Hugo Joncour - 40139130
Zié Coulibaly - 40046155

ECON 422: Econometrics 2 - Machine Learning and Economics

Assignment 2

1. Introduction

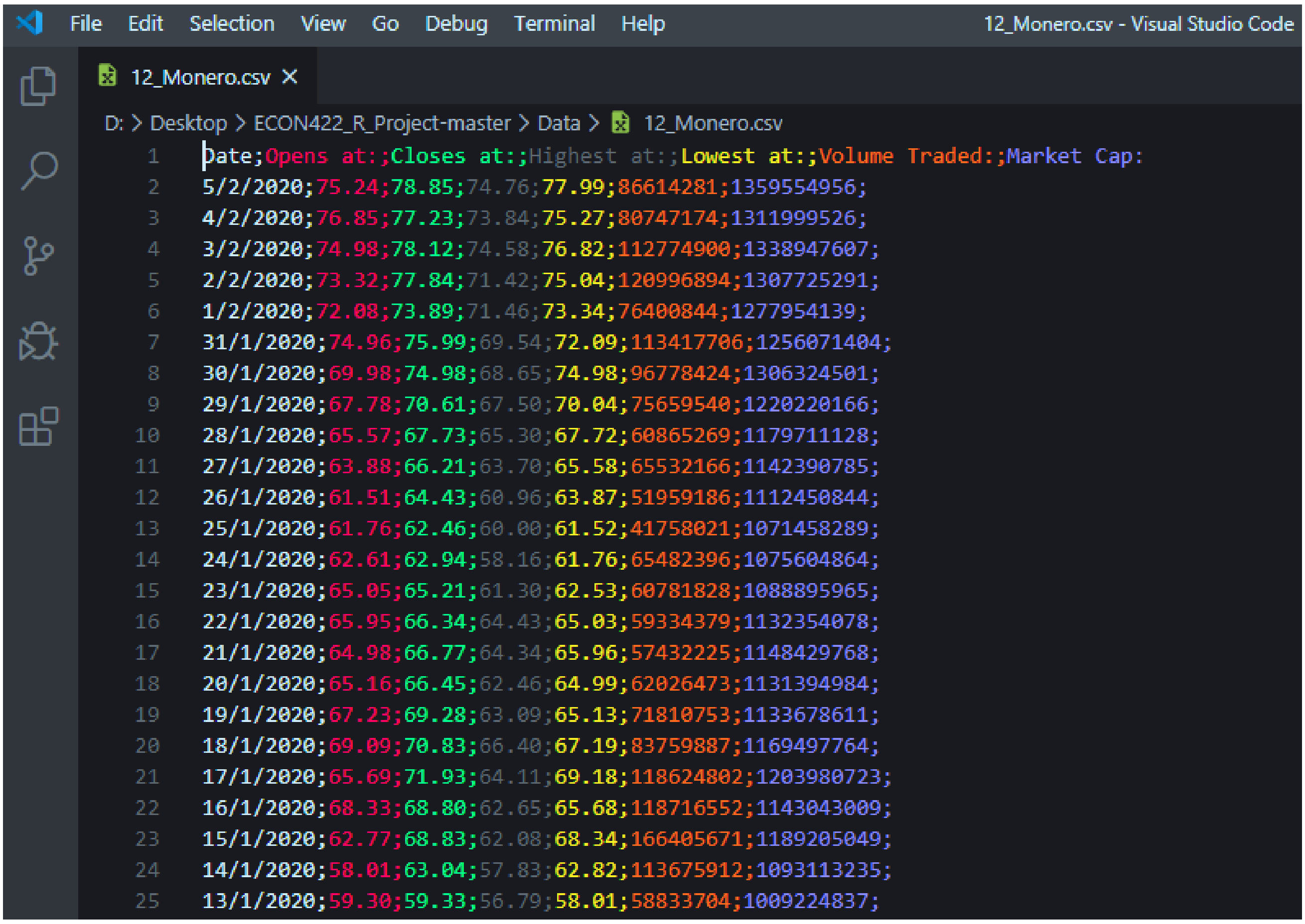
Research Project:
Using Machine Learning to analyze cryptocurrency trends:
Can we predict Tether's market cap evolution using other cryptocurrencies's demand ?

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.

2. 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.




```
Date;Opens at:;Closes at:;Highest at:;Lowest at:;Volume Traded:;Market Cap:
5/2/2020;75.24;78.85;74.76;77.99;86614281;1359554956;
4/2/2020;76.85;77.23;73.84;75.27;80747174;1311999526;
3/2/2020;74.98;78.12;74.58;76.82;112774900;1338947607;
2/2/2020;73.32;77.84;71.42;75.04;120996894;1307725291;
1/2/2020;72.08;73.89;71.46;73.34;76400844;1277954139;
31/1/2020;74.96;75.99;69.54;72.09;113417706;1256071404;
30/1/2020;69.98;74.98;68.65;74.98;96778424;1306324501;
29/1/2020;67.78;70.61;67.50;70.04;75659540;1220220166;
28/1/2020;65.57;67.73;65.30;67.72;60865269;1179711128;
27/1/2020;63.88;66.21;63.70;65.58;65532166;1142390785;
26/1/2020;61.51;64.43;60.96;63.87;51959186;1112450844;
25/1/2020;61.76;62.46;60.00;61.52;41758021;1071458289;
24/1/2020;62.61;62.94;58.16;61.76;65482396;1075604864;
23/1/2020;65.05;65.21;61.30;62.53;60781828;1088895965;
22/1/2020;65.95;66.34;64.43;65.03;59334379;1132354078;
21/1/2020;64.98;66.77;64.34;65.96;57432225;1148429768;
20/1/2020;65.16;66.45;62.46;64.99;62026473;1131394984;
19/1/2020;67.23;69.28;63.09;65.13;71810753;1133678611;
18/1/2020;69.09;70.83;66.40;67.19;83759887;1169497764;
17/1/2020;65.69;71.93;64.11;69.18;118624802;1203980723;
16/1/2020;68.33;68.80;62.65;65.68;118716552;1143043009;
15/1/2020;62.77;68.83;62.08;68.34;166405671;1189205049;
14/1/2020;58.01;63.04;57.83;62.82;113675912;1093113235;
13/1/2020;59.30;59.33;56.79;58.01;58833704;1009224837;
```

Hugo Joncour - 40139130
Zié Coulibaly - 40046155

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

3. Preliminary Results

For now, we only used a sample of the data available (10

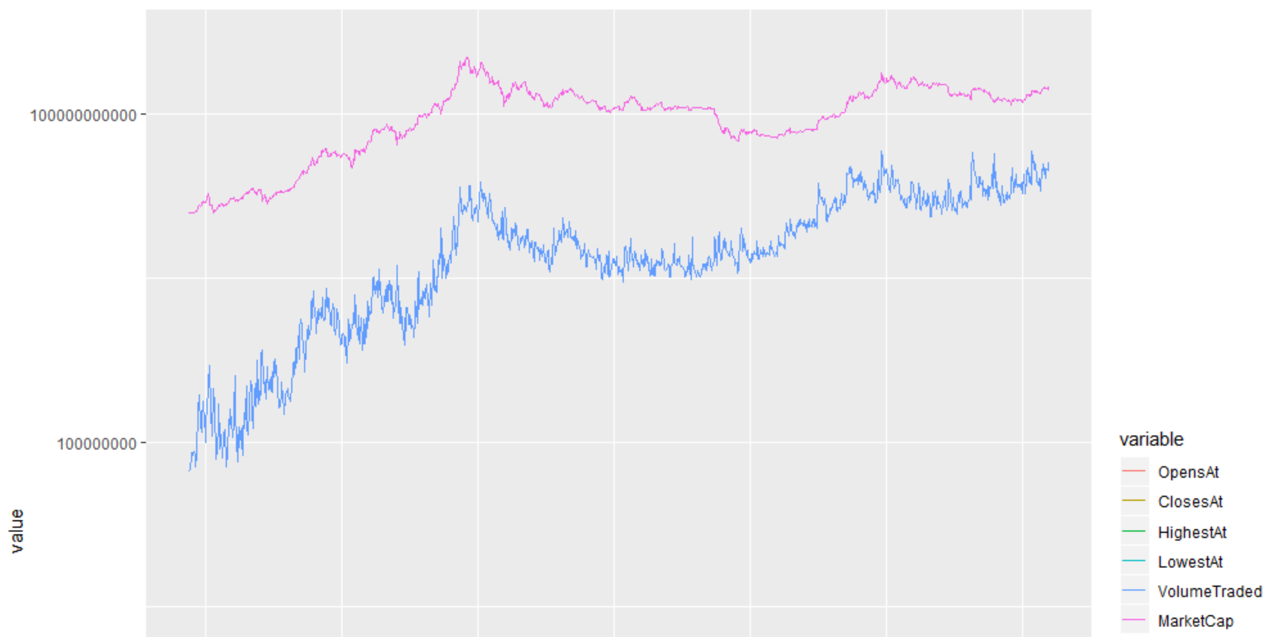
```
#Graph Creation (Normal & Log)


#Scientific to numerical notation
options(scipen = 999)

data <- Bitcoin #Coin file
data$Date <- dmy(data$Date)

#Normal portion
meltdata <- melt(data,id="Date")
ggplot(meltdata,aes(x=Date,y=value,colour=variable,group=variable)) +
  geom_line()

#Log portion
lg <- ggplot(meltdata,aes(x=Date,y=value,colour=variable,group=variable)) +
  geom_line()

lg + scale_y_continuous(trans = log10_trans())
```

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:

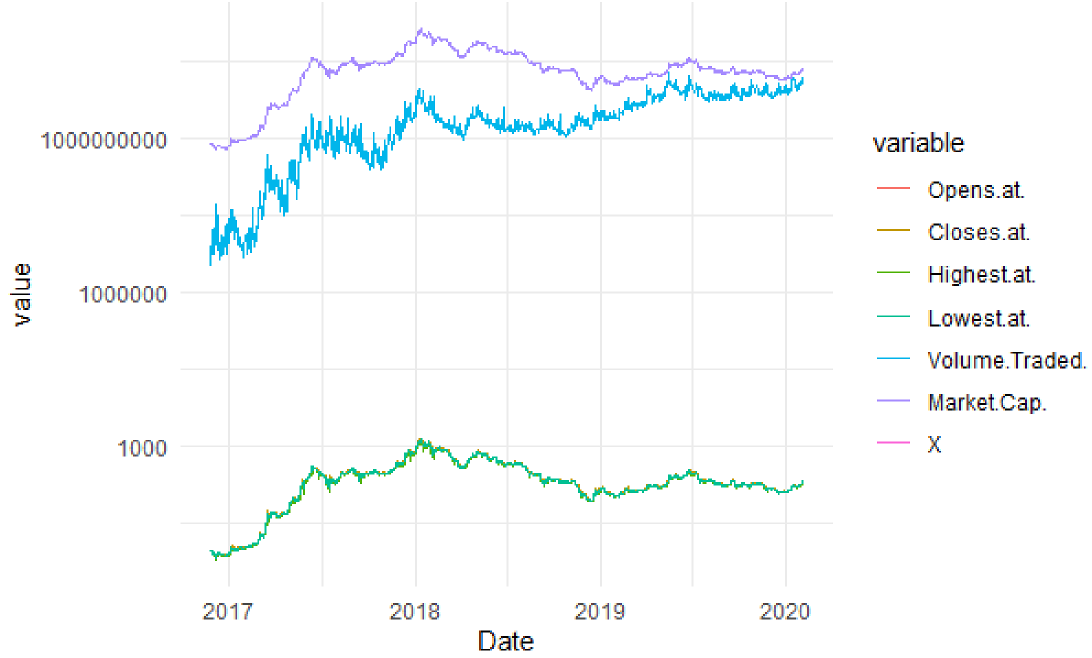
Bitcoin is the biggest and most important cryptocurrency by most metrics. It is worth 63% of the global market cap of all cryptocurrencies.



Hugo Joncour - 40139130
Zié Coulibaly - 40046155

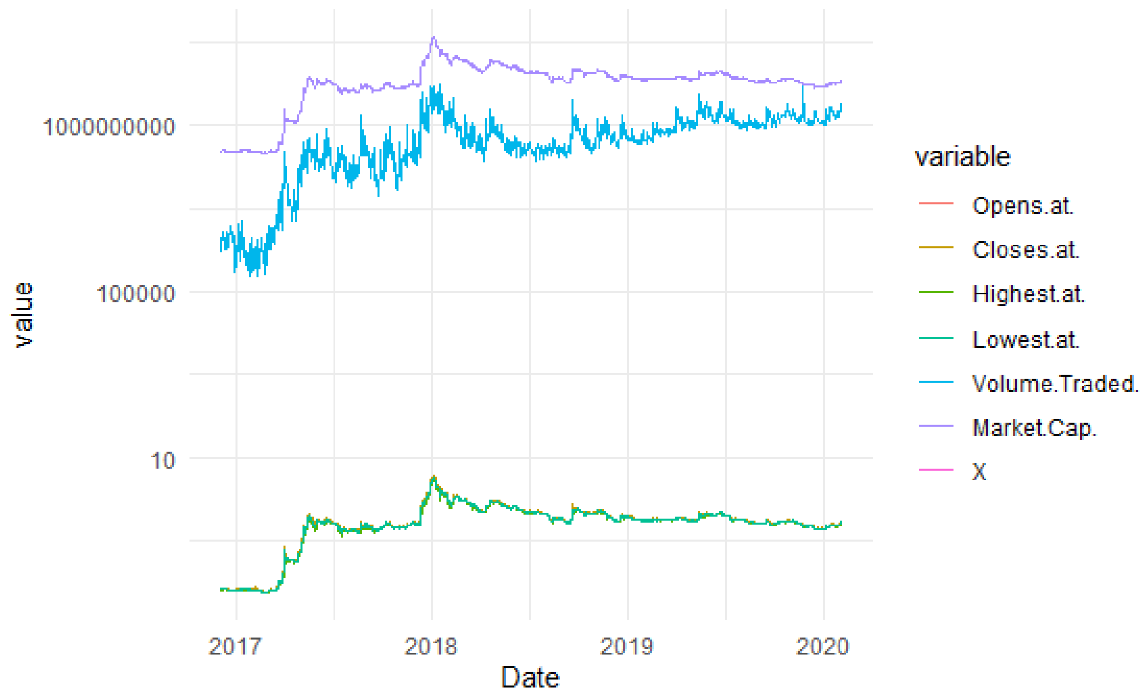
Ethereum:

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

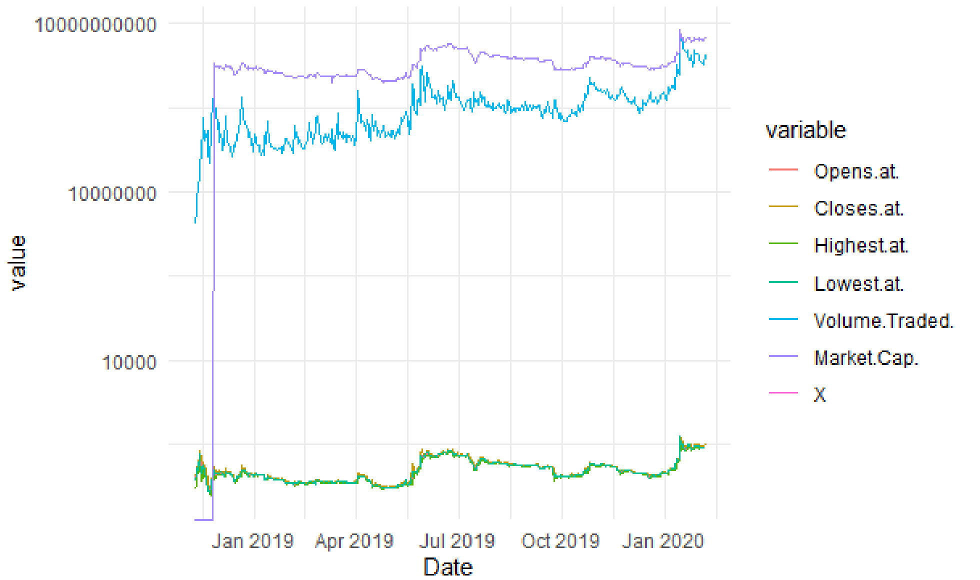
Ripple was created to facilitate financial settlement and money transfers through a cryptocurrency protocol providing security through encryption, cost efficiency and no third-party transaction fee.
It is used by many financial institutions for real-time gross settlement, and unlike most cryptocurrencies, has no limit to the number of units that can be emitted.



Hugo Joncour - 40139130
Zié Coulibaly - 40046155

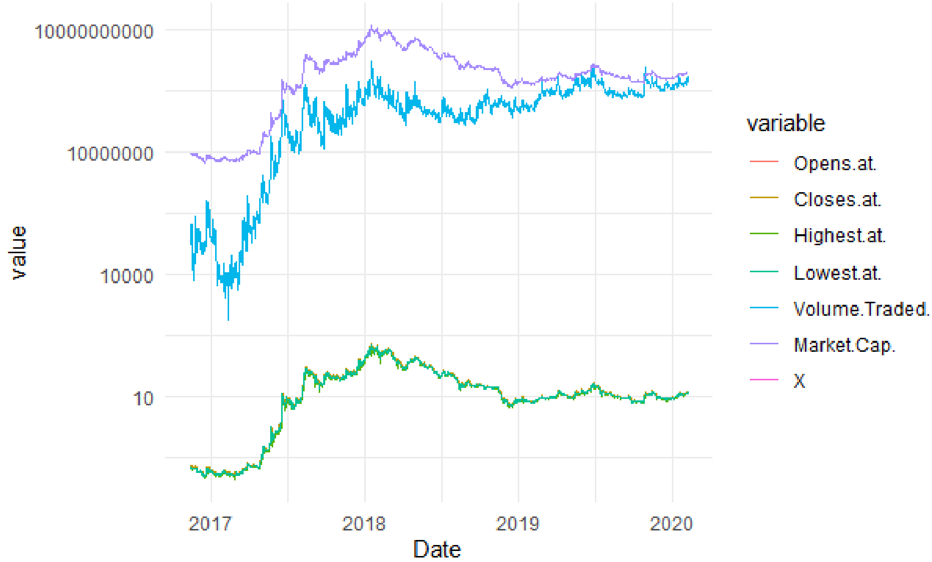
Bitcoin SV[1]:

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

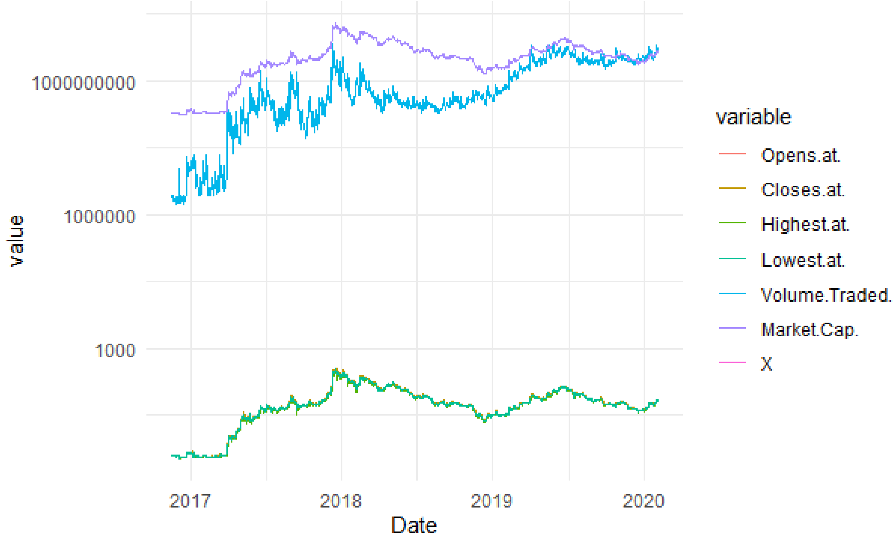
Cryptocurrency used for a protocol allowing cross border transactions between real world currencies.
Most notably used by Deloitte for its Deloitte Digital Bank and IBM

[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.

Hugo Joncour - 40139130
Zié Coulibaly - 40046155

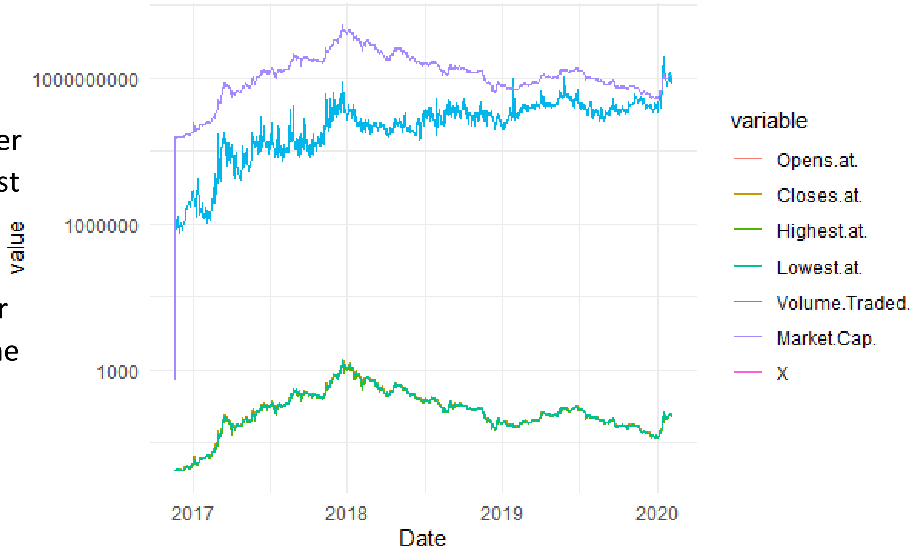
Litecoin

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



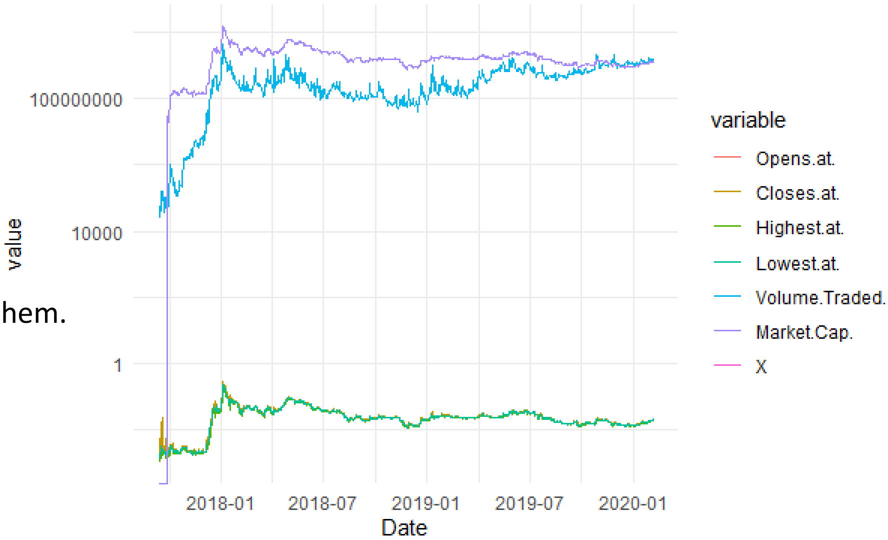
Dash

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.



Hugo Joncour - 40139130
Zié Coulibaly - 40046155

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:

```
              [,1]          [,2]          [,3]          [,4]          [,5]          [,6]
 [1,] 2.384102e+21  6.410551e+20  3.354263e+20  2.953990e+20  1.266103e+19  1.134423e+20
 [2,] 6.410551e+20  5.899837e+20  2.837552e+20  2.085586e+20 -8.731644e+18  6.326402e+19
 [3,] 3.354263e+20  2.837552e+20  2.125945e+20  1.110426e+20 -4.232475e+18  3.377801e+19
 [4,] 2.953990e+20  2.085586e+20  1.110426e+20  9.566095e+19 -4.390558e+18  2.539125e+19
 [5,] 1.266103e+19 -8.731644e+18 -4.232475e+18 -4.390558e+18  1.196380e+18 -5.632953e+17
 [6,] 1.134423e+20  6.326402e+19  3.377801e+19  2.539125e+19 -5.632953e+17  9.733276e+18
 [7,] 3.354481e+19  4.110833e+19  1.920425e+19  1.326382e+19  1.180160e+17  4.468126e+18
 [8,] 2.047979e+19 -7.719876e+18 -2.792810e+18 -3.962501e+18  1.053517e+18  2.121627e+17
 [9,] 3.498544e+19  4.849201e+19  2.914468e+19  1.650922e+19 -7.398377e+17  4.748859e+18
[10,] 2.535342e+19  2.621885e+19  1.787643e+19  9.102720e+18 -1.703787e+17  2.973410e+18
[11,] 4.656583e+19  5.071007e+19  2.427180e+19  1.689448e+19 -8.050920e+17  5.220948e+18
[12,] 5.819478e+19  4.142639e+19  2.247455e+19  1.853952e+19 -1.081557e+18  5.197601e+18
              [,7]          [,8]          [,9]         [,10]         [,11]         [,12]
 [1,] 3.354481e+19  2.047979e+19  3.498544e+19  2.535342e+19  4.656583e+19  5.819478e+19
 [2,] 4.110833e+19 -7.719876e+18  4.849201e+19  2.621885e+19  5.071007e+19  4.142639e+19
 [3,] 1.920425e+19 -2.792810e+18  2.914468e+19  1.787643e+19  2.427180e+19  2.247455e+19
 [4,] 1.326382e+19 -3.962501e+18  1.650922e+19  9.102720e+18  1.689448e+19  1.853952e+19
 [5,] 1.180160e+17  1.053517e+18 -7.398377e+17 -1.703787e+17 -8.050920e+17 -1.081557e+18
 [6,] 4.468126e+18  2.121627e+17  4.748859e+18  2.973410e+18  5.220948e+18  5.197601e+18
 [7,] 7.834798e+18  4.175219e+17  4.704941e+18  3.053226e+18  3.152210e+18  1.872131e+18
 [8,] 4.175219e+17  1.670466e+18 -8.256872e+17  2.998361e+16 -8.618919e+17 -9.348365e+17
 [9,] 4.704941e+18 -8.256872e+17  5.657571e+18  2.860813e+18  4.278559e+18  3.147959e+18
[10,] 3.053226e+18  2.998361e+16  2.860813e+18  2.169485e+18  2.123465e+18  1.650647e+18
[11,] 3.152210e+18 -8.618919e+17  4.278559e+18  2.123465e+18  4.731930e+18  3.473734e+18
[12,] 1.872131e+18 -9.348365e+17  3.147959e+18  1.650647e+18  3.473734e+18  3.872977e+18
```

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

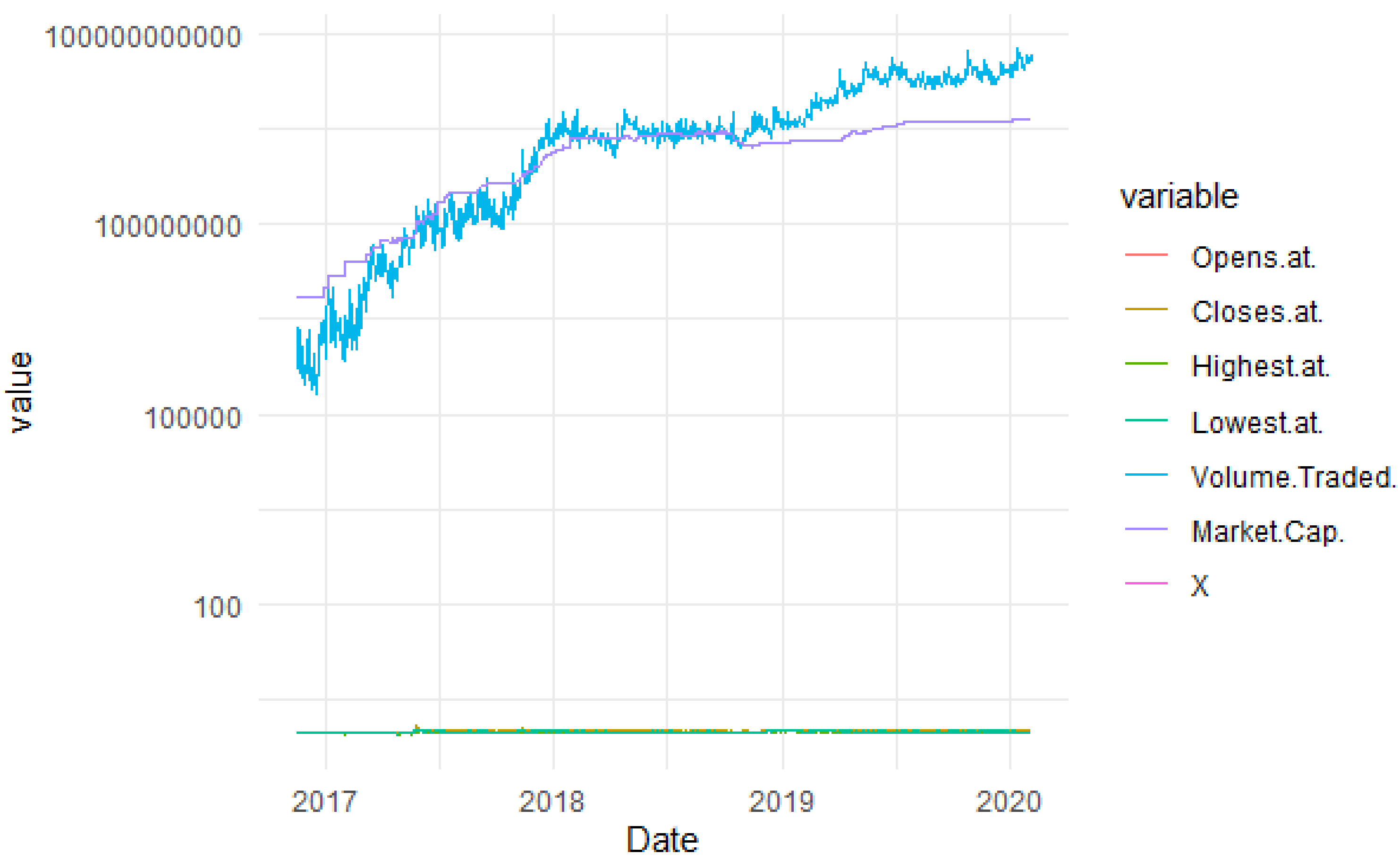
Volume Traded:

```
              [,1]         [,2]         [,3]         [,4]         [,5]         [,6]
 [1,] 7.065249e+19 2.727197e+19 5.605474e+18 6.884347e+18 8.419076e+19 1.175717e+19
 [2,] 2.727197e+19 1.157185e+19 2.148614e+18 2.548551e+18 3.440961e+19 4.801735e+18
 [3,] 5.605474e+18 2.148614e+18 1.184311e+18 5.989699e+17 5.074413e+18 8.328215e+17
 [4,] 6.884347e+18 2.548551e+18 5.989699e+17 1.203901e+18 7.750228e+18 1.109931e+18
 [5,] 8.419076e+19 3.440961e+19 5.074413e+18 7.750228e+18 1.147136e+20 1.462879e+19
 [6,] 1.175717e+19 4.801735e+18 8.328215e+17 1.109931e+18 1.462879e+19 2.333193e+18
 [7,] 6.921795e+18 2.978955e+18 5.128867e+17 6.506201e+17 8.982691e+18 1.231165e+18
 [8,] 8.338763e+17 3.501490e+17 6.611941e+16 7.029641e+16 9.906209e+17 1.603951e+17
 [9,] 9.306191e+17 3.792120e+17 1.189279e+17 9.370451e+16 9.833316e+17 1.610314e+17
[10,] 2.936515e+18 1.159380e+18 2.981960e+17 2.912081e+17 3.431081e+18 4.554602e+17
[11,] 1.264104e+18 5.148565e+17 1.265599e+17 1.146096e+17 1.446519e+18 2.132992e+17
[12,] 1.066638e+18 4.349768e+17 7.212786e+16 1.371825e+17 1.409108e+18 1.822314e+17
              [,7]         [,8]         [,9]        [,10]        [,11]        [,12]
 [1,] 6.921795e+18 8.338763e+17 9.306191e+17 2.936515e+18 1.264104e+18 1.066638e+18
 [2,] 2.978955e+18 3.501490e+17 3.792120e+17 1.159380e+18 5.148565e+17 4.349768e+17
 [3,] 5.128867e+17 6.611941e+16 1.189279e+17 2.981960e+17 1.265599e+17 7.212786e+16
 [4,] 6.506201e+17 7.029641e+16 9.370451e+16 2.912081e+17 1.146096e+17 1.371825e+17
 [5,] 8.982691e+18 9.906209e+17 9.833316e+17 3.431081e+18 1.446519e+18 1.409108e+18
 [6,] 1.231165e+18 1.603951e+17 1.610314e+17 4.554602e+17 2.132992e+17 1.822314e+17
 [7,] 1.072607e+18 9.445412e+16 1.018221e+17 3.097807e+17 1.371556e+17 1.254581e+17
 [8,] 9.445412e+16 1.679515e+16 1.275584e+16 3.688852e+16 1.713980e+16 1.168910e+16
 [9,] 1.018221e+17 1.275584e+16 2.349918e+16 4.387738e+16 2.025082e+16 1.514114e+16
[10,] 3.097807e+17 3.688852e+16 4.387738e+16 2.036202e+17 5.557382e+16 4.483645e+16
[11,] 1.371556e+17 1.713980e+16 2.025082e+16 5.557382e+16 3.713641e+16 1.842158e+16
[12,] 1.254581e+17 1.168910e+16 1.514114e+16 4.483645e+16 1.842158e+16 4.057914e+16
```

Hugo Joncour - 40139130
Zié Coulibaly - 40046155

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.

Hugo Joncour - 40139130
Zié Coulibaly - 40046155

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

```
Call:
lm(formula = X5_Tether$MarketCap[1:852] ~ bnb_vt + bcc_vt + btc_vt +
    dash_vt + eos_vt + eth_vt + ltc_vt + neo_vt + trx_vt + xlm_vt +
    xrp_vt + xtz_vt + bnb_mc + bcc_mc + btc_mc + dash_mc + eos_mc +
    eth_mc + ltc_mc + neo_mc + trx_mc + xlm_mc + xrp_mc + xtz_mc)

Residuals:
        Min          1Q      Median          3Q         Max
-1018198832  -185561366    35292130   216960097  1835945144

Coefficients:
                   Estimate        Std. Error t value             Pr(>|t|)
(Intercept) 1150826119.240497  55002733.764391  20.923 < 0.0000000000000002 ***
bnb_vt              -0.836247         0.231205  -3.617             0.000316 ***
bcc_vt              -0.100375         0.024604  -4.080    0.000049477621656 ***
btc_vt               0.005754         0.007290   0.789             0.430127
dash_vt              0.578256         0.098400   5.877    0.000000006069663 ***
eos_vt              -0.200295         0.033878  -5.912    0.000000004933627 ***
eth_vt               0.144303         0.018691   7.720    0.000000000000033 ***
ltc_vt              -0.082940         0.034230  -2.423             0.015607 *
neo_vt              -0.938465         0.128484  -7.304    0.000000000000656 ***
trx_vt               0.495821         0.072424   6.846    0.000000000014790 ***
xlm_vt              -0.557734         0.169072  -3.299             0.001013 **
xrp_vt               0.025602         0.021747   1.177             0.239431
xtz_vt               9.266279         1.306541   7.092    0.000000000002834 ***
bnb_mc               0.215938         0.034523   6.255    0.000000000637337 ***
bcc_mc               0.053982         0.008892   6.071    0.000000001940629 ***
btc_mc               0.008762         0.000737  11.889 < 0.0000000000000002 ***
dash_mc             -0.708892         0.049094 -14.439 < 0.0000000000000002 ***
eos_mc               0.046847         0.015316   3.059             0.002295 **
eth_mc              -0.016050         0.003408  -4.710    0.000002907358752 ***
ltc_mc               0.064160         0.015839   4.051    0.000055857066752 ***
neo_mc               0.225566         0.033619   6.710    0.000000000036195 ***
trx_mc              -0.135210         0.031969  -4.229    0.000026043743289 ***
xlm_mc               0.041753         0.025396   1.644             0.100539
xrp_mc               0.006856         0.003634   1.887             0.059524 .
xtz_mc               0.117685         0.063185   1.863             0.062880 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 360000000 on 827 degrees of freedom
Multiple R-squared:  0.8928,	Adjusted R-squared:  0.8897
F-statistic: 286.9 on 24 and 827 DF,  p-value: < 0.00000000000000022
```

A 0.8928 may look like a very good R^2 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 R^2

```
> cor(bnb_mc,bnb_vt, method="pearson" )
[1] 0.8422204
> cor(bcc_mc,bcc_vt, method="pearson" )
[1] 0.1578615
> cor(btc_mc,btc_vt, method="pearson" )
[1] 0.46154
> cor(dash_mc,dash_vt, method="pearson" )
[1] -0.1457533
> cor(eos_mc,eos_vt, method="pearson" )
[1] 0.1900146
> cor(eth_mc,eth_vt, method="pearson" )
[1] -0.2457992
> cor(ltc_mc,ltc_vt, method="pearson" )
[1] 0.06909263
> cor(neo_mc,neo_vt, method="pearson" )
[1] 0.03783648
> cor(trx_mc,trx_vt, method="pearson" )
[1] 0.3389256
> cor(xrp_mc,xrp_vt, method="pearson" )
[1] 0.4611258
> cor(xlm_mc,xlm_vt, method="pearson" )
[1] 0.02233195
> cor(xtz_mc,xtz_vt, method="pearson" )
[1] 0.4873249
> confint(model1, conf.level=0.95)
```

Hugo Joncour - 40139130
Zié Coulibaly - 40046155

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

```
> vt_cormatrix
           [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]
 [1,] 1.0000000 0.9533657 0.6107862 0.7451675 0.9349577 0.9152362 0.7932031
 [2,] 0.9533657 1.0000000 0.5781646 0.6810888 0.9442972 0.9236925 0.8440918
 [3,] 0.6107862 0.5781646 1.0000000 0.4995241 0.4327788 0.4986136 0.4519312
 [4,] 0.7451675 0.6810888 0.4995241 1.0000000 0.6578682 0.6605604 0.5700474
 [5,] 0.9349577 0.9442972 0.4327788 0.6578682 1.0000000 0.8935938 0.8087066
 [6,] 0.9152362 0.9236925 0.4986136 0.6605604 0.8935938 1.0000000 0.7767212
 [7,] 0.7932031 0.8440918 0.4519312 0.5700474 0.8087066 0.7767212 1.0000000
 [8,] 0.7639867 0.7929525 0.4663537 0.4918769 0.7121308 0.8091594 0.7017438
 [9,] 0.6017302 0.5920356 0.2589850 0.4223605 0.6919995 0.5049835 0.4248093
[10,] 0.7201857 0.7251695 0.7115440 0.5547982 0.5965788 0.6857634 0.6385880
[11,] 0.7726417 0.7535738 0.6053737 0.5860719 0.7082858 0.6587315 0.6603982
[12,] 0.7791901 0.7842220 0.6017955 0.5399221 0.6993334 0.7231629 0.6853410
[13,] 0.6274156 0.6322589 0.3258486 0.6187917 0.6512090 0.5898520 0.5985444
           [,8]      [,9]     [,10]     [,11]     [,12]     [,13]
 [1,] 0.7639867 0.6017302 0.7201857 0.7726417 0.7791901 0.6274156
 [2,] 0.7929525 0.5920356 0.7251695 0.7535738 0.7842220 0.6322589
 [3,] 0.4663537 0.2589850 0.7115440 0.6053737 0.6017955 0.3258486
 [4,] 0.4918769 0.4223605 0.5547982 0.5860719 0.5399221 0.6187917
 [5,] 0.7121308 0.6919995 0.5965788 0.7082858 0.6993334 0.6512090
 [6,] 0.8091594 0.5049835 0.6857634 0.6587315 0.7231629 0.5898520
 [7,] 0.7017438 0.4248093 0.6385880 0.6603982 0.6853410 0.5985444
 [8,] 1.0000000 0.3167181 0.6399410 0.6286411 0.6846571 0.4446369
 [9,] 0.3167181 1.0000000 0.3025148 0.5669557 0.4093915 0.5212708
[10,] 0.6399410 0.3025148 1.0000000 0.6320091 0.6838172 0.4872347
[11,] 0.6286411 0.5669557 0.6320091 1.0000000 0.6371455 0.4902712
[12,] 0.6846571 0.4093915 0.6838172 0.6371455 1.0000000 0.4717779
[13,] 0.4446369 0.5212708 0.4872347 0.4902712 0.4717779 1.0000000
> mc_cormatrix
           [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]
 [1,] 1.0000000 0.9533657 0.6107862 0.7451675 0.9349577 0.9152362 0.7932031
 [2,] 0.9533657 1.0000000 0.5781646 0.6810888 0.9442972 0.9236925 0.8440918
 [3,] 0.6107862 0.5781646 1.0000000 0.4995241 0.4327788 0.4986136 0.4519312
 [4,] 0.7451675 0.6810888 0.4995241 1.0000000 0.6578682 0.6605604 0.5700474
 [5,] 0.9349577 0.9442972 0.4327788 0.6578682 1.0000000 0.8935938 0.8087066
 [6,] 0.9152362 0.9236925 0.4986136 0.6605604 0.8935938 1.0000000 0.7767212
 [7,] 0.7932031 0.8440918 0.4519312 0.5700474 0.8087066 0.7767212 1.0000000
 [8,] 0.7639867 0.7929525 0.4663537 0.4918769 0.7121308 0.8091594 0.7017438
 [9,] 0.6017302 0.5920356 0.2589850 0.4223605 0.6919995 0.5049835 0.4248093
[10,] 0.7201857 0.7251695 0.7115440 0.5547982 0.5965788 0.6857634 0.6385880
[11,] 0.7726417 0.7535738 0.6053737 0.5860719 0.7082858 0.6587315 0.6603982
[12,] 0.7791901 0.7842220 0.6017955 0.5399221 0.6993334 0.7231629 0.6853410
[13,] 0.6274156 0.6322589 0.3258486 0.6187917 0.6512090 0.5898520 0.5985444
           [,8]      [,9]     [,10]     [,11]     [,12]     [,13]
 [1,] 0.7639867 0.6017302 0.7201857 0.7726417 0.7791901 0.6274156
 [2,] 0.7929525 0.5920356 0.7251695 0.7535738 0.7842220 0.6322589
 [3,] 0.4663537 0.2589850 0.7115440 0.6053737 0.6017955 0.3258486
 [4,] 0.4918769 0.4223605 0.5547982 0.5860719 0.5399221 0.6187917
 [5,] 0.7121308 0.6919995 0.5965788 0.7082858 0.6993334 0.6512090
 [6,] 0.8091594 0.5049835 0.6857634 0.6587315 0.7231629 0.5898520
 [7,] 0.7017438 0.4248093 0.6385880 0.6603982 0.6853410 0.5985444
 [8,] 1.0000000 0.3167181 0.6399410 0.6286411 0.6846571 0.4446369
 [9,] 0.3167181 1.0000000 0.3025148 0.5669557 0.4093915 0.5212708
[10,] 0.6399410 0.3025148 1.0000000 0.6320091 0.6838172 0.4872347
[11,] 0.6286411 0.5669557 0.6320091 1.0000000 0.6371455 0.4902712
[12,] 0.6846571 0.4093915 0.6838172 0.6371455 1.0000000 0.4717779
[13,] 0.4446369 0.5212708 0.4872347 0.4902712 0.4717779 1.0000000
```

Hugo Joncour - 40139130
Zié Coulibaly - 40046155

To end this model, here is the confidence interval:

```
> confint(model1, conf.level=0.95)
                      2.5 %           97.5 %
(Intercept) 1042864737.9851156473 1258787500.495879173
bnb_vt             -1.2900646194        -0.382429840
bcc_vt             -0.1486690514        -0.052081331
btc_vt             -0.0085543808         0.020062917
dash_vt             0.3851128695         0.771399812
eos_vt             -0.2667918811        -0.133797901
eth_vt              0.1076149649         0.180991580
ltc_vt             -0.1501290317        -0.015751448
neo_vt             -1.1906576983        -0.686272770
trx_vt              0.3536651540         0.637977677
xlm_vt             -0.8895946611        -0.225872715
xrp_vt             -0.0170837374         0.068286852
xtz_vt              6.7017519354        11.830805707
bnb_mc              0.1481748025         0.283700416
bcc_mc              0.0365276637         0.071436542
btc_mc              0.0073153562         0.010208521
dash_mc            -0.8052558289        -0.612527420
eos_mc              0.0167831779         0.076910527
eth_mc             -0.0227386040        -0.009360936
ltc_mc              0.0330704488         0.095248918
neo_mc              0.1595784658         0.291553951
trx_mc             -0.1979593389        -0.072460375
xlm_mc             -0.0080950594         0.091602004
xrp_mc             -0.0002759275         0.013987833
xtz_mc             -0.0063367267         0.241706541
```

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

Hugo Joncour - 40139130
Zié Coulibaly - 40046155

4. Annex

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

```
public static void scrapeHistoricalCurrency(Currency currency){
    try {
        int iterator = 1;
        final Document coinMarketCapCurrencies = Jsoup.connect(currency.getHistoricsURL()).get();
        for(Element row: coinMarketCapCurrencies.select(".cmc-tab-historical-data tr")){
            if(iterator > 4) {
                String date = row.select(".cmc-table__cell--sticky.cmc-table__cell").text();
                String priceOpen = row.select(".cmc-table__cell:nth-of-type(2)").text();
                String priceClose = row.select(".cmc-table__cell:nth-of-type(3)").text();
                String dayHigh = row.select(".cmc-table__cell:nth-of-type(4)").text();
                String dayLow = row.select(".cmc-table__cell:nth-of-type(5)").text();
                String volumeTradedUSD = row.select(".cmc-table__cell:nth-of-type(6)").text();
                String marketCapUSD =  row.select(".cmc-table__cell:nth-of-type(7)").text();
            currency.addHistorics(currency.getName(),
                    date, priceOpen, priceClose, dayHigh, dayLow, volumeTradedUSD, marketCapUSD);
            } iterator++;
            }
        } catch(Exception exception) {exception.printStackTrace(); }
    }
}
```

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

```
package prog.output.csv;

import java.io.File;
import java.io.PrintWriter;
import prog.algo.types.*;

public class Printer {
    public static void createFullCSV(Currency currency){
        try {
                PrintWriter pw = new PrintWriter(new File("D:\\Desktop\\Projets\\TestDefaultOutput\\"
                        +currency.getRank()+"_"+currency.getName()+".csv"));
                StringBuilder sb = new StringBuilder();
                sb.append("Date");
                sb.append(";");
                sb.append("Opens at:");
                sb.append(";");
                sb.append("Closes at:");
                sb.append(";");
                sb.append("Highest at:");
                sb.append(";");
                sb.append("Lowest at:");
                sb.append(";");
                sb.append("Volume Traded:");
                sb.append(";");
                sb.append("Market Cap:");
                sb.append("\r\n");
                pw.write(sb.toString());
                for(int i = 0; i < currency.getHistory().size(); i++) {
                    StringBuilder line = new StringBuilder();
                    line.append(currency.getHistory().get(i).getDay()+"/"+
                            currency.getHistory().get(i).getMonth()+"/"+
                            currency.getHistory().get(i).getYear()+";");
                    line.append(currency.getHistory().get(i).getPriceOpen()+";");
                    line.append(currency.getHistory().get(i).getPriceClose()+";");
                    line.append(currency.getHistory().get(i).getDayHigh()+";");
                    line.append(currency.getHistory().get(i).getDayLow()+";");
                    line.append(currency.getHistory().get(i).getVolumeTraded()+";");
                    line.append(currency.getHistory().get(i).getMarketCap()+";\n");
                    pw.write(line.toString());
        }
                pw.close();
        } catch (Exception e) {
            e.printStackTrace();
        }
    }
}
```

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.

Hugo Joncour - 40139130
Zié Coulibaly - 40046155

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

```
X1_Bitcoin <- read_csv("10 first/1_Bitcoin.csv")
X2_Ethereum <- read_csv("10 first/2_Ethereum.csv")
X3_XRP <- read_csv("10 first/3_XRP.csv")
X4_Bitcoin_Cash <- read_csv("Data/4_Bitcoin_Cash.csv")
X5_Tether <- read_csv("10 first/5_Tether.csv")
X6_BitcoinSV <- read_csv("10 first/6_Bitcoin SV.csv")
X7_Litecoin <- read_csv("10 first/7_Litecoin.csv")
X8_EOS <- read_csv("10 first/8_EOS.csv")
X9_BinanceCoin <- read_csv("10 first/9_Binance Coin.csv")
X10_Tezos <- read_csv("10 first/10_Tezos.csv")
X11_Chainlink <- read_csv("10 first/11_Chainlink.csv")
X12_Cardano <- read_csv("Data/12_Cardano.csv")
X13_Monero <- read_csv("Data/13_Monero.csv")
X14_Stellar <- read_csv("Data/14_Stellar.csv")
X15_TRON <- read_csv("Data/15_TRON.csv")
X19_Neo <- read_csv("Data/19_Neo.csv")
X20_Dash <- read_csv("Data/20_Dash.csv")

btc_mc <- X1_Bitcoin$MarketCap[1:857]
eth_mc <- X2_Ethereum$MarketCap[1:857]
xrp_mc <- X3_XRP$MarketCap[1:857]
bcc_mc <- X4_Bitcoin_Cash$MarketCap[1:857]
usdt_mc <- X5_Tether$MarketCap[1:857]
bsv_mc <- X6_BitcoinSV$MarketCap[1:857]
ltc_mc <- X7_Litecoin$MarketCap[1:857]
eos_mc <- X8_EOS$MarketCap[1:857]
bcb_mc <- X9_BinanceCoin$MarketCap[1:857]
xtz_mc <- X10_Tezos$MarketCap[1:857]
xlm_mc <- X14_Stellar$MarketCap[1:857]
trx_mc <- X15_TRON$MarketCap[1:857]
neo_mc <- X19_Neo$MarketCap[1:857]
dash_mc <- X20_Dash$MarketCap[1:857]

market_cap_comatrix <- data.frame(btc_mc, eth_mc, xrp_mc, bcc_mc, usdt_mc, ltc_mc, eos_mc, bcb_mc, xlm_mc, trx_mc, neo_mc, dash_mc)
View(market_cap_comatrix)
mc_covmatrix = matrix(c(cov(market_cap_comatrix)), nrow=12, ncol=12)

mc_covmatrix
btc_vt <- X1_Bitcoin$VolumeTraded[1:857]
eth_vt <- X2_Ethereum$VolumeTraded[1:857]
xrp_vt <- X3_XRP$VolumeTraded[1:857]
bcc_vt <- X4_Bitcoin_Cash$VolumeTraded[1:857]
usdt_vt <- X5_Tether$VolumeTraded[1:857]
ltc_vt <- X7_Litecoin$VolumeTraded[1:857]
eos_vt <- X8_EOS$VolumeTraded[1:857]
bcb_vt <- X9_BinanceCoin$VolumeTraded[1:857]
xtz_vt <- X10_Tezos$VolumeTraded[1:857]
xlm_vt <- X14_Stellar$VolumeTraded[1:857]
trx_vt <- X15_TRON$VolumeTraded[1:857]
neo_vt <- X19_Neo$VolumeTraded[1:857]
dash_vt <- X20_Dash$VolumeTraded[1:857]

volume_traded_comatrix <- data.frame(btc_vt, eth_vt, xrp_vt, bcc_vt, usdt_vt, ltc_vt, eos_vt, bcb_vt, xlm_vt, trx_vt, neo_vt, dash_vt)
View(volume_traded_comatrix)
vt_covmatrix = matrix(c(cov(volume_traded_comatrix)), nrow=12, ncol=12)
vt_covmatrix
```

Code for the multilinear regression:

```
#rm(list=ls(all=TRUE))#to remove all the items

colnames(markc) <- c("BinanceCoin MC", "BitcoinCash MC", "Bitcoin MC", "Dash MC", "EOS MC", "ETH MC", "Litecoin MC", "NEO MC",
                     "TRON MC", "Stellar MC", "Ripple MC", "Tezos MC")
colnames(volt) <- c("BinanceCoin VT", "BitcoinCash VT", "Bitcoin VT", "Dash VT", "EOS VT", "ETH VT", "Litecoin VT", "NEO VT",
                    "TRON VT", "Stellar VT", "Ripple VT", "Tezos VT")

volt <- cbind(bnb_vt, bcc_vt,btc_vt, dash_vt, eos_vt, eth_vt, ltc_vt, neo_vt,
              trx_vt, xlm_vt, xrp_vt, xtz_vt)
markc <- cbind(bnb_mc, bcc_mc,btc_mc, dash_mc, eos_mc, eth_mc, ltc_mc, neo_mc,
               trx_mc, xlm_mc, xrp_mc, xtz_mc)

names <- cbind("BinanceCoin", "BitcoinCash", "Bitcoin", "Dash", "EOS", "ETH", "Litecoin", "NEO",
               "TRON", "Stellar", "Ripple", "Tezos")


predictor_data <- cbind(volt, markc)

model1 <- lm(X5_Tether$MarketCap[1:852] ~ bnb_vt+ bcc_vt+btc_vt+ dash_vt+ eos_vt+ eth_vt+ ltc_vt+ neo_vt+
               trx_vt+ xlm_vt+ xrp_vt+ xtz_vt+ bnb_mc+ bcc_mc+btc_mc+ dash_mc+ eos_mc+ eth_mc+ ltc_mc+ neo_mc+
               trx_mc+ xlm_mc+ xrp_mc+ xtz_mc)
summary(model1)
```

Hugo Joncour - 40139130
Zié Coulibaly - 40046155

Code for the plots:

```
#Graph Creation (Normal & Log)


#Scientific to numerical notation
options(scipen = 999)

data <- Bitcoin #Coin file
data$Date <- dmy(data$Date)

#Normal portion
meltdata <- melt(data,id="Date")
ggplot(meltdata,aes(x=Date,y=value,colour=variable,group=variable)) +
  geom_line()

#Log portion
lg <- ggplot(meltdata,aes(x=Date,y=value,colour=variable,group=variable)) +
  geom_line()

lg + scale_y_continuous(trans = log10_trans())
```