ECON 422: Econometrics 2

Machine Learning and Economics

Project:

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

Index:

1. Introduction:

* Cryptocurrencies
* Tether

1. Research Proposal[[1]](#footnote-1)

* The Data
* Method
* Why

1. First Results[[2]](#footnote-2)

* Observations
* First Conclusions

1. Machine Learning Models

* Regularization

- > Lasso

- > Ridge

* Prediction

-> Support Vector Regression

1. Conclusion
2. Annex
3. Sources
4. Introduction
5. Cryptocurrencies

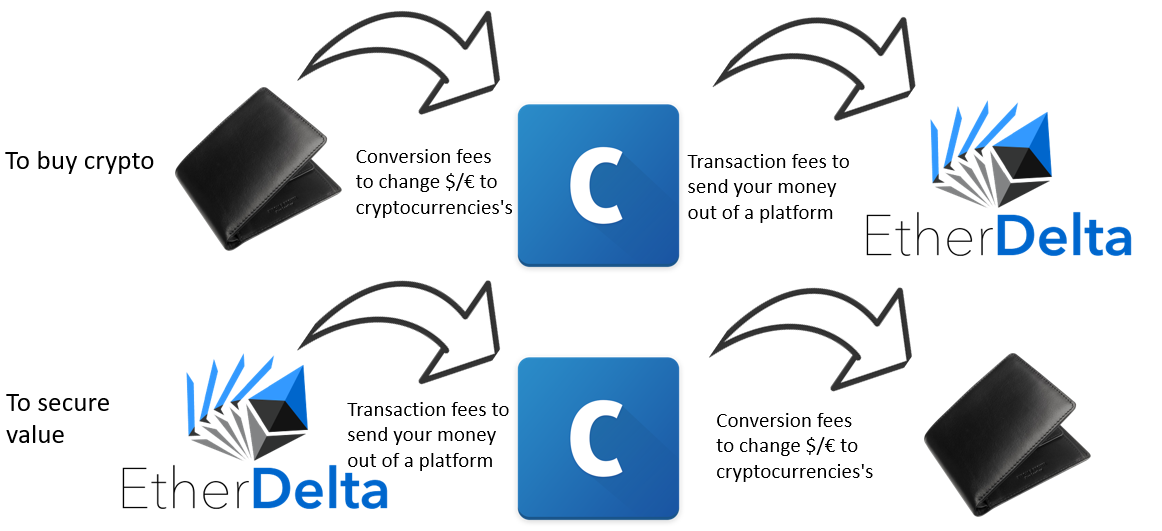
Cryptocurrencies are digital assets used as currencies for online transactions, using technology to secure money creation, money movement, limit money emission, or add features such as smart contracts. Their price is determined only by offer, demand and trust that people have in it; thus, their value is very volatile, which makes them both a digital currency that can be used for shopping (if accepted) and a very speculative asset that can be traded against any other cryptocurrency on online platforms.

Every unit of cryptocurrency is assigned to a wallet, an address composed of many letters and numbers. Many online exchanges offer people to trade one crypto against another, but few accept normal currencies such as USD or Euros to buy cryptos, so in order to buy any cryptocurrency, you first have to buy it on one platform that accept regulated currencies, then send it to a wallet on an exchange platform where you can trade it, which causes transaction costs.

Anyone wanting to trade cryptos must go through this process:

1. Create an account on a platform that sells cryptocurrencies against real currencies[[3]](#footnote-3)
2. Create an account on a platform where you can trade cryptocurrencies
3. Send what you bought from the first platform to the second one, which causes transaction costs[[4]](#footnote-4)

And the process is the same in the other direction for anyone who would like to exchange his cryptocurrency to get real money: send the crypto from the trading platform to the conversion platform (which causes transaction costs) and convert it back (which causes conversion costs).



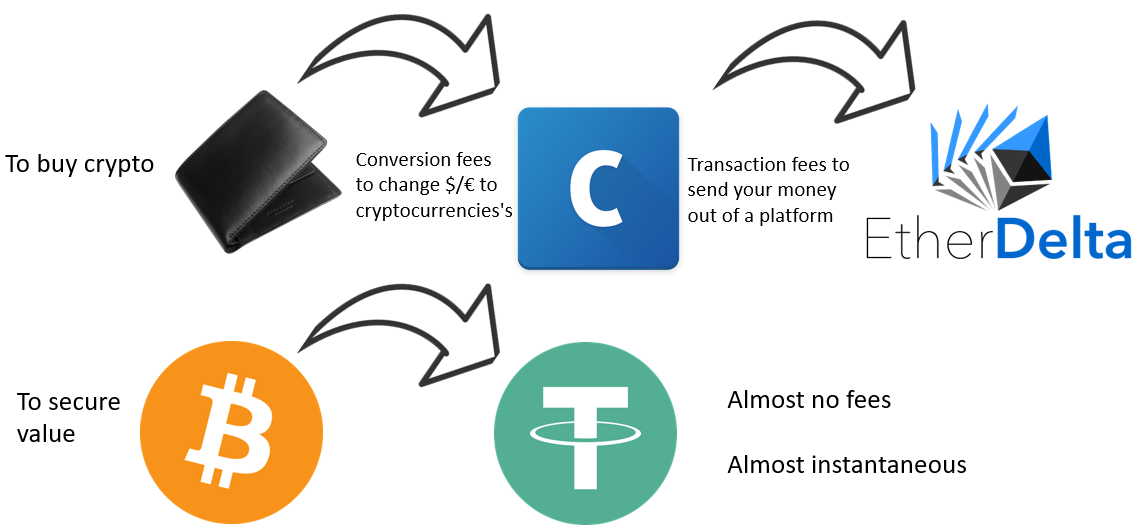
The process isn’t immediate and even if the amount of bitcoin sent from one platform to another doesn’t change, the value in dollar of the asset may have changed by the time it is finally converted to dollars.

These intermediate steps make it long and costly to exchange a cryptocurrency asset against a safe and stable value such as a normal currency, so a platform thought about a way that would make it possible for people to move much faster from a “speculative” asset to a stable value asset and opposite: Tether

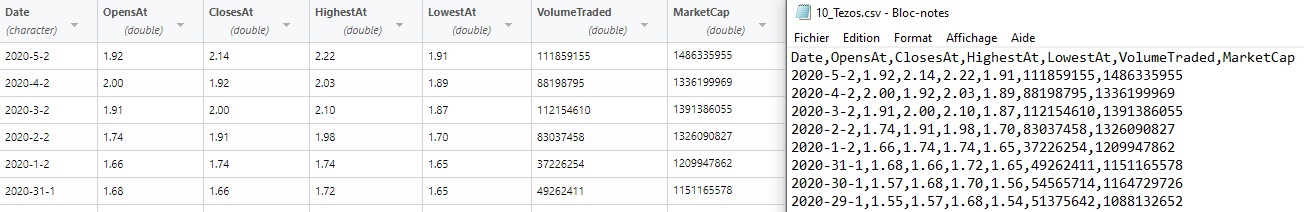
1. The Tether

The Tether[[5]](#footnote-5) (USDT) is a cryptocurrency created by Bitfinex, one of the biggest cryptocurrency trading platform, that (supposedly) has every unit of tether backed by its original currency (for the USDT, the original currency is the US Dollar), and guarantees a fixed rate of:

1 USDT = 1USD

This not only makes it much faster to secure an investment, but also avoid both conversion and transaction costs. The Tether is called a “stable-coin” (its value is stable), and thus the ultimate safe-haven when people feel like other cryptocurrencies’s values are about to go down[[6]](#footnote-6).

1. Research Proposal
2. The Data

We built a scraping program in java that can scrape Coinmarketcap.com and a parser in Python that can reorganize it in a CSV containing for each asset: the day, the price at the opening, the price at the closure, the highest price, the lowest price, the volume traded and the market cap:

1. Method

We know that the Tether is a safe haven for anyone who wants to freeze or secure the value of his investment in a cryptocurrency. If someone exchanges an asset for something stable, he expects the value of the asset going down.

We can assume if the demand in Tether goes up, the market is going down and if the demand in Tether goes down, people are selling their Tethers for riskier assets. Since the price of the Tether is 1USD$ we can only rely on its market cap to know if its demand goes up or down, however for other cryptocurrencies we can look at their price, market cap and volume traded to know how the market evolves.

Using this data and what we know about the Tether we want to predict the evolution of the Tether’s Market Cap: If we can predict the Tether’s Market Cap given the trend of other cryptocurrencies we will know whether the Tether is bullish or bearish and from this we will know if we should buy or sell.

1. Why ?

We think this question may be more than just a way to predict markets on this domain but a general behavior about risk.

We think this this topic is interesting because even though cryptocurrencies are unregulated and for many of them heavily influenced by big actors, they follow human behaviors that most of them replicated in regulated market finance and other domains:

We will learn through this project if we can predict Tether’s value, but we will learn about people’s behavior toward a “safe-value” which is more than just a way to predict markets on this domain but a general behavior about risk aversion.

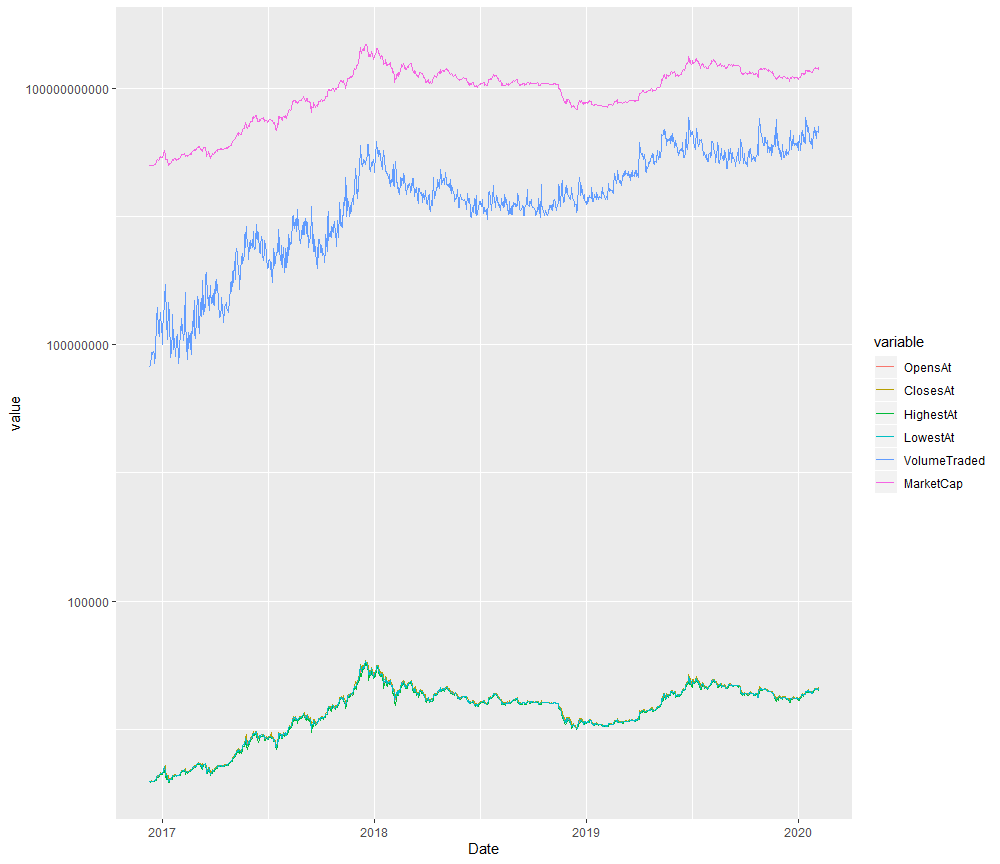
This topic is also interesting because the US Justice Department is investigating Bitfinex for illegally manipulating the price of Bitcoin using the Tether, and because Bitfinex failed to provide an audit showing reserves correctly backing every Tether “printed” or emitted, meaning that this cryptocurrency could be the trigger of a crash if Bitfinex failed at some point to guarantee the fixed rate.

1. First Results
   1. Observations

For this part we only used a sample of all the data we could gather to plot and compare the different trends through time. These graphs allow us to observe correlation between the different variables and the currencies themselves.

To plot these graphs, we used a logarithmic scale on the y axis for clarity purpose and limited the time windows from 2017 to January 2020 (included) so we could compare the most important cryptocurrencies[[7]](#footnote-7):

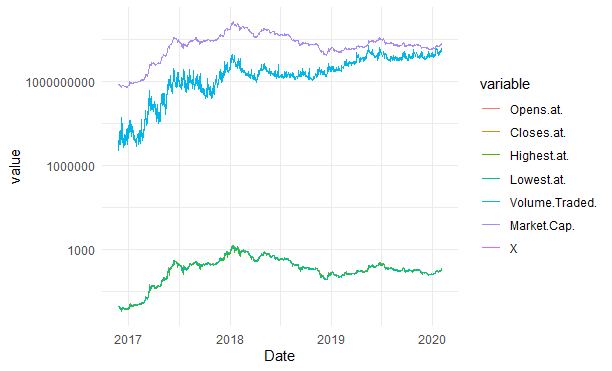
* Bitcoin (BTC):

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

* Ethereum (ETH):

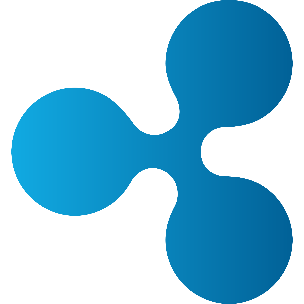
Une image contenant lumière

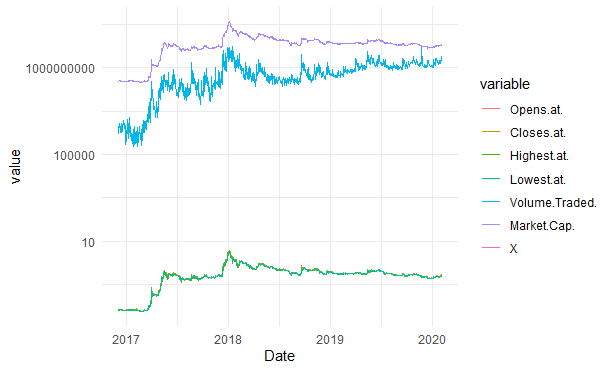
Description générée automatiquementEthereum is the second biggest cryptocurrency, similar in many aspects to the Bitcoin it differs mostly by the possibility to program smart-contracts (example: allowing transactions without third party).



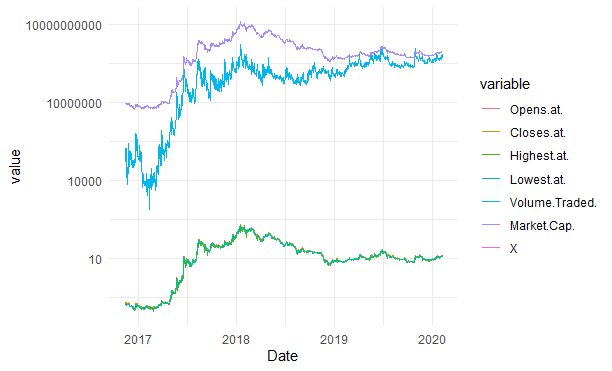
* Ripple (XRP):

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.



* NEO (NEO):

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

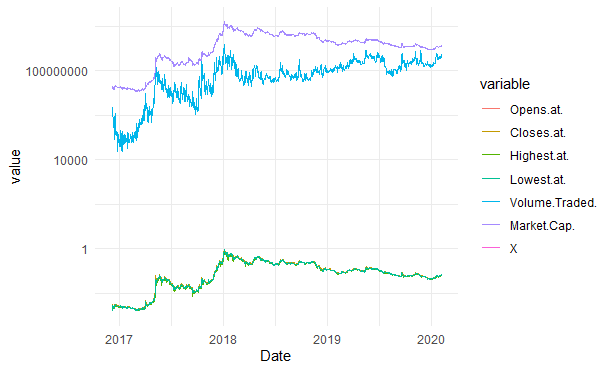
Une image contenant ordinateur

Description générée automatiquement

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

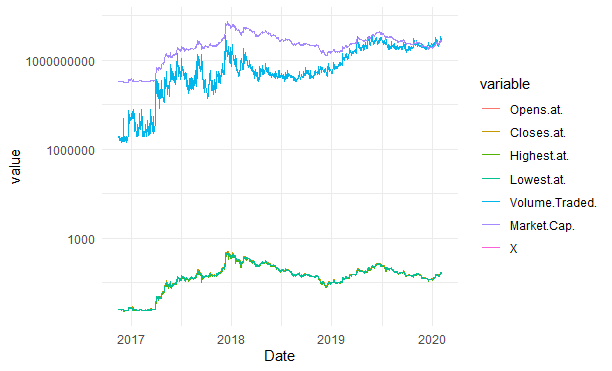
Une image contenant assis

Description générée automatiquement

* Litecoin (LTC):

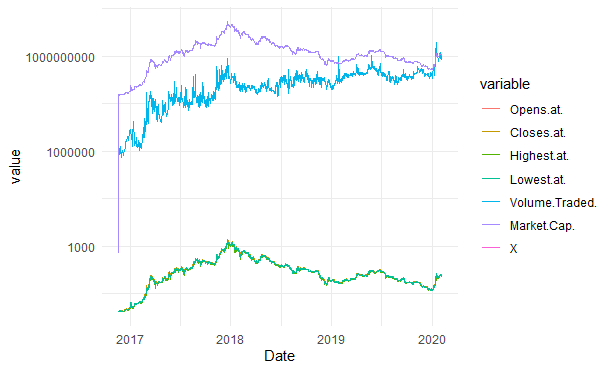
Une image contenant dessin

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



* Dash (DASH):

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

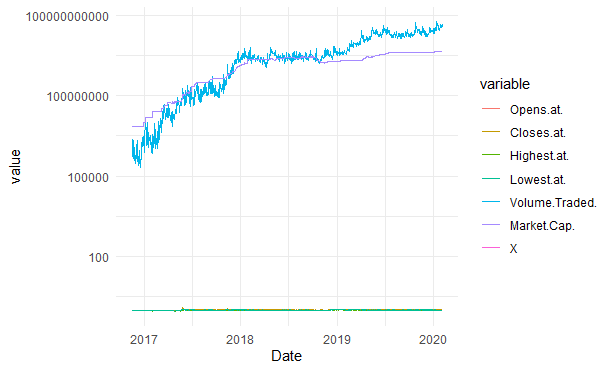


* Tether (USDT):

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.

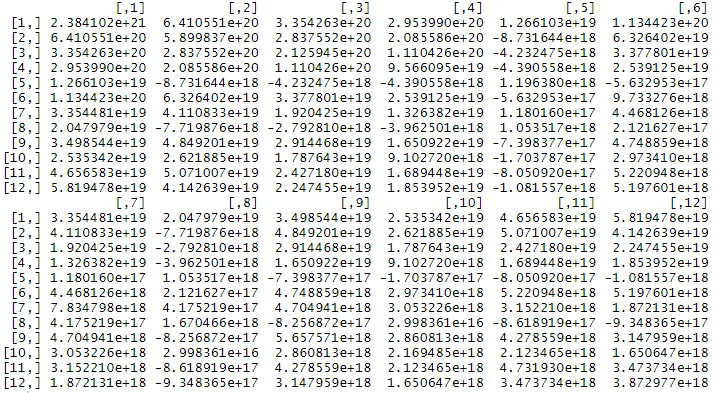
As we can see on this graph, its volume traded is by far the highest of all cryptocurrencies, and we can also observe that unlike other cryptocurrencies its demand is much less affected by exogenous shocks: even when other cryptocurrencies’s market cap goes down (because of the price going down), Tether 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 part.



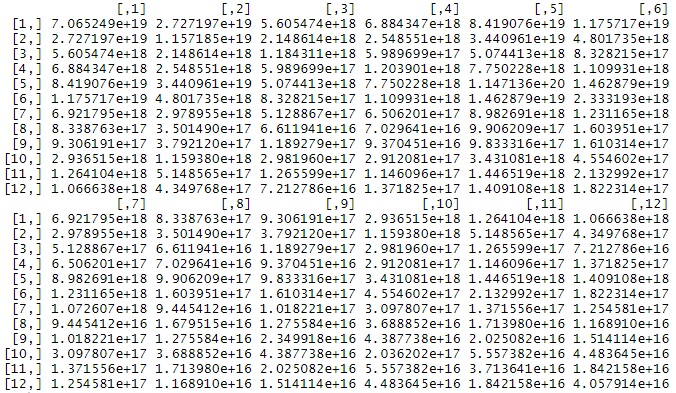
* 1. First Conclusions

Cryptocurrencies can be created with a specific purpose in mind but as we can see on the graphs, their prices, volume traded, and market cap have the same spikes at the same events: they seem to be equally affected by exogenous shocks.

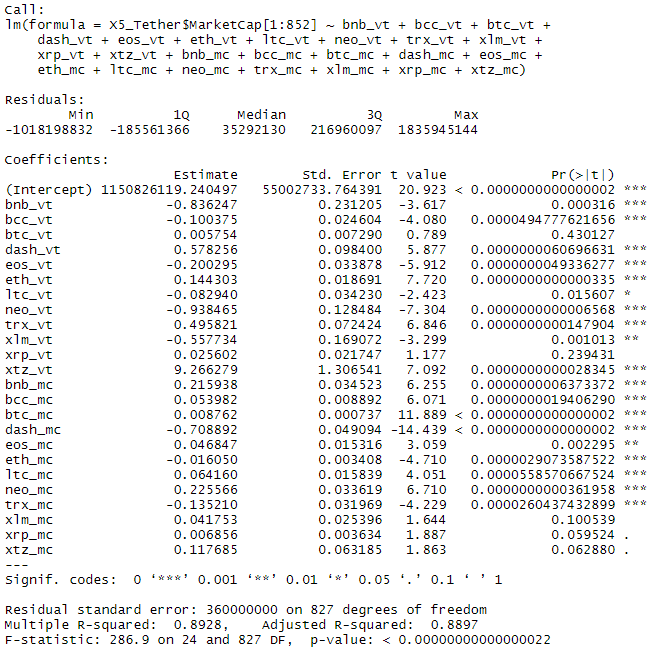
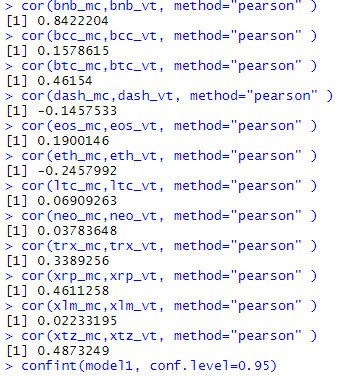
We used covariance matrices to observe this:

Market Cap:

On 144 observations, 28 are 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 at the same time in similar ways, however we cannot determine yet how much

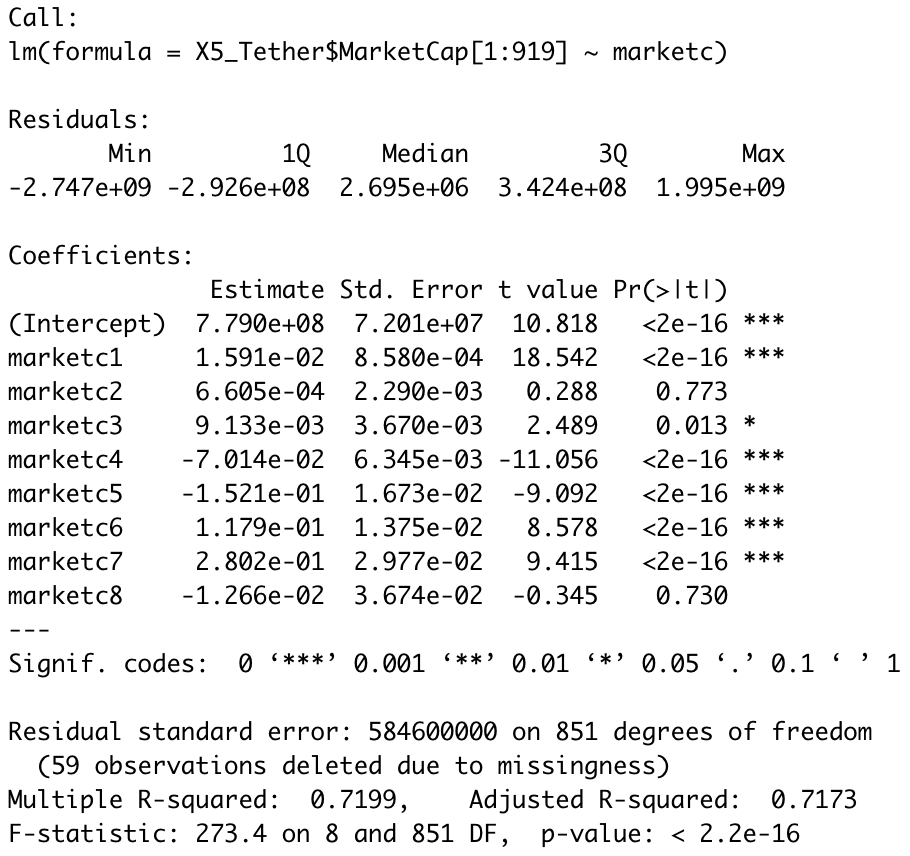
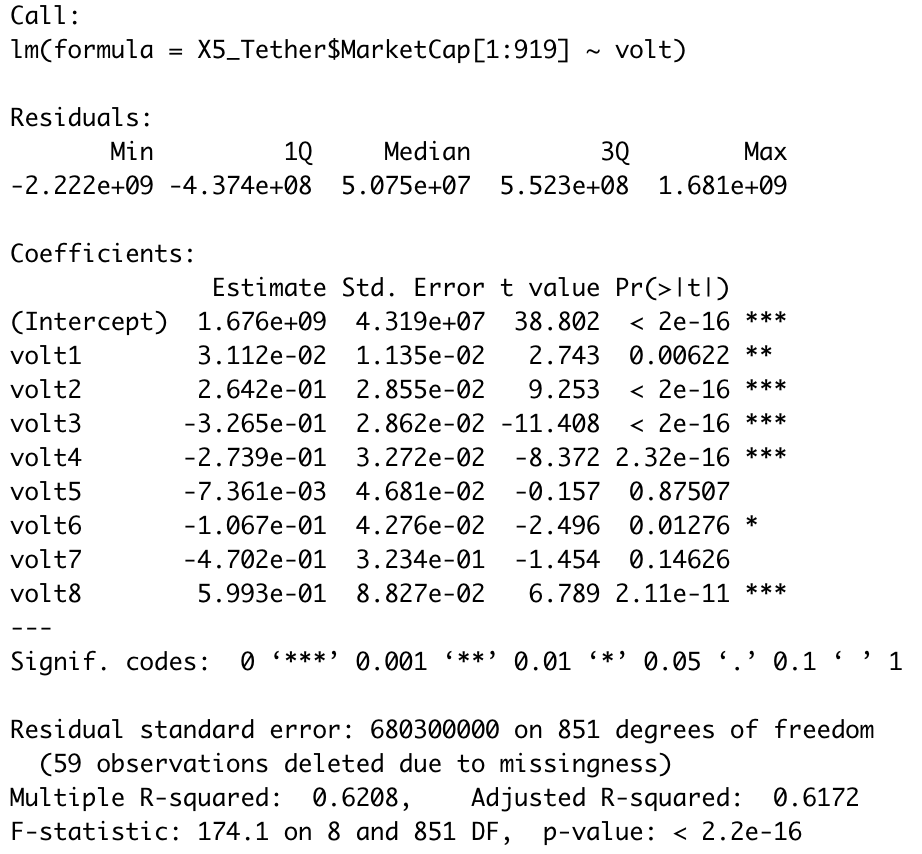
We tested the Multilinear Regression Model on a limited sample using their volume traded and market cap:

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 doubt the ability of this model to efficiently predict the market cap of the Tether:

it might simply be overfitting.

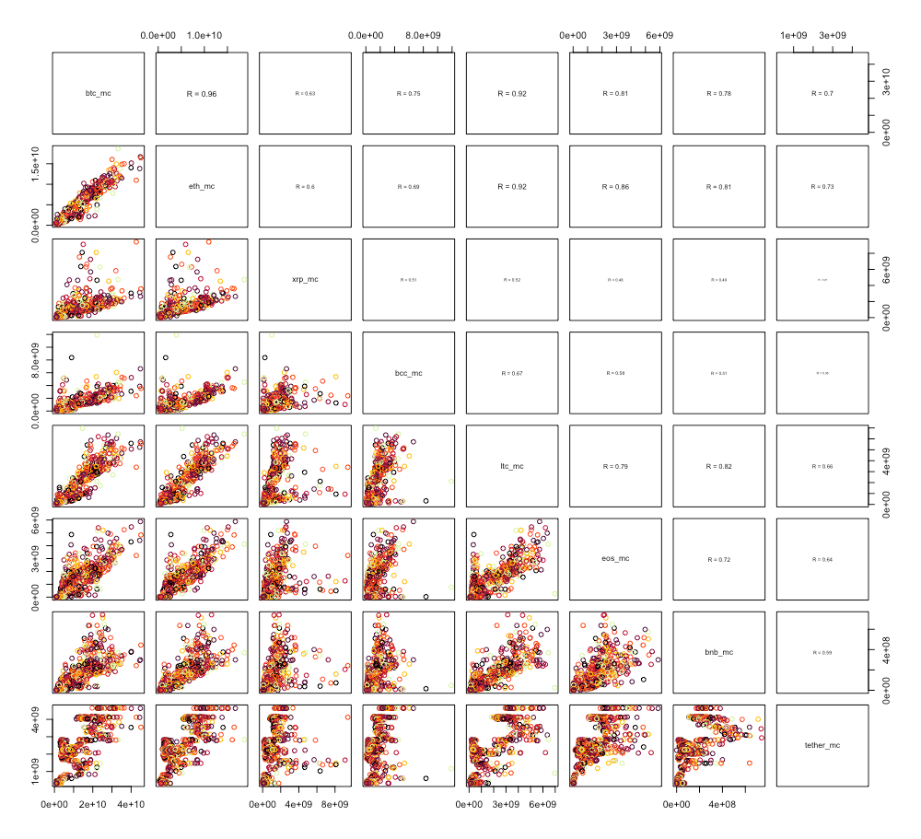
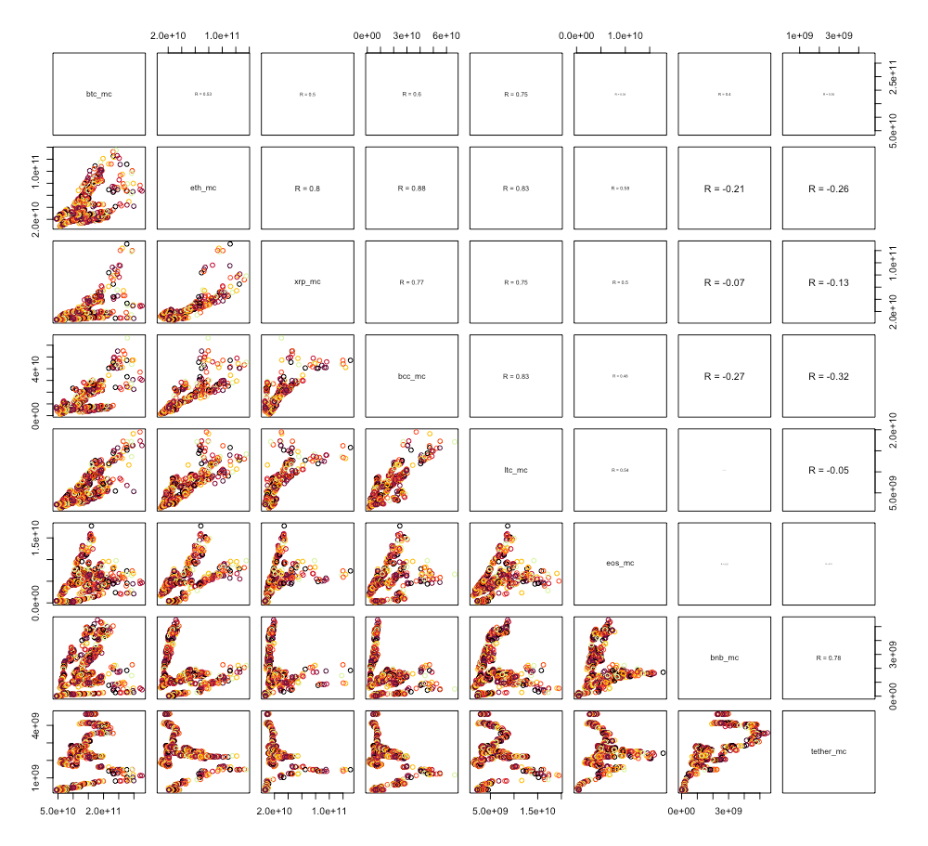
From what we saw in the previous part a regression using many variables, both market cap and volume traded would overfit so we decided to compare which of the two is the most correlated reducing the number of  
 variables, but on a longer window of time (not all data have an historic that long):



Color palet:

|  |  |
| --- | --- |
| * Bitcoin      * Ethereum      * Ripple      * Bitcoin Cash | * Litecoin      * EOS      * Binance Coin      * Tether |

(hyperlink to full size pictures of the matrices included)

[Volume Traded](https://ibb.co/L8wbWhB): [Market Cap](https://ibb.co/1KRRzk0):

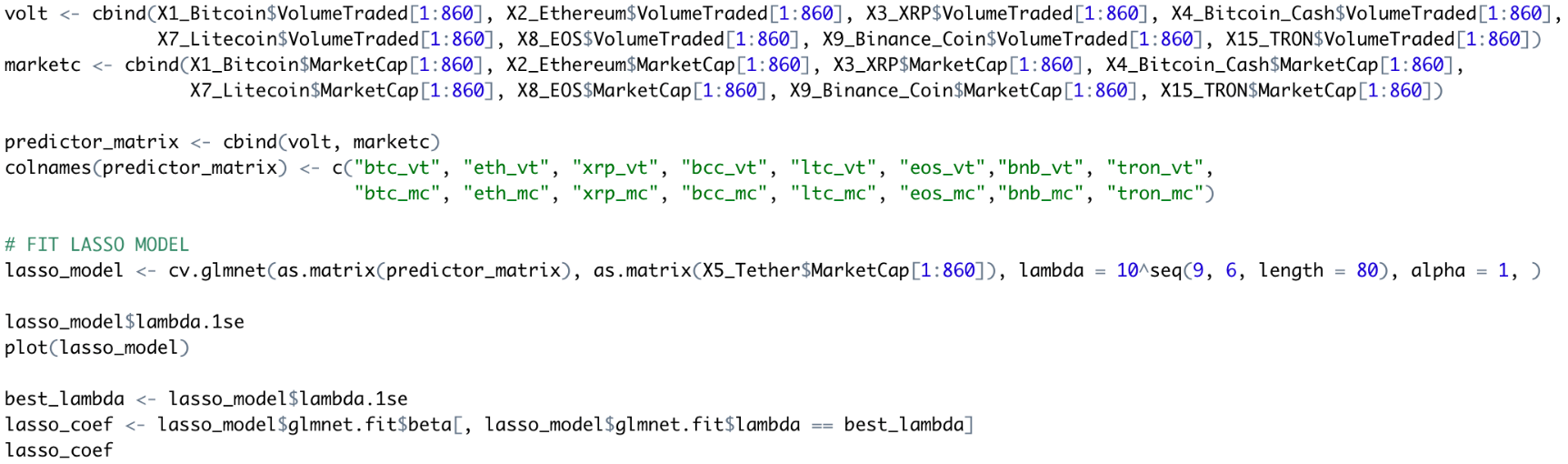
From these matrices we learn that the Volume Traded is much more correlated than the Market Cap, which confirms our previous assumption about the exogeneity of demand, and from the comparison of the linear models we also learn that Market Cap is a slightly better estimator than the Volume Traded. More important, the R² is much lower than it was with the previous linear model, which might indicate that the high R² was caused more because of the high number of variables than because of an accurate model.

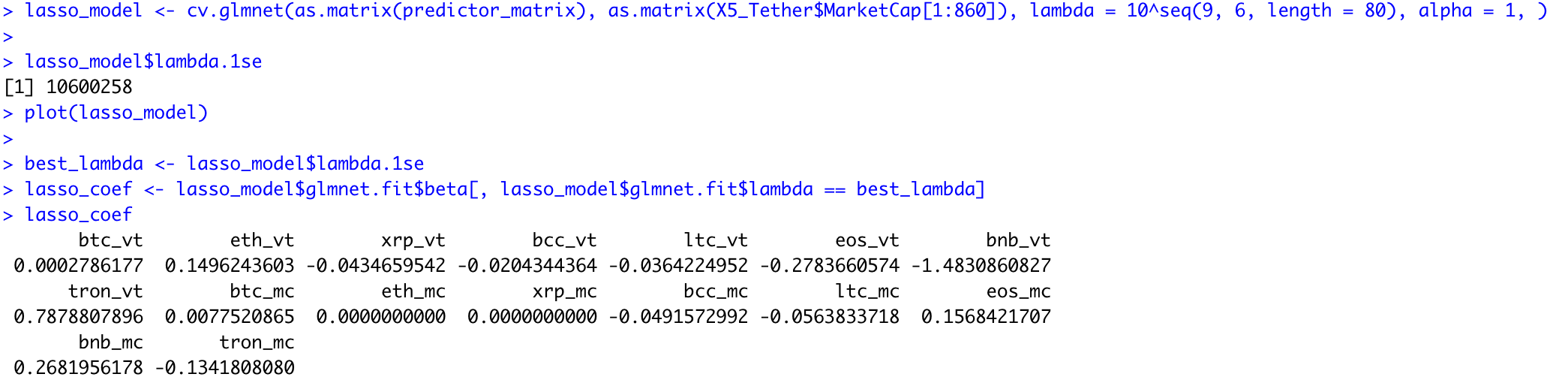
In the next part we will proceed to the regularization and prediction of the data.

1. Machine Learning Models

From what we saw in the previous part a regression using many variables, both market cap and volume traded would overfit. In the next part we will compare different Regularization methods to select which variable should we use in the prediction part.

1. Regularization:
2. Lasso

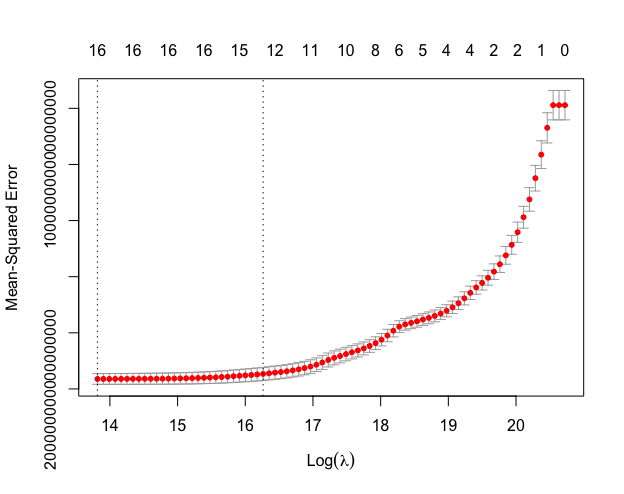
We want to select among our variables the best predictors for the Tether’s market cap, so we perform a lasso regression:

With output:

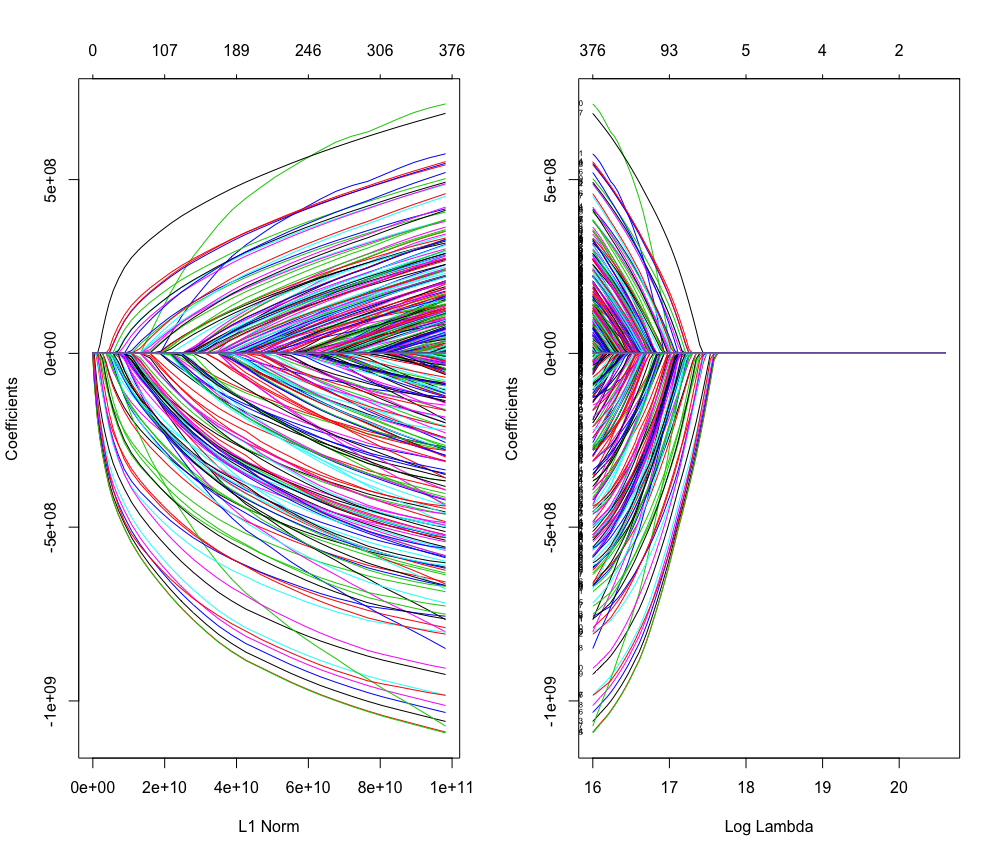
With lasso\_model$Lambda.1se we receive the best λ : 10600258

The absolute value penalty places a constant penalty on deviation from 0,

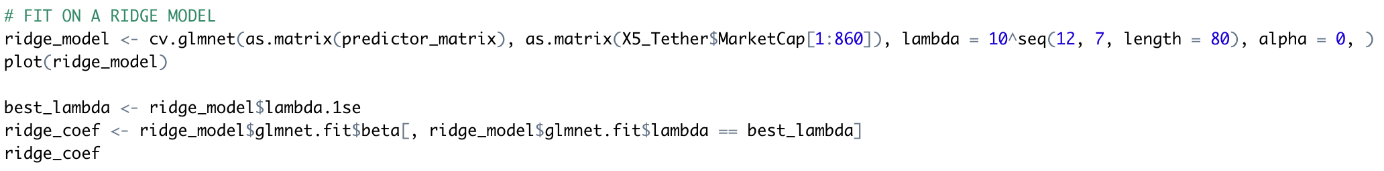
And lasso\_coef gives us the corrected variable coefficients, and as expect some variable’s effect on the Tether’s market cap is null: The Ethereum and the Ripple’s market cap has no effect on the Tether. The optimal Log-Lambda lying between 16 and 17.

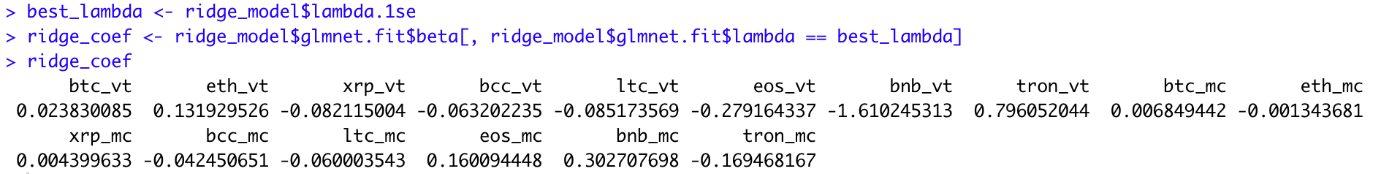


And here is the plot of the Mean-Squared Error



1. Ridge

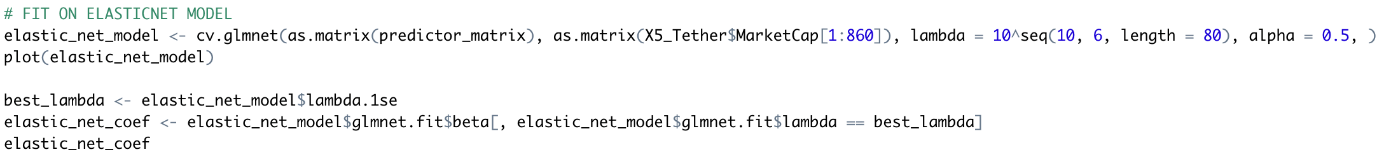
With the same data we performed a regularization using the ridge model:

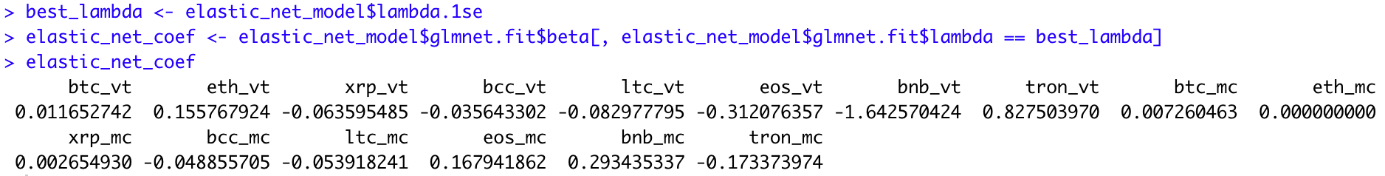
The coefficients:

The model is working well, small impact variables aren’t set to 0 like the lasso does, moderated marginal impact of the biggest variables

1. Elastic net

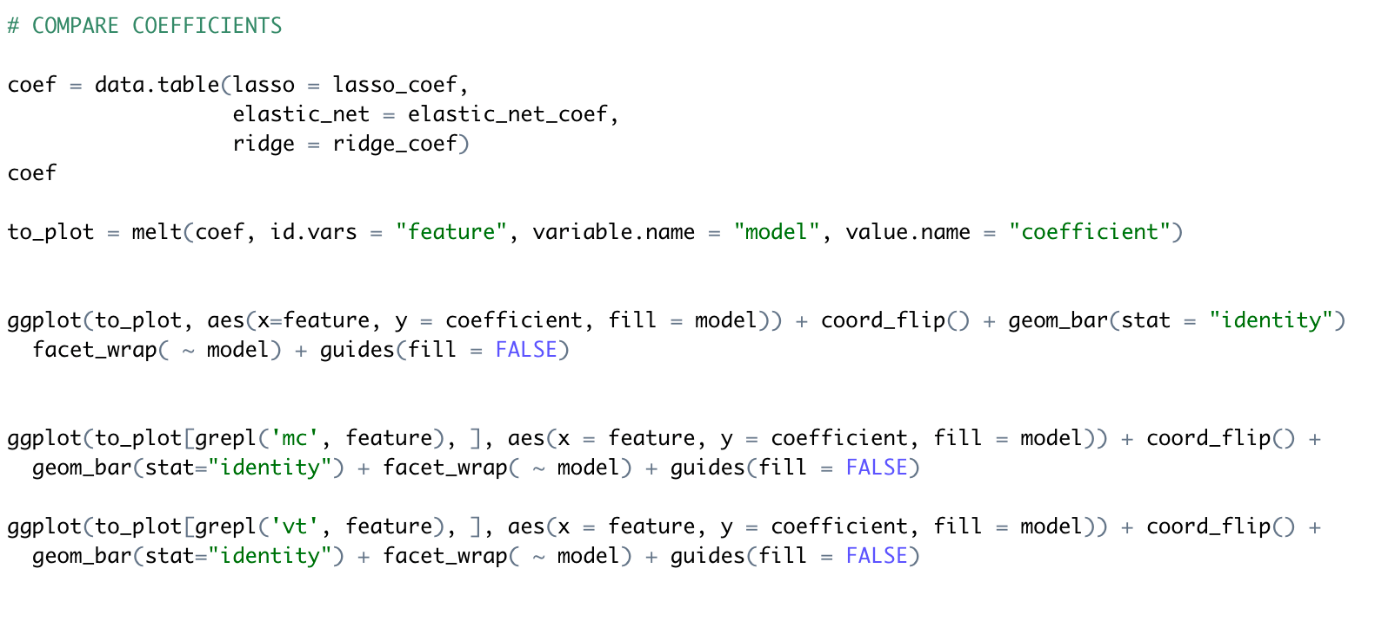
The elastic net is a simple combination of ridge and lasso:

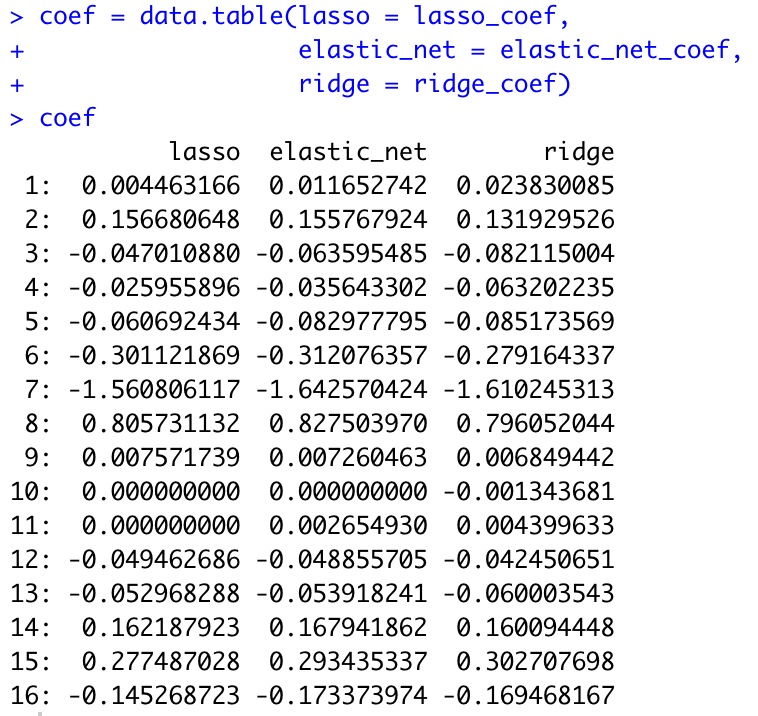


The coefficients:

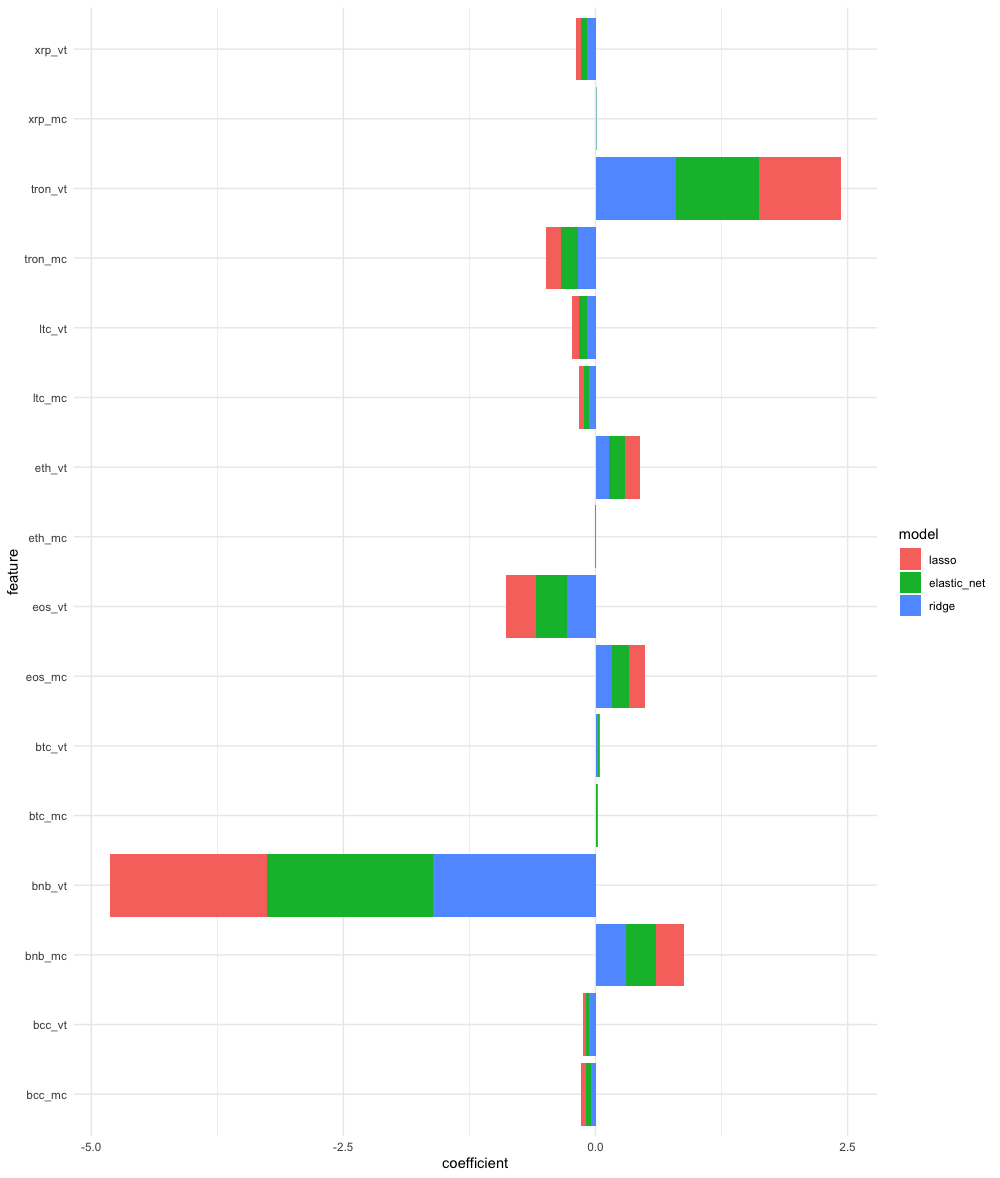
It confirms what we thought: it is ‘less severe’ than the Lasso because the Ethereum’s Market Cap is screened out, but the Ripple’s Market Cap isn’t.

1. Comparison:

These 3 models return different coefficients and mean-squared error we will see below:

Coefficients:

Upon a quick glance, it might seem that the three methods don’t agree much. However, when taking a deeper dive, it would seem that they are in alliance. In fact, we can see that they move in the same direction and differ by very little; some more negligible than others. Despite this, it seems to be the case that the Lasso and the Elastic Net methods agree amongst themselves more than they do with the Ridge method. Although it is close, the Ridge method is not as consistent as the other two methods. As a result, the data confirms our assumptions and thoughts.



When it comes to the comparison between Market Cap and Volume Traded, we can see something quite impressive regarding the three methods; lasso method, elastic net method, and ridge method. It would seem that within a single cryptocurrency, it would seem that the three methods are in agreement. In fact, they move in the same direction. However, it is nonetheless important that they vary very slightly. It would seem that in terms of coefficients, we observe that the Market Cap of Ripple, Tron, Litecoin and Bitcoin Cash considering the three methods, tend to yield negative coefficients. Whereas the Market Caps of EOS, Bitcoin and Binance Coin tend towards a positive coefficient. This shows that these coins either negatively or positively affect the Market Cap of Tether, depending on the coin.

On the other hand, when it comes to the Volume traded, we observe a different outcome. Indeed, it would seem that the Volume traded of Ripple, Litecoin, EOS, Binance coin and Bitcoin Cash have a negative tendency with respect to correlation Tether. Alternatively, it looks as though the Volume traded of Tron and Ethereum are of positive coefficients.

1. Prediction:

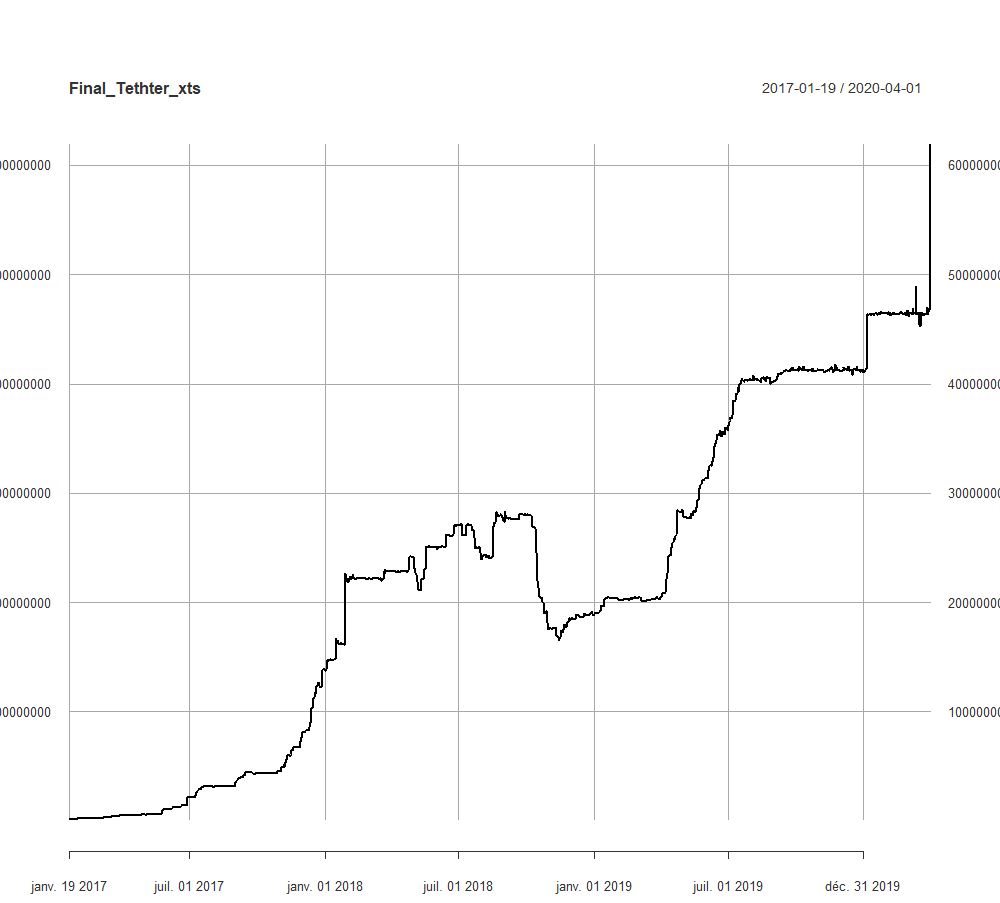
Having identified which variables would be ideal in this model we decided to use a Support Vector Regression to predict how will the Tether evolve in consequence.

Support Vector Regression uses the same idea as Support Vector Machine studied in class but applies it to real values rather than a class: It doesn’t try to predict a class but a value.

We will first predict values for the parameters of the regression:

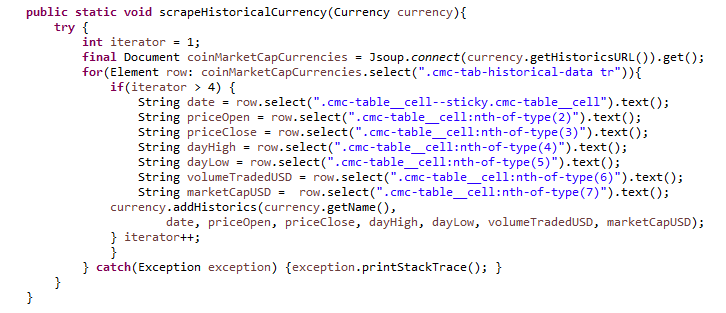
We will use these predicted values (above is just a sample, the full code is in annex) to predict the value of the Tether:



1. Conclusion
2. Annex

Here is most of the code we used for this project:

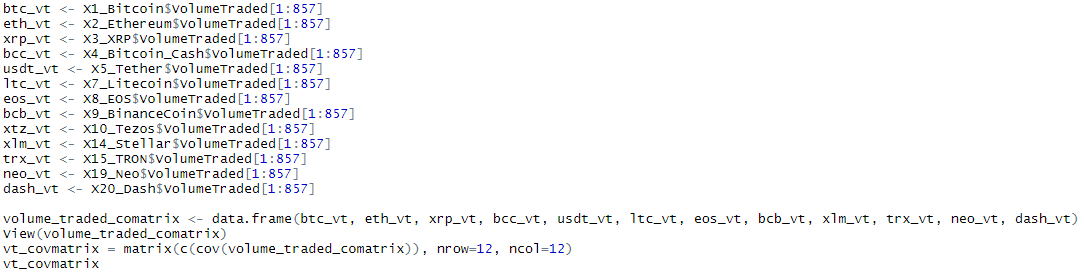
* 1. Java

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.

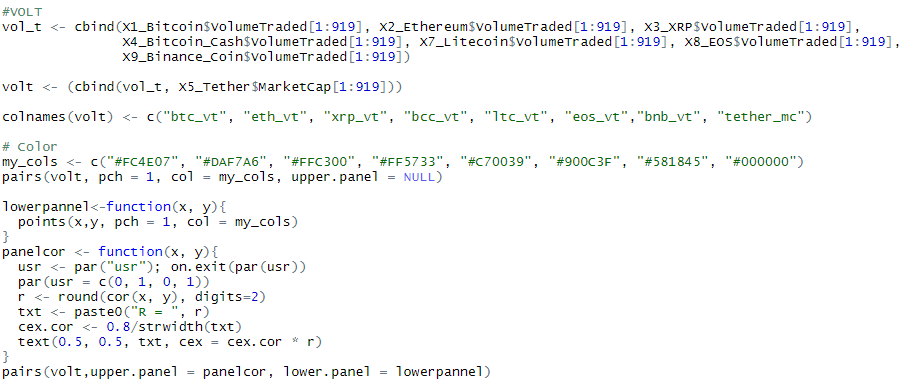


* 1. R

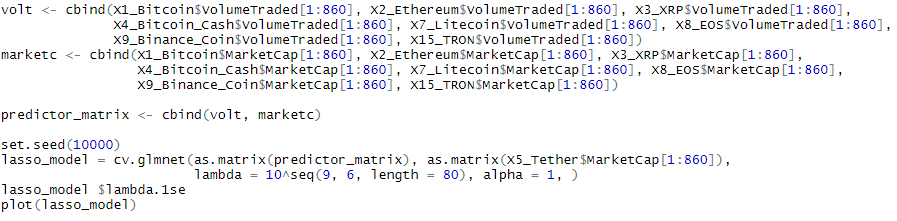
Code we used for the covariance matrices of the market cap and volume traded

Code for the crytprocurrencies plots ‘First Results’:

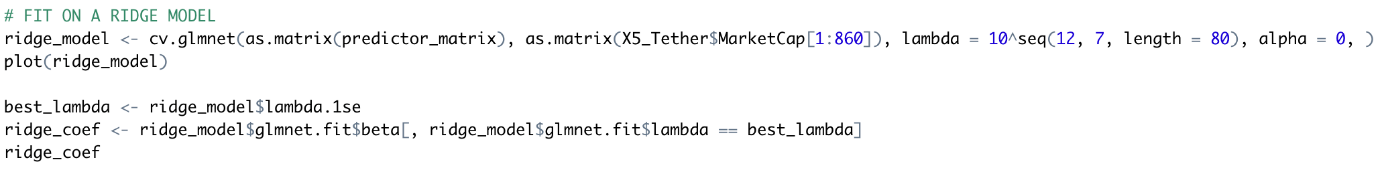
Code for the linear models and matrices (for Volume Traded, Market Cap is the same):



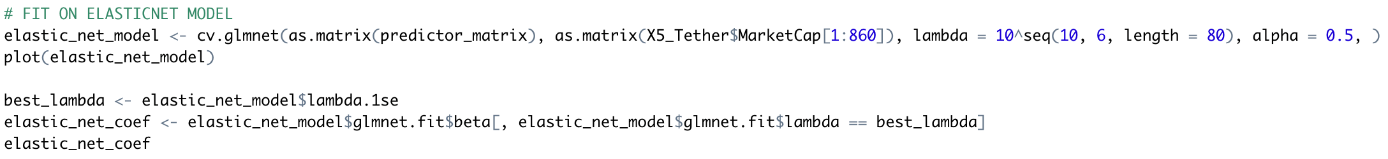
Lasso Regression:

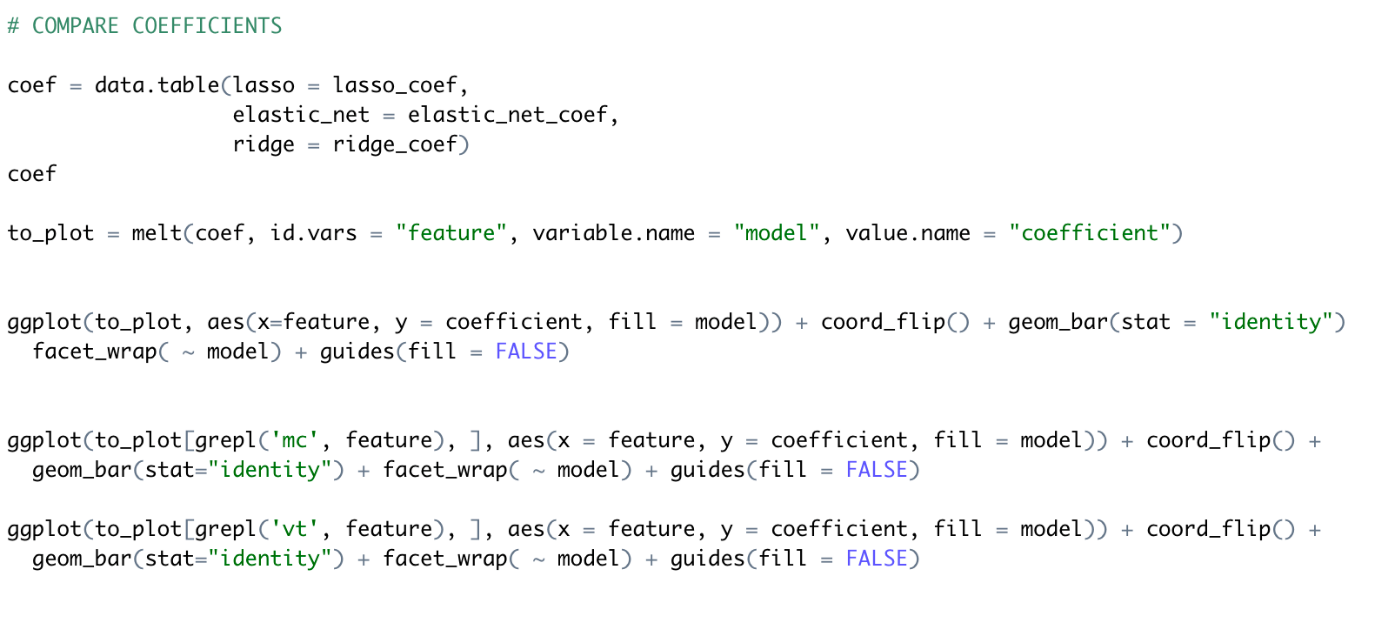


Ridge Regression:



Elastic net Regression:



Coefficient Comparison & Plots:

Support Vector Regression:

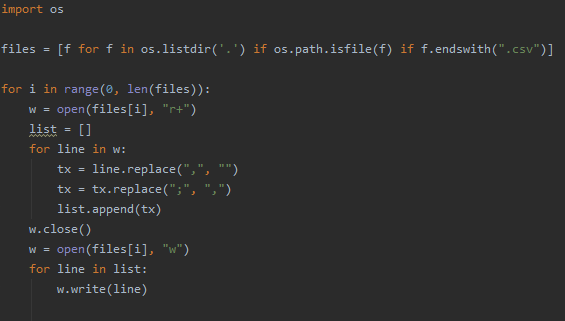
Une image contenant capture d’écran

Description générée automatiquement





* 1. Python

Python code used to arrange the .csv so he can be correctly used by R (comma is used to separate thousands) in the scraped data.

1. Sources
   1. Data:

* <https://coinmarketcap.com/currencies/>
  1. R:
* <https://github.com/TaddyLab/BDS>
* <http://www.sthda.com/english/wiki/scatter-plot-matrices-r-base-graphs>
* <http://www.sthda.com/english/articles/32-r-graphics-essentials/128-plot-time-series-data-using-ggplot/>
* <https://www.r-bloggers.com/plotting-time-series-data-using-ggplot2/>
* <http://www.sthda.com/english/articles/32-r-graphics-essentials/128-plot-time-series-data-using-ggplot/>
* <https://www.svm-tutorial.com/2014/10/support-vector-regression-r/>
* <https://www.kdnuggets.com/2017/03/building-regression-models-support-vector-regression.html>
* <http://www.semspirit.com/artificial-intelligence/machine-learning/regression/support-vector-regression/support-vector-regression-in-r/>
* <https://medium.com/coinmonks/support-vector-regression-or-svr-8eb3acf6d0ffv>
* <https://web.stanford.edu/~hastie/glmnet/glmnet_alpha.html>
* <https://www.rstatisticsblog.com/data-science-in-action/lasso-regression/>
* <https://www.r-bloggers.com/ridge-regression-and-the-lasso/>
* <https://www.youtube.com/watch?v=FWCPFUwZkn0>
* Business Data Science – Matt Taddy
  1. Diverse:
* <https://www.bloomberg.com/news/articles/2018-11-20/bitcoin-rigging-criminal-probe-is-said-to-focus-on-tie-to-tether>
* <https://arstechnica.com/tech-policy/2018/02/tether-says-its-cryptocurrency-is-worth-2-billion-but-its-audit-failed/>

1. Research Proposal is mostly composed of the first assignment [↑](#footnote-ref-1)
2. First Results is for the most part the content of the second assignment [↑](#footnote-ref-2)
3. Few platforms accept to sell cryptocurrencies against real money (called fiat). Most platform accepting fiat make money on conversion rate euro/bitcoin, usd/btc while exchange platforms, where people trade cryptocurrencies make money on transaction fees [↑](#footnote-ref-3)
4. Every time money is moved from a wallet to another, a transaction fee must be paid to secure the transaction. A computer will record the movement through a secured process and add the amount, the origin, the destination, the date and the time to a public register that will make sure this transaction can’t be undone. [↑](#footnote-ref-4)
5. We will be talking about the USDT for “United State Dollar Tether”, the full name of this asset [↑](#footnote-ref-5)
6. The Tether is not the only stable coin, but its market cap and volume traded everyday are respectively 10 times and 100 times more important than the second biggest stable coin, so I didn’t judge necessary to make it part of the research proposal [↑](#footnote-ref-6)
7. many of the 30 biggest were created after 2018, this is why we only used a few ones [↑](#footnote-ref-7)