Cryptocurrencies

Carlos Echeverri

Objective

Much like internet in the 2000's, cryptocurrencies have exploded in popularity these last few years. Everything started circa 2009 with the most popular cryptocurrency of all, Satoshi Nakamoto's Bitcoin. Nowadays, we there are over a thousand coins to choose from and the influx of new coins is not likely to slow down anytime soon as people is getting to know about blockchain and these speculative assets become mainstream. The objective of this project is to gather all available data for the coins of our interest, clean it and analyse it with the purpose of seeing patterns that may help us have a better understanding of the behaviour of these assets. The maing aspects that we want to know about these coins are: Liquidity, Correlation and Performance.

Liquidity: What is the purpose of holding an asset that you cannot turn back into fiat currency, which you can spend at any place? It would be interesting to know if increase in liquidity makes a coin more desirable, or even less desirable.

Correlation: We need to know if the coin that we are interested in moves in the same way as other assets, like another cryptocurrency or fiat, perhaps.

Performance: Volatility is extremely important and so is the expected return of the asset that we are interested in. We need to evaluate how the price performs in a year, or within a month. How far away the top and bottom prices are, etc.

Getting Started

We begin by loading the required libraries for our analysis. Some libraries that we may have not encountered before are the ones used after tidyverse. This is what each of them does:
1) jsonlite - We use it to parse JSON responses into a data frame.
2) httr - It gives us the functionality to get a JSON from an API.
3) htmltab - Scrapes the source code of a website and converts the specified table into a data frame that we can use in R.
4) Quandl - Provides us with access to the Quandl database that has the forex data.
5) padr - Provides us with a function to add missing dates to time series.

This is the definition of some formulas that we will use in our analysis.

```
#### Formulas:
pct <- function(x) {(x/lead(x)) - 1}
value <- function(buy, sell, inv) {(inv * sell / buy)}
```

Connect to the coinmarketcap API and get the list of the top 10 currencies by market capitalization. Convert numbers to millions to avoid running into problems with R due to their length.

```
#### API: Top 10 Coins Load, Convert Market Cap to Millions
api <- GET("https://api.coinmarketcap.com/v1/ticker/?limit=10")
top <- as.tbl(fromJSON(content(api, "text", encoding = "UTF-8")))
top <- top %>%
  rename(mktcap = market_cap_usd) %>%
  mutate(mktcap = sub("*\\.[0-9]", "", mktcap)) %>%
  mutate(mktcap = substr(mktcap, 1, nchar(mktcap) - 3)) %>%
  mutate(mktcap = as.double(mktcap) / 1000000)
```

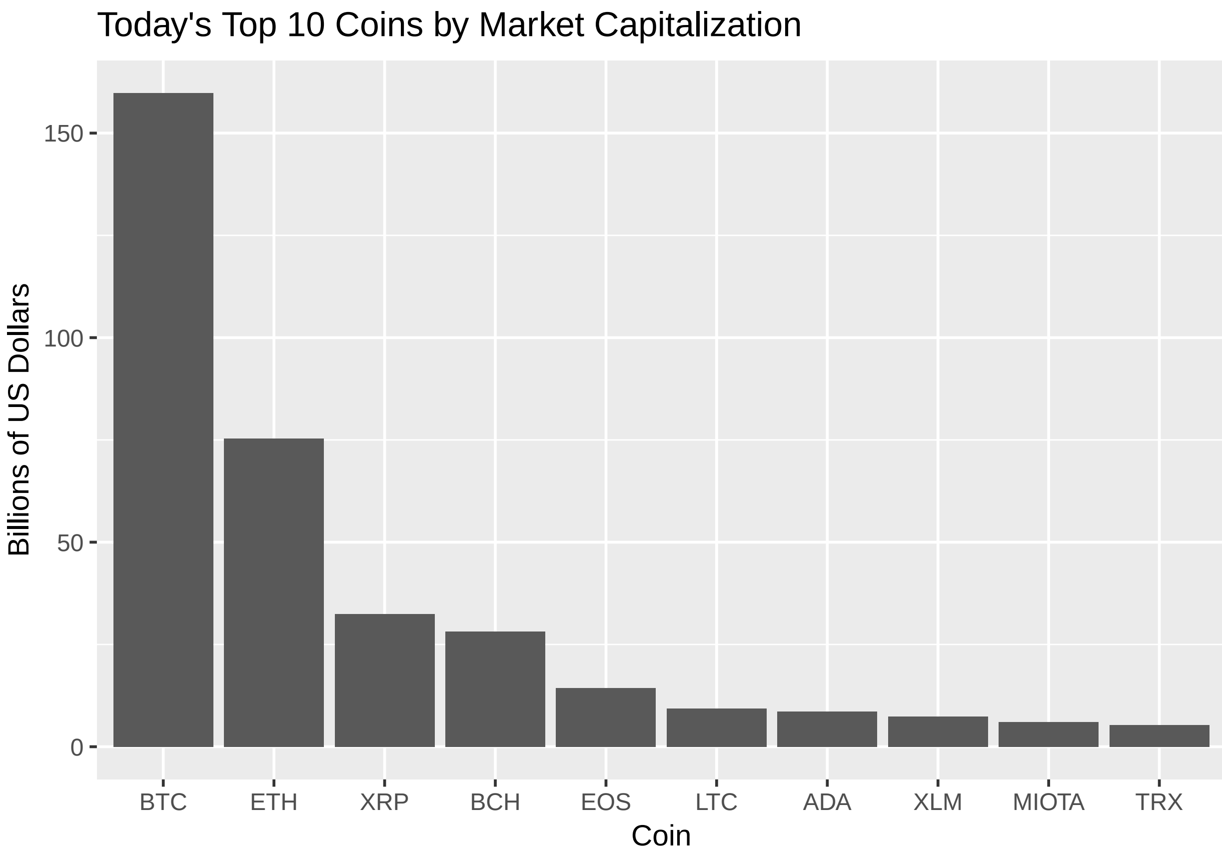
Use a loop to scrape the website for each of the top 10 currencies in coinmarketcap.com and make calculations to add a Percentage Change column, bind all currencies in a single tibble.

```
#### Load Crypto Data
url1 <- "https://coinmarketcap.com/currencies/"
url2 <- "/historical-data/?start=20130428&end="
today <- str_replace_all(Sys.Date(), "-", "")

N <- 10

for (i in 1:N){
  url <- str_c(url1, top$id[i], url2, today, sep = "")
  temp <- htmltab(doc = url, which = 1)
  temp <- temp %>%
    mutate(Close = as.double(Close)) %>%
    mutate(PctChange = pct(Close) * 100) %>%
    mutate(Coin = top$symbol[i])

  if (i == 1) {
    data <- temp
  } else {
    data <- rbind.data.frame(data, temp)
  }

  if (i == N) {
    rm(temp, url1, url2, today)
  }
}
```

Load some popular stocks and forex data and convert to the same format as the cryptocurrencies. For this purpose, we will use Iextrading's API to pull the historical data of some popular blue chip companies' stocks such as: APPL, TSLA and GOOGL, and we will also include some lesser known stocks that offer higher volatility such as: LMT, CEIX and NKTR. To get the data of forex currencies we will use Quandl's database. We will focus on major currencies such as the Swiss Franc, the Euro, the British Pound Sterling, etc.

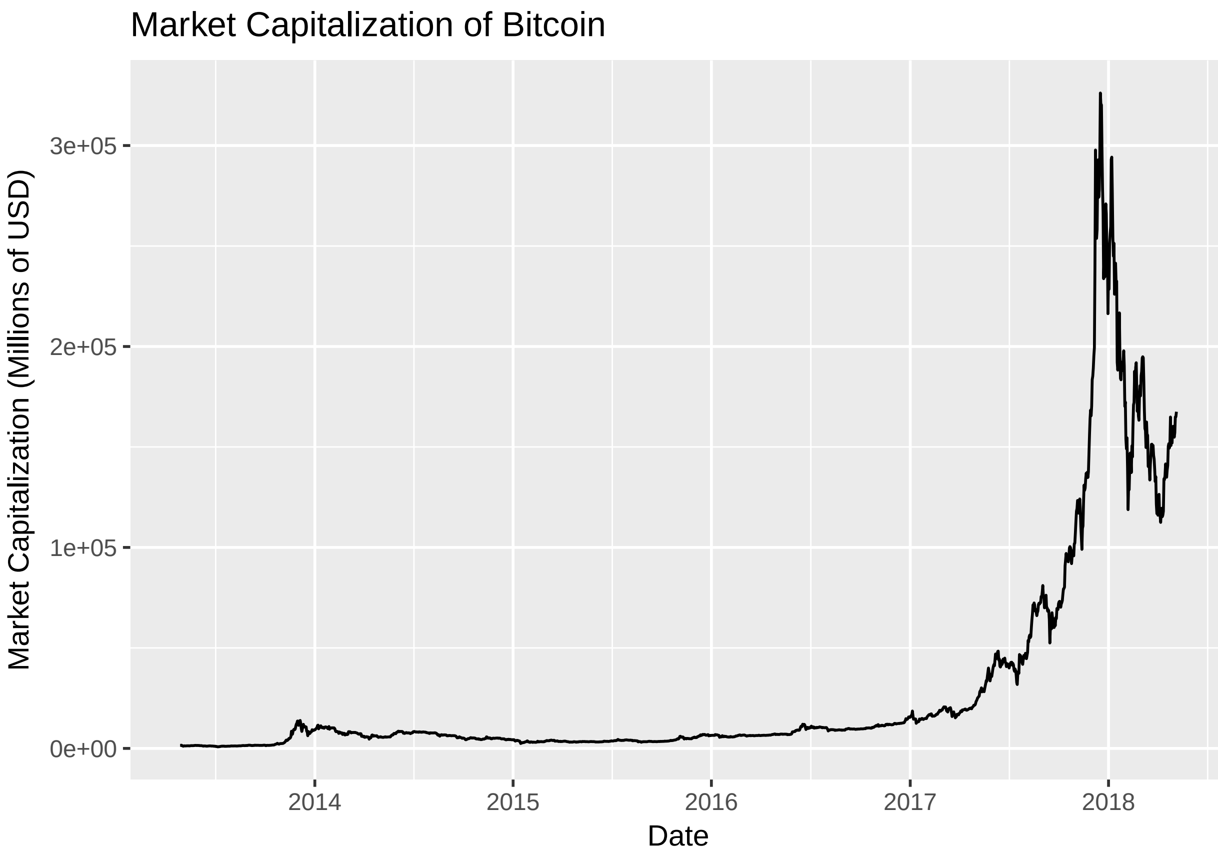
Liquidity

Use the market cap value of each of the top 10 coins to plot them in descending order. We can see how BTC greatly outperforms the rest of the coins in the top 10.

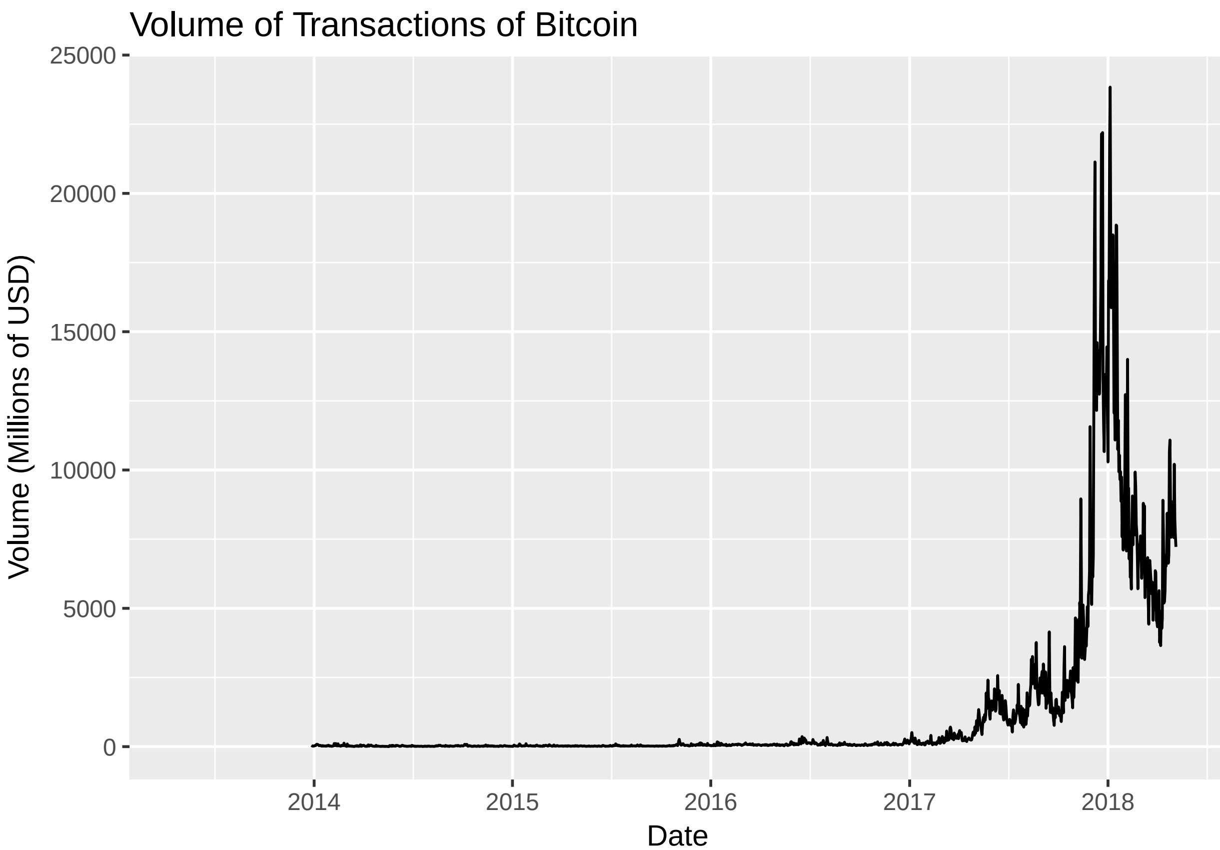


Most of the rest of the functions take as input one or two coins, in the Shiny app there will be a dropdown menu allowing the user to choose which coin to analyse; however, we will focus mostly on the most popular coins: Bitcoin (BTC) and Ether (ETH).

Filter the data by the selected coin and use the MktCap column as an aesthetic to zoom into any coin's market capitalization.

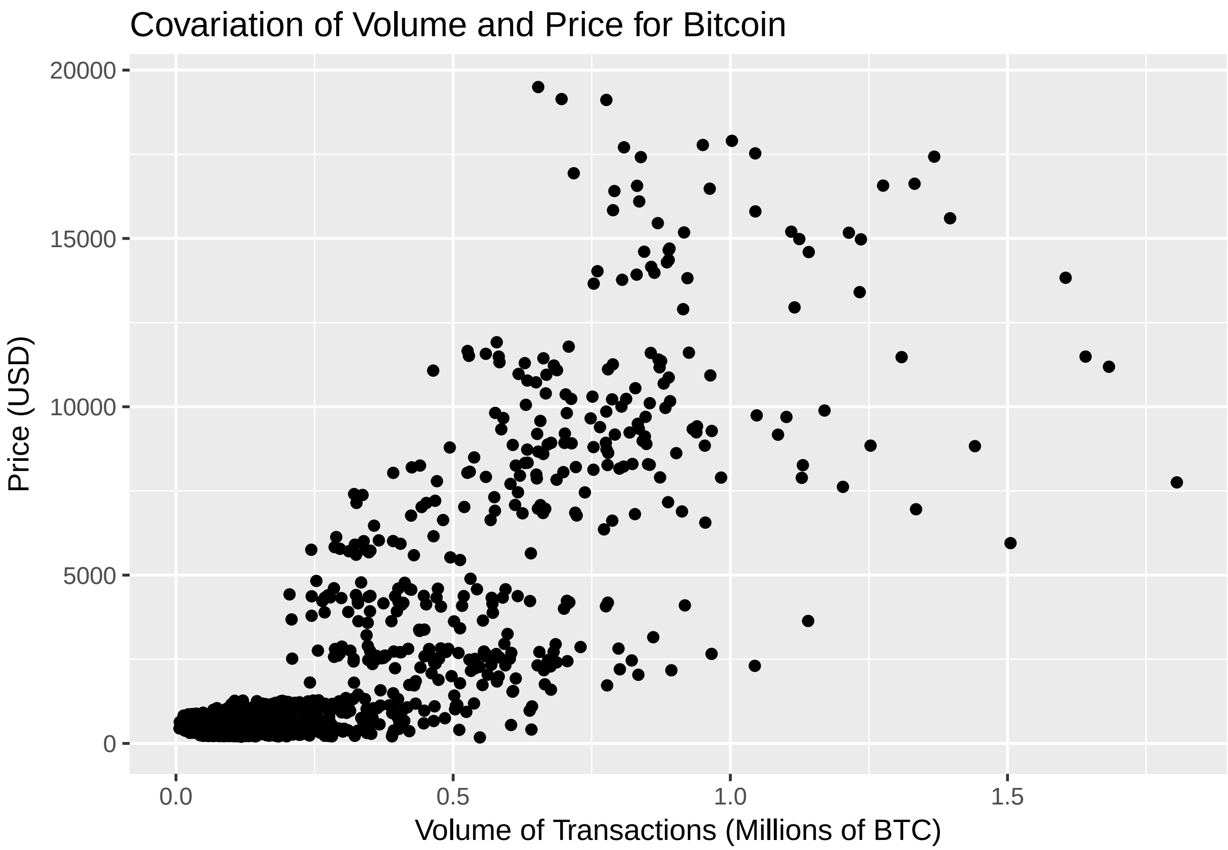


Filter the data by the selected coin and use the Volume column as an aesthetic to zoom into any coin's volume of transactions. This helps us see more clearly how Bitcoin becomes more popular as an investment vehicle as people are making more trades with it.

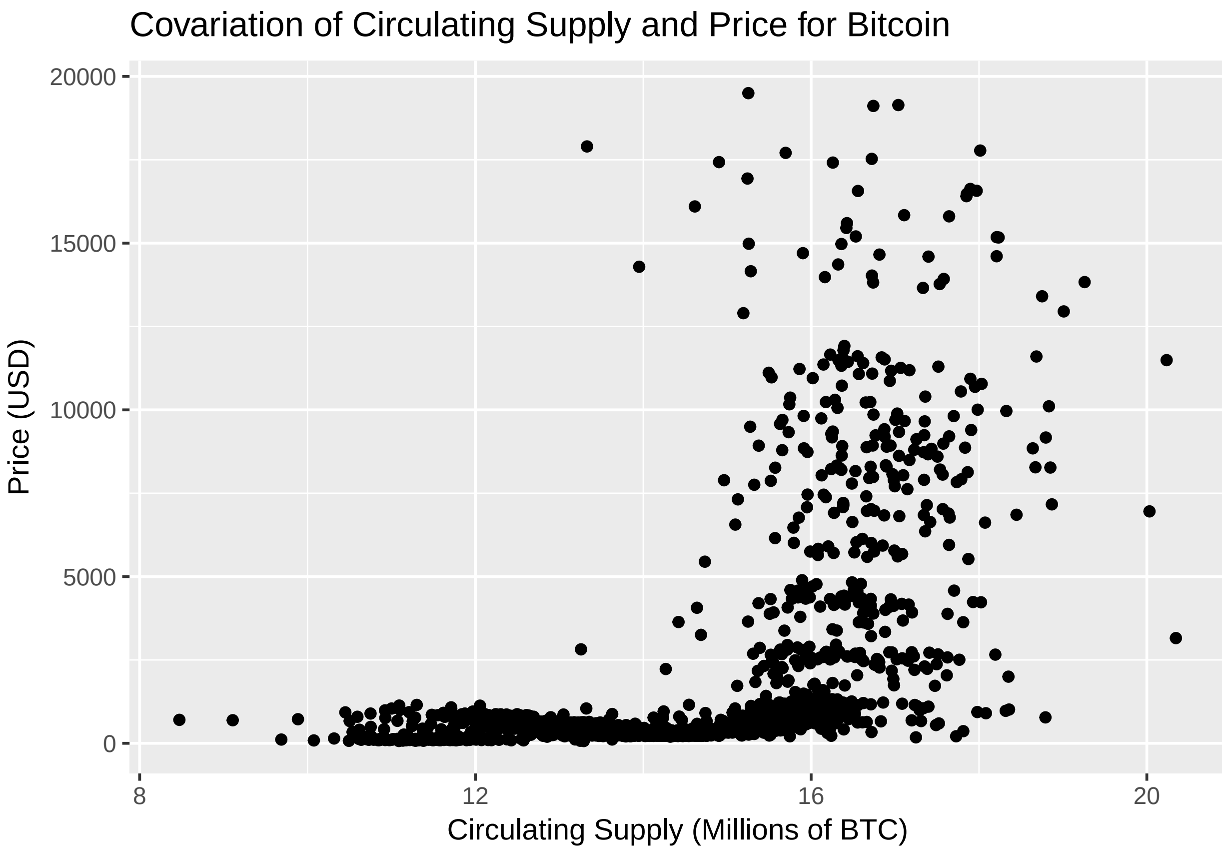


Correlation

We plot the volume of transactions (denominated in BTC) against the Price of the coin to see if there is any kind of relation between these two factors. The scatterplot does not gives us a conclusive answer because at first prices seem to be increasing with volume, but in the higher ends of volume we can observe low prices as well.

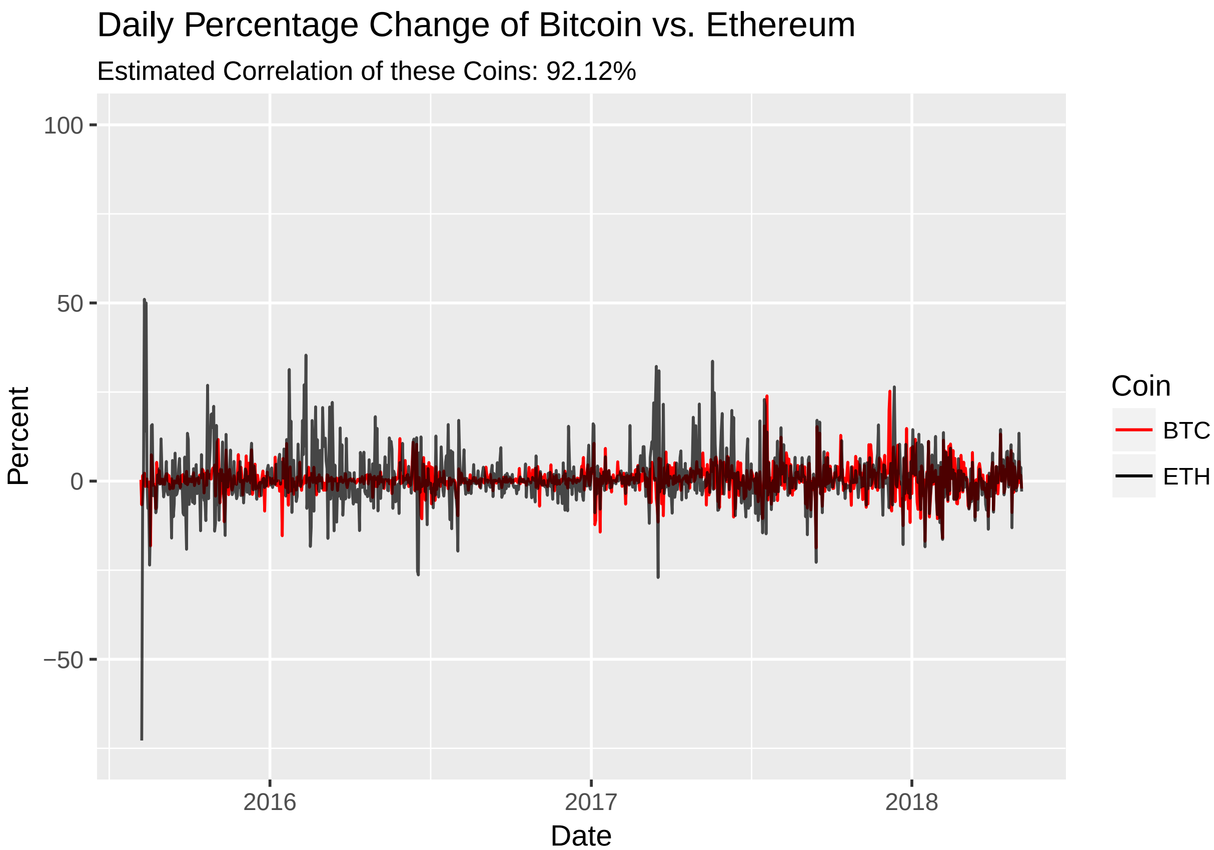


We plot the circulating supply (denominated in BTC) against the Price of the coin to see if there is any kind of relation between these two factors. We cannot conclude either that there is a linear relationship between the factors because before the 16 million BTC supply mark, the price of bitcoin was practically flat and explodes after this point.

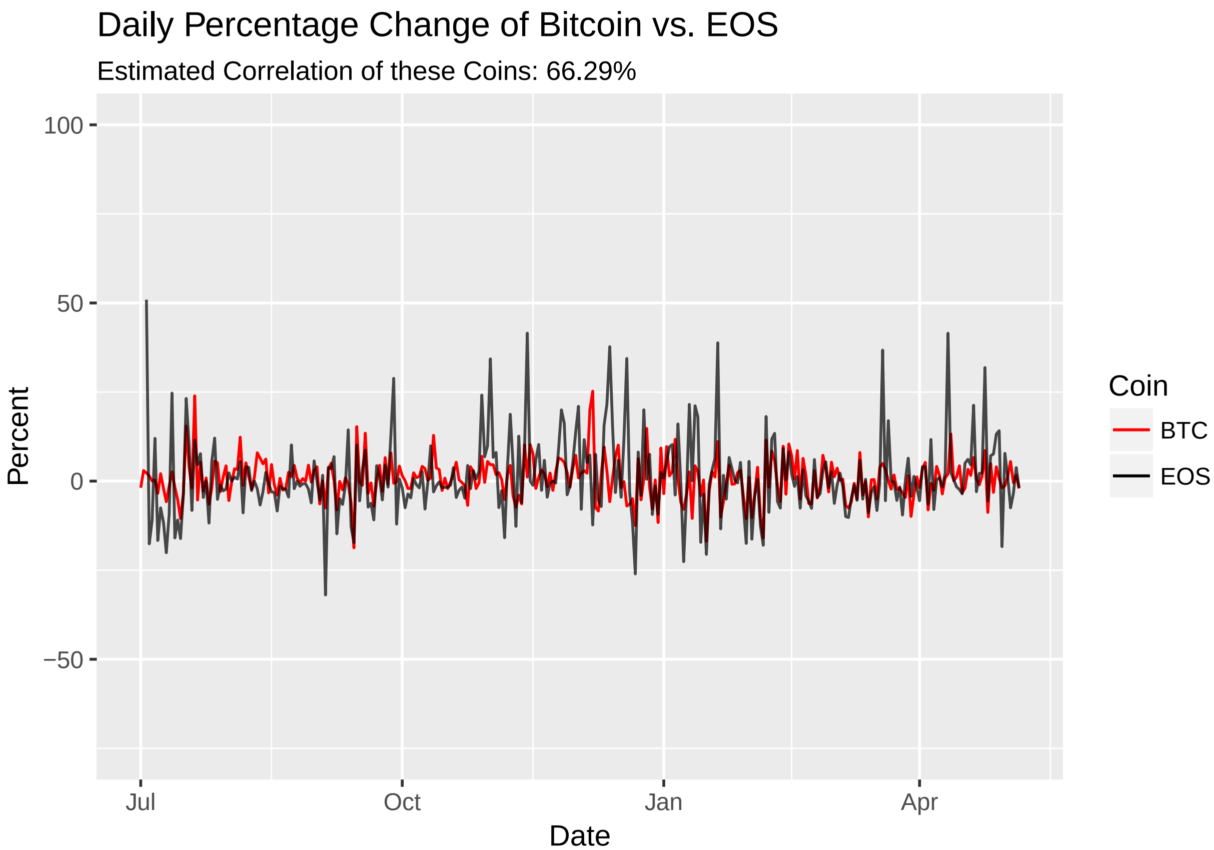


Furthermore, we want to explore the relationship between two coins. A common belief is that Bitcoin leads the way in the crypto market and altcoins follow. We standardize our measure by using percentage changes instead of price changes as it is easier to compare visually. We use R's built-in correlation function to evaluate the daily percentage changes of two coins of our choice. Since not all currencies have been around for as long as bitcoin, we introduce a *min_date* variable that corresponds to the maximum of the first coin's first day of data and the second coin's first day of data, we use this variable to evaluate only the period on which both coins have co-existed.

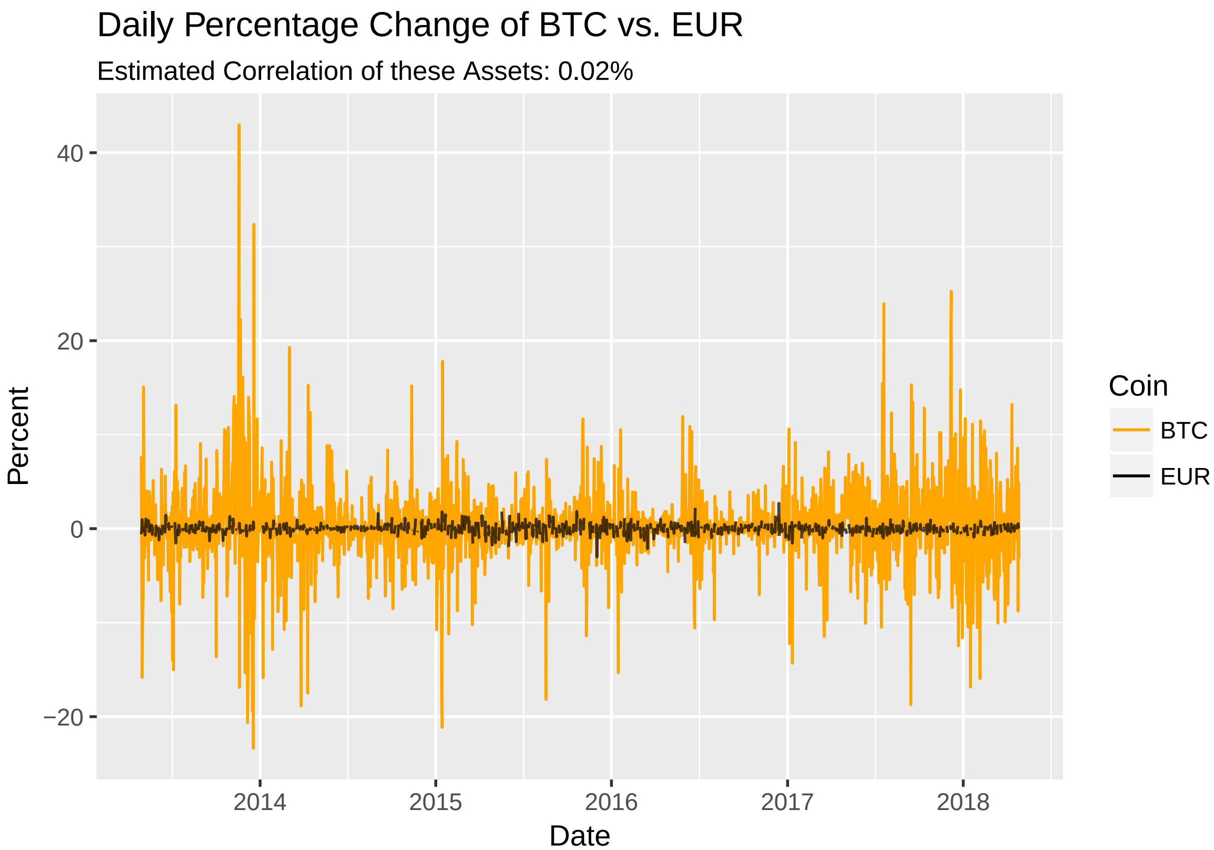
From this plot we get that the estimated correlation between Bitcoin and Ether is 92.08%, which we can assess graphically.



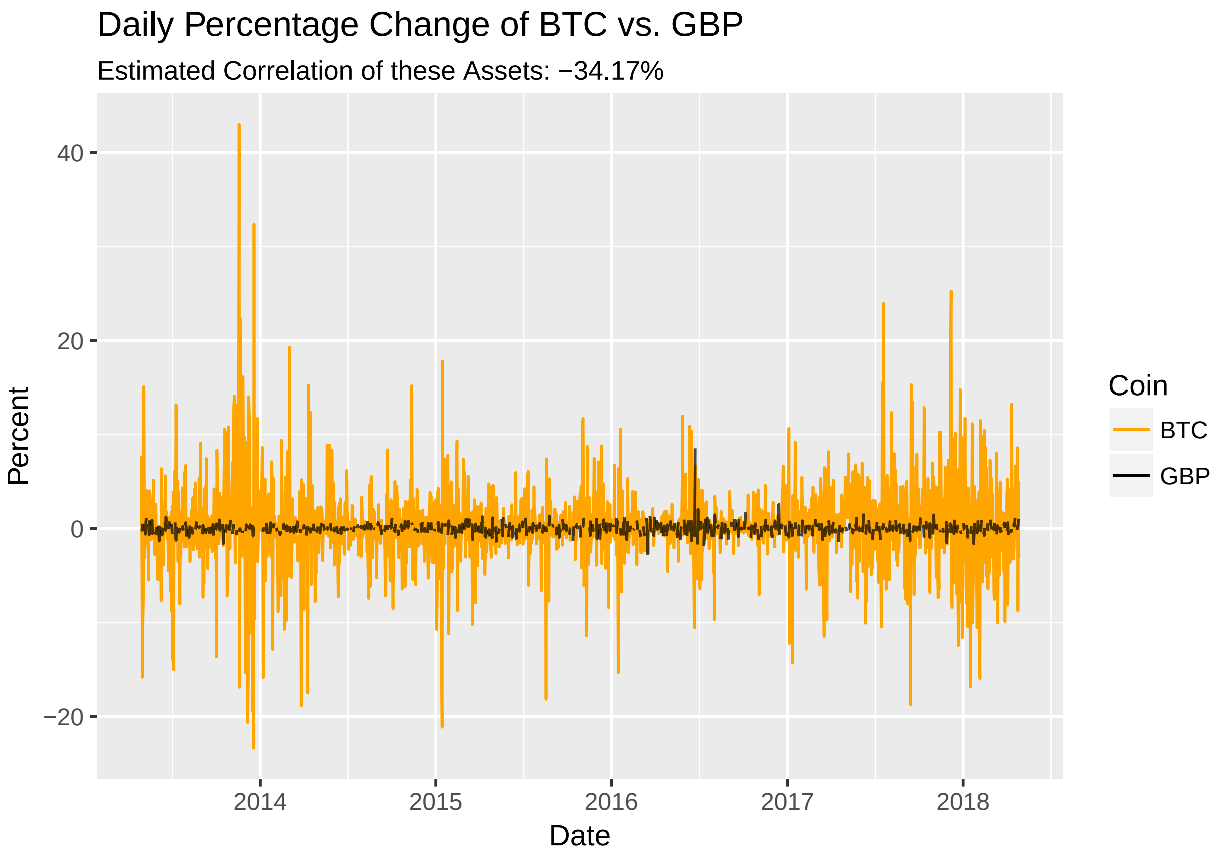
Other coins, such as EOS, show a much smaller correlation with BTC (67.01%). Note that this estimate is affected by the length of the data, which in the case of EOS is not lengthy.



Another market of interest is the foreign exchange market. The perk of forex data is that all currencies have existed since long before the inception of cryptocurrencies, so we can use the entire life of a coin as the period to compare. We proceed in the same way as in our previous plot but using a fiat currency instead of a cryptocurrency. As one can expect, we can see the cryptocurrencies are extremely more volatile than their fiat counterparts. Correlation decreases dramatically, being 0.02% in the case of bitcoin and the euro, this implies that we could not use the movements of one to correct our expectations about the other.

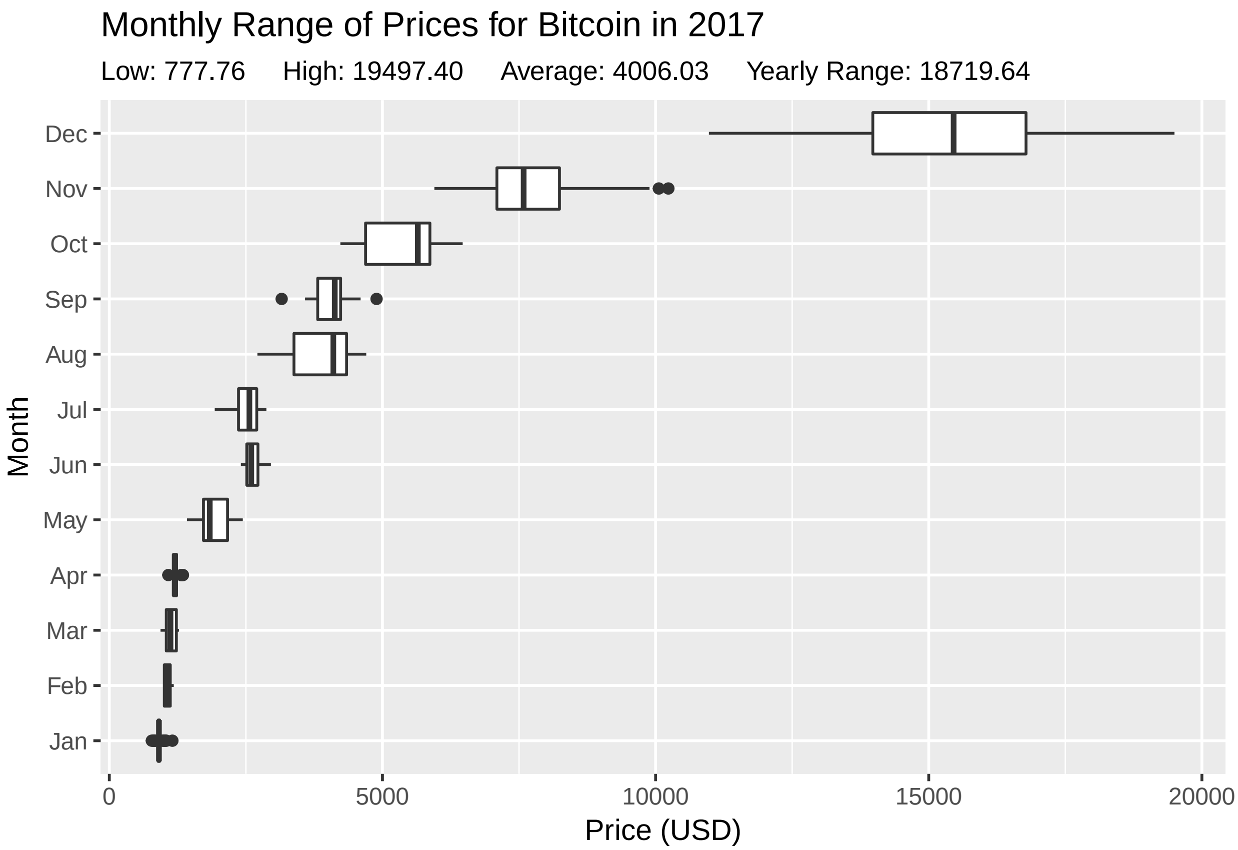


In fact, take a look at another major fiat currency, the British Pound Sterling. The negative correlation between these coins seems to indicate a contraian relationship between the two of them.

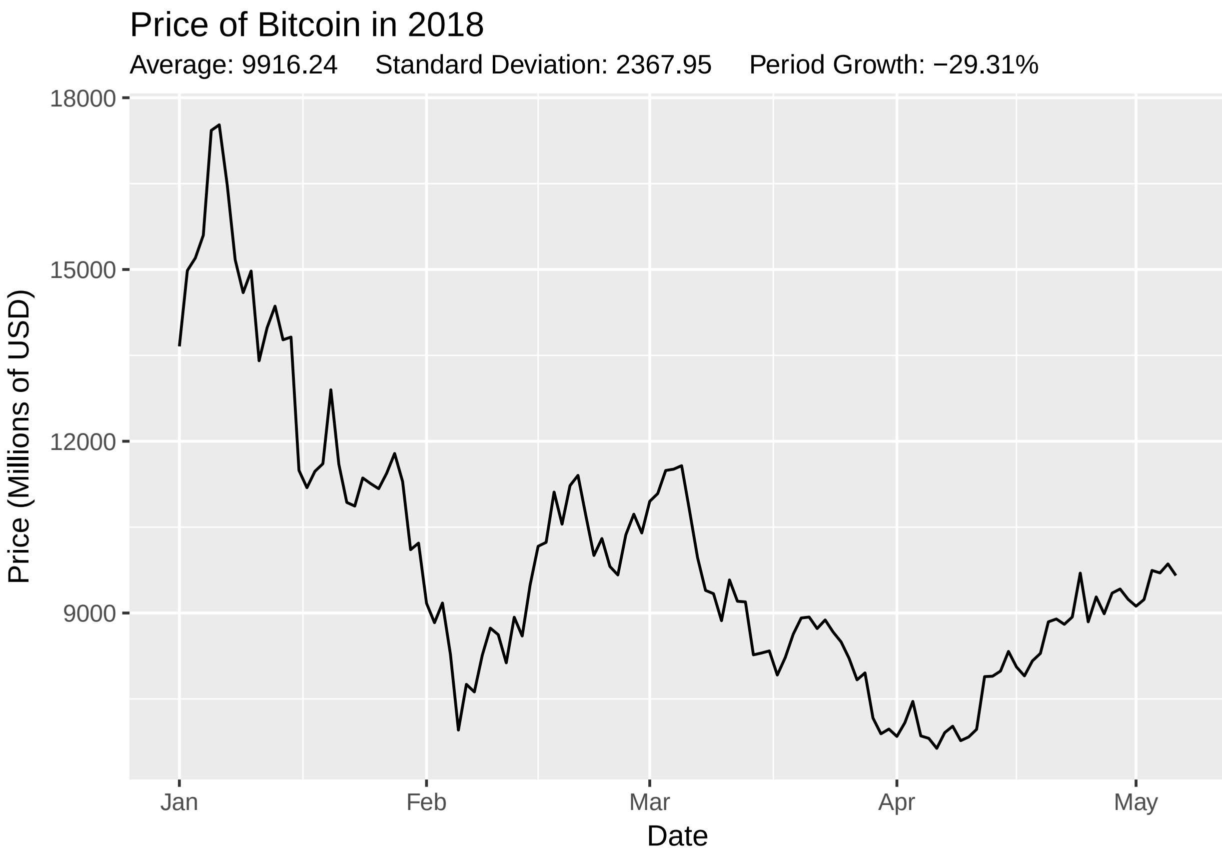


Performance

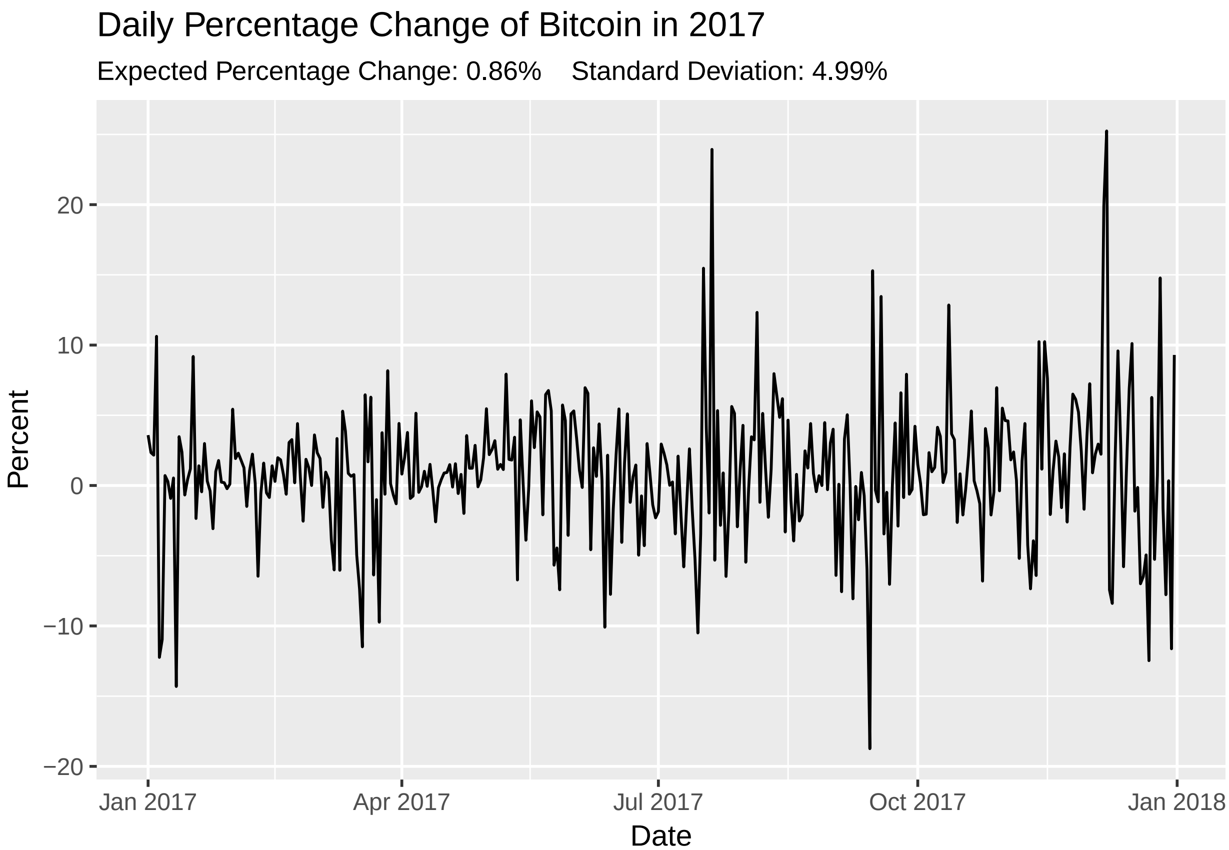
To explore how prices are changing within a given year, we filter the data to get only our coin and year of choice. We using the closing prices in this year to get the mean, lowest and highest prices. We then do a boxplot of the prices for each month in the year using daily observations of the closing prices. It is easy to see that the range of the prices is substantially different for each month.



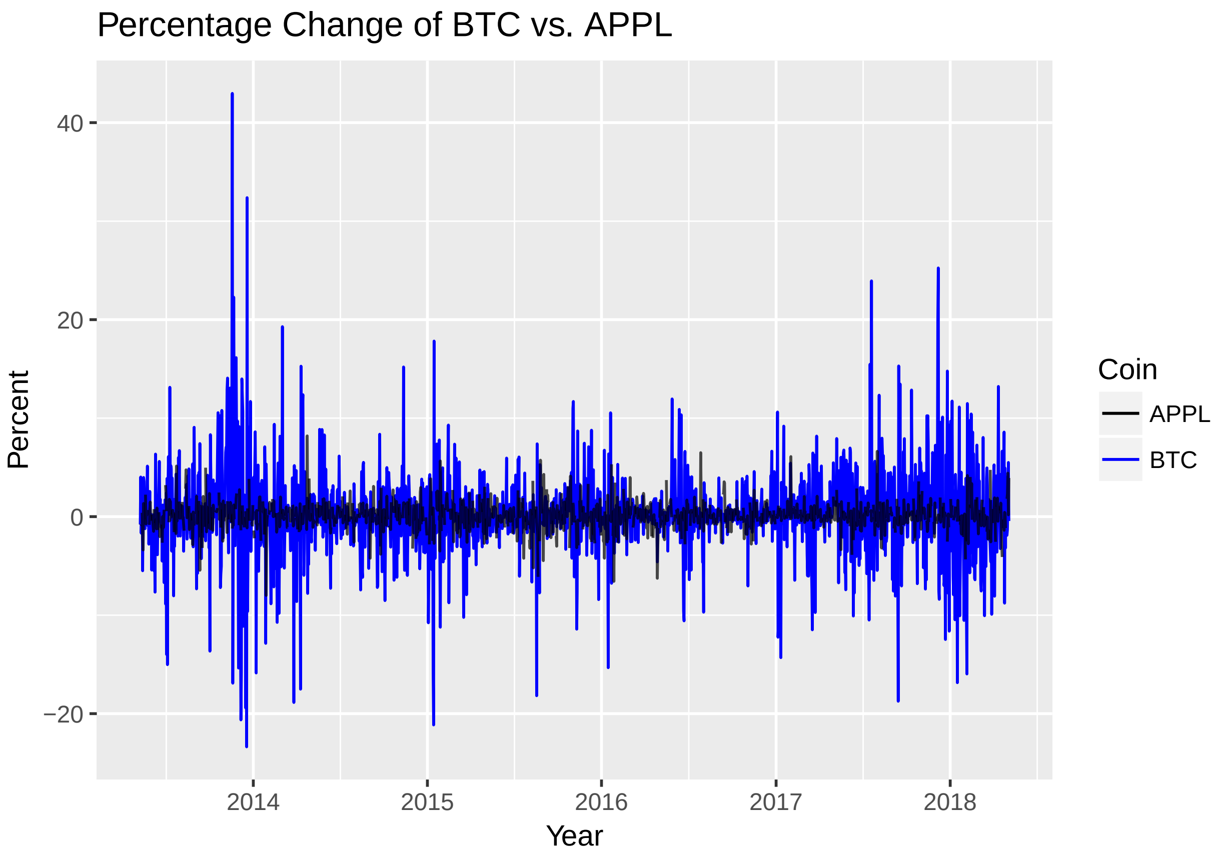
We zoom into a single year to be able to appreciate more carefully the change in price along this period. We start by spreading our date column into three columns containing year, month and day. Then we use built-in functions to get useful indicators such as the average price and the standard deviation. We define our function to calculate the percentage change from the first day to the last day in the period. Now we can look closer at the yearly prices of any coin in the top 10 and try to see any patterns in price.



We reshape the code of the above plot to reflect changes in term of percentages instead of prices. This will provide us with a plot that will make comparisons between cryptocurrencies easier. One surprising discovery that we make from this plot and we could have hardly made looking at our previous plot, is that even if we have very wild price swings in the short run, when we take the long run into account, we can see that all this changes in price even out and give us and expected daily percentage change of just about 1%. This should make bitcoin look far less volatile than it usually is looked at.

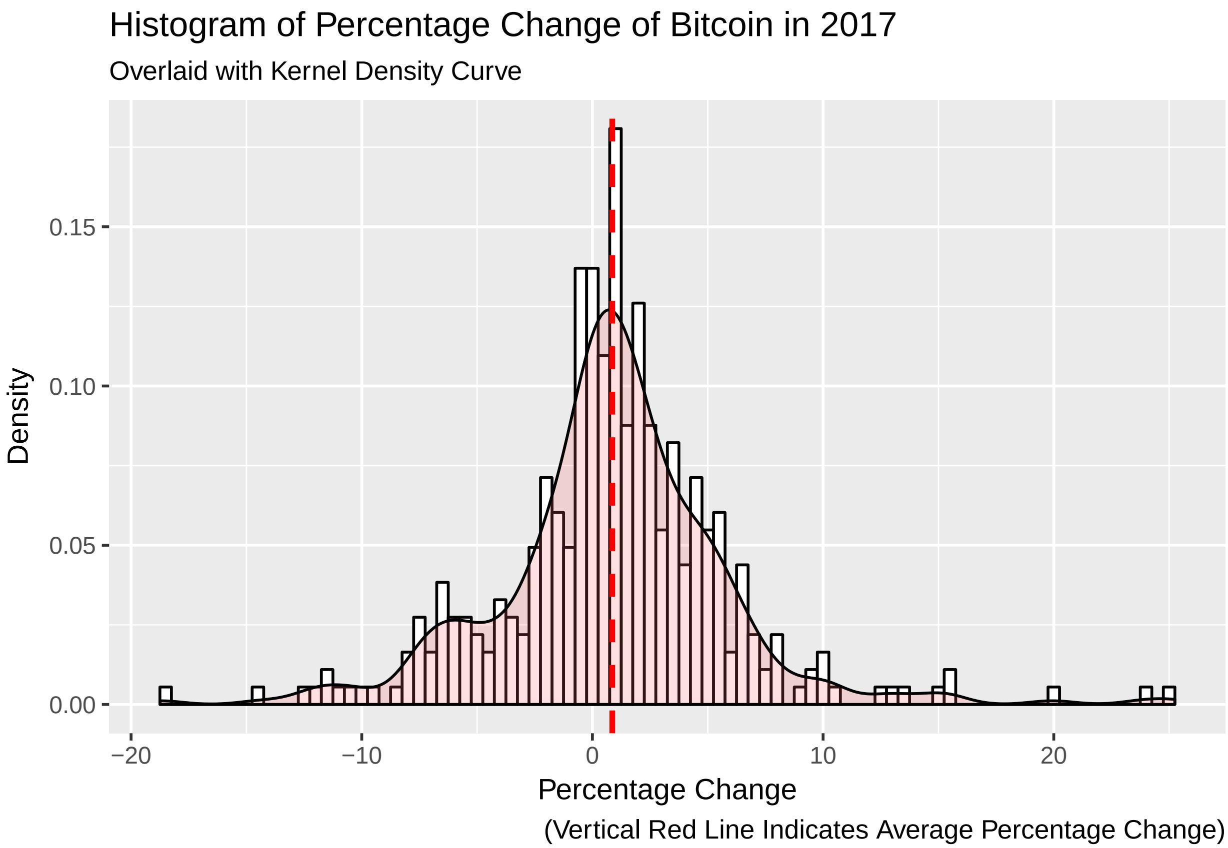


It would also be interesting to know how cryptocurrencies perform against other "risky" assets such as stocks. We plot the percentage chain of bitcoin on top of one of Wall Street's favorite company's stock, Apple. We choose this stock in particular because it is one of the stocks with the highest market capitalization of its respective market.



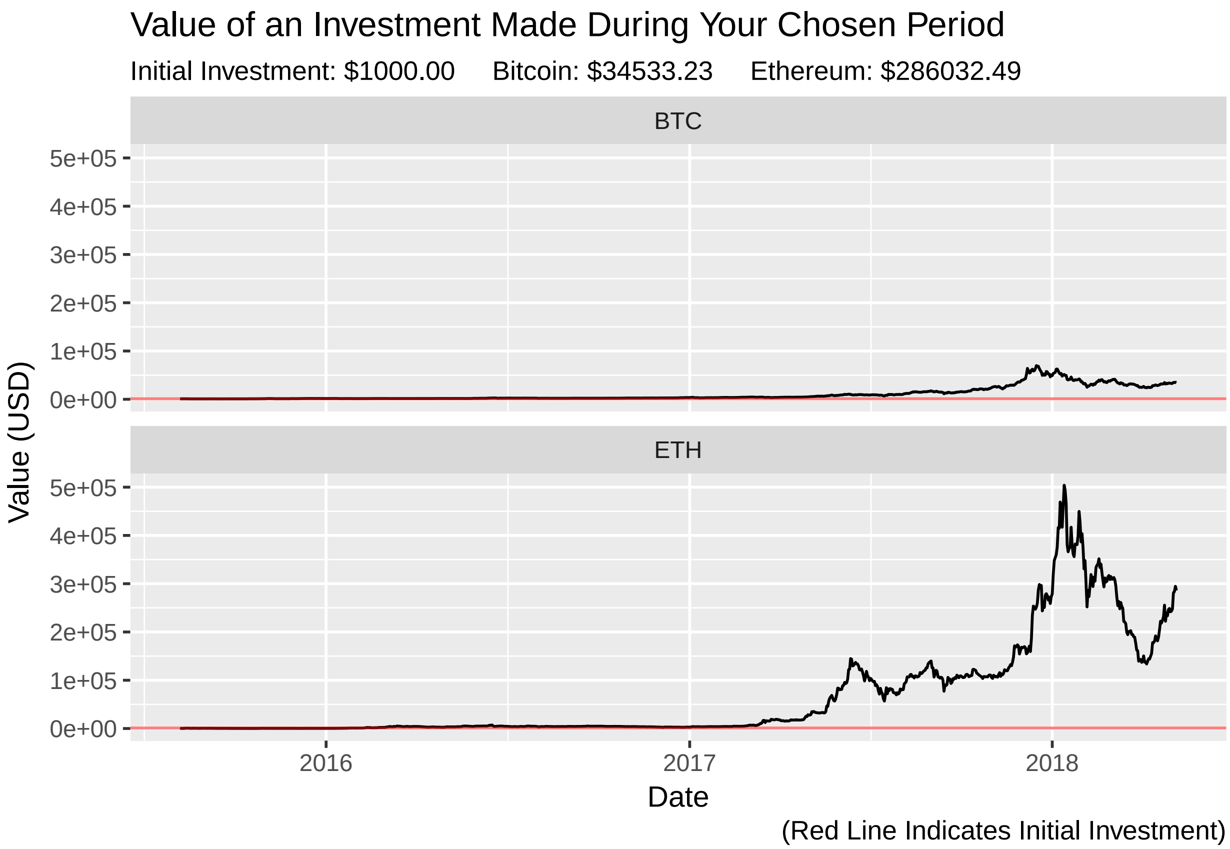
In our yearly performance (%) plot we noticed that in the long run, rises and falls in prices cancel each other and we get an expected daily percentage change of about 1% for BTC. Nevertheless, we do not know how these changes behave, do we have mostly small falls in prices but a few really big rises? Is it the other way around? One of the major assumptions of these speculative assets is that the price follows a random walk, because if it did not we could be able to accurately forecast its price and make a safe profit. Therefore, the price of this asset is a random variable, and so is its percentage change, so we can make a density plot of percentage changes in the year, and overlay it with a kernel density curve in order to see if we can get an idea about the probability density function of the variable.

The distribution of this variable looks almost like a normal distribution. Making the assumption that the daily percentage change of bitcoin is in fact a normal random variable with $\mu = 0.86$ and $\sigma = 4.99$ for any given year, we can make some inferences about it. For example, consider we have invested in bitcoin and we have profited, but we want to know the probability of facing tomorrow a price that is at least 5% lower than today's price. Well, we could easily calculate the area under the curve using our previous assumptions and come to the conclusion that $P(X <= \text{-5}) = 0.1201$



Investment Calculator

Finally we create a tool that lets us select a date range and calculate and compare the amount of money we would have if we had invested a specified amount of money during that period. We need to use again a *min_date* variable since not all currencies have for as long as others, and we want to compare between equal amounts of time. We filter each coin by year and apply the value formula that we defined in the beginning using our investment amount as an input. We then create a plot of this value column to see graphically how our investment changes from day to day. We include a red line at the amount of our original investment so we can always know on which days our investment was in the money and out of the money.



Conclusion

The lack of an option to get historical data from coinmarketcap's API gives us a small delay in our data collection process; however, this website contains information on almost every cryptocurrency in circulating, in contrast to any other crypto exchange or free database. Some final conclusions that we can make is that most of the cryptocurrencies are highly correlated, but we cannot say with certainty that Bitcoin is the one that leads the way. The forex market appears to be totally disconnected from the crypto market, as evidenced by the low correlation between some of the major coins of these markets. Although some forex currencies seem to behave opposite as the cryptocurrencies.

Another very interesting fact is that the percentage change in bitcoin seems to be normally distributed in the last years. Some other currencies like Ripple and Ether appear to share this characteristc, whereas newer currencies seem to have another distribution. The reason for this is that probably since these coins are on their first years, they are more likely to behave in an unpredictable way, but as their popularity increases it is highly possible that its distribution will converge to a normal distribution like most of the other popular coins that we dealt with.