An Exploration on Crypto Currencies

Emine Doganay, Hale Nur Kaza?esme, 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 …

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

```
##   X.                                          Name Symbol       Market.Cap
## 1  1      BTC\n                            Bitcoin    BTC $136,973,772,570
## 2  2     ETH\n                            Ethereum    ETH  $35,147,802,012
## 3  3 BCH\n                            Bitcoin Cash    BCH  $20,008,323,026
## 4  4       XRP\n                           Ripple    XRP   $9,103,758,162
## 5  5        DASH\n                           Dash   DASH   $4,073,795,924
## 6  6     LTC\n                            Litecoin    LTC   $3,812,420,255
##        Price
## 1  $8204.78
## 2   $366.58
## 3  $1189.88
## 4 $0.235709
## 5   $528.84
## 6    $70.68
##
Circulating.Supply
## 1
16,694,387
```

```
## 2
95,878,997
## 3
16,815,413
## 4 38,622,870,411\n                        \n                        \n
*
## 5
7,703,297
## 6
53,938,708
##      Volume..24h.  X..1h X..24h  X..7d
## 1 $3,314,020,000  0.44%  1.78% 18.16%
## 2   $663,017,000  0.37%  1.60%  8.80%
## 3   $508,538,000 -0.08%  0.86% -7.64%
## 4   $105,741,000  0.16% -0.20% 12.11%
## 5   $197,615,000  2.47% 12.33% 22.03%
## 6   $147,617,000  0.23%  0.26% 11.00%
```

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:

```
## Loading required package: jsonlite

##              id         name symbol   price_usd  price_btc 24h_volume_usd
## 1       bitcoin      Bitcoin    BTC 8204.780000 1.00000000     3314020000
## 2      ethereum     Ethereum    ETH  366.585000 0.04471740      663017000
## 3  bitcoin-cash Bitcoin Cash    BCH 1189.880000 0.14514600      508538000
## 4        ripple       Ripple    XRP    0.235709 0.00002875      105741000
## 5          dash         Dash   DASH  528.838000 0.06450950      197615000
## 6      litecoin     Litecoin    LTC   70.680600 0.00862188      147617000
## 7          iota         IOTA  MIOTA    0.926316 0.00011299       82304700
## 8           neo          NEO    NEO   36.322700 0.00443077       87355100
## 9       monero       Monero    XMR  144.870000 0.01767170       81963500
## 10          nem          NEM    XEM    0.206186 0.00002515        7380160
##    market_cap_usd available_supply total_supply  max_supply
## 1    136973772570         16694387     16694387 2.10000e+07
## 2     35147802012         95878997     95878997          NA
## 3     20008323026         16815413     16815413 2.10000e+07
## 4      9103758162      38622870411  99993173757 1.00000e+11
## 5      4073795924          7703297      7703297 1.89000e+07
## 6      3812420255         53938708     53938708 8.40000e+07
## 7      2574723374       2779530283   2779530283 2.77953e+09
## 8      2360975500         65000000    100000000          NA
## 9      2227807402         15377976     15377976          NA
## 10     1855674000       8999999999   8999999999          NA
##    percent_change_1h percent_change_24h percent_change_7d last_updated
## 1               0.44               1.78             18.16   1511337258
## 2               0.37               1.60              8.80   1511337253
```

```
## 3                 -0.08                0.86              -7.64   1511337270
## 4                  0.16               -0.20              12.11   1511337241
## 5                  2.47               12.33              22.03   1511337247
## 6                  0.23                0.26              11.00   1511337242
## 7                  0.94                4.09              42.96   1511337266
## 8                  0.62                0.54              21.35   1511337258
## 9                 -0.65                6.01              17.43   1511337244
## 10                 0.41               -0.25               5.12   1511337249
```

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$Name column
df_cryptocurrencies$Name <- lapply(df_cryptocurrencies$Name, gsub, pattern =
"\n", replacement = "-")

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

# 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", "change_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
## 1     BTC-Bitcoin    BTC 136973772570  8204.78    16694387 3314020000
## 2    ETH-Ethereum    ETH  35147802012   366.58    95878997  663017000
## 3 BCH-BitcoinCash    BCH  20008323026  1189.88    16815413  508538000
## 4      XRP-Ripple    XRP   9103758162 0.235709 38622870411  105741000
## 5       DASH-Dash   DASH   4073795924   528.84     7703297  197615000
## 6    LTC-Litecoin    LTC   3812420255    70.68    53938708  147617000
##   change_1h change_24h change_7d
## 1      0.44       1.78     18.16
## 2      0.37       1.60      8.80
## 3     -0.08       0.86     -7.64
## 4      0.16      -0.20     12.11
## 5      2.47      12.33     22.03
## 6      0.23       0.26     11.00
```

For mathematical analysis variable types should be numeric.

```r
# 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.

```
## Warning in lapply(df_cryptocurrencies[-c(1:2)], as.numeric): Zorlamadan
## dolayı ortaya çıkan NAs

## [1] "double"
```

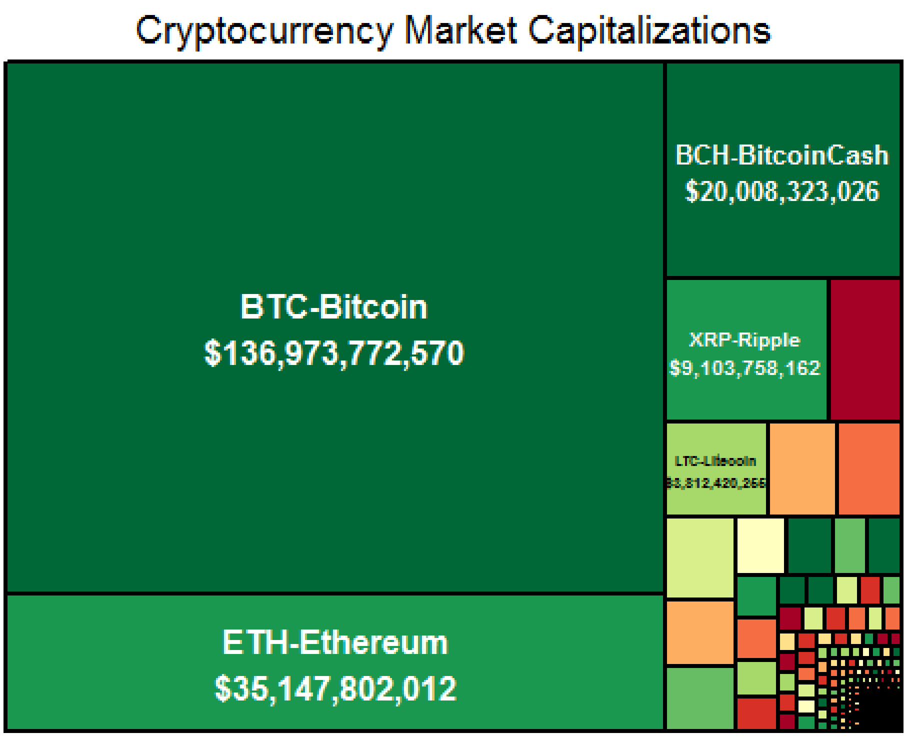
Let's see how our data frame looks like now:

```
##                   name symbol    marketcap       price       supply
## 1          BTC-Bitcoin    BTC 136973772570 8204.780000     16694387
## 2         ETH-Ethereum    ETH  35147802012  366.580000     95878997
## 3      BCH-BitcoinCash    BCH  20008323026 1189.880000     16815413
## 4           XRP-Ripple    XRP   9103758162    0.235709 38622870411
## 5            DASH-Dash   DASH   4073795924  528.840000      7703297
## 6         LTC-Litecoin    LTC   3812420255   70.680000     53938708
## 7           MIOTA-IOTA  MIOTA   2574723374    0.926316   2779530283
## 8              NEO-NEO    NEO   2360975500   36.320000     65000000
## 9           XMR-Monero    XMR   2227807402  144.870000     15377976
## 10             XEM-NEM    XEM   1855674000    0.206186   8999999999
## 11 ETC-EthereumClassic    ETC   1763882506   18.060000     97678730
## 12            LSK-Lisk    LSK   1141165060    9.910000    115156368
## 13           QTUM-Qtum   QTUM   1039095501   14.100000     73679040
## 14           ZEC-Zcash    ZEC    790154285  295.540000      2673631
## 15         ADA-Cardano    ADA    746035898    0.028774 25927070538
##        volume change_1h change_24h change_7d
## 1  3314020000      0.44       1.78     18.16
## 2   663017000      0.37       1.60      8.80
## 3   508538000     -0.08       0.86     -7.64
## 4   105741000      0.16      -0.20     12.11
## 5   197615000      2.47      12.33     22.03
## 6   147617000      0.23       0.26     11.00
## 7    82304700      0.94       4.09     42.96
## 8    87355100      0.62       0.54     21.35
## 9    81963500     -0.65       6.01     17.43
## 10    7380160      0.41      -0.25      5.12
## 11  100393000      0.46       1.11      2.22
## 12   24635600      1.62      -0.63     28.26
## 13   71997300     -0.29      -1.13     21.17
## 14   60160100      0.41       1.14      7.18
## 15    4285500      0.51       1.08      7.14

## [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 via using 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
## -49.1400   0.0000   0.3500   0.6793   0.4300 126.8500       70

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

## [1] 7.577462
```

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 
##  -95.900   -4.882    0.720    9.282    5.145 5061.760       68
```

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

```
## [1] 178.1998
```

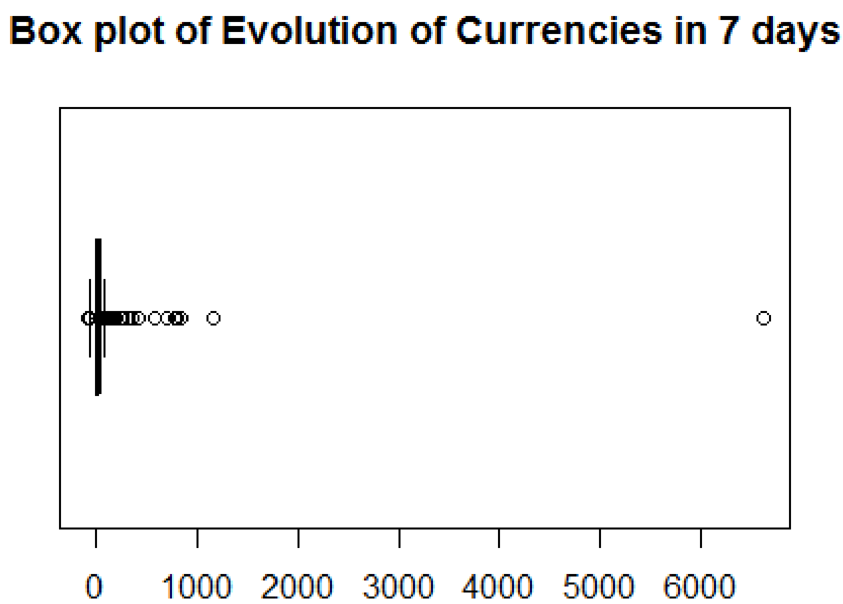
Summary of statistics and standard deviation of market change in 7 days are as follows:

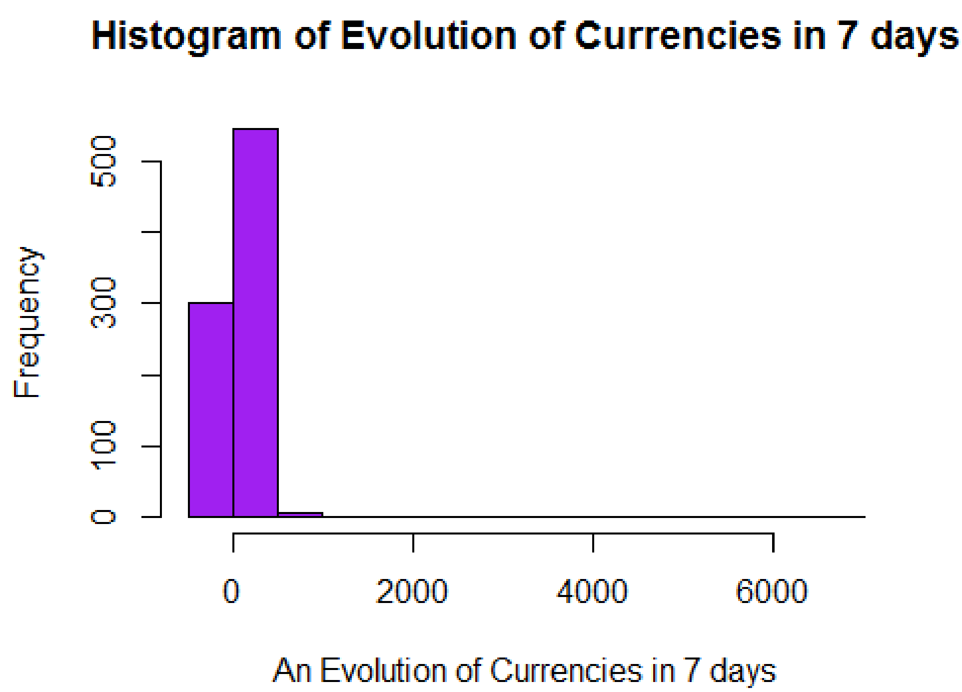
```r
summary(df_allcryptocurrencies$change_7d)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's 
##  -92.210   -7.748    8.720   26.463   23.405 6618.870       60
```

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

```
## [1] 240.2387
```





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 $ 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
##  [4] ONX-Onix             SUPER-SuperCoin      BCO-BridgeCoin
##  [7] GRE-Greencoin        UNIT-UniversalCur... XSH-SHIELD
## [10] ORB-Orbitcoin        XGR-GoldReserve      BCF-BitcoinFast
## [13] BXT-BitTokens        BTCR-Bitcurrency     4CHN-ChanCoin
## [16] 888-OctoCoin         GB-GoldBlocks        USDE-USDe
## [19] VEC2-VectorAI        XRC-Rawcoin
## 914 Levels: -Money 020-020LondonCoin 1337-1337 2GIVE-2GIVE ... 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
##  [4] DIME-Dimecoin     XGOX-XGOX         FIMK-FIMKrypto
##  [7] ZEIT-Zeitcoin     1337-1337         RNS-Renos
## [10] HNC-Helleniccoin  LDOGE-LiteDoge    MAO-MaoZedong
## [13] XIOS-Xios         LUX-LUXCoin       CRM-Cream
## [16] BLAS-BlakeStar    SPRTS-Sprouts     MNC-Mincoin
## [19] ATOM-AtomicCoin   EMB-EmberCoin     808-808Coin
## [22] EL-Elcoin         EUC-Eurocoin      TOR-Torcoin
## [25] ITZ-Interzone     XCS-CybCSec       CTIC2-Coimatic2.0
## [28] FAL-Falcoin
## 914 Levels: -Money 020-020LondonCoin 1337-1337 2GIVE-2GIVE ... 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] "$233,697,977,326"
```

and Bitcoin's share is:

```
## [1] "0.5861145%"
```

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

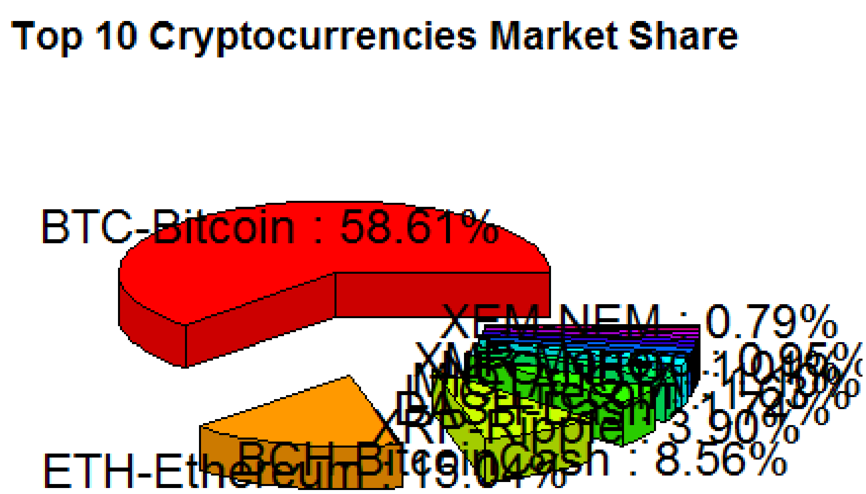
Let's focus on top 10 currrencies by market cap.

```
##               name symbol    marketcap       price      supply     volume
## 1      BTC-Bitcoin    BTC 136973772570 8204.780000    16694387 3314020000
## 2     ETH-Ethereum    ETH  35147802012  366.580000    95878997  663017000
## 3  BCH-BitcoinCash    BCH  20008323026 1189.880000    16815413  508538000
## 4       XRP-Ripple    XRP   9103758162    0.235709 38622870411  105741000
## 5        DASH-Dash   DASH   4073795924  528.840000     7703297  197615000
## 6     LTC-Litecoin    LTC   3812420255   70.680000    53938708  147617000
## 7       MIOTA-IOTA  MIOTA   2574723374    0.926316  2779530283   82304700
## 8          NEO-NEO    NEO   2360975500   36.320000    65000000   87355100
## 9       XMR-Monero    XMR   2227807402  144.870000    15377976   81963500
## 10         XEM-NEM    XEM   1855674000    0.206186  8999999999    7380160
##    change_1h change_24h change_7d          formatted_market_cap
## 1       0.44       1.78     18.16    BTC-Bitcoin\n$136,973,772,570
```

```
## 2        0.37        1.60       8.80     ETH-Ethereum\n$35,147,802,012
## 3       -0.08        0.86      -7.64 BCH-BitcoinCash\n$20,008,323,026
## 4        0.16       -0.20      12.11         XRP-Ripple\n$9,103,758,162
## 5        2.47       12.33      22.03          DASH-Dash\n$4,073,795,924
## 6        0.23        0.26      11.00      LTC-Litecoin\n$3,812,420,255
## 7        0.94        4.09      42.96        MIOTA-IOTA\n$2,574,723,374
## 8        0.62        0.54      21.35           NEO-NEO\n$2,360,975,500
## 9       -0.65        6.01      17.43        XMR-Monero\n$2,227,807,402
## 10       0.41       -0.25       5.12           XEM-NEM\n$1,855,674,000
##    marketcap_percent
## 1        0.586114498
## 2        0.150398401
## 3        0.085616158
## 4        0.038955229
## 5        0.017431884
## 6        0.016313450
## 7        0.011017311
## 8        0.010102678
## 9        0.009532848
## 10       0.007940480
```

Market shares of top 10 coins are:

```
## [1] "93.34229%"
```



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    marketcap       price      supply     volume
## 1     BTC-Bitcoin    BTC 136973772570 8204.780000    16694387 3314020000
## 2    ETH-Ethereum    ETH  35147802012  366.580000    95878997  663017000
## 3 BCH-BitcoinCash    BCH  20008323026 1189.880000    16815413  508538000
## 4      XRP-Ripple    XRP   9103758162    0.235709 38622870411  105741000
## 5       DASH-Dash   DASH   4073795924  528.840000     7703297  197615000
##   change_1h change_24h change_7d           formatted_market_cap
## 1      0.44       1.78     18.16    BTC-Bitcoin\n$136,973,772,570
## 2      0.37       1.60      8.80    ETH-Ethereum\n$35,147,802,012
## 3     -0.08       0.86     -7.64 BCH-BitcoinCash\n$20,008,323,026
## 4      0.16      -0.20     12.11       XRP-Ripple\n$9,103,758,162
## 5      2.47      12.33     22.03        DASH-Dash\n$4,073,795,924
##   marketcap_percent
## 1        0.58611450
## 2        0.15039840
## 3        0.08561616
## 4        0.03895523
## 5        0.01743188
```

We import the historical data for these currencies.

Bitcoin

```
## Warning in lapply(X = X, FUN = FUN, ...): Zorlamadan dolayı ortaya çıkan
## NAs

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

```
## Warning in lapply(X = X, FUN = FUN, ...): Zorlamadan dolayı ortaya çıkan
## NAs

##       Date                 Open              High        
##  Min.   :2015-08-07   Min.   :  0.4316   Min.   :  0.483  
##  1st Qu.:2016-03-03   1st Qu.:  7.1350   1st Qu.:  7.357  
##  Median :2016-09-28   Median : 11.4500   Median : 11.825  
##  Mean   :2016-09-28   Mean   : 73.0548   Mean   : 76.140  
##  3rd Qu.:2017-04-25   3rd Qu.: 53.0425   3rd Qu.: 54.867  
##  Max.   :2017-11-21   Max.   :397.5900   Max.   :414.760  
##                                                           
##       Low               Close              Volume         
##  Min.   :  0.4209   Min.   :  0.4348   Min.   :1.021e+05  
##  1st Qu.:  6.7575   1st Qu.:  7.1700   1st Qu.:5.850e+06  
##  Median : 11.1750   Median : 11.4850   Median :1.671e+07  
##  Mean   : 70.0287   Mean   : 73.4710   Mean   :2.187e+08  
##  3rd Qu.: 51.5400   3rd Qu.: 53.1250   3rd Qu.:1.691e+08  
##  Max.   :383.4700   Max.   :401.4900   Max.   :2.883e+09  
##                                                           
##    Market.Cap          Percentage       
##  Min.   :3.197e+07   Min.   :-0.510344  
##  1st Qu.:5.880e+08   1st Qu.:-0.034190  
##  Median :9.600e+08   Median : 0.001083  
##  Mean   :6.806e+09   Mean   :-0.009241  
##  3rd Qu.:4.810e+09   3rd Qu.: 0.025597  
##  Max.   :3.676e+10   Max.   : 0.728042  
##  NA's   :1
```

BitcoinCash

```
## Warning in lapply(X = X, FUN = FUN, ...): Zorlamadan dolayı ortaya çıkan
## NAs

##       Date                 Open             High              Low        
##  Min.   :2017-07-23   Min.   : 212.2   Min.   : 223.7   Min.   : 201.0  
##  1st Qu.:2017-08-22   1st Qu.: 343.4   1st Qu.: 371.7   1st Qu.: 316.5  
##  Median :2017-09-21   Median : 438.6   Median : 471.8   Median : 414.4  
##  Mean   :2017-09-21   Mean   : 517.5   Mean   : 577.2   Mean   : 476.1  
##  3rd Qu.:2017-10-21   3rd Qu.: 601.1   3rd Qu.: 642.5   3rd Qu.: 569.1  
##  Max.   :2017-11-21   Max.   :1381.8   Max.   :2477.7   Max.   :1215.6  
##                                                                          
##      Close            Volume            Market.Cap       
##  Min.   : 213.2   Min.   :8.501e+04   Min.   :3.497e+09  
##  1st Qu.: 343.0   1st Qu.:1.645e+08   1st Qu.:5.664e+09  
##  Median : 439.6   Median :2.993e+08   Median :7.396e+09  
##  Mean   : 521.9   Mean   :7.065e+08   Mean   :8.784e+09
```

```
##  3rd Qu.: 606.2   3rd Qu.:7.658e+08   3rd Qu.:1.027e+10
##  Max.   :1388.9   Max.   :8.371e+09   Max.   :2.321e+10
##                                       NA's   :10
##    Percentage
##  Min.   :-0.539691
##  1st Qu.:-0.052810
##  Median : 0.004763
##  Mean   :-0.016847
##  3rd Qu.: 0.048317
##  Max.   : 0.359841
##
```

Ripple

```
## Warning in lapply(X = X, FUN = FUN, ...): Zorlamadan dolayı ortaya çıkan
## NAs

##       Date                 Open               High
##  Min.   :2013-08-04   Min.   :0.002809   Min.   :0.003082
##  1st Qu.:2014-08-31   1st Qu.:0.006023   1st Qu.:0.006126
##  Median :2015-09-28   Median :0.007760   Median :0.007924
##  Mean   :2015-09-28   Mean   :0.036805   Mean   :0.038760
##  3rd Qu.:2016-10-24   3rd Qu.:0.014436   3rd Qu.:0.015033
##  Max.   :2017-11-21   Max.   :0.394931   Max.   :0.435170
##
##       Low               Close              Volume
##  Min.   :0.002802   Min.   :0.002810   Min.   :8.316e+03
##  1st Qu.:0.005933   1st Qu.:0.006027   1st Qu.:2.829e+05
##  Median :0.007543   Median :0.007772   Median :6.692e+05
##  Mean   :0.034998   Mean   :0.036947   Mean   :2.950e+07
##  3rd Qu.:0.014032   3rd Qu.:0.014435   3rd Qu.:2.307e+06
##  Max.   :0.332222   Max.   :0.394132   Max.   :1.770e+09
##                                        NA's   :145
##    Market.Cap          Percentage
##  Min.   :2.196e+07   Min.   :-1.793669
##  1st Qu.:1.431e+08   1st Qu.:-0.020750
##  Median :2.341e+08   Median : 0.002604
##  Mean   :1.302e+09   Mean   :-0.005443
##  3rd Qu.:3.157e+08   3rd Qu.: 0.020879
##  Max.   :1.513e+10   Max.   : 0.460047
##
```

Litecoin

```
## Warning in lapply(X = X, FUN = FUN, ...): Zorlamadan dolayı ortaya çıkan
## NAs

##       Date                 Open             High             Low
##  Min.   :2013-04-28   Min.   : 1.15   Min.   : 1.34   Min.   : 1.11
##  1st Qu.:2014-06-19   1st Qu.: 3.09   1st Qu.: 3.14   1st Qu.: 3.02
##  Median :2015-08-10   Median : 3.86   Median : 3.92   Median : 3.81
```

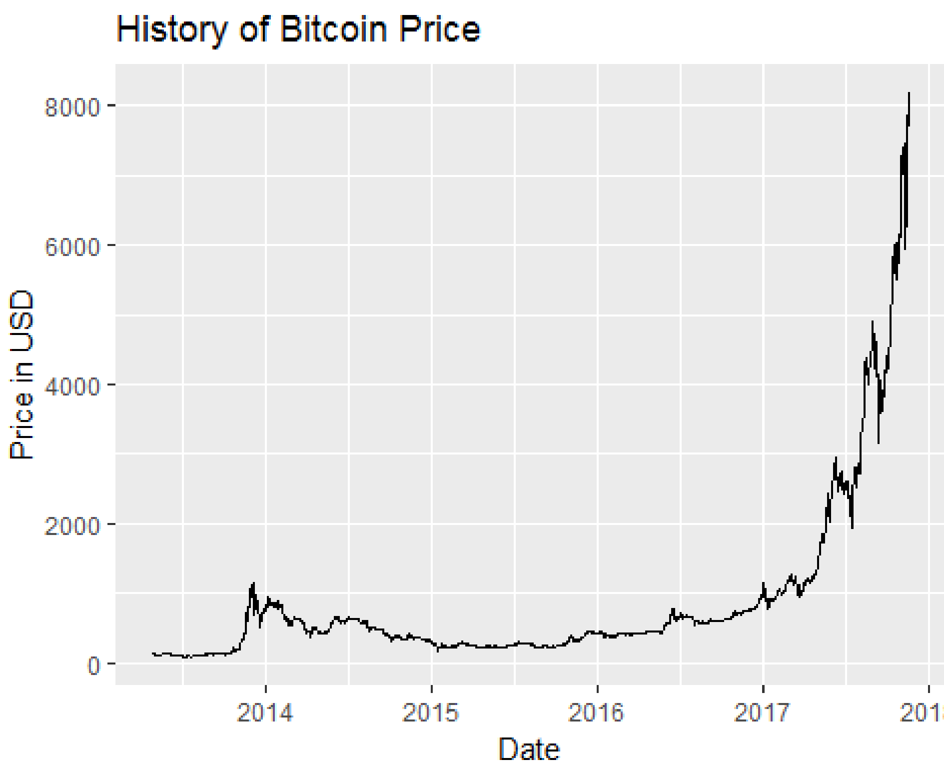
```
##  Mean   :2015-08-10   Mean   :10.75   Mean   :11.23   Mean   :10.25
##  3rd Qu.:2016-09-30   3rd Qu.:10.30   3rd Qu.:10.55   3rd Qu.: 9.80
##  Max.   :2017-11-21   Max.   :85.83   Max.   :92.07   Max.   :75.59
##
##      Close           Volume            Market.Cap
##  Min.   : 1.16   Min.   :4.817e+05   Min.   :3.794e+07
##  1st Qu.: 3.08   1st Qu.:1.755e+06   1st Qu.:1.205e+08
##  Median : 3.86   Median :3.162e+06   Median :1.760e+08
##  Mean   :10.79   Mean   :5.047e+07   Mean   :4.698e+08
##  3rd Qu.:10.31   3rd Qu.:9.418e+06   3rd Qu.:2.959e+08
##  Max.   :86.04   Max.   :1.731e+09   Max.   :4.527e+09
##                  NA's   :243
##    Percentage
##  Min.   :-1.290954
##  1st Qu.:-0.016491
##  Median : 0.000000
##  Mean   :-0.004071
##  3rd Qu.: 0.017422
##  Max.   : 0.401857
##
```

Prepare date format of top 5 crypto currencies

```
df_bitcoin$Date <- as.Date(df_bitcoin$Date)
df_ethereum$Date <- as.Date(df_ethereum$Date)
df_bitcoincash$Date <- as.Date(df_bitcoincash$Date)
df_ripple$Date <- as.Date(df_ripple$Date)
df_litecoin$Date <- as.Date(df_litecoin$Date)
```

Let’s examine the percentage of price changes for Bitcoin:

```
# Plot bitcoin price versus time
p_01 <- ggplot(data = df_bitcoin)
p_01 + labs(title = "History of Bitcoin Price") + labs(x = "Date") + labs(y =
"Price in USD") + geom_line(aes(x = Date, y = Close))
```

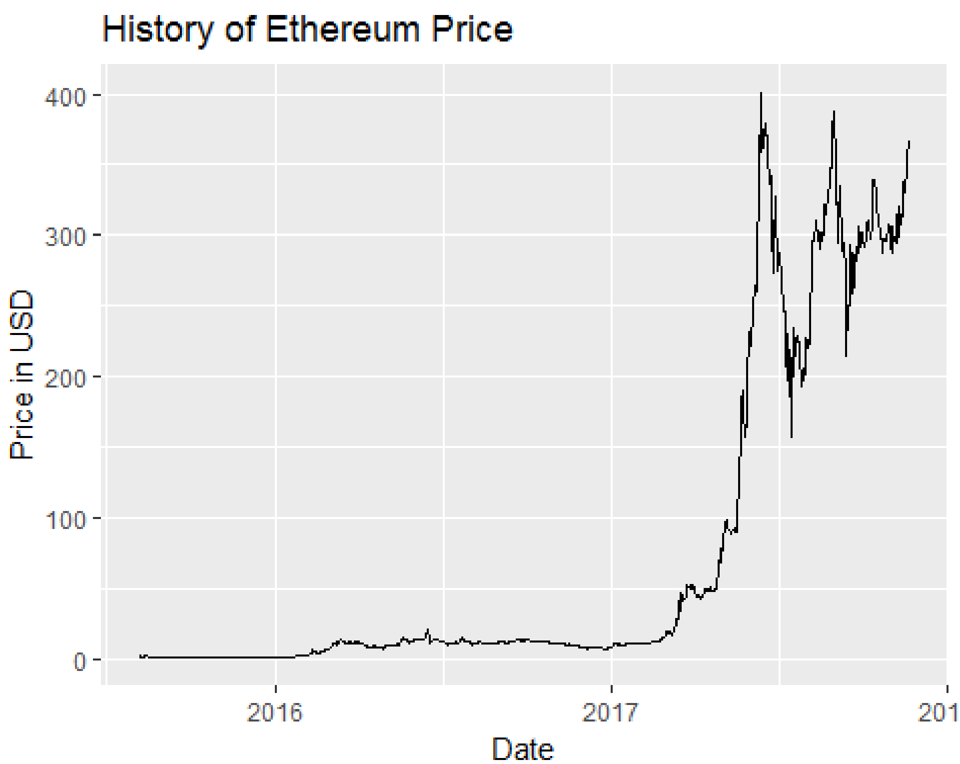



```
summary(df_bitcoin$Percentage)

##       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 very 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:

```
# Plot Ethereum price versus time
p <- ggplot(data = df_ethereum)
p + geom_line(aes(x = Date, y = Close)) + labs(x = "Date") + labs(y = "Price
in USD") + labs(title = "History of Ethereum Price")
```



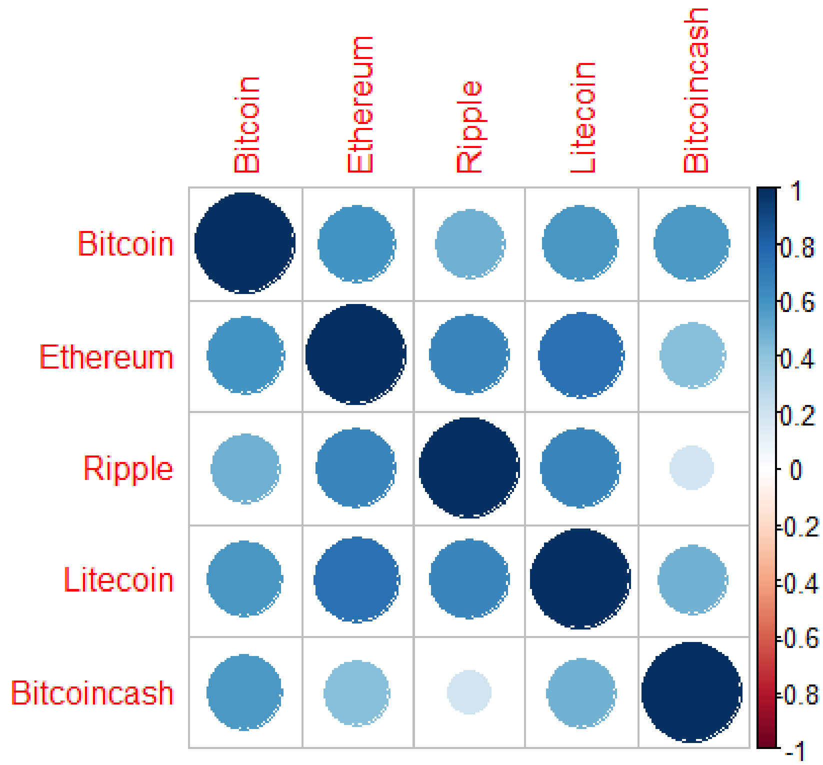

```r
summary(df_ethereum$Percentage)
```

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

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

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

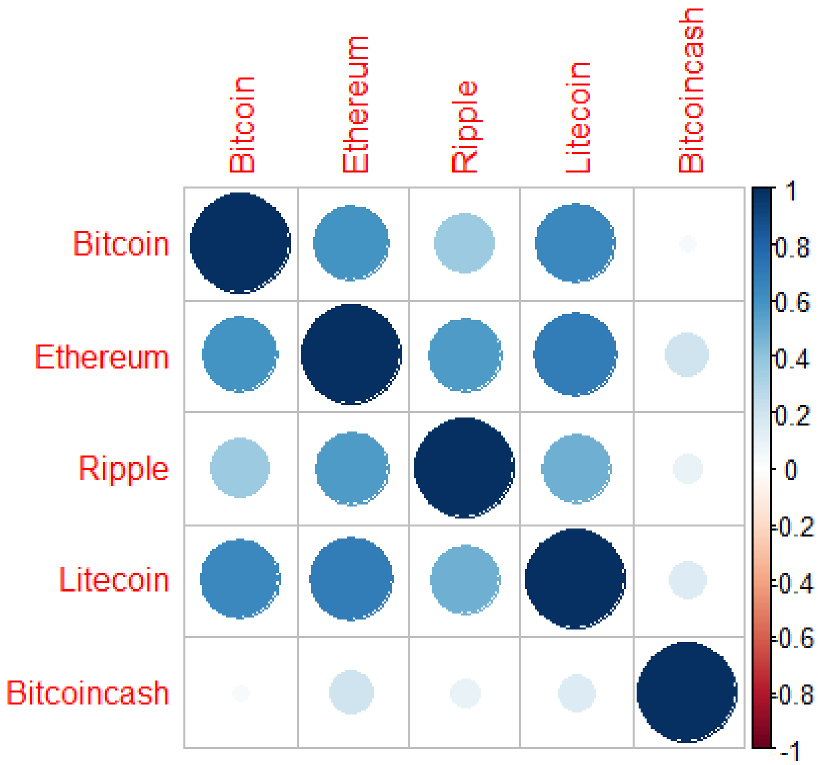
Ethereum is also extremely volatile for a currency: it has gained up to 72% or lost up to 51% in only one day.




```
##                Bitcoin  Ethereum    Ripple  Litecoin Bitcoincash
## Bitcoin      1.0000000 0.5977615 0.4830272 0.5872964   0.5739115
## Ethereum     0.5977615 1.0000000 0.6510914 0.7441854   0.4290277
## Ripple       0.4830272 0.6510914 1.0000000 0.6542968   0.1902201
## Litecoin     0.5872964 0.7441854 0.6542968 1.0000000   0.4797128
## Bitcoincash  0.5739115 0.4290277 0.1902201 0.4797128   1.0000000
```

The market capitalizations of these five cryptocurrencies are highly correlated.

```
# Correlation of volatility between cryptocurrencies
correlation <- cor(cbind(Bitcoin = df_bitcoin[df_bitcoin$Date >
as.Date('2017-07-23'),]$Percentage, Ethereum = df_ethereum[df_ethereum$Date >
as.Date('2017-07-23'),]$Percentage, Ripple = df_ripple[df_ripple$Date >
as.Date('2017-07-23'),]$Percentage, Litecoin = df_litecoin[df_litecoin$Date >
as.Date('2017-07-23'),]$Percentage, Bitcoincash =
df_bitcoincash[df_bitcoincash$Date > as.Date('2017-07-23'),]$Percentage))
corrplot(correlation)
```



The correlations are much weaker except between Bitcoin and Bitcoincash .It is interesting to trade cryptocurrencies on exchanges. The market is highly volatile so it can be dangerous.

Correlation with well known Asset, Gold. Bitcoin may be accepted as the gold of the digital world. Are gold and Bitcoin correlated ?

```
correlation <- function(series1, series2, k = 0) {
  len = length(series1)
  return(cor(cbind(series1[1:(len-k)], series2[(1+k):len]), use =
'na.or.complete')[1,2])
}
# The daily prices of gold on Quandl are downloaded as CSV format,
gold_daily.csv
Gold_Data <- read.csv('gold_daily.csv', header = T, stringsAsFactors = F)
df_Gold <- data.frame(Date = as.Date(Gold_Data$Date), Close =
Gold_Data$Value)
correlation(df_bitcoin$Close,
df_Gold[is.element(df_Gold$Date,df_bitcoin$Date),]$Close)

## [1] 0.1037629
```

It can be seen that there is no correlation. Bitcoin is only a tiny fraction of gold in terms of market capitalization. Bitcoin and gold are two separate markets

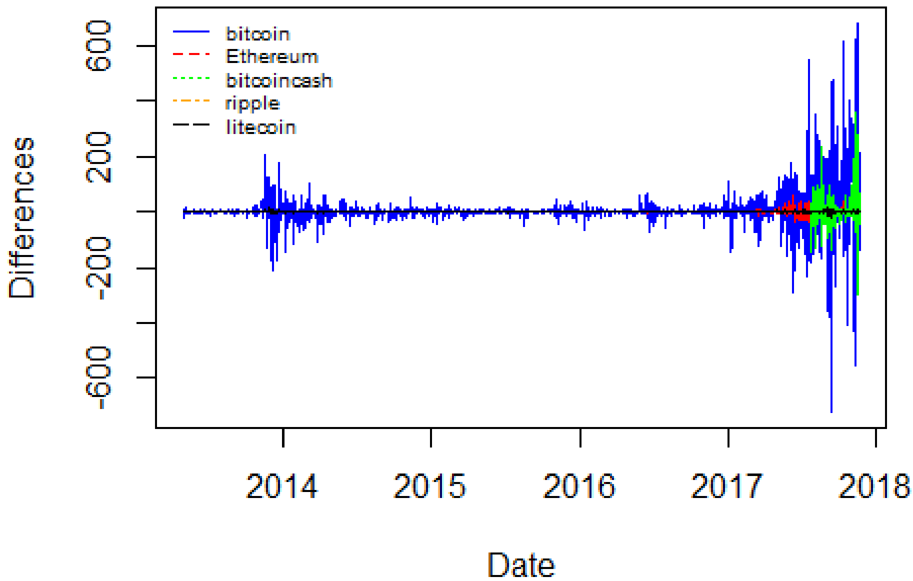
Correlation: Bitcoin vs Google In here, we try to define the correlation between bitcoin and google search: Import Google search historical price datas.

```r
df_GoogleSearch <- read.csv('multiTimeline.csv', header = F, stringsAsFactors
= F)[c(-1,-2),]
colnames(df_GoogleSearch) <- c('Date', 'Volume')
df_GoogleSearch$Date <- as.Date(df_GoogleSearch$Date)
df_GoogleSearch$Volume <- as.numeric(df_GoogleSearch$Volume)
row.names(df_GoogleSearch) <- NULL
df_GoogleSearch <- df_GoogleSearch[nrow(df_GoogleSearch):1,]
correlation(df_GoogleSearch$Volume, df_bitcoin[is.element(df_bitcoin$Date,
df_GoogleSearch$Date),]$Close)

## [1] 0.9228585
```

In here; open and close price graphs for 5 top curriencies is given in the same plot.

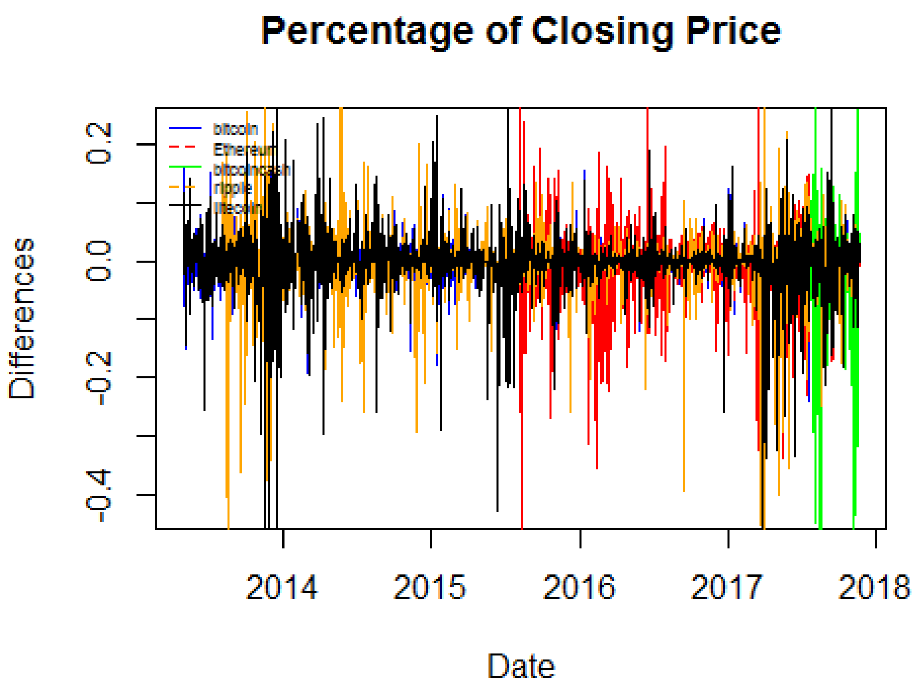
```r
plot(df_bitcoin$Date,(df_bitcoin$Close-df_bitcoin$Open), type = 'line',
col="blue", main = 'Daily Differences between Opening and Closing Price',
xlab ='Date', ylab = 'Differences')

## Warning in plot.xy(xy, type, ...): çizim tipi 'line' in ilk karakteri
## kullanılacak

lines(df_ethereum$Date,(df_ethereum$Close-df_ethereum$Open), col="red")
lines(df_bitcoincash$Date,(df_bitcoincash$Close-df_bitcoincash$Open),
col='green')
lines(df_ripple$Date,(df_ripple$Close-df_ripple$Open), col="orange")
lines(df_litecoin$Date,(df_litecoin$Close-df_litecoin$Open), col='black')
legend("topleft", legend=c("bitcoin", "Ethereum", "bitcoincash", "ripple",
"litecoin"),
       col=c("blue", "red", "green", "orange","black"), lty=1:5, bty ="n" ,
cex=0.6)
```

Daily Differences between Opening and Closing Pri



Plot Daily Percentage Change of the top 5 curriencies:

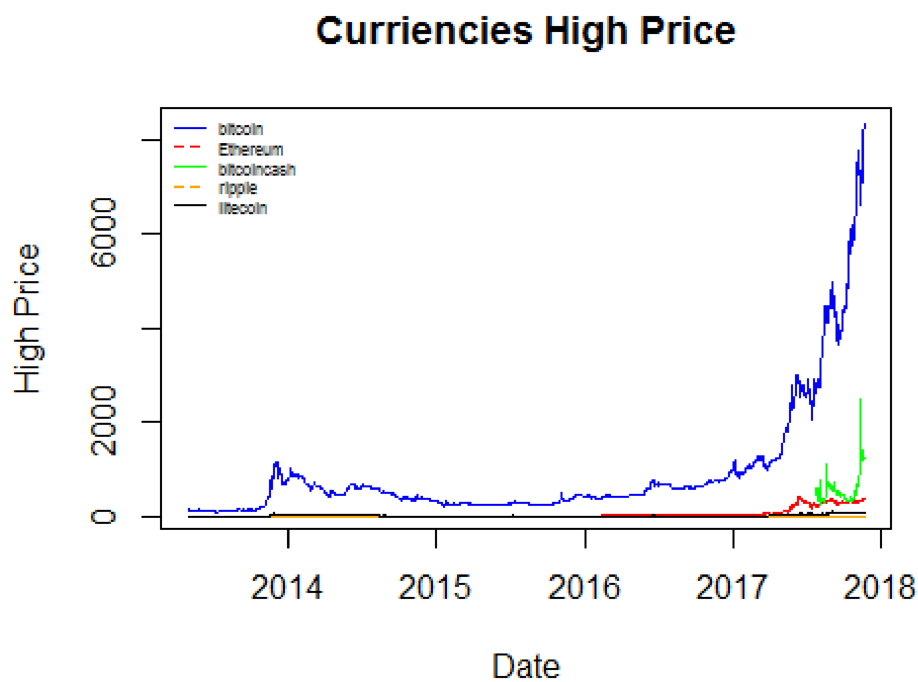
```
plot(df_bitcoin$Date,df_bitcoin$Percentage, type = 'line', col="blue", main =
'Percentage of Closing Price', xlab ='Date', ylab = 'Differences')

## Warning in plot.xy(xy, type, ...): çizim tipi 'line' in ilk karakteri
## kullanılacak

lines(df_ethereum$Date,df_ethereum$Percentage, col="red")
lines(df_bitcoincash$Date,df_bitcoincash$Percentage, col='green')
lines(df_ripple$Date,df_ripple$Percentage, col="orange")
lines(df_litecoin$Date,df_litecoin$Percentage, col='black')
legend("topleft", legend=c("bitcoin", "Ethereum", "bitcoincash", "ripple",
"litecoin"),
       col=c("blue", "red", "green", "orange","black"), lty=1:2, bty ="n" ,
cex=0.5)
```



How did the historical prices / market capitalizations of top 5 currencies (in terms of market capitalization) change over time? a. plot price vs time, each and all b. market caps vs time

Plot Currencis Change over time

```
plot(df_bitcoin$Date,df_bitcoin$High, type = 'line', col="blue", main =
'Curriencies High Price', xlab ='Date', ylab = 'High Price')

## Warning in plot.xy(xy, type, ...): çizim tipi 'line' in ilk karakteri
## kullanılacak

lines(df_ethereum$Date,df_ethereum$High, col="red")
lines(df_bitcoincash$Date,df_bitcoincash$High, col='green')
lines(df_ripple$Date,df_ripple$High, col="orange")
lines(df_litecoin$Date,df_litecoin$High, col='black')
legend("topleft", legend=c("bitcoin", "Ethereum", "bitcoincash", "ripple",
"litecoin"),
       col=c("blue", "red", "green", "orange","black"), lty=1:2, bty ="n" ,
cex=0.5)
```



Add Candle Plot for bit coin And Ethereum Candle Plot for Bitcoin Last 30 days

```
i <- list(line = list(color = '#FFD700'))
d <- list(line = list(color = '#0000ff'))
df_bitcoin_Last30 <- head(df_bitcoin, n = 30)
p <- df_bitcoin_Last30 %>%
  plot_ly(x = df_bitcoin_Last30$Date, type="candlestick",
          open = df_bitcoin_Last30$Open, close = df_bitcoin_Last30$Close,
          high = df_bitcoin_Last30$High, low = df_bitcoin_Last30$Low,
          increasing = i, decreasing = d) %>%
  layout(title = "Candlestick Chart For Bitcoin: Last 30 Days")
print(p)
```

Candle Plot for Whole Data sets for Bitcoin

```
p <- df_bitcoin %>%
  plot_ly(x = df_bitcoin$Date, type="candlestick",
          open = df_bitcoin$Open, close = df_bitcoin$Close,
          high = df_bitcoin$High, low = df_bitcoin$Low) %>%
  add_lines(y = df_bitcoin$Open, line = list(color = 'black', width = 0.75))
%>%
  layout(title= "Bitcoin Candlestick Chart", showlegend = TRUE)
print(p)
```

```
## Warning: 'scatter' objects don't have these attributes: 'open', 'close',
'high', 'low'
## Valid attributes include:
## 'type', 'visible', 'showlegend', 'legendgroup', 'opacity', 'name', 'uid',
```

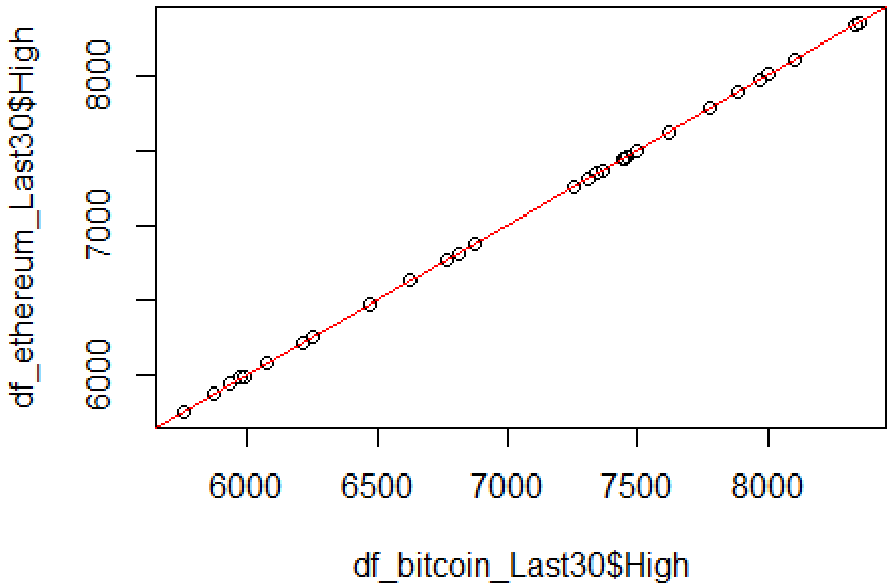
```
'ids', 'customdata', 'hoverinfo', 'hoverlabel', 'stream', 'x', 'x0', 'dx',
'y', 'y0', 'dy', 'text', 'hovertext', 'mode', 'hoveron', 'line',
'connectgaps', 'cliponaxis', 'fill', 'fillcolor', 'marker', 'textposition',
'textfont', 'r', 't', 'error_y', 'error_x', 'xaxis', 'yaxis', 'xcalendar',
'ycalendar', 'idssrc', 'customdatasrc', 'hoverinfosrc', 'xsrc', 'ysrc',
'textsrc', 'hovertextsrc', 'textpositionsrc', 'rsrc', 'tsrc', 'key', 'set',
'frame', 'transforms', '_isNestedKey', '_isSimpleKey', '_isGraticule'
```

Linear regression analysis between bitcoin Hugh Price and Ethereum High Price for Last 30 days.

```
library(dplyr)
df_bitcoin_Last30 <- head(df_bitcoin, n = 30)
df_ethereum_Last30<-head(df_bitcoin, n=30)
df_bitcoin_RegModel3 <- lm(formula = df_bitcoin_Last30$High ~
df_ethereum_Last30$High, data=c(df_bitcoin_Last30,df_ethereum_Last30)) #
regression formula
summary(df_bitcoin_RegModel3) # show regression coefficients table
```

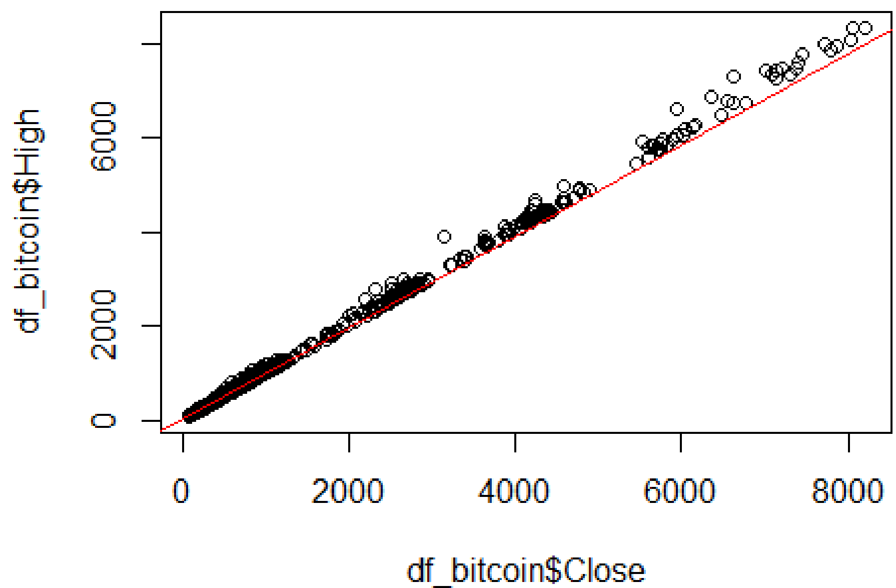
```
## Warning in summary.lm(df_bitcoin_RegModel3): essentially perfect fit:
## summary may be unreliable

##
## 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
```

```
plot(df_bitcoin_Last30$High, df_ethereum_Last30$High)
abline(df_bitcoin_RegModel3,col ='red')
```



As it is seen figure; there is linear relationships between ethereum high price and bitcoin high prices. Linear Regression between bitcoin Closing price and high price

```
df_bitcoin_RegModel <- lm(formula = Close ~ High, data=df_bitcoin) #
regression formula
summary(df_bitcoin_RegModel) # show regression coefficients table

##
## 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

plot(df_bitcoin$Close, df_bitcoin$High)
abline(df_bitcoin_RegModel,col ='red')
```




```
# Summarize and print the results
df_bitcoin_RegModel2 <- lm(Volume ~ Market.Cap + High + Percentage, data =
df_bitcoin)
summary(df_bitcoin_RegModel2) # show regression coefficients table
```

```
##
## 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
```

```r
plot(df_bitcoin_RegModel2, which= 1)
```