Time Series

Presented by
Mr Gourav Manjrekar
M.Sc. (Statistics) Gold Medalist
www.youtube.com/GouravManjrekar
Andriod App: Gourav Manjrekar Academy

Meaning

Time series refers to a series in which one variable is time.

Values of variables are chronologically arranged.

Introduction

Time-Series Data

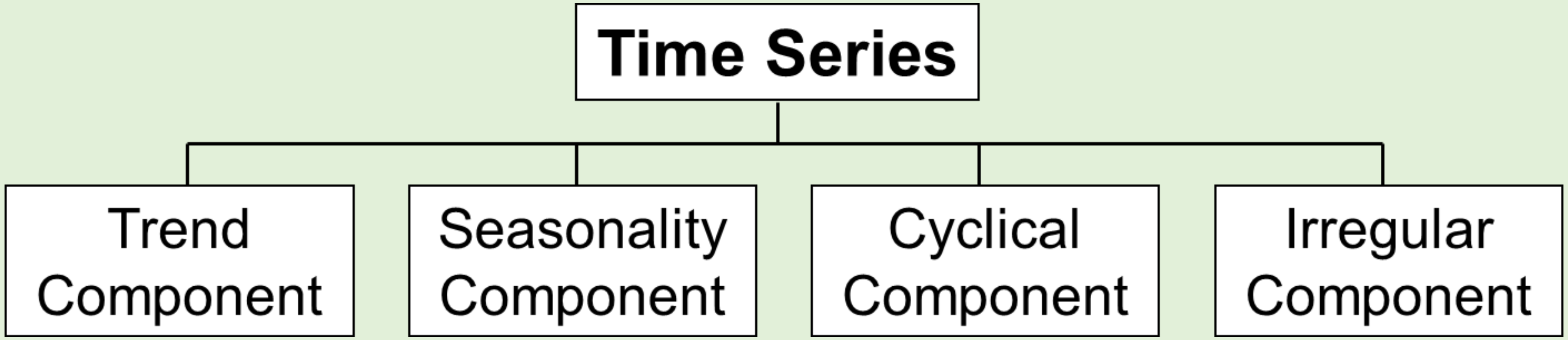
- Numerical data ordered over time
- The time intervals can be annually, quarterly, daily, hourly, etc.
- The sequence of the observations is important
- Example:

Year:	2008	2009	2010	2011	2012
Sales:	75.3	74.2	78.5	79.7	80.2

Purpose

- Making forecast for future
- For evaluating past performances

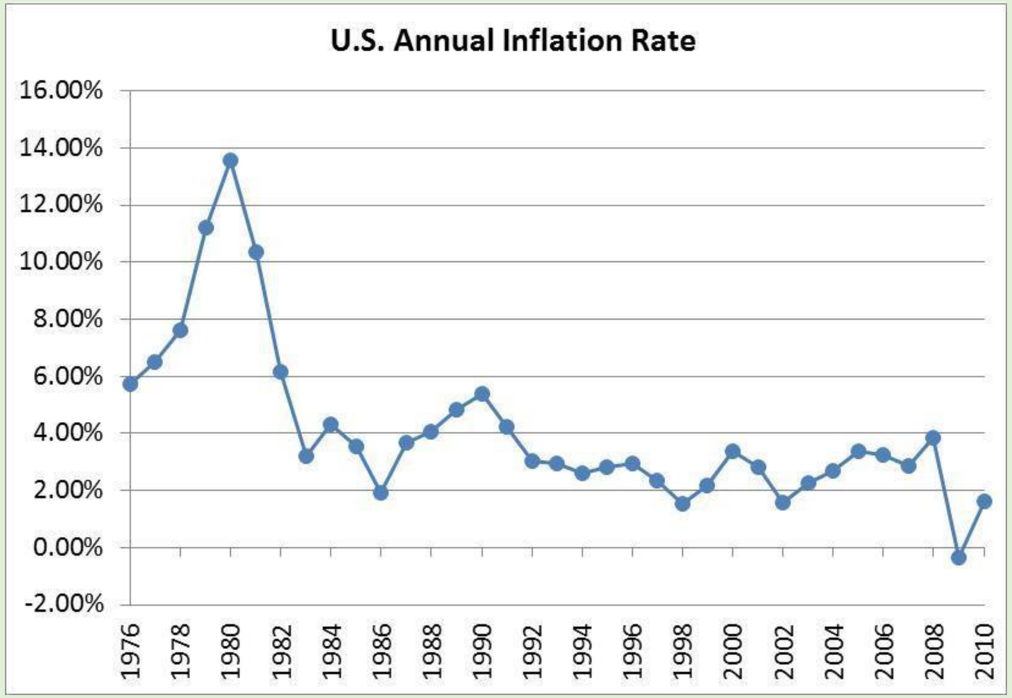
Components of a Time Series



Time-Series Plot

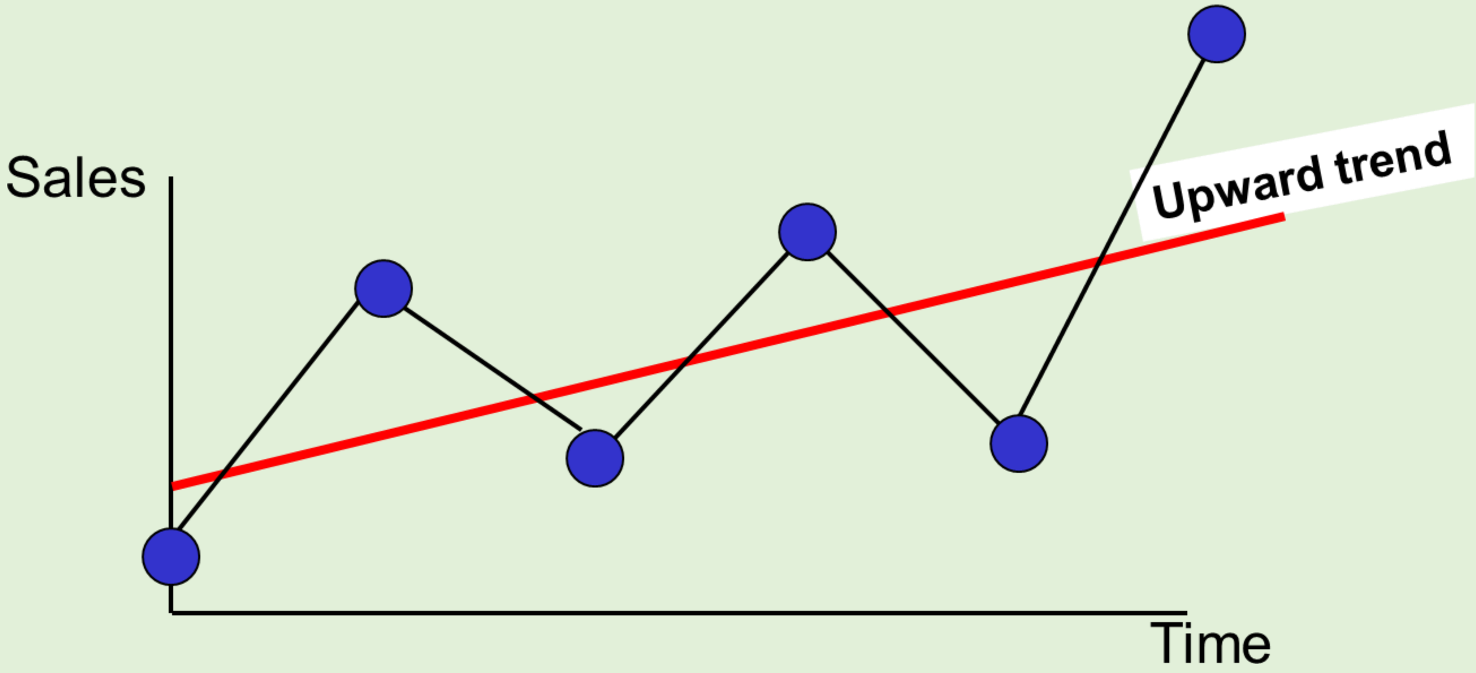
A time-series plot is a two-dimensional plot of time-series data

- the vertical axis measures the variable of interest
- the horizontal axis corresponds to the time periods



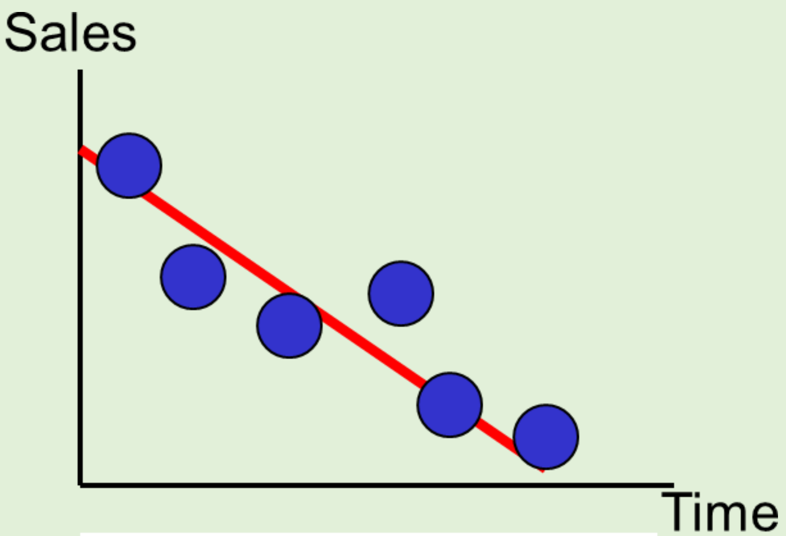
Trend Component (1 of 2)

- Long-run increase or decrease over time (overall upward or downward movement)
- Data taken over a long period of time

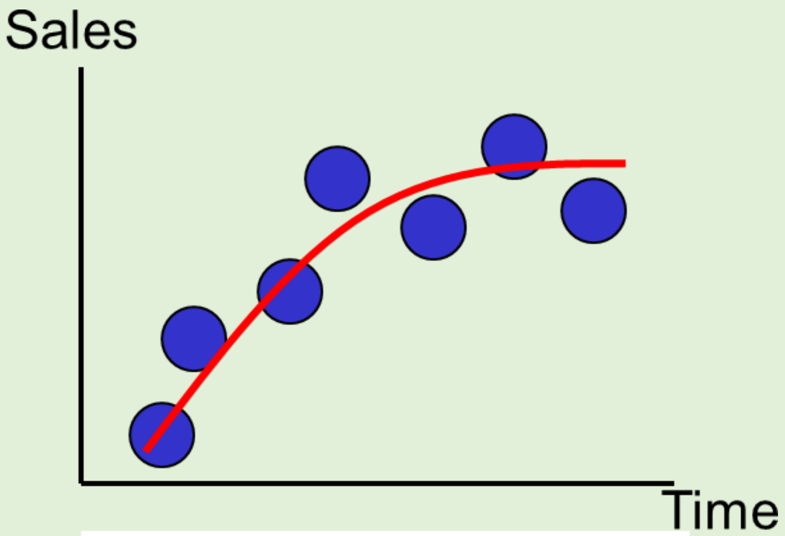


Trend Component (2 of 2)

- Trend can be upward or downward
- Trend can be linear or non-linear



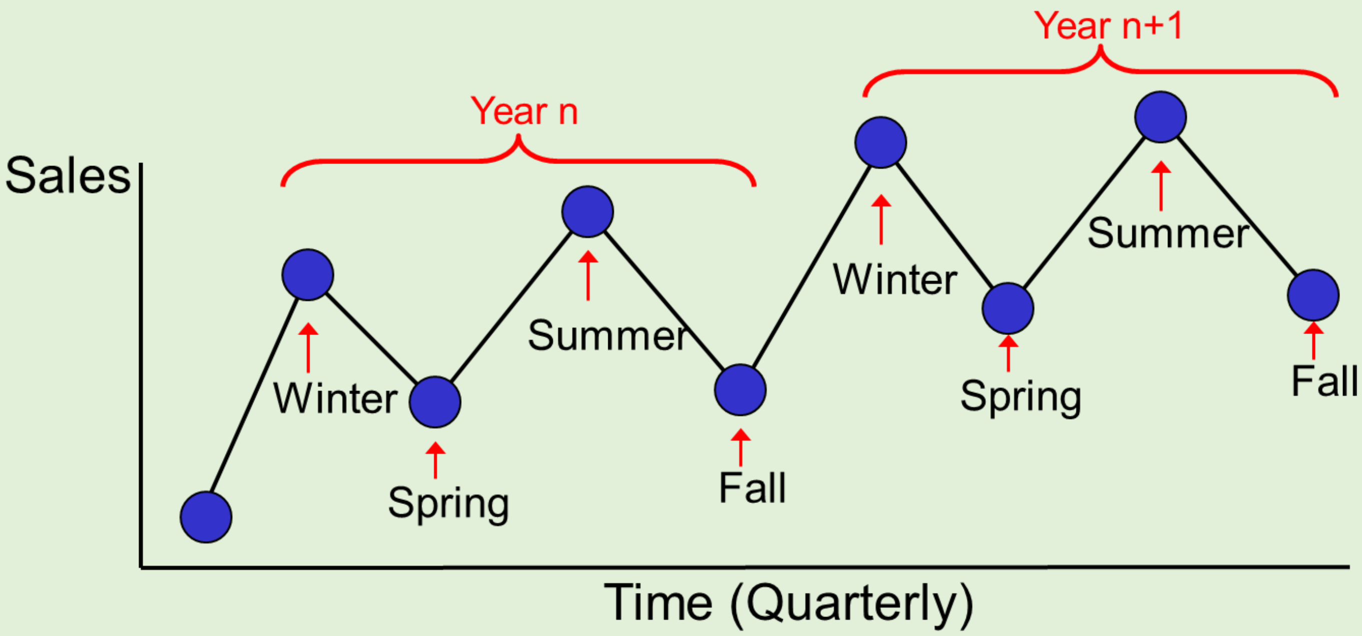
Downward linear trend



Upward nonlinear trend

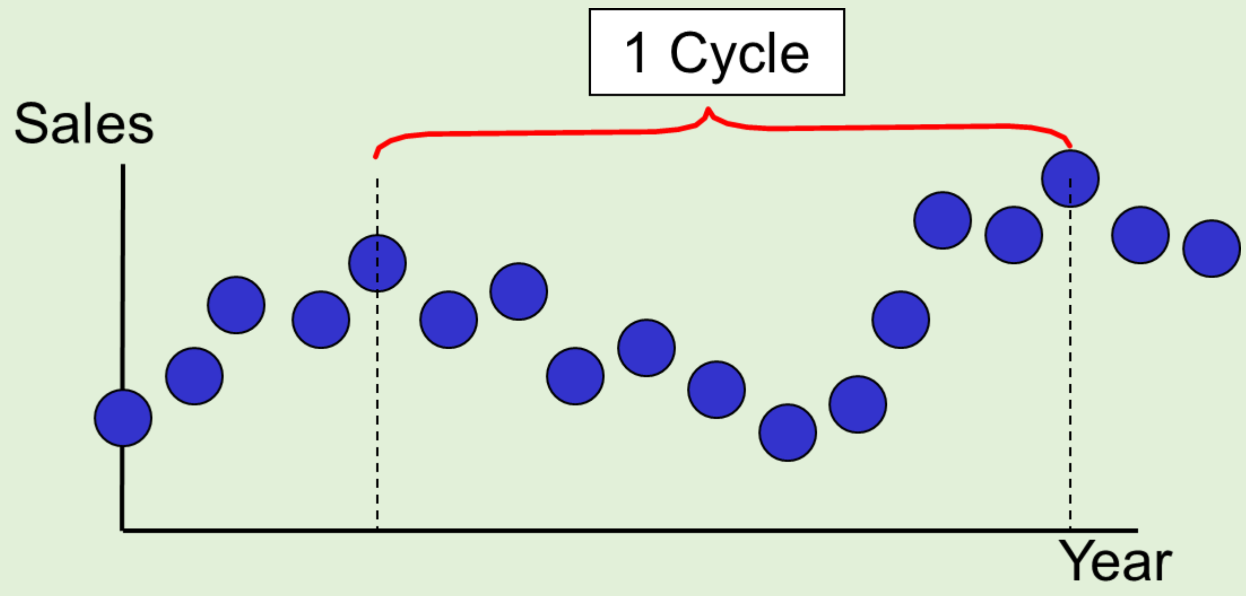
Seasonal Component

- Short-term regular wave-like patterns
- Observed within 1 year
- Often monthly or quarterly



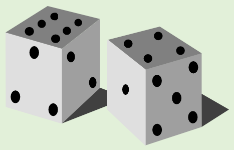
Cyclical Component

- Long-term wave-like patterns
- Regularly occur but may vary in length
- Often measured peak to peak or trough to trough



Irregular Component

- Unpredictable, random, “residual” fluctuations
- Due to random variations of
  - Nature
  - Accidents or unusual events
- “Noise” in the time series



Time-Series Component Analysis

- Used primarily for forecasting
- Observed value in time series is the sum or product of components
- Additive Model

$$X_t = T_t + S_t + C_t + I_t$$

- Multiplicative model (linear in log form)

$$X_t = T_t S_t C_t I_t$$

where X_t = value of the time series at time *t*

T_t = Trend component at period *t*

S_t = Seasonality component for period *t*

C_t = Cyclical component at time *t*

I_t = Irregular (random) component for period *t*

Least Square method

- This is one of the most popular methods of fitting a mathematical trend.
- The fitted trend is termed as the best in the sense that the sum of squares of deviations of observations, from it, is minimized.
- This method of Least squares may be used either to fit linear trend or a nonlinear trend (Parabolic and Exponential trend).

FITTING OF LINEAR TREND

- Given the data (Y_t, t) for n periods, where t denotes time period such as year, month, day, etc. We have to the values of the two constants, ‘a’ and ‘b’ of the linear trend equation:

$$Y_t = a + b * t$$

- Where the value of ‘a’ is merely the Y-intercept or the height of the line above origin. That is, when X=0, Y= a.
- The other constant ‘b’ represents the slope of the trend line.
- When b is positive, the slope is upwards, and when b is negative, the slope is downward.
- This line is termed as the line of best fit because it is so fitted that the total distance of deviations of the given data from the line is minimum. The total of deviations is calculated by squaring the difference in trend value and actual value of variable. Thus, the term “Least Squares” is attached to this method.

FITTING OF LINEAR TREND

using least square method, the normal equation for obtaining the values of a and b are :

$$\sum Y = na + b\sum t$$

$$\sum t * Y_t = a\sum t + b\sum t^2$$

$(Y = a + bt) * t$

$Yt = at + bt^2$

Let X = t – A, such that $\sum X = 0$, where A denotes the year of origin.
The above equations can also be written as

$$\sum Y = na + b\sum X$$

$$\sum XY = a\sum X + b\sum X^2$$

$n =$ odd $\rightarrow$ Mid Y

$n =$ even $\rightarrow$ Avg of 2 Mid Y

Since $\sum X = 0$ i.e. deviation from actual mean is zero.

FITTING OF LINEAR TREND

$\sum X = 0,$
So

$$a = \frac{\sum Y}{n}$$

$$\sum XY = b \sum X^2$$ or

$$b = \frac{\sum XY}{\sum X^2}$$

FITTING OF LINEAR TREND

$\sum X = 0,$
So

$$\text{a} = \frac{\sum Y}{n}$$

$\sum XY = b \sum X^2$ or

$$b = \frac{\sum XY}{\sum X^2}$$

Fitting a Straight line trend

Given below is the time series data on production (in thousand units) of a certain firm. Fit a straight line trend to the above data. Also estimate the trend for the year 2011

years	2004	2005	2006	2007	2008	2009	2010
Production	42	49	62	75	92	122	158

$n = 7$

$x = t - A$

Fitting a Straight line trend

years	Production			X^2
t	Yt Or Y	X = t - 2007	XY	Xsqaure
2004	42	-3	-126	9
2005	49	-2	-98	4
2006	62	-1	-62	1
2007	75	0	0	0
2008	92	1	92	1
2009	122	2	244	4
2010	158	3	474	9
Total	600	$\Sigma X =$ 0	524	ΣX^2 28

n=7 ΣXY

$$a = \frac{\sum Y}{n} = 600/7 = 85.71$$

$$b = \frac{\sum XY}{\sum X^2} = 524/28 = 18.71$$

X = 2011 - 2007

$$Y_t = a + b * X$$
$$Y_t = \underline{85.71} + \underline{18.71} * \underline{X}$$

2011

Moving Averages

- Calculate moving averages to get an overall impression of the pattern of movement over time
- This smooths out the irregular component

Moving Average: averages of a designated number of consecutive time-series values

Moving Average

- A series of arithmetic means over time
- Result depends upon choice of *m* (the number of data values in each average)
- Examples:
  - For a 5 year moving average, $m = 2$
  - For a 7 year moving average, $m = 3$
  - Etc.
- Replace each x_t with

$$x_t^* = \frac{1}{2m+1} \sum_{j=-m}^{m} x_{t+j} \quad (t = m+1, m+2, \ldots, n-m)$$

Moving Averages

- Example: Five-year moving average
  - First average:

$$x_3^* = \frac{x_1 + x_2 + x_3 + x_4 + x_5}{5}$$

  - Second average:

$$x_4^* = \frac{x_2 + x_3 + x_4 + x_5 + x_6}{5}$$

  - etc.

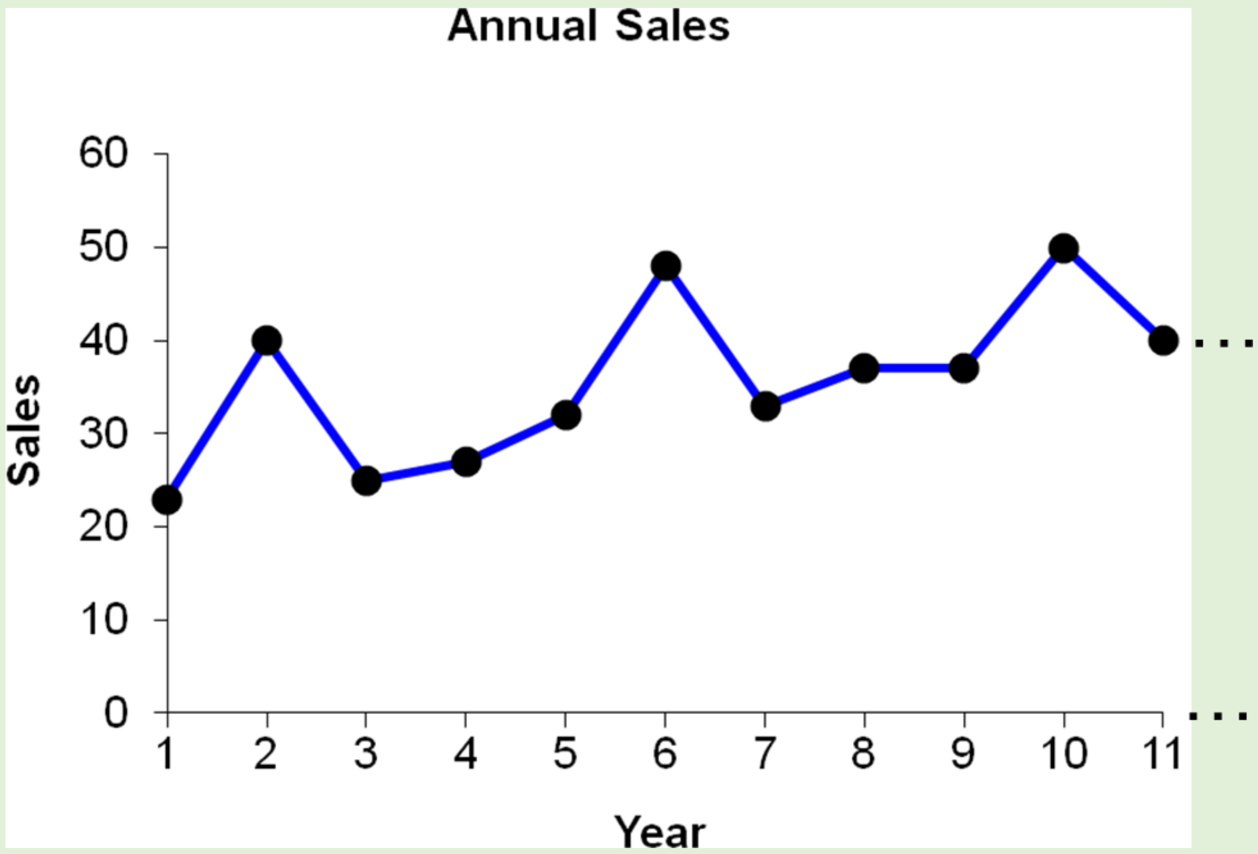
3 –yearly Moving Average Table

trend Value

YEAR	OUTPUT	3- YEARLY MOVING TOTAL	3- YEARLY MOVING AVERAGE
y1	x1		
y2	x2	T1 =(x1+ x2 + x3)	A1= T1/3
y3	x3	T2 =(x2 + x3 +x4)	A2= T2/3
y4	x4	T3=(x3 + x4 + x5)	A3 = T3/3
y5	x5	T4 = (x4 + x5 + x6)	A4 = T4/ 3
y6	x6	T5 = (x5 + x6 + x7)	A5 = T5 / 3
y7	x7		

Example: Annual Data

Year	Sales
1	23
2	40
3	25
4	27
5	32
6	48
7	33
8	37
9	37
10	50
11	40
etc…	etc…



Calculating Moving Averages

- Let $m = 2$

Year	Sales
1	23
2	40
3	25
4	27
5	32
6	48
7	33
8	37
9	37
10	50
11	40

etc…

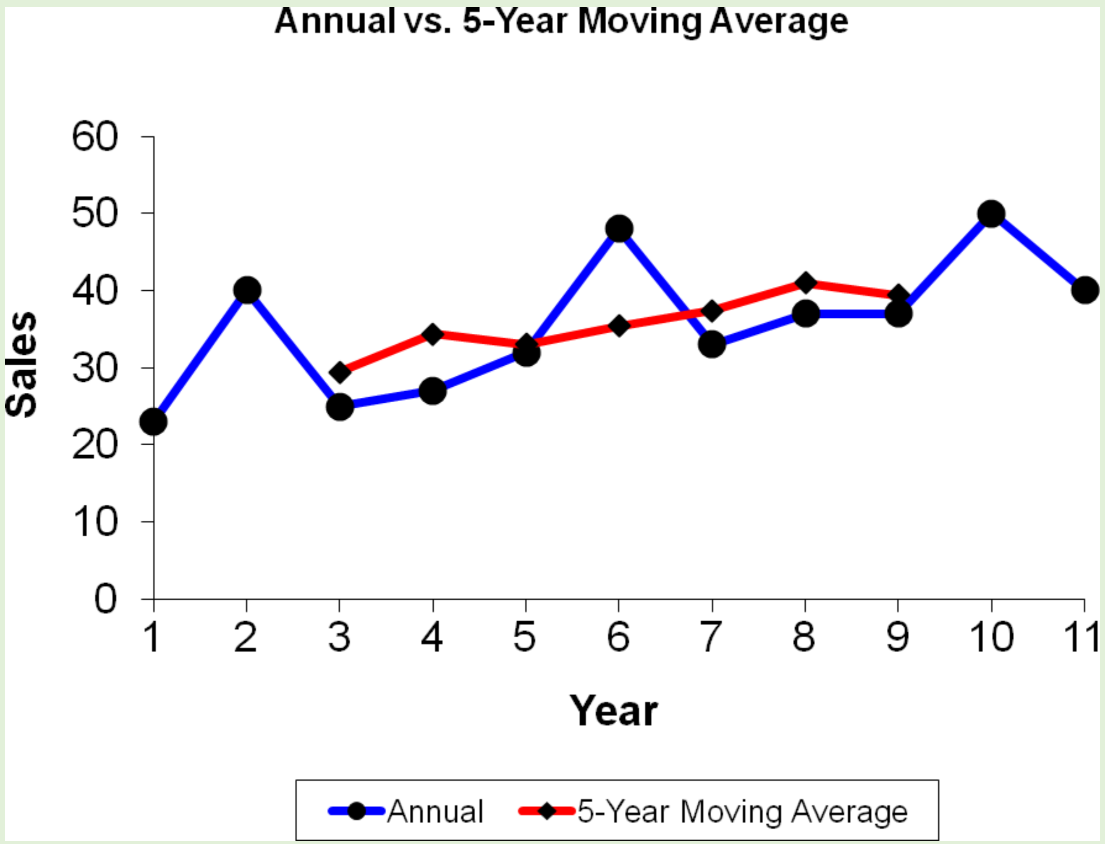
Average Year	5-Year Moving Average
3	29.4
4	34.4
5	33.0
6	35.4
7	37.4
8	41.0
9	39.4
…	…

$$29.4 = \frac{23+40+25+27+32}{5}$$

- Each moving average is for a consecutive block of $(2m+1)$ years

Annual vs. Moving Average

- The 5-year moving average smoothes the data and shows the underlying trend



Centered Moving Averages (1 of 2)

- Let the time series have period *s*, where *s* is even number
  - i.e., *s* = 4 for quarterly data and *s* = 12 for monthly data
- To obtain a centered *s*-point moving average series x_t^* :
  - Form the *s*-point moving averages

$$x_{t+0.5}^* = \frac{\sum_{j=-\left(\frac{s}{2}\right)+1}^{\frac{s}{2}} x_{t+j}}{s} \qquad \left(t = \frac{s}{2}, \frac{s}{2}+1, \frac{s}{2}+2, \ldots, n-\frac{s}{2}\right)$$

  - Form the centered *s*-point moving averages

$$x_t^* = \frac{x_{t-0.5}^* + x_{t+0.5}^*}{2} \qquad \left(t = \frac{s}{2}+1, \frac{s}{2}+2, \ldots, n-\frac{s}{2}\right)$$

Centered Moving Averages (2 of 2)

- Used when an even number of values is used in the moving average
- Average periods of 2.5 or 3.5 don't match the original periods, so we average two consecutive moving averages to get centered moving averages

Average Period	4-Quarter Moving Average
2.5	28.75
3.5	31.00
4.5	33.00
5.5	35.00
6.5	37.50
7.5	38.75
8.5	39.25
9.5	41.00

etc…

Centered Period	Centered Moving Average
3	29.88
4	32.00
5	34.00
6	36.25
7	38.13
8	39.00
9	40.13

4 –yearly Moving Average Table

YEAR	OUTPUT	4- YEARLY MOVING TOTAL	4 YEARLY MOVING AVERAGE	4 YEARLY MOVING AVERAGE CENTERED
y1	x1			
y2	x2			trend
		T1 =(x1+ x2 + x3 +x4)	A1= T1/4	
y3	x3			C1 = (A1 + A2)/2
		T2 =(x2 + x3 +x4 + x5)	A2= T2/4	
y4	x4			C2 = (A2 + A3) /2
		T3=(x3 + x4 + x5+ x6)	A3 = T3/4	
y5	x5			C3 = (A3 + A4)/2
		T4 = (x4 + x5 + x6 + x7)	A4 = T4/ 4	
y6	x6			
y7	x7			