

# Looking at the Past and Predicting the Future of Bitcoin

---

By Brian Nzuki, Iona Clark, Joana Avrami, and Stephanie Sanjines

---



*Image of multiple Bitcoins*

*GitHub repo with raw and cleaned data: [ds10 final project data](#)*

## **Crypto Basics: What is the cryptocurrency market and where is it heading?**

Cryptocurrency is decentralized digital money, based on blockchain technology that can be traded for regular goods and services but has recently become a novel and exciting asset class. From novice investors to experts in the market, those who have invested in cryptocurrencies such as Bitcoin, Ethereum, and Dogecoin have seen their money grow exponentially throughout the course of the pandemic, and these trends will likely continue as buying volumes increase.

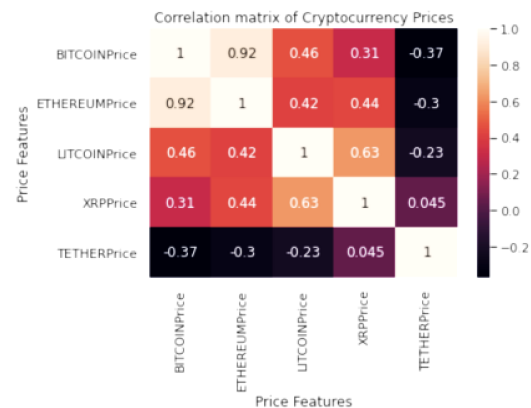
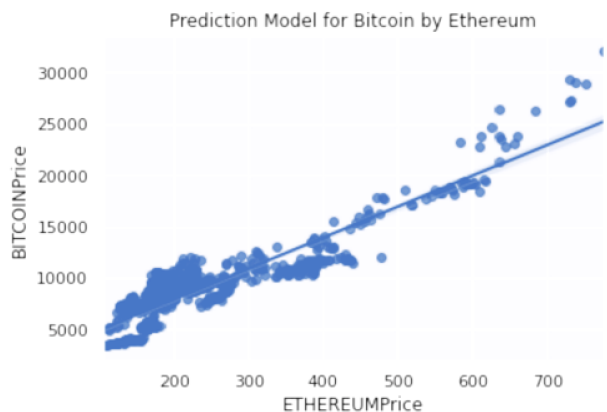
A pioneer of its industry, Bitcoin was the very first cryptocurrency to hit the market back in 2009 at a starting price of \$0. Today, Bitcoin is worth over \$50K a coin and has been making headlines as its price continues to rise and as renowned buyers like Elon Musk drive its popularity upwards. Like other cryptocurrencies, Bitcoin is known for being a high risk investment due to its seemingly volatile nature but can we derive patterns from historical data to predict the future of Bitcoin and perhaps give investors insight as to whether or not investing in Bitcoin is suitable to their risk portfolio. While there are a number of potential external factors that can influence

the price of Bitcoin, we chose to explore markers of the economy, namely the NASDAQ Index and the price of gold, selling prices of other prominent cryptocurrencies like Ethereum, and important current events such as the start of the COVID-19 pandemic.

## Correlation Between the Prices of Bitcoin and Other Cryptocurrencies

While Bitcoin dominates the cryptocurrency market, it is by no means the only coin traded on the market. Other popular currencies include Ethereum, Litecoin, XRP, and Tether. Despite the volatility of each of these individual currencies on their own, can these cryptos be used in conglomeration to predict the price of Bitcoin? Using datasets found online containing the daily volume, closing, opening, high, and low prices for each of these cryptos, we set out to explore this question.

When creating simple linear regressions for the correlations between Bitcoin and each of the 4 other cryptos mentioned, we discovered that the closing price of Ethereum is by far the best predictor of the Bitcoin closing price. Interestingly, combining all of the cryptos in a multiple linear regression model yielded a slightly more accurate prediction of Bitcoin prices.



Actual and Predicted Closing Price of Bitcoin using Multiple Regression Model using other Cryptocurrencies (Ethereum, Litecoin, Tether, XRP)



Figure 1: *Exploring other cryptocurrencies as predictors of Bitcoin price*

**Top Left:** Regression model with line of best fit showing the correlation between the price of Ethereum and Bitcoin. **Right:** Correlation matrix showing the correlation coefficients between Bitcoin, Ethereum, Litecoin, XRP, and Tether. The highest correlation for Bitcoin is with Ethereum, with a correlation coefficient of 0.92, which indicates a large positive correlation. The lowest correlation for Bitcoin is with XRP, with a coefficient of only 0.31, indicating weak correlation. **Bottom Left:** Line Graph comparing the actual values of Bitcoin prices since 2019 to the prices predicted by our multiple linear regression model using the daily closing prices of the 4 other cryptos.

$$\text{BTC Closing Price} = 30.3729(\text{Ethereum-Close}) + 37.0805(\text{Litecoin-Close}) - 11350(\text{XRP-Close}) - 26510(\text{Tether-Close}) + 29300$$

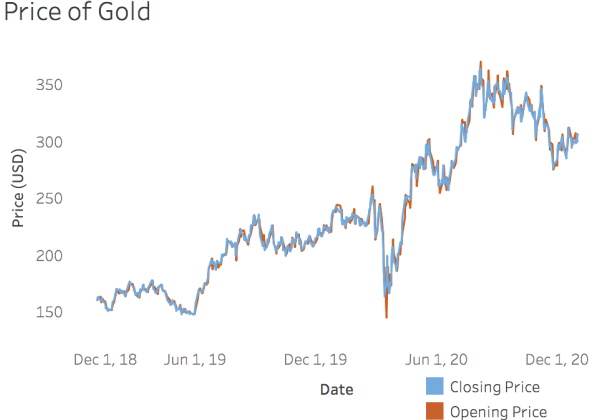
### Correlation Between the Price of Bitcoin and Various Economic Markers

The NASDAQ Composite is a stock market index that includes almost all stocks listed on the NASDAQ stock exchange, including giant corporations like Amazon, Apple, and Tesla. It is one of most-followed stock market indices in the United States and is used as an indicator for the health of the economy, as stock market values tend to perform well when the economy is strong. The price of gold is another important economic indicator that usually decreases when the economy is strong and increases when the economy is weak, as investors shift their money towards more stable investments like cash and gold. To measure the correlation between these economic markers with the price of Bitcoin, we first created separate simple linear regression models using each marker. Individually, the NASDAQ Price Index is a better predictor of Bitcoin prices when compared to the price of gold. For both markers, the opening price did a better job of predicting Bitcoin price. We later found that using a combinatory approach yielded more accurate predictions in a multiple linear regression model with the opening prices of the NASDAQ Index and the price of gold as input values.

Given the success of our multiple regression model, we decided to look further at additional markers of the economy like the S&P 500 Index and the Dow Jones Index which measure the risks and returns of the largest 500 and 30 American companies on the NASDAQ and the New York Stock Exchange respectively. Of the stock market indices, the NASDAQ index has the highest correlation to the price of Bitcoin, which validates our choice from before to use this as an input variable. We do see, however, that Ethereum still has the highest correlation value (0.91) to Bitcoin price.

Opening and Closing Price of Gold & NASDAQ Index

Price of Gold



NASDAQ Index



Actual vs. Predicted Price of Bitcoin with Multiple Regression Model using NASDAQ Index and Price of Gold

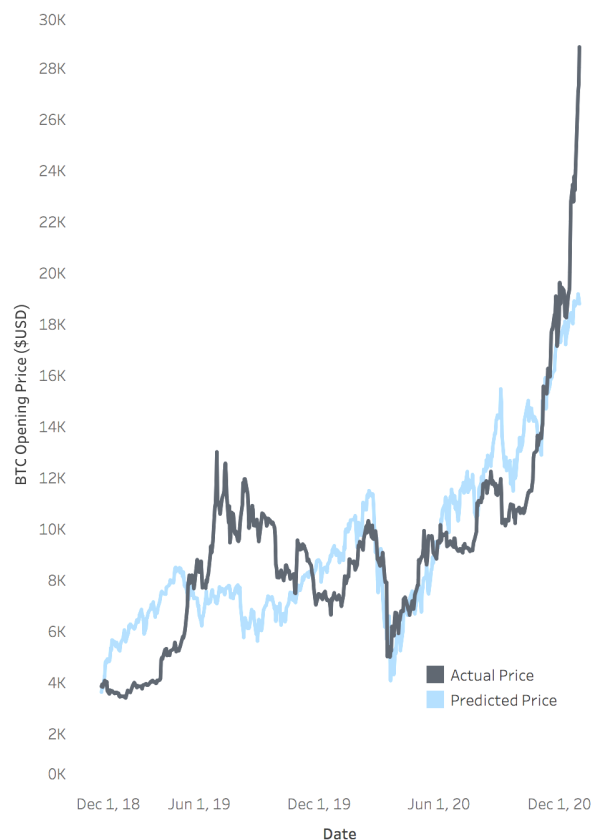


Figure 2: Opening and closing prices of Gold and NASDAQ index from 1/1/19-12/31/20 and actual closing vs. Bitcoin closing price predictions using a multiple regression model with the opening prices of gold and the NASDAQ index.

$$\text{BTC Opening Price} = 2.9313(\text{NASDAQ-Open}) - 23.9912(\text{Gold-Open}) - 1160; R^2 = 0.683$$

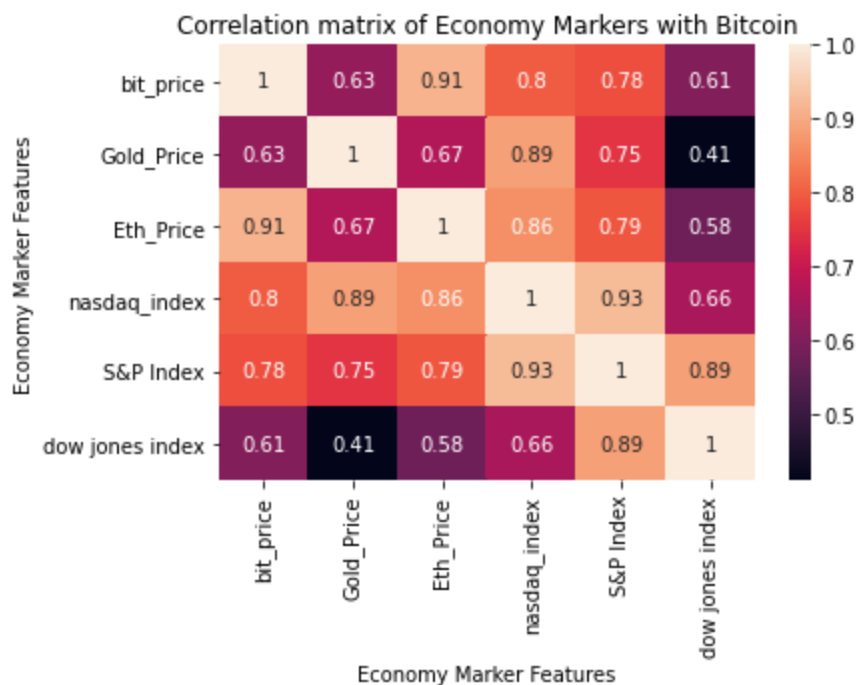


Figure 3: *Correlation matrix of economy markers with Bitcoin.* Above, we see the correlation values from this model between Bitcoin, the price of gold, the NASDAQ Index, the S&P 500 Index, the Dow Jones Index, and Ethereum. We included Ethereum as a means of comparison to our previous model using different cryptocurrencies to predict the price of Bitcoin.

### Zooming in at Bitcoin trends throughout the COVID-19 pandemic

Seeing as the correlation of other cryptos and economic markers with Bitcoin were not perfect, we were wondering what other factors have influenced the price of Bitcoin in the recent past? To investigate this, we zoomed in on the Bitcoin data from the dates of the COVID-19 pandemic, starting from December 2019 to the present day. We decided to isolate days for which the high price of Bitcoin either increased or decreased by at least 10 % from the previous day. There were about 10 of such days, 4 of which were particularly interesting.

Notable Bitcoin Plunges and Surges (12/1/19 - 03/31/21)

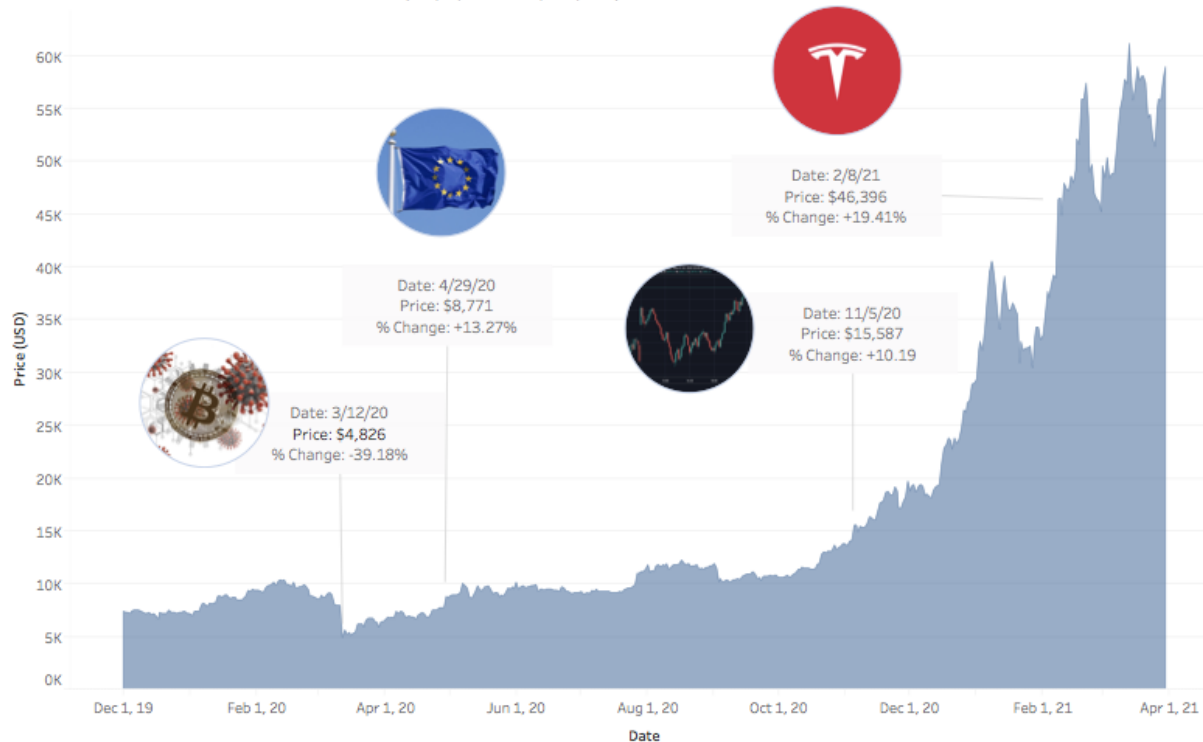


Figure 4: 4 of the major Bitcoin plunges and price surges since December, 2019 to the present day. From left to right: 1) The first plunge took place the day after the WHO announced the start of the global COVID-19 pandemic. The price of Bitcoin crashed nearly 40% as economic turmoil swept the globe. 2) Analysts attributed the rise in Bitcoin price to "bitcoin halving" that takes place on May 12th. In addition, market factors that drove the price upwards of close to 15% included the beginning of the removal of quarantine in the EU and the most powerful emission of the Central Bank of the EU and the Federal Reserve. 3) The 10.19% price gains happened as global equities rallied; notably, the European stock indexes rose around 1% and U.S. stock indexes such as the S&P 500 increased by over 2%. 4) Tesla says it bought \$1.5 billion of Bitcoin, increasing the price of the cryptocurrency by 19.41% to a record high of \$46K at the time.

As we can see, important current events, ranging from pandemics to major companies purchasing Bitcoin, all can significantly influence the price of Bitcoin. These events are much harder to predict, which could throw off our models from above.

### What's next for Bitcoin?

For this report, we have only looked at a few factors that we thought would be useful in predicting the price of Bitcoin. The factors were mainly the price of other cryptocurrencies and indicators of the economy like NASDAQ Index and the price of Gold. Hopefully knowing that

these factors influence Bitcoin prices can help even novice investors make more informed decisions on whether and when to purchase coins. With that being said, to make future predictions more robust, we recommend incorporating other strong covariates like the cost of producing Bitcoin through the mining process. The mining process involves solving complicated cryptographic Math problems by miners, which makes them incur a huge cost of production due to the high computing power demanded by their machines and also the cost of electricity. It is important to consider this variable because this is what influences the supply of Bitcoin, since the supply of Bitcoin is regulated and only certain quantities can be mined within a certain period of time. For instance, the Bitcoin mining algorithm only allows one Bitcoin to be found or mined once every ten minutes. Furthermore, research has shown that the price of Bitcoin is closely related to its marginal cost of production, and we postulate that considering mining costs in the model would produce more robust results.

Future analyses can look more into how we can incorporate current events into our models as well as exploring the possibility of creating a model that combines both cryptocurrencies *and* economic markers. Will Bitcoin continue to rise as we exit out of the COVID-19 pandemic economy or will the hype behind its trading die down as investors move towards more stable assets like gold.

*Joana Avrami, Iona Clark, Brian Nzuki, and Stephanie Sanjines, students at Harvard College enrolled in Data Science 10*



### *Datasets Explained and Troubleshooting*

- We used datasets that included the opening and closing prices of each of our covariates on a daily basis with the exception of the stock indices and the price of gold, for which there is only data available for trading hours when the markets are open.
- Figure 2: To do this analysis, we first had to clean up our dataset since Bitcoin is traded on a daily basis at all times while the stock market runs only on weekdays from 9:30 AM to 4 PM EST. A potential issue with this is that there could have been spikes or plunges of Bitcoin on the weekends that were not accounted for in this model. However, the other 4 cryptocurrencies, similarly to Bitcoin, are traded 24/7, so we were able to look at the daily trends, not just on the weekdays, when creating create Figure 1 Next, we had to decide whether we would use the opening or closing prices of each index to predict the price of Bitcoin and we did so by creating two models and found that the model using opening prices was a better fit for our dataset ( $R^2$  (closing) = 0.678;  $R^2$  (opening) = 0.683).
- Figure 3: The parameters from this correlation matrix show a strong positive correlation between the price of gold, NASDAQ Index, Dow Jones Index, S&P 500 Index, Ethereum Price and the price of Bitcoin. This is quite promising. For all the Economy Markers, the NASDAQ Index and the S&P 500 Index have the strongest positive correlation with the price of Bitcoin, while the Dow Jones and the price of Gold have almost the same correlation with Bitcoin.
  - Despite it not having the highest correlation to Bitcoin price, we decided to keep the Price of Gold as one of our input variables given that it is a more stable measure of the economy that does not rely on stock prices.
- With our strongest predictors, we created a regression model that calculates the price of Bitcoin as follows:  $\text{Bitcoin Closing Price} = 1.4(\text{Gold Closing Price}) + 32.6(\text{Ethereum close}) - 4.2(\text{NASDAQ Close}) + 34(\text{S\&P Close}) - 2.2 (\text{Dow Close}) - 4749.89$  R-squared=0.88. For all our models, this was the best R-squared value that we obtained hence confirming that this combination produces the best predictors.
- Initially, we explored the correlation of daily volume, or the number of coins traded in a day, to the price of Bitcoin and found that there was essentially no correlation between these two variables, which is why we moved forward in exploring other covariates.

Link to Notebook:

<https://colab.research.google.com/drive/103OUUpBeX-OrAGgtdw5mGqQScERlZE3b?usp=sharing>