International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2018; 3(3): 120-125

www.mathsjournal.com
Received: 15-03-2018
Accepted: 16-04-2018

Suni Ajay Kumar
Department of Statistics,
Kristu Jayanti College,
Bangalore, Karnataka, India

Correspondence
Suni Ajay Kumar
Department of Statistics,
Kristu Jayanti College,
Bangalore, Karnataka, India

A regression model for crypto-currency price

Suni Ajay Kumar

Abstract
The aim of the paper is to fit a regression model which can be used commonly for the four important crypto currencies: Bitcoin, Litecoin, Ethereum and Ripple to predict the prices. The data has information over the past six years regarding price, transaction volume, transaction count, exchange volume, generated coins etc of these currencies. Understanding the dynamics of crypto currency market can help to a certain extent to take wise investment decisions. Among the variables under consideration the study revealed that transaction volume can be used as an influencing variable to fit a quadratic regression model and predict the prices of the crypto currencies.

Keywords: Bitcoin, litecoin, ethereum, ripple

Introduction
Good money should be durable, portable, divisible and must have intrinsic value. Even though the preferred medium of exchange was gold originally, governments were forced to create more accessible medium of exchange. Thus fiat money came into existence. Major disadvantages of fiat currency are the unlimited ability of government to print up money whenever they want, loss of value over time and its potential to drop to zero value. In order to fix some of these issues, crypto currencies began to emerge in 2009 leveraging a disruptive technology called block chain. A crypto currency is a digital currency that uses cryptography for security. Block chain specifically deals with the way in which data is structured and allows for the existence of decentralized digital ledgers where single organizations are not able to effect transactions. Block chain can be used to overcome the disadvantages of any data transfer. It is transparent, fast, cheaper and more easy to use. When block chain technology is used to send some valuable information, transmission cannot be changed or falsified since it is confirmed by thousands of computers around the world. This network also contains multiple copies of your information which can be verified by any user at any time. Block chain provides information of who owns what. So any participant can make sure of your financial solvency any time. All this information is protected by heavy duty encryption. Crypto currencies are physical pre computed files utilizing a public key/private key pairs generated around a specific encryption algorithm. These key pairs are stored in a file named “wallet.dat” which resides in a default hidden directory on the hard drive of the owner. The private keys are sent to users using dynamic wallet addresses generalized by the users engaged in transactions. The public key of the crypto currency key pair is the destination payment address. The supply/demand of the coins and the fluctuating difficulty levels required for mining each coin, assign the value of each unit. If a user loses their wallet.dat file the currency is lost. The transaction volume in crypto currencies is estimated to exceed 100 million USD per day. A block chain is a decentralized, digitalized, public ledger of all crypto currency transactions. Block chain technology can be used to solve other complex problems. There is also a need to balance privacy regulatory requirements.

Types of cryptocurrencies
Bitcoin (BTC) is the first crypto currency to emerge based on the SHA-256 algorithm. This virtual commodity was conceptualized in a white paper written in 2009 by a pseudonymous author who went by the name Satoshi Nakamoto. The market price of Bitcoin fluctuated from below 0,01 USD to over 250 USD over the first four years. The market volatility made it less

attractive for long term investors and daily users. But highly volatile price made it an attractive investment alternative. The smallest increment of a Bitcoin is known as a Satoshi. The increment can be as small as 0.00000001 BTC per transaction. There is a limit to the total amount of Bitcoin will ever be created, hence the rise in the value of BTC is anticipated. Once the Bitcoin block chain is completed the users can only circulate the coin that still exists on the network. In increasing the value of Bitcoin the principles of supply and demand economics will come into play. Bitcoin will only have 21 million coins mined. Bit coin can be interpreted as being the "gold standard" of crypto currency. All other crypto currency market prices are matched to the price of BTC.

Litecoin (LTC) can be considered as the "silver standard" of crypto currency. It has been the second most adopted crypto currency. Litecoin is based on encryption algorithm. The goals of Litecoin are to have transaction faster than Bitcoin and to make use of an algorithm that was resistant to accelerated hardware mining technologies such as ASIC. The total amount of Litecoin that is available for mining and circulation is four times as the amount of BTC. Litecoins will have only 84 million coins mined. It was launched in 2011.

Altcoins (alternative crypto currencies) are built up on the basic frame work provided by BTC or LTC. They make use of SHA-256 or Script encryption algorithm and feature their own unique properties. Eg: Ethereum, Ripple, Feather coin, P2 Pcoin, Terra coin.

Ethereum (ETH) is a decentralized computing platform for the building and development of decentralized application, featuring smart contract functionality. Ether is a fuel for operating the distributed application platform. It is a unique code that can be used as a way for running the application or a program. Ethereum is officially launched in 2015 and has attracted interest from many developers and institutional actors. Ether can be transferred between accounts and used to compensate participant mining nodes for computations performed. Ethereum allows to create smart contracts, autonomous applications stored in block chain.

Ripple (XRP) is the only crypto currency that does not have a block chain but uses a global consensus ledger. It is a real time gross settlement system, currency exchange and remittance network. Ripple is based on a shared public database that makes use of a consensus process between those validating servers to ensure integrity. Those validating servers can belong to individuals or banks. Ripple facilitates transactions all over the world and transfer fees are less with no typical confirmation time. This network was launched in 2012.

Mining Crypto currency

Mining is the use of computational power to process transactions for a crypto currency block chain in order to receive a reward of crypto currency for the effort. Solo miners are able to use the computational powers of their CPU or GPU to process transaction for the crypto currency network on their own. They would receive a full payment for a completed block chain. But the chances of repeatedly completing a block with a valid share submission are minimal due to increasing difficulty rate.

A pool is software hosted on a web server usually a VPS or dedicated server. Pool miners create accounts on the pool server and then add pool authentication credentials to the configuration files of their mining client software on local mining equipment. Once the mining client authenticates it is able to share resources in the distributed computing network that will allow for a more efficient use of hardware. The mining pool server will receive reward payments from the crypto currency network and distribute the payment to miners based on the amount of the miner's computational effort accepted by the network.

Stratum is used to synchronize the computational effort of multiple mining rigs to reduce the chances of duplicate share submission, there by maximizing efficiency of the miner's combined resources. CPU mining became a non-profitable method as the difficulty rates increased. GPU accelerated mining is the most common method of majority of crypto currencies. One single GPU is able to process equivalent of 44.64 CPUs approximately. ASIC mining have been developed for Bitcoin. ASIC mining is having exponentially more computational processing power using significantly less resources such as hardware and electricity than GPU mining.

Economic activity with crypto currency

Crypto currencies have provided a way for businesses to engage in e-commerce on an international level. Merchant processing solutions such as Bit pay provide a solution for merchants who wish, to be able to convert their payments into USD and withdraw to a bank account while maintaining compliance with government regulatory authorities. Services like Gyft provide a way for miners and traders to use their Bit coin by purchasing gift cards for major retail stores, restaurants, hotels and many other types of merchants. The gift cards are sold as electronic codes and are instantly redeemable on cell phones or through print outs. The only way to purchase card from Gyft website is via Bitcoin.

Attacks against crypto currency

The target of any organized criminal will be location of money. In the case of crypto currencies the location of money are in the form of mining pool servers, trading platforms, third party wallet services and end user computers. Each value location has experienced multiple forms of attack that resulted in direct theft of coins. Exfiltration of wallet.dat file through physical access which is most often contributed malware is another area of attack against crypto currency. Unskilled criminals will end up being burned by improper use of features, while more sophisticated criminals will use the properties of anonymity to their advantage, while still bearing the risk of market fluctuations. Crypto currency has a long way to go before it is refined as a commodity suitable for daily commercial use. RBI had cautioned the users, holders and traders of virtual currencies about the financial, operational, legal customer protection and security related risks that they are exposing themselves

Literature review

Adam Hayes (2015) [9] aims to identify the likely sources of value that crypto currencies exhibit in the market place using cross sectional empirical data examining 66 of the most used coins. He estimated a regression model that points to three variables: the difficulty in mining for coins, the rate of unit production and the cryptographic algorithm employed. Using the analysis a cost of production model is proposed for valuing bit coin. They concluded that cost of production drives value and anything that serves to reduce the cost of bit coin production will tend to have a negative influence on its price. Also when the Bitcoin block reward halves, it will effectively increase the cost of production over night.

Ryan Farell (2015) [6] in his paper provides an overview of the industry, brief history of digital currencies, in-depth analysis of coin economics, an overview of factors affecting industry growth. He concluded by acknowledging the industry's rapid growth in the number of coins in circulation and its creativity in implementing workable solutions to deficiencies in the development of new coins.

Alexander D'Alfonso *et al*. (2016) [3] analysed two highly disruptive crypto currencies to develop the ideal investment strategy for a $1000,000 investment for five years without any additional trading. They extrapolated the values to form a five year projection. They ran simulations to predict the expected values. They concluded by saying that the approximate investment ratio is 69:31 for Bitcoin and Ethereum respectively.

Jan Lansky (2016) [8] created a database which includes 60917 records with prices of 1278 crypto currencies from 28 April 2013 to 3 July 2016 and analyzed the crypto currencies by their length of existence, the biggest price drops and highest price increases of crypto currencies. When they compiled an overview of crypto currencies they realized that the current price of crypto currency is at least 10 times of their minimum price

Shipra Sarawat *et al*. (2017) [5] in their paper has a comprehensive knowledge of crypto currency. They concluded that this industry has shown growth in the section of its price and in the number of coins that are currently in circulation.

Darius Tajanko *et al*. (2017) [7] investigates various applications of machine learning, econometrics and general data science within the crypto currency space. They predicted returns from data driven analysis in market finance, legislative policy and online content management.

Analyzed statistical properties of the largest crypto currencies. They characterized their exchange rates versus the US dollar by fitting parametric distributions to them They clearly found that no single distribution fits well jointly to all the crypto currencies returns.

Peter M Krafft *et al*. (2018) [10] conducted an online experiment to study how susceptible traders in these markets are to peer influence from trading behavior. They created bots that executed over one hundred thousand trades costing less than a penny each in 217 crypto currencies over the course of six months. They found that individual "buy" actions led to short- term increases in subsequent buy-side activity hundreds of times the size of their interventions.

Data

The data used in this paper is from coinmatrics.io website. They provide time series data of crypto currencies related to transaction volume, transaction count, exchange volume, generated coins, price etc from 1st May 2013 to 3rd May 2018. Summary statistics of the prices of the currencies are given below.

Table 1: Summary statistics of daily exchange rates of Bitcoin, Litecoin, Ethereum, and Ripple versus the U.S. Dollar from 1 May 2013 until 3 May 2018.

Statistics	BTC	LTC	ETH	XRP
Minimum	68.5	1.15	0.431589	0.002809
Q1	281.99	3.15	8.44	0.00613
Median	527.48	3.96	12.755	0.008035
Mean	1769.707037	25.88737562	180.9329073	0.124877842
Mode	418.42	3.84	1.22	0.007857
Q3	1003.52	16.28	297.35	0.024441
Maximum	19475.8	359.13	1397.48	3.36
Skewness	2.829906584	3.173556573	1.851237109	4.753492944
Kurtosis	7.936252211	10.29425504	2.87815708	28.8283462
S.D	3234.709475	53.81598848	284.2431916	0.340562848
C.V	182.78	207.8850683	157.0986704	2.727167943
Range	19407.3	357.98	1397.048411	3.357191
IQR	721.53	13.13	288.91	0.018311

From table 1, it can clearly be seen that all the four crypto currencies are positively skewed with XRP being the most skewed. In terms of kurtosis, ETH rates are showing less peakedness than a normal distribution, and others are highly peaked than normal distribution with XRP showing the maximum peakedness. The exchange rates of XRP are showing more consistency than the other coins, indicating that their low volatility can perhaps be explained by the low values of the exchange rates and the fact that their range and interquartile ranges are very limited. LTC is showing the most inconsistent behavior followed by BTC and ETH. The non-normal behavior of the data is evident from these summary statistics.

An attempt is made to fit a common regression model which can be used for predicting the exchange rates of the four selected crypto currencies BTC, LTC, ETH and XRP. Linear regression is ruled out since the data is nonrandom, auto correlated and hetroscedastic in nature.

Run test is used to test for randomness of the data: Code, values above the median as positive and values below the median as negative. A run is defined as a series of consecutive positive or negative values. The runs test is defined as:

H_0: The sequence was produced in a random manner

H_1: The sequence was not produced in a random manner

Test statistic:

$$Z=\frac{(R-\bar{R})}{S_R}$$

Where R is the observed number of runs, $\bar{R}$ is the expected number of runs, and S_R is the standard deviation of the number of runs. The values of $\bar{R}$ and S_R are computed as follows:

$$\bar{R} = \frac{2n_1n_2}{n_1+n_2} + 1 \text{ and } S_R^2 = \frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}$$

Where $n_1$and n_2 are the number of positive and negative values. The runs test rejects the null hypothesis if $|\boldsymbol{Z}| > \boldsymbol{Z}_{1-\alpha/2}$, α is the level of significance.

Run test for the data showed that the prices of all the four currencies under study are random. The results are listed below.

Table 2: BTC

Runs Test	
	Price (USD)
Test Value[a]	527.4800
Cases < Test Value	914
Cases > = Test Value	915
Total Cases	1829
Number of Runs	10
Z	-42.358
Asymp. Sig. (2-tailed)	.000
a. Median	

LTC

Runs Test	
	Price (USD)
Test Value[a]	3.9600
Cases < Test Value	911
Cases >= Test Value	918
Total Cases	1829
Number of Runs	65
Z	-39.785
Asymp. Sig. (2-tailed)	.000
a. Median	

XRP

Runs Test	
	Price (USD)
Test Value[a]	.0080
Cases < Test Value	865
Cases >= Test Value	866
Total Cases	1731
Number of Runs	76
Z	-38.011
Asymp. Sig. (2-tailed)	.000
a. Median	

ETH

Runs Test	
	Price (USD)
Test Value[a]	12.7550
Cases < Test Value	499
Cases >= Test Value	499
Total Cases	998
Number of Runs	22
Z	-30.277
Asymp. Sig. (2-tailed)	.000
a. Median	

Autocorrelation is the correlation of signal with a delayed copy of itself as a function of delay. Durban-Watson test is used to find the presence of first order serial correlation in the data. The Durban-Watson statistic is a number that tests for autocorrelation in the residuals from a statistical regression analysis.

Let the regression model be

$Y_t = \beta_1 + \beta_2 X_{t2} + \beta_3 X_{t3} + \beta_4 X_{t4} + \ldots\ldots B_k X_{tk} + U_t$

Where $U_t = \rho U_{t-1} + \varepsilon_t; -1 < \rho < 1$; $\widehat{U_t} = Y_t - \widehat{\beta_1} - \widehat{\beta_2} X_{t2} - \widehat{\beta_3} X_{t3} - \widehat{\beta_4} X_{t4} - \ldots\ldots \widehat{\beta_k} X_{tk}$

H_0: The error terms are not autocorrelated.

H_1: The error terms are positively or negatively auto correlated.

Compute $d = \frac{\sum_{t=2}^{n}(\widehat{U_t} - \widehat{U_{t-1}})^2}{\sum_{t=1}^{n} U_t^2}$

If $d < dL_a$, we reject the null hypothesis,

If $d > dU_a$, we do not reject the null hypothesis

If $dL_a < d < dU_a$, the test is inconclusive.

Since, $d \approx 2(1-\hat{\rho})$, When d=2, ρ =0, When d = 4, ρ = -1, and when d=0, ρ =1

The tests confirmed that there is first order serial correlation present in the data.

Table 3: Durbin-Watson test results

	DW	p-value
BTC	0.28944	< 2.2e-16
LTC	0.28755	< 2.2e-16
ETH	0.22576	< 2.2e-16
XRP	0.24167	< 2.2e-16

Heteroscedasticity is a systematic change in the spread of the residuals over the range of measured values. Heteroscedasticity is a problem because ordinary least square (OLS) regression assumes that all residuals are drawn from a population that has a constant variance (homoscedasticity). Breusch-Pagan test is used for testing hetroscedasticity in the data used here.

For a regression model $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots\ldots \beta_p X_p + U$

If the data is homoscedastic then Var (U/X) $= \sigma^2$,

If the data is hetroscedastic then Var (U/X) $= \sigma^2 f(X)$

$= \sigma^2 (\alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \ldots\ldots \alpha_p X_p$

$\widehat{U}^2 = \delta_0 + \delta_1 X_1 + \delta_2 X_2 + \ldots\ldots \delta_p X_p$

H_0: $\delta_1 = \delta_2 = \delta_3 = \ldots\ldots = \delta_p = 0$

H_1: At least one of them is not equal to zero

$$F = \frac{R^2/p}{(1-R^2)/(N-P-1)} \sim F_{P,\, N-P-1}$$

Or, $LM = NR^2 \sim \chi_P^2$

Test revealed that there is hetroscedasticity present in the data.

Table 4

Breusch Pagan test			Koenkar test	
Crypto currency	**Chi-square**	**Significance level**	**Chi square**	**Significance level**
BTC	3099.422	0.0000	510.012	0.0000
LTC	3098.934	0.0000	483.564	0.0000
ETH	343.437	0.0000	97.576	0.0000
XRP	7200.684	0.0000	468.803	0.0000

Since linearity assumptions are not satisfying for the data, decided to proceed for fitting a polynomial regression. Among the variables considered for fitting a regression model, on-chain transaction volume which is a measure of the total value of outputs on the block chain on a given day seemed to be appropriate for fitting a polynomial regression. This variable is the value of USD circulates on the blockchain in a day.The statistical analysis and testing conducted here revealed that the other variables like Transaction count which is the number of transactions happening on the public blockchain in a day, Exchange volume which is the dollar value of the volume at exchanges like GDAX and Bitfinex, and generated Coins which is the number of new coins that have been brought into existence on that day, are not capable of fitting a single regression model jointly to all the four selected crypto currencies.

Correlation coefficients between transaction volume and exchange rates for the currencies are found to be significant and the values are given in the following table.

Table 5

BTC	LTC	ETH	XRP
0.868899	0.6784897	0.7641732	0.3469863

Empirical results of regression analysis

Second degree polynomial regressions out puts for the four crypto currencies are given below.

BTC

Coefficients

	Estimate	Std. Errort	value	Pr (>|t|)
(Intercept)	9.061e+01	4.080e+01	2.221	0.0265 *
Tx Volume	7.925e-07	1.452e-08	54.562	<2e-16 ***
(tx Volume)^2	-1.002e-17	4.846e-19-	20.669	<2e-16 ***

Residual standard error: 1442 on 1826 degrees of freedom
Multiple R-squared: 0.8014, Adjusted R-squared: 0.8012
F-statistic: 3685 on 2 and 1826 DF, p-value: < 2.2e-16

LTC

Coefficients

	Estimate	Std. Error	t value	Pr (>|t|)
(Intercept)	9.217e+00	.516e-01	9.686	<2e-16 ***
Tx Volume	8.154e-08	2.217e-09	36.774	<2e-16 ***
(tx Volume)^2	-5.508e-18	3.456e-19	-15.936	<2e-16 ***

Residual standard error: 37.06 on 1826 degrees of freedom
Multiple R-squared: 0.5262, Adjusted R-squared: 0.5257
F-statistic: 1014 on 2 and 1826 DF, p-value: < 2.2e-16

ETH

Coefficients

	Estimate	Std. Error	t value	Pr (>|t|)
(Intercept)	4.703e+01	6.603e+00	7.122	2.04e-12 ***
Tx Volume	9.096e-08	3.120e-09	29.154	< 2e-16 ***
(tx Volume)^2	-1.743e-18	1.811e-19	-9.622	<2e-16 ***

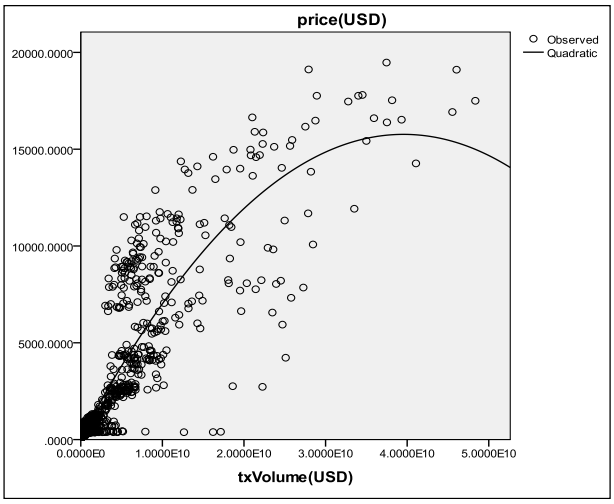
Residual standard error: 175.5 on 995 degrees of freedom
Multiple R-squared: 0.6194, Adjusted R-squared: 0.6186
F-statistic: 809.6 on 2 and 995 DF, p-value: < 2.2e-16

XRP

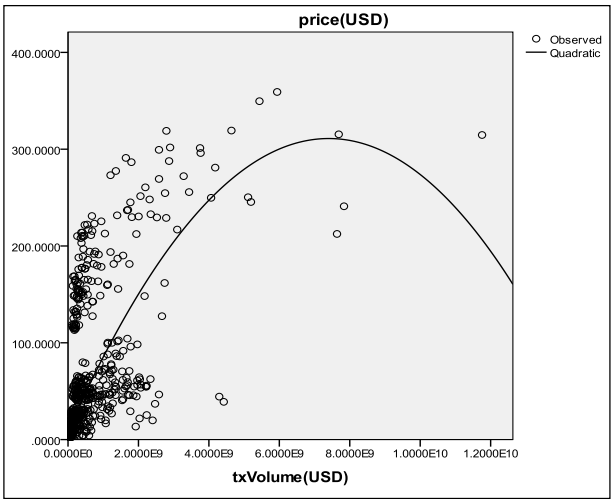
Coefficients

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	9.648e-02	6.779e-03	14.23	<2e-16 ***
Tx Volume	5.395e-12	1.899e-13	28.41	<2e-16 ***
(tx Volume)^2	-1.715e-24	7.393e-26	-23.19	<2e-16 ***

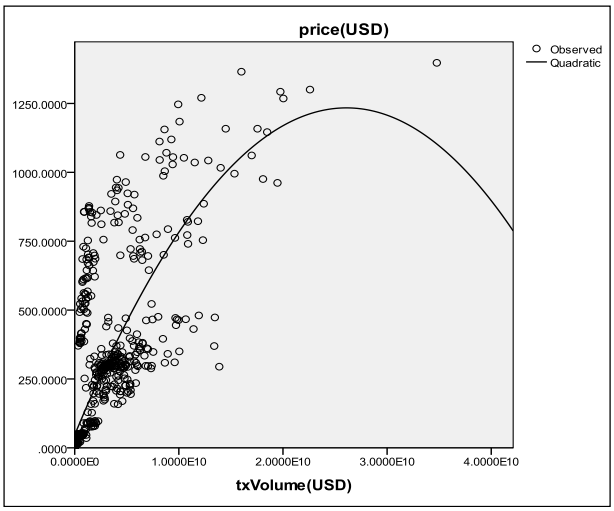
Residual standard error: 0.2791 on 1728 degrees of freedom
Multiple R-squared: 0.3292, Adjusted R-squared: 0.3284
F-statistic: 424 on 2 and 1728 DF, p-value: < 2.2e-16



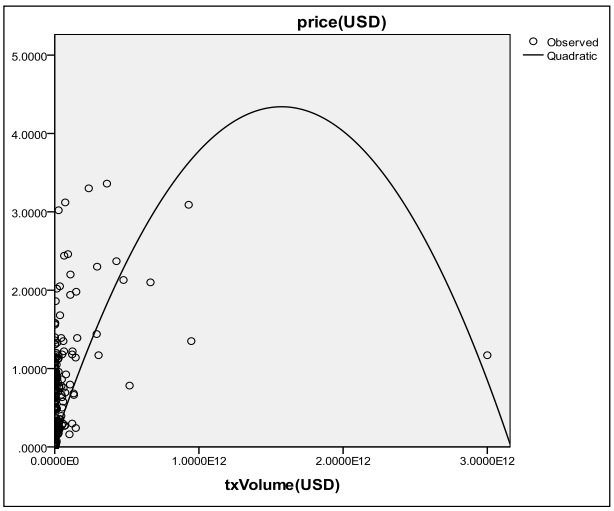
BTC



LTC



ETH



XRP

The polynomial regression outputs and the visual presentation of the model fit clearly reveals that a second degree polynomial regression model can be considered as one of the best fits collectively for the four currencies under consideration.

The presented work takes only few variables into consideration, but there are various other factors that determine the price and eventually the trend of crypto currencies. Also it is essential to check whether the model fits thousands of crypto currencies existing currently in the market.

Conclusion

Data used here for the regression analysis contains historical data related to price (in USD), transaction volume, exchange volume, generated coins, transaction count etc, is publicly available. Four important crypto currencies (BTC, LTC, ETH and XRP) historical data is used for analysis. The descriptive statistics of the data showed non-normal characteristics for the prices of all the four currencies. Several statistical tests are conducted to rule out linear regression model since the assumptions of linearity are not met for the data. After examining each of the variables under study, concluded that a quadratic regression model using transaction volume can be used collectively for predicting price of all the four crypto currencies.

References

1. Alex Heid. Analysis of the Cryptocurrency Marketplace, URL: http://www.hackmiami.org/whitepapers/.
2. Dr. Garrick Hileman, Michel Rauchs. Global crypto currency benchmarking study, 2017.
3. Alexander D'Alfonso, Peter Langer, Zintis Vandelis. The future of cryptocurrency-an investor's comparison of Bitcoin and Ethereum, Ryerson University, 2016.
4. Stephen Chan, Jeffrey Chu, Saralees Nadarajah, Joerg Osterrieder. A Statistical Analysis of Cryptocurrencies "Journal of risk and financial management (31st May), 2017.
5. Shipra Saraswat, Vishal Singh Chauhan, Neetu Faujdar. Analysis on crypto-currency, International Journal of Latest Trends in Engineering and Technology, 2017, 9(1).
6. Farell Ryan. An Analysis of the Cryptocurrency Industry", Wharton Research Scholars, 2015, 130. http://repository.upenn.edu/wharton_research_scholars/130.
7. Darius Tajanko Initio Ventures. Inc: Data Science and Cryptocurrencies-A Survey of Data Science Applications in the Cryptocurrency Space, 2017. dtajanko@initioventures.com (August 23).
8. Jan Lánský. Analysis of Cryptocurrencies Price Development, Acta Informatica Pragensia, 2016.
9. Adam Hayes: Cryptocurrency Value Formation: An Empirical Analysis Leading to a Cost of Production Model for Valuing Bitcoin, (MCIS proceedings), 2015.
10. Peter M, Krafft Nicolás, Della Penna. Alex Sandy Pentland: An Experimental Study of Cryptocurrency Market Dynamics. CHI, 2018.
11. Abhyudit Bisht, Puru Agarwal. Analysis of Bitcoin using Linear Regression and Data Mining Techniques, International Journal of Advanced Research in Computer and Communication Engineering ISO Certified. 2017; 6(11): 3297-2007.