Transforming Non-Stationary Series into a Stationary Series by removing the Trend and Seasonality

Practical 3
Ananya Kaushal
2022-01-26

INTRODUCTION:

A static relationship requires inputs and outputs with constant parameters such as mean, median, and variance. In other words, algorithms perform best when the inputs and outputs are stationary. This is not the case in time series forecasting. Distributions that change over time can have unique properties such as seasonality and trend. These, in turn, cause the mean and variance of the series to fluctuate, making it hard to model their behavior. So, making a distribution stationary is a strict requirement in time series forecasting. In this article, we will explore several techniques to detect non-stationary distributions and convert them into stationary data.

AIM:

The aim of this experiment is to choose two appropriate datasets –

one with seasonality and one without and then convert them into stationary series.

PROCEDURE:

Importing the package 'astsa' and 'tseries' which contains our required datasets.

```
#Practical 3
library(astsa)
library(tseries)

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

About The Dataset:

The dataset 'Quarterly adjusted US GDP' is the Seasonally adjusted quarterly U.S. GDP from 1947 to 2018.

S No.	GDP
1	2033.061
2	2027.639
3	2023.452
4	2055.103
5	2086.017
6	2120.45
7	2132.598
8	2134.981
9	2105.562
10	2098.38

Table 1.1 containing a sample of values of the dataset ‘Quarterly adjusted US GDP’

The dataset ‘Johnson and Johnson Quarterly Earnings Per Share’ describes the Johnson and Johnson quarterly earnings per share, 84 quarters (21 years) measured from the first quarter of 1960 to the last quarter of 1980.

S No.	Earning
1	0.71
2	0.63
3	0.85
4	0.44
5	0.61
6	0.69
7	0.92
8	0.55
9	0.72
10	0.77

Table 1.2 containing sample values of the dataset ‘Johnson and Johnson Quarterly Earnings Per Share’

```
#Elimination of trend in the absence of seasonality
plot(gdp, main = "Quaterly US GDP")
```

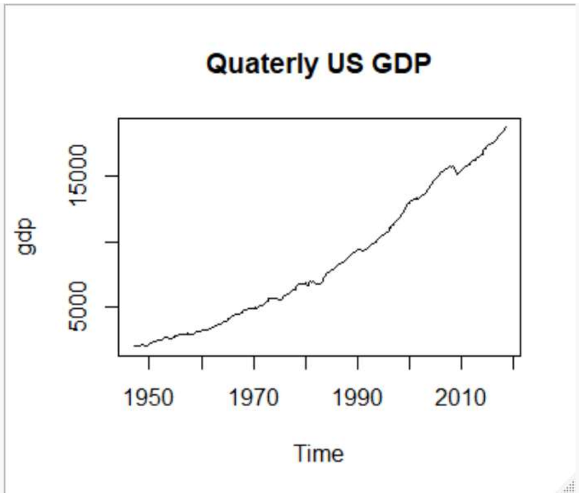

Figure 1: Quarterly GDP vs time

```
#We have a trend Component but it is free from a seasonal component, hence it is non-stationary.
#We can transform this series into a Stationary series using Method of differencing

dd=diff(gdp)
plot(dd)
```

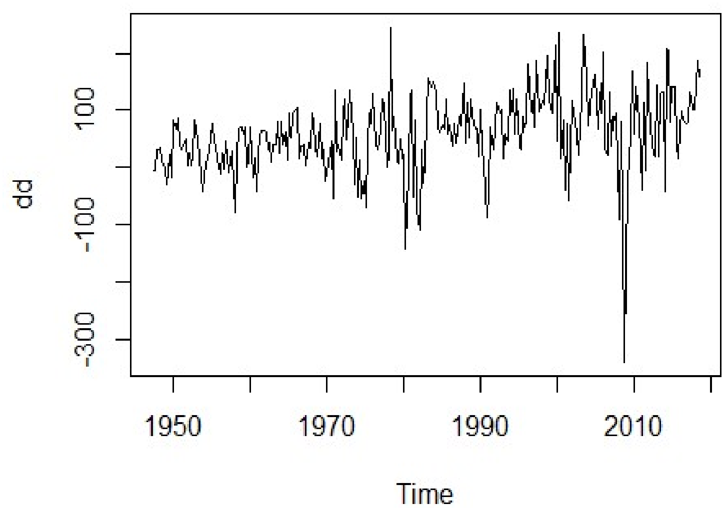

Figure 2: Plot after using method of differencing

```
adf.test(dd)

## Warning in adf.test(dd): p-value smaller than printed p-value

##
##  Augmented Dickey-Fuller Test
```

```
##
## data:  dd
## Dickey-Fuller = -5.7575, Lag order = 6, p-value = 0.01
## alternative hypothesis: stationary

#Since the p-value is less than 0.5 we can reject the null hypothesis and
#conclude that our differenced data is Stationary.


#Elimination of trend and seasonality

plot(jj, main = "Johnson and Johnson quarterly earnings per share")
```

Figure 3

```
#Transoforming Multiplicative to Additive
data=jj
datal=log(data)
plot(datal)
```

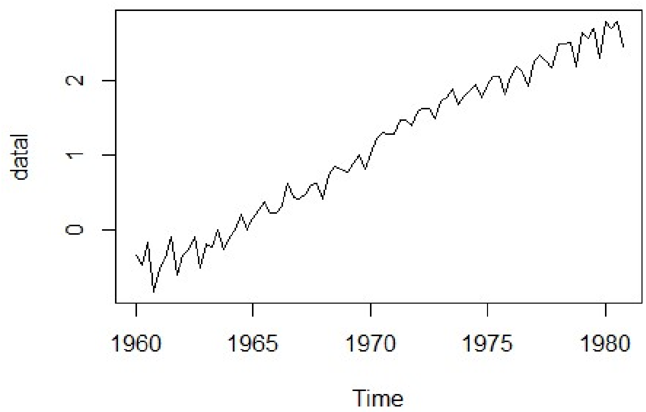

Figure 4

```
#Performing seasonal differencing on additive data
d=diff(datal,lag=12)
plot(d)
```

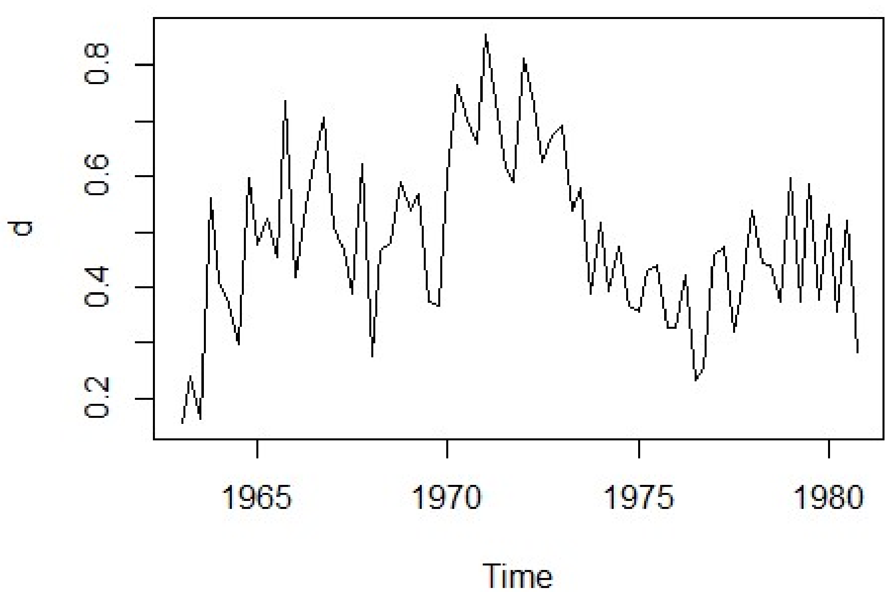

Figure 5

```
acf(d)
```

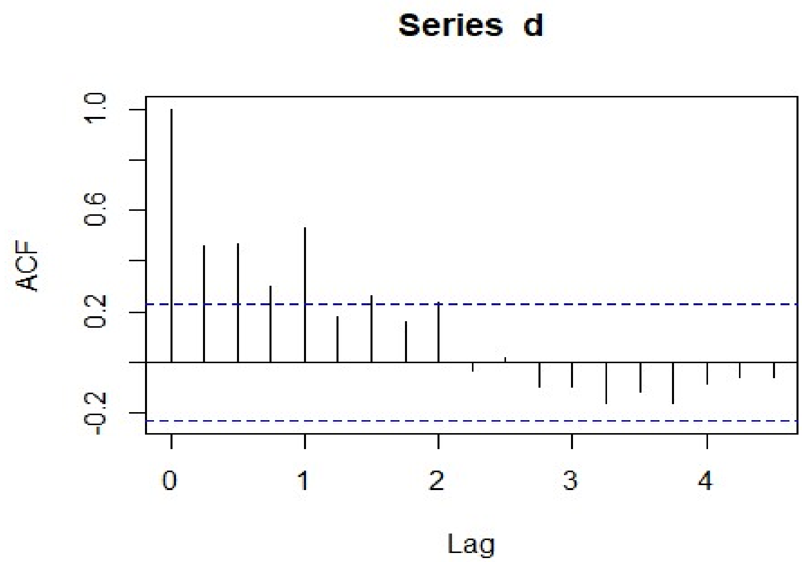

Figure 6

```
#We can see that no seasonality is present anymore.

#Removing the trend component
d2=diff(d)
plot(d2)
```

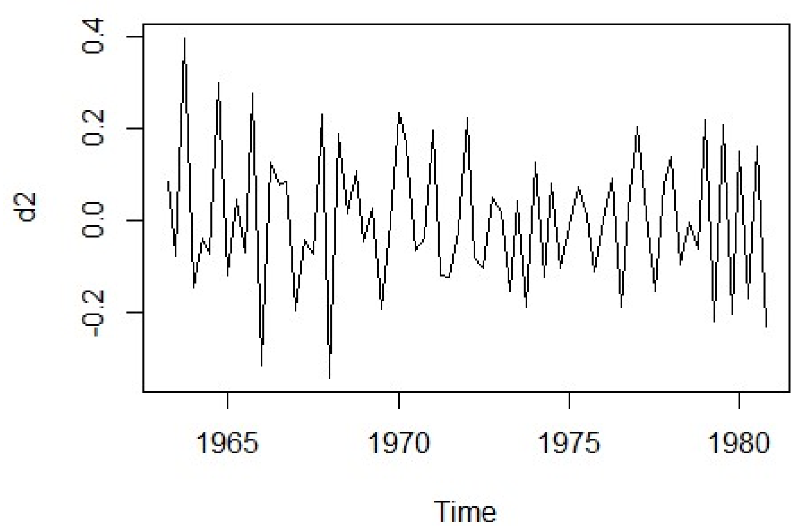

Figure 7

```r
adf.test(d2)

## Warning in adf.test(d2): p-value smaller than printed p-value

##
##  Augmented Dickey-Fuller Test
##
## data:  d2
## Dickey-Fuller = -4.2086, Lag order = 4, p-value = 0.01
## alternative hypothesis: stationary

#Since the p-value is less than 0.5 we can reject the null hypothesis and
#conclude that our differenced data is Stationary.
```

CONCLUSION:

From the above result we can see that after performing the adf test we get the p-values as less than 0.05 in dicating that our dataset are stationary. Hence, the datasets have successfully been made Stationary and th e Trend and Seasonality components have been removed.