

Computer Assignment 1

Julia Holmgren & Filip Axelsson

2021-04-11

Uppgift 1

I första uppgiften läser vi in data från två aktier, de utvalda aktierna blev som kommer arbetas med är Apple Inc. och Microsoft Corp. Detta är två teknikatwier, där Apple är ett amerikanskt data- och hemelektronikbolag som desginar och producerar mobiler, datorer, surfplattor och klockor. De mest kända varumärknen är till exempel Macbook, Iphone, Ipad och Apple watch. Den andra aktien, Microsoft är ett globalt IT-bolag som utvecklar, desginar och producerar ett utbud av programvaruprodukter, både hård och mjukvara för den globala datamarknaden.

Genom att först undersöka Apple-aktien kan det från datamängden observeras att den sensat stängde på en kurs som motsvarar 133.00 dollar datumet 2021-04.11 och hade ATH ("all time high") i denna period 2021-01-26 vilket motsvarade en kurs på 143.16 dollar. I *Figur 1* kan stängningspriset för Apple analyseras under 10 års perioden, utifrån grafen kan det observeras att stängningspriset tenderar att stiga, kurvan liknar en exponentiell funktion.

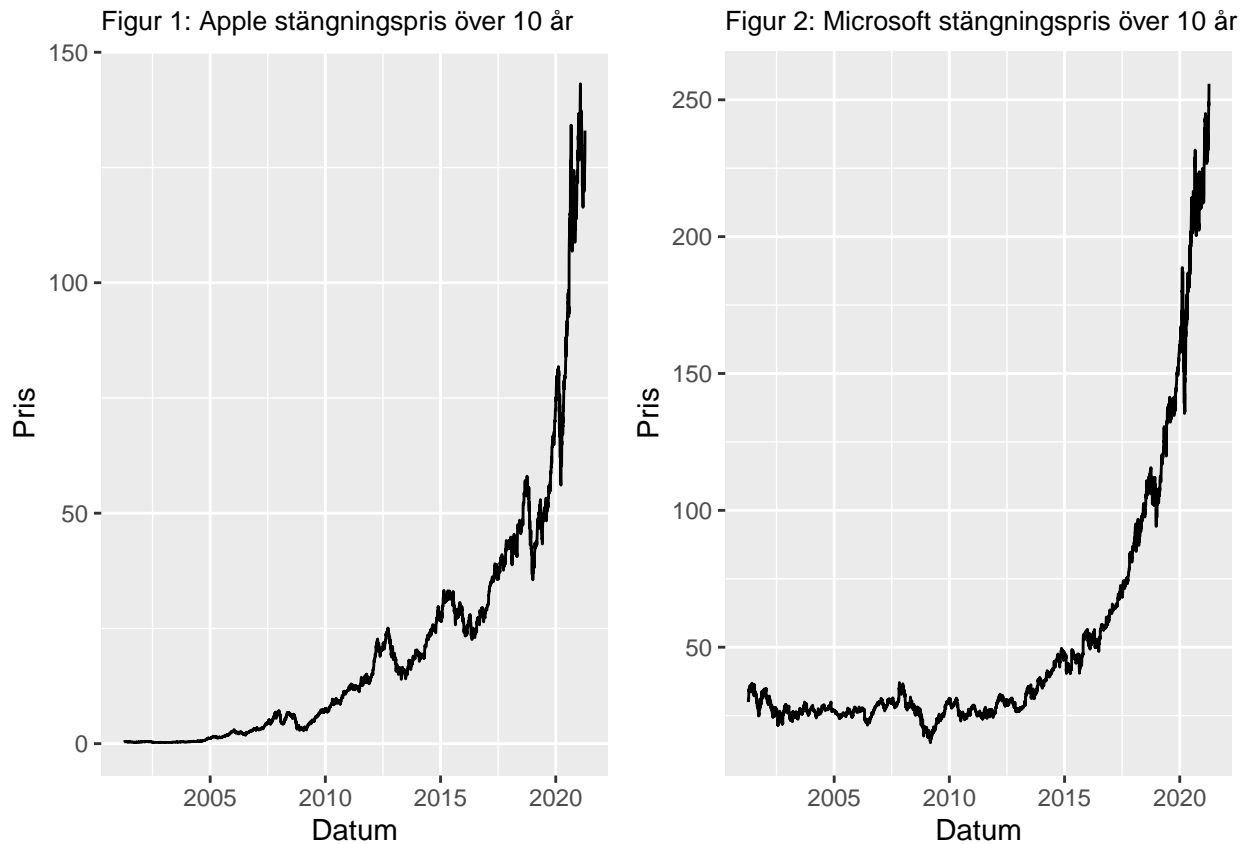
Vidare genom att analysera Microsoft-aktien kan det observeras den sensaste stängningskursen 2021-04-09 motsvarade en kurs på 255.85 dollar, vilket även för ATH aktien under denna perioden. I *Figur 2* visas stängningspriset för Microsoft under en 10 års period, genom analysera denna graf kan slutatsen att trenden för stängningspriser verkar vara positivt, det vill säga priset på aktien stiger över tid. Precis som för Apple liknar även kurvan för Microsoft en exponentiell funktion.

```
apple<-read_csv("apple_stock.csv")
MSFT<-read_csv("MSFT_stock.csv")

plot_A<-ggplot(apple,aes(x=Date,y=Close)) +
  geom_line() +
  ggtitle("Figur 1: Apple stängningspris över 10 år") +
  labs(
    x= "Datum",
    y= "Pris"
  ) +
  theme(plot.title = element_text(size = 10))

plot_M<-ggplot(MSFT,aes(x=Date,y=Close)) +
  geom_line()+
  ggtitle("Figur 2: Microsoft stängningspris över 10 år") +
  labs(
    x= "Datum",
    y= "Pris"
  ) +
  theme(plot.title = element_text(size = 10))
```

```
cowplot::plot_grid(plot_A, plot_M)
```



Uppgift 2

Genom att använda formeln $\log\left(\frac{S(T)}{S(0)}\right)$ kan en daily log return beräknas, detta genom att utgå ifrån att dag 1 får en daily log-return på 0. Denna beräkning kommer utföras på aktien Apple. I *Figur 3* visas resultatet utav daily log-return för Apple, utifrån detta kan det analyseras att aktien verkar ha en stabil uppgång med vissa perioder där det finns tecken på frekventa svängingar i priset.

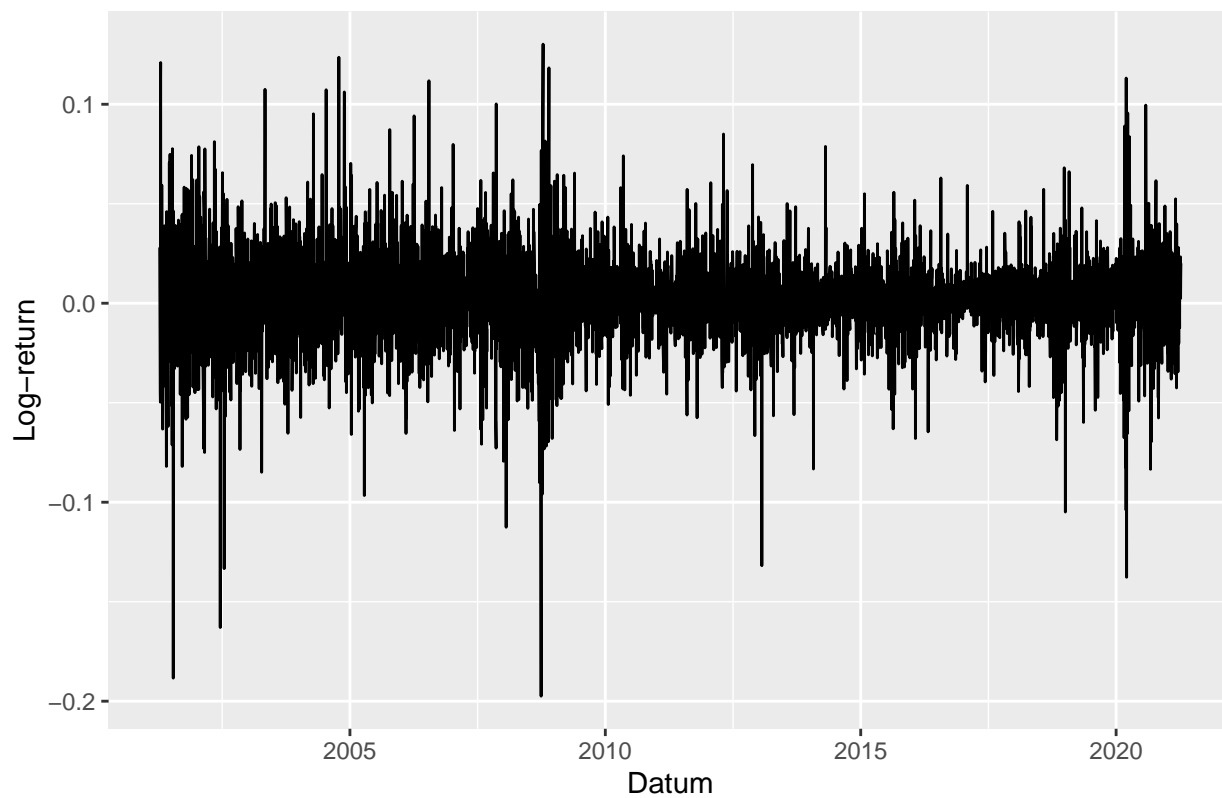
Nästa steg blir att beräkna one-step autocorrelation utav daily log-returns, detta visas i *Figur 4*, genom att analysera denna graf kan slutsatsen att värdet för lag 1 blir -0.0242 och detta innebär att R_t är svårt att bestämma med hjälp utav historisk data. Det går även göra slutsatsen att när antalet lags ökar konvergerar autocorrelationen mot 0. Detta innebär när antalet lags ökar genererar det att det nuvarande värdet beror mindre på det dåvarande värdet, med andra ord, aktiens historiska värden påverkar mindre vad värdet på aktien är idag.

```
apple <- apple %>%
  mutate(daily_return = log((Close)/lag(Close)), default = first(log(Close)),
         share_name = "apple")
apple <- apple[-1,]

ggplot(apple, aes(x=Date, y=daily_return)) +
  geom_line() +
  labs(title = "Figur 3: Apple's daily log-returns över 10 år",
       x = "Datum",
```

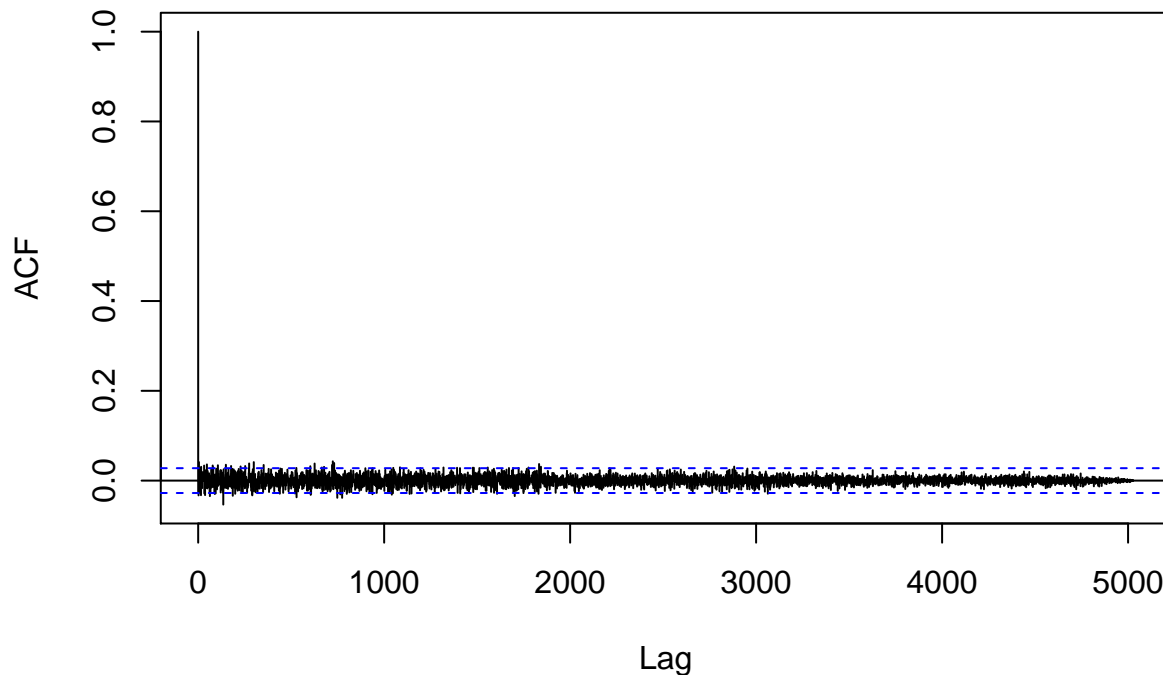
```
y= "Log-return"  
)
```

Figur 3: Apple's daily log-returns över 10 år



```
acf(apple$daily_return, lag.max = length(apple$daily_return), main = "Figur 4: Autocorrelation för Appl
```

Figur 4: Autocorrelation för Apple



Uppgift 3

Genom att använda den givna koden nedanför kan QQ-plots för både t-fördelningen och normalfördelningen produceras. Det kan även konstueras ett histogram där respektive täthetsfunktion är anpassad, den blåa linjen representerar täthetsfunktionen för t-fördelningen och den svarta linjen representerar täthetsfunktionen för normalfördelningen. Genom att analysera hur dessa grafer förändrar sig givet antalet frihetsgrader kan slutsatsen att redan vid ett lågt antal frihetsgrader så är normalfördelningen bättre anpassad till data (frihetsgrader mellan 1 och 3). T-fördelningen börjar sedan att bli bättre anpassad mot data när antalet frihetsgrader är till exempel mellan 3 och 15. Dock ska det noteras att när antalet frihetsgrader ökar kommer t-fördelningen succesivt konvergera mot att likna mer och mer normalfördelningen. Utifrån detta anses normalfördelning vara den mest lämpliga fördelningen vid höga frihetsgrader och när frihetsgraderna är under 3. Tillfället vid ett lägre antal frihetsgrader men som är större än 3 är t-fördelningen bättre anpassad mot data.

```
library(MASS)

##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##   select
x = apple$daily_return

#t-distribution is fitted to the data, with good initial guesses given

df = 7
```

```

param=fitdistr(x, "t", df=df, start = list(m=mean(x),s=sd(x)), lower=c(-1, 0.001,1))
#c(-1, 0.001,1)

#mean, scaling parameter and degrees of freedom are identified from the list "param"
mean=param[[1]][1]
sp=param[[1]][2]

#a histogram is plotted, scaled so that it forms a probability distribution
#note that here we specify how much of the y-axis is shown by specifying ylim.
# This is not necessary for the code to work.
hist(x,prob=TRUE,ylim=c(0,30), xlab = "Log-return", main = "Histogram för Log-returns
      \n med täthetsfunktionen för t-fördelningen och normalfördelningen")

#a suitable x-axis for the density function is created
x_axis = seq(min(x),max(x),length.out=100)

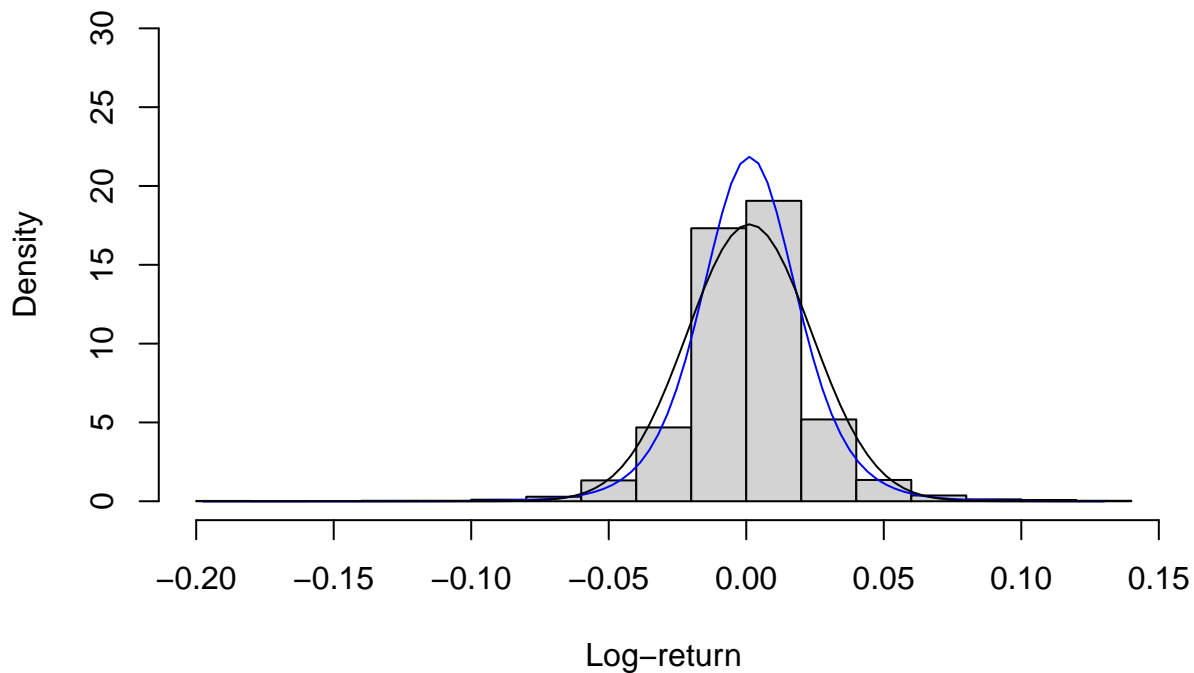
#density function with suitable parameters is plotted (r function for t-density
#function is dt() )
lines(x_axis,(1/sp)*dt((x_axis-mean)/sp,df),col="blue")

# Normal distribution
normal_param <- fitdistr(x, "normal")
normal_mean <- normal_param[[1]][1]
normal_sp <- normal_param[[1]][2]
lines(x_axis,(1/normal_sp)*dnorm((x_axis-normal_mean)/normal_sp),col="black")

```

Histogram för Log-returns

med täthetsfunktionen för t-fördelningen och normalfördelningen



#Run the code below to make a qq-plot using simulated data from t-distribution:

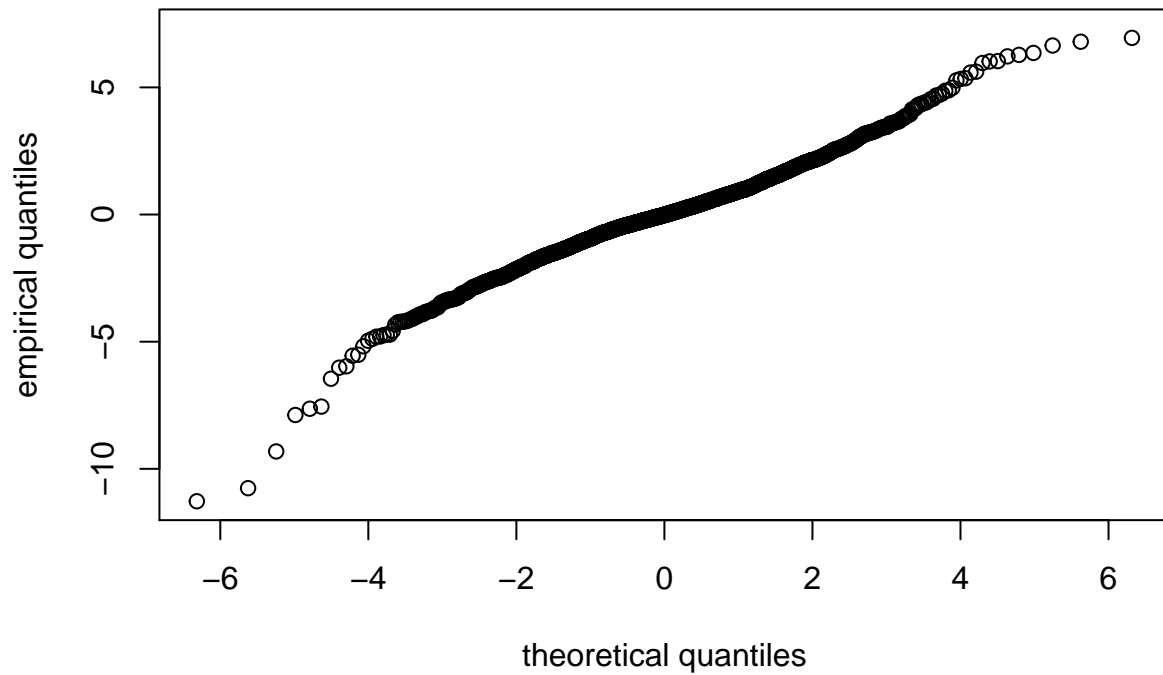
```
#normalize x to just a t-distribution with degrees of freedom df:  
x_normalized = (x-mean)/sp
```

```
#plot the theoretical quantiles against the ordered data sample to obtain a qq-plot:  
x_normalized_quantiles = sort(x_normalized)
```

```
#attain theoretical quantiles with given degrees of freedom df  
p = (1:length(x_normalized))/length(x_normalized)  
theoretical_quantiles = qt(p,df)
```

```
plot(theoretical_quantiles, x_normalized_quantiles, xlab="theoretical quantiles",  
      ylab= "empirical quantiles",main="t distribution Q-Q Plot")
```

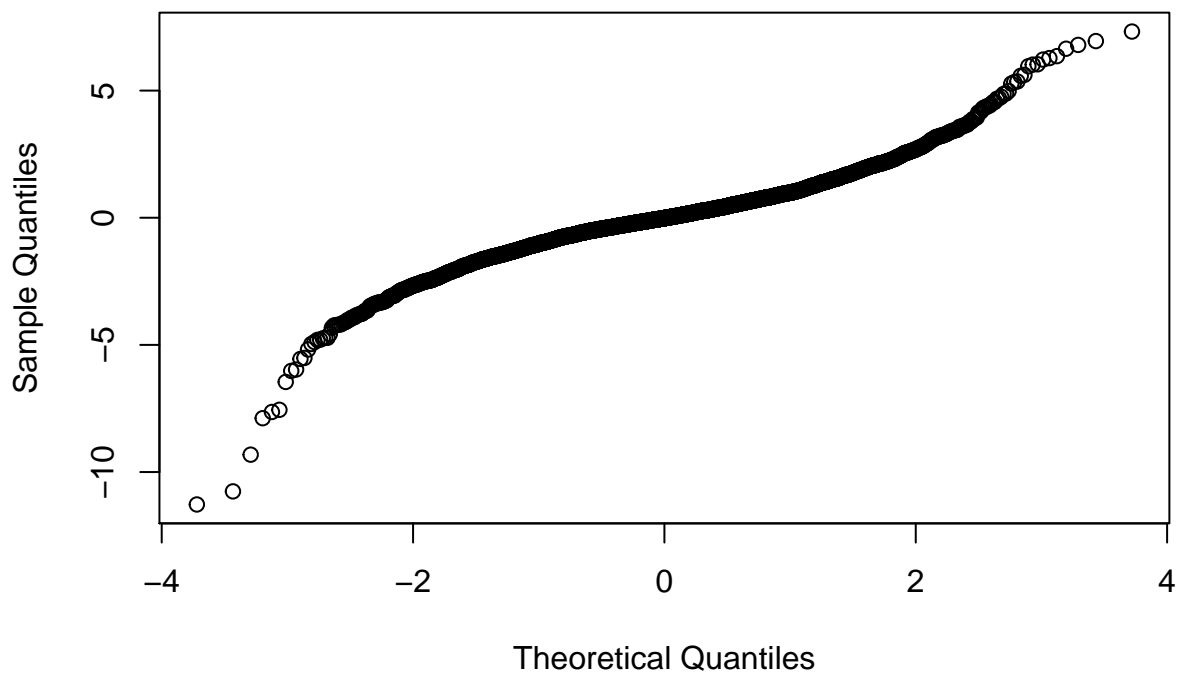
t distribution Q–Q Plot



*#Making a qq-plot for a normal distribution is significantly easier (normalization
#should really be done with the mean and standard deviation from the estimation for
#the normal distribution):*

```
qqnorm(x_normalized)
```

Normal Q–Q Plot

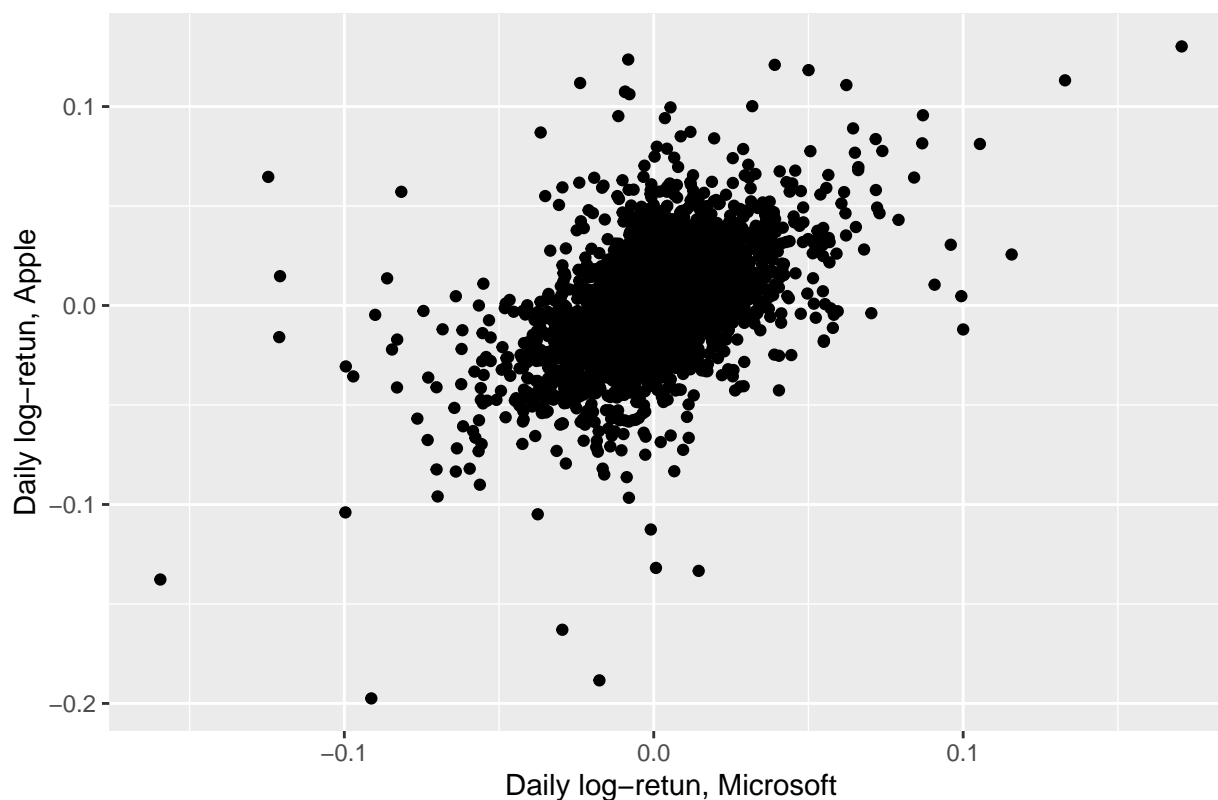


Uppgift 4

Utifrån *Figur 5* som visar den daily log-return för respektive aktie, så kan slutasten att båda aktierna tenderar att följa varandra relativt bra, det finns några enskilda dagar och punkter som varierar mer än de resterande men överlag finns samband mellan dessa två aktier.

```
MSFT <- MSFT %>%  
  mutate(daily_return = log((Close)/lag(Close)), default = first(log(Close)),  
         share_name = "MSFT")  
MSFT <- MSFT[-1,]  
  
combined_data <- MSFT %>%  
  full_join(apple, c("Date"="Date"))  
  
ggplot(combined_data, aes(x=daily_return.x, y=daily_return.y)) +  
  geom_point() +  
  ggtitle("Figur 5: Scatterplot för daily log-return för respektive aktie") +  
  labs(x= "Daily log-retun, Microsoft", y = "Daily log-retun, Apple")
```

Figur 5: Scatterplot för daily log-return för respektive aktie



Uppgift 5

Genom koden nedanför kunde en korrelation mellan de två aktiernas daily log-return beräknas till ungefär 0.5 vilket ändå är ett ganska förväntat resultat med tanke på *Figur 5* som visade att aktierna tenderade att följa varandras utveckling relativt bra.


```
cor(MSFT$daily_return,apple$daily_return)
```

```
## [1] 0.4965347
```