Decision Support System for Managers
Prof. Kunal Kanti Ghosh
Vinod Gupta School of Management
Indian Institute of Technology, Kharagpur

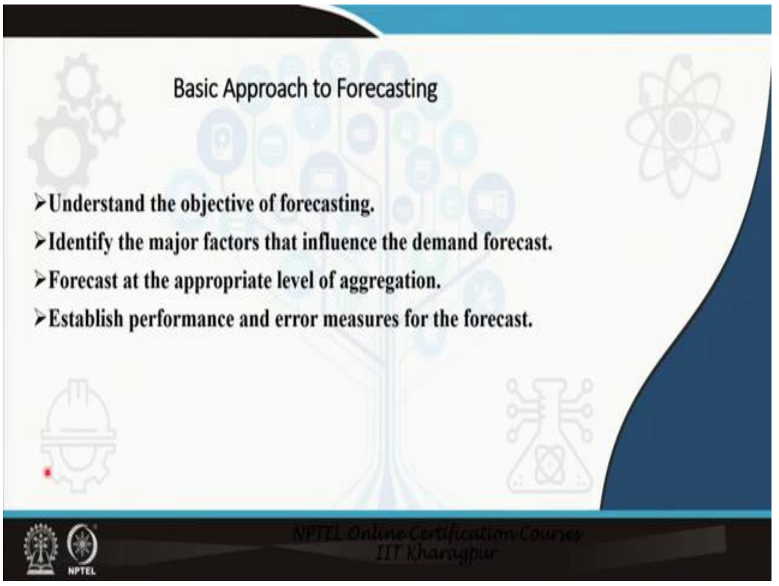
Week - 03
Module - 02
Lecture - 12
Decision Support Systems for Forecasting (Contd.)

Hi, welcome to our 2nd module of week 3 on our course "Decision Support Systems"! Today we are going to continue from where we had left. It is all related to forecasting techniques, forecasting model which are extensively deployed if we want to develop a decision support systems for forecasting particularly, the quantitative models.

(Refer Slide Time: 00:49)

Today we are going to cover the forecasting process itself, elementary concepts of linear regression, then some basics on time series forecasting and finally, we will talk about the static method of forecasting, the principle of which will be utilized in subsequent modules.

(Refer Slide Time: 01:21)

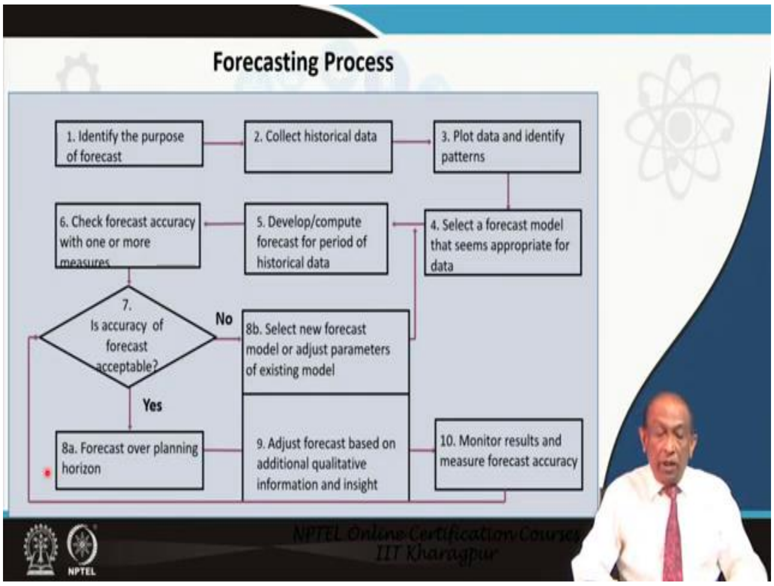


So, first let us start with the basic approach to forecasting. We have discussed in the 1st module that, first of all we need to understand the objective of any forecasting system. In relation to that we have to first identify the major factors that influences the demand forecast. Second we had already mentioned that forecast at the aggregate level is much more accurate rather than forecasting at each component level.

So, appropriate level of aggregation is another prerequisite for design and development of a decision support forecasting system and then last of all we have to establish the performance measures for error in a forecast, because a forecast is just a forecast it cannot be 100 percent accurate.

So, there will be some error and that error how do we quantify it, what are the different performance measures for forecast error which is basically the difference between whatever we forecast and the actual demand. So, how do we really quantify it? What are the measures? That we need to understand.

(Refer Slide Time: 03:45)



If we look at the forecasting process, it is absolutely clear that we need to identify the purpose of the forecast as the first step. Then collect that historical data related to the item for which we are interested in the forecast and here there is a basic assumption that the historical pattern the first trend we will continue in the future.

The third step is to plot that data and try to identify if there is any detectable pattern in that data. We are discussed in our previous module that if it is a time series; that means, when the demand is plotted against time that series might identify either a random fluctuations around an average level of demand or there can be some trend either upward or downward superimposed on the level along with the random fluctuations.

So, in here that is a trend along with random fluctuations or there can be seasonality over this average level of demand and there can be combination of all these; that means, there can be trend, along with seasonality, clubbed with random fluctuations. So, we have for each such situation different forecasting model.

So, we have to select a forecast model that seems appropriate in that context and having selected some two or three appropriate forecast model we develop or compute forecast for a given period of time which we are calling the forecasting planning horizon and then we check the accuracy of that forecast with different kinds of performance measures to quantify the forecast error.

Then we ask a question that is that accuracy of forecast acceptable to us? If it is acceptable to us if yes, then we will do the forecast over the selected planning horizon; ok.

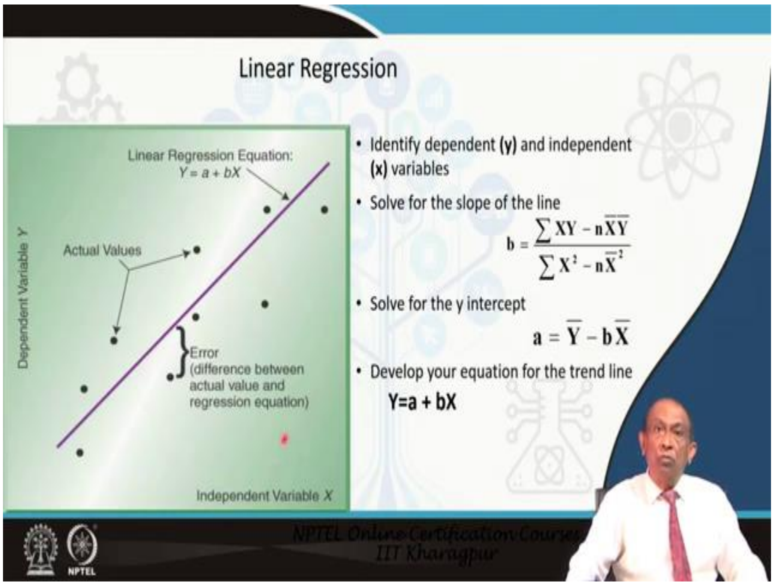
And then adjust that forecast based on additional quantitative information and insight this modification is based on managers experience, wisdom and judgement and there is the manager to compute an interaction which is one of the sense of any decision support system, because quantitative models cannot take into consideration all those factors that influence the forecast of an item.

So, some kind of modification is required based on all those unaccounted factors, then we monitored the results and measure the forecast accuracy over a given period of time and after some point in time we again go back and check that whether the accuracy of that forecast is acceptable or not.

Now, in the very first instance when we come from this block 6 to block 7, if the desired accuracy level of forecast is not reached, then what we do? We try to select a new forecasting model or the managers they adjust parameters of the existing model in the decision system and then again this particular loop is followed where is wherein we compute the forecast based on those revised parameter values.

And again go and come back and check whether the desired accuracy of forecast has been reached or not. So, this is the overall forecasting process that gets incorporated into any decision support systems used for forecasting the demand for any item.

(Refer Slide Time: 09:49)



Now, we talk off the elementary concepts related to linear regression, which is one of the quantitative techniques popularly used for time series analysis. Why we would like to discuss here?

Because the subsequent models they require the deployment of this concept of linear regression and since the examples that we have chosen in future discussion relates to only two variables Y being the dependent variable and X the independent variable, we will talk about only bivariate linear regression.

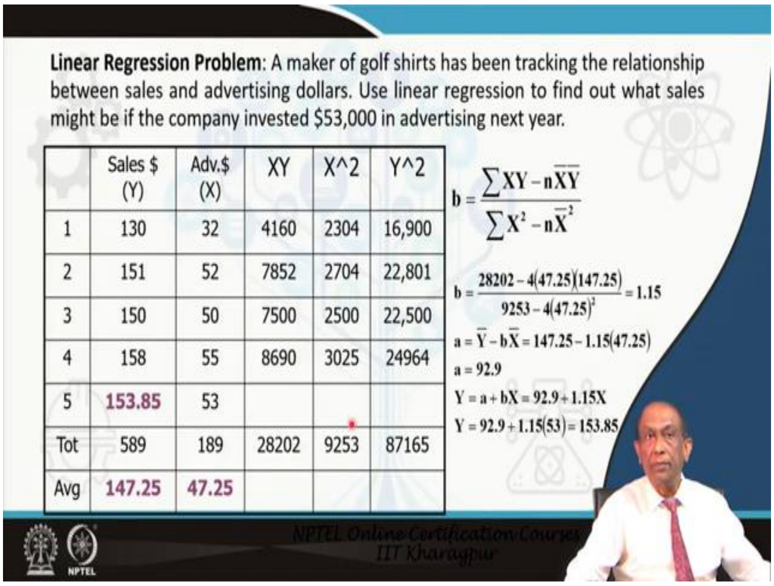
Now in a bivariate linear regression if the independent variable is X and the dependent variable is Y, then the scatter plot depicts a diagram like this, wherein the actual values of the dependent variable corresponding to some value of the independent variable are these black dots; ok.

And having plotted the scatter diagram we try to fit in a particular line which is called a regression line in such a way that the error which is the difference between actual values and the plotted line is minimized. In fact, we try to minimize the squared errors the sum of squared errors considering all the points.

And that approach helps us to identify the slope of the regression line which in this case is given by b and the Y intercept which is the value a and the expression for computing b and a is given by these equations, which basically comes from differentiating the erector with respect to these two parameters and these are very elementary things given in any statistics book.

So, having done that the equation for the trend line can be represented by Y equal to a plus bX, we know the value of a, we know the value of b. So, these regression line equation can be easily found out.

(Refer Slide Time: 13:14)

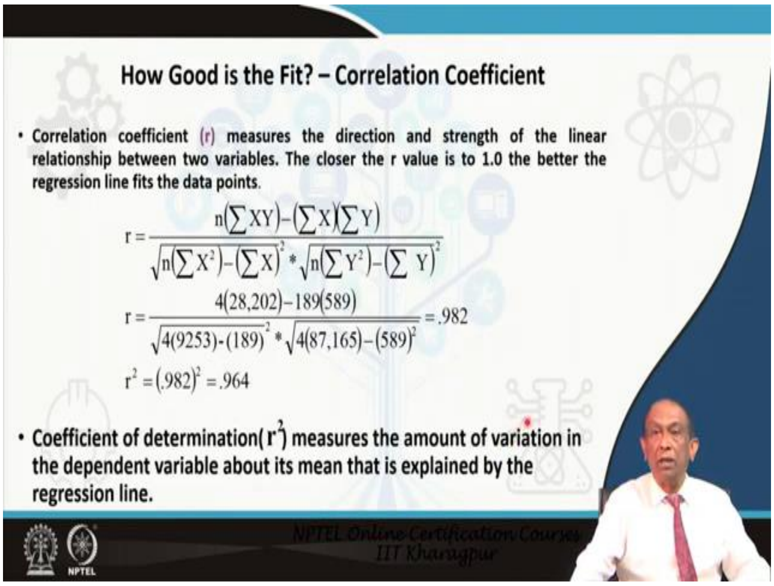


Here is one example problem, wherein a maker of golf shirts has been tracking the relationship between sales and advertising dollars. The problem is to use linear regression to find out what sales might be if the company invested so much in advertising next year.

Now, this table gives as all the required input values wherein this expression 32 is advertisement expenses. So, all these figures are in thousands of dollars. So, having plotted or having tabulated the values of X and Y, we find XY values, X square, Y square, then we do all the summations, Y square, X square, sum of XY, sum of X, sum of Y. We compute X bar which is the average value of X and average value of Y, Y bar all these are needed in here in computing these expression.

So, after having done this computation, we compute the value of b which in this case is this much and having computed b, we compute a, and then find out the equation of the regression line. Having obtained the equation of the regression line for any corresponding value of X, we can find out the value of Y. For example, in this case if the advertising expense is in terms of thousands of dollars say 53,000 dollars, then the corresponding sales that is we achieved will be 153,000 dollars, it is absolutely simple.

(Refer Slide Time: 15:27)

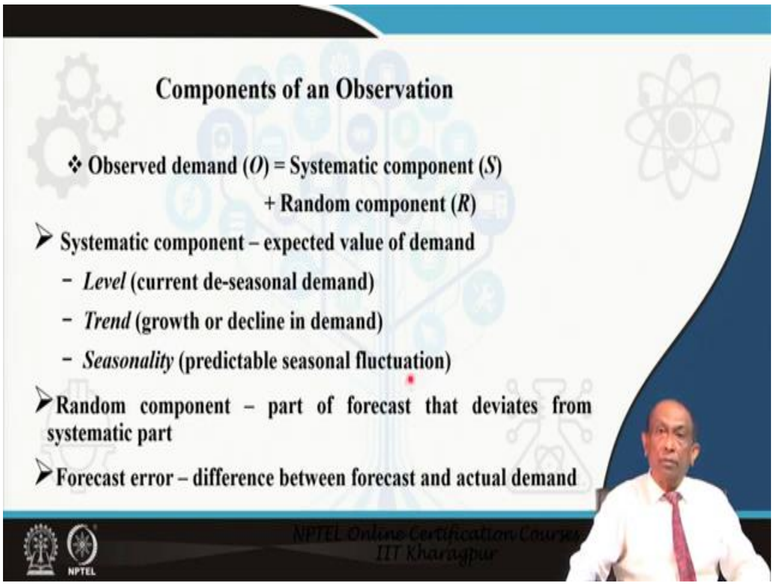


Having done that if we are interested to find out the degree of association between the sales revenue and the amount spent on advertisement, then we find out a measure which is called the correlation coefficient measure for which this is the equation, this is the formula.

And from the values that we have been given in this case, the correlation coefficient value comes out to be 0.982 which shows a very high degree of correlation that means the degree of association between these two variables is very high significant. In fact, any correlation coefficient value which is greater than say 0.5, 0.6 seems to be acceptably significant; ok.

And then having computed r, we compute another coefficient which is called coefficient of determination which is quantified by the square of the correlation coefficient that is r square. This coefficient of determination measures the amount of variation in the dependent variable about its average which is explained by the regression line. So, this is all about the regression technique and these will be utilized in subsequent models.

(Refer Slide Time: 17:24)

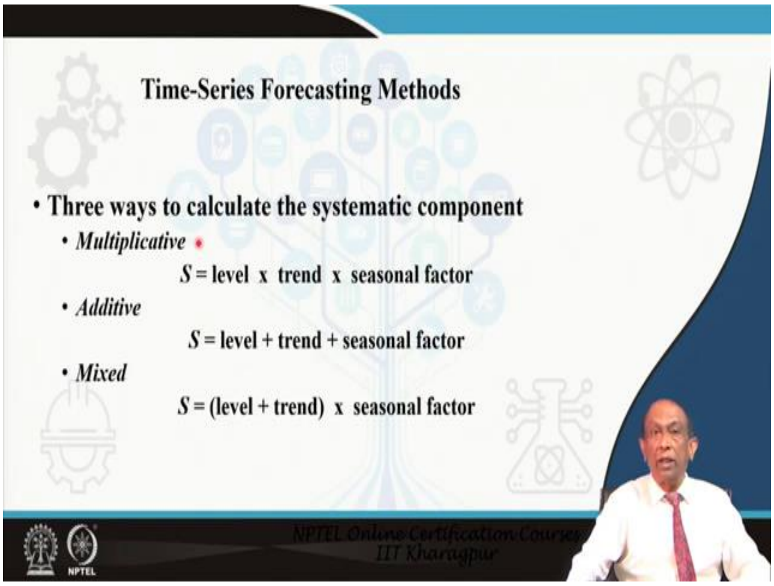


Now, if we look at the observed demand, then we will find that; that means, the demand data plotted over time whatever pattern you are observing that observed demands consists of two components. One is the systematic component S and the other one is the random component R.

The systematic component is the expected value of demand average value of demand, which has got a level which is basically de-seasonalized demand, then there is a trend aspect which basically represents the growth or decline in demand and there can be seasonality.

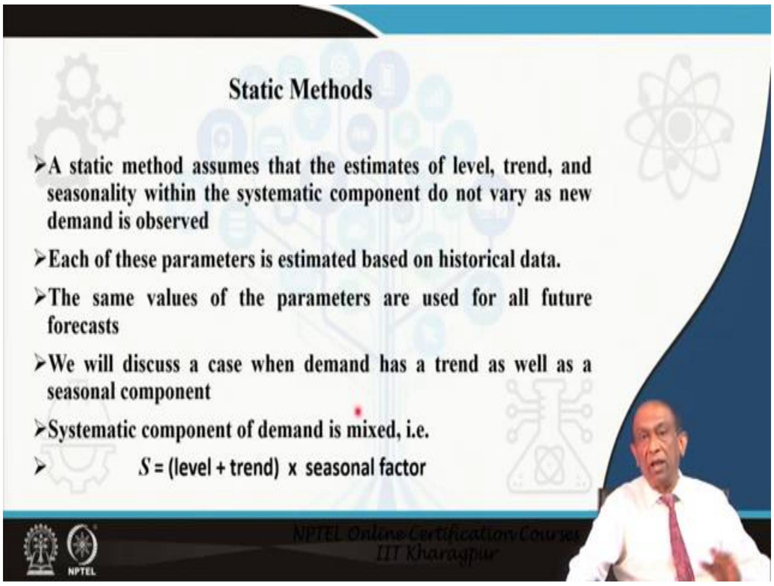
And there is the random component which is the part of forecast that deviates in a random manner from the systematic part and we define forecast error as the difference between forecast and actual demand. So, these particular concept we will utilize in forecasting models that we are going to explain.

(Refer Slide Time: 18:57)



So, there are three ways to calculate the systematic component: one is the multiplicative model, wherein we express the systematic component as the product of these three factors level into trend into seasonal factor. In the additive model it is all plus in case of this multiplication sign it is all plus. And the most widely deployed model is a mixed model, where the systematic component is the sum of level plus trend multiplied by seasonal factor.

(Refer Slide Time: 19:37)

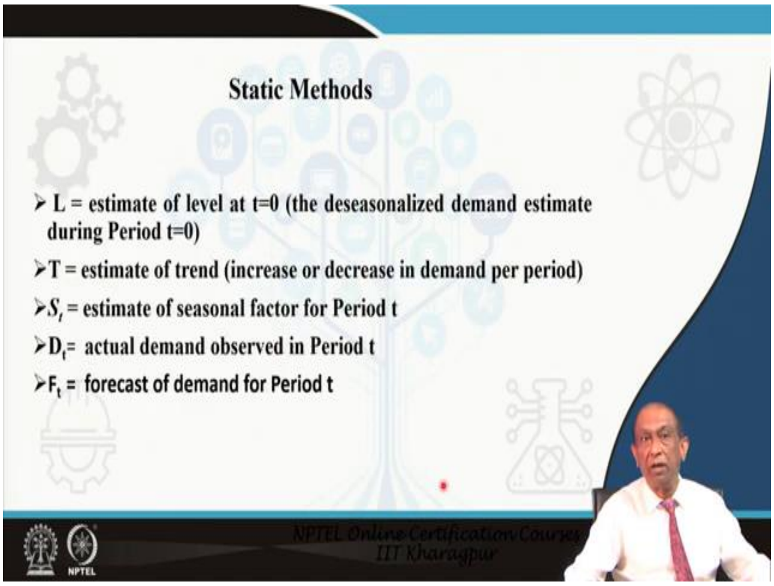


We start with the static methods of forecasting. In fact, we will discuss only one static method and that to for a mixed model case. So, a static method assumes that the estimates of level, the estimate of trend and the estimate of seasonality within that systematic component do not vary if you observe any new demand.

That means, once you compute these estimates it will remain static and these parameters that is the estimate of level, trend and seasonality is computed on the basis of historical data that we collect. And once that estimate is found out the same values of those parameters will be used for all future forecasts.

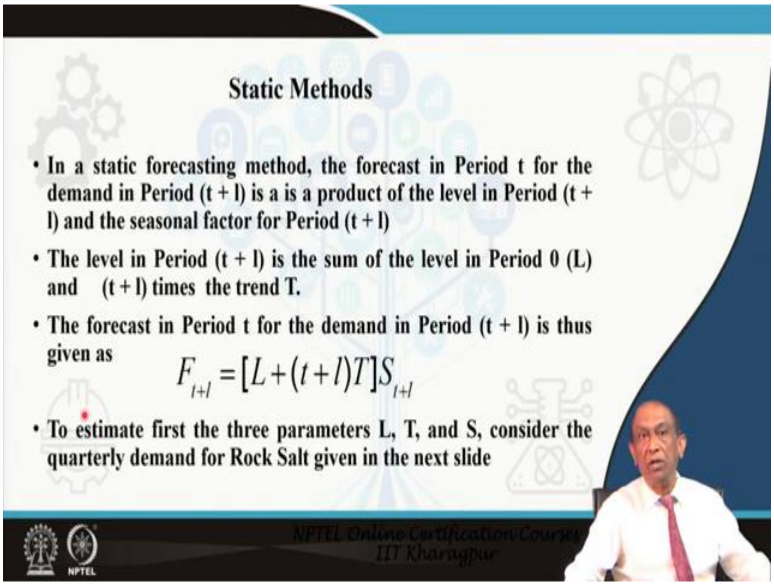
So, there are various cases, but we will discuss only one case where the mixed model will apply that is the case when demand has a trend as well as a seasonal component. So, as mentioned in the previous slide in this case the systematic component is equals the sum of level plus trend; that means, the estimate of level plus estimate of trend multiplied by the seasonal factor.

(Refer Slide Time: 21:33)



So, now let us see how we have deployed it. We will assume that capital L is the estimate of level at a point t equal 0, which is the de-seasonalized demand estimate during the period t equals 0. T is the estimate of trend that is increase or decrease in demand per period, S t is the estimate of seasonal factor for period t and D t is actual demand observed in period t, F t is a forecast of demand for period t ok, these are all subscripts.

(Refer Slide Time: 22:22)

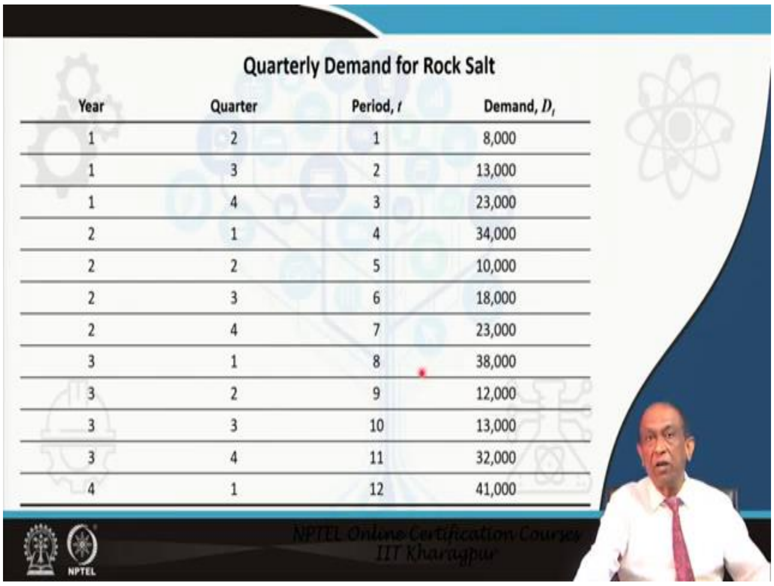


So, in a static forecasting method, the forecast computed in period t for the demand in period t plus l; that means, standing at the point in time t we are trying to forecast the demand l periods ahead of t, what is that? That is a product of the level in period t plus l and the seasonal factor for period t plus l.

The level estimate of level in period t plus l is basically the sum of the level in period 0; that means, the initial period which is computed as a L and then t plus l times the trend. And hence the forecast in period t for the demand in period t plus l is given as F t plus l is L plus t plus l into T this whole thing multiplied by the estimate S t plus l which is the seasonal factor for this particular period t plus l.

Now, when we take an example, it will be much more clear. So, to estimate first, the three parameters L, T and S, let us consider the quarterly demand for rock salt which is given in the next slide.

(Refer Slide Time: 24:11)

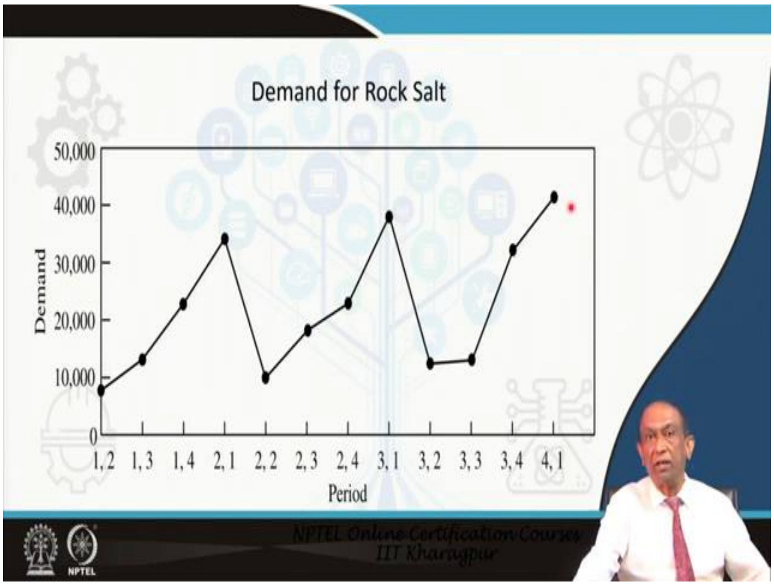


Quarterly Demand for Rock Salt

Year	Quarter	Period, t	Demand, D_t
1	2	1	8,000
1	3	2	13,000
1	4	3	23,000
2	1	4	34,000
2	2	5	10,000
2	3	6	18,000
2	4	7	23,000
3	1	8	38,000
3	2	9	12,000
3	3	10	13,000
3	4	11	32,000
4	1	12	41,000

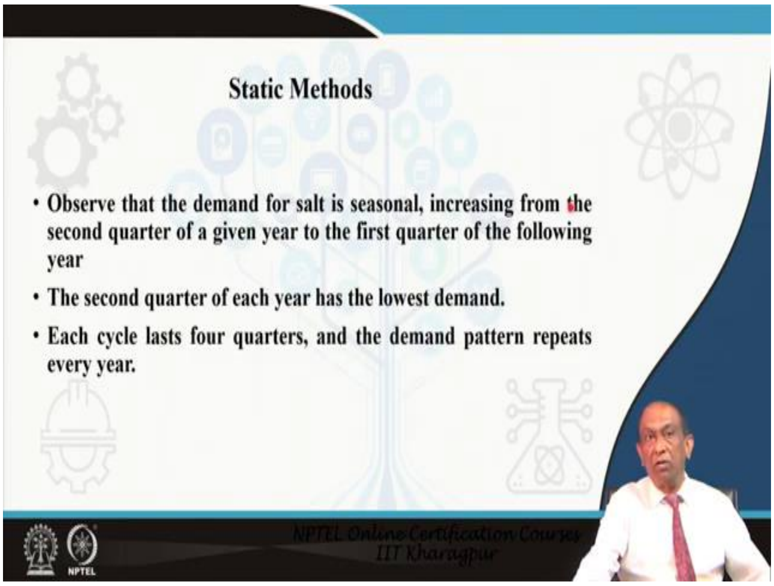
You see if you look at this data there is seasonality, it is a quarterly demand data we start from the second quarter of the 1st year, the demand in here is 8000 and you see you come here, then you come for the 3rd year, again second quarter, you see this pattern gets repeated. If you plot this data over a period of time, then we see that there is an effect of seasonality.

(Refer Slide Time: 24:50)



So, we have data for three years and we have plotted; we observed that there is trend as well as seasonality.

(Refer Slide Time: 25:10)



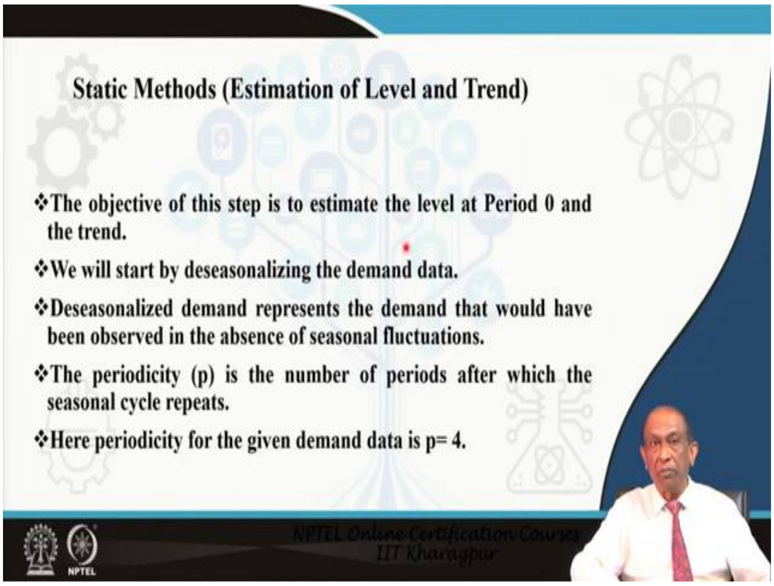
So, observe in that particular figure again for your convenience that the demand for salt is seasonal, increasing from the second quarter of a given year to the first quarter of the following year and if you look at this figure then you will find that the second quarter of each year has the lowest demand. We start from here and each cycle lasts four quarters and the demand pattern repeats every year.

(Refer Slide Time: 26:11)

And if you look at the data carefully you will also find that there is also a growth trend in the demand, with sales growing over the past three years. It is assumed that growth will continue in the coming year at the historical rates; that means, the pattern will continue.

The first thing what we have to do in order to estimate those three parameters level, trend and seasonal factors? We have to de-seasonalized the demand and run linear regression to estimate level and trend and then estimate the seasonal factors. And for de-seasonalizing demand there are various techniques available in any statistics book we have chosen one particular technique you can choose any one of them.

(Refer Slide Time: 27:15)

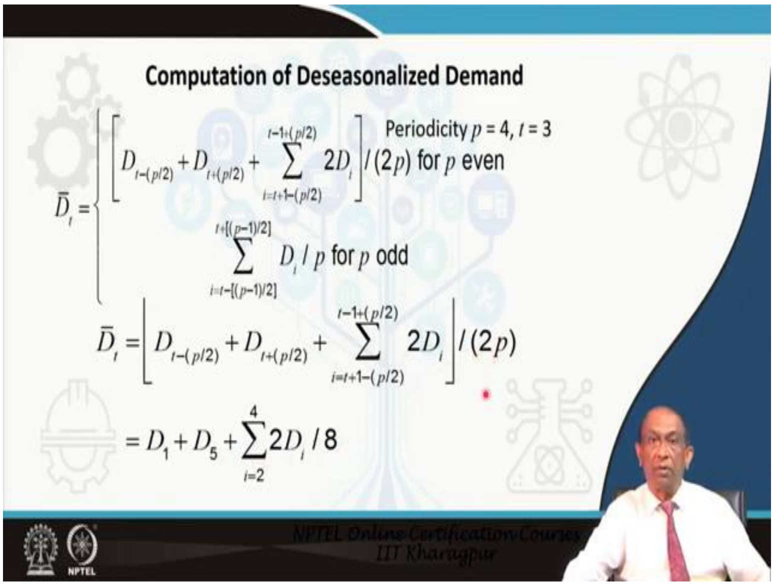


So, the objective of this step first step is to estimate the level at the initial period that is at period 0 and to estimate the trend at the initial period 0 first. So, we will start by de-seasonalizing the demand. De-seasonalized demand represents the demand that would have been observed if there would have been no seasonal fluctuations. With respect to the data that is given the periodicity of the data is p equals 4, where periodicity is the number of periods after which the seasonal cycle repeats, in this case it is very clearly 4.

(Refer Slide Time: 28:12)

To ensure that each season is given the equal weight at the time of de-seasonalizing demand, we will take the average of p consecutive periods of demand.

(Refer Slide Time: 28:43)

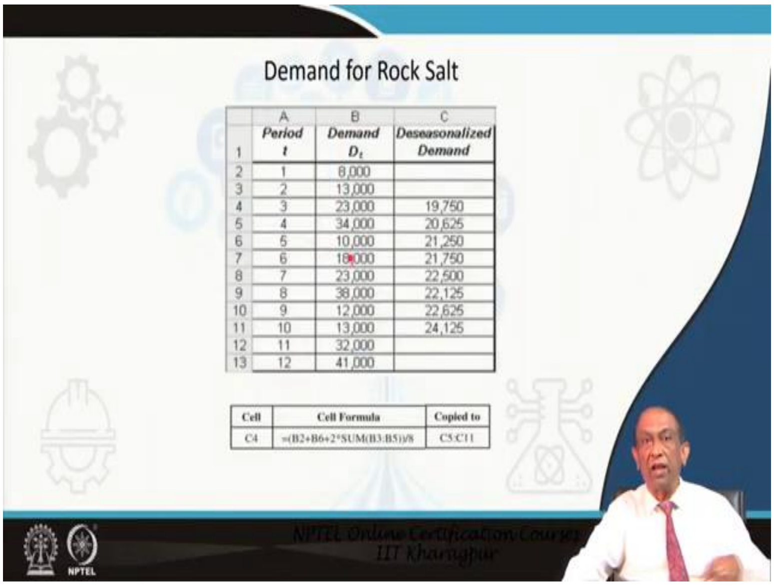


So, in our example p equals 4. So, first we will compute the estimate for t equals 3 using a formula which is very widely used. So, for t equal to 3, D 3 bar, D t bar is a de-seasonalized demand. So, D 3 bar is D 1 plus D 5 plus this with this particular expression.

(Refer Slide Time: 29:12)

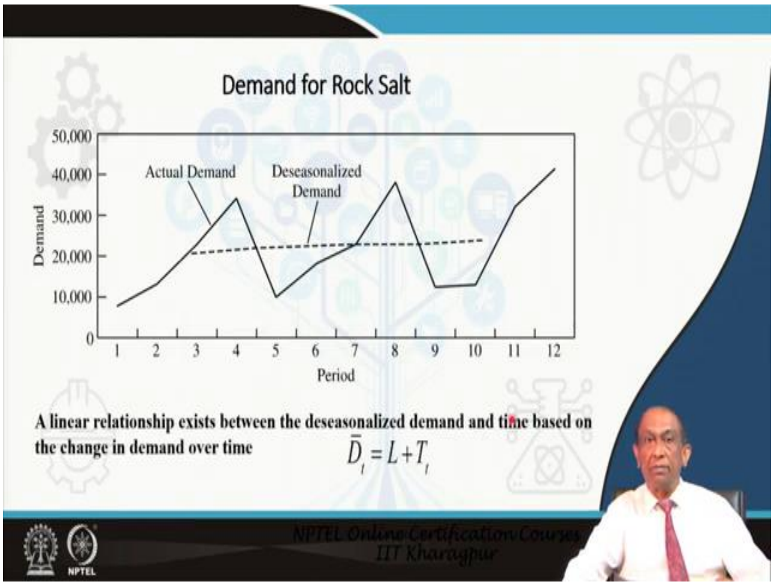
You need not worry about it because I have very clearly explained that through a spreadsheet. And using the same formula, we can obtain deseasonalized demand between periods 3 and 10 as given in the next slide.

(Refer Slide Time: 29:27)



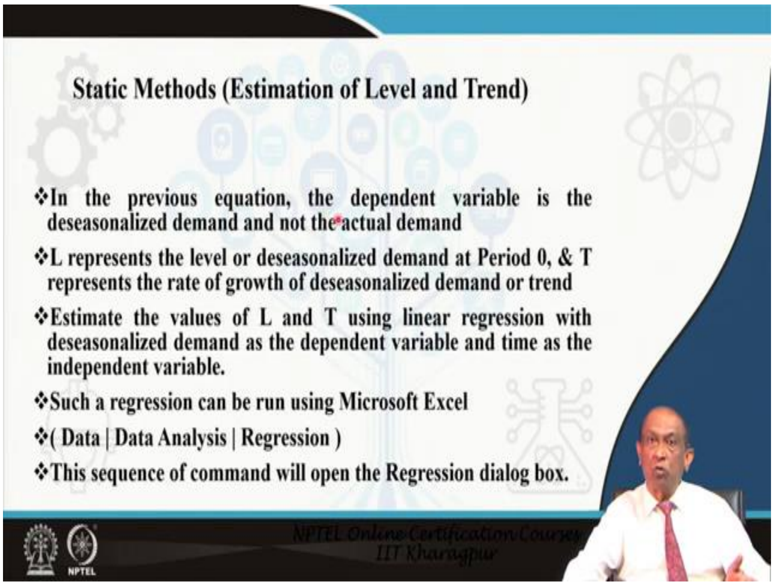
This is the cell formula that we have used based on the equation that we have given.

(Refer Slide Time: 29:42)



Having done that if we plot the actual demand and the de-seasonalized demand over a period of time, then we observes that a linear relationship exists between the de-seasonalized demand D t bar and time given by the expression D t bar equals L plus T t.

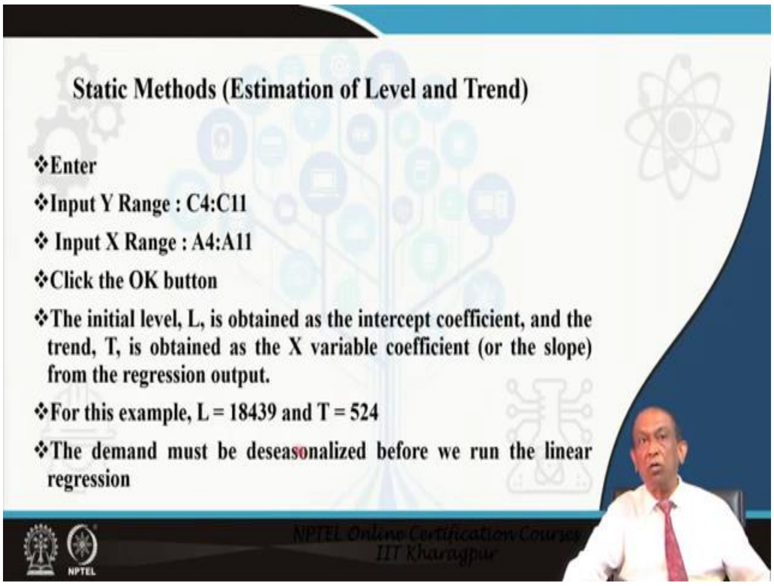
(Refer Slide Time: 30:18)



So, in this equation previous equation that we have just shown, the dependent variable D t bar if the de-seasonalized demand and not the actual demand. We have to always de-seasonalized demand before we apply this model. L represents the level of de-seasonalized demand at period 0 and T represents the rate of growth of de-seasonalized demand.

We will estimate the values of L and T using linear regression techniques for which we will be using Microsoft Excel we will be using the data option within data, we will go to data analysis. And then we will typing the command or choose from the menu regression and this sequence of command will open the regression dialogue box.

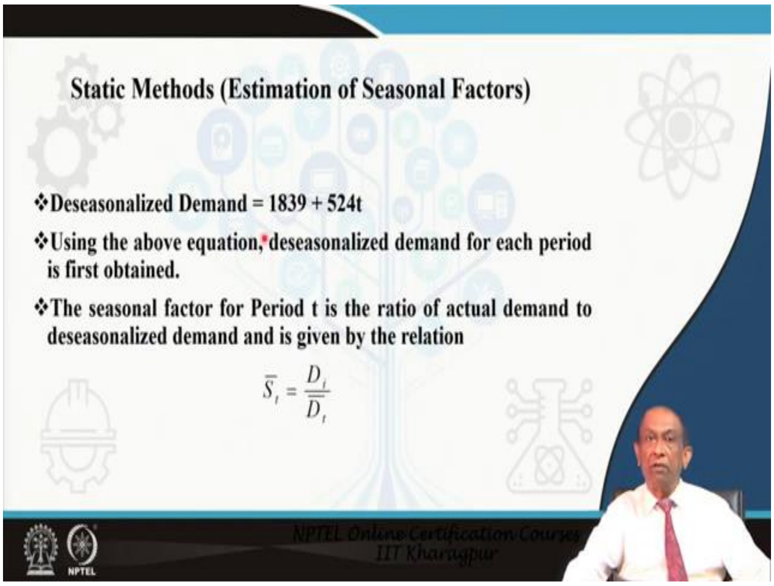
(Refer Slide Time: 31:12)



Wherein we will be inputting the Y range that is the dependent variable from the spreadsheet that we have shown it is C 4 to C 11. We give the input range and then we get the initial level is obtained at the intercept coefficient and the trend T is obtained as the X variable coefficient or the slope from the regression output.

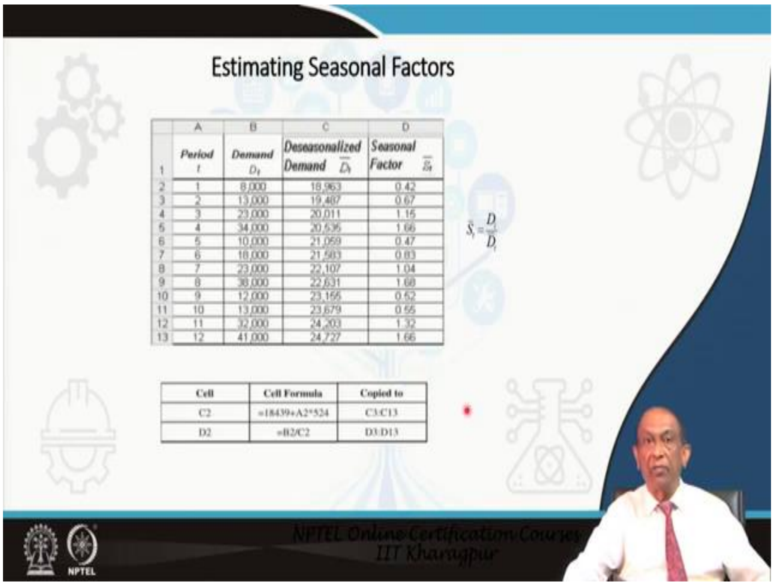
And for this example the value of L will be 18439 and the value of trend or estimate of trend will be 524 ok. So, always remember that you need to de-seasonalize the demand before you run the linear regression.

(Refer Slide Time: 31:59)



Now, having obtained L and T you can find the equation of the de-seasonalized demand equals this. Using that above equation, de-seasonalized demand for each period is first obtained and then we compute the seasonal factor for period t as the ratio of actual demand to de-seasonalized demand and is given by this expression.

(Refer Slide Time: 32:27)

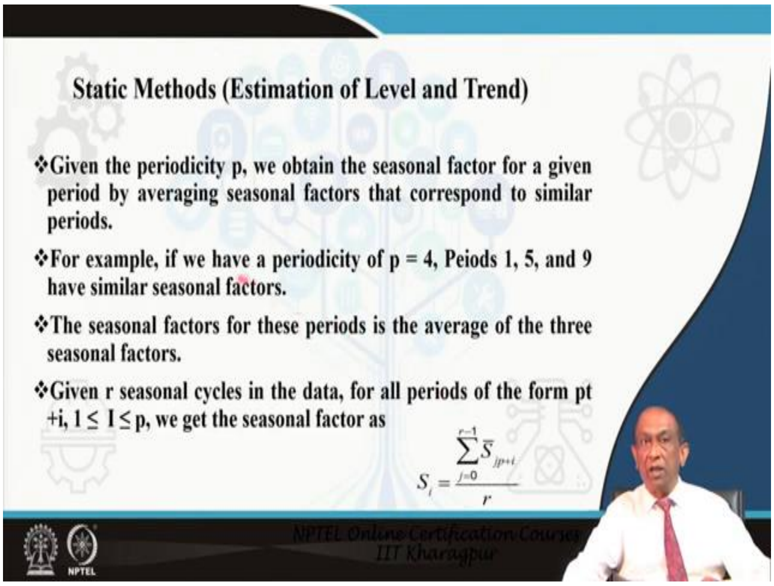


	A	B	C	D
1	Period t	Demand D_t	Deseasonalized Demand $\bar{D}_t$	Seasonal Factor $\bar{S}_t$
2	1	8,000	18,963	0.42
3	2	13,000	19,487	0.67
4	3	23,000	20,011	1.15
5	4	34,000	20,535	1.66
6	5	10,000	21,059	0.47
7	6	18,000	21,583	0.83
8	7	23,000	22,107	1.04
9	8	38,000	22,631	1.68
10	9	12,000	23,155	0.52
11	10	13,000	23,679	0.55
12	11	32,000	24,203	1.32
13	12	41,000	24,727	1.66

Cell	Cell Formula	Copied to
C2	=18439+A2*524	C3:C13
D2	=B2/C2	D3:D13

So, if you look at this, you see de-seasonalized demand by actual demand; we compute the seasonal factors for all these periods; the cell formulas are given having done that; ok.

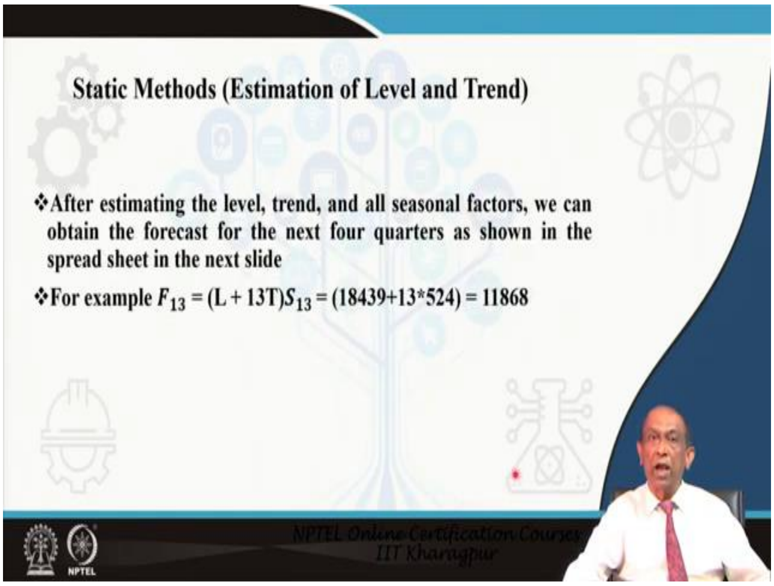
(Refer Slide Time: 32:47)



Given the periodicity p because it is repeating after every 4 inter time intervals, we obtain the seasonal factor for a given period by averaging seasonal factors that correspond to similar periods.

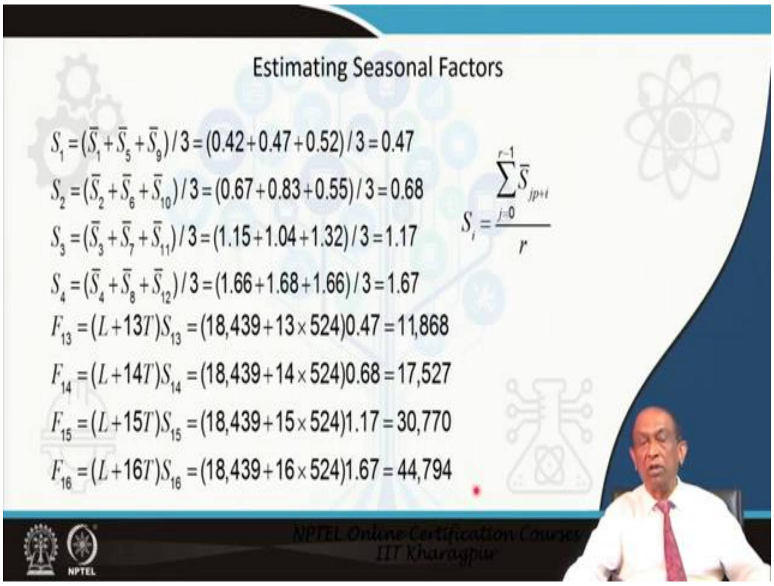
For example if we have a periodicity of p equals 4, then what we will find at periods 1, 5 and 9 have almost similar seasonal factors. Then what we have to do the seasonal factors for these periods if the average of the three seasonal factor that we have already computed. Given r seasonal factors in here it is 3, for all periods of the form pt plus i we will be computing the seasonal factor with this kind of expression.

(Refer Slide Time: 33:44)



If you look at the next computation, then it will be absolutely clear and then after estimating the level, trend and all seasonal factors we can obtain the forecast for the next four quarters for example, if you look at it.

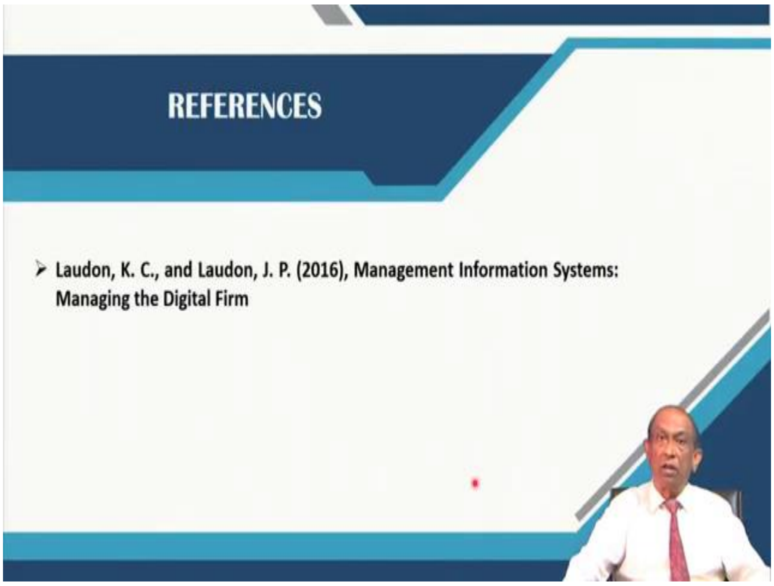
(Refer Slide Time: 34:10)



You see S 1 will be S 1 bar plus S 5 bar plus S 9 bar 4 periodicity 4, 1 5 9, then 2 6 10 like this, we compute S 1 as this, S 2 as this, S 3 as this, S 4 as this, having done that the forecast for period 13. Because three years data we have taken periodicity 4. So, for all 12 months we have got the data, for 12 quarters we have got the data.

So, for the 13th quarter, it is L plus 13 T substitute the value 13 here in place of L and take the corresponding seasonal factor for the 13th period which is nothing but 1 5 9 then 13. So, take these value 0.47 so, the forecast for the 13th period is 11868. Similarly, forecast for the 14th period is 17527, in here the seasonal factor for S 14 will be the same as what we have computed here as S 2 like this for this remaining 2. So, this is all about the static method for a mixed model.

(Refer Slide Time: 36:02)



This is the reference that we have used.

Thank you!