Analysis of Cryptocurrency

* Introduction:

Introduced and first documented by Satoshi Nakamoto in 2009,Bitcoin is a form of cryptocurrency –“an electronic payment system based on cryptographic proof”, instead of traditional trust. Interest in bitcoins has increased over the years. At the end of August 2013, the total available Bitcoins were valued at over 1.5 billion USDs. It has the property which could make it important in commerce. Bitcoins as an international payment standard has its benefits,but its volatile price suggests that it may still suffer from problems of transitional currencies. Therefore, Bitcoins could be considered as an exchange rate.

* Cryptography: **“crypt”** means “**hidden” .**It is a method of protecting informations and communications through the use of codes, so that only those for whom the information is intended can read and process it. They are not secured by people or trust but by math.
* Cryptocurrency: A digital asset designed to work as a medium of exchange wherein individual coin ownership records are stored in a *digital ledger* or a computerized database using strong cryptography.
* *Blockchain:* In colonial version it means chain of chains, but in actual context it indicates the way in which digital information (the “block”) stored in a public database (the “chain”). A blockchain is a public ledger of information collected through a network that sits on top of the internet. It is a digital record of transactions. The name comes from its structure, in which individual records, called blocks, are linked together in single list, called a chain. Blockchains are used for recording transactions made with cryptocurrencies, such as Bitcoin, and have many other applications.
* Data Source:

<https://www.kaggle.com/philmohun/cryptocurrency-financial-data>

* Description of the data set:
* VARIABLES:

1. Date: Date of the particular day when stock price is noted.
2. Open: It is price at which the financial security opens in the market when trading begins.
3. High: It is the highest price during that time window, here it is that particular day.
4. Low: It is the lowest price during that time window, here it is that particular day.
5. Close: The price of a financial security at the end of the day’s business in a financial market.
6. Volume: Amount of currency transacted in time window. It is given for different times of the day.
7. Market cap: Market value.

* Goals:

Estimation of missing values.

Statistical analysis of the closing prices of bitcoins through visual as well as summary measures.

To fit a proper distribution to the log-returns.

To model the volatility and forecast the future Value-at-Risks and volatility.

* Code:

# Libraries Used:

library(dplyr)

library(imputeTS)

# Data:

data = read.csv(file.choose())

str(data) # Checking the data types

'data.frame': 28944 obs. of 8 variables:

$ Currency : chr "tezos" "tezos" "tezos" "tezos" ...

$ Date : chr "Dec 04, 2019" "Dec 03, 2019" "Dec 02, 2019" "Dec 01, 2019" ...

$ Open : chr "1.29" "1.24" "1.25" "1.33" ...

$ High : chr "1.32" "1.32" "1.26" "1.34" ...

$ Low : chr "1.25" "1.21" "1.20" "1.25" ...

$ Close : chr "1.25" "1.29" "1.24" "1.25" ...

$ Volume : chr "46,048,752" "41,462,224" "27,574,097" "24,127,567" ...

$ Market.Cap: chr "824,588,509" "853,213,342" "817,872,179" "828,296,390" ...

# Changing the data types:

data$Currency = as.character(data$Currency)

data$Date = as.Date(data$Date, format = "%b %d, %Y")

data$Open = as.numeric(data$Open)

data$High = as.numeric(data$High)

data$Low = as.numeric(data$Low)

data$Close = as.numeric(data$Close)

data$Volume = as.numeric(gsub(",","",data$Volume))

data$Market.Cap = as.numeric(gsub(",","",data$Market.Cap))

data\_str = data[order(data$Date),]

summary(data\_str)

Currency Date Open High Low

Length:28944 Min. :2013-04-28 Min. : 0.0014 Min. : 0.0015 Min. : 0.0012

Class :character 1st Qu.:2014-12-21 1st Qu.: 0.1203 1st Qu.: 0.1252 1st Qu.: 0.1179

Mode :character Median :2016-08-15 Median : 2.7500 Median : 2.8400 Median : 2.6600

Mean :2016-08-15 Mean : 61.2307 Mean : 61.9915 Mean : 60.4507

3rd Qu.:2018-04-10 3rd Qu.: 14.2900 3rd Qu.: 14.8075 3rd Qu.: 13.7400

Max. :2019-12-04 Max. :998.8900 Max. :998.9100 Max. :999.7500

NA's :1243 NA's :1278 NA's :1203

Close Volume Market.Cap

Min. : 0.0014 Min. :0.000e+00 Min. :0.000e+00

1st Qu.: 0.1213 1st Qu.:2.419e+05 1st Qu.:6.345e+07

Median : 2.7400 Median :5.213e+06 Median :3.454e+08

Mean : 61.2191 Mean :8.133e+08 Mean :7.195e+09

3rd Qu.: 14.2900 3rd Qu.:1.555e+08 3rd Qu.:3.422e+09

Max. :999.1800 Max. :5.351e+10 Max. :3.265e+11

NA's :1244

# Missing Values:

apply(is.na(data\_str), 2, any) # Columns containing missing data

Currency Date Open High Low Close Volume Market.Cap

FALSE FALSE TRUE TRUE TRUE TRUE FALSE FALSE

round(sum(complete.cases(data\_str))\*100/nrow(data\_str),4) # Percentage of Data that is free of NA

[1] 95.5846

data\_list = list()

for(i in 1:length(unique(data\_str$Currency)))

data\_list[[i]] = data\_str %>% filter(Currency == unique(data\_str$Currency)[i])

names(data\_list) = unique(data\_str$Currency)

for(i in 1:length(unique(data\_str$Currency)))

if(any(is.na(data\_list[[i]]$Close))==T)

cat(unique(data\_list[[i]]$Currency), i,"\n") # Which currencies have missing data

bitcoin 4

bitcoin-cash 7

ethereum 10

# Estimating missing values:

x = c(1,2,3,5,6,8,9,11,12)

y = c(4,7,10)

for(i in y)

data\_list[[i]]$Close.1 = na\_interpolation(data\_list[[i]]$Close, option = "linear")

for(i in x)

data\_list[[i]]$Close.1 = data\_list[[i]]$Close

# Computation of logreturns and study of the closing prices:

for(i in 1:length(unique(data\_str$Currency)))

for(j in 2:nrow(data\_list[[i]]))

data\_list[[i]]$logreturns[j]=log(abs(data\_list[[i]]$Close[j]-data\_list[[i]]$Close[j-1]))

par(mfrow=c(1,2))

for(i in 1:length(unique(data\_str$Currency))){

hist(data\_list[[i]]$Close, main=paste("Histogram of closing price of",unique(data\_str$Currency)[i]), xlab="Closing Price")

hist(data\_list[[i]]$logreturns, main=paste("Histogram of logreturns of",unique(data\_str$Currency)[i]), xlab="Logreturns")

cat("For",unique(data\_str$Currency)[i],":\n")

print(summary(data\_list[[i]]$logreturns))

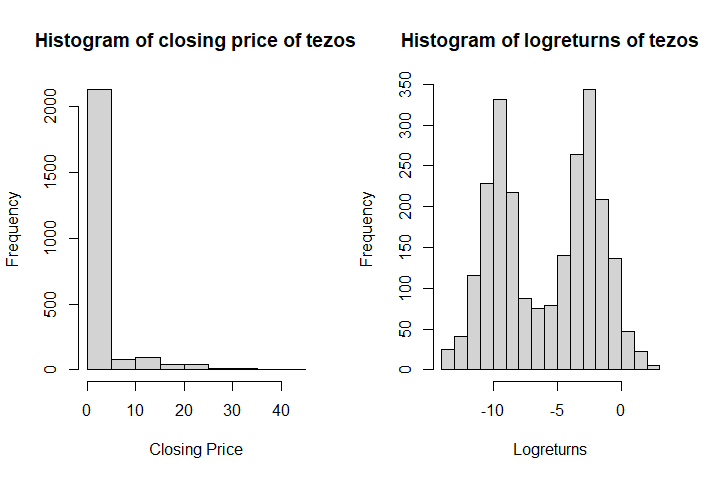
cat("\n\n")

}

For tezos :

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

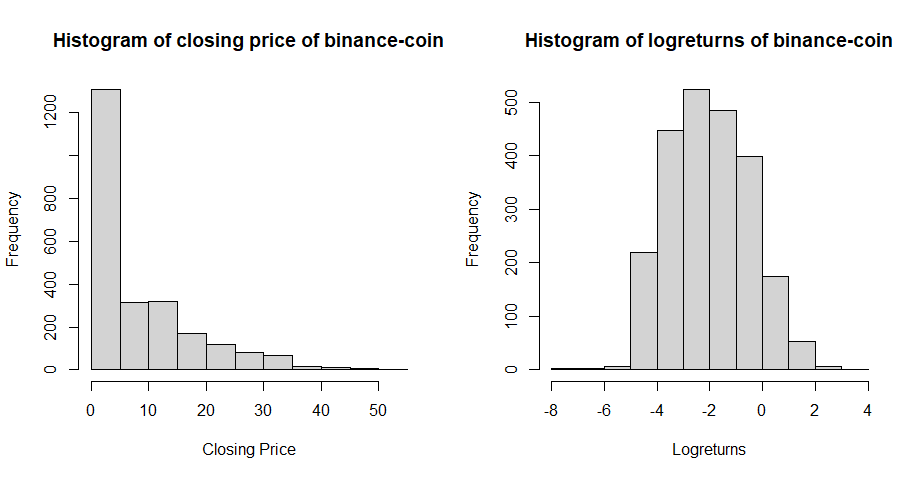
-Inf -9.567 -5.515 -Inf -2.659 2.928 1



For binance-coin :

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

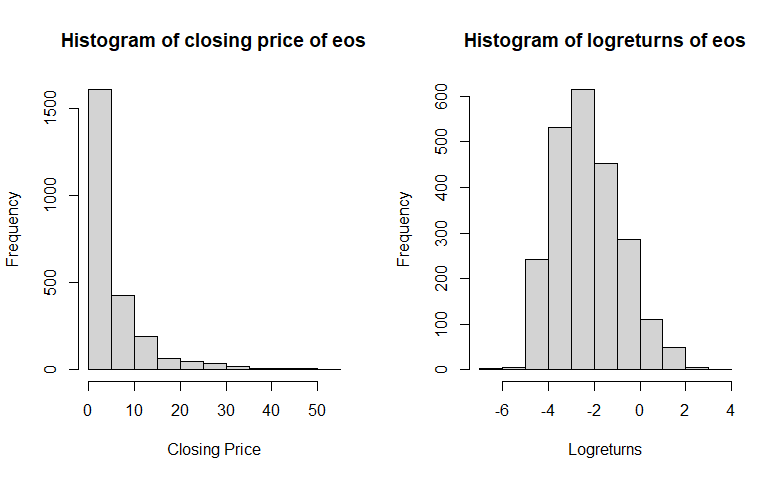
-Inf -3.5066 -2.2073 -Inf -0.9416 3.6722 1



For eos :

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

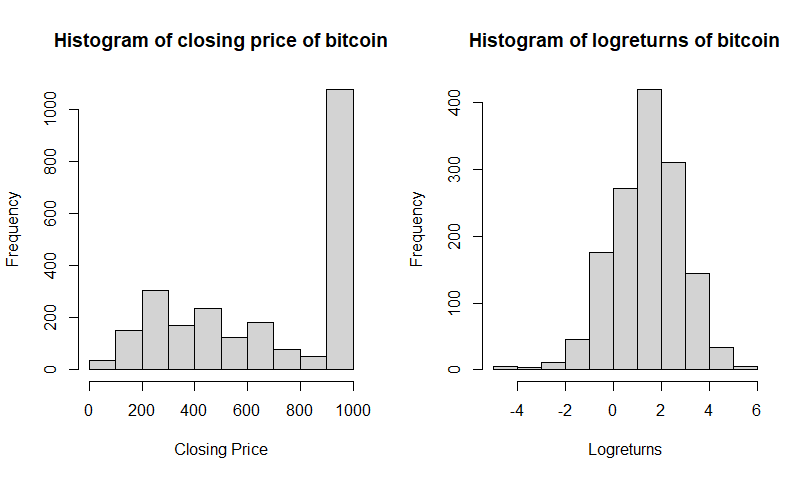
-Inf -3.507 -2.526 -Inf -1.386 3.672 1



For bitcoin :

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

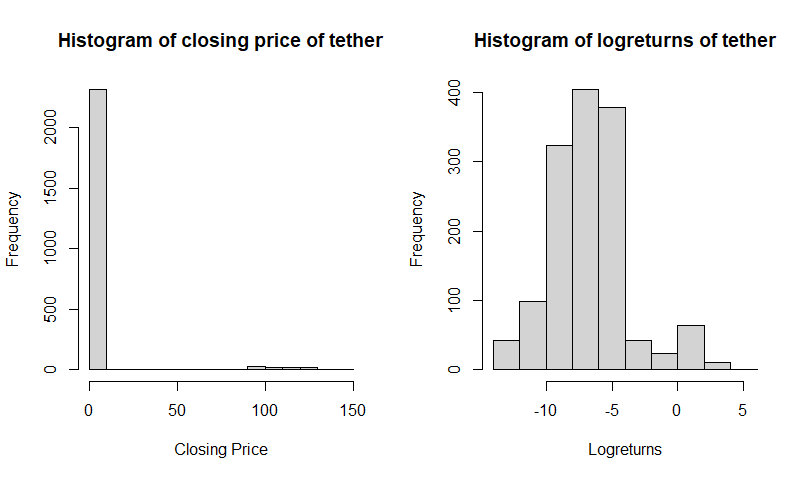
-Inf -Inf -0.08338 -Inf 1.75181 5.35399 1



For tether :

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

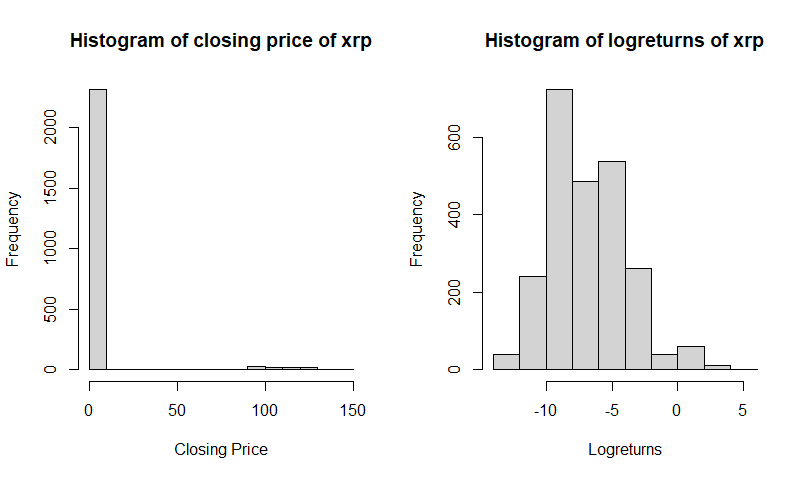
-Inf -Inf -9.567 -Inf -6.573 4.644 1



For xrp :

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

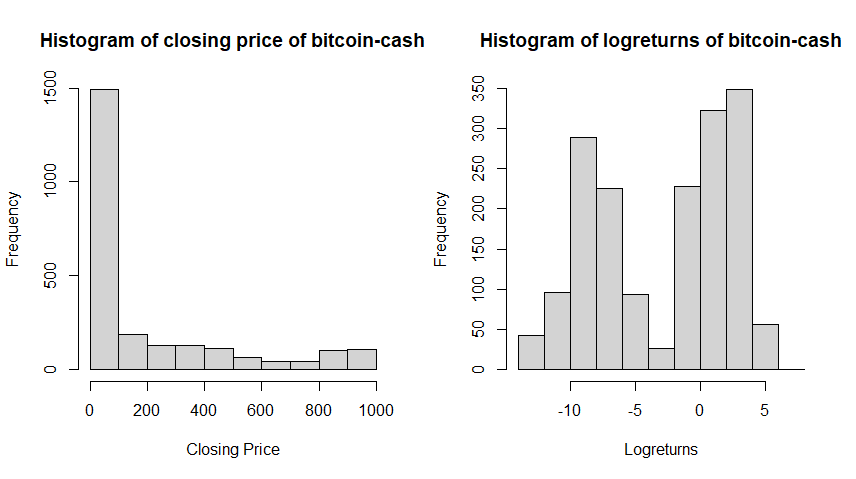
-Inf -9.071 -7.346 -Inf -4.856 4.644 1



For bitcoin-cash :

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

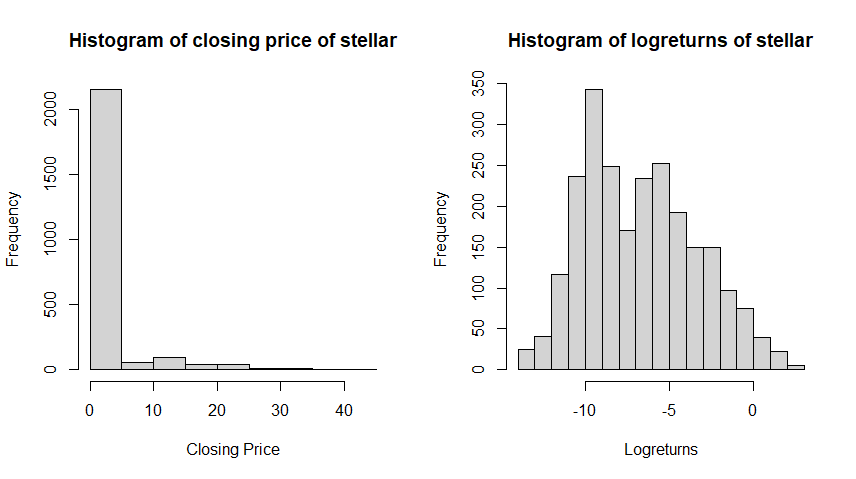
-Inf -Inf -7.376 -Inf 1.017 6.021 1



For stellar :

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

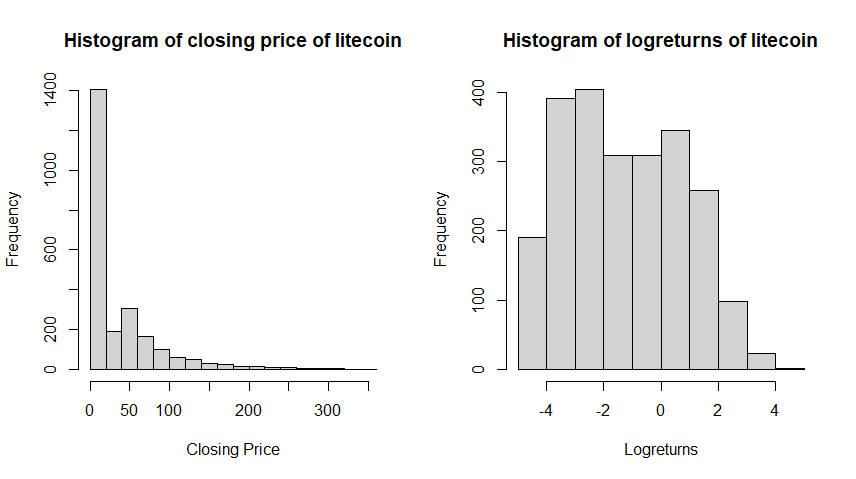
-Inf -9.525 -6.971 -Inf -4.464 2.928 1



For litecoin :

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

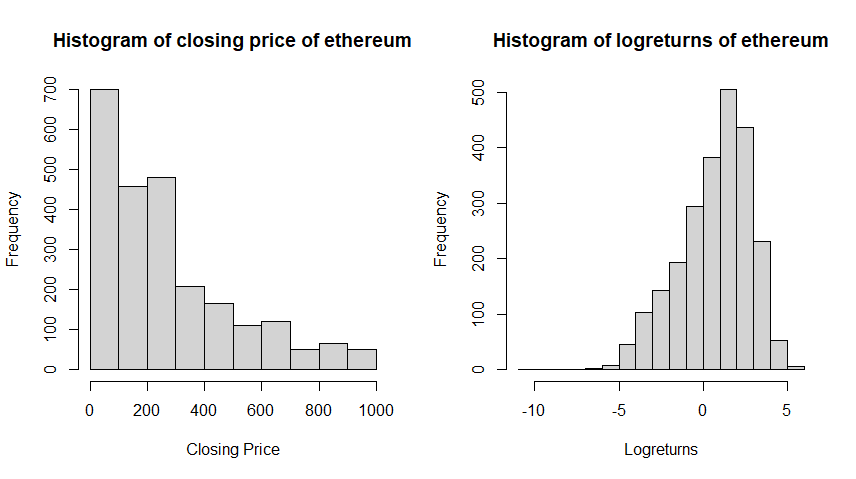
-Inf -3.2189 -1.5606 -Inf 0.3988 4.6246 1



For ethereum :

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

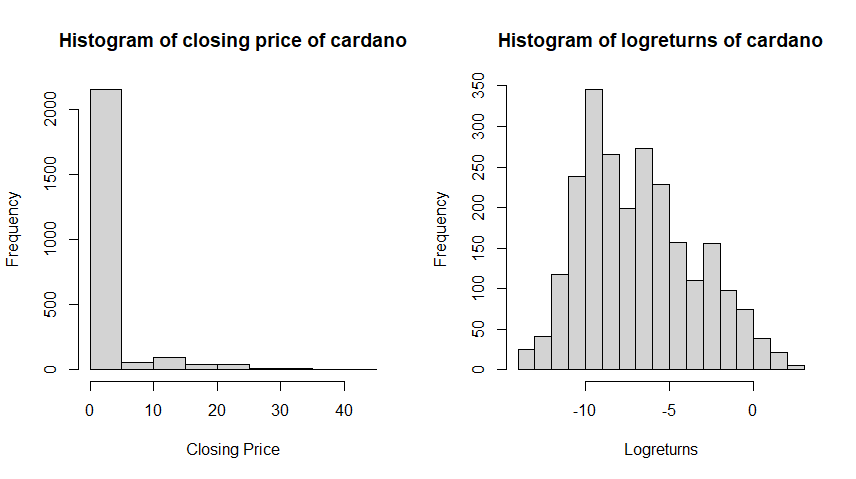
-Inf -0.6733 1.0473 -Inf 2.2370 5.6197 1



For cardano :

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-Inf -9.539 -7.252 -Inf -4.626 2.928 1



For bitcoin-sv :

Min. 1st Qu. Median Mean 3rd Qu. Max. NA's

-Inf -3.5066 -2.0402 -Inf -0.4862 4.5679 1

