Homework1

Code

Hide

```
library("quantmod")
library("e1071")
require("quantmod")
library(ggplot2)
```

Hide

```
SP500 <-read.csv("SP500constituentsSecid2016sector.csv", stringsAsFactors = FALSE, strip.white = TRUE)
SP500['Ticker']
```

Hide

```
SP500 <- read.csv("SP500constituentsSecid2016sector.csv", header=TRUE, sep=",")
sp<- new.env()
listEnterprise <- c()
for(i in SP500$Ticker){
  cat("Downloading time series for symbol '", i, "' ...\n", sep = "")
  status <- tryCatch(getSymbols(i, env = sp, src = "yahoo", from = as.Date("1960-01-01")), error = identity)
  if(inherits(status, "error")) cat("Symbol '", i, "' not downloadable!\n", sep = "")
  else{
    listEnterprise <- c(listEnterprise, i)
  }
}
```

Question 1 Building of the matrix with the mean, the variance, the skewness and kurtosis

Hide

```
moyenne <- c()
variance <- c()
skew <- c()
kurto <- c()
for(i in listEnterprise){
  if (i!= "BHI"){
    a = getSymbols(i, auto.assign = FALSE, from = as.Date("1960-01-01"))
    adjusted = Ad(a)
    return = exp(diff(log(adjusted))) - 1
    return = return[2:length(return)]
    moyenne <- c(moyenne,mean(return))
    variance <- c(variance,sqrt(var(return)))
    skew <- c(skew, skewness(return))
    kurto <- c(kurto, kurtosis(return))
  }
}
```

Nonparametric density estimates

Hide

```
cat("The mean of the mean of comapanys is",mean(moyenne),"whereas its medianne is equal to", median(moyenne)
)
```

```
The mean of the mean of comapanys is 0.0008793301 whereas its medianne is equal to 0.0008074429
```

Hide

```
cat("The mean of the variance of comapanys is",mean(variance),"whereas its medianne is equal to", median(var
iance))
```

```
The mean of the variance of comapanys is 0.02410942 whereas its medianne is equal to 0.02173297
```

Hide

```
cat("The mean of the skewness of comapanys is",mean(skew),"whereas its medianne is equal to", median(skew))
```

```
The mean of the skewness of comapanys is 0.8347739 whereas its medianne is equal to 0.2738966
```

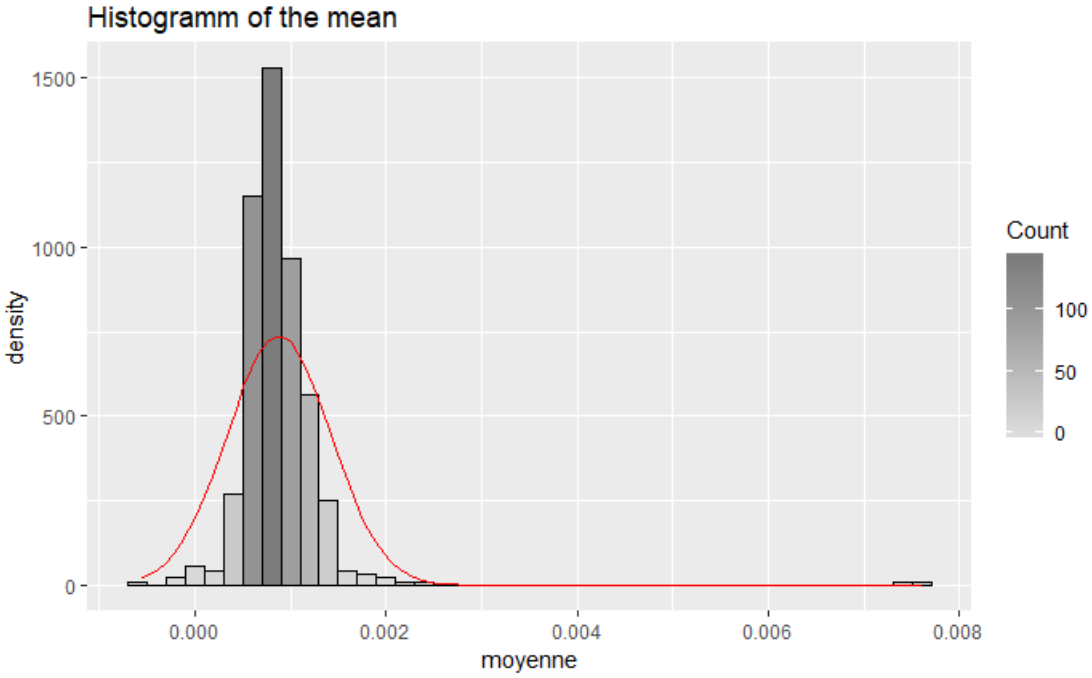
Hide

```
cat("The mean of the kurtosis of comapanys is",mean(kurto),"whereas its medianne is equal to", median(kurto)
)
```

```
The mean of the kurtosis of comapanys is 53.15956 whereas its medianne is equal to 11.87869
```

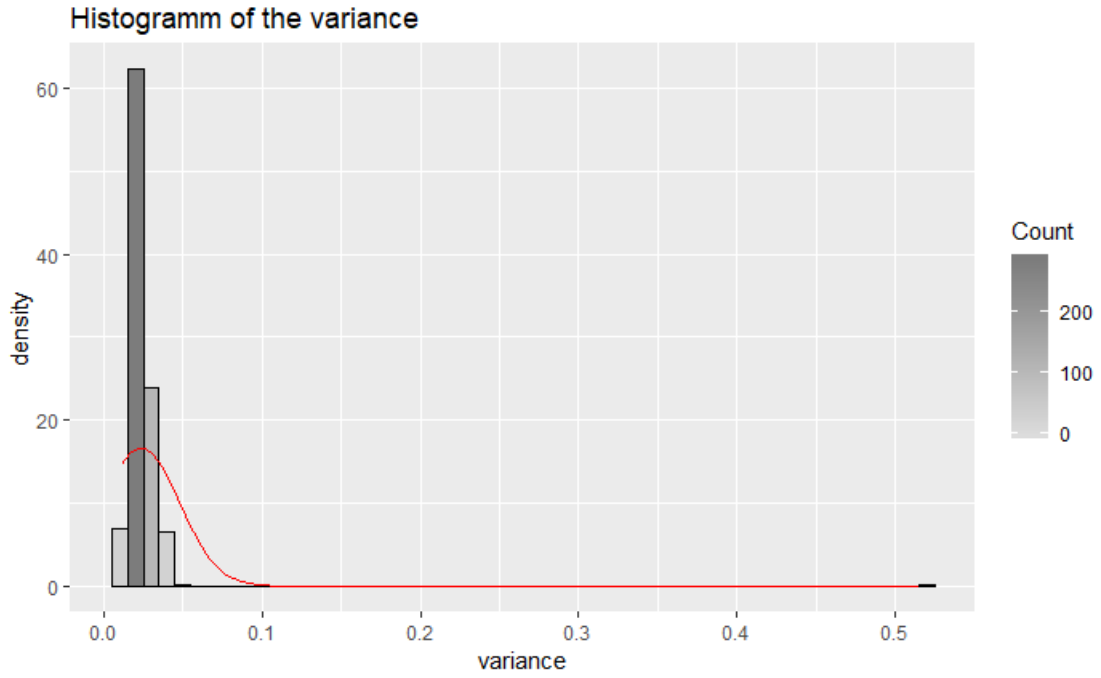
Hide

```
moyennedf = data.frame(moyenne)
gg = ggplot(data =moyennedf, aes(x=moyenne)) + geom_histogram(binwidth=0.0002, colour="black", aes(y=..densi
ty.., fill=..count..))
gg = gg + scale_fill_gradient("Count", low="#DCDCDC", high="#7C7C7C")
gg = gg + stat_function(fun=dnorm, color="red",args=list(mean=mean(moyenne), sd=sd(moyenne)))
gg = gg + ggtitle("Histogramm of the mean")
gg
```



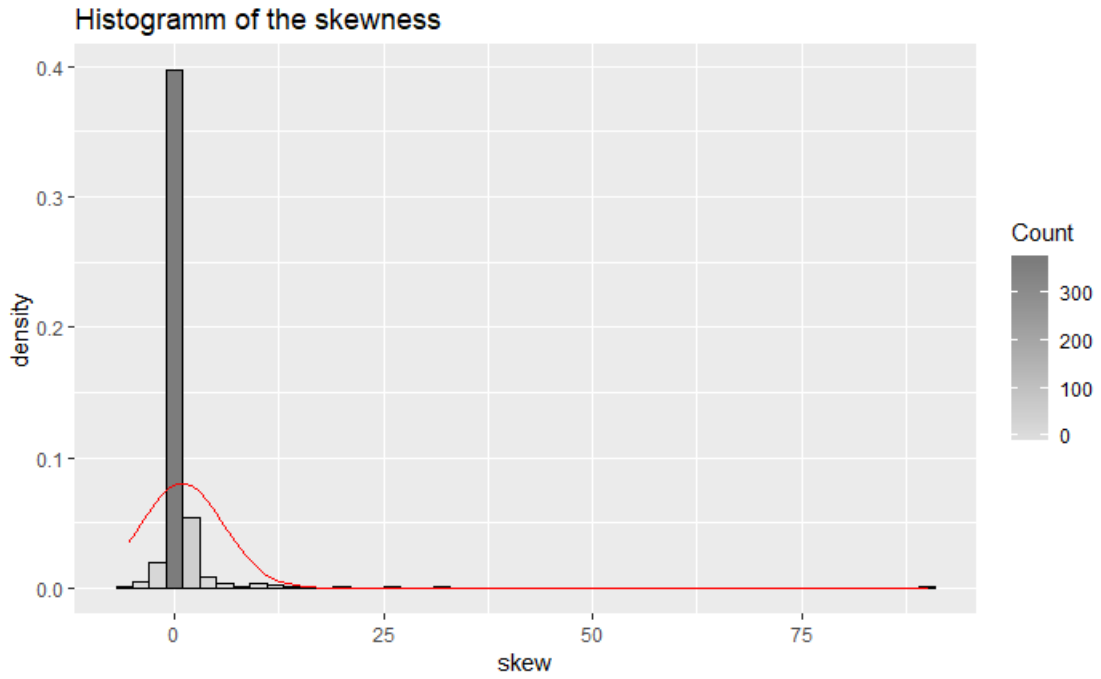
Hide

```
variancedf = data.frame(variance)
gg = ggplot(data =variancedf, aes(x=variance)) + geom_histogram(binwidth=0.01, colour="black", aes(y=..densi
ty.., fill=..count..))
gg = gg + scale_fill_gradient("Count", low="#DCDCDC", high="#7C7C7C")
gg = gg + stat_function(fun=dnorm, color="red",args=list(mean=mean(variance), sd=sd(variance)))
gg = gg + ggtitle("Histogramm of the variance")
gg
```



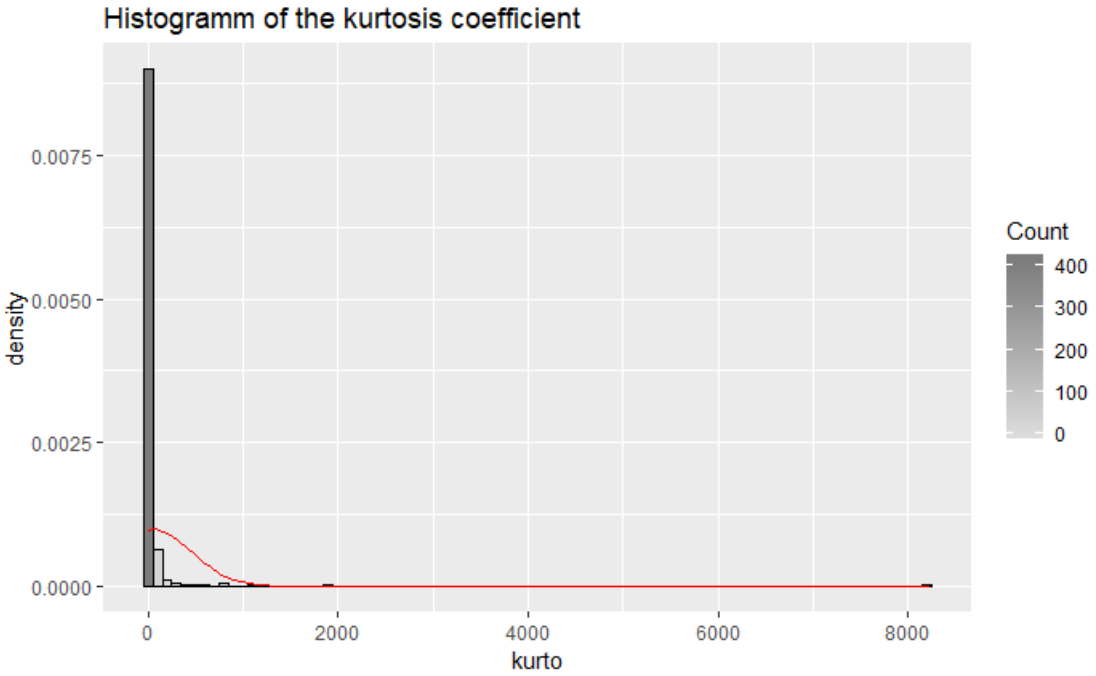
Hide

```
skewdf = data.frame(skew)
gg = ggplot(data =skewdf, aes(x=skew)) + geom_histogram(binwidth=2.0, colour="black", aes(y=..density.., fil
l=..count..))
gg = gg + scale_fill_gradient("Count", low="#DCDCDC", high="#7C7C7C")
gg = gg + stat_function(fun=dnorm, color="red",args=list(mean=mean(skew), sd=sd(skew)))
gg = gg + ggtitle("Histogramm of the skewness")
gg
```



Hide

```
kurtodf = data.frame(kurto)
gg = ggplot(data =kurtodf, aes(x=kurto)) + geom_histogram(binwidth=100, colour="black", aes(y=..density.., f
ill=..count..))
gg = gg + scale_fill_gradient("Count", low="#DCDCDC", high="#7C7C7C")
gg = gg + stat_function(fun=dnorm, color="red",args=list(mean=mean(kurto), sd=sd(kurto)))
gg = gg + ggtitle("Histogramm of the kurtosis coefficient")
gg
```



Skewness evaluates the asymmetry of a distribution here there are 3 left-skewed distributions mean variance and kurtosis meaning a negative skewness. Skewness as a variable has on the contrary a null skewness fitting the normal distribution. Kurtosis evaluates the spread of extreme values with respect to the normal distribution and is regarded as the flatness compared to it. the kurtosis of the four variables are hence all negative as it is seen that they have a higher spike.

Question2

Hide

```
moyenne2 <- c()
variance2 <- c()
skew2 <- c()
kurto2 <- c()
for(i in listEnterprise[1:7]){
  a = getSymbols(i, auto.assign = FALSE, from = as.Date("2007-01-01"))
  adjusted = Ad(a)
  return = exp(diff(log(adjusted))) - 1
  moyenne2 <- c(moyenne,mean(return))
  variance2 <- c(variance,sqrt(var(return)))
  skew2 <- c(skew, skewness(return))
  kurto2 <- c(kurto, kurtosis(return))
}
moyenne2 = na.omit(moyenne2)
variance2 = na.omit(variance2)
skew2 = na.omit(skew2)
kurto2 = na.omit(kurto2)
SP500Matrix2 <- c(moyenne2, variance2, skew2, kurto2)
SP500Matrix2 <- matrix(SP500Matrix2, byrow=TRUE, nrow = 4)
```

Question3

Hide

```
moyenne3 <- c()
variance3 <- c()
skew3 <- c()
kurto3 <- c()
for(i in listEnterprise){
  if(i !="ABBV" & i !="ALLE" & i !="AWK" & i !="BHI" & i !="AVGO" & i !="CBOE"){
    a = getSymbols(i, auto.assign = FALSE, from = "1960-01-01", to = "2007-01-01")
    adjusted = Ad(a)
    return = exp(diff(log(adjusted))) - 1
    return = return[2:length(return)]
    moyenne3 <- c(moyenne3,mean(return))
    variance3 <- c(variance3,sqrt(var(return)))
    skew3 <- c(skew3, skewness(return))
    kurto3 <- c(kurto3, kurtosis(return))
  }
}
```

```
恫拖getSymbols恫柞 currently uses auto.assign=TRUE by default, but will
use auto.assign=FALSE in 0.5-0. You will still be able to use
恫拖loadSymbols恫柞 to automatically load data. getOption("getSymbols.env")
and getOption("getSymbols.auto.assign") will still be checked for
alternate defaults.

This message is shown once per session and may be disabled by setting
options("getSymbols.warning4.0"=FALSE). See ?getSymbols for details.

AAPL contains missing values. Some functions will not work if objects contain missing values in the middle o
f the series. Consider using na.omit(), na.approx(), na.fill(), etc to remove or replace them.BAX contains m
issing values. Some functions will not work if objects contain missing values in the middle of the series. C
onsider using na.omit(), na.approx(), na.fill(), etc to remove or replace them.
```

Hide

```
moyennedf3 = data.frame(moyenne3)
gg = ggplot(data =moyennedf3, aes(x=moyenne3)) + geom_histogram(binwidth=5, colour="black", aes(y=..density.
., fill=..count..))
gg = gg + scale_fill_gradient("Count", low="#DCDCDC", high="#7C7C7C")
gg = gg + stat_function(fun=dnorm, color="red",args=list(mean=mean(moyenne3), sd=sd(moyenne3)))
gg = gg + ggtitle("Histogramm of the mean before 2007")
gg
```

Hide

```
variancedf3 = data.frame(variance3)
gg = ggplot(data =variancedf3, aes(x=variance3)) + geom_histogram(binwidth=10, colour="black", aes(y=..densi
ty.., fill=..count..))
gg = gg + scale_fill_gradient("Count", low="#DCDCDC", high="#7C7C7C")
gg = gg + stat_function(fun=dnorm, color="red",args=list(mean=mean(variance3), sd=sd(variance3)))
gg = gg + ggtitle("Histogramm of the variance before 2007")
gg
```

Hide

```
skewdf3 = data.frame(skew3)
gg = ggplot(data =skewdf3, aes(x=skew3)) + geom_histogram(binwidth=0.1, colour="black", aes(y=..density.., f
ill=..count..))
gg = gg + scale_fill_gradient("Count", low="#DCDCDC", high="#7C7C7C")
gg = gg + stat_function(fun=dnorm, color="red",args=list(mean=mean(skew3), sd=sd(skew3)))
gg = gg + ggtitle("Histogramm of the skewness before 2007")
gg
```

Hide

```
kurtodf3 = data.frame(kurto3)
gg = ggplot(data =kurtodf3, aes(x=kurto3)) + geom_histogram(binwidth=0.2, colour="black", aes(y=..density..,
fill=..count..))
gg = gg + scale_fill_gradient("Count", low="#DCDCDC", high="#7C7C7C")
gg = gg + stat_function(fun=dnorm, color="red",args=list(mean=mean(kurto3), sd=sd(kurto3)))
gg = gg + ggtitle("Histogramm of the kurtosis coefficient before 2007")
gg
```

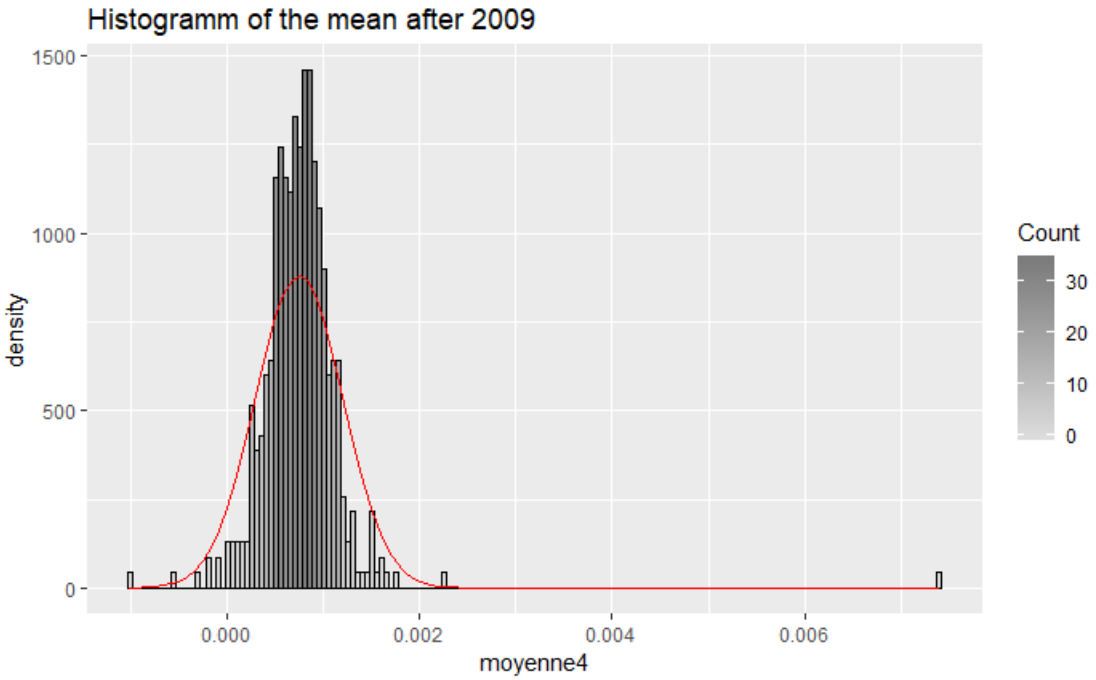
Question4

Hide

```
moyenne4 <- c()
variance4 <- c()
skew4 <- c()
kurto4 <- c()
for(i in listEnterprise){
  a = getSymbols(i, auto.assign = FALSE, from = as.Date("2009-01-01"))
  adjusted = Ad(a)
  return = exp(diff(log(adjusted))) - 1
  return = return[2:length(return)]
  moyenne4 <- c(moyenne4,mean(return))
  variance4 <- c(variance4,sqrt(var(return)))
  skew4 <- c(skew4, skewness(return))
  kurto4 <- c(kurto4, kurtosis(return))
}
```

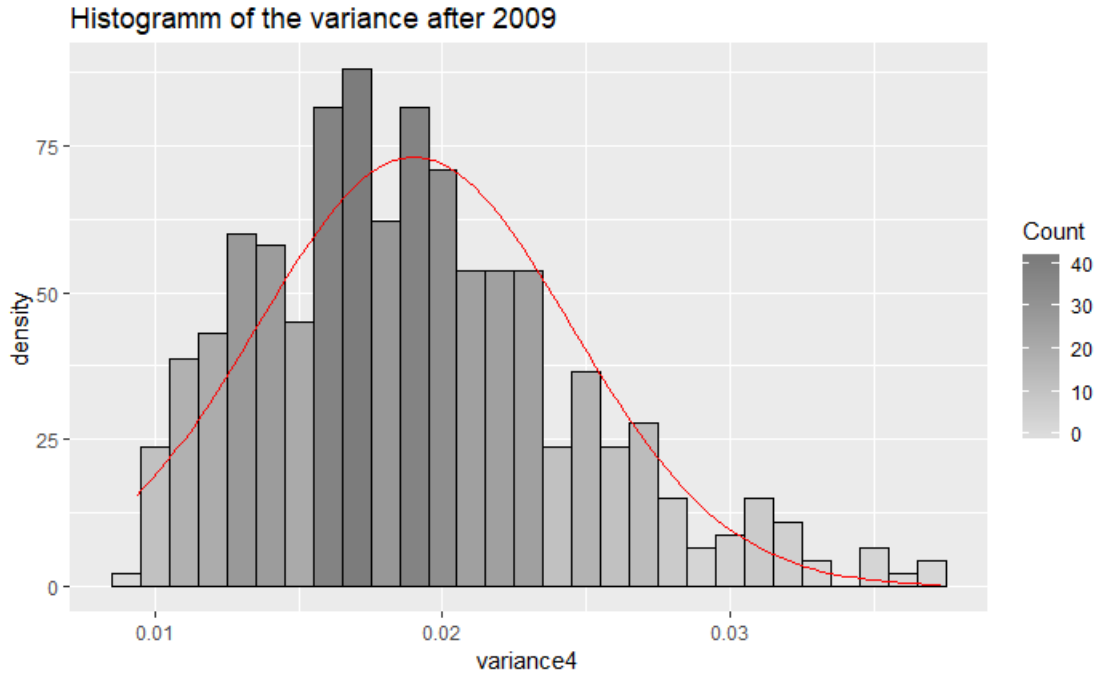
Hide

```
moyennedf4 = data.frame(moyenne4)
gg = ggplot(data =moyennedf4, aes(x=moyenne4)) + geom_histogram(binwidth=0.00005, colour="black", aes(y=..de
nsity.., fill=..count..))
gg = gg + scale_fill_gradient("Count", low="#DCDCDC", high="#7C7C7C")
gg = gg + stat_function(fun=dnorm, color="red",args=list(mean=mean(moyenne4), sd=sd(moyenne4)))
gg = gg + ggtitle("Histogramm of the mean after 2009")
gg
```



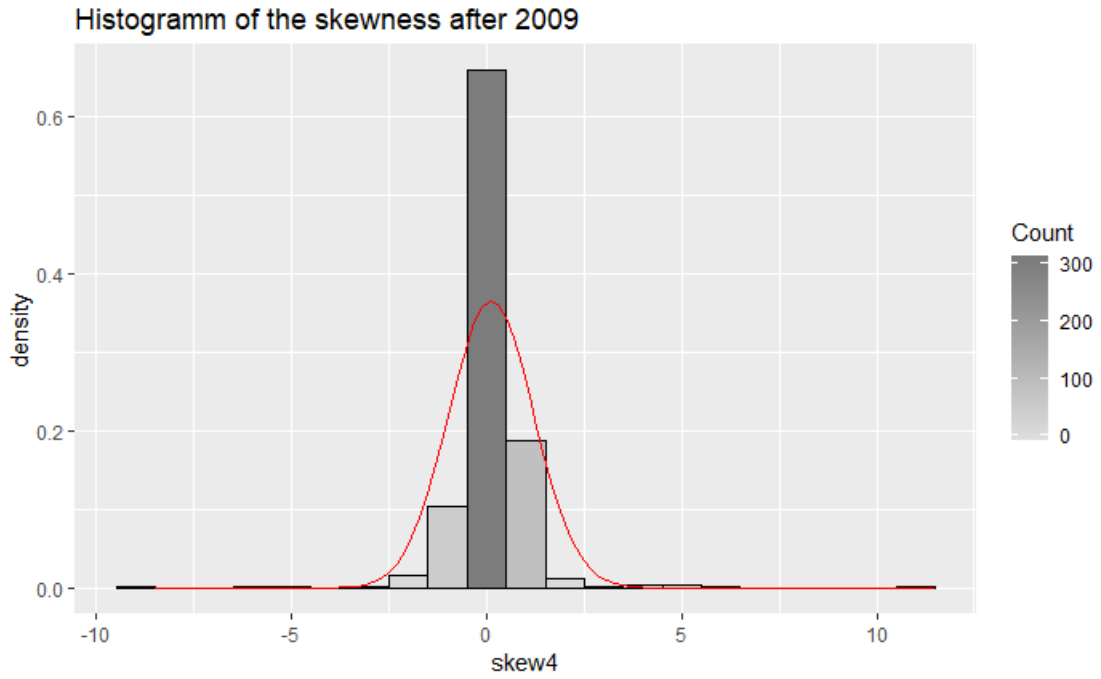
Hide

```
variancedf4 = data.frame(variance4)
gg = ggplot(data =variancedf4, aes(x=variance4)) + geom_histogram(binwidth=0.001, colour="black", aes(y=..de
nsity.., fill=..count..))
gg = gg + scale_fill_gradient("Count", low="#DCDCDC", high="#7C7C7C")
gg = gg + stat_function(fun=dnorm, color="red",args=list(mean=mean(variance4), sd=sd(variance4)))
gg = gg + ggtitle("Histogramm of the variance after 2009")
gg
```



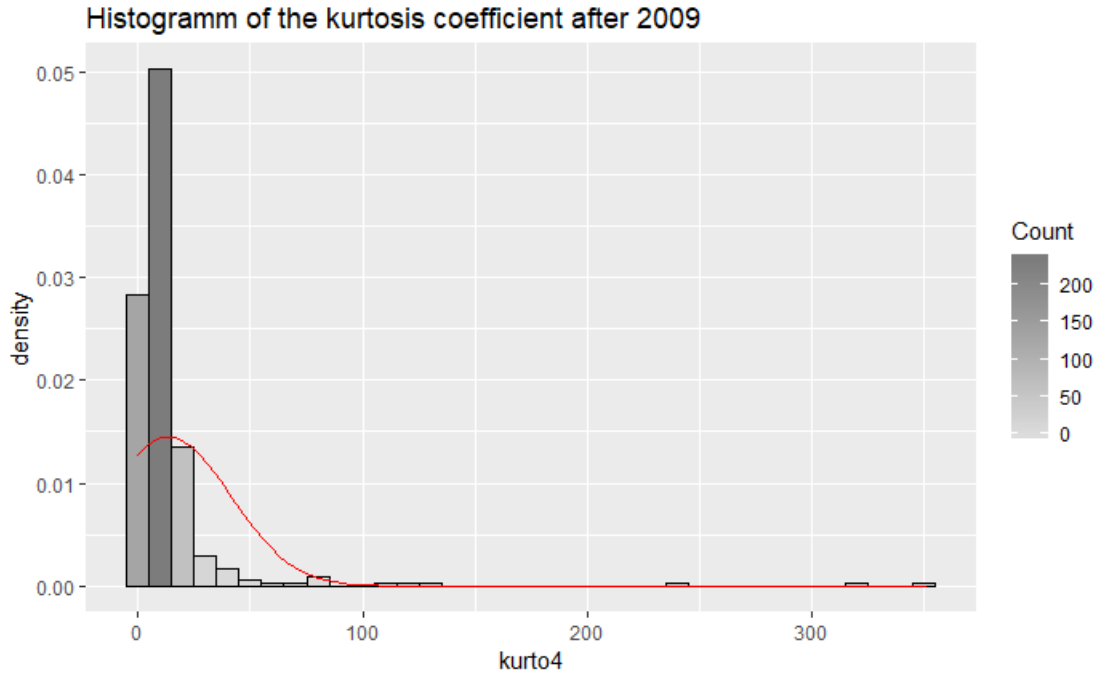
Hide

```
skewdf4 = data.frame(skew4)
gg = ggplot(data =skewdf4, aes(x=skew4)) + geom_histogram(binwidth=1.0, colour="black", aes(y=..density.., f
ill=..count..))
gg = gg + scale_fill_gradient("Count", low="#DCDCDC", high="#7C7C7C")
gg = gg + stat_function(fun=dnorm, color="red",args=list(mean=mean(skew4), sd=sd(skew4)))
gg = gg + ggtitle("Histogramm of the skewness after 2009")
gg
```



Hide

```
kurtodf4 = data.frame(kurto4)
gg = ggplot(data =kurtodf4, aes(x=kurto4)) + geom_histogram(binwidth=10, colour="black", aes(y=..density..,
fill=..count..))
gg = gg + scale_fill_gradient("Count", low="#DCDCDC", high="#7C7C7C")
gg = gg + stat_function(fun=dnorm, color="red",args=list(mean=mean(kurto4), sd=sd(kurto4)))
gg = gg + ggtitle("Histogramm of the kurtosis coefficient after 2009")
gg
```



We see that we have We see that log returns have similar variances after 2009 and on all the period.

Question5