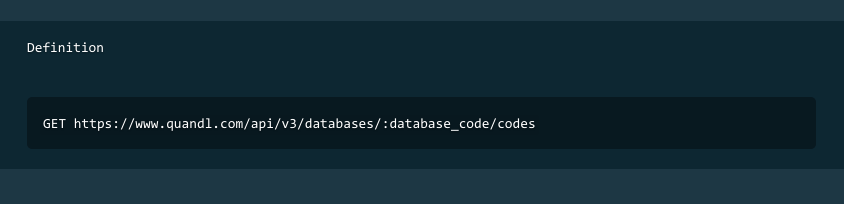
**Quandl database FinSentS Web News Sentiment vs Selected S&P 500 Stock Price Analysis**

Aim

Find the correlation between the sentiment value with stock price using Python language and related libraries

Parsing Database from Quandl



I first look through Quandl api documents and downloads both google finance and FinSentS database list to parse stock ticket symbol in the later procedure.

I had chosen 20 mainly S&P 500 US stock because of the completeness in sentiment signals and diversity from different industries to fully representation the trend of the current US market.

1. Car Manufacturing : Tesla, General Motor
2. Manufacturing : 3M
3. Construction machinery and equipment ： Catepillar Inc
4. IT ： Apple,Facebook, Amazon, Microsoft, Ebay, Yahoo
5. Oil & Energy : Exxon Mobil
6. Electronics : General Electrics, Intel Corp
7. Agriculture : Syngenta, Monsanto
8. Pharmacy : Allergan Inc, Abbott Laboratories
9. Banking and Finance : JP Morgan, Bank of America, Master Card

Using Quandl API library, I had parsed the selected stock sentiment data and history market price from FinSents database and google finance respectively.

For the ease of future analysis, I saved the parsed data into pickle module.

Analysis

After setting up the database, I parse out the sentiment signal as well as opening price in the same date. The opening price will be the bought price of the stock. The next 1,3,5,7,9,13,15 days of closing price are also stored in the same row along with our opening price and sentiment signal.

Next, we calculate the profit or loss margin and log return for normalization purpose.

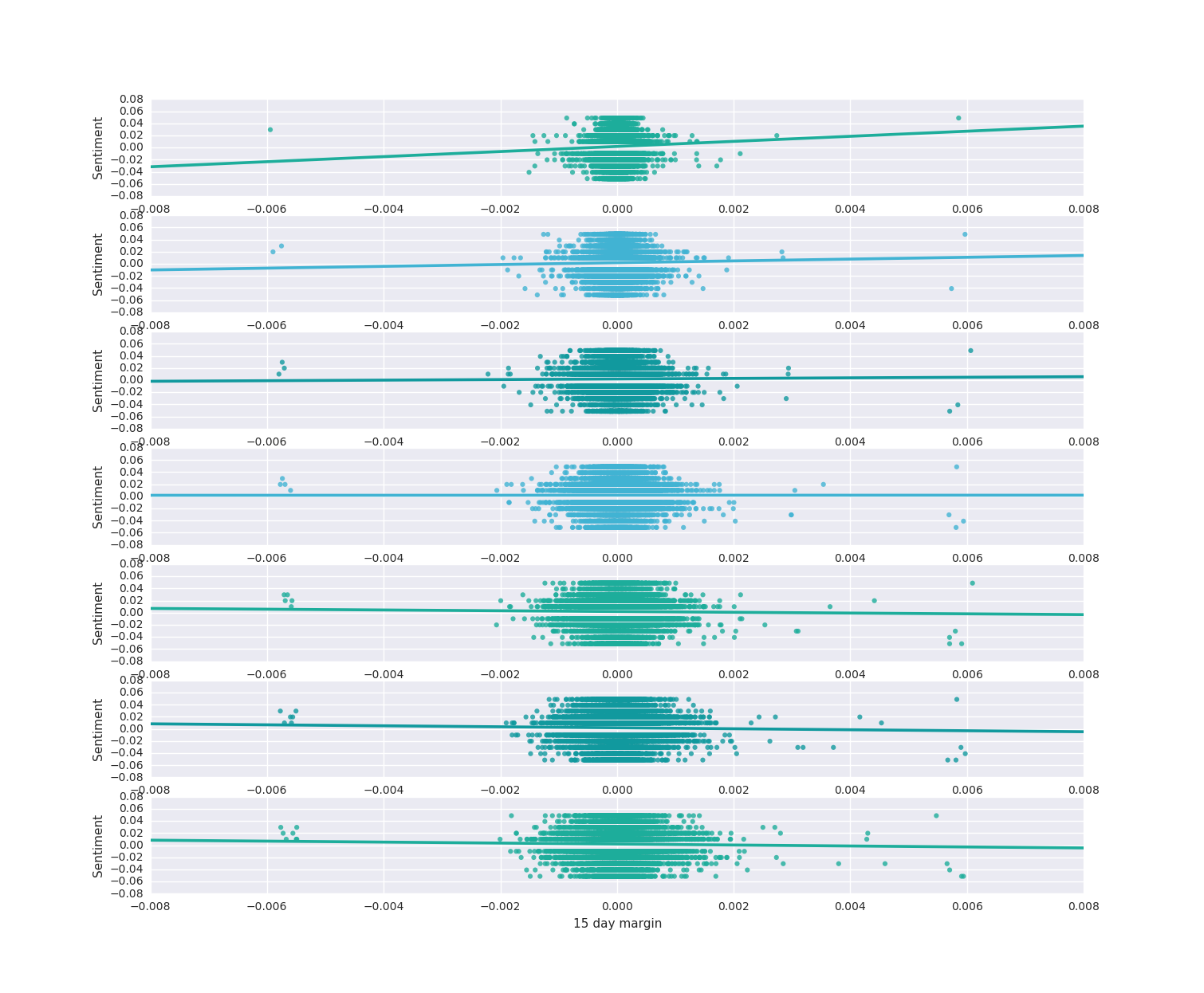
I use seaborn ploting library to plot out the linear regression of profit margins of each closing day with sentiment signal.

On the other hand I also calculate the pearson correlation coefficient of each closing day for better understanding of the relationship between loss or profit margin, closing day and sentiment signal.

Results

The below plots shows the regression plot of each closing day starting by holding the stock for 1,3,5,7,9,13,15 days and sold it during closing price.

Top plot is the shortest holding time of 1 day and the last plot holds the stock for 15 days before selling.



Numerical data of the respective plot are shown below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Days | Slope | R Value | P Value | Pearson correlation coefficient |
| 1 day margin | 4.194 | 0.033539 | 5.95755 | 0.033540 |
| 3 day margin | 1.479 | 0.016189 | 0.05293 | 0.016190 |
| 5 day margin | 0.469 | 0.006218 | 0.457565 | 0.006219 |
| 7 day margin | 0.008 | 0.000115 | 0.98900 | 0.000115 |
| 9 day margin | -0.638 | -0.010787 | 0.197499 | -0.010787 |
| 13 day margin | -0.821 | -0.014969 | 0.07395 | -0.014969 |
| 15 day margin | -0.804 | -0.015631 | 0.062312 | -0.015631 |