PES UNIVERSITY ONLINE

DATA ANALYTICS

Unit 3: Introduction to Time series data

Jyothi R.
Department of Computer Science and Engineering

INTRODUCTION TO FORECASTING

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

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

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

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

- 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

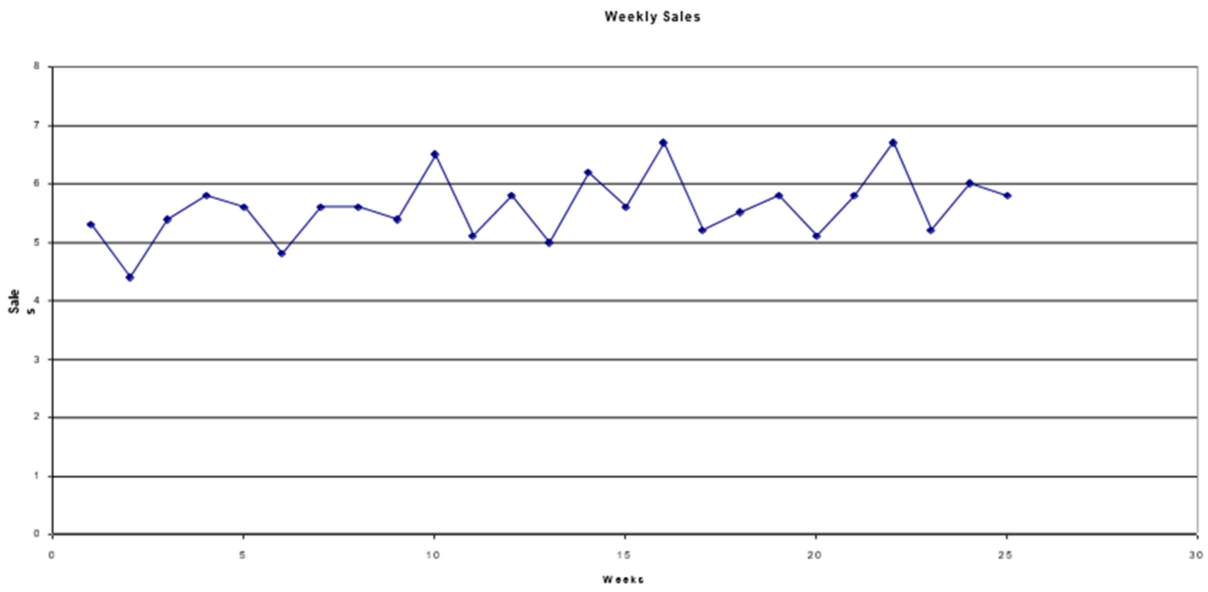
- Equal weights are assigned to each observation used in the average

- 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

Example: Weekly Department Store Sales

- The weekly sales figures (in millions of dollars) presented in the following table are used by a major department store to determine the need for temporary sales personnel



Week (t)	Sales (y)
1	5.3
2	4.4
3	5.4
4	5.8
5	5.6
6	4.8
7	5.6
8	5.6
9	5.4
10	6.5
11	5.1
12	5.8
13	5
14	6.2
15	5.6
16	6.7
17	5.2
18	5.5
19	5.8
20	5.1
21	5.8
22	6.7
23	5.2
24	6
25	5.8

Example: Weekly Department Store Sales

- Use a three-week moving average (k=3) for the department store sales to forecast for the week 24 and 26.

Week	Sales
21	5.8
22	6.7
23	5.2
24	6
25	5.8

$$\hat{y}_{24} = \frac{(y_{23} + y_{22} + y_{21})}{3} = \frac{5.2 + 6.7 + 5.8}{3} = 5.9$$

- The forecast error is
  - $e_{24} = y_{24} - \hat{y}_{24} = 6 - 5.9 = 0.1$

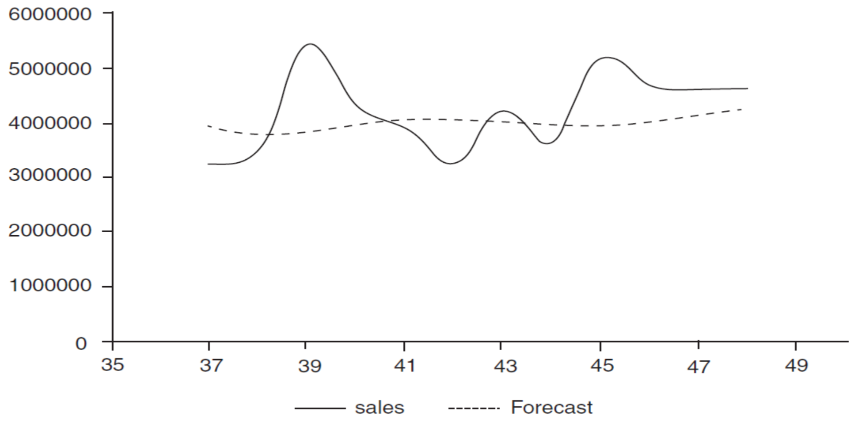


FIGURE 13.2 Plot of actual sales forecasted sales using moving average.

Weighted Average

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

Week	Sales
21	5.8
22	6.7
23	5.2
24	6
25	5.8

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

Exponential smoothing methods

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

- Holt’s method makes use of two different parameters and allows forecasting for series with trend

- 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 $1- \alpha$

Why ‘exponential’?

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

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

Influence of the exponential factor

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

When is alpha small and when large?

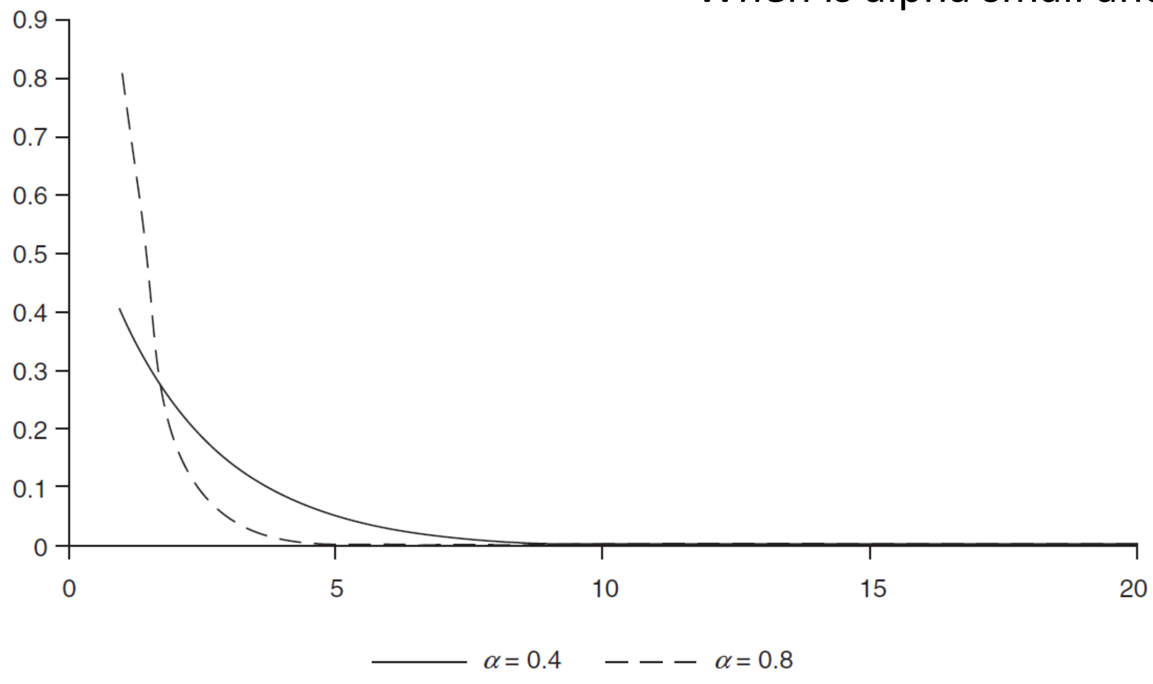


FIGURE 13.3 Exponential decay of weights to older observations.

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.

Level (or Intercept) equation (L_t):

$$L_t = \alpha \times Y_t + (1 - \alpha) \times F_t \tag{13.12}$$

The trend equation is given by (T_t)

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

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

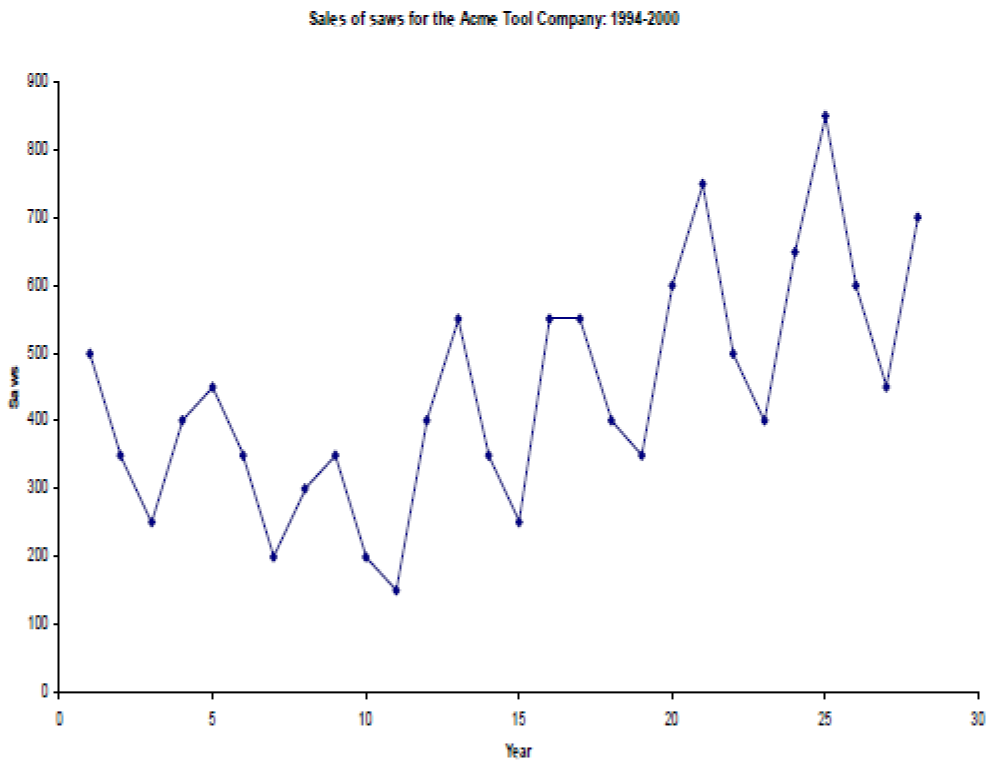
$$F_{t+1} = L_t + T_t \tag{13.14}$$

$$F_{t+n} = L_t + nT_t \tag{13.15}$$

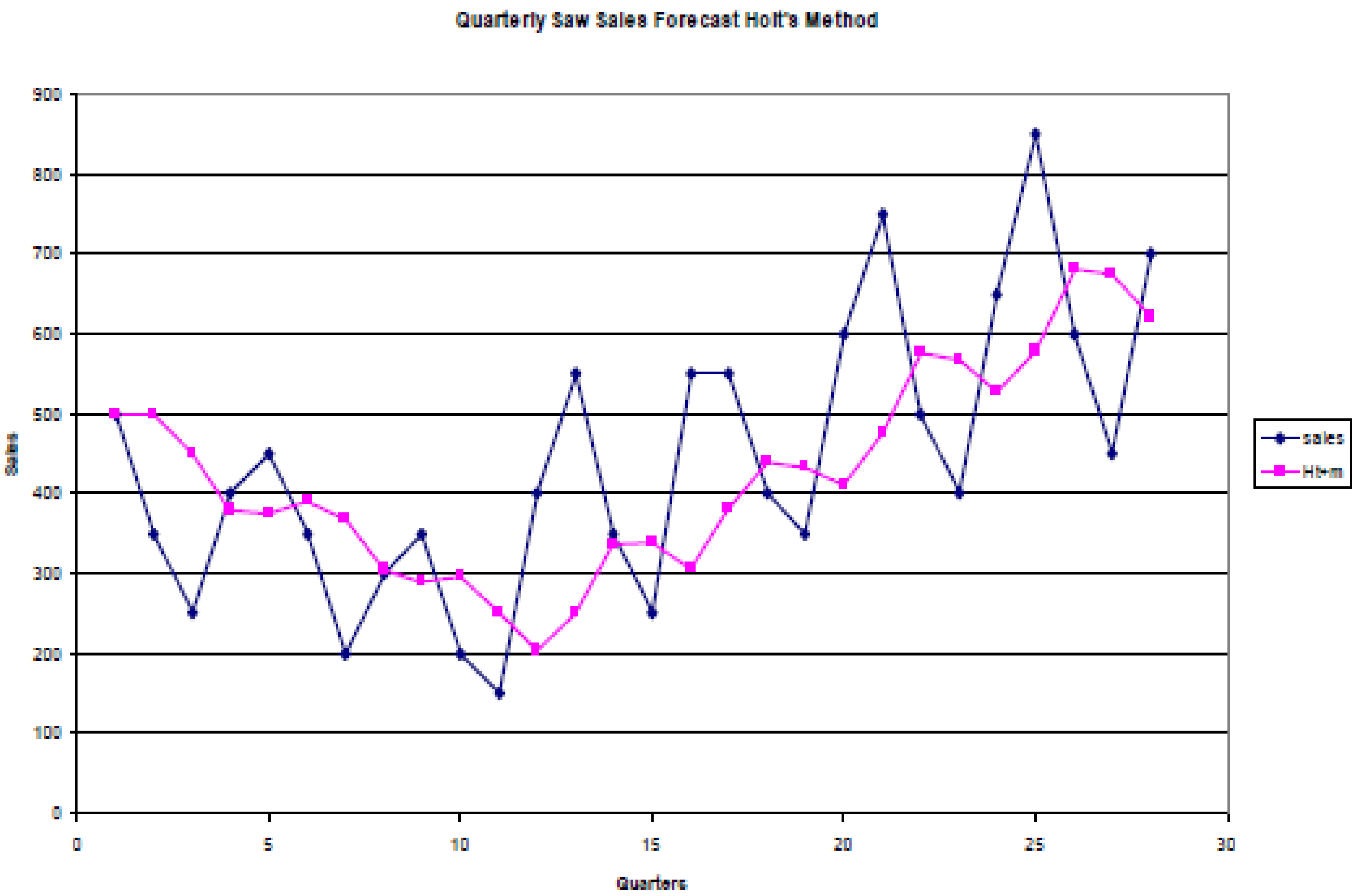
where L_t is the level which represents the smoothed value up to and including the last data, T_t is the slope of the line or the rate of increase or decrease at period t, n is the number of time periods into the future.

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 $(Y_t - Y_1)/(t - 1)$.

Holt’s exponential smoothing - example



Holt’s exponential smoothing - example



Alpha = 0.3
Beta = 0.1

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

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

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

Note: this is a multiplicative model

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

Initializations for Holt Winter's Method

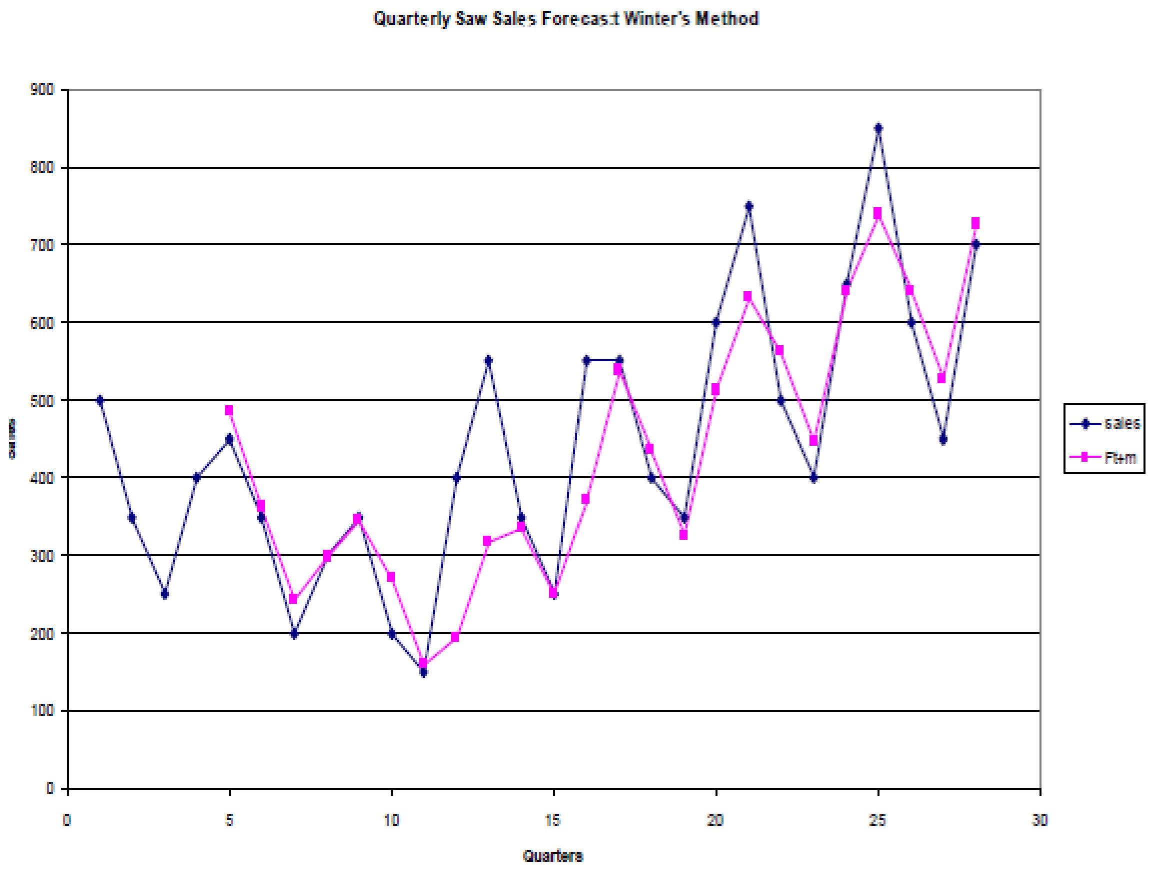
$$L_t = Y_t$$

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

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

Example of Holt Winter's Method

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



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

THANK YOU

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