Time series : It is a finite sequence of numbers represented in chronological correspondence.

Eg. Time (in years) (x): 2010 2011 2012 2013 2014 2015 2016

Export ( ‘000 rs.) (y): 250 240 280 285 270 290 250

* X is an independent variable – time.
* Y is the actual export amount for that year. Y is mutually independent within and **of the time x also**. i.e., Y is mutually independent, and Y is independent of X.

// the Y values are not pure numbers, they are mixture of T,S,C,I variations. From Y, wehave to find these variations.

time series is a sequence of numbers which are in correspondance to equal interval of time.

each value of time -> equal interval.

time series values are not dependent on time, it is in correspondance to time.

there are many factors which individually or collectively affect the given variable and what we finally observe is a **joint effect**. There are many important factors which jointly, fully or partly impress the variable in problem. We study those factors and effects.

Importance-> to know the pattern.trend in which the data flows/appears.

for forcasting

1. To study the nature of variable in accordance with time.
2. To identify and segregate the different components in the given data.
3. To make forecast.

We know that the main purpose of study of the time series is to identify all the different components which to some or more extent are responsible to generate the numerical data which are in accordance with time and also the important feature is to make, so far as possible, forecast which finally helps understand the natural trend/ flow of the data

Components:

1. Long term components

* Secular Trend/Trend->
  + Basic tendency inherent in the variable
  + Steady Movement of data over long period of time, shows general direction.
  + Those values we obtain after removal of other variations.
  + Inherent in the given chronological data are the driving forces which build up the sequence; they do mix-up with the other forces giving rise to the time series data. ( Actual set of observations)
  + Price of gold, population, and national income.. Etc; observations in this directions are the real facts but they ae the resultant of all the four factors– Long term and short term.
  + Study is important as it can (1) Identify general propensity (2) Help explain the major factors comprising the time series. (3) Forecasting over a reasonable period of time using different models (4) Extrapolation
* Cyclical Component ->Cyclical variations are the forces which exhibit their effects on data items/ actual price/ rate **over a period of time greater** than one year. It may be a cycle of more than one year but generally less than or equal to four years.
  + You cannot predict when and why they happen.
  + These changes are very slowly dependent of political situation of some major countries, steady climatic changes, impacts of modernization, and implementation of scientific researches and development in the society.
  + This cyclical changes are not at a regular interval. They do appear but between two occurrences the time slot is not constant.

1. Short term components

* Seasonal Component ->These variation are, as name suggests, seasonal. They are to appear with a span of one year. Also its impact on the given data is only for a particular span of the given year.

Eg price of gold, other items increase during festive season, and then drop during end and obtain a steady state

Both C and S are temporary changes but in some cases they attain a new height or fall which remain slowly attaining its trend.

* Erratic / irregular ->Such changes are the not well defined in advance. The time and intensity of their occurrence are never pre- known
  + Unpredicatble risks, not planned before, rare.
  + Situations like devastating flood, earth-quake, cyclone, and war between two or more countries etc. are the major causes and these are the forces which are responsible for extreme recession or hike in price of any commodities under study.
  + Once these forces are reflected in the data, it can be understood very distinctly. It is very important to understand to identify all such variations from the given graph.

All these may or may not be present fully or partially. Combined effect of all these factors is the given set of time correspondent values of some variable; which we call time series.

Time Series Models:

We have, on broad classification, three types of models to express a given situation more accurately.

1. Additive Model: In this case the actual observation has been expressed as y = T + S + C + I. We understand that S, C, and I are absolute deviations about the trend. All these components are independent.
2. Multiplicative Model: In this case the actual data is expressed as the product of these factors. Y = T. S.C. I where S, C, and I are relative variations.( rate, %, Index numbers) 625 = 600 x 1.05x 0.99 x 1.002

Also, log y = log ( Tx Cx S x I )

1. Mixed model are according to the flow of data.Y = T + S + C x I

**Methods of Finding Trend**

1. Graphic Method

In this method the key point is assumption of linearity.

We assume simple approach. We are given the graphical presentation of the data. We take two different points on the graph. This becomes subjective; but we do consider such two points that the line joining those two point remains very close to all the points showing the data points. This can be done as follows.

1. Statistical Method
   * selected Points

The central idea is the inserting arithmetic mean between two terms.

Time: t1 t2 t3 t4 t5 t6 t7 t8

Yield : y1 y2 y3 y4 y5 y6 y7 y8

Say we select (t2, y2) and (t7, y7) as fixed points.

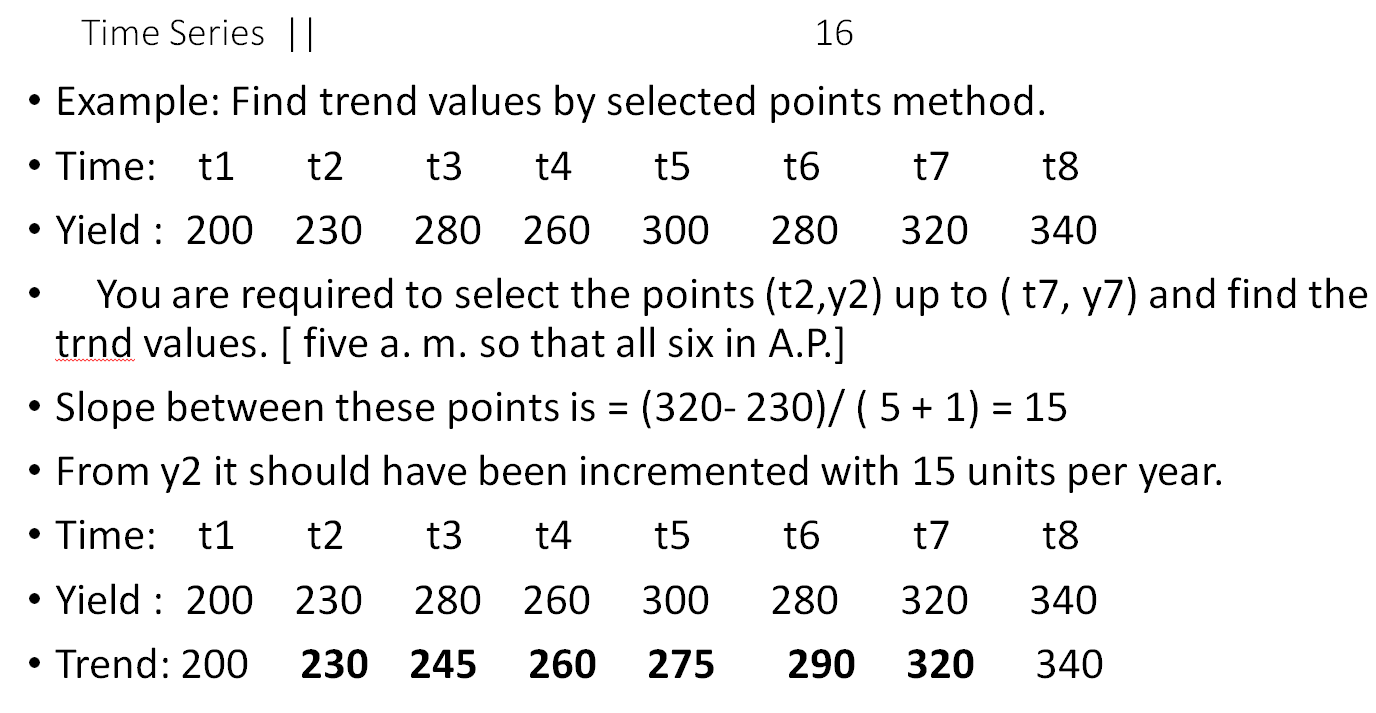
Y2 ,……., ………, ………., ………, y7. In between there are four terms to beinserted that all the terms from y2 to y7 takes up the format of an **A.P.**

**Let d**  = ( y7 – y2)/ (n+1) then we have assumption that there is linearity.

Time: t1 t2 t3 t4 t5 t6 t7 t8

Predicted Yield : y1 y2=a a+d a+2d a +3d a +4d y7 =a +5d y8

// d = (l-a)/(n+1) //n+1 = total number of terms including a,l



* + **Semi Averages Method: Better than above**

The central idea is same; insert in arithmetic means between two given numbers. There are two types 1) when there are odd number of observation and 2) there are even number of observations.

* Time : 1 2 3 4 5 6 7 **8** 9 10 11 12 13 14 15
* Yield: 10 11 13 8 14 12 9 **14** 13 10 12 16 14 16 17
* Mean: ………… ..….11………………….. ……………………..14 ……………………..
* Trend **11 12.5 14**
* **With incremental factor = ( 14 – 11)/8 = 3/8 [ 11** and 14 remain fixed.] seven terms so that all eight terms in A. P.

// for odd, same as below, either omit middle term, or include it into both sides

* **Semi – Average Method**----- There are even number of observations in the series. Study the table and concentrate on the flow. N = No. of years =10
* Time: 01 02 03 04 05 06 07 08 09 10
* Price: 16 17 18 19 20  **20 22 23 19 21**
* Average: ………………90/5 =18………. …………….105/5= 21…………
* . we have S1 = 90 . And S2 = 105 We assume linear form : y = a + bx
* a = intercept on y axis and b is the slope A of the line. **a = 2 (S1)/ n**
* a = 2 x 90 / 10 = 18 and **b = 4 (S2- S1) / n^2**= 4 ( 105 - 90) / 10^2 = 0.6

// n is total numbe of observations

* We have the linear form y = a + bx. ; where x is an independent variable
* Taking up different values. A + 18 and b = 0.6; y = 18 + 0.6 x
* X : --2 --1 0 1 2 3 4 5 6
* Y : 16.8 17.4 18 18.6 19.2 19.8 20.4 21 21.6 ----Trend
* Given : 16 17 18 19 20 20 22 23 19

For even -> s1 = sum of first n/2 values, s2 next n/2

1. Algebraic Method
   * **Natural method (Algebraic Method)**; It is with an assumption of linearity, or assuming flow of data with quadratic rule, or some other rule– like y = a. b^x

1 y = a + bx 2 y = a + bx + c.x^2 3 y = a. b^x

Where a , b, and c are constants and x is an  **independent variable**

// This is regression analysis

1. Moving Average

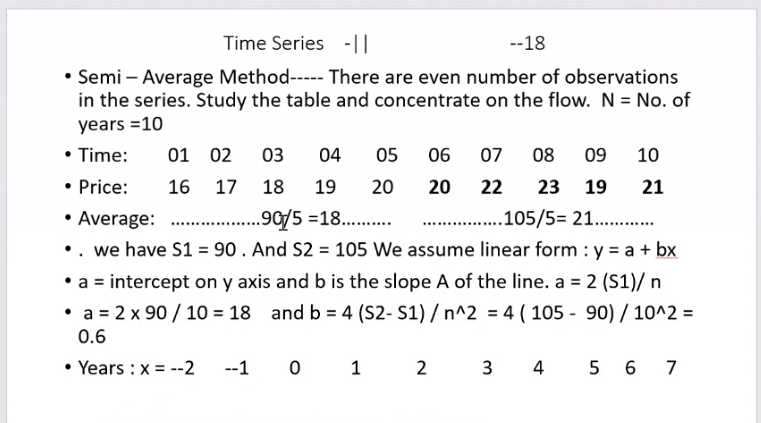
27/7

methods -> graphical method -> not perfect

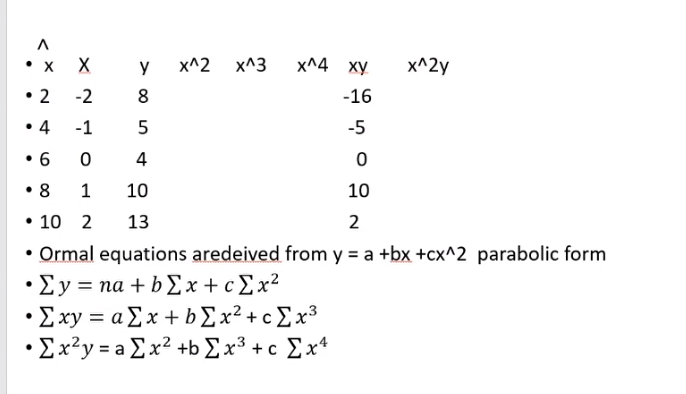
algebric method -> regression analysis

semi-averages (moving averages) -> no. of time needs to be odd

time -> can be adjusted as -3,-2,-1,0,1,... but you cannot adjust Y



* Fit linear regression
* Fit quadratic: Y = a+bx+cx^2

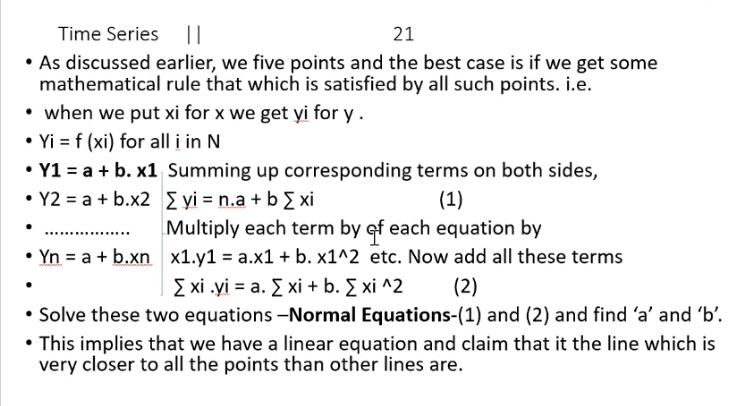


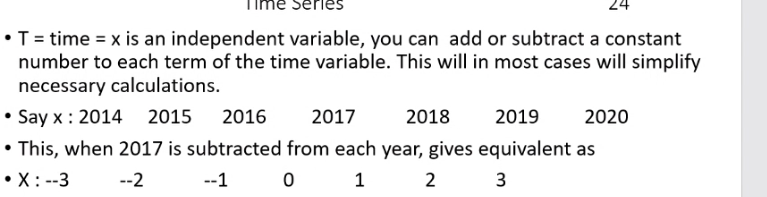
Taking centred X -> makes trend fitting technique easier

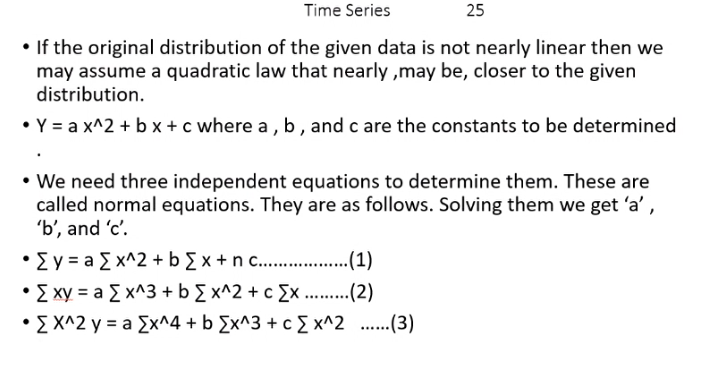
Given a question:

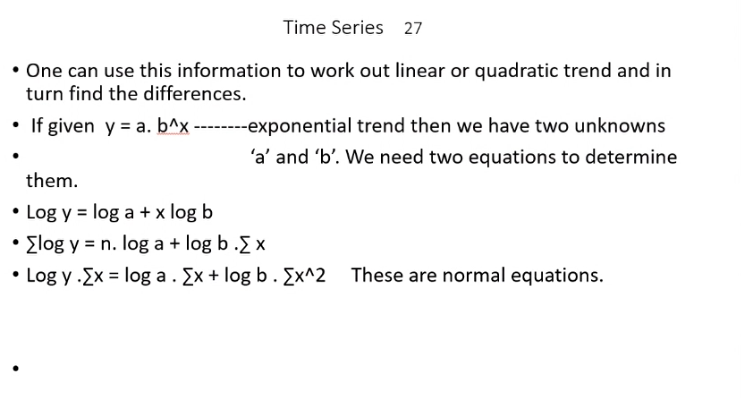
* Draw a scatter diagram
* If points in line -> linear assumption
* If quadratic -> quadratic assumptions
* Normal equations -> equations to find out the coefficients

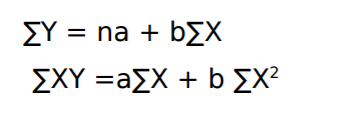
Once a,b,c are identified, you can forecast.

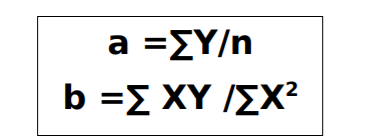


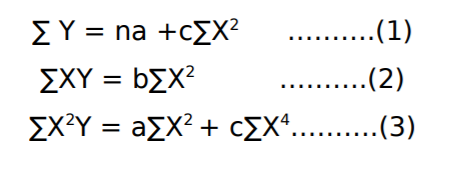


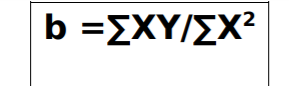


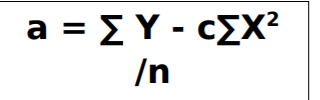


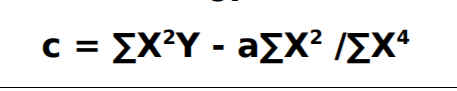




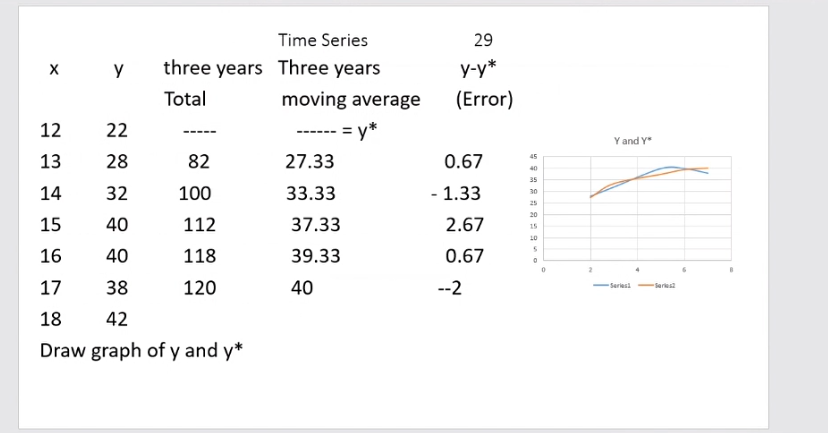


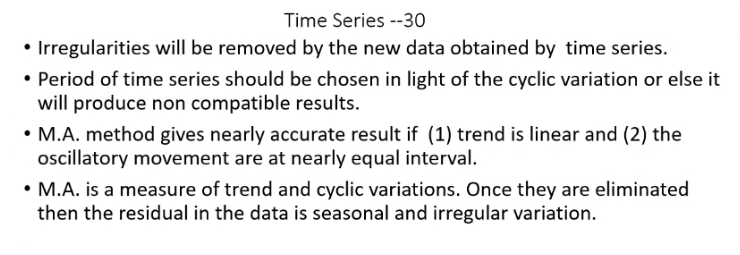






Focus -> find trend -> in general, over long period of time, what is the flow of data

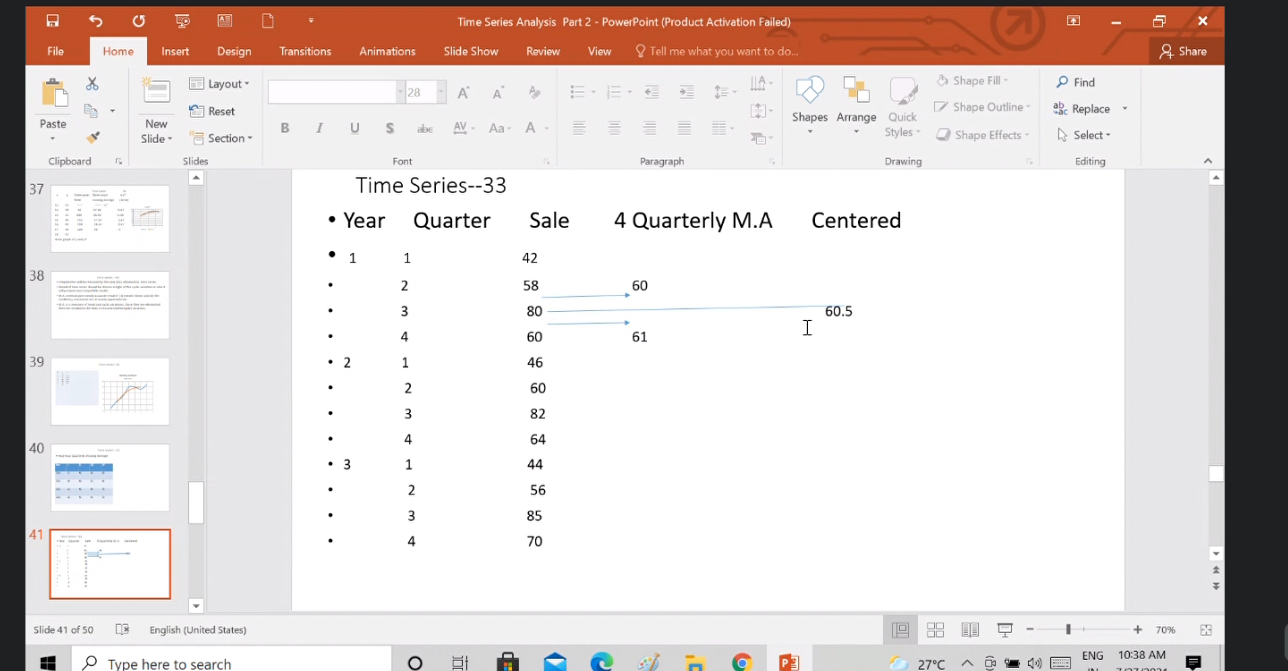




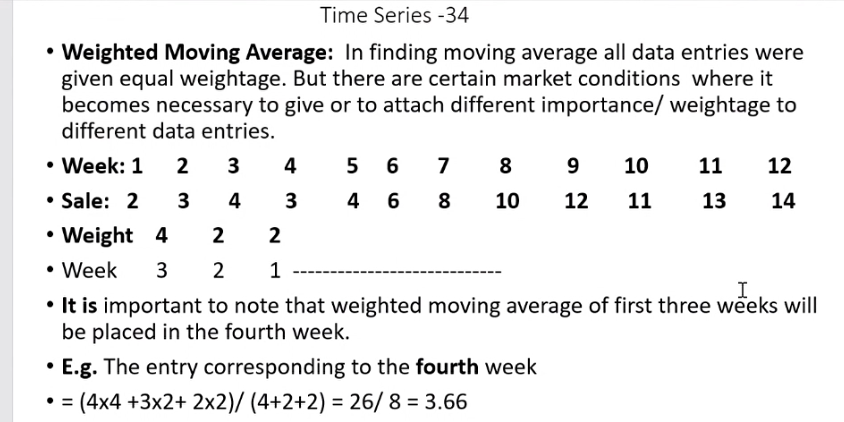
* Usually time period is 3-4. Because using higher periods -> if some time, seasonal or cyclic variation is high, then it’s effect on average will be higher.
* Good only when trend is linear, not good in paralbolic trends.

4 Quaterly moving average:

* Average of 4’s -> write between second and third..
* Leave on above, ladd on bellow -> write b/w 3,4th
* Centre it further -> write across 3rd quarter…
* Continue. For each quarter

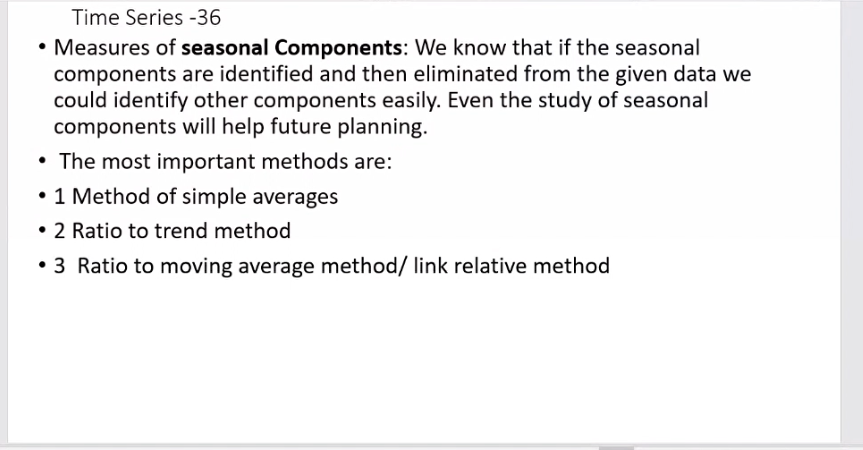


Weighted moving average -> better than above:

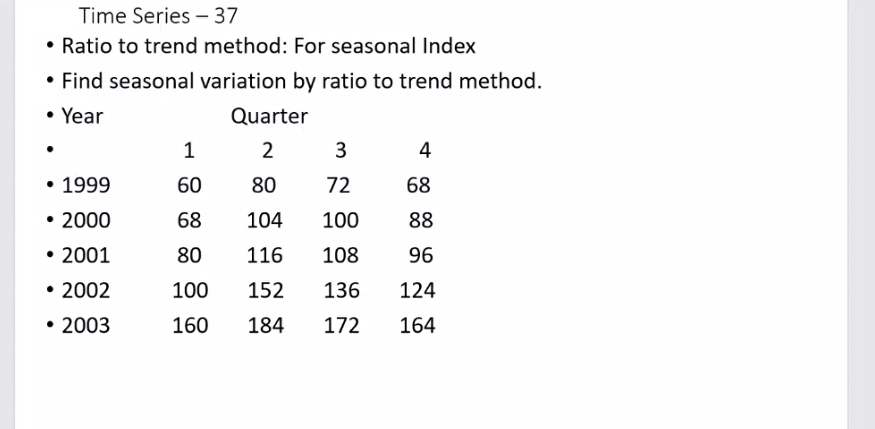


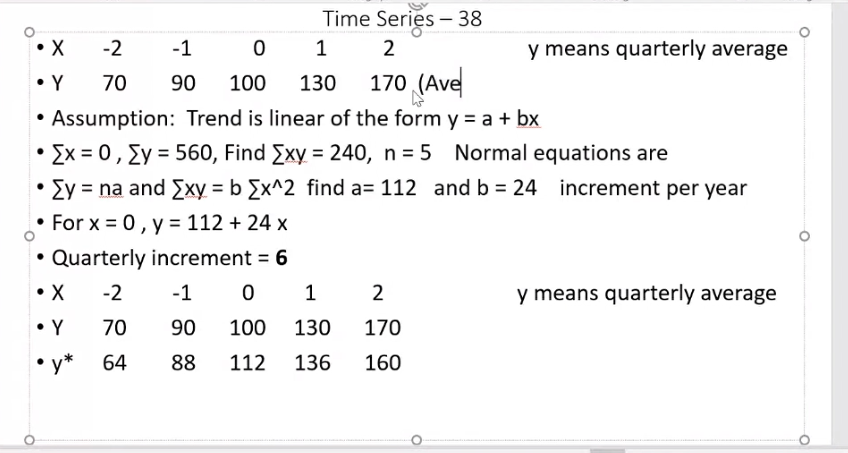
(2,2,4) -> 4 for the just last one.

sum(Y\*weight)/sum(weight)



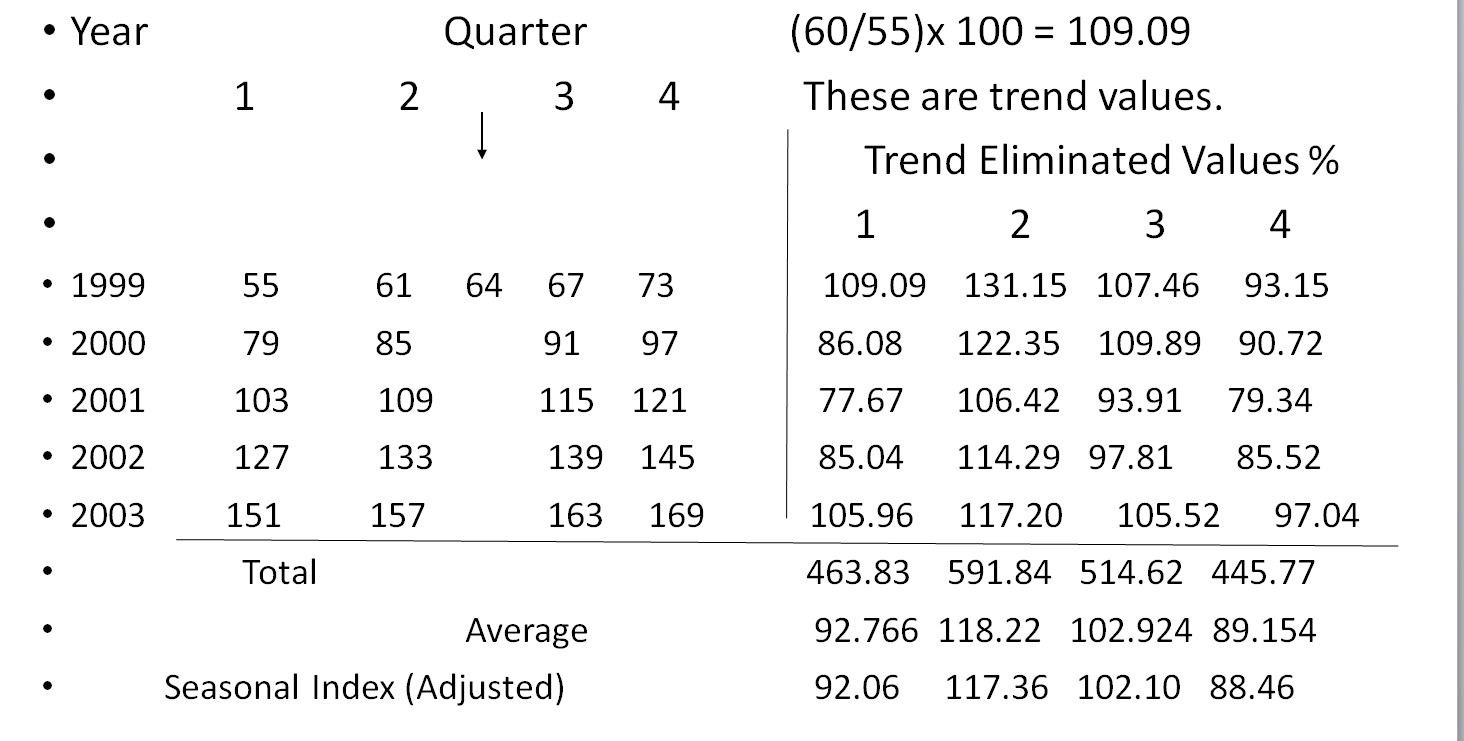
2. -> for seasonal index.





Y is average of all quarters of the year

quarterly increment -> b/4



?? Here, 64 is the value of Y\* in year 1999 -> predicted value for 1st july

Finding Trend Eliminated % values:

109.09 -> 60/55\*100 -> original value/obtained value X 100

Average = avg of column/ quarter

Adjusted seasonal -> find the average of average = grand average = 100.78

Then correction factor, CF = 100/grand average = .992

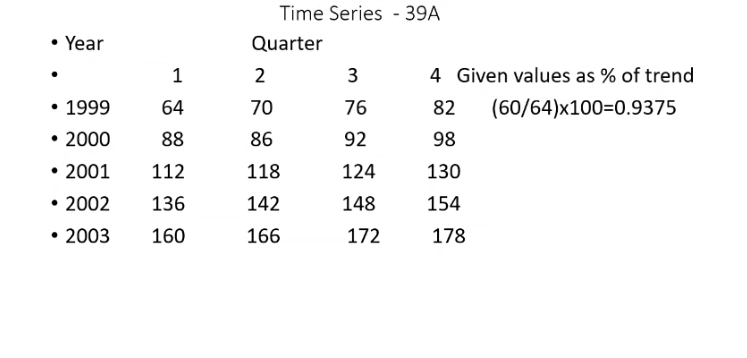
Seasonal index = CF\*Average

For verification -> add seasonal indices, = 400. Must be 400. If CF not applied, answer would be graeater than 400.

These seasonal indices/factors have been derived using multiplicative model of the time series. i.e, assuming original data is in multiplicative model

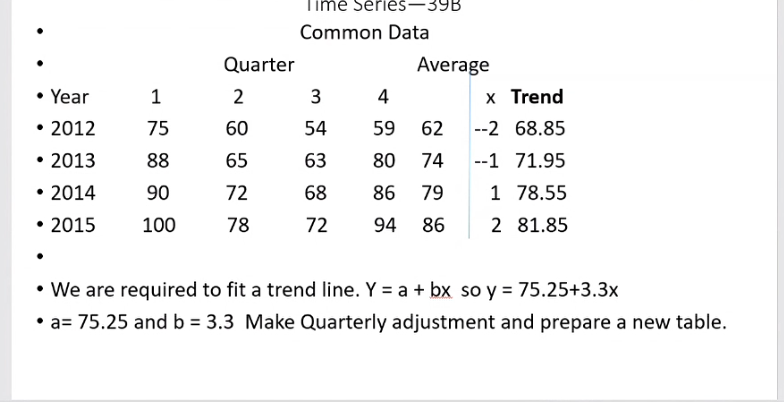
/////Ignoree

// predicted quarters



//////////////////

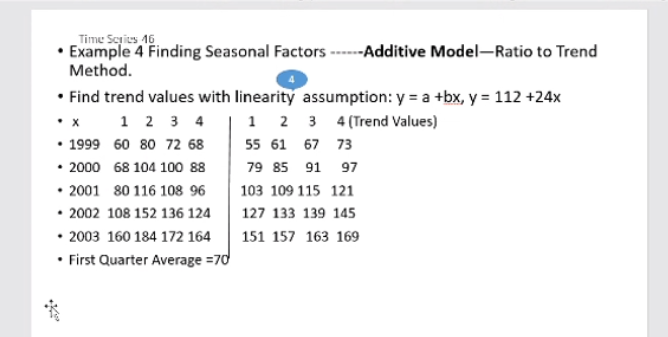
// another ques

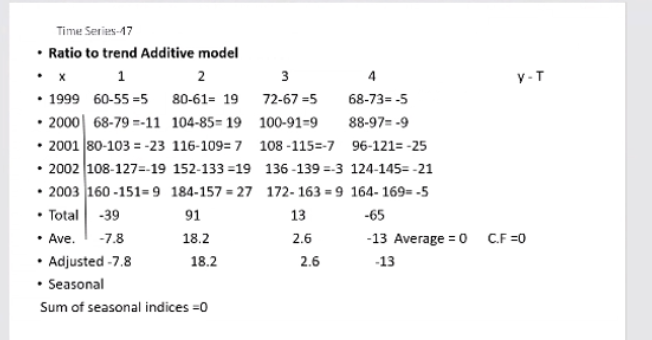




Additive model method:

Do everything till before taking ratios



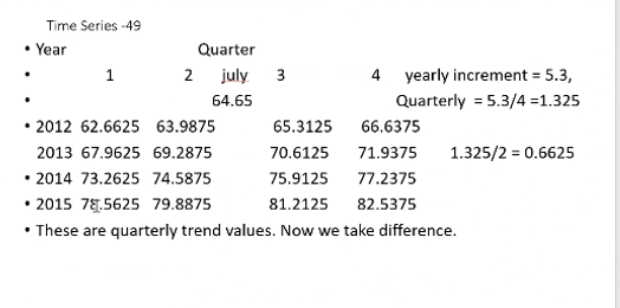


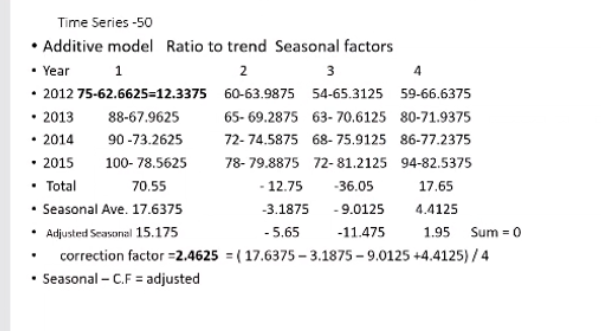
Find total, average, correction factor, and get seasonal indices = avg – CF

CF = average of average values

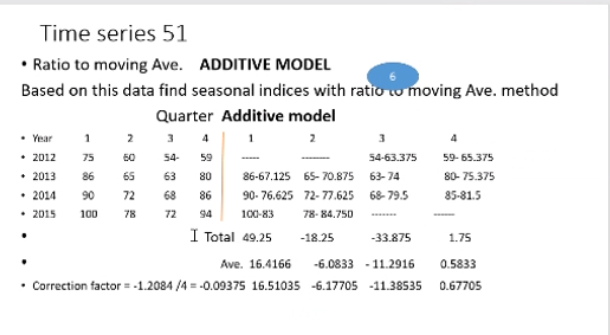
Sum of seasonal indices = 0



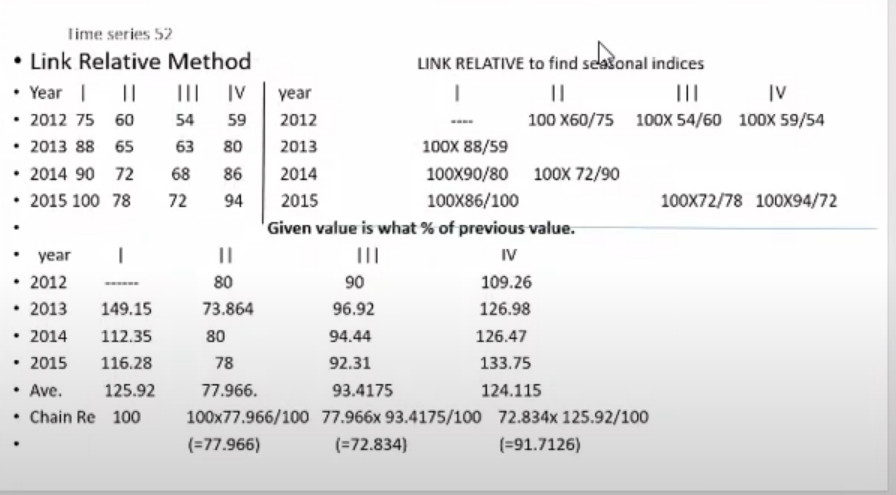




Ratio to moving average:

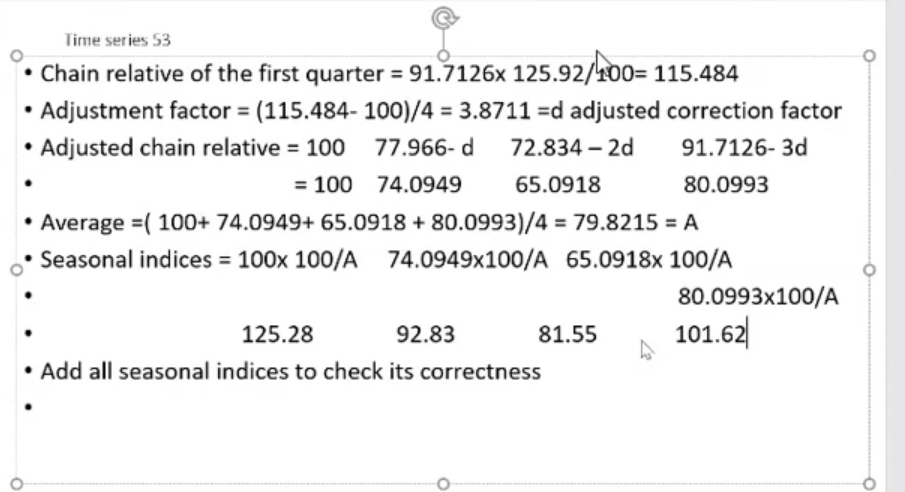


Link relative method



Find percentages of today than yesterday ( right above table shows calculation, table below its result)

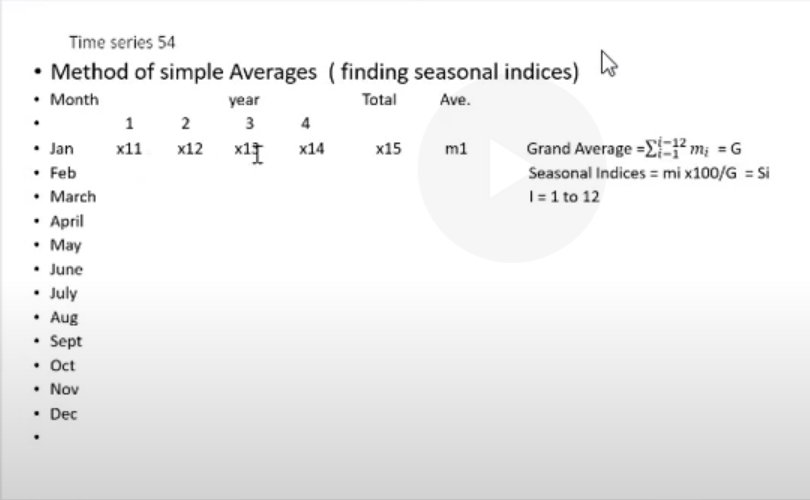
I.e. given value is what percentage of previous values

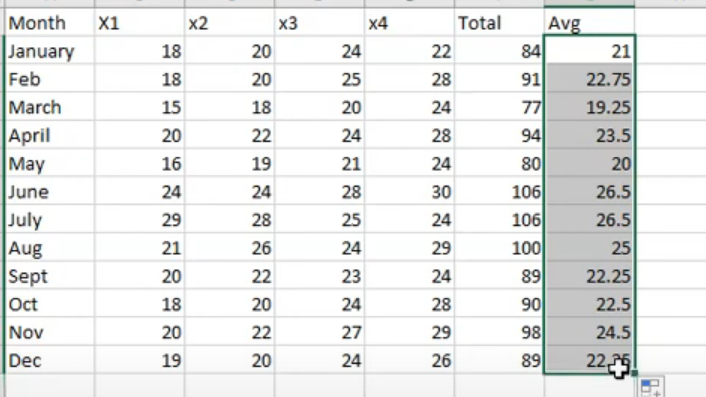


Sum of seasonal = 400

1. Make the table..
2. Find the averages (/3 for first, /4 for others)
3. Take chain relative -> Multipliy current avg/100 with last chain relative
4. Find chain relative to first quarter again
5. Find adjustement factor
6. And ahjusted chain relative
7. Take avg of above factors. -> A
8. 100/A = Correction factor..
9. Seasonal indices = CF \* adj chain relative

Method of seasonal averages:





------ ------------ ----------------------- ----------------------------------------

* **Forecast**: Any ongoing system needs planning for the future activities.
* Planning for Production, revenue to be generated, procurement of raw material.
* The total amount of resources we need for immediate future activities( short term planning) and activities spread over one year( long term planning) we need to have projection(forecast).
* Planning can be done if there is a forecast for demand. Production depends on forecast. In addition to this there are many factors which involve forecast.
* **There are** two types of forecasts.

1. Qualitative Forecast and
2. 2 Quantitative Forecast

Qualitative Forecast:

* There are three types of qualitative forecasts.
* 1 Delphi Method: This method depends on opinions and experiences of an expert panel in the particular area. This is often called jury of expert opinion. Forecasts done by this panel are known as executive forecasts. Each panel member engaged in the same type of business and being highly experienced makes forecast depending on factual data available. All the different opinions collected together and we arrive at the most common forecasts. No doubt, there is subjectivity associated with this but in most of the cases panel decision is better and reliable.
* 2 Sales force Composite method: As the name indicates, the forecasts purely depend on the marketing intuation and also marketing situation of sales and future demands on the team members of the sales executive engaged in the sales activities.
* 3 **Consumer Panel Survey Method:** This may prove the best one as it is the end users who are responsible to continue with the product over a long time period We need to make survey being in contact with the end users. Their suggestions are highly welcome and if required may be considered. If the product needs feasible changes in the size, packing, warranty, and availability of spare parts in the free market then such suggestions are given priorities.
* Al these features taken in to account help make forecasts over a long period of time. Such variations in the sale calls for cyclic variations.
* **Quantitative Forecasts:** There are two types forecasts in this category.
* 1 Regression Analysis: This is called the natural method as it totally gives importance to the available data. It is known that the data involves many factors like trend, seasonal , cyclical and irregular effects, judgement or forecast depending upon this algebraic calculations may not be perfect. In addition to this, it depends on the flow of data distribution. As we know that, there may be linearity, parabolic nature , and exponential distribution of the data and hence before making any forecast it is necessary to study the flow of the data. Once it is identified and studied this method gives better results for forecasts.
* 2 **Exponential Smoothing Method:** This method is used widely in the business decisions. The main importance of this method is that it fundamentally depends upon the actual observation and the prediction for that period in the past.

Exponential smoothing

------- ----------------- ------------------- ---------------

* A time series is a sequence of real values/ numbers that flow according to time but not dependent on time. The time series data , as we know, depends on four main components like trend, seasonal variation, cyclic variation, and irregular variations ( if any).
* If we plot the graph then we may note many important points. According to the given data, **over a period of time** may be such that
* 1 Mean value changes
* 2 Variance changes and
* 3 seasonal variations do appear.
* Such a series makes it impossible for the forecast purpose.
* We have to see that such factors are minimized or totally eliminated.

Stationarity:

* You can forecast only when data is stationary
* Stationarity of Time- Series: We define time series as **stationary** if a shift in time doesn’t cause a change in the shape of the distribution.
* The basic of distribution we are talking about is **mean, variance and covariance. For** stationarity absolute values of characteristic equation must be less than 1. It should remain within the unit circle.
* **First-order stationarity**implies that the series have means that never change with time. Any other moment like variance can change.
* **Second-order/ weak stationarity**means that the series has a constant mean, variance, and auto covariance that doesn’t change with time. Other moments are free to change over time.
* A stationary time series is **one whose properties do not depend on the time at which the series is observed**. 14. Thus, time series with trends, or with seasonality, are not stationary — the trend and seasonality will affect the value of the time series at different times.
* A common assumption in any time series techniques is that the data are stationary.
* A stationary process has the property that the mean, variance and autocorrelation structure do not change over time. Stationarity can be defined in precise mathematical terms, but for our purpose we mean a flat looking series, without trend, constant variance over time, a constant autocorrelation structure over time and no periodic fluctuation

Tests for stationarity:

* **Unit root tests**
* The Dickey-Fuller Test. The Dickey-Fuller test was the first statistical test developed to test the null hypothesis that a unit root is present in an autoregressive model of a given time series, and that the process is thus not stationary. ...
* The KPSS Test. ...
* The Zivot and Andrews Test. ...
* Variance Ratio Test.

--- ------- --------------------------------- ----------------------------

AR Models:

The models in which the current values are most likely to depend on past values

* Auto regressive models.
* There are two types of models. These models are planned according to given data and its type. To be more clear on this, these models are designed for forecasting purpose. It happens in many cases that two models in coordination of each other can operate to give better results.
* A.R. models are the ones in which the current values are most likely understood to depend on past values.
* These past values of the observation may be , from last ( previous) year to say ‘p’ years. It is quite obvious that the value of ‘p’ my be one, two , or more years. It is traditional to consider this ‘p’ so long as it gives more and more accurate predictions.

LHS = actual observation

RHS first two/3 terms is the prediction of AR(1)/AR(2) and e is error = LHS-RHS

## mathematical parts

MA Model:

Fudamental difference: for MA, Today’s predicition Depends on past errors, for AR, today’s prediction depends on past actual values.