Engineering Econometrics
Prof. Rudra P. Pradhan
Vinod Gupta School of Management
Indian Institute of Technology, Kharagpur

Lecture – 41
Time Series Modelling - Basics

Hello everybody; this is Rudra Pradhan here, welcome to Engineering Econometrics. Today we will start with a new component; that is Time Series Modelling. In the last couple of lectures, we have discussed various types of econometric modelling starting with simple regression modelling, multiple regression modelling, again the cluster between linear regression modelling, and non-linear regression modelling. Again the kind of you know modelling where dependent variable will be qualitative in nature, again independent variable will be qualitative in nature. And we have gone through various you know test structures starting with; specification test, goodness of fit test, diagnostics tests, and that to check the validity of the estimated model.

And then over the time we have used several kind of you know examples and then we use different kind of this kinds of you know a model to you know apply the engineering econometrics tools; to solve some of the you know kind of you know engineering problems. And as per the kind of you know engineering requirement or the kind of you know organisational requirement and the kind of you know industrial requirement. However, the thing is that you know whatever we have discussed till now mostly it is the kind of you know cross sectional modelling, where we have use cross sectional data to estimate the model and use the model for the kind of you know prediction and the kind of you know forecasting.

That means, any kind of you know engineering predictions and any kind of you know engineering forecasting. So, the cross sectional data may be like you know individual specific, organisational specific, then you know sectoral specific, country specific, then you know kind of you know institutional specific. So, likewise we can have you know various forms of you know cross sectional data. So that means, you know basically it brings you know different kind of you know set ups and through which you can actually generate the data. And with the basis of you know theory problems and the kind of you

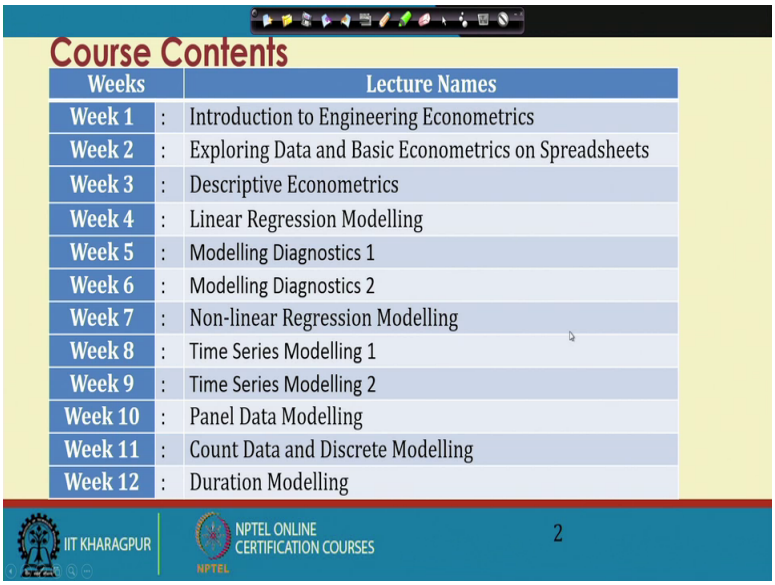
know requirement; we like to you know you like to use this data and you know kind of you know econometric model.

And then we estimate and after doing estimations we can go for the testing. And then once we you know get the clearance from the testing through like you know; specification test, goodness fit test, then the diagnostic test, then will you go for the kind of you know engineering you know forecasting requirement or the engineering prediction requirement. So, what is happening here is you know it means in this kind of you know time series modelling; one of the most important is the data structure. So; obviously, you know the model itself will give you the indication what is the a kind of you know data we have used, for the estimation process and the kind of you know prediction process. The modelling structure will be in three different forms so starting with you know cross sectional modelling, then time series modelling, and again the kind of you know panel data.

So, these modelling's are exclusively on the basis of you know data structure. If the data in say if you use time series data for a kind of you know estimation process then we simply called as you know time series modelling. And, when we use cross sectional type of you know data then we called as you know time series modelling. And then when we pull both cross sectional and time then it will come as a you know pull data and later we will have a discussion on the kind of you know concept called as you know panel data and that too panel data modelling. So, whatever we have discussed earlier that is mostly with respect to cross sectional modelling. And right now we will discuss the time series modelling and after knowing the time series structure and time series modelling we will move to panel data modelling.

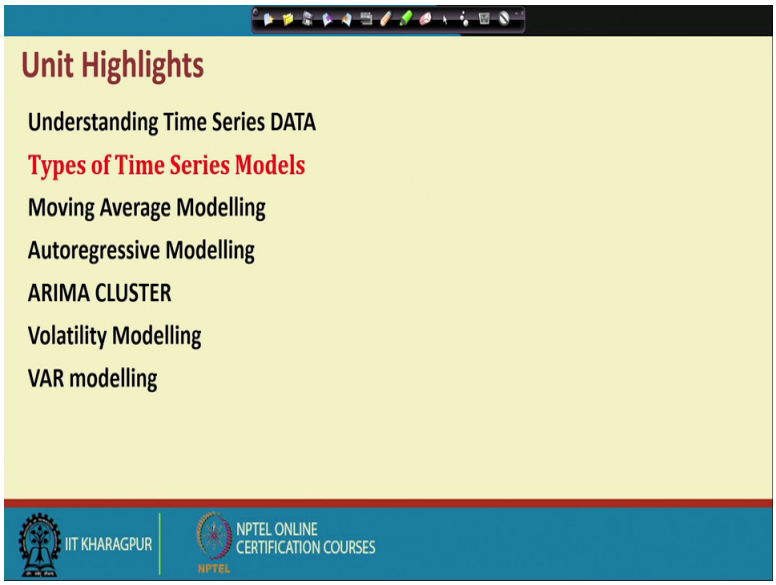
So, let us see what is the kind of you know time series structure and the kind of you know time series modelling. And what are the you know kind of you know problems through which we can actually use time series modelling to predict the a kind of you know engineering requirement or something like you know a kind of you know profession requirement or something like that.

(Refer Slide Time: 05:21)



So, let us see how is this you know structure.

(Refer Slide Time: 05:25)

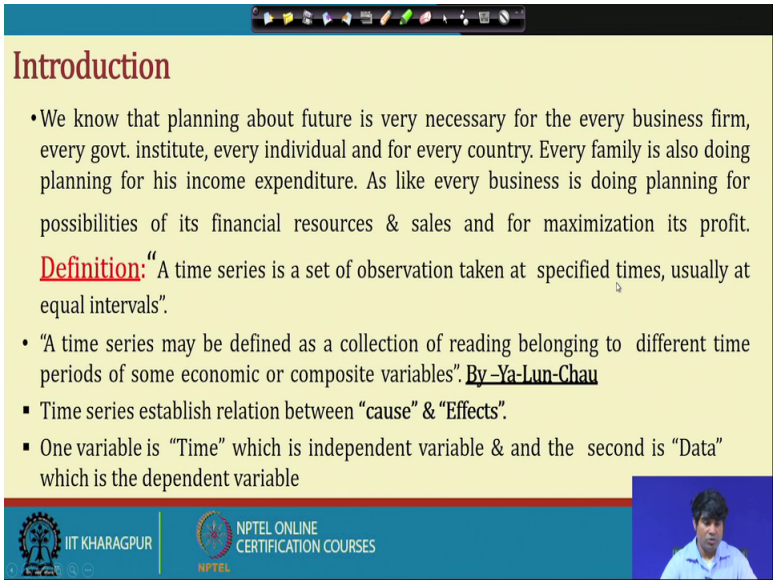


And first structure is the Understanding of the Times Series you know structure. So, in this particular unit we like to understand first time series data and then we like to go through you know various time series model. And then you know starting with the like you know Moving Average Modelling, Autoregressive Modelling then, ARIMA Clusters, Volatility Modelling and VAR modelling. So, we have lots of you know kind of you know components through which we need actually understanding. And we need actually

kind of you know clarity to apply some of the engineering problems and that can be solved through these models.

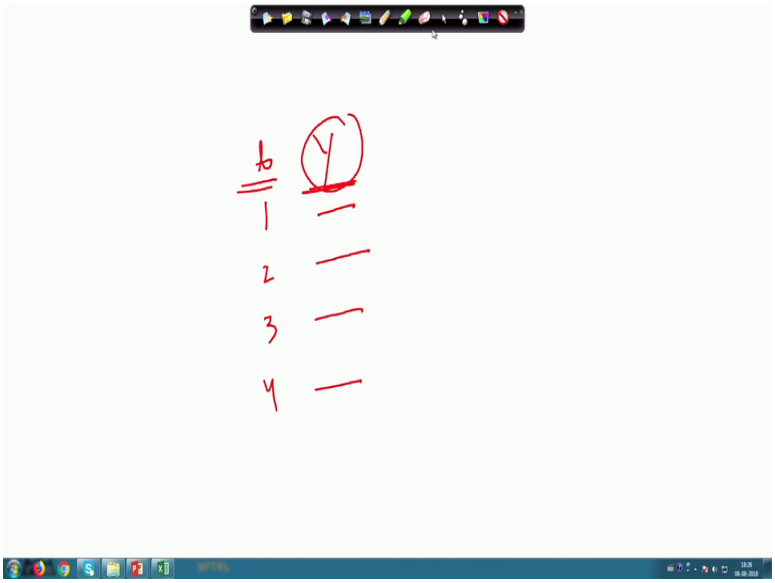
So, basically you know understanding these models are very important because all these you know engineering problems may not be fitted. Because, this particular models are exclusively used for any kind of you know engineering forecasting. So, provided the data structure must be time series you know in character. So, that is how the first requirement of this modelling is to understand the data structure. Once you understand the data structure then we can go for the modelling right.

(Refer Slide Time: 06:38)



So, what will it do here we first know what is exactly the time series. So, time series means the data will be available you know mostly over the time; for instance we will have like this you know we can have like this.

(Refer Slide Time: 06:56)



Let us say you know Y is a variables and you know we have a time specification like this. And this is a variable and this may be a kind of you know information let us say production of a an industry. Then we like to know: what is the production of that industry over the time. So, it maybe annually, it maybe monthly, it may be weekly, it may be kind of you know day wise. So, the so, it will vary like this 1 2 3 4 and so on then against there will be you know production unit like this. So, that means, there is a need of you know consistency.

If it is a you know annual data then the starting and ending must be clear. For instance let us say the production data of a industry since 1991 to 2001 or you know 2018. So, you will get you know in a annual data and that itself gives you know give some kind of you know consistency. And so, you know means the data structure itself give you similar kind of you know structure through which you can do the trend analysis then the kind of you know prediction.

So, compared to cross sectional data you know, time series data has more structure because it maintains consistency while you know cross sectional data there is no such you know consistency. So, all the industry can be you know moved here and there, but in the time series data, data must be in a chronological orders. If it is monthly then for you know starting of the month, and ending of the month, starting of the week, ending of the week, starting of the day, ending of the day, starting of the year, ending of the year. So,

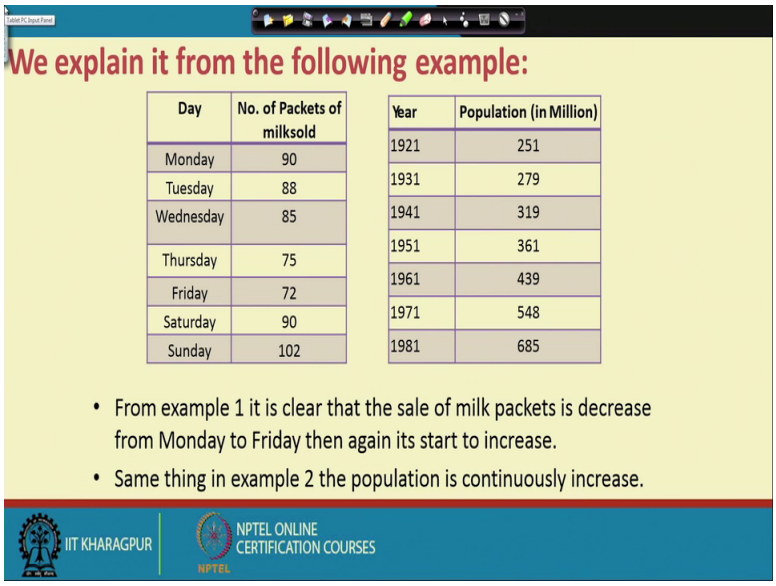
these all are very essential and the data structure must be fitted on that angle only. Then it you know it clarifies the structure of the time series data. Otherwise you know the beauty of you know time series will lead simply actually lose its importance. So, what we can do so you we first note what is this time series structure then we will go through different components and different models which is connected to time series you know data and then the kind of you know engineering problems.

So, any engineering problems and when we actually exactly identify the problem and the kind of you know variable structure. So, we like to see whether the data can be available in a time series format. If the data will be available in time series format then by default we can use time series modelling to predict that engineering requirement and the kind of you know industry requirement. So, let us see how is this particular you know structure and ultimately to know this once we must actually see the a time series structure. So, this is what actually you know representation of time series we can you know define in different way.

But what we have already actually highlighted it is like you know a time series a set of observation taken at specified times usually at you know equal intervals for instance you know we can define in other way time series may be defined as a collection of you know reading belonging to different time period. Or so, periods of you know some economic or you know composite variables or some kind of you know engineering variables. And basically time series establish relationship between cause and effects

There is a high chance that you know the information of a particular time period will be closely linked to the previous time periods. And again it may be closely linked to the you know post time period. So that is why there is a slightly you know cause and effect relationship within the data structure. And that is what the you know you can say the beauty of the times series structure.

(Refer Slide Time: 10:54)



So, let us see how is this you know time series data altogether. So, now, if you look here in this table; and here there is device informations and then over the you know day so we have actually you know packets of you know milk sold. So, starting with you Monday, Tuesday, then Wednesday, then we have here we have here Thursday, Friday, Saturday, and Sunday. So, this is actually just you know 7 days you know data structure and by default starting with the Monday to Sunday. So, the that itself maintains you know some kind of you know consistency.

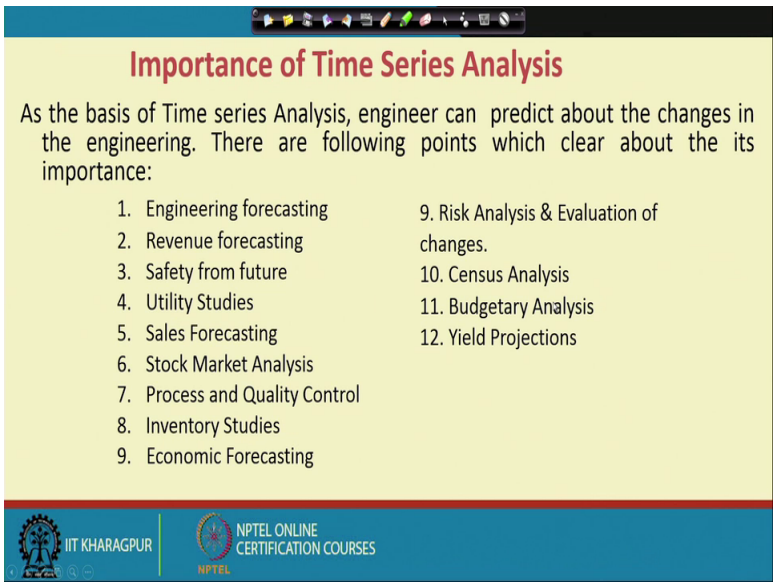
So; that means, one way you can say that you know there is a some cause and effective relationship for instance if you start with a let us sat Thursday data you see here a 75, then you know we expect that you know this 75 structure has a connection with you know previous data structure say let us say 85 or somehow 88 or 90. Either exactly you know that to with 85 or with the help of all these three. Again this 72 item may be connected with you know 75 so; that means, there is a some link actually.

Of course, you know it is not a kind of you know increasing order or a kind of you know decreasing orders. So, here some ups and downs happening and it is not the ups and downs because of you know the device impact. It may it may happens that you know it may happen that you know there are some other factors as a result the data's you know structure is a changing.

For instance the packets of milk sold day wise of course, you know the next day a selling; obviously, it depends upon the previous selling. But ultimately if something going wrong in the market or due to something external pursues like you know the weather is not so good or you know, there is there is some kind of you know crises, or some kind of you know something external attacks, or something like that. Then by default it will affect the market and then it will affect the selling process. So, this is how the first structure is the device information and that itself gives the indication about the time series data. And similarly this is the yearly data you know that to decadal data starting with the 21, 1921, 1931 that is with respect to population.

So; that means, you know it is also time series data where the particular interval is on the basis of you know 10. So, when we will start with you know let us say 1921, 1922, 1923 then the interval will be in between 1 only. But the here we have the interval 10. So, that itself also you know justifies the times structure. It is not that you know one day interval it may be you know in between you know means we need actually what we can called as you know balance you know gap. If it is actually in between balance gap balance interval then that itself will you know represent the time structures and as a result we can actually use actually time series modelling to produce as per the particular you know engineering requirement.

(Refer Slide Time: 14:11)



So, for the importance of time series modelling is concerned. So, we have actually you multiple use; that means, this itself is a multi dimensional character. And mostly a we can you know having the time series you know data an engineer can predict about the changes in the engineering or an economist can predict the business environment something like that. So; that means, some of the problems connecting to time series modellings will be like this.

Engineering forecasting, revenue forecasting, safety forecasting, utility forecasting, then sales forecasting, stock market forecasting, quality control forecasting, inventory studies, economic forecasting, and then budgetary analysis, risk analysis; so that means, multiple you know areas where you know we can frequently use this you know time series data to do the predictions and as per the particular you know engineering requirement. So that means, you know briefly I have just highlight some of the areas where you know time series modelling are very frequent. Because you know these areas the information availability or what we can called as you know data or actually in a time series format.

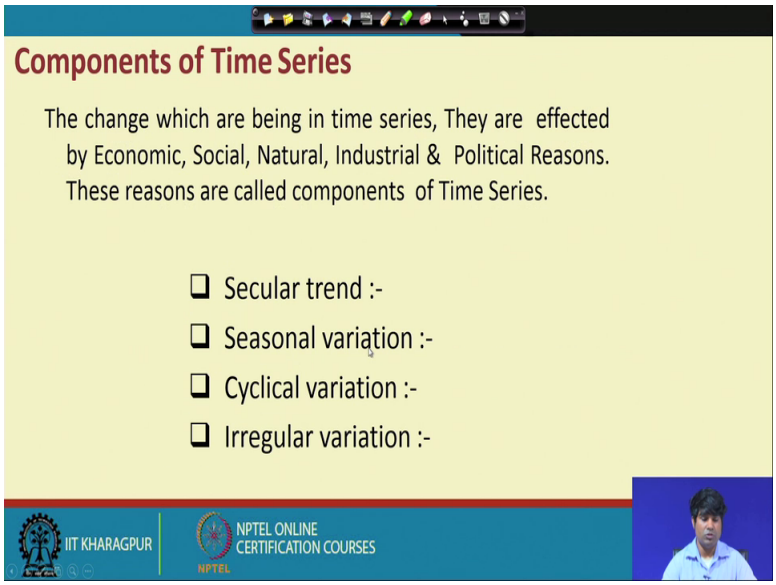
So, when a problem is having data in a time series format; then there is a high chance that you know time series modelling can be applied. Of course, you know that particular problems you know needs that requirement on. If that problem you know not required at all for the kind of you know time series modelling, then you can skip. But there is a high chance if the particular variable information is in a time series format then you can easily apply the time series data. And if not then you can again you know through cross sectional and time series you can use the panel data.

But you know time series has you know unique kind of you know identity and unique kind of you know structuring through which we can you know you do the modelling. So; that means, with a particular variable we can create a multiple variables by introducing the a lag variables which is the that is you can say you know beauty of this you know time series structures. So; that means, if you have a variable say y t you can create y t minus 1, y t minus 2, just you know same data. It will be you know copied and pasted you know one point you know behind. Then you create you know data structure with respect to y t, y t minus 1, and y t minus 2.

So; that means, that is that is the process which we can actually create with the help of only time series data otherwise you cannot just copy paste and create you know lag

variable like t minus 1 and t minus 2. If the data maintains consistency then you can create a lag variable if the data does not maintain the consistency. So, you cannot you know create a you know time series you know the variable. So, you know or the kind of you know lag variable and that to the use of you know time series modelling.

(Refer Slide Time: 17:19)



In the case of you know time series modelling so we have actually understanding of the data which we have already discussed in details. And basically in the time series we have you know specific you know components. The first component will be secular trend; that means, you know the process how the data is you know structured or you know maintained the consistency. Then there is a seasonal variation, cyclical variations, and irregular variations. So; that means, since it is actually recorded and you know collected over the times maybe with annually or with a kind of you know day wise or week wise or monthly and then quarterly.

But it has a kind of you know pattern and when we have the data so that particular pattern may be secular in nature and seasonal in natures cyclical in natures and irregular in natures. So; that means, there is a variations means say when you know moved from one point to another point. So, there is a variation if there is no variation then by default we cannot use this data for any kind of you know in the modelling that too engineering econometrics. So; obviously, the data will have some kind of you know variation. Now the issue is in the time series data when there is a variations that variation may be lead to

secular trend maybe seasonal variations, maybe cyclical variations, and maybe irregular variations.

(Refer Slide Time: 18:54)

So, we will see how is the kind of you know structure and then we will discuss as per the particular you know requirement. The first one is the understanding of the secular trend if you say that you know time series data follows a secular trend so what is exactly that. So that means, it is the increase and decrease in the moments of time series and that is what the term is called as you know secular trend. The time series data may show a upward trend or downward trend for a for a period of you know years and this may maybe due to you know multiple factors.

Like you know maybe due to population increase may be due to technological advancement or due to in a large shift of you know consumers demands or something like you know high growth rate of you know industrial base or the kind of you know money market. That means, you know money flow of money supply itself will generate some kind of you know. So, there is a lots of in a interdependency and connections through which you know the time series data have been ups and downs. If there is no ups and downs then by default what we have what we have already said that you know this particular you know information or this particular data cannot be used for you know modelling.

So, we need some variations, but ultimately we like to know how these variations you know happening. For example, a population increases over a period of time, then by default price increased over a period of years production of goods on a capital market of a country increases. So that means, if something happens then the other things by default will happens. So, that is how we are actually very much interested to know the kind of you know trend structure and how it affect reference actually. Obviously, when there is a kind of you know trend some factor may be responsible for that that is how it brings a kind of you know trend. If that is not happening then you know it may be a kind of you know more or less same for instance, if there is no change of populations, or there is no change of you know income, and there is no technological advancement.

So, whatever demand in the market s o that will be remain there ok. So, the by default so there will be no change. So, it will be it will be simply a kind of you know stagnant. So, if there is a stagnant and the variation of data is you know almost all 0 then; that means, technically we are not in a position to do the modelling and in a reality that cannot happens. Even if you know technological progress or population change so these are all actually long term impact, but still in the short run there may be some kind of you know change. And that may not be very large one it in a small interval, but, but obviously so, you will find some variation will be happening.

You got to any market in the world or any engineering you know area. So, whatever you know informations are recorded you know for a particular variable most of the instances most of the instances you know there is a kind of you know variation. And also there are instances where you know information base is more or less constant. For instance suppose there is kind of you know engineering experiment and the means very specifically let us say it is a kind of you know production process. And the entire production process is going through a kind of you know machinery process only.

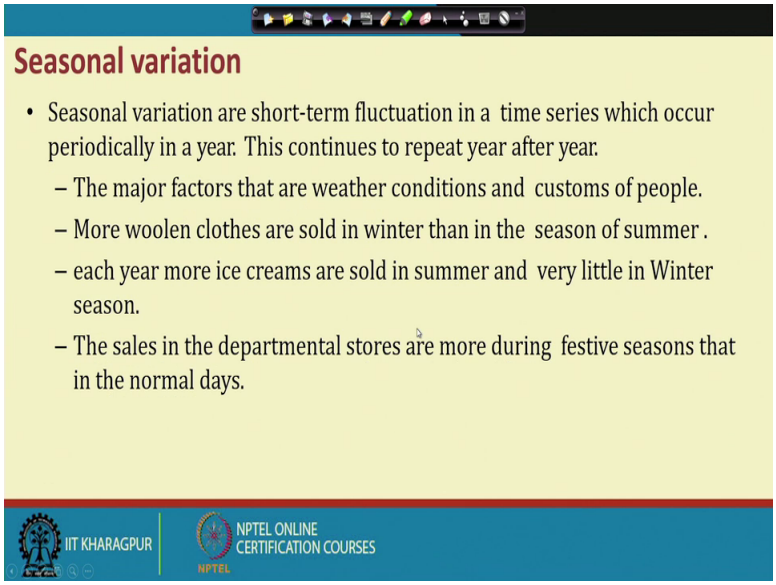
So, now if the machine is and everything is and obviously, the production quantity maybe same over the time but, after a particular point of time it may not actually you know true, the thing is that you know machine itself has as you know some kind of you know depreciation. So, when machine becomes you know old and old by default the particular production capacity may not maintain. So, there is a you know some kind of you know deviations. So; that means, up to a particular point of time there may be you know the data structure or production structure may be stagnant. But after a particular

point of time and definitely there will be variations; that means, technically a you know a in some instances the change maybe takes time and sometimes the change may take actually very less time.

So, but there will be definitely you know ups and downs because nobody can you know maintain a particular you know stagnant of you know in a kind of you know infinite framework. So, it will be definitely have a change. So, means in the real life scenario you know talk about any variable the moment of that variables cannot be actually you know simply say you know constant over the time. So, what we say that you know constant means it maintains like you know horizontally simply ok. So, that is 10 10 10 like this 20 20 20 like this. So, that is the stagnant is all about in a kind of you know time series format.

And in reality that is very difficult for a industry for a individual or a for a country to maintain that kind of you know things It definitely there will be up ands downs even if you maintain you have a lot lots of you know safety measures to have this kind of you know maintenance. But stills there are certain external factors through which you know this change may happen. So, that is why we have a different forms of you know time series understanding starting with this secular trend.

(Refer Slide Time: 24:31)



Then against we will go through the other components like you know seasonal variations then cyclical variations and then irregular variations. So, now, like the trends secular

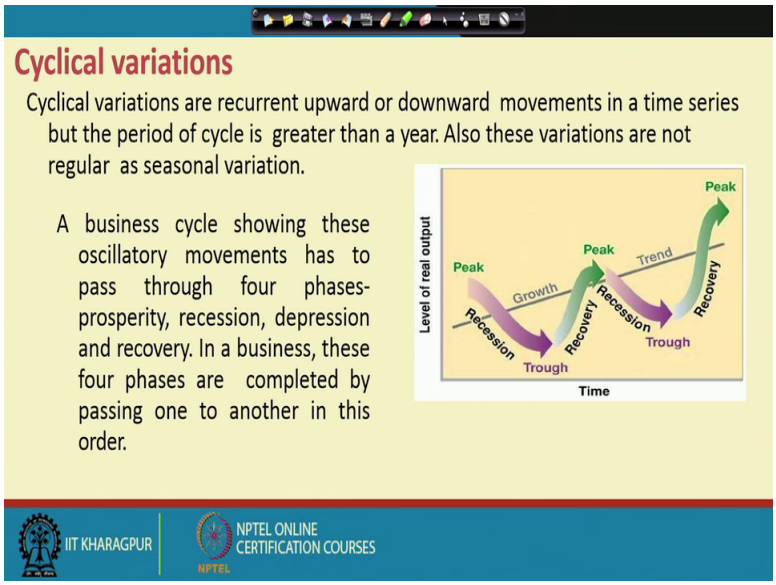
trends so the variations may be seasonal in character. So, that is again slightly in a longer intervals, but the seasonal maybe 3 months interval or 4 months interval 6 months interval like this. That means, here the change happening over this you know in seasons only; for instance let us say the production of you know woollen cloths.

So; obviously, if it is actually summers then the trend maybe constant over a time. So, the moment it will be going to winter then; obviously, there will be high demand and then by defaults you know sales of these or production of these woollen clothes will start you know increasing. So, against it will increasing increasing. So, when the means the moment winter will start you know increasing it is volume. Then it will again reach a peak and again the winter will start down, down, down and then the production will be down, down, down and again it may be maintained like you know the original ones.

So; that means, you know this kind of you know change you know is with respect to you know seasonal variations. So, like you know each year more ice creams are sold in summer. So; that means, mostly you know the time series data which is a having variations with respect to this you know difference you know seasons. So, there are you know certain engineering areas where. So, the information contents you know information contents are you know changing or you know may change you know with respect to different point of time or that too with you know a kind of you know seasonality structures. It may be with respect to 2 months interval 3 months interval 4 months interval but ultimate they that does not you know matters

Like you we have discussed earlier they the day wise variations and then the decadal variations. Now here what is happening day wise variation then the you know month seasonality means it is a month wise variations. Maybe the interval with 2 to 3 months or 4 months like that. But the change will be happening or you know ups and downs happening after a certain you know a month's only and that too for the structure of the seasonal variation.

(Refer Slide Time: 26:59)



So, likewise we have a cyclical variations and that is mostly because of you know business cycle operations. So; that means, you know whatever you know things are there in the real life scenario. It ultimately depends upon you know economic conditions starting with any engineering productions and the sales forecasting or something kind of you know revenue structure whatever maybe you even talk you know talk there will be a kind of you know change and that change is a kind of you know cyclical.

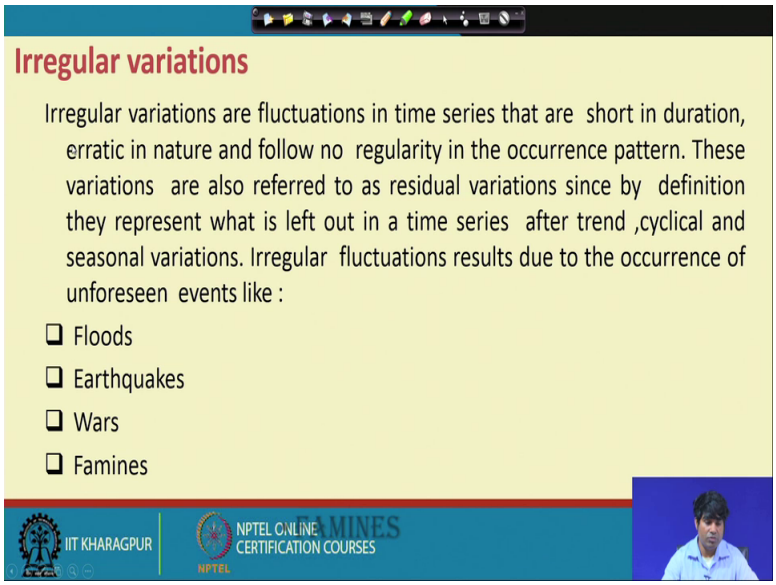
And it is mostly because of you know the operation of you know business cycle, if you talk about business cycles then there are you know two different phases altogether. So, like this so if you start with you know up and then move into down. So; that means, when at the movement will be up to down. So, it will not actually continuously down down and that too not the case of you know infinite. So, it will be down and after a particular point of time it will start increasing. Again when it will start increasing then it will not again continuously over that you know in a infinite kind of you know structure, again after a particular point of time it will start you know get down.

So; that means, this ups and downs are you know very phenomenal in the life scenario. And some of the variables data information's will be because of you know cyclical variation only. For instance if something you know good environment, business environment, then the sales will good. If the business environment is down then sales will lay down so that is what the kind of you know. So; that means, here the information

about a particular variable exclusively depends upon you know cyclical patterns that too the business mostly because of you know business cycles.

So, upswing, down swing, normal then the you know the data structure will changed frequently. So; that means, this is the regions through which you know there is the variation of the data. And again every point of the every point in the data we will have a slight connection to the previous data and then again the a you know upwards data. So, that is why you know there is a some kind of you know linkage is there. And that is what the beauty of the time series data because it is very consistent with you know time framework.

(Refer Slide Time: 29:24)

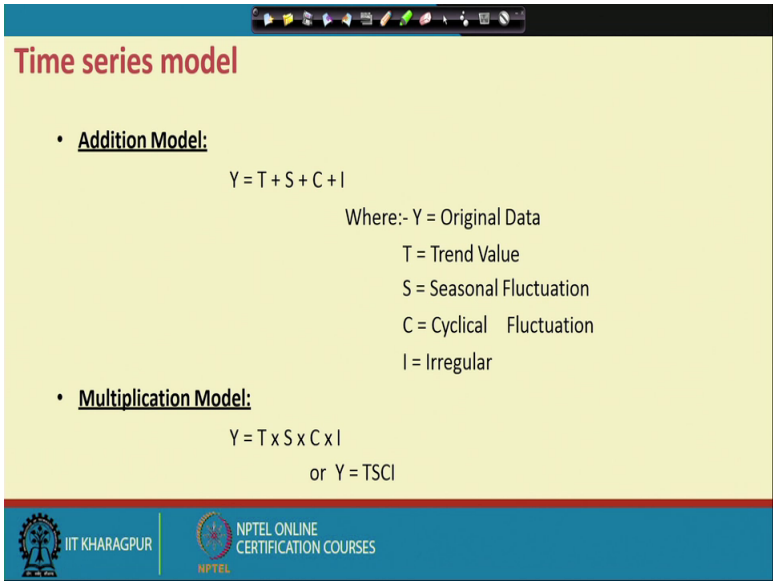


So, then likewise you know there is another case called as you know irregular variation. And irregular variations are fluctuations in time series that are short in duration and erratic in nature and follow no irregularity in the occurrence patterns. And since it is actually very irregulars like, you know let us say seasonality it is a regular every year it will be come into this particular you know structure. But these are all you know variations happening in the time series you know very you know occasionally you can say; floods, earthquakes, wars, like these are you know classic example crises. So, if it happens then you know they there will be very you know very huge decrease or you know huge increase. These are all negatives, but you know sometimes you know there are is you know good policy or something kind of a good technology

So, then you know there will be you know drastic change happening. So, irregular variations it means you can it means technology change we will not be very frequent. So, it will be you know coming over the time only so like you know floods and the earthquakes. So, as a result so there will be in a change positive, but it in a long interval and very occasionally again. So, like you know floods it will be down the structure data structure, but again it is very occasional it may it may not depending upon the external factors.

So, that means, technically we are still you know means we are in a position to say that you know the time series data how the variation is coming exactly right. That is the way we are you know discussing this irregular pattern.

(Refer Slide Time: 31:01)



And like this you know we have actually discussed some of the variations starting with the circular trend, seasonal variations, cyclical variations and then irregular variations. So, these are the a various regions through which you know the time series data you know behaves ups and downs over the time. It cannot be stagnant it cannot be constant so there is a movement. So, these are the various you know factors which starting with you know seasonal trend and then you know then secular trends, seasonal variations then the kind of you know cyclic variations and then irregular variations.

So, that that itself justifies means logically it justifies that you know yes; there are ways you know the data will be you know a having ups and downs over the time. And that

itself gives some kind of you know clue to the future predictions as per the particular you know engineering requirement or the kind of you know industry requirement.

And with this you know we will close this you know lectures. And in the next lectures we will discuss about the, you know time series models and how these data can be used to predict you know engineering requirement as per the availability of you know time series data.

Thank you, very much. Have a nice day.