

Volatility in crypto assets and factors affecting buying decision

Surbhi Kumar

Harrisburg University

Volatility in crypto assets and factors affecting buying decision

Investors are capitalizing on the greatest investment opportunity with the help of bitcoin and blockchain technology. Bitcoin is a buzzword nowadays. It was the first crypto asset, and now there are more than 800 crypto assets available for purchase like Ripple and Litecoin. The most popular crypto assets break all the unthinkable records repeatedly. India and Japan have recognized Bitcoin as official currency and US recognized Bitcoin as an asset. In 2018 even Venezuela plans to float crypto assets called Petro in the market. In short, Crypto assets is the future of investment markets. Crypto asset market is extremely fast changing. It is tough to navigate in this market as fluctuation in these assets are so high in a day that it can be compared to a year's pattern of the stock market. One of major risks posed by cryptocurrencies is their use in illegal activities. It can be used to bypass government authorities' laws and regulations of any country (Nica, 2017).

People investing in these assets are still considered brave because of the risky nature. These assets are evolving in the current block chain revolution. Crypto asset constitutes of three taxonomies - cryptocurrencies, crypto commodities, and crypto tokens. Cryptocurrency is a digital currency in which encryption techniques are used to regulate the generation of units of currency and verify the transfer of funds, operating independently of a central bank.

Decentralized cryptocurrencies such as bitcoin now provide an outlet for personal wealth that is beyond restriction and confiscation. Crypto commodities - One of the well-known crypto commodities is Ethereum. It is based on blockchain technology. Ethereum is a platform where currencies like Bitcoin could be used, and more could be added. It is a framework and provides protocols to create other cryptocurrencies. The creators of commodities wanted to abolish corruption and establish a dynamic and decentralized network like the internet as a framework

and set up of protocols to create websites. Crypto tokens are tokens which can represent specific assets or utilities that resides on top of a different blockchain. They are loyalty points and can represent any assets. Multiple studies have been conducted to understand and define this market disrupting class of assets. For many analysts the market is an interesting new market with many players because of its digital nature presence of high quality data and volumes. This helps in excellent research data support (Gandal, 2016).

Currently, there has been a total of 445 Billion Dollars invested in crypto assets. Compared to other mature asset classes this is a small amount, but on its own, it is a significant sum of money. Some studies have been conducted to understand the factors affecting the volatility of this asset class and give investors some advice (Osterrieder, Lorenz, & Strika, 2016), while other studies have shown that investors relegate this asset class to a ‘bubble category’ (Blumenthal, 2017). While some other analysis suggests to have a portfolio in cryptomarket with some stable growth coins like Dash and Monero (McCormack, 2017). However, none of the studies dwell into the impact that eastern countries have on the buy/sell patterns of the western countries. The current studies have mapped the impact that major cryptocurrencies have on the market, but none of them have documented the impact that different crypto asset categories have on each other. Moreover, the studies have not characterized the investors and have not documented the impact the socio-economic factors have on the long-term sustainability of this asset class. We have conducted this research to fill the gaps in our current understanding of the crypto assets and help the investors make intelligent buy/sell decisions by having a better understanding of the volatility factors and having a clear sense of the sustainability of this asset class.

Literature Review

To more efficiently understand the crypto assets, its relation to the class of risky assets, the pricing volatility, and buy/sell decision factors, I have selected a subset of literature relevant to the following questions.

1. Are crypto assets a bubble or a class of risky assets?
2. What are the factors contributing to the volatility of crypto assets?
3. How can an investor decide when to buy/sell the crypto asset?
4. Do people value crypto assets and do they consider making it part of their portfolio?
5. What is the long-term view on crypto assets?

Multiple websites and magazines that have taken up this topic and have published various expert opinions, but for this review, I focused on the academic work and journal entries.

Are crypto assets a bubble or a class of risky assets?

Heyde (1999) establishes using (Black-Scholes) model for pricing volatility that any asset with the price volatility of 3% is considered risky. The primary fear about cryptocurrencies is the existence of a pricing bubble (Shaen, 2018). Heyde (1999) furthers his case by plotting the returns and showing that the returns if defy logarithmic progression, are considered risky. However, the author fails to identify a reasonable timeline for which the returns should be mapped. Some stocks, when mapped for shorter timelines, fall into the category of risky assets, like Facebook, Apple, but when returns are plotted over longer timelines are considered safe assets. Bakar & Rosbi (2017) calculated the pricing volatility of Bitcoin and other crypto assets and found them to vary from 4% to 13%. Bakar & Rosbi (2017) further calculated the pricing

volatility for Bitcoin and found it to be 4.48%, which according to Heyde (1999) will be considered as “risky.” But again, the authors conveniently chose the timeline that suited them, i.e., from the year 2014 to 2016. Also, to keep in mind is the factor of macroeconomics news to cryptocurrencies return (Shaen, 2018). According the author’s findings crypto currencies majorily Bitcoin is subject to effects of same macroeconomic factors are traditional financial instruments.

During this time period, there were several thefts of cryptocurrencies and also the exchanges like MtGOX defaulted, which had an adverse effect on the pricing of crypto assets. I would argue that the analysis would not hold true if the time period is elongated or if other mature crypto assets like Ethereum are considered. There is a fine distinction between an asset considered as “risky” versus a “bubble.” Garber (1990) through his analysis of the history’s most famous bubbles, namely, the Dutch Tulipmania, the Mississippi Bubble, and the South Sea Bubble, explains that while the bubbles are caused by uninformed investors, fraud, and erroneous news, and market speculations, not all market speculations can be relegated to the bubbles category. The author continues to say that some assets with strong fundamentals can also be dragged into market speculations from time to time.

Blumenthal (2017) presents a strong case in favor of the underlying technology and usefulness of crypto assets. According to the author, a lot of investors are calling the crypto assets a bubble because they either do not fully understand the underlying technology or do not possess the vision of its utility. The author summarizes some of the benefits of the cryptocurrencies as - “Minimal fees for transactions as compared to charges and fees from banks. Enjoy sole ownership of your own cryptocurrency. Ensure privacy and ensure anonymity with your transactions. Get easy access to your cryptocurrency - anytime and anywhere. No risk of

loss for sellers of goods and cryptocurrency. Absolutely no chance of counterfeit. Instant processing, verification and completion of transactions.” Although I agree with most of the benefits, I am not fully on board with the research the author conducted to dissipate the reasons for which people are calling crypto assets a ‘bubble.’ Furthermore, I believe that the author could have done a better literature review when it comes to presenting examples of assets that were considered speculative at first but then lived long enough to become stable.

What are the factors contributing to the volatility of crypto assets?

Sams (2015) discusses the apparent flaws in the cryptocurrencies supply of coins, which has led to high price volatility. The author says that because the coin generation and dissemination rules are deterministic, they are unable to account for the unanticipated changes in the market dynamics, for example, the sudden popularity of bitcoin has led to a sharp increase in the pricing volatility. I believe that the author’s recommendations for an elastic supply make sense, however, when the demand decreases, the additional currency will cause value dilation and will lead to a sudden fall in value. Also, some of the newer currencies like Ripple and Neo came up with a static number of coins from the so, so analyzing those currencies will help us understand if the coin generation algorithm is responsible for volatility. As per author Barker (2017) specifically Bitcoin’s rate of adoption is hampered by bad press however Bitcoin friendly investors took it as signs of maturing market. As per the author Bitcoin as an investment can be compared with gold as Bitcoin fluctuates as per it’s perceived value.

Osterrieder, Lorenz, & Strika (2016) talk about the close dependence of the smaller currency on the larger more dominant ones. According to the authors, the volatility of the smaller currencies is driven because of the selloff behavior of the crypto assets that have more market share like Bitcoin, Ethereum, Litecoin. I have kept close track of these selloff behaviors and

agreed with the authors. However, there are other factors at play here. There have been instances when less dominant crypto assets like Neo and Ark performed consistently well because of they had an active user base for their platforms. Moreover, if we start breaking apart the crypto assets into its three native categories – cryptocurrencies, crypto platforms, and crypto tokens, the authors Burniske & Tatar (2018), found that there is a very vague correlation between the three. Often, when there is a general fall in the category of cryptocurrencies, the crypto platform and crypto token assets registered an uptrend. Guides (2018) also supports the argument I am making because he says that there is a strong correlation between the overall market performance and the top performing crypto assets but weak correlation between smaller market dominant crypto assets and major market dominant crypto assets. But the author did not do a thorough correlation analysis to prove his point.

Bitcoin was created to stand as anti-establishment, anti-state way of transferring value. It is a decentralized, peer to peer virtual currency outside of the regulatory systems and current power structures. Kutiš (2014) states Crypto assets as the grand new hope of the technological utopians and cyber-libertarian pundits who, in the Internet Revolution of the mid-1990s, promised us the death of the established order and the birth of a new self-governing virtual realm with vastly increased human freedom and empowerment. As Casey & Vigna (2015) state that since its inception in 2009, Bitcoin has cemented a reputation as a currency of volatility, insecurity, and criminality. The author further writes that in the public discourse surrounding Bitcoin, there has been an emphasis on its speculative nature, the associated risks and its alleged use in illegal activities, including its role in money laundering and illicit commerce. I agree that these reasons have definitely hindered the widespread adoption of crypto assets and have added to the volatility of crypto assets. But many countries, including Iceland and Venezuela, have taken steps

to invest in this concept by launching their versions of the cryptocurrencies, whereas other countries, like India and China, are taking a stance against this asset class. I believe that more research is needed to see how the political support/rejection impacts the volatility of the crypto assets. As government regulations can play a very important role in involvement of a countries investors and pricing of cryptocurrency. For example, Venezuela which launched arrest of Bitcoin miners. This lack of freedom in financial sector forced for miners to go into hiding (Oreso, 2018). Many still feel negatively about cryptocurrencies as they for the first time allowed smooth transactions and emergence of illegal market places in the cyberspace like Silk Road, founded in 2011 and closed in 2013 by US authorities (Stroukal & Nedvědová, 2016).

In research, Böhme, Christin, Edelman, & Moore (2015) highlighted the fact about cryptocurrencies market concentration resulting in high volatility. Bitcoin has an enormous market in China, which is home to 90 percent of the world's Bitcoin trades and 70 percent of world mining production. Any market regulations or political upheaval in these core countries results in huge selloff trades lowering the price of the cryptocurrency. In 2015, the US and China were the market dominators for crypto assets, but the playing field has changed significantly since then. The authors at the time of writing the paper ignored that the fact that the growth in this asset class is exponential and that its awareness and eagerness to buy among consumers is spreading rapidly. Today, cryptocurrencies and are more spread with over 190 exchanges. The research did not study how the increased coverage of crypto exchanges has decreased the volatility caused by certain market upheavals.

In more recent research, Yeoh (2017), wrote about regulatory challenges impacting blockchains, innovative distributed technologies, in the European Union (EU) and the USA. The

blockchain is the underlying technology powering crypto assets. Yeoh (2017) stated that the future volatility of crypto assets is closely intertwined with the technological improvements related to scaling. The author addressed a key issue in the research but failed to address several other key issues regarding the regulations impacting energy consumption. The crypto asset generation and maintaining transactions ledger require enormous amounts of computing power, which requires large amounts of electricity consumption. Regulations impacting energy consumption in key crypto markets can lead to price volatility, but the widespread of crypto exchanges should minimize the impact.

How can an investor decide when to buy/sell the crypto asset?

Urquhart (2017) examined the volatility of Bitcoin using the GARCH and HAR models and forecasts the pricing using the models. The author finds that HAR models more accurately forecasted the pricing volatility. However, the author did not take into account the long run conditional variance. Moreover, the error rate with the HAR models is upwards of 10%, which is too much in this sensitive market. Bakar & Rosbi (2017) refined the models by implementing forecasting using the autoregressive integrated moving average (ARIMA). This study performed autocorrelation function (ACF) and partial autocorrelation function (PACF) analysis in determining the parameter of ARIMA model. According to the authors, the result shows the first difference of Bitcoin exchange rate is a stationary data series. Also, the forecasting error value decreased to 5.76%, much better than the HAR modeling methodology. Additionally, the ARIMA modeling has shown reliability in highly volatile commodity trading environments like Oil and Gas sector in the Malaysian stock market. The authors also validated the model using the high volatile stocks in the stock market. The results highlight the limitations in using the conventional

approach in order to identify the best specified ARIMA model in sample, when the purpose of the analysis is to provide forecasts. The results show that the ARIMA models can be useful in anticipating broad market trends; there are substantial differences in the forecasts obtained using alternative specifications. Katsiampa (2017) also proposed a parallel forecasting model - AR-CGARCH to account for the highly sensitive nature of the cryptocurrency market. Although users can use the forecasting models by Bakar & Rosbi (2017) and Katsiampa (2017) to make long-term buying decisions, none of these models are used to make intraday or short-term buying decision which are more relevant given the current market behavior and volatility. However, as per author Gandal (2016) research for a short time period shows that Bitcoin appreciate or deprecate as per the ups and downs of USD market. But the same is not true for other currencies as if USD appreciates other currencies tend to depreciate.

Fry & Cheah (2016) draw upon the close relationship between statistical physics and mathematical finance to develop a suite of models for financial bubbles and crashes. The study did a detailed analysis on two currencies – Bitcoin and Ripple and presented the investment strategies relevant for investing in the rival cryptocurrencies. The study is interesting because it detailed ways for hedging investment bets; however, the premise on which the study is based is biased because it assumes that crypto market is in a bubble because it is extremely volatile and speculative.

All of these models ignore the fact that crypto assets are traded 24x7 all year around, and that the buying/selling pattern of eastern countries heavily influence the buying/selling patterns of the eastern countries. Moreover, none of the qualitative research has been conducted to account for the market sentiments surrounding the trading of the crypto assets.

Conley (2017) lays out the principles grounded in monetary theory, financial economics, and game theory to identify Blockchain startups that have embraced initial coin offerings (ICOs) as a vehicle to raise early capital. The crypto-tokens offered in these sales are intended to fill a widely varied set of roles on different platforms. Some tokens are similar to currencies, others are more like securities, and others have properties that are entirely new. Each company's technological vision calls for a token with unique properties and uses. However, the author does not detail the effect of ICO offering location, which plays a key role because of rules and regulations surrounding the crypto assets. Currently, Switzerland is considered as the ICO haven because of its open crypto policies.

McCormack (2017) has his own recommendations which includes various suggestions like to getting used zooming in on trend graphs to understand the market. The authors other suggestions include to not chase assets which are in uptrend and to not buy a falling asset until the down trend is broken. However other tries to give some handy tools to the readers but they are not specific enough for readers to get the que and make purchase and sale of crypto assets.

Do people value crypto assets and do they consider making it part of their portfolio?

Osterrieder, Lorenz, & Strika (2016) outlined the extreme value behavior of cryptocurrencies, their correlations and tail dependencies as well as their statistical properties for the financial institutions to start looking into this asset class to start making trading decisions. Chan, Chu, Nadarajah, & Osterrieder (2017) generated the Gaussian and Laplace distributions for the top cryptocurrencies for investment and management purposes. Burniske, & Tatar (2018) confirmed that wall street and other established banks and financial institutions are either already investing in crypto assets or devising strategies to get into the trades. Catania, Grassi, &

Ravazzolo (2018) wrote that the Cryptocurrencies have recently gained a lot of interest from investors, central banks and governments worldwide. According to Jeong and Kevin (2018) cryptocurrency can be viewed as a rival to traditional currency that governments and banks currently control. Also, to remember is Bitcoins are effectively irrelevant when compared to larger payment system (Whelan, 2018). But I will either be naïve or too optimistic to think governments will let the cryptocurrency market function without any taxes or regulations for long.

Although there is a need for some sort of regulatory body for widespread adoption of crypto assets, the current research suggests that the adoption by prominent financial institutes is underway. Bresett (2017) detailed the portfolio of some hedge funds dealing in crypto assets and they align closely with the strategies detailed by Brennan (1975) for a portfolio with both safe and risky assets. This analysis confirms the treatment of crypto assets as a risky asset class. However, none of the above researchers focus on the individual buying behavior and their appetite to hold a risky asset in the form of crypto assets.

What is the long-term view on crypto assets?

Li & Wang (2017) conducted a study on Bitcoin Exchange rate against USD taking into consideration both technology and economic factors. The sole reason behind this article is to understand blockchain based crypto currencies and understanding exchange rates in comparison to traditional currencies. The authors found that in the short term the Bitcoin prices were sensitive to technology, hacking, etc. but in the long term, the prices were more sensitive to the economic fundamentals. They also found that the impact of mining algorithms on pricing has decreased significantly. The authors applied the autoregressive distributed lag (ARDL) model on

different time periods to account for significant hack and closure event of one of the more prominent crypto exchanges MtGOX. The authors concluded that Bitcoin exchange rate relates more with economic fundamentals and less with technology factors as Bitcoin evolves and that the impact of computational capacities on Bitcoin is decreasing as technology progresses. The authors kept this study extremely narrow focus which has both pros and cons. The pros are that the authors were able to make strong correlations and findings, but they were unable to generalize the results to other crypto assets.

Bohr & Bashir (2014) in their research found that the characteristics of the users investing in crypto assets like Bitcoin were different than the characteristics of the users who invest in the risky assets. A lot of proportion of the users were technology enthusiasts who were investing because of their affinity to the underlying technology. The authors mapped the research data over the period of 4 years and found that the number of investors with tech background grew steadily. I found the study by Bohr & Bashir (2014) extremely encouraging. The crypto assets are considered risky, but a large group of users is loyal to the asset because of the linkages to technology. However, in the study, the authors did not detail the age groups and ethnic backgrounds of the users. If the age groups of these tech enthusiasts are on the younger side, there are strong chances that they belong to lower spending capacity group, which would undermine the authors claim that the loyal tech followers would exponentially increase the growth of crypto assets.

Böhme, Christin, Edelman, & Moore (2015) found that among the buyers of crypto assets there is a strong reliance on the underlying technology and a strong acceptance of how this class of assets came into being. The authors summarize the top positives as, “Bitcoin's rules were designed by engineers with no apparent influence from lawyers or regulators. Bitcoin

is built on a transaction log that is distributed across a network of participating computers. It includes mechanisms to reward honest participation, to bootstrap acceptance by early adopters, and to guard against concentrations of power. Bitcoin's design allows for irreversible transactions, a prescribed path of money creation over time, and a public transaction history. Anyone can create a Bitcoin account, without charge and without any centralized vetting procedure—or even a requirement to provide a real name.” The authors did express concerns regarding the regulations and rules when Bitcoin interacts with the conventional financial system and the real economy. In this research also, the authors did not do a detailed analysis of the demographics. If the positive sentiments run only across the younger generation who do not have enough money to sustain the growth, the crypto assets’ growth and prospects will be impacted.

In summary, while there has been a multitude of literature concerning the crypto assets volatility, relationship to risky assets, and factors affecting buy/sell decision, little attention has been placed on establishing the influence the buying patterns of eastern countries on that of the western countries. Also, profiling the user base of cryptocurrency buyers and determining the long-term sustainability of this asset based on the socio-economic factors has not been part of any study. Moreover, little to none research has been done establishing the impact of the number of crypto currency exchanges on the volatility of the crypto assets. Neither has there been efforts to document the influence of various types of crypto assets (currencies, tokens, and platforms) on each other.

In this research, we will focus on three hypotheses. First, India and China’s buying and selling of crypto assets like Bitcoin, Ethereum, and Litecoin during daytime impacts the daily prices of the US crypto market. Next, there is no significant relationship between the individuals who buy and sell crypto assets and their demographics (age, gender, income). Lastly, the prices

of major crypto assets ‘cryptocurrencies, crypto commodities and crypto tokens’ do not positively or negatively influence each other’s prices.

Method

In this research, I have three hypotheses. Therefore, I have divided the method section into three parts as per the hypothesis.

The first hypothesis is daily prices of Bitcoin does not impact the purchase of Bitcoin. For this hypothesis we have used the Transaction Data for Bitcoin available on Bitstamp.net website. The variables are ‘date’ in Unix timestamp, transaction id as ‘tid’, ‘price’ of bitcoin and ‘amount’ of Bitcoin either purchased or sold. The last variable is the categorical dependent variable called ‘type’. When type=0 it depicts a buy and 1 is for a sell. For the purpose of the analysis of this hypothesis we are using regression analysis and the details and outcomes are shared in the results section in detail. I have changed the ‘type’ variable to be treated as a factor variable for the logit regressions after running the descriptive statistics for this dataset. Will use price and amount as numeric variable therefore coded them to be treated as numeric variables.

	date	tid	Price	type	amount
1	1529551851	68754331	6768.07	0	0.00267800
2	1529551836	68754325	6768.08	0	0.02974930
3	1529551831	68754324	6768.18	0	0.16290587
4	1529551810	68754320	6768.29	0	0.03304641
5	1529551805	68754319	6768.29	0	0.91798468
6	1529551779	68754315	6765.18	1	0.07819757

Table 1. This is a small portion of Bitcoin transaction dataset used in the hypothesis

After converting the variables to either factor or numeric I have coded the logit regression model to understand the significance of the hypothesis.

The second hypothesis is that there is no significant relationship between the individuals’ profitability and their demographics (age, gender, income) in the cryptocurrency market. I have conducted a convenience sample survey of 10 participants who have purchased and sold crypto

assets in the past. The age group ranges between 25 to 55 years. These individuals are of any gender, sexual orientation, race, and ethnicity. Individuals with a documented history of crime are avoided from this survey as we did not want to include people who might use crypto assets for illegal purposes. The participants were contacted via phone and belong to my social circle. Before starting of the procedure, I have contacted them and asked about whether they are dealing in crypto assets or not and whether they would like to be part of crypto assets related survey.

Participants were first surveyed about their basic demographics - age, gender, race and income. Then I asked them whether they are currently trading in crypto currency. After this I asked the participants to name a mainstream cryptocurrency they trade in and whether their crypto assets profile is profitable or not. This survey altogether has 7 questions as shown in Figure 1. Participants took approximately fifteen minutes to complete the survey.

The age variable is divided into three groups constituting of 10 years' age range each. The income group falls into four categories an annual income of less than \$25,000, between \$25,000 to less than \$50,000 and the third category is between \$50,00 to less than \$75,00 and the last category is more than \$75,000. For the gender, the variable male is 0 and female are 1 as shown in Figure 1.

Survey

1. Age Group
 - 25 - 35 years - 0
 - 35 - 45 years - 1
 - 45 - 55 years - 2

2. Gender
 - Male - 0
 - Female - 1

3. Income
 - 0 - 24,999 - 1
 - 25,000 - 49,999 - 2
 - 50,000 - 74,999 - 3
 - 75,000+ - 4

4. Race
 - Asian - 0
 - African American - 1
 - White - 2
 - Hispanic - 3
 - Another Ethnicity - 4

5. Have you ever traded any Crypto Currencies?
 - Yes - 0
 - No - 1

6. Name one major Cryptocurrencies you have traded with?

7. Is the trading Profitable?
 - Yes - 0
 - No - 1

Figure 1. Shows the survey questions and the values given to the nominal variables to easily create factors in r programming.

Since all the variables are categorical, we have again used Logit regression technique to understand targetable characteristics and scoring our prospects to understand whether they are profitable or not.

The third hypothesis is that the prices of bitcoin is not influenced by dates as in not impacted by either seasonal or cyclical factors. For this hypothesis, I have collected data from the Bitcoin charts website. I have downloaded daily data from September 2011 to May 2018. The variables are Date in UNIX timestamp, price and amount.

In time series analysis we don't consider data that we know will change patterns in the future because of external factors. The time series can either be trend, seasonal, cyclical or random behavior.

For the purpose of creating a predictive time series model I have created two new variables called count and value. The count will calculate the number of Bitcoins purchased on a daily basis. The value will be calculated by multiplying the price and amount. Then will create a time series dataset with date and value field. Once the 5 variables are present in the data set we create a 2 set of time series object. One called 'myts' for the date range of January 2012 to December 2017. The other dataset 'daily_ts' is for the date range of September 2011 to June 2018. This is done intentionally so that by forecasting myts dataset for January, February and March 2018 data; we can see whether the forecast is accurate or not by comparing it to the daily_ts time series plot 2018 data.

The GitHub Project Link: <https://github.com/surbhikr89e/finalproject699>

RMarkdown URL on RPubS: <http://rpubs.com/surkumar/ANLY699FinalProject>

Results

In this research, I have three hypotheses. Therefore, I have divided the results section into three parts as per the hypothesis.

From the first hypothesis the transaction data set we are using is a data frame with 100 rows and 5 variables. As the summary output in Table2 shows date, price and amount are character variables. For the purpose of regression analysis, we will be converting other variables. Price and Amount to numeric variables and type to a factor variable.

date	tid	price	type	amount
Length:100	Min. :68754024	Length:100	Min. :0.00	Length:100
Class:character	1st Qu.:68754083	Class :character	1st Qu.:0.00	Class :character
Mode:character	Median :68754150	Mode :character	Median :0.00	Mode :character
	Mean :68754158		Mean :0.23	
	3rd Qu.:68754199		3rd Qu.:0.00	
	Max. :68754331		Max. :1.00	

Table2. The Table shows the summary statistics for the bitcoin transaction data without data cleaning. All the variables except transaction id 'tid' is showing up as character variable.

After converting the variables to numeric and factor type we run the summary function again as shown in Table 3.

date		tid	price	type	amount
1529461078: 4	Min.	:68714515	Min. : 1.00	0:64	Min. : 1.00
1529460594: 4	1st Qu.:	:68714591	1st Qu.: 9.75	1:36	1st Qu.:19.75
1529461113: 3	Median	:68714646	Median :17.50		Median :44.50
1529461102: 2	Mean	:68714653	Mean :21.83		Mean :44.98
1529461067: 2	3rd Qu.:	:68714724	3rd Qu.:36.00		3rd Qu.:69.25
1529460978: 2	Max.	:68714794	Max. :52.00		Max. :94.00
(Other)		:83			

Table 3. shows the change in the summary of variables after data type change

As the summary output shows in Table 3, the price variable varies from a minimum of 1 to a maximum of 52. When it comes to amount, it ranges from 1 to 94 with a mean of 44.98. In the variable type 64 buys and 36 sells have happened.

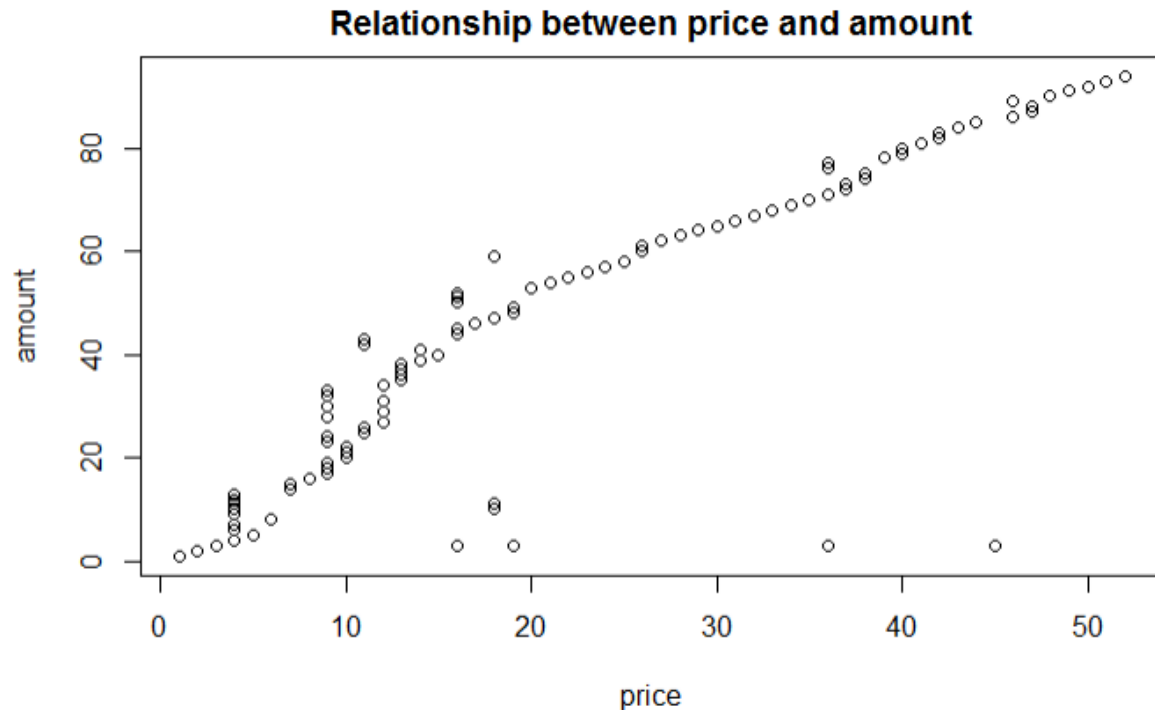


Figure2. The graph that price and amount variables have strong relationship

As shown in Figure 2 a scatter plot depicting the relationship between the 2 variables price and amount. This means the higher the price goes more number of Bitcoins will be purchased.

The Table 4 shows the standard deviation for the bitcoin transaction data set. Since type variable is a factor, therefore, the standard deviation is not applicable.

price	type	amount
14.6066645	0.4824182	28.2899775

Table4. The Table shows the standard deviation of the variables price, type and amount

We will now use the logistic regression model using the glm (generalized linear model) function. In order to get the result of the regression we use the summary command and get the result shown in Table 5.

```
call:
glm(formula = type ~ amount + price, family = binomial(link = "logit"),
    data = bs_df2)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.2415  -0.9029  -0.8150   1.3075   1.6430

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.077478    0.412915  -2.609  0.00907 **
amount      -0.005228    0.014604  -0.358  0.72038
price        0.033043    0.028259   1.169  0.24228
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 130.68  on 99  degrees of freedom
Residual deviance: 127.66  on 97  degrees of freedom
AIC: 133.66

Number of Fisher Scoring iterations: 4
```

Table5. The Table shows the summary statistics for the logit regression model

As can be seen in Table 5 the first item shown is the formula R programming used to fit the data. The formula shows the predictor variables price and amount. Along with that it shows the dependent/ target variable called type with the bitcoin transaction data type.

The next item in the output of Table 5 is deviance residuals which helps in measuring model of fit. The part shows the distribution of the deviance residuals in relation to individual cases of the model.

The next part of the summary output shows coefficients including the standard error, z-statistics and associated p-values. Theoretically coefficients in terms of regression are of two unknown constant that show a linear model of intercept and slope. If we wanted to predict how many bitcoin and what prices should help in buying of a Bitcoin we use a training data set and predict the values as coefficients by using the formula. Ultimately the goal is to get a fitted line that is as close to the other 100 observations of the data set. Both coefficient of amount and price statistically insignificant. For every one unit change in amount the log of odds for sell (versus buy) decreases by 0.005. For every one unit change in price the log of odds of buying a bitcoin increases by 0.03.

The coefficient estimate from Table 5 contains of three rows and the first one is the intercept. In short in our example the value of price and amount it takes for a buyer to purchase a bitcoin. The intercept for this data set is -1.077 means the user wants to buy the Bitcoin in case of surge in price and amount. The second and third row shows the slope of the coefficients or in our example the value of price and amounts. It means with every 0.005 decrease in amount user wants to buy one more Bitcoin. Similarly, for price for every 0.03 increase in price the Bitcoin buyer would want to buy 1 more Bitcoin.

The standard error measure the average amount the estimate of coefficients is different from the response variable actual value. Needless to say, ideally, we would want this error value to be as lower as possible as otherwise it shows the issue with coefficients. In the present Table 5 I would like to conclude that the current coefficient for amount is 0.005 and that can vary by 0.14. Similarly, for price the current coefficient is 0.03 and it can vary by 0.02.

The p-value shown in the Table 5 as column $\Pr(>|z|)$ is the probability of measuring any value equal to or larger than the z value. Typically, the p-value to be less than equal to 5% to show any significance. In our scenario all p-values are greater than 5% therefore we reject the reject the alternate hypothesis.

The deviance measures in the Table 5 the goodness of fit of a generalized linear model (Gupta, 2017). As in the article the higher the deviance goes the poorer the model fit is. The null deviance is 130.68 on 99 degrees of freedom. The null deviance only shows the performance of the model when it is governed by the intercept. The residual deviance in the next sentence shows the model when the independent variable price and amount are included in the model. For our model the residual deviance is 127.66 on 97 degrees of freedom. The lower level of residual deviance than null deviance shows that the model is better after adding the two variable price and amount. For the degrees of freedom in Null deviance is 99 which is $N-1$. So, for this model $N=100$ and therefore the Degree of freedom for Null Deviance is $100 - 1 = 99$. For calculating the residual deviance, the formula is $N-K-1$. K being the number of variables. Since in this model we have two variable that is why $100-2-1 = 97$ degrees of freedom for the residual deviance. Degrees of freedom with Null Deviance and Residual deviance as can be calculated differs by 2 ($99-97$) as the model has two variable amount and price. This means only 2 additional parameters were used to calculate this model. Therefore, only two degrees of freedom are used.

The next is the AIC (Akaike Information Criterion) value at 133.6. Lower the AIC the better the model.

For the second hypothesis from Table 6 we can see all the data collected in the survey. For this hypothesis we will be using the profitability variable as the dependent variable. Since the whole purpose of this research is to understand whether demographics impact the profitability in Crypto currency market or not. So, the independent variables will be age group, gender, income and race.

AgeGroup	Gender	Income	Race	TradedCrypto	NameCurrency	Profitability
0	0	3	0	1	Ethereum	0
0	0	3	0	1	Ripple	0
1	0	4	2	1	Ethereum	0
1	0	3	0	1	Ethereum	0
1	1	3	0	1	Litecoin	0
0	1	3	0	1	Ethereum	0
0	0	4	0	1	Ripple	0
0	0	3	0	1	Litecoin	1
0	1	4	0	1	Litecoin	1
0	1	4	0	1	Bitcoin	0

Table 6. Shows the survey result conducted for this project. The number of participants is 10

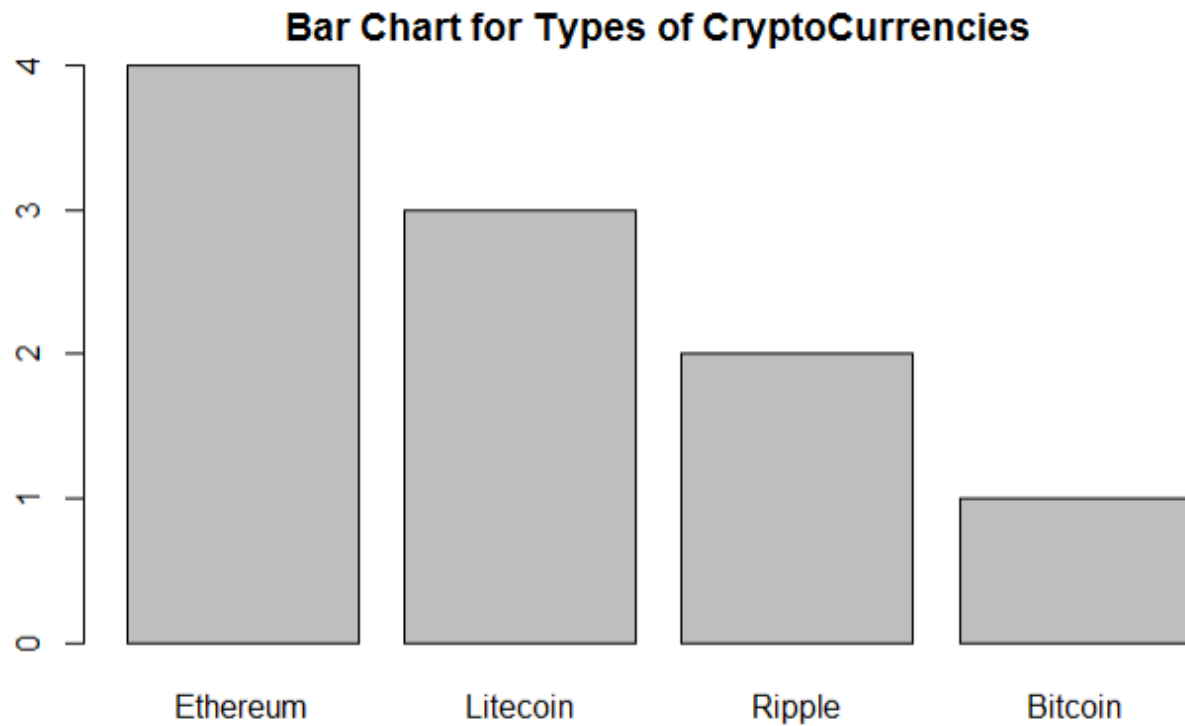


Figure 3. The chart shows the data from the survey conducted. It shows that the major crypto currency investment done by the 10 participants. Unlike the market trend for Bitcoin which is only purchased by one of the participant for 4 participants the most mainstream share was Ethereum.

AgeGroup	Gender	Income	Race	TradedCrypto	NameCurrency	Profitability
0:7	0:6	3:6	0:9	Min. :1	Bitcoin :1	Min. :0.0
1:3	1:4	4:4	2:1	1st Qu.:1	Ethereum:4	1st Qu.:0.0
				Median :1	Litecoin:3	Median :0.0
				Mean :1	Ripple :2	Mean :0.2
				3rd Qu.:1		3rd Qu.:0.0
				Max. :1		Max. :1.0

Table 7. The Table shows the summary of all the variables from the survey conducted.

The survey data set as shown in Table 7 includes 7 variables. Starting from the gender variable there are 6 male and 4 female participants. For the income variable 6 participants belong to the income range of '\$50,000 - \$74,999' and 4 participants to the income range of '\$75,000 plus'. In race variable 9 of my participants are Asian and one is Caucasian. 7 participants are from 25-35 years age group and the remaining are from 35-45 years age group. In all it can be said that general age of the participants is younger.

For the purpose of logit regression, I have converted age group, income race and gender variables to factor for ease of calculating the regression analysis as these are categorical independent variables.

We will now use the logistic regression model using the glm (generalized linear model) function. In order to get the result of the regression we use the summary command and get the result shown in Table 8.

```

call:
glm(formula = Profitability ~ AgeGroup + Gender + Income + Race,
     family = binomial(link = "logit"), data = survey)

Deviance Residuals:
    1      2      3      4      5      6      7
-0.73395 -0.73395 -0.00008 -0.00007 -0.00009 -0.83043 -0.83043
    8      9
 1.69907  1.44080
   10
-0.93513

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -1.1741     1.2645  -0.928   0.353
AgeGroup     -18.5412    7587.8338  -0.002   0.998
Gender1         0.2867     1.8511   0.155   0.877
Income4         0.2867     1.8511   0.155   0.877
Race2         -0.1374    13161.4597   0.000   1.000

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 10.0080  on 9  degrees of freedom
Residual deviance:  8.2938  on 5  degrees of freedom
AIC: 18.294

Number of Fisher Scoring iterations: 18

```

Table 8. Table shows logit regression summary for the survey data set

As can be seen in Table 8 the first item shown is the formula R programing used to fit the data.

The formula shows the predictor variables ageGroup, Gender, Income and Race. Along with that it shows the dependent/ target variable called Profitability showing whether participants crypto currency trading is profitable or not.

The next item in the output of Table 5 is deviance residuals which helps in measuring model of fit. The part shows the distribution of the deviance residuals in relation to individual cases of the model.

The next part of the summary output shows coefficients including the standard error, z-statistics and associated p-values. Theoretically coefficients in terms of regression are of two unknown constant that show a linear model of intercept and slope. Ultimately the goal is to get a fitted line that is as close to the other 10 observations of the data set. Coefficients of all the

independent ageGroup, Gender, Income and Race are statistically insignificant. For every one unit in ageGroup 1 (range 35-45 years) versus other age groups changes odds of profitability by -18.54. Similarly, for gender for male versus female the odds ratio is 0.0286 and same is odds ration for Income greater than \$75,000+.

The coefficient estimate from Table 8 contains of five rows and the first one is the intercept. In short in our example the value of ageGroup, Gender, Income and Race it takes for a profitable crypto market profile. The intercept for this data set is -1.174 means the participants have profitable portfolio is there is age group 1 (range 35-45 years), gender 1(female), income 4(\$75,000+) and race 2 (Caucasian). The second row onwards shows the slope of the coefficients or in our survey hypothesis the value of ageGroup, Gender, Income and Race. It means with odds of male profitably is 0.028.

The standard error measure the average amount the estimate of coefficients is different from the response variable actual value. Needless to say, ideally, we would want this error value to be as lower as possible as otherwise it shows the issue with coefficients. In the present Table 8 I would like to conclude that the current coefficient for age is very high at -18.54 and that can vary by 7587. Similarly, for race the current coefficient is - 0.137 and it can vary by 13161.

The p-value shown in the Table 8 as column $\Pr(>|z|)$ is the probability of measuring any value equal to or larger than the z value. Typically, the p-value to be less than equal to 5% to show any significance. In our scenario all p-values are greater than 5%. Hence, we accept the null hypothesis as can be inferred none of the demographics are impacting the profitability of a crypto currency profile. I am guessing because the number of participants is so low that is why it is impossible to derive and valuable conclusions from this logistic regression.

The deviance measures in the Table 8 the goodness of fit of a generalized linear model (Gupta, 2017). As in the article the higher the deviance goes the poorer the model fit is. The null deviance is 10.008 on 9 degrees of freedom. The null deviance only shows the performance of the model when it is governed by the intercept. The residual deviance in the next sentence shows the model when the independent variable age, race income and gender are included in the model. For our model the residual deviance is 8.293 on 5 degrees of freedom. The lower level of residual deviance than null deviance shows that the model is better after adding the two variable price and amount. For the degrees of freedom in Null deviance is 9 which is $N-1$. So, for this model $N=10$ and therefore the Degree of freedom for Null Deviance is $10 - 1 = 9$. For calculating the residual deviance, the formula is $N-K-1$. K being the number of variables. Since in this model we have 4 variables that is why $10-4-1 = 5$ degrees of freedom for the residual deviance. Degrees of freedom with Null Deviance and Residual deviance as can be calculated differs by 4 ($9-5$) as the model has four variable age, race gender and income. This means only 4 additional parameters were used to calculate this model. Therefore, only four degrees of freedom are used. The next is the AIC (Akaike Information Criterion) value at 18.294. Lower the AIC the better the model.

In the third hypothesis is time does not impact the bitcoin values. For this analysis we get our data from bitcoin charts website. The data set that we get from the website contains 3 variables. One is a UNIX time column and other two are Bitcoin price and amount as can be seen in Table 9. Then by converting the UNIX time we get date column. Also create a count function to number of times a date showed up in the data set. The next variable we create is the value which is a calculated field. The formula for value is $\text{price} * \text{amount}$. All the variables and 2 rows

of their values are shown in Table 10. In total there are 2453 observations and 5 variables as shown in Table 11.

	unixtime	price	amount
1	1315922016	5.80	1
2	1315922024	5.83	3

Table 9. Shows the 3 variables in the data set obtained from the bitcoin charts website.

	date	count	m_value	m_price	m_amount
	<date>	<int>	<dbl>	<dbl>	<dbl>
1	2011-09-13	12	5.87	4.86	28.8
2	2011-09-14	14	5.58	4.37	24.4

Table 10. Shows the 5 variables in the data set. The 3 new variables created are date, count and value. Unixtime variable is dropped

Classes 'tbl_df', 'tbl' and 'data.frame':	2453 obs. of 5 variables:
\$ date	: Date, format: "2011-09-13" "2011-09-14" "2011-09-15" ...
\$ count	: int 12 14 6 4 1 8 1 17 5 5 ...
\$ m_value	: num 5.87 5.58 5.12 4.83 4.87 ...
\$ m_price	: num 4.86 4.37 13.36 9.98 0.3 ...
\$ m_amount	: num 28.84 24.42 68.04 48.44 1.46 ...

Table 11. Shows the details and variable type of the data set and its variables

Then by using the date and value variable we create a time series blot as shown in Figure 4.

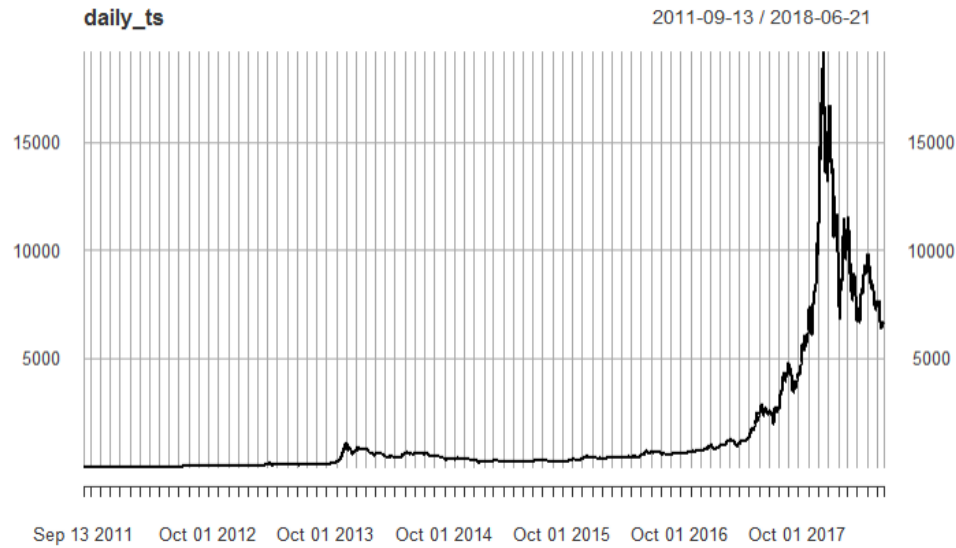


Figure 4. Showing time series for date range 13/09/2013 to 21/06/2018. The time series plot is using date and value variables.

Next, we have taken log of the time series with date range 13/09/2013 to 21/06/2018 as shown in Figure 5. We can take log of a time series by using the log function. One way to correct problems created by linear regression is log. Log time series is helpful in seeing percentage change of the values. As can be seen there is a lot of fluctuation in the beginning of the time series and the later on the trend is not that fluctuating in comparison to early time period.

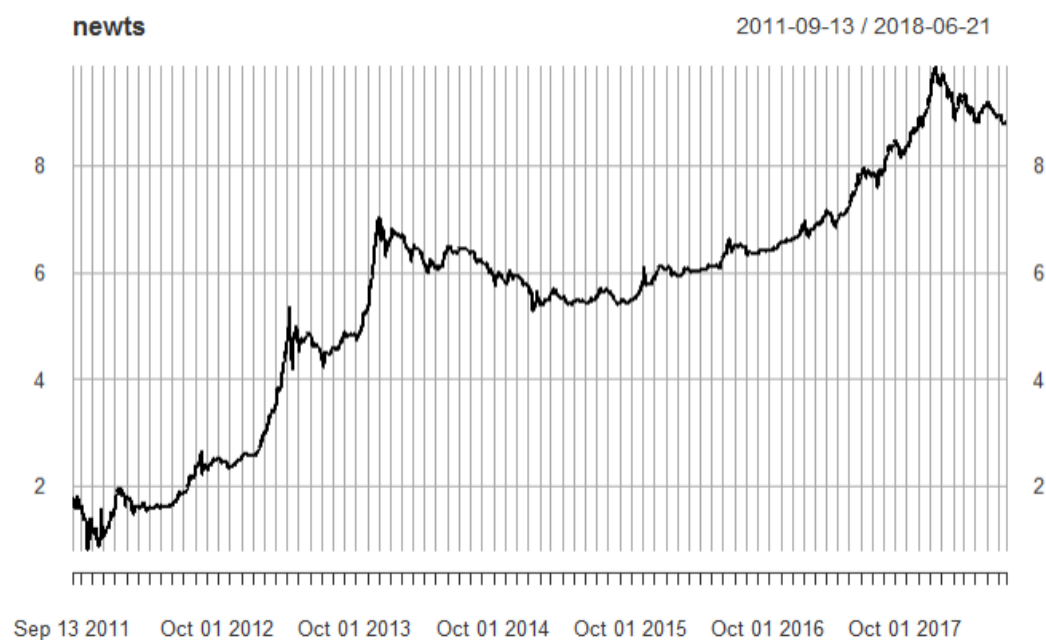


Figure 5. Show the log of the time series.

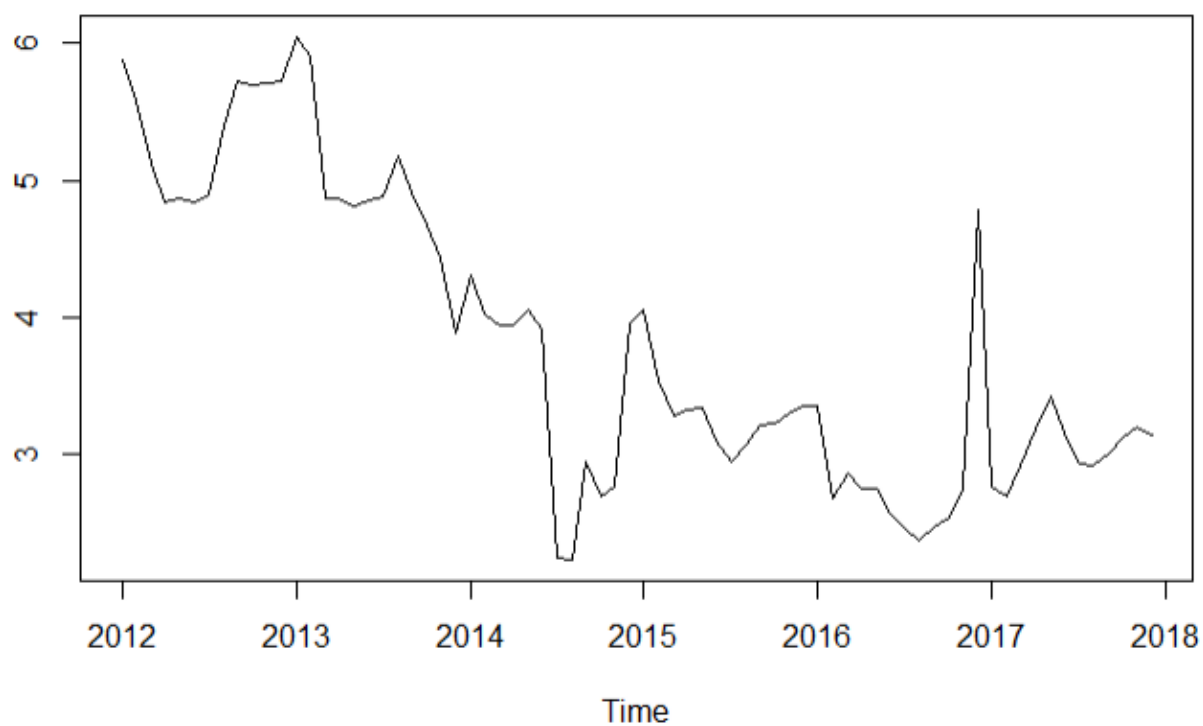


Figure 6. Shows an early time series for the date range January 2012 to December 2017.

The time series in Figure 6 only runs from January 2012 to December 2017.

Whereas time series for Figure 4 runs from September 2011 to June 2018. This is done intentionally so that when forecasting using the first dataset we can see whether the forecast is accurate or not by comparing it to the second dataset time series plot. Since the time series is showing an exponential trend that is why using HoltWinter() function in the installation.

Using the forecasting package, we try and forecast the next three time periods for the January 2012 to December 2017 data set as shown in Table 12.

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2018	2.672097	1.997656	3.346538	1.640629	3.703565
Feb 2018	2.628337	1.844973	3.411702	1.430285	3.826390
Mar 2018	2.571115	1.692224	3.450006	1.226967	3.915263

Table 12. Shows the forecast of January, February and March values column of 2018

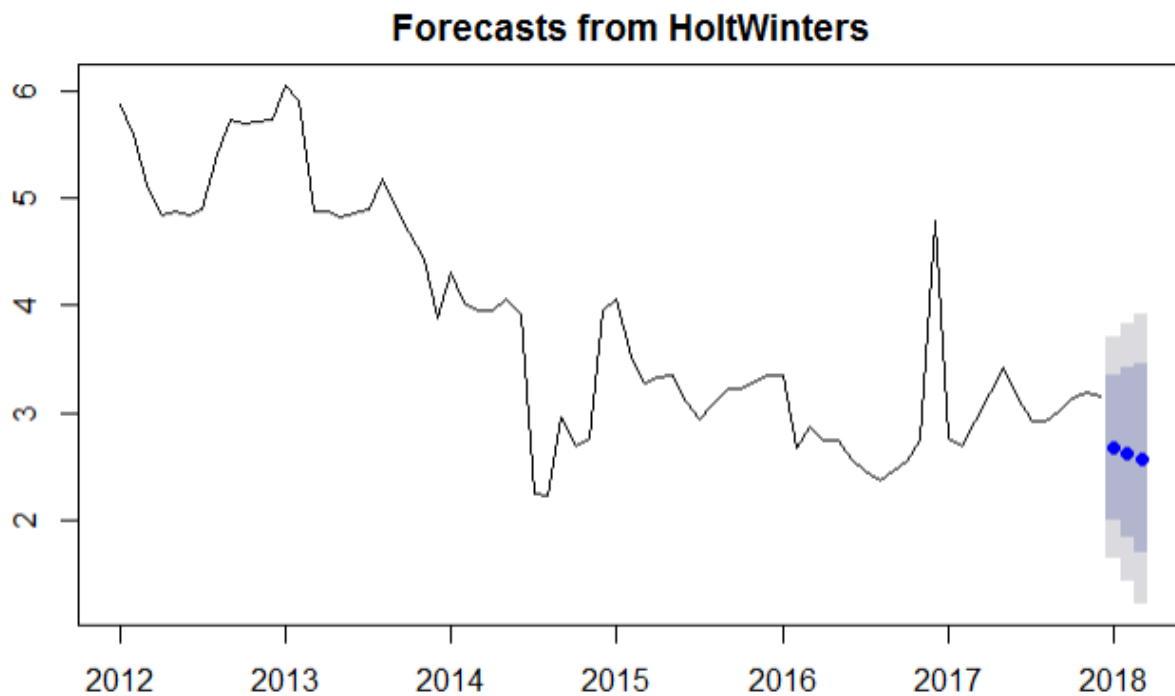


Figure 7. The three blue dots show the forecast for the first 3 months of the year 2018

By visually comparing Figure 4 and Figure 7 it can be inferred that first three months of the 2018 has an upward trend in both the time series plots. At the end as can be seen bitcoin value is not impacted by seasonal or cyclical trends since the growth is exponential.

References

Arpan Gupta (2017). Deviance and AIC for Logistic Regression in R. *Indian Institute of Roorkee, India*

Bakar, N. A., & Rosbi, S. (2017). High Volatility Detection Method Using Statistical Process Control for Cryptocurrency Exchange Rate: A Case Study of Bitcoin. *The International Journal of Engineering and Science*, 6(11), 39-48.

Bakar, N. A., & Rosbi, S. (2017). Autoregressive Integrated Moving Average (ARIMA) Model for Forecasting Cryptocurrency Exchange Rate in High Volatility Environment: A New Insight of Bitcoin Transaction. *International Journal of Advanced Engineering Research and Science*, 4(11), 130-137.

—[Barker, J. T. \(2017\). *Why Is Bitcoin's Value So Volatile?*](#)

Blumenthal, S. (2017). *Cryptocurrency: Understanding The Concept Of Cryptocurrency, Blockchain And Bitcoin-The Simple Introduction To Internet Money, It's Benefits And What You Need To Know About Investing*. New York, NY: CreateSpace Independent Publishing Platform

Bohr, J., & Bashir, M. (2014, July). Who uses bitcoin? an exploration of the bitcoin community. *Privacy, Security and Trust (PST), 2014 Twelfth Annual International Conference, IEEE*, 94-101.

Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, technology, and governance. *Journal of Economic Perspectives*, 29(2), 213-38.

Brennan, M. J. (1975). The optimal number of securities in a risky asset portfolio when there are fixed costs of transacting: Theory and some empirical results. *Journal of Financial and Quantitative analysis*, 10(3), 483-496.

Bresett, M. (2017). *Cryptocurrency: Bitcoin, Ethereum, Blockchain The Ultimate Guide to Understanding the Cryptocurrency Revolution*. New York, NY: CreateSpace Independent Publishing Platform

Burniske, C., & Tatar, J. (2018). *Cryptoassets: The Innovative Investor's Guide to Bitcoin and Beyond*. San Francisco, CA: McGraw-Hill.

Casey, M. J., & Vigna, P. (2015). Bitcoin and the digital-currency revolution. *The Wall Street Journal*, 23, 28-37.

Catania, L., Grassi, S., & Ravazzolo, F. (2018). Predicting the Volatility of Cryptocurrency Time-Series. *BI Norwegian Business School, Centre for Applied Macro- and Petroleum Economics*, 3, 8-15.

Chan, S., Chu, J., Nadarajah, S., & Osterrieder, J. (2017). A statistical analysis of cryptocurrencies. *Journal of Risk and Financial Management*, 10(2), 12.

Conley, J. P. (2017). *Blockchain and the Economics of Crypto-tokens and Initial Coin Offerings*. Vanderbilt University Department of Economics, VUECON-17-00008.

Corbet, S., Larkin, C., Lucey, B., & Yarovaya, L. (2018). *KODAKCoin: a blockchain revolution or exploiting a potential cryptocurrency bubble?*

Corbet, S., Larkin, C., Lucey, B., Meegan, A., & Yarovaya, L. (n.d.). *The volatility generating effects of macroeconomic news on cryptocurrency returns*, 15

DeVries, P. D. (2016). An Analysis of Cryptocurrency, Bitcoin, and the Future. *International Journal of Business Management and Commerce*, 1(2), 2-9.

Fishburn, P. C., & Burr Porter, R. (1976). Optimal portfolios with one safe and one risky asset: Effects of changes in rate of return and risk. *Management Science*, 22(10), 1064-1073.

Fry, J., & Cheah, E. T. (2016). Negative bubbles and shocks in cryptocurrency markets. *International Review of Financial Analysis*, 47, 343-352.

Gandal, N., & Halaburda, H. (2016). *Can We Predict the Winner in a Market with Network Effects? Competition in Cryptocurrency Market*, 2, 10

Garber, P. M. (1990). Famous first bubbles. *Journal of Economic perspectives*, 4(2), 35-54.

Guides, T. S. (2018, Feb 22). The Best Bitcoin Trading Strategy—5 Steps to Profit (Updated). Retrieved from <https://tradingstrategyguides.com/best-bitcoin-trading-strategy/>

Heyde, C. C. (1999). A risky asset model with strong dependence through fractal activity time. *Journal of Applied Probability*, 36(4), 1234-1239.

Jeong Hun Oh, K. N. (2018). *The Growing Role of Cryptocurrency: What Does It Mean for Central Banks and Governments?*

Katsiampa, P. (2017). Volatility estimation for Bitcoin: A comparison of GARCH models. *Economics Letters*, 158, 3-6.

Kutis, R. (2014). Bitcoin-Light at the End of the Tunnel for Cyber-Libertarians. *Masaryk UJL & Tech.*, 8, 209.

Li, X., & Wang, C. A. (2017). The technology and economic determinants of cryptocurrency exchange rates: The case of Bitcoin. *Decision Support Systems*, 95, 49-60.

McCormack, P. (2017). *An Introduction to Crypto Technical Analysis — Using Trend Lines.*
Retrieved

Nakamoto, S. (2008). *Bitcoin: A peer-to-peer electronic cash system* [PDF document]. Retrieved from
https://s3.amazonaws.com/academia.edu.documents/32413652/BitCoin_P2P_electronic_cash_system.pdf

Nica, O., Piotrowska, K., & Schenk-hoppé, K. R. (2017). *Cryptocurrencies : Economic benefits and risks, 1–56.*

Oreso, T. (2018). *Six Factors That Can Influence Cryptocurrency Price*

Osterrieder, J., Lorenz, J., & Strika, M. (2016). Cryptocurrencies, Their Statistical Properties and Extreme Tail Behaviour. *Advanced Risk & Portfolio Management Paper*, 6, 86-97

Raymaekers, W. (2015). Cryptocurrency Bitcoin: Disruption, challenges and opportunities. *Journal of Payments Strategy & Systems*, 9(1), 30-46.

Sams, R. (2015). *A Note on Cryptocurrency Stabilisation: Seigniorage Shares*. London, UK: Clearmatics

Stroukal, D., & Nedvěďová, B. (2016). Bitcoin and Other Cryptocurrency As an Instrument of Crime in Cyberspace. *Proceedings of 4th Business & Management Conference*, 12

Urquhart, A. (2017). The volatility of Bitcoin. *Centre for Digital Finance*, 3, 12-18

Whelan, Karl (2018). *How Is Bitcoin Different From The Dollar?*

Witheridge, G. (2018, Feb 21). Once Bit (coined), No longer cryptocurrency shy. Retrieved from <https://social.shorthand.com/GivernyW/32S3Mp7iK6/once-bitcoined-no-longer-cryptocurrency-shy>

Yeoh, P. (2017). Regulatory issues in blockchain technology. *Journal of Financial Regulation and Compliance*, 25(2), 196-208.