Pacific Ex-Japan AlphaDEX Fund
         FBZ             First Trust Brazil AlphaDEX Fund
         FCAN            First Trust Canada AlphaDEX Fund
         FCA             First Trust China AlphaDEX Fund
         FDT             First Trust Developed Markets Ex-US AlphaDEX Fund

Have you found your symbol yet ? Y/N :
```

```
Have you found your symbol yet ? Y/N : Y

Enter the final symbol : GOOGL
```

Step 4: Following is the available list of activities that a user can do with finalized company symbol. The user needs to select one of the numbers from 1-12 and hit enter. After selecting the option, the system will come back to the same menu, so that users can visualize other aspects of data for the company or select another company symbol or exit the program altogether.

```
Enter the final symbol : GOOGL
/-----------------------------------------------------\
              Stock Analysis MENU of GOOGL
|-----------------------------------------------------|
| 1.  Current Data                                    |
| 2.  Summary Statistic                               |
| 3.  Raw time-series                                 |
| 4.  Linear trend line                               |
| 5.  Moving Averages                                 |
| 6.  Predict close price for a day                   |
| 7.  Enhance Prediction                              |
| 8.  Predict close price for N-future days           |
| 9.  Compare 2 companies using candlestick chart     |
| 10. Analyse with new start and end date             |
| 11. Search New Company                              |
| 12. Exit                                            |
\-----------------------------------------------------/

Enter your choice [1-12]:
```

Step 5: For most of the options selected by the user, there will be timeframe asked from the user. Start date and end date are dependent on the analyst, how much data needs to be analyzed. User needs to provide these dates as and when asked by the system. There are few other analyses which require specific inputs: - Window (in case of moving averages) & Particular date (prediction on that day).
Inputs for Date range: Start Date and End Dates are inputs for option 2 through 9. If user wants to start the analysis with new date range, he/she can select option 10.

```
Enter your choice [1-12]: 2

Please enter Start Date in YYYY-MM-DD format: 2017-01-01

Please enter End Date in YYYY-MM-DD format: 2019-01-01
```

Input for Window Period:

```
Enter your choice [1-12]: 5

Please enter Start Date in YYYY-MM-DD format: 2017-01-01

Please enter End Date in YYYY-MM-DD format: 2019-01-01

Please enter window period in days less than 730: 10
```

Step 6: The tool also allows the user to compare closing data of the two companies so that the analyst can advise which company is better for investment. The Graph would show data of both companies on a common timescale.

```
/-----------------------------------------------------\
              Stock Analysis MENU of GOOGL
|-----------------------------------------------------|
| 1.  Current Data                                    |
| 2.  Summary Statistic                               |
| 3.  Raw time-series                                 |
| 4.  Linear trend line                               |
| 5.  Moving Averages                                 |
| 6.  Predict close price for a day                   |
| 7.  Enhance Prediction                              |
| 8.  Predict close price for N-future days           |
| 9.  Compare 2 companies using candlestick chart     |
| 10. Analyse with new start and end date             |
| 11. Search New Company                              |
| 12. Exit                                            |
\-----------------------------------------------------/

Enter your choice [1-12]: 9

Are you sure of company symbol you want to search? (Y/N/E(Exit)) : Y

Enter the final symbol : AAPL
```

Step 7: User can select another company symbol without exiting the program and continue slicing and dicing just like from step 2 – step 5. To perform this functionality, user can select option 11 for analyzing new company.

```
/-----------------------------------------------------\
              Stock Analysis MENU of GOOGL
|-----------------------------------------------------|
| 1.  Current Data                                    |
| 2.  Summary Statistic                               |
| 3.  Raw time-series                                 |
| 4.  Linear trend line                               |
| 5.  Moving Averages                                 |
| 6.  Predict close price for a day                   |
| 7.  Enhance Prediction                              |
| 8.  Predict close price for N-future days           |
| 9.  Compare 2 companies using candlestick chart     |
| 10. Analyse with new start and end date             |
| 11. Search New Company                              |
| 12. Exit                                            |
\-----------------------------------------------------/

Enter your choice [1-12]: 11
```

Step 8: User can exit the program from option 12.

```
/-----------------------------------------------------\
              Stock Analysis MENU of GOOGL
|-----------------------------------------------------|
| 1.  Current Data                                    |
| 2.  Summary Statistic                               |
| 3.  Raw time-series                                 |
| 4.  Linear trend line                               |
| 5.  Moving Averages                                 |
| 6.  Predict close price for a day                   |
| 7.  Enhance Prediction                              |
| 8.  Predict close price for N-future days           |
| 9.  Compare 2 companies using candlestick chart     |
| 10. Analyse with new start and end date             |
| 11. Search New Company                              |
| 12. Exit                                            |
\-----------------------------------------------------/

Enter your choice [1-12]: 12
Thank you for using KBC trading... Have a great day!!
```

USP's of the Product: -

1. More personalized approach to design, to match NY time, system automatically handling time zone changes, used UTC time.
2. Personalized Welcome Message on the landing page.
3. For Database, we use CSVs, same file is reused, instead of hitting the network again and again since these data is stored in csv files, data can be shared with other platforms easily. In cases of file of company being lost, up-to date data can be fetched almost immediately by the system automatically.
4. Search function of a company, pattern search using regular expression.
5. Case in-sensitive search everywhere
6. Flexible user menu designed with minimalistic approach; repetitive questions avoided
7. Highly interactive graphs, users can download and check data at low & High granular level data, giving a better user experience
8. Elements of artificial intelligence introduced in prediction, with concepts of hyper parameter tuning.

References:

1. https://stackoverflow.com/questions/44290635/python-how-to-convert-datetime-data-using-toordinal-considering-the-time
2. https://stackoverflow.com/questions/16453644/regression-with-date-variable-using-scikit-learn
3. https://scikit-learn.org/stable/modules/ensemble.html#voting-classifier
4. https://plot.ly/python/subplots/
5. https://www.learndatasci.com/tutorials/python-finance-part-3-moving-average-trading-strategy/
6. https://towardsdatascience.com/implementing-macd-in-python-cc9b2280126a
7. https://towardsdatascience.com/time-series-forecasting-with-prophet-54f2ac5e722e
8. https://www.dummies.com/programming/big-data/data-science/how-to-use-python-to-plot-time-series-for-data-science/