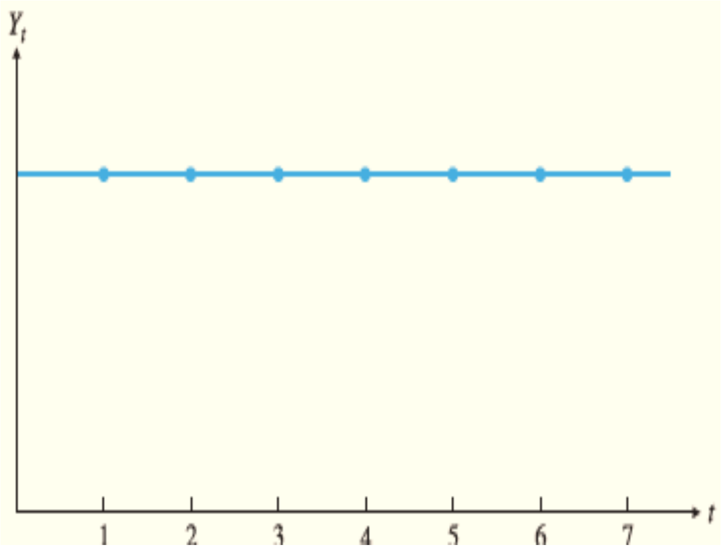
DATA ANALYTICS

Unit 3: Introduction to Time series data-Forecasting techniques and accuracy(MA,Exponential smoothing Holt’s and Holt Winter's model)

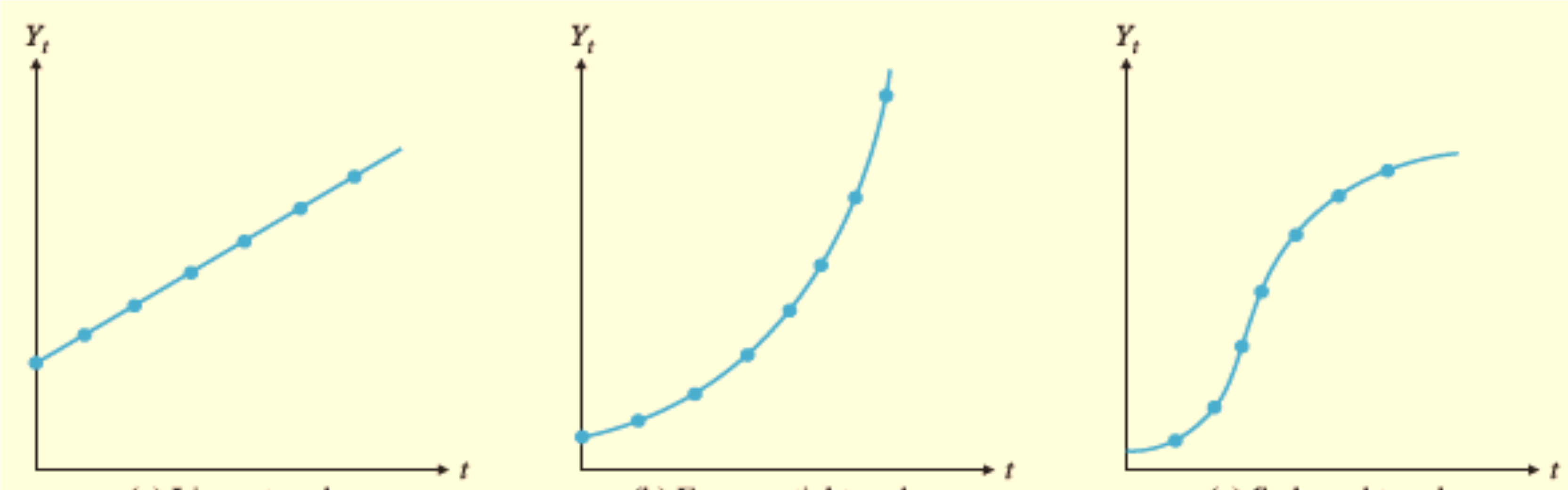
Jyothi R., Bharathi R
Department of Computer Science and Engineering

A revisit to Time Series data

The Base Series



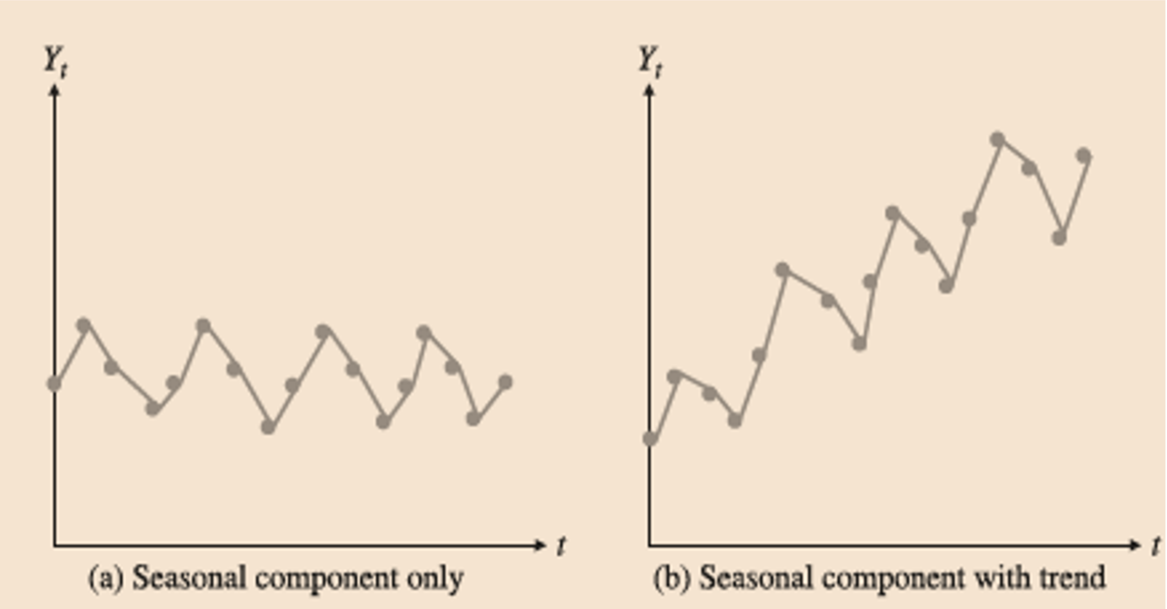
Series with Trends



(a) Linear trend (b) Exponential trend (c) S-shaped trend

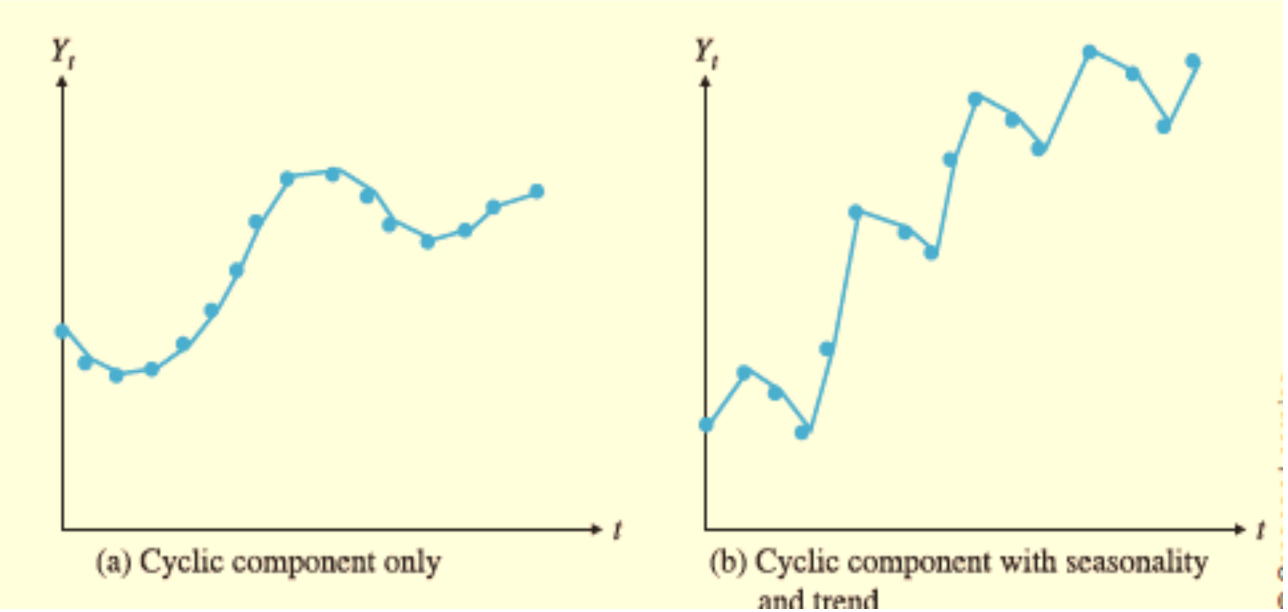
A revisit to Time Series data

Series with Seasonality



(a) Seasonal component only (b) Seasonal component with trend

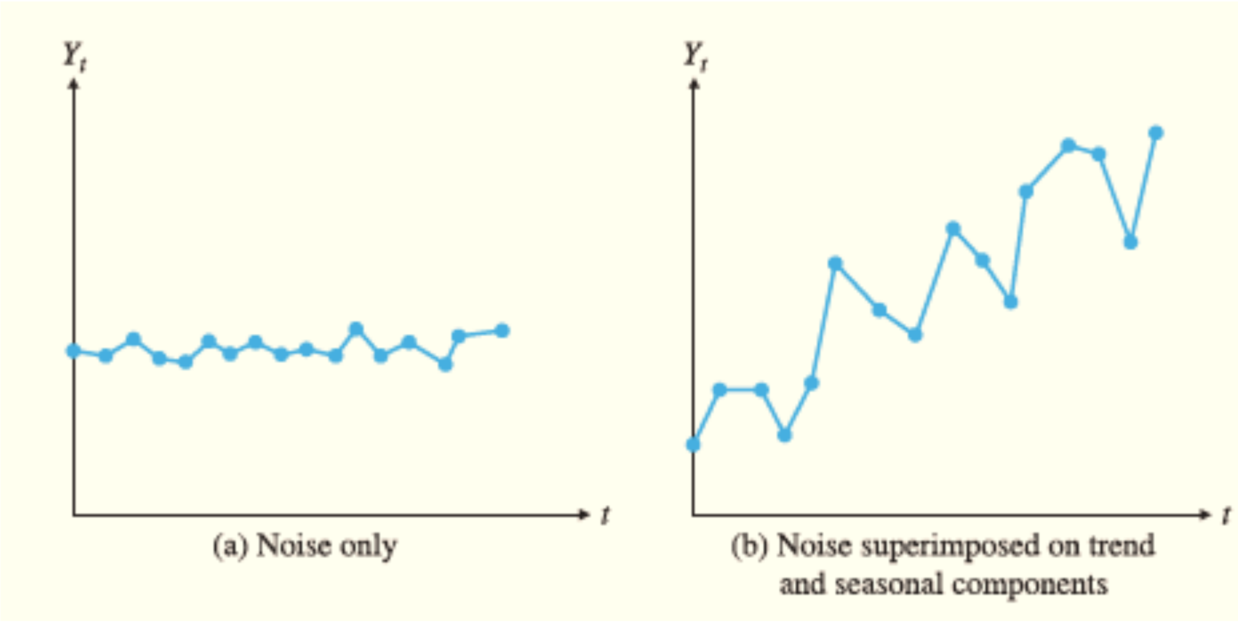
Series with Cyclic Component



(a) Cyclic component only (b) Cyclic component with seasonality and trend

A revisit to Time Series data

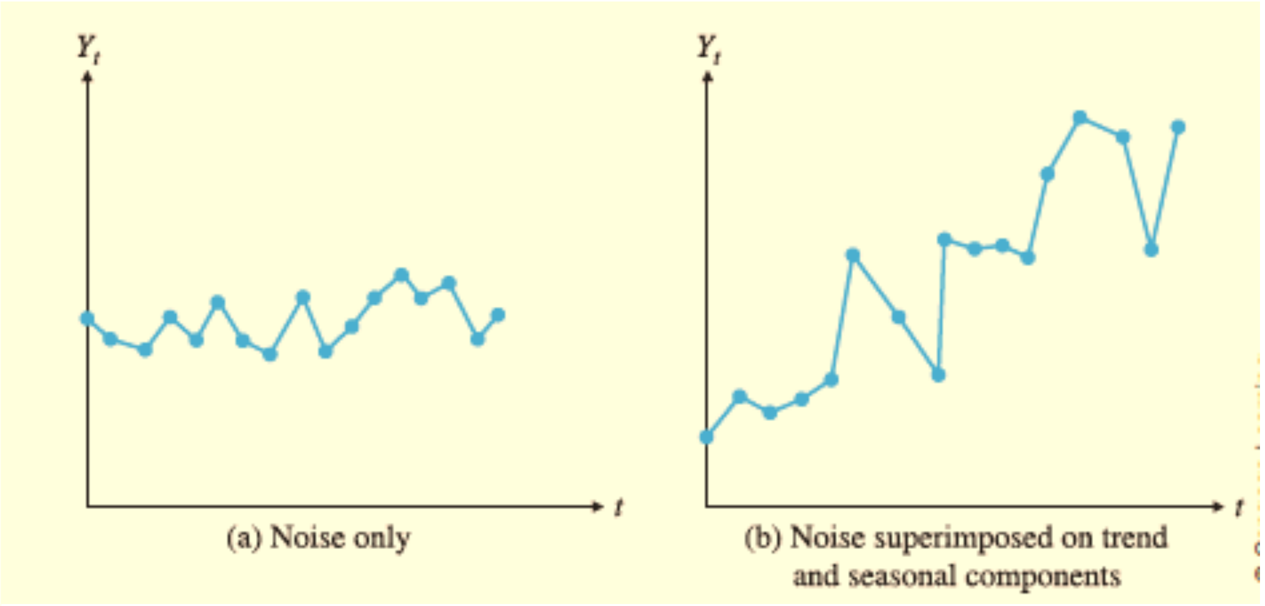
Series with Noise



(a) Noise only

(b) Noise superimposed on trend and seasonal components

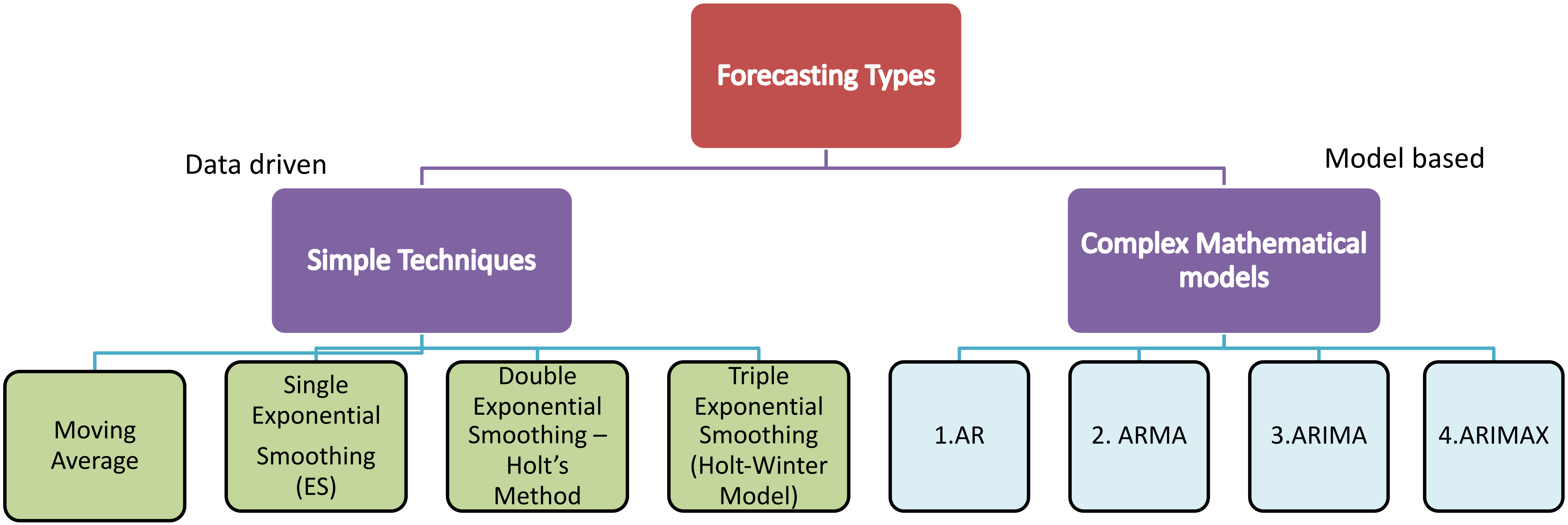
Series with More Noise



(a) Noise only

(b) Noise superimposed on trend and seasonal components

INTRODUCTION TO FORECASTING



INTRODUCTION TO FORECASTING – Simple Techniques

1. Moving Average
2. Single Exponential Smoothing (ES)
3. Double Exponential Smoothing – Holt's Method
4. Triple Exponential Smoothing (Holt-Winter Model)

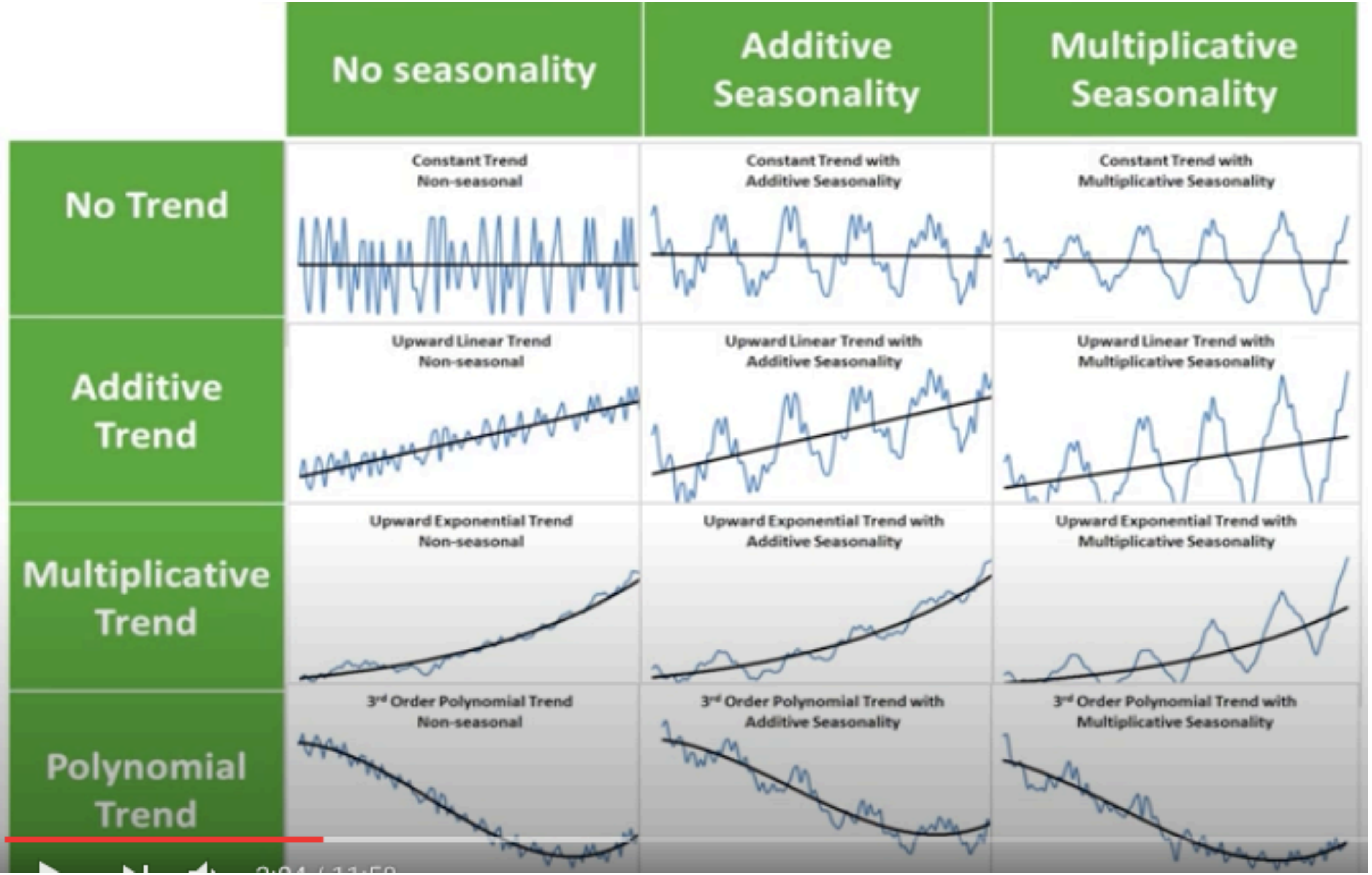
Introduction to Forecasting

- These models are applicable to time series data with seasonal, trend, or both seasonal and trend component and stationary data

- Forecasting methods discussed in this chapter can be classified as:
  - Averaging methods
  - Equally weighted observations
  - Exponential Smoothing methods
  - Unequal set of weights to past data, where the weights decay exponentially from the most recent to the most distant data points

- All methods in this group require that certain parameters to be defined
  - These parameters (with values between 0 and 1) will determine the unequal weights to be applied to past data

Types of Time Series Data



Moving Average

- If a time series is generated by a constant process subject to random error, then mean is a useful statistic and can be used as a forecast for the next period
- Averaging methods are suitable for stationary time series data where the series is in equilibrium around a constant value (the underlying mean) with a constant variance over time

Mean: Uses the average of all the historical data as the forecast

$$F_{t+1} = \frac{1}{t}\sum_{i=1}^{t} y_i$$

- When new data becomes available , the forecast for time t+2 is the new mean including the previously observed data plus this new observation

$$F_{t+2} = \frac{1}{t+1}\sum_{i=1}^{t+1} y_i$$

- This method is appropriate when there is no noticeable trend or seasonality

Averaging Methods

- The moving average for time period t is the mean of the “k” most recent observations
- The constant number k is specified at the outset
- The smaller the number k, the more weight is given to recent data points
- The greater the number k, the less weight is given to more recent data points

$$F_{t+1} = \hat{y}_{t+1} = \frac{(y_t + y_{t-1} + y_{t-2} + \square + y_{t-k+1})}{K}$$

$$F_{t+1} = \frac{1}{k} \sum_{i=t-k+1}^{t} y_i$$

- A large k is desirable when there are wide, infrequent fluctuations in the series

- A small k is most desirable when there are sudden shifts in the level of series

Moving Averages

- For quarterly data, a four-quarter moving average, MA(4), eliminates or averages out seasonal effects
- For monthly data, a 12-month moving average, MA(12), eliminate or averages out seasonal effect
- Equal weights are assigned to each observation used in the average
- Each new data point is included in the average as it becomes available, and the oldest data point is discarded
- The moving average model does not handle trend or seasonality very well although it can do better than the total mean

DATA ANALYTICS

Example: GASOLINE SALES TIME SERIES PLOT AND THREE-WEEK MOVING AVERAGE FORECASTS

MOVING AVERAGE FORECAST OF ORDER k

$$F_{t+1} = \frac{\sum(\text{most recent } k \text{ data values})}{k} = \frac{Y_t + Y_{t-1} + \cdots + Y_{t-k+1}}{k}$$

where

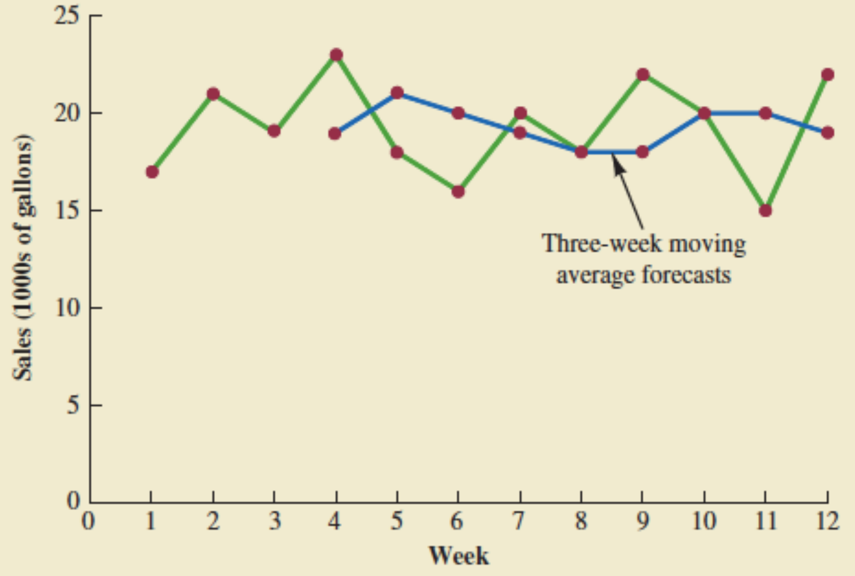
F_{t+1} = forecast of the times series for period $t + 1$
Y_t = actual value of the time series in period t

$$F_4 = \text{average of weeks 1–3} = \frac{17 + 21 + 19}{3} = 19$$

$$F_5 = \text{average of weeks 2–4} = \frac{21 + 19 + 23}{3} = 21$$

$$F_{13} = \text{average of weeks 10–12} = \frac{20 + 15 + 22}{3} = 19$$

Week	Time Series Value	Forecast
1	17	
2	21	
3	19	
4	23	19
5	18	21
6	16	20
7	20	19
8	18	18
9	22	18
10	20	20
11	15	20
12	22	19
		Totals



Weighted Average

$$F_{t+1} = \sum_{k=t+1-N}^{t} W_k \times Y_k$$

$$\sum_{k=t+1-N}^{t} W_k = 1$$

In the moving averages method, each observation in the moving average calculation receives the same weight. one variation, known as weighted moving averages, involves selecting a different weight for each data value and then computing a weighted average of the most recent k values as the forecast. In most cases, the most recent observation receives the most weight, and the weight decreases for older data values. Let us use the gasoline sales time series to illustrate the computation of a weighted three-week moving average. We assign a weight of 3/6 to the most recent observation, a weight of 2/6 to the second most recent observation, and a weight of 1/6 to the third most recent observation. Using this weighted average, our forecast for week 4 is computed as follows:

$$\text{Forecast for week 4} = \tfrac{1}{6}(17) + \tfrac{2}{6}(21) + \tfrac{3}{6}(19) = 19.33$$

note that for the weighted moving average method the sum of the weights is equal to 1.

Exponential smoothing methods – 3 methods

1. The simplest exponential smoothing method is the single smoothing (SES) method where only one parameter needs to be estimated

2. Holt's method makes use of two different parameters and allows forecasting for series with trend

3. Holt-Winters' method involves three smoothing parameters to smooth the data, the trend, and the seasonal index

Simple Exponential Smoothing

- Formally, the exponential smoothing equation is

$$F_{t+1} = \alpha y_t + (1-\alpha)F_t$$

  - F_{t+1} = forecast for the next period.
  - α = smoothing constant.
  - y_t = observed value of series in period t.
  - F_t = old forecast for period t.

- The forecast F_{t+1} is based on weighting the most recent observation y_t with a weight α and weighting the most recent forecast F_t with a weight of 1- α

Why 'exponential'?

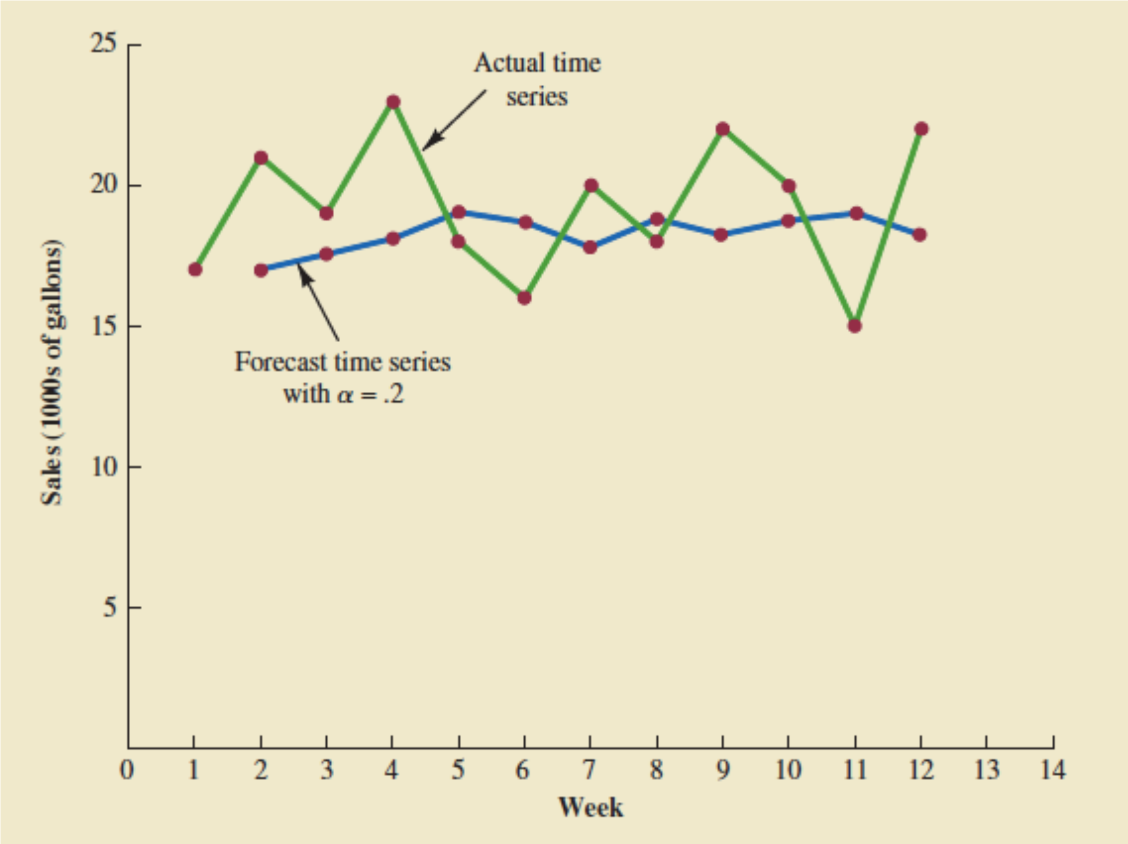
$$
\begin{aligned}
F_{t+1} &= \alpha\, y_t + (1-\alpha)F_t \\
&= \alpha\, y_t + (1-\alpha)[\alpha\, y_{t-1} + (1-\alpha)F_{t-1}] \\
&= \alpha\, y_t + \alpha(1-\alpha)y_{t-1} + (1-\alpha)^2 F_{t-1}
\end{aligned}
$$

$$F_{t+1} = \alpha\, y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \alpha(1-\alpha)^3 y_{t-3} + \cdots + \alpha(1-\alpha)^{t-1} y_1$$

Example

Summary of the exponential smoothing forecasts
With smoothing constant a = .2

Week	Time Series Value	Forecast
1	17	
2	21	17.00
3	19	17.80
4	23	18.04
5	18	19.03
6	16	18.83
7	20	18.26
8	18	18.61
9	22	18.49
10	20	19.19
11	15	19.35
12	22	18.48
		Totals



Example : Choosing a value for α

Summary of the exponential smoothing forecasts
With smoothing constant α = .2

Week	Time Series Value	Forecast	Forecast Error	Squared Forecast Error
1	17			
2	21	17.00	4.00	16.00
3	19	17.80	1.20	1.44
4	23	18.04	4.96	24.60
5	18	19.03	−1.03	1.06
6	16	18.83	−2.83	8.01
7	20	18.26	1.74	3.03
8	18	18.61	−0.61	0.37
9	22	18.49	3.51	12.32
10	20	19.19	0.81	0.66
11	15	19.35	−4.35	18.92
12	22	18.48	3.52	12.39
		Totals	10.92	98.80

Smoothing constant α = .3

Week	Time Series Value	Forecast	Forecast Error	Squared Forecast Error
1	17			
2	21	17.00	4.00	16.00
3	19	18.20	0.80	0.64
4	23	18.44	4.56	20.79
5	18	19.81	−1.81	3.28
6	16	19.27	−3.27	10.69
7	20	18.29	1.71	2.92
8	18	18.80	−0.80	0.64
9	22	18.56	3.44	11.83
10	20	19.59	0.41	0.17
11	15	19.71	−4.71	22.18
12	22	18.30	3.70	13.69
		Totals	8.03	102.83

The value of the sum of squared forecast errors is 102.83; hence MSE = 102.83/11 = 9.35. With MSE = 9.35, we see that, for the current data set, a smoothing constant of α = .3 results in less forecast accuracy than a smoothing constant of α = .2. Thus, we would be inclined to prefer the original smoothing constant of α = .2. Using a trial-and-error calculation with other values of α, we can find a "good" value for the smoothing constant.

Influence of the exponential factor

Alpha in (0,1) and not equal to either 0 or 1

When is alpha small and when large?

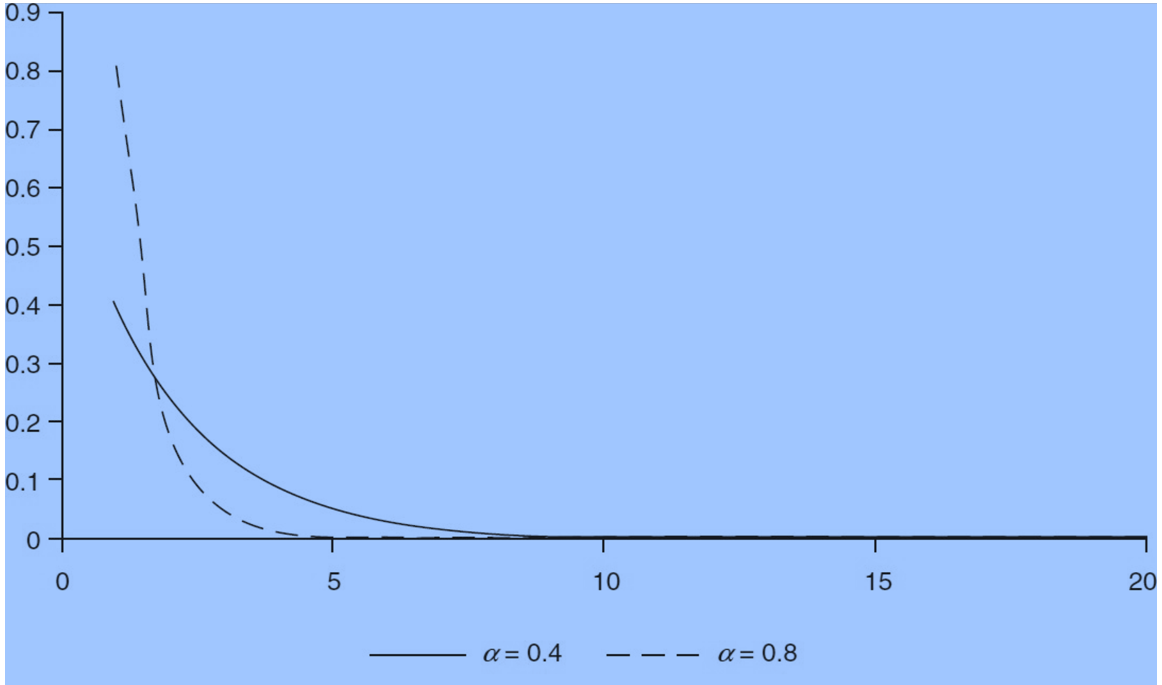


FIGURE 13.3 Exponential decay of weights to older observations.

Some pros and cons of Single Exponential Smoothing

Advantages:

1. It uses all the historic data unlike the moving average where only the past few observations are considered to predict the future value.
2. It assigns progressively decreasing weights to older data.

Some disadvantages of smoothing methods are:

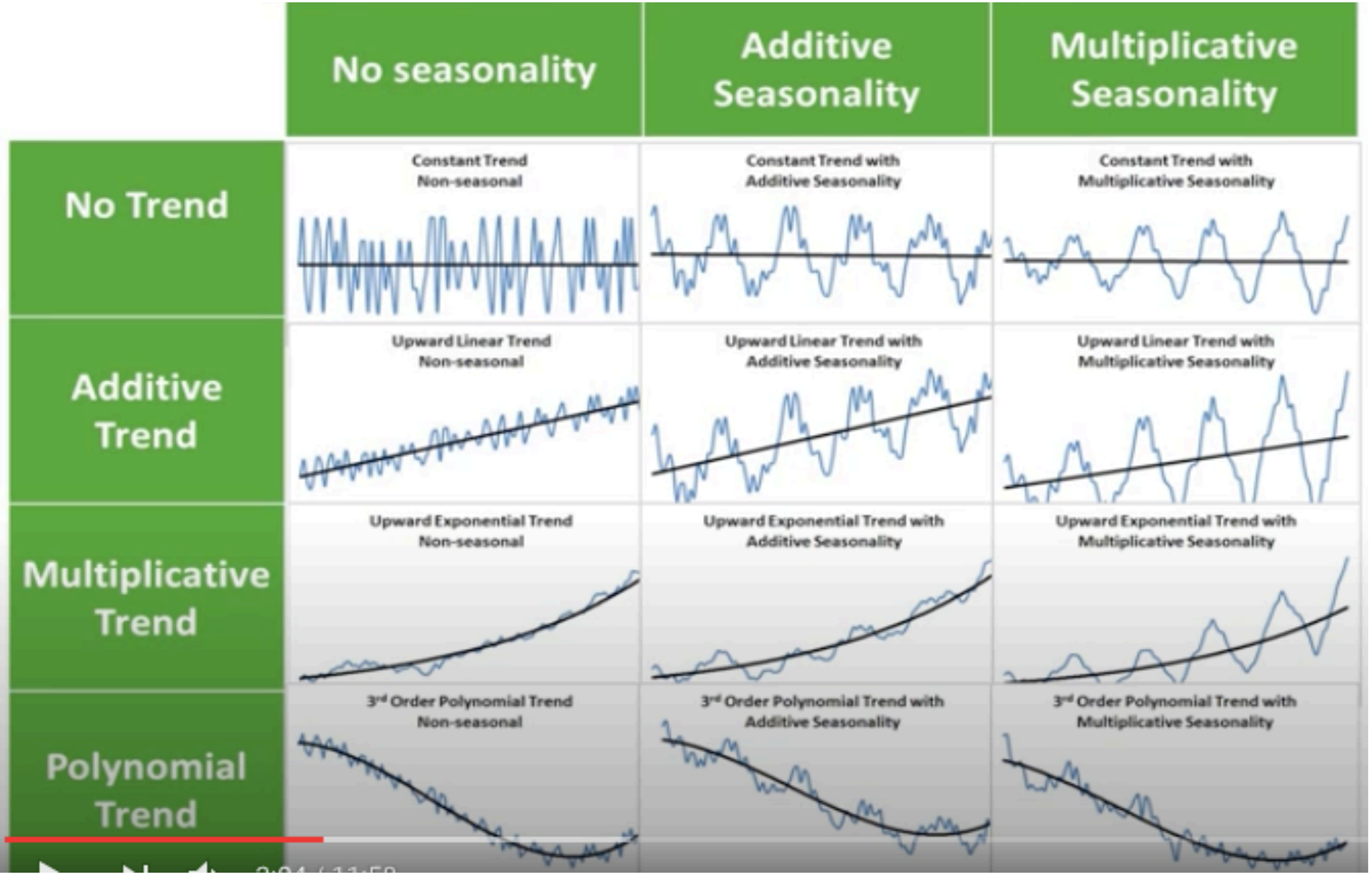
1. Increasing *n* makes forecast less sensitive to changes in data.
2. It always lags behind trend as it is based on past observations. The longer the time period *n*, the greater the lag as it is slow to recognize the shifts in the level of the data points.
3. Forecast bias and systematic errors occur when the observations exhibit strong trend or seasonal patterns.

Holt’s two parameter exponential smoothing

Holt’s two parameter exponential smoothing method is an extension of simple exponential smoothing.

It adds a growth factor (or trend factor) to the smoothing equation as a way of adjusting for the trend.

Types of Time Series Data – Revisit Additive and Multiplicative trend



Holt's two parameter exponential smoothing

Level (or Intercept) equation (L_t):

$$L_t = \alpha \times Y_t + (1 - \alpha) \times F_t \quad (13.12)$$

The trend equation is given by (T_t)

$$T_t = \beta \times (L_t - L_{t-1}) + (1 - \beta) \times T_{t-1} \quad (13.13)$$

α and β are the smoothing constants for level and trend, respectively, and $0 < \alpha < 1$ and $0 < \beta < 1$.

The forecast at time $t + 1$ is given by

$$F_{t+1} = L_t + T_t \quad (13.14)$$

$$F_{t+n} = L_t + nT_t \quad (13.15)$$

where L_t is the level which represents the smoothed value up to and including the last data,
T_t is the slope of the line or the rate of increase or decrease at period t,
n is the number of time periods into the future.

Holt's two parameter exponential smoothing- Additive and Multiplicative Trend

Level (or Intercept) equation (L_t):

$$L_t = \alpha \times Y_t + (1 - \alpha) \times F_t \quad (13.12)$$

The trend equation is given by (T_t)

$$T_t = \beta \times (L_t - L_{t-1}) + (1 - \beta) \times T_{t-1} \quad (13.13)$$

Additive Trend	Multiplicative Trend
Forecast = most recent *estimated level* + *estimated trend*	Forecast = most recent *estimated level* x *estimated trend*
$F_{t+k} = L_t + k\,T_t$	$F_{t+k} = L_t \times (T_t)^k$

Holt's two parameter exponential smoothing

TABLE 13.5 Forecasted values using double exponential smoothing ($\alpha = 0.0328$ and $\beta = 0.9486$)

Month	Actual Demand	L_t	T_t	$F_t (= L_{t-1} + T_{t-1})$	$(Y_t - F_t)^2$	$\lvert Y_t - F_t \rvert / Y_t$
37	3216483	3678293	66894.6916	3693955	2.27979E+11	0.148445
38	3453239	3735612	57810.9617	3745188	85234318285	0.084544
39	5431651	3847157	108782.913	3793423	2.68379E+12	0.301608
40	4241851	3965318	117678.771	3955940	81745109031	0.067402
41	3909887	4077319	112292.624	4082997	29966946139	0.044275
42	3216438	4157691	82013.2329	4189611	9.47066E+11	0.302562
43	4222005	4239124	81462.532	4239704	313269245.9	0.004192
44	3621034	4297641	59696.6025	4320586	4.89374E+11	0.193191
45	5162201	4383737	84739.1839	4357338	6.47805E+11	0.155915
46	4627177	4473682	89677.0074	4468476	25185883916	0.034298
47	4623945	4565346	91562.092	4563359	3670690475	0.013103
48	4599368	4655021	89771.7849	4656908	3310862728	0.01251

Holt's two parameter exponential smoothing

Initial value of L_t is usually taken same as Y_t (that is, $L_t = Y_t$). The starting value of T_t can be taken as $(Y_t - Y_{t-1})$ or the difference between two previous actual values of observations prior to the period for which forecasting is carried out. Another option for T_t is $(Y_t - Y_1)/(t-1)$.

The value of

$$L_1 = Y_1 = 3002666$$

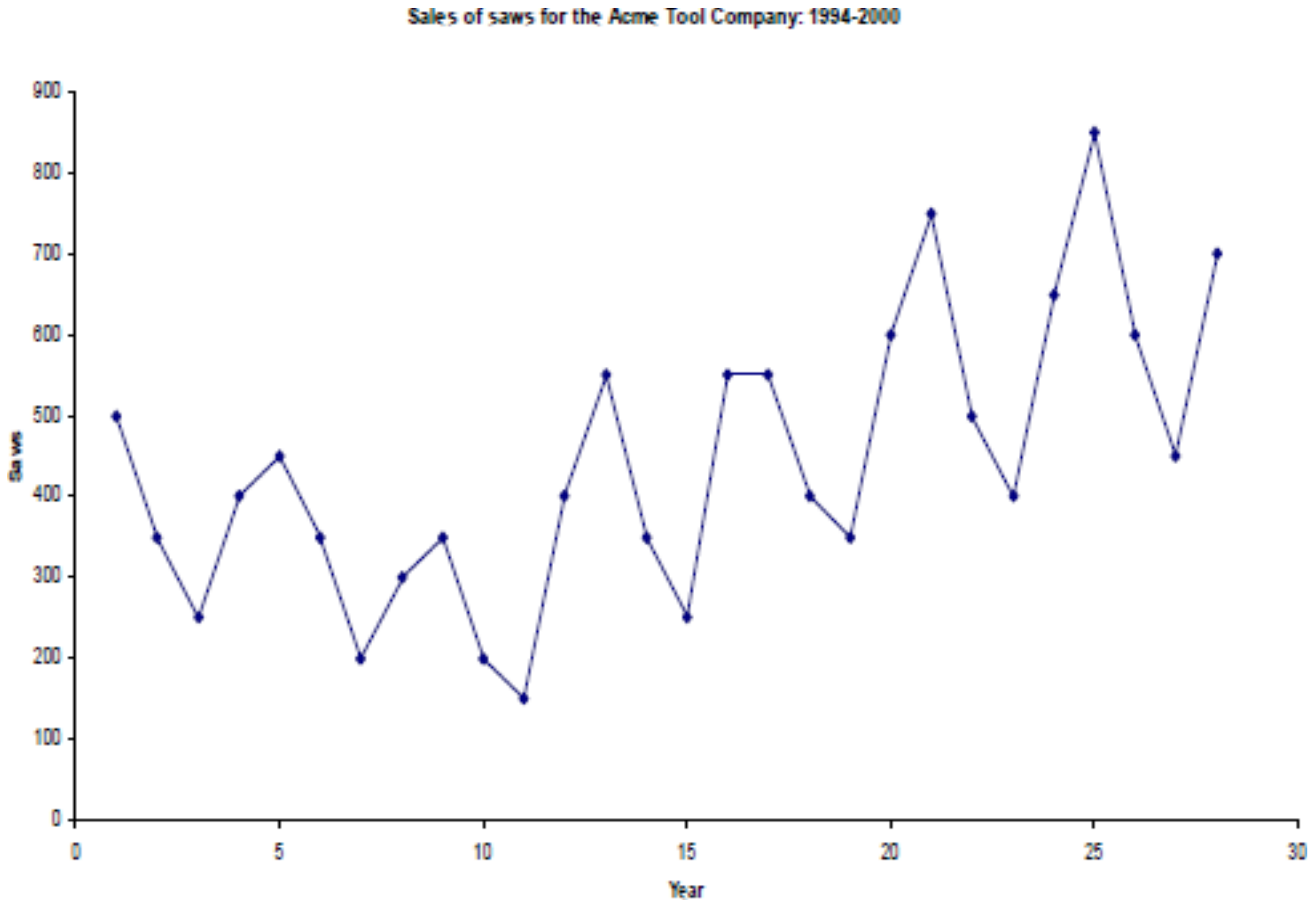
and $$T_1 = (Y_{36} - Y_1)/35 = (4732677 - 3002666)/35 = 49428.8857$$

The value of $$F_2 = L_1 + T_1 = 3002666 + 49428.8857 = 3052095$$

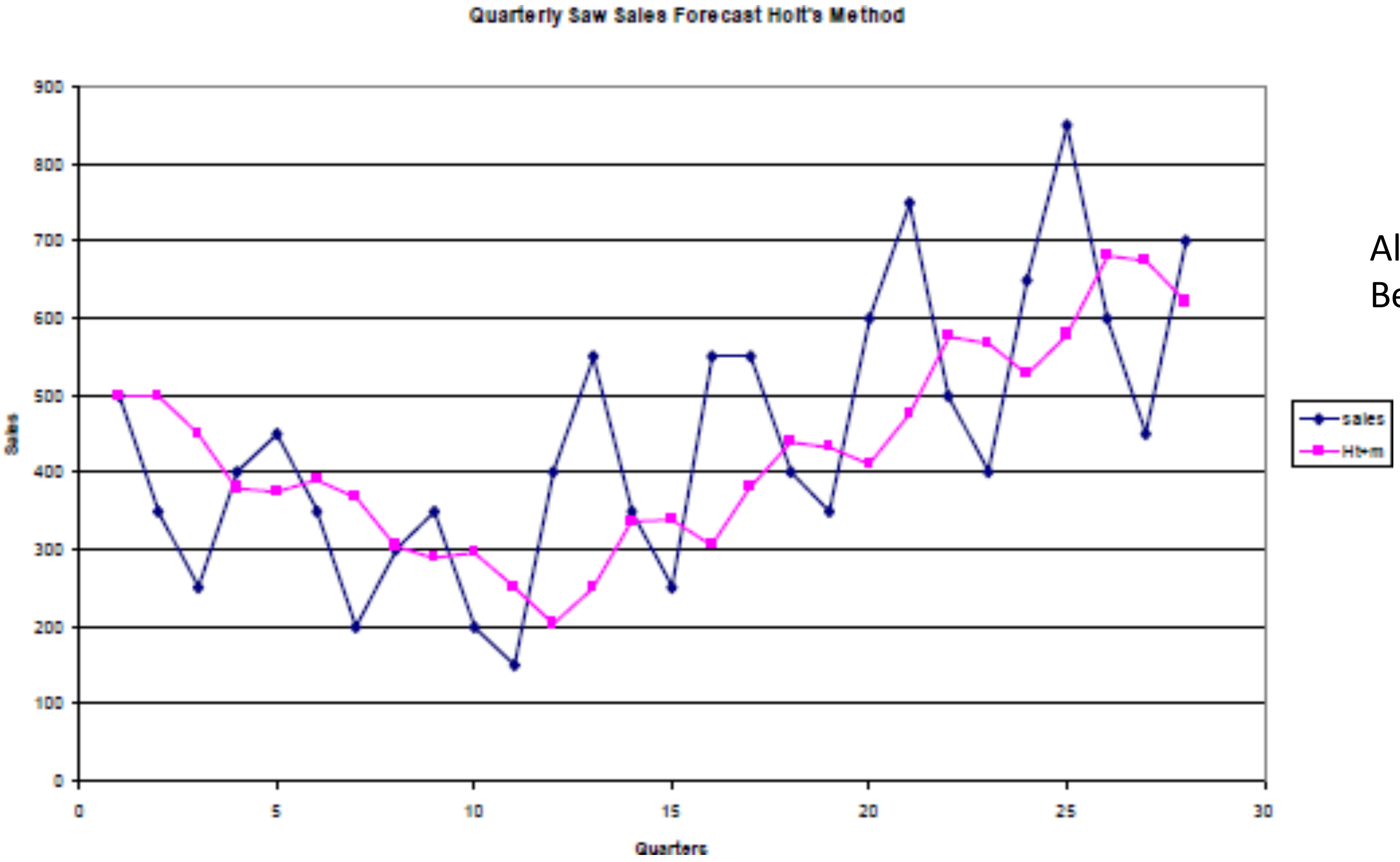
The forecasted values for periods 37 to 48 are shown in Table 13.5 ($\alpha = 0.0328$ and $\beta = 0.9486$).

The RMSE and MAPE of the forecast using double exponential smoothing is given by 659888.9554 and 0.1135 (11.35%). The values of α and β used in Table 13.5 are optimized values of α and β that minimize the root mean square error.

Holt’s exponential smoothing - example



Holt's exponential smoothing - example

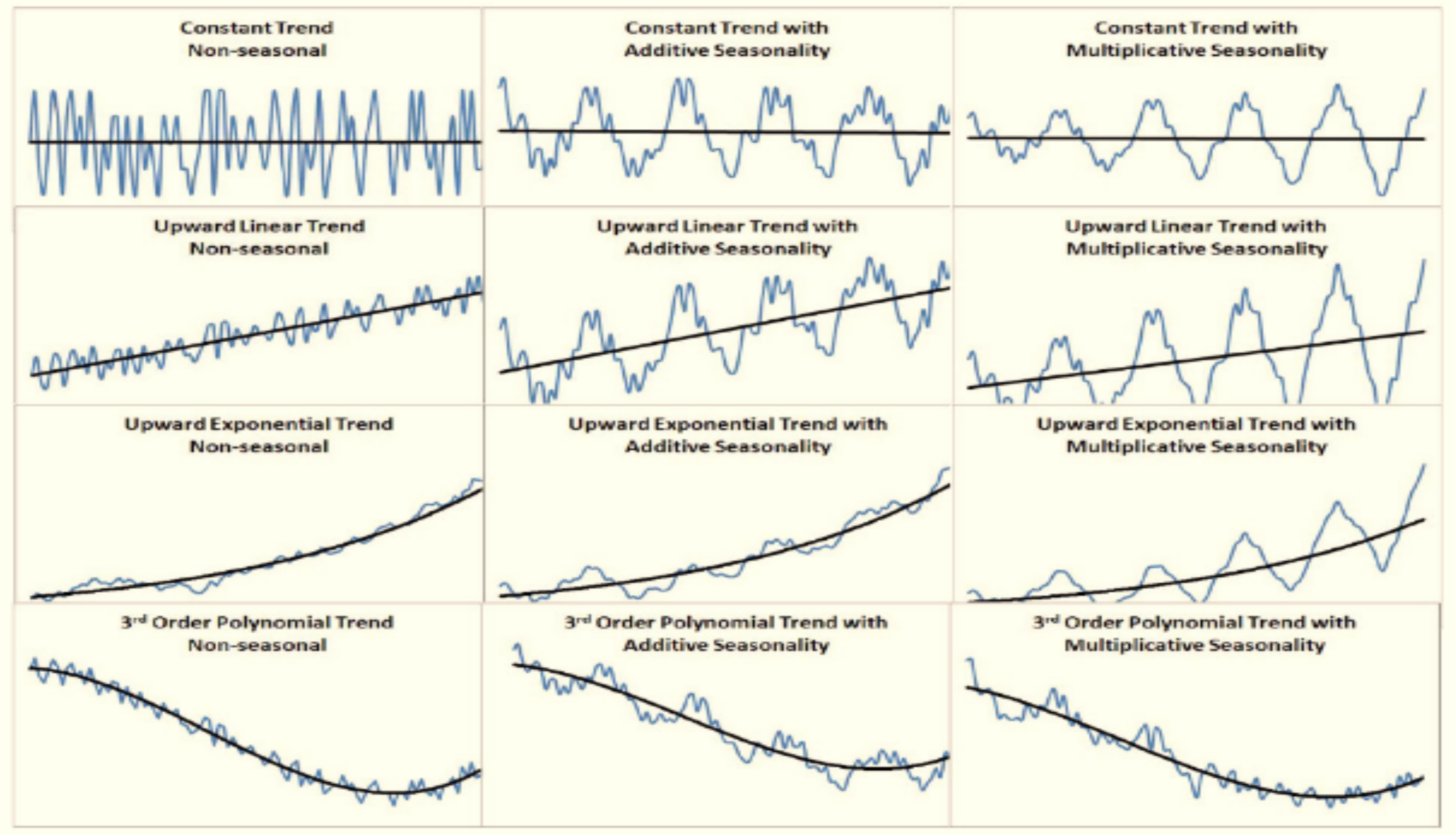


Alpha = 0.3
Beta = 0.1

Triple Exponential Smoothing (Holt Winter's Method)

DATA ANALYTICS

Types of Time Series Data – Revisit to Trend with Additive seasonality and Multiplicative seasonality



TRIPLE EXPONENTIAL SMOOTHING (HOLT-WINTER MODEL)

- Moving averaging and single and double exponential smoothing techniques can handle data as long as the data do not have any seasonal component associated with it.
- When there is seasonality in the time-series data, techniques such as moving average, exponential smoothing, and double exponential smoothing are no longer appropriate.
- In most cases, the fitted error values (actual demand minus forecast) associated with simple exponential smoothing and Holt's method will indicate systematic error patterns that reflect the existence of seasonality.
- For example, presence of seasonality may result in all positive errors, except for negative values that occur at fixed intervals.
- Such pattern in error would imply existence of seasonality.
- Such time series data require the use of a seasonal method to eliminate the systematic patterns in error.

Triple Exponential Smoothing (Holt Winter's Method)

Level (or Intercept) equation:

$$L_t = \alpha \frac{Y_t}{S_{t-c}} + (1-\alpha)\left[L_{t-1} + T_{t-1}\right]$$

Trend equation:

$$T_t = \beta \times (L_t - L_{t-1}) + (1 - \beta)\, T_{t-1}$$

Seasonal equation:

$$S_t = \gamma \frac{Y_t}{L_t} + (1-\gamma) S_{t-c}$$

The forecast F_{t+1} using triple exponential smoothing is given by

$$F_{t+1} = \left[L_t + T_t\right] \times S_{t+1-c}$$

Holt-Winter's Model

Additive Model

- The seasonal component in Holt-Winters' method.
- The basic equations for Holt's Winters' additive method are:

$$L_t = \alpha(y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + T_{t-1})$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$

$$S_t = \gamma(y_t - L_t) + (1-\gamma)S_{t-s}$$

$$F_{t+m} = L + mT_{t-1} + S_{t+m-s}$$

- The initial values for L_s and T_s are identical to those for the multiplicative method.
- To initialize the seasonal indices we use

$$S_1 = y_1 - L_s, \quad S_2 = y_2 - L_s, \cdots, S_s = Y_s - L_s$$

Triple Exponential Smoothing (Holt Winter's Method)

Level (or Intercept) equation:

$$L_t = \alpha \frac{Y_t}{S_{t-c}} + (1-\alpha)\left[L_{t-1} + T_{t-1}\right]$$

Trend equation:

$$T_t = \beta \times (L_t - L_{t-1}) + (1 - \beta)\ T_{t-1}$$

Seasonal equation:

$$S_t = \gamma \frac{Y_t}{L_t} + (1-\gamma) S_{t-c}$$

The forecast F_{t+1} using triple exponential smoothing is given by

$$F_{t+1} = \left[L_t + T_t\right] \times S_{t+1-c}$$

Additive Seasonality	Multiplicative Seasonality
Forecast = most recent *estimated level* *+ trend* *+ seasonality*	Forecast = most recent *estimated* *(level + trend) x seasonality*
$F_{t+k} = L_t + k\,T_t + S_{t+k-M}$	$F_{t+k} = (L_t + k\,T_t)\,S_{t+k-M}$

Initializations for Holt Winter's Method

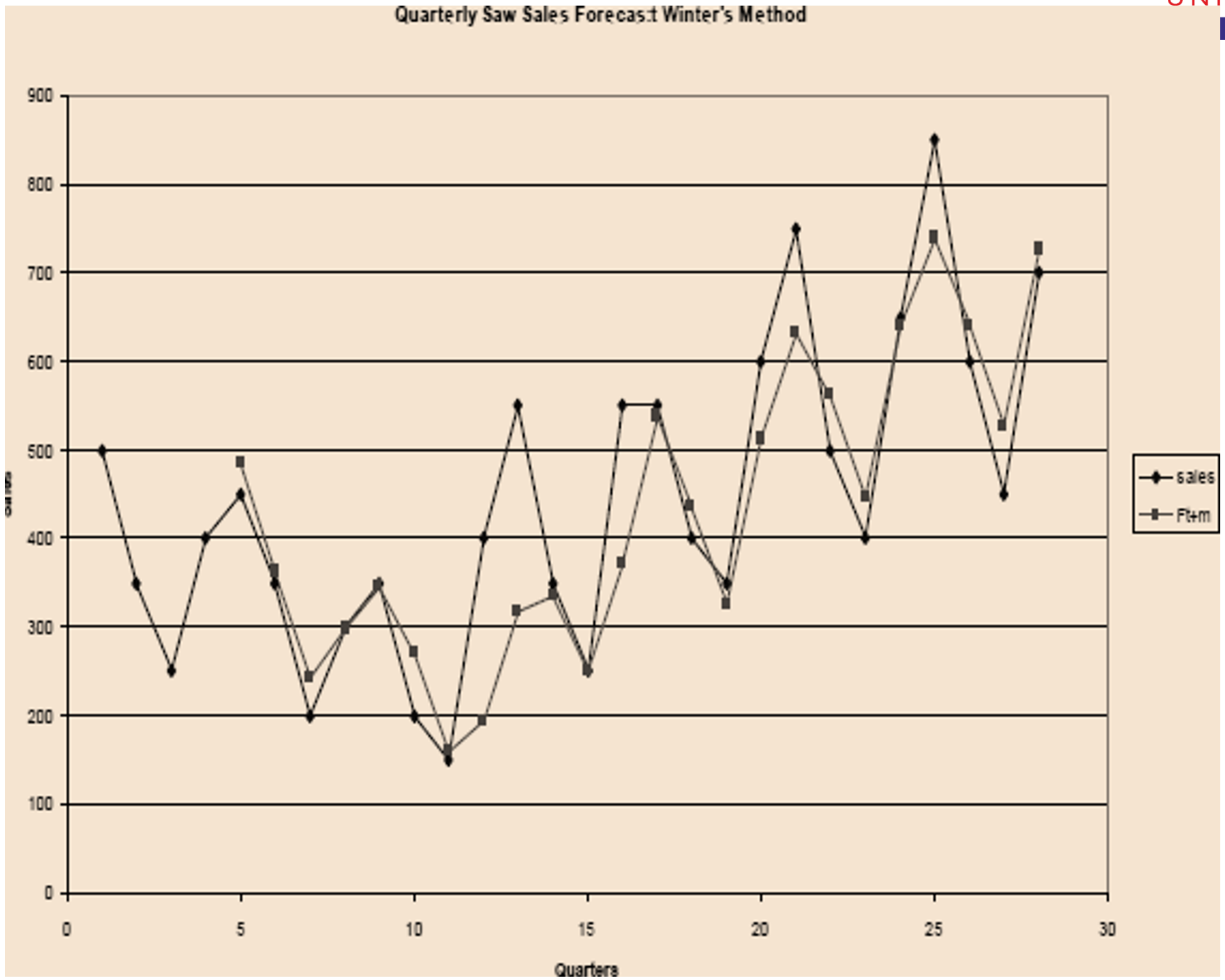
$$L_t = Y_t$$

$$L_t = \frac{1}{c}\left(Y_1 + Y_2 + \cdots + Y_c\right)$$

$$T_t = \frac{1}{c}\left[\frac{Y_t - Y_{t-c}}{12} + \frac{Y_{t-1} - Y_{t-c-1}}{12} + \frac{Y_{t-2} - Y_{t-c-2}}{12} + \ldots + \frac{Y_{t-c+1} - Y_{t-2c+1}}{12}\right]$$

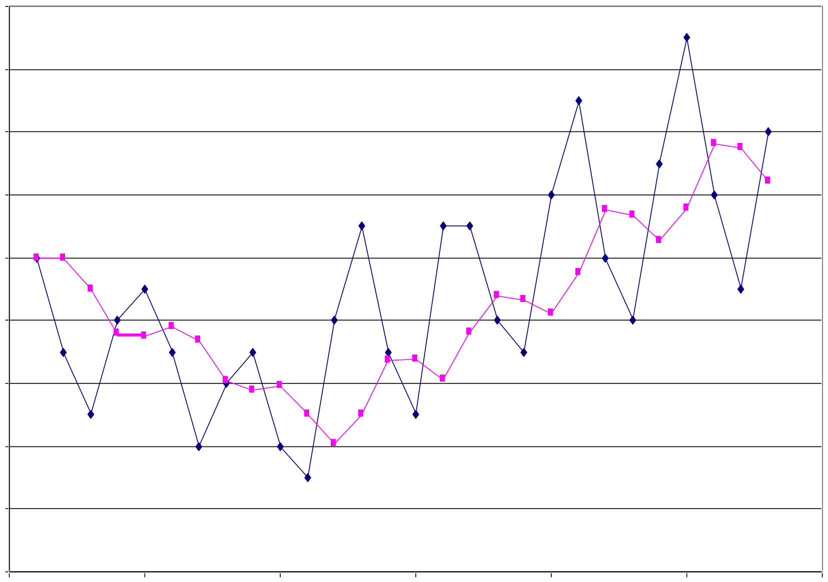
Example of Holt Winter's Method

$$\alpha = 0.4,\ \ \beta = 0.1,\ \ \gamma = 0.3$$
$$\text{and RMSE} = 83.36$$



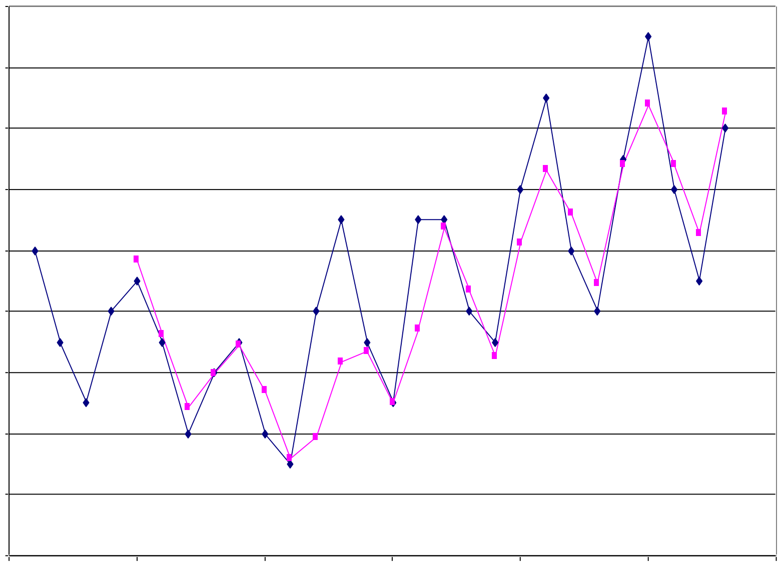
Exponential Models

Quarterly Saw Sales Forecast: Holt's Method



Quarters

Quarterly Saw Sales Forecast: Winter's Method (Multiplicative seasonality)



Quarters

- RMSE for this application is: $\alpha = .3$ and $\beta = .1$

  RMSE = 155.5

- The plot also showed the possibility of seasonal variation that needs to be investigated.

- RMSE for this application is:

  $\alpha = 0.4, \quad \beta = 0.1, \quad \gamma = 0.3$ and

  RMSE = 83.36

- Note the decrease in RMSE.

Computing the Seasonality Index

- Several techniques exist to calculate the initial seasonality index (Winters, 1961; Makridakis *et al.*, 1998;Taylor 2011).
- The initial seasonality index can be calculated using a technique called method of simple averages.
- Several variations to the procedure discussed in next section exist, such as ratio-to-moving average.

Predicting Seasonality Index Using Method of Averages

- The following steps are used for predicting the seasonality index using method of averages:
- **STEP 1**
  1. Calculate the average of value of *Y* for each season that is, if the data is monthly data, then we need to calculate the average for each month based on the training data.
  2. Let these averages be $\bar{Y}_1, \bar{Y}_2, \bar{Y}_3 \ldots\ldots \bar{Y}_c$
- **STEP 2**

  Calculate the average of the seasons' averages calculated in step 1 (say $\bar{Y}$).
- **STEP 3**
  1. The seasonality index for season *k* is given by the ratio $\bar{Y}_k / \bar{Y}$.
  2. to the procedure explained above is first divide the value of Y_t with its yearly average and calculate the seasonal average
  3. We will use first 3 years of data to calculate the seasonality index for various months.

Predicting Seasonality Index Using Method of Averages- Example

TABLE 13.1 Data on sales of shampoo, promotion expenses (in 1000 of rupees), and dummy variable for promotion by competition

Month	Sale Quantity	Promotion Expenses	Competition Promotion	Month	Sale Quantity	Promotion Expenses	Competition Promotion
1	3002666	105	1	25	4634047	165	0
2	4401553	145	0	26	3772879	129	1
3	3205279	118	1	27	3187110	120	1
4	4245349	130	0	28	3093683	112	1
5	3001940	98	1	29	4557363	162	0
6	4377766	156	0	30	3816956	140	1
7	2798343	98	1	31	4410887	160	0
8	4303668	144	0	32	3694713	139	0
9	2958185	112	1	33	3822669	141	1
10	3623386	120	0	34	3689286	136	0
11	3279115	125	0	35	3728654	130	1
12	2843766	102	1	36	4732677	168	0
13	4447581	160	0	37	3216483	121	1
14	3675305	130	0	38	3453239	128	0
15	3477156	130	0	39	5431651	170	0
16	3720794	140	0	40	4241851	160	0
17	3834086	167	1	41	3909887	151	1
18	3888913	148	1	42	3216438	120	1
19	3871342	150	1	43	4222005	152	0
20	3679862	129	0	44	3621034	125	0
21	3358242	120	0	45	5162201	170	0
22	3361488	122	0	46	4627177	160	0
23	3670362	135	0	47	4623945	168	0
24	3123966	110	1	48	4599368	166	0

Predicting Seasonality Index Using Method of Averages

- The seasonality index based on first 3 years of data using method of averages.

- Seasonality index can be interpreted as percentage change from the trend line.
- For example, the seasonality index for January is approximately 1.088 or 108.8% (textbook example).

- This implies that in January, the demand will be approximately 8.8% more from the trend line. The seasonality index for March is 0.8885 or 88.85% (textbook example).

Predicting Seasonality Index Using Method of Averages

- TABLE 1: Seasonality index using method of averages

Month	Sale Quantity (2012)	Sale Quantity (2013)	Sale Quantity (2014)	Monthly Average $\overline{Y}_K$	Seasonality Index $\overline{Y}_k / \overline{\overline{Y}}$
January	3002666	4447581	4634047	4028098.00	1.087932
February	4401553	3675305	3772879	3949912.33	1.066815
March	3205279	3477156	3187110	3289848.33	0.888541
April	4245349	3720794	3093683	3686608.67	0.9957
May	3001940	3834086	4557363	3797796.33	1.02573
June	4377766	3888913	3816956	4027878.33	1.087872
July	2798343	3871342	4410887	3693524.00	0.997568
August	4303668	3679862	3694713	3892747.67	1.051375
September	2958185	3358242	3822669	3379698.67	0.912808
October	3623386	3361486	3689286	3558053.33	0.960979
November	3279115	3670362	3728654	3559377.00	0.961337
December	2843766	3123966	4732677	3566803.00	0.963342
Average of monthly averages				3702528.22	

Predicting Seasonality Index Using Method of Averages

- This implies that the demand in March will be 11.15% less from the trend line.
- Note that, multiplicative model is used in this example.
- To start the triple exponential smoothing, we need to set the starting values of level and trend.

  $L_{36} = Y_{36}/S_{36} = 4732677/0.9633 = 4912983.494$

- The initial value of trend (T_{36}) can be calculated based on second and third year by using

$$T_{36} = \frac{1}{12}\left[\frac{Y_{36}-Y_{24}}{12} + \frac{Y_{35}-Y_{23}}{12} + \frac{Y_{34}-Y_{22}}{12} + \cdots + \frac{Y_{25}-Y_{13}}{12}\right]$$

$$T_{36} = \frac{1}{12}\left[\frac{4732677-3123966}{12} + \frac{3728654-3670362}{12} + \ldots + \frac{4634047-4447581}{12}\right] = 21054.35$$

Predicting Seasonality Index Using Method of Averages

- The forecast for period 37 using triple exponential smoothing is given by

$$F_{37} = [L_{36} + T_{36}] \times S_{37-12} = [L_{36} + T_{36}] \times S_{25}$$

- The seasonal index S_{25} (seasonality index for January) is 1.088.
- Substituting the values of L_{36}, T_{36} and S_{25}, we get

$$F_{37} = [4912983.494 + 21054.35] \times 1.088 = 5368233.2$$

Predicting Seasonality Index Using Method of Averages

- TABLE 3: Forecasting using triple exponential smoothing (values differ for different round off values of parameters)

Month t	Actual Demand	L_{t-1}	T_{t-1}	S_t	F_t	$(Y_t - F_t)^2$	$\|Y_t - F_t\|/Y_t$
37	3216482	4912983.49	21054.35	1.09	5367895.97	4.62858E+12	0.668872
38	3453239	4301229.28	−295349.93	1.07	4273531.48	6.7288E+11	0.237543
39	5431651	3759825.78	−418376.71	0.89	2969014.38	6.06458E+12	0.453386
40	4241851	4228345.39	25071.45	1.00	4235127.90	45200134.5	0.001585
41	3909887	4255577.53	26151.79	1.03	4391900.21	2.32337E+11	0.123281
42	3216437	4131354.31	−49035.71	1.09	4441041.44	1.49966E+12	0.380733
43	4222004	3722098.55	−229145.74	1.00	3484457.63	5.43975E+11	0.174691
44	3621034	3729543.06	−110850.61	1.05	3804603.81	33697874118	0.050695
45	5162201	3562820.55	−138786.56	0.91	3125486.18	4.14821E+12	0.394544
46	4627176	4138038.05	218215.47	0.96	4186269.09	1.94399E+11	0.095286
47	4623945	4503072.73	291625.07	0.96	4609319.06	213918263.4	0.003163
48	4599368	4799566.34	294059.34	0.96	4906905.01	94579010434	0.066865

Predicting Seasonality Index Using Method of Averages

- The forecast for the period 37 to 48 for the data in Table 2 is given in Table 3 .
- Note that the values such as seasonality index are rounded to two decimals, the forecast values will be different if the actual seasonality index values are used.
- The RMSE and MAPE using triple exponential smoothing are 1228588.29 and 0.2208 (22.08%), respectively.
- The values of a = 0.32, b = 0.5, and γ = 1 are used for calculating the level, trend, and seasonal components.

- It is important to note that the exponential smoothing techniques are very sensitive to initial values of level, trend, and seasonal index

Choosing a Method

Table 1. Choosing a Classic Time-series Forecasting Method

No Trend or Seasonality	Trend Only, No Seasonality	Seasonality Only, No Trend	Both Trend and Seasonality
Single exponential smoothing	Double exponential smoothing	Seasonal additive	Holt-Winters' additive
Single moving average	Double moving average	Seasonal multiplicative	Holt-Winters' multiplicative
	Damped trend smoothing		Damped trend additive
			Damped trend multiplicative

Summary

- Simple Exponential(series with no trend or seasonality)
- Double Exponential(for seris with trend)
- Triple Exponential(for series with trend and seasonality)

Summary

Forecast =
estimated level at most recent time point

$$F_{t+k} = L_t$$

Additive Trend	Multiplicative Trend
Forecast = most recent *estimated level* + *estimated trend* $F_{t+k} = L_t + k\,T_t$	Forecast = most recent *estimated level* x *estimated trend* $F_{t+k} = L_t \times (T_t)^k$

Additive Seasonality	Multiplicative Seasonality
Forecast = most recent *estimated level* + *trend* + *seasonality* $F_{t+k} = L_t + k\,T_t + S_{t+k-M}$	Forecast = most recent *estimated (level + trend) x seasonality* $F_{t+k} = (L_t + k\,T_t)\,S_{t+k-M}$

Class Project

- Choice of problem
- Choice of data
- Literature review
- Outcome

References

Text Book:

- "Business Analytics, The Science of Data-Driven Making", U. Dinesh Kumar, Wiley 2017 Chapter 13.4-13.7

Additional reference (for the interested reader)

•"Introduction to Time Series and Forecasting", Second Edition
Peter J. Brockwell, Richard A. Davis Springer 2002

Image Courtesy

https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Time+Series+Analysis:+The+Basics

Business Analytics, Data Analysis and Decision making by Albright and Winston

THANK YOU

Jyothi R.
Assistant Professor,Department of Computer Science
jyothir@pes.edu