Springboard Data Science Intensive Program – Capstone Project #1

Bitcoin & the Digital Gold Rush: Mining the data for Traditional Finance

Garrick Chu (2018)

Table of Contents

"Bull markets are born on pessimism, grow on skepticism, mature on optimism and die on euphoria."

- Sir John Templeton

1. Introduction

Bitcoin was created in 2009 out of a white-paper published under a nom-de-plume, Satoshi Nakamoto. Though the actual identity of the creator of Bitcoin and Satoshi are still unknown, the recent media attention and astronomic rise of Bitcoin are very real. Bitcoin and the underlying Blockchain technology have been both praised as a revolutionary decentralized ledger system and criticized as an outright scam. However, there is no denying that Bitcoin has begun to gain the attention of the public, mainstream media outlets and Wall Street. In recent months, news feeds and headlines have been filled with stories of rags-to-riches, statements from prominent Wall Street leaders and calls for regulation.

From Wall Street to Main Street, this begs the question: Does Bitcoin have a purpose in traditional finance? Do we define Bitcoin as a medium of exchange, a globally accepted currency, or a conduit for fraud and shady transactions? Is Bitcoin another asset bubble that will turn out badly as we have seen in recent past? And can we formally identify driving factors in the Bitcoin market, uncover patterns and make predictions?

Though an obvious question, the answers prove to be much more elusive. The search for answers has become more urgent as Wall Street and individual investors fear missing their golden opportunity but are equally cautious of being the “greater fool.”

I intend to take a data-driven approach using Data Science tools to uncover insights on these questions.

2. Data Acquisition

I acquired datasets from 4 online sources:

- Quandl.com for Bitcoin and other cryptocurrency market data
- Yahoo! Finance for stock market data (S&P 500 and Nasdaq Indices)
- Chicago Board of Exchange for Volatility Index (VIX) data
- Federal Bank of St. Louis for S&P/Case-Shiller Home Price Index data

Quandl’s Bitcoin specific data also contains technical characteristics such as Number of Users (on the My Wallet service), Market Price and Transaction data. This data set spans from 2009 to December 2017.

I also obtained market data on other Cryptocurrencies including Ripple, Litecoin, Ether and Iota from Quandl and spans from 2017 through January 2018. I chose these sets because they are the next biggest coins by market capitalization after Bitcoin.

For my analysis on whether Bitcoin is similar to previous bubbles, I used: S&P/Case-Shiller Home Price Index data from 1997 to 2006 (known as the US Housing Bubble), Nasdaq Composite Index prices from 1997 to 2001 (also known as the Dot-Com Bubble) and a smaller data set on the Tulip Price Index spanning from 1636 to 1642 (Tulip Mania). I choose these sets as they are the most ubiquitous asset bubbles in modern finance.

For the analysis of Bitcoin as a medium-of-exchange, I acquired BTC Volume data (across exchanges) and Daily Number of Transactions from data.bitcoinity.org.

3. Data Wrangling, Cleaning and Exploration

3a) Exploring the Bitcoin Market and Technical Data

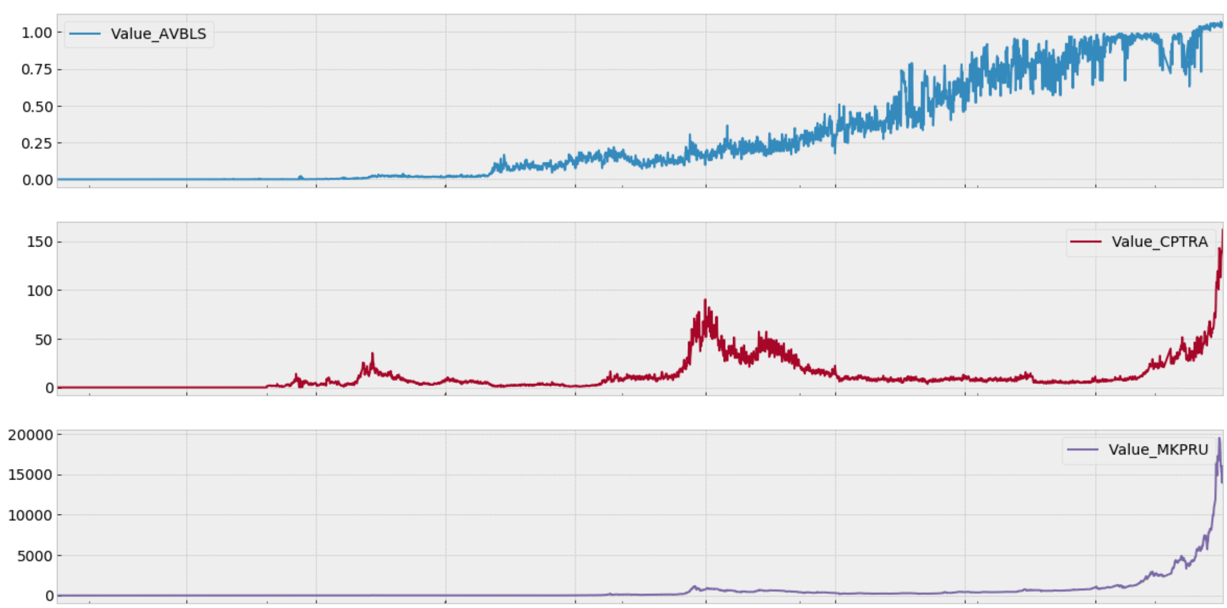
The Bitcoin dataset is an amalgamation of CSV files from Quandl with each variable/measure having its own file. Each file contained a date column and a measure column, where each date had its own row. Here, I used a ‘glob’ function along with the Pandas library to read in the CSV files and combine into a single dataframe sharing the same time-series index. The result is a Pandas Dataframe containing 10 columns and 3259 rows:

```
            Value_AVBLS  Value_CPTRA  Value_MKPRU  Value_MWNTD  Value_MWNUS  \
Date
2017-12-27         1.05       161.68     15999.04      45559.0   21272882.0
2017-12-26         1.05       146.59     14119.02      42867.0   21249422.0
2017-12-25         1.06       138.78     13949.17      49434.0   21204476.0
2017-12-24         1.06       139.69     15360.26      40957.0   21165559.0
2017-12-23         1.06       137.02     15190.94      43222.0   21100453.0

             Value_NADDU  Value_NTRAN  Value_NTRAT  Value_NTRBL   Value_TRVOU
Date
2017-12-27      605853.0     247440.0  286214316.0      1742.53  1.745062e+09
2017-12-26      565074.0     228926.0  285966876.0      1695.74  1.134113e+09
2017-12-25      652209.0     279523.0  285737950.0      1838.96  1.798168e+09
2017-12-24      729637.0     308211.0  285458427.0      2217.34  1.967537e+09
2017-12-23      890731.0     380648.0  285150216.0      2455.79  5.352016e+09
```

I begin by taking an initial view at the raw time-series data in Fig. 1:

Note: The 4th subplot from the top indicates we have 0 values for the “Number of My Wallet Users per day”. This could be due to 0 values in the data or periods of service outage. However, as I progressed in my analysis, I can omit this data series from further analysis.



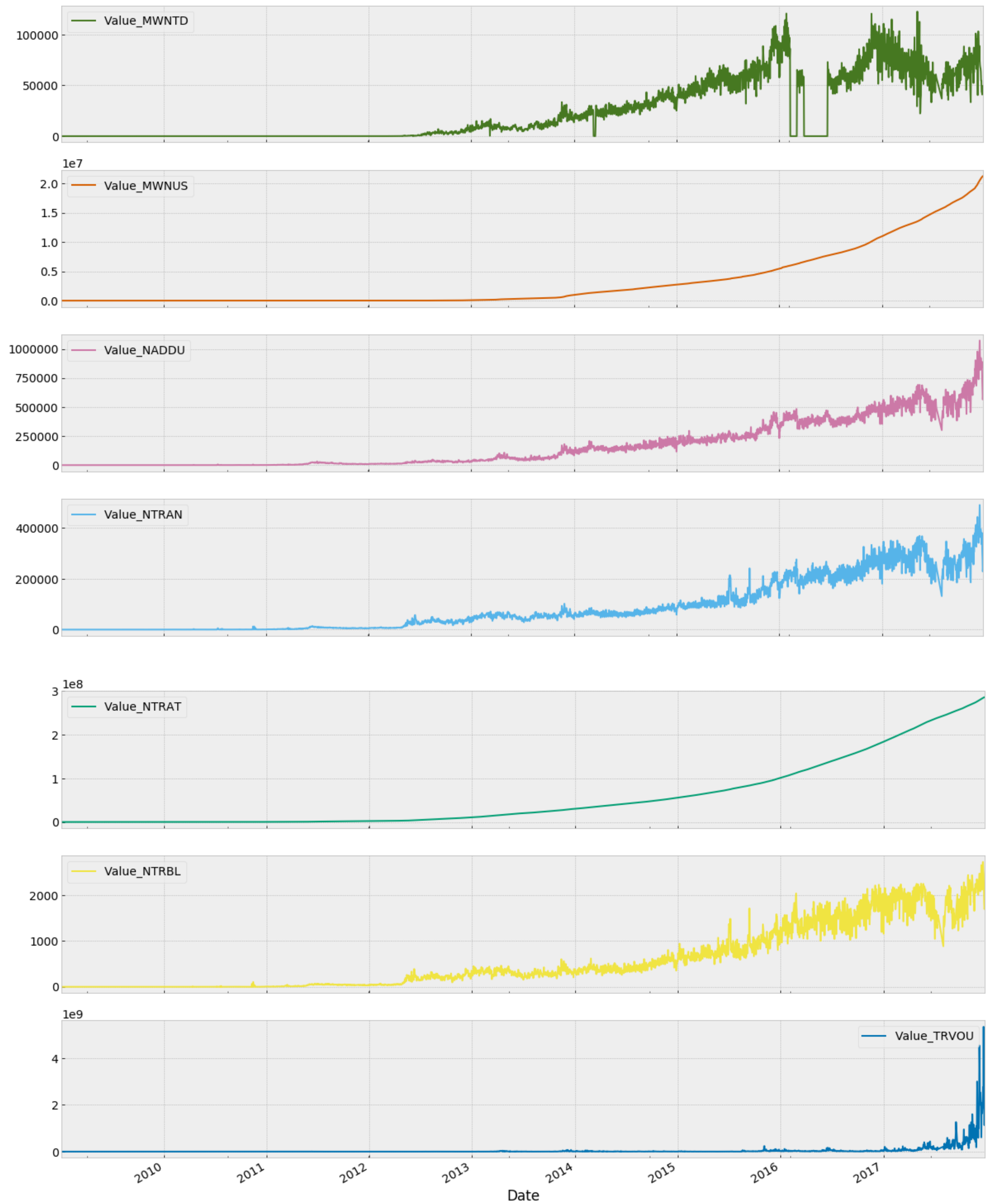
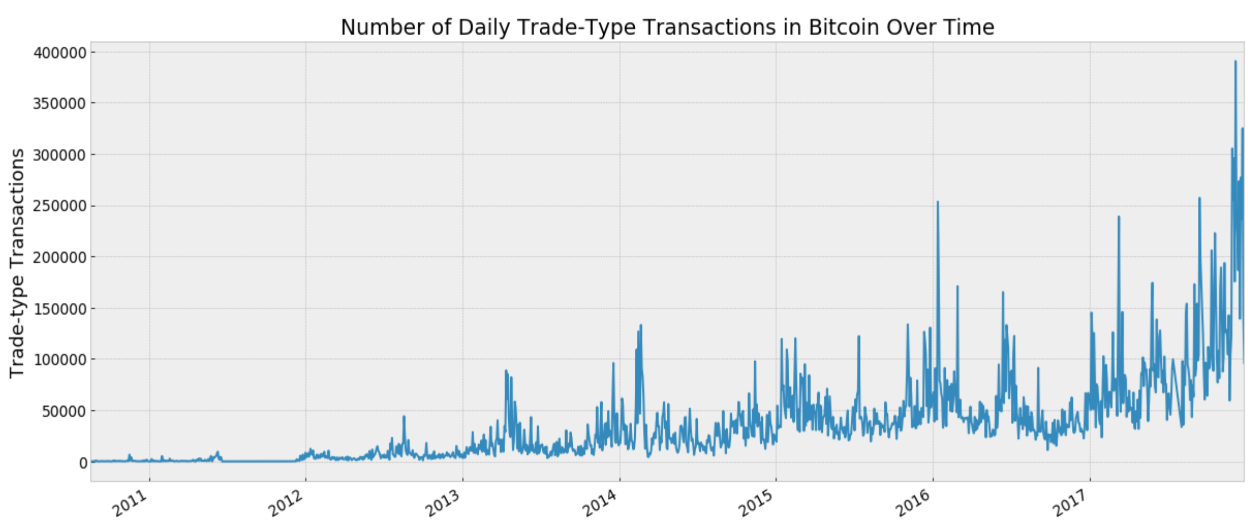


Fig 1 – Visual exploration of the Historical Bitcoin Data

3b) Assessing Bitcoin as Medium-of-Exchange or Storage-of-Value

At this stage, I wanted to assess if Bitcoin was increasingly being used as a medium of exchange. To accomplish this, I was inspired by a white paper from the Federal Reserve Bank of Boston (link here) from 2014. Part of the paper had set out to also determine if Bitcoin was increasingly becoming adopted as a medium-of-exchange. In their analysis, the authors calculated “Trade Volume to Transaction Volume Ratio” with the idea that of all transactions recorded to the blockchain are mostly comprised of either trade transactions (such as buying/selling on an exchange) or transfer transactions (such as paying for merchandise). By examining the ratio over time, this may be indicative of increasing or decreasing adoption as a medium of exchange. The Boston Fed paper only spans from 2012 to its year of publication, 2014.

For my analysis, I aim to replicate this measure armed with more time-series data. Here I loaded in my data from Bitcoinity.org, synthesized a new column summing up volume across exchanges, added in column “Average BTC per Transaction” from a previous calculation and used as denominator to output a new column, “Number of Trades.” We arrive at the below chart which plots the raw count of daily trade transactions over time. The key takeaway here is that trade transactions have been increasing over time with spikes in activity and has been notably elevated in 2017. I will be using the “Number of Daily Trade Transaction” time-series data as the numerator to recreate the “Trade Volume to Transaction Volume Ratio.”



Secondly, I loaded into the Number of Total Transactions data and added to my existing dataframe to calculate the “Ratio of Trades to Transactions” (calculated as a percentage of Total Transactions) and plotted as in Fig. 2:

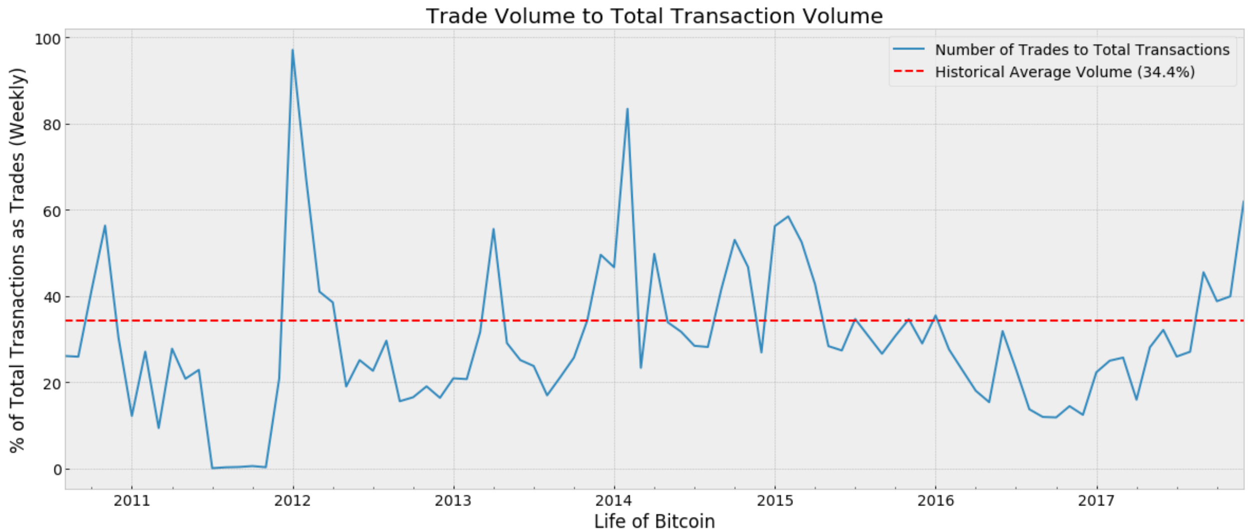


Fig. 2 – Historical Ratio of Trades to Total Transactions

Fig. 2 fails to present an obvious historical trend that suggests an increased rate of adoption of Bitcoin as a medium-of-exchange. If this had been the case, we would have seen the ratio decrease over time. Of note, we see a rise of Trade transactions in 2017 (peaking at double the historical average) which corresponds with the rally in Bitcoin prices.

3c) Finding "Stages" of Bitcoin

The second objective with this dataset was to explore if the measures can classify distinct "eras" of Bitcoin's life. That is, I wish to explore whether it is possible to classify Bitcoin's history into "phases" as characterized by the data and to classify future instances of Bitcoin as belonging to or being reminiscent of a prior "era." For this purpose, I with use a method of unsupervised machine learning called clustering. I chose this method given that the data is unlabeled and that I wish to find an unknown set of patterns in the data. Moreover, I am using the clustering method for classification as opposed to classification methods with known labels (such as SVM or Decision Trees).

In effort to prepare the data for the clustering model, I employed a Scaling function to normalize the data and eliminate the potential distortion of scale in the data. This is followed by applying Principal Component Analysis (PCA) to distill the most descriptive components within the data to cluster by. Finally, I used two Clustering methods, K-Means and Spectral Clustering and compare the results. Spectral Clustering resulted in overfitting our data and thus will highlight only the K-Means results as follows.

In the initial data preparation step, the scaler function normalizes the data and takes the unit-measure element out of each data series:

0	1	2	3	4	5	6	7	8	9
2.021026	8.454132	8.069559	3.365756	2.028827	1.397554	2.75473	1.533849	6.11432	-0.240011
2.050119	7.141669	6.987264	3.352554	2.259346	1.712912	2.748698	1.67529	6.306733	-0.249939
2.050119	7.040798	7.642896	3.332476	3.445465	2.706911	2.741255	2.580043	19.182948	-0.253659
2.079211	5.665864	8.083936	3.309207	3.262128	2.60497	2.732186	2.971424	9.765444	-0.250324
1.991934	7.077478	9.63387	3.278738	3.60356	2.840503	2.722703	2.155191	7.114602	-0.251639

In the following step, I employed PCA to reduce the number of dimensions (from our 10 variables) and found that I could reduce our data to just two variables/dimensions. I suspect the two variables to be Bitcoin's price and volume. I determined this by choosing the value in which explains the majority of the variance across all the data (Fig. 3).

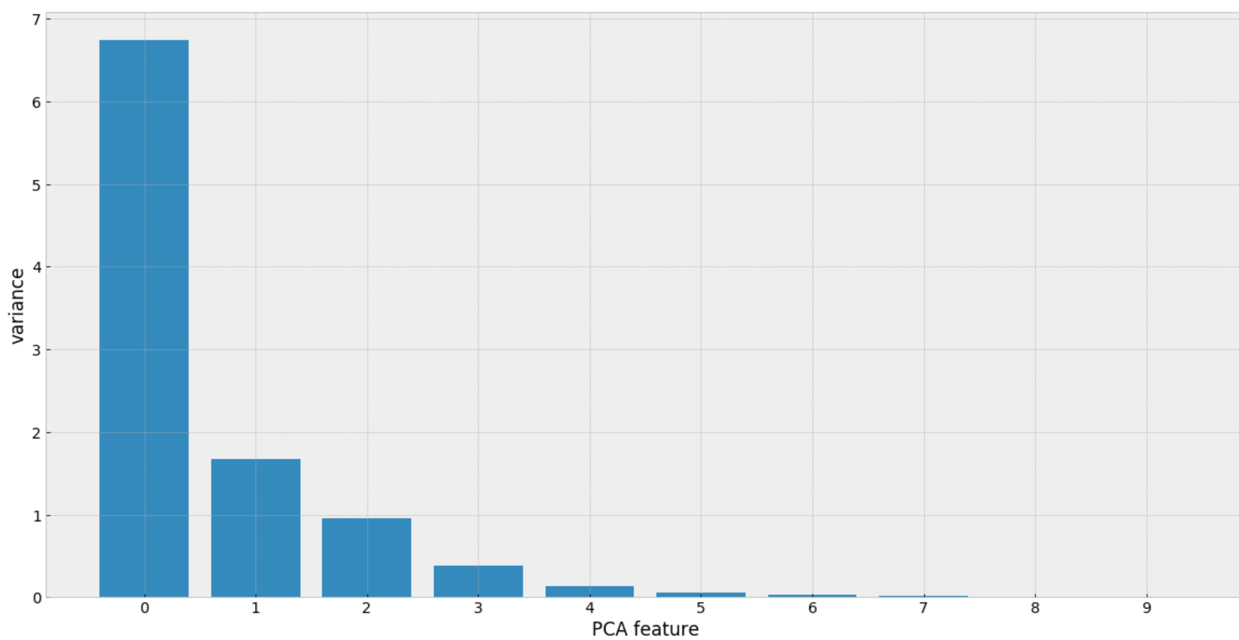


Fig. 3 – # of Principal Component Features vs. Explained Variance.

Given these two PCA features, we assess an ideal number of clusters in which to fit the data. Here, I chose 5 clusters using the "elbow-method" in Fig. 4 where the slope begins to flatten out.

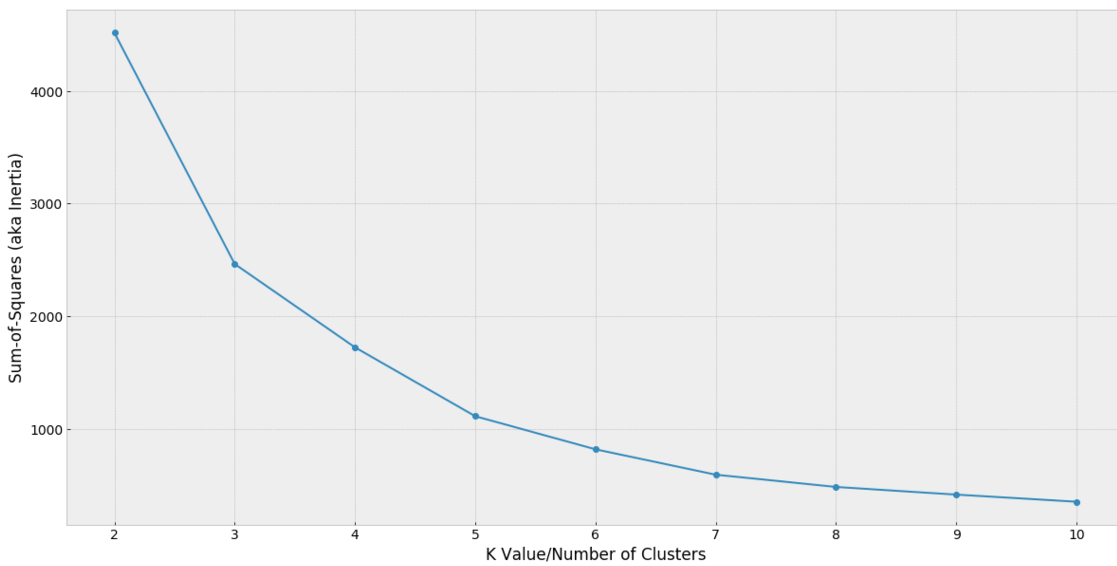


Fig. 4 – Number of Clusters in K-Means Model vs. Sum-of-Squares (Inertia).

In applying this is our optimized K-Means classification model, we can visualize the resultant clusters in a scatter plot of Bitcoin Prices to Bitcoin Trade Volume (Fig. 5).

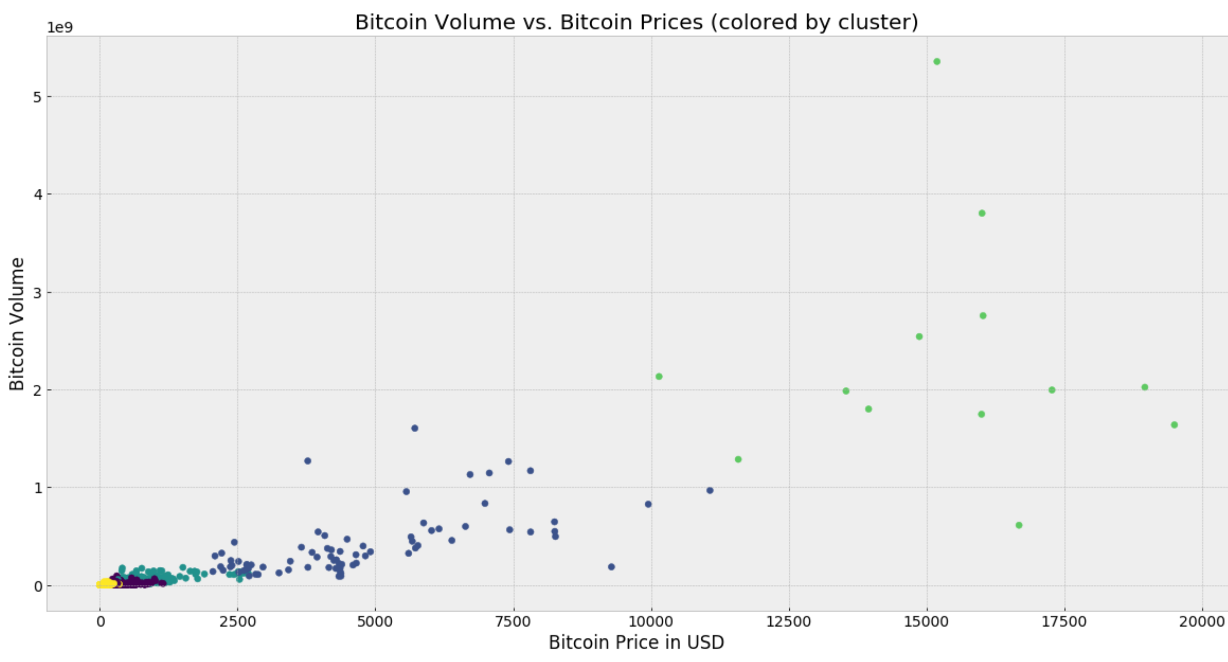


Fig. 5 – Scatterplot displaying Bitcoin Prices and Volume, data plots colored by cluster classification.

In a scatterplot of Bitcoin Prices and Volume, the clusters appear discrete and distinct just as I had suspected. Near the origin, a cluster may be defined as Early Bitcoin when volume and prices were low. On the other end, we see a cluster defined by high prices and high volumes.

Looking at Bitcoin's history with distinct time periods (determined by cluster), we see the Bitcoin has gone through "eras" as I had initially hope to find (Fig. 6). I made an effort to characterize and explain these clusters in further detail below.

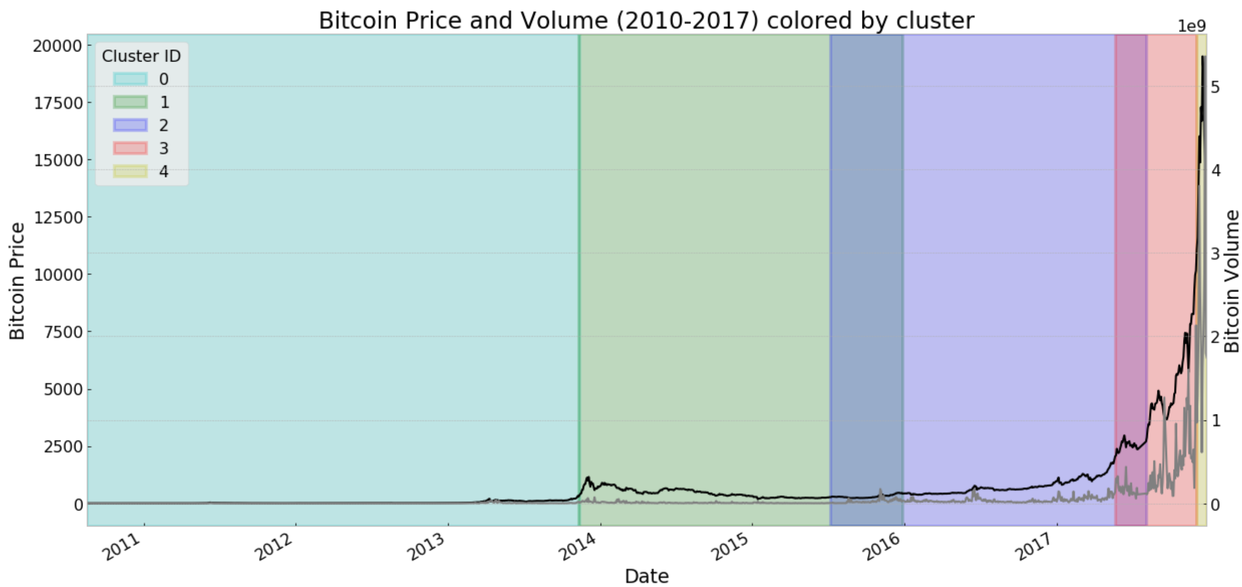


Fig. 6 – Historical Bitcoin Price and Volume with date-ranges color-coded with respective cluster

- The Cyan-shaded period is Bitcoin's early growth stage where the market dynamics of Bitcoin remained largely homogenous. The end of the yellow period and the start of the Cyan period is marked by news of retailers such as Overstock, Newegg and Dell beginning to accept Bitcoin as method of payment.
- The Green period from late-2013 to late 2015 is characterized as when Coinbase raised Series C Funding of $75 million and 21 Inc. raises $116 million (the largest of any amount for digital currency companies).
- The Purple section is when academic research around Bitcoin increases and more businesses begin to accept Bitcoin (including major online gaming platform, Steam). In late 2016, the Chinese Renminbi depreciates against the USD, correlating with a surge in price from $600 to $780 USD.
- The Red section is the pre-cursor to the December 2017 rally and when Bitcoin gains legitimacy among lawmakers and legacy financial institutions. This period saw early indications of volatility as several price records were broken.
- The last Yellow section is solely December 2017 and that it is a standalone cluster may be indicative that in this era of time, Bitcoin was/is in an asset bubble (and is fueled by the "greater fool theory"[1]).
- I suspect the periods where the clusters overlap are periods of profit-taking where long-term holders liquidated their stakes and contributed to elevated price/volume activity.

[1]The **greater fool theory** states that the price of an object is determined not by its intrinsic value, but rather by irrational beliefs and expectations of market participants. A price can be justified by a rational buyer under the belief that another party is willing to pay an even higher price.

3d) Bitcoin as a Portfolio Management Tool

As with the Bitcoin data, I acquired time-series data from different sources. I utilized Pandas to load in the various CSV files and constructed/combined into dataframes for analysis.

With the analysis of Bitcoin as a source of diversification/portfolio management, I needed to further process static values of Bitcoin and traditional assets to find their respective returns. The main tenant of diversification is to create a portfolio of assets with uncorrelated returns. To get this, I created columns for each asset's "total return" (indexed to the first observation in their respective data series). Lastly, I determined the Pearson Coefficients of each asset's percentage return which will show the strength of the correlation between each asset's return. This results in the below table:

```
             BTC_TR    VIX_TR  SP500_TR  Gold_Return  Oil_Return
BTC_TR      1.000000 -0.316832  0.542559    -0.334866   -0.250007
VIX_TR     -0.316832  1.000000 -0.618065     0.653232    0.205093
SP500_TR    0.542559 -0.618065  1.000000    -0.785429   -0.730141
Gold_Return -0.334866  0.653232 -0.785429     1.000000    0.724267
Oil_Return -0.250007  0.205093 -0.730141     0.724267    1.000000
```

And visualized with a heatmap:

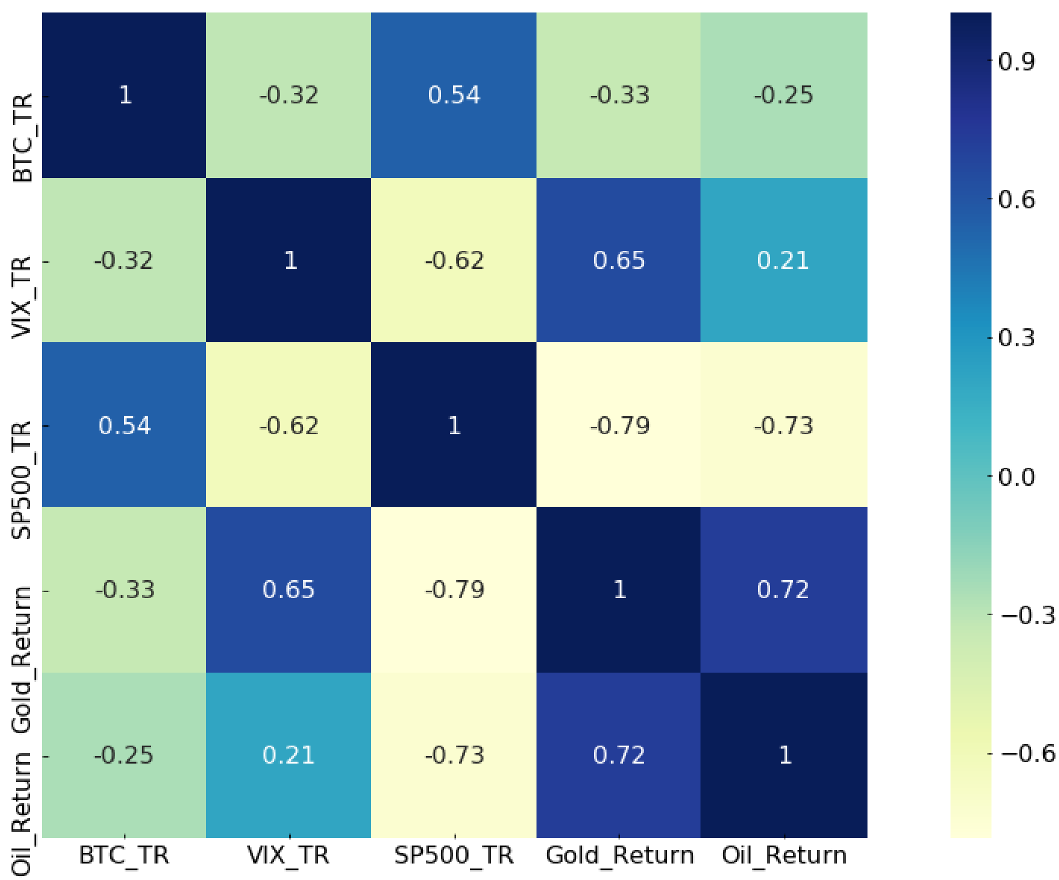


Fig. 5 – Correlation Coefficient Matrix Heatmap for the Total Return for: Bitcoin (BTC_TR), Volatility Index (VIX_TR), S&P 500 Index (SP500_TR), Gold (Gold_Return), and Oil (Oil_Return).

In Fig. 5's heatmap, the top row illustrates the correlations between Bitcoin's returns and the returns of other assets. As the coefficients range from -0.33 to 0.54, this suggests that Bitcoin's returns are not strongly correlated with the returns of Gold, Oil, the Volatility Index and the broader stock market. Thus, one can infer that a traditional portfolio can benefit from diversification by incorporating Bitcoin.

For additional insight, I wanted to view how differently a traditional portfolio has performed since 2011 vs. a portfolio with 5%, 10% and 25% allocated to Bitcoin (Fig. 6)

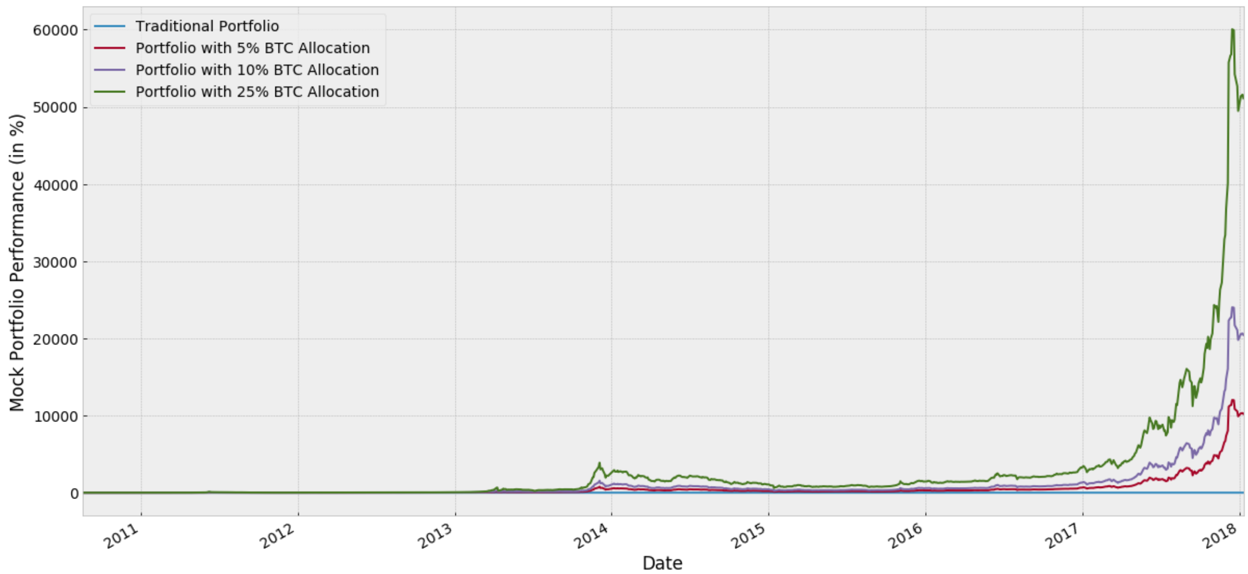


Fig. 6 – Comparison of Hypothetical Portfolio Returns

We find that a portfolio with just 5% of total capital invested in Bitcoin would have outperformed the standard portfolio by nearly 10,000%. In the case of a 10% allocation to Bitcoin, that portfolio would have outperformed by 20,000%. Lastly with a 25% allocation, the number jumps to nearly 50,000%.

An outperformance of 50,000% in one's portfolio is certainly both impressive and desirable but does not take "risk" into account. I refer to "risk" within textbook finance terminology as "variation in expected return". Using an analogy, an asset that has returns between -2% and 2% is seen as less risky than an asset which has returns between 25% and 50%.

An appropriate measure to compare assets would be the Sharpe Ratio. The Sharpe Ratio characterizes how well the return of an asset compensates the investor for the risk taken. It can also be thought of as an assets excess return for every additional unit of volatility. Generally, values of 1 or higher are good and values of 3 or higher are most desirable.

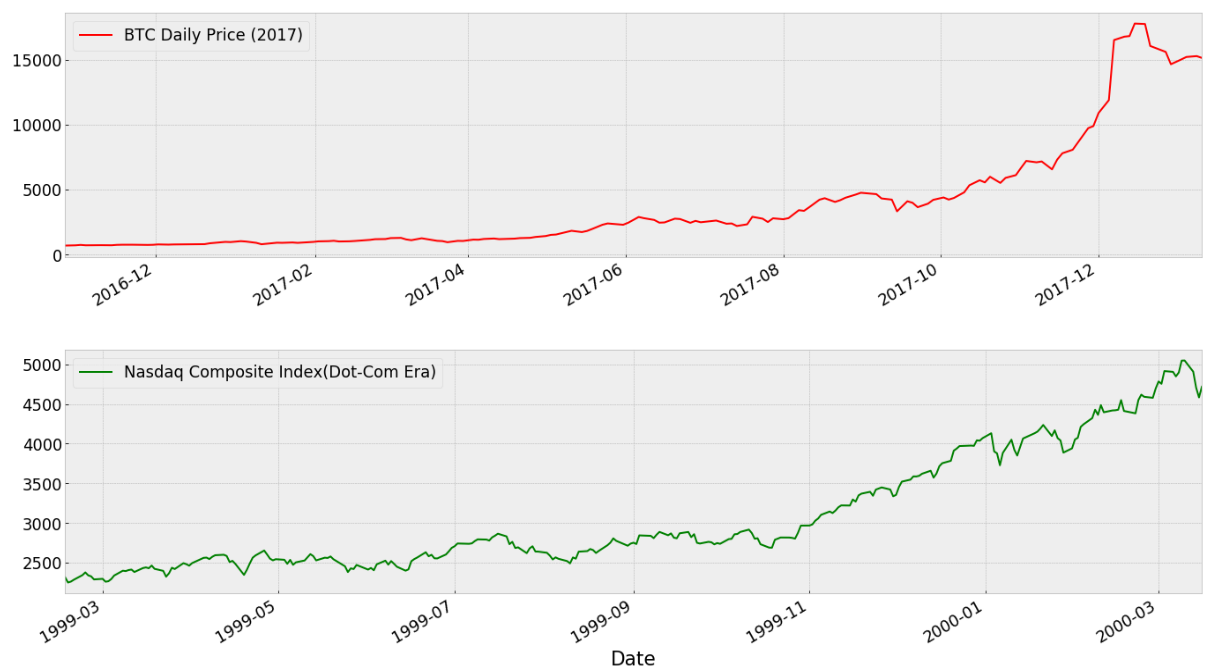
Using the same dataframe, I computed Excess Return as the Total Return of Bitcoin less the Total Return of the S&P 500 and return the values as its own series. Using the formula for Sharpe Ratio as:

$$S = \left(\frac{R_p - R_f}{\sigma_p}\right)$$

I computed the mean excess return and divided by the standard deviation of Bitcoin's Total Return and obtained a **Sharpe Ratio of 0.37**. Therefore, adding Bitcoin to a portfolio adds diversification but does not adequately compensate for taking on the additional risk.

3e) Bitcoin and Asset Bubble Analysis

The second objective in this area was to explore the similarities, if any, between Bitcoin and the most widely recognized asset bubbles. Here I used the S&P/Case-Shiller National Home Price Index from 1996 to 2007 (widely known as the US Housing Bubble timeline), Tulip Price Index Data from 1634 – 1642, Nasdaq Composite Index data from a 12-month period leading up to the March 10, 2000 peak. The Housing data uses a different time scale as the data is recorded monthly and is subject to slower market dynamics and price discovery mechanisms than that of equities or Bitcoin and thus a longer time-series is needed for an appropriate comparison. With respect to the Tulip Index, the data contains only 14 observations and thus I favored more data points with an inconsistent timescale over matching the time scale of the other data series.



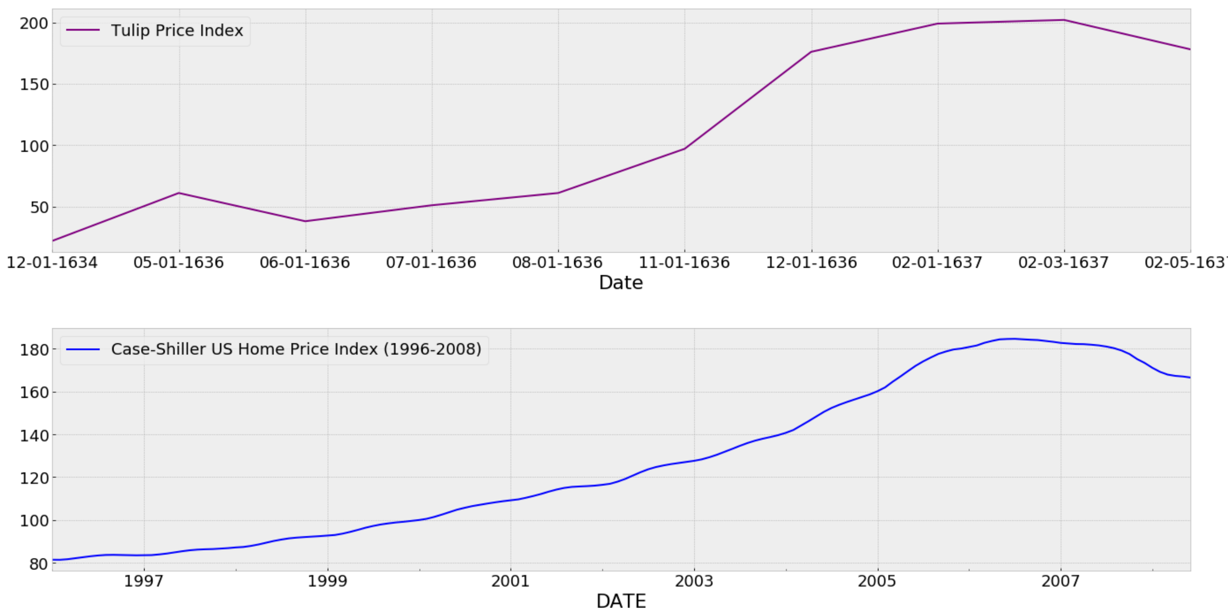


Fig. 7 – Comparison of Bitcoin and Asset Bubbles. Bitcoin Prices (2017), Nasdaq Composite Index (1997-2001), Tulip Prices in Netherlands (1634-1642), Case-Shiller Home Price Index (1997-2007).

The plotted price data among all 4 plots show strikingly similar trends but requires further analysis to confirm. In the case of the Tulip data, Pandas uses a 64-bit integer for its Time-Series functionality and is limited in scope to years 1677 to 2220. As my Tulip data is out of this range, I was unable to manipulate the data for further analysis and had to exclude from the next analytical step.

To quantify the above visual plots, I utilized statistical analysis again to determine Pearson Coefficients between the Bitcoin, Nasdaq and Housing data. Additionally, I computed the R-Squared values to quantify the “fit” of the correlation within the data itself. With respect to the Nasdaq data, I resampled the data using the data’s weekly mean to ensure the data was of equal size. After processing the data, I obtained the following results:

```
Pearson R correlation coefficient for Bitcoin (2017) and the Dot-Com Bubble:
0.8527488585004591
R-squared of BTC and Dot-Com Bubble: 0. 727180615674

Pearson R correlation coefficient for Bitcoin (2017) and US Housing Bubble:
0.7525625687801886
R-squared of BTC and US Housing Bubble: 0.566350419929
```

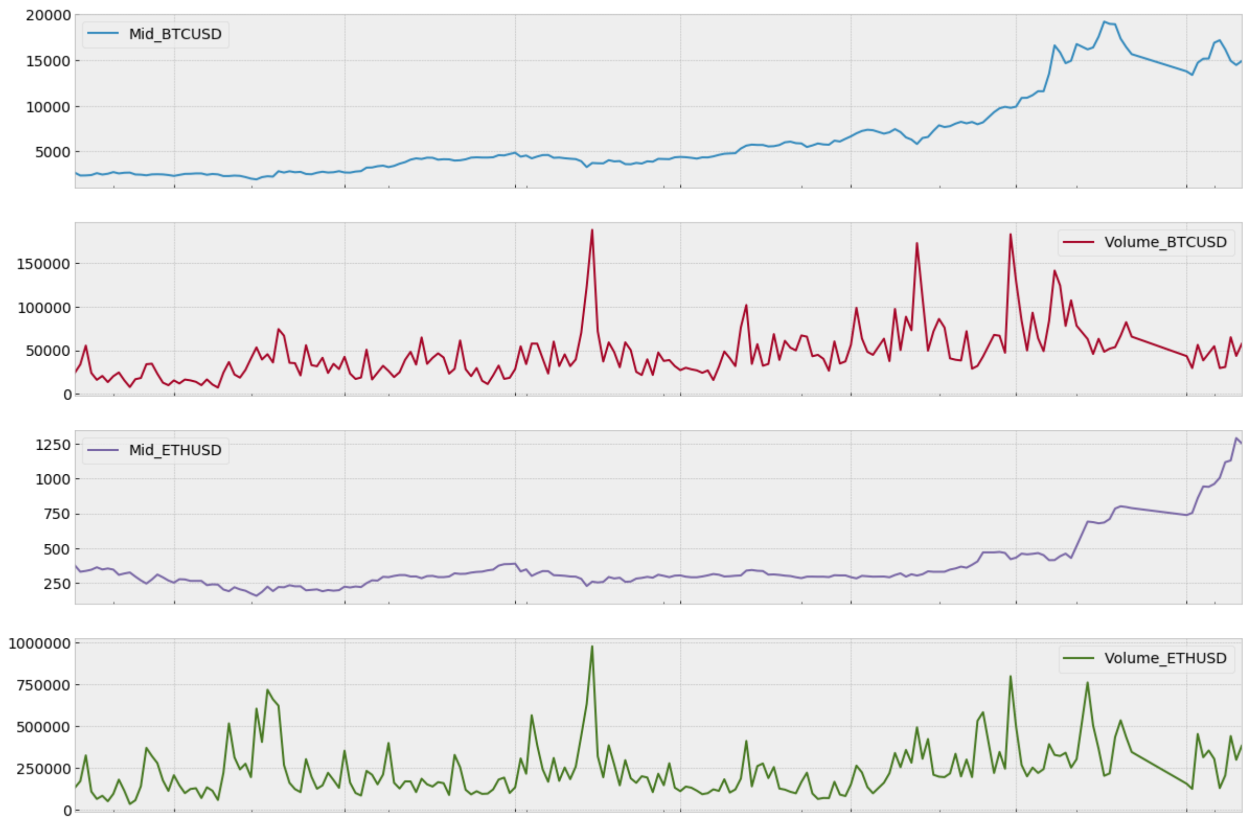
Given a strong correlation of 0.85 paired with an assumed fitted-line which explains 72% of the variance in the data, one can infer that Bitcoin is experiencing an economic bubble very similar to the Dot-Com Bubble. In the 30 months following March 2000 peak, the Nasdaq Index fell 78%. Similarly, Bitcoin lost nearly 50% of its value shortly after this data was collected in late January 2018.

The US Housing Bubble appears to also have a strong correlation with Bitcoin and is further evidence that Bitcoin is experience an economic bubble. However, this correlation is not as robust as with Bitcoin and the Dot-Com Bubble.

3f) Comparison of Bitcoin and Biggest Alternative Coins (Ethereum, Litecoin, Ripple and Iota) and Predicting Bitcoin Prices

Like the Bitcoin data, I acquired the data on the other larger cryptocurrencies from Quandl. Each file contains time-series data from 2017 through January 2018 and High, Low, Mid and Closing prices in addition to Daily Volume for each coin. Iota was introduced in 2017 and thus the combined/merged dataframe excludes data pre-June 2017. As a result, the dataframe for this analysis had 35 columns and 201 observations (days).

For a simple visualization, I plotted the Mid-Daily Price of each coin and their volume over this time period:



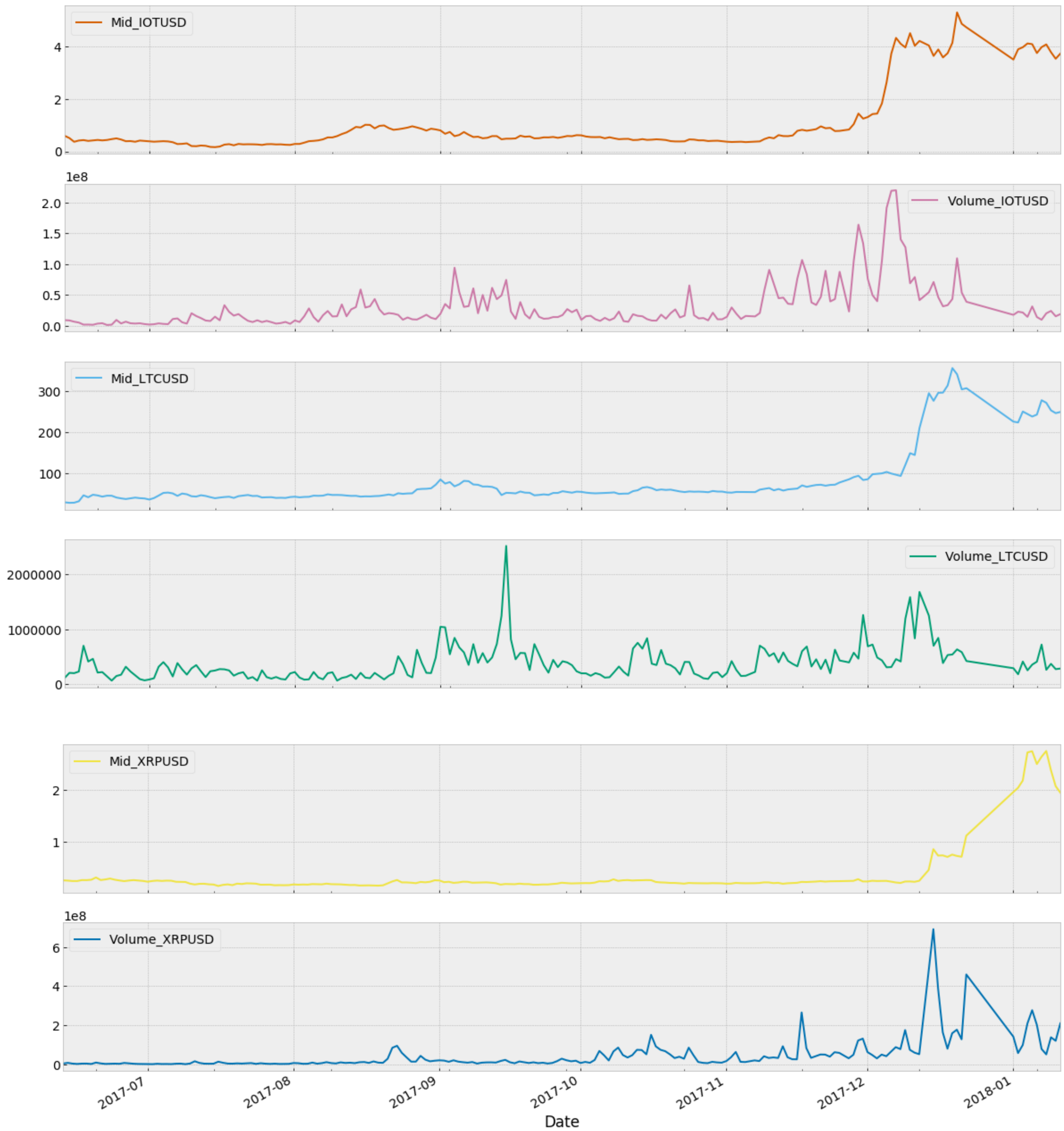


Fig. 8 – Average Daily Prices of Coins and respective volumes. From Top to Bottom: Bitcoin, Ethereum, Iota, Litecoin and Ripple.

Despite a relatively short time period of less than a year, we see some similarities in price trends in November – January. Furthermore, one can see more similarities in the volumes of Bitcoin, Ethereum and Litecoin (but in neither the Ripple nor Iota data).

I also investigated these potential relationships using a Scatter Matrix:

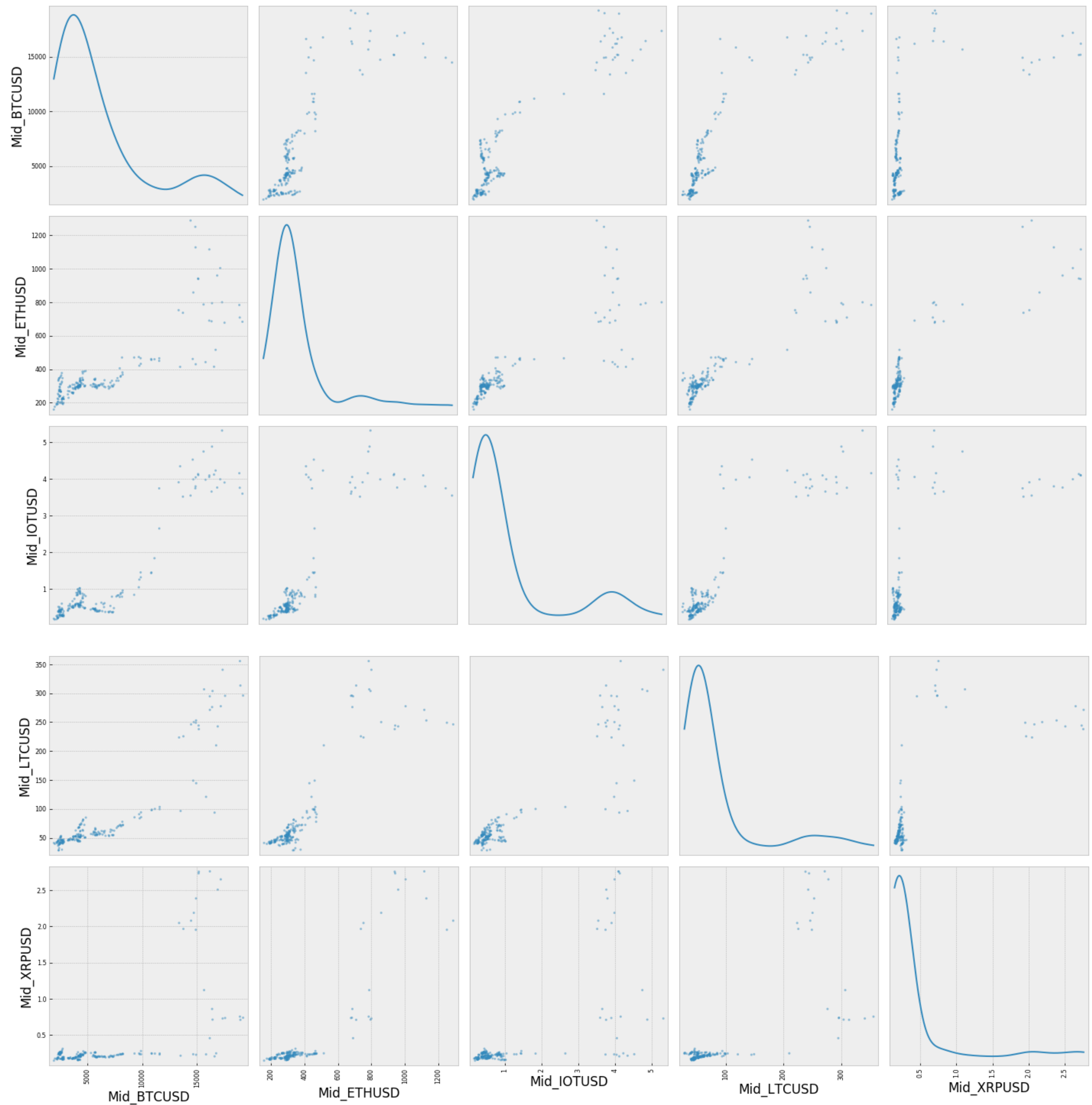


Fig 9 – Scatterplot Matrix of correlations between coins and their Daily Mid Prices.

While this visualization was not helpful to me, a closer look at their correlations coefficients uncover a high degree of correlation that was not initially obvious to me:

	Mid_BTCUSD	Mid_ETHUSD	Mid_IOTUSD	Mid_LTCUSD	Mid_XRPUSD
Mid_BTCUSD	1.000000	0.831611	0.921704	0.899781	0.625674
Mid_ETHUSD	0.831611	1.000000	0.840506	0.891562	0.879821
Mid_IOTUSD	0.921704	0.840506	1.000000	0.906808	0.672821
Mid_LTCUSD	0.899781	0.891562	0.906808	1.000000	0.737545
Mid_XRPUSD	0.625674	0.879821	0.672821	0.737545	1.000000

Moreover, I plotted the above correlations in a correlation heatmap as a visual aid in Fig. 10:

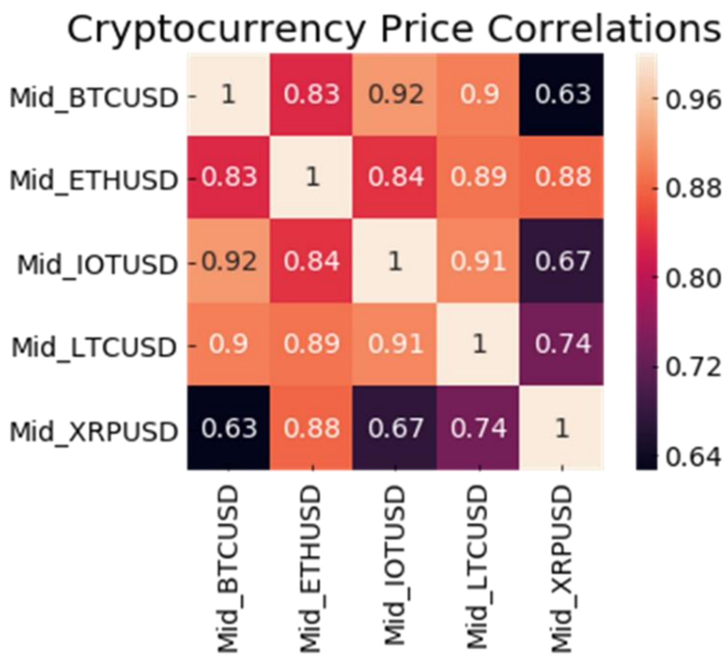


Fig 10- Correlation Heatmap of Daily Mid Prices of Coins. We see strong correlations among many of the covariates.

This was a similar case with volumes:

	Volume_BTCUSD	Volume_ETHUSD	Volume_IOTUSD	Volume_LTCUSD	Volume_XRPUSD
Volume_BTCUSD	1.000000	0.618239	0.556679	0.629883	0.262252
Volume_ETHUSD	0.618239	1.000000	0.392737	0.558947	0.344295
Volume_IOTUSD	0.556679	0.392737	1.000000	0.392757	0.290177
Volume_LTCUSD	0.629883	0.558947	0.392757	1.000000	0.296894
Volume_XRPUSD	0.262252	0.344295	0.290177	0.296894	1.000000

I suspect these covariate correlations were substantial enough use them in a model for predicting Bitcoin prices. With the objective of estimating Bitcoin prices (a continuous variable) given the price and volume of other major coins, I concluded using a Linear Regression model would be the best method. Here I used both a standard Linear Regression model (all covariates equal-weighted but pre-processed using a scaler function) and a Lasso model (allowing the model to apply weightings to the covariates which have also been pre-processed using a scaler function).

After normalizing the data and further splitting into a training and test set, I utilized fitted both Lasso and Linear Regression model. I found that both models determined similar weightings as illustrated in Fig. 11. Of note, Litecoin prices and Bitcoin volumes were overweight while Ripple prices and Ethereum Volumes were underweight:

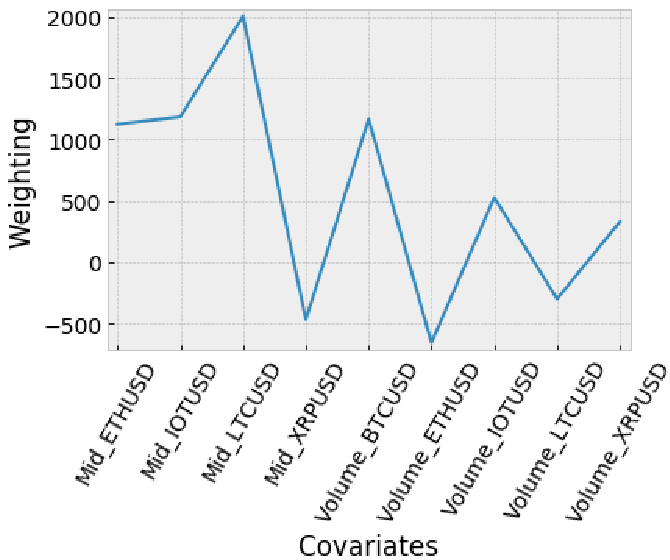


Fig. 11 – Covariates vs. Model determined weightings (both Lasso and Linear Regression).

With the Lasso model, I achieved and accuracy score of 89.39%. With the Linear Regression model, I achieved a similar accuracy of 89.37%. In using newer and unseen data, the model overestimate's Bitcoin prices consistently by roughly $3,500. Using "lagged" variables (shifting the predictor values into prior time periods), the model's accuracy to predicting values one-day ahead decreased to 70.2%. In predicting values one-month ahead, the accuracy diminished even further to 40.2%.

Lastly, I experimented with Facebook's recently released forecasting tool, Prophet. The tool aims to be a more flexible additive model to account for seasonality and changes in trend. I found that even this tool, could not accurately capture Bitcoin's historical price and volatility.

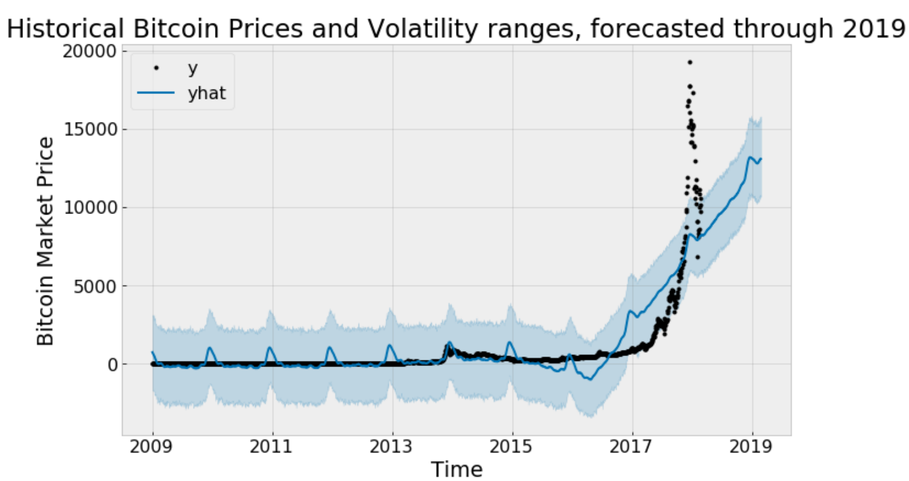


Fig 12 – Prophet's Historical and Forecast of Bitcoin Prices vs. Actual

4. Limitations and further research

4a) Technical Limitations

Timestamp Limitations

Although minor in the broader analysis, I faced a technical limitation with Pandas and its Datetime functionality. This proved an obstacle with making the comparison with Bitcoin 2017 prices and the Tulip price data as the data falls out of range. Since pandas represents timestamps in nanosecond resolution, the time span that can be represented using a 64-bit integer is limited to approximately 584 years:

```
In [66]: pd.Timestamp.min
Out[66]: Timestamp('1677-09-21 00:12:43.145225')

In [67]: pd.Timestamp.max
Out[67]: Timestamp('2262-04-11 23:47:16.854775807')
```

Productization Limitations

I initially set out to analyze certain asset classes and events for demonstrative purposes and did not consider farther-reaching applications of the structure and code I used for this project. Therefore, there is opportunity for implementing a product which can take in user/client data for their own economic bubble analysis or predictive modeling.

4b) Data Limitations

The tulip data itself was also a concern as I pulled 14 data points of inconsistent frequency from an academic paper (a few sparse monthly data points and then a grouping of daily prices). I could not locate the actual index itself. This may be worth further exploration to conduct a proper comparison.

The Bitcoin data is also only from the My Wallet Users and may not be an accurate representation of the broader population of market participants.

The Bitcoin and cryptocurrency data was also at such a scale that made visual analysis a challenge. For example, attempting to visualize Bitcoin in 2010 (at $0.07) and in 2017 (at $14,165.57) on the same plot.

4c) Further research

The Quandl Bitcoin data used for my analysis is only a sub-section of all the data available on Bitcoin metrics. There may be other features worth exploring to improve the Bitcoin Price Regression models.

I also suspect other features exogenous to Bitcoin and the Blockchain may also improve this model including counts of articles in major publications, frequency of tweets or trending cryptocurrency hashtags, instances of political or economic events, or other capital markets metrics.

5. Recommendations and Actions

Bitcoin can be a source of diversification for a risk-tolerant traditional portfolio but may not be worth the added risk.

In my analysis, Bitcoin can provide added diversification in a traditional portfolio as it historically had returns uncorrelated with traditional asset classes that make up a standard portfolio (such as a composition of equities and commodities). We also found that a portfolio gains only the added benefit of diversification but does substantially benefit the return of a portfolio. Recall that we found the Sharpe Ratio of Bitcoin to be 0.37. Therefore, it is not a risk-reward trade-off a majority of investors are willing to accept.

Bitcoin is presently in treacherous territory, likely seeing in "Greater Fool" market dynamics.

My analysis supports the idea that recent (read: December 2017 to present) Bitcoin activity is entirely distinctive from the rest of Bitcoin's life. This current distinctive stage corresponds with volatile price movements and unprecedented media attention surrounding Bitcoin and other cryptocurrencies. The model further suggests that we remain in this life-cycle stage despite the massive January 2018 correction.

Diversifying digital asset portfolio with other cryptocurrencies is doubtful.

We see a level of correlation among Bitcoin and other coins such that achieving diversification may be very difficult. Rather, I propose treating Bitcoin and other coins as a homogenous digital asset class and supplementing with a portfolio of traditional asset classes (taking into consideration one's risk tolerance and investment horizon). In other words, buying a variety of different coins isn't diversification but adding Bitcoin and other coins to a portfolio with traditional classes can be beneficial.

Bitcoin is likely a storage-of-value than a medium of exchange.

From an asset-class perspective, Bitcoin should be treated as an investment vehicle and not currency. Would not recommend using Bitcoin for purposes of Cross-Currency exchange (for portfolio management/hedging) and better suited as a holding.

Similar to traditional investments, Bitcoin prices and returns are cannot be consistently modeled and predicted.

Using cryptocurrency trading data as predictors/features, I found that modeling and predicting near-term and future Bitcoin prices to be difficult and unfeasible. Although I'm able to somewhat predict real-time values of Bitcoin based on these features, the real value of such a model would be to predict direction and trends in the future (as opposed to just intra-day).