An Exploration on Crypto Currencies

Emine Doganay, Hale Nur Kazacesme, Hasan Palaz

November 18, 2017

The notion of cryptocurrency has entered our lives with the introduction of Bitcoin on Jan 2009. Crypto currencies are decentralized digital currencies, meaning the payment system works without a central repository or administrator. The transactions between users, creation of additional units and verification of transfers are secured by cryptography and logged into the transaction database; blockchain. In most cryptocurrencies, transactions are verified through a process called mining that requires computer processing power to find blocks of transactions and miners are rewarded with newly created crypto coins for their finds. Coins are stored in cryptocurrency wallets, which are essentially a collection of cryptographic keys. Yes, it is hard to digest the mechanics behind cryptocurrencies.

With the uncontrollable climb of Bitcoin and all these buzzwords flying around, cryptocurrencies became a hype. Part of the hype is created by the mystery caused by the enigmatic nature of cryptocurrencies and the obscurity of the workings of blockchain. The suspicion surrounding the future of cryptocurrencies were not only founded on the fear of the unknown but also on the fact that they can be lost forever from their markets due malware or data loss or even an accident as in a recent incident with an Etherium wallet that caused approximately $280m worth of Ether to be frozen.

All this said, our aim is to understand and relate cryptocurrencies to shed some light onto this conundrum. We will try to understand the correlation between crytocurrencies, correlation with gold and how google trends are related to Bitcoin price.

We read the data directly from coinmarketcap.com using package rvest. The data looks like:

```
##   X.                                    Name Symbol       Market.Cap     Price
## 1  1      BTC\n                      Bitcoin    BTC $136,994,306,666  $8206.01
## 2  2     ETH\n                      Ethereum    ETH  $35,191,906,350   $367.05
## 3  3 BCH\n                      Bitcoin Cash    BCH  $20,028,501,521  $1191.08
## 4  4       XRP\n                     Ripple    XRP   $9,105,109,962 $0.235744
##                                                                              Circulating.
Supply   Volume..24h.  X..1h X..24h  X..7d
## 1                                                                                    16,6
94,387 $3,319,270,000  0.50%  1.84% 18.27%
## 2                                                                                    95,8
78,997   $665,355,000  0.54%  1.75%  8.94%
## 3                                                                                    16,8
15,413   $512,279,000 -0.01%  0.95% -7.55%
## 4 38,622,870,411\n                       \n                        \n
     *   $105,769,000  0.19% -0.18% 12.15%
```

We can see that the data contains a few unwanted characters such as new line, , $,*, spaces, %. In order to conduct analysis, we need to remove these characters.

An alternative way to retrive data is to use the JSON API at api.coinmarketcap.com which provides clean data. The data from the JSON API looks like:

```
##             id         name symbol   price_usd  price_btc 24h_volume_usd market_cap_usd avail
able_supply total_supply max_supply percent_change_1h
## 1      bitcoin      Bitcoin    BTC 8199.620000 1.00000000     3312340000   136887629533
   16694387     16694387    2.1e+07              0.41
## 2     ethereum     Ethereum    ETH  366.740000 0.04474300      664158000    35162663256
   95878997     95878997         NA              0.45
## 3 bitcoin-cash Bitcoin Cash    BCH 1190.460000 0.14523900      510238000    20018075965
   16815413     16815413    2.1e+07             -0.04
## 4       ripple       Ripple    XRP    0.235734 0.00002876      105844000     9104723733
38622870411  99993173757    1.0e+11              0.19
##   percent_change_24h percent_change_7d last_updated
## 1               1.74             18.13   1511336958
## 2               1.66              8.85   1511336969
## 3               0.91             -7.59   1511336969
## 4              -0.19             12.13   1511336941
```

which is much cleaner aside from NA's. But to demonstrate what we have learned in this class we chose to work with the data collected with rvest package and cleaned it.

```
# Cleaning the data
# Before cleaning, "\n" in the name strings should be replaced with "-" in df_cryptocurrencies$N
ame column
df_cryptocurrencies$Name <- lapply(df_cryptocurrencies$Name, gsub, pattern = "\n", replacement =
 "-")

# In order to conduct analysis, we need to remove the unwanted characters such as \n, $,*, space
s, %.
df_cryptocurrencies[] <- lapply(df_cryptocurrencies, gsub, pattern = "\\\n|\\s|[%*$,?]", replace
ment = "")

# remove the first column X. (coins' ranking of coinmarketcap web site)
df_cryptocurrencies$X. <- NULL

# Update the column names short, lowercase and meaningful
names(df_cryptocurrencies) <- c("name", "symbol", "marketcap", "price", "supply", "volume", "cha
nge_1h", "change_24h", "change_7d")

# Let's see how our data frame, df_cryptocurrencies looks after cleaning process
head(df_cryptocurrencies)
```

```
##              name symbol    marketcap    price      supply     volume change_1h change_24h ch
ange_7d
## 1     BTC-Bitcoin    BTC 136994306666  8206.01    16694387 3319270000      0.50       1.84
  18.27
## 2    ETH-Ethereum    ETH  35191906350   367.05    95878997  665355000      0.54       1.75
   8.94
## 3 BCH-BitcoinCash    BCH  20028501521  1191.08    16815413  512279000     -0.01       0.95
  -7.55
## 4      XRP-Ripple    XRP   9105109962 0.235744 38622870411  105769000      0.19      -0.18
  12.15
## 5       DASH-Dash   DASH   4072563397   528.68     7703297  195967000      2.72      12.12
  22.00
## 6    LTC-Litecoin    LTC   3811983352    70.67    53938708  147851000      0.28       0.26
  11.01
```

For mathematical analysis variable types should be numeric.

```
# Check type of marketcap
typeof(df_cryptocurrencies$marketcap)
```

```
## [1] "character"
```

However some are not. So we transform the type of variables except name and symbol to numeric and assign into a new data frame. And after transformation data type of marketcap becomes:

```
## [1] "double"
```

Let's see how our data frame looks like now:

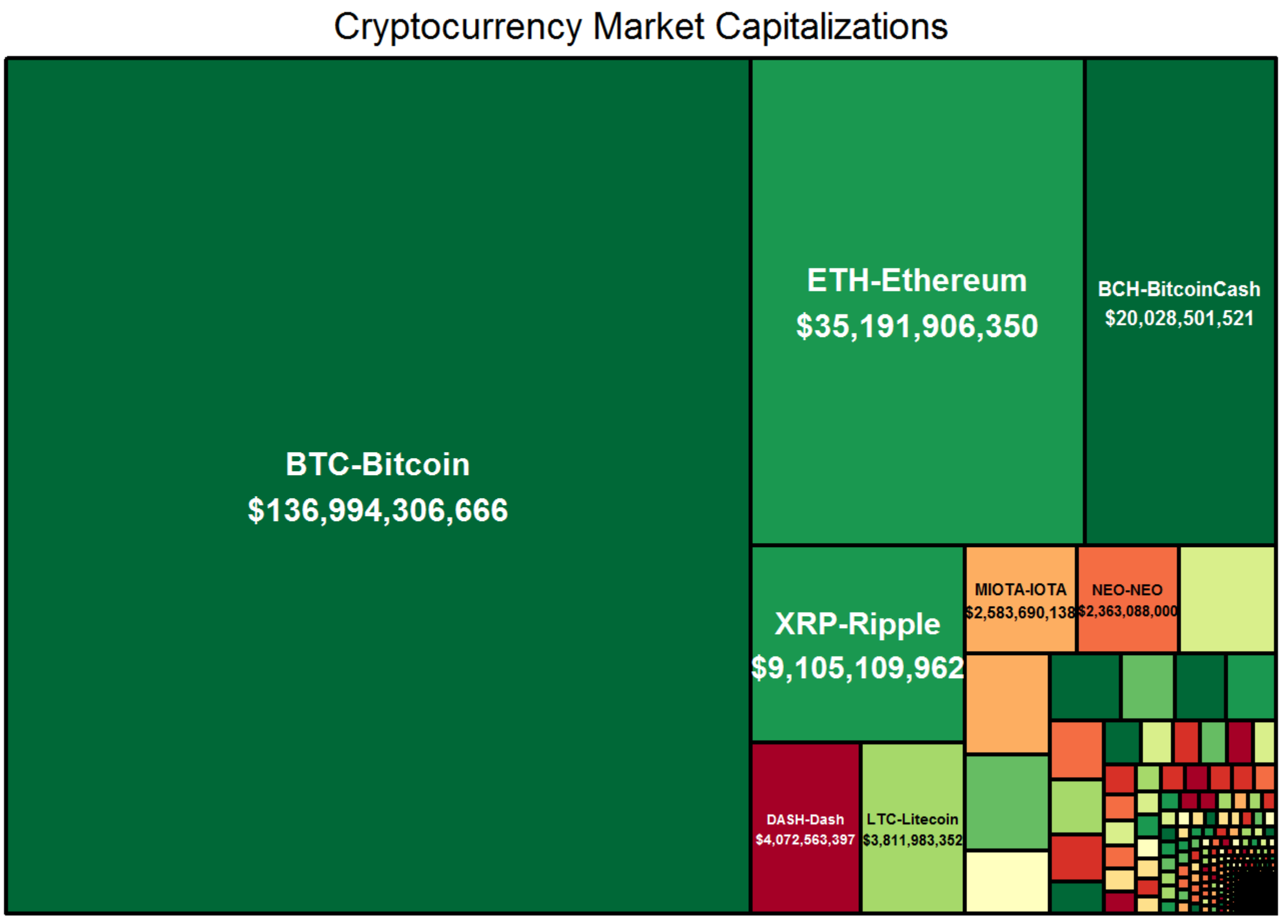
```
##              name symbol    marketcap       price      supply     volume change_1h change_24h
change_7d
## 1     BTC-Bitcoin    BTC 136994306666 8206.010000    16694387 3319270000      0.50       1.84
     18.27
## 2    ETH-Ethereum    ETH  35191906350  367.050000    95878997  665355000      0.54       1.75
      8.94
## 3 BCH-BitcoinCash    BCH  20028501521 1191.080000    16815413  512279000     -0.01       0.95
     -7.55
## 4      XRP-Ripple    XRP   9105109962    0.235744 38622870411  105769000      0.19      -0.18
     12.15
```

```
## [1] 914   9
```

and how many coins are there in the market?

```
## [1] 914
```

When you consider there are only 180 recognized currencies in the world this is a huge number. However not all of them are big players. Let's draw the market share of all crypto currencies as a treemap to get a better understanding:



Normal currencies are dependable because they have a relatively stable standing against each other and their market only gets effected from big political or economical events. However with a decentralized non-physical currency it all works differently. With no mechanism to act on a financial policy to regulate sudden price changes, price of the crypto-currencies are literally defined in the market.
Let's see how big these changes can be by exploring the evolution of the market in 1 hour, 24 hours and 7 days for all crypto currencies.
Summary of statistics and standard deviation of market change in 1 hour are as follows:

```
summary(df_allcryptocurrencies$change_1h)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's 
## -39.6100   0.0850   0.6500   0.6711   0.7000  72.5500       71
```

```
sd(df_allcryptocurrencies$change_1h, na.rm=TRUE)
```

```
## [1] 6.567036
```

Summary of statistics and standard deviation of market change in 24 hours are as follows:

```
summary(df_allcryptocurrencies$change_24h)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's 
##  -95.920   -5.165    0.960    9.357    5.317 4967.040       68
```

```
sd(df_allcryptocurrencies$change_24h, na.rm=TRUE)
```

```
## [1] 175.0159
```

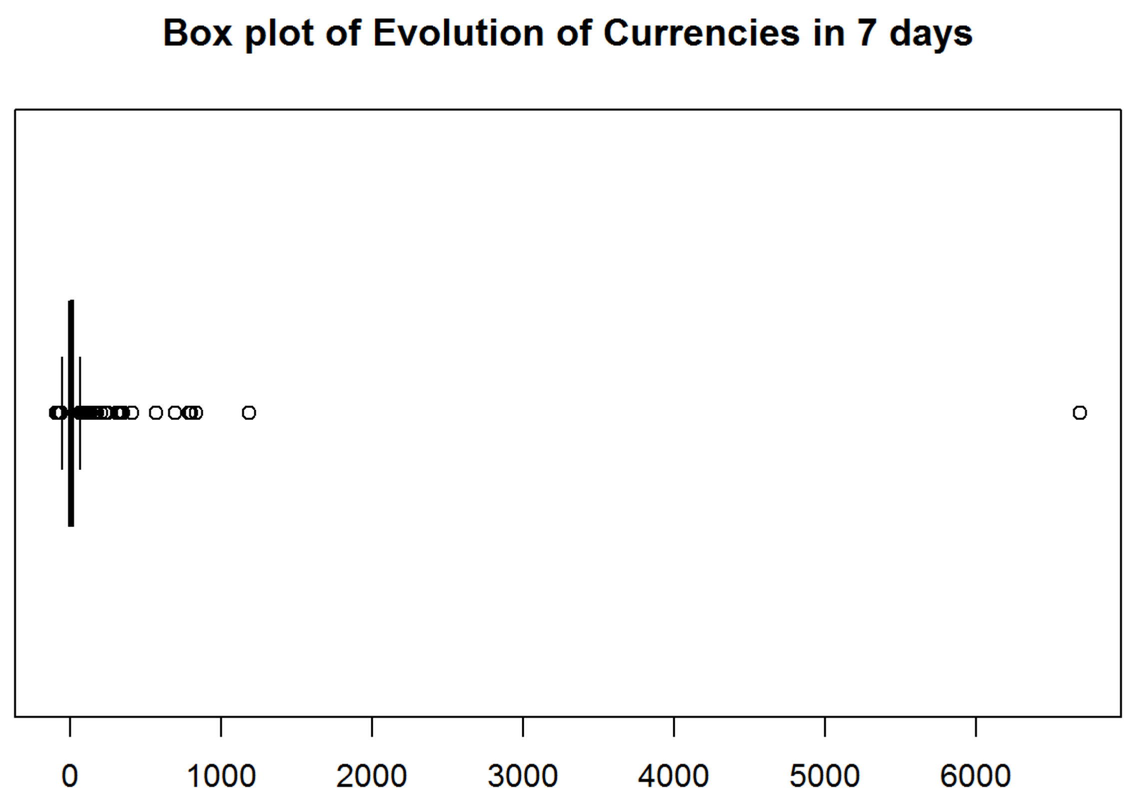
Summary of statistics and standard deviation of market change in 7 days are as follows:

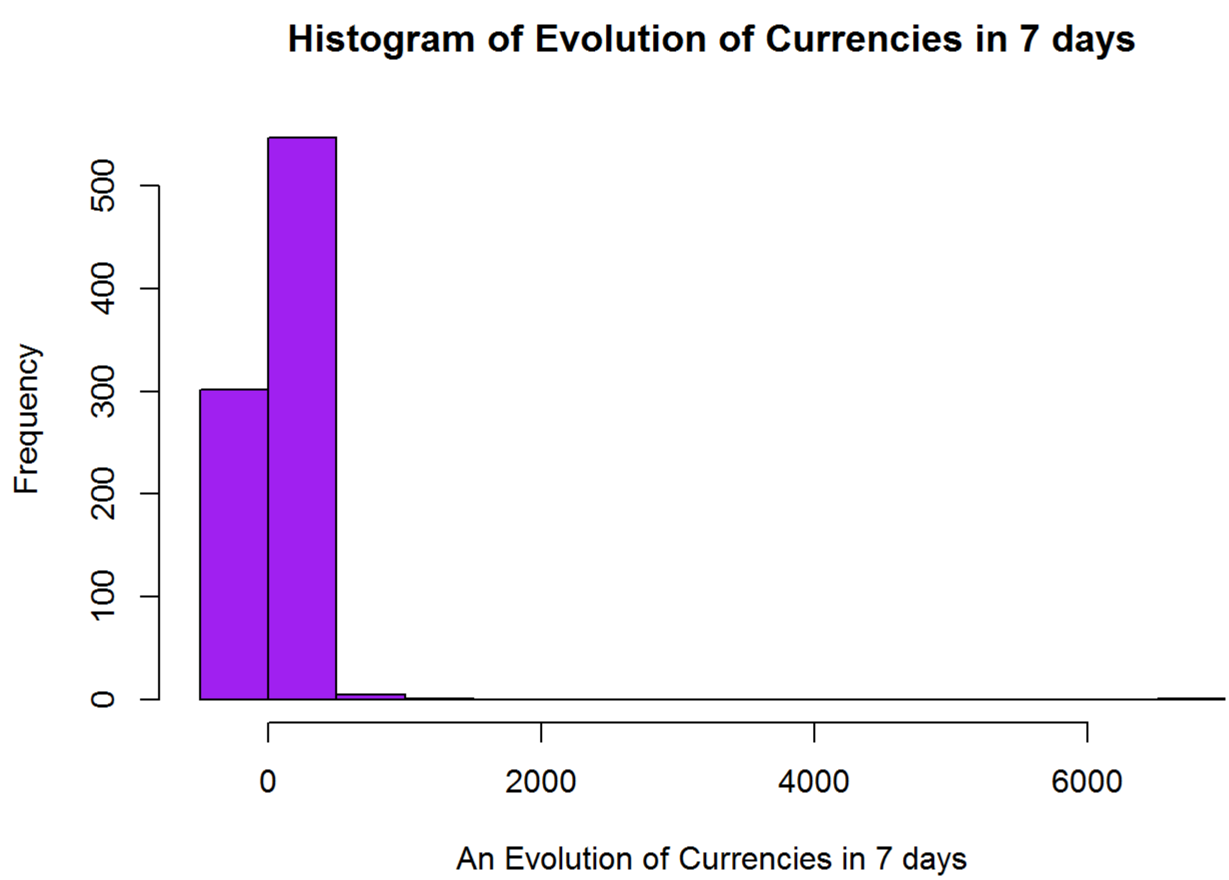
```
summary(df_allcryptocurrencies$change_7d)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's 
##  -92.030   -7.532    8.900   26.812   23.345 6688.660       60
```

```
sd(df_allcryptocurrencies$change_7d, na.rm=TRUE)
```

```
## [1] 242.7385
```


Box plot of Evolution of Currencies in 7 days
0
1000
2000
3000
4000
5000
6000


Histogram of Evolution of Currencies in 7 days
Frequency
0
100
200
300
400
500
0
2000
4000
6000
An Evolution of Currencies in 7 days

However these price changes don't mean much without the context. Surely an $100 increase of Bitcoin (which is valued about $8000 at the time of this writing) is not the same an $100 increase of Ethereum (which is valued about $300 at the time of this writing). Let's find all the cryptocurrencies whose value has more than doubled in 7 days:

```
##  [1] NXT-Nxt              FTC-Feathercoin      KRB-Karbowanec       ONX-Onix             SUPE
R-SuperCoin      BCO-BridgeCoin
##  [7] GRE-Greencoin        UNIT-UniversalCur... XSH-SHIELD           ORB-Orbitcoin        XGR-
GoldReserve      BCF-BitcoinFast
## [13] BXT-BitTokens        BTCR-Bitcurrency     4CHN-ChanCoin        888-OctoCoin         GB-G
oldBlocks        USDE-USDe
## [19] VEC2-VectorAI        XRC-Rawcoin
## 914 Levels: -Money 020-020LondonCoin 1337-1337 2GIVE-2GIVE 42-42-coin 4CHN-ChanCoin 611-SixEl
even 808-808Coin 888-OctoCoin 8BIT-8Bit ... ZZC-ZoZoCoin
```

and let's find all the cryptocurrencies which lost more than 25% of its value in 7 days:

```
##  [1] B3-B3Coin         BDL-Bitdeal       EQT-EquiTrader    DIME-Dimecoin     XGOX-XGOX
  FIMK-FIMKrypto    ZEIT-Zeitcoin     ARC-ArcticCoin
##  [9] 1337-1337         RNS-Renos         INSN-InsaneCoin   HNC-Helleniccoin  LDOGE-LiteDoge
  MAO-MaoZedong     XIOS-Xios         LUX-LUXCoin
## [17] CRM-Cream         BLAS-BlakeStar    SPRTS-Sprouts     MNC-Mincoin       ATOM-AtomicCoin
  EMB-EmberCoin     808-808Coin       EL-Elcoin
## [25] EUC-Eurocoin      TOR-Torcoin       ITZ-Interzone     XCS-CybCSec       CTIC2-Coimatic2.
0 FAL-Falcoin
## 914 Levels: -Money 020-020LondonCoin 1337-1337 2GIVE-2GIVE 42-42-coin 4CHN-ChanCoin 611-SixEl
even 808-808Coin 888-OctoCoin 8BIT-8Bit ... ZZC-ZoZoCoin
```

By the tree chart above we know that Bitcoin is the dominant player in the market.
The total market capitalization currently is:

```
## [1] "$233,804,325,168"
```

and Bitcoin's share is:

```
## [1] "0.5859357%"
```

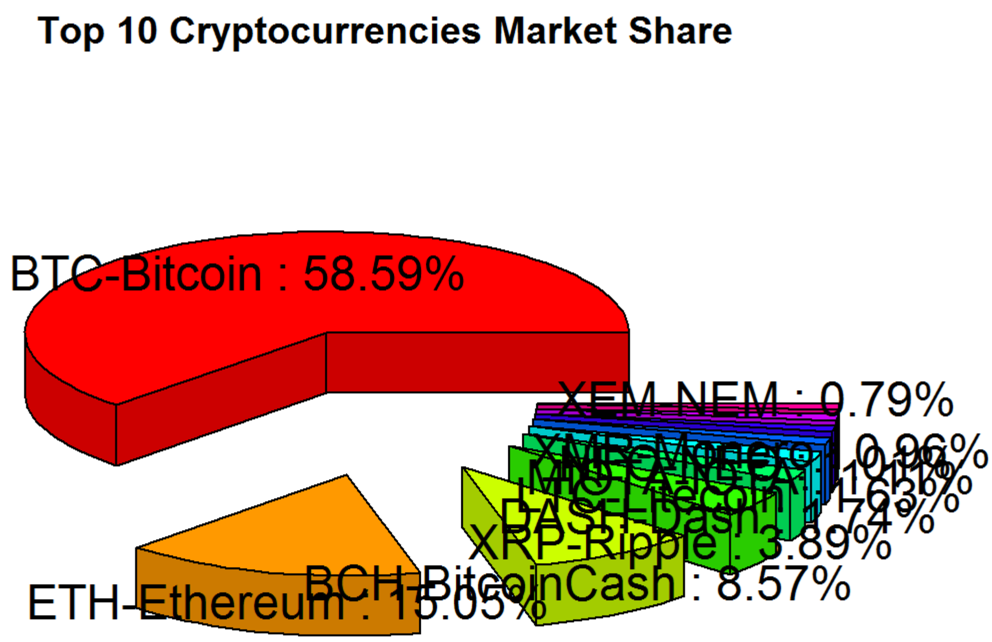
Even though Bitcoin is sharing the market with more and more crypto currencies each day, it remains the biggest player.

Let's focus on top 10 currrencies by market cap.

```
##               name symbol    marketcap       price      supply     volume change_1h change_24
h change_7d             formatted_market_cap
## 1      BTC-Bitcoin    BTC 136994306666 8206.010000    16694387 3319270000      0.50       1.8
4     18.27    BTC-Bitcoin\n$136,994,306,666
## 2     ETH-Ethereum    ETH  35191906350  367.050000    95878997  665355000      0.54       1.7
5      8.94    ETH-Ethereum\n$35,191,906,350
## 3  BCH-BitcoinCash    BCH  20028501521 1191.080000    16815413  512279000     -0.01       0.9
5     -7.55 BCH-BitcoinCash\n$20,028,501,521
## 4       XRP-Ripple    XRP   9105109962    0.235744 38622870411  105769000      0.19      -0.1
8     12.15       XRP-Ripple\n$9,105,109,962
## 5        DASH-Dash   DASH   4072563397  528.680000     7703297  195967000      2.72      12.1
2     22.00        DASH-Dash\n$4,072,563,397
## 6     LTC-Litecoin    LTC   3811983352   70.670000    53938708  147851000      0.28       0.2
6     11.01     LTC-Litecoin\n$3,811,983,352
## 7       MIOTA-IOTA  MIOTA   2583690138    0.929542  2779530283   82883900      1.62       4.5
3     43.57       MIOTA-IOTA\n$2,583,690,138
## 8          NEO-NEO    NEO   2363088000   36.360000    65000000   87469400      0.71       0.6
0     21.44          NEO-NEO\n$2,363,088,000
## 9       XMR-Monero    XMR   2233389607  145.230000    15377976   82114100     -0.46       6.2
1     17.71       XMR-Monero\n$2,233,389,607
## 10         XEM-NEM    XEM   1856286000    0.206254  8999999999    7281700      0.50      -0.2
2      5.16          XEM-NEM\n$1,856,286,000
##    marketcap_percent
## 1        0.585935725
## 2        0.150518628
## 3        0.085663520
## 4        0.038943291
## 5        0.017418683
## 6        0.016304161
## 7        0.011050652
## 8        0.010107118
## 9        0.009552388
## 10       0.007939485
```

Total market share of the top 10 coins is:

```
## [1] "93.34337%"
```



Since we are focusing on the top currencies we gather historical data from CoinMarketCap web site as the top curriencies go way back than some other small currencies. HTML data is converted into a dataframe because there is no API to get the historic data of cryptocurrencies. We will study the daily evolution of top 5 cryptocurrencies:

```
##              name symbol
## 1     BTC-Bitcoin    BTC
## 2    ETH-Ethereum    ETH
## 3 BCH-BitcoinCash    BCH
## 4      XRP-Ripple    XRP
## 5       DASH-Dash   DASH
```

We import the historical data for these currencies.

Bitcoin

```
##       Date                     Open              High              Low               Close
    Volume            Market.Cap
##  Min.   :2013-04-28   Min.   :  68.5   Min.   :  74.56   Min.   :  65.53   Min.   :  68.43
Min.   :2.858e+06   Min.   :7.793e+08
##  1st Qu.:2014-06-19   1st Qu.: 265.5   1st Qu.: 270.83   1st Qu.: 260.95   1st Qu.: 265.66
1st Qu.:2.181e+07   1st Qu.:3.774e+09
##  Median :2015-08-10   Median : 454.9   Median : 460.15   Median : 446.11   Median : 454.99
Median :5.067e+07   Median :6.600e+09
##  Mean   :2015-08-10   Mean   : 875.6   Mean   : 901.99   Mean   : 852.00   Mean   : 880.30
Mean   :3.211e+08   Mean   :1.362e+10
##  3rd Qu.:2016-09-30   3rd Qu.: 758.7   3rd Qu.: 774.70   3rd Qu.: 738.92   3rd Qu.: 758.70
3rd Qu.:1.298e+08   3rd Qu.:1.074e+10
##  Max.   :2017-11-21   Max.   :8205.7   Max.   :8348.66   Max.   :7949.36   Max.   :8200.64
Max.   :8.957e+09   Max.   :1.370e+11
##
NA's   :243
##    Percentage
##  Min.   :-0.429680
##  1st Qu.:-0.018626
##  Median :-0.001956
##  Mean   :-0.003381
##  3rd Qu.: 0.011286
##  Max.   : 0.233712
##
```

Ethereum

```
##       Date                     Open                High              Low                Close
        Volume             Market.Cap
##  Min.   :2015-08-07   Min.   :  0.4316   Min.   :  0.483   Min.   :  0.4209   Min.   :  0.434
8   Min.   :1.021e+05   Min.   :3.197e+07
##  1st Qu.:2016-03-03   1st Qu.:  7.1350   1st Qu.:  7.357   1st Qu.:  6.7575   1st Qu.:  7.170
0   1st Qu.:5.850e+06   1st Qu.:5.880e+08
##  Median :2016-09-28   Median : 11.4500   Median : 11.825   Median : 11.1750   Median : 11.485
0   Median :1.671e+07   Median :9.600e+08
##  Mean   :2016-09-28   Mean   : 73.0548   Mean   : 76.140   Mean   : 70.0287   Mean   : 73.471
0   Mean   :2.187e+08   Mean   :6.806e+09
##  3rd Qu.:2017-04-25   3rd Qu.: 53.0425   3rd Qu.: 54.867   3rd Qu.: 51.5400   3rd Qu.: 53.125
0   3rd Qu.:1.691e+08   3rd Qu.:4.810e+09
##  Max.   :2017-11-21   Max.   :397.5900   Max.   :414.760   Max.   :383.4700   Max.   :401.490
0   Max.   :2.883e+09   Max.   :3.676e+10
##
                        NA's   :1
##    Percentage
##  Min.   :-0.510344
##  1st Qu.:-0.034190
##  Median : 0.001083
##  Mean   :-0.009241
##  3rd Qu.: 0.025597
##  Max.   : 0.728042
##
```

BitcoinCash

```
##       Date                  Open             High             Low             Close             
Volume            Market.Cap          Percentage      
##  Min.   :2017-07-23   Min.   : 212.2   Min.   : 223.7   Min.   : 201.0   Min.   : 213.2   Mi
n.   :8.501e+04   Min.   :3.497e+09   Min.   :-0.539691  
##  1st Qu.:2017-08-22   1st Qu.: 343.4   1st Qu.: 371.7   1st Qu.: 316.5   1st Qu.: 343.0   1st 
Qu.:1.645e+08   1st Qu.:5.664e+09   1st Qu.:-0.052810  
##  Median :2017-09-21   Median : 438.6   Median : 471.8   Median : 414.4   Median : 439.6   Med
ian :2.993e+08   Median :7.396e+09   Median : 0.004763  
##  Mean   :2017-09-21   Mean   : 517.5   Mean   : 577.2   Mean   : 476.1   Mean   : 521.9   Mea
n   :7.065e+08   Mean   :8.784e+09   Mean   :-0.016847  
##  3rd Qu.:2017-10-21   3rd Qu.: 601.1   3rd Qu.: 642.5   3rd Qu.: 569.1   3rd Qu.: 606.2   3rd 
Qu.:7.658e+08   3rd Qu.:1.027e+10   3rd Qu.: 0.048317  
##  Max.   :2017-11-21   Max.   :1381.8   Max.   :2477.7   Max.   :1215.6   Max.   :1388.9   Ma
x.   :8.371e+09   Max.   :2.321e+10   Max.   : 0.359841  
##                                                                                              
                NA's   :10
```

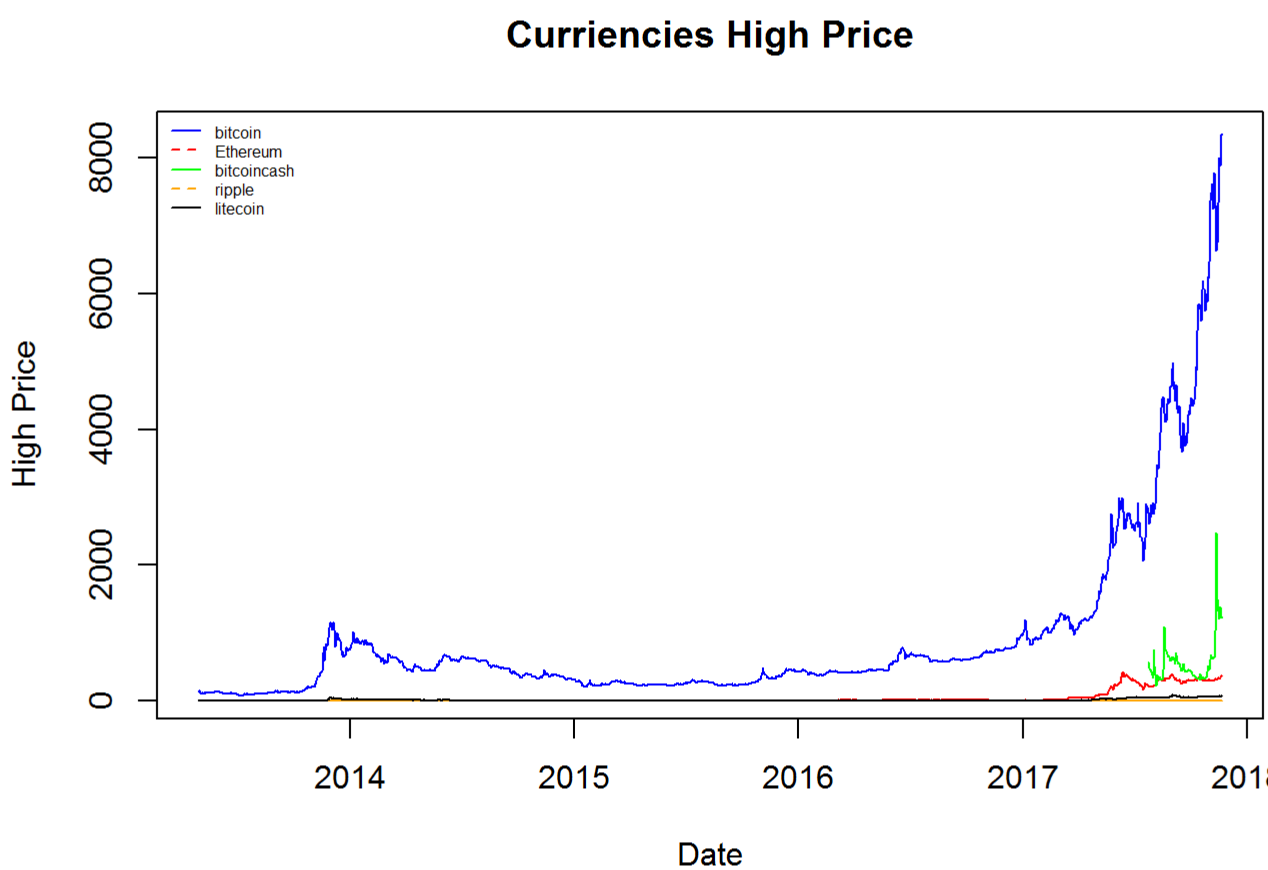
Ripple

```
##       Date                  Open               High                 Low                 Close          
         Volume            Market.Cap         
##  Min.   :2013-08-04   Min.   :0.002809   Min.   :0.003082   Min.   :0.002802   Min.   :0.0028
10   Min.   :8.316e+03   Min.   :2.196e+07  
##  1st Qu.:2014-08-31   1st Qu.:0.006023   1st Qu.:0.006126   1st Qu.:0.005933   1st Qu.:0.0060
27   1st Qu.:2.829e+05   1st Qu.:1.431e+08  
##  Median :2015-09-28   Median :0.007760   Median :0.007924   Median :0.007543   Median :0.0077
72   Median :6.692e+05   Median :2.341e+08  
##  Mean   :2015-09-28   Mean   :0.036805   Mean   :0.038760   Mean   :0.034998   Mean   :0.0369
47   Mean   :2.950e+07   Mean   :1.302e+09  
##  3rd Qu.:2016-10-24   3rd Qu.:0.014436   3rd Qu.:0.015033   3rd Qu.:0.014032   3rd Qu.:0.0144
35   3rd Qu.:2.307e+06   3rd Qu.:3.157e+08  
##  Max.   :2017-11-21   Max.   :0.394931   Max.   :0.435170   Max.   :0.332222   Max.   :0.3941
32   Max.   :1.770e+09   Max.   :1.513e+10  
##                                                                                              
     NA's   :145                            
##    Percentage      
##  Min.   :-1.793669  
##  1st Qu.:-0.020750  
##  Median : 0.002604  
##  Mean   :-0.005443  
##  3rd Qu.: 0.020879  
##  Max.   : 0.460047  
## 
```

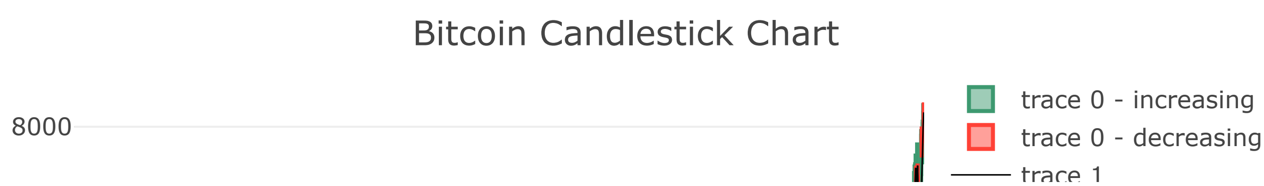
Litecoin

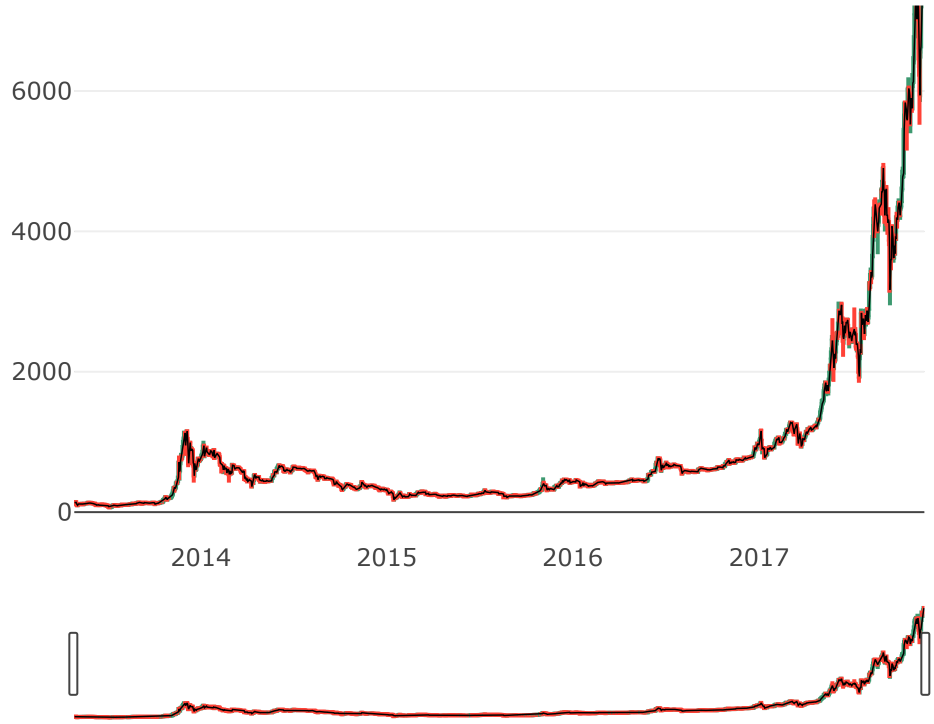
```
##       Date                     Open             High             Low              Close            Vol
ume             Market.Cap             Percentage
##  Min.   :2013-04-28   Min.   : 1.15   Min.   : 1.34   Min.   : 1.11   Min.   : 1.16   Min.
:4.817e+05   Min.   :3.794e+07   Min.   :-1.290954
##  1st Qu.:2014-06-19   1st Qu.: 3.09   1st Qu.: 3.14   1st Qu.: 3.02   1st Qu.: 3.08   1st Q
u.:1.755e+06   1st Qu.:1.205e+08   1st Qu.:-0.016491
##  Median :2015-08-10   Median : 3.86   Median : 3.92   Median : 3.81   Median : 3.86   Median
:3.162e+06   Median :1.760e+08   Median : 0.000000
##  Mean   :2015-08-10   Mean   :10.75   Mean   :11.23   Mean   :10.25   Mean   :10.79   Mean
:5.047e+07   Mean   :4.698e+08   Mean   :-0.004071
##  3rd Qu.:2016-09-30   3rd Qu.:10.30   3rd Qu.:10.55   3rd Qu.: 9.80   3rd Qu.:10.31   3rd Q
u.:9.418e+06   3rd Qu.:2.959e+08   3rd Qu.: 0.017422
##  Max.   :2017-11-21   Max.   :85.83   Max.   :92.07   Max.   :75.59   Max.   :86.04   Max.
:1.731e+09   Max.   :4.527e+09   Max.   : 0.401857
##                                                                                        NA's
:243
```

If we plot the high prices of these 5 currencies together we get.

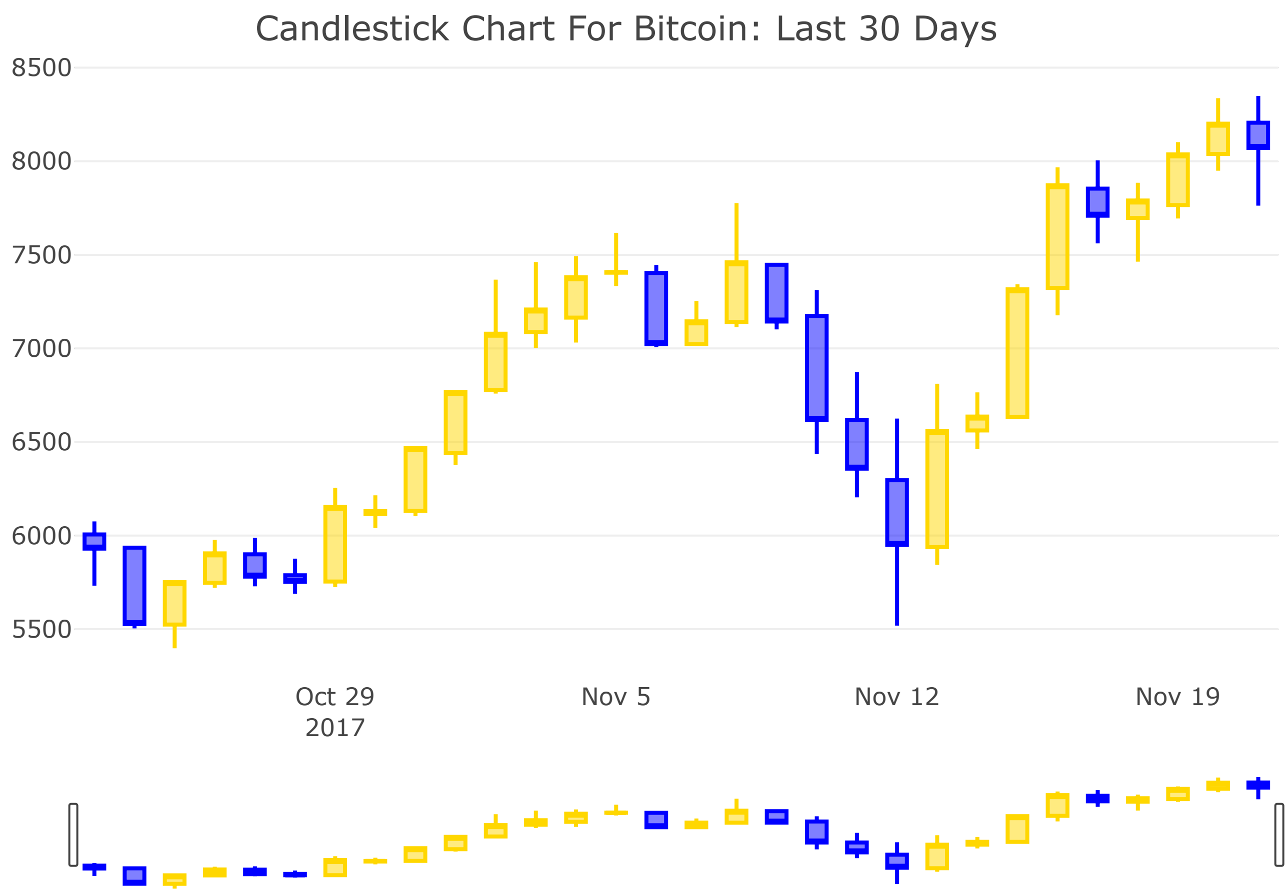


Our motivation to work with cryptocurriencies actually comes from this; look at that steep increase in a year. But when you look closely it's not a smooth increase. With a candlestick plot we can see that the price is always alternating between increasing and decreasing.





Focusing on only the last 30 days.



This means between the day we decided on this project and today we could’ve increased our money by 33% ! That is a huge amount. Let’s examine the percentage of price changes for Bitcoin to see if this is normal.

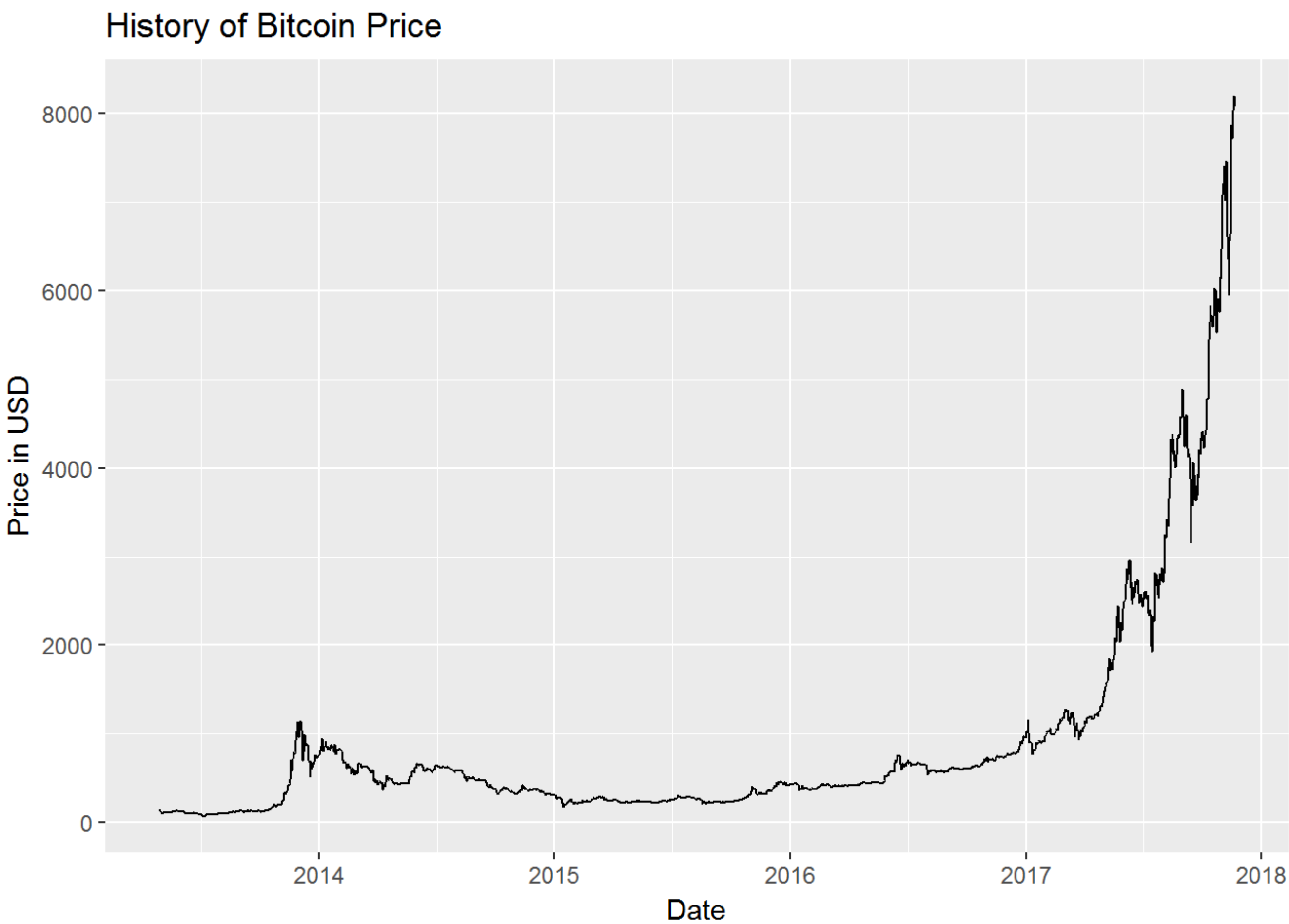
```
summary(df_bitcoin$Percentage)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## -0.429680 -0.018626 -0.001956 -0.003381  0.011286  0.233712
```

```
sd(df_bitcoin$Percentage)
```

```
## [1] 0.04312892
```

Bitcoin is highly volatile; it has gained up to 23% or lost up to 43% in only one day.



Let's examine the percentage of price changes for Ethereum to see if a similer phenomenon occurs.

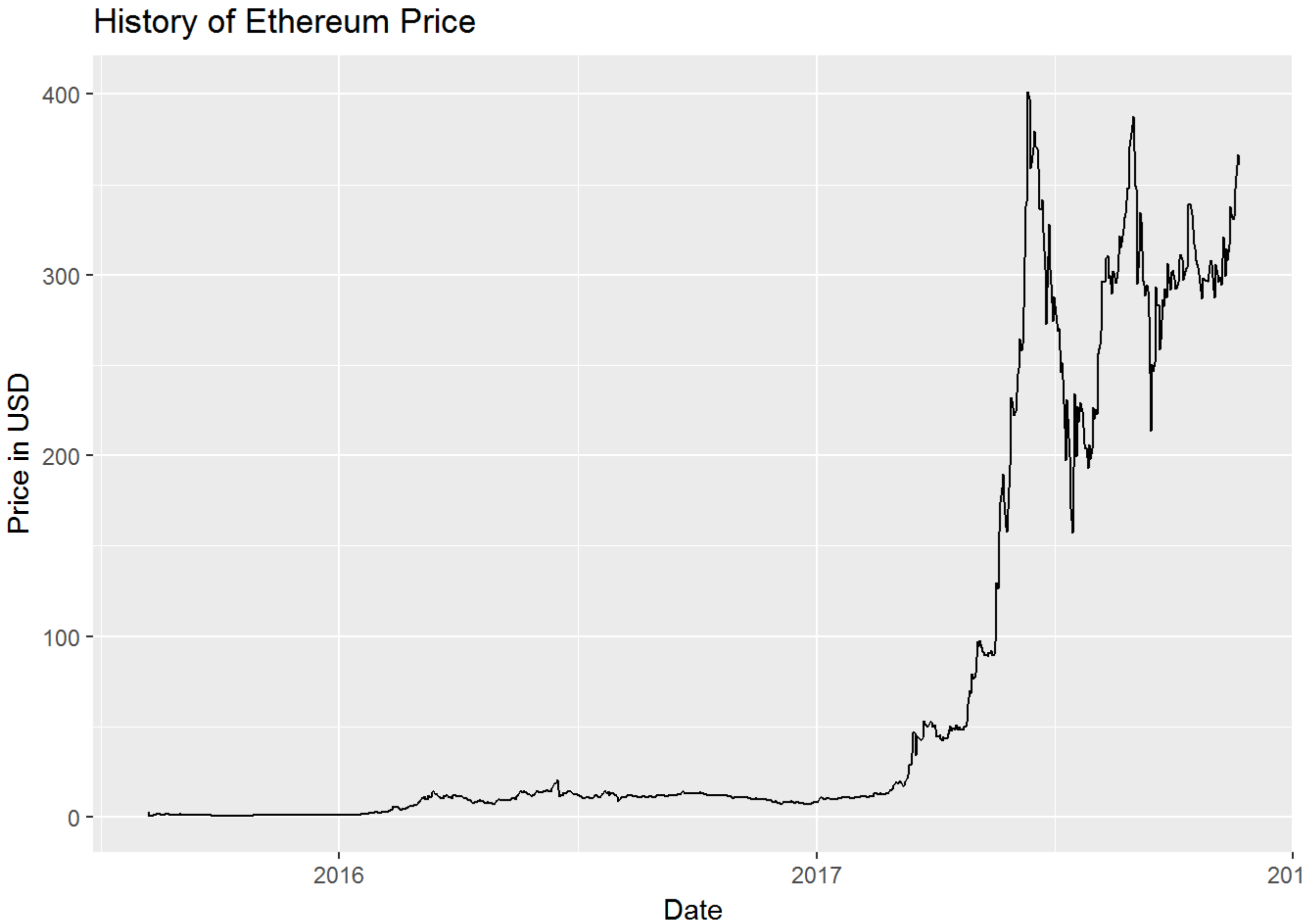
```
summary(df_ethereum$Percentage)
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
## -0.510344 -0.034190  0.001083 -0.009241  0.025597  0.728042
```

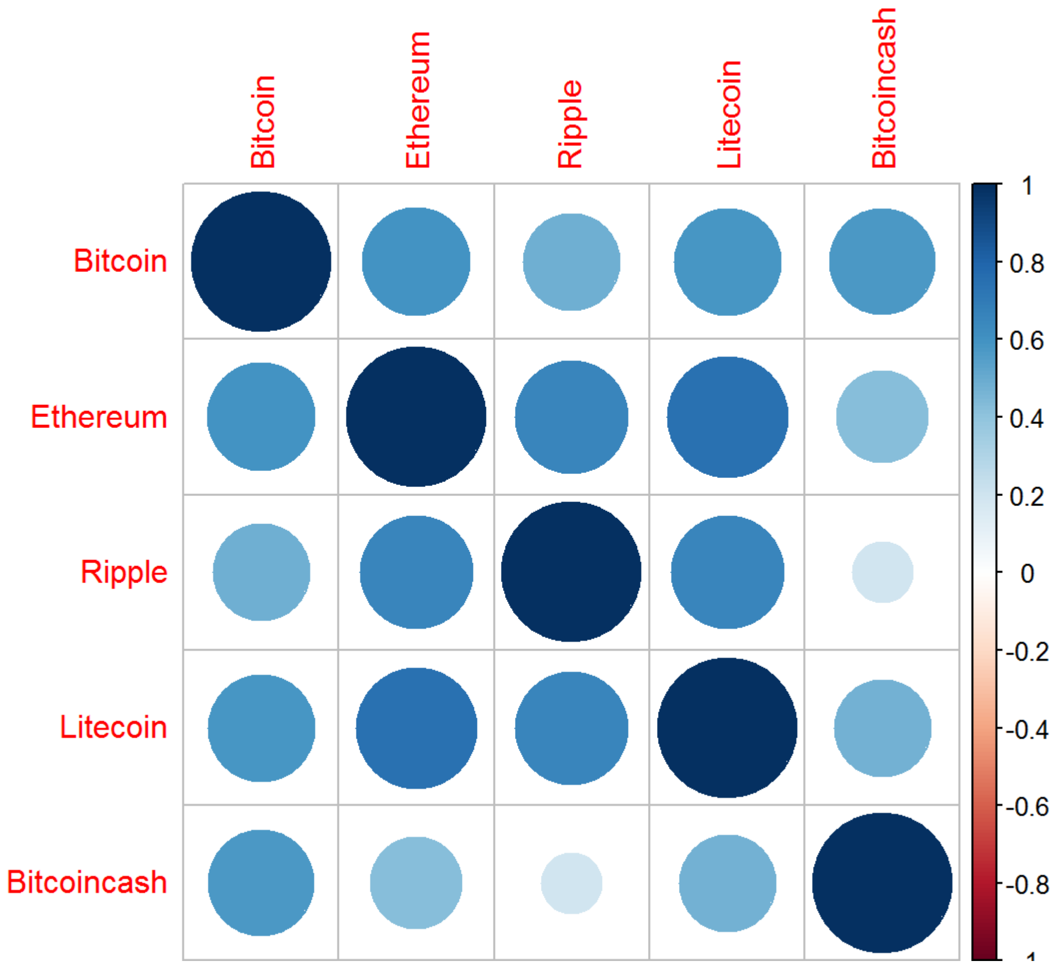
```
sd(df_ethereum$Percentage)
```

```
## [1] 0.08004012
```

Ethereum is also extremely volatile for a currency; it has gained up to 72% or lost up to 51% in only one day.

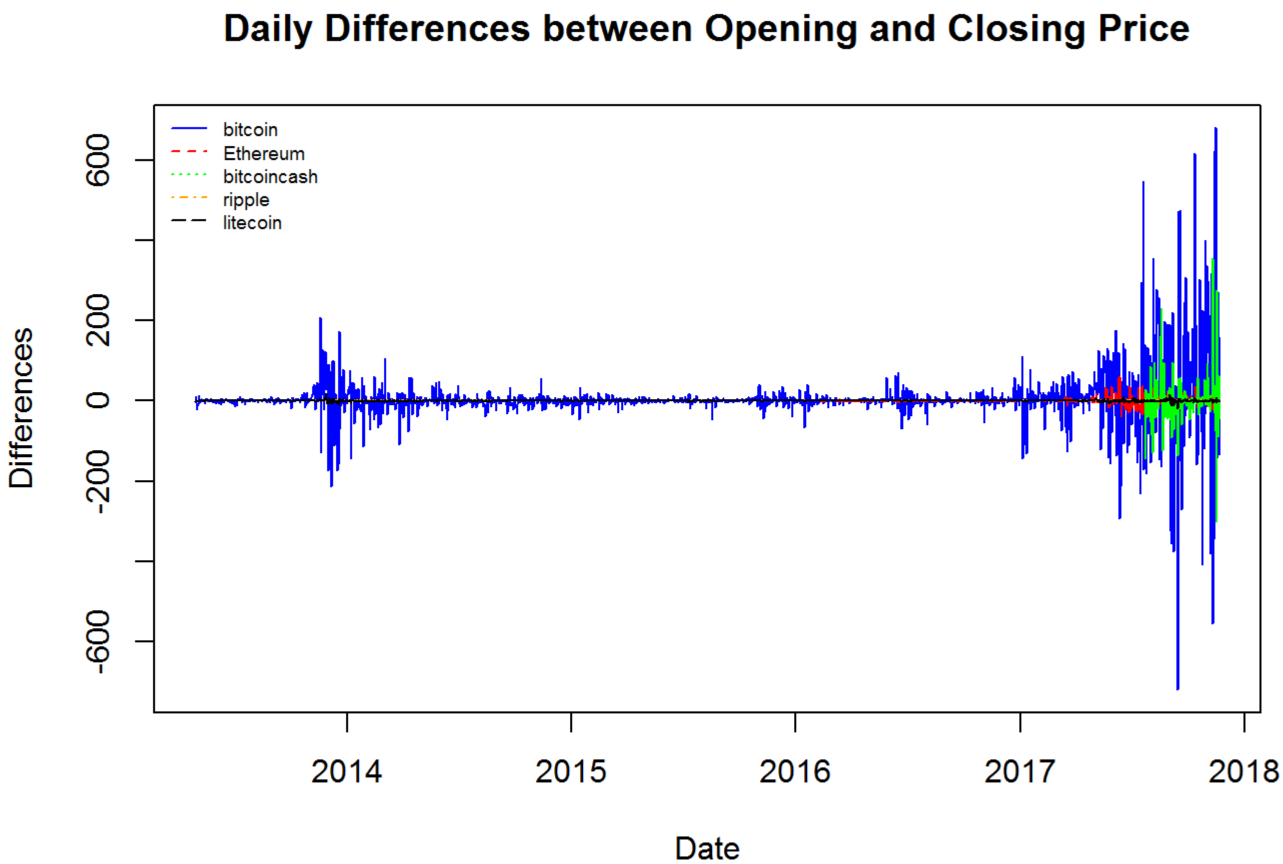


Correlation between cryptocurrencies' closing prices gives us an idea if they behave similarly to normal currencies or if people treat them basically being the same since they are not affiliated with anything but code chunks and blockchains.

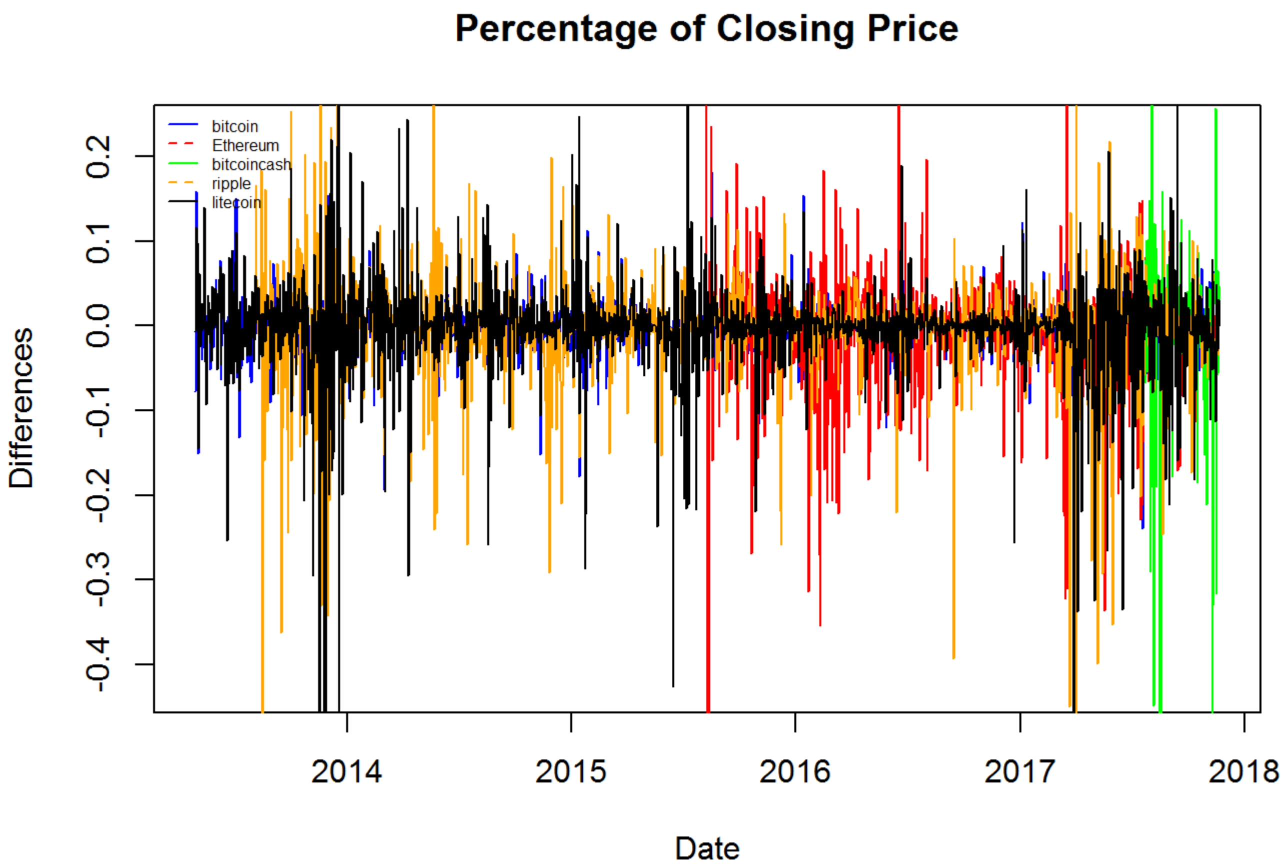


We can see that, at least the top five cryptocurrencies are correlated. This means that a negative news about one might affect the others negatively even if not at the same magnitude.

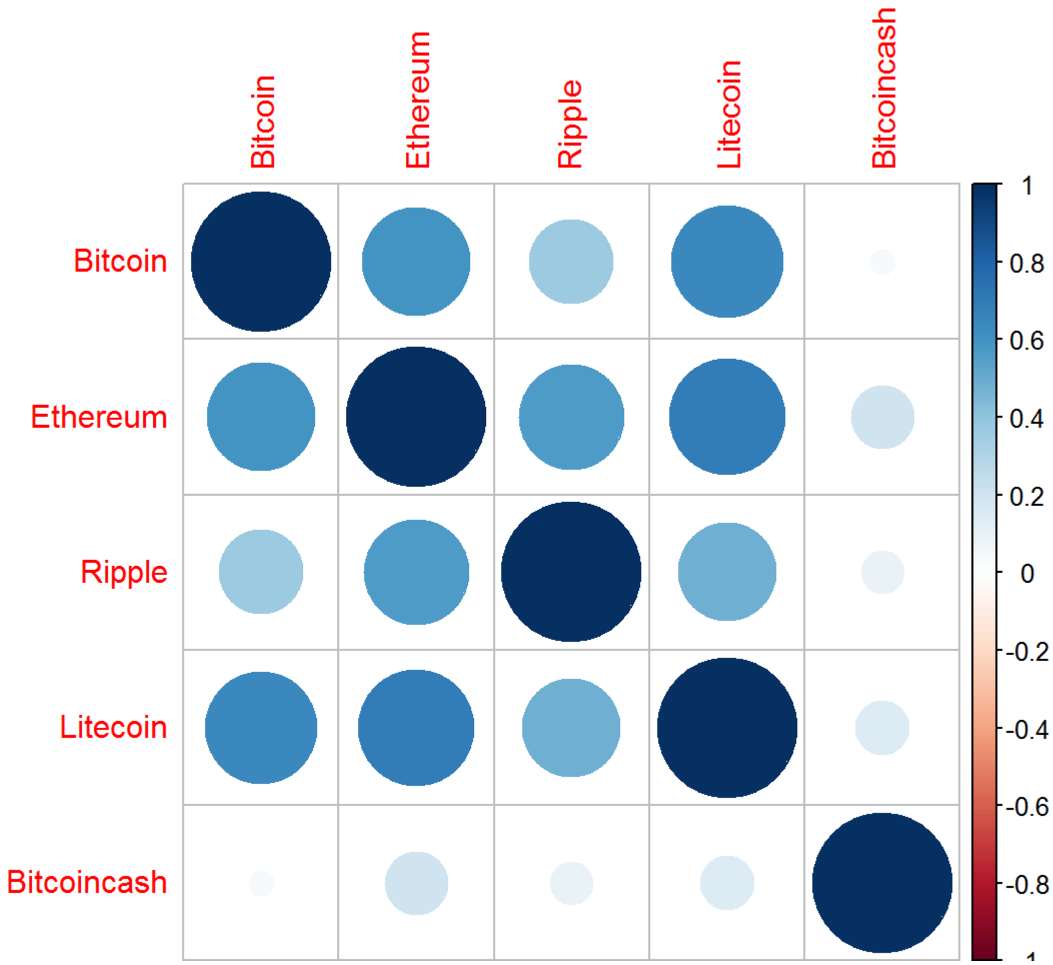
Wait a second! Are we jumping to conclusions here? Remember, these all sail at different price levels so we need some sort of normalization. Currently we trying to correlate a data like this:



You can’t even see Ripple and Litecoin in the graph. But after normalizing by turning them into percentage changes the data we are trying to correlate becomes like this:



Let’s look at the correlation between the percentage changes:



The correlations have weakened a little bit and BitcoinCash seems to be acting independently. So, good news! You might still bet on some crypto currencies while shorting others and make money on both exchanges! Obligatory Warning: The market is highly volatile so it can be too risky.

Since Bitcoin was first to market, has limited supply and unregulated a natural question to ask is “how similar it is to gold”? Correlation with the well known asset and the gold of the digital world is:

```
## [1] 0.1037629
```

It can be seen that the correlation is nothing of importance. Bitcoin is only a tiny fraction of gold in terms of market capitalization and gold’s physical vs Bitcoin’s digital makes them two separate markets even though you can mine Bitcoins, too.

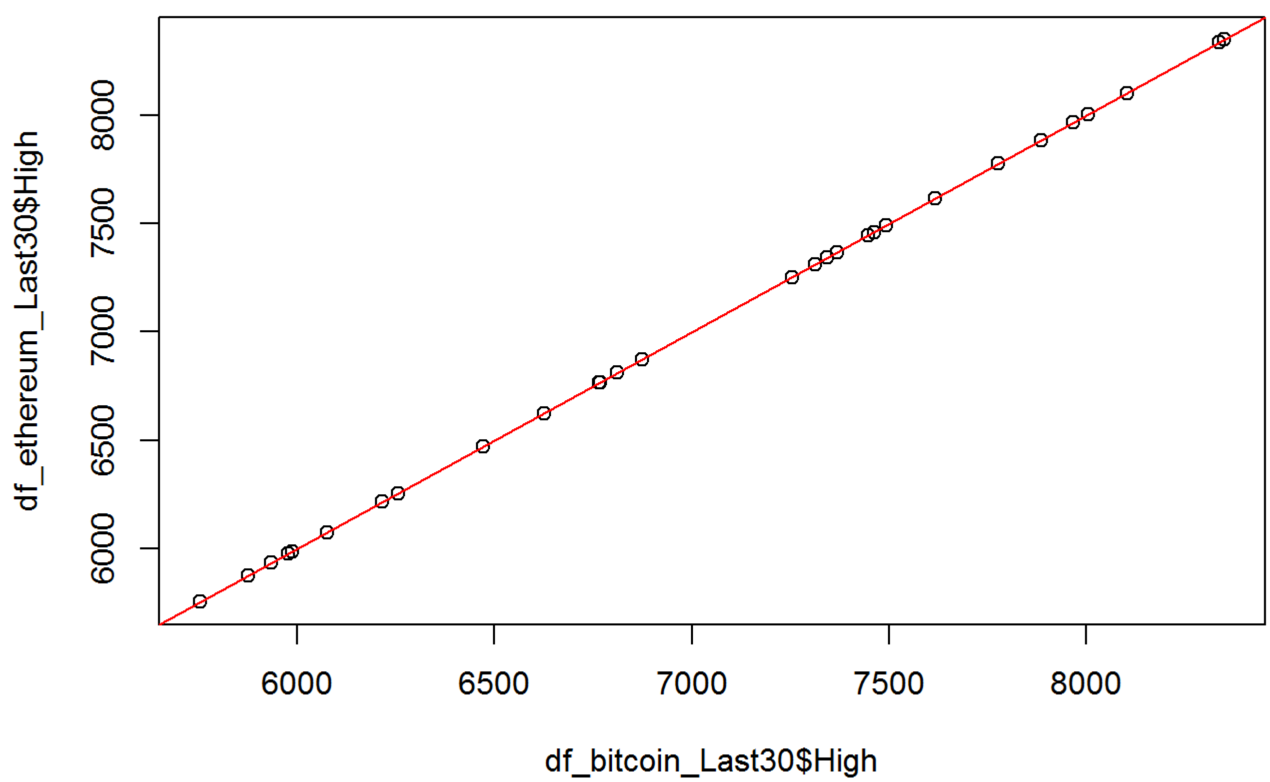
We have talked about the speculative aspect of crypto currencies, let’s see if our intuition was correct on this one. In here, we try to define the correlation between bitcoin and google search by comparing historical price datas to Google search trend of Bitcoin.

```
## [1] 0.9228585
```

Yes, they are extremely correlated! This could mean that the price of crypto currencies can be affected from negative or positive publicity and/or vice versa; a huge change in prices might spark interest and sequentially people might be searching for more information on them.

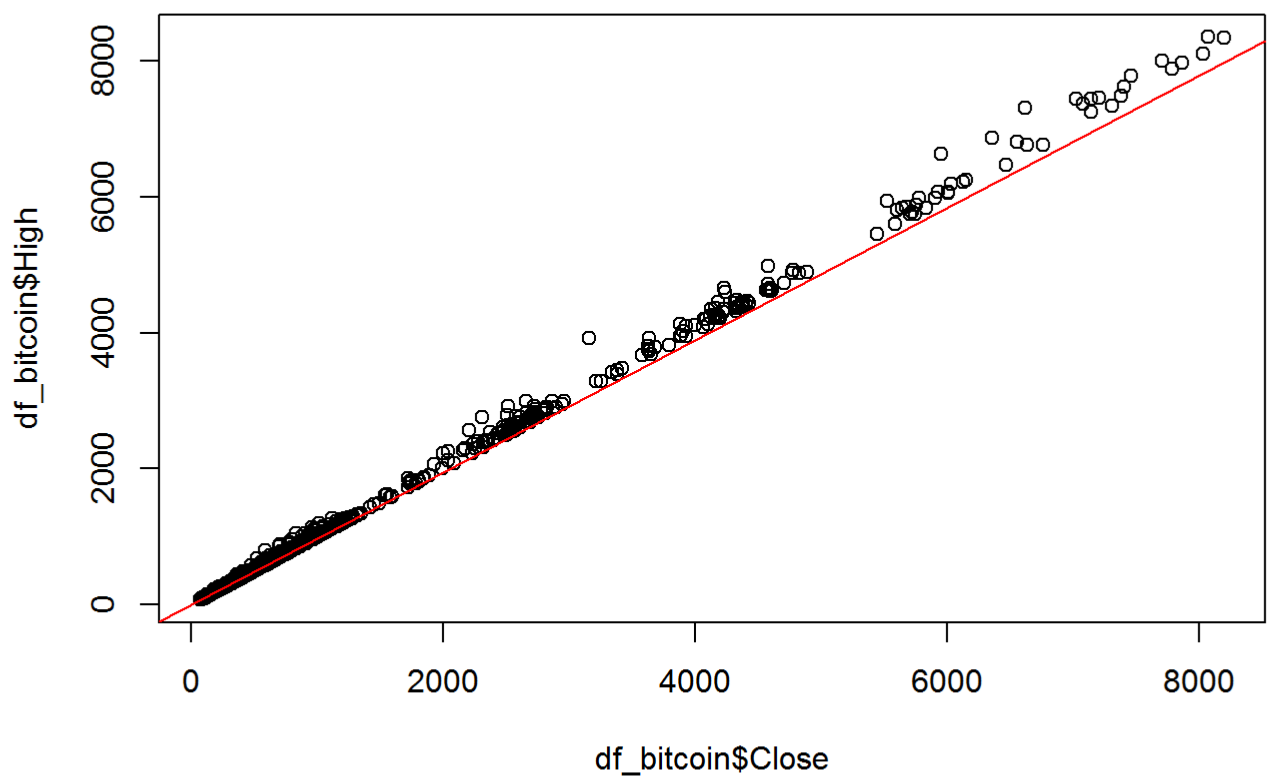
Linear regression analysis between bitcoin high price and Ethereum high price for last 30 days.

```
##
## Call:
## lm(formula = df_bitcoin_Last30$High ~ df_ethereum_Last30$High,
##     data = c(df_bitcoin_Last30, df_ethereum_Last30))
##
## Residuals:
##    Min     1Q Median     3Q    Max
##      0      0      0      0      0
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)             -1.328e-12  0.000e+00    -Inf   <2e-16 ***
## df_ethereum_Last30$High  1.000e+00  0.000e+00     Inf   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0 on 28 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:      1
## F-statistic:   Inf on 1 and 28 DF,  p-value: < 2.2e-16
```

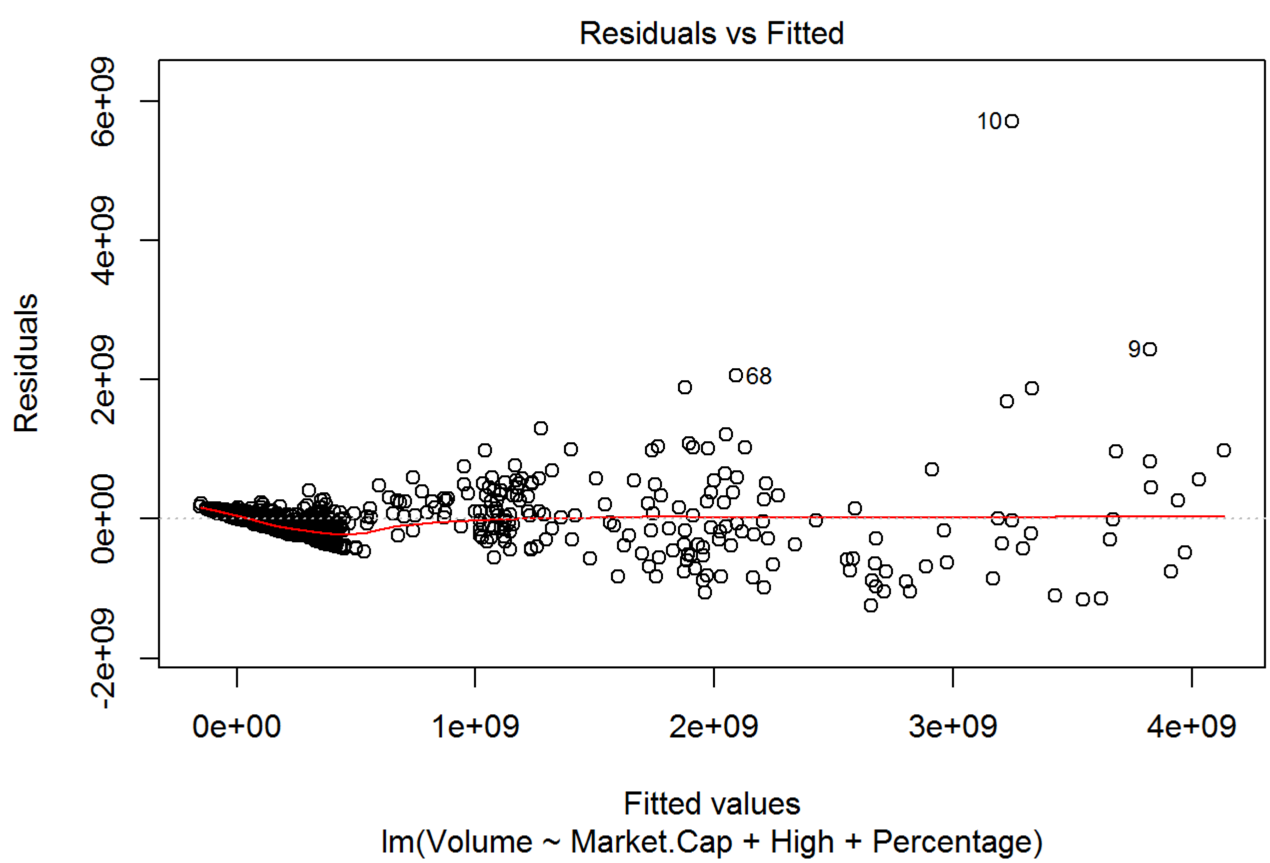


As it is seen in the figure; there is linear relationships between Ethereum and Bitcoin high prices. Linear Regression between bitcoin closing price and high price gives us

```
##
## Call:
## lm(formula = Close ~ High, data = df_bitcoin)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -660.21   -2.91    2.38    9.44  184.43
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept) 3.3411415  1.3357430    2.501   0.0125 *
## High        0.9722529  0.0008459 1149.388   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 44.79 on 1667 degrees of freedom
## Multiple R-squared:  0.9987, Adjusted R-squared:  0.9987
## F-statistic: 1.321e+06 on 1 and 1667 DF,  p-value: < 2.2e-16
```



```
##
## Call:
## lm(formula = Volume ~ Market.Cap + High + Percentage, data = df_bitcoin)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.250e+09 -9.321e+07  1.872e+07  9.518e+07  5.712e+09
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.395e+08  1.164e+07 -20.567   <2e-16 ***
## Market.Cap  -5.834e-02  6.939e-03  -8.408   <2e-16 ***
## High         1.444e+06  1.123e+05  12.862   <2e-16 ***
## Percentage  -3.951e+08  2.147e+08  -1.840    0.066 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 298100000 on 1422 degrees of freedom
##   (243 observations deleted due to missingness)
## Multiple R-squared:  0.843,  Adjusted R-squared:  0.8427
## F-statistic:  2545 on 3 and 1422 DF,  p-value: < 2.2e-16
```



These analyses helped us to understand cryptocurrencies a little better and lifts the veil of mystery around them a bit. Of course every mind asks different questions. Our minds wandered to the question of whether we can use machine learning techniques to make money of off this. Hence we applied 4 methods on python to understand the data better and see how they performed.

Summary of results; we get 58% accuracy on KNN and SVM for Bitcoin price predictions trained on past 30 days of prices. However this result is not that great considering the upward trajectory of Bitcoin. The best predictor predicted if it will go up with 55% accuracy and it will go down with 53% accuracy. That means we won’t be rich tomorrow. But we didn’t lose hope, we will try again.