DATA ANALYTICS

Unit 3: Time Series Analysis

Jyothi R.
Department of Computer Science and Engineering

1. Forecasting → predictive analytics to find significant impact of bottom & top line of org.

2. Time Series data

↳ Data is time dependent

analog
A/D Convertor

Quantization or discrete

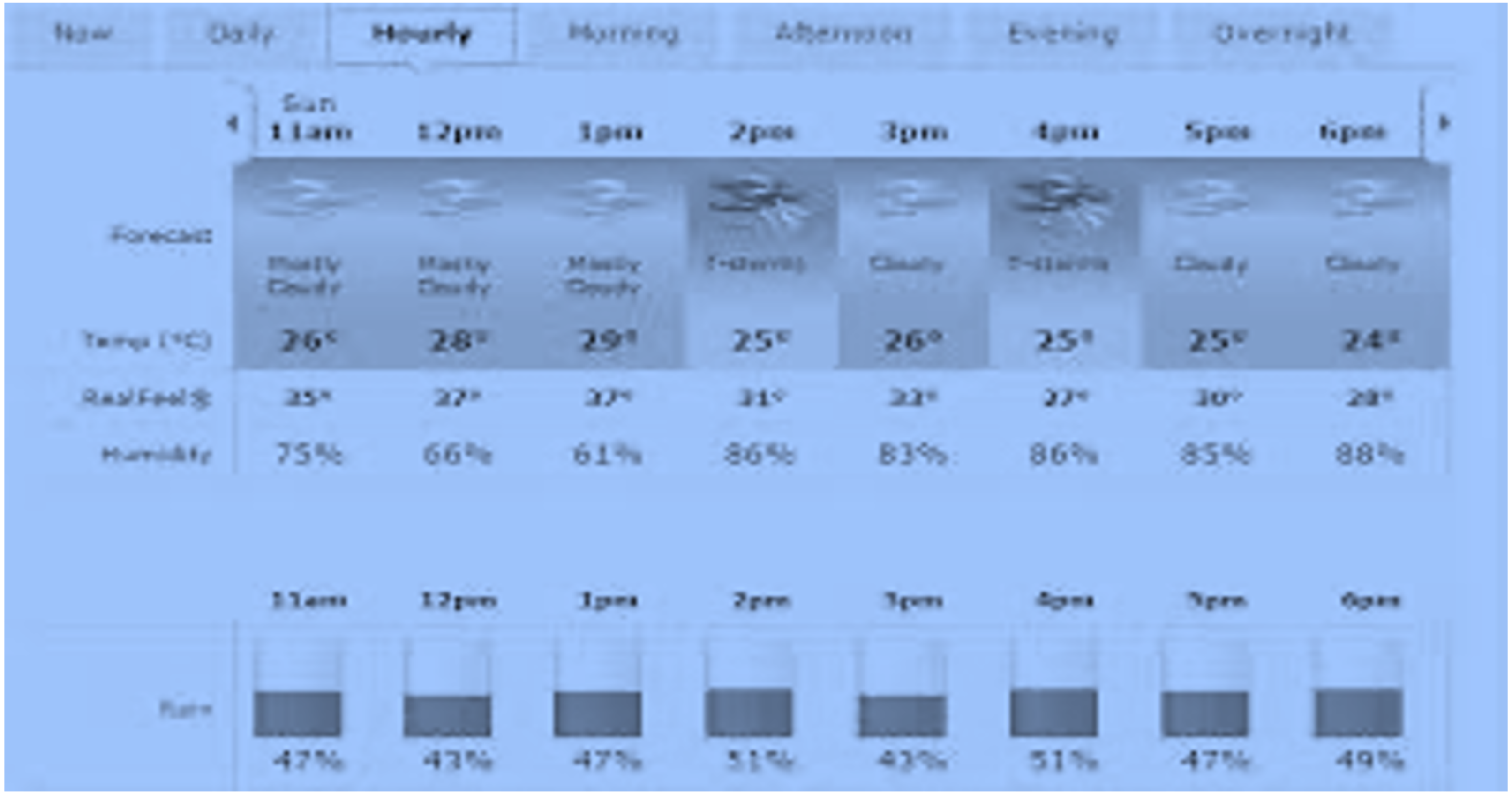
DATA ANALYTICS

Unit 3: Introduction to Time Series Data

Jyothi R., Gowri Srinivasa

Department of Computer Science and Engineering

Weather Forecast



StockMarket Forecast

DATA ANALYTICS

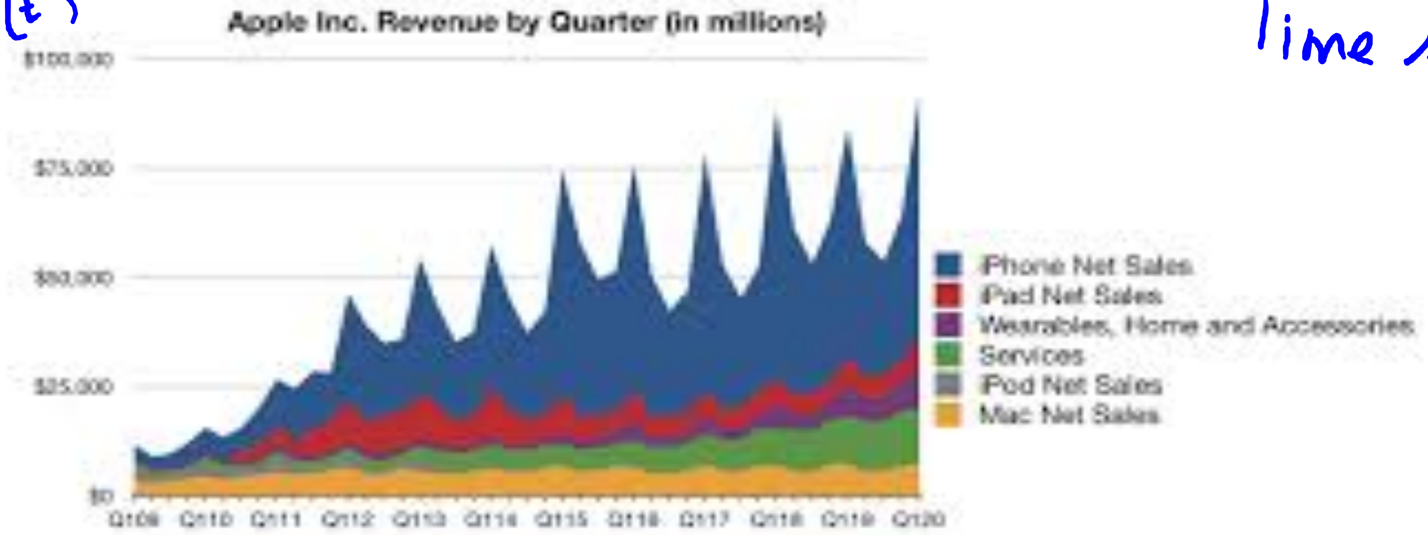
Apple Inc –Quarterly sales

TSF → science of predicting events or data measurements in future.

data "patterns", "trends", & "cycles".

"Time series decomposition"

Revenue(t)



t

What is Time series Forecasting ?

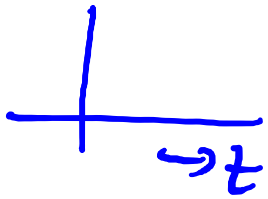
PES UNIVERSITY ONLINE

Time series analysis comprises methods for analyzing **time series** data in order to extract meaningful statistics and other characteristics of the data.

Time series forecasting is the use of a model to predict future values based on previously observed values.

predictive analytics

Time series data: A set of observations on the values that a variable takes at different times.



Chapter: 13

INTRODUCTION TO FORECASTING

(Amazon)

- Forecasting can be very challenging with stock keeping units (SKUs) running into several millions.

1. Boeing 747-400 has more than 6 million parts and several thousand unique parts (Hill, 2011). Forecasting demand for spare parts is important since non-availability of mission critical parts can result in aircraft on ground (AOG) which can be very expensive for airlines.

forecasting

2. Amazon.com sells more than 350 million products through its E-commerce portal. Amazon itself sells about 13 million SKUs and has more (about 2 million) retailers selling their products through Amazon (Ali, 2017).

3. Walmart sells more than 142,000 products through their supercenters. Being a brick-and-mortar retail store, Walmart has to maintain stock for each and every product sold and predict demand for the products as accurately as possible.

INTRODUCTION TO FORECASTING

demand on products + services. ↓ manpower p

- Forecasting - important and frequently addressed problems in analytics
- Inaccurate forecasting has a significant impact
- For example
  - non-availability of product → customer dissatisfaction
  - too much inventory → erodes the organization's profit
- Necessary to forecast the demand for a product and service as accurately as possible.
- Every organization prepares long-range and short-range planning
  - forecasting demand for product and service is an important input for both long-range and short-range planning
- Budget allocation, manpower, warehouse capacity, machine resource planning, etc., based on forecast of demand for a product

Time Series data — two types

1. Univariate time series ⇓ Single variable

y_t → t

2. Multivariate time Series data — multiple

1) demand of product
2) price of product
3) promotion of product

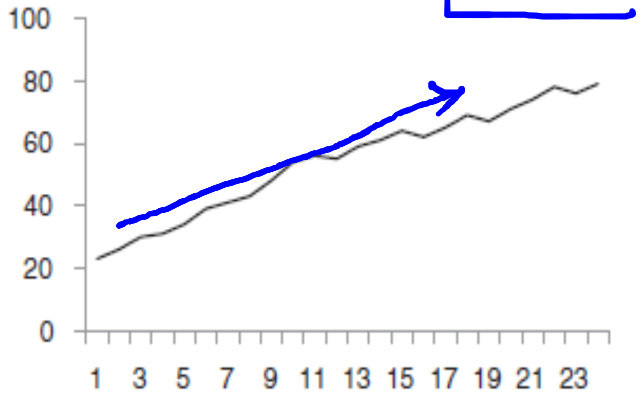
→ t

1. Trend, 2) Seasonal 3) Cyclical 4) Irregular
T_t S_t C_t I_t
Random error in Reg model

COMPONENTS OF TIME-SERIES DATA

From a forecasting perspective, a time-series data can be broken into the following components

1. **Trend Component (T_t):** Trend is the consistent long-term upward or downward movement of the data over a period of time.



(a) Trend

Bull

Bear

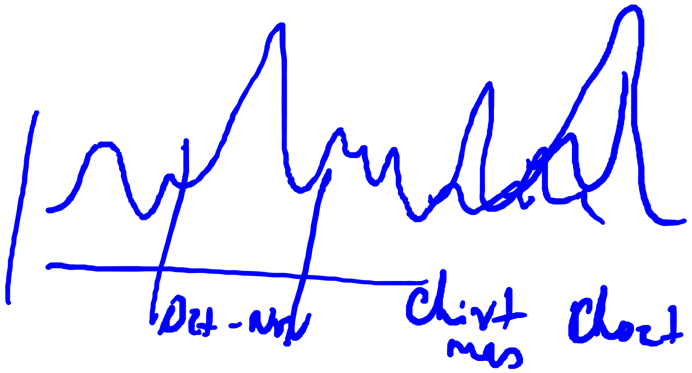
https://school.stockcharts.com/doku.php?id=chart_analysis:trend_lines data pattern that repeats over a certain period of time.

COMPONENTS OF TIME-SERIES DATA Contd.

Sales of ice cream { Summer, Winter }

2. Seasonal Component (S_t): Seasonal component is the repetitive upward or downward movement (or fluctuations) from the trend that occurs within a calendar year such as seasons, quarters, months, days of the week, etc.

- The upward or downward fluctuation may be caused due to festivals, customs within a society, school holidays, business practices within the market such as 'end of season sale', and so on.
- For example, in India demand for many products surge during the festival months of October - December.
- Seasonal fluctuation occurs at fixed intervals (such as months, quarters) known as periodicity of seasonal variation and repeats over time.



Seasonal Component (S_t): Contd.

The seasonal component consists of effects that are reasonably stable with respect to timing, direction and magnitude. It arises from systematic, calendar related influences such as:

- **Natural Conditions:** Weather fluctuations that are representative of the season (uncharacteristic weather patterns such as snow in summer would be considered irregular influences) x
- **Business and Administrative procedures:**
  Start and end of the school term – academic
- **Social and Cultural behavior:**
  Christmas

Seasonal Component (S_t):

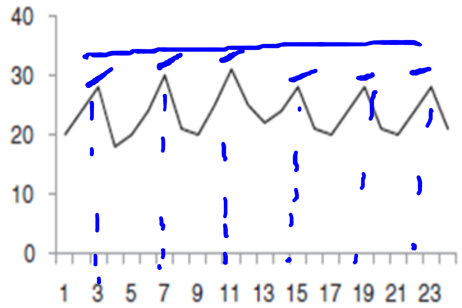
It also includes calendar related systematic effects that are not stable in their annual timing or are caused by variations in the calendar from year to year, such as:

- **Trading Day Effects**

the number of occurrences of each of the day of the week in a given month will differ from year to year
- There were 4 weekends in March in 2000, but 5 weekends in March of 2002

- **Moving Holiday Effects**

holidays which occur each year, but whose exact timing shifts
- Diwali, Easter, Ramadan



(b) Seasonality (fixed periodicity)

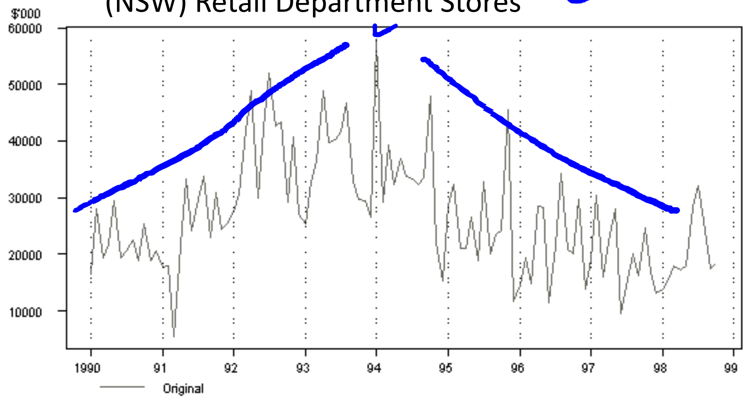
Identifying seasonal components:
Regularly spaced peaks and troughs which have a consistent direction and approximately the same magnitude every year, relative to the trend

Seasonal Component (S_t):

- Seasonality in a time series can be identified by regularly spaced peaks and troughs which have a consistent direction and approximately the same magnitude every year, relative to the trend.
- In this example, the magnitude of the seasonal component increases over time, as does the trend.

Monthly Retail Sales in New South Wales

(NSW) Retail Department Stores



Obvious large seasonal increase in December retail sales in New South Wales due to Christmas shopping

Seasonality : smaller duration

COMPONENTS OF TIME-SERIES DATA

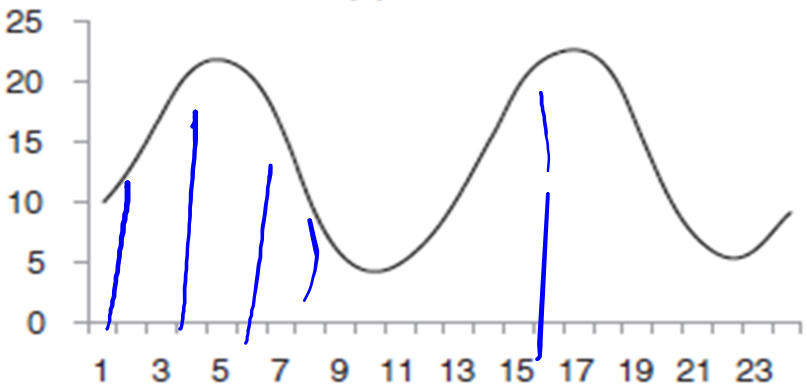
longer duration

3. Cyclical Component (C_t): Cyclical component is fluctuation around the trend line that happens due to macro-economic changes such as recession, unemployment, etc.

- Cyclical fluctuations have repetitive patterns with a time between repetitions of more than a year

Recession: 2002, 2008, 2012

data pattern that occur over several years of time.

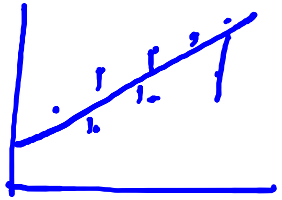


2002 (c) Cyclicality 2008

Economic expansion & contraction

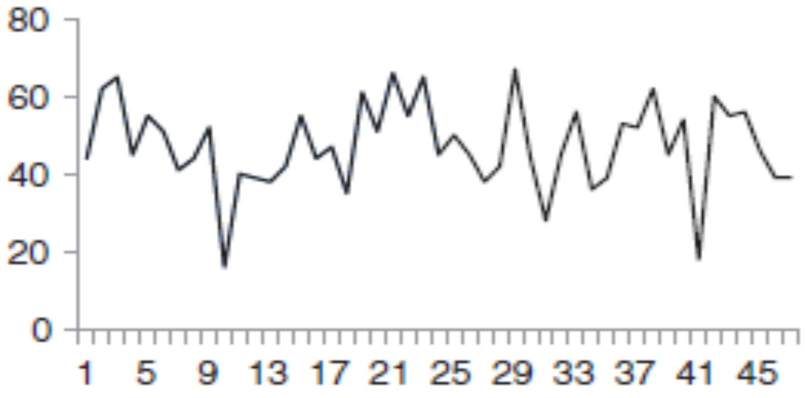
- A major difference between seasonal fluctuation and cyclical fluctuation is that seasonal fluctuation occurs at fixed period within a calendar year, whereas cyclical fluctuations have random time between fluctuations.
- That is, periodicity of seasonal fluctuations is constant, whereas the periodicity of cyclical fluctuations is not constant.

COMPONENTS OF TIME-SERIES DATA contd.

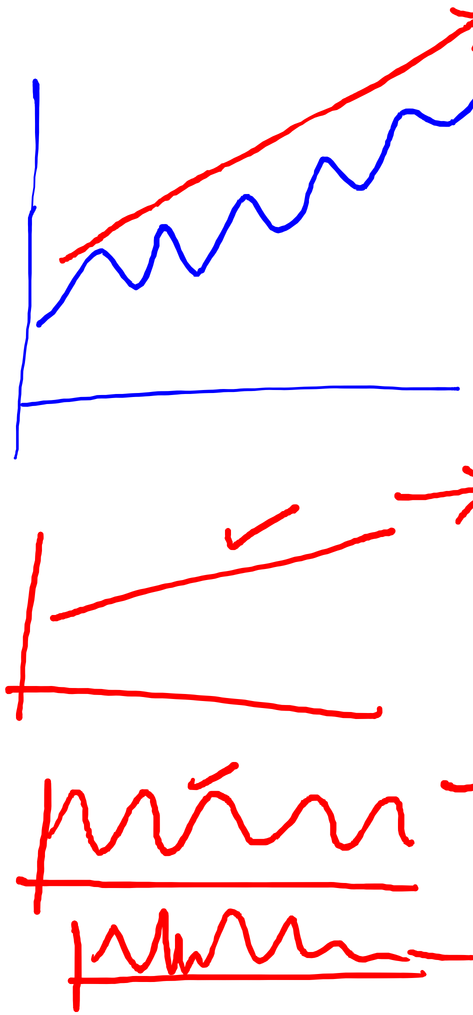


$e \sim N(0, \sigma^2)$

4. Irregular Component (I_t): Irregular component is the white noise or random uncorrelated changes that follow a normal distribution with mean value of 0 and constant variance.



(d) Irregular



trend

t

T_t

S_t

Noise $\sim N(0, \sigma^2)$

Time Series decomposition

- What remains after the seasonal and trend components of a time series have been estimated and removed.
- It results from short term fluctuations in the series which are neither systematic nor predictable

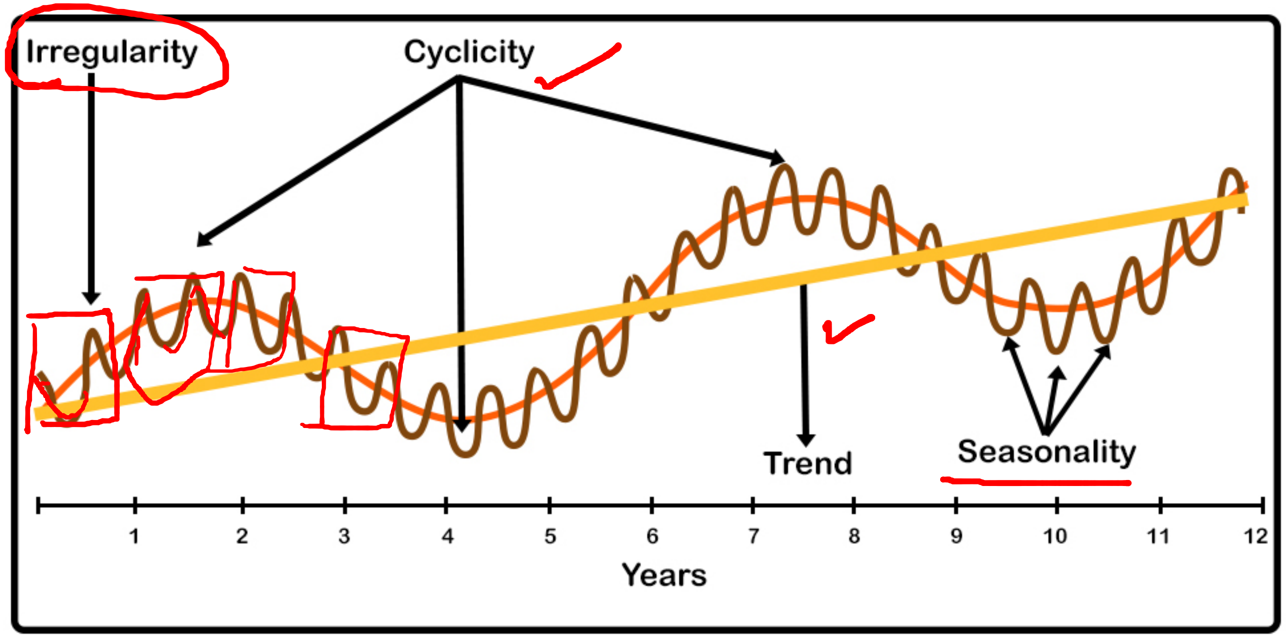
Time Series Decomposition

Trend → T_t
Seasonal → S_t
Cyclic → C_t
Irregular → I_t

Time series data can exhibit a variety of patterns, and it is often helpful to split a time series into several components, each representing an underlying pattern category.

$T_t + S_t + I_r$
$T_t + S_t + C_t + I_r$
$T_t + C_t + I_r$

Time Series forecasting helps us recognize and adjust data patterns, trends and Cycles. These patterns are called "Time Series Decomposition"

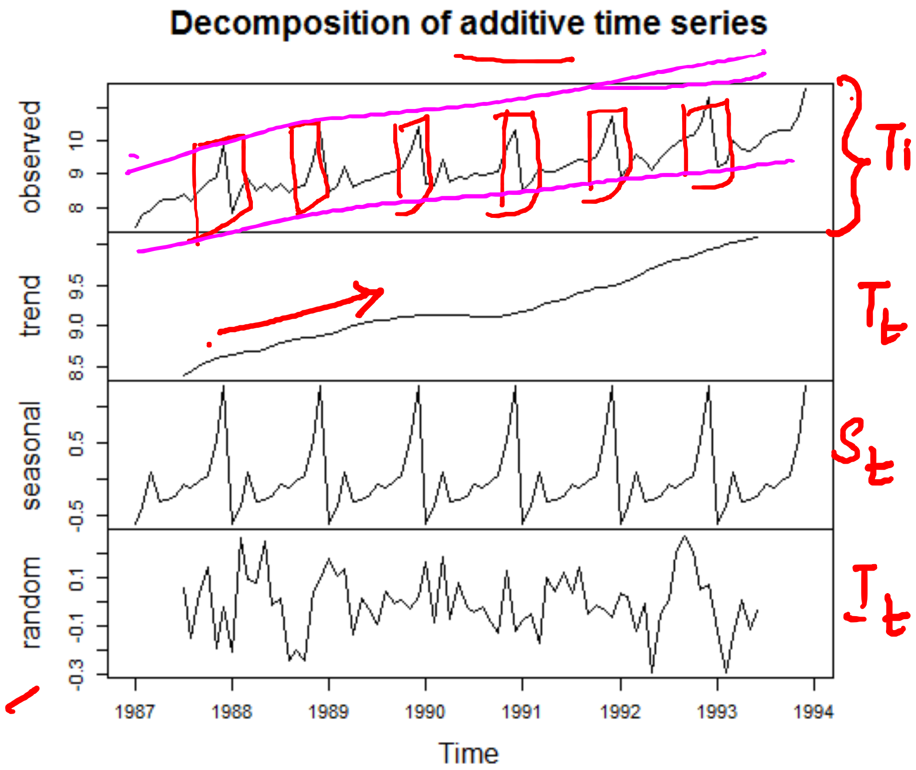


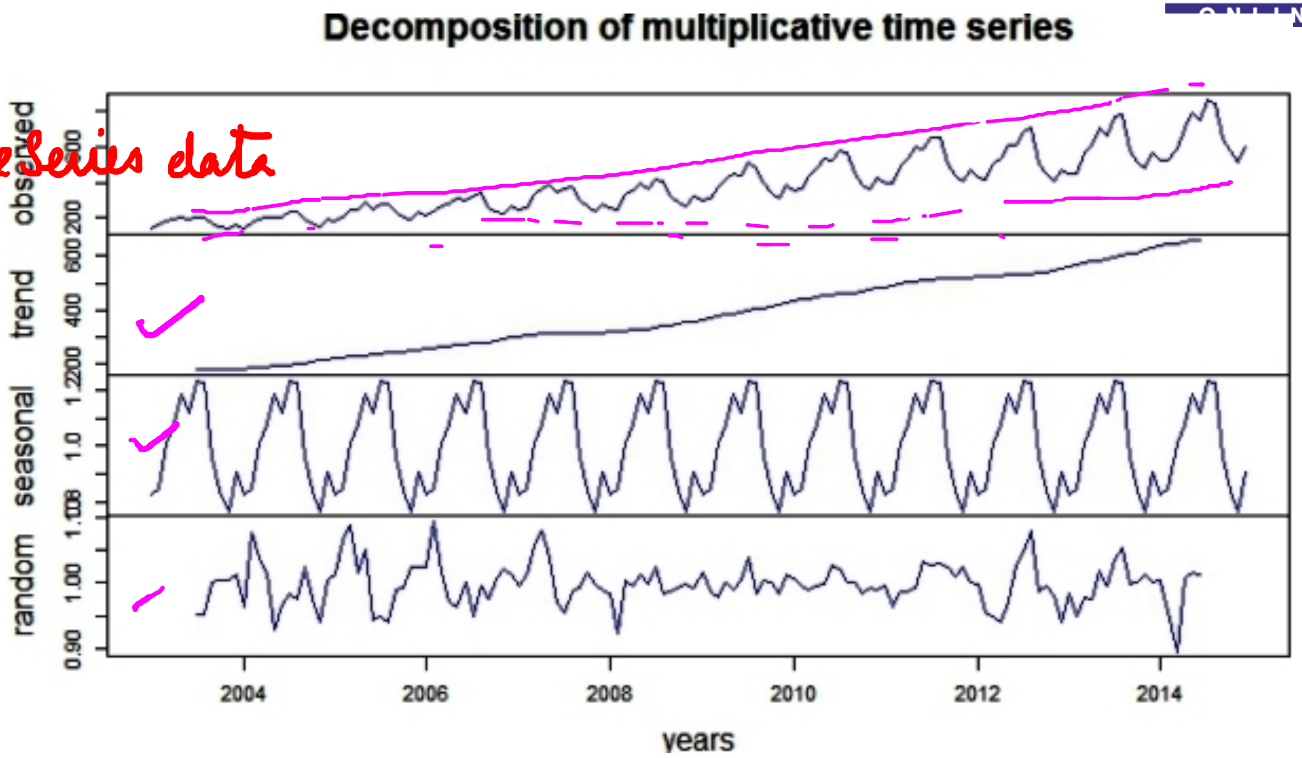
DATA ANALYTICS

Additive and Multiplicative Time Series Data

		Bonus		Bonus 10%
S1	70	+5	75	77
S2	80	+5	85	88
	40	+5	45	44
	90	+5	95	99
S100	55	+5	60	60

— Multi





$$Y_t = T_t + S_t + C_t + I_t$$

0

$$Y_t = T_t \times S_t \times C_t \times I_t$$

Additive and Multiplicative Time Series Revisited

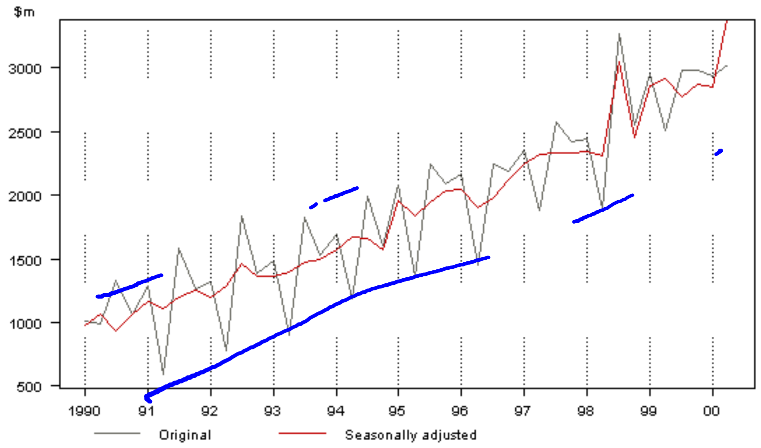
- The additive time-series model is given by
$$Y_t = T_t + S_t + C_t + I_t$$
- The additive models assume that the seasonal and cyclical components are independent of the trend component.
- Additive models are not very common since in many cases the seasonal component may not be independent of trend.
- The additive model is appropriate if the seasonal component remains constant about the level (or mean) and does not vary with the level of the series.

- The multiplicative time-series model is given by
$$Y_t = T_t \times S_t \times C_t \times I_t$$
- Multiplicative models are more common and are a better fit for many data sets.
- In many cases, we will use the form
$$Y_t = T_t \times S_t$$
- To estimate the cyclical component we will need a large data set.
- The multiplicative model is more appropriate, if seasonal variation is correlated with level (local mean).

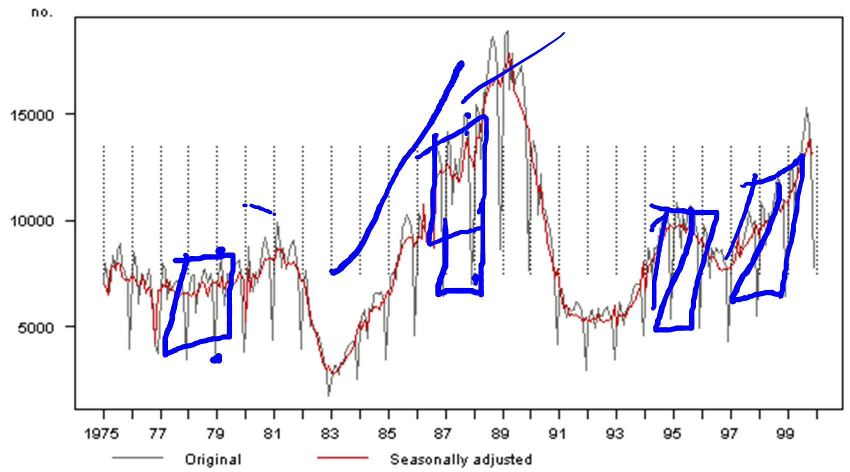
Additive and Multiplicative Time Series Revisited

- The additive time-series model is given by $Y_t = T_t + S_t + C_t + I_t$
- The multiplicative time-series model is given by $Y_t = T_t \times S_t \times C_t \times I_t$

General Government and Other Current Transfers to Other Sectors



Monthly NSW ANZ Job Advertisements



The underlying level of the series fluctuates but the magnitude of the seasonal spikes remain approximately stable

The trend has the same units as the original series, but the seasonal and irregular components are unitless factors, distributed around 1

Decomposition of Time Series Data - Additive

- Decomposition models are typically additive or multiplicative, but can also take other forms such as pseudo-additive.

Additive Decomposition

In some time series, the amplitude of both the seasonal and irregular variations do not change as the level of the trend rises or falls. In such cases, an additive model is appropriate.

In the additive model, the observed time series (O_t) is considered to be the sum of three independent components: the seasonal S_t, the trend T_t and the irregular I_t.

Observed series = Trend + Seasonal + Irregular

$$O_t = T_t + S_t + I_t$$

$\hat{S}_t$ $\hat{T}_t$ $\hat{O} - \hat{T} = S + I$

Seasonally adjusted series = Observed-Seasonal

$$SA_t = O_t - \hat{S}_t$$

$$= T_t + I_t$$

COMPONENTS OF TIME-SERIES DATA contd.

- **Multiplicative Decomposition**

  In many time series, the amplitude of both the seasonal and irregular variations increase as the level of the trend rises. In this situation, a multiplicative model is usually appropriate.

  In the multiplicative model, the original time series is expressed as the product of trend, seasonal and irregular components.

- Observed series = Trend x Seasonal x Irregular

$$O_t = T_t \times S_t \times I_t$$

Seasonally Adjusted series = Observed ÷ Seasonal

= Trend × Irregular

$$SA_t = \frac{O_t}{\hat{S}_t} = T_t \times I_t$$

COMPONENTS OF TIME-SERIES DATA contd.

Pseudo-Additive Decomposition:

- The multiplicative model cannot be used when the original time series contains very small or zero values
- This is because it is not possible to divide a number by zero
- In these cases, a pseudo additive model combining the elements of both the additive and multiplicative models is used
- This model assumes that seasonal and irregular variations are both dependent on the level of the trend but independent of each other.

The original data can be expressed in the following form:

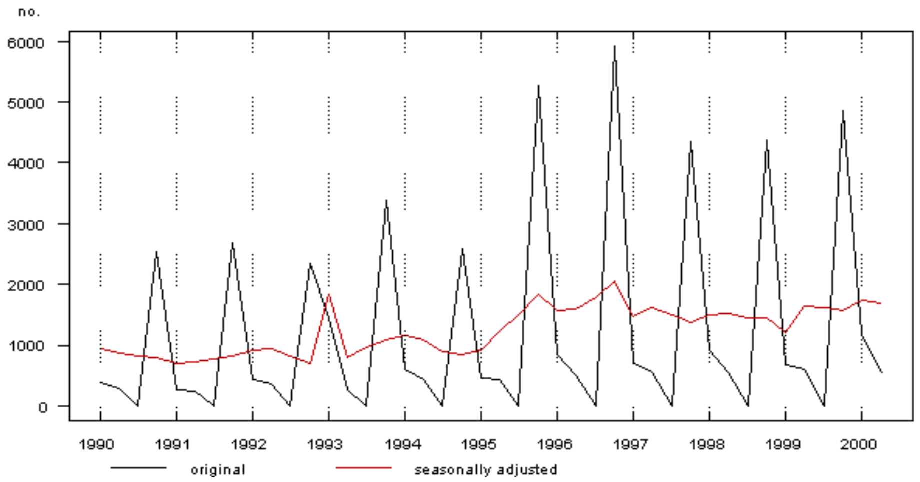
$$O_t = T_t + T_t \times (S_t - 1) + T_t \times (I_t - 1)$$
$$= T_t \times (S_t + I_t - 1)$$

- Both the seasonal factor S_t and the irregular factor I_t centered around one
- We need to subtract one from S_t and I_t to ensure that the terms $T_t \times (S_t - 1)$ and $T_t \times (I_t - 1)$ are centered around zero.
- These terms can be interpreted as the additive seasonal and additive irregular components respectively; the original data O_t will be centered around the trend values T_t .

COMPONENTS OF TIME-SERIES DATA contd.

- An example of series that requires a pseudo-additive decomposition model is shown below.
- This model is used as cereal crops are only produced during certain months, with crop production being virtually zero for one quarter each year.

Quarterly Gross Value for the Production of Cereal Crops



This model is used as cereal crops are only produced during certain months, with crop production being virtually zero for one quarter each year.

References

Text Book:

“Business Analytics, The Science of Data-Driven Making”, U. Dinesh Kumar, Wiley 2017 (Chapter 13.1-13.2)

Additional reference and image courtesy:

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

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

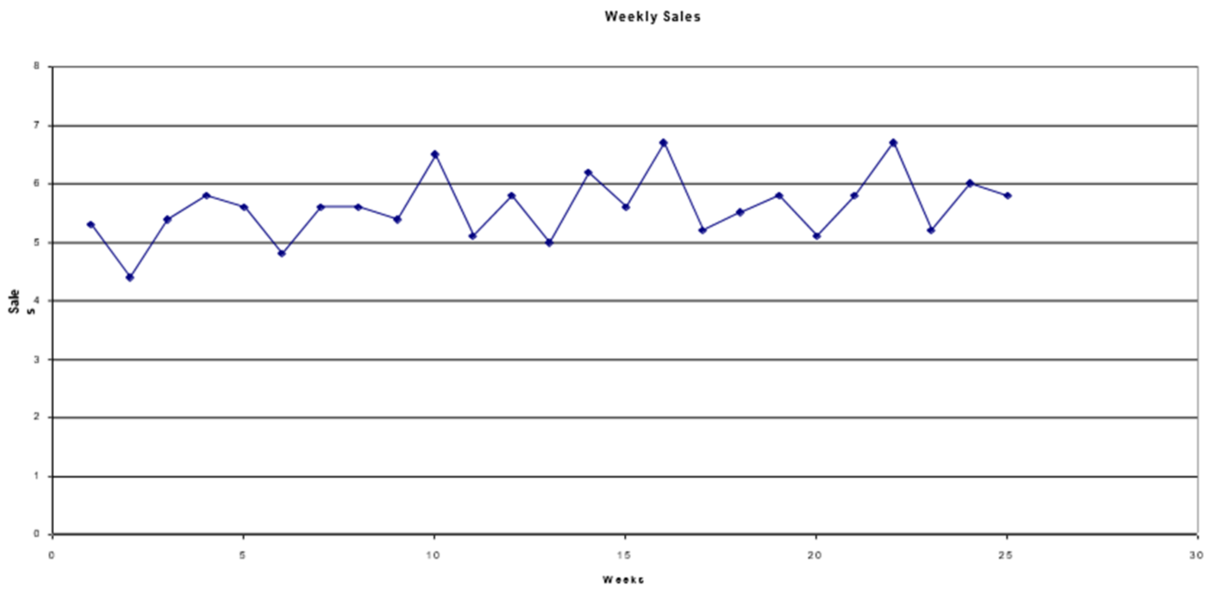
- A small k is most desirable when there are sudden shifts in the level of series

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

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

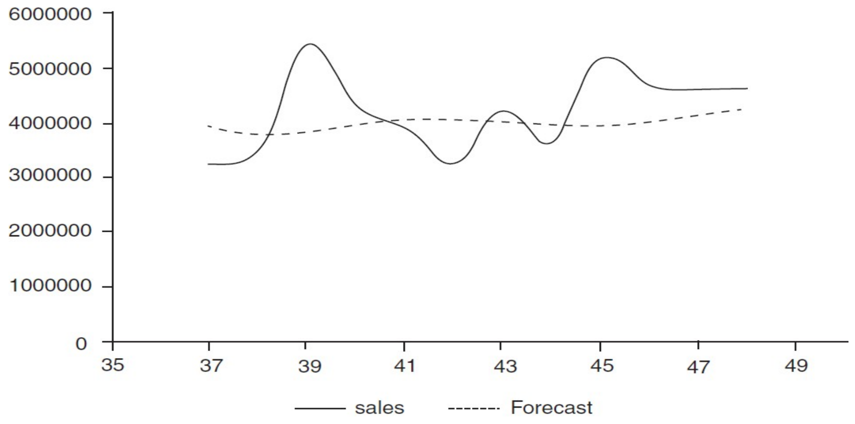


FIGURE 13.2 Plot of actual sales forecasted sales using moving average.

- The forecast error is
  - $e_{24} = y_{24} - y_{24} = 6 - 5.9 = 0.1$

Weighted Average

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

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

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

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?

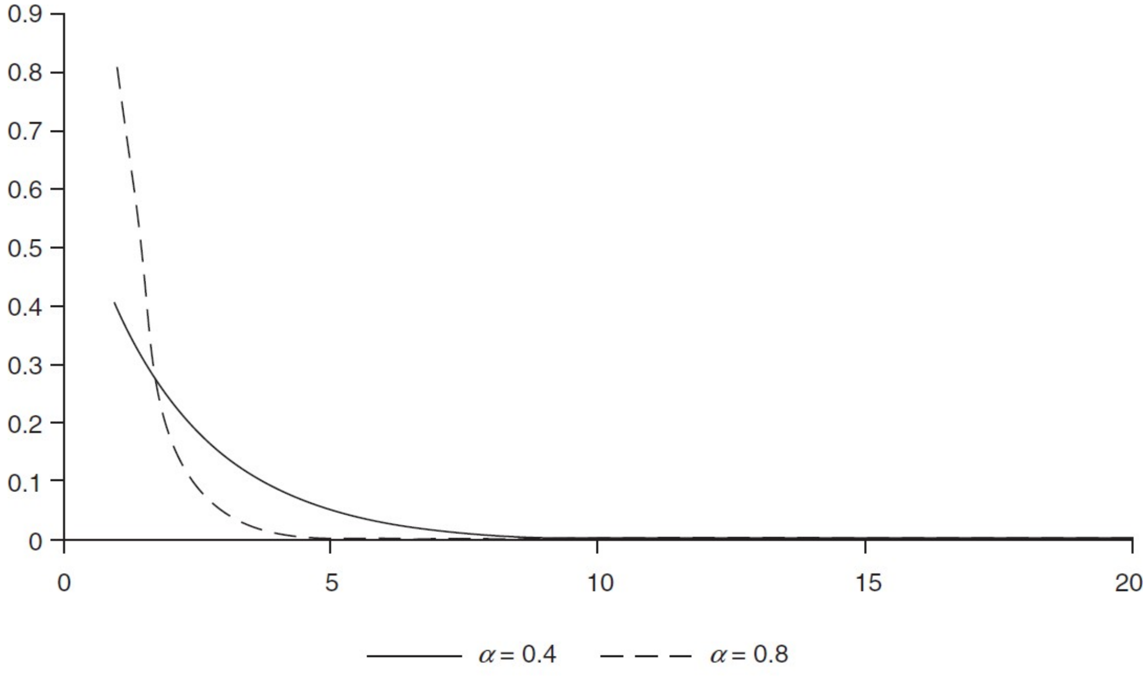


FIGURE 13.3 Exponential decay of weights to older observations.

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

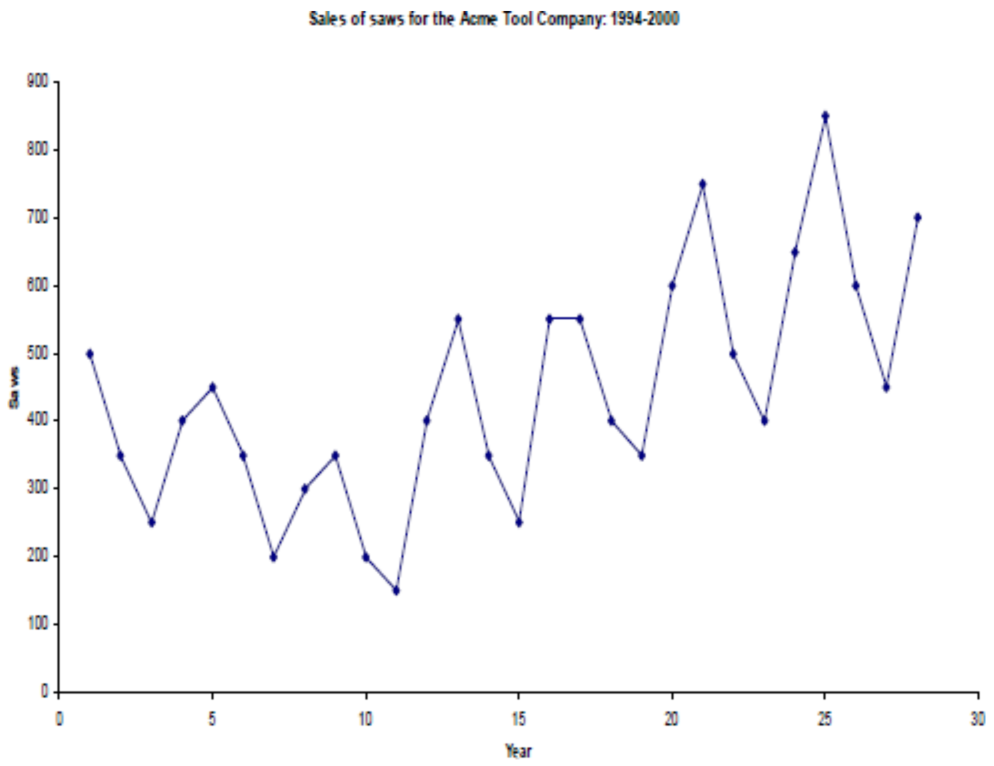
$$F_{t+1} = L_t + T_t \tag{13.14}$$

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

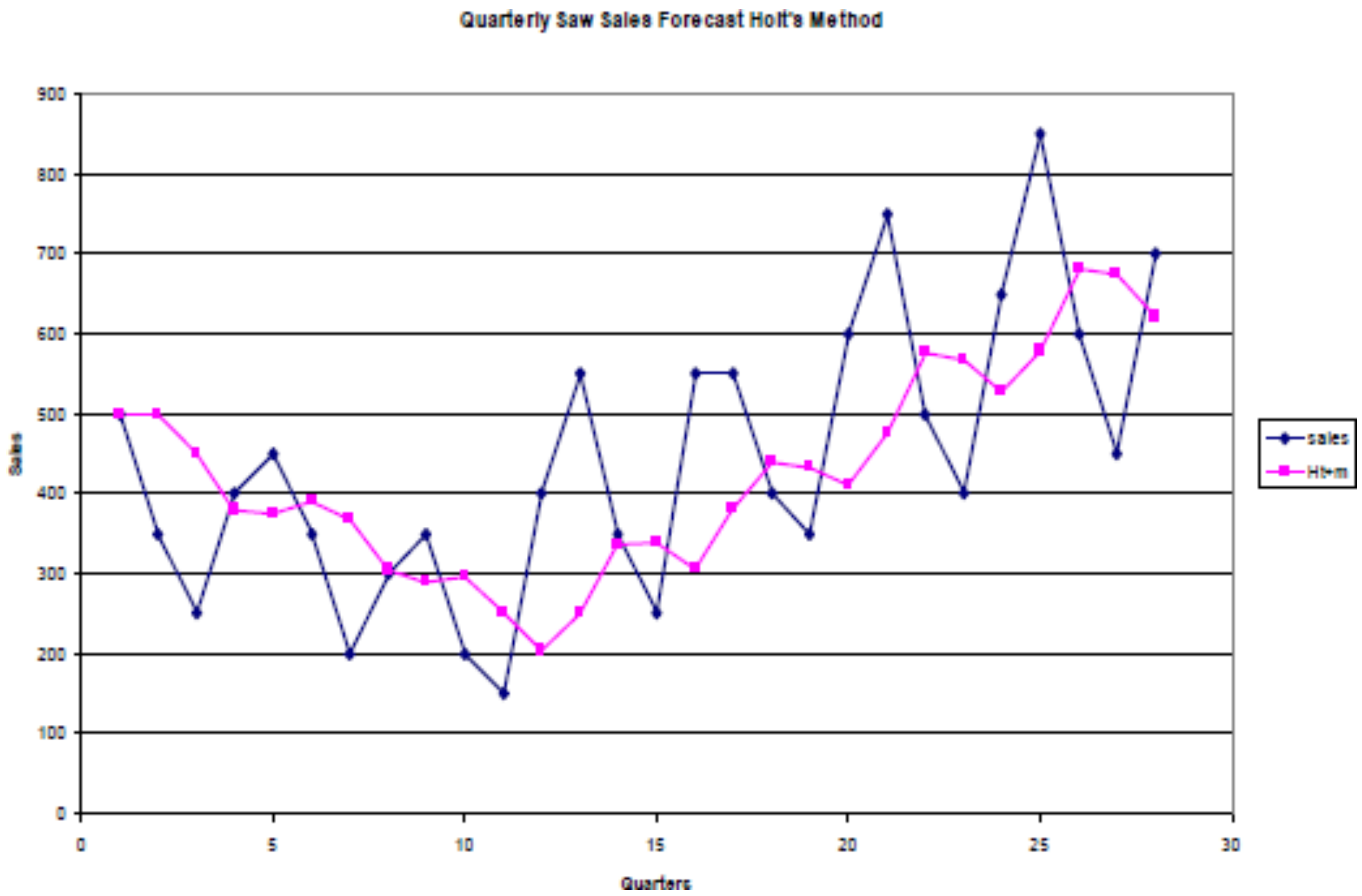
where L_t is the level which represents the smoothed value up to and including the last data, T_t is the slope of the line or the rate of increase or decrease at period t, n is the number of time periods into the future.

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

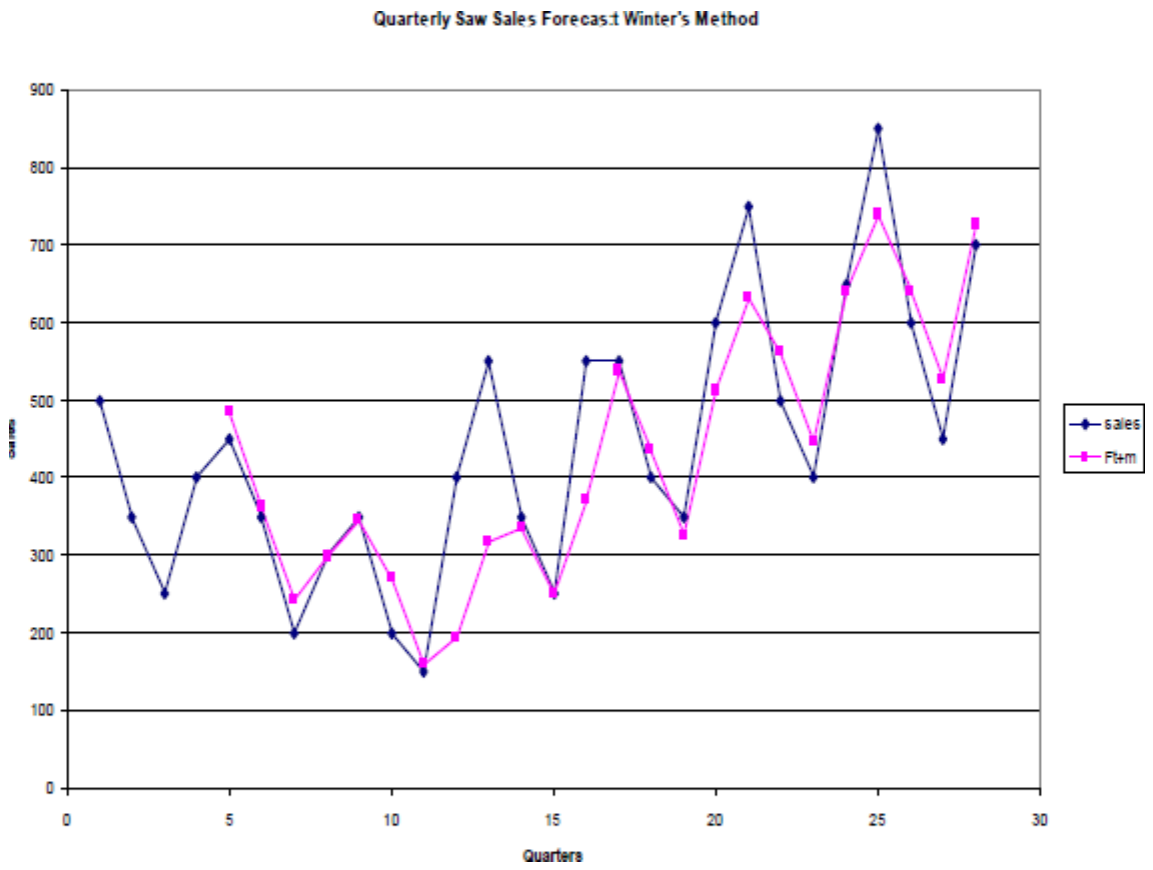
$$L_t = Y_t$$

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

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

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

THANK YOU

Dr. Mamatha H R
Professor, Department of Computer Science
mamathahr@pes.edu

+91 80 2672 1983 Extn 834

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