

How Does Google Search Volume Index Affect Stock Trading Volume and Stock Price

Xinyu Zhang

University of Oklahoma

May 8 2019

1 ABSTRACT

This paper intends to determine how does Google Search Volume Index (SVI) affect the stock price and trading volume. In order to identify the result, in this paper, I will use five of the most variable companies from five different industries in the American as my study sample. The data I used is the weekly search volume index in 2018 from google trends and stock price and trading volume will be collected from NASDAQ. Basically, I will use the weekly search volume index, stock price and trading volume generating linear regressions to see the effect between them. After the research, I found that google search volume index will have a positive effect on the stock trading volume. We could not decide how does google search volume index affect stock price since the movement of stock price is undefinable.

2 INTRODUCTION

So far, the stock is one of the most popular financial assets either for people who are trying to make the profit from the financial market or for those of who are trying to make a saving plan for future. Because stocks are not like the certificates of deposit which only generate a fixed return or commodities, such as gold and silver, have the highest risk. In the stock market, people will have diverse options for their investment, either high risk high return or low risk low return. As these reasons, there are more and more people are attracted by stocks. Especially, for those of who have the knowledge of computer science and finance. Such as, financial engineers. they have ability to use both existing public information and personal knowledge identifying the factors that affect the stock price and predict the future price. Such as, the factors of search volume index, which is a kind of new theory that people mentioned affecting the stock price.

Search volume index is an index which represent search interest for a certain term in a certain period and the range of index is between 0 and 100. The number of 100 is the highest number of a term which means that this term has a highest popularity during that period. The number of 50 means that the popularity of this term is in the middle range. A number of 0 means that the information is not enough to generate the index [3]. This volume index is not only used in the financial market, but also people will use it in many different industries. Such as, in the paper of Predicting the Present with Google Trends, the author used the search volume index in google predicted the sales in automobile, unemployment rate, consumer confidence, and travel destination. [2]. As long as people are searched that term, we will be able to use that data to either find the factors or predict future.

Even there are many people using the search volume index to predict future, in this paper, I am not going to predict the stock price. Because I think that the stock price is too sensitive for us to predict. It is true that there is many forecasting of stock price, but the range of the forecasting is very broad, which means that everything in the future could affect the price of a stock. Just as the result from exploring the relationship between google trends data and stock price data, the author concluded that even the stock price is highly correlated with some other variables, it is not a good idea for them to invest base on what they found [7]. As these reasons, in this paper, I will only determine how does the search volume index affects the stock price and trading volume.

3 LITERATURE REVIEW

Just as I mentioned in the above, there are many people have done research which related with the stock price and search volume index. However, the paper I found is very interesting is not directly related with search volume index, it is related with peoples attention and advertisement. In the paper of *Advertising, Attention, and Stock Returns*, the authors had mentioned that more advertisements of a company means that more people will pay attention on that company and more attention of that company will bring a large contemporary return and a small future return for that stock [1]. Even this paper did not mention anything related with search volume index, it had pointed out that peoples attention will affect the stock return. As the development of digitization, I think that search volume index is a better way to represent peoples attention since in our current life, if we are interest in something, we will use internet learning about that. As more and more people search that term the search volume will be larger.

As an example of proving the direction of my paper is correct, I found that in paper which is called *Google Search Volume Index: Predicting Returns, Volatility and Trading Volume of Tech Stocks* determined that there was a strong positive correlation between google search volume index and stock trading volume, stock return, and weekly volatility on an aggregate level [6]. As the author summarized, we realize that as the increase in the google search volume index, there will be an increase in stock trading volume, stock return, and weekly volatility. There is not only one study which concluded the result as this. the same result shows in the paper of Google searches and stock market activity: Evidence from Norway. the author concluded that as the increase in google searches, the volatility and trading volume will be increased [4].

The same conclusion has showed in study of *google search intensity and its relationship with returns and trading volume of Japanese stocks*. However, the difference is that in Japanese stocks market the relationship between stock return and search volume index is weakly positive. Furthermore, the author pointed out that the probability of increased trading volume rises the stock price is not high [8](Japanese stock,2014). The difference in a strong positive relationship and a weak positive relationship is not able to prove which of the result is better than another one or the correctness of the model. It may cause by the difference in samples and time of the data collected since the sample will directly affect the result of a study.

According to the author, the stock samples of the study on Japanese stocks were 189 Japanese stocks in between 2008 and 2011 But the stock samples of XUs paper were technology company in the United States and the time range is from 2004 to 2015. As this reason, we may understand why

there is slightly different between Xus and Fumikos paper. Reason of low probability in increased trading volume causes an increase in price may be that the lags may or may not existing for each of related news. For example, for those of news which directly related with that stock should affect the stock price immediately. However, for those of having an indirect effect on a certain stock, the lags or reaction time may take longer we expect. So, as this reason, it is hard for us to define whether an increase in stock price is driven by today's positive news or the news published a couple of weeks before.

In the study of *the different impacts of news-driven and self-initiated search volume on stock price*, the author concluded the relationship between stock price and search volume index based on a novel view, which they distinguished whether the increase in search volume is caused by public news or caused by the investor themselves. They summarized that if the increase in search volume is driven by the investor themselves, the stock price will increase and if the increase is driven by public news, the stock price will decrease [5]. In this paper, the author defined the relationship more detailed, and defined the reason of increase in search volume index, because the limitation of the data, I am unable to include these variables in my regressions.

4 Data

Due to the number of public traded companies in the US is too large, I am unable to include all the companies in this study. I will select top five of the most valuable companies from different industries in the United States and the value defined as the market value at the end of 2018. The companies in table 1 will be the sample of this study. The reason why I used five of the most

valuable companies in the US as my sample is that first, avoiding the potential effects which only exist in one certain industry. Second, the potential problem of liquidity, since I samples which I selected have the highest value in the US, there will be no problem for the stock liquidity.

The primary data of this study will be the weekly search volume index in 2018 in google trends of these five companies because google search has already became the most popular search engine over the world, the volume index in google trends will be better representing the search preference for a period, and the location of people search that term will be in the United States because the sample are American companies. The stock price and stock trading volume of these five companies will be the weekly price in NASDAQ in 2018.

When I collect the search volume index from google trends, I realized that google search engine will recognize the wrong typo and automatically show the potential correct result. Furthermore, the result from google search will be the same as long as the meaning of searching is the same. For example, in google search, the result of searching apple price and the stock symbol of AAPL is the same, which shows the current stock price of Apple. However, the result of search volume index in google trends for these two inputs will be different. To solve this issue, in this study, I will only use the stock symbol as an input in the google trends, because the purpose of people who search stock symbols must try to find the stock price for a certain company. However, purpose is unsure for those of who search apple price. I also will include the stock symbols for these five companies in table 1 and use trading volume, search index and stock price generating the figures for these companies to have a general view of these data.

5 Empirical Methods

In order to find the result of how does search volume index affect the trading volume and stock price. I will basically generate two kind of equations. First, I use the stock price as the dependent variable and current search volume index and previous search volume index as my independent variable generating a regression to see the relationship between stock price and search volume index. Second, the trading volume will be the dependent variable and current search volume index and previous search volume index as the independent to see how google search volume index affect trading volume. The full equations will be in the below and the description will follow by the equations.

$$\log APL.V = \beta_0 + \beta_1 APL.INDEX + \beta_2 APL.INDEX_{t-1} + U \quad (1)$$

$$\log BRKB.V = \beta_0 + \beta_1 BRKB.INDEX + \beta_2 BRKB.INDEX_{t-1} + U \quad (2)$$

$$\log XOM.V = \beta_0 + \beta_1 XOM.INDEX + \beta_2 XOM.INDEX_{t-1} + U \quad (3)$$

$$\log JNJ.V = \beta_0 + \beta_1 JNJ.INDEX + \beta_2 JNJ.INDEX_{t-1} + U \quad (4)$$

$$\log FB.V = \beta_0 + \beta_1 FB.INDEX + \beta_2 FB.INDEX_{t-1} + U \quad (5)$$

where $\log APL.V$, $\log BRKB.V$, $\log XOM.V$, $\log JNJ.V$, and $\log FB.V$ are the weekly trading volume for each company. $APL.INDEX$, $BRKB.INDEX$, $XOM.INDEX$, $JNJ.INDEX$, and $FB.INDEX$ are the weekly google search volume index for each company. $APL.INDEX_{t-1}$, $BRKB.INDEX_{t-1}$, $XOM.INDEX_{t-1}$, $JNJ.INDEX_{t-1}$, and $FB.INDEX_{t-1}$ are the lag of weekly google search volume

index.

$$\log AAPL.PRICE = \beta_0 + \beta_1 AAPL.INDEX + \beta_2 AAPL.INDEX_{t-1} + U \quad (6)$$

$$\log BRKB.PRICE = \beta_0 + \beta_1 BRKB.INDEX + \beta_2 BRKB.INDEX_{t-1} + U \quad (7)$$

$$\log XOM.PRICE = \beta_0 + \beta_1 XOM.INDEX + \beta_2 XOM.INDEX_{t-1} + U \quad (8)$$

$$\log JNJ.PRICE = \beta_0 + \beta_1 JNJ.INDEX + \beta_2 JNJ.INDEX_{t-1} + U \quad (9)$$

$$\log FB.PRICE = \beta_0 + \beta_1 FB.INDEX + \beta_2 FB.INDEX_{t-1} + U \quad (10)$$

where $\log AAPL.PRICE$, $\log BRKB.PRICE$, $\log XOM.PRICE$, $\log JNJ.PRICE$, and $\log FB.PRICE$ are the weekly trading volume for each company. $AAPL.INDEX$, $BRKB.INDEX$, $XOM.INDEX$, $JNJ.INDEX$, and $FB.INDEX$ are the weekly google search volume index for each company and $AAPL.INDEX_{t-1}$, $BRKB.INDEX_{t-1}$, $XOM.INDEX_{t-1}$, $JNJ.INDEX_{t-1}$, and $FB.INDEX_{t-1}$ are the lag of weekly google search volume index.

The reason of using natural logarithm in trading volume and stock price is that the numerical number of trading volume is huge. If I use the numerical number to generate the regression, the result of changes in trading volume be hard to define. The reason of using log price is that the change in percentage of the stock price will be clearer than the change in numerical. Even the numerical number for stock price is not as large as the trading volume, it will be clearer for us to see the change in percentage.

6 Research Finding

The main result of this study will show in table 1 and table 2. The table 1 is the statistical summary showing how does google search volume index affect the stock price for these five different companies. The coefficient of AAPL index is 0.003 which means one percent increase in search volume index will result 0.3 percent of increase in trading volume. The estimate of lagged AAPL index is even larger compare with the AAPL index since the coefficient is 0.013 the R square of AAPL index is about 21 percent, which is in the middle range compare with other four equations.

The coefficient of BRKB index, XOM index, JNJ index and FB index is 0.002, 0.003, 0.013 and 0.002, which all of these four companies have a positive correlation between search volume index and trading volume. As an interpretation of those numbers is that as one unit increase in google search volume index will increase the trading volume by 0.2 percent, 0.3 percent, 1.3 percent, and 0.2 percent for these companies. Furthermore, even the level of percentage change of lagged search volume has a slightly different compare with the original index, it still shows the positive effect to the trading volume. The R square for these four companies is in the range from 11 percent to 46 percent.

The information in table 2 shows the results of regressions of stock price and search volume index for these five different companies. The estimate of AAPL index is about -0.0002, as this result, we can say that as one unit increase in AAPL index, the stock price of AAPL will decrease about 0.02 percent which is a very small effect on the stock price of AAPL. But the main objective of the paper is finding how google search volume index affect the stock price, we will ignore

whether there is large effect on the stock price and trading volume or small effect.

The estimates of BRKB index, XOM index, JNJ index, and FB index are 0.0002, -0.001, 0.0002, and 0.0004 which interpret as one unit increase in google search volume index will bring a 0.02 percent of increase to the stock price of BRKB, a 0.1 percent of decrease to the stock price of XOM, a 0.02 percent of increase to JNJ and a 0.04 percent of increase to the stock price of FB. The effect of lagged index also is not in the same direction, for the lagged XOM index and JNJ index have a negative effect on the stock price, but the positive effects are existed in lagged AAPL, lagged BRKB and lagged FB.

Moreover, if we check the R square in the regression of stock price and google search volume index, we will discover that some of R square in table 2 are extremely small. For example, the R square for AAPL index is about 0.3 percent and R square for FB index is about 0.4 percent. They seem like extremely small if we compare with 13.1 percent for BRKB index and 11.5 percent for XOM index. The potential reason of extremely small in R square may cause by the broad range of google search volume index.

7 Conclusion

To summarize the result above, I found that google search volume index will have a positive effect on the stock trading volume and the relationship between google search volume index and stock price is unable to determine since the regression unable to determine whether there is positive

change or negative change on the stock price.

It is not hard for us to understand why we get this result. Since the main factor which rise the google search volume is the news. Whether the news is positive or negative, all news will attract investors attention and on the basis of these news, investors will decide either buy the stocks or sell the stocks. Both of the action will increase the trading volume for a stock. However, the stock price will dependent on whether the news will bring profit to a company or have a disadvantage for that company, and in my case, there is no way to define the increase in google search volume is cause by a positive news or a negative news. So, obviously, there is no way I can determine how search volume affect stock price.

References

- [1] Thomas J Chemmanur and An Yan. Advertising, attention, and stock returns. *Quarterly Journal of Finance*, page 1950009, 2019.
- [2] Hyunyoung Choi and Hal Varian. Predicting the present with google trends. *Economic Record*, 88:2–9, 2012.
- [3] JOSEPH Engelberg and Pengjie Gao. In search of attention. *The Journal of Finance*, 66(5):1461–1499, 2011.
- [4] Neri Kim, Katarína Lučivjanská, Peter Molnár, and Roviell Villa. Google searches and stock market activity: Evidence from norway. *Finance Research Letters*, 28:208–220, 2019.
- [5] Xianwei Liu and Qiang Ye. The different impacts of news-driven and self-initiated search volume on stock prices. *Information & Management*, 53(8):997–1005, 2016.
- [6] Xu Rui. Google search volume index: Predicting returns, volatility and trading volume of tech stocks. *working paper*, 32:1–32, 2015.
- [7] Dartanyon Shivers. Exploring the relationship between google trends data and stock price data. *Working paper*, 26:1–26, 2011.
- [8] Fumiko Takeda and Takumi Wakao. Google search intensity and its relationship with returns and trading volume of japanese stocks. *Pacific-Basin Finance Journal*, 27:1–18, 2014.

Company name	stock symbol	value in dollars
Apple	AAPL	926.9
Facebook	FB	541.5
Berkshire Hathaway	BRKB	491.9
ExxonMobil	XOM	344.1
Johnson and Johnson	JNJ	341.3

Table 1:

	<i>Dependent variable:</i>				
	logAAPL.V	logBRKB.V	logXOM.V	logJNJ.V	logFB.V
	(1)	(2)	(3)	(4)	(5)
AAPL.INDEX	0.003 (0.004)				
lag(AAPL.INDEX, 1)	0.013*** (0.004)				
BRKB.INDEX		0.002 (0.003)			
lag(BRKB.INDEX, 1)		0.007* (0.003)			
XOM.INDEX			0.003 (0.005)		
lag(XOM.INDEX, 1)			0.015*** (0.005)		
JNJ.INDEX				0.013*** (0.004)	
lag(JNJ.INDEX, 1)				0.017*** (0.004)	
FB.INDEX					0.002 (0.007)
lag(FB.INDEX, 1)					0.025*** (0.007)
Constant	16.727*** (0.197)	14.761*** (0.259)	15.309*** (0.382)	14.514*** (0.210)	15.531*** (0.401)
Observations	51	51	51	51	51
R ²	0.217	0.108	0.186	0.456	0.270
Adjusted R ²	0.185	0.071	0.152	0.434	0.239
Residual Std. Error (df = 48)	0.406	0.365	0.366	0.351	0.389
F Statistic (df = 2; 48)	6.663***	2.908*	5.477***	20.134***	8.860***

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 2:

	<i>Dependent variable:</i>				
	logAAPL.PRICE	logBRKB.PRICE	logXOM.PRICE	logJNJ.PRICE	logFB.PRICE
	(1)	(2)	(3)	(4)	(5)
AAPL.INDEX	−0.0002 (0.001)				
lag(AAPL.INDEX, 1)	0.0004 (0.001)				
BRKB.INDEX		0.0002 (0.0004)			
lag(BRKB.INDEX, 1)		0.001** (0.0004)			
XOM.INDEX			−0.001 (0.001)		
lag(XOM.INDEX, 1)			−0.001 (0.001)		
JNJ.INDEX				0.0002 (0.001)	
lag(JNJ.INDEX, 1)				−0.001 (0.001)	
FB.INDEX					0.0004 (0.002)
lag(FB.INDEX)					0.001 (0.002)
Constant	5.226*** (0.057)	5.238*** (0.031)	4.513*** (0.057)	4.905*** (0.034)	5.066*** (0.136)
Observations	51	51	51	51	51
R ²	0.003	0.131	0.115	0.016	0.004
Adjusted R ²	−0.038	0.095	0.078	−0.025	−0.037
Residual Std. Error (df = 48)	0.117	0.043	0.055	0.057	0.132
F Statistic (df = 2; 48)	0.075	3.633**	3.112*	0.398	0.101

Note:

*p<0.1; **p<0.05; ***p<0.01

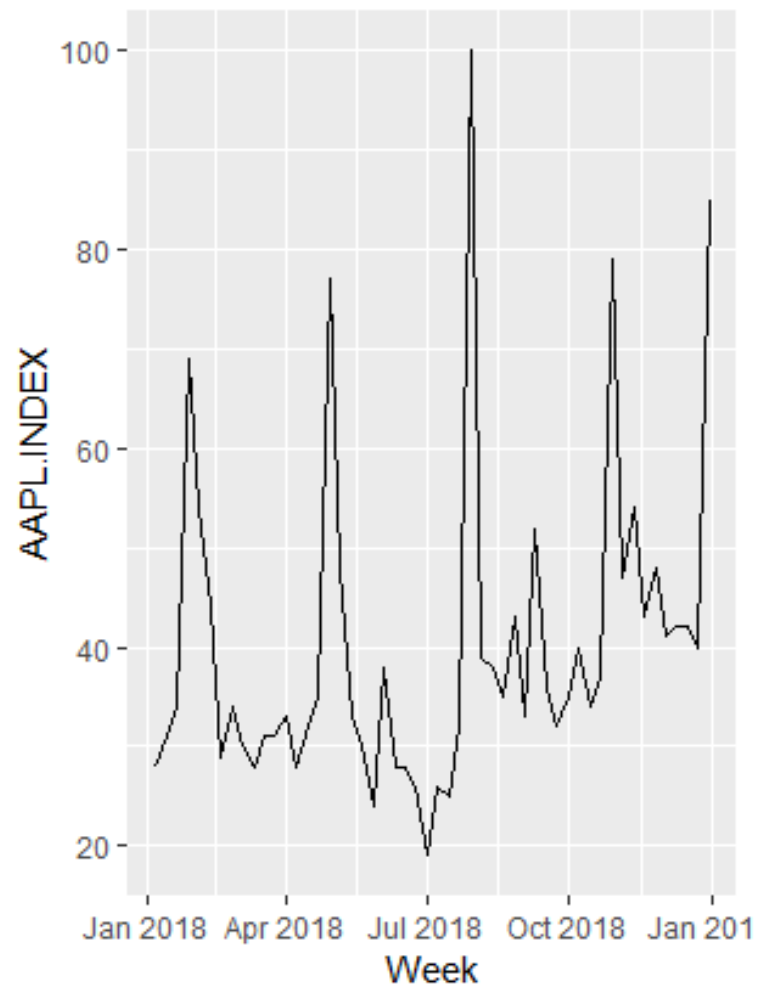


Figure 1: Search volume index of Apple

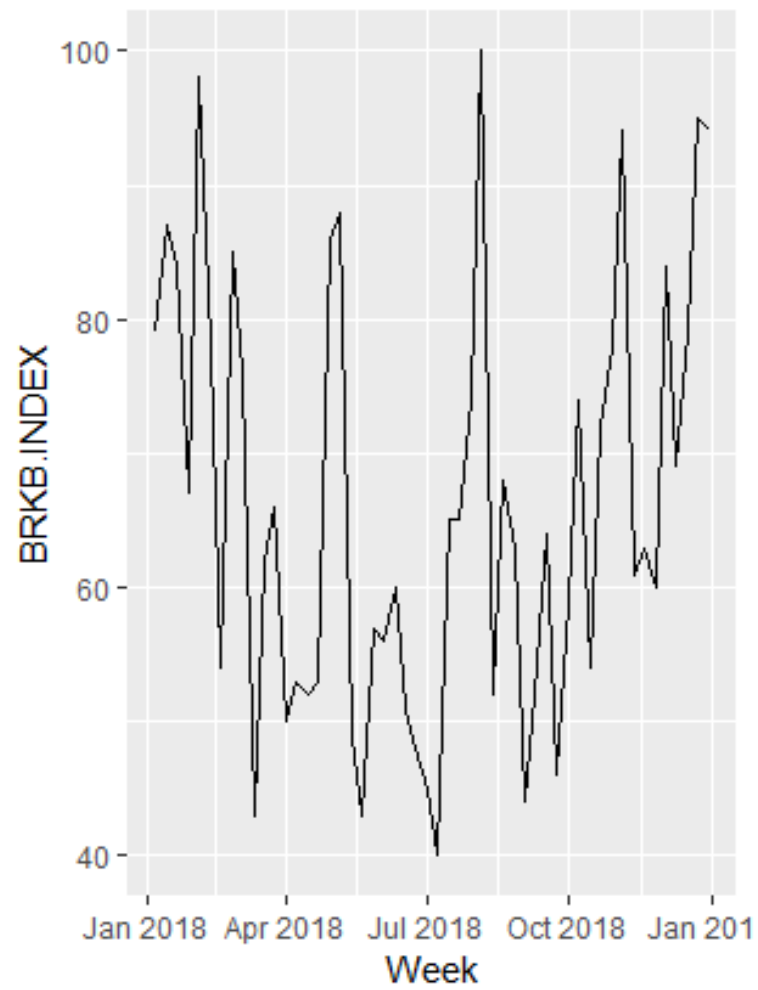


Figure 2: Search volume index of BRKB

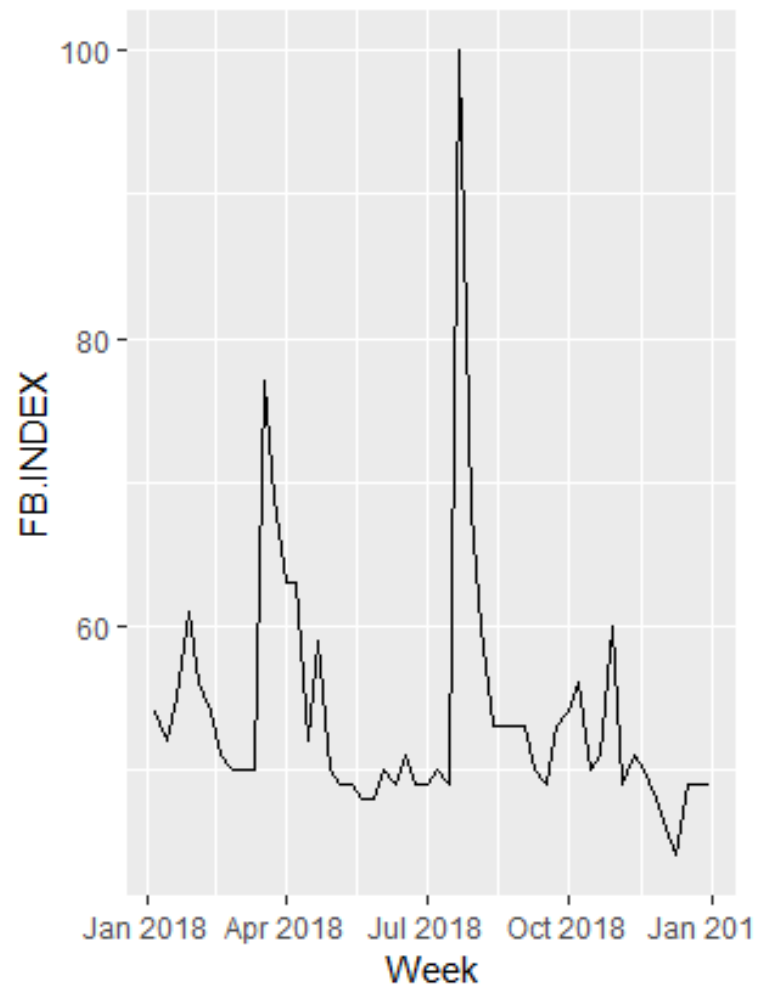


Figure 3: search volume index of FB

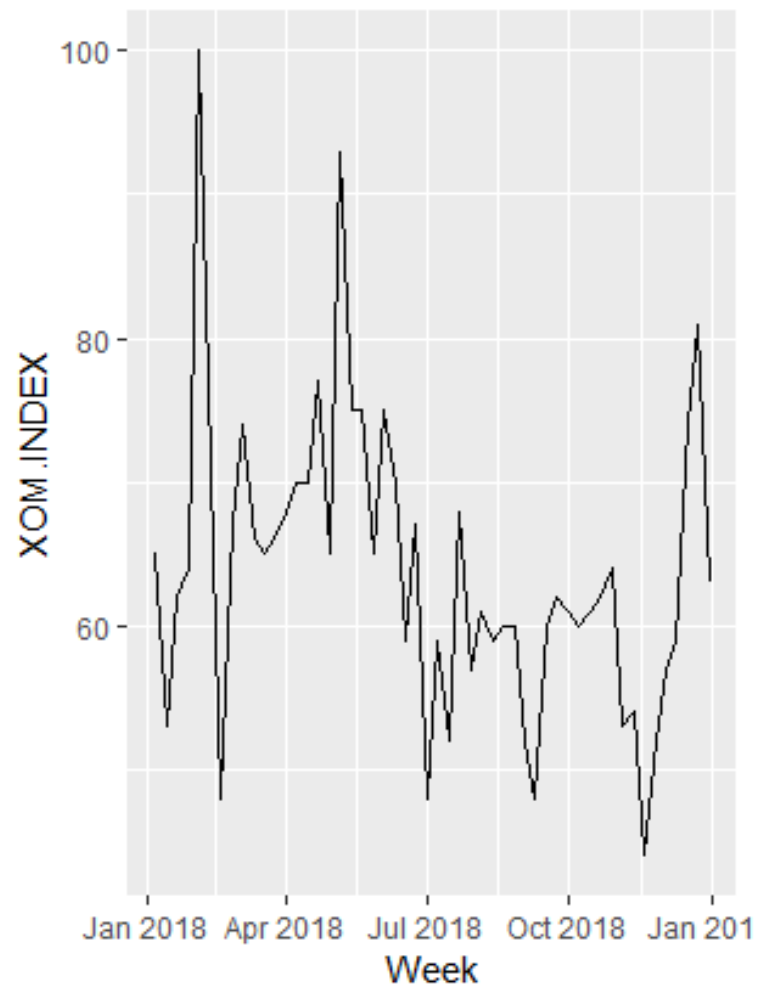


Figure 4: Search volume index of xom

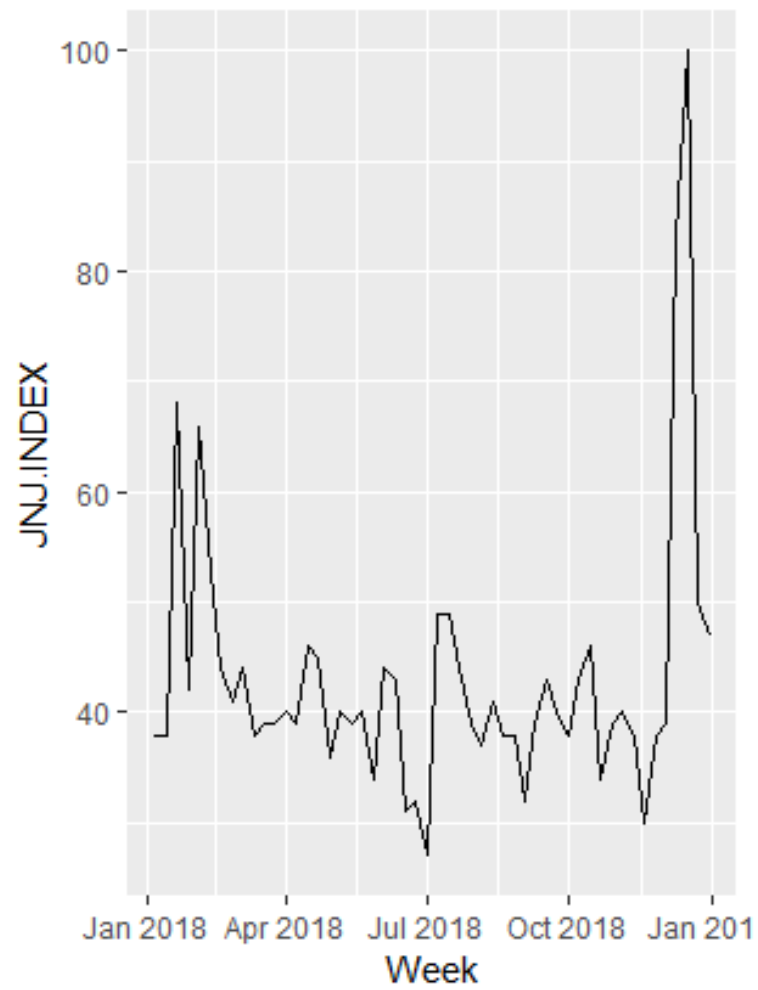


Figure 5: Search volume index of JNJ