Analysis of time series

Time-Series Data

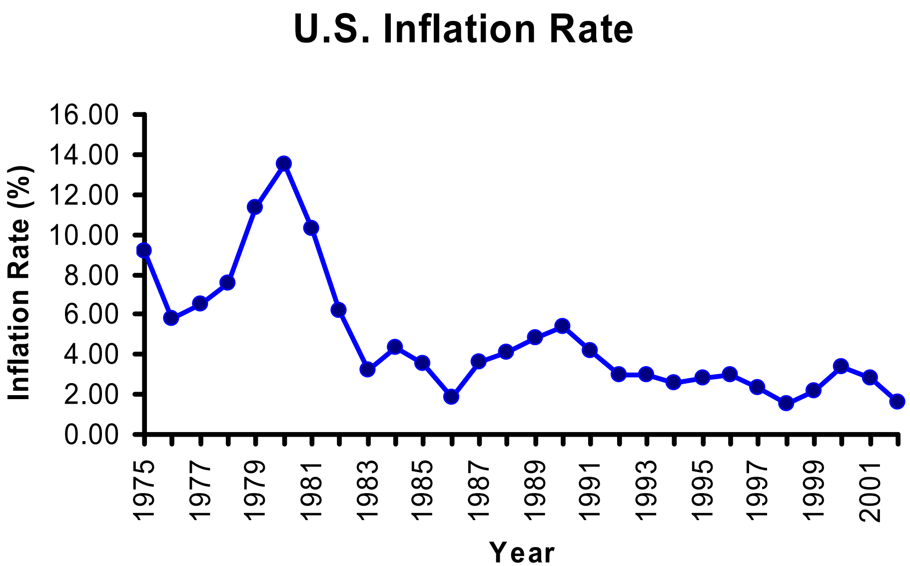
- Numerical data obtained at regular time intervals
- The time intervals can be annually, quarterly, daily, hourly, etc.
- Example:

Year:	2000	2001	2002	2003	2004
Sales:	75.3	74.2	78.5	79.7	80.2

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



A **time series** is a sequence of quantitative observations collected typically at uniform time intervals such as hourly, daily, weekly, monthly, quarterly, annually

It presents the evolution of a given phenomenon in time. For this reason such data are typically not independent.

Time series are of interest mainly for economic and social analysis as they present the trend of the phenomenon.

They allow to reveal and analyse the patterns of behavior over time based on the past to forecast the current estimates and forecasts.

The aims of time series analysis are:

- to describe and summarize time series data,
- fit models, and make forecasts

- A *time series variable* (Y) consists of data observed over n periods of time.

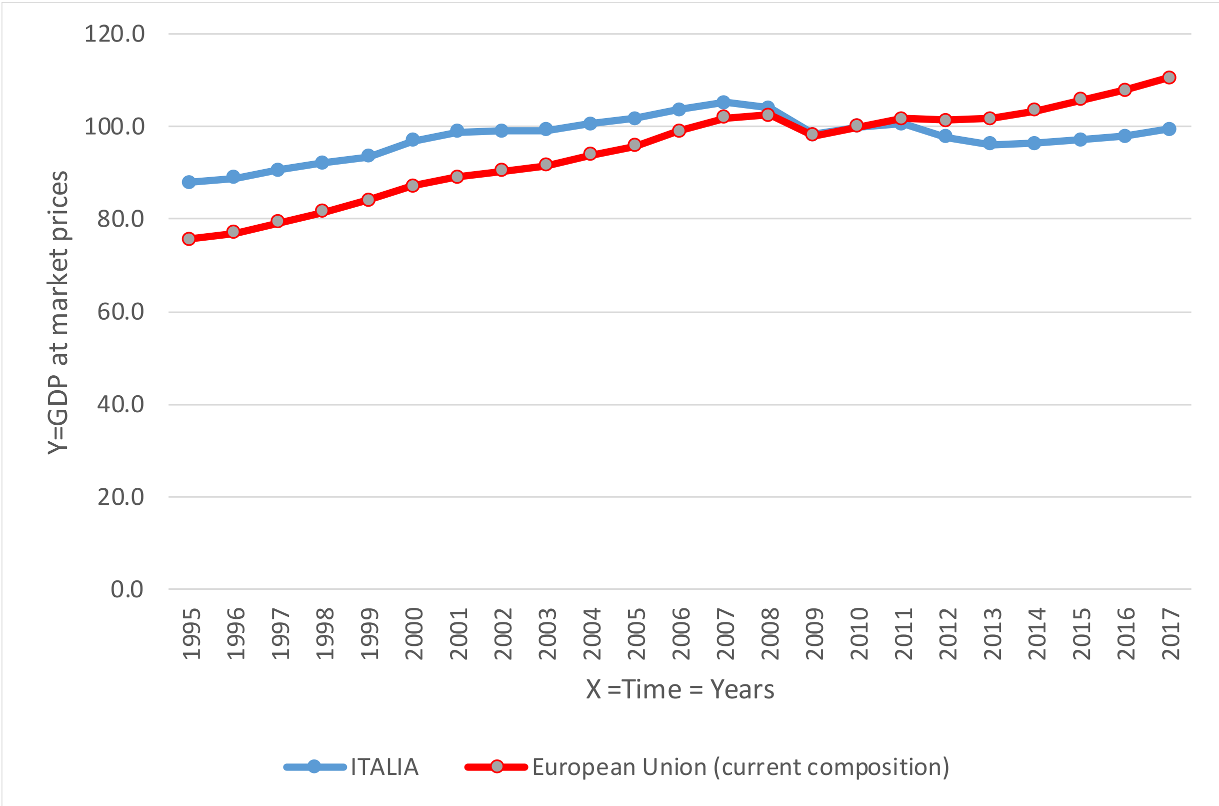
- Time series data can also be used to understand economic, population, health, crime, sports, and social problems…..
- Businesses, for example, use time series data
  - to monitor a process to determine if it is stable
  - to predict the future (forecasting).
- Time series data are usually plotted as a line or bar graph.
- Time is on the horizontal axis variable (X), while the variable in observation is on the vertical axis (Y)

  This reveals how a variable changes over time.

  Fluctuations are easier to see on a line graph.

- The following notation is used:

  y_t is the value of the time series in period t (t is an index denoting the time period t = 1, 2, …, n); n is the number of time periods y_1, y_2, …, y_n is the data set for analysis. Time is denoted with **x=t**


120.0
100.0
80.0
60.0
40.0
20.0
0.0
Y=GDP at market prices
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
X =Time = Years
ITALIA
European Union (current composition)

Time-Series Components

Time Series

- **Trend Component**: Overall, persistent, long-term movement
- **Seasonal Component**: Regular periodic fluctuations, usually within a 12-month period
- **Cyclical Component**: Repeating swings or movements over more than one year
- **Irregular Component**: Erratic or residual fluctuations

TREND COMPONENT

The trend is the long-term pattern of a time series. A trend can be positive or negative depending on whether the time series exhibits an increasing long term pattern or a decreasing long term pattern.

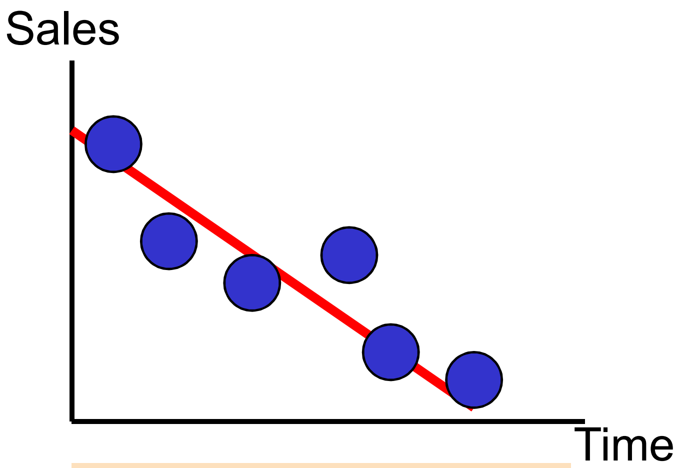
If a time series does not show an increasing or decreasing pattern then the series is stationary in the mean

- Trend: the long-term patterns or movements in the data.
- Overall or persistent, long-term upward or downward pattern of movement.
- Due to different causes as technology, market conditions, etc.
- Several years duration

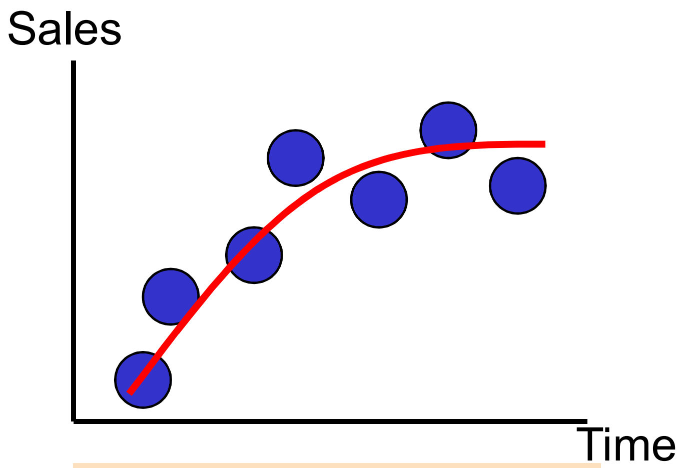
Trend Component

(continued)

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



Downward linear trend



Upward nonlinear trend

Seasonal variation

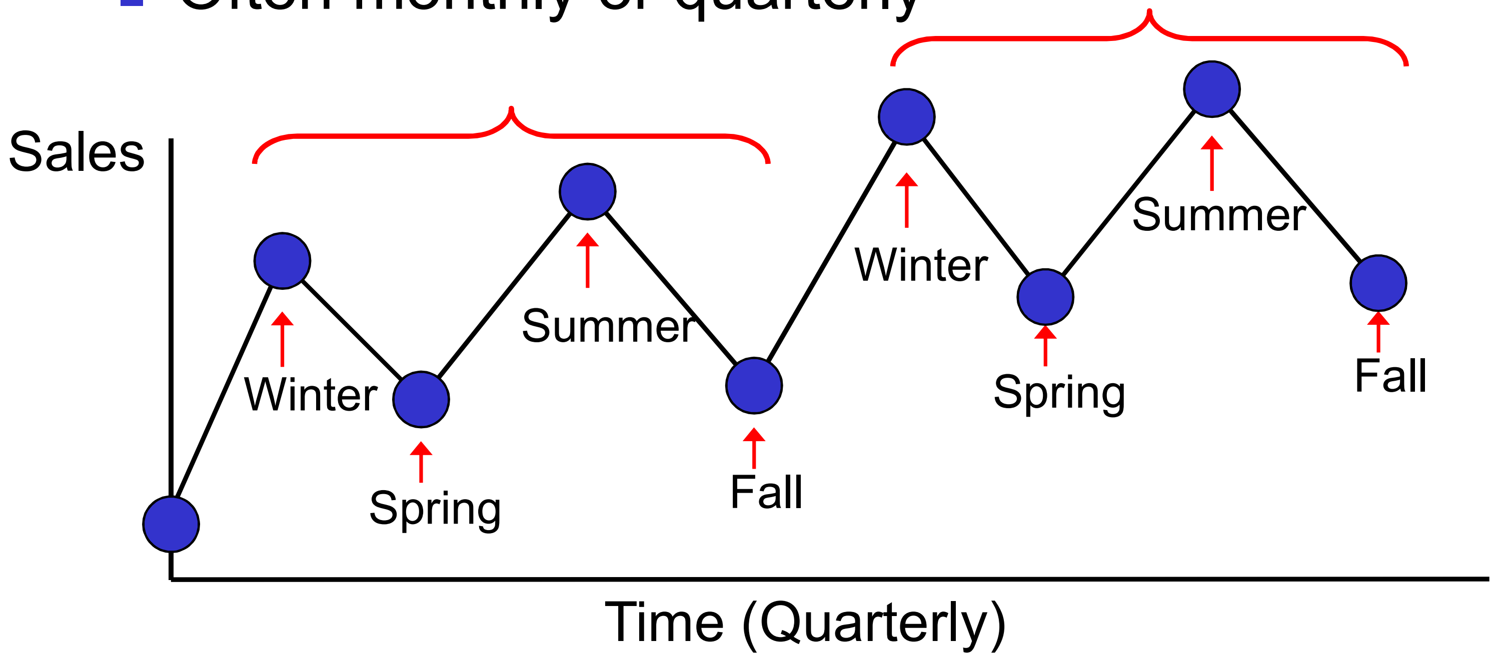
Seasonality occurs when the time series exhibits regular fluctuations during the same month (or months) every year, or during the same period every year.

→ Regular periodic fluctuations that occur within year.

For instance, retail sales peak during the month of December or ice cream sales increase in August and so on.

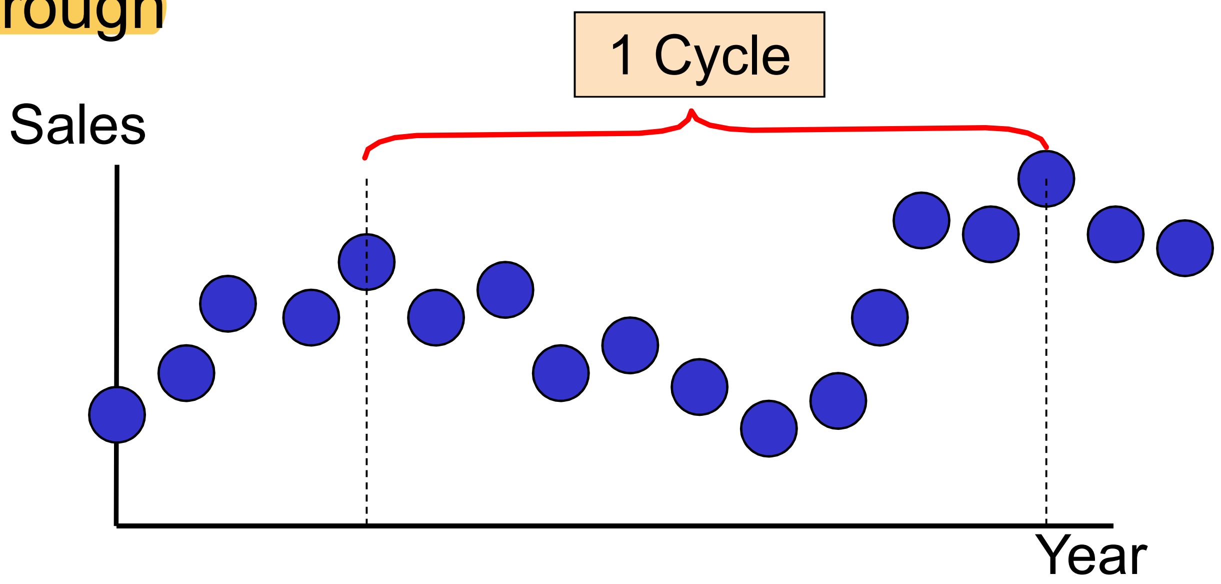
Seasonal Component

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



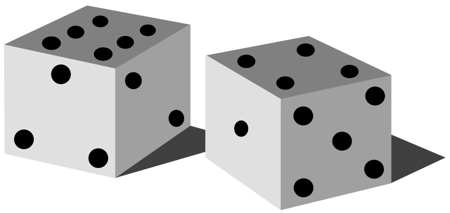
Cyclical Component

- 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



Additive versus Multiplicative Models

- Time series *decomposition* seeks to separate a time series Y into four components:
  - Trend (T)
  - Cycle (C)
  - Seasonal (S)
  - Irregular (I)

- These components are assumed to follow either an additive or a multiplicative model.

Model	***Components***	***Used For***
Additive	$Y = T + C + S + I$	Data of similar magnitude (short-run or trend-free data) with constant *absolute* growth or decline.
Multiplicative	$Y = T \times C \times S \times I$	Data of increasing or decreasing magnitude (long-run or trended data) with constant *percent* growth or decline.

- The multiplicative model becomes additive if logarithms are taken (for non-negative data):

$$\log(Y) = \log(T \times C \times S \times I) = \log(T) + \log(C) + \log(S) + \log(I)$$

Multiplicative Time-Series Model for Annual Data

- Used primarily for forecasting
- Observed value in time series is the product of components

$$\boxed{Y_i = T_i \times C_i \times I_i}$$

where T_i = Trend value at year i

C_i = Cyclical value at year i

I_i = Irregular (random) value at year i

Multiplicative Time-Series Model with a Seasonal Component

- Used primarily for forecasting
- Allows consideration of seasonal variation

$$\boxed{Y_i = T_i \times S_i \times C_i \times I_i}$$

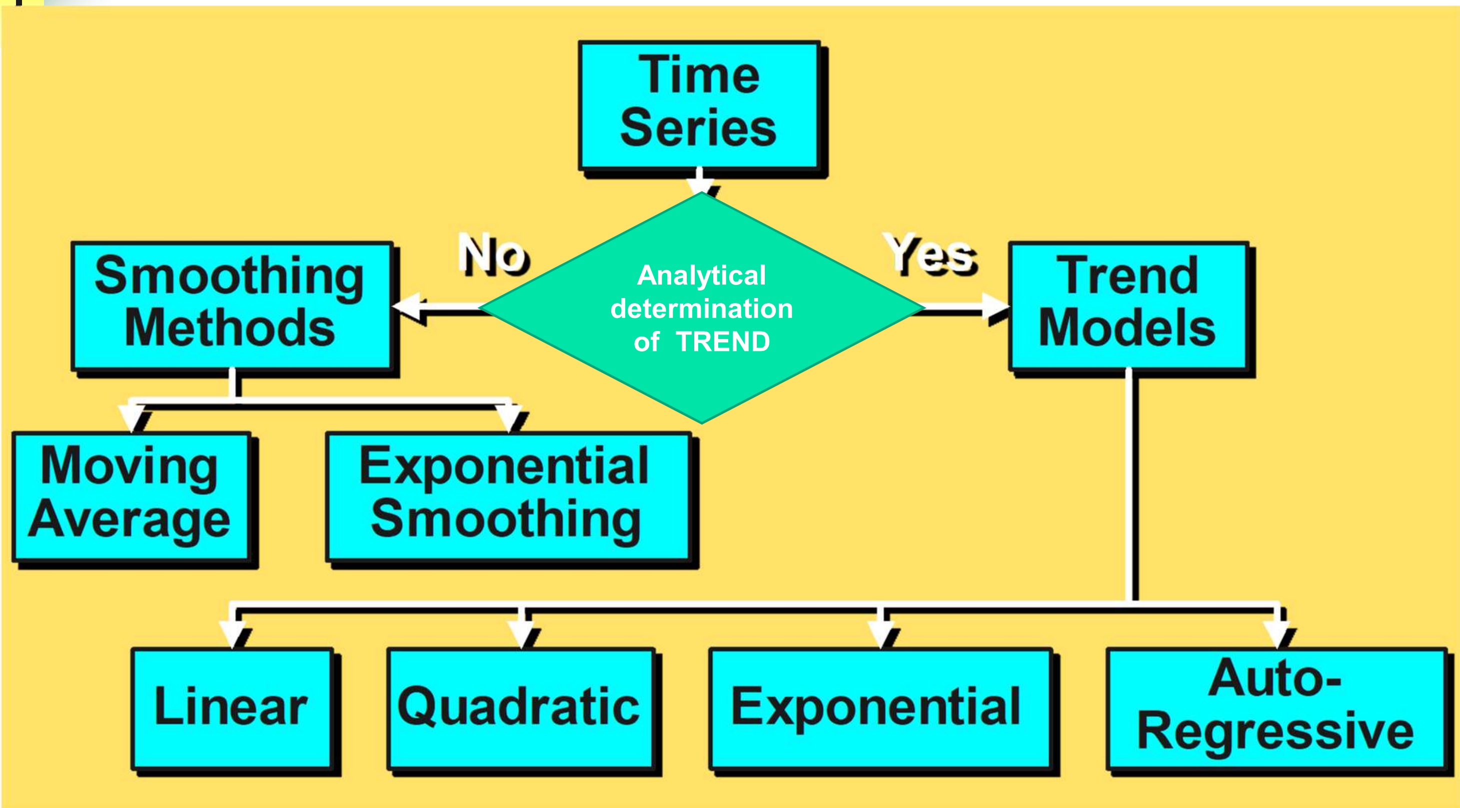
where T_i = Trend value at time i

S_i = Seasonal value at time i

C_i = Cyclical value at time i

I_i = Irregular (random) value at time i

Time Series Forecasting



TREND MODELS

LINEAR AND NON LINEAR TIME-SERIES FORECASTING MODEL

Trend-Based Forecasting

- Estimate a trend line using regression analysis

Year	Time Period (X)	Sales (Y)
1999	0	20
2000	1	40
2001	2	30
2002	3	50
2003	4	70
2004	5	65

- Use time (X) as the independent variable:

$$\hat{Y} = b_0 + b_1 X$$

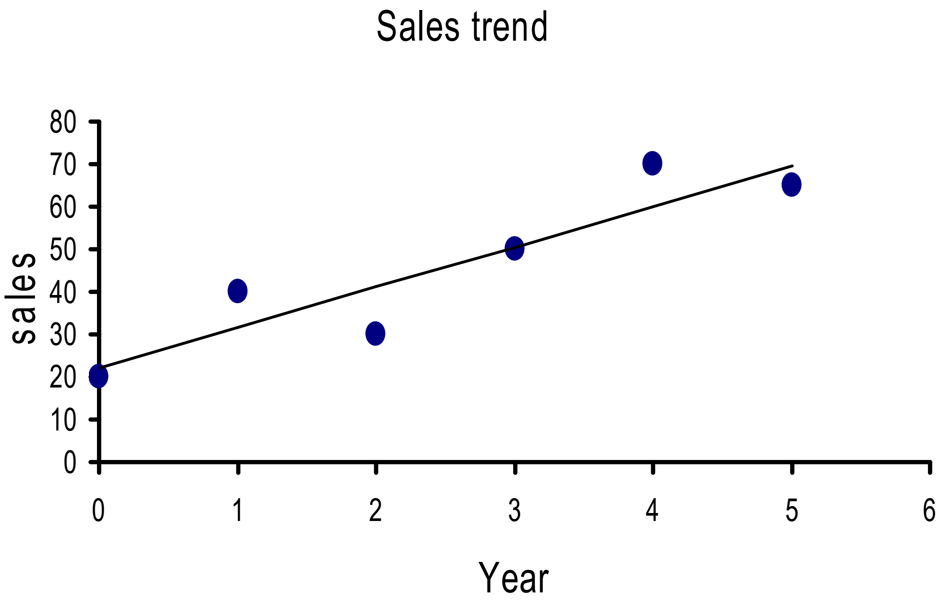
Trend-Based Forecasting

(continued)

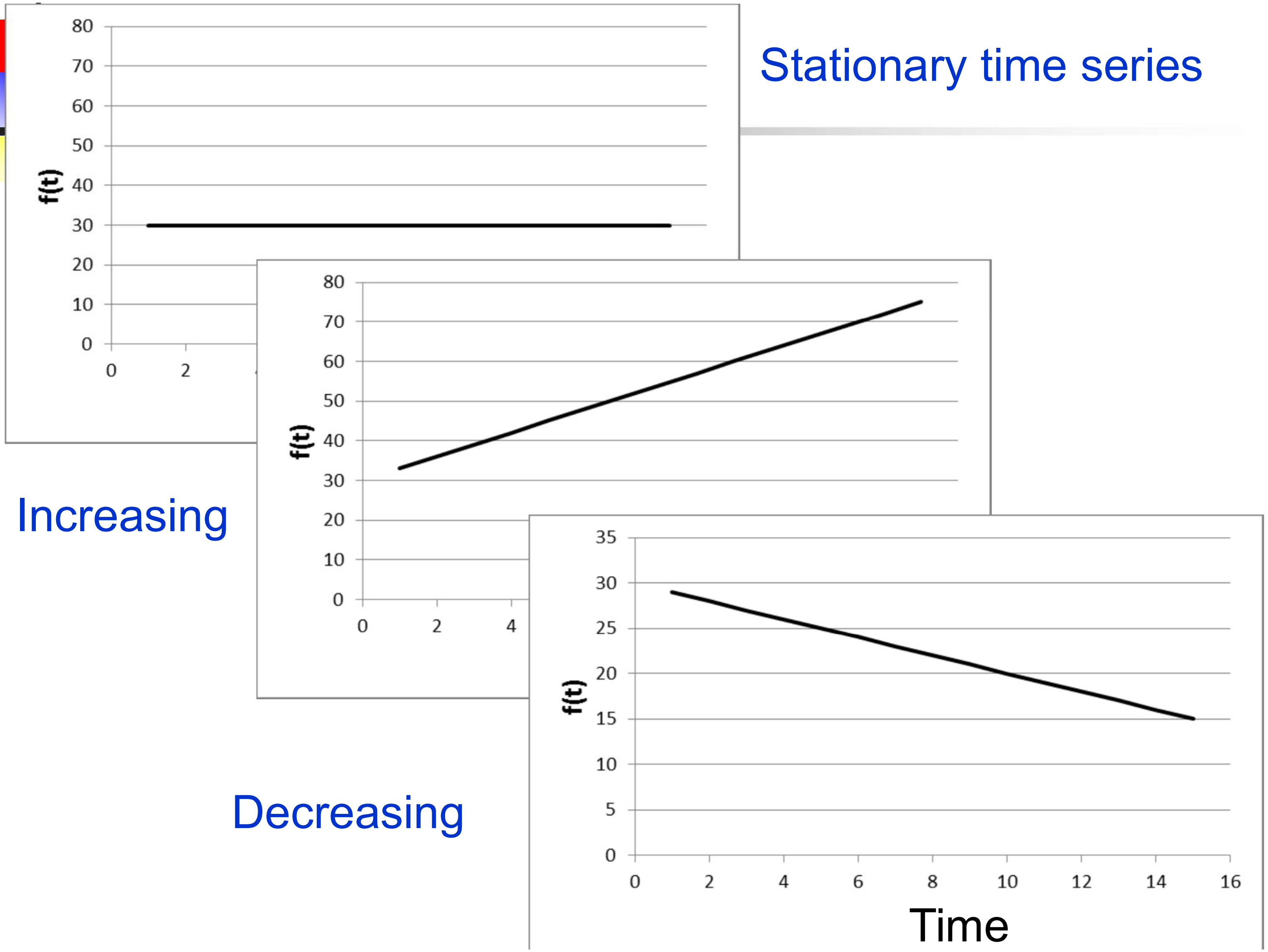
Year	Time Period (X)	Sales (Y)
1999	0	20
2000	1	40
2001	2	30
2002	3	50
2003	4	70
2004	5	65

- The linear trend forecasting equation is:

$$\hat{Y}_i = 21.905 + 9.5714\, X_i$$



Stationary time series



Increasing

Decreasing

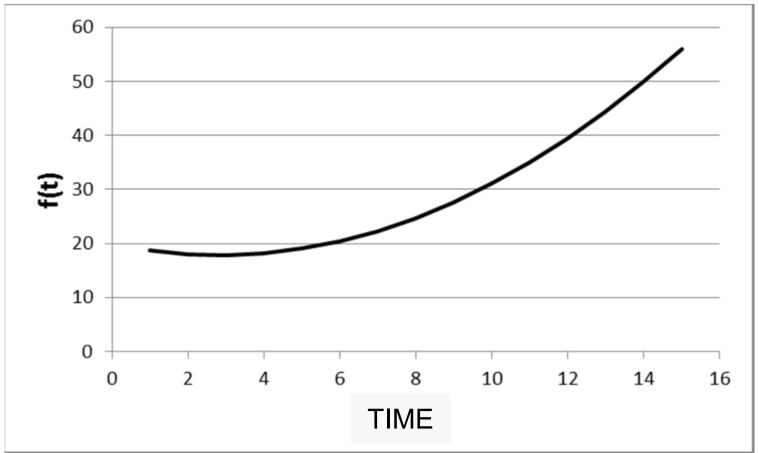
Nonlinear Trend Forecasting

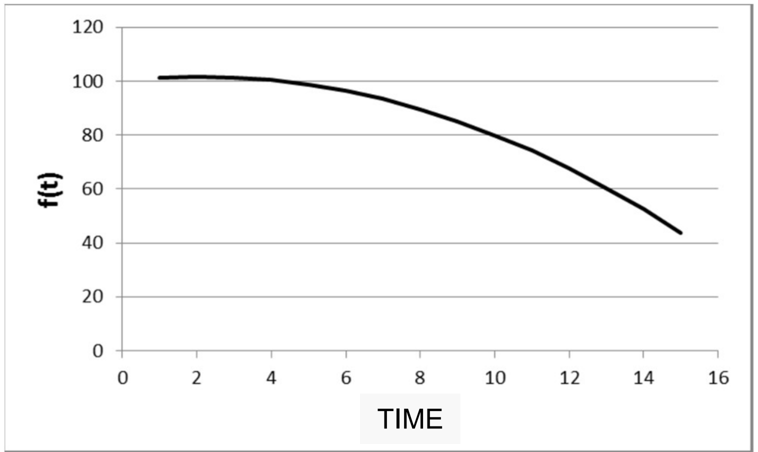
- A nonlinear regression model can be used when the time series exhibits a nonlinear trend
- Quadratic form is one type of a nonlinear model:

$$\boxed{Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i}$$

- Compare adj. r^2 and standard error to that of linear model to see if this is an improvement
- Can try other functional forms to get best fit

Quadratic Trend





Quadratic Trend Model

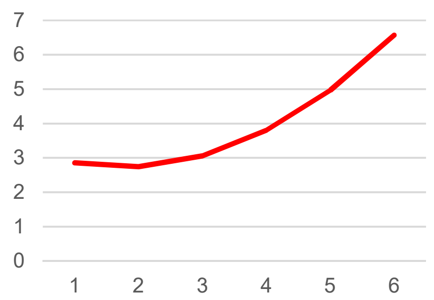
Year	Coded	Sales
94	0	2
95	1	5
96	2	2
97	3	2
98	4	7
99	5	6

$$\hat{Y}_i = b_0 + b_1 X_i + b_2 X_i^2$$

	Coefficients
Intercept	2.85714286
X Variable 1	-0.3285714
X Variable 2	0.21428571

Excel Output

$$\hat{Y}_i = 2.857 - 0.33 X_i + .214 X_i^2$$



Exponential Trend Model

- Another nonlinear trend model:

$$\boxed{Y_i = \beta_0 \beta_1^{X_i} \varepsilon_i}$$

- Transform to linear form:

$$\boxed{\log(Y_i) = \log(\beta_0) + X_i \log(\beta_1) + \log(\varepsilon_i)}$$

Exponential Trend Model

(continued)

- Exponential trend forecasting equation:

$$\log(\hat{Y}_i) = b_0 + b_1 X_i$$

where b_0 = estimate of $\log(\beta_0)$

b_1 = estimate of $\log(\beta_1)$

Interpretation:

$(\hat{\beta}_1 - 1) \times 100\%$ is the estimated annual compound growth rate (in %)

Exponential Time-Series Model

$$\hat{Y}_i = b_0 b_1^{X_i} \quad \text{or} \quad \log \hat{Y}_i = \log b_0 + X_1 \log b_1$$

Year	Coded	Sales
94	0	2
95	1	5
96	2	2
97	3	2
98	4	7
99	5	6

	Coefficients
Intercept	0.33583795
X Variable	0.08068544

Excel Output of Values in

antilog(.33583795) =	2.17
antilog(.08068544) =	1.2

$$\hat{Y}_i = (2.17)(1.2)^{X_i}$$

Model Selection Using Differences

- Use a linear trend model if the first differences are approximately constant

$$\boxed{(Y_2 - Y_1) = (Y_3 - Y_2) = \cdots = (Y_n - Y_{n-1})}$$

- Use a quadratic trend model if the second differences are approximately constant

$$\begin{aligned}[(Y_3 - Y_2) - (Y_2 - Y_1)] &= [(Y_4 - Y_3) - (Y_3 - Y_2)] \\ &= \cdots = [(Y_n - Y_{n-1}) - (Y_{n-1} - Y_{n-2})]\end{aligned}$$

Model Selection Using Differences

(continued)

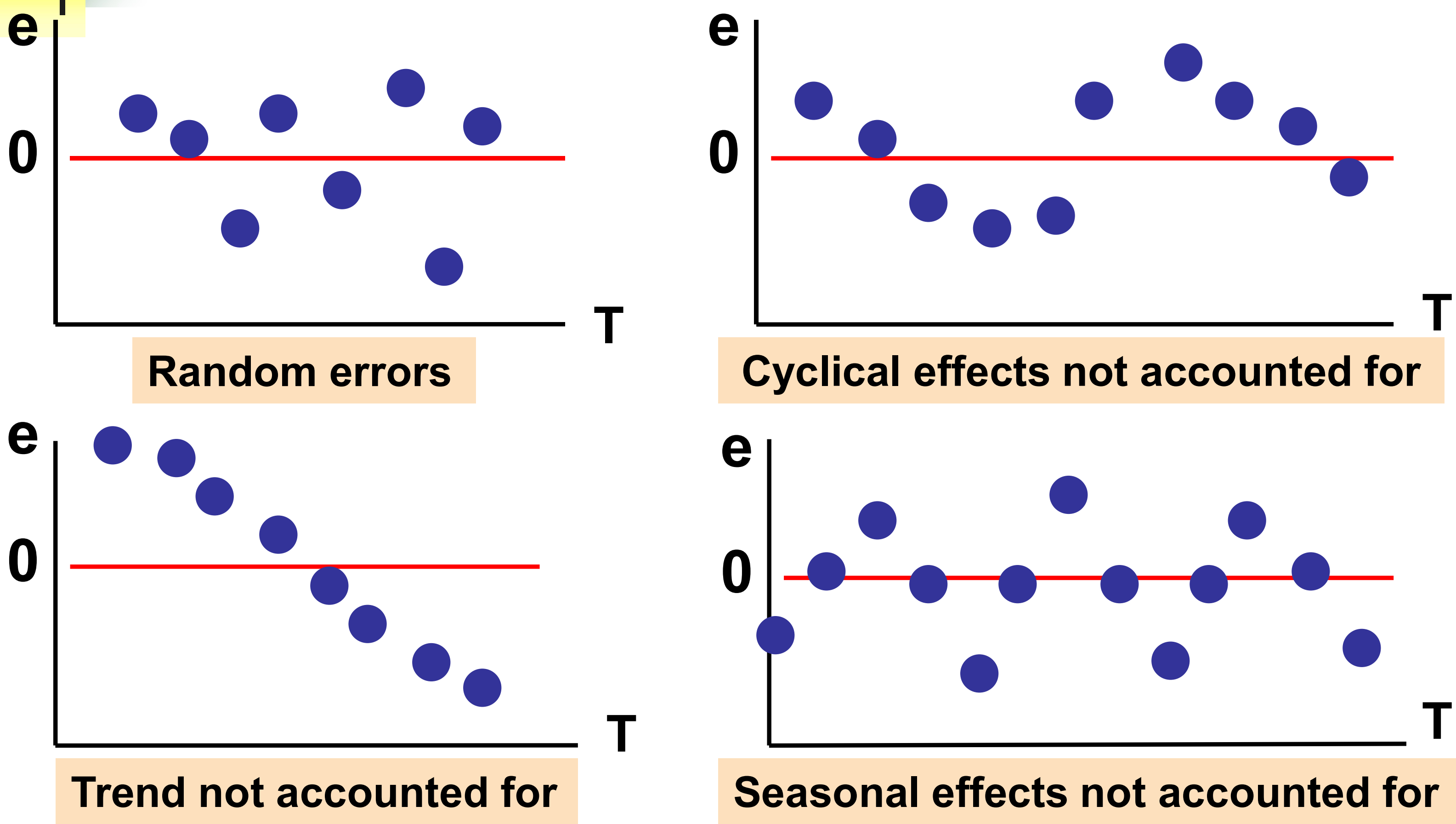
- Use an exponential trend model if the percentage differences are approximately constant

$$\frac{(Y_2 - Y_1)}{Y_1} \times 100\% = \frac{(Y_3 - Y_2)}{Y_2} \times 100\% = \cdots = \frac{(Y_n - Y_{n-1})}{Y_{n-1}} \times 100\%$$

Choosing A Forecasting Model

- Perform a residual analysis
  - Look for pattern or direction
- Measure magnitude of residual error using squared differences
- Measure residual error using absolute differences
- Use simplest model
  - Principle of parsimony

Residual Analysis



Measuring Errors

- Choose the model that gives the smallest measuring errors

- Sum of squared errors (SSE)

$$SSE = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

  - Sensitive to outliers

- Mean Absolute Deviation (MAD)

$$MAD = \frac{\sum_{i=1}^{n} \left| Y_i - \hat{Y}_i \right|}{n}$$

  - Not sensitive to extreme observations

Assessing Fit: "Fit" refers to how well the estimated trend model matches the observed historical past data.

TABLE 14.9 Five Measures of Fit

Statistic	Description	Pro	Con
(14.4) $R^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y}_t)^2}$	Coefficient of Determination	1. Unit-free measure. 2. Very common.	1. Often interpreted incorrectly (e.g., "percent of correct predictions").
(14.5) $MAPE = \frac{100}{n}\sum_{t=1}^{n}\frac{\lvert y_t - \hat{y}_t\rvert}{y_t}$	Mean Absolute Percent Error (*MAPE*)	1. Unit-free measure (%). 2. Intuitive meaning.	1. Requires $y_t > 0$. 2. Lacks nice math properties.
(14.6) $MAD = \frac{1}{n}\sum_{t=1}^{n}\lvert y_t - \hat{y}_t\rvert$	Mean Absolute Deviation (*MAD*)	1. Intuitive meaning. 2. Same units as y_t.	1. Not unit-free. 2. Lacks nice math properties.
(14.7) $MSD = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2$	Mean Squared Deviation (*MSD*)	1. Nice math properties. 2. Penalizes big errors more.	1. Nonintuitive meaning. 2. Rarely reported.
(14.8) $SE = \sqrt{\sum_{t=1}^{n}\frac{(y_t - \hat{y}_t)^2}{n-2}}$	Standard Error (*SE*)	1. Same units as y_t. 2. For confidence intervals.	1. Nonintuitive meaning.

Higher values
are better
values

Lower values
are better
values

Lower values
are better
values

Coefficient of determination: $$R^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2}$$

An R^2 close to 1 would indicate a good fit to the data.

However, more information is needed since the forecast is simply a projection of current trend assuming that nothing changes.

Principal of Parsimony

- Suppose two or more models provide a good fit for the data
- Select the simplest model
  - Simplest model types:
    - Least-squares linear
    - Least-squares quadratic
    - 1st order autoregressive
  - More complex types:
    - 2nd and 3rd order autoregressive
    - Least-squares exponential

Autoregressive Modeling

- Used for forecasting
- Takes advantage of autocorrelation
  - 1st order - correlation between consecutive values
  - 2nd order - correlation between values 2 periods apart
- p^{th} order Autoregressive model:

$$\boxed{Y_i = A_0 + A_1 Y_{i-1} + A_2 Y_{i-2} + \cdots + A_p Y_{i-p} + \delta_i}$$

Random Error

Autoregressive Model: Example

The Office Concept Corp. has acquired a number of office units (in thousands of square feet) over the last eight years. Develop the second order Autoregressive model.

Year	Units
97	4
98	3
99	2
00	3
01	2
02	2
03	4
04	6

Autoregressive Model: Example Solution

- **Develop the 2nd order table**
- **Use Excel to estimate a regression model**

Year	Y_i	Y_{i-1}	Y_{i-2}
97	4	--	--
98	3	4	--
99	2	3	4
00	3	2	3
01	2	3	2
02	2	2	3
03	4	2	2
04	6	4	2

Excel Output

	Coefficients
Intercept	3.5
X Variable 1	0.8125
X Variable 2	-0.9375

$$\hat{Y}_i = 3.5 + 0.8125Y_{i-1} - 0.9375Y_{i-2}$$

Autoregressive Model Example: Forecasting

Use the second-order equation to forecast number of units for 2005:

$$\hat{Y}_i = 3.5 + 0.8125 Y_{i-1} - 0.9375 Y_{i-2}$$

$$\begin{aligned}
\hat{Y}_{2005} &= 3.5 + 0.8125(Y_{2004}) - 0.9375(Y_{2003}) \\
&= 3.5 + 0.8125(6) - 0.9375(4) \\
&= 4.625
\end{aligned}$$

Autoregressive Modeling Steps

1. Choose p   (note that df = n – 2p – 1)
2. Form a series of "lagged predictor" variables

   Y_{i-1} , Y_{i-2} , … ,Y_{i-p}
3. Use Excel to run regression model using all p variables
4. Test significance of A_p
   - If null hypothesis rejected, this model is selected
   - If null hypothesis not rejected, decrease p by 1 and repeat

Regression withSeasonality

- Many time series have distinct seasonal pattern. (*For example room sales are usually highest around summer periods, or ice cream sales …)*
- In case of seasonality multiple regression models can be used to forecast a time series with seasonal components.
- Multiple regression means only one dependent variable Y and several independent variables X
- The use of dummy variables for seasonality is common. A dummy variable has only two values: 0 or 1
  - Dummy variables needed = total number of seasonality –1
  - For example: Quarterly Seasonal: 3 Dummies are needed, Monthly Seasonal: 11 Dummies needed, and so on
  - The load of each seasonal variable (dummy) is compared to the one which is *hidden* in intercept.

Forecasting With Seasonal Data

- Recall the classical time series model with seasonal variation:

$$\boxed{Y_i = T_i \times S_i \times C_i \times I_i}$$

- Suppose the seasonality is quarterly
  - Define three new dummy variables for quarters:

Q_1 = 1 if first quarter, 0 otherwise

Q_2 = 1 if second quarter, 0 otherwise

Q_3 = 1 if third quarter, 0 otherwise

(Quarter 4 is the default if $Q_1 = Q_2 = Q_3 = 0$)

Exponential Model with Quarterly Data

$$Y_i = \beta_0 \beta_1^{X_i} \beta_2^{Q_1} \beta_3^{Q_2} \beta_4^{Q_3} \varepsilon_i$$

$(\beta_1 - 1) \text{x} 100\%$ is the quarterly compound growth rate

β_i provides the multiplier for the i^{th} quarter relative to the 4th quarter (i = 2, 3, 4)

- Transform to linear form:

$$\log(Y_i) = \log(\beta_0) + X_i \log(\beta_1) + Q_1 \log(\beta_2) + Q_2 \log(\beta_3) + Q_3 \log(\beta_4) + \log(\varepsilon_i)$$

Estimating the Quarterly Model

- Exponential forecasting equation:

$$\boxed{\log(\hat{Y}_i) = b_0 + b_1 X_i + b_2 Q_1 + b_3 Q_2 + b_4 Q_3}$$

where b_0 = estimate of $\log(\beta_0)$, so $10^{b_0} = \hat{\beta}_0$

b_1 = estimate of $\log(\beta_1)$, so $10^{b_1} = \hat{\beta}_1$

etc…

Interpretation:

$(\hat{\beta}_1 - 1) \times 100\ \%$ = estimated quarterly compound growth rate (in %)

$\hat{\beta}_2$ = estimated multiplier for first quarter relative to fourth quarter

$\hat{\beta}_3$ = estimated multiplier for second quarter rel. to fourth quarter

$\hat{\beta}_4$ = estimated multiplier for third quarter relative to fourth quarter

Quarterly Model Example

- Suppose the forecasting equation is:

$$\log(\hat{Y}_i) = 3.43 + .017X_i - .082Q_1 - .073Q_2 + .022Q_3$$

b_0 = 3.43, so $\quad 10^{b_0} = \hat{\beta}_0 = 2691.53$

b_1 = .017, so $\quad 10^{b_1} = \hat{\beta}_1 = 1.040$

b_2 = -.082, so $\quad 10^{b_2} = \hat{\beta}_2 = 0.827$

b_3 = -.073, so $\quad 10^{b_3} = \hat{\beta}_3 = 0.845$

b_4 = .022, so $\quad 10^{b_4} = \hat{\beta}_4 = 1.052$

Quarterly Model Example

(continued)

Value:	Interpretation:
$\hat{\beta}_0 = 2691.53$	Unadjusted trend value for first quarter of first year
$\hat{\beta}_1 = 1.040$	4.0% = estimated quarterly compound growth rate
$\hat{\beta}_2 = 0.827$	Ave. sales in Q_2 are 82.7% of average 4^{th} quarter sales, after adjusting for the 4% quarterly growth rate
$\hat{\beta}_3 = 0.845$	Ave. sales in Q_3 are 84.5% of average 4^{th} quarter sales, after adjusting for the 4% quarterly growth rate
$\hat{\beta}_4 = 1.052$	Ave. sales in Q_4 are 105.2% of average 4^{th} quarter sales, after adjusting for the 4% quarterly growth rate

Smoothing the Annual Time Series

- Calculate moving averages to get an overall impression of the pattern of movement over time

Moving Average: averages of consecutive time series values for a chosen period of length L

Moving Averages

- Used for smoothing
- A series of arithmetic means over time
- Result dependent upon choice of L (length of period for computing means)
- Examples:
  - For a 5 year moving average, L = 5
  - For a 7 year moving average, L = 7
  - Etc.

Moving Averages

(continued)

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

$$\boxed{MA(5) = \frac{Y_1 + Y_2 + Y_3 + Y_4 + Y_5}{5}}$$

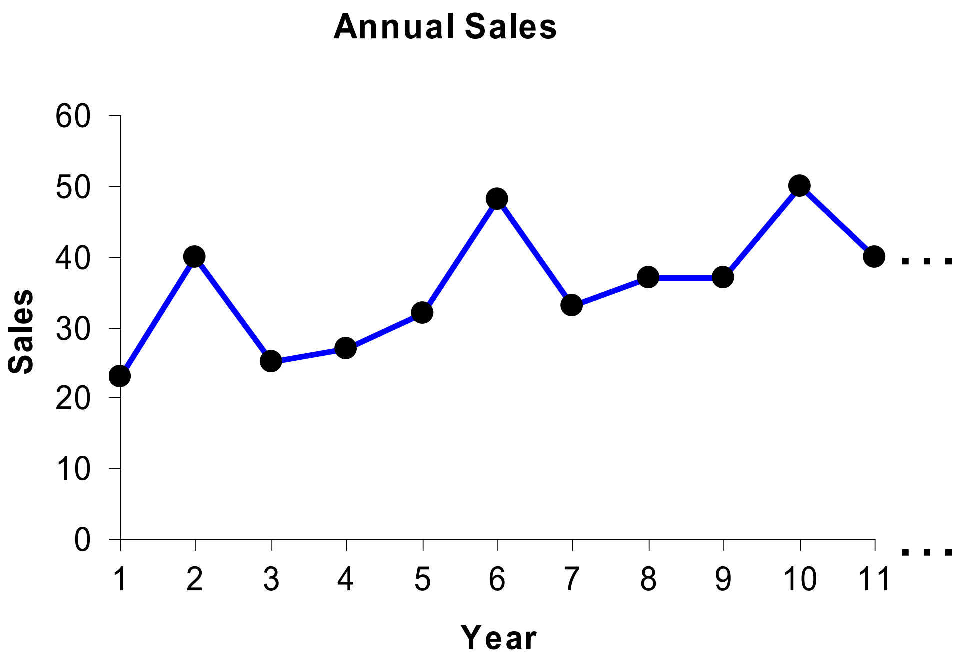
  - Second average:

$$\boxed{MA(5) = \frac{Y_2 + Y_3 + Y_4 + Y_5 + Y_6}{5}}$$

  - etc.

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

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
…	…

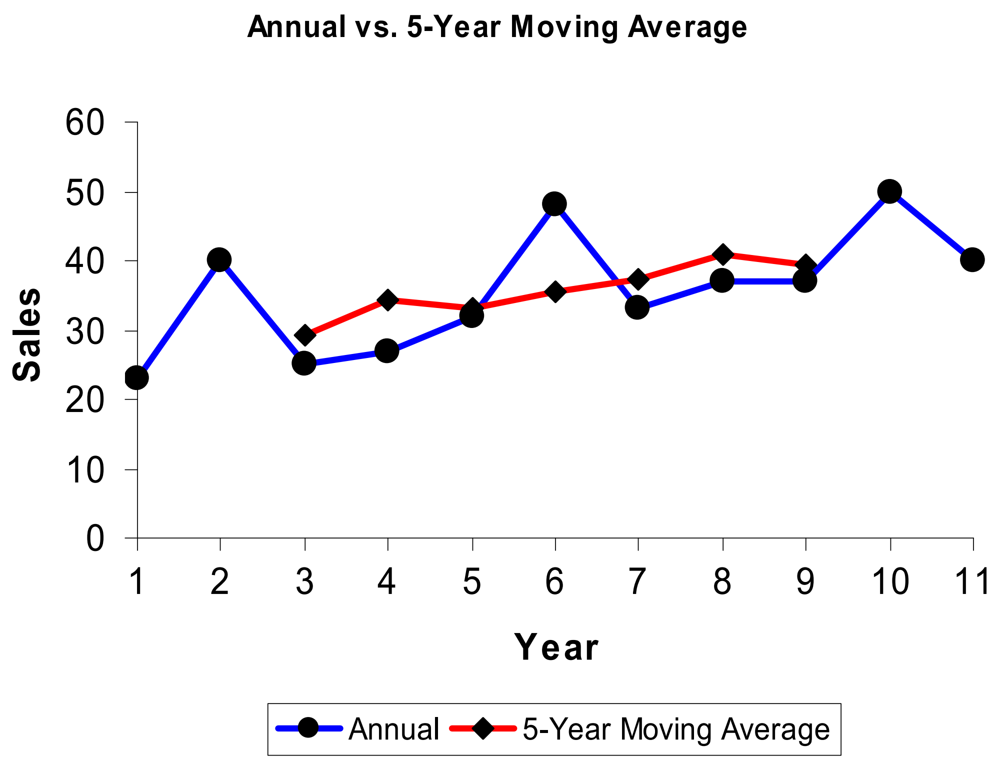
$$3 = \frac{1+2+3+4+5}{5}$$

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

- Each moving average is for a consecutive block of 5 years

Annual vs. Moving Average

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



We need even numbered MA for seasonal adjustments

eg: 4 – quarterly data

12 – monthly data

Even number of points.

Two stages:

1. Obtain MA, centered halfway between t and t-1.

2. To get a trend take the average of two successive estimates. Estimate centered halfway between t and t-1.

For example for k=4.

Step 1

$$MA_{1,1} = \frac{(y_1 + y_2 + y_3 + y_4)}{4}$$

$$MA_{1,2} = \frac{(y_2 + y_3 + y_4 + y_5)}{4}$$

Step 2

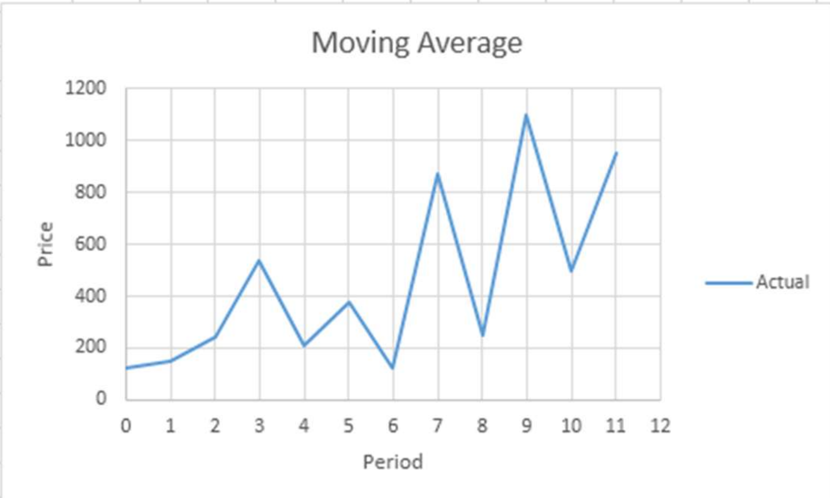
$$\overline{MA_1} = \frac{MA_{1,1} + MA_{1,2}}{2}$$

…. and so on …..

This example teaches you how to calculate the **moving average** of a time series in **Excel**. A moving average is used to smooth out irregularities (peaks and valleys) to easily recognize trends.

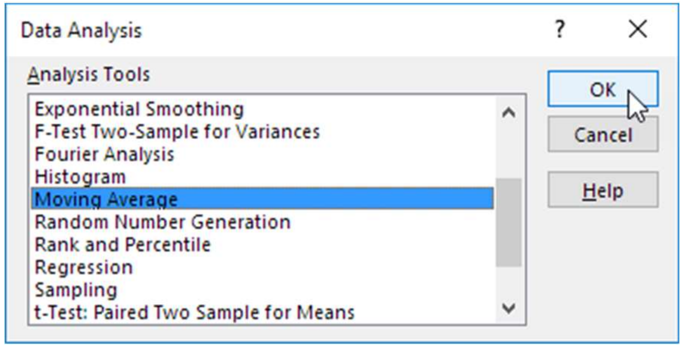
1. First, let's take a look at our time series.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Period	0	1	2	3	4	5	6	7	8	9	10	11	
2	Actual	120	150	240	540	210	380	120	870	250	1100	500	950	

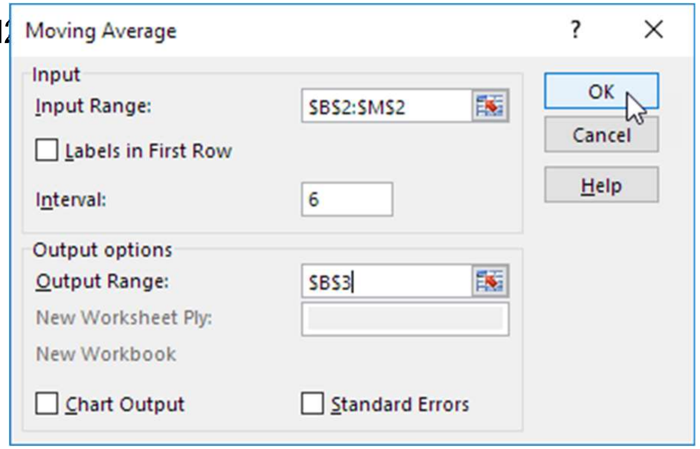


2. On the Data tab, in the Analysis group, click Data Analysis.

3. Select Moving Average and click OK.



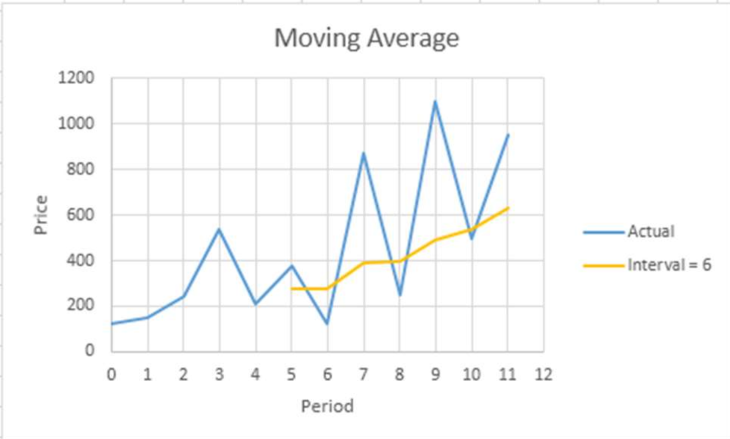
4. Click in the Input Range box and select the range B2:M
5. Click in the Interval box and type 6.
6. Click in the Output Range box and select cell B3.
7. Click OK



8. Plot a graph of these values.

COUNTIF =AVERAGE(B2:G2)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Period	0	1	2	3	4	5	6	7	8	9	10	11	
2	Actual	120	150	240	540	210	380	120	870	250	1100	500	950	
3	Interval = 6	#N/A	#N/A	#N/A	#N/A	#N/A	=AVER	273	393	395	488	537	632	



Explanation: because we set the interval to 6, the moving average is the average of the previous 5 data points and the current data point.

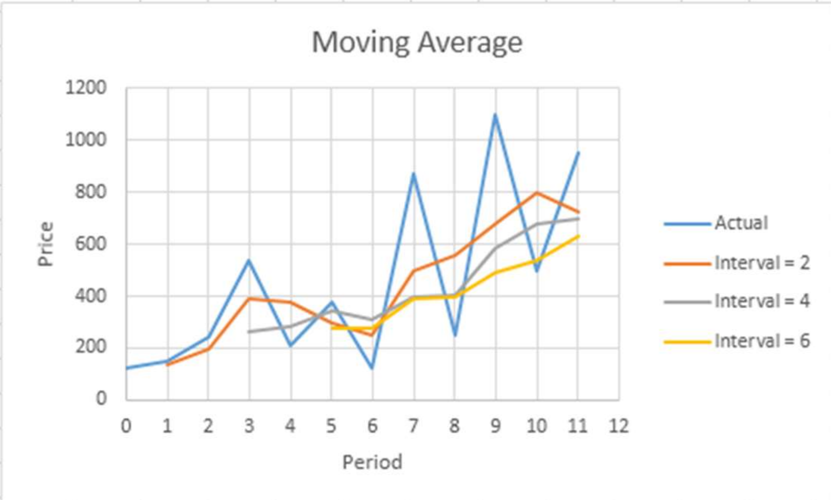
As a result, peaks and valleys are smoothed out.

The graph shows an increasing trend. Excel cannot calculate the moving average for the first 5 data points because there are not enough previous data points.

9. Repeat steps 2 to 8 for interval = 2 and interval = 4.

COUNTIF | =AVERAGE(B2:C2)

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Period	0	1	2	3	4	5	6	7	8	9	10	11
2	Actual	120	150	240	540	210	380	120	870	250	1100	500	950
3	Interval = 2	#N/A	=AVER	195	390	375	295	250	495	560	675	800	725
4	Interval = 4	#N/A	#N/A	#N/A	263	285	343	313	395	405	585	680	700
5	Interval = 6	#N/A	#N/A	#N/A	#N/A	#N/A	273	273	393	395	488	537	632



Conclusion: The larger the interval, the more the peaks and valleys are smoothed out. The smaller the interval, the closer the moving averages are to the actual data points.

Exponential Smoothing

- A weighted moving average
  - Weights decline exponentially
  - Most recent observation weighted most

- Used for smoothing and short term forecasting (often one period into the future)

Exponential Smoothing

(continued)

- The weight (smoothing coefficient) is W
  - Subjectively chosen
  - Range from 0 to 1
  - Smaller W gives more smoothing, larger W gives less smoothing
- The weight is:
  - Close to 0 for smoothing out unwanted cyclical and irregular components
  - Close to 1 for forecasting

Exponential Smoothing Model

- Exponential smoothing model

$$\boxed{E_1 = Y_1}$$

$$\boxed{E_i = WY_i + (1 - W)E_{i-1}}$$

For i = 2, 3, 4, …

where:

E_i = exponentially smoothed value for period i

E_{i-1} = exponentially smoothed value already computed for period i - 1

Y_i = observed value in period i

W = weight (smoothing coefficient), 0 < W < 1

Exponential Smoothing Example

- Suppose we use weight W = 0.2

Time Period (i)	Sales (Y_i)	Forecast from prior period (E_{i-1})	Exponentially Smoothed Value for this period (E_i)
1	23	--	23
2	40	23	(.2)(40)+(.8)(23)=26.4
3	25	26.4	(.2)(25)+(.8)(26.4)=26.12
4	27	26.12	(.2)(27)+(.8)(26.12)=26.296
5	32	26.296	(.2)(32)+(.8)(26.296)=27.437
6	48	27.437	(.2)(48)+(.8)(27.437)=31.549
7	33	31.549	(.2)(48)+(.8)(31.549)=31.840
8	37	31.840	(.2)(33)+(.8)(31.840)=32.872
9	37	32.872	(.2)(37)+(.8)(32.872)=33.697
10	50	33.697	(.2)(50)+(.8)(33.697)=36.958
etc.	etc.	etc.	etc.

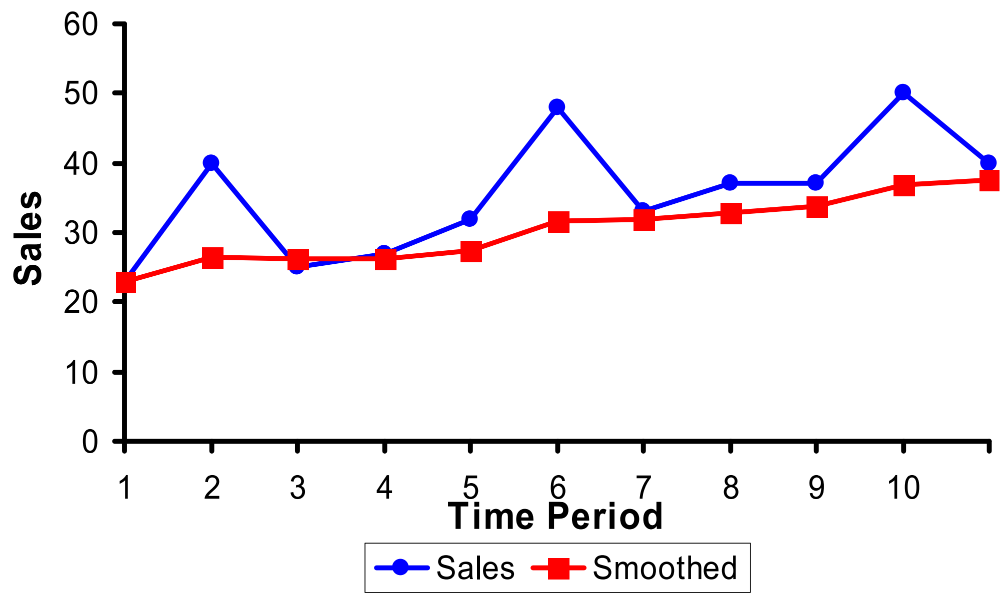
$E_1 = Y_1$ since no prior information exists

$$\boxed{E_i = WY_i + (1-W)E_{i-1}}$$

Sales vs. Smoothed Sales

- Fluctuations have been smoothed

- NOTE: the smoothed value in this case is generally a little low, since the trend is upward sloping and the weighting factor is only .2



Forecasting Time Period i + 1

- The smoothed value in the current period (i) is used as the forecast value for next period (i + 1) :

$$\boxed{\hat{Y}_{i+1} = E_i}$$

Exponential Smoothing in Excel

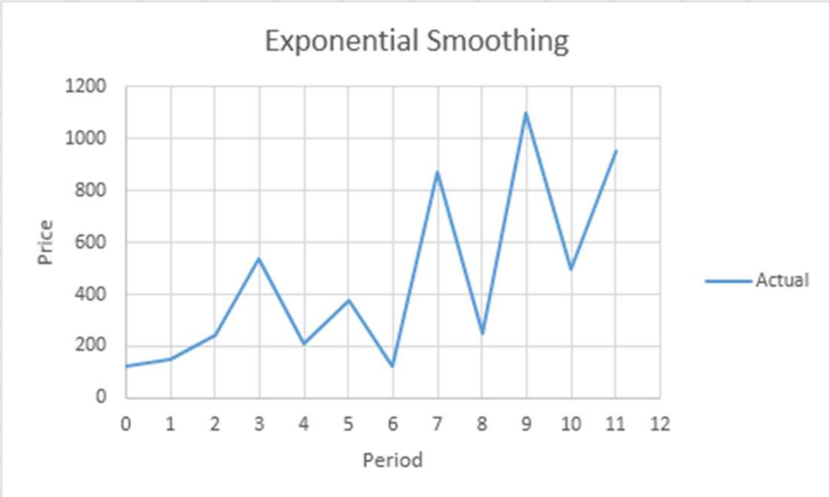
- Use tools / data analysis / exponential smoothing
  - The “damping factor” is (1 - W)

This example teaches you how to apply **exponential smoothing** to a time series in **Excel**. Exponential smoothing is used to smooth out irregularities (peaks and valleys) to easily recognize trends.

1. First, let's take a look at our time series.

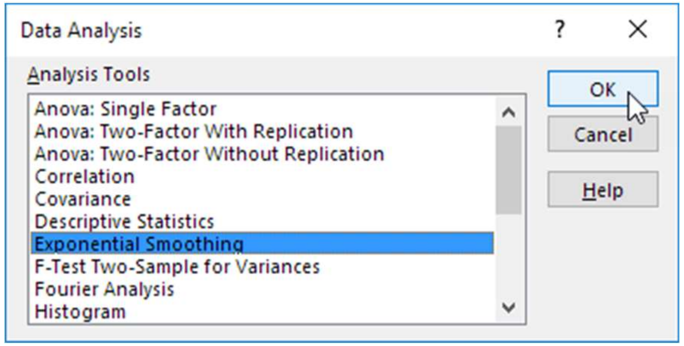
Q24

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Period	0	1	2	3	4	5	6	7	8	9	10	11	
2	Actual	120	150	240	540	210	380	120	870	250	1100	500	950	

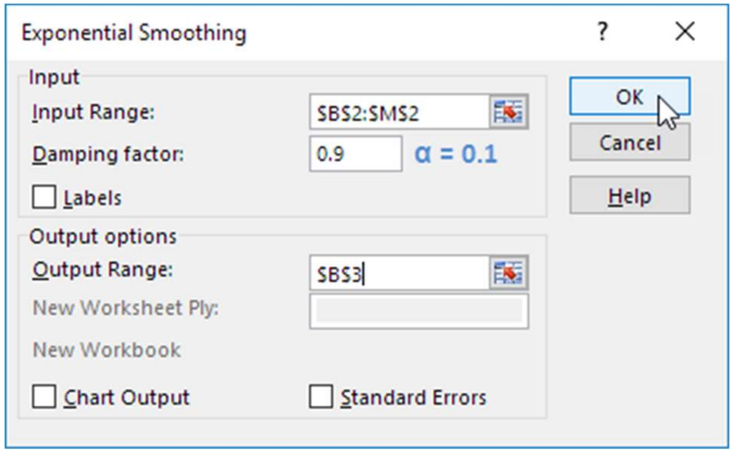


2. On the Data tab, in the Analysis group, click Data Analysis.

3. Select Exponential Smoothing and click OK.



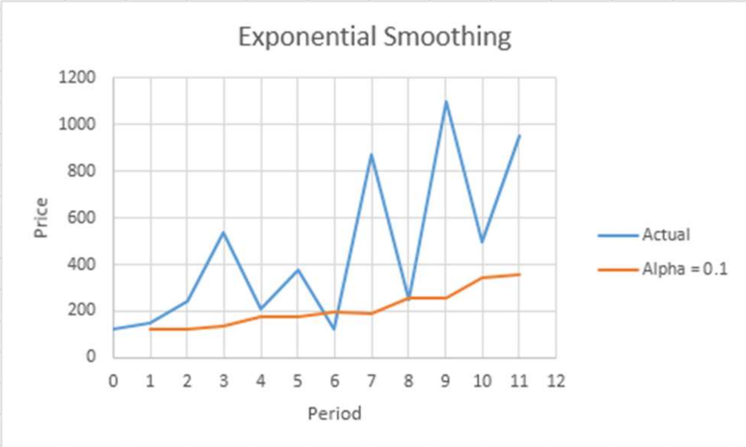
4. Click in the Input Range box and select the range B2:M2.
5. Click in the Damping factor box and type 0.9.
Literature often talks about the smoothing constant α (alpha).
The value (1- α) is called the damping factor.
6. Click in the Output Range box and select cell B3.
7. Click OK.



8. Plot a graph of these values.

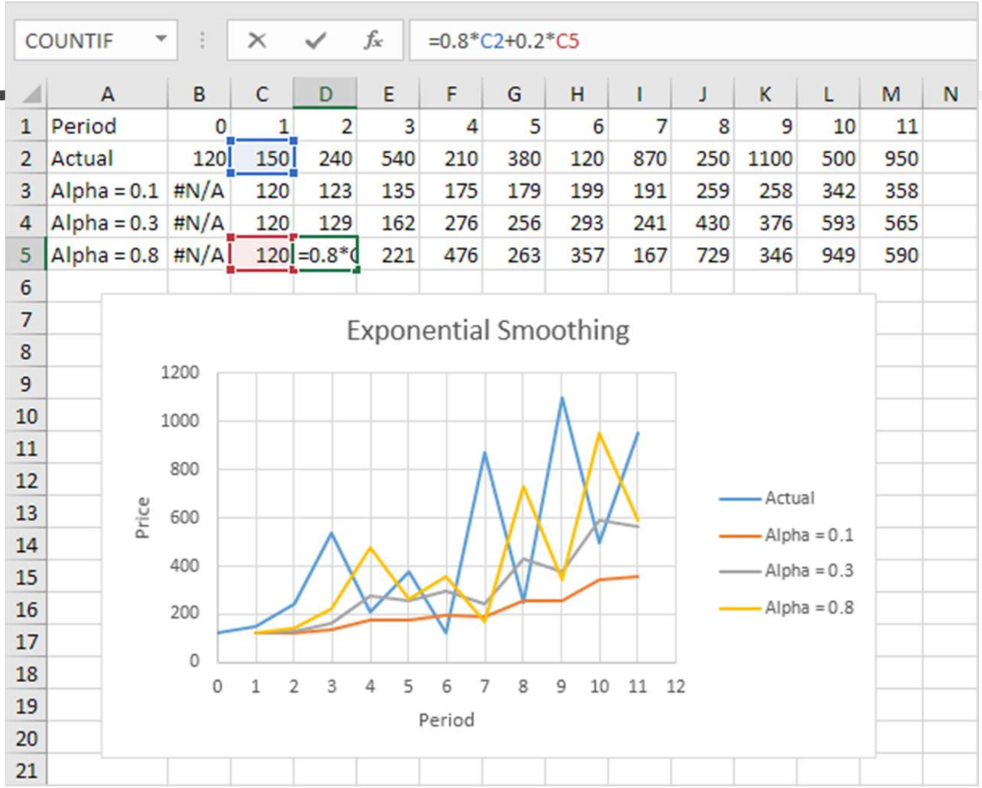
COUNTIF | =0.1*C2+0.9*C3

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Period	0	1	2	3	4	5	6	7	8	9	10	11	
2	Actual	120	150	240	540	210	380	120	870	250	1100	500	950	
3	Alpha = 0.1	#N/A	120	=0.1*(	135	175	179	199	191	259	258	342	358	



Explanation: because we set alpha to 0.1, the previous data point is given a relatively small weight while the previous smoothed value is given a large weight (i.e. 0.9). As a result, peaks and valleys are smoothed out. The graph shows an increasing trend. Excel cannot calculate the smoothed value for the first data point because there is no previous data point. The smoothed value for the second data point equals the previous data point.

9. Repeat steps 2 to 8 for alpha = 0.3 and alpha = 0.8.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Period	0	1	2	3	4	5	6	7	8	9	10	11	
2	Actual	120	150	240	540	210	380	120	870	250	1100	500	950	
3	Alpha = 0.1	#N/A	120	123	135	175	179	199	191	259	258	342	358	
4	Alpha = 0.3	#N/A	120	129	162	276	256	293	241	430	376	593	565	
5	Alpha = 0.8	#N/A	120	=0.8*C	221	476	263	357	167	729	346	949	590	

<u>Conclusion</u>: The smaller alpha (larger the damping factor), the more the peaks and valleys are smoothed out. The larger alpha (smaller the damping factor), the closer the smoothed values are to the actual data points.