

Oil's Effect on S&P 500 Prices

Colin Macy, Jeewaan Khadka, Kellan Meyer, Eric Lu

Introduction

With the S&P 500 and oil price representing the general economic trend, the prosperity of both indexes greatly affects the quality of our life. The relationship between oil prices and the S&P 500 index is important because oil prices are a key driver of inflation, impacting interest rates and the economy. Additionally, the relationship between oil prices and the S&P 500 index can provide insights into broader economic trends and market sentiment, making it a valuable tool for analysts and investors to monitor.

We look to examine the relationship between the S&P 500 and oil prices over time. With our analyzed data we will decide if there is a direct correlation between the two and whether or not it is possible to predict the price of the S&P 500 when given oil price data. Our analysis would be beneficial towards Wall Street analysts and investors in order to provide a possible edge in market timing. Furthermore, it could be useful towards college students as higher S&P 500 prices represent more job opportunities. Economic boom is beneficial for all of society and S&P 500 and crude oil prices represent the general success of the economy more than other indexes.

Thesis

Based on our analysis between crude oil and S&P 500 prices, we believe there will be a strong positive correlation between the two hence, if oil prices go up, S&P 500 prices should rise as well. According to our findings, S&P 500 price is moderately predictable when given crude oil prices.

Questions of interest

We will be using two datasets, one of which shows the S&P 500 stock market share prices and another that shows Oil prices over time. We are looking to predict growth and its correlation with the S&P 500. Using this information we will predict S&P 500 data and prices based on current/future oil prices.

Background/ Data

The Standard' and Poor(S&P) 500 index includes 500 leading companies in the leading economies of the US Market. This is regarded as the gauge of the US Equity Market and it contains 75% of the US equities market. The S&P 500 index is typically interpreted as a measure of the overall health and performance of the U.S. stock market. As the index includes 500 large publicly traded companies from a diverse range of industries and sectors, it provides a broad and representative snapshot of the U.S. economy. We collected the data of S&P 500 from FRED which had an agreement with S&P Dow Jones Indices LLC. Our data of S&P 500 is a price index not a return index as it does not contain the dividends. The observations in the index data from FRED has daily index value at market close. We took that index value and calculated the monthly average index value of S&P 500. The level of the S&P 500 index can also be used to gauge investor sentiment and market trends. For example, a rising S&P 500 is generally seen as an indicator of positive investor sentiment and economic growth, while a falling S&P 500 is often interpreted as a sign of negative sentiment and potential economic contraction. S&P 500 doesn't have a unit.

For Crude oil prices, the data set contains the first oil purchase price for each month. We collect the monthly oil price data from U.S. energy information administration(eia), who have calculated the monthly oil price for us. The units of measurement are dollars per barrel. On the x-axis is time since 1975. "Dollars per barrel" (or "\$/bbl") is a unit of measurement used to express the price of crude oil. It represents the cost, in U.S. dollars, of a single barrel of oil. The price of crude oil is typically expressed in terms of dollars per barrel because this allows for easy comparison between different grades of oil and different market conditions. The price of oil can fluctuate widely depending on a variety of factors, including global supply and demand, geopolitical tensions, and economic conditions. The dollar per barrel price of oil is a key determinant of gasoline prices and can have a significant impact on the broader economy.

Crude Oil has been an important commodity in the international and domestic market affecting most of the sectors including the stock market. When the price of crude oil rises, the cost of goods sold for most businesses tends to go up. We asked if there is any correlation between the price of crude oil and the S&P 500 index. If so, how strong of a relationship it is and can we predict the index value given the price of crude oil at any time.

Our data has the monthly price of crude oil and the averages of daily closing values per month of the S&P 500 index. This is related to our main question of finding the correlation between S&P 500 and crude oil. There is couple of NA in the data. However, we dropped them during our data cleaning process.

Some unusual factors which may affect the interpretation of the results include the following the Federal reserve's action on adjusting the economical period. Also, the time frame we took contains few extreme events creating more buzz in the prices. For example, the effects of the Ukraine War on oil prices or the effects of Covid-19 in 2020. Few other events that might have influenced the data are New Energy policy published in targeted countries and New discovery of the oil field and the improvement of renewable energy during that time frame.

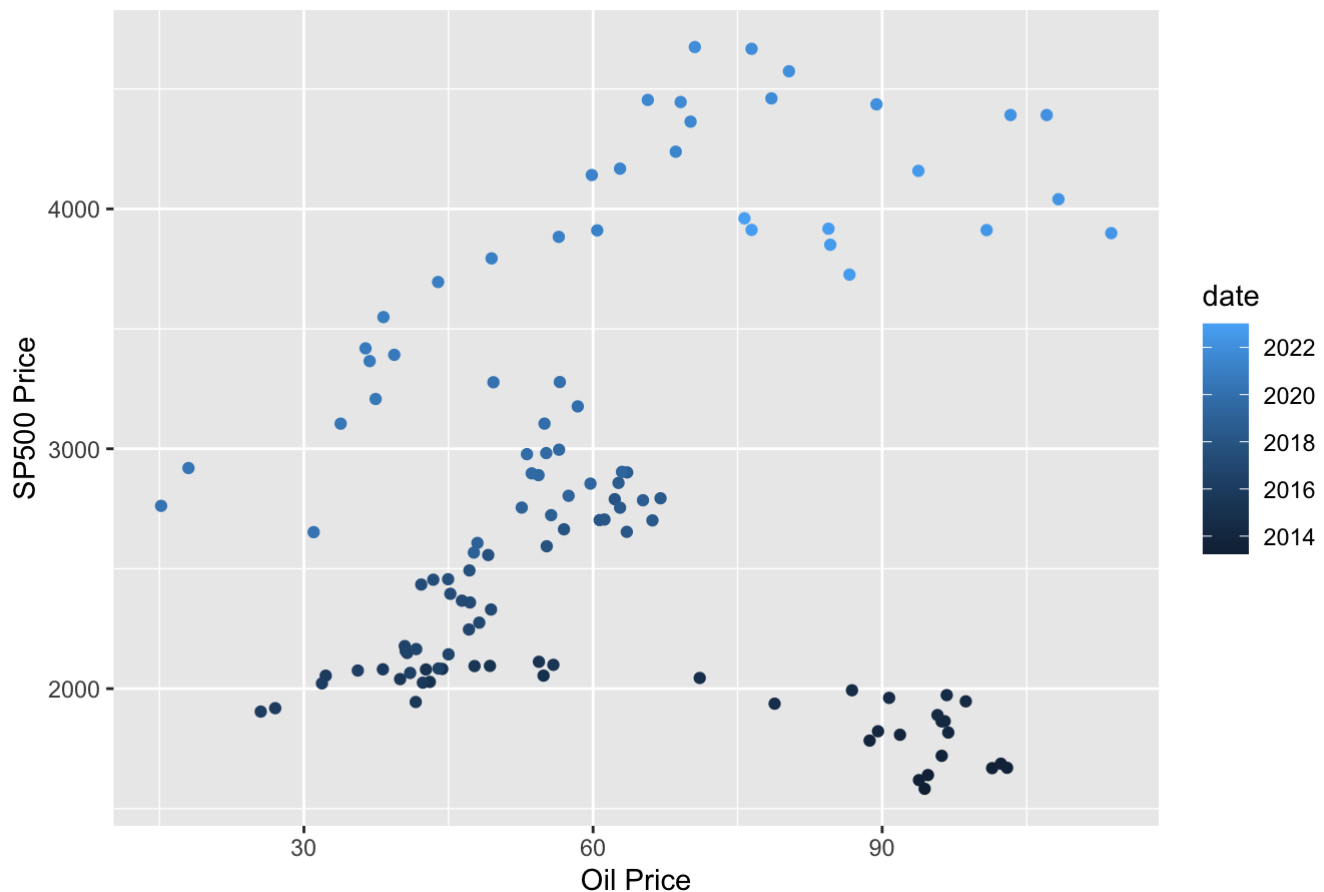
We intend to use our datasets for a regression analysis in order to explore the correlation between S&P 500 and oil price. Our model will allow us to explore the correlation and predict various S&P 500 prices based on given oil prices.

Graphs/ Analysis

Scatterplot Graph

We first look to plot our Oil Price and S&P500 price together. This will allow us to examine if there is a general trend between the two and get an idea for the datasets we are working with.

Oil Price vs S&P 500 Over Time

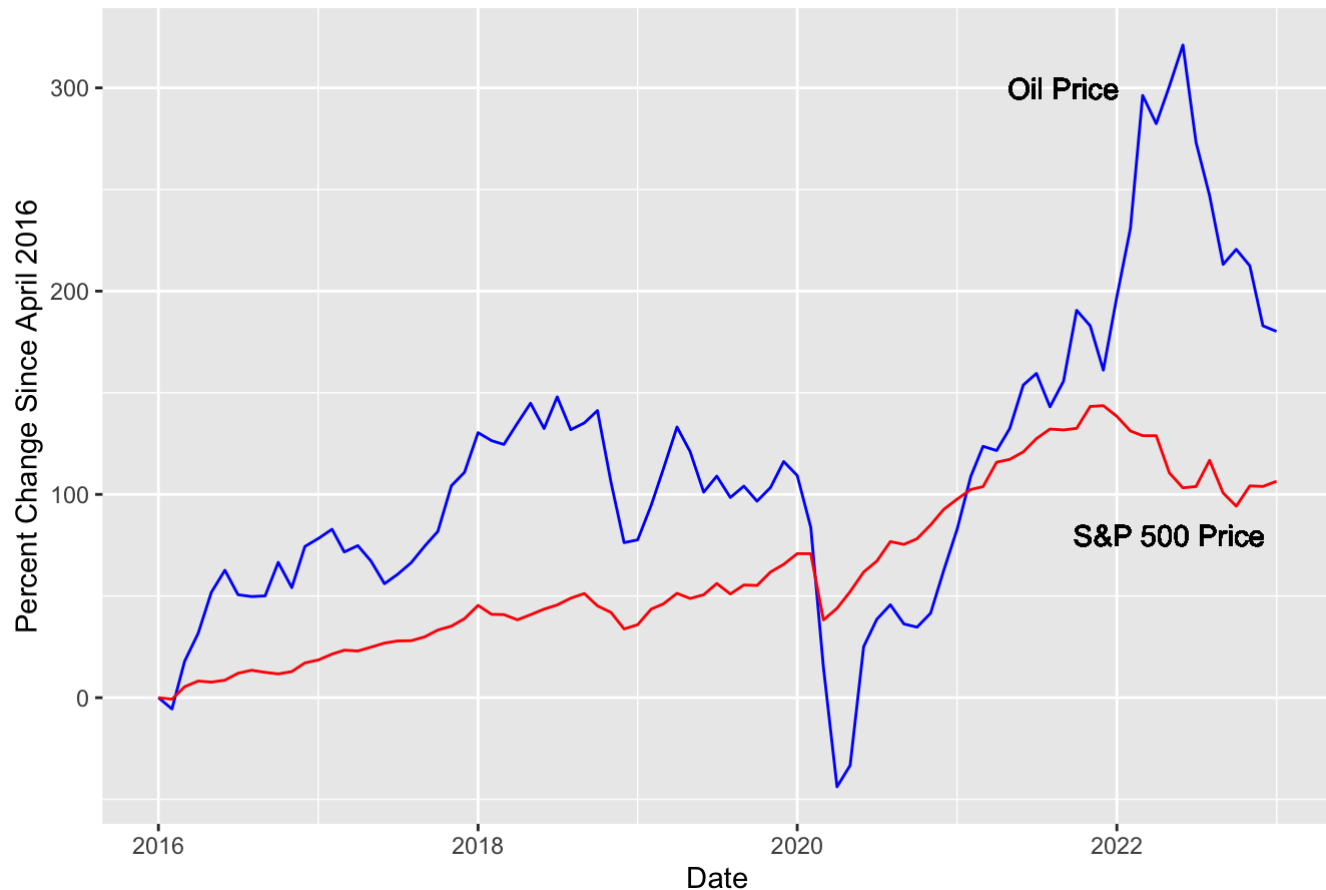


This graph shows the relationship between oil prices and S&P 500 Prices over time. The color scale is done in terms of years with the lightest colors representing more recent data. Generally, as the oil price increased over time, the monthly average value of S&P 500 increased as well. This may tell us there is a correlation between the two, but we cannot confirm this quite yet. We notice in the bottom right corner of our plot a large cluster that does not represent the rest of the trend throughout our data. After doing some deeper research, we noticed that from 2014 to the start of 2016 there was a drastic oil plunge that was due to a large spike in US oil production and a large demand plummet in China and Russia. Based on this information we decided to remove all data points up until 2016 to get a more accurate model later on.

Oil/S&P 500 Time Graph

Next, we want to examine the general price of oil and S&P500 over time to see if there are any trends and/or changes that stand out. We can calculate the percent change since the starting point for each data set to examine how the prices have changed from 2016 to 2023. This model allows us to see the two prices on the same scale so we can see how the prices affect one another relatively.

Percent Change in Oil Prices and S&P 500 Since April 2016



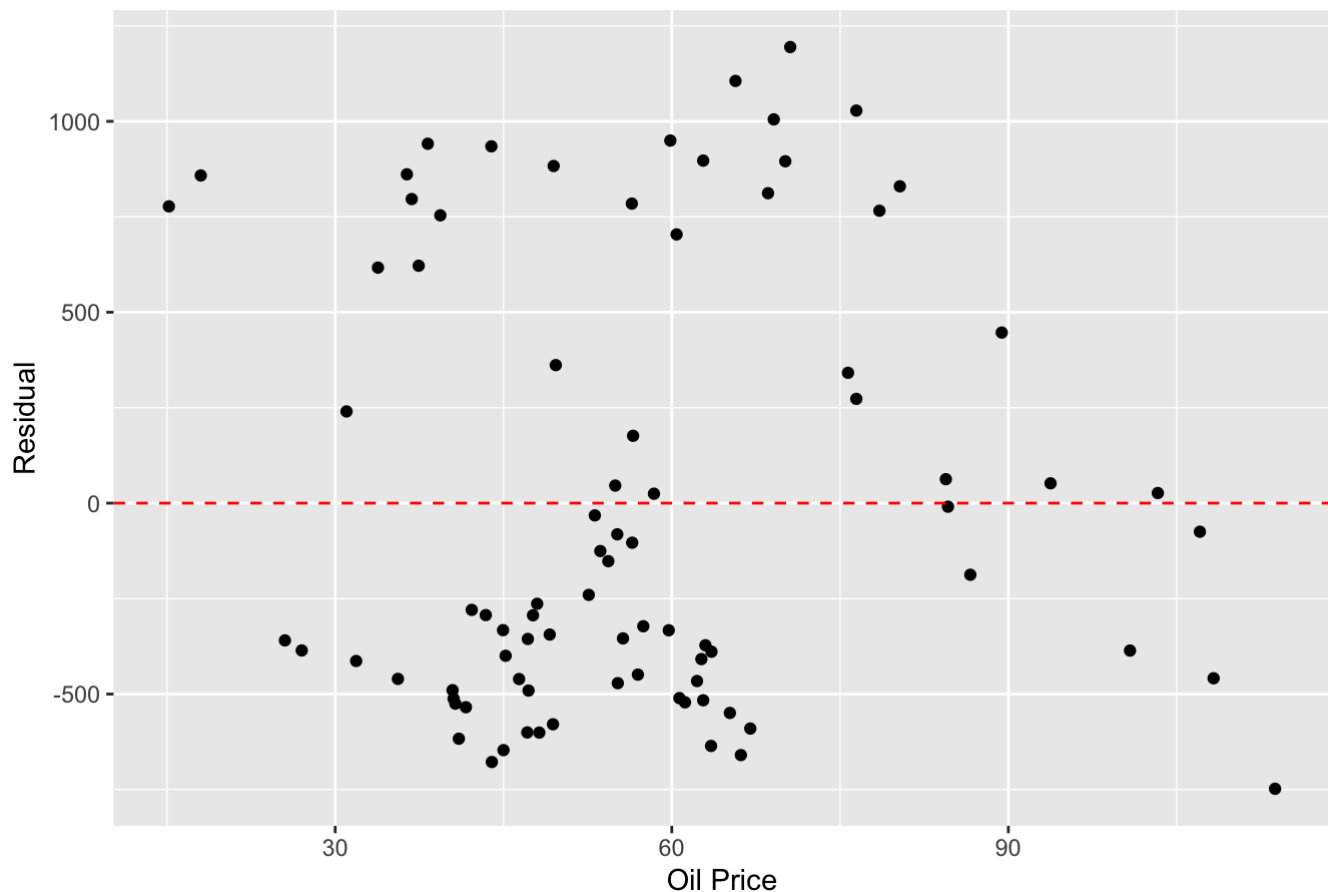
From the graph we can see that there was a clear drop in price around 2014 which, after further research, was due to a drastic increase in the supply of oil. This continued throughout 2016 where it leveled out shortly after. After a short increase in oil prices, 2020 saw another drastic fall off in price due to COVID-19 followed by an extreme increase in oil prices due to the Russian/Ukrainian war since Russia was a big supplier. These drastic changes in oil prices are historically significant and hence they may influence our model leading to higher variance.

Over time our graph shows a consistent, fairly linear growth in terms of S&P 500 prices. The only noticeable drop was in 2020 due to the effects of COVID-19. There were small margins of drop in price which was extremely similar to when oil prices dropped. Besides that, the price has continued to increase since it is not very reactive to social and political inconsistency.

Residual Plot

Before building our model, the first thing we want to do is create a residual plot that shows what the appropriate model choice would be. We look to examine the distribution of the data points in order to confirm the acceptable model for the predictions of S&P 500 prices.

Residual vs Oil Prices



The residual plot has no clear pattern, as there is no clustering and the points are scattered randomly throughout the graph. This suggests that it would be appropriate to use a simple linear regression model to predict monthly S&P 500 prices from oil price.

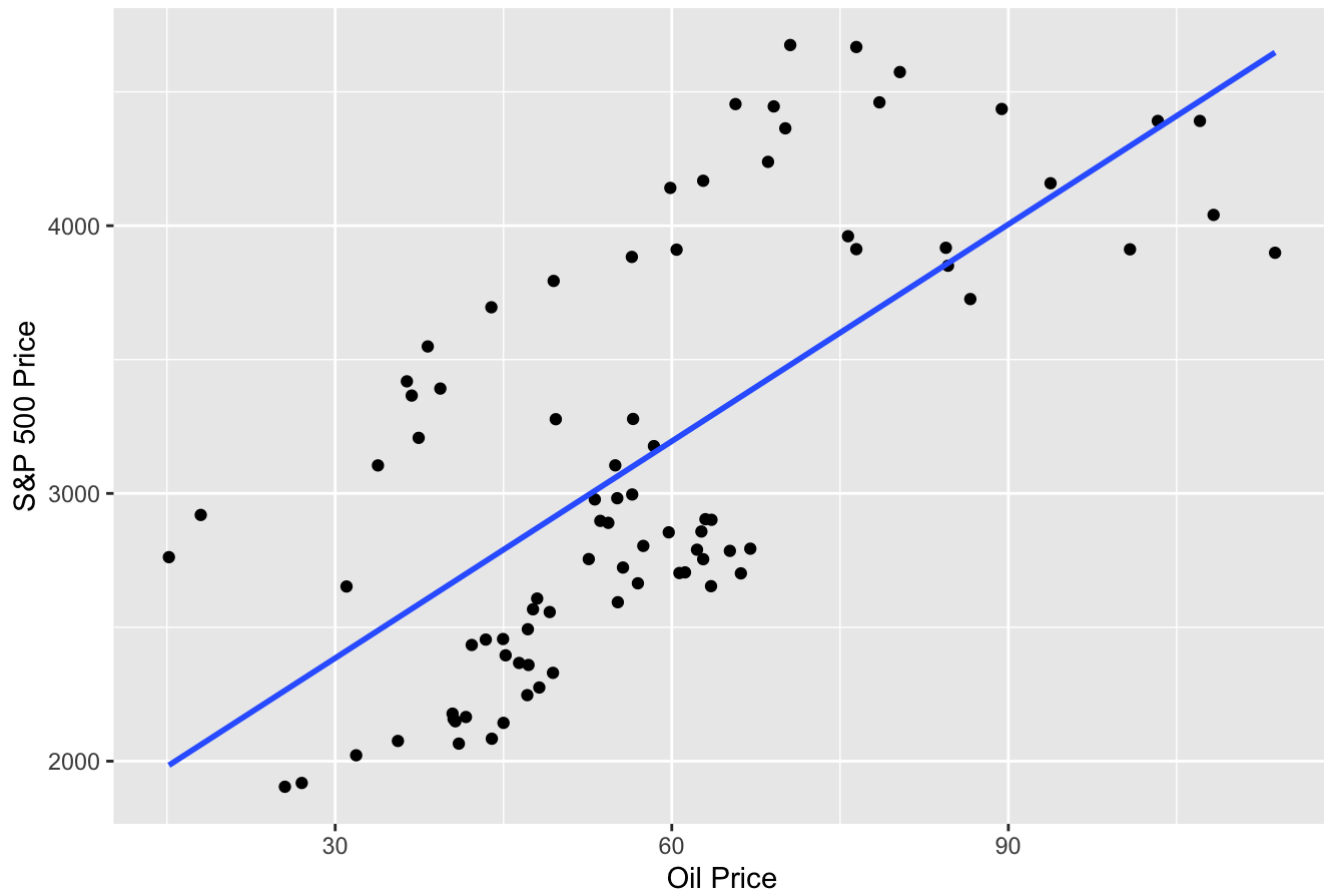
Regression Model

Finally we can build our regression model that will give us a prediction of S&P 500 prices based on each oil price point. The data below plots each of the points since 2016 along with the new predicted points labeled with a blue linear regression line.

Since we were looking to see if S&P 500 prices were heavily correlated with oil price, our independent variable, we wanted to use a linear regression model as it can allow us to predict our responsive variable, S&P 500, based on our predictor, oil price.

Our assumption is that oil prices are independent of S&P 500 and totally derived from pure market demand and supply. Another assumption we made was that the S&P may be influenced by oil prices and external factors. Considering these two assumptions, it is acceptable to use a regression model.

S&P 500 Price vs Oil Price



Using a linear regression plot, our data shows a positive linear trend between the oil price and S&P 500 price. The slope is about \$27 which shows that as the price of crude oil increases by a dollar, S&P 500 prices jump by about \$27. The y-intercept sits at \$1575 which represents the S&P 500 price when the oil price is \$0. Now that we can see our two prices are linearly correlated, we examine the strength of this correlation.

Our correlation coefficient is 0.68 which shows a moderately strong positive correlation between our two explored variables. Due to the strength of our correlation coefficient, it is a safe assumption to say oil price does in fact have an influence the S&P 500 price.

Based on this data we were able to calculate the confidence intervals for both the y-intercept and the slope of the predictive linear regression line. We are 95% confident that the true slope is between \$20.7 and \$33.3, and we are 95% confident that the true y-intercept is between \$1194 and \$1956.

Interpretation

The first thing our data makes clear is that there is a relationship between oil price and S&P 500 price. We observed a 0.68 correlation coefficient which shows the correlation is positive and moderately strong, and the S&P 500 index is influenced by oil prices. However, this model does not allow us to account for external effects of the economy. There is a correlation between oil price and the S&P 500, but it is impossible to confirm that it is not just general economic success that brings both the prices up.

Our linear regression model makes sense because oil is a major indicator of the global economy. Oil can be used in different approaches from jet fuel to plastic products. The scarcity and functionality of oil make it a huge parameter for the modern economy. As a result, movements in the price of oil can provide insights into the overall health of the economy and the stock market. If the price of oil rises, it may indicate strong global demand and

economic growth, which could boost profits for many companies in the S&P 500. Conversely, if the price of oil is falling, it may signal weaker demand and slower growth, which could weigh on the earnings of many S&P 500 companies.

Oil prices are easily affected by the international political situation whereas the S&P 500 are not significantly affected, which is a potential shortcoming of this analysis. However, the S&P 500 is easily affected by the United State government like the Fed, while oil prices receive little impact from the adjustment policy inside the United States. As a result, one shortcoming could be that the unstable international situation would greatly affect the outcome and the interpretability of our result.

For future directions for additional work, one question we might ask is, "Can we build a model that can adjust to the effect of international instability?" With sufficient knowledge, we can potentially use machine learning to build stronger predictions. This approach involves training algorithms on historical data to identify patterns and predict future price movements based on a range of variables.

If there is any other indicator could help us understand more about the economy, that could be GDP, OPEC oil production data, gold price, other major stock indicator for U.S stock market. Because these indicators are a very good estimate and measurement of the global economy status, Thus, these indicators like GDP for the United States would be a very good substitution for our analysis.

Summary

From our analysis, we concluded that there is a positive, moderately strong correlation between S&P 500 prices and crude oil prices. When we removed the 2013-2015 price plunge we could observe a more consistent linear trend upwards throughout the data. Even with this, we noticed S&P 500 data rarely drops drastically regardless of oil price volatility. Regardless, it is worth taking a deeper dive into exploring with more advanced machine learning tools as economic stability is extremely beneficial to all US citizens, as it provides for financial stability and job security.

Based on our linear regression model created, we can be 95% confident the slope for S&P 500 price vs Oil Price is within the interval (20.8, 33.3) and the y-intercept is within the interval (1194, 1956). Based on this data we believe we can predict S&P 500 prices to some extent when given crude oil prices.

References

<https://fred.stlouisfed.org/series/SP500> (<https://fred.stlouisfed.org/series/SP500>)

https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=pet&s=f000000__3&f=m
(https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=pet&s=f000000__3&f=m)

<https://blogs.worldbank.org/developmenttalk/what-triggered-oil-price-plunge-2014-2016-and-why-it-failed-deliver-economic-impetus-eight-charts> (<https://blogs.worldbank.org/developmenttalk/what-triggered-oil-price-plunge-2014-2016-and-why-it-failed-deliver-economic-impetus-eight-charts>)

<https://www.bls.gov/opub/btn/volume-4/pdf/the-2014-plunge-in-import-petroleum-prices-what-happened.pdf>
(<https://www.bls.gov/opub/btn/volume-4/pdf/the-2014-plunge-in-import-petroleum-prices-what-happened.pdf>)