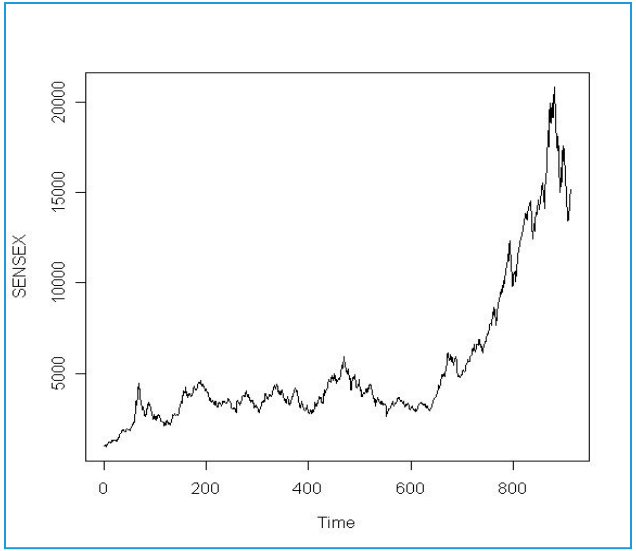
Introduction to

Time Series Analysis

Contents

What is Time Series ?

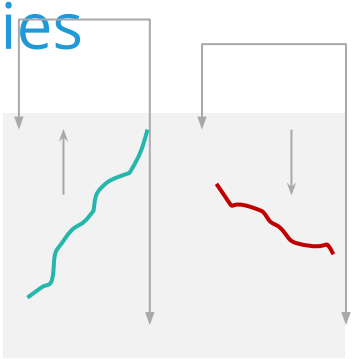
Time Series is a sequence of values observed over time

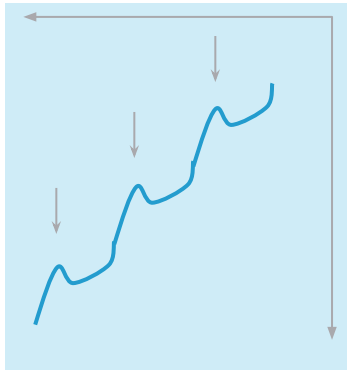


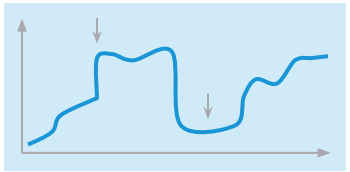
Types of Time Scale	
Discrete	Value changes after jumping from one time period to other. Example: Dow Jones Index -End of Day Values
Continuous	Value changes within an infinitely short amount of time. Example. Temperature, Dow Jones Index Tracked Real Time

Components of Time Series

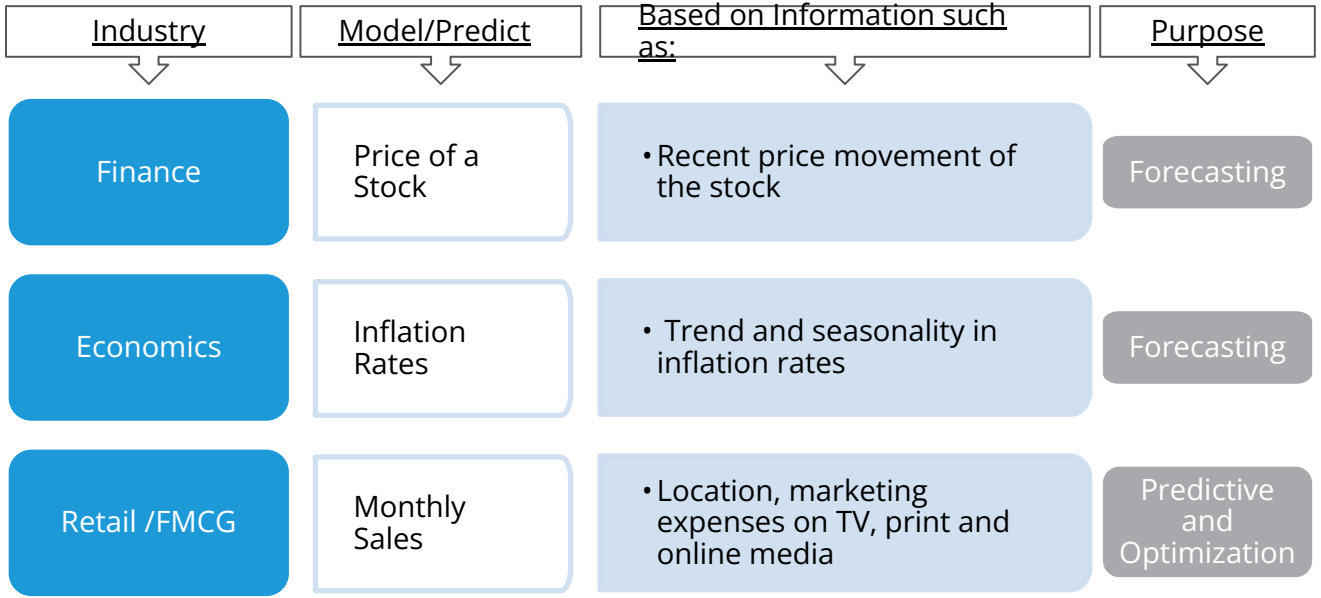
Trend	Long-term increase or decrease in the time series. There may be increase/decrease in short term but overall trend in the long term can be increasing or decreasing.
Seasonality	**Predictable and recurring trends and patterns** over a period of time, normally a year. An example of a seasonal time series is retail data, which sees spikes in sales during holiday seasons like Christmas. Seasonality is reflected only when data is available for more than one year
Cyclic Pattern	Exists when data exhibit rises and falls that are not of fixed period. The duration of these fluctuations is usually of at least 2 years







Application Areas



R packages

R has robust infrastructure for handling and managing time series data. The base R has function “ts” to handle time series data. However, there are many other packages in R to work with time series data.

Package	Features
zoo	• R package for regularly and irregularly spaced time series
xts	• Based on zoo and provides uniform handling of R's different time-based data classes
timeSeries	• Recommended for managing financial time series objects

Case Study

Background

- Annual Sales for a specific company from year 1961 to 2017

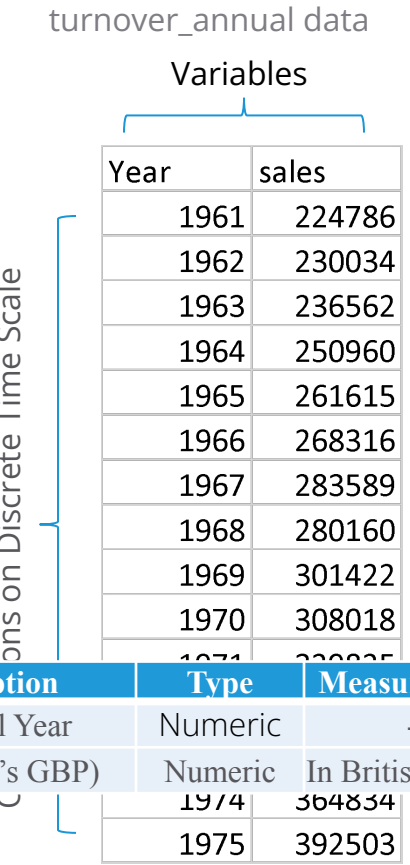
Objective

- To plot a time series object

Available Information

- Number of cases: 57
- Variables: Year, sales(in 10's GBP)

Data Snapshot



Year	sales
1961	224786
1962	230034
1963	236562
1964	250960
1965	261615
1966	268316
1967	283589
1968	280160
1969	301422
1970	308018
1974	364834
1975	392503

Columns	Description	Type	Measurement	Possible values
Year	Financial Year	Numeric	-	-
sales	sales(in 10's GBP)	Numeric	In British Pound	Positive values

Time Series in R

```
# Import turnover_annual Data
turnover_annual <- read.csv("turnover_annual.csv",header=TRUE)
```

- Performing any time series related operations require the data type to be time-based.
- We will use the most basic function **ts()** for converting our data frame object into time series object.

```
#Creating a Time Series Object
salesseries<-ts(turnover_annual$sales,start=1961,end=2017)
```

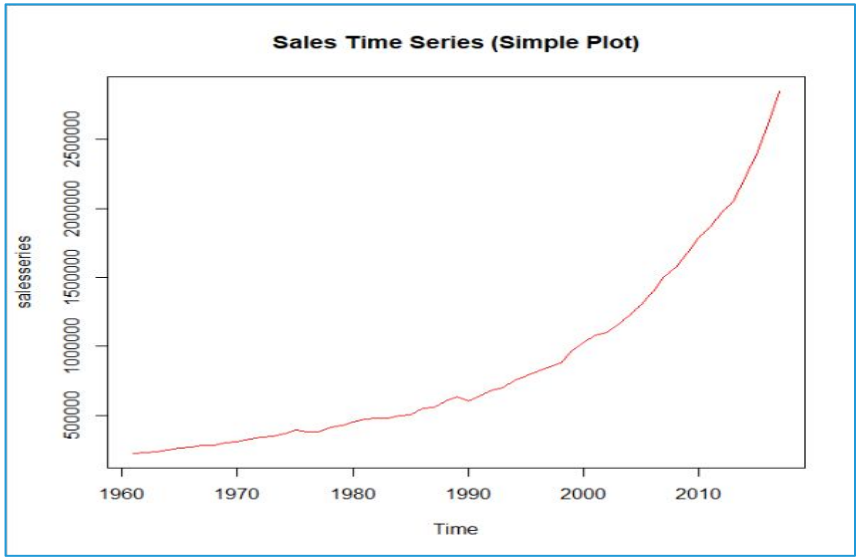
ts() from package **stats** in R, is a very important function and is used to create time series object. We use **ts()** to convert a column from our data frame **turnover_annual** to time series object named **salesseries.**
The **start=** and **end=** arguments specify first and last period
When the time series has seasonal component, argument **frequency =** can be included in **ts()**. It denotes number of observations per unit of time. Ex. If data is quarterly: **frequency = 4**, if data is monthly: **frequency = 12**.
Frequency argument is not required for annual data.

Plotting Time Series in R

```
# Plotting a Time Series Object
plot(salesseries,col="red",main="Sales Time Series (Simple Plot)")
```

plot() generates a simple line chart.

```
# Output
```



Interpretation :

- The time-series clearly shows upward trend.

Subsetting Time Series in R

- Large volumes of data are required for most real world analytics, time series is no exception.
- Subsetting is an important tool as it facilitates partitioning the data within R for micro-level specific analysis.

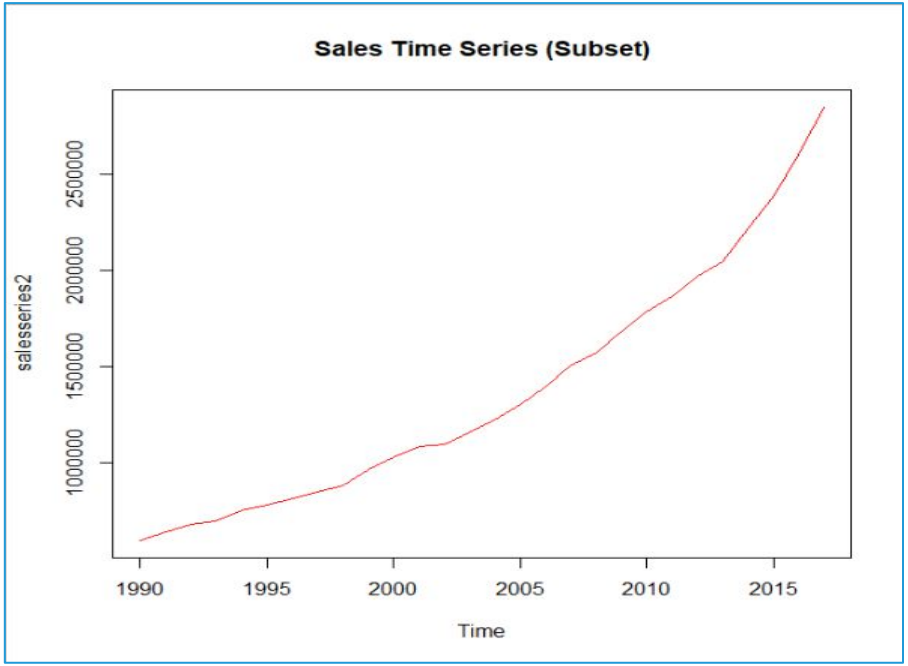
```
# Subsetting a Time Series Object
salesseries2<-window(salesseries,start=1990,end=2017)
```

window() is a generic function which extracts the subset of the object x observed between the times **start** and **end**.

Subsetting Time Series in R

```
plot(salesseries2, col="red", main="Sales Time Series (Subset)")
```

```
# Output
```



Interpretation :

- Upward trend is observed for period 1990 to 2017

Irregular Time Series

- The time series considered so far is ”Regular” time series. It is observed after regular time interval of one year.
- **However, not all time series are evenly spaced**. For instance, ATM withdrawal data of a customer is a time series, but the transactions may not happen after uniform intervals like daily, weekly, monthly.
- **Irregular time series** can be found in a number of industries and domains like ecology (natural disasters), finance (stock prices), clinical research (patient's reactions), astronomy (movement of celestial objects), to name a few.
- But with advancements in computer technology, such uneven time series can be analyzed in their original form.

Handling Irregular Time Series in R :

- It is impossible to coerce irregular time series into **`ts()`**, because the arguments of start, end and frequency are irrelevant.
- R has two highly effective packages for handling unevenly spaced time series, zoo and xts.

Quick Recap

Time Series	• Sequence of values measured over time • Time scale can be discrete or continuous
Time Series in Analytics	• Used for simple forecasting or building predictive model on time series data
Time Series in R	• R has different packages and functions for handling different types of time series • **`ts()`** in the basic stats package is the most commonly used function • A time series object can also be subsetted using **`window()`** from stats package
Irregular Time Series	• Time Series which are not evenly spaced. • R has two **highly effective packages** for handling unevenly spaced time series, **zoo and xts.**