Project 1

1)

a) Simulate e^X with your random vector. Find an estimate for θ.

```
set.seed(141)

y<-sqrt(3)

y

# [1] 1.732051

x<-rnorm(75,1,y)

x

# [1]  1.89089717  0.80466386  1.11144026 -0.13491729  1.86901181 -0.446694
# 53 -0.16908540
# [8]  0.70603867  1.85298749  0.39393357  0.55416849 -0.40947759  3.37217
# 598  4.35145740
# [15]  0.48164696 -0.12171518  1.96415306  1.42972711 -0.44766438  1.02809
# 959 -0.67337728
# [22]  0.97397550  4.43382777  3.80396920  1.75837967  0.08679560 -1.63558
# 502  1.87787071
# [29] -0.01888620  0.73788885  1.93748693 -1.62852341  2.18766857  1.23282
# 557  0.97584997
# [36]  3.74390955 -1.87761610  1.31965046  0.26217604 -1.90277996  1.11691
# 514  2.49449149
# [43]  2.95978096 -0.63706706  2.69206425  1.99343912  2.25288691  1.58027
# 176 -1.51612212
# [50]  0.40629491  1.33726476  2.59836172 -1.28579732  3.35011578  0.64882
# 151 -0.08130412
# [57] -0.52616545  3.05183992 -1.07256105 -1.09861231  2.53275523 -2.89137
# 035 -1.06624935
# [64]  0.32587146 -0.79627007 -0.11552752  1.33948114  1.15293773  2.00960
# 878  3.50547232
# [71]  0.56901740  2.69037352  3.08963279  3.88833355 -0.78842797


q<-exp(x)

q

# [1]  6.62531008  2.23594477  3.03873181  0.87378818  6.48188788  0.639739
# 30  0.84443678
#  [8]  2.02594989  6.37884781  1.48280205  1.74049314  0.66399704 29.14187
# 033 77.59146245
# [15]  1.61873821  0.88540051  7.12887228  4.17755900  0.63911915  2.79574
# 771  0.50998331
# [22]  2.64845247 84.25330262 44.87896502  5.80302693  1.09067373  0.19483
# 835  6.53956539
# [29]  0.98129103  2.09151535  6.94128511  0.19621910  8.91440557  3.43091
# 015  2.65342159
# [36] 42.26289666  0.15295430  3.74211312  1.29975533  0.14915340  3.05541
# 414 12.11557106
```

```
# [43] 19.29374516  0.52884121 14.76211723  7.34073611  9.51516565  4.85627
# 536  0.21956167
# [50]  1.50124522  3.80861178 13.44169868  0.27643009 28.50603392  1.91328
# 472  0.92191327
# [57]  0.59086634 21.15423073  0.34213118  0.33333333 12.58814161  0.05550
# 011  0.34429744
# [64]  1.38523730  0.45100806  0.89089606  3.81706246  3.16748445  7.46039
# 813 33.29716703
# [71]  1.76653041 14.73717948 21.96900918 48.82944703  0.45455881
```

```
estimate<-mean(q)

estimate

# [1] 8.952487
```

b) Construct a 95% confidence interval for θ.

```
low<- qnorm(0.025,1,y)

low

# [1] -2.394757

lowerlimit<-exp(low)

lowerlimit

# [1] 0.09119482

up<- qnorm(0.975,1,y)

up

# [1] 4.394757

upperlimit<-exp(up)

upperlimit

# [1] 81.02496

ConfidenceInterval<-c(lowerlimit, upperlimit)

ConfidenceInterval

 [1]  0.09119482 81.02495570
```

c) Using your knowledge from probability theory course, find the exact value of θ (Hint: Solve the integral for the expected value).

$$\mathrm{E}(e^X) = \mathrm{E}(e^{\mu+\sigma Z}) = \int_{-\infty}^{\infty} e^{\mu+\sigma z}\varphi(z)\,dz$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\mu+\sigma z}e^{-z^2/2}\,dz = \frac{1}{\sqrt{2\pi}}e^{\mu}\int_{-\infty}^{\infty} e^{\sigma z}e^{-z^2/2}\,dz.$$

We have $\sigma z - \frac{z^2}{2}$ so of course we complete the square:

$$\frac{1}{2}(z^2 - 2\sigma z) = \frac{1}{2}(z^2 - 2\sigma z + \sigma^2) - \frac{1}{2}\sigma^2 = \frac{1}{2}(z-\sigma)^2 - \frac{1}{2}\sigma^2.$$

Then the integral is

$$\frac{1}{\sqrt{2\pi}}e^{\mu+\sigma^2/2}\int_{-\infty}^{\infty} e^{-(z-\sigma)^2/2}\,dz$$

This whole thing is

$$e^{\mu+\sigma^2/2}.$$

```
mean<-1
#sigma^2=variance
variance<-3
exactvalue<-exp(mean+(variance/2))
exactvalue
# [1] 12.18249
```

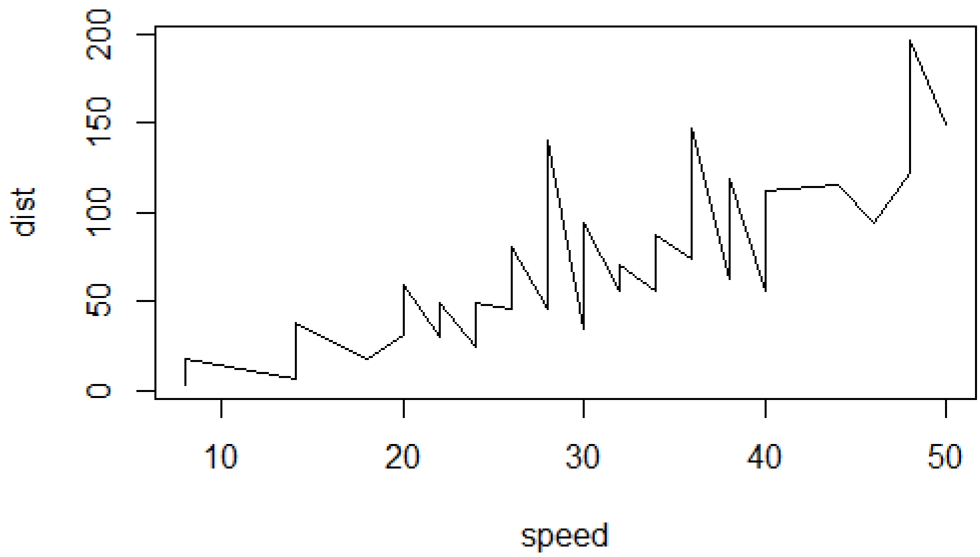
d) Is the solution of (c) inside the confidence interval you have found in (b) ?

Yes, it is in the interval.

2)

a) Draw a scatterplot to show the relationship between two sets of data.

```
setwd("C:/Users/gülce/Desktop/3-2/IE360/Assignment 1")
question2<-read.table("cars.txt",header=TRUE)
plot(question2, type="l")
```



b) Calculate the correlation between the speed and the distance.

question2.b<-cor(question2$speed,question2$dist)

question2.b

[1] 0.8124657

c) Comment on your results in parts a and b.

Correlation is too high which means there is a strong relation between two sets of data. Also correlation is positive which means that distance is increasing when speed is increasing. It also means that distance is decreasing when speed is decreasing. It can be seen at scatterplot also.

3) a) Draw a time series plot.

setwd("C:/Users/gülce/Desktop/3-2/IE360/Assignment 1")

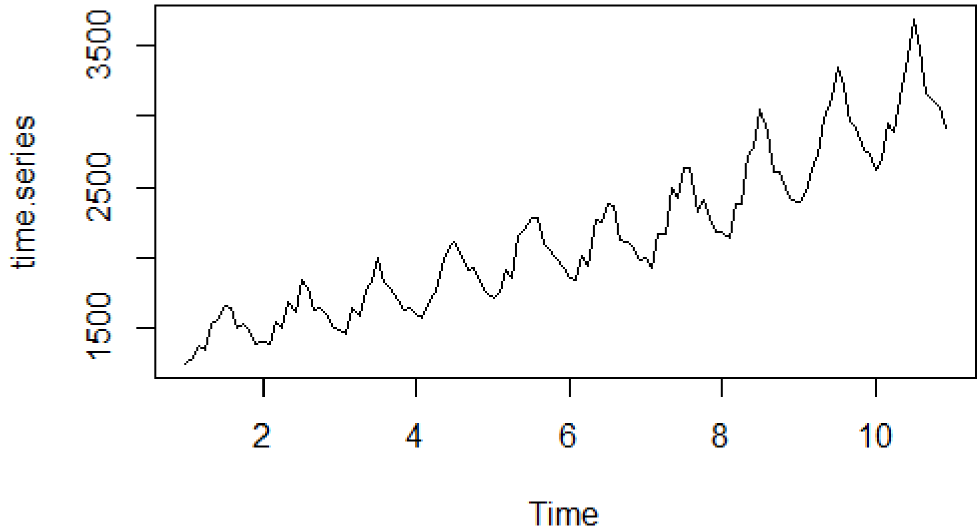
question3<-read.table("electricity.txt",header=FALSE)

time.series<-ts(question3[1:120,1], frequency = 12)

time.series

```
# Jan  Feb  Mar  Apr  May  Jun  Jul  Aug  Sep  Oct  Nov  Dec
# 1   1254 1290 1379 1346 1535 1555 1655 1651 1500 1538 1486 1394
# 2   1409 1387 1543 1502 1693 1616 1841 1787 1631 1649 1586 1500
# 3   1497 1463 1648 1595 1777 1824 1994 1835 1787 1699 1633 1645
# 4   1597 1577 1709 1756 1936 2052 2105 2016 1914 1925 1824 1765
# 5   1721 1752 1914 1857 2159 2195 2287 2276 2096 2055 2004 1924
# 6   1851 1839 2019 1937 2270 2251 2382 2364 2129 2110 2072 1980
# 7   1995 1932 2171 2162 2489 2424 2641 2630 2324 2412 2284 2186
# 8   2184 2144 2379 2383 2717 2774 3051 2891 2613 2600 2493 2410
# 9   2390 2463 2616 2734 2970 3125 3342 3207 2964 2919 2764 2732
# 10  2622 2698 2950 2895 3200 3408 3679 3473 3154 3107 3052 2918
```

plot(time.series)

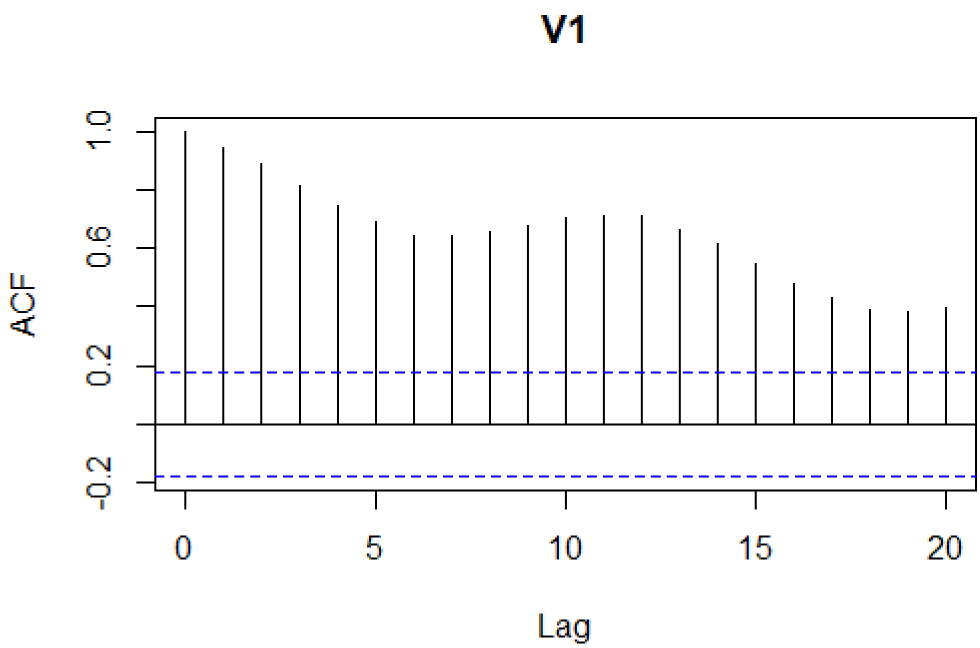


b) Draw an autocorrelation plot.

Since it is not a suitable data set for additive model, log of the data should be considered.

logquestion3<-log(question3)

acf(logquestion3, lag.max = NULL)



c) Comment on your results in parts a and b.

There is not an immediate decrease anormality in autocorrelation which means data points are correlated. Also, fluctuations in autocorrelation are not exponentially, they are more likely linear which means there is a trend in this time interval.

4)

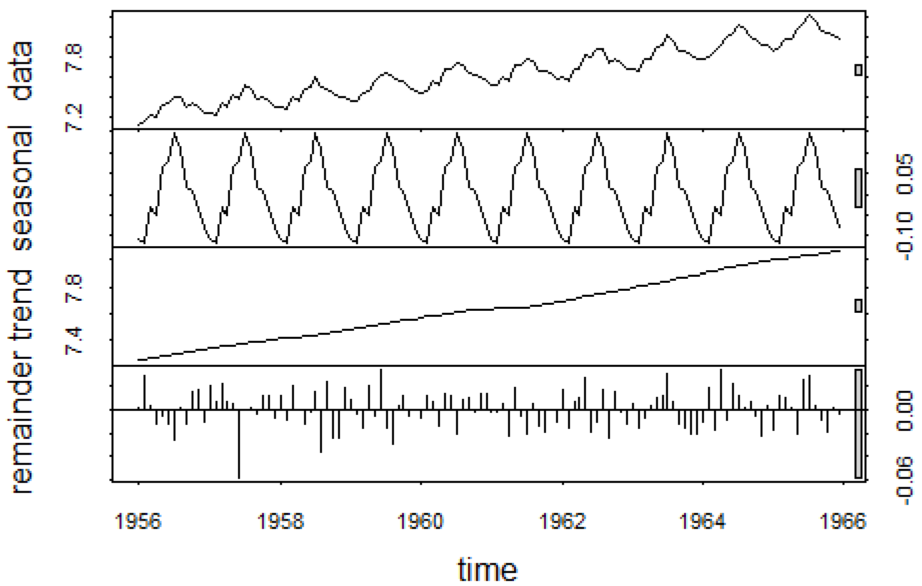
a) Construct a decomposition model. (If needed, you can do transformation on your time series).

```
setwd("C:/Users/gülce/Desktop/3-2/IE360/Assignment 1")

question3<-read.table("electricity.txt",header=FALSE)

q3ts<-ts(question3, frequency = 12,start =c(1956,1),end = c(1965,12))

# Since the model is multiplicative, log of the data should be used in decomposition in order to make
# it suitable for an additive model.

time.series<-log(q3ts)

decomposition<-stl(time.series[,c(1)] , s.window="periodic")

decomposition

# Call:
#  stl(x = time.series[, c(1)], s.window = "periodic")
#
# Components
#              seasonal    trend     remainder
# Jan 1956 -0.108794878 7.241939  9.494076e-04
# Feb 1956 -0.116705226 7.249903  2.919934e-02
# Mar 1956 -0.031780419 7.257868  3.026724e-03
# Apr 1956 -0.048495738 7.266074 -1.268586e-02
# May 1956  0.067450618 7.274281 -5.445594e-03
# Jun 1956  0.078864693 7.282542 -1.217585e-02
# Jul 1956  0.146530024 7.290803 -2.577706e-02
# Aug 1956  0.108230328 7.298940  1.966333e-03
# Sep 1956  0.017437860 7.307076 -1.129371e-02
# Oct 1956  0.006983257 7.315568  1.568699e-02
# Nov 1956 -0.037822987 7.324060  1.760665e-02
# Dec 1956 -0.081897534 7.331354 -9.523864e-03
# Jan 1957 -0.108794878 7.338648  2.078198e-02
# Feb 1957 -0.116705226 7.344743  6.861123e-03
# Mar 1957 -0.031780419 7.350837  2.242764e-02
# Apr 1957 -0.048495738 7.356379  6.669466e-03
# May 1957  0.067450618 7.361922  4.885191e-03
# Jun 1957  0.078864693 7.367174 -5.832954e-02
# Jul 1957  0.146530024 7.372427 -8.924370e-04
# Aug 1957  0.108230328 7.377838  2.224901e-03
# Sep 1957  0.017437860 7.383250 -3.739235e-03
# Oct 1957  0.006983257 7.389309  1.163250e-02
# Nov 1957 -0.037822987 7.395367  1.142623e-02
# Dec 1957 -0.081897534 7.401720 -6.601873e-03
# Jan 1958 -0.108794878 7.408072  1.194083e-02
# Feb 1958 -0.116705226 7.413220 -8.270545e-03
# Mar 1958 -0.031780419 7.418368  2.073022e-02
# Apr 1958 -0.048495738 7.422593  5.316368e-04
```

```
# May 1958  0.067450618 7.426818 -1.158711e-02
# Jun 1958  0.078864693 7.431579 -1.656928e-03
# Jul 1958  0.146530024 7.436340  1.502744e-02
# Aug 1958  0.108230328 7.441857 -3.528778e-02
# Sep 1958  0.017437860 7.447374  2.348172e-02
# Oct 1958  0.006983257 7.454040 -2.322856e-02
# Nov 1958 -0.037822987 7.460707 -2.470984e-02
# Dec 1958 -0.081897534 7.468256  1.913717e-02
# Jan 1959 -0.108794878 7.475805  8.871898e-03
# Feb 1959 -0.116705226 7.483425 -3.440071e-03
# Mar 1959 -0.031780419 7.491045 -1.560054e-02
# Apr 1959 -0.048495738 7.498781  2.050884e-02
# May 1959  0.067450618 7.506517 -5.588046e-03
# Jun 1959  0.078864693 7.514051  3.365407e-02
# Jul 1959  0.146530024 7.521586 -1.604548e-02
# Aug 1959  0.108230328 7.528970 -2.832959e-02
# Sep 1959  0.017437860 7.536354  3.159135e-03
# Oct 1959  0.006983257 7.543804  1.189431e-02
# Nov 1959 -0.037822987 7.551254 -4.643618e-03
# Dec 1959 -0.081897534 7.558772 -9.681257e-04
# Jan 1960 -0.108794878 7.566289 -6.833809e-03
# Feb 1960 -0.116705226 7.574021  1.119712e-02
# Mar 1960 -0.031780419 7.581753  6.977712e-03
# Apr 1960 -0.048495738 7.588939 -1.372617e-02
# May 1960  0.067450618 7.596126  1.382415e-02
# Jun 1960  0.078864693 7.602650  1.242276e-02
# Jul 1960  0.146530024 7.609174 -2.070791e-02
# Aug 1960  0.108230328 7.614332  7.612810e-03
# Sep 1960  0.017437860 7.619489  1.085895e-02
# Oct 1960  0.006983257 7.623412 -2.364300e-03
# Nov 1960 -0.037822987 7.627335  1.338834e-02
# Dec 1960 -0.081897534 7.630325  1.373462e-02
# Jan 1961 -0.108794878 7.633314 -1.037788e-03
# Feb 1961 -0.116705226 7.635634 -1.951347e-03
# Mar 1961 -0.031780419 7.637954  4.184421e-03
# Apr 1961 -0.048495738 7.640055 -2.266315e-02
# May 1961  0.067450618 7.642155  1.792900e-02
# Jun 1961  0.078864693 7.645549 -5.284315e-03
# Jul 1961  0.146530024 7.648943 -1.977771e-02
# Aug 1961  0.108230328 7.654294  5.585990e-03
# Sep 1961  0.017437860 7.659645 -1.367488e-02
# Oct 1961  0.006983257 7.666896 -1.943583e-02
# Nov 1961 -0.037822987 7.674147 -5.433072e-05
# Dec 1961 -0.081897534 7.682327 -9.577221e-03
# Jan 1962 -0.108794878 7.690507  1.668737e-02
# Feb 1962 -0.116705226 7.699225 -1.620925e-02
# Mar 1962 -0.031780419 7.707944  6.779450e-03
# Apr 1962 -0.048495738 7.716693  1.059198e-02
# May 1962  0.067450618 7.725441  2.674432e-02
# Jun 1962  0.078864693 7.733697 -1.938687e-02
# Jul 1962  0.146530024 7.741952 -9.568795e-03
# Aug 1962  0.108230328 7.749729  1.677942e-02
# Sep 1962  0.017437860 7.757507 -2.389981e-02
# Oct 1962  0.006983257 7.765741  1.548685e-02
# Nov 1962 -0.037822987 7.773976 -2.469266e-03
# Dec 1962 -0.081897534 7.783568 -1.184151e-02
# Jan 1963 -0.108794878 7.793160  4.548605e-03
# Feb 1963 -0.116705226 7.802884 -1.575000e-02
# Mar 1963 -0.031780419 7.812608 -6.391953e-03
# Apr 1963 -0.048495738 7.821096  3.515415e-03
# May 1963  0.067450618 7.829584  1.024927e-02
# Jun 1963  0.078864693 7.837564  1.161696e-02
```

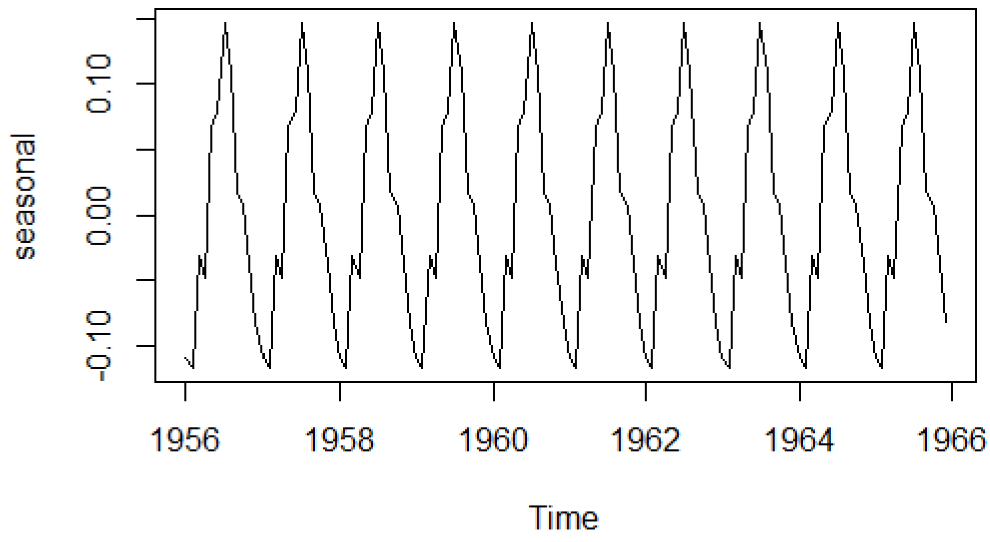
```
# Jul 1963  0.146530024 7.845544  3.115048e-02
# Aug 1963  0.108230328 7.853990  7.137545e-03
# Sep 1963  0.017437860 7.862436 -1.161915e-02
# Oct 1963  0.006983257 7.871179 -1.489518e-02
# Nov 1963 -0.037822987 7.879922 -2.085667e-02
# Dec 1963 -0.081897534 7.889011 -1.973189e-02
# Jan 1964 -0.108794878 7.898101 -1.025764e-02
# Feb 1964 -0.116705226 7.908001  1.783989e-02
# Mar 1964 -0.031780419 7.917900 -1.671818e-02
# Apr 1964 -0.048495738 7.927905  3.411207e-02
# May 1964  0.067450618 7.937909 -9.042436e-03
# Jun 1964  0.078864693 7.946696  2.162916e-02
# Jul 1964  0.146530024 7.955482  1.231232e-02
# Aug 1964  0.108230328 7.963075  1.785583e-03
# Sep 1964  0.017437860 7.970668  6.188911e-03
# Oct 1964  0.006983257 7.977379 -5.366247e-03
# Nov 1964 -0.037822987 7.984091 -2.183334e-02
# Dec 1964 -0.081897534 7.990935  3.751822e-03
# Jan 1965 -0.108794878 7.997779 -1.729181e-02
# Feb 1965 -0.116705226 8.004731  1.223991e-02
# Mar 1965 -0.031780419 8.011683  9.657512e-03
# Apr 1965 -0.048495738 8.017718  1.518127e-03
# May 1965  0.067450618 8.023753 -2.029717e-02
# Jun 1965  0.078864693 8.029779  2.523750e-02
# Jul 1965  0.146530024 8.035805  2.806148e-02
# Aug 1965  0.108230328 8.041793  2.750375e-03
# Sep 1965  0.017437860 8.047782 -8.793044e-03
# Oct 1965  0.006983257 8.053553 -1.912361e-02
# Nov 1965 -0.037822987 8.059325  2.050798e-03
# Dec 1965 -0.081897534 8.064911 -4.360105e-03
```

plot(decomposition)



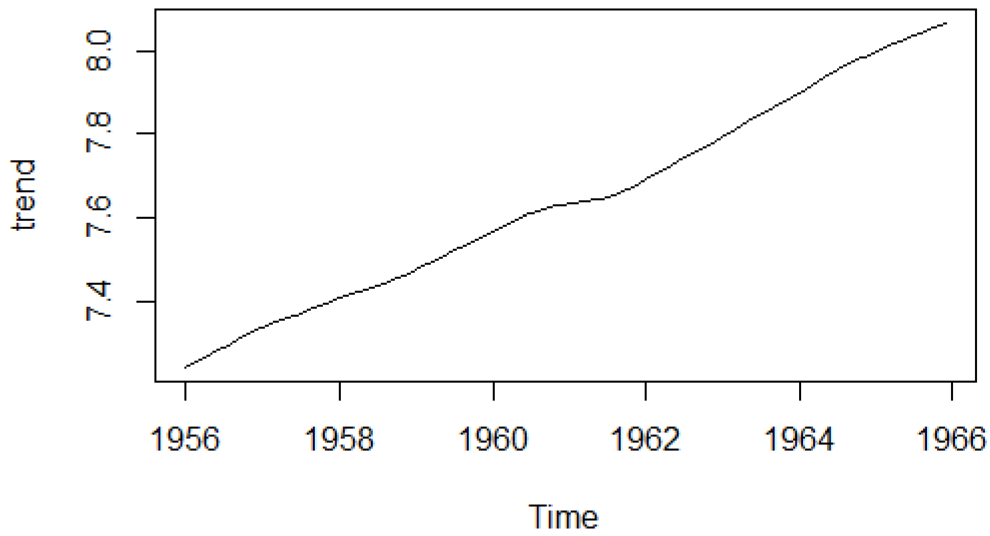
seasonal<-decomposition$time.series[,1]

plot(seasonal)



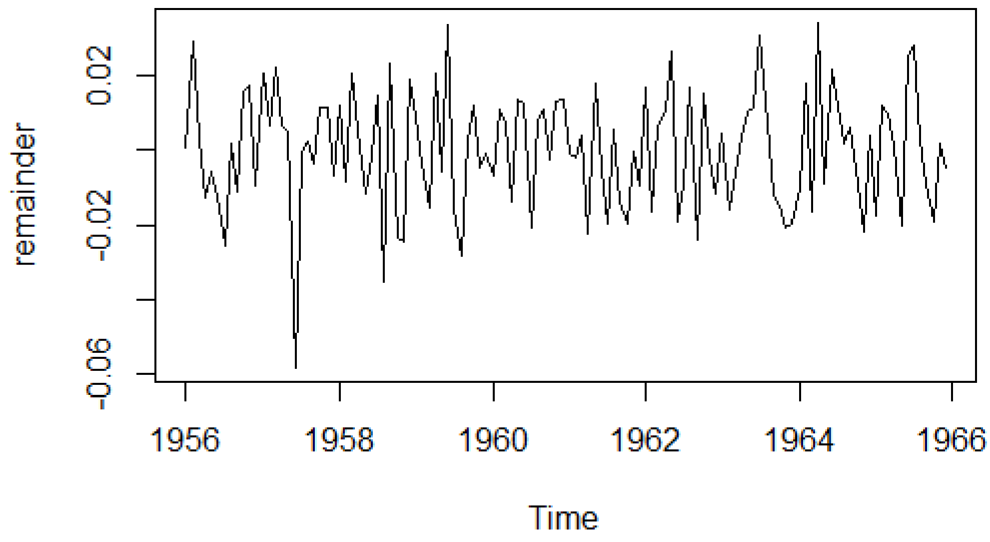
trend<-decomposition$time.series[,2]

plot(trend)



remainder<-decomposition$time.series[,3]

plot(remainder)



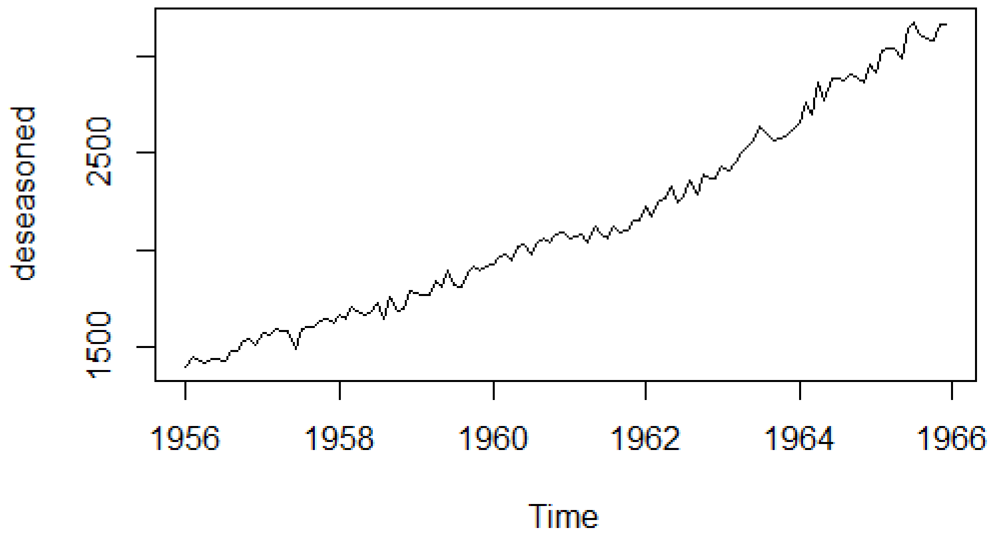
b) Remove the seasonality from your data. Draw a time series plot and autocorrelation plot for the deseasonalized data.

exceptseasonality<-trend+remainder

Since the log of the data is considered for decomposition, exponential of it should be taken back in
order to get real time series.

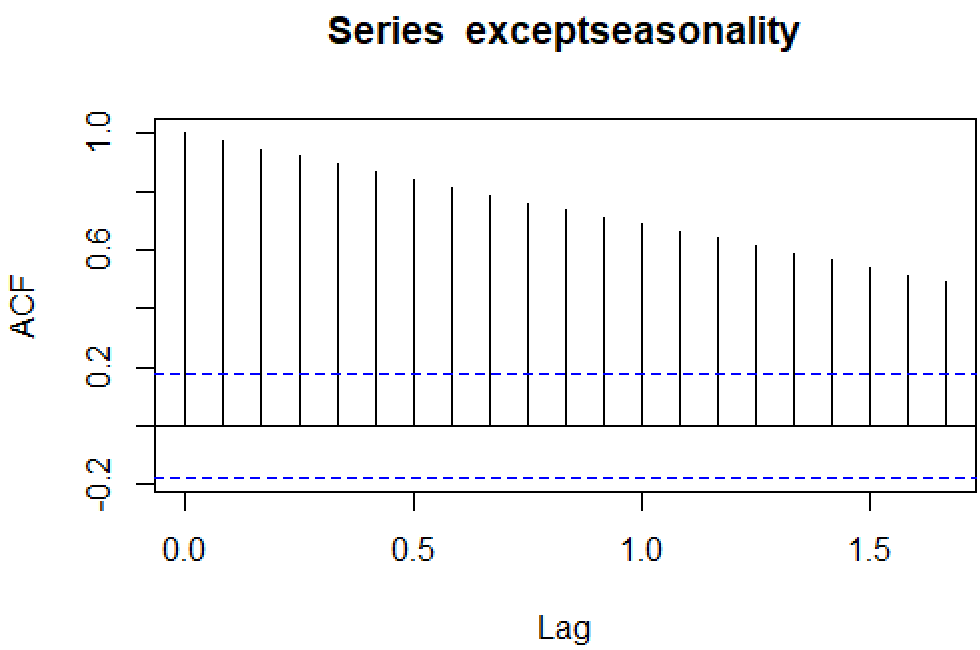
deseasoned<-exp(exceptseasonality)

plot(deseasoned)



log of the data should be used for autocorrelation since it is a multiplicative model.

acf(exceptseasonality)

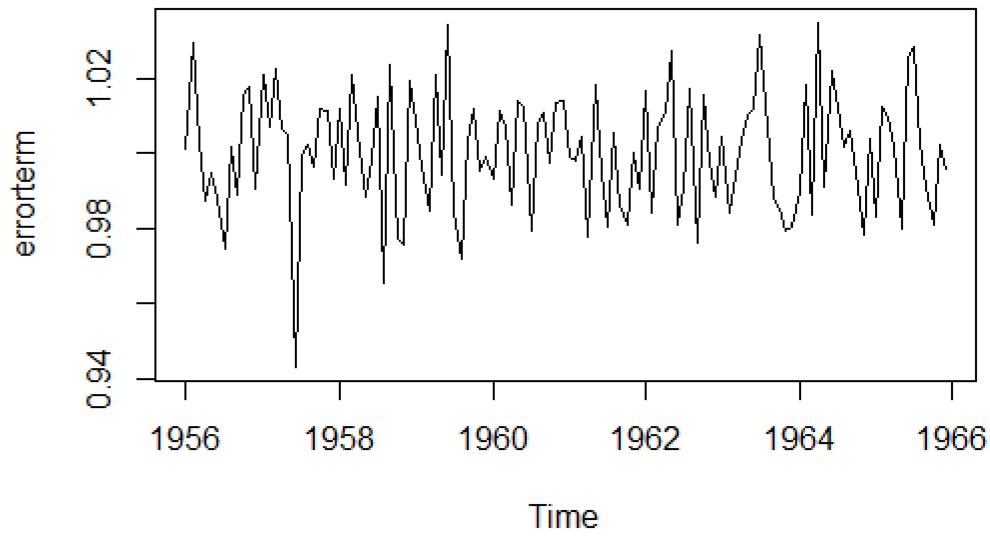


c) Remove the trend in your deseasonalized data. Draw a time series plot and autocorrelation plot without trend.

Since the log of the data is considered for decomposition, exponential of it should be taken back in
order to get real time series.

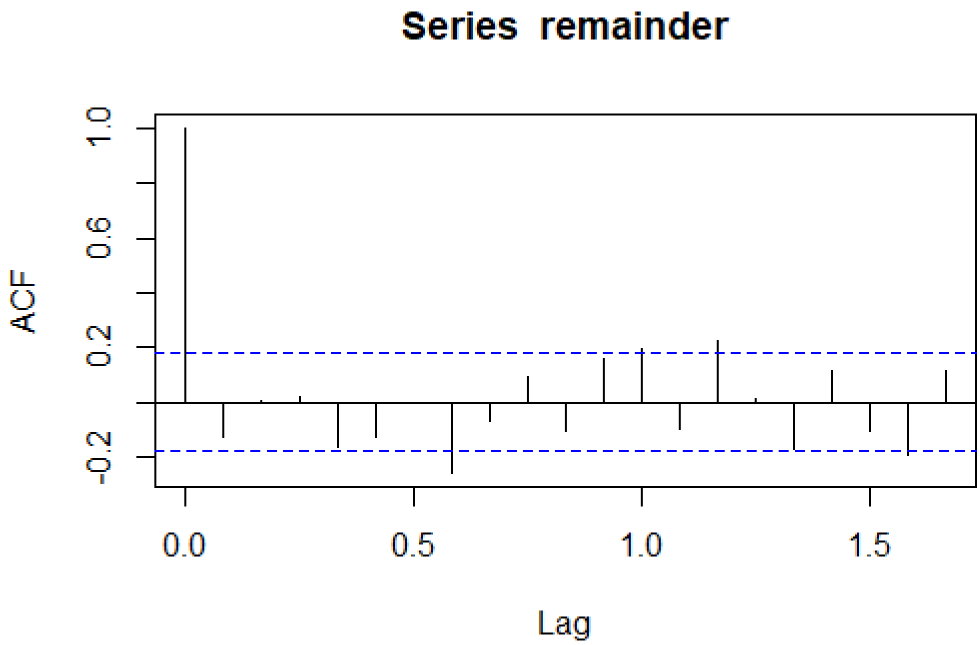
errorterm<-exp(remainder)

plot(errorterm)



log of the data should be used for autocorrelation since it is a multiplicative model.

acf(remainder)



d) Comment on your results in parts b and c.

In part b);

It can be seen that except the effect of seasonality, there is a strong increasing trend in the data.

Autocorrelation signs that there is a strong correlation. The spikes are statistically important.

In part c);

Except for one major downturn, there is not any particular pattern. Fluctuations tend to be around same level. So, returns are independent of each other.

The returns are not highly correlated, the spikes are not statistically important.