Dhruv Shah/Tom Duong
12/2/20
MA346

Bitcoin Exploratory Analytics

We are attempting to look at a few different bitcoin metrics to help us uncover any correlations that may exist. We will also investigate hypotheses that are widely accepted by bitcoiners.

We sourced our data from two different websites. First, we sourced our bitcoin CSV data from CoinMetrics, one of the leading data providers in crypto. There are legends told amongst bitcoiners that bitcoin price follows the hash rate, which we attempt to verify. Secondly, we sourced the total assets on the Federal Reserve's balance sheet from FRED, the official Federal Reserve website. As you may know, bitcoin is referred to as digital gold. Gold has historically been an inflation hedge thanks to its limited supply. Not only are we exploring bitcoin metrics, but we price bitcoin as a proportion of the Fed's balance sheet. Overall, our sources will help us justify the hypotheses.

We begin by graphing bitcoins price as a proportion of the Federal Reserve balance sheet. First, we import our CSV files from CoinMetrics and Fred. We then merge the data frames to create a united one that includes dates, Fed balance sheet amount (reported weekly), bitcoins price (USD), and bitcoins price divided by the Fed balance sheet amount. This minuscule fraction will show us how proportionally "cheap" bitcoin is to the dollars being pumped into the economy by the Federal Reserve. This metric is ever important as the US has seen multiple rounds of stimulus and quantitative easing over the past decade.

Dhruv Shah/Tom Duong
12/2/20
MA346

Bitcoin priced against Fed Balance Sheet
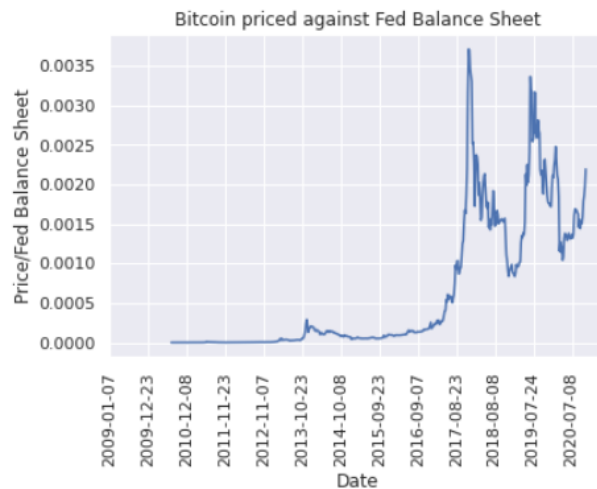
In late 2017, bitcoin hits a high of $19.5k. Now in 2020, bitcoin has gone a run from about $10k to 19.7K over the past month. Although media outlets keep printing "Bitcoin All-Time High!", we know that bitcoin is still cheaper compared to when it was trading at similar levels in 2017.

Next, we attempt to authenticate the Bitcoin difficulty adjustment mechanism. Bitcoin adjusts the difficulty to mine blocks after each block is mined. It will be raised or lowered depending on the time the last block took. There is a Bitcoin block mined about every 10 minutes. To check if the difficulty is correctly adjusting, we plot the number of blocks mined each day in 2017. We chose 2017 because by this time, Bitcoin has matured as a chain and should be running smoothly.

Q-Q Plot

$R^2 = 0.971$

Dhruv Shah/Tom Duong
12/2/20
MA346
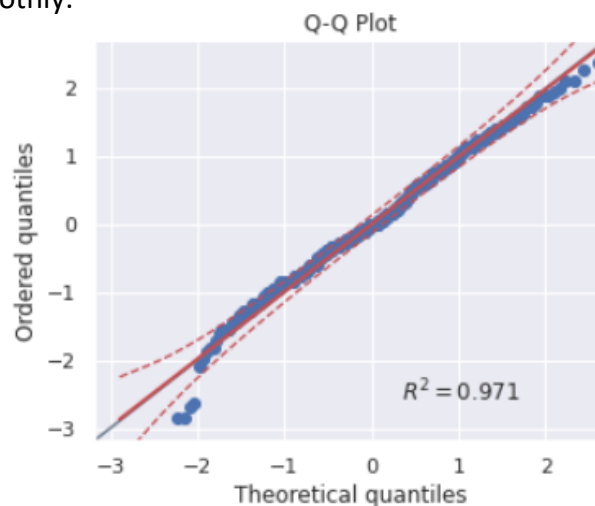
We created a Q-Q plot with each dot representing the number of blocks mined in a day. With an R^2 of 0.971, the plot's distribution is almost perfectly distributed. This test is a further proof that Bitcoin's difficulty adjustment is working accurately.

Then, we attempt to see if there are any correlations between bitcoins price and hash rate. Bitcoin's hash rate tells us the amount of computing power being spent to mine blocks. As the hash rate increases, the work of mining new blocks is becoming more competitive. A higher hash rate usually implies higher security for the chain which has led people to believe that price follows the hash rate. To test whether this was true, we first graphed the hash rate on top of a line graph of bitcoin's price from 2017-2020.
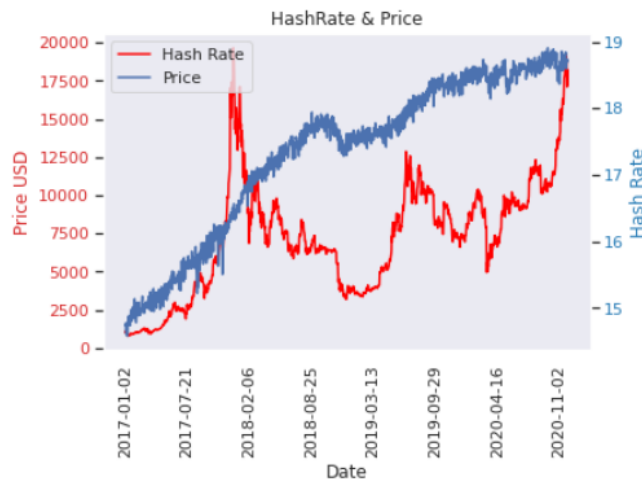
From what we observe, there does not seem to be a correlation, though it is difficult to assess. Therefore, we decided to run a regression analysis to see if we could uncover anything new.

	names object	coef float64	se float64	T float64	pval float64	r2 float64	adj_r2 float64
0	Intercept	4431.134907053668	1.2703985501040701e-14	348798800714206800	0	0.3371711985858824	0.3367063817545962
1	HashRate	0.000049620380721537754	0.000001132098802688725	43.83043300080326	1.7496478864613933e-266	0.3371711985858824	0.3367063817545962

Dhruv Shah/Tom Duong
12/2/20
MA346

Upon further analysis with the linear regression, we found that there is no linear correlation for bitcoin price and hash rate. The nearly 0 correlation coefficient between the two variables implies there is no linear regression. A low R^2 further solidifies our findings. With a very low standard error, we can be confident that the results are accurate. The belief that prices follow the hash rate has turned out to be a myth (at least from 2017 till 2020).

Here we will evaluate a few other miner metrics. For each block mined there is a block subsidy as an incentive for miners to mine blocks. When the chain was founded, the block subsidy was 50 BTC per block. In order to keep bitcoin scarce like gold, the block subsidy is cut in half every four years. Not only are miners paid by the subsidy, but there are also fees paid by the people transacting on the Bitcoin network that will be given to miners. The subsidy and fees are imperative to miner profits which keep them mining bitcoin. This is important because the more miners mining, the higher the hash rate. Therefore, we hypothesize there should be a correlation between the subsidy and/or fees to the hash rate. To test it, we ran a multivariate regression analysis. FeeTotUSD is the total fees paid in a day and IssTotUSD is the total bitcoin issued by the chain. Both variables are measured in USD.

names object	coef float64	se float64	T float64	pval float64	r2 float64	adj_r2 float64
0 Intercept	16.165481830287128	0.056452163908552636	286.3571688141793	0	0.3232423941025835	0.3222925588662362
1 FeeTotUSD	-2.9345049143484194e-7	1.6098747722471564e-8	-18.228156406552465	6.788279747884966e-67	0.3232423941025835	0.3222925588662362
2 IssTotUSD	1.1880835181106159e-7	4.677857712592212e-9	25.398026000501098	1.0783755487403193e-117	0.3232423941025835	0.3222925588662362

We were thoroughly surprised to see an extremely low correlation coefficient and R^2. As previously stated, considering how important the monetary incentive of subsidy and fees are

you would imagine there would be a correlation, but our regression tells us otherwise. We must

reject our original hypothesis.

We moved on to create a grid of pair plots to see quickly if there were any correlations,

we could scope from the naked eye. We chose to plot the following: Average Block Size

(BlkSizeByte), Transaction Count (TxCnt), Median Transaction Transfer Value

(TxTfrValMedUSD), and The Median Fees Paid (FeeMedUSD). All of these are calculated on a

daily basis for the 2017 calendar year.

Dhruv Shah/Tom Duong
12/2/20
MA346
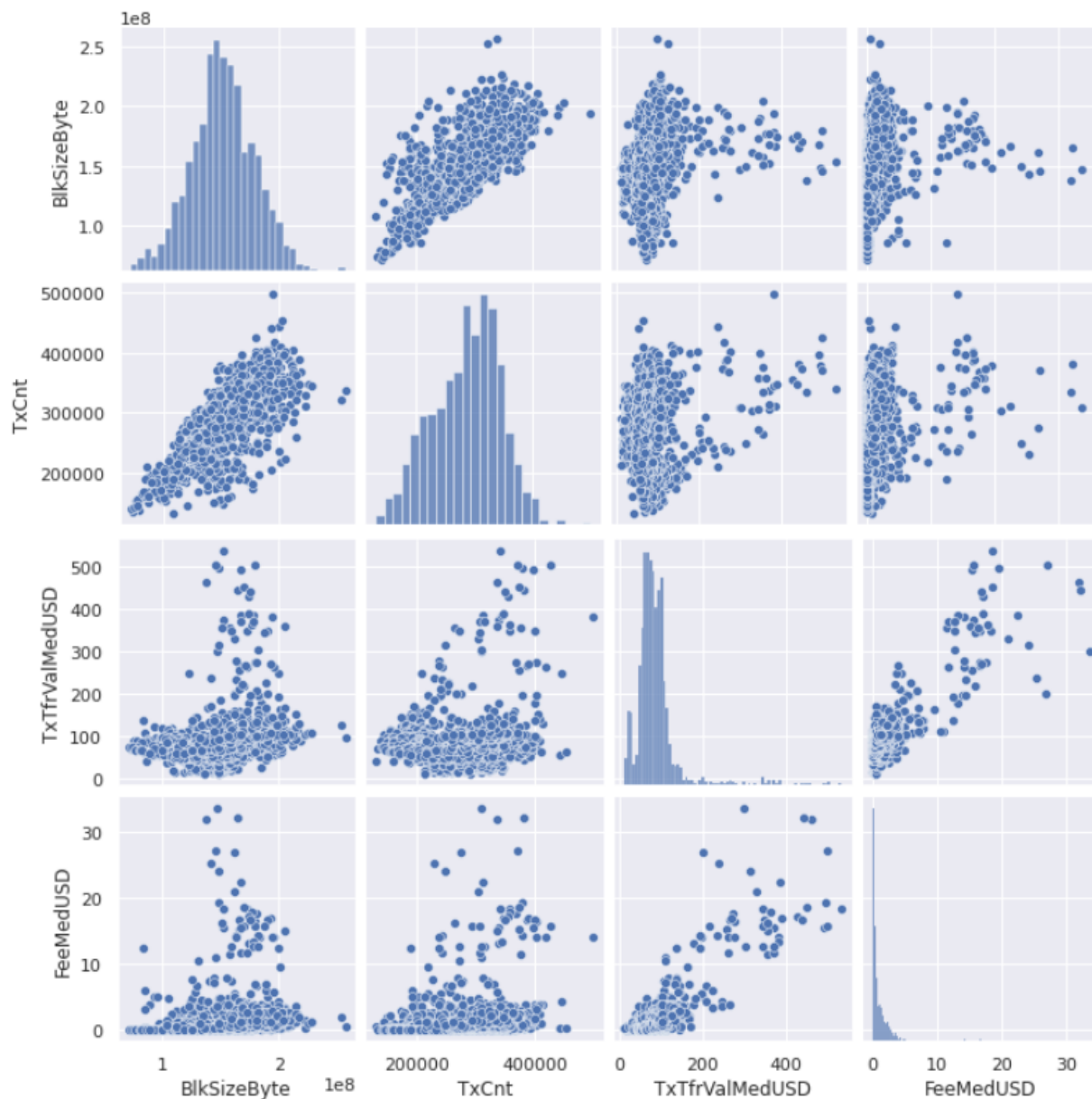
We can immediately notice that there is a correlation between the total number of bytes in the blocks and the transaction count. This would make sense considering the more transactions the denser blocks would become. Median transaction transfer value and fees also seem to have a positive correlation. That is odd because the transaction value (in terms of dollars) should not affect the fees paid per transaction. Fees are determined by the size in bytes of the transaction. That being said, we are unable to discover a correlation between block size byte and fees.

Last but not least, we created a pair plot of the price and market Cap. Considering the market cap is the price multiplied by total supply, this plot should be a near-perfect positive correlation.
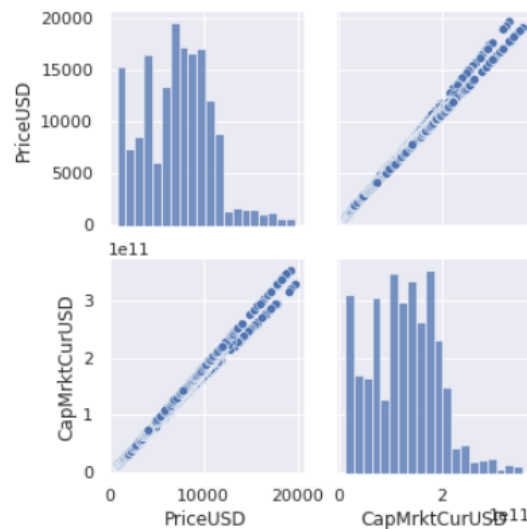
Our hypothesis was confirmed by the pair plot. Price and market cap have a strong positive correlation.

Dhruv Shah/Tom Duong
12/2/20
MA346

In conclusion, our ability to price bitcoin against the Federal Reserve Balance Sheet helps us price bitcoin proportionally. We were also able to debunk bitcoin prices following the hash rate hypothesis and uncover correlations that currently exist.