

Let's first start with installing the required packages.

```
> .libPaths("E:\\Frequently Used\\Research\\Current Research\\R\\library")
> #.libPaths("/home/nabeel/myStuff/library")
> #.libPaths("C:\\Users\\rushadf\\Dropbox\\library")
> require(forecast)
> require(fpp)
```

1 Chapter 6

Here we learn about Trend, Seasonality and Cyclic behavior. Trend: long term increase or decrease in data. Seasonality: When a series is influenced by seasonal factors. Cyclic: when data exhibits rise and fall that are not of fixed period. Duration of fluctuation is at least two years. Now the following R code shows different type of these patterns.

This monthly data on housing sales. We can have some fair idea about the nature of the time series data by the following simple command.

```
> str(hsales)

Time-Series [1:275] from 1973 to 1996: 55 60 68 63 65 61 54 52 46 42 ...
```

This tells us that it is a time series object of 275 observation from 1973 to 1996. Now we want to know the periodicity of the data.

```
> frequency(hsales)
```

```
[1] 12
```

It produces the number 12 which means it is monthly data. We can also know about the frequency from the following command.

```
> tsp(hsales)

[1] 1973.000 1995.833 12.000
```

To know about the exact time of start and end, we can use the following command

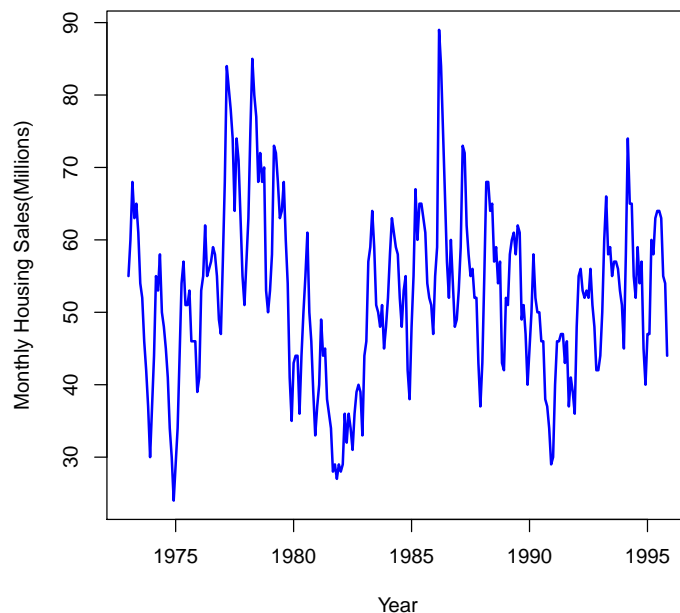
```
> start(hsales)
```

```
[1] 1973 1
```

```
> end(hsales)
```

```
[1] 1995 11
```

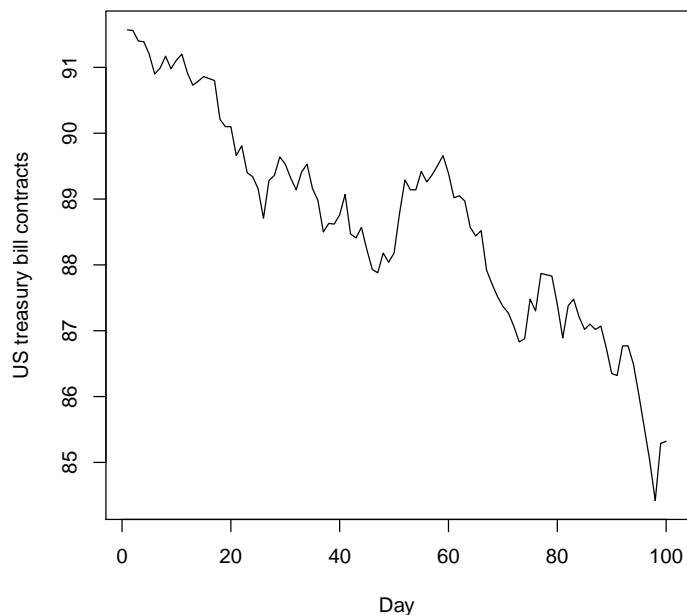
```
> plot(hsales, col="blue",lwd=2,xlab="Year", ylab="Monthly Housing Sales(Millions)")
```



One thing obvious is that there is no trend. First I thought there is no apparent cycle in there. But here we see that from mid 1970s to early 80s there is some upward trend, then again downward trend. So we can say every 6/10 years there is some cyclical pattern in the housing sale. The book says that here is strong seasonality. I don't know why it says that. The seasonality might result from the strong spikes shown almost every year. That might mean that within each year, there is some months when the housing sales are higher. That indicates that seasonal factors are affecting the housing sales. May be during Summer time this is happening.

Now we draw a graph of US treasury bill contracts in 1981 for 100 consecutive days. I am not sure what the number is saying. Is it the amount of volume contracts in dollar or what?

```
> plot(ustreas, xlab="Day", ylab="US treasury bill contracts")
```



We see a clear downward trend here. Since it is daily data, no apparent seasonality is obvious here. If we had longer data, maybe we would have seen that this downward trend is actually part of a large cycle.

Now we come to Australian monthly electricity consumption data.

```
> plot(elec,xlab="Year",ylab="Australian monthly electricity production")
```

We clearly see an upward trend. We also see spikes in there. This indicates that there is strongly seasonality within each year. Again maybe during summer months the electricity consumption goes up.

2 Time Series Decomposition

We can think of any time series y_t is composed of three components:

- Seasonal
- Trend-Cycle
- Remainder random error

If we think of an additive model, then we write

$$y_t = S_t + T_t + E_t$$

Part I

Chapter 8: ARIMA Models

Apart from Exponential smoothing, ARIMA models provide another way to forecast time-series data. While exponential smoothing tries to use the trend and cycle pattern, ARIMA models use autocorrelation in the data.

3 Stationarity and Differencing

A stationary time series is which properties does not depend on the time at which the series is observed. As a result, the series which is trending and has seasonal pattern is not stationary. On the other hand, a white noise series is stationary, since it does look the same at any point of time.

There might be some confusion regarding the cyclical pattern. Trend might be a part of the cycle. Therefore it is difficult to predict at certain point the nature of the time series if it is showing cyclical pattern. Therefore cyclical time series can be stationary.

Therefore a stationary time series will have no predictable pattern in the long run. Time plot will roughly showing horizontal pattern with constant variance. There are few plots which gives us a fairly good idea about the different types of plot.

3.1 Differencing

While Dow-Jones index itself is non-stationary, it's change is stationary. It is vindicated by the autocorrelation function (ACF). But the point is that if we have non-stationary series, then it can be made stationary by taking difference. Then it is called difference stationary process (DSP).

3.1.1 Random Walk Model

A very common example of difference stationary process is Random Walk. Here

$$y_t = y_{t-1} + \epsilon_t$$

is Random walk model where $\epsilon_t \sim N(0, \sigma^2)$. Now if we take the first difference

$$\Delta y_t = y_t - y_{t-1} = \epsilon_t$$

is white noise which is obviously stationary.

A random walk model is very widely used in economics and finance literature. It is characterized by two things:

- A long upward or downward trend.
- Sudden change in the direction of the trend.

Author says that best forecast that is available from random walk model is equal to value from last observation. That is because the next observation is equally likely to go up or down. Therefore the best prediction is the observation from the last time period.

3.1.2 Random Walk Model with a drift

We have seen that first difference in a random walk model has zero mean. A closely related model is that when difference has a non-zero mean. It can be expressed as:

$$\begin{aligned}y_t &= c + y_{t-1} + \epsilon \\y_t - y_{t-1} &= c + \epsilon \\E(\Delta y_t) &= c + E(\epsilon) \\E(\Delta y_t) &= c \quad \text{since} \quad E(\epsilon) = 0\end{aligned}$$

3.2 Second-order Differencing

$$\begin{aligned}y_t'' &= y_t' - y_{t-1}' \\&= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \\&= y_t - 2y_{t-1} + y_{t-2}\end{aligned}$$

In this case y_t'' has $T - 2$ observations. Here we are modelling "change in the changes". Usually, it is hardly necessary to go beyond second differences in empirical models.

3.3 Seasonal Difference

A seasonal difference is the difference between an observation and the corresponding observation in the same season last year. We can write it in the following way:

$$y_t' = y_t - y_{t-m}$$

where m = no. of seasons in a year