

Opgave 1

Input i R:

```
library(forecast)
auto.arima(Data_aktiekurs)
forecast(auto.arima(Data_aktiekurs),h=10)
plot(forecast(auto.arima(Data_aktiekurs)))
```

Output fra R:

```
> library(forecast)
> auto.arima(Data_aktiekurs)
Series: Data_aktiekurs
ARIMA(1,0,0) with non-zero mean
```

Coefficients:

	ar1	mean
	0.3847	12741.3638
s.e.	0.1378	24.4429

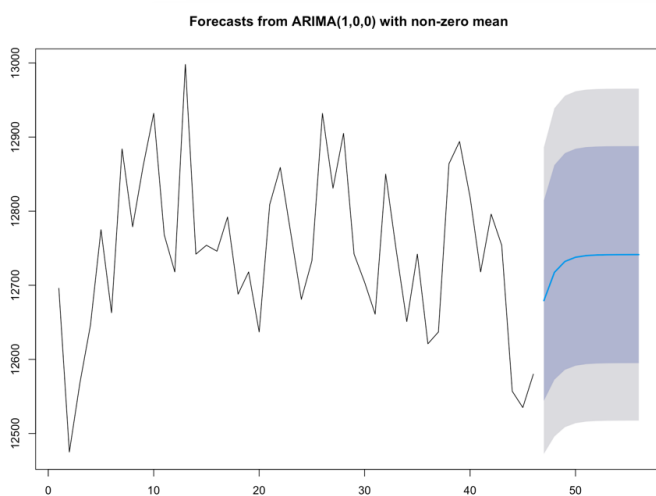
sigma² estimated as 11125: log likelihood=-278.62

AIC=563.24 AICc=563.81 BIC=568.72

```
> forecast(auto.arima(Data_aktiekurs),h=10)
```

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
47	12679.28	12544.11	12814.45	12472.56	12886.01
48	12717.48	12572.65	12862.31	12495.98	12938.98
49	12732.17	12585.97	12878.38	12508.57	12955.78
50	12737.83	12591.42	12884.24	12513.92	12961.74
51	12740.00	12593.57	12886.44	12516.05	12963.96
52	12740.84	12594.40	12887.28	12516.88	12964.81
53	12741.16	12594.72	12887.61	12517.20	12965.13
54	12741.29	12594.84	12887.73	12517.32	12965.25
55	12741.33	12594.89	12887.78	12517.37	12965.30
56	12741.35	12594.91	12887.80	12517.39	12965.32

```
> plot(forecast(auto.arima(Data_aktiekurs)))
```



Efterregning i hånden:

$$(1-0,3847) * 12.741,3638 = 7839,7611$$

Periode 47: $7839,7611 + 0,3847 * 12.580 = 12.679,28$

Periode 48: $7839,7611 + 0,3847 * 12.679,287 = 12.717,48$

Periode 49: $7839,7611 + 0,3847 * 12.717,483 = 12.732,17$

Opgave 2

Goodness of fit	
Forudsætninger	De forventede celleværdier skal være større end 5, for at sikre kvaliteten af testen, hvilket er opfyldt.
Hypotese	H_0 : Børs i stikprøvefordelingen stemmer overens med antal listede aktier på børsene H_1 : Børs i stikprøvefordelingen stemmer ikke overens med antal listede aktier på børsene.
Teststørrelse	295,74
P-værdi/signifikanssandsynlighed	0,000%
Signifikansniveau	5%
Teknisk konklusion	Da p-værdien er 0,00 % er mindre end signifikansniveauet på 5% forkaster vi nulhypotesen.
Ikke-teknisk konklusion	Da vi forkaster H_0 , er der altså ikke repræsentativitet mht. antal listede aktier på børsene. Vi kan af Chi cellebidragene se, at det skyldes at især London stock Ex. aktie er overrepræsenterede på listen. Der er altså meget færre observerede end forventet. Samtidig er der langt flere observerede aktier på New York stock end forventet.

Goodness of fit				
	Antal observeret	Sandsynligheder forventet	Chi square bidrag	Forventet antal
Deutsche Boerse AG	16	0,0485	3,086185095	24,73754153
London Stock Exh	32	0,2750	83,53786329	140,2358804
Nasdaq stock market	130	0,3668	17,40352302	187,0564784
new york stock ex	332	0,3097	191,7223973	157,9700997
Række 5				
Række 6				
Række 7				
Række 8				
Række 9				
Række 10				
Række 11				
Række 12				
Række 13				
Række 14				
Række 15				
Række 16				
Række 17				
Række 18				
Række 19				
Række 20				
Sum	510	1	295,7499687	
Frihedsgrader	3 Kritisk værdi		7,815	
Chi teststørrelse	295,7499687			
p-værdi	0,0000000000%			
Signifikans niveau	5,00%			

H_0 : Datafordelingen er identisk med den angivne fordeling
 H_A : Datafordelingen er ikke identisk med den angivne fordeling

P-værdien/signifikanssandsynligheden 0 er mindre end signifikansniveauet 5%, Det betyder teststørrelsen 295,75 er større end den kritiske værdi 7,81. Vi afviser nulhypotesen om at datafordelingen er identisk med den angivne fordeling.

Opgave 3

Chi i anden test	
Forudsætninger	De forventede celleværdier skal være større end 5, for at sikre kvaliteten af testen. Denne forudsætning er opfyldt.
Hypotese	H_0 : Der er uafhængighed mellem risiko vurdering Fremtidige afkast forventninger. H_1 : Der er afhængighed mellem risiko vurdering og Fremtidig afkast forventninger.
Teststørrelse	20.62
P-værdi/signifikanssandsynlighed	3,75%
Signifikansniveau	5%
Teknisk konklusion	Da p-værdien er 3,75% og mindre end signifikansniveauet på 5% forkaster vi nulhypotesen.
Ikke-teknisk konklusion	Der er sammenhæng mellem risiko vurdering og fremtidige afkast forventninger. Afhængigheden skyldes især at folk med lave forventninger tager en højere risiko end forventet. Og at folk med høje forventninger tager en lavere risiko end forventet.

Chi i anden ordnede data

Observeret	Lav forventning	Medium forve	Høj forventnin
Lav risiko	30	43	142
Medium risiko	53	49	140
Høj risiko	22	8	23

Række 5			
Række 6			
Række 7			
Række 8			
Række 9			
Række 10			
Sum	105	100	305

Forventet	Lav forventning	Medium forve	Høj forventnin
Række 1	44,26470588	42,1568627	128,578431
Række 2	49,82352941	47,4509804	144,72549
Række 3	10,91176471	10,3921569	31,6960784
Række 4	0	0	0
Række 5	0	0	0
Række 6	0	0	0
Række 7	0	0	0
Række 8	0	0	0
Række 9	0	0	0
Række 10	0	0	0
Sum	105	100	305

Chi i anden bidrag	Lav forventning	Medium forve	Høj forventnin
Row 1	4,596931796	0,01686275	1,40100095
Row 2	0,202514063	0,05056717	0,15429388
Row 3	11,26755985	0,55064743	2,38584026
Row 4			

Eksamen 27.10.2021
Kvantitativ metode

Frihedsgrader	4	Kritisk værdi	9,488	Signifikans niveau	5,00%
Chi-teststørrelse	20,62621815				
p-værdi	0,037553949%				
H_0: Der er uafhængighed mellem række- og søjlekriterie H_A: Der er afhængighed mellem række- og søjlekriterie P-værdien/signifikanssandsynligheden 0,0004 er mindre end signifikansniveauet 5%. Det betyder Chi-teststørrelsen 20,63 er større end den kritiske værdi 9,49. Vi afviser nulhypotesen, der er afhængighed mellem række- og søjlekriterie.					

Opgave 4.

Chi i anden test	
Forudsætninger	De forventede celleværdier skal være større end 5, for at sikre kvaliteten af testen. Denne forudsætning er opfyldt.
Hypotese	H_0: Der er uafhængighed mellem stiftet år og Fremtidige afkast forventninger. H_1: Der er afhængighed mellem stiftet år og Fremtidig afkast forventninger.
Teststørrelse	0,83
P-værdi/signifikanssandsynlighed	65,98%
Signifikansniveau	5%
Teknisk konklusion	Da p-værdien er 65,98% og større end signifikansniveauet på 5% forkaster vi ikke nulhypotesen.
Ikke-teknisk konklusion	Der er ikke sammenhæng mellem stiftet år og fremtidige afkast forventninger.

Chi i anden ordnede data 			
Observeret 	Lav forventning	Medium forve	Høj forventnin
Før år 2000	74	72	206
Efter år 2000	31	28	99
Række 5			
Række 6			
Række 7			
Række 8			
Række 9			
Række 10			
Sum	105	100	305
Forventet 			
Række 1	72,47058824	69,0196078	210,509804
Række 2	32,52941176	30,9803922	94,4901961
Række 3	0	0	0
Række 4	0	0	0
Række 5	0	0	0
Række 6	0	0	0
Række 7	0	0	0
Række 8	0	0	0
Række 9	0	0	0
Række 10	0	0	0
Sum	105	100	305
Chi i anden bidrag			
Row 1	0,032276547	0,12869875	0,09661465
Row 2	0,071907244	0,28672127	0,21524277

Frihedsgrader	2	Kritisk værdi	5,991	Signifikans niveau	5,00%
Chi-teststørrelse	0,831461234				
p-værdi	65,985800096%				
H_0 : Der er uafhængighed mellem række- og søjlekriterie					
H_A : Der er afhængighed mellem række- og søjlekriterie					
P-værdien/signifikanssandsynligheden 0,6599 er større end signifikansniveauet 5%, Det betyder Chi-teststørrelsen 0,83 er mindre end den kritiske værdi 5,99. Vi kan ikke afvise nulhypotesen om uafhængighed mellem række- og søjlekriterie					

Opgave 5.

Svar: De vil kunne benytte en klyngeanalyse til at undersøge hvilke virksomheder der minder om hinanden. En klyngeanalyse er god til at gruppere respondenter, som i dette tilfælde er virksomhederne.

Opgave 6.

Input i R:

```
km3<-kmeans(scale(London),5,nstart=25)
pacman::p_load(factoextra)
fviz_cluster(km3, data = London, main ="Klyngeplot opdelt i 5 klynger", repel=TRUE)
```

Klyngeplot med 5 klynger:

