**IE 360 / HOMEWORK – 1 REPORT**

#Generate a normal random variable set distributed with µ = 1, σ = Sqrt(3) by using seed 330

**set.seed(330)**

**X<-rnorm(75,1,sqrt(3))**

**X**

[1] 6.88338870 4.57045542 1.53566845 -1.24732062 1.11615047 2.12772989 0.14959550 3.30464017

[9] -0.39185126 -1.03760059 2.79500184 0.25843459 -0.80230293 2.32036716 2.91318955 -0.98683596

[17] -1.22535802 3.36852044 1.03853399 -1.24547799 -1.20686943 1.26743769 5.78745173 -0.64408692

[25] 2.26900403 0.77022745 -2.28604972 1.28087693 1.59910804 0.70014369 1.03374516 -1.32152270

[33] 0.21026232 -2.26638382 -0.50895750 -0.96911467 1.19442508 1.35533101 1.94737562 -1.29905549

[41] 0.98054419 4.25745743 -0.90643285 1.61478965 2.45256563 0.09893742 -1.10077146 0.61011813

[49] 3.49688441 2.11006394 -0.80534651 3.60095537 -1.03818472 0.08296019 2.60610452 -0.29507132

[57] 2.61933746 2.21733007 0.33997733 0.33050958 0.74886589 2.44644487 2.39951949 0.66348328

[65] 0.54871143 0.91213136 3.62995872 -0.82589148 6.11086295 -1.59785739 1.80862619 0.75810329

[73] -2.63982854 2.61957328 1.70653846

# Generate Exp(X) random variable sample using above X values

**a<-exp(X)**

**a**

[1] 39.94886445 48.53328841 3.80685807 7.29059679 7.42300623 2.05309848 0.13942052 1.66304375

[9] 6.25948435 0.53417057 10.22905564 0.53733378 2.22447120 7.01111848 3.69324872 1.77352787

[17] 20.21423740 3.05130680 0.44227301 2.30510843 16.39741650 3.35320433 1.26796330 3.76260856

[25] 0.51432439 242.65832419 0.75219019 6.28607594 1.37456366 0.01620074 14.97541375 1.16607242

[33] 1.09368928 0.64237275 11.30746831 4.00306731 2.05281173 0.38857928 1.39049643 2.43003746

[41] 1.10494871 0.16980578 7.19511597 94.80008024 0.38239211 3.75212875 4.04007951 2.26637871

[49] 38.29428944 0.40759216 0.14244307 35.44785070 0.41873928 1.49639162 1.32962591 1.33901085

[57] 6.31483046 0.40833016 1.47585811 1.55955944 3.79880623 3.74467754 2.02087650 0.43843957

[65] 6.47627150 12.04415867 26.73233614 3.66741089 0.49595308 0.81932193 1.36958800 3.75258056

[73] 5.92838269 0.46849984 0.29690894

#find an estimation for expectation of , the mean of the sample is used:

**a1<-mean(a)**

**a1**

#[1] 10.12488

#Construct a 95% confidence interval for 𝜃:

**b=t.test(a,conf.level = 0.95)**

**b**

#One Sample t-test

#data: a

#t = 2.8615, df = 74, p-value = 0.005479

#alternative hypothesis: true mean is not equal to 0

#95 percent confidence interval:

# 3.074564 17.175197

#sample estimates:

#mean of x

# 10.12488

E,

* E, , E = 0, Var = 1
* E, =
* E
* E =
* = = =
* E=
* E= = =
* When X ̴ N(1,2) µ = 1, = 2
* E = = = = 12.18249396

1. The solution which is obtained in part c is in the interval that found in part b:

The interval: [3.074564,17.175197]

3.074564<12.18249396>17.175197

**a)**

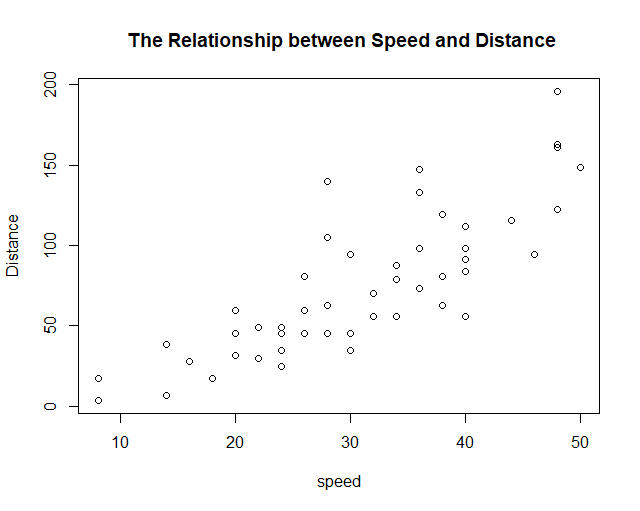
**getwd()**

**setwd("C:/Users/Sinan/Desktop/Assignment 1 360")**

**cardata <- read.table(file="cars.txt", header=TRUE)**

**cardata**

**plot(cardata,xlab="speed",ylab="Distance",main=" The Relationship between Speed and Distance")**



**b)**

**cor(cardata$speed,cardata$dist)**

#[1] 0.8124657

**c)**

The correlation shows us that there is high relationship between two data.This means that when one data is increasing then the other one will increase as nearly with the same rate.

#Getting the required data from txt file:

**getwd()**

**setwd("C:/Users/Sinan/Desktop")**

**netdata <- read.table(file="electricity.txt", header=FALSE)**

#to draw time series plot of data convert to time series model the data

**netdatats<-ts(netdata,frequency = 12,start=c(1956,1))**

**netdatats**

# Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

#1956 1254 1290 1379 1346 1535 1555 1655 1651 1500 1538 1486 1394

#1957 1409 1387 1543 1502 1693 1616 1841 1787 1631 1649 1586 1500

#1958 1497 1463 1648 1595 1777 1824 1994 1835 1787 1699 1633 1645

#1959 1597 1577 1709 1756 1936 2052 2105 2016 1914 1925 1824 1765

#1960 1721 1752 1914 1857 2159 2195 2287 2276 2096 2055 2004 1924

#1961 1851 1839 2019 1937 2270 2251 2382 2364 2129 2110 2072 1980

#1962 1995 1932 2171 2162 2489 2424 2641 2630 2324 2412 2284 2186

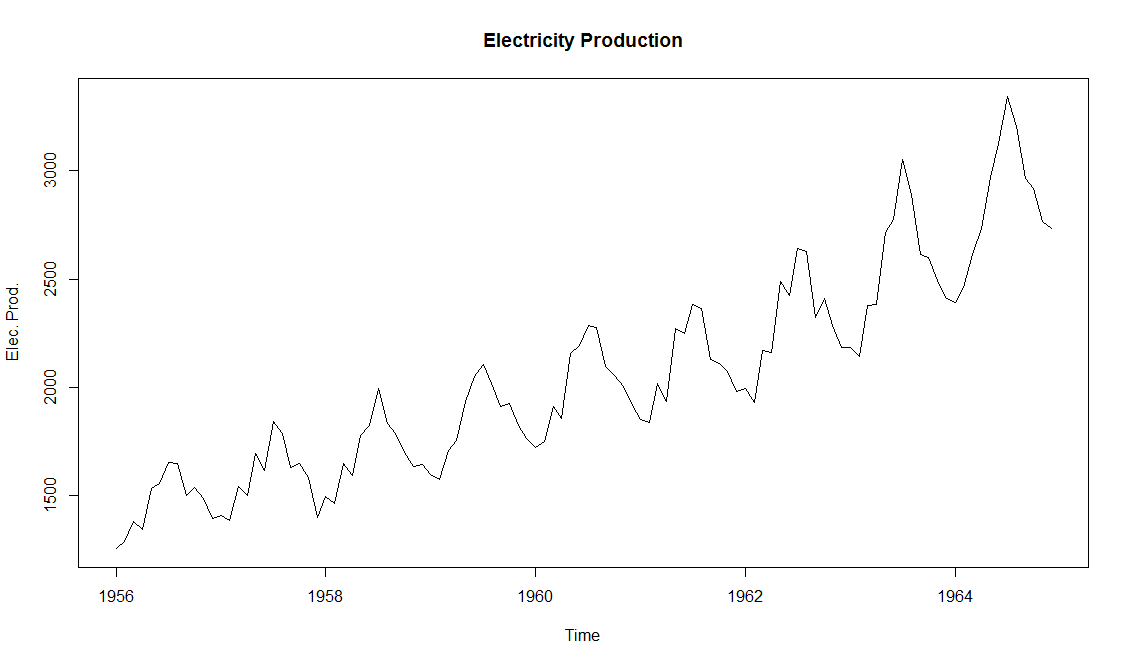
#1963 2184 2144 2379 2383 2717 2774 3051 2891 2613 2600 2493 2410

#1964 2390 2463 2616 2734 2970 3125 3342 3207 2964 2919 2764 2732

#1965 2622 2698 2950 2895 3200 3408 3679 3473 3154 3107 3052 2918

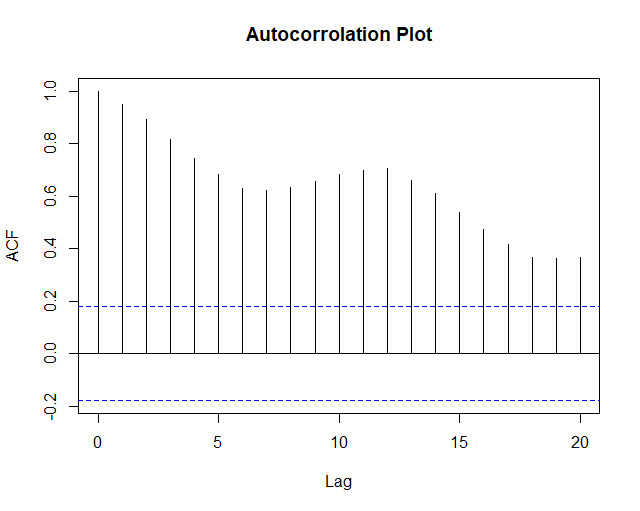
#Drawing time series plot of data

**plot(netdatats, main="internet usage")**



#Drawing autocorrelation plot of data

**acf(netdata,main="Autocorrolation Plot")**



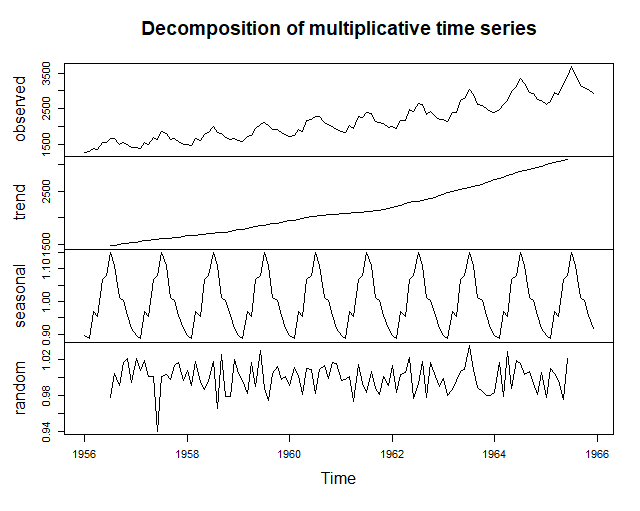
1. In part a, it can be seen that there are seasonality and an increasing trend between years 1956 and 1965.Also,variance is small in the first years according to the last years, that is,variance is increasing by time.Because of this reason, data shows us that it changes multiplicatively.

In part b, ACF is decreasing very slowly and remains well above the significance range (dotted blue lines). This is indicative of a non-stationary series and an indicative of sufficiently high correlation.

#Since variance increases with time, to decompose multiplicative type should be used

**netdatats\_dec<-decompose(netdatats,type = "multiplicative")**

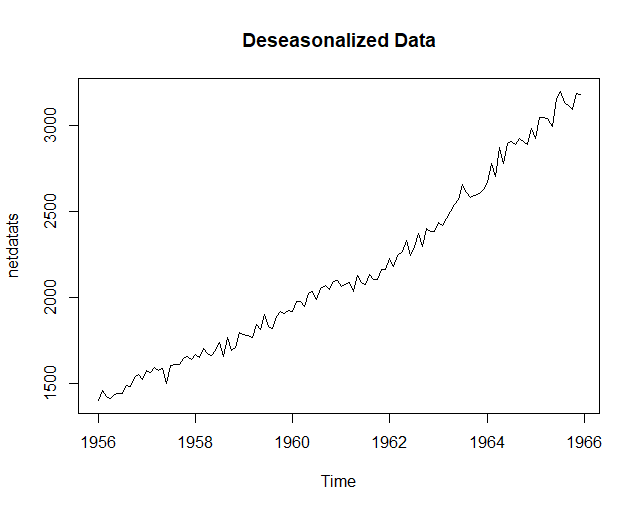
**plot(netdatats\_dec)**



#Deseasonalized data

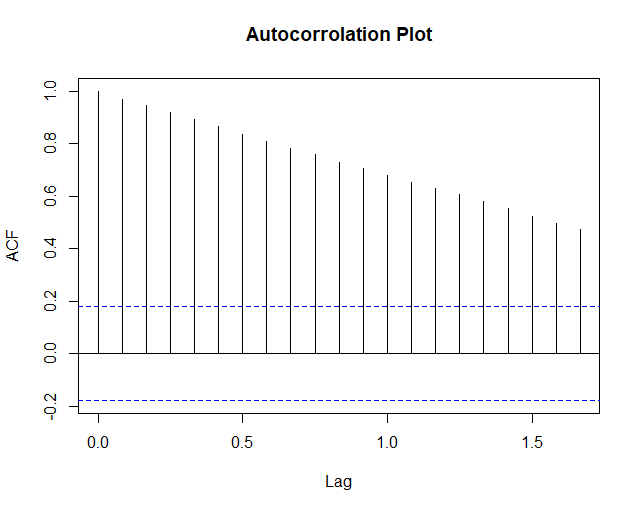
**desea\_data<-netdatats/netdatats\_dec$seasonal**

**plot(desea\_data, main="Deseasonalized Data")**



#Autocorrelation plot of deseasonalized data

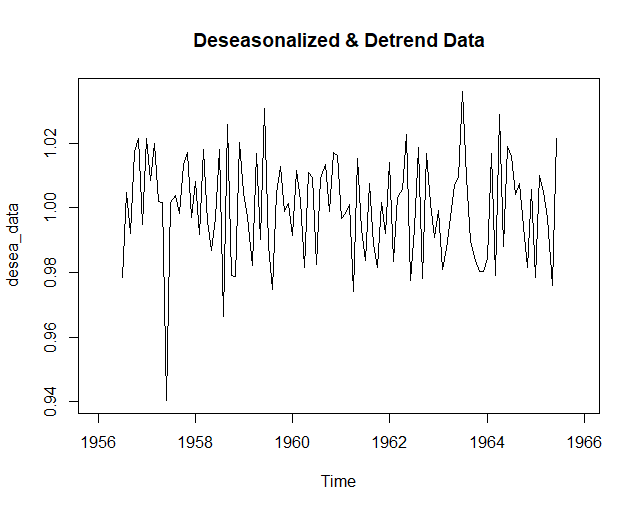
**acf(desea\_data,main="Autocorrolation Plot")**



#Remove trend from the deseasonalized data

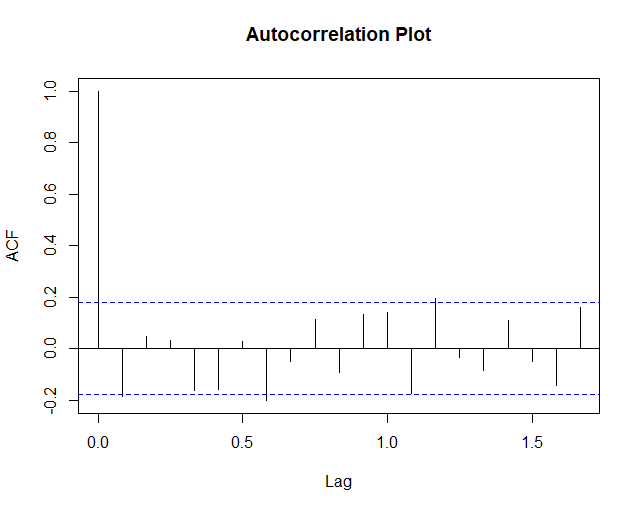
**detrend\_and\_desea\_data<-desea\_data/netdatats\_dec$trend**

**plot(detrend\_and\_desea\_data,main="Deseasonalized & Detrend Data")**



#Autocorrelation plot of deseasonalized and detrend data

**acf(detrend\_and\_desea\_data,na.action=na.pass,main="Autocorrelation Plot")**



1. When we extract seasonality from the data, the trend stays only and because of the reason that trend is highly correlated it is can be easily seen from the deseasonalized autocorrelation function.

When the trend is extracted from the data, the acf of deseasonalized and detrend data stays below the significance range. This is indicative of a stationary series.Therefore, the numbers are purely random and the correlation can be theoretically zero, or at least insignificant.

SİNAN DEMİRHAN 2016402330