Twitter's Effect on Dogecoin and Bitcoin

John Tsirigotis
tsirigoj@umich.edu
School of Information
University of Michigan
Ann Arbor, MI 48109

Abstract

During the COVID19 pandemic, the world began to see the rise of alternative investing methods, including the emergence of several cryptocurrencies. Most notably, as meme investing culture took off, a small cryptocurrency called Dogecoin really began to emerge. With a huge social media following, including several Twitter mentions by Tesla CEO, Elon Musk, the cryptocurrency made many average people quite a bit of money. In this paper, I investigate if there truly exists a correlation between Twitter mentions from verified accounts and Dogecoin and Bitcoin markets.

1 Introduction

During the height of the COVID19 pandemic in 2020, the entire world was issued a stay-at-home order as people were forced to socially isolate from each other as a public health measure. Since many organizations lost business and many workers were out of work, suddenly, people had a lot of time on their hands.

With the economy essentially shutting down, the United States government issued stimulus checks to most families throughout the country. With this, there was a surprising increase in interest in alternate investing methods, such as sports cards, meme stocks (such as GameStop and AMC), and cryptocurrencies. Specifically, Dogecoin, a cryptocurrency that started as a joke by two software engineers as a satire to its predecessor, Bitcoin, became so popular that its worth grew to be more than the word of Ford Motor Company, BP, and Tesco (Partington).

Between January 2020 and May 2021, the cryptocurrency grew by more than 14,000% (Partington). Many celebrities showed their support for the cryptocurrency, including rapper Snoop Dogg and KISS bassist, Gene Simmons, but perhaps no celebrity showed (and continues to show) more interest in Dogecoin than Tesla CEO, Elon Musk (Partington).

Submitted as the final report for SI 671: Data Mining – Methods and Applications

Known for his off-beat Twitter commentary, Musk often showed his support for the no-longer-a-joke cryptocurrency. Seemingly, whenever he would voice his support for Dogecoin on his Twitter page, the value of the cryptocurrency would go up as investors continued to buy into the stock.

1.1 Related Work

In 2018, entrepreneurs Dr. Vytautas Karalevicius, Niels Degrande, and Associate Professor Dr. Jochen De Weerdt of The Katholieke Universiteit Leuven in Leuven, Belgium, systematically analyzed whether sentiment can be used as one of the predictors of securities' price movements (Karalevicius). In their work, the trio focused on whether the sentiment of the media effected interday trading of Bitcoin by using sentiment analysis using a lexicon-based approach and creating reaction patterns to maximize the expected return based on interday price patterns. Their finding was that the market almost always goes in the direction of the sentiment, but often overreacts towards the sentiment, and then corrects itself (i.e., returns to its previous position away from the sentiment of the news) (Karalevicius).

Given the nature of equity markets and the relative ease that any individual could make a trade within the market, it makes sense that the news can be triggering for any given market. The news and media can very clearly steer the market in a certain direction with their reporting, but can a high-powered individual do the same on Twitter?

1.2 The Data

There were three data sets from where I pulled for regression analysis. The first data set was the Twitter Streaming API, where I queried for Tweets from only verified Twitter accounts that contained any iteration of the words 'Bitcoin' or 'Dogecoin.' I excluded any Retweets, replies to Tweets, and Tweets that contained links. Included in the data were the date and time of the Tweet, the contents of the Tweet, and the author of the Tweet. After cleaning the Twitter data, there were 1,332 cases of Dogecoin and/or Bitcoin mentions between midnight of November 29, 2021, and 4 PM on December 7, 2021, Eastern Standard Time. A binary variable was added to detail whether there was a mention of Dogecoin, Bitcoin, or both (Twitter API).

The Bitcoin dataset used was from the Gemini Exchange and included hourly data for Bitcoin in United States Dollars. Included in the data set was the date and time, the opening price, closing price, high price, low price, and volume (number of buyers and sellers) for the hour (Gemini Exchange Data).

The Dogecoin dataset used was from Coin Telegraph and, like the Bitcoin data, was also hourly. It had all the same variables as the Bitcoin data (Coin Telegraph).

The Twitter API and Dogecoin dataset had times with various 'minutes' and 'seconds,' so they were all rounded down to the hour. Dogecoin was only off by two or three minutes while Twitter API could have been off by as much as fifty-nine minutes. This

decision was made because I wanted the Tweets to serve as somewhat of a leading indicator (predicts the action of the cryptocurrency), and in order to be able to merge all three datasets together on date, time, and cryptocurrency type.

Figure 1 A Bitcoin data point. Note the blue checkmark to designate a verified account.

Figure 2 A Dogecoin data point

2 Experimentation

2.1 Approaches

Multiple linear regression was the main method used to see if there were any correlations between the number of mentions on Twitter about the respective cryptocurrency (Bitcoin or Dogecoin).

2.2 Data set

For every hour of the data set, the difference in price between the current hour and the previous hour was divided by the previous hour to get the percent change in price by hour. The same was done for volume. Additionally, a variable for the number of mentions for every hour was created. In order to conduct the regression, any missing values were treated with the median values for the given variable (the mean was considered, but given the volatility of volume, the median seemed to fit the data more).

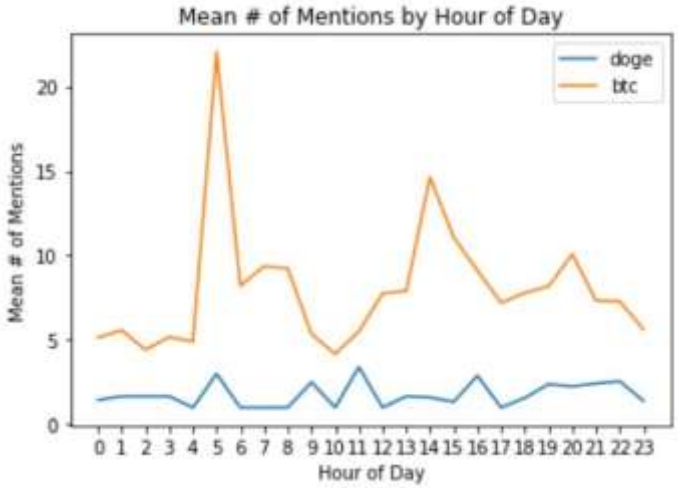

Figure 3 Average number of mentions by hour of day (Eastern Standard Time).

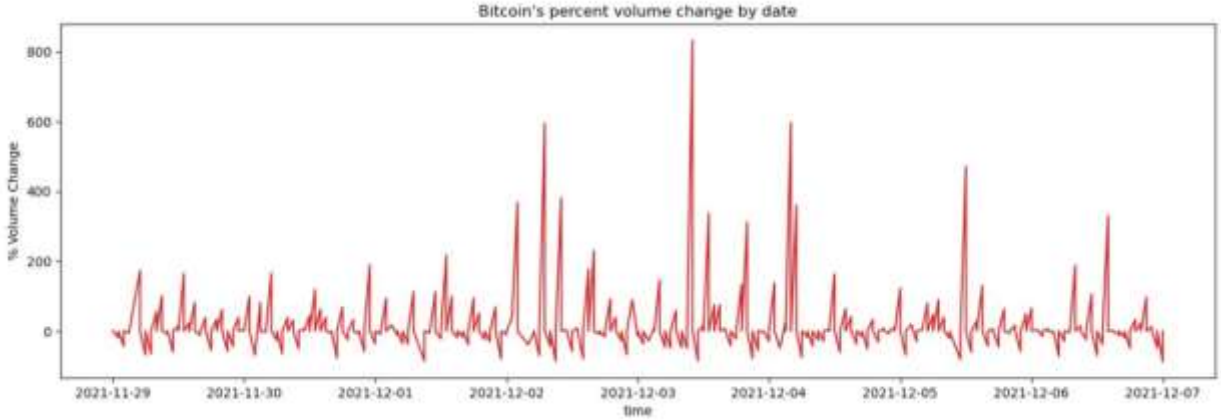

Figure 4 Percent hour-to-hour change in volume for Bitcoin.

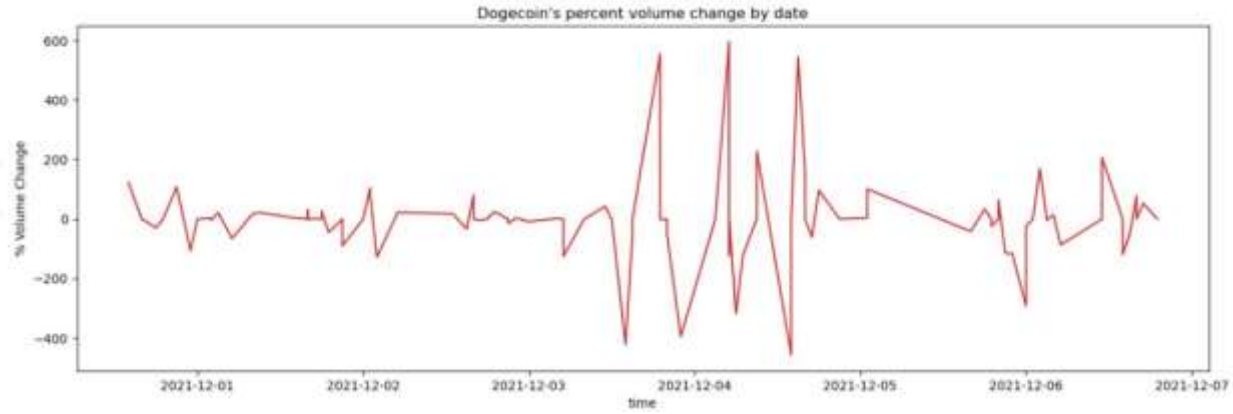

Figure 5 Percent hour-to-hour change in volume for Dogecoin

3 Evaluation

3.1 Regression

Initially, I wanted to investigate whether there was any correlation between the movement in price and volume in Dogecoin and Bitcoin. Along with having a binary variable for Dogecoin and another for Bitcoin (*bitcoin* = 1 if bitcoin mention, *bitcoin* = 0 if not), I included both Bitcoin and Dogecoin in my first model along with an interaction variable (Dogecoin*Bitcoin), and the resulting model was underwhelming.

Given that I focused more on hourly charts (i.e., intraday trades), I decided to focus solely on number of mentions while controlling for the hour of the day and seeing how this affected both volume and price movement for both cryptocurrencies. The results are as follows:

Effect	Estimate	*SE*	95% CI		*p*
			LL	UL	
Intercept	-0.0312	0.023	-0.077	0.015	0.182
Btc_mentions	-0.0014	0.001	-0.004	0.001	0.288
Hour	0.0027	0.001	~0.00	0.005	0.048

Table 1 Regression on change in hourly price, Bitcoin.

Effect	Estimate	*SE*	95% CI		*p*
			LL	UL	
Intercept	9.0900	3.458	2.306	15.874	0.009
Btc_mentions	-0.0972	0.201	-0.492	0.297	0.629
Hour	-0.2680	0.204	-0.669	0.133	0.190

Table 2 Regression on change in hourly volume, Bitcoin.

Effect	Estimate	*SE*	95% CI		*p*
			LL	UL	
Intercept	-0.0777	0.084	-0.243	0.087	0.356
Doge_mentions	0.0178	0.039	-0.059	0.095	0.650
Hour	0.0045	0.002	~0.000	0.009	0.057

Table 3 Regression on change in hourly price, Dogecoin.

Effect	Estimate	*SE*	95% CI		*p*
			LL	UL	
Intercept	-27.3967	5.831	-38.836	-15.957	0.000
Doge_mentions	14.2981	2.728	8.946	19.650	0.000
Hour	-0.0969	0.164	-0.419	0.226	0.556

Table 4 Regression on change in hourly volume, Dogecoin.

4 Discussion & Conclusion

4.1 Discussion

Based on my regression analysis, the null hypothesis of there being no difference in volume and price of Dogecoin and/or Bitcoin due to the volume of Twitter mentions by verified account cannot be rejected, as very little in terms of statistical significance has been shown. However, the data could be somewhat skewed due to a relatively large drop in prices for both cryptocurrencies along with a huge *increase* in volume as a direct result of the sell-off. Because of this, the statistical results in such a small sample could, essentially, be artificial. Also, Tweets that likely occurred during and after the sell-off (more action is likely correlated with more Tweeting about the cryptocurrency) would be difficult to correlate to our models directly, as we intended for Tweets to be treated as a *leading* indicator, rather than a *lagging* indicator with Tweets about 'the sky is falling' when a cryptocurrency goes down significantly.

This is a very robust data set, and a few things could be done to get more precise and accurate results. First, having access to all historic Twitter data, Bitcoin data, and Dogecoin data would make for more robust analysis. The only conclusions that could be drawn based on this project are very limited to the dates that this project pertains to (late November to early December).

Sentiment analysis could further allow researchers to control for the sentiment of each Tweet. One could find whether certain kinds of Tweets correlate with an increase (or decrease) in price and volume and, perhaps, use this data to predict future trends of the cryptocurrency, which could be very profitable and useful for financial organizations and independent investors.

Lastly, researchers could incorporate other important data that correlates to market movement, such as major news events. Major news events are one of the most significant (if not, *the* single most significant) influence on the markets, and it would be incredibly important to control for that in regression analysis.

4.2 Conclusion

Even though one cannot generalize about the direction or volume of these cryptocurrencies given the results, one can clearly see based on the data that these cryptocurrencies are incredibly volatile, which is very appealing to many investors. Generally, there may, indeed, be a correlation between Twitter mentions and market movement, however, given the small timeframe for when this project was completed, we cannot draw any significant conclusions.

Figure 6 Hourly chart of Bitcoin, 26 Nov 2021 - 9 Dec 2021

Figure 7 Hourly chart of Dogecoin, 26 Nov 2021 - 9 Dec 2021

References

$DOGEUSD - Dow Jones industrial average. StockCharts.com. (n.d.). Retrieved December 10, 2021, from https://stockcharts.com/h-sc/ui.

Bukele, N. [@nayibbukele]. (2021, Dec 4). *El Salvador just bought the dip! 150 coins at an average USD price of ~$48,670. #Bitcoin* [Tweet]. Twitter. https://twitter.com/nayibbukele/status/1467000621354135555

C. (2021, December 5). *Dogecoin | DOGE Price Index & Live Chart by*. Cointelegraph. Retrieved December 5, 2021, from https://cointelegraph.com/doge-price-index

Crypto Data Download. (n.d.). Gemini Exchange Data. Retrieved December 5, 2021, from https://www.cryptodatadownload.com/data/gemini/

Karalevicius, V., Degrande, N., and De Weerdt, J. (2018), "Using sentiment analysis to predict interday Bitcoin price movements", *Journal of Risk Finance*, Vol. 19 No. 1, pp. 56-75. https://doi-org.proxy.lib.umich.edu/10.1108/JRF-06-2017-0092

Partington, R. Guardian News and Media. (2021, May 5). *Dogecoin's record-breaking rise shoots 'joke' cryptocurrency to wider attention*. The Guardian. Retrieved December 9, 2021, from https://www.theguardian.com/technology/2021/may/05/dogecoins-record-breaking-rise-shoots-the-cryptocurrency-to-wider-attention.

Twitter API. (2021, December 5). Twitter. Retrieved from http://developer.twitter.com.

Wallace, M. [@MattWallace888]. (2021 Dec 5). *We are only days away from 2022… the year Elon Musk IS GOING TO PUT A #DOGECOIN ON THE MOON!!!* [Tweet]. Twitter. https://twitter.com/MattWallace888/status/1467573187168845825