**Engineering Metrology**
Prof. J. Ramkumar
Dr. Amandeep Singh Oberoi
Department of Mechanical Engineering & Design Programme
Department of Industrial & Production Engineering
Indian Institute of Technology, Kanpur
National Institute of Technology, Jalandhar

Lecture - 41
Statistics for metrology, Fundamental concepts (Part 3 of 3)

Good morning. Welcome back to the course on Engineering Metrology. So, this is a third part of the lecture on Statistics for Metrology Fundamental Concepts. So, I took the fundamentals of statistics as requested by some students in the feedback. And, in this lecture I would just try to explain take the data visualization that I showed you in the previous lecture. I would just take that to take that to the excel file and I like to show you that how does it work.

(Refer Slide Time: 00:47)
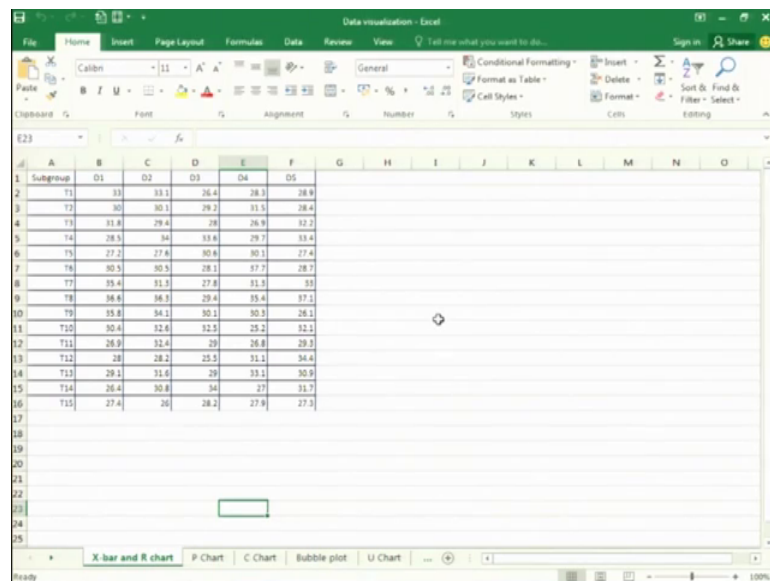
So, let us come to the excel file. So, these are a few data that I have selected for X-bar and R-chart, P-chart chart C-chart these charts would be covered when I will cover the quality control.

However, how to plot the charts, the diagrams or the figures or the illustrations that are discussed in the previous lecture; so, I discussed about a few data visualization or a few charts, such as bar charts, histograms, then pie charts, scatter plots, line charts, bubble

plot, then box plot and so on these are discussed. So, I like to show you how these are plotted here. This is just a data set in which these are the subgroup say T 1 to T 15 are the subgroups. And, 4 D 5; 5 readings are taken this is the diameter of a shaft and 5 readings are taken at 5 5 different locations and this is the output that has come.

So, I will explain the data set more when I will discover the quality control or the control charts you know you can see the name it is made for the X bar and R chart. So, first I like to say I like to show you how do we construct the bar chart. So, you know the first thing is I go to insert tab in my excel window. And, you can see the various options here we can insert in the charts and filters and so many things can be inserted here. First I will go to charts here there is certain recommended charts, those are just I would put here in the front face here.

(Refer Slide Time: 02:37)
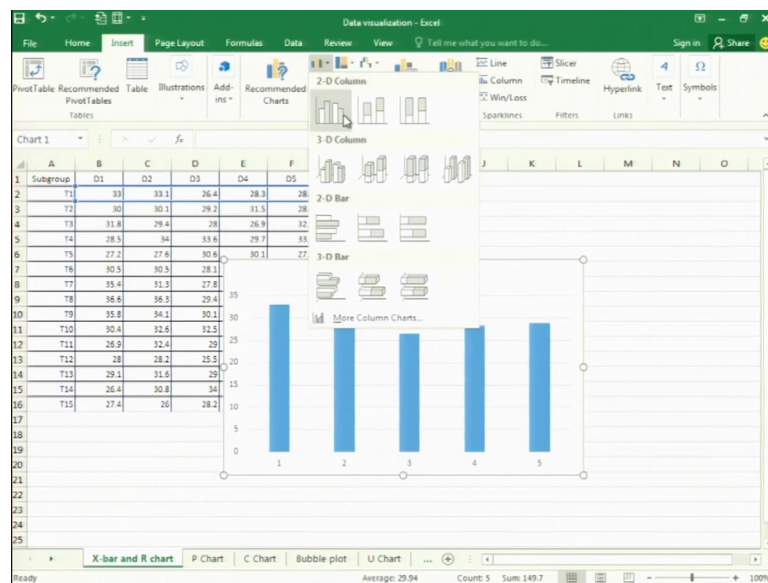
So, I will just click this bar or the columns. So, it is showing 2 D columns 3 D columns 2 D then horizontal and vertical columns. So, as I said that 2 kinds of bars one is stacked bar, another is spread bar. Spread bar is when it is spread it is like I showed the height and weight in the previous lecture. Stacked bar is when it is stacked over each other, when 1 height can be added to the other height. So, those kinds of plots can be made here.

(Refer Slide Time: 03:07)
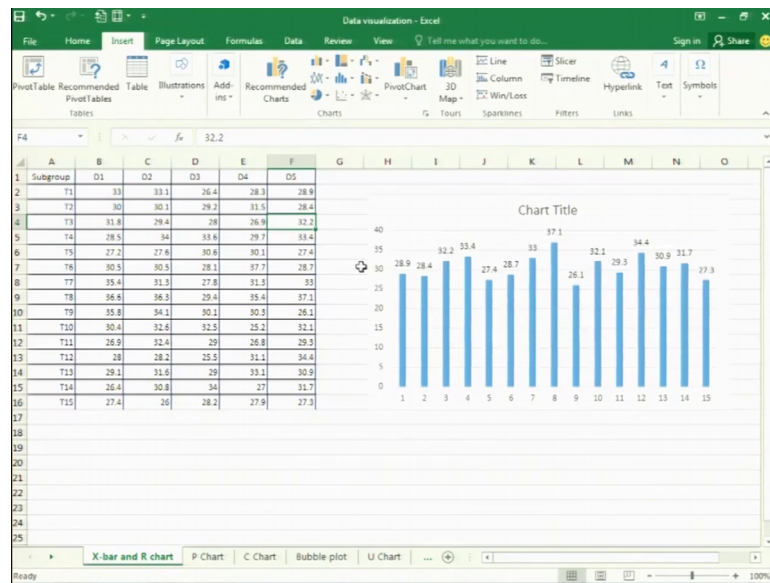
So, let me pick this column I just select the column and go to my chart and just click this thing. Now, this chart is selected it has plotted these values, it is plotted these values this 1 to 15 is the number of observations, you can see there are 15 observations. 1 to 15 is the number of observations and this is the value. For the first second third if I go here I can see series 1.1 value this one.

So, if I like to added this thing I can if I like to see the values here, I can just say add data labels. Data labels is the value of the data that would be represented on the chart itself. So, data labels can be added at data labels. So, it is not showing the value 20 8.9 is the first observation 28.4 32.2 these labels are being shown here. So, this is just 1 series I have selected 1 column in test just spread that parts.
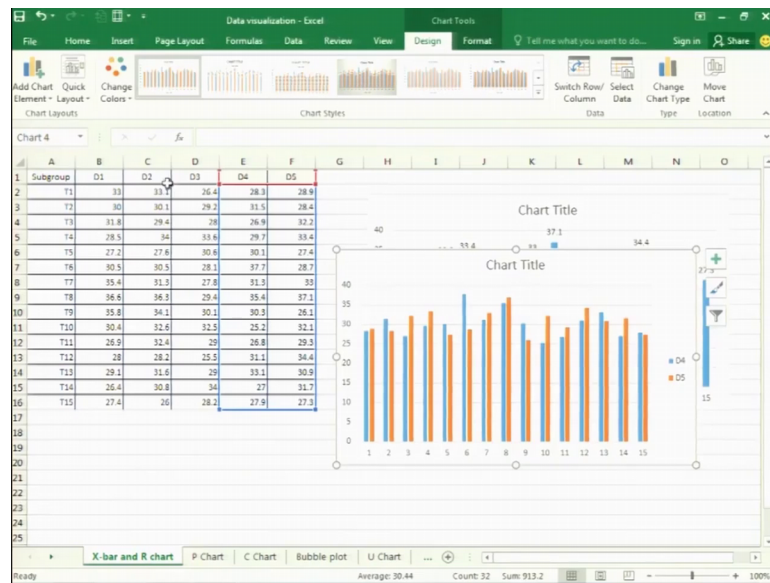
Now, I can select 2 series let me select the first row as well D 4 and D 5. Now, it will try to show the name of the series as well you can see here, it is trying to represent it in 2 different colours blue and orange. So, blue is D 4 it is showing here this is what we call in general we call it as a key of the chart or maybe of the map sometimes. So, this is known as legend here this D 4 D 5 is known as legend here.

So, we can even edit the legend if you like format legend we can change colours, we can keep it at top we can keep it at bottom left right. If look instance, if I just keep it at right and I can even put some colour receptor like I showed 2 requirements are there in data visualization number one is functionality. Functionality is which kind of chart should be select to properly fit the data number 1 it is function ok.

Number 2 is aesthetics in aesthetics we can think of different colours and the whether the chart looks appealing to the eye or not those things can also be thought of whether it does it fit properly in the report or not. So, this 2 legends sorry this legend is 1 legend is there and there 2 data series are there D 4 and D 5 D 4 is this column and D 5 is this column.

(Refer Slide Time: 05:43)
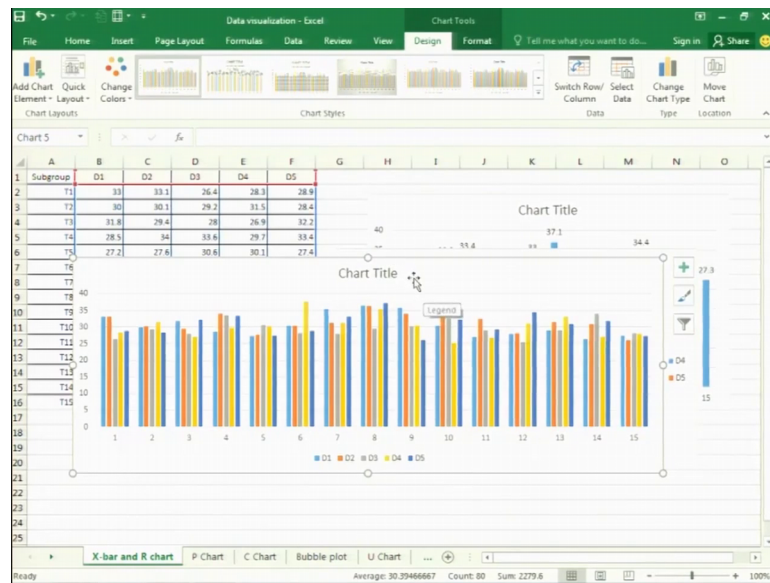
Similarly, I can plot all; the 5 insert. So, this is my bar chart. If I expand it you can see it is also expecting. So, this is a kind of a spread bar ok. This is a kind of a spread bar also I can has stack bar. For instance if I pick this excel sheet we actually share with you, and if you like to see you can see the charts are actually paste over each other you can see this charts ok.

(Refer Slide Time: 06:23)
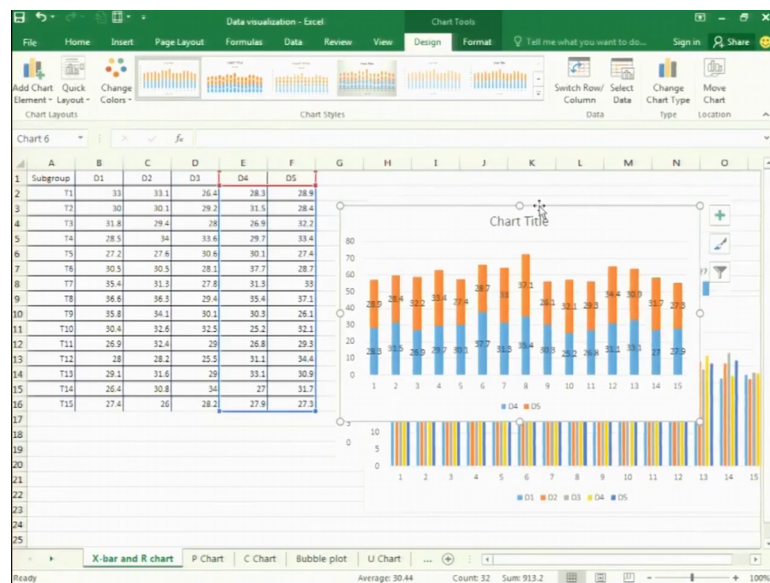
So, if I select these 2 columns and then insert chart, and I select this thing, this is a kind of a stacked bar chart. So, 1 value or 1 series is stacked over the other series. So, this is

the data series this whole this column. So, I can even show the labels her for the specific these values add data labels and I can even get the data labels for the blue 1 as well. So, D 4 and D 5 I have been shown as they stacked over each other.

So, again the stacked and the spread kind of bar charts, also we can have these in the horizontal direction as well. So, that is your choice where if you will where, but the way you would like to the present the data. Say like in time series sometimes if you need to show that this is a time plot as well, it is the first reading second reading third reading, it is better to keep the time or the order in the X coordinate only ok.

(Refer Slide Time: 07:33)
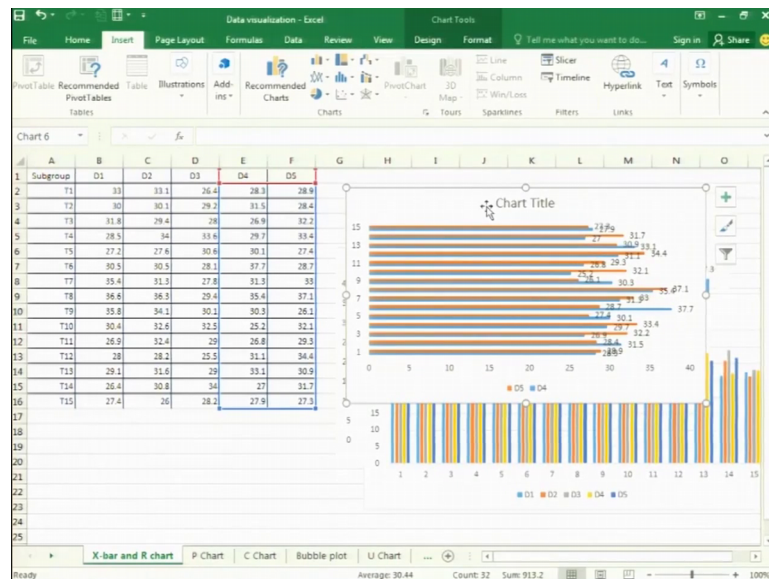
So, otherwise you can pick these as well ok. So, these are the values these are the spread bars in horizontal direction. So, this is the bar charts we will use the X bar and R chart X bar chart actually we would X bar and R charts are control chart, but we will plot in line diagram there. So, I will come to line diagram again. Next chart I will like to show similar to bar charts is the histogram as I showed. Histogram what does that show if you could recall histogram shows the frequency of the occurrence of some specific range or some specific value ok?

So, if I try to plot histogram here this is how we have histogram here. Also, we can see all the charts. So, if I open this one you see I have click this button, this button here it is opening charts all charts it is showing bar, area, scatter plot, surface and I will just cover a few out of this and I can pick histogram from here ok.
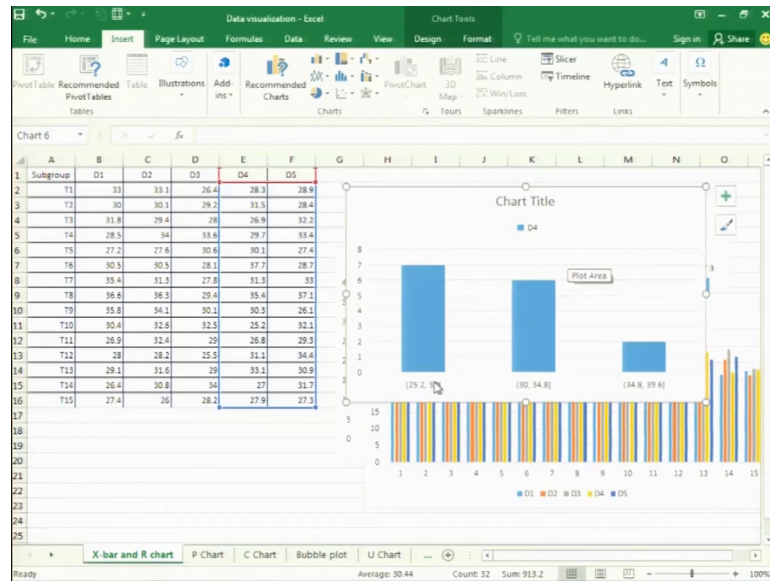
It is not showing the values between certain ranges from 25.2 to 30, 30 to 34.8, 34.8 to 39.6 it is the value for this series D 4 ok. So, it is showing that from 25.2 to 30 there are about 7 observations there total 15 observations. So, 7 plus 6 plus 2 is 15.

So, it is showing the frequency of occurrence. So, this is the way histogram is plot as I said histogram when you print when you have to have to print it for instance this chart, sometimes you can find this kind of chart this chart stacked bar chart, this actually this spread bar (Refer Time: 09:21) is there. So, there were 15 observation had it been maybe 50 observation 50, it would not look very elegant when it is presented on a paper.

So, as I said there are certain tips when we plot a chart or a histogram it has to have about from 3 to 13 bars. It can have more bars, but 3 to 14 is a good number to be selected. It has to be fit into the paper when you print it out it should be legible to be read. Then certain other observations that I said the other tips are important for selecting the chart for like intervals are to be defined.

For instance I did not pick any intervals it pick the intervals by itself, but we can also define the intervals here. So, it has to be comparable with other graphs these things were important as I said. So, this is the histogram chart.

(Refer Slide Time: 10:16)
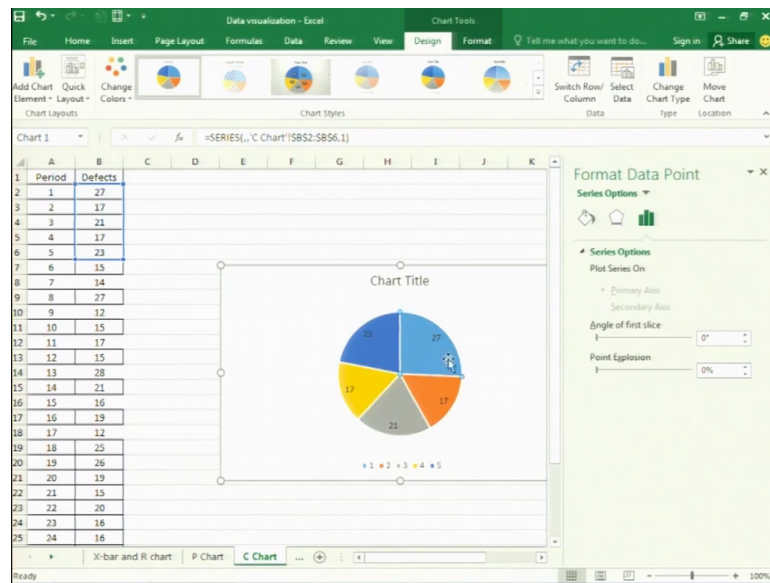
Next, I will come to the I will pick the pie chart now, let me pick some other data this is the data set for the proportion of defective sets. So, I will try to plot now C Chart I will try to plot the pie chart now. Let me pick first 5 values ok, these 5 values if I go to insert here pie chart is directly here over the front end. If, I just put the pie chart it is showing the different it is showing that this is 1 2 3 4 and 5. It is showing one is this much area 2 is this much area you can see then induction number is 1 27. So, this one is this much area ok, let me add labels to see the actual size of each pi.

So, this is 27 units this is 23 units 17 20 and 17 units. So, I can edit this pie chart for instance if I even select all the points or the full complete chart I have put pie explosion. Pie explosion will just expand the chart to some extent ok, 55 percent 8 percent or even if I need to show some specific reading this is the maximum reading.

(Refer Slide Time: 11:31)
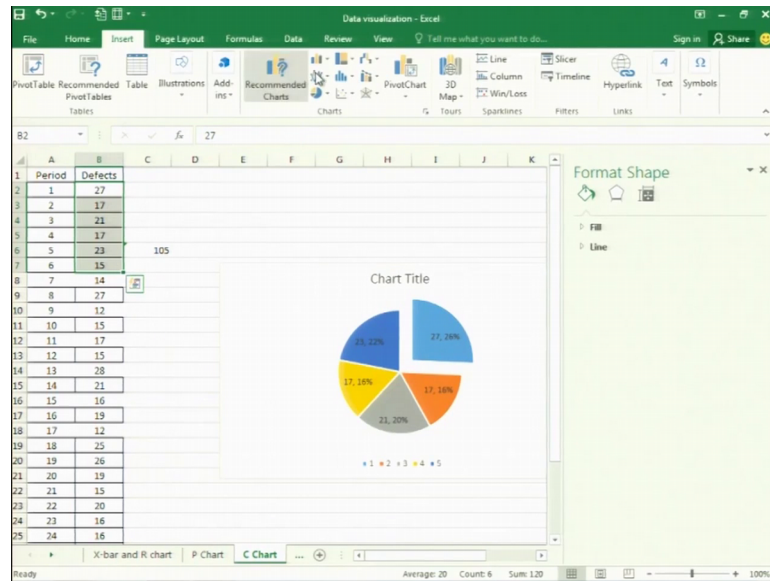
I can just pick 1 and then exclude that. Now, it is excluding only this reading 27 ok. This is a maximum number of units. Even in place of the number of units we can also have the percentages here. This is actually pie chart is the 100 percent complete values are there. The first 5 values I have picked and the whole values are there as I said in the previous lecture that we cannot cheat with the axis we cannot cheat with the scales in case of pie chart.

So, this is actually some percentage is there that as some people like to represent the data in the in percentages. So, I can even edit the data labels. So, I will go to format data labels and add percentages. So, these are now percentages actually it is very close to I think this whole sum would be very close to 100. So, this is this percentage 27 it was the number was 27. So, it is the percentage is also close to that. So, let me say what is the sum of these things let me say this is equal to sum of the sum is 105 that is close to 100.

(Refer Slide Time: 12:47)
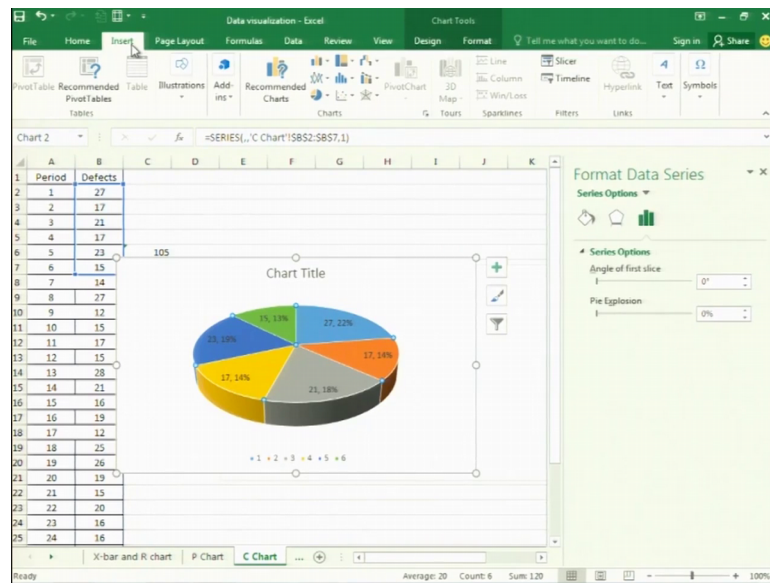
Now, let me pick the 6 values and then plot a pie chart ok. Let me have 3 D pi this time. So, let me add labels a data labels. So, it is showing the numbers here. So, I can edit or format the data labels and add percentages. Now, it is all showing the percentages here ok. So, this is pie chart, how do we plot the pie chart. Similarly, we can pick any kind of chart and plot the next chart I would like to pick a few very important chart is scatter plot. So, let me pick some data and try to see this scatter plot that I will use here only.

So, scatter plot is very important as I said the correlation or the trend is shown in this scatter plot, when we have two different values like height and weight. Then, we had hardness, and wear those can be compared the scatter plot is very important to compare let these are actually diameters.

(Refer Slide Time: 13:48)
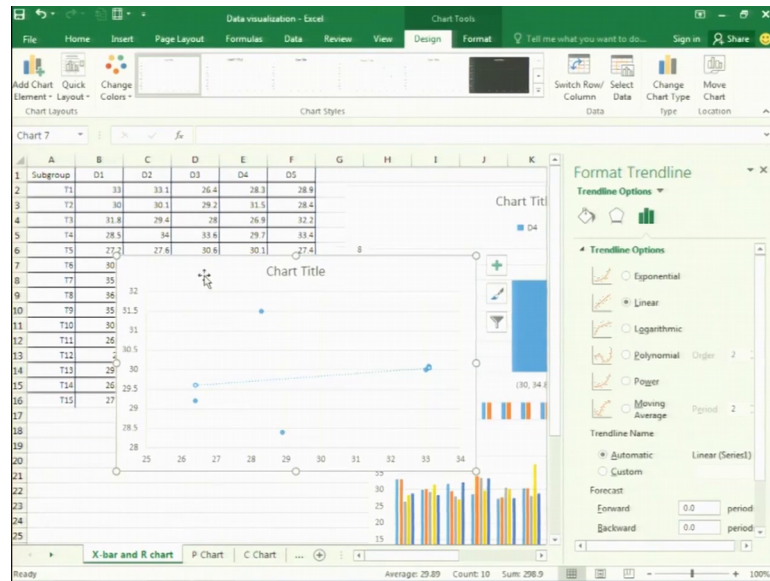
So, let me see whether the diameters of the test 1 and test 2; how do they correlate? So, this is a scatter plot that is also kept here. So, I will just plot the scatter plot here. So, these are 5 points. So, it has taken T 1 on this side edit axis format axis ok. So, one thing is this axis as the X co-ordinate has started from 0 to 35.

However, no value is lower than 25 as so, I can vary this value as well. So, let me edit this axis it is 0 to 25. I will take it from 25 to 35 now it is showing better. So, I can even add a trend line here add trend line it is kind of showing a trend with this kind of trend a positive trend is there. However, this is if you actually see the correlation this is not a very great or very strong correlation, so but some trend is there some trend it is a positive trend is there.
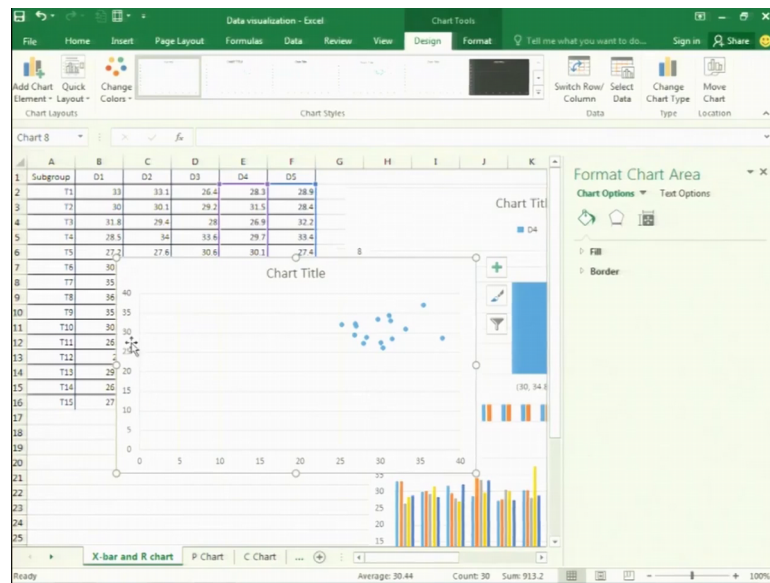
So, let me try to plot some more value little bit try to plot these values here in a scatter plot. So, you can see the values all the values are greater than 20 for maximum all the values are greater than 25 here. So, I can edit the axis format axis this is 0 to 40 I will start it from 25 ok.
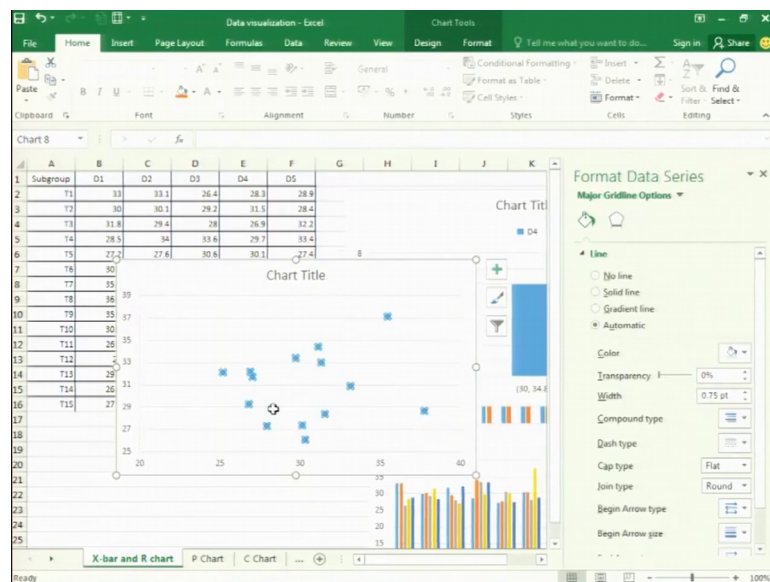
(Refer Slide Time: 15:41)
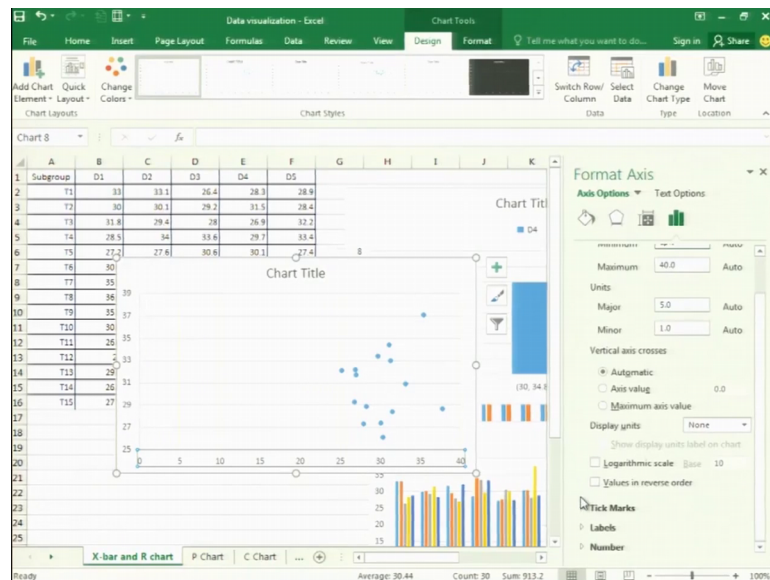
Now this, axis I will also format right click ok, select axis first right click format axis I will start this from 20 maybe ok.
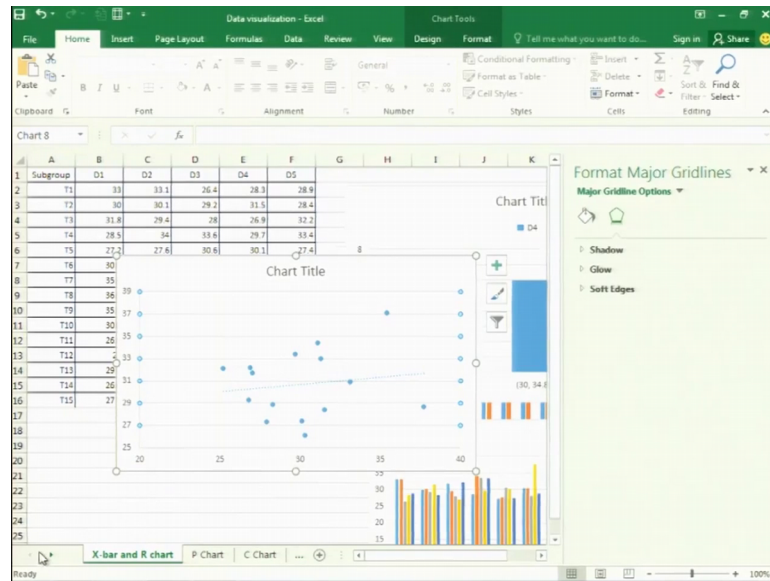
(Refer Slide Time: 16:06)
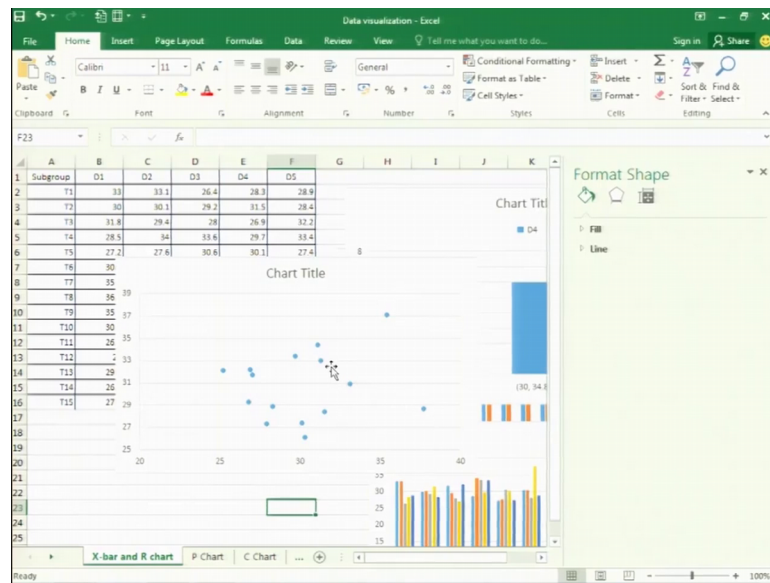
So, this is also kind of a very poor correlation, but let us see whether there exist some that trend or not.

(Refer Slide Time: 16:18)

I will the select the data and add a trendline: yes, it is again showing the trendline also if I need to add a trendline.

If a trendline can also be forms at this is actually linear trendline I can add exponential. I can add some other function some other mathematical relation for the trendline can be moving average like those are the things actually which are out of scope of this course.

But, yes exponential if I add exponential it is show it will take exponential function and then add a trendline. So, the best way at this point of time it is better to just plot a linear trend line. So, it is also showing this thing similar to the scatter plot.
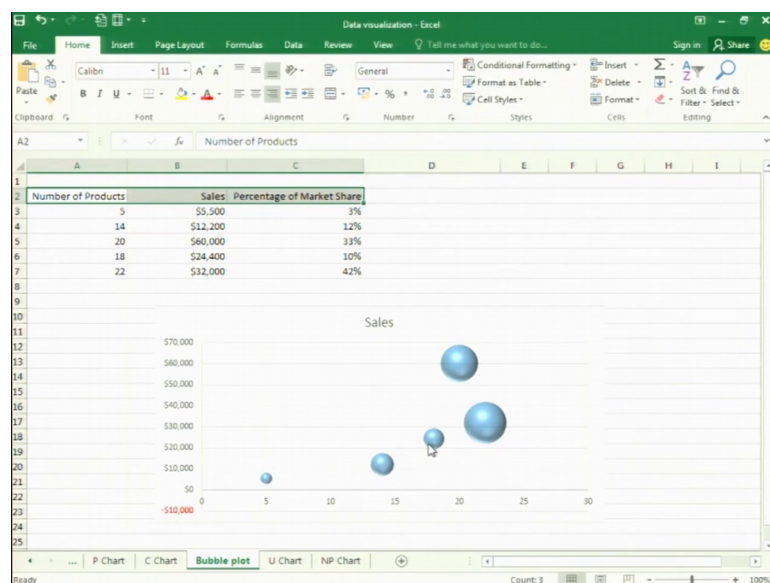
If, we have a third dimension I have picked some data from the available data set of Microsoft to represent the bubble plot. So, we can plot the bubble plot as well. So, bubble plot is not kept on the front face, because those they do not vary frequently used for the people. So, I can just go to all the charts all charts and pick bubble plot here, yes bubble plot is there in the category of scatter plots only. So, I will pick these 3 D bubbles ok. Now, what is this chart representing? So, this is actually number of products which are produced by company this is the sales. So, this is the percentage of market share this specific product has ok.

Now, it is showing 3 dimensions here 0 to 30 to 30 is the number of products ok, 5 products, 14 products ok, then 20 18 22 5 14 20 18 22, then we have the this sales and the sales here are from 55 500 dollars to 32 thousand dollars ok. The sales are here what is this size of the bubble representing? It is representing the market share of the specific product that is having ok. In metrology we can pick let me say, if I am trying to compare the length and diameter and number of pieces which are inspected. So, a number of species they would show the size of the bubble and the length and diameter can be represented in the simple X and Y direction, in the both these axis.

So, this is when we need to show or when we need to present the 3 dimensions in 1 of the 2 D plots. So, this is bubble plot ok. I can even a box plot was also discussed in the previous lecture. Box plot is if I have this let me first plot the line diagram. Let me pick any chart here let pick any chart.

(Refer Slide Time: 19:27)
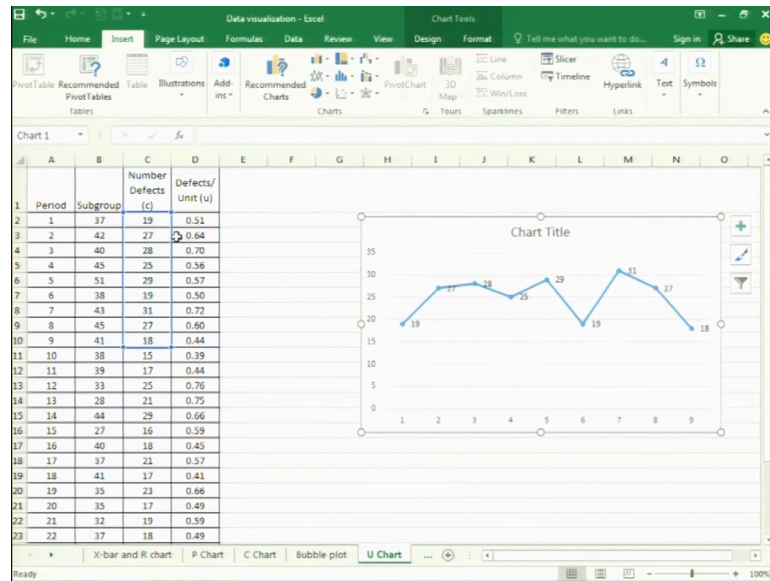
And I will first plot I will pick first trend readings and add a line diagram. So, it is showing the if this is the data is continuous, ok. So, it is better to represent in a line diagram and this we will use very much in our control charts.

So, the data of the first data point is 19 27 28, it is being representing in a line. So, I will just add the data labels here ok, it is values are showing. So, this is our line diagram ok. So, we can have various variables in line diagram we can even have this point say for instance.

(Refer Slide Time: 20:08)
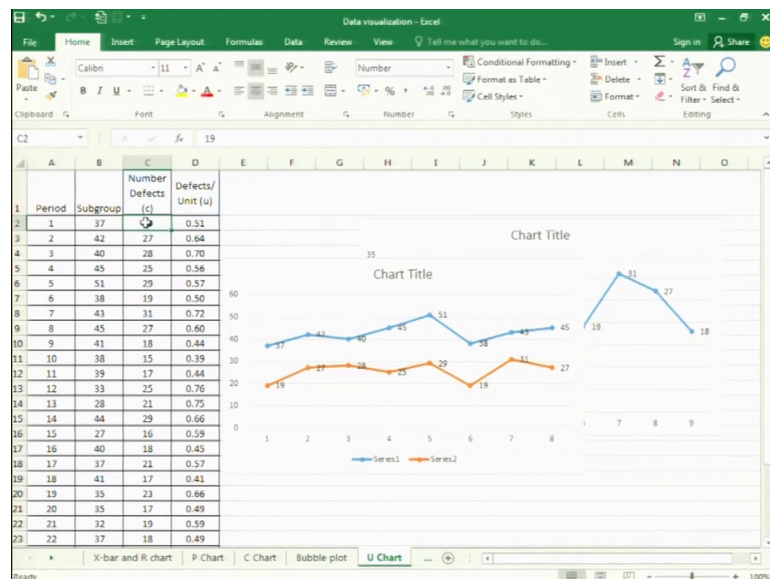
If I have 2 lines, let me plot these 2 lines here. And, you can see that these 2 lines represented even these points the data labels data labels are T varied number 1 ok. I can, even if I have these lines number 1 it is differentiated by the colour blue and orange colour ok. The last chart I would like to show here is the box and whisker plot.

(Refer Slide Time: 20:50)
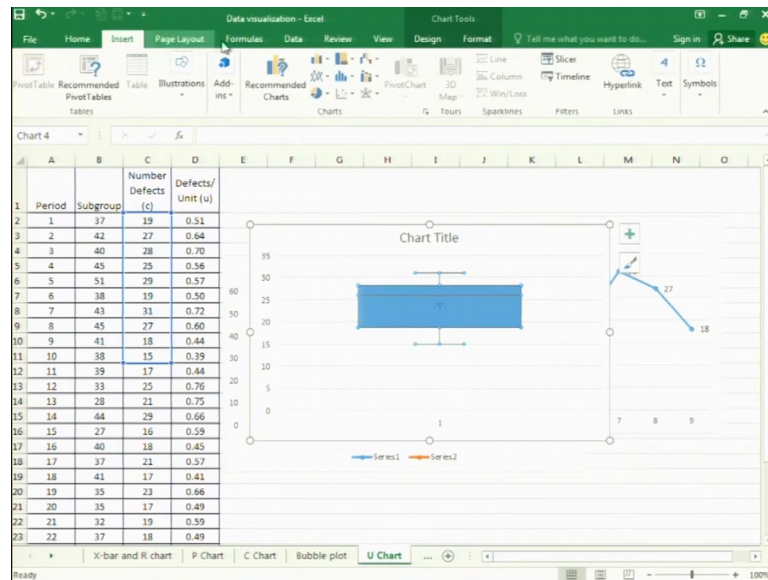
So, let me pick these 11 points here, I picked 11 points got I will like to show is median here, so insert go to all charts recommended charts in all charts. So, then all charts we have box and whisker. So, this box and whisker plot is there ok. So, I can have many box in this plots like this. If you remember box and whisker plot the canter line was our median and we have quartiles, first quartile then median then second quartile. And, we have maximum and the minimum value you can see out of these points 19 and from 19 to from point number for period number 1 to period number 11 I have selected actually. So, the minimum value is 15 and the maximum value where is 31.

So, it is presenting 31 here and 15 here the minimum value. This is the median value median value is the centre value centre value is close to 27 or something ok. So, this is box and whisker plot these are 2 whiskers this is quartile 1, this is quartile 2 over median, this is quartile 3, this is the minimum value this is maximum value.

So, I can have you know if you see the all charts here there are number of charts number of charts that we can plot. We can have even have the surface diagrams like, in bubble plot we have the third dimension, in surface diagram also we can we have the third

dimension the 2 dimensions are there. So, this is a 3 D representation. The number of charts which are there an excel is having a great capability and it can help us to plot many there are other design software's as well. When we design for experiment it is known as design of experiments, and we select the specific number of experiment those are to be conducted in metrology when we to measurements. So, there number of plots that we can use to represent out data.

So, this was the about the fundamentals of statistics in metrology. So, in next week we will meet and discuss more on statistics in metrology.

Thank you.