Course Name - Operations and Revenue Analytics

Professor Name - Prof. Rajat Agrawal

Department Name - Department of Management Studies

Institute Name - IIT, Roorkee

Week - 01

Lecture - 04

Welcome friends, in our last class we discussed some basics about descriptive analytics. We discussed how, just by simple data visualization using tools like Excel, we can see what the important characteristics of your data are. For those characteristics, we basically relied on two important things, one of which, if you recall, is the measure of central tendency. We saw that we can calculate mean, median, mode, geometric mean, etc. These are different ways of identifying the central tendency of your data.

We also discussed the variation in your data, which we explored with the help of range. We discussed standard deviation, variance, and the coefficient of variation. All these different things represent the variation in your data. So, by applying central tendency and variation, we are able to see many important characteristics of our data. Going further, we now know about predictive analytics. Predictive analytics is like expecting that what has happened in the past will also happen in the future.

Many decision-making processes in operations management involve identifying a particular pattern in your data, and if you identify that pattern, you may proceed with some kind of decision-making. Descriptive analytics is just about a kind of postmortem of your historical data. But using that same historical data for decision-making is predictive analytics. So, it is one step ahead of descriptive analytics. So, in this particular session, we are going to discuss a very popular use case of predictive analytics in operations management.

In operations management, as we discussed in the very first class, most of the decisions are taken on the basis of forecasting. That is, what will be the possible demand for your product in the future? And when I say in the future, the future may be the immediate next period, the future may be some period ahead, or the future may be many periods ahead. So, there can be different classifications of the future as well. So, predictive analytics can help us in getting some numbers, some data for my future demand, and that is what we are going to discuss in this particular session. Where we are primarily, in this particular session, going to focus on this time series forecasting.

Time series forecasting means you have historical data. These are periods 1, 2, 3, 4, 5. Now, based on these past data, can I predict for the 6th period or the 7th period? These are my immediate next periods. Can I predict for the 8th, 9th period, slightly further, or can I predict for the 12th period, which is even further? So, all those things are possible with the help of predictive analytics. So, this is the purpose of this particular session, that we are going to discuss one use case of forecasting, but there may be many such cases where you can see how you have performed in the past and accordingly, you have to use that historical data for your future predictions. So, time series forecasting, as I just

Time Series Forecasting

- **Multiplicative:** Systematic component = level × trend × seasonal factor
- **Additive:** Systematic component = level + trend + seasonal factor
- **Mixed:** Systematic component = (level + trend) × seasonal factor

mentioned, is the use of historical demand data for knowing a particular forecast.

If you only talk of forecasting, there are so many different methods through which you can do the forecasting, and one of the most common methods, which is based on this predictive analytics, is the time series analysis where we are using historical data. And this historical data, historical data means past data. So, whenever you have any kind of past data, you will see that, let us say, I am giving you some indication that this is, let us say, period 1, period 2, period 3, period 4, period 5. So, it may appear to you that this data is like this way, and then now you have identified that pattern of this data that there is some increase in data, some increase in demand happening from one period to another period, and you are expecting that the same pattern will continue for the next period as well. And therefore, in the sixth period, this will be the possible demand.

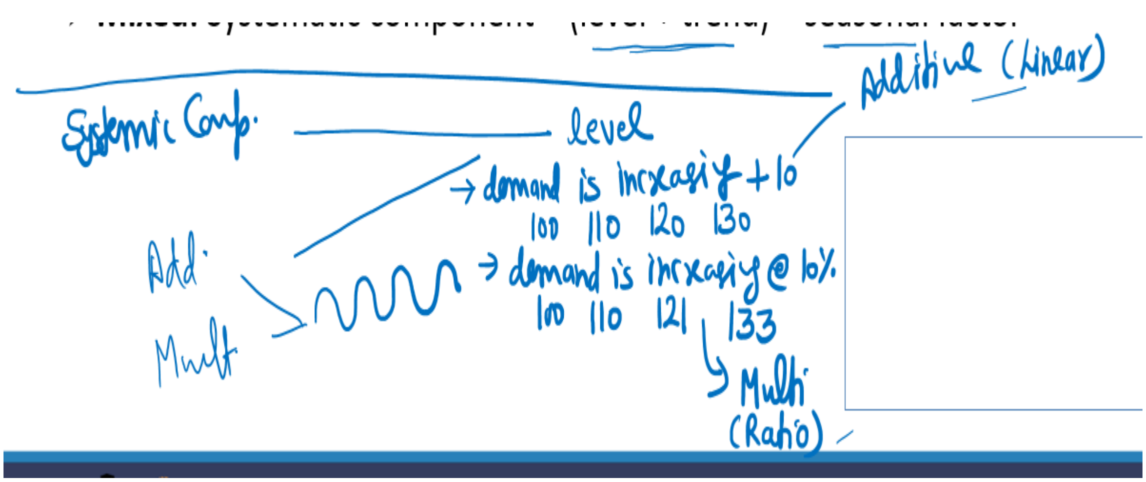
But actually, what happens is this is what I am saying theoretically, but actually, what happens is the demand is moving like this. If you see this green line, the demand is moving like this; it is not a straight line like the red line. It is actually a zigzag line, this green line. And therefore, it is not so easy; it is not so easy to predict how this zigzag line will move, whether it is like this or it is also possible like this. All these are the possible values which may be there in the sixth period because there is a random variation happening in your demand data. Now, you cannot determine any kind of randomness in your data. But, when we are doing predictive analytics, we divide our data into two components.

Out of this past data, which is moving like this, you can make a line which is a straight line, which best represents this zigzag movement. And let us say that the best line is this line, which best represents this zigzag movement of the data. Now, in this actual data, there are two things. One is the systemic component, and another is the random variation. So, the systemic component plus the random component makes your actual data.

So, what we do with the help of predictive analytics is we are basically predicting this systemic component. And then we also determine what the limits are in which this random component may vary because, based on historical knowledge, we know what the limits are in which this random component will vary, and based on that, we may say that this will be the possible forecast in the future, and these are the possibilities of the variation. So, that gives us more complete information. So, now when we are going for this systemic component of the demand data, this systemic component may also have some kind of characteristics, and these characteristics are possible in a variety of ways. The systemic component may have a straight line like this also.

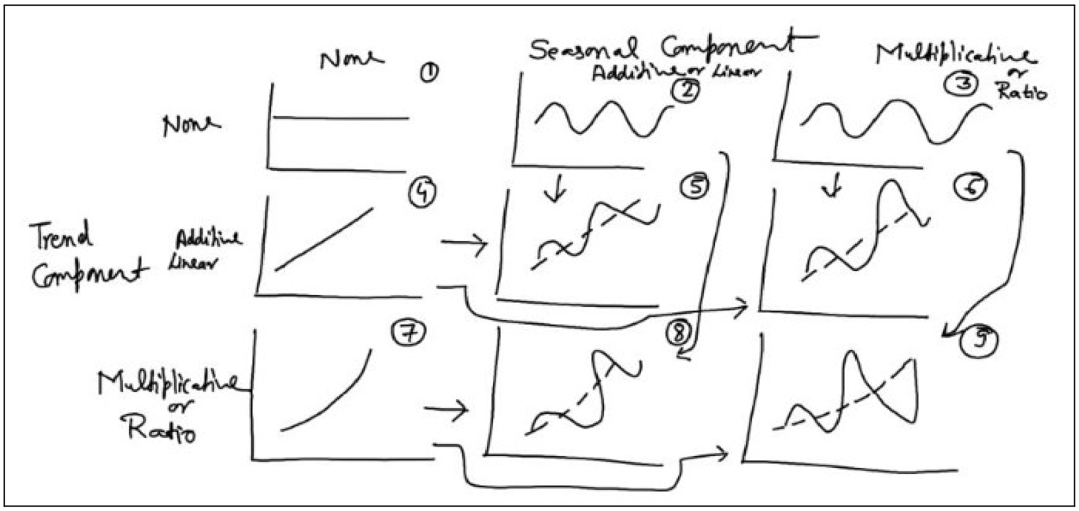
The systemic component may itself have a trend, and a systemic component may also have some kind of cyclic movement. So, all these different types of characteristics are possible in your systemic component as well. And when you see the systemic component of your demand data, there is a level where we say that the data is moving horizontally. It is parallel to the x-axis. The systemic component may have a trend where, from one period to another period, there is an increase in demand. Now, that increase in demand from one period to another period may be possible in two ways.

One is that demand is increasing. Let us say by plus 10 from one period to another period. Last period, the demand was 100, then it becomes 110, then it becomes 120, and 130 like that, so from one period to another period, demand is increasing by plus 10. Another way is that demand is increasing at the rate of 10 percent, which is also possible. So, initially, if demand is 100, it becomes 110, and then it becomes 121, and then again it becomes 133, and so on. So, now demand is increasing at the rate of 10 percent.



So, you may have your trend either as an additive trend or a multiplicative trend; both these things are possible. This is additive; this is multiplicative. Similarly, seasonality means that in a year, demand increases to a high level in a particular season, such as the winter season, rainy season, or festival season, and then again demand comes to a lower level. And this increase in seasonality from one cycle to another cycle may be on an additive term or a multiplicative term. So, based on all these things, you may have your systemic component, which may have a level, and if all these are multiplicative, the trend is also multiplicative.

One more thing, you can call multiplicative activities as ratio and additive as linear. So, these are other terms which are popular in literature for additive and multiplicative, linear and ratio. So, if all the components are multiplicative in nature. So, label multiplied by trend multiplied by seasonal factor will give you your systematic component. If all these are additive, then level plus trend plus seasonal factor will give you your systematic component.



Our objective is to determine only the systematic component. Random variations are beyond our control, and if one of them is additive or multiplicative, then appropriately, like in this particular case, we have considered trend as additive and seasonal factor as multiplicative. So, in this way, label plus trend multiplied by the seasonal factor will be the formula for calculating your systematic component. Now, here you see that we may have various different combinations, like just we saw, you have label, trend, seasonality, and trend can be additive or multiplicative, seasonality can also be additive and multiplicative. So, in this way, you may have a variety of combinations of trend and seasonality with level.

So, we have made nine different types of diagrams where you have different combinations of level, trend, and seasonality. So, if you see in this picture, figure number 1, here neither trend is there nor seasonality is there. So, you see this horizontal line which is parallel to the x-axis; on the y-axis, we have the systematic component, and on the x-axis, we have time. So, here you see that the systematic component of demand is almost horizontal, parallel to the x-axis in diagram number 1. In diagram number 2, we have a seasonal component which is linear, but there is no trend.

Before we go to diagram 2, let us go to diagram 4 first, which is easier to understand. Now, in diagram 4, there is no seasonal component. Only a linear trend is present, and you can see that this is an additive trend. So, a straight line like this is created. Similarly, in diagram number 7, we have a multiplicative trend, and it looks like this. In 1, 4, and 7, there is no seasonal component. Now, coming to diagram numbers 2, 5, and 8, where you have a linear seasonal component, an additive seasonal component.

So, diagram 2, diagram 5, and diagram 8 are linear seasonal components with no trend, additive trend, and multiplicative trend, respectively, and the shapes of the systematic components are given here. Like this is the shape, this is the shape which you can see in blue color, and similarly, if you see diagram numbers 3, 6, and 9. Diagram numbers 3, 6, and 9, in this case, we have a multiplicative trend which is combined with no trend, multiplicative seasonality which is combined with no trend, additive trend, and multiplicative trends, and accordingly, these three diagrams. So, you see that in all, we have nine combinations, or you can say nine types of characteristics in historical data. And based on these, we may have different types of predictive analytics models for handling each of these different cases.

One formula may not be possible for all the cases. Depending on the characteristic your historical data has, you need to require a particular algorithm, a particular method which is suitable to handle that type of characteristic of data. Now, let us consider the simplest type of example, which is case number one: no seasonality, no trend. And when there is no seasonality and no trend, you actually go with the simplest type of method, which is the moving average method. Now, the moving average method is where you have only a level in your data, no trend, and no seasonality.

And, in this case, the simple process is that whatever historic data you have, you need to take the average of some recent number of data. So, some recent data is known as N periods, and with that N period data, you will be able to get the forecast for the next period. Now, obviously, the critique of this system is that for having a better forecast, you need to have data of a large number of periods, and that will, if you remember in our previous class of descriptive analytics, we discussed that the average is not a very good indicator and is very much affected by outliers. So, to minimize the impact of these outliers for any kind of average activity, you need to take a large amount of data, and in this example, I am just showing you how we are calculating this moving average. For example, we have the weekly sales of a particular product that is 38,000 tons, 35,000 tons, 77,000 tons, and 90,000 tons. These are the weekly sales of the April month.

So, week 1, week 2, week 3, week 4, you have data 38, 35, 37, and 90. Now, if I have to determine the forecast for the first week of May, these are all April, to determine the forecast for the first week of May. Now, if I take the moving average period equal to 1, if I take the moving average period equal to 1, the forecast for the first week of May will be 90, if N is equal to 1. Now, if I want to again see whether my forecast is correct, can we have a better forecast? The answer is yes.

You can have a better forecast, but you need to increase the period of the moving average from N equal to 1 to N equal to 2, taking the average of the most recent two periods. You can further improve your forecast; now take N equal to 3, and this N equal to 3 will be for

Example

The forecast of demand for Period 5, is expressed as $F_5 = S_4 = 60$ thousand tons

As the sale in Period 5, D_5 is 80, we have a forecast error for Period 5 of

$E_5 = F_5 - D_5 = 60 - 80 = -20$

After observing demand in Period 5, the revised estimate of level for Period 5 is given by

$S_5 = (D_2 + D_3 + D_4 + D_5)/4 = (35 + 77 + 90 + 80)/4 = 70.5$

the most recent three periods, W2, W3, W4: 35 plus 77 plus 90 divided by 3. And in my slide, I have already solved that if I take the moving average period equal to 4, this will be 38, 35, 77 plus 90 divided by 4, and you get the answer equal to 60.

So, you see that as I am increasing the period of the moving average, my estimations are also changing, and therefore, for having a better forecast, you should have more periods in my historical data. For if I use N equal to 5 in this case, if I use N equal to 5, I will be requiring March week 4 data also.

Only when I have March week 4 data, then I will be able to calculate the forecast using N equals to 5. So, this is the moving average method. And the moving average method has the limitation of requiring more data. Moving average requires and is very sensitive to any kind of outlier. Therefore, sometimes people say that your demand should be more dependent on the recent data. In this case, when I am taking N equals to 4, I am giving equal weightage to all the past periods. But, it may not be desirable to give equal weightage to all the past periods.

So, we therefore go to the weighted moving average method. In the weighted moving average method, we may have the same N equals to 4, but here what we do is this is week 1, week 2, week 3, week 4 of April. So, if I am looking for a forecast for the first week of May, in this case earlier it was simply the sigma wi, i equals to 124 by 4. But now I will have W4 and a weight alpha4 plus W3 alpha3 plus W2 alpha2 plus W1 alpha1. These are the weights alpha4, alpha3, alpha2, alpha1, and I have taken these weights in such a manner that alpha4 is higher than or equal to alpha3, higher than or equal to alpha2, higher than or equal to alpha1, and alpha4 plus alpha3 plus alpha2 plus alpha1 equals to 1 also.

$$\text{April} \left\{ \begin{array}{l} w_1 \\ w_2 \\ w_3 \\ w_4 \end{array} \right. \qquad N=4$$

$$\text{May} \quad F_{w_1} = \underbrace{\frac{\sum_{i=1}^{4} w_i}{4}}_{\text{Now}}$$

$$\left.\begin{array}{l} w_4\alpha_4 + w_3\alpha_3 + w_2\alpha_2 + w_1\alpha_1 \\ \alpha_4 \geq \alpha_3 \geq \alpha_2 \geq \alpha_1 \\ \text{and} \quad \alpha_4+\alpha_3+\alpha_2+\alpha_1 = 1 \end{array}\right\}$$

So, these are the two conditions, where we get the liberty and flexibility to give more weightage to the most recent period, and as we are moving away from the most recent period, these weights can reduce in a particular manner. So, you will not be affected by a particular outlier. If there is a period which is slightly abnormal in your demand data, historic demand data, you can reduce the weightage of that period by choosing a particular suitable weight. So here, you see that these types of weights etc. are possible for our demand data of the previous case, and you can use the same data for these weights and see what your new forecast is. So, after these moving average methods where we have the limitation that we need a huge amount of historic data and we are not able to set the importance of our historic data, because there may be outliers which may affect our decision-making in a big way because of a static method of decision-making.

It will have limited use in predictive analytics where we are looking for regular calculations, and we are getting more real-time data. So, how can we use that real-time data for improved forecasting purposes? Then, we have another category of methods which are part of exponential smoothing methods. In these exponential smoothing methods, we may be using different types of smoothing constants: alpha, beta, gamma, etc. As you may recall, I discussed that in our historical data, we have two important components.

We actually have our data like this, and if I smooth this data, it becomes my level diagram. The dotted line represents the level, and the zigzag line, the continuous line, is the actual data. So, there are fluctuations like these. Now, you see my diagram; the diagram is moving like this, and if I compress these lines in such a manner that I smooth these lines. So, it will then become a straight line, which is represented by—let me use a different color so that you can understand this line.

This is that line which is the result of the smoothing impact on the actual demand data. This is known as the exponential smoothing method. Now, smoothing methods will give you the flexibility to use different types of methods, different types of formulas, depending upon the nine different models which we discussed at the beginning of this particular session. You may have examples like when there is no trend or seasonality. We will be discussing in our next class how different types of exponential smoothing methods are possible.

With the start of a very simple case of exponential smoothing where there will be no trend, no seasonality, and slowly we can build complexity so that you may have a trend also, you may have seasonality also, additive multiplicative trend, multiplicative seasonality. And that will give you the power of analytics to understand how different types of trends in your data, different types of characteristics of data, can be used for modeling predictive analytics forecasting systems. So, with this, we come to the end of this particular session, and we will be discussing various types of exponential methods in our next session. Thank you very much.