

Exploratory Data Analysis for Cryptocurrency

Jeffrey Adams, Pow Chang, Sweta Bhattacharya, Matt White

W209 - 1 - Summer 2021

Introduction:

Our research is focused on the cryptocurrency market which is still in its infancy. Although Bitcoin, the leading cryptocurrency, has been in circulation since 2009, in the past year there has been a paradigm shift where companies, governments, investors and hobbyists have started to take a real serious look at cryptocurrencies as a vehicle for their investments. With companies like Microstrategy and Tesla purchasing billions of dollars of bitcoins. However the crypto market is unregulated and is very sensitive to events that highly influence the movement of the value of coins. Our research and exploratory analysis are centered around how the crypto market moves, whether it's in conjunction with traditional markets and factors that affect traditional markets, or whether it's influenced by other factors that have little to no effect on traditional trading markets, stocks, ETFs and commodities. Our 3 hypotheses will help guide our research through exploration of the data in our datasets, to identify patterns, correlations and outliers in our data.

Observations in the data (without visualizations):

We analyzed several datasets including Yahoo Finance, Google searches, Twitter trends, Elon Musk tweets, and a collation of relevant world events and news articles. The Yahoo Finance dataset is referenced by API with configurable parameters so we could restrict our retrieval of data to specific stocks and time periods. In looking at the Yahoo Finance dataset we observe there are NaN entries where there is a holiday, or weekend when the markets are closed, this is true of all of the publicly traded stocks. For Bitcoin, Ether and cryptocurrencies we do not see this as these markets do not close.

We also had to transform the columns in the dataframe returned by Yahoo! Finance since we were unable to get Altair to work with nested columns.

Our dataset for Google searches and Twitter trends was built from their respective APIs and was seeded for keywords, “bitcoin”, “btc”, “ether”, “cryptocurrency” as a measure of how popular crypto currencies were on a day-by-day basis.

Our events database was constructed with data that we retrieved independently, including Elon Musk's tweets, government announcements that impacted the crypto market as well as other world events that may have an influence on crypto market performance.

In viewing this data it is a little tough to see patterns because it's time series data over a long period and there is what we believe high correlation with the data we have from the Yahoo! Finance dataset.

From our initial analysis of the datasets we were able to come up with the following hypotheses:

Hypothesis 1: Cryptocurrency market trends mirror traditional markets

Hypothesis 2: Alt coins are highly correlated with and follow Bitcoin price.

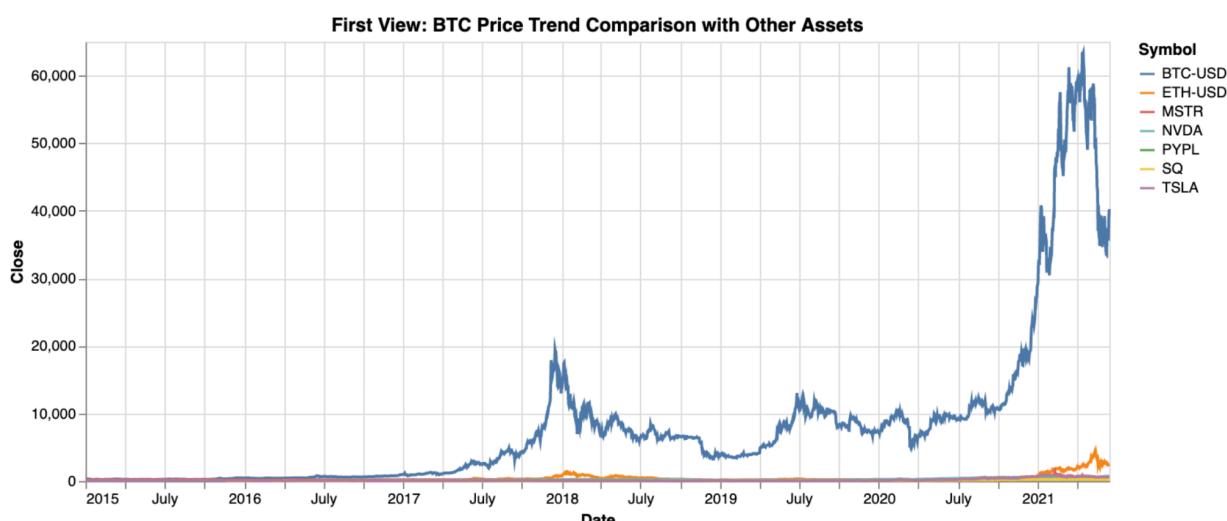
Hypothesis 3: Cryptocurrency market is highly influenced by non-market factors that don't affect traditional markets

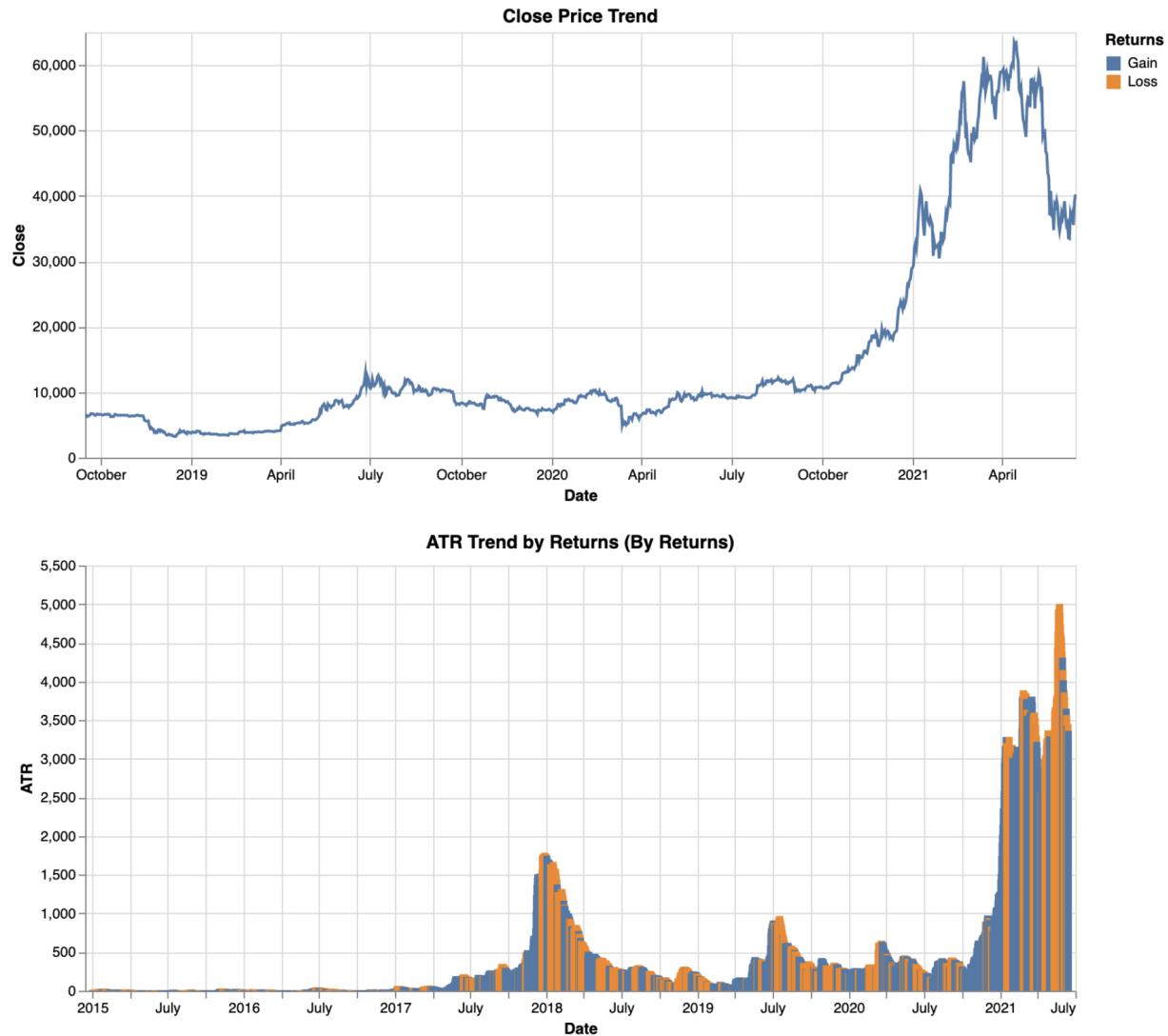
Hypothesis 1: Cryptocurrency Market Trends Mirror Traditional Markets.

Our hypothesis 1, which is our primary hypothesis for our exploratory analysis is that the crypto market mirrors the behavior of traditional markets, specifically NASDAQ and the NYSE. In the following section we set out to find evidence in support of and against our hypothesis to ascertain what is most probable by looking at the historical movements of traditional stocks, mutual funds, ETFs and crypto currencies.

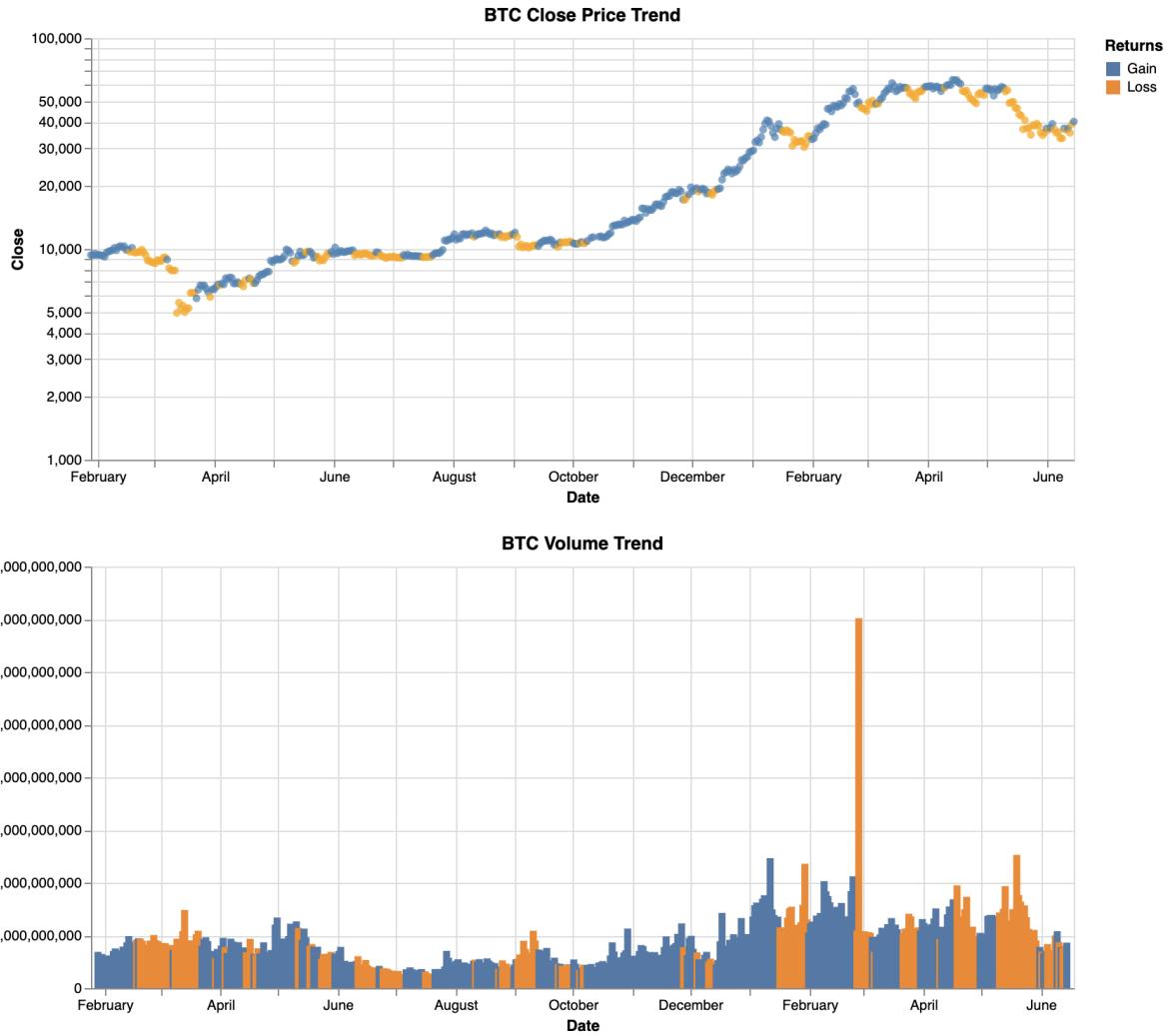
Exploratory Analysis:

Here we look at the values of stocks, ETFs, mutual funds and compare those to crypto values over time. We start with a look at some stocks and crypto currencies separately, then start to plot them together and contrast and compare them as we incrementally improve our visualization.



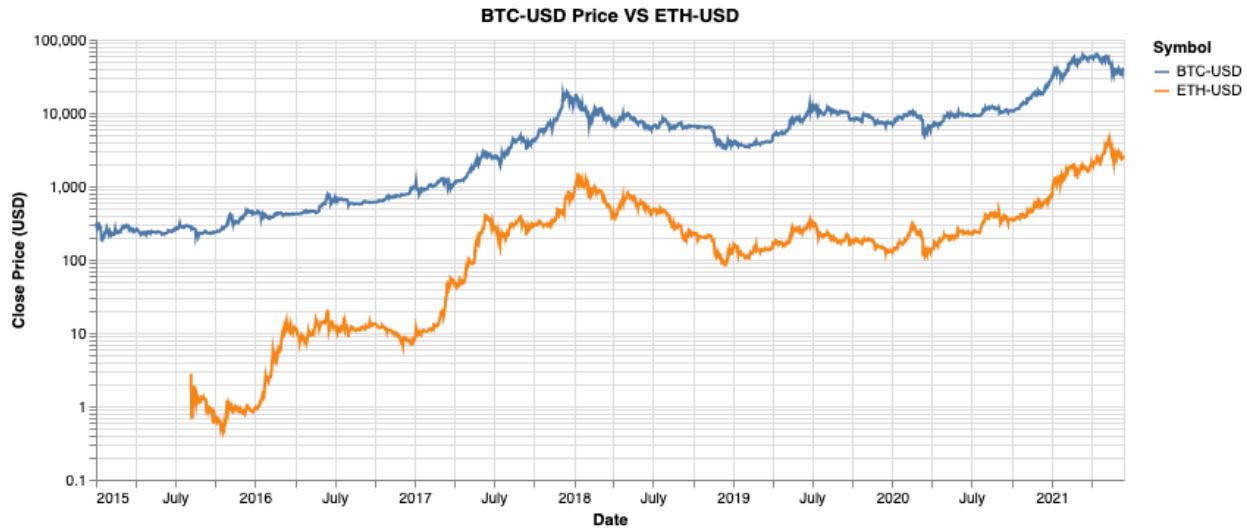


We first took a look at trends from February 2020 to June 2021 for Bitcoin alone. We can see that there is a definite upward trend starting in late 2020 and in January of 2021 a huge increase in ATR and price of Bitcoin.



We also try a slightly modified version of our first graph to visualize Bitcoin close prices over the same time period, which improves readability a little. Our second graph looks at trading volumes of Bitcoin, we observe that in March of 2021 there was a huge spike in trading volume. We will want to explore the cause of that in our 3rd hypothesis.

We will now add more assets to the graphs to see if there is correlated movement that would support our hypothesis that crypto markets move with traditional markets.



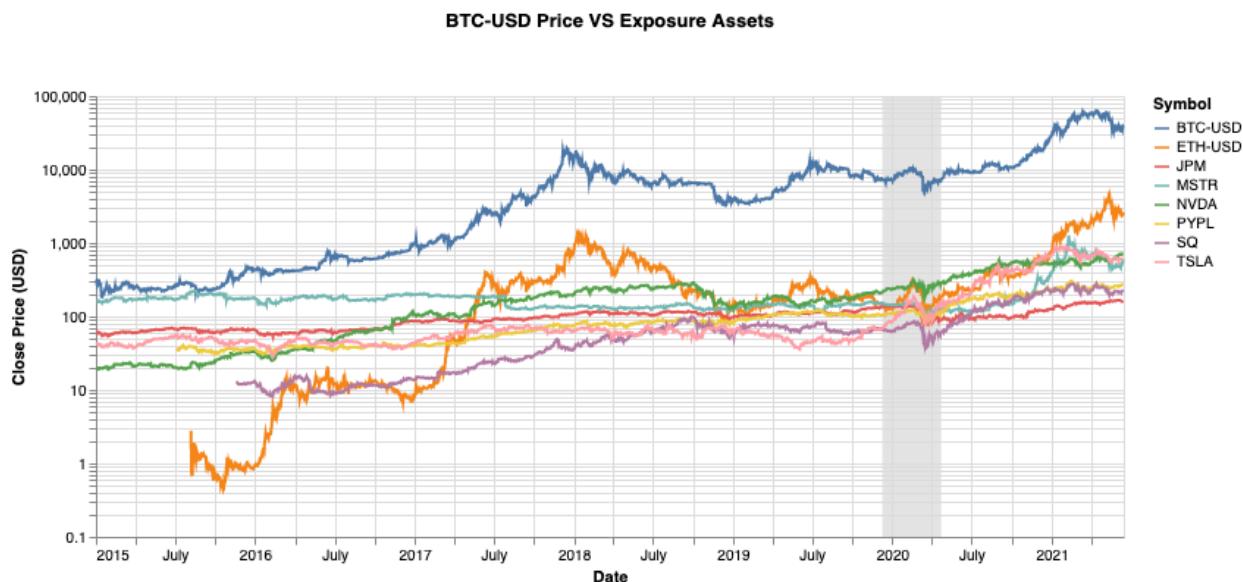
Here we look at the movements of Bitcoin compared to Ether, which are both cryptocurrencies. We can see that when Ether first comes onto the market in August 2016 that is quickly moves up in value and around July 2018 starts to mirror the movements of Bitcoin.



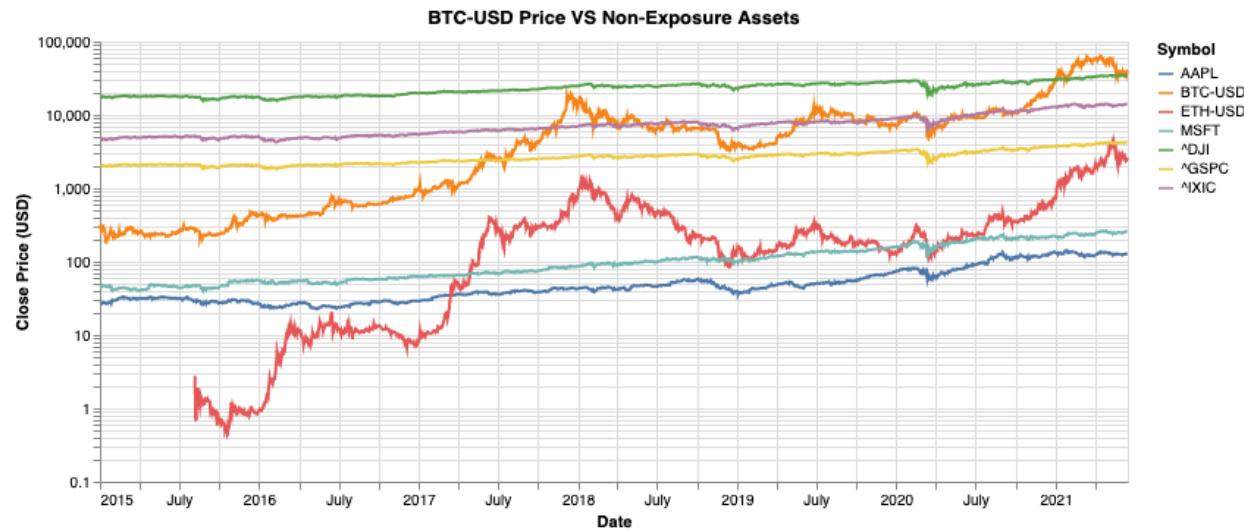
We first start by comparing close values in USD, but we see that many are very close in value. To improve, let's reduce our data points and modify the scale to log scale



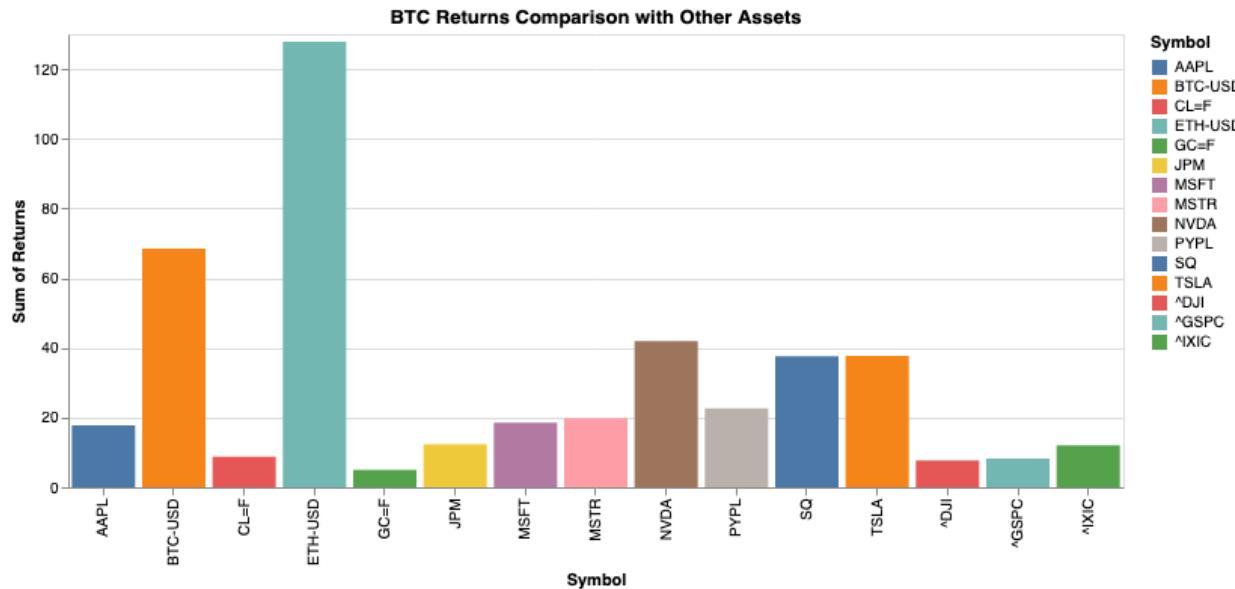
Here we have successive line graphs comparing the traditional markets with cryptocurrencies over the period from Jan 2015 to June 2021. What we see are very loose correlations, Bitcoin and Ether prices fluctuate drastically in comparison to traditional stocks. The peaks and valleys with cryptocurrencies are very dramatic and their rise beginning in late 2020 is meteoric. We were able to improve the graph through using colors from the same family for traditional markets and another color family for cryptocurrencies.



In this view you can see a few major index funds, stocks and cryptocurrencies Bitcoin and Ether. It can be observed that there might be some correlation between BTC and traditional markets but the overall movements of Ethereum and Bitcoin do move more erratically, especially Ether. You can see that all stocks and crypto currencies dip in Feb 2020. But we would attribute this to the beginning of the pandemic. These rough correlations can be seen in a few spots, but are not conclusive. We did improve this graph with log scale and highlighted comparison period from our experience working with the first set of graphs.

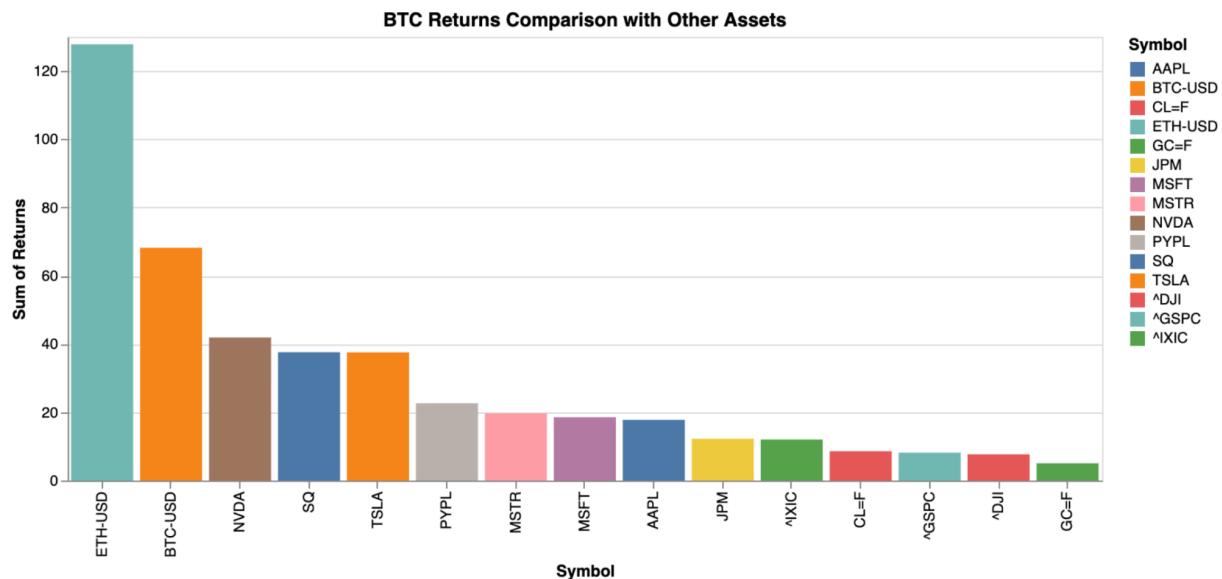


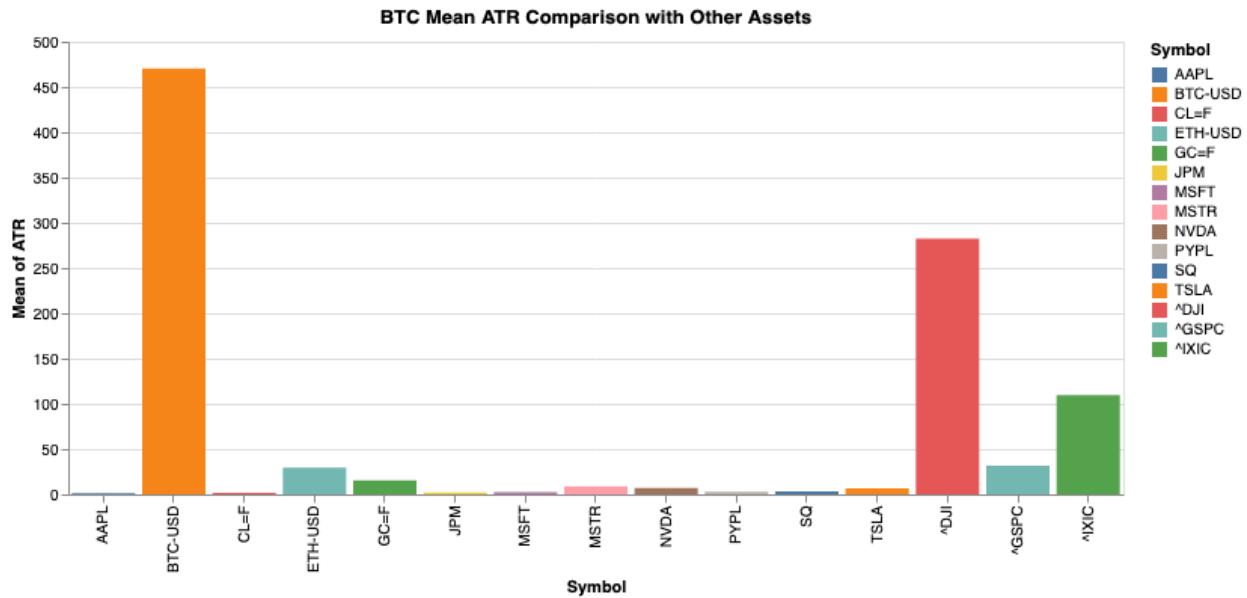
It is difficult to see any exact correlations between BTC and the others and most of the tickers are so low on the Y axis that any fluctuations in price can't be seen. A daily returns plot could be an improvement because it will move all of the tickers on to a similar scale.



We can sum our returns and create a bar chart that compares Bitcoin and Ethereum against non crypto assets. It's clear that Ethereum has some very high returns, then Bitcoin while the rest of the stocks and indexes have much smaller returns. We can improve on this graph by ordering from highest value to lowest to improve readability, shifting the highest returns to the left.

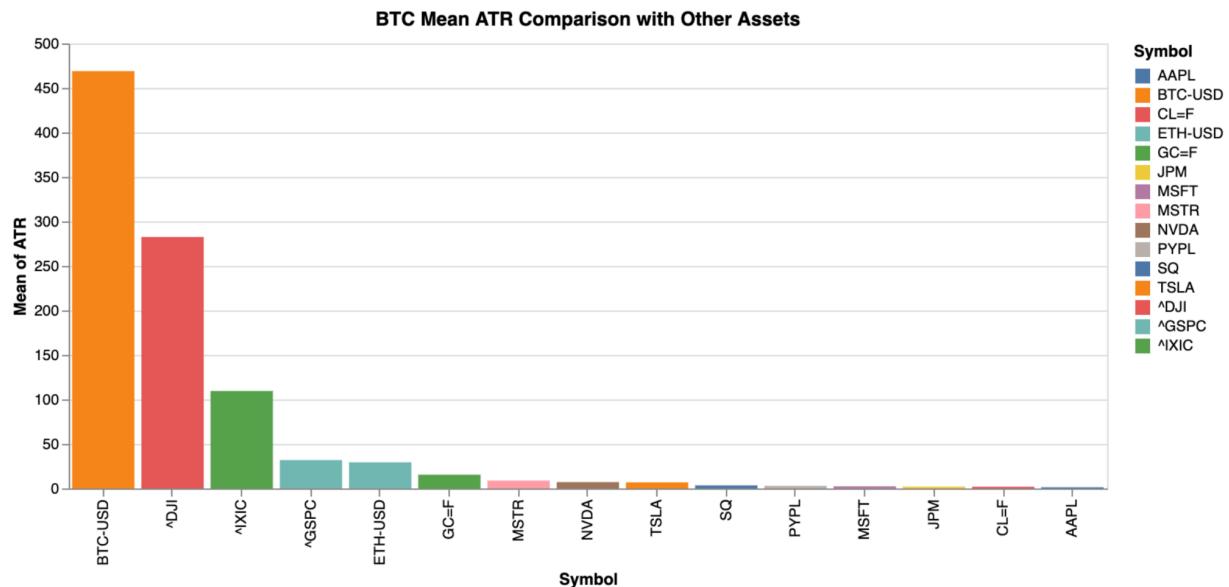
We further improved the view to be an ordered barchart as follows:

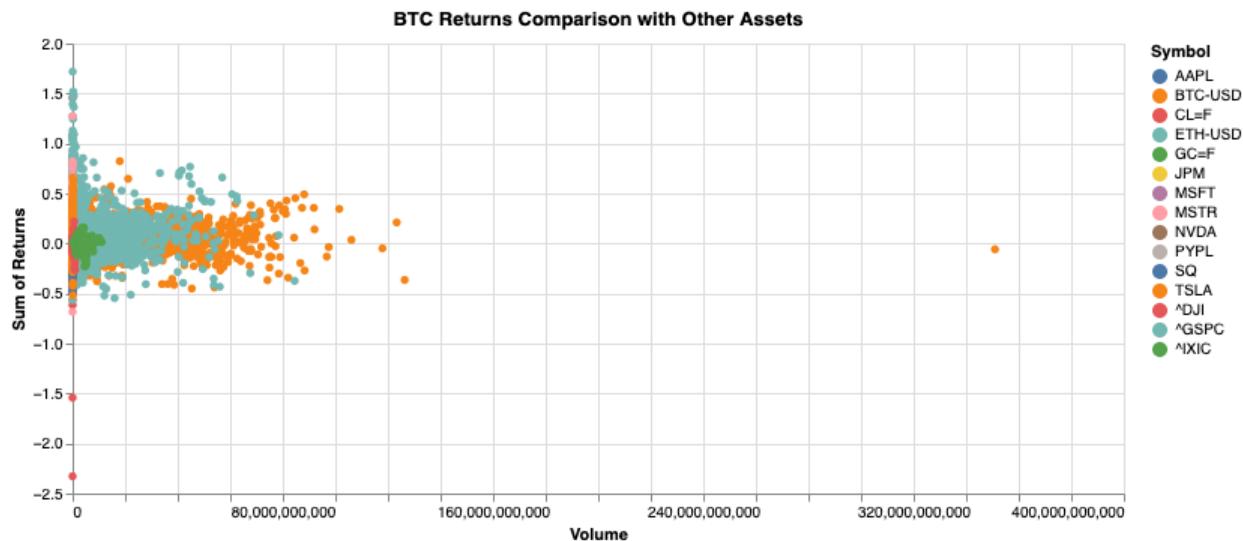




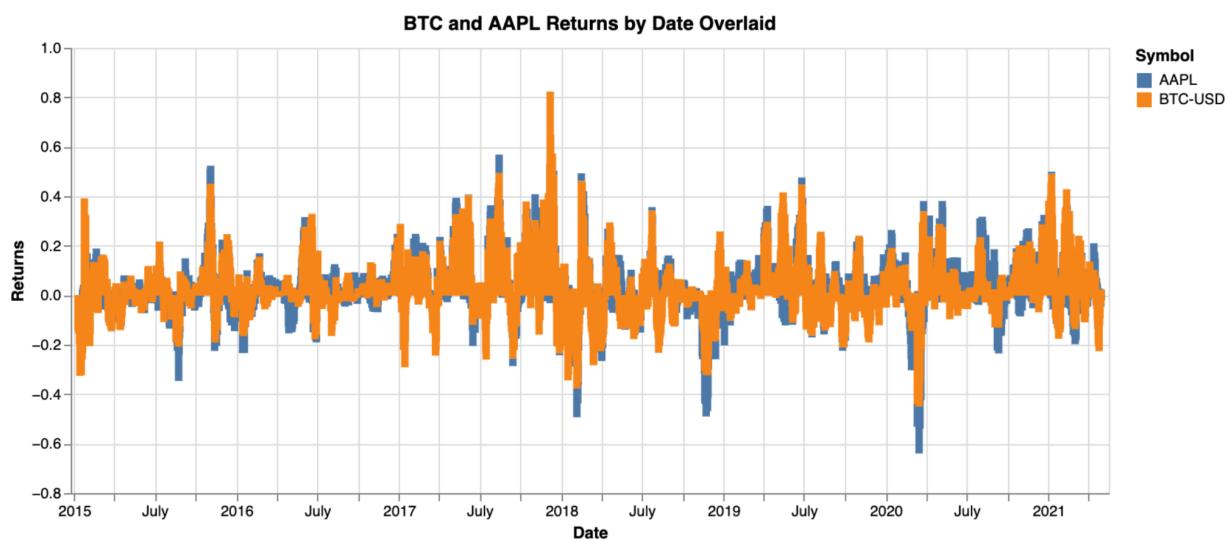
Here we have a quick look at the ATR (Average True Range) which is effectively a 14 day moving average of the true ranges of an asset. What this shows is that Bitcoin and DJI are highly volatile compared to other assets. We can improve this graph by ordering from highest to lowest ATR.

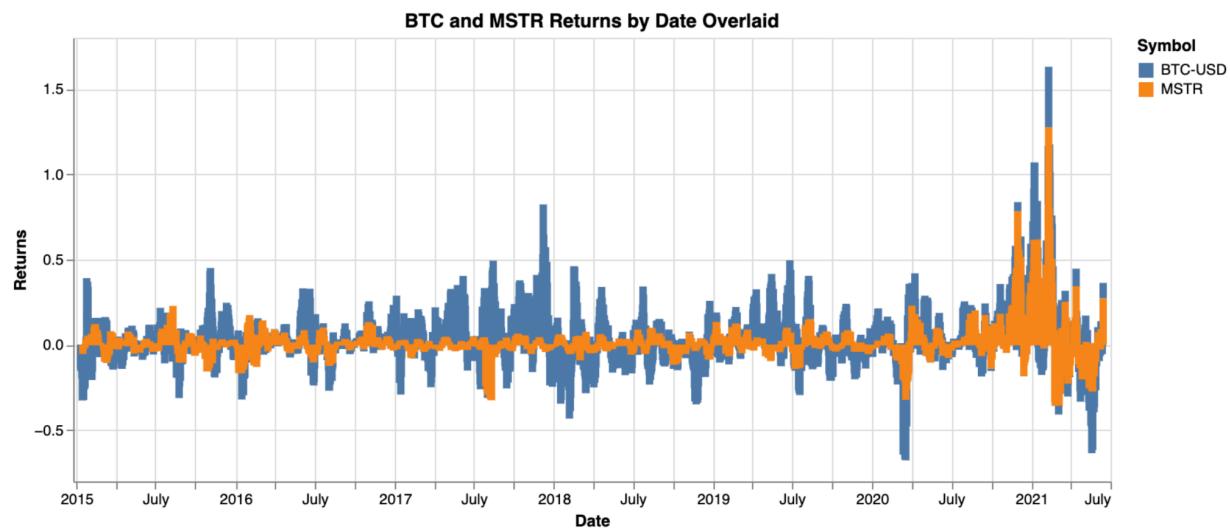
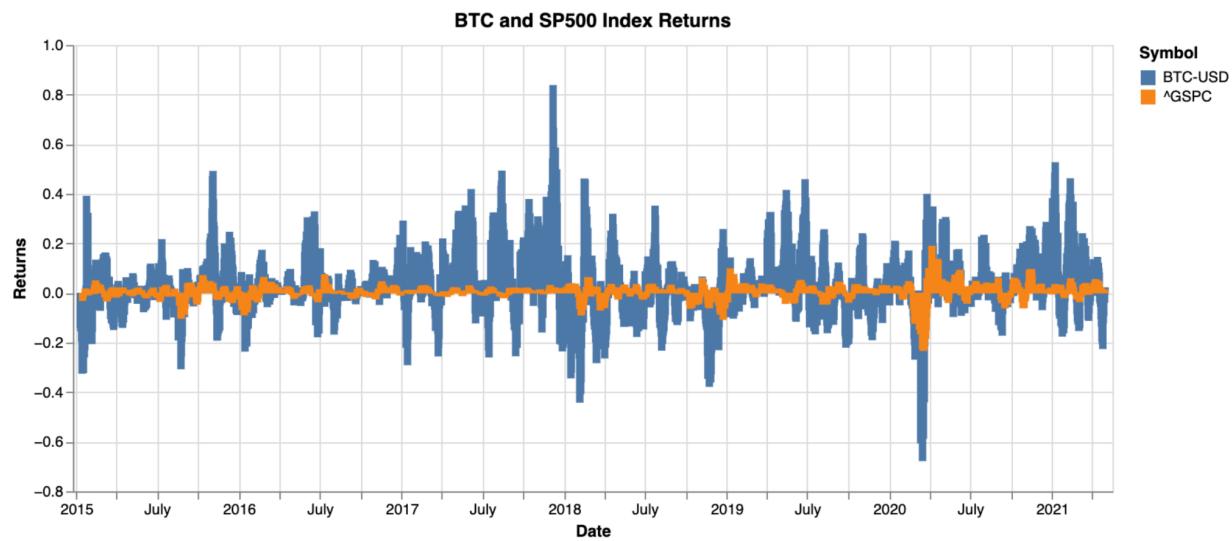
We improve the view further by ordering the barchart in descending order here.

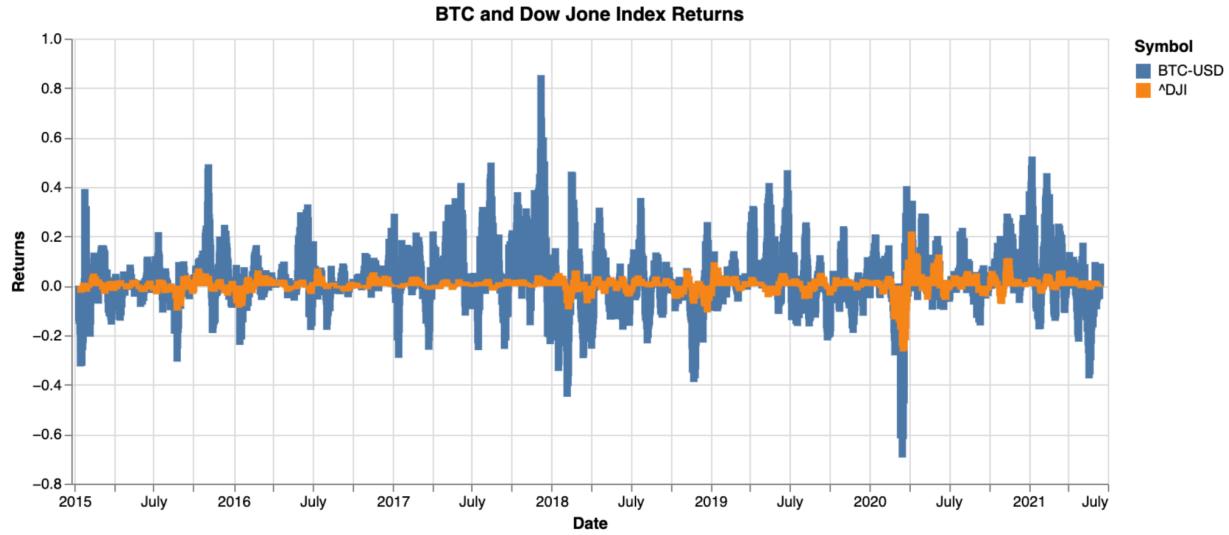




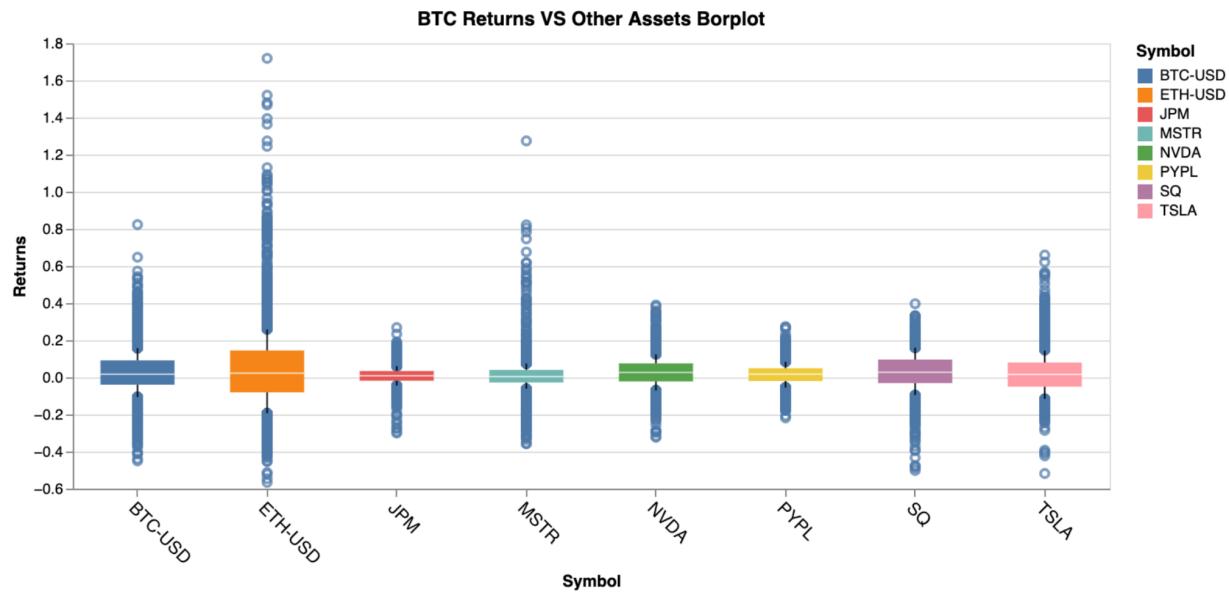
Here we tried a different graph to look at returns vs volume, and we can see that both Bitcoin and Ether have higher returns with the traditional markets really see that most stocks have much less volume and lower values with smaller market capitalization.



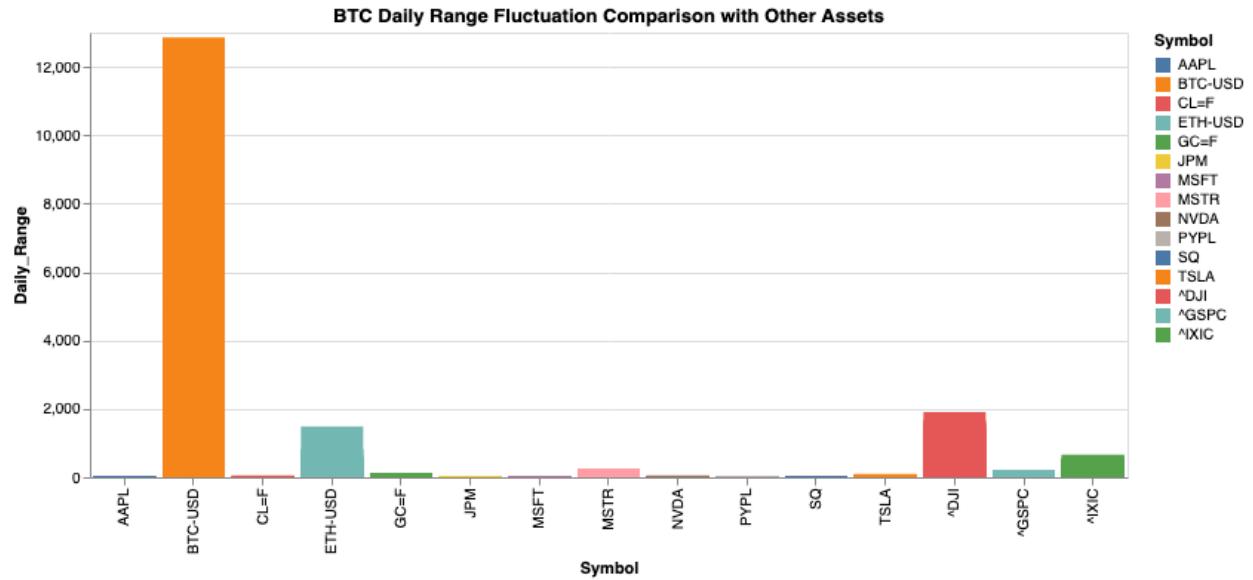




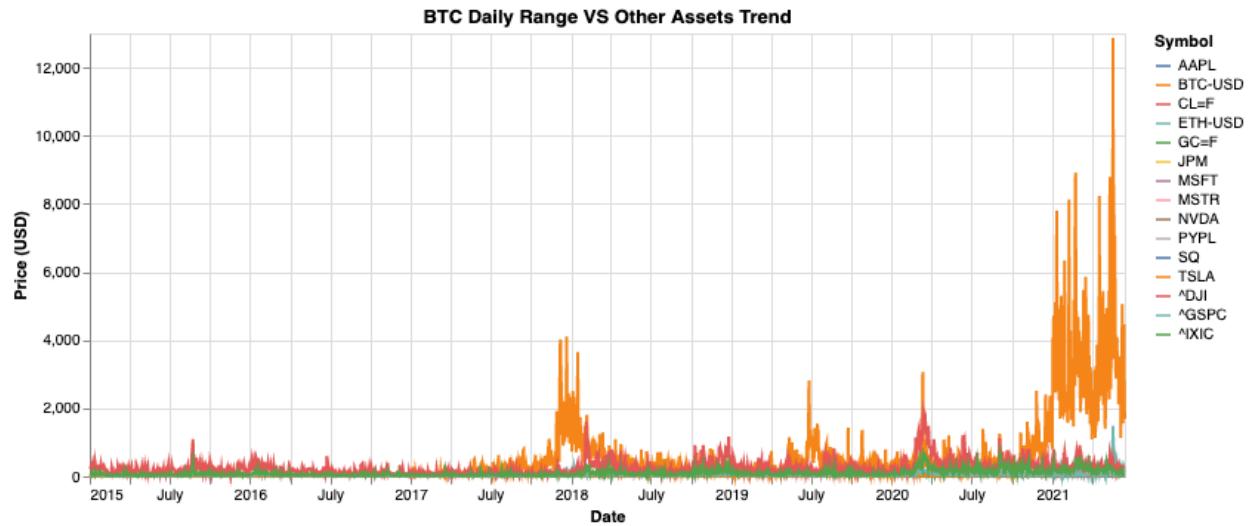
The last series of graphs we compare returns of Bitcoin to a number of stocks and indexes, namely Apple, the Dow Jones Industrial, S&P 500 and Microstrategy (known to have substantial Bitcoin holdings.) Apple seems to have very similar returns to Bitcoin, where the other stocks and indexes do not, with some exception of returns with Microstrategy in July 2020 mirroring the returns of Bitcoin.



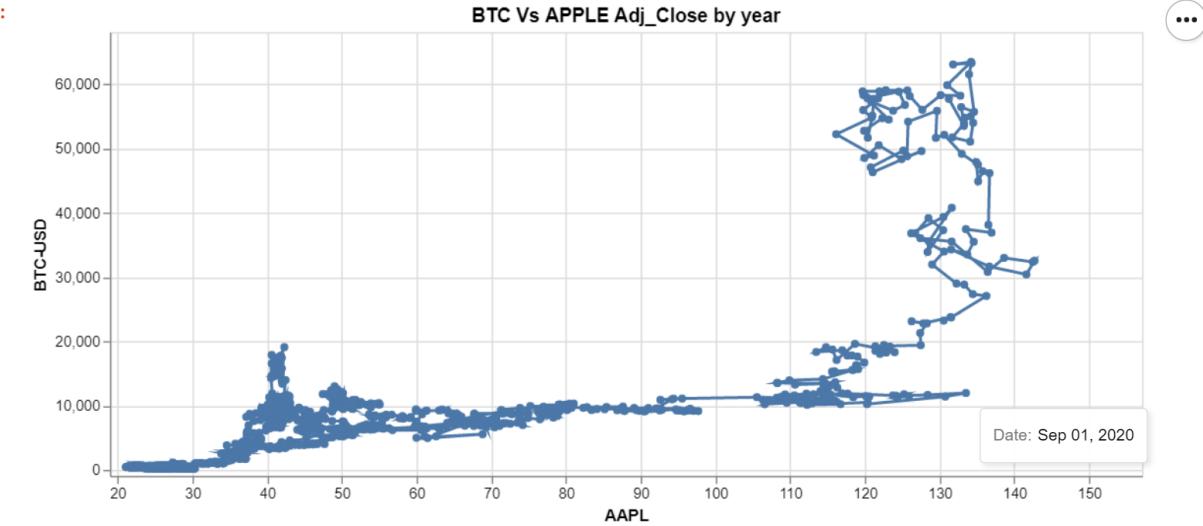
We now try a box plot to see if we can improve readability of overall returns. We do see that the median of the crypto currencies are higher than traditional stocks with higher variance. This would affirm our past observations that crypto currencies do have higher returns.



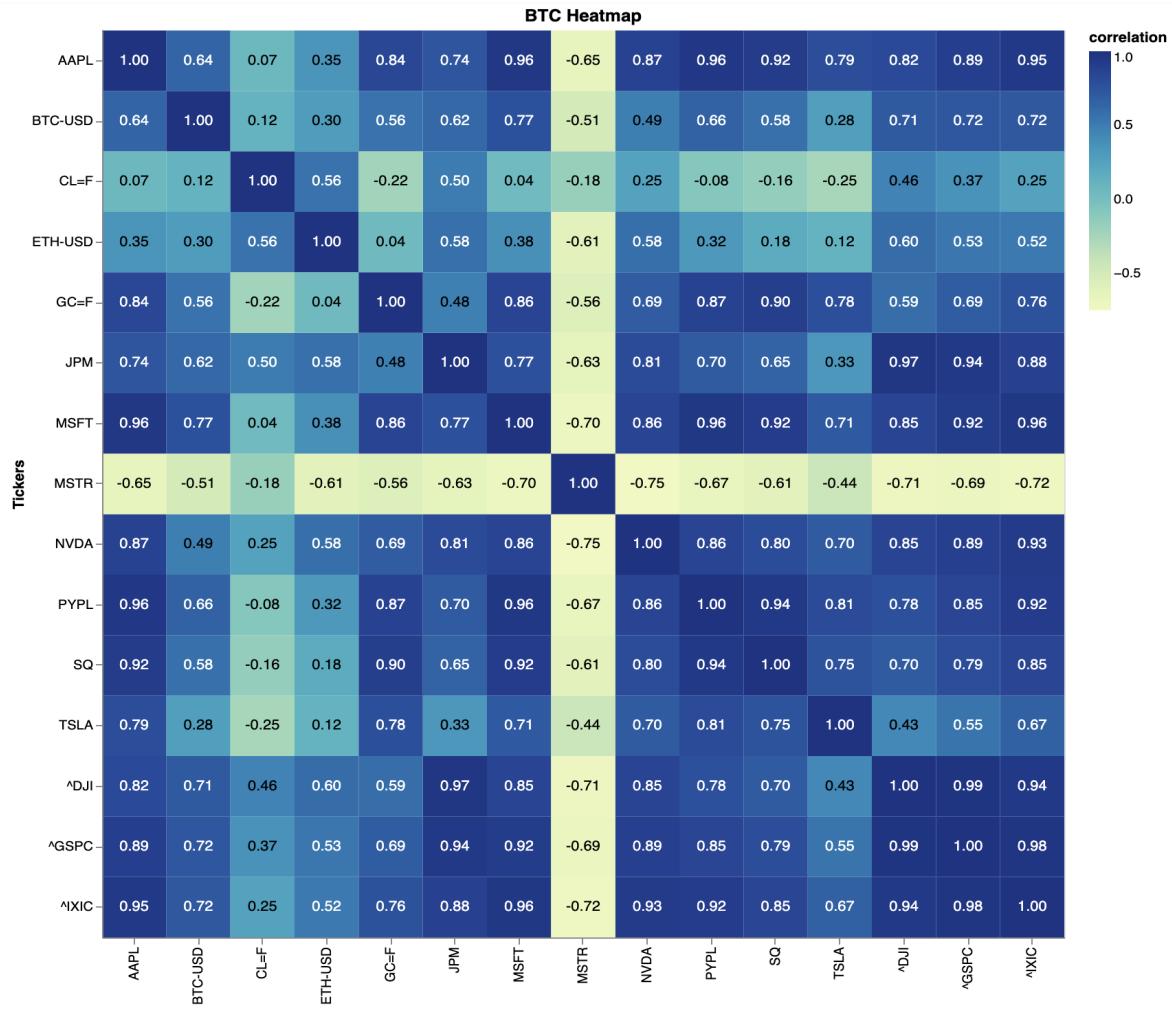
Here we try bar charts again to visualize the daily fluctuations of stocks, crypto currencies and indexes. We can see that by leaps and bounds Bitcoin fluctuates the most, then the Dow Jones Industrial (which is price indexed and is expected) then Ethereum with the others not seeing as much volatility.



A play on the previous bar chart but over our overall timeline we see that Bitcoin again doesn't seem to move with the rest of the market, especially starting in January 2021.



In the above interactive plot, we compare Bitcoin's Adjusted price with Apple's using a connected scatter plot where the dots are representing years. With the tooltip option we have shown that on September 1, 2020 Apple was nearing an all time high whereas Bitcoin was quite low. Hovering the mouse we noticed that Bitcoin has risen quite a bit in the year 2021 compared to Apple. Here, the connected dots tell a story of the two tickers quite well on a yearly basis when the visualization is viewed with its animation. We believe this plot will give some insights on any unusual event-driven price volatility, this graph should be interpreted together with twitter data at the later stage.



Here we have a correlation matrix that looks at the movement of stocks, indexes and cryptocurrencies with respect to one another. The indexes seem to all move in a highly correlated manner with one another, which would be expected. We don't see that any other security moves with Bitcoin, it's clearly on its own.

Conclusion:

Our hypothesis was that the crypto market mirrors the movements of traditional markets. What we observed is that crypto currencies are highly volatile, with their values fluctuating drastically even over the course of the same day. That the patterns of Bitcoin and Ether appear closely linked but their movements, returns, market capitalization, and other indices do not appear to mirror traditional markets.

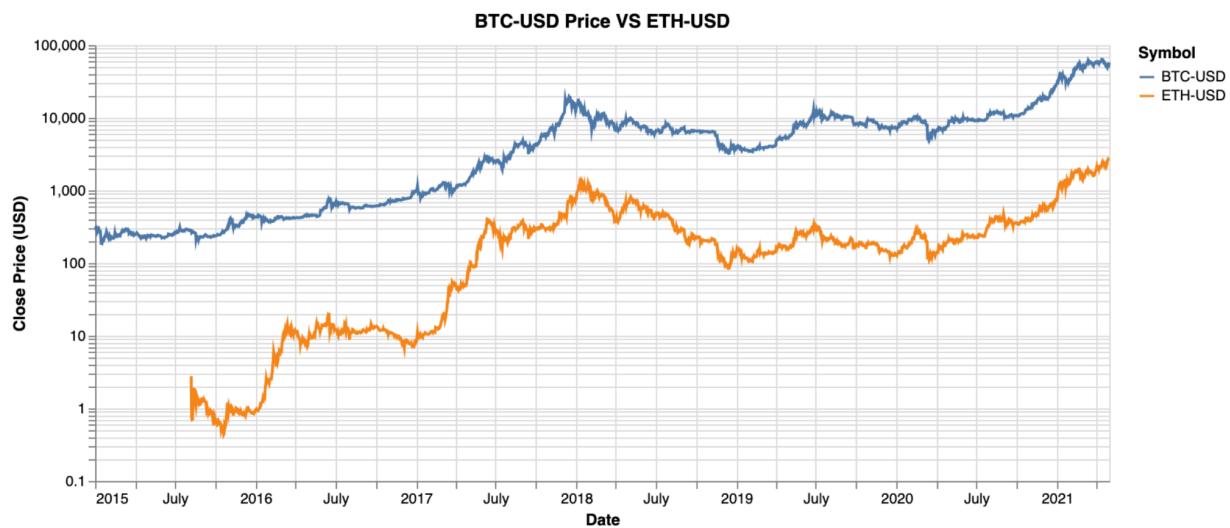
We know that the crypto market is always open, unlike NASDAQ, NYSE and traditional markets trading happens every day and night. We also know that crypto markets are unregulated, and anyone can create a crypto currency. We wonder if those who trade in crypto are different from

those who trade in conventional assets (we know to some degree this to be true but also changing.) As we move further with our project we will have to try to ascertain the underlying factors that influence the movement of cryptocurrencies, but I think it is safe to say that we can reject our hypothesis that the crypto market mirrors the movements of traditional markets.

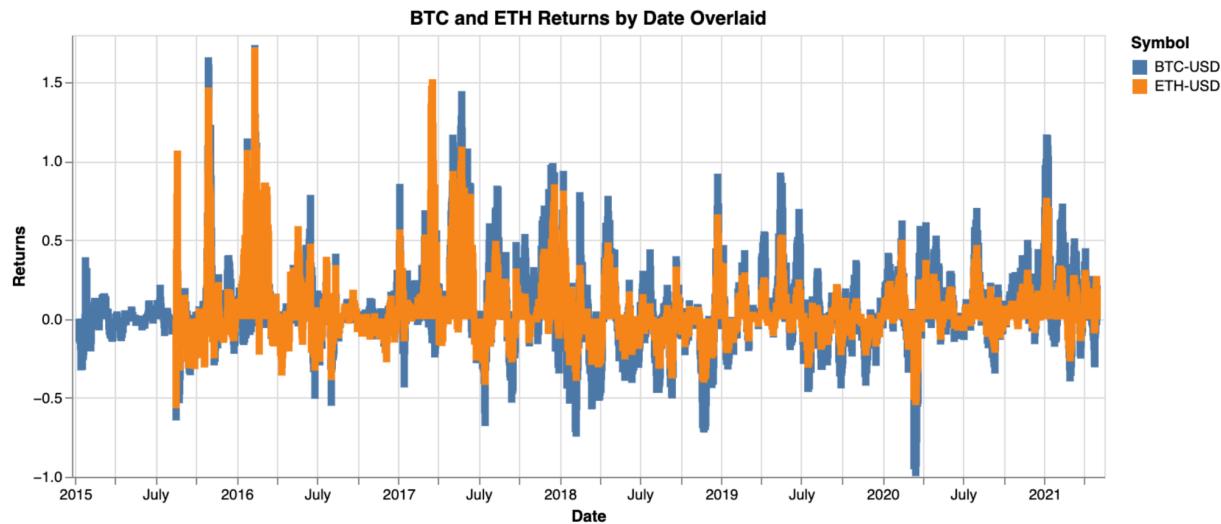
Hypothesis 2: Alt Coins are Highly Correlated With and Follow Bitcoin Price.

Our second hypothesis is that alt (alternative) coins, which is a catch-all term for cryptocurrencies that are not Bitcoin or Ether, are highly correlated with the price of Bitcoin. So effectively the entire crypto market shadows the movements of Bitcoin. When Bitcoin goes up, so does everything else, and likewise when the value of Bitcoin goes down so do the alt coins.

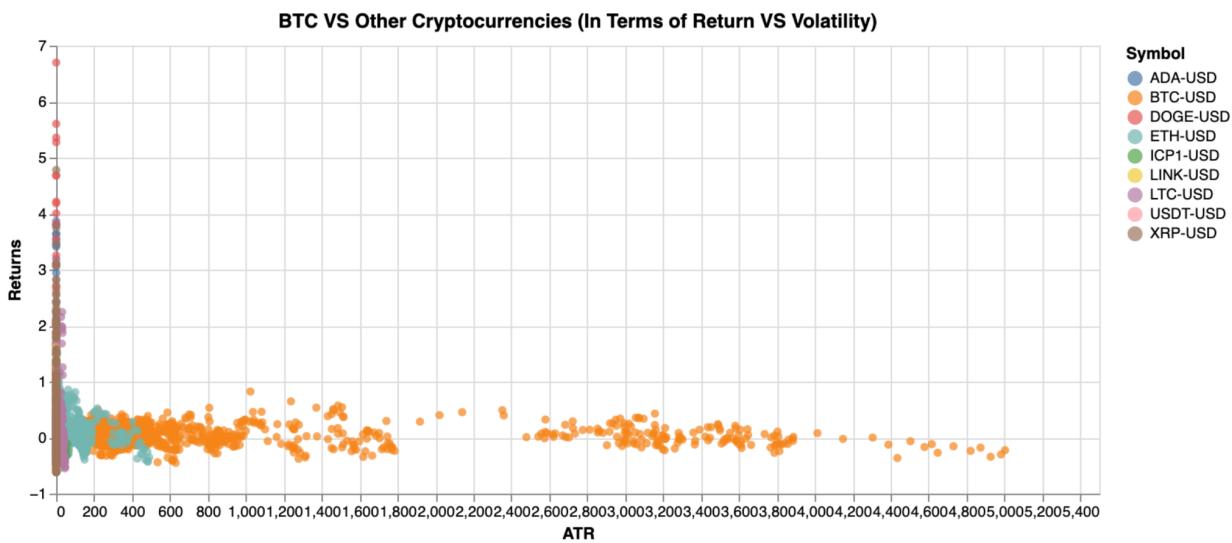
Exploratory Analysis:



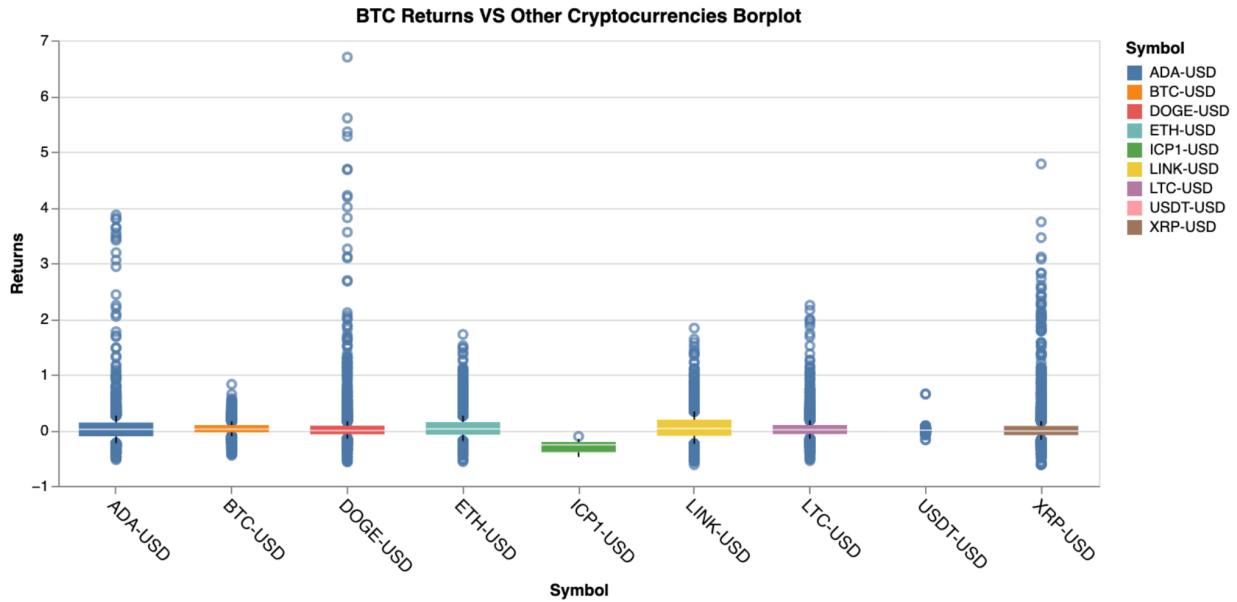
First we revisit our plot of Bitcoin vs. Ether by value over time to see if we see consistent movement. We do see that after around March 2017 Ether starts to more closely mimic the movements of Bitcoin, which we know to be the market leader.



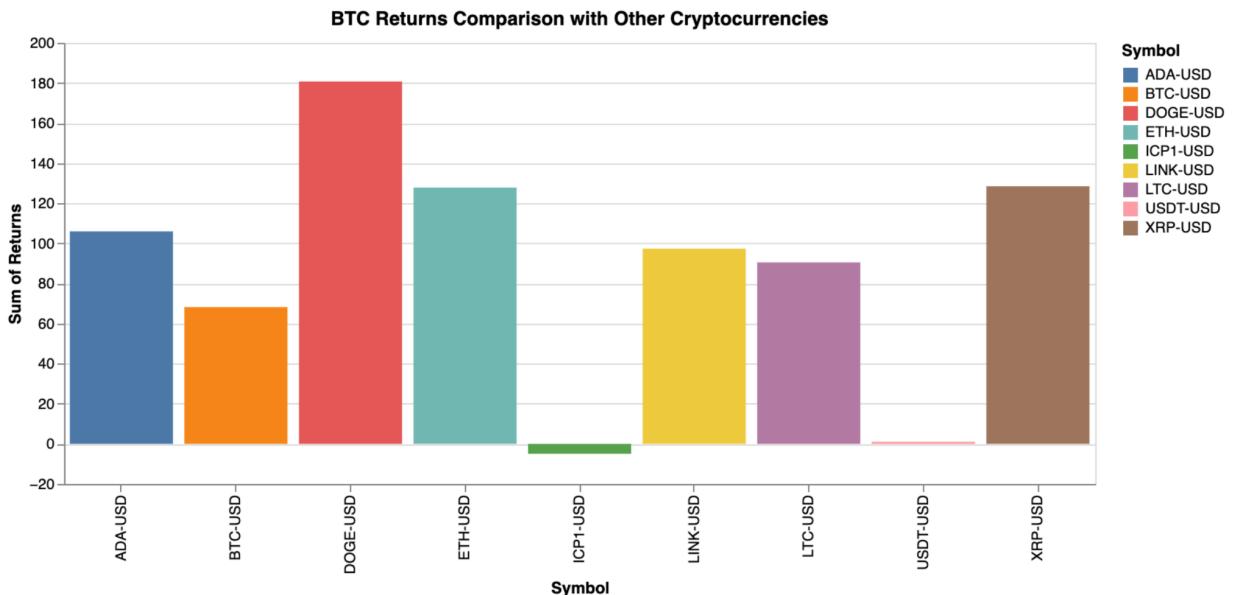
Here we are overlaying returns from Bitcoin and Ether and we can see they are very similar in movement which at least supports the notion that either Ether follows Bitcoin or Bitcoin follows Ether.



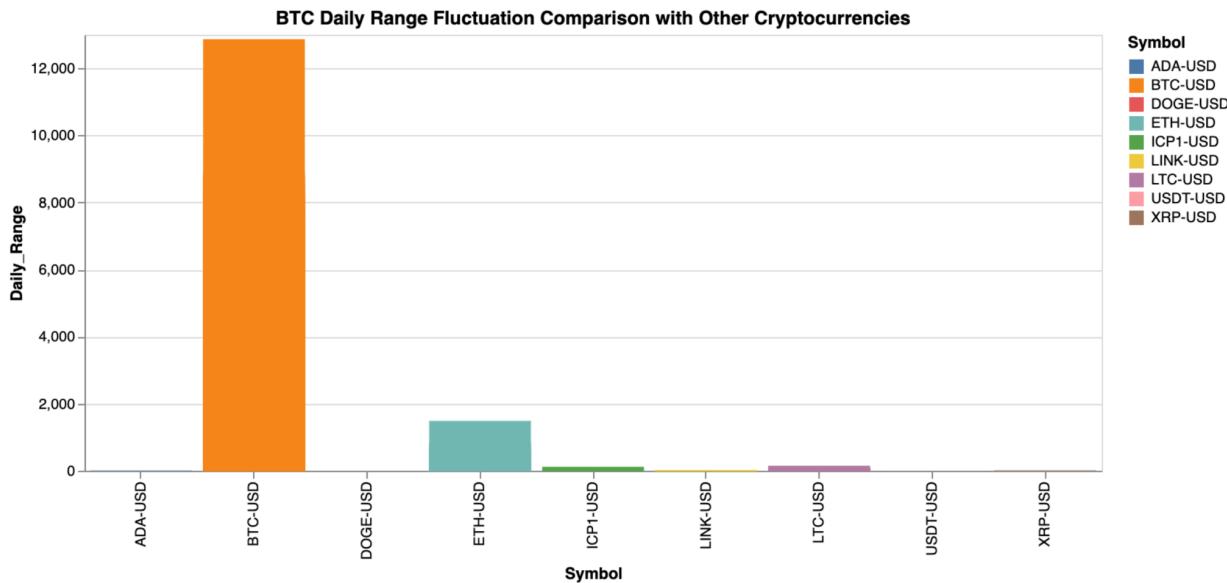
Here we look at the returns vs ATR to see if there are any patterns or correlations between Bitcoin, Ether and alt coins. We don't observe that alt coins mirror the returns or volatility of Bitcoin or Ether.



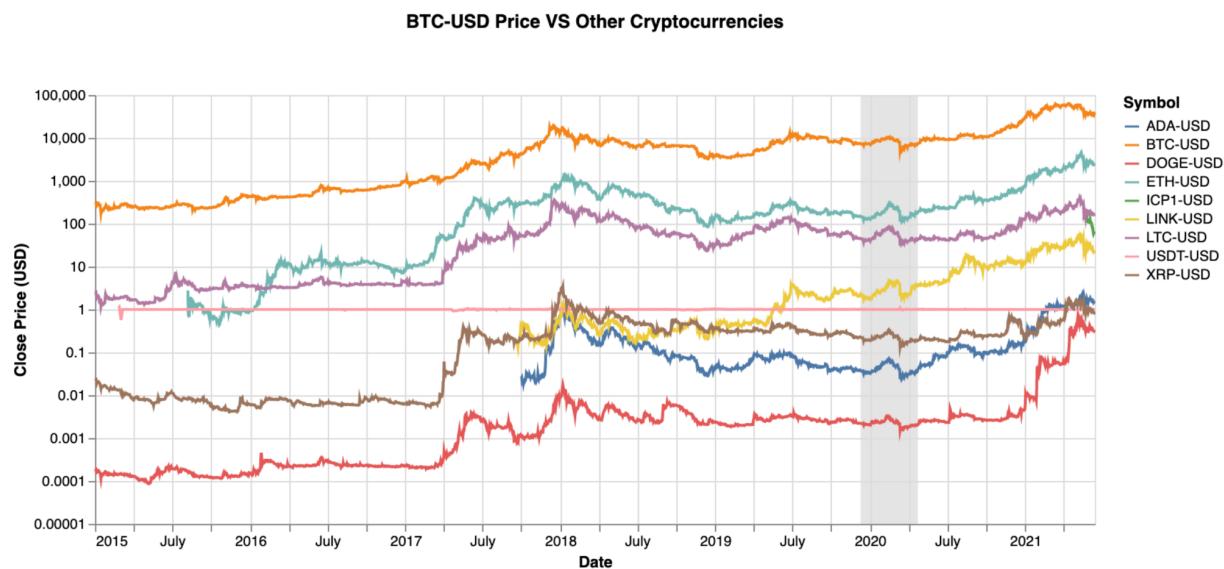
Here we use a BoxPlot to show how returns for various crypto currencies. Where previously we saw Bitcoin had very high returns compared to conventional stocks, here we see that in comparison to other high performing alt coins, the gains are not substantial. DOGE, ADA, XRP and nearly all others outperform Bitcoin.



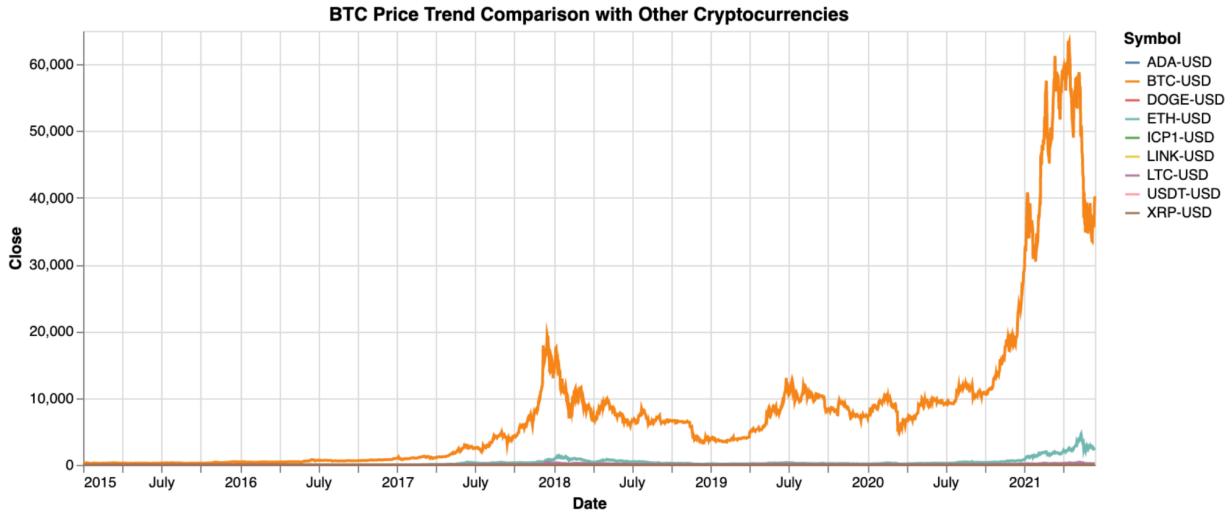
Here we use bar charts to compare returns, which makes it a little more clear that DOGE, ETH, ADA and XRP have the highest returns.



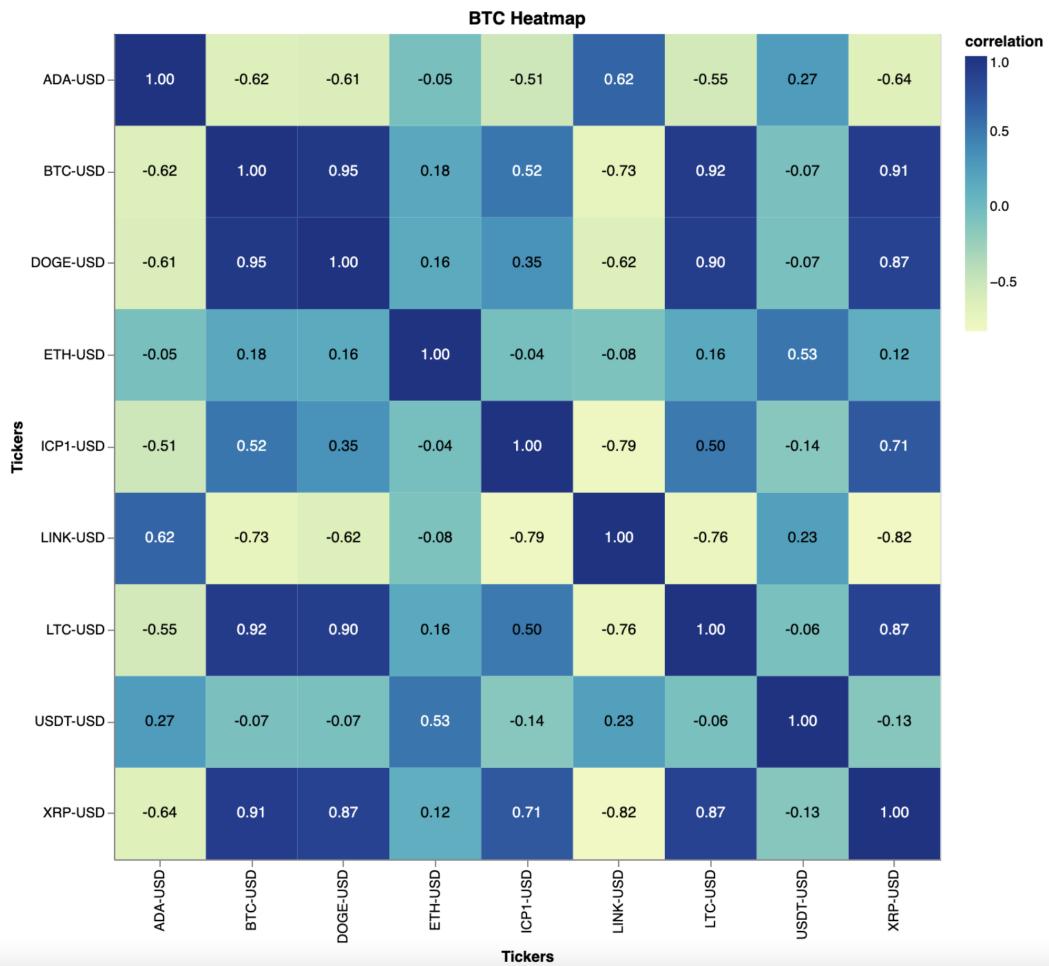
From this bar chart we see that Bitcoin is much more volatile than other crypto currencies, which leads us to believe that alt coins do not move with Bitcoin and Ether. The alt coins seem pretty stable in movement while Bitcoin fluctuates drastically, ethereum much less but more than other alt coins. But volatility doesn't tell the whole story.



When we look at Bitcoin vs alt coins we do see that the market actually moves together overall. One anomaly we observe is that DOGE in January 2021 which we know is when Elon Musk started to tweet about DOGE which drove up its value.



A look at overall market value we can see that most alt coins do not register on the graph as the market value of Bitcoin is just so much higher on a per coin basis.



Again we use a correlation matrix to see what crypto currencies are correlated with Bitcoin with respect to market movement. Since this plot covers data for the last 6 months it shows a strong correlation of Bitcoin with DOGE and Litecoin which were market movers recently. This graph is not the strongest representation of correlation as we know that Bitcoin is highly volatile and moves up and down with a high degree of variance.

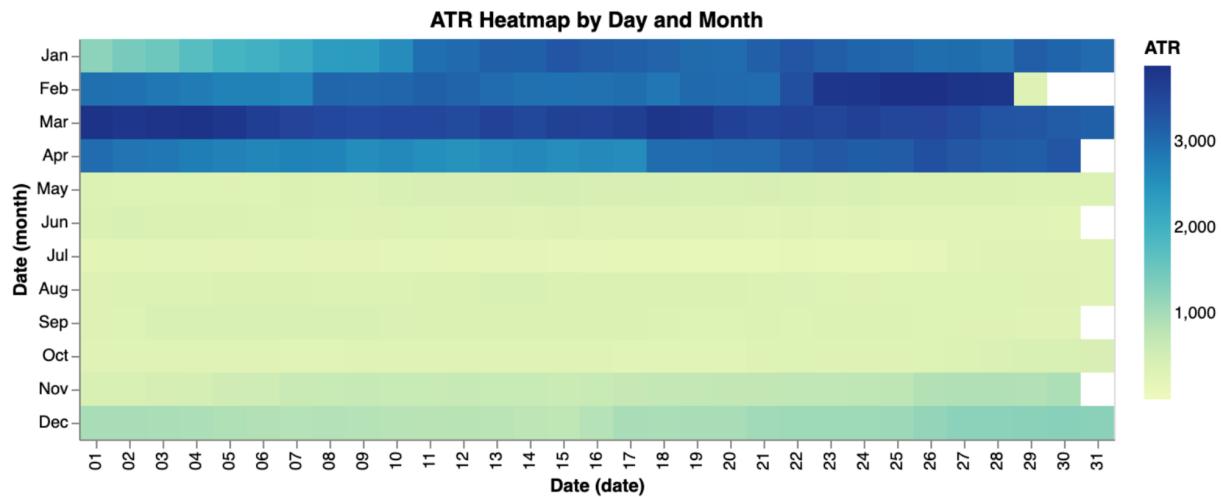
Conclusion:

Through incremental visualizations we are able to see that movement of principle alt coins are correlated for the most part with the movement of Bitcoin, although not to the degree in terms of value or variance, but over longer time periods we see that the whole market appears to be affected by the same factors. Discovering what those factors are requires a more in-depth analysis.

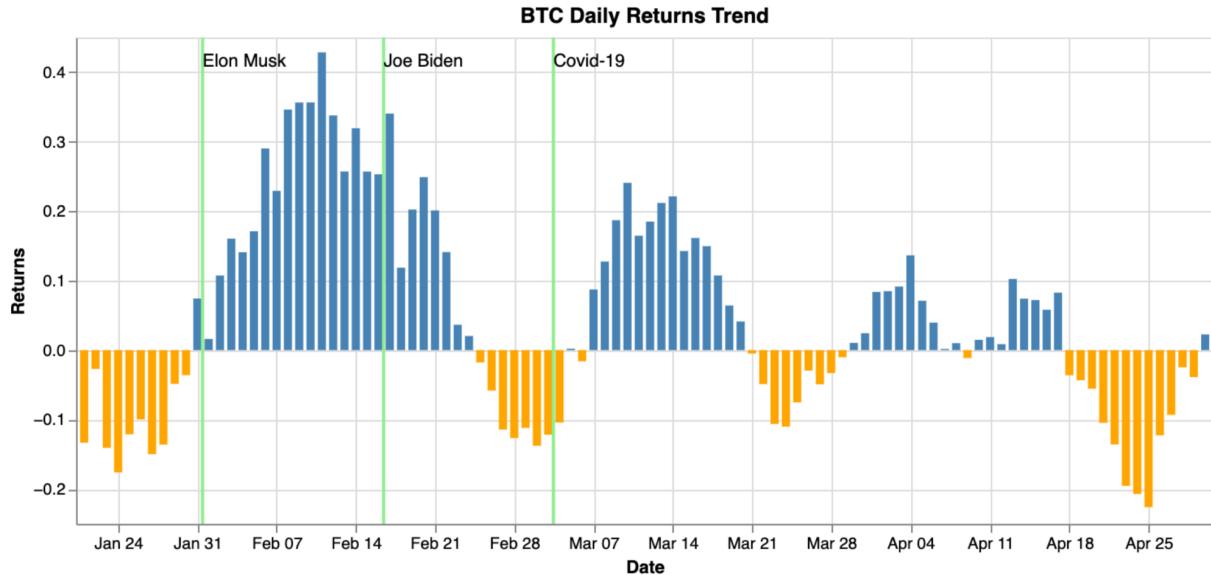
Hypothesis 3: Cryptocurrency market is highly influenced by non-market factors that don't affect traditional markets (ie. tweets, corporate investment, individual opinions, google searches).

What we are trying to discover here is whether or not the crypto market, being without regulation, is susceptible to factors that don't affect conventional markets like NASDAQ and NYSE. Tweets and comments made by Elon Musk have affected the overall crypto market, as well as the price of Bitcoin and the price and demand for Doge. Here we want to see just what factors have the biggest impact on movements in the crypto market, whether positive or negative.

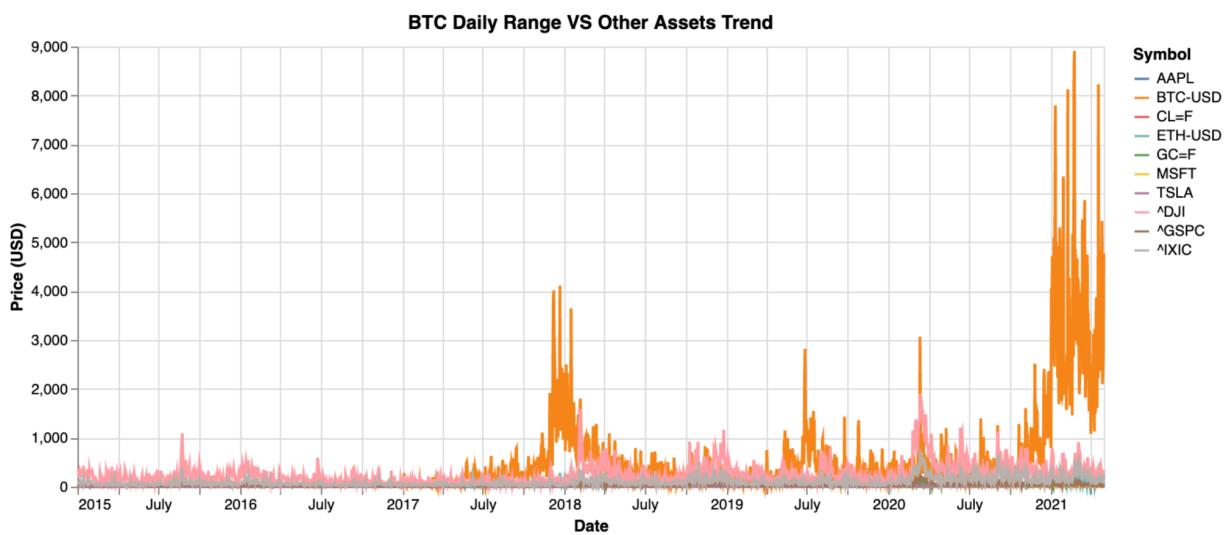
Exploratory Analysis:



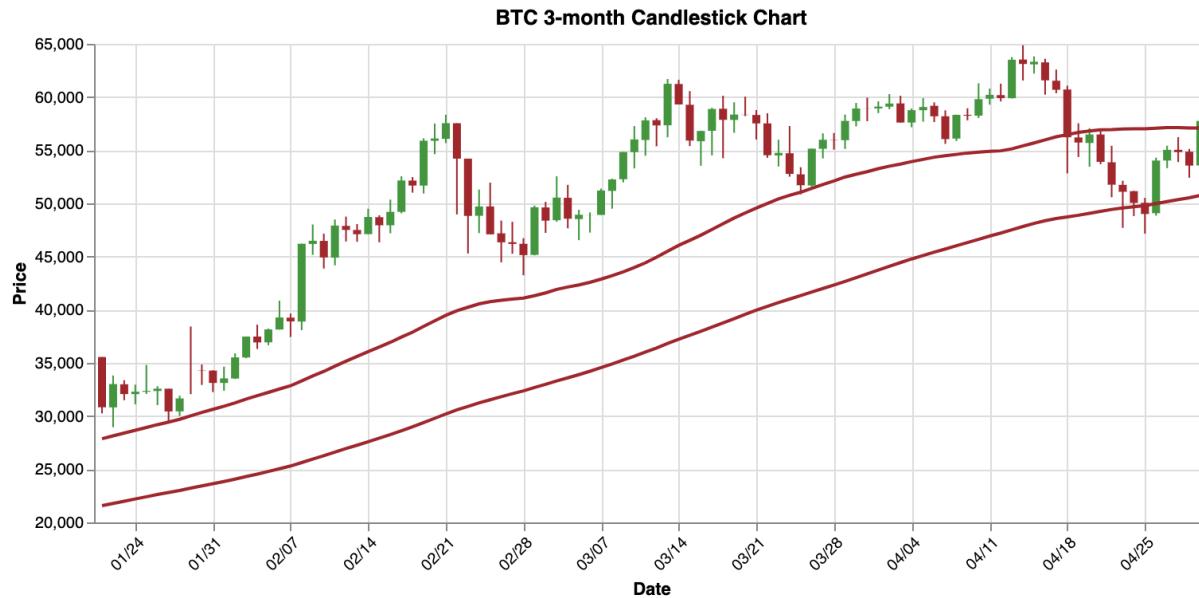
Here we have a heatmap, which in our notebook is interactive, we are able to move over the blocks and see details of each day in each month for the year 2020. We see that there is high ATR in the months January to April.



Working off our last visualization, we then plot Bitcoin daily returns to see how Bitcoin moves with some key events. Principally Elon Musk purchasing 100s of millions of Bitcoins and tweeting frequently about Bitcoin and Doge Coin. We see Bitcoin prices jumping when Elon Musk creates positive press and it continues through a period of time where the Biden administration decides not to impose any regulations in the crypto market. We see a dip in February which coincides with news that China is banning Bitcoin mining, then we see as COVID-19 vaccines reach a daily high and the US begins to return to normal that the value of Bitcoin makes another positive jump. We will need to explore further to understand the other movements.



This graph shows us fluctuations principally with Bitcoin, which we would like to dig into further and look at smaller windows of time to see what we can attribute to each movement.



Over a shorter 3 month time window we see the positive and negative movements of Bitcoin which as a part of our final visualization we will explore further to be able to plot influential events that seem to correlate with movements in Bitcoin.

Conclusion:

We are able to see that some events have a positive and negative impact on the price of Bitcoin and the overall crypto market. We saw from our previous sections in Hypothesis 1 and 2 that although alt coins move with Bitcoin, Bitcoins movement does not mirror the movement of traditional markets. So by consequence we can say that events that impact the movement of Bitcoin generally impact the entire crypto market as it is the 13,000 lb elephant and everything moves with it. However when compared to traditional markets, tweets by Elon Musk, regulatory announcements that only affect crypto and other crypto market-related news does not affect traditional markets. However we will pursue this hypothesis in further detail with our final project where we hope to get much more granular with respect to events.