Study of nature and stationarity of time series data of U.K. imports: goods and services (Pound millions) from 1960-1970.

**INTRODUCTION**

***Problem Description:*** Here in this problem we are interested in

1. Demonstrating the various steps that we need to understand about the nature of the given data.
2. Discussing the mathematical model for the given data and transform the data into a stationary series.
3. And we also want to check if the logarithmic transformation helps in achieving stationarity of the data set and justify our answer based on the result obtained.

***Objective:*** The main objective of this problem is to identify the nature of the given timeseries dataset and analyse the stationarity of the timeseries, take the required steps to convert into stationary series, if it is non stationary, and comment about the same.

*#Setting and getting the current working directory.*  
**setwd**("E:/M.Sc/SEM III/TIME\_SERIES\_ANALYSIS(MST371)/Practical Labs")  
**getwd**()

## [1] "E:/M.Sc/SEM III/TIME\_SERIES\_ANALYSIS(MST371)/Practical Labs"

***Data Description:***

The data set consists of Quarterly series of U.K. imports: goods and services (Pound millions) from 1960 – 1970. The data set has total 45 records of UK imports and and the year when the data is recorded. Also it is observed that in each year the data has been recorded in four quarters. There two variables in the the dataset i.e. time and imports: goods and services (Pound millions)

*#Loading the package required to load the dataset.*  
**library**(readxl)  
  
*#Loading the UK import dataset.*  
data <- **read\_excel**("E:/M.Sc/SEM III/TIME\_SERIES\_ANALYSIS(MST371)/data.xlsx")

## New names:  
## \* `` -> ...2

*#Obtaining the first few records of the dataset.*  
**head**(data)

## # A tibble: 6 x 2  
## Quarter ...2  
## <chr> <dbl>  
## 1 1960 Q1 1382  
## 2 1960 Q2 1417  
## 3 1960 Q3 1432  
## 4 1960 Q4 1438  
## 5 1961 Q1 1457  
## 6 1961 Q2 1403

**ANALYSIS:**

*#Extracing the data for the imports variable which we are interested in.*  
imports=data**$**...2  
  
*#Now converting the it into a time series data.*  
imports1=**ts**(imports)  
  
*#Here. we are checking if the dataset has been concverted into a timeseries plot.*  
**class**(imports1)

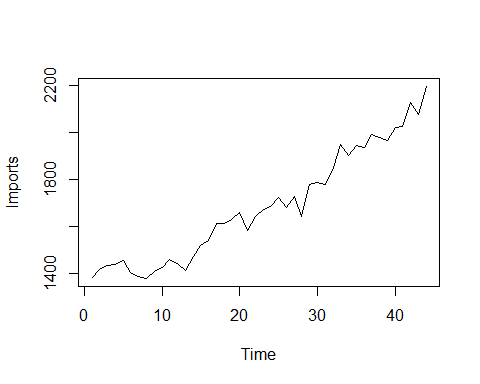
## [1] "ts"

Hence, now the dataset we are interested in is a timeseries data.

***Question 1***

***Demonstrate the various steps that you have learned to understand about the nature of the given data.***

*#Now we will first obtain the time series plot of the data to understand the nature of*  
**ts.plot**(imports1, gpars = **list**(xlab="Time",ylab="Imports",lty=**c**(1**:**20)))





***Interpretation:*** From the above time series plot (figure 1) we observe that there exists a trend component in the dataset since there is observed a increase or decrese pattern for a longer period of time. Also we observe that there is some kind of irregularity in the dataset hence we can say that there also exists a error component in the dataset.

*#loading the 'tseries' package.*   
**library**(tseries)

## Warning: package 'tseries' was built under R version 4.0.5

## Registered S3 method overwritten by 'quantmod':  
## method from  
## as.zoo.data.frame zoo

*#Now we want to check for the stationarity of the time series dataset using Augmented Dickey Fuller(ADF test) test.*  
**adf.test**(imports1)

##   
## Augmented Dickey-Fuller Test  
##   
## data: imports1  
## Dickey-Fuller = -2.5343, Lag order = 3, p-value = 0.3623  
## alternative hypothesis: stationary

***Interpretation:*** From the above test we observe that p value = 0.3623 > 0.5, hence we fail to reject the null hypothesis and conclude that the time series obtained is non stationary.

***Question 2***

***Discuss the mathematical model for the given data and transform the data in to a stationary series.***

In the above time series plot we observed that there exists trend and error component in the dataset hence we can directly go for a additive model since there does not exists any seasonal componenet in the dataset. Thus, the mathematical model is given by,

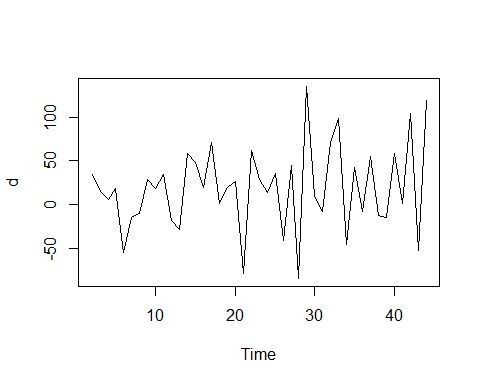
**Zt = f(mt, et) = mt + et**

where,

mt denotes the trend component.

et denotes the error component.

*#Now since our timeseries model is non stationary hence we try to covert it to stationary by the method of differencing.*  
d=**diff**(imports1)  
  
*#obtaining the time series plot for the new dataset.*  
**ts.plot**(d)



*#Now we want to check for the stationarity of the new time series dataset using Augmented Dickey Fuller(ADF test) test.*  
**adf.test**(d)

##   
## Augmented Dickey-Fuller Test  
##   
## data: d  
## Dickey-Fuller = -3.8136, Lag order = 3, p-value = 0.02772  
## alternative hypothesis: stationary

***Interpretation:*** From the above statistical test it is observed tha pvalue = 0.02772 < 0.05, thus we accept the null hypothesis and conclude that the timeseries have been converted into stationary series.

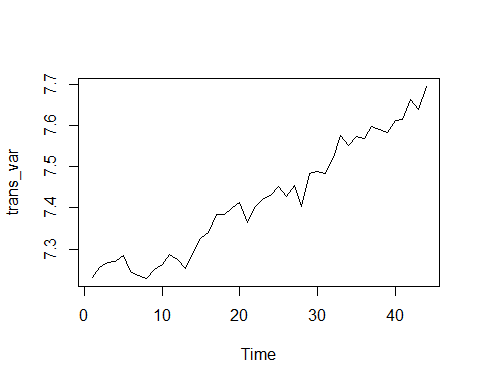
***Question 3***

***Will the logarithmic transformation helps in achieving stationarity of the data set? Justify your answer.***

*#Tranforming the variable by doing logarathmic transformation of the dataset.*  
trans\_var=**log**(imports1)  
trans\_var

## Time Series:  
## Start = 1   
## End = 44   
## Frequency = 1   
## [1] 7.231287 7.256297 7.266827 7.271009 7.284135 7.246368 7.236339 7.229114  
## [9] 7.249926 7.262629 7.286192 7.273786 7.254178 7.294377 7.326466 7.339538  
## [17] 7.384610 7.385231 7.397562 7.413970 7.365813 7.404279 7.421776 7.430114  
## [25] 7.451242 7.427144 7.453562 7.403670 7.482682 7.488294 7.483807 7.522941  
## [33] 7.574558 7.551187 7.573017 7.568896 7.596894 7.590852 7.583756 7.612831  
## [41] 7.613819 7.663877 7.639161 7.694848

*#Obtaining the timeseries plot of the transformed dataset.*  
**ts.plot**(trans\_var)



*#Now we want to check for the stationarity of the new time series dataset using Augmented Dickey Fuller(ADF test) test.*  
**adf.test**(trans\_var)

##   
## Augmented Dickey-Fuller Test  
##   
## data: trans\_var  
## Dickey-Fuller = -3.3173, Lag order = 3, p-value = 0.08181  
## alternative hypothesis: stationary

***Interpretation:*** From the above time series plot (figure 3) we observe that even after doing the logarathmic transformation the time series still exhibit the same nature and also on performing the Augmented Dickey-Fuller Test to check the stationarity we fail to reject the null hypothesis and conclude that the series is still non stationary. Thus making the logarathmic transformation does not make any difference in the stationarity of the timeseries.

**CONCLUSION:**

The final conclusion is that the timeseries data se of U.K. imports was having a trend as well as error component and it was non-stationary in nature. The mathematical model of the timeseries is additive in nature which is given by,

**Zt = f(mt, et) = mt + et**

where,

mt denotes the trend component.

et denotes the error component.

It was converted to stationary time series using the method of differencing in the 1st differencing.

We also observed that on making the logarithmic transformation in the original non- stationary time series data set does not make it a stationary dataset.