ECON 422: Econometrics 2 - Machine Learning and Economics

1. Introduction

Research Project:

Using Machine Learning to analyze cryptocurrency trends:

Can we predict Tether’s market cap evolution using other cryptocurrencies’s demand ?

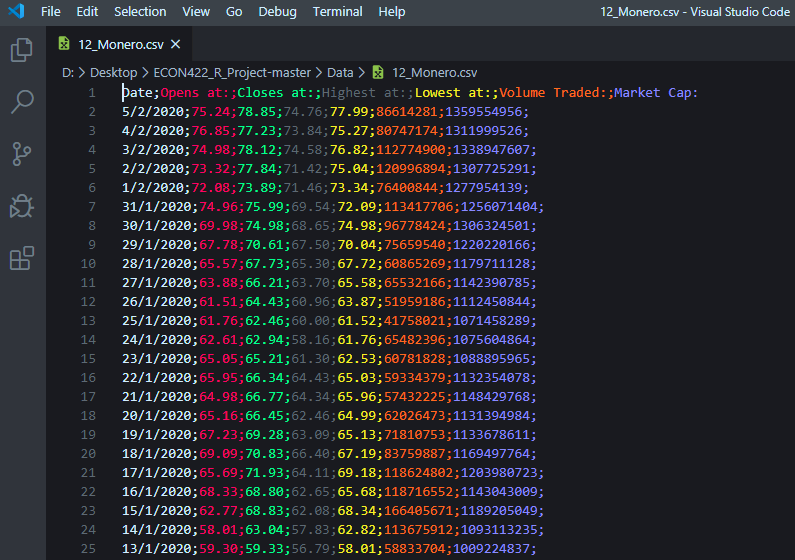
Context:

Cryptocurrencies are very volatile assets, but one, the Tether, has a guaranteed exchange rate of 1 dollar per unit, making it one of the most exchanged of all since it becomes possible to exchange instantly a speculative asset against a very stable one whose value is guaranteed in dollars. Knowing this, we can expect this asset’s market cap to be positively correlated to pessimism on markets. Analyzing demand and price of the most exchanged cryptocurrencies may tell us how the Tether’s market cap will evolve.

1. Data

We will be using the data from the website coinmarketcap.com that we gathered using a scraping tool we built in java available at <https://github.com/SanteauX/histDataCmcJava> to which we added a special class to build csv available in the annex part.

We are using .csv files such as the following to obtain the history of price, market cap and volume of transaction of the biggest cryptocurrencies.



Using this data, we can see the evolution of a cryptocurrency’s value and demand. For the final part we will be using the 30 biggest cryptocurrencies, but for this part we will use at most half of them.

Using these data, we are able to tell if demand for one goes up, down or stay constant. Among these we know that one, the Tether has an exchange rate guaranteed of 1 USDT = 1 USD, which makes it a safe haven

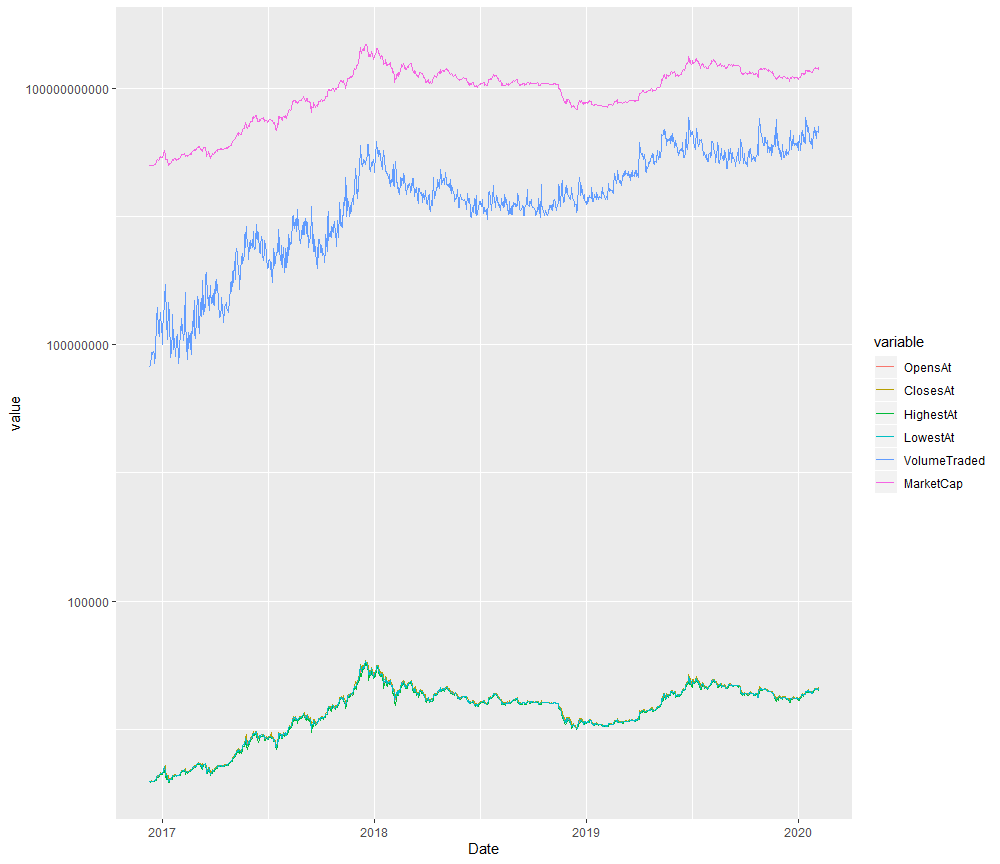
1. Preliminary Results

For now, we only used a sample of the data available (9 cryptocurrencies on the 30 we scraped) to plot graph and compare trends of the different currencies’s variables evolution through time.  
These graphs allow us to observe correlation between the different variables and the currencies themselves.

For these graphs we used a logarithmic scale on the y axis for clarity purpose and limited the time window from 2017 to January 2020 (included) so we could compare most of the most important currencies.

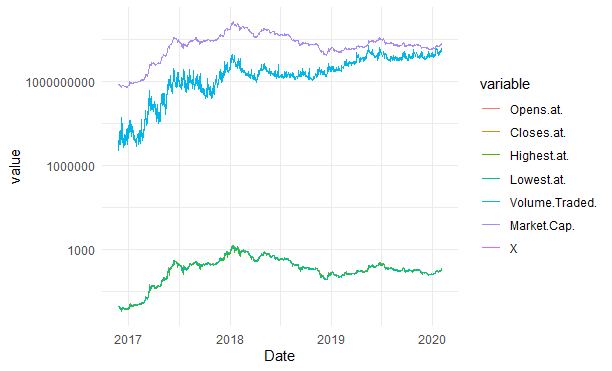
Bitcoin:

Bitcoin is the biggest and most important cryptocurrency by most metrics. It is worth 63% of the global market cap of all cryptocurrencies.



Ethereum:

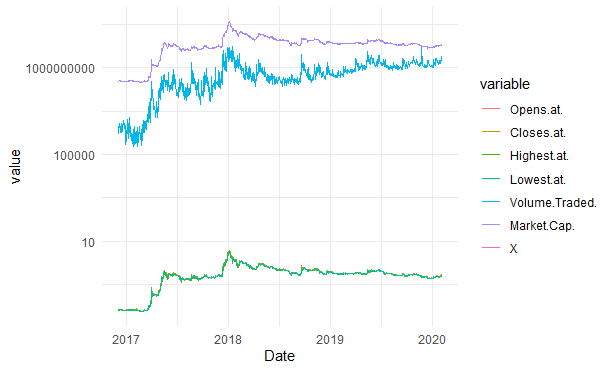
Ethereum is the second biggest cryptocurrency, similar in many aspects to the Bitcoin it differs mostly by the possibility to program smart-contracts (example: allowing transactions without third party).

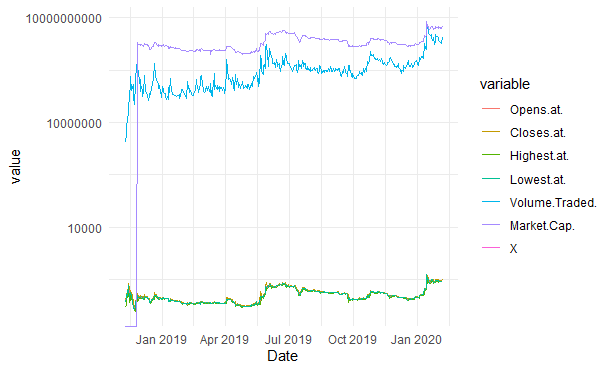


XRP

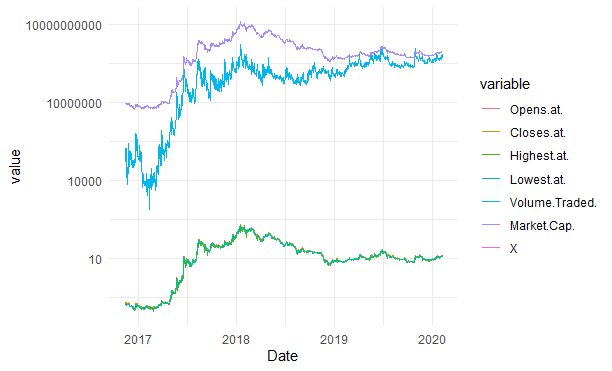
Ripple was created to facilitate financial settlement and money transfers through a cryptocurrency protocol providing security through encryption, cost efficiency and no third-party transaction fee.

It is used by many financial institutions for real-time gross settlement, and unlike most cryptocurrencies, has no limit to the number of units that can be emitted.



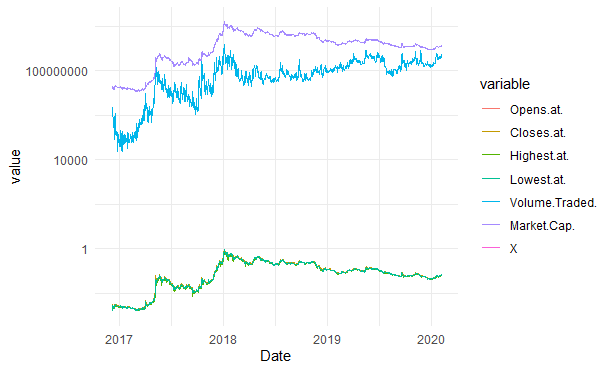
Bitcoin SV[[1]](#footnote-1):

Uses a different encryption  
than Bitcoin. Used mostly  
because transactions are  
secured faster.



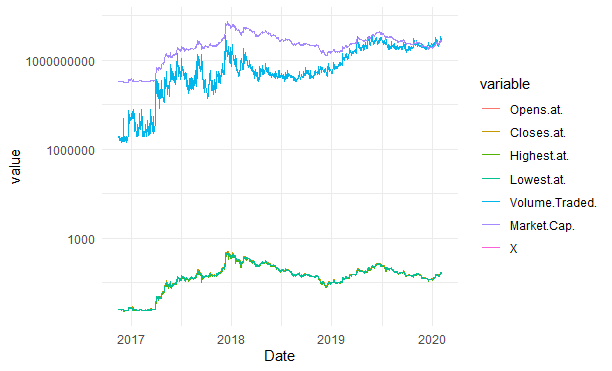
Neo

Regulator friendly cryptocurrency   
created in China. Deploys smart   
contract applications and helps   
manage digitized assets.

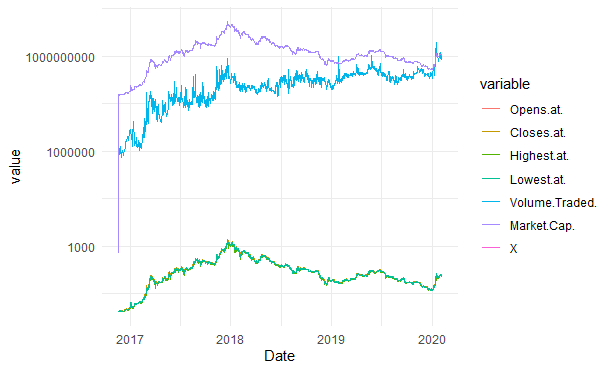


Stellar

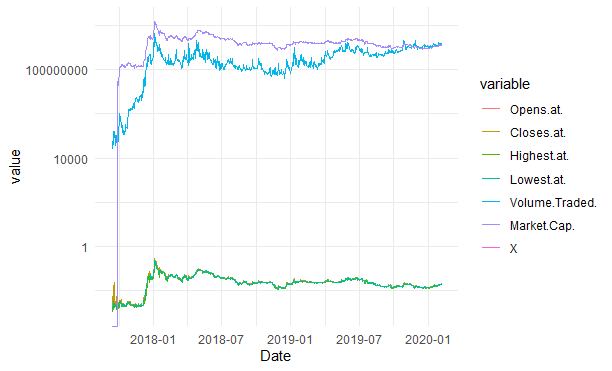
Cryptocurrency used for a   
protocol allowing cross border   
transactions between real world   
currencies.   
Most notably used by Deloitte  
for its Deloitte Digital Bank and  
IBM

Litecoin

Litecoin (LTC) is very similar to  
the Bitcoin but confirms   
transactions must faster  
because of a different system   
of encryption

  
Dash

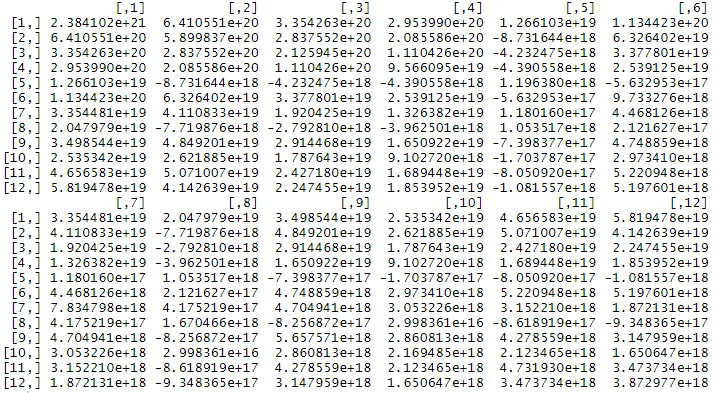
Started as a copy of the Bitcoin,   
it is most notably different from other  
cryptocurrencies because unlike most  
other cryptocurrencies,   
transactions in Dash are   
untraceable. It was the most popular  
cryptocurrency in Venezuela after the  
Bolivar’s value collapsed.



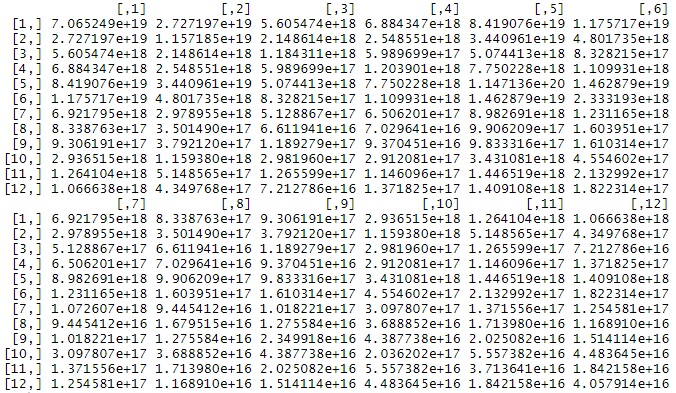
TRON

Protocol created to concurrence   
Ethereum, boomed in market   
capitalization after the creator also   
bought the biggest torrent  
peer-to-peer file sharing network   
(BitTorrent) to associate the two of them.

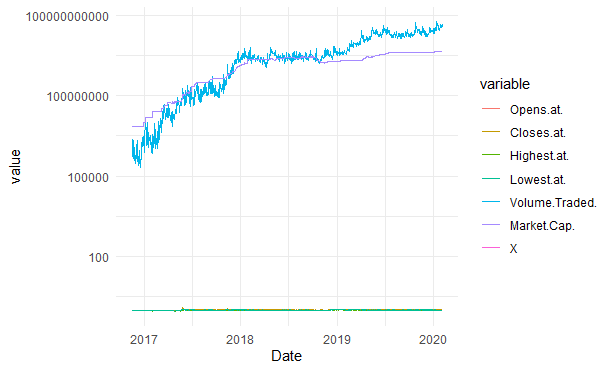
As we saw in the three previous pages, cryptocurrencies may have very different uses but as we can see on the graphs, their prices, volume traded and market cap seems to be highly correlated, which is confirmed with these covariance matrices on the data from the last 852 days (code in the annex):

Market Cap:

It appears that on the 144 observations, 28 were negatively correlated while the rest is positively correlated.

Volume Traded:

The 144 observations are all positively correlated. We can observe from this that demand in cryptocurrencies is exogenous and affects all of them, however we cannot determine yet how much

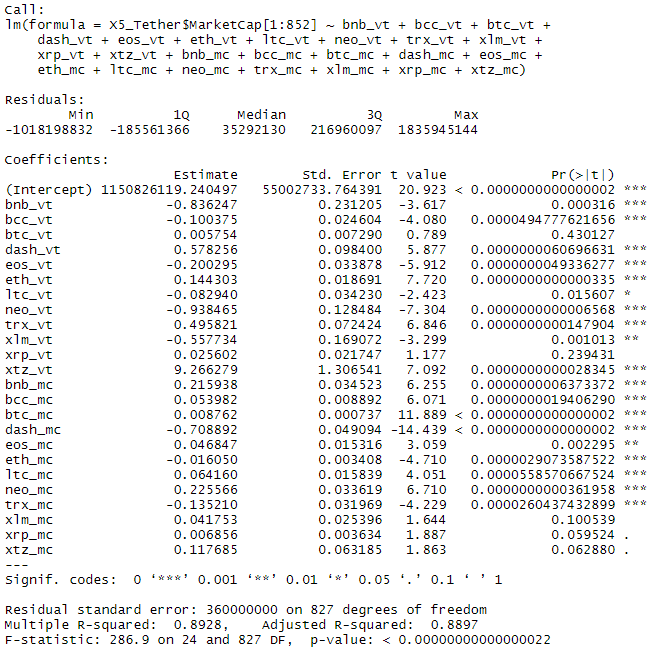
Tether

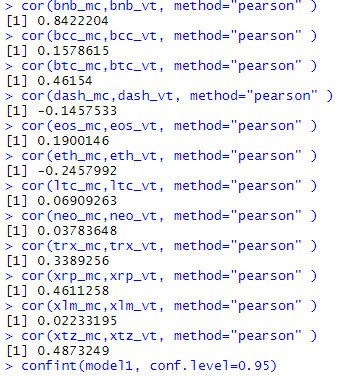
The Tether is our variable of interest. As explained early, it has parity with the dollar, every Tether emitted is supposed to be backed by a dollar and can be exchanged at any time for a 1$USD on the Bitfinex platform.

What makes it special is that since its price can’t go down, its market cap can only go down if people exchange it against a dollar (a Tether exchanged for a dollar is deleted), and it can only go up if people buy newly printed Tether.

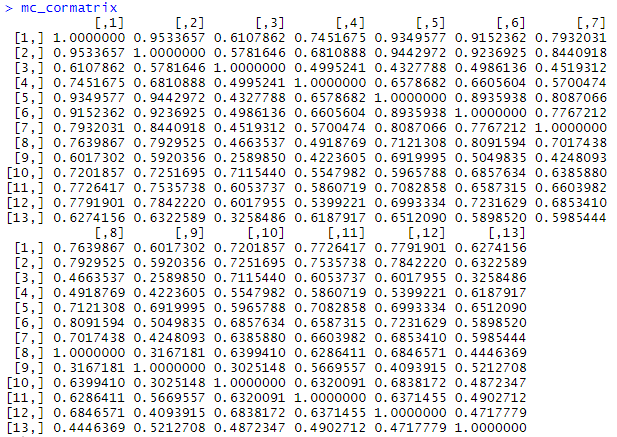
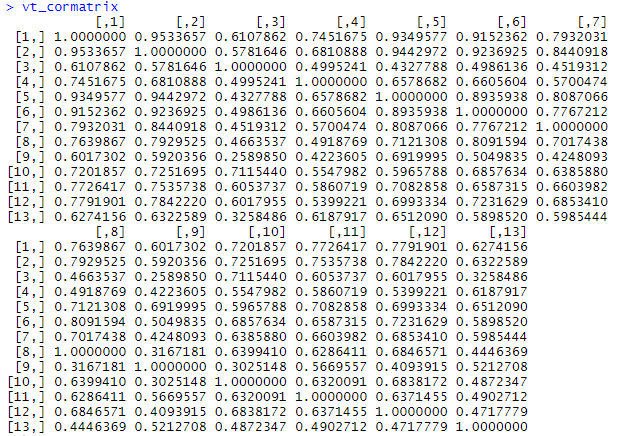
As we can see on this graph, its volume traded is by far the highest of all cryptocurrencies, and we can also observe that unlike other cryptocurrencies its demand is much less affected by exogenous shocks: even when other cryptocurrencies’s market cap goes down (because of the price going down), Tether demand goes up or doesn’t change because people want to secure the value of their assets which confirms the hypothesis we had in the previous assignment.

First model:

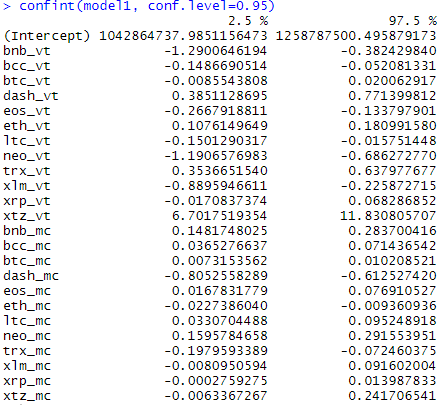
For now, the “most efficient” (but flawed) prediction model we used with the limited sample of data we collected is the multilinear regression using the different cryptos’s market cap and volume

A 0.8928 may look like a very good R² but for some  
currencies the market cap and the volume traded are highly   
correlated to each other:

Adding to this that there are 24 explanatory variables   
we have good reasons to seriously doubt the validity of  
such a high R²

And the same goes for the traded volume and the market cap for whom the correlation is pretty high 

To end this model, here is the confidence interval:

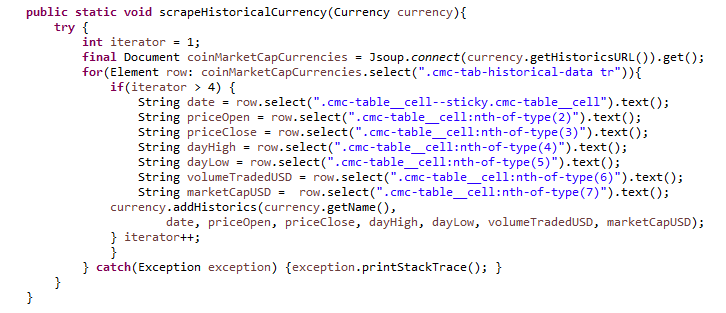


Conclusion :

We showed in this assignment that this data is quite complex because of many factors: exogenous demand, high correlation between variables, complex parameters (fixed price of the Tether), huge dominance of the Bitcoin, this is why simple models such as the multilinear regression don’t seem to be a good fit to predict such phenomenon.

We will in the next part compare different models of machine learning algorithms, their efficiency to predict result and compare them to the real-life data

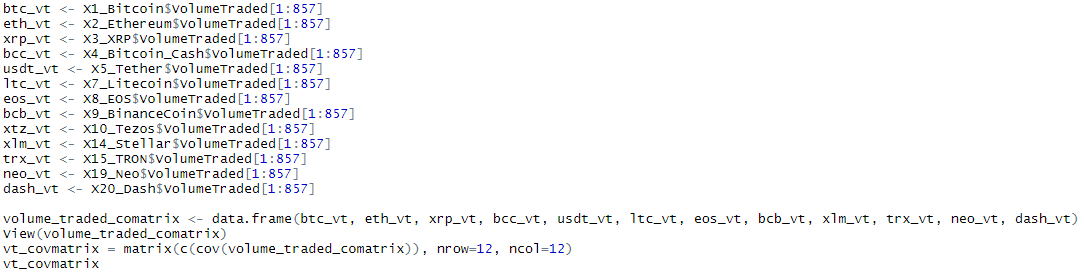
1. Annex

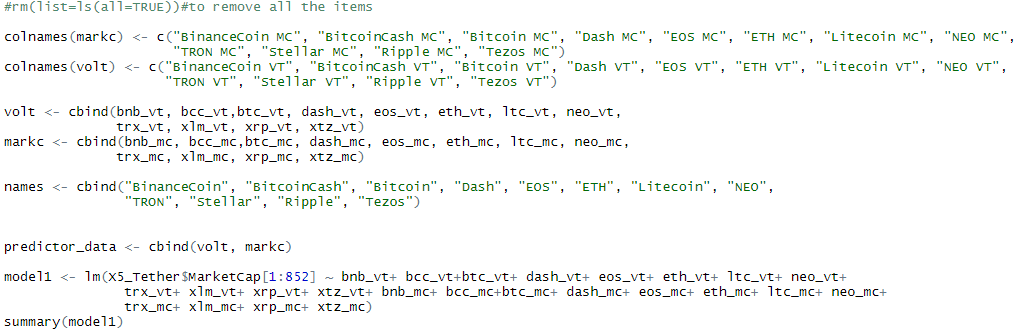
Here are the most important parts of the code that we used to get the data for this part.

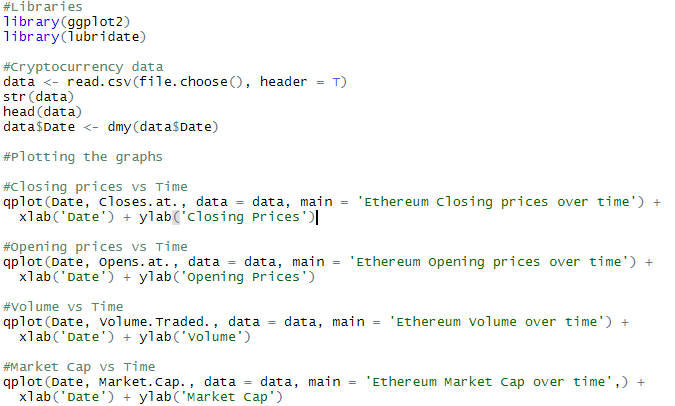
This method was used to identify data on a website’s page and make a java object out of it through the Currency Class.



This method was used to make a CSV for every currency we had data on. For parsing reasons, we used semicolons because commas were used for decimals on the site.

Code we used for the covariance matrices of the market cap and volume traded

Code for the multilinear regression:

Code for the plots:

1. The reason why the market cap seems to be low at the beginning of the chart is because the Bitcoin SV comes from a “hard fork” of the Bitcoin Cash. A forced split of the asset into two different assets that occurred in November 2018, because of a disagreement on technical issues dividing the community. [↑](#footnote-ref-1)