Business Analytics for Management Decision
Prof. Rudra P Pradhan
Vinod Gupta School of Management
Indian Institute of Technology, Kharagpur

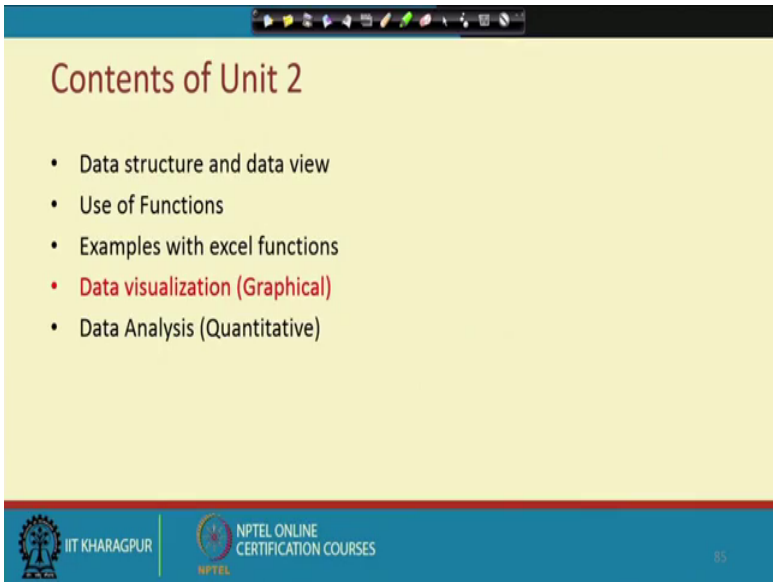
Lecture – 09
Exploring Data and Analytics on Spreadsheets (Contd.)

Hello everybody, this is Rudra Pradhan here. Welcome to Business Analytics for Management Decision and we are still in second unit and the topic of this discussion is exploring data and analytics on spreadsheet.

So, we have already discussed couple of lectures on this unit and the basic idea about this particular unit structure is to understand the data then to know more details about the kind of data by a visualization process and little bit you know basics from the a data structures. So, the idea is that you know once you identify a particular problem and before you start the investigations, one of the particular requirements is the data which we have already highlighted in many different ways. So, here is the particular structure is you know just enter the data in the excel sheet then you try to extract or you know get some kind of you know insights and that that insights may help you lot to understand the problems or to get some kind of you know inference to solve the problems or in other way you can also you know in a position to use the particular data for you know particular techniques.

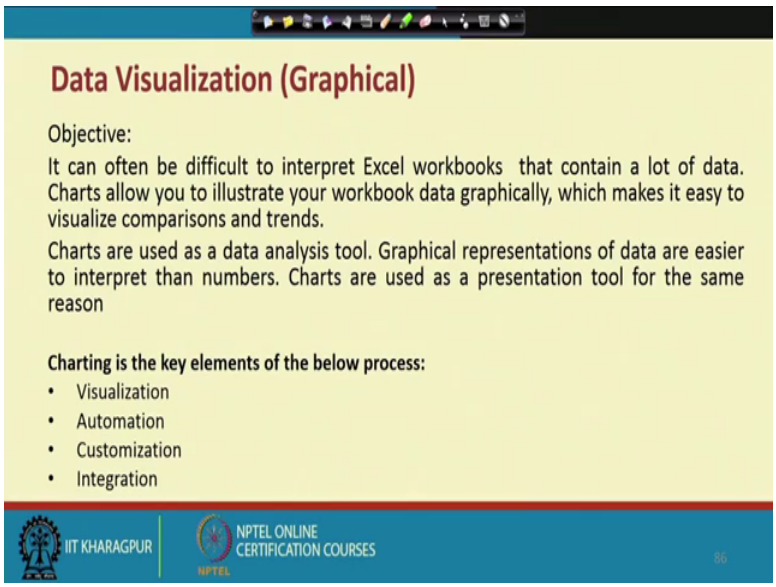
So, the idea is that you know just enter the data in the excel sheet and try to get some kind of insights. So, the insights may be in a quantitative angles or it may be in a kind of qualitative angles by using some kind of you know graphical visualization process. So, in this lecture typically we like to understand the data in more attractive way by using some kind of graphical structures or you know data graphical data visualization process.

(Refer Slide Time: 02:16)



So, the topic of discussion of today's is data visualizations and that through some kind of you know graphs.

(Refer Slide Time: 02:24)



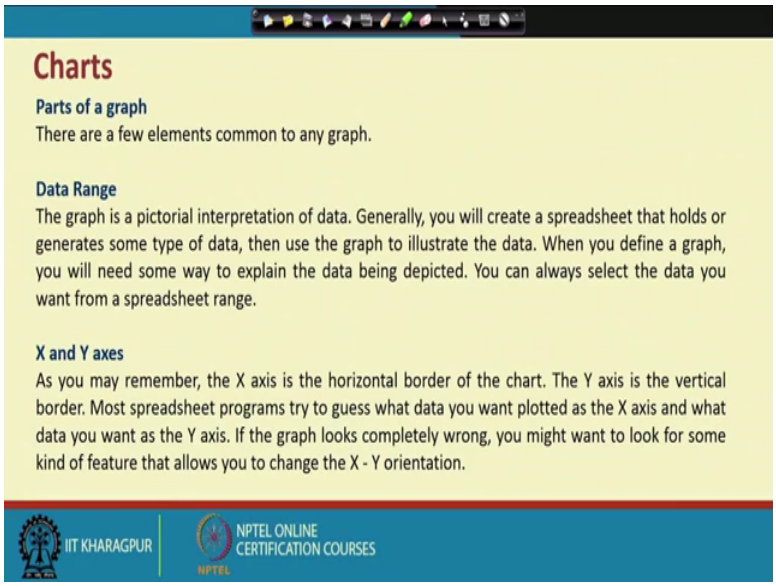
So, it is a kind of you know against you know understanding of the data and the idea is that you know sometimes when you enter the data. So, particularly when we have a big data set, it is very difficult to understand the particular structure of a particular you know variables or in a particular problem. So, entering data to get some kind of you know basics. So, that is the one part of the story, but the other part of the stories you know you

should know you know one in to another and how it how the data is all about, what is the kind of you know flow, what kind of you know structures.

So, the requirement is that you know of course, you know it will give you more insights by knowing you know the outliers problems the kind of you know difference or the kind of you know maximum the kind of you know minimum; that means, of course, in the excel sheet which we have already discussed earlier. So, having data you can actually try to get what is the maximum number, minimum number, average, sum all these things you can count with a kind of you know if and by using if functions something like that, but still. So, it will be actually a specific kind of requirement, but when you use these data and you do some kind of plotting and then and that actually the idea is here the graphical inspections, so by doing the kind of no plotting. So, it will give you much more insights than what we have already discussed in a quantity format right. So, it will give you some kind of snap shots from one in to another end. So, and that itself will give you some kind of you know basic setups through which you can you know pick up a particular technique and then solve for this particular you know business problem. So, this will give you clear cut you know idea what is the kind of you know structures of these data with respect to a particular variables.

So, now, I let me little bit you know give you hint about here the idea. So, you can do visualization and then some automation process, customizations and then integration these are the four differently items we can actually target in the kind of you know graphical representations or you know data visualization process by graphs.

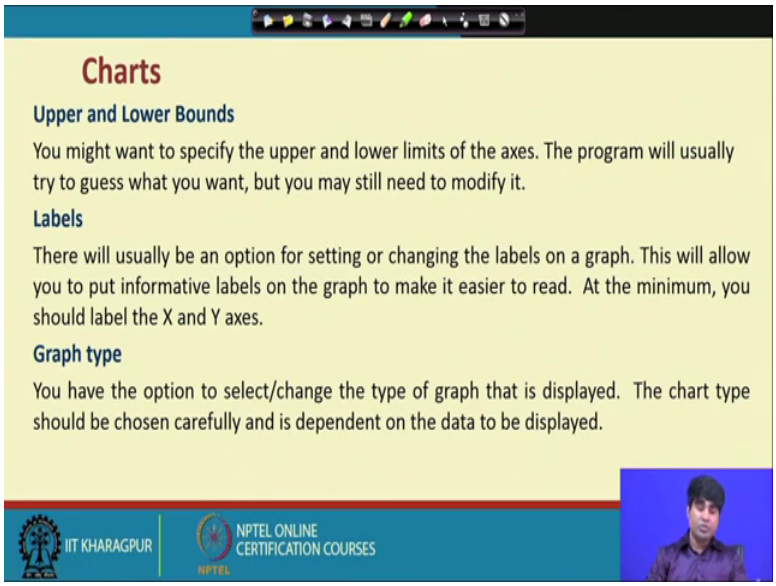
(Refer Slide Time: 05:02)



So, basically, we have a couple of you know graphs or we can call as you know a chart. So, various you know graphs and charts are there and you can use actually you know one after another you know to know more about the particular means insights and you have the data, but you know when you plot in a different way or different structure and in shapes then it will get it or you will get more insight on this particular setup.

So, that is the beauty of this kind of you know graph structures. So, in the requirement in the chart or you know graph is the just to check the range and then the kind of you know as usual graphical process you have to see the x axis the structures and y axis structure then accordingly after doing the kind of you know requirement automatically it will give you some kind of better visualizations and that too you can understand the data you know a more attractive way and you can also easily get some kind of better insights what we have not done in the kind of quantitative framework.

(Refer Slide Time: 06:17)



So, the charts you may have actually you can fix the upper bound and lower bounds you can do the labeling. So, all these things will be there in the particular process.

(Refer Slide Time: 06:34)

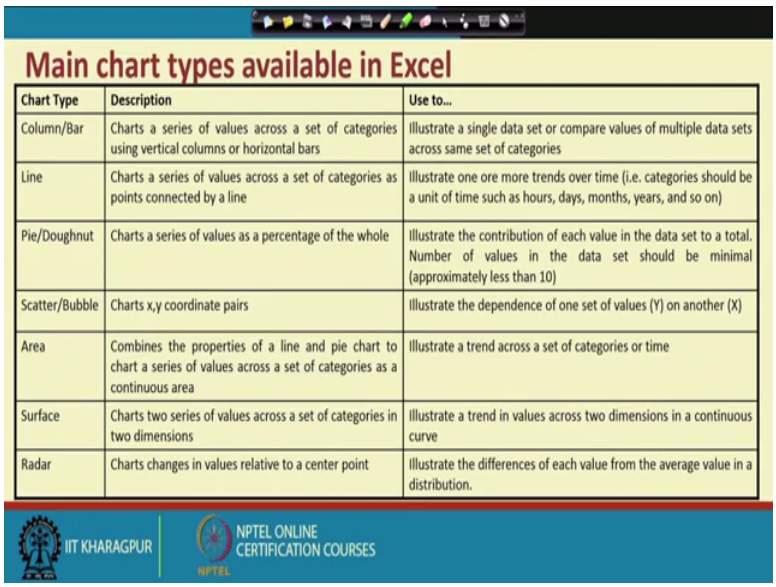


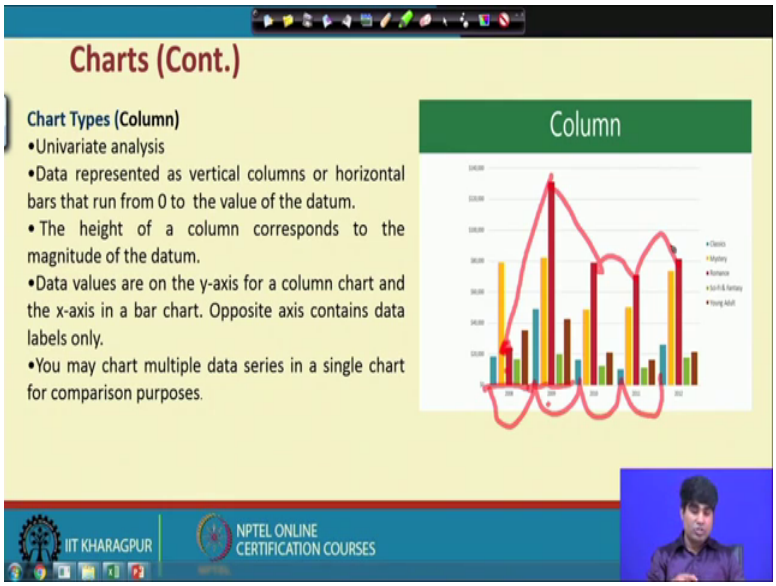
Chart Type	Description	Use to...
Column/Bar	Charts a series of values across a set of categories using vertical columns or horizontal bars	Illustrate a single data set or compare values of multiple data sets across same set of categories
Line	Charts a series of values across a set of categories as points connected by a line	Illustrate one ore more trends over time (i.e. categories should be a unit of time such as hours, days, months, years, and so on)
Pie/Doughnut	Charts a series of values as a percentage of the whole	Illustrate the contribution of each value in the data set to a total. Number of values in the data set should be minimal (approximately less than 10)
Scatter/Bubble	Charts x,y coordinate pairs	Illustrate the dependence of one set of values (Y) on another (X)
Area	Combines the properties of a line and pie chart to chart a series of values across a set of categories as a continuous area	Illustrate a trend across a set of categories or time
Surface	Charts two series of values across a set of categories in two dimensions	Illustrate a trend in values across two dimensions in a continuous curve
Radar	Charts changes in values relative to a center point	Illustrate the differences of each value from the average value in a distribution.

So, let me give you some kind of examples you know what are the types of examples in that kind of you know data visualization in the graphic mode and then you can use a particular graphic modes and then to get more kind of you know insights. So, actually like you know we have a plenty of you know analytics tools.

So, here in the graphical visualization process you have been you know plenty of you know graphs and the same data can be plotted or you know by using different graphs and the whole idea is exactly every time to know the exact structure of the data to get some kind of insights. So, ultimately you have to understand the data very perfectly. So, if you understand the data perfectly and the plus minus in the data then you may be in a process to you know use the data as per a particular you know requirements. Sometimes some data may not be in good structures, but after knowing the particular you know visualization process. So, then you can try to bring more structured data or do some kind of transformations do some kind of you know manipulations by using some kind of you know function or something like that and so that you know the data can be in a proper format or you know in a proper structure and then you can use the particular data for the further analysis and as per you know your requirement.

Until unless you do the data processing properly, this will not give you better results as per your particular requirement. And so far as a graphical inspection is concerned we have a plenty of options starting with you know bar diagrams, line, line charts, then pie diagrams and then scatter plotting, area graphs, surface graphs, radar graphs. So, many you know varieties are there you know same data can be actually visualized through column diagrams or you know bar diagrams, line diagrams, then you know pie charts, then scatter you know scatter plotting, then area plottings surface plottings and radar plotting. So, I will just briefing you every case, then I will take you a particular examples and connect you how the same data set can be plotted differently and you will get some kind of better insights compared to the previous discussions through a particular tools right.

(Refer Slide Time: 08:45)

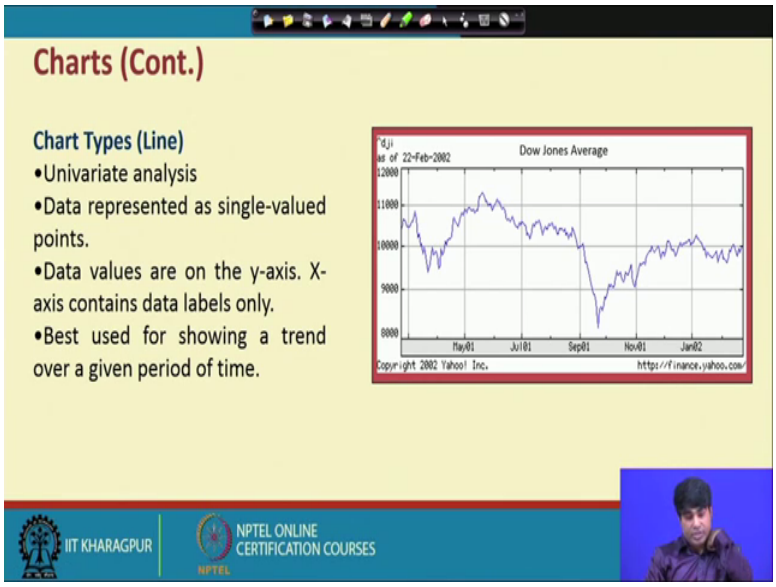


Here in the first you know column you know, column chart this is the classic example of you know column charts and here. So, obviously, actually you must have some kind of you know data with respect to some either in a time framework or in the kind of cross sectional framework then you know; that means, you must have you know excess indication then y axis indications then automatically if you actually connect them automatically it will give you some kind of you know plotting.

For instance in this case, this plotting will give you actually. So, there are actually for you know how see here just let me give you some highlights here and we have actually 5 different components and that will be indicated with different colors. So, 1 2 3 4 5, this is for you know a particular year set 2008, again the same for you know 5 category in 2009, then 2010, 2011, 2012; that means, actually it gives you know clear cut you know you know indications of what is happening actually. Let us say you know classics right. So, the classics the trends of classics you know 2008, then 2009, then 2010, then 2011, then 2012. Similarly in the case of you know romance let us say this is actually is this third case. So, what is the trend here in 2008 to then you know 2009, then you know again 2010, then 2011, then 2012; that means, actually see here. So, it clears gives you it clearly gives you know better insights to know actually at a particular point of time how these items are you know varying with respect to different, different points of time. So, so this is actually one of the you know kind of you know beauty of you know graphical inspection and that too through excel sheet right.

Similarly we have actually you know other kind of units diagrams or charts.

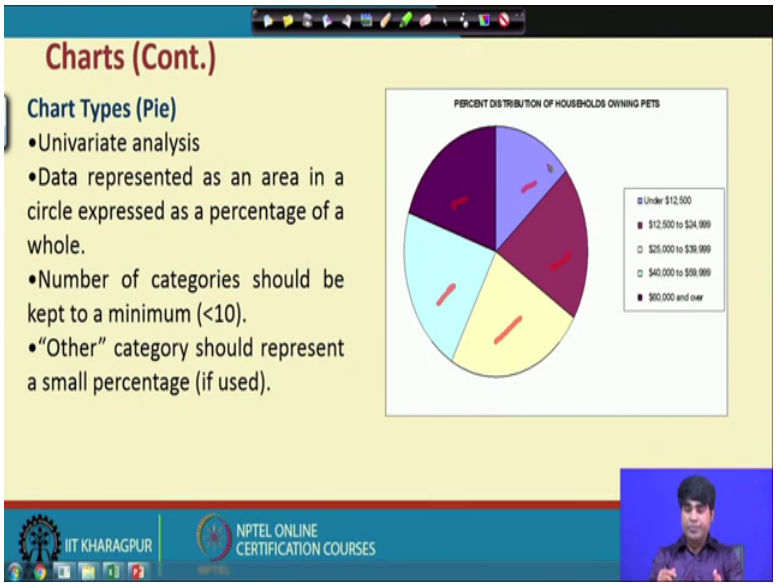
(Refer Slide Time: 11:23)



So, this is a kind of line charts and the classic example is here again. So, we have actually a time series data, monthly time series data and then we like to actually plotting some kind of you know financial figures. So, you see here so; that means, against you know time. So, you have actually information and then against you just you know give the command and automatically it will give you some kind of structure you know with respect to that particular you know variables.

So, here, we can actually just you can just get to know the trend of this particular you know series and whether it is a kind of you know increasing kind of you know situations or decreasing kind of a situation, whether it is a kind of you know linear movement or non-linear movement. So, many things we can actually easily actually predict from this particular you know inspections.

(Refer Slide Time: 12:22)

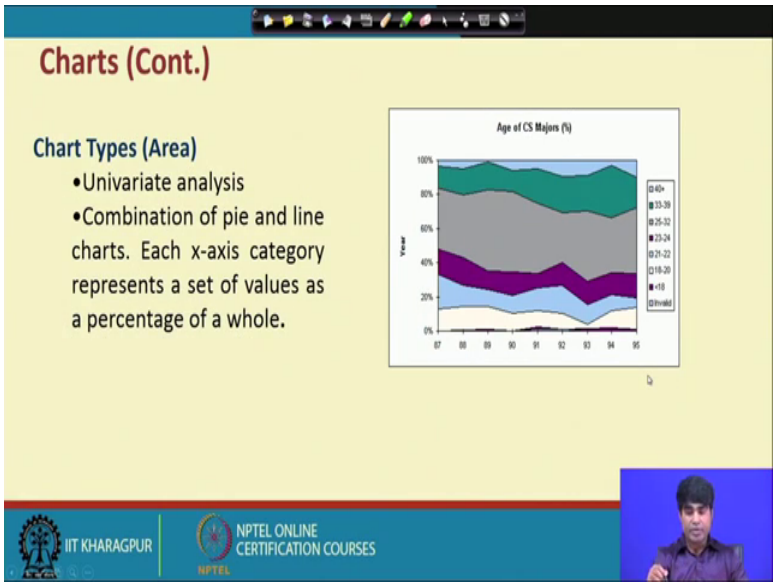


So, likewise there is another kind of you know types called as a pie diagrams and this you know here actually five different divisions. So, this will give you better you know indications you know how it, what is the kind of percentage contributions if one particular range to another particular range. And in fact, it is a kind it is all these are you know examples of you know univariate structure and means this is a kind of you know single variables and the variation is happening with respect to different ranges right. And how is this particular you know weightage with respect to different you know range, this pie diagram can help you a lot to understand the particular structure. This clearly you know visualize is something that and by default you will get some kind of insights out of this particular structure structures.

So, you see here is just you know our see here is the kind of differentiations so; that means, automatically it gives kind of in a comparative analysis and if you go by see numerical you can also calculate and report, but here in a one particular structures it will it will reflect everything in details.

So, that is actually the beauty of this you know graphical inspections right.

(Refer Slide Time: 13:48)

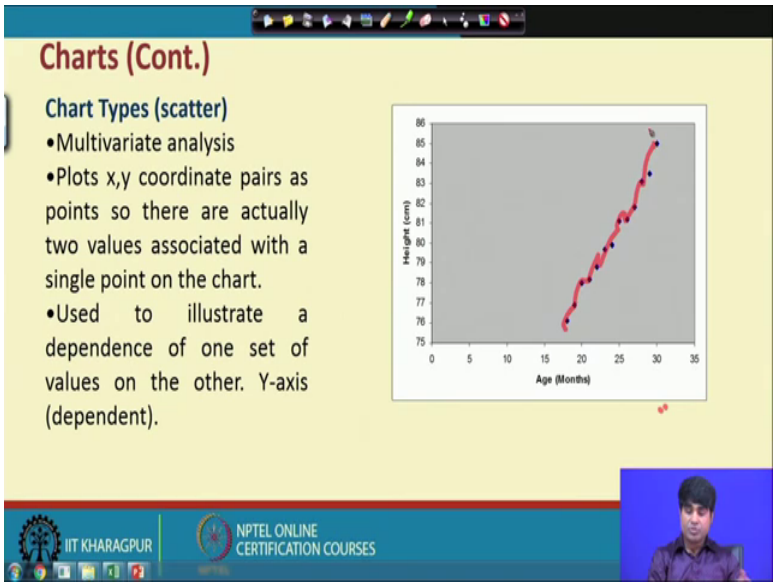


So, likewise you can also have another kind of you know chart that is called as area chart and here same things actually. So, you need to have a x indications and y indications and then you just see how is this kind of you know movement. So, this is actually again on the basis of you know intervals and it will give you some kind of you know differentiations with respect to different intervals. Like you know, previous examples you know pie diagrams this is all the similar kind of you know situation and here you see here, the kind of you know structures means say all right.

So, see here is the kind of examples it is something going wrong is (Refer Time: 14:37). So, we are actually here and I am just highlighting something you know just specifying that this by default you can find out some kind of you know difference a among these particular groups.

So, the difference can be reported in a quantitative way by calculating the mean of the particular structure or range of this particular structure, but still. So, the graphical plotting will give you much more or can say that you know better insights and then the quantity visualizations.

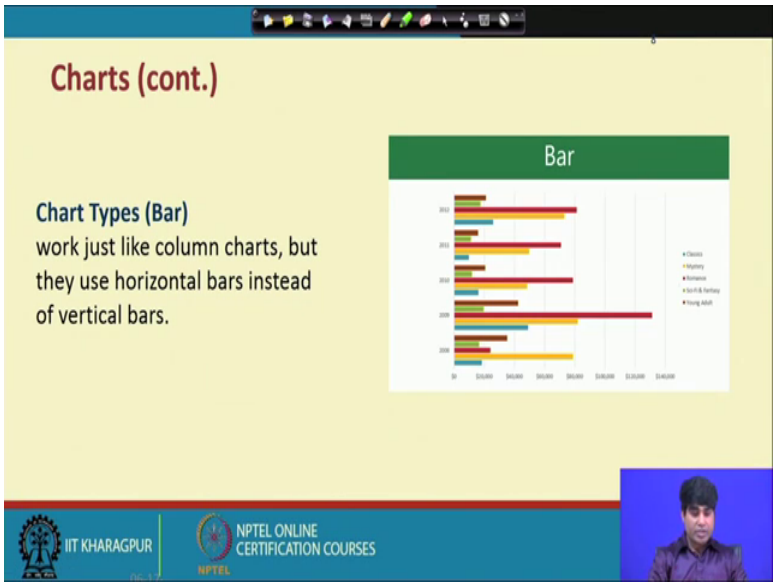
(Refer Slide Time: 15:22)



So, this is another kind of you know chart and that to you know just some kind of you know scattered plotting and here again you know a month wise kind of information and you know the game between age and height and you see that you know the plotting behavior here. So, see here even if you can join these points, you can join these points you know you will get some kind of you know line diagram also.

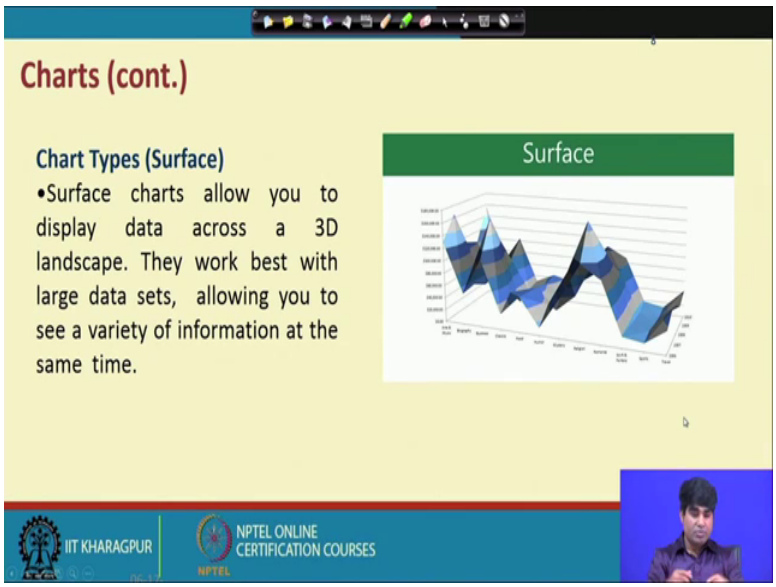
So, this can be also give some kind of you know movement it is a increasing kind of you know trend, but its it gives you know some kind of a better visualization which you can actually have from this with the help of some kind of quantitative tools.

(Refer Slide Time: 16:17)



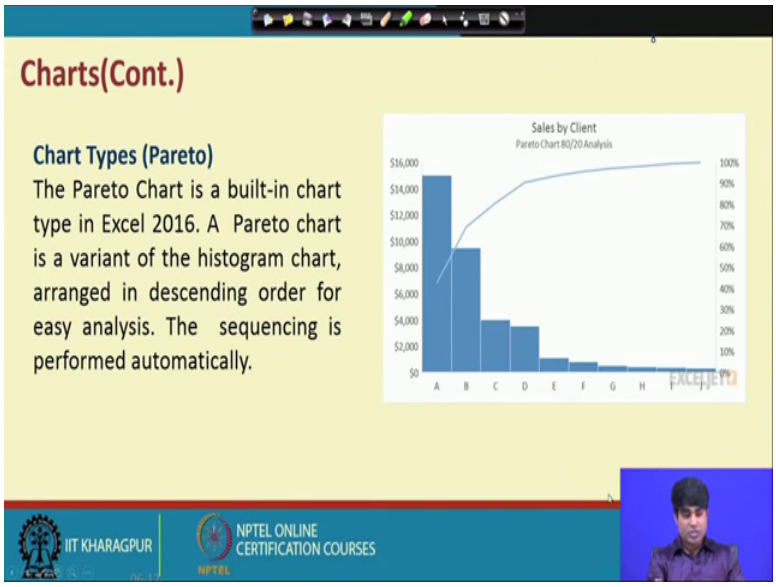
So, like this then you can you can have actually similarly again another kind of you know bar chart. So, previously it was vertical kind of visualization now it is a horizontal kind of visualization just same things, but it will be actually look in a different way altogether the same way, but you are just you not changing the angles and you know seeing how is this particular data structure with respect to these variations. So, the same example classic (Refer Time: 16:48) all these things about you know it is a give you some kind of you know better insights as per the particular requirement.

(Refer Slide Time: 16:58)



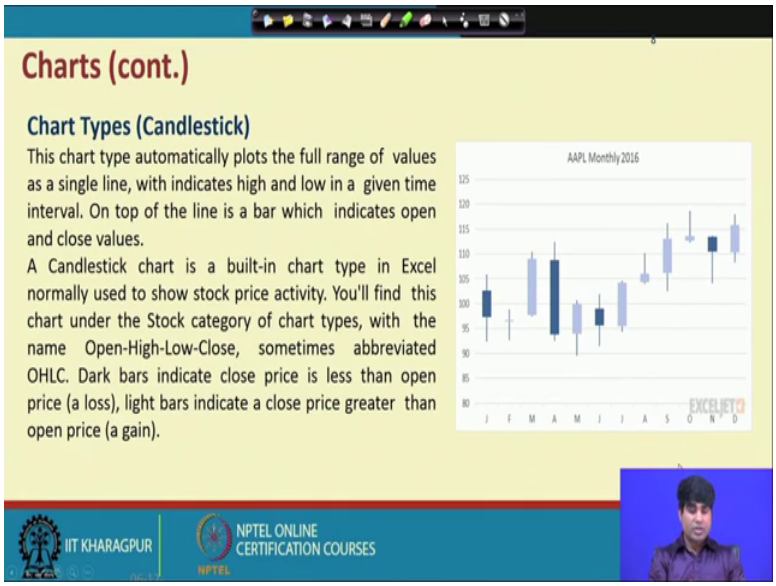
And this is another kind of you know charts called again a surface chart and again this will give you some kind of visualization how it is actually happening over the you know or different kind of situations. So, means over the time how it is actually is changing actually shaped all together. So, this is another kind of visualization.

(Refer Slide Time: 17:21)



And similarly, pareto chart and this gives you know another kind of you know indications to a see how these particular figures are changing over the time. And in fact, you know you can actually see the visualize visualization in a kind of absolute scale and also you can actually see this change structures that is actually in a relatives relative scale and that is actually a again one of the beauty of this you know graphical visualization process.

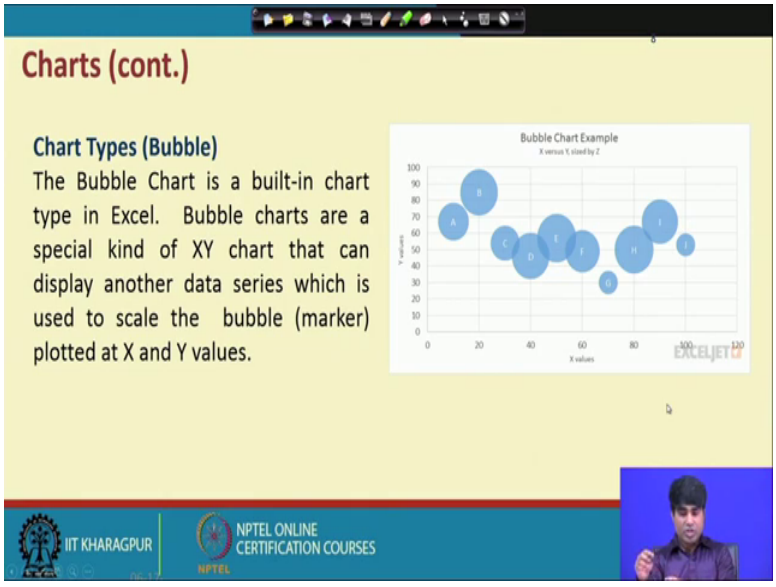
(Refer Slide Time: 17:55)



And this is another kind of you know plotting and this is called as a candlestick and this gives you again some kind of you know indication how is the movement of you know variables over different you know or different point of times.

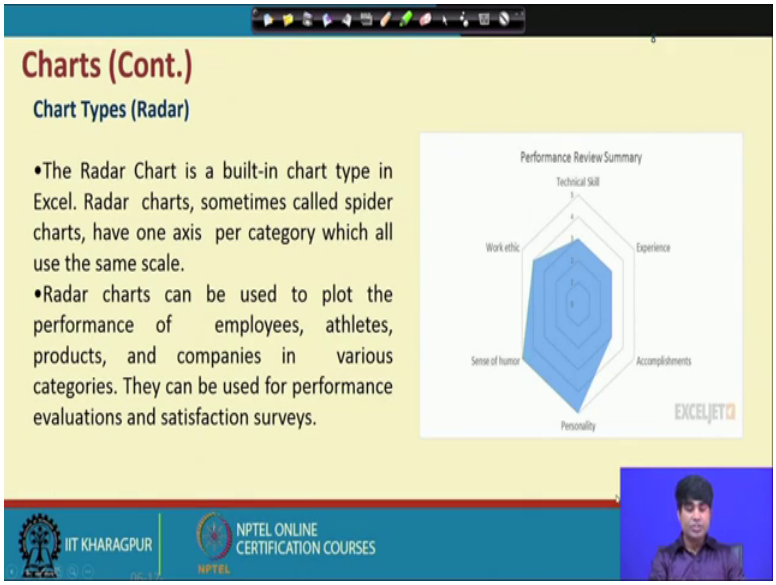
So, you can also you know same line diagram or line plotting you can do, but you know still and this gives some kind of you know better kind of visualization process. But you know so the thing is that you know we just like to know what are the ways you can actually may try to understand or try to visualize the data in a more attractive way.

(Refer Slide Time: 18:27)



So, we have a you know a couple of kind of you know instances or you know couple of you know options and you can use any of such options. So, that you can try to understand the data and you get some kind of insights and on the basis of these insights and understanding data you can predict certain things or you can use in a much better way to solve the problems.

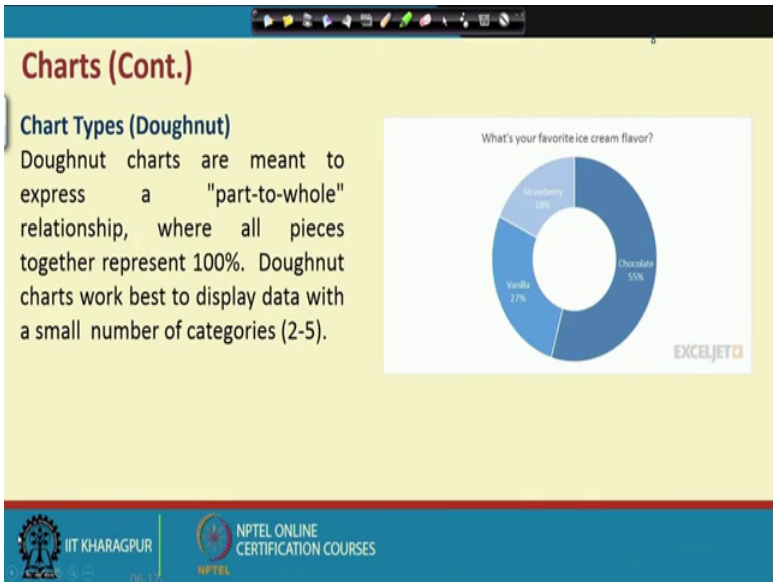
(Refer Slide Time: 19:04)



So, this is you know another kind of chart called as you know radar chart and again this gives you know some kind of insights to understand the happenings you know subject to different kind of attributes.

So, the idea is that you know all these plottings have a different kind of you know structures and, but ultimately the objective is a just to know the kind of you know happenings and you know you means try to understand the data properly and get some kind of you know insights and try to find out any kind of you know obstacles in that data like you know getting you know there may be chance that there will be out layer or something like huge variations or different functional forms. So, these are all you know some of the kind of requirements for further kind of investigations so far as you know advanced analytics tools are in a concerned. So, this is another kind of you know chart which you can actually also help to do some kind of you know visualization process.

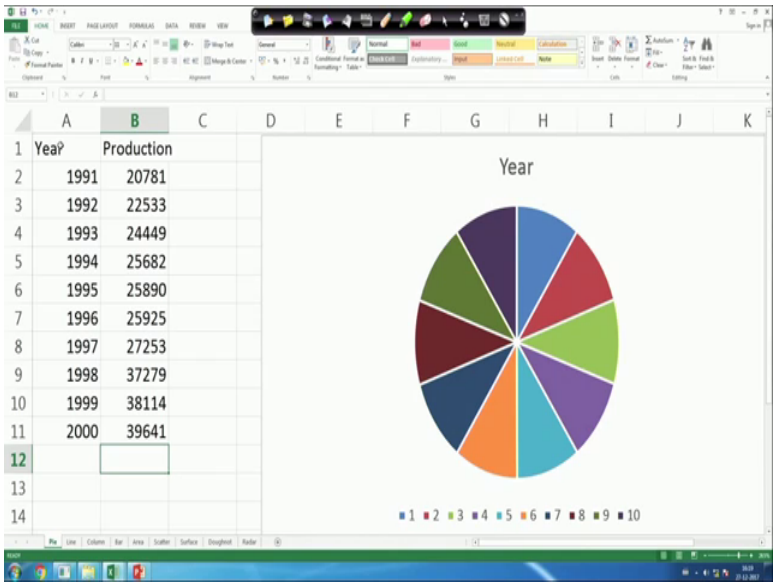
(Refer Slide Time: 20:22)



So, before that you know. So, what I will do, what I have done already. So, the whole idea is exactly here to know the data structures and understanding data, get some kind of you know insights and that is how we have actually gone through so many visualizations you know tools; that means, basically various charts are you know in the excel sheet and just to have you know better understanding of the data and to get some kind of you know insights something negative or positive kind of in the things. Positive in the sense you know that is already in a structure and negative means some that some data has some kind of problems like huge variations having out layers or you know the range is very high something like that. So, this itself will give you some kind of in a better inference for you know further kind of an investigation.

In order to you know understand the kind of you know exact happening then let me you know come to a particular examples and then I will highlight the particular structures. So, here, whatever in graphs we have already discussed and then you know, I will come to here the particular visualization process.

(Refer Slide Time: 21:35)



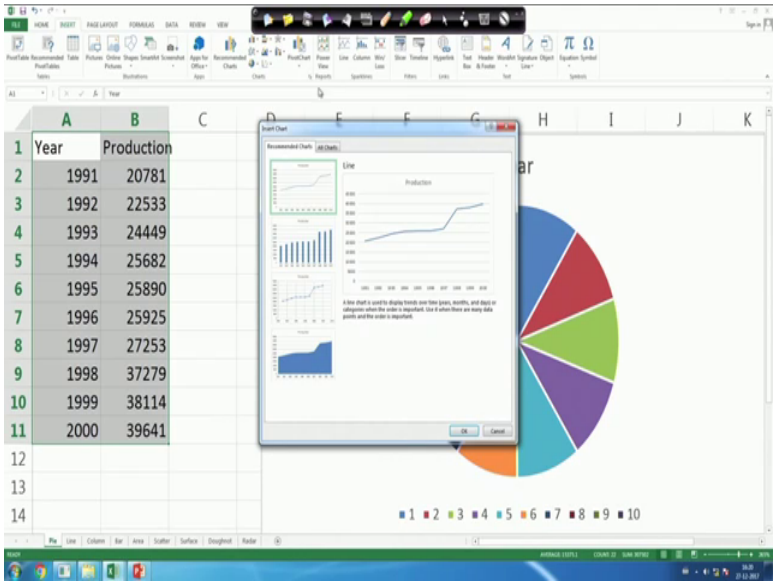
So, here, let us say l information is there and this is a production data which we have already highlighted and with the help of you know this production data you can find out here anything you like to target you can you know find out for instance put you know equal to you find out to maximum, you can find out to minimum, so you know see you will find maximum of the series, you can also find out to minimum of the series, you can find you know for instance you put here it is not coming. So, put here its minimum of the series, minimum of the series minimum of the series this here.

So, just end in a fix the range and close the loop, close the loop and this will give you some kind of you know so; that means, these are all actually a absolute kind of you know understanding, but by default here in the graphical plotting it will give you much better insights. For instance with these particular figures you can have some kind of description or you know basic statistics to understand this particular data or to understand this particular variables, but the idea of this particular lectures or you know this graphical visualization is to get more insights in this spread sheet or you know data set and that too this production data set with respect to times that is from 1991 to 2000. So, now, whatever you know plottings we have already discussed the same you know plottings can be visualized here on the basis of this particular data.

So, go to this particular structure and a click here insert you will find you know plenty of order does data structure is there you just you click here and then you will find you know

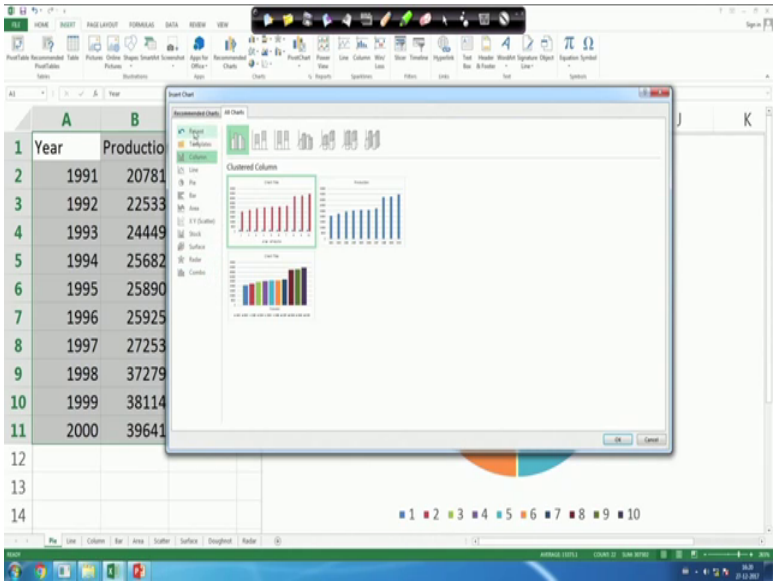
plenty of you know structure here. So, you just you know little bit make it big so that you know you can look into properly look into this particular process.

(Refer Slide Time: 23:35)



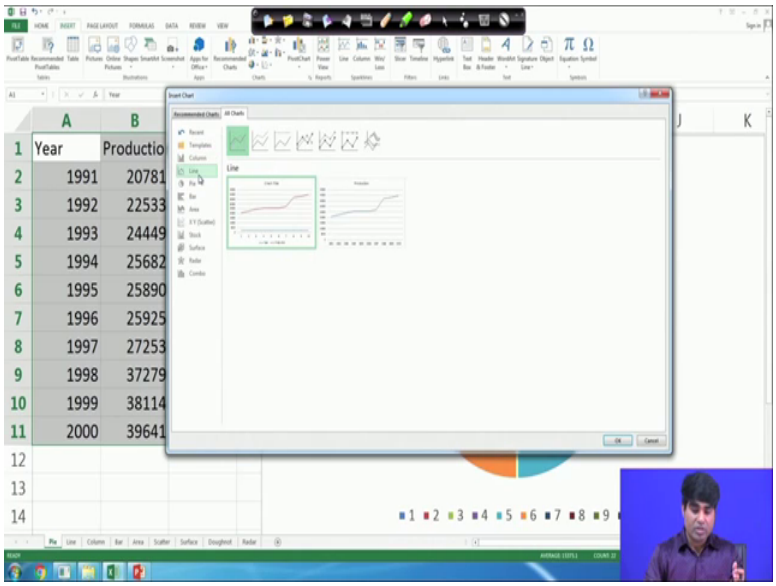
So, this is what actually the particular you know you know box and give the option of you know all charts.

(Refer Slide Time: 23:55)



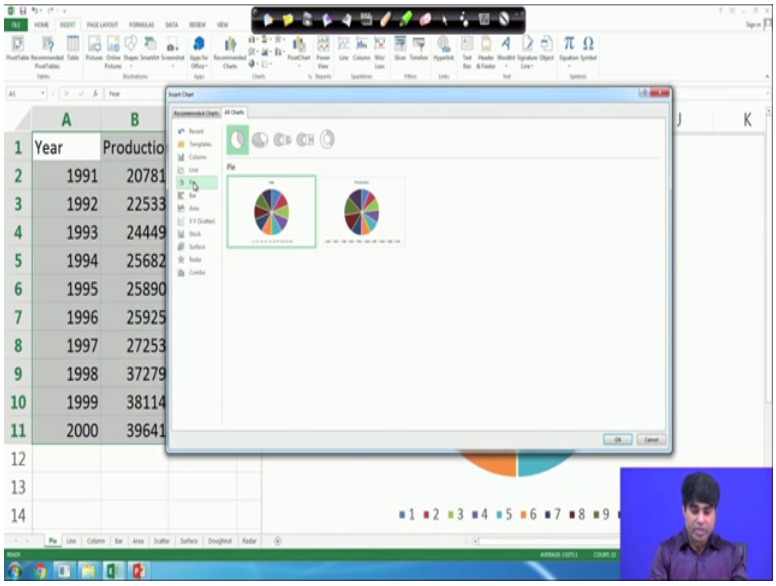
So, whatever you know charts we have already discussed and then the same data can be visualized with the help of you know all these possible options right.

(Refer Slide Time: 24:08)



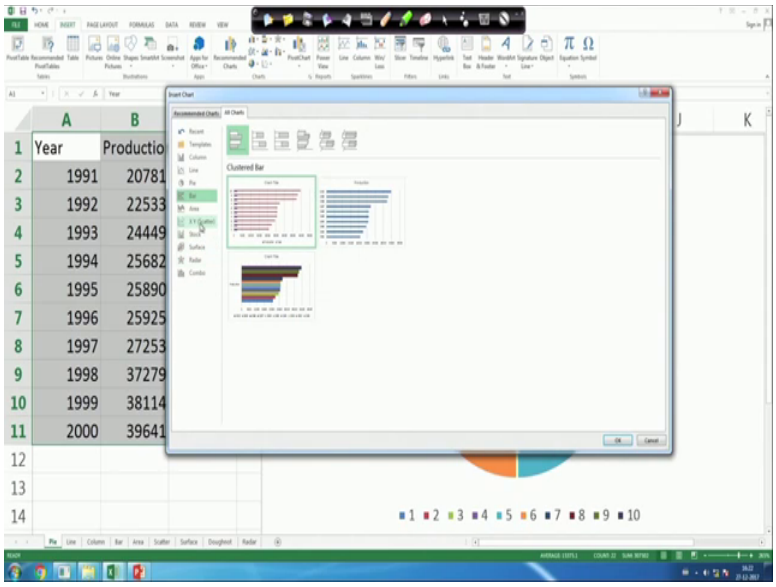
So, for instance you know the you put line a line kind of in options then you can try to know how the production behaves over the times right. So, this is a kind of know classic example line chart and this is another angle of you know production behavior with respect to different point of time. Then again and you can transfer into the column you know representation see here is, so again yearly kind of you know indications and the production movement.

(Refer Slide Time: 24:44)



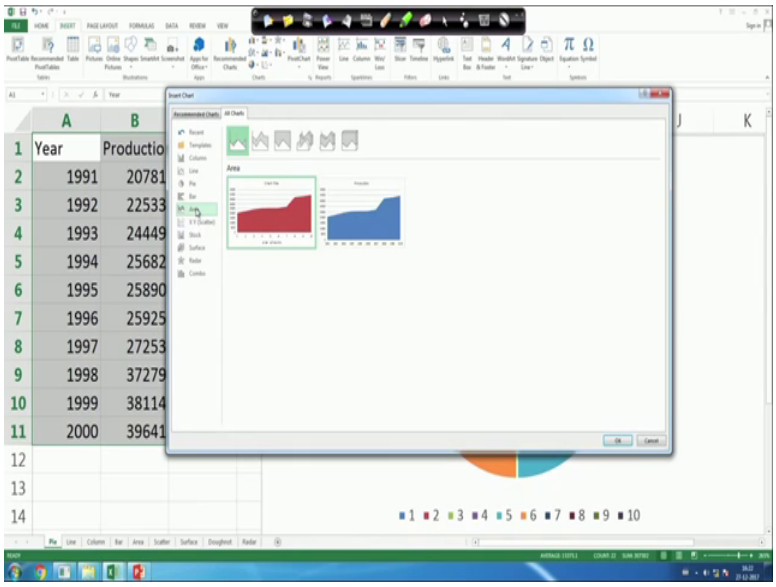
So, this is you know give you some kind of you know better insights. Again you know you go to the pie, pie diagram; that means, it will give you yearly breakup actually which year is high, which year is low, then can compare it with various sums with respect to differ different point of time right. So, this is another look or the same data set.

(Refer Slide Time: 25:03)



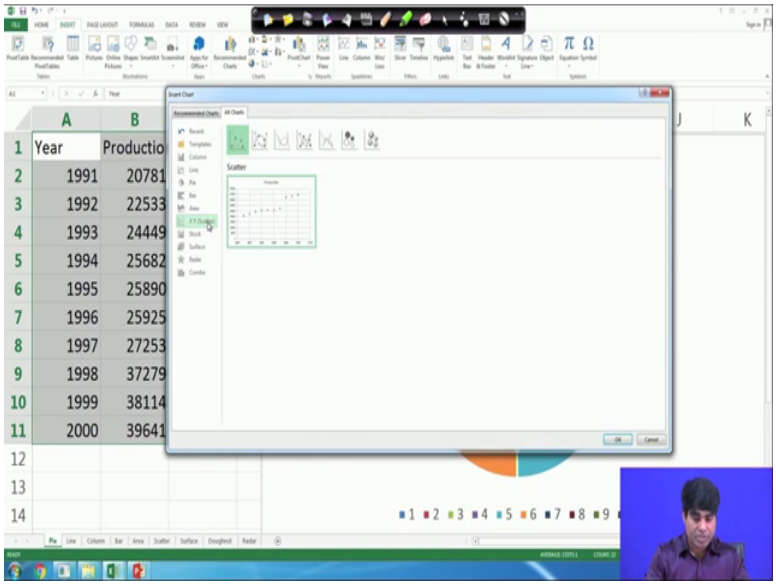
Then again you can have a bar kind of you know that then you can get to know the inside seconds. So, this is another kind of you know look.

(Refer Slide Time: 25:15)



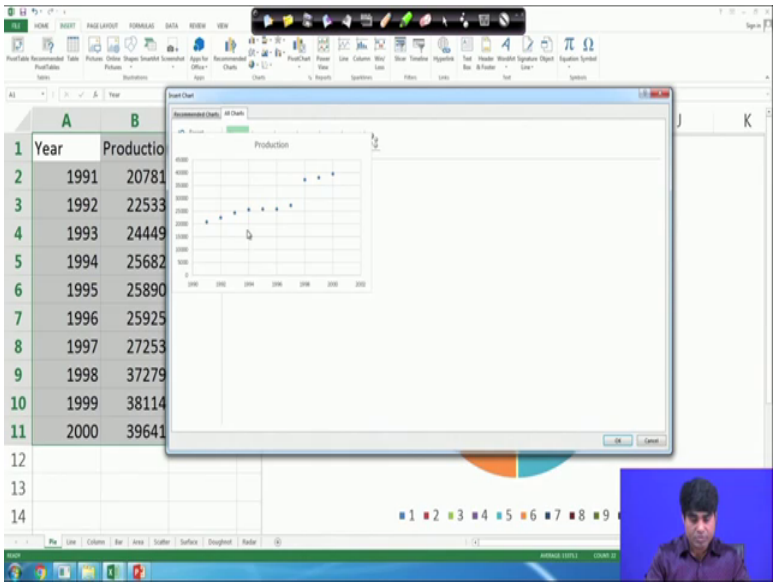
Similarly the same data you can put in a kind of area format and how it is actually moving from one point of time to another point of times.

(Refer Slide Time: 25:26)



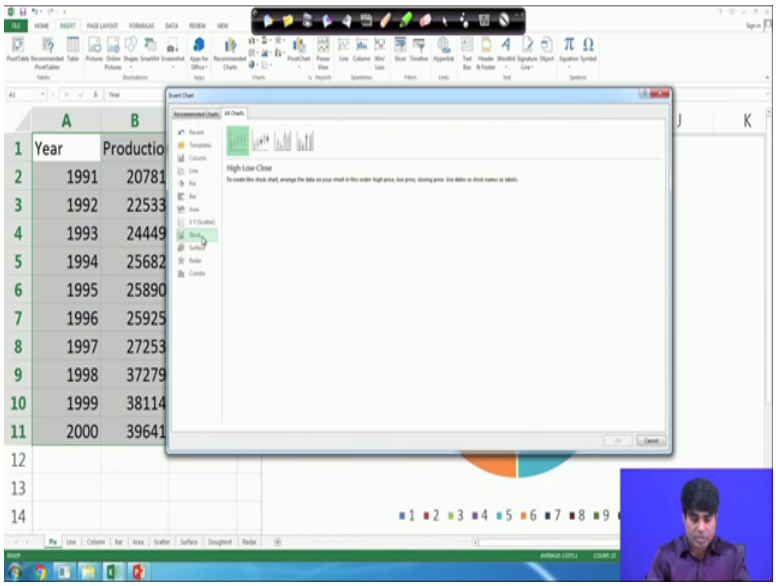
Then again you can go for in a scatter plottings.

(Refer Slide Time: 25:29)



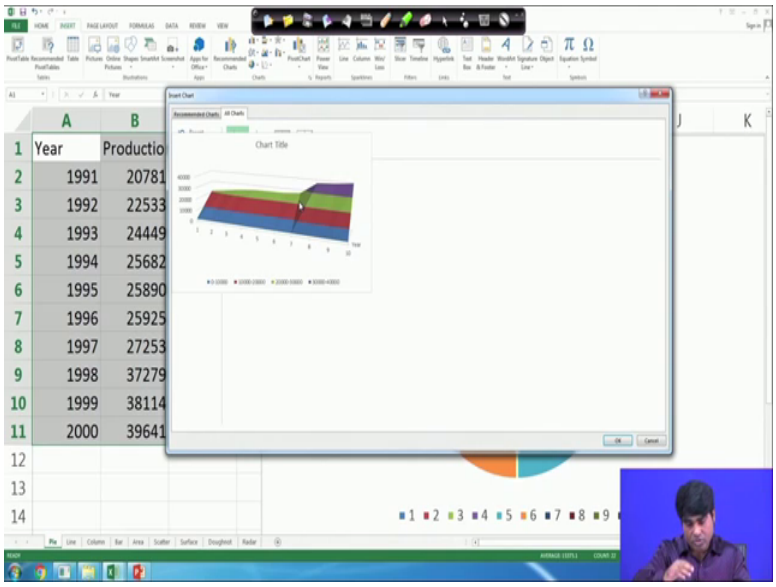
So, this movement you know see this is actually again giving you this some kind of you know understanding about the data visualization process and to know how the production behaves over the time.

(Refer Slide Time: 25:43)



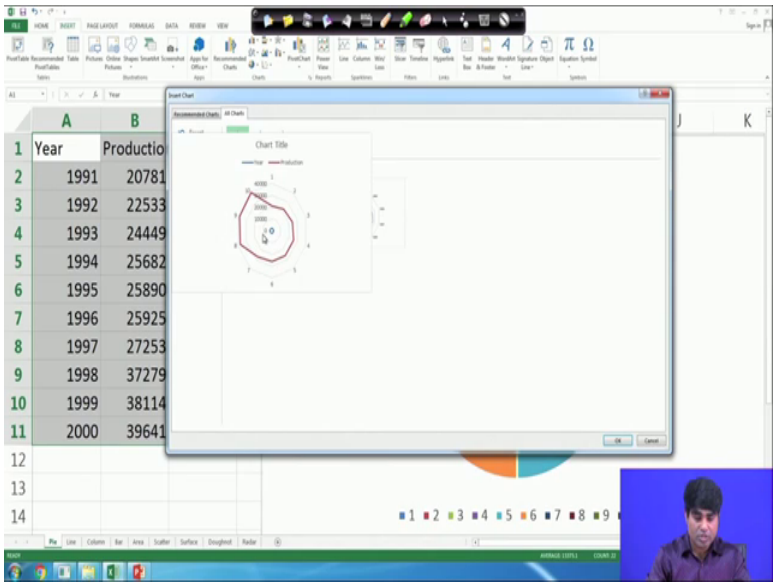
So, then again you know this is another kind of moment and, anyways.

(Refer Slide Time: 25:50)



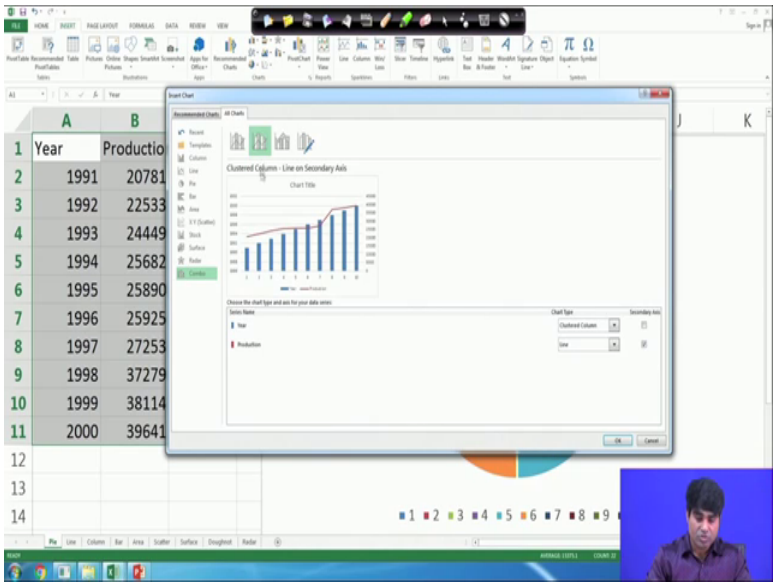
So, this is again scat a kind of you know surface chart. So, this will give us again you know some kind of you know yearly variations with respect to in our production behaviors right.

(Refer Slide Time: 26:05)



So, then same things it can be having you know radar representations. So, again it will be yearly year wise variations over the time and that too how to kind of showing the particularly structures.

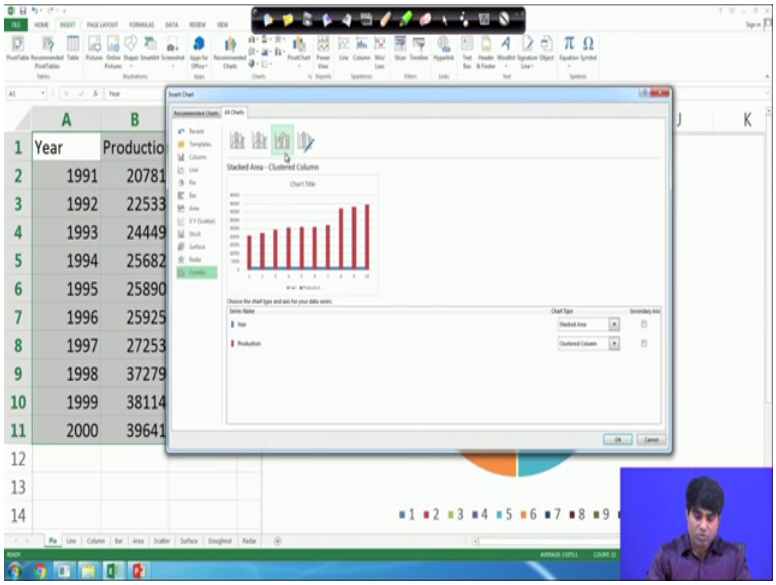
(Refer Slide Time: 26:19)



So that means, again another kind of you know structure here how the production is behaving over you know different point of time; that means, actually the same data, but you know we are actually representing differently. So, this is another kind of you know

movement and just you know checking the trend with the indication that you know how these trend is moving over the time. So, this is another loop.

(Refer Slide Time: 26:43)



This is another loop. So, likewise you know there are you know many different options are there in the excel sheet and just to know actually or just to know what is happening with this data you know over the time or over the cross sectional unit.

So that means, technically any kind of you know data visualization process. So, you must have a data with respect to variables and then you must have some kind of you know entity whether time series entity or you know cross sectional entity. That means, obviously, data cannot data will be meaningless until unless you specify whether it is actually with respect to time or with respect to any kind of in a cross sectional units. So, the cross sectional unit may be individual, may be organization, may be institute, may be anything, like it is a similar kind of you know structure and similarly a similar it can be reported with respect to time it may be annually, it may be monthly, it may be weekly, it may be daywise, it may be hour wise, but some kind of you know structure is required actually. So, when you are reporting a data with respect to variables, then you know, you must be you must be very a careful or you must specify that what exactly it this particular you know data whether you know it is with respect to any cross sectional unit with a different point of time or with respect to time with different cross sectional units.

So, until unless you clear about to all these things then you are not in a position to you know visualize properly and you are not in a position to understand the data and you are not in a position to get any kind of insights. So, in order to get better understanding better insights or you know better kind of you know use, it is better to you know have the data in a kind of in a structured format so that means, the simple structure is like that you know you must have the data that is the information corresponding to a particular variables or you know corresponding to few variables with respect to you know particular structures, maybe time series structure, cross sectional structures, or you know pull structures.

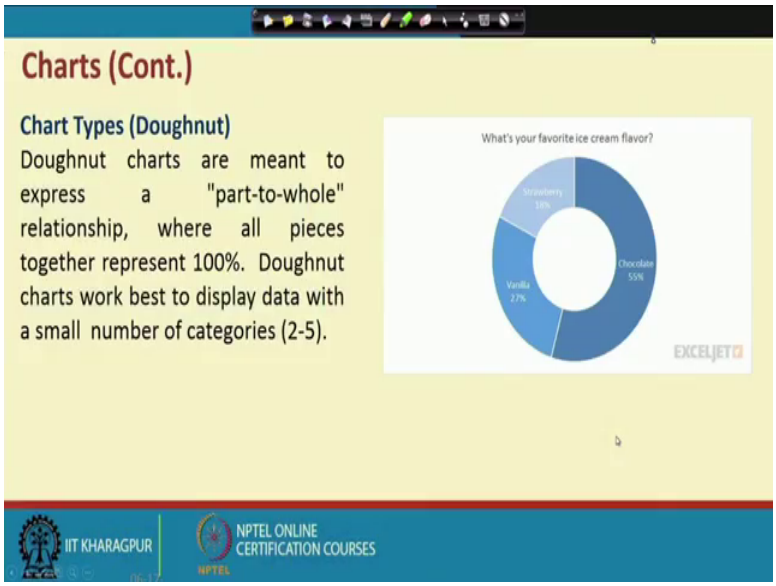
But actually pull and panel is the advanced kind of an understanding, but in this particular stage so far as you know data visualization is concerned that is quantity visualization or you know graphical visualization. So, here the idea is you know every data must have a kind of you know link with you know cross sectional unit and you know time series unit. So, here idea is actually you like to check actually how is this kind of you know variation altogether with respect to different point of time and that too for a particular cross sectional unit or few cross sectional units or else with a particular cross sectional unit how these particular variables behaves, particular variable behaves with respect to a different point of time.

So, in order to know much better insights about this particular data and that too for these particular variables, you have a several kind of plotting options or several kind of quantitative strategy; that means, by using maximum or minimum sum average accounting kind of any kind of logic. So, the idea is you know you try to extract the data or you know get maximum insights by using any kind of functions right. So, that is how the beauty of you know excel functions. So, if you are very familiar with you know excel function in all angles, then you know you can understand the data a very you know effective way and sometimes you can actually you know transfer the data or in some kind of you know further structuring you can do with the help of you know excel functions or something like that and then you transfer the data in a kind of structure where it will be give you better insights and you know better kind of processing as per your you know particular requirement.

But all these plottings we are doing here with respect to the original data, but once you understand the data at the original levels or you know these are all original level or we

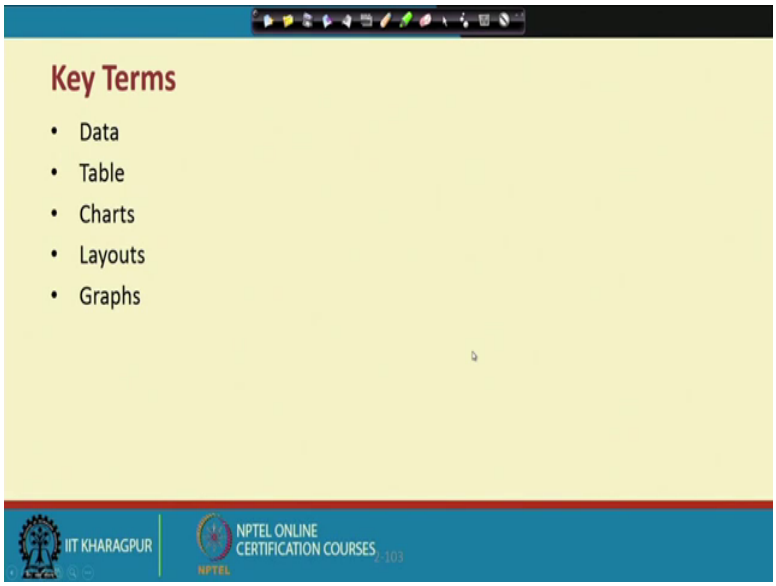
can say that you know it is a kind of you know actual data right, but when you actually use these data for any kind of investigation by using any kind of analytics. So, that times the data may be in a different kind of you know structures. So, you may use actually transport data or some kind of you know process data right means because I mean, what I can say that you know the structure and processing data is not a kind of you know constant kind of nothing. There are many different ways you can process the data or many different transformation you can apply and ultimately we need all these transformation and processing so that you know it will be it will be well connected with a particular technique and well connect with a particular you know modeling. So, that you know we can predict the particular environment or you can get the better insights as for a particular requirement to take some kind of you know management decisions. This is what actually and the kind of you know structure.

(Refer Slide Time: 31:57)



And then, the items which you have discussed today is actually the data, the various tables and you know charts and layouts, graphs, and the whole idea is to just to understand the data and to get some kind of insights as per the availability and as per the requirement.

(Refer Slide Time: 32:00)



And in the next lectures we will discuss something actually quantitative kind of you know structure or you know more in depth little bit about the data understanding and the kind of requirement. With this, we well stop here.

Thank you very much, have a nice time.