The Roadrunners of Wall St

Mike Mahoney - 01701652

Surendra Thapa - 01715340

****

**Background**

          Is Cryptocurrency the next booming opportunity or the next looming bubble?  Is it a Ponzi scheme built on a house of cards or the next big thing to hit the financial markets?  Although these are interesting and tough questions, we cannot directly answer, we can speculate on where the market is going.  So, what is cryptocurrency really? How is it mined?  For explanation we will focus on the blockchain currency, in particular Bitcoin.  Bitcoin was created as a peer to peer currency system designed to circumnavigate traditional Central Banks (like the Federal Reserve and Bank of England).  It was designed to be an alternative to trading in fiat currency, instead using an untraceable currency that anyone could trade under the guise on anonymity.

          Bitcoin in particular is mined by “verifying transactions”, by cracking complex cryptographic mathematical computational problems, on a master ledger called blockchains.  Each member of the Bitcoin community has access to the master ledger (blockchains), and with the right software (and enough computational power and graphics cards) and can verify these transactions.   Each entry on the ledger is encrypted with a 256-bit algorithm, which is called the SHA-256 Hash (developed by the NSA).  Also, each member has their own Private Key and Public Key, which is used with the SHA-256 Hash which creates a unique non-replicable (in theory) encrypted entry on the blockchain.  Now each entry on this master ledger is added to the block and when the block is full, it is chained to the succeeding block which forms a chain, thus why its called a blockchain.

          Now for an example, let’s say Jack and Dave decide to play poker.  Neither Jack nor Dave brought any fiat currency, and both decide they will assess the winner at the end.  So, at the end of each hand both Jack and Dave keep their own ledger to keep track of the money.  Now, both decided to submit each entry in their ledger (after each hand) to a block in the Bitcoin system.  At the end, Jack’s ledger shows he owes Dave $100, but Dave’s ledger shows Jack owes him $200.  This is where the verifying transactions takes place; each entry is verified by the Bitcoin community (a process called mining), in which these miners attempt to crack each encrypted entry on Jack and Dave’s ledgers. After 10 minutes all the codes are cracked by a miner named Bob, who indeed shows that Jack owes Dave $100.  Now, Bob’s reward for cracking the code, using his hardware, server space and energy cost is given in Bitcoin, which is a winner take all system resulting in only Bob being rewarded for cracking the code.  Now we can not only take Bob’s word, after cracking the code several other miners also verify Bob’s findings; and it is only after multiple other miners confirm Bob’s findings, is he rewarded.

          Currently the reward for solving one block is 3.125 Bitcoin.  This number originated at 50 Bitcoin and is cut in half roughly every 4 years.  Now a crucial piece to the whole process is there are only a finite number of bitcoin in existence, and that number is 21 million Bitcoin which is expected to be exhausted in the year 2140.  This creates scarcity, but also holds to the principle that each block is marginally harder to crack than the previous block.  It has also been stated that the last Bitcoin to be mined will be so computationally expensive, that the reward of Bitcoin will be worth less than the resources used to extract it (almost like real gold mining).

**Motivation**

Cryptocurrencies have gained quite the media attention, as they have soared in value; Bitcoin reached $19498.63 on December 18, 2018, but dropped to $16064.44 just 10 days later, resulting in an 18% loss in just a week and a half.  Cryptocurrencies are extremely volatile, in which the entire market is susceptible to 20% swings from each day to the next.  What makes these cryptocurrencies unique is they are not tied to any fiat currency (except a select few) or backed by any governments (not yet).  So, when I buy my $10 lunch and I hand the cashier two $5 bills, the Federal Reserve and the vendor have an agreement that no matter what the US Dollar is exchange rate is, they will still honor my two $5 bills.  Cryptocurrency does not have that guarantee, and Mr. Kevin O’Leary (aka Mr. Wonderful) says it best (referring to Bitcoin), "The fact is, it is so unstable — volatility is both directions, it's up and it's down — that nobody in a substantive transaction will take that risk." He is right, how can a currency truly be a currency if it’s exchange rate moves substantially during the transaction.

Now what really grabbed our attention to the crypto markets, was essentially the second Black Tuesday (1/9/2018), in which the entire crypto markets had a market capitalization (total outstanding shares \* price per share) of $814 billion. By 2/7/2018, the entire crypto markets market capitalization sunk to only $333 billion, a loss of 59% in almost a month.  The abrupt loss of value resulted in widespread panic (not as drastic as 1929), with some questioning if these assets are even currencies at all.  By 03/01/0217 the entire market cap had increased 15% from almost a week prior.  This chart represents the market capitalization as of 03/01/2018 for all currencies.



Despite the extreme volatility, we decided to try and find a pattern and even try to predict where a select few currencies will go.

**Analysis Goals**

We plan on accurately predicting a select few cryptocurrencies, despite the extreme volatility and unpredictability.  We expect that Tether (for reasons we will discuss in the results) can be accurately predicted, but Bitcoin with is extreme overvaluation and lavish swings will be an even greater challenge.  Our true goal is to expose the rampant swings in not just Bitcoin but all of our selected currencies.

**Data Sources**

          Our original data set was from Kaggle, in which a web scrape was conducted on the popular website [www.coinmarketcap.com](http://www.coinmarketcap.com/) which is a historical data set of cryptocurrency which is updated in daily interval; but the updates were slow and unreliable from Kaggle.  So we decided to run our own web scrape in python, in which we also scraped the coinmarketcap.com, to get the exact same information as we needed it.  We got our data earliest form 04/28/2013 and latest we got to 02/23/2018. Some of the features in our dataset are, symbol which is symbol of different cryptocurrency. XRP, LTC, NEO,MKR, USDT, ETH are some of symbol used for the coin. Date is the date of the Bitcoin when the data was updated for open, high, low, and close value. It is time series so, date variable shows when the information was updated. Rank is the rank of the cryptocurrency on which rank they are kept. Open is the opening value of the cryptocurrency for particular day, and high is the maximum value it reached for that date, low is the lowest value it reached for that date and close is the closing value of the it for that day. We also, have important features named spread which is the difference between high and low. This gives the spread per day. How much difference it occur in each day can be seen from the features. We can see kind of trend going on the dataset taking this variable in account.

 We instead decided to web scrape “coinmarketcap” ourselves, and instead only focus on a select few currencies and conduct daily appends to our scraped dataset.  With our web scrape, we decided to get data from 01/01/2017 to 02/23/2018 for our selected currencies; except for Bitcoin and Maker.  We decided because of its place as the first released currency, that Bitcoin deserved to be scraped from 04/28/2013 to 02/23/2018 and Maker because it is relatively new from 08/27/2017 to 02/23/2018.

For each selected currency we scrapped we have uniform variables.  We grabbed the slug, which is the name of the currency.  We of course have the date, so we can perform our time series analysis.  Next, we grabbed the four vital stock market variables: open, high, low, close.  Open is price the currency opened or started for that day. High is the highest value that currency reached for that particular day.  Low is the lowest point that currency plunged to for that day. Close is what the currency ended at for the day.  Now not unlike the NYSE and other Financial markets, there usually is a slight difference between the close of one day to the opening of the next, its not uncommon.  While stocks usually differ by a few cents to a dollar, cryptocurrencies can have substantial differences between the opening and close between those two days.  Now a currency like Dogecoin can differ by 0.0001-.0001 cents, Ethereum and Bitcoin can differ by $20 and $400 respectively from one close to the succeeding opening day.

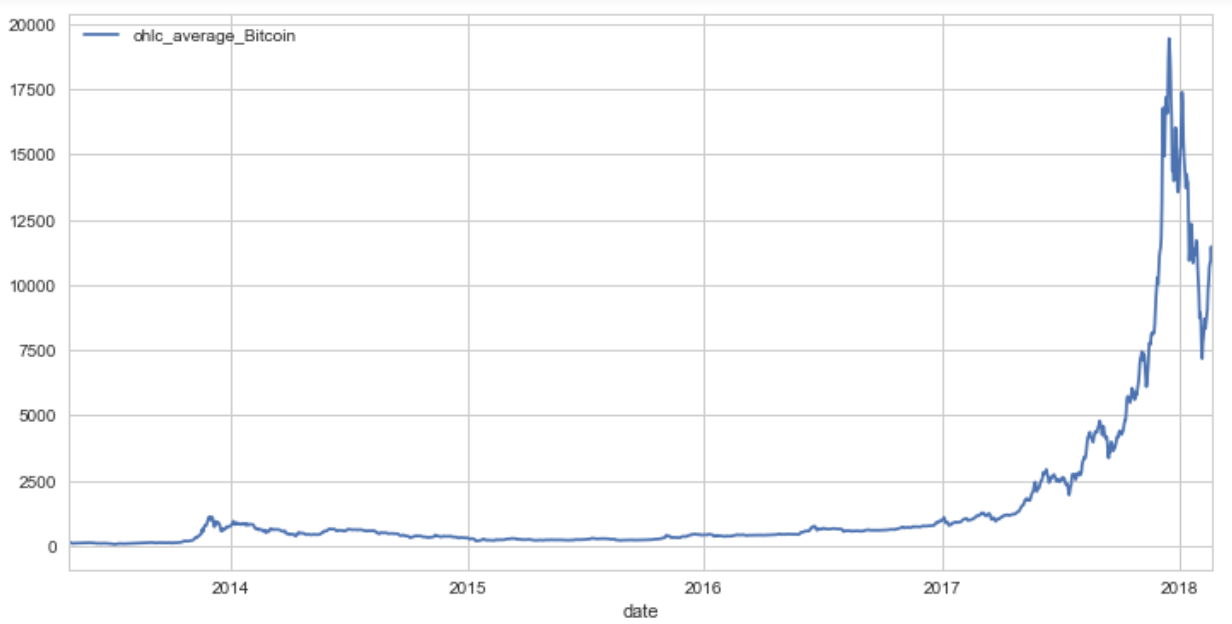
We calculated the infamous spread by taking the difference of the high and low for each day for each particular currency.  Spread gives us an accurate depiction of how wild each of our currencies performed for that said day.  Currencies like Dogecoin have seen it’s highest spread of below $0.005, and Tether had a spread high of $0.21 despite allegedly being tied to the USD.  Currencies like Bitcoin have seen spreads as high as $4110.40, Ethereum $417.09 and Maker $747.82 in one day.

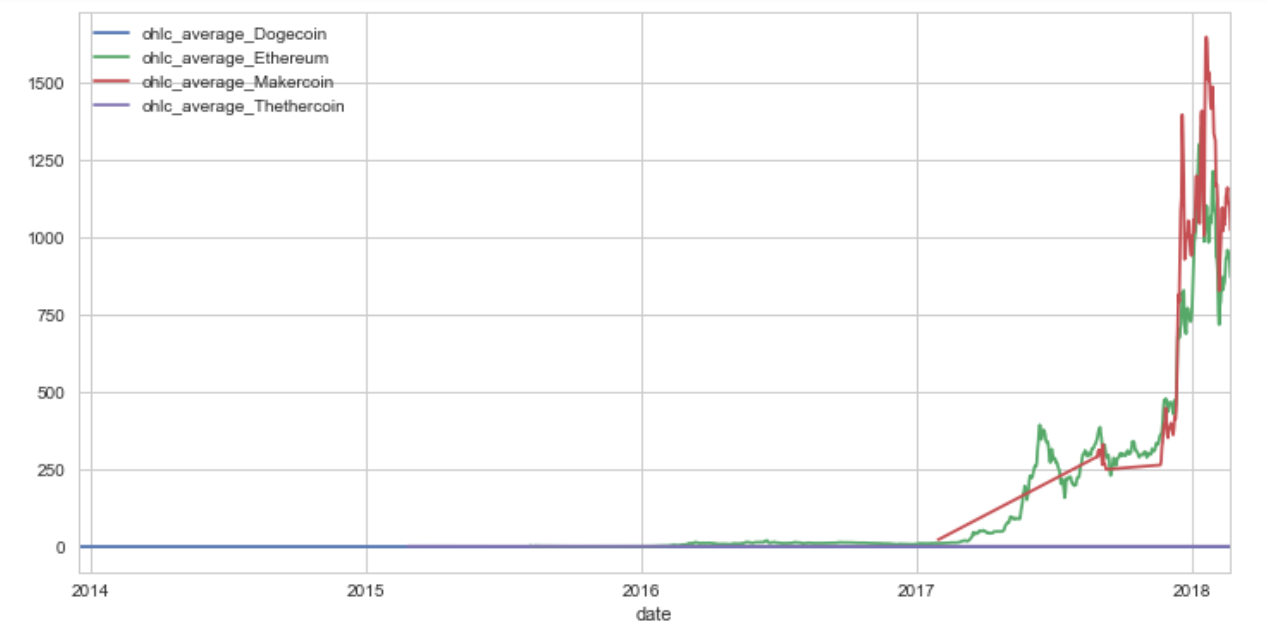
**Date Preprocessing**

From summary statistics, we have only one variable  name close\_ratio,which has missing values which is less than two percent. Also, this variable is not our interest so, we decided to remove that variable from our dataset. There are no any impossible values, only we can see some kind of shocks or sharp peaks in the data point but those values are existing values. All the variable are in same metrics so we don’t need to standardized and transform the variables. We added two of the variables, OHLC and Day of the week. OHLC is the average of open, high, low, and close, and is seen a measure a coin (or stock’s) average change throughout the day. We added this variable to see the trend and spread so that we can predict the model and forecast with low Confidence Interval for correct forecast in future. Also days of the week, we added this variable to predict and forecast in daily basis.

**Data Exploration**

So to explore our data we plotted our OHLC variable for the coins we planned to analyze. We knew cryptocurrencies we volatile but we wanted to see changes from the beginning of 2017.  This top graph represents Bitcoin from 04/28/2013 to 02/23/2018.



We can see that Bitcoin stays relatively non volatile with very little change until the 2nd quarter of 2017, with the exception of a spike in 2014. Then it begins to climb.

In the above graph we have plotted Dogecoin, Ethereum, Maker and Tether. All of these coins are only collected from 01/01/2017 to 02/23/2018 (with the exception of Maker which was released around 01/29/2017).  Dogecoin and Tether stay relatively the same, with the former floating below $0.01 and the latter firmly at $1.  We notice that Maker has relatively unpredictable spikes and drops, with Ethereum although still appearing very volatile, following less linear path still adhering to the same shocks and spikes of Maker.

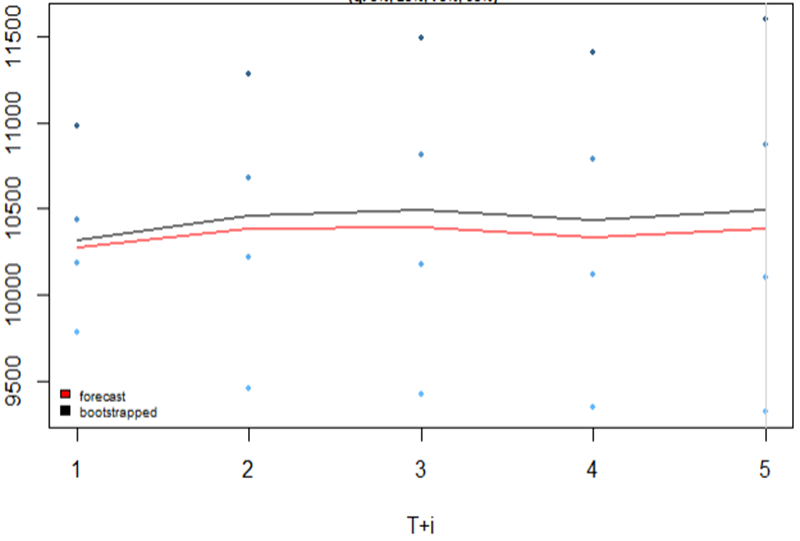
**Data Analysis**

We are doing time series analysis, but our variance is constantly changing with time like most financial assets.  We used the GARCH model or the General AutoRegressive Conditional Heteroskedastic model, which is a modified form of the ARMA model but does not need constant variance.  The GARCH model preparation for each currency is the same with the exception of the “armaOrder”, which is changed according to how many days we want use for our prediction.

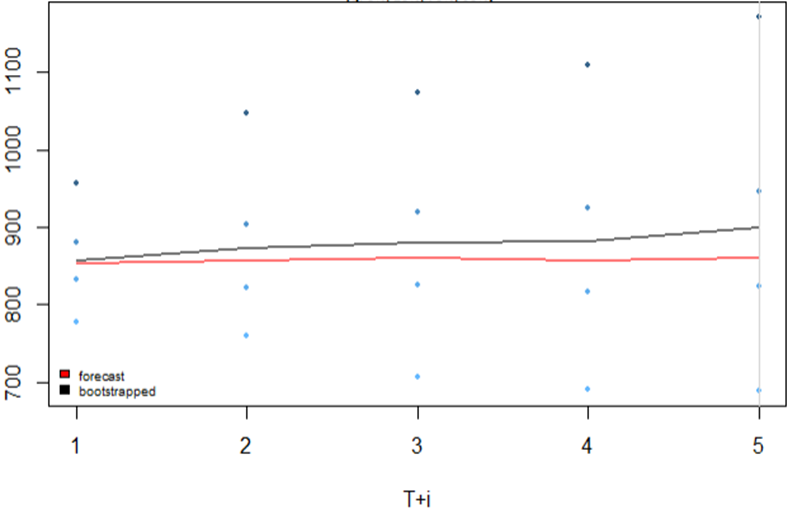
Our entire analysis is performed in the R language, specifically using the “rugarch” package to run our GARCH models.  We would build are models using “ugarchspech” function, we would fit our data to the model using the “ugarchfit” function and run our predictions using the “ugarchboot” function.  Web scraping was performed with python (anaconda actually, so you might need to install some libraries) using the “BeautifulSoup” and “Urllib2” libraries.

**Model Result**

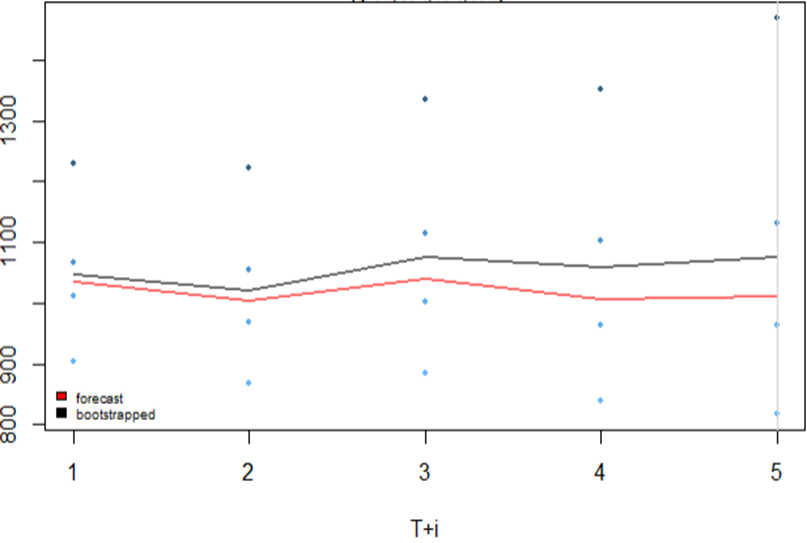
The first coin we tried to predict was bitcoin, using 9 days to predict the next 5 days closing price.  The average price for 5 days is just under $10,500, with confidence intervals a high of $11500 and low of $9500 ($2,000 margin of error).  Our close on the 5th day (02/28/2018) was $10393.10.  Our prediction function also runs a bootstrapped estimate as well as the prediction, labeled forecast.  For each graph the dark blue circles indicate the optimistic estimate (High C.I.) and the light blue the pessimistic (Low C.I.).



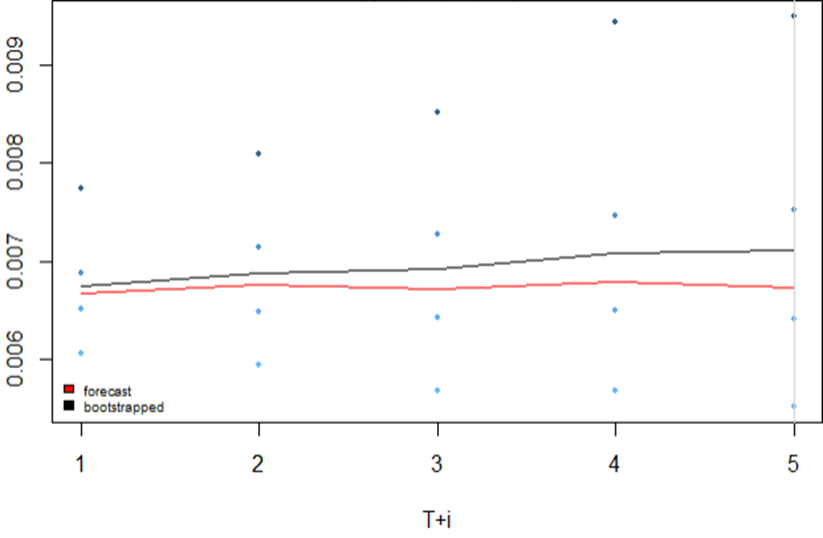
The next coin we analyzed was Ethereum, where we used 7 days to predict the closing price of the next 5 days.  The average closing price predicted in the model was around $857.74 with a positive linear trend to $900.42, and included confidence intervals just above $700 and slightly above $1200.  The close on the 5th day (02/28/2018) was $855.20.



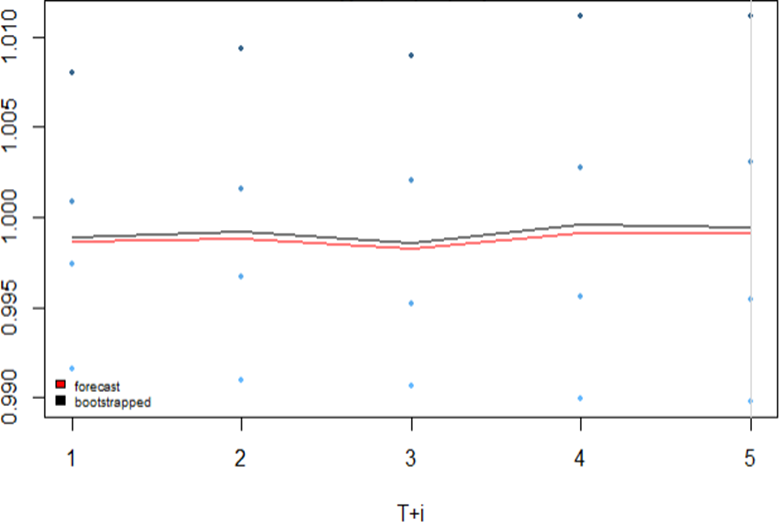
Next we ran Maker, a new currency that has seen rapid growth despite only being 14 months old.  We were only able to use 6 previous days to predict the next 5.  The average closing price predicted is non linear line with its slope moving slightly up and down with each passing day and is roughly $1046 - $1056.3 with confidence intervals as high as $1436 and a low of $865. The close on the 5th day (02/28/2018) was $1028.82



Now we get to the fun currency, Dogecoin which remains below one penny. To predict those fractional pennies we used 9 days to get the next 5. Our average close was $0.00752 and a confidence interval high of possibly 0.015 and low 0.005.  Dogecoin’s close for the 5th day (02/28/2018) was a mere $0.006.

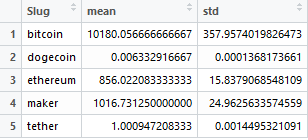


Our final coin we analyzed was the plagued Tether, which claims to be tied to the USD.  Our model predicted an average close of around $0.998, with a confidence interval high of $1.015 and low of $0.990.  The tether close for the 5th day (02/28/2018) was a shocking $0.998 because of its relation to the USD.



From our models we actually get relatively close to the actual predictions.  Bitcoin is the currency the has everyone’s interest, but prone to so rapid swings on the trading day and across sequent days.  We were relatively close to the predicted mean close for the 5 days.  Ethereum is the second largest market capitalization behind Bitcoin and is another highly sought after currency. Our model’s average close prediction was slightly higher than the actual close but well within the confidence intervals.  Maker is the second most expensive currency behind Bitcoin despite still being in infancy, but our model predicted almost exactly the closing price.  Dogecoin, known for it’s cute Shiba Inu mascot, was created off an internet meme meant to be a joke, but has soared in both market capitalization and popularity.  Although its price is a bit of a joke, our model overstated the averages by more than 0.0005 (it's so much) but still well with in our lowest confidence interval.  Our final currency is Tether, which is usually perceived to be $1 or very close.  Our model was very accurate to the closing price of 0.998.

Despite our models performing quite well, I am skeptical of our results because the cryptomarket was relatively stable during our GARCH prediction. All of our currencies (except Bitcoin) had means that were relatively close (each one is one a different scale) to the actual closing price; indicating our currencies closing price in each prediction day, never really deviated from the mean or had returned close to the mean by the 5th day.  Bitcoin however closed $213.05 higher than its mean of those 5 days and had a standard deviation of 358.  This chart is the mean and standard deviations of the OHLC for each currency during our 5 prediction days, I chose the OHLC because it is a more accurate depiction of how each currency performed throughout the day.



It really is not surprising that the most overvalued and volatile currency would deviate substantially from the mean throughout those 5 days.

**Recommendations**

When running the model, we could not consistently use the same number of days to calculate the 5 days because often the model would not converge; so it became trial and error to find the number of days that worked.  I feel the GARCH model plot although is fairly accurate, could be greatly improved.  I found it nearly impossible to apply axis labels, graph labels, and even custom labels.  The output is a little bland, and could highly benefit with more numbers for reference on the y-axis.  Despite the aesthetics shortcomings, I found the GARCH model to be highly useful and plan on using it for further analysis.  I would be interested to use the other GARCH model packages in R, to see if I can get better results with more days used.

**Conclusions**

We were actually surprised with our model’s ability to accurately predict: Maker, Ethereum and Bitcoin as all three are extremely volatile and notorious for their sporadic daily behavior.  Tether and Dogecoin were not really surprising as the former is a fun currency known as a joke to the whole market (in several ways).  Tether however did act as a bit of a control just to make sure the model was working as we already had a good idea of where the price would be.  Tether could be a catalyst for some currencies collapsing as there are rumors the company truly does not keep reserves equal o outstanding “shares”.  If indeed Tether is a “Ponzi” scheme with no actual value or exchange rate, it might blow open the market revealing other crypto-frauds.  Tether is also a highly speculated to be tied to money laundering for its owners and to inflate other currencies.  Perhaps the most alarming sign of all cryptocurrencies is the creations of options for the market by the Chicago Board of Options and Exchanges (CBOE).  Now an anonymous highly volatile asset traded on open markets is highly susceptible to price manipulation, without any transparency on the options buyer’s themselves.  This is a huge red flag and shows the markets (and the CBOE) are getting uncontrollably greedy.  Perhaps the next smackdown to the cryptomarket is regulation, as several governments, such as South Korea, are considering banning trading within their jurisdictions and are imposing or about to impose regulations.  The United States government has placed a ban on trading Bitfinex, a currency tied to Tether, based out of the Virgin Islands a known tax haven to the wealthy.

I personally do not advise trading cryptocurrencies out of the idea of how volatile they are but are susceptible to manipulation and how difficult it is to sell them.  I will say that Dogecoin despite its paltry value, is just that a joke and just for fun.  If you approach this market just to play, and not attempt to get rich quick, you will not lose your life savings.  Furthermore, I do not think cryptocurrencies are going anywhere, but some will simply fall out of popularity.  However, I do feel that consistent regulation is needed in the United States and worldwide, to protect consumers and banks from getting scammed and left holding the bag when the market collapses.  Warren Buffett says it best when referring to investors, “They should try to be fearful when others are greedy and greed only when others are fearful.”

**References**

<https://www.coindesk.com/information/how-bitcoin-mining-works/>

<https://99bitcoins.com/price-chart-history/>

<https://www.cnbc.com/2017/12/07/kevin-oleary-bitcoin-is-an-asset-not-a-currency.html>

<https://coinmarketcap.com/>

<http://www.businessinsider.com/dogecoin-cryptocurrency-has-market-cap-above-2-billion-2018-1>

<http://www.cboe.com/>