[Music]
[Music]
welcome to the next lecture on the course descriptive statistics with our software who may recall that in the last lecture we had considered some graphics and we had considered the construction of bar diagram now he will continue with the topics on graphic and we will try to learn some other types of graphs in this lecture we are essentially going to discuss about subdivided bar diagram pie diagrams and say three dimensional pie diagrams okay so let me start our discussion at the first topic that is subdivided bar diagram and this is also called as component bar diagram what does this mean and what it interprets you have seen that when we created the bar diagram we had created these types of bar right and every bar indicating a class a1 a2 and so on but these bar diagrams are going to indicate only one value at a time now suppose there is a situation where the value inside this bar or this bar is also subdivided and it depends upon some other values then what we will try to do that I will try to create here the bar and I will try to subdivide it like s this is the component of first aspect and here for example this is the component of second aspect similarly if I try to take the second aspect then I can say here well this is the component of second aspect and similarly if I try to stay here take the third aspect can I can see here this is my here third aspect so now you can see here inside these classes a1 and a2 you also able to compare different things for example I can compare that the contribution of this part it is here so I can see here that in the class a1 the contribution of this green diagonal lines is less similarly if you want to compare these orange lines they can be compared by the height like this and if you want to compare the third category you can compare by the red lines like this so what happens that this subdivided bar diagrams they try to divide the total magnitude of variables into different parts in various parts let me try to take an example and try to show you how these things are going to work and then how we are going to do it on the our software suppose I try to take here our data and this data is on three shops shop number one shop number two and shop number three and the data is recorded on the number of customers who are visiting say for example between 10:00 to 11:00 a.m. in the morning on four consecutive days which I am denoting by day 1 day 2 day 3 day 4 so this is sort of two way table in which the rows are indicating the shops and columns are indicating the days and the interpretation of this is like this suppose if I take the value here – this means there are two customer who visited shop one on day one whereas there are 20 customers who've Estate shop - and there are 30 customers who visited shop 3 on the first day and similarly if you try to take any other number suppose if I take here this 15 here this is indicating that there are 15 customer who visited on day 3 the shop - and so on so now you can see here that here are two aspects one is here shop and another is here day and these two values are going to determine the number of customers visiting the shops during 10:00 to 11:00 a.m. in the morning right so now how to do it or how to plot a subdivided bar diagram in this case what I would like here is the following you can see here that in case if you try to make a simple bar diagram then it is not very convenient or it is not very informative because that will be giving you the information either on shops or on the days but these data values are depending on two aspects shop and the days so this is the advantage of using the subdivided by diagram that I can control or I can represent both the aspects together okay so what I would like here is the following suppose I want to create here three bars so this is indicating the shop number one first bar second bar is indicating the shop number two and third bar is indicating the shop number three so on the x-axis I am trying to denote here the shops and on the y-axis I will denote here the DS so for example in case if I say here on day one is representing this thing here this thing here and this thing here similarly if I try to take it here day

to day two might be indicating here somewhere here like this this orange lines and similarly if I try to take here day three day three I can take here like this and if you go for day four this is some this dotted area so you can see here that this height this indicating the d1 and this height orange height here here and here this is indicating d2 and so on so now looking at this type of graphic you can have the information that how many people visited a particular shop on a particular day in a single glance and this is called as subdivided bowed diagram now we want to construct it but before you use the command to plot this subdivided bar diagram you have to think that how you are going to input the data in your are command Y if you remember in the bar diagram you had input the data using the C command just as a simple data vector but in this case it is not a data vector but the data is given in two dimension so I can use the aspect of matrix theory and I can use the matrix command to when put my data and you can see here I am trying to give here data if you try to write down this matrix here I can represent this data as to 2030 26:53 forty forty two fifteen twenty five thirty seventy five and hundred so this is going to provide us a matrix so what I would like to do here that first I try to create the data matrix so now you can see here that in this matrix there are one two three and four there are four rows and there are three columns so now you may recall that we already have discussed the use of matrix theory or how we do you would like to provide the data inside the matrix so I use the same command here and I try to store the outcome in the data vector or the data variable say here caste which is indicating the customers so I will use here the command matrix now as per the rules of this command I will try to provide the number of rows by the parameter n Rho equal to for number of columns by the parameter n called equal to 3 and now I have to give the data which I want to insert inside the matrix so this data is given row wise so I'm trying to give here this command by Rho is equal to 3 that means true and data is given in the form it like here 2 to 20 to 30 and then here 26 and then 53 and then 40 and so on so you can see here now I have given here this data and once I try to see the outcome of this command you can see here I get here a matrix of order 4 by 3 data where this column is denoting the shop number 1 shop number 2 by the second column and shop number 3 by the third column and what about this Rose these rows are converting the DS so now you can see here the data what is here in this matrix it is the same as data given in this two-by-two table right once you enter the data after this you have to use the command bar plot and inside the arguments you have to give the data or the name of the variable that is containing the data in the matrix format and this command will create a subdivided or component bar diagram where the columns of the matrix are going to denote the bars of the diagram so this bar will denote the columns and there will be some sections here and these sections are going to denote the frequency in cumulative format form it what does this mean for example if you try to look in this data matrix the first column is here 226 42:30 so if I try to denote here this value here say had to 26:42 and 30 right so you will see here this these are my frequencies and now they will be denoted in the cumulative format how to do it how it will look like I will just try to plot it and when I will try to explain you right so remember one thing that in the subdivided bar diagrams the frequency on the bars they are essentially the cumulative frequencies and in case if you want to find out the frequencies by looking at the bar diagrams or the subdivided bar diagrams it is pretty simple try to subtract the two cumulative frequencies to get the difference and that will be indicating the values of that particular class suppose if I try to take the cumulative frequency of two classes and I try to subtract it by the cumulative frequency of the class one then whatever is the different that is going to provide the value of the frequency for the class - okay so now just for your given here the data right and when I try to execute the

calm here barplot cast cast was the name of the variable in which I have given the data in the matrix format then I get this type of subdivided bar diagram or the component bar diagram what we need to do here that first thing let us try to understand what is this showing us you can see here there are four sections here one here is black second here is see here dark gray and then here the lighter gray and then here is more lighter gray so these are four different colors which are used inside this bar to divide it into four different components what is here your bar bar here is like this and what are your hair components first component is here one second component is here two third component is here say like here three and this is your hair fourth component so you can see here as the name suggests the bar of the diagram is subdivided now if you try to see what is happening on the x-axis this is trying to denote here the sharp one second bar is denoting shop two and third body's denoting shop three well the basic command that is the bar plot will not give you all this information but in the further slides I will show you that how you can insert these lesions on the x-axis y-axis and how you can add titles and how you can provide different types of colors to the bars right but in this light I am simply trying to explain you that what is the interpretation of a bar and its component now if you try to see over here on the y-axis these values are 0-3 fifty hundred hundred fifty and so on so these values are the values of cumulative frequency how let me try to explain it by taking the first bar of shop one you can see here here where I am denoting this is your a very small bar of black color so in case if you try to move from bottom to up on this y axis the height of the bar diagram is actually here - this is given here so this height if I try to make it here this is here - what is it because – this - is actually this value and now whatever is the boundary of this dark gray and light gray component where I am trying to make a cross this boundary is the cumulative sum of first two classes what are the first two classes if you try to look at this data table in this first column in the shop one first frequency is two second frequencies twenty-six third frequency is 42 and fourth frequency is 30 so this border line is indicating the cumulative frequency of two classes that is the first and second class so you can see here the frequency of the first class is here - and the frequency of the second class is here 26 which is given here so their sum becomes 28 and this value here is actually 28 and similarly if you come on the next partition I am trying to make here a small circle so that you can see on the screen what is this point this point is again representing the cumulative frequency cumulative frequency of what cumulative frequency of the class first second and third classes what are this thing the frequency of first class is to frequency of second class is 26 and frequency of third class is 42 how you can see here like this this is the 42 value so this value at this circle this is indicating the value 2 plus 26 plus 42 which is equal to here 70 so this is the cumulative frequency of all the observations and if you come to the last border we're I'm trying to make it here a square what is this point this point is the sum of all the frequencies so all the frequencies are here first class has sequency to 2nd class has frequency 2006 third class has frequency 42 and the fourth class has frequency 30 that you can see here and their sum is going to be hundred and this is what is being read you noted here that's a hundred and the same story goes for the shop two and shop three similarly you can create partitions and you can create the component bar diagram for shop two and shop three now what is the advantage of creating this type of bar diagram so I couldn't try to have a look on this bar diagram if you try to see here if I try to compare here the peaks like this one or if I try to compare here the height of this particular components what are the indicating the height of shop number one first component is smallest the height of thus bar 2 which is indicating the shop 2 has more height can the height of the shop 1 that is the was Bar and third bar has the highest height so that is indicating that the

number of customers visiting show up one show up two and shop three so one can very clearly see from this graphic that the number of customer visiting the shop number three they are the highest and for that you don't need to look into the data and now in case if you want to find out that on a particular day which of the shop has more number of customer what you have to do you simply have to just compare the component with respect to that day in this bar for example in case if you want to see that on day four which of the shop was visited most by the customers so you can see here in this bar number three height of this component and try to look in the height of the this part in the second component you can see that here this component is smaller then this one so I can save very clearly by looking in the last component of these three bar that the number of customers who visited on day four were the highest in shop number three then in shop number two and the lowest was in shop number one because this height is the smallest similarly if you want to see what really happens on that day two so you can see here by comparing the dark gray part this part in the three bars you can simply compare and can look into the heights of the components and whichever height is more you can say very clearly that the number of customers going to that shop they are the highest now let me try to first show you this graphic on the our console so that you get more convinced so firstly let me try to with this data vector here so you can see here I have created here a data matrix like this and after that you can see here my command was bar plot and name of the variable in this case so into this case my name is cust so I can write down here bat plot C here C UST and you can see here this is the same graph which we have just obtained right come back to our slide and try to do something more now you could see my objective is that I would like to add some colors and I would like to add some regions on X and y axis I want to add some labels so how to get it done you see adding colors will definitely make the components more informative they will be more easy to visualize the choice of colors depends on you and in our software there is a particular code for or every color well I'm trying to use here the simple colors like a red green orange brown for that they have the same spelling but in case if you want to use any particular you please look into the help menu of our software and can decide what color you want and what is the correct spelling of the command to give that color so I'll try to write down here the command and I will explain you what is really happening so you can see here first 2 bar plot cuz that that is the same command to have the bar plot this subdivided bar diagram now I want to add here these labels please look into the diagram shop one sharp two and shop three and I would like to add here that this is my x-axis which is indicating the shops you know how to do it in order to add these names you have a command here or a parameter in the bar plot command which is names dot Arg n am E s dot Arg and then you have to give the name of the bass which you would like to put inside the double quotes separated by comma so suppose I have here three bars and I want to give it the name shop one sharp two and sharp 3 so I have enclose it with double quotes and I have separated it with here comma and all these values they are converted into a data vector using the C command now in order to put a legion on the x-axis for example here I am using her shop so this I am doing by using the command here X lab X lab is going to give you the idea that what should be the label on the x-axis so here I have the same thing that I am trying to put the name inside the double quotes and I want here the shop so these names are user defined and a completely depends on you similarly on the y-axis also I want to give here our name days so this is given here by Y lab right and this is here the days inside the double quotes and now you can see here in this bar I have given here first component as a hair red second component it has a green third component as see here orange and fourth component here as a brown so I need to give these colors in the same sequence in which I want to put from bottom to top so I'm trying to

make here a data vector of red green orange and brown colors and each of the color is written inside the double quotes and they are separated by comma and then all these colors are put into a greater vector using the command C and the name of the parameter under which I'm going to store this data is C o L which is the short-form of color and once you try to do it and then you try to execute this command you will get the same outcome so I can show you here on the our console also that how these things are happening so on this our console I try to copy and paste this command and I try to execute it so you see here you are getting the same outcome which I have shown you right so this is here red color this is green color this is orange this is brown and on the x axis we have a label shops and different bars have got the name shop one shop to shop 3yx is you have the liberties and so on no let me come back to our this slide now in this slide or in this graphic in case if you want to make interpretation you can also do it for example just by comparing the height of the brown component you can compare that how many people visited shop on day four and you can compare that with shop had more number of customer and by simply by comparing the heights of say orange component you can once again compare that which of the shop was restate more by the customer so the height of the component is simply proportional to the number of customers visiting a shop right and on this y-axis as I told you this is giving you the value of cumulative frequency right so this is how you can create the subdivided bar component and yeah there are many other options available here and if you want to explore them more I would ask you to look into the help menu okay and you can also see here I have given you different aspect means if you want to add labels if we want to add colors so now you can see that this graphic is almost the same which you use to OP by any software that was an expensive paid software the same thing can be obtained in the our software without any cost and it is not that difficult the only thing is this yes you'd need to study the commands you but that is also not difficult help menu is always there you simply have to look into the help menu and then just type the commands okay now after this I try to come on another chart which is the pie diagram pie diagrams also are used to visualize the absolute and relative frequency and what happens in the pie diagram that a circle is created and circle is divided into different segments and these segments will denote a particular category like a category 1 category to category 3 category 4 and the size of these sections like as here this one or say here the size of this here the category 2 or the size of this category 3 actually this depends upon the relative frequency and the size of this segment is controlled by the angle well I can use a head of pet color to make it more clear this is the angle which is going to determine the size of that category 1 similarly this is the angle which is going to determine the size of the category 3 and this size is determined by the angle relative frequency multiplied by 360 degree so whatever is the frequency that you have obtained just multiply it by 360 and whatever the angle you get here we need to create this angle over here and that will give you the segment of the pie diagram and this type of diagram is called as pie diagram now this pie diagram can be created into two-dimensional and three-dimensional here I'm trying to make it here in the two-dimensional plot but the same plot can also be made it like this like this something like this and more beautiful and so on so here you can see this is the height and so on so I will try to discuss two dimension and three dimensional pie diagram both here in order to construct the pie diagram in are the command here is PI and inside the arguments you have to give the the data this data is given by here a vector called as here X now I will be more often using the symbol X to denote the data vector and after that there is a long list of the arguments or the parameters which can be used here to give labels control the size control the colors and so on right so but in our case I have chosen some popular aspects for example first aspect is here X which is giving the data vector

then the second parameter I will show you that so the labels which is giving a description to the slices then third parameter is radius which is indicating the radius of the circle of the pie chart and then another parameter here is mean mean is going to indicate the title of the chart cuo el colors that is going to indicate the colors of the slices that we can choose and last option which I will show you here that is the clockwise clockwise means this is used to indicate that if the slices are drawn clockwise or same anti-clockwise and for that you can use here the command here logical say true or logical false by writing true and false in capital in letters so and if you want to have here more idea I will request you that you please go to your AR and try to look into the help for example I can show you here if you want to help you the PI simply try to give it here help inside the double quotes if you go to the ply PI and you can see here you will come to the website of the art software where they have given here all the details so but for this you need an internet connection right you can see here there are many many options so d'être definitely I am NOT going into those details but I will try to continue with these things so now I would try to show you or explain you this thing as an example suppose 10 percents are asked whether they are graduate and or non graduate and their data is recorded as G for graduate and and for non graduate like is here graduate G non graduate n and then in order to convert it into a numerical value we will use the symbol 1 or the number 1 to denote a graduate and number 2 to denote a non graduate percent so that data on the third person which is here G can be converted or can be written as 1 the data on the 4^{th} person which is non graduate can be written s or can be denoted as 2 so if we have now here this data vector and we want to create a pie chart for this thing ok now I try to collect this data using the C command under a variable named quali this is a short form of qualification so this is the data which I have stored here and this is a screenshot now in case if I want to create here the pie chart you can see here that there are now two categories categories 1 & 2 indicating the graduate and indicating the non graduate so in case if I want to create here a pie diagram I would simply use here pie and then here quale
and as soon as you do it you will get here a graph like this one but now my question is do you want this think about it if you try to look into this graphic this pie diagram this is giving you 1 2 3 4 and so on up to here 10 categories but just now you indicated that there are only two categories 1 & 2 for the graduate and non graduate then what is this happening now you may recall in the earlier lectures while creating the bar plot I explained this aspect that whenever we are trying to plot the bar plot or the pie chart we are essentially plotting the frequencies so whatever is the data that has to be converted first into the frequencies and then I have to create this charge on the frequency right so first I try to use the table command and I try to convert this data into frequency table so you can see here there are two categories 1 & 2 and this is indicating that there are seven percent in category 1 & 3 persistent category 2 and then I try to make it here the pie diagram you can see here now this is giving us a pie chart or a pie diagram that we wanted so this white is indicating that there are seven percents and this blue is indicating that there are 3 percent so by looking at this angle you can see that this segment is much larger than this segment so this is giving us a clear idea that the number of graduates are higher than the number of non graduates and this is the screenshot here but I would like to show you here first on the our console that how the things are happening so first I try to create the data so you can see here this is my data on qualification quail and then I try to make it here a frequency table of this data quail which is like this and then I would try to use the PI command over table quali and you can see here you are getting here the same outcome that we had in the slides now I will come on the say this pin X aspect of this pie diagram that if you want to make it more beautiful by adding colors and labels etc so you can see here this is the same pie diagram but here I have

added a label a title and I have added here the labels graduate and non graduate and I have used different colors red color and blue color so how to do it now I have to use the different options different options are means if I want to give this graduate and non graduate labels I have to use I have to give it here by using the parameter labels L a b LS equal to graduate inside the double quotes and the second label non graduate inside the double quotes separated by this comma and both this graduate and non graduate labels are combined using the C command and this title the persons with qualification this is given here by the parameter main main is used to indicate the the title and then whatever title I want I'm trying to give it inside the double quotes you can see here and after this I am trying to give up vector of colors as I did earlier that colors red and blue they are written again inside the double quotes and separated by comma and they are combined with the C command and they are stored in the parameter Co L and if you try to do it here you can see here that you are getting the same thing so I would try to show you here on the our console so you can see here you are getting the same graphics over here no I would just take a quick example to show you that what really will happen when we have large amount of data for example here you can see I am taking a simple example where there are hundred customers who are visiting a shop and they are attended by three sales percent what we call as one two and three and it is recorded that which of the customer was attended by which of the salesperson liked his first customer was attended by salesperson one second customer was attended by salesperson two and so on right so this is the data and I try to collect all the data inside this data vector sales / that is indicating the sales person and they can I try to create here a frequency table you can see here there are three categories which identicating the salesperson one two and three now you can see here by looking at this data you may not have an idea that what is the number of ones twos and threes but by looking at this frequency table I can say very clearly that first sale person has attended a 28% second has attended forty three percent and third attended 29 percents and if I try to create a why diagram this is now given over here so this is indicating the category one this is indicating the category two and this is attending the category three and similarly if you want to make it here more beautiful by adding label more informative then I simply have to use the same command labels means and colors and then I have to define what colors I want what heading I want and what labels I want to give it here for example I am giving here sp1 sp2 and sp3 to the salesperson one two and three right so let me quickly show you here that what will happen here so I try to use this data store this data and then I try to make it here table of your salesperson this is here Christine and then I try to create here a pie diagram of you can see here okay so the same pie diagram that you obtain now I would like to add here the scholarship etc so I can make it here you can see here I'm going to get the same outcome which I just shown you there now I would like to stop in this lecture we have learned how to create the pie diagram which is essentially a two-dimensional pie diagram now in the next lecture I will continue and I will show you that how to create the three-dimensional pie diagram so now I would request you please try to take some data from the books and try to create such diagrams on the our console and try to experiment it and now and my suggestion will be that please don't restrict yourself only to the parameters that I have used in showing it I am doing it because of the limitation of the time but please go through with the help menu try to read what are the different interpretation of different parameters and try to use them inside the our software under this diagrams and that will give you more confidence and you will become much better in producing more beautiful and more informative graphics so you practice and we will see you in the next lecture till then, Good Bye