

## TimeSeries\_AirPassengers

```
## Main Source and inspiration == http://www.maths.adelaide.edu.au/emacs2009/  
#Introductory Time Series with R  
#by Paul S.P. Cowpertwait and Andrew Viggo Metcalfe  
#Springer. ISBN: 978-0-387-88697-8  
# All Text below within the QUOTE and UNQUOTE blocks is from this SPRINGER text mentioned above
```

```
library(ggfortify)
```

```
## Loading required package: ggplot2
```

```
library(tseries)  
library(forecast)  
data(AirPassengers)  
ts_AirPassengers <- AirPassengers  
class(ts_AirPassengers);head(ts_AirPassengers);tail(ts_AirPassengers);dim(ts_AirPassengers)
```

```
## [1] "ts"
```

```
##      Jan Feb Mar Apr May Jun  
## 1949 112 118 132 129 121 135
```

```
##      Jul Aug Sep Oct Nov Dec  
## 1960 622 606 508 461 390 432
```

```
## NULL
```

```
# Dimensions = NULL ??
```

```
# Check for Missing or NA
```

```
sum(is.na(ts_AirPassengers)) # No Missing values - No NA
```

```
## [1] 0
```

```
#
```

Check Frequency of TimeSeries and the Cyclic part of the TS

Check the summary of TS data

Plot Raw TimeSeries using the inbuilt base PLOT

...

...

```
frequency(ts_AirPassengers);cycle(ts_AirPassengers)
```

```
## [1] 12
```

```
##      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
```

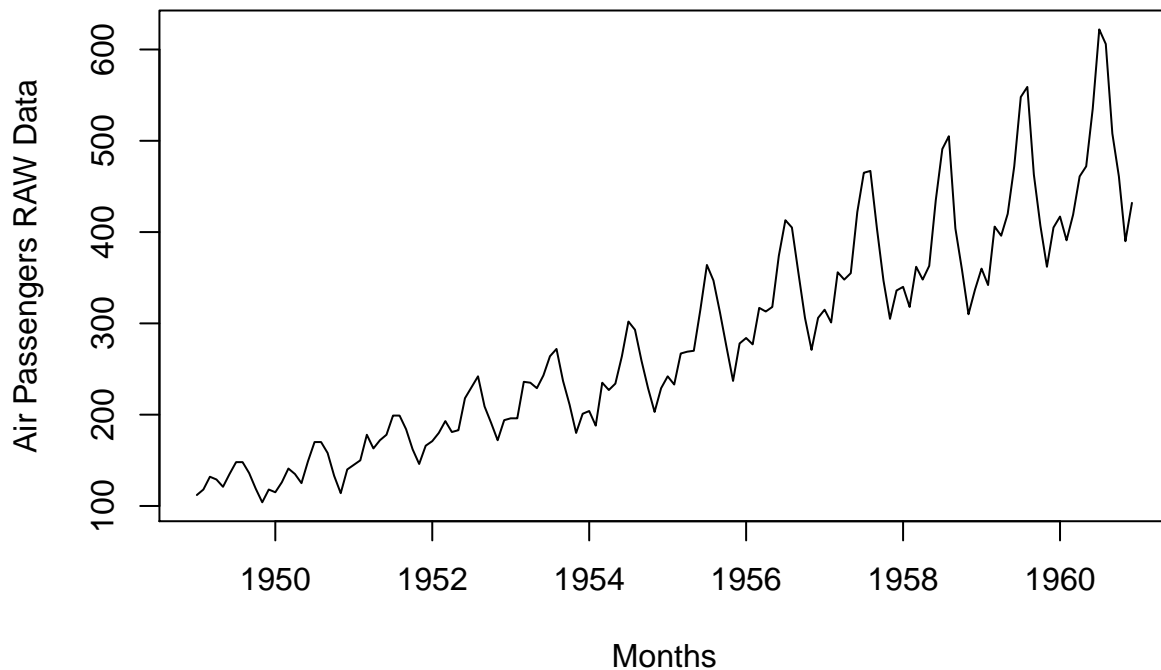
```
## 1956  1  2  3  4  5  6  7  8  9 10 11 12
## 1957  1  2  3  4  5  6  7  8  9 10 11 12
## 1958  1  2  3  4  5  6  7  8  9 10 11 12
## 1959  1  2  3  4  5  6  7  8  9 10 11 12
## 1960  1  2  3  4  5  6  7  8  9 10 11 12
```

```
summary(ts_AirPassengers)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  104.0   180.0   265.5   280.3   360.5   622.0
```

```
plot(ts_AirPassengers,xlab="Months", ylab = "Air Passengers RAW Data",
     main=("data(AirPassengers) Raw Data TimeSries Plot"))
```

### data(AirPassengers) Raw Data TimeSries Plot

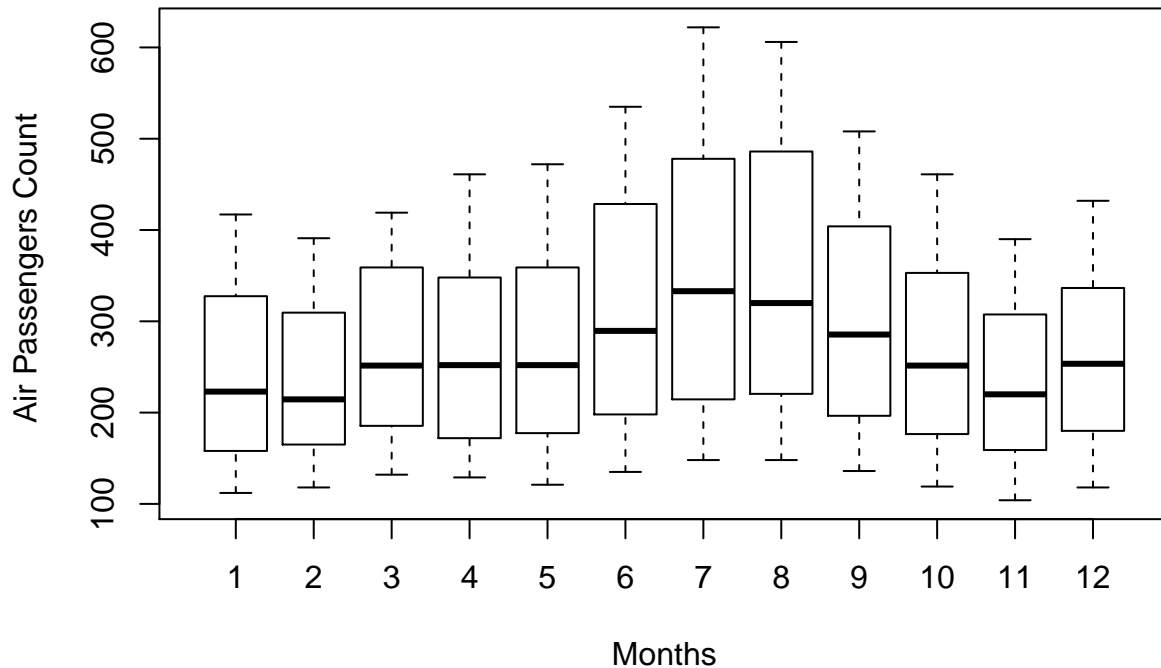


#QUOTE – A systematic change in a time series that does not appear to be PERIODIC is known as a TREND – UNQUOTE

Seasonality Plot – Air Passengers

```
boxplot(ts_AirPassengers~cycle(ts_AirPassengers),xlab="Months", ylab = "Air Passengers Count " ,
      main = "Seasonality Plot - Air Passengers ")
```

## Seasonality Plot – Air Passengers



A repeating pattern within each year is known as a - Seasonal Variation , or seasonality .  
This is true even if the period is not a year - seasonality can be within a time span of a month also.

QUOTE - Random or stochastic trends are common in Financial TimeSeries -  
thus Regression model is not appropriate for Fin TS . Further forecating is based  
on extrapolation (assumption that existing trends will continue )

Also in absence of external shock - the linear trend can be easily extrapolated.

Outliers are to be managed - robust methods of fitting models and imputation  
of missing values may be used .

- UNQUOTE

For imputation refer here = <https://www.kaggle.com/rohitdhankar/multiple-imputation-of-missing-values>

Package MICE - Multiple Imputation with MCA

Source URL's <http://juliejosse.com/wp-content/uploads/2016/06/user2016.pdf>

<https://arxiv.org/pdf/1606.05333v2.pdf>

[http://www.ats.ucla.edu/stat/r/faq/R\\_pmm\\_mi.htm](http://www.ats.ucla.edu/stat/r/faq/R_pmm_mi.htm)