**Weather Effects on the Stock Market**

Hyeoncheol Baik, Visswanath Venkataraman and Xing Wang

Although many researchers and practitioners has investigated the stock market behavior, it is still a challenge to predict stock prices. They identify that the stock market are strongly intertwined with economic indicators. However, some exogenous factors should be considered because they sometimes make the stock market respond irrationally. Thus, in this paper, we present the relationship between weather (the degree of sunshine) and stock market (stock price). Another examination by different industries is given to see from micro-level perspective. We also develop some visualizations on the web-based interactive platform, hoping to give better understanding of the results.

# Introduction

Over the past decade, much research has been done to identify important factors on the stock market behavior and make accurate stock market forecasts. Most causes are economic indicators such as unemployment rate, gross domestic product growth and consumer price index. However, we have to take an account into exogenous factors such as scandal and natural disasters because the stock market sometimes shows an irrational manner responding to these factors. One of reasons for this phenomenon might be because the, as a human, the investor’s mood is apt to influence significantly on decision-making behavior.

Among these external factors, we are interested in weather effects on the stock market. As increased amounts of sunshine, more people start to feel happier and more energetic. We believe that many people agree with this statement based on psychological findings and human intuition. This fact also implies that investors have same effect such that they would judge the future stock market more positively, resulting in betting higher stock price. On the other hand, investors would be reluctant to take more risk or tend to make more mistake on gloomy days, resulting in decreasing stock price.

These observations lead us to examine whether sunshine really makes an effect on stock price or not and by how much if the case is right. Therefore, in this project, we aim to identify the relationship between the degree of sunshine and stock market behavior (more specifically stock price trend). Another objective through this project is to present the findings when we look at different industries.

The remainder of this paper is organized as follows. Next section shows relevant literature. Then, we describe how we have coded the data into the interactive visualizations and analysis results. Finally, conclusions and future work are uncovered.

# Relevant Literature

This is a relevant literature section.

# Development of the Visualizations

In this section, we start with providing the data collection and processing, then show how the visualization has been developed.

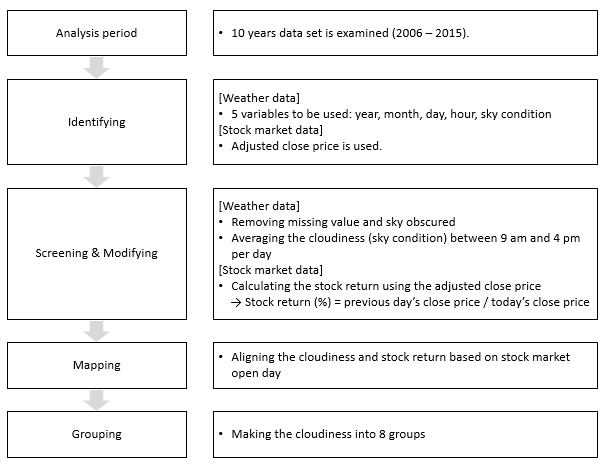
## Data Collection

There are two types of data set used in this project: weather and stock market data. First, in order to get weather information, we explore online data from the National Environmental Satellite, Data, and Information Service (NESDIS) of the National Oceanic and Atmospheric Administration which is operated under the U.S. Department of Commerce: <http://www7.ncdc.noaa.gov/CDO/cdoselect.cmd>. The service website provides a variety of weather data set over worldwide. We use the data set observed at LaGuardia airport in New York City where is the closest weather station to the New York Stock Exchange (NYSE). This is because we assume that Wall Street investors have most power to drive stock market behavior.

Then, in order to get stock market information, we use Yahoo Finance website: <http://finance.yahoo.com>. The website offers historical stock data of major stock indexes and each companies in the form of spreadsheet. This allows us to manipulate the data easily in Microsoft Excel software. In particular, we deal with the S&P 500 and Nasdaq Composite index to compare between large and whole range of companies. The S&P 500 consists of 500 large companies having common stock listed on the NYSE or Nasdaq while the Nasdaq Composite is based on more than 3000 equities traded on the Nasdaq exchange.0

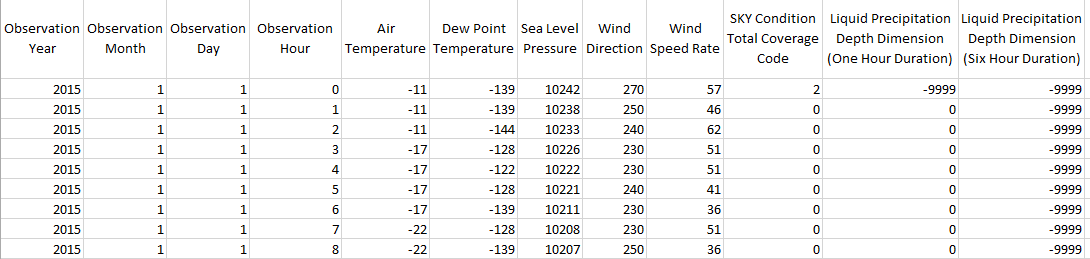
## Data Processing

Since the weather and stock market data set have lots of variables, we have to manipulate them in order to conduct appropriate analysis for our purposes. The Figure 1 shows the data processing flow chart. First, we decide to examine 10 years between 2006 and 2015.



**Figure 1 Data Processing Flow Chart**

Second, it is necessary to identify appropriate variables for the analysis that we are looking for. For the weather data, we obtain year by year information at particular station (here, LaGuardia airport) through FTP access. The data has 12 fields such as day, air temperature, wind speed rate and liquid precipitation depth.0 The Figure 2 illustrates raw data that are composed of 12 variables. Among them, we select 5 variables: year, month, day, hour and sky condition. In meteorology, the amount of cloud cover is evaluated in terms of how many eighths of the sky are covered in cloud, ranging from 0 (completely clear sky) to 8 (completely overcast).0 In this project, we call this metric as the cloudiness for easy understanding. For the stock market data, we make use of adjusted close price because this is a normalized variable that can access all stock price on the one dimension.



**Figure 2 Raw Weather Data Set at LaGuardia Airport**

Third, we go through both screening and modifying process. The cloudiness are measured hourly. However, we are not able to see all the time because the missing data exists and sometimes the sky is too obscured to observe the amount of cloud. Those cases are screened, then we take the average cloudiness number per one day between 9 am and 4 pm of which time window is aligned with the stock market opening hours (9:30 am and 4 pm). The stock return is calculated by dividing previous day’s close price by today’s close price in the unit of percentage.

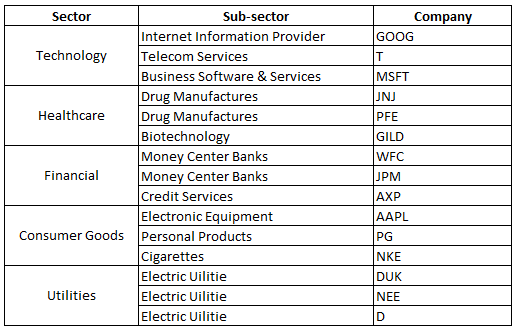
Fourth, mapping is necessary among two data set because all 365 days have the cloudiness data while only official days have the stock market data. Thus, after this work, each combined data set includes stock return and corresponding cloudiness. Lastly, we classify the cloudiness into 8 groups in order to convey sky condition information easily. That is, the user can observe whether the first group represents the sunny day and the last group shows the overcast day.

## Visualization Platform

To understand the stock return change by each cloudiness group, we choose to use a bar chart. This is because (i) the bar chart is appropriate for comparing 8 groups, (ii) adding several gridlines would be help to read the stock return value approximately, and (iii) the range of stock return is enough to use zero-based axis. On the other hands, we make a line chart to examine the stock return change associated with each cloudiness group over years and by different industries. This is because the line chart is well known to represent the trend of data.

We use two types of visualization: Shiny and D3. The Shiny is a web application framework for R, and D3 is also web visualization based on JavaScript. Two filtering options are implemented on the shiny visualization. For the first part result, the user can observe the data associated with some particular or all years through the filtering option. Similarly, the user can look at the trend over years for some particular or all cloudiness groups through the filtering option for the second part result.

The D3 visualization provides three interactive functions. It is noted that (i) the user can adjust opacity of the line color in the graph, (ii) the user can export the whole data set as a spreadsheet, and (iii) the user can switch y-axis (each cloudiness group) toward left or right. In addition, Figure 3 shows industries of interest are 5 sectors that composed of 3 representative companies by market capitalization.



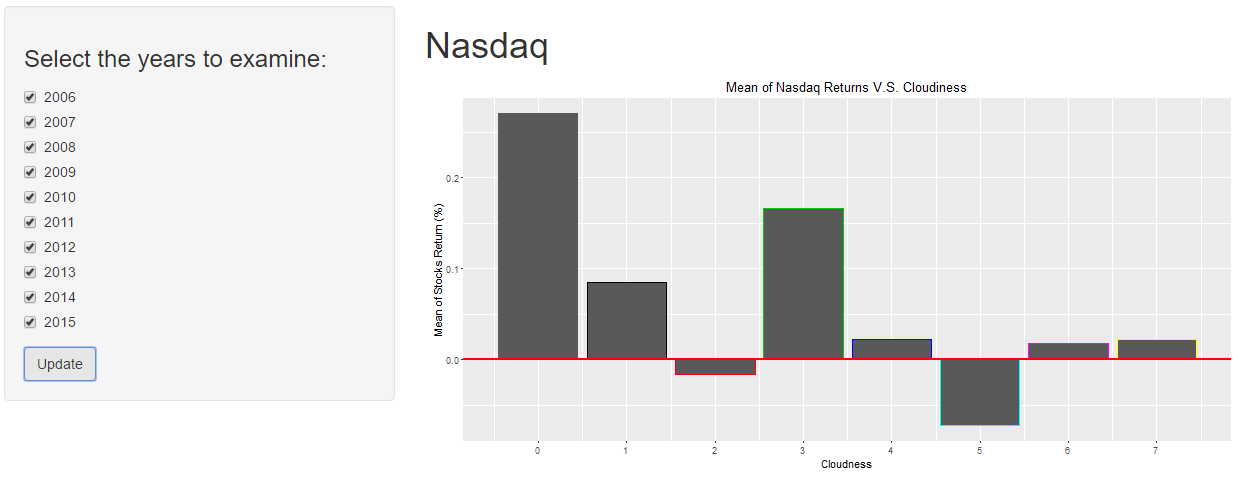
**Figure 3 Five Industries of Interest for Third Part Analysis**

# Analysis Results

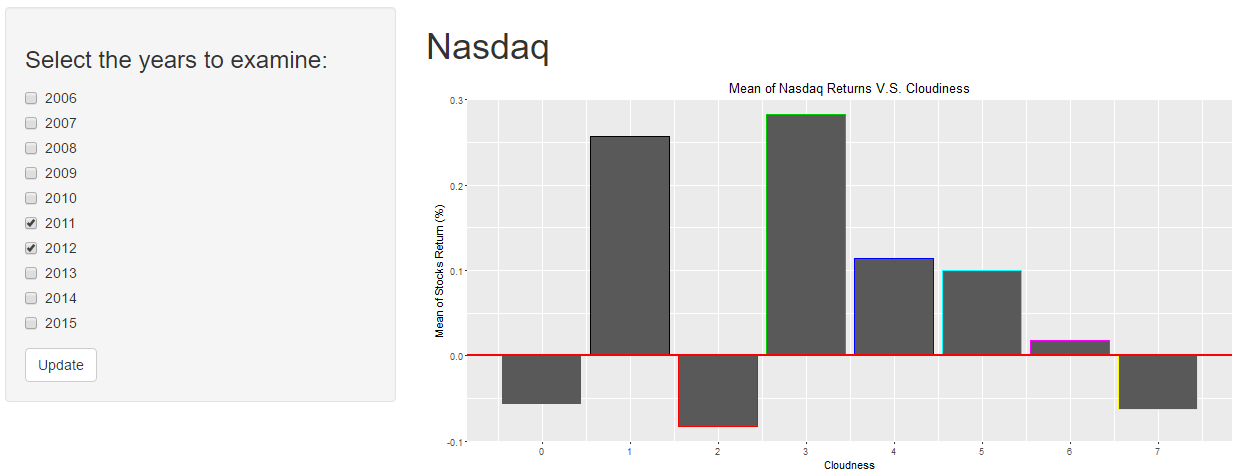
We describe analysis results with each research question and its related visualization followed by our findings and some implications. The result has three parts. In the first two parts, we use the Shiny which is a web application framework for R. In the third part, we use the D3 visualization.

## Cloudiness and Stock Market Indices

1. *How has the stock market behaved associated with varying cloudiness?*

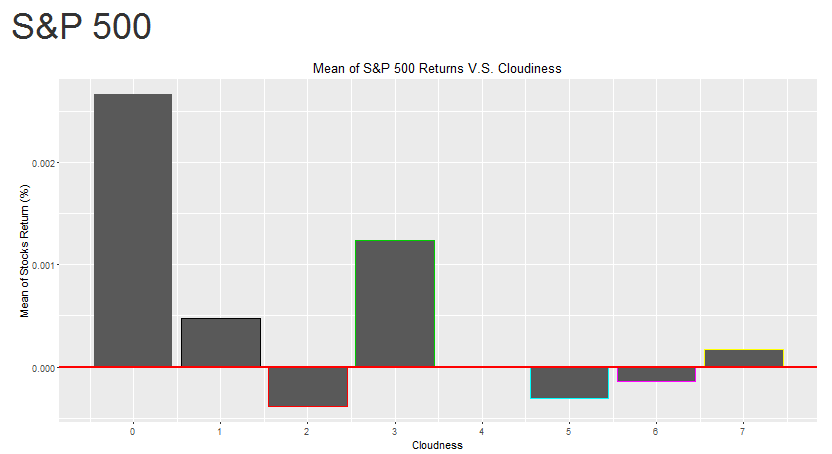


On sunny day (cloudiness: 0), the Nasdaq stock return has much higher performance rather than other days as shown in Figure 0. We observe that sunny day’s stock return (0.27%) is over thirty-four times higher than the worst case of which cloudiness is 5, and over thirteen times higher than the cloudiest day. The overall figure also indicates that the stock return seems like follow a wave pattern.



As shown in Figure 0, it is noted that the stock return has negative number in 2011 and 2012 while it has positive one on all other years.

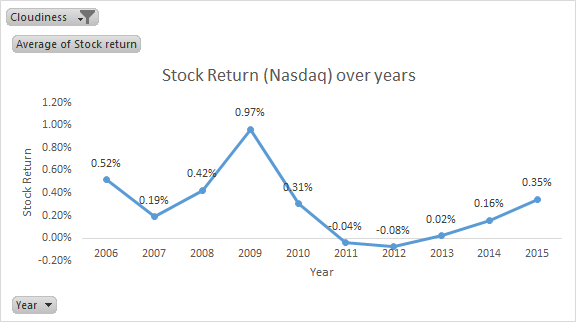
1. *Is there any difference between two stock market exchanges (Nasdaq and S&P 500) associated with varying cloudiness?*



The S&P 500 shows quite similar tendency of Nasdaq as illustrated in Figure 0. On sunny day, investors receive much higher return (0.27%) that is over thirty-one times higher than the worst case of which cloudiness is 2, and over thirteen higher than the cloudiest day.

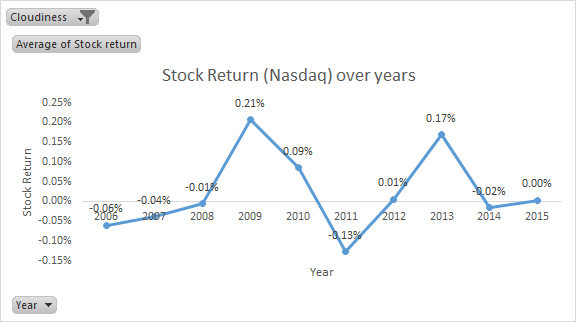
1. *For the Nasdaq, how has the stock return changed over years on some particular cloudiness conditions?*

* Stock returns when cloudiness is 0:

****

Finding from the graph(figure 3) above for Nasdaq show that on sunny days, investors did receive a very high return in the year 2009 (0.97%) which was way over 1% higher than the worst case in 2012(-0.08%). However after 2012 we can that the stock returns have followed an upward trend. Another observation is that during the recession of 2008, we can see that the stock return was 0.42%.

* Stock returns when cloudiness is 7:

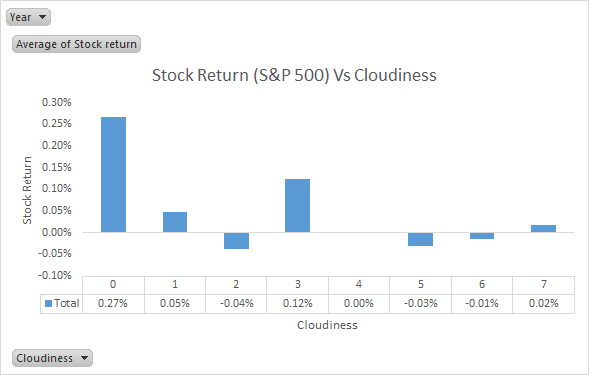


Findings from the graph(figure 4) above for Nasdaq show that on overcast days, investors did observe an up down trend over the years. One key observation to note is that in the year 2008 during recession the returns is -0.01% as compared to 0.42% when the cloudiness is 0. We can also observe that in the recent years (2014,2015) the stock returns have been minimal when the weather conditions are overcast.

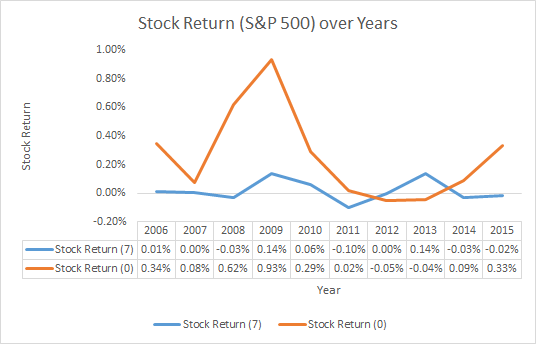
* Stock returns when the cloudiness vary during the day:

Findings from the graph(figure 5) above, we can see that recession affected the year 2008 very badly when the stock returns was -0.17%. Also post the recession, the market picked up quickly & has since then followed an up down trend. As we can see the stock returns in the recent years have been declining at a slow pace.

1. *For the S&P 500, how has the stock return changed over years on some particular cloudiness conditions?*



Finding from graph(figure 6) for the S&P 500 is similar to NASDAQ. We can observe that investors are more prudish & conscientious when it is sunny as compared to when it is overcast. The maximum stock return of 0.25% during sunny days is more than 13 times higher than when the stock returns are the worst (-0.04% when cloudiness level is 2).



Findings from the graph(figure 7), very clearly show that during sunny days the year 2009 had a very good stock yield however it wasn’t that good in the year 2013. The stock returns when overcast has however maintained a constant trend despite some slight rises over the 10 year tenure. We can however visually observe the fact that the stock returns during sunny days is starting to follow an upward trend from 2014.

## Cloudiness and Stock Return by Different Industries

1. *Which industry has been the most affected by weather (cloudiness)?*

State here.

# Conclusions and Future Work

This is a conclusion and future work section.

# References

This is a reference section.

[0] <http://www.visualcapitalist.com/whats-difference-dow-sp-500-nasdaq>

[0] National Centers for Environmental Information, 2006, “Integrated surface data – lite format documentation”

[0] <https://en.wikipedia.org/wiki/Okta>

[\*] Chen, C-T., Lin, C-T., Huang, S-F., 2006, “A fuzzy approach for supplier evaluation and selection in supply chain management”, *International Journal of Production Economics* 102, 289-301

# Appendices

## Shiny Visualization Screen Capture

Attached here

## D3 Visualization Screen Capture

Attached here