UCD MIS41110 Business Analytics
Assignment: Group Project

Miguel Nicolau

October 2018

1 Introduction

The purpose of this assignment is to apply the knowledge you have acquired so far in this module to a real-world scenario. This is a group assignment. It facilitates application of several techniques, namely control flow, I/O, and the use of typical analytics packages in Python 3.

2 Financial Data

Buying and selling of stocks (or bonds and other securities) is a potentially profitable business, but can also lead to big financial losses. Business Analytics can provide insights to better inform choices in this industry.

When designing a stock portfolio, some of the decisions to take are:

- The capital to invest;
- The risk the investor is willing to take;
- The time the investor will wait for his/her returns.

The use of analytics methods can help with some of these decisions.

3 Assignment Details

Your task is to write a computer program, which allows users to consult/analyse/model stock times-series. It will source its data online, but optimise its use of downloaded data, to avoid excessive network traffic.

Users will have the possibility of searching for specific stocks, and query specified time ranges, along with associated analysis, such as statistical descriptions of prices and/or volume (mean, median, range, etc), technical indicators, visualisation (of the raw data, but also of transformations, such as moving averages), and even basic modelling (such as regression). Data will be restricted to daily frequency.

The program should provide a comprehensive but easy to use text interface (i.e. no graphical interface), based around a carefully designed structure of menus and sub-menus. It should be easy to use, and require no programming knowledge from the end user.

3.1 Data Gathering

Your first task will be to load a list of active trading companies. There are several sources of such data. For example, the NASDAQ stock exchange offers a direct CSV download:

`http://www.nasdaq.com/screening/companies-by-name.aspx?letter=0&exchan0ge=nasdaq&render=download`

You should also provide the ability to source data for any company. Trading data is widely available; see for example:

`https://www.quantshare.com/sa-43-10-ways-to-download-historical-stock-quotes-data-for-free`

3.2 Description

Your program should be able to provide descriptive analytics of the trading price for any company the user selects. These include (but are not limited to):

- Mean;
- Quartiles;
- Range;
- Standard variation;
- Coefficient of variation;
- etc.

It should also provide a way to graphically visualise data related to any company, such as:

- Raw time-series;
- Trend lines;
- Moving Averages (e.g. MA(n), with user-selectable n);
- Weighted Moving Averages;
- Moving Average Convergence/Divergence (MACD);
- etc.

3.3 Prediction

Your program should provide closing price predictions for companies specified by the user, with regression models, using the following process (or similar):

1. the user specifies the modelling period (i.e. the training data);
2. They also specify a date for which they require a prediction;
3. A linear model is built, using the specified period;
4. The prediction is produced, along with the model's RMSE and R^2 value (coefficient of determination).

Any other prediction approaches you may choose to implement (e.g. non-linear regression models) will add credit to your project.

linear regression model

4 Deliverables

The following deliverables are required:

- A PDF report, comprised of:
  - An appropriate cover page;
  - A statement outlining the contribution of each group member;
  - A user-manual for your program (between 1 and 3 pages);
  - An upper level UML Activity Diagram, describing the main functionality and flow control of the program.
- Code submission.

Both deliverables will be submitted through BrighSpace.

5 Practicalities

- This is a group assignment, for groups of three students;
- The PDF Report must state the name and ID number of all students on the first page (the cover page), and the title of the project;
- Plagiarism, including accidental plagiarism, may result in penalties including grade penalties, failure of the module, and disciplinary action:

  `http://www.ucd.ie/registry/academicsecretariat/plagiarism.htm`

6 Submission

- **Deadline for submission is Friday 23rd November 2018 at 5:00pm**. Both your PDF and your source code files must be submitted through BrightSpace by the deadline.
- You are free to discuss your assignment with other students, but not to look at any other group's work, or show your work to any student outside your group. If you use quotes or paraphrases, images, or ideas from outside sources, you must give a citation. Plagiarism, including accidental plagiarism, may result in penalties including grade penalties, failure of the module, and disciplinary action.
- Please be aware that your submissions will be thouroughly checked for plagiarism; this includes the use of BrightSpace's own plagiarism tool (which checks online code), and also a proprietary tool that compares all submitted projects (including from past years).

7 Evaluation

The projects will be evaluated in week 12. This will consist of a short 10 minute interview with all group members, to demonstrate the final software, and discuss and justify the design and programming choices made. Each student will be asked questions in turn, concerning both the functionality of the program, and the Python source code.

The main criteria for evaluation will be:

- The ability to read online data;
- Graceful handling of erroneous data, such as:
  - Invalid menu options;
  - Non-existing ticker codes;
  - Illegal time-ranges.
- General user-friendly nature of program interface;
- Good program structure and efficient implementation:
  - The use of packages such as numpy, pandas, matplotlib, scipy and datetime is not obligatory, but it is recommended;
  - Dividing your program into several modules, if it helps improve readability, maintainability and/or reuse, is also recommended.
- Sensible variable naming and extensive (but clear) use of comments.