

Name: Varun Sudhir

Reg No: 21BDS0040

Exploratory Data Analysis Lab Experiment 7.1

Write program for the following questions using Time Series Analysis

- 1) Load the in-built AirPassengers dataset from R. (data())

Code:

```
# Varun Sudhir 21BDS0040

# Load the AirPassengers dataset
data("AirPassengers")
```

- 2) Check for the Structure and the Data type of *AirPassengers*. (Str, Class)

Code:

```
# Varun Sudhir 21BDS0040

# Check the structure of the dataset
str(AirPassengers)

# Check the class (data type) of the dataset
class(AirPassengers)
```

Output:

```
> # Varun Sudhir 21BDS0040
> str(AirPassengers)
  Time-Series [1:144] from 1949 to 1961: 112 118 132 129 121 135 148 148 136 119 ...
> class(AirPassengers)
[1] "ts"
> |
```

- 3) Check for missing values in the dataset.

Code:

```
# Varun Sudhir 21BDS0040
```

```
# Check for missing values
any(is.na(AirPassengers))
print("Varun Sudhir 21BDS0040")
```

Output:

```
> # Varun Sudhir 21BDS0040
> # Check for missing values
> any(is.na(AirPassengers))
[1] FALSE
> |
```

4) Check for the Starting date and Ending date. (start, end)

Code:

```
# Varun Sudhir 21BDS0040

start(AirPassengers)
end(AirPassengers)
```

Output:

```
> # Varun Sudhir 21BDS0040
> start(AirPassengers)
[1] 1949    1
> end(AirPassengers)
[1] 1960   12
~ |
```

5) Check for the frequency of the dataset. (frequency)

Code:

```
# Varun Sudhir 21BDS0040

frequency(AirPassengers)
```

Output:

```
> # Varun Sudhir 21BDS0040
> frequency(AirPassengers)
[1] 12
```

6) Check for the summary of the dataset.

Code:

```
# Varun Sudhir 21BDS0040  
summary(AirPassengers)
```

Output:

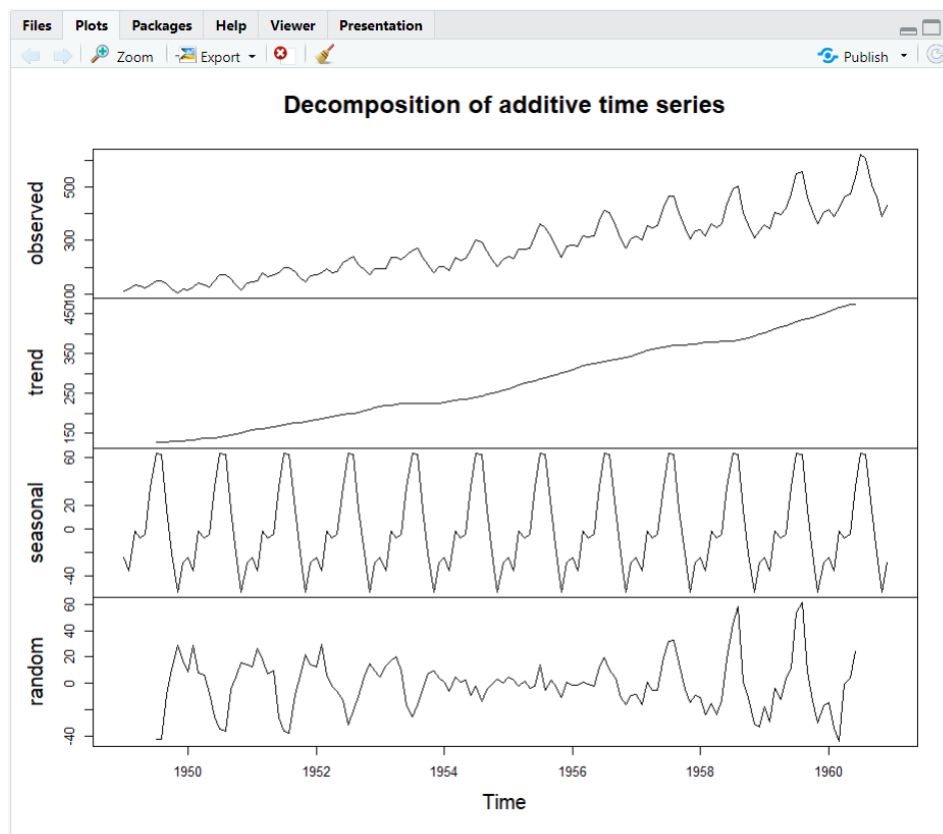
```
>  
> # Varun Sudhir 21BDS0040  
> summary(AirPassengers)  
      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
 104.0   180.0   265.5   280.3   360.5   622.0  
> |
```

7) Plot the decomposition of the dataset – Break data into trend, seasonal, and random.
(Simple plot function with decompose)

Code:

```
# Varun Sudhir 21BDS0040  
decomposed_data <- decompose(AirPassengers)  
plot(decomposed_data)
```

Output:



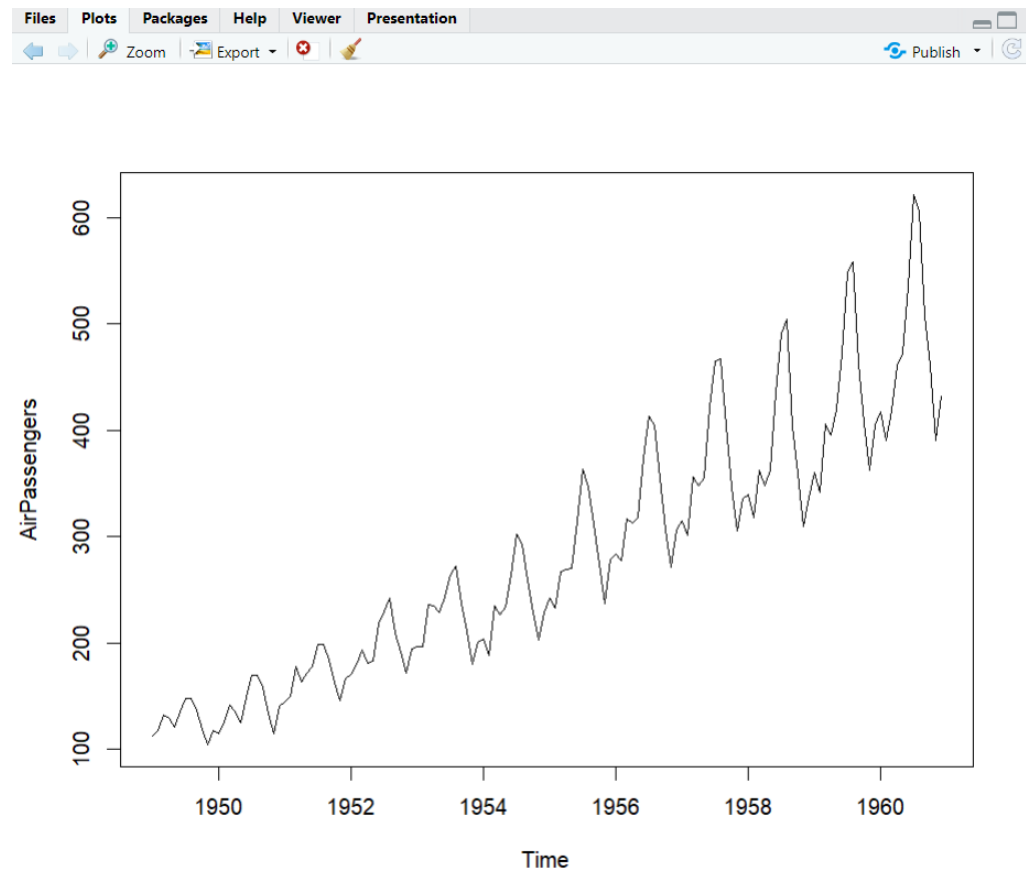
8) Plot the dataset (plot())

Code:

```
# Varun Sudhir 21BDS0040
```

```
plot(AirPassengers)
```

Output:



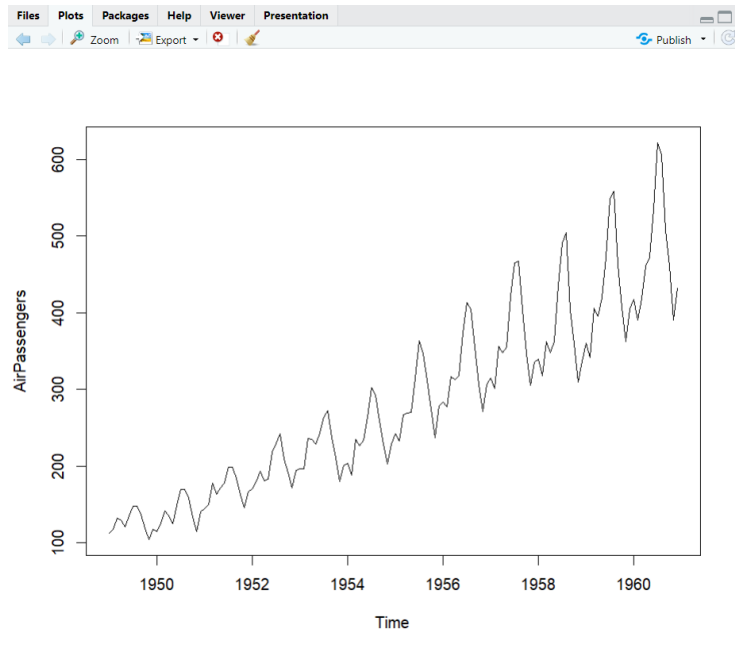
9) Plot the time series of the dataset (plot.ts())

Code:

```
# Varun Sudhir 21BDS0040
```

```
plot.ts(AirPassengers)
```

Output:

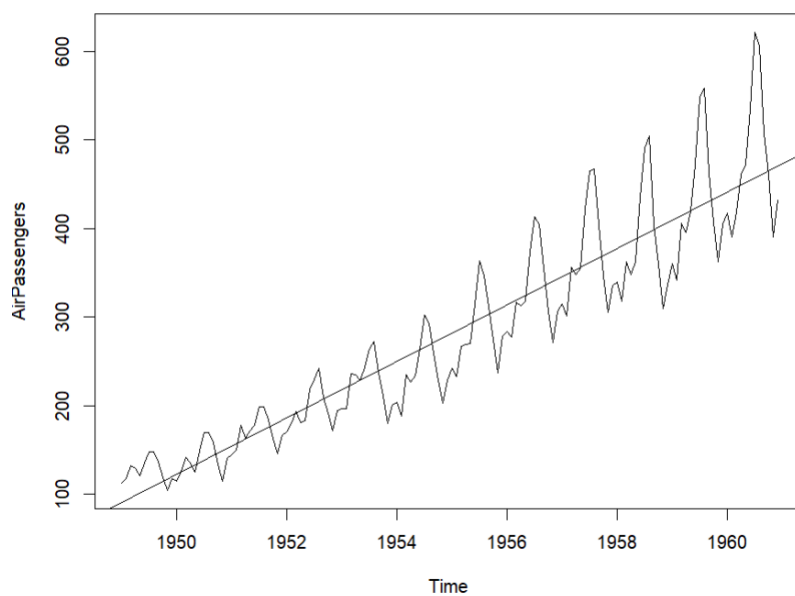


10) Draw the regressor line for the Q9. (`abline(lm(AirPassengers~time))`)

Code:

```
# Varun Sudhir 21BDS0040  
abline(lm(AirPassengers ~ time(AirPassengers)))
```

Output:



11) Print the cycle across the years for the dataset. (cycle())

Code:

```
# Varun Sudhir 21BDS0040
```

```
cycle(AirPassengers)
```

Output:

```
> cycle(AirPassengers)
      Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
1949   1   2   3   4   5   6   7   8   9  10  11  12
1950   1   2   3   4   5   6   7   8   9  10  11  12
1951   1   2   3   4   5   6   7   8   9  10  11  12
1952   1   2   3   4   5   6   7   8   9  10  11  12
1953   1   2   3   4   5   6   7   8   9  10  11  12
1954   1   2   3   4   5   6   7   8   9  10  11  12
1955   1   2   3   4   5   6   7   8   9  10  11  12
1956   1   2   3   4   5   6   7   8   9  10  11  12
1957   1   2   3   4   5   6   7   8   9  10  11  12
1958   1   2   3   4   5   6   7   8   9  10  11  12
1959   1   2   3   4   5   6   7   8   9  10  11  12
1960   1   2   3   4   5   6   7   8   9  10  11  12
```

12) Make the dataset to stationary ie. constant mean and variance and plot it.

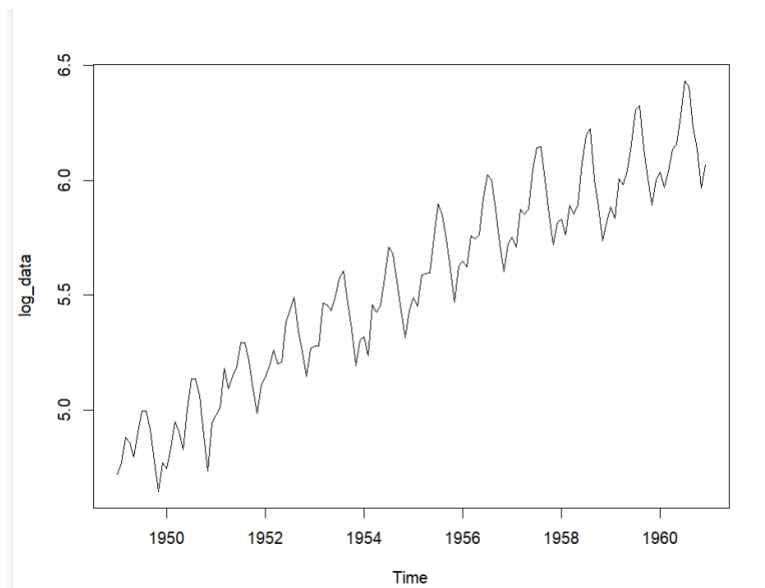
a) log(dataset)

Code:

```
# Varun Sudhir 21BDS0040
```

```
log_data <- log(AirPassengers)
plot(log_data)
```

Output:



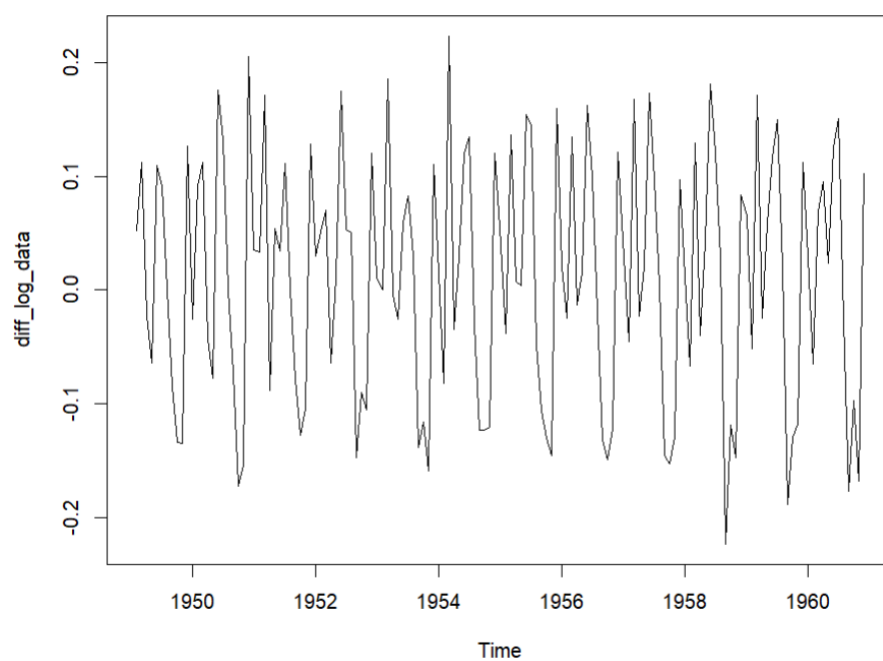
b) for Stationary mean ($\text{diff}(\log(\text{dataset}))$)

Code:

```
# Varun Sudhir 21BDS0040
```

```
diff_log_data <- diff(log_data)
plot(diff_log_data)
```

Output:



13) Plot box plot across months for seasonal effect(boxplot(dataset~cycle(dataset))

Code:

```
# Varun Sudhir 21BDS0040
```

```
boxplot(AirPassengers ~ cycle(AirPassengers),  
xlab="Month",ylab="Passengers")
```

Output:

