Video 31

Data visualisation techniques are a very effective way of presenting the results of data analysis but it is important to remember that the effectiveness of the analysis also depends on making sure that we choose the right set of visualisation. What does it means to talk about the right set of visualisation techniques?

There are many types of visualisation options available. We can create line charts, bar charts, pie charts, bubble charts and so on but not all the charts are appropriate for all types of data. So, its important to understand appropriate chart types given some data or a problem statement.

In this section, we will look at what are the different kinds of commonly used visualisation types or chart types and what sort of data is appropriate to use these types for. We will use excel to demonstrate a lot of these chart types but a lot of what we are going to cover is obviously translatable to any visualisation tool that we use.

Excel has a wide variety of charts and graphs options. And so as we create charts in excel, it is also important to remember some fundamental principles when creating visualisations. One, obviously as we have mentioned before, it is very important to choose the appropriate type of chart

for the problem or the data that we have. But also, its important to label and format the chart clearly so that the person looking at the chart or the audience is easily able to understand what the analysis is saying. The less is more rule will apply here which is that even though there are many options available for formatting, labelling, colors etc., we want to keep the chart as clear as possible so that there is no confusion in understanding what the chart is actually saying. But let's first look at some commonly used types of chart options and what sort of data or situations they are appropriate for.

Let's start with line charts which are very very simple charts and are very commonly used. And line charts typically are used when you want to show a trend over a time period. So for example, changes in share prices, changes in revenue for a company. All of these are essentially trends over time. Another commonly used type of charts are bar charts and bar charts are used when you want to compare changes at a fixed point in time, for example, what is the different... what is the difference in sales by Geography for a particular company or what is the % of R&D spent to total revenues by product type. Remember, all of these are happening at a fixed point in time and we are comparing changes across dimensions, so for example, Geography or Product or location etc.

Column charts are similar to bar charts. The difference is that column charts are vertical, bar charts are horizontal. Column charts are also looking at comparisons or changes but they are also looking at it over time. So, sales by Geography across quarters. Pie charts are again a very commonly used type of charting option and used when we want to show proportion to 100%, so for example, what is the allocation of disposable income say across entertainment, food and rent. They will add upto 100% and we want to say for example that disposable income 30% goes for entertainment, 40% goes for rent and the remaining 30% goes for food, so when they add upto 100%. Now these four types of charts are in fact some of the most commonly used types of visualisations.

There are other charts as well that we can look at. One very popular type of charts used when we want to look at relationships between two variables are scatter plots. So for example, if we wanted to understand is there a relationship between income of a person and the years of their education, then we will create what is called as X Y plot or a scatter plot, then there are area charts which show relative importance of values over time and then other advanced charting options like bubble charts or doughnut charts or stock charts etc. But for now, let's focus on the five most commonly used type of charts which are line charts, bar charts, column

charts, pie charts and scatter plots and see, how we actually bill these in excel.

In excel, creating a chart is very easy. You simply have to go to the insert menu, the insert ribbon and click on the type of chart that you want to use. So you can see here, column, line, pie, bar, area, scatter etc. Of course, the chart is always created on the basis of data. So typically, you will go to excel, highlight the data that you want to use and then, go to the insert ribbon and click the type of chart that you want to create. Let's start by looking at how to create the most simple chart which is a line chart.

So we have some data here which is by month usage for a particular customer. This is a trend over time, Month 1 to Month 6, therefore, a line chart is appropriate here. How do we do this in excel? So this is the data, we can highlight the data, go to the insert ribbon and then we can see there is a line chart here, we simply click on insert line chart and pick one of these line charts. Now all of these charts... now when we do that, you can see that there are many kinds of line charts, but these are simply design options, so for example, instead of picking this option, if I pick this, I will get the same line chart but now I have these circles. The choice is totally upto you in terms of formatting ease. Remember, a line chart is used when we want to show trend over time. So you can see that there is

an increasing trend in usage for a customer over time. So that's how line charts are created. If for example we had two data series, we had usage and we had revenue, then we could put revenue here and then we would get 2 lines.

What about bar charts? Again, bar charts remember are comparing at a fixed period in time. The example that we are going to use is keywords that people use to come to a certain website. So when people type in fluevog shoes, fluevog boots, unique shoes, funky shoes etc., we are recording the number of visits to this particular website based on these keywords. Now this is all happening at a certain fixed period of time, may be the last ten days or the last quarter. So if I wanted to visualise this information, I will use a bar chart. So you can see this is the data, we are interested in visualising the visits information, so we go to insert and here we choose bar chart and its as simple as this in excel. You can see the visualisation of visits. Now what if I also wanted to look at revenue, so, I am looking at visits but I also have associated revenue, so I want to add revenue as well. Its very simple, we can simply take this data, highlight revenue in and visits and then click on bar chart. So now, I have created a second bar chart which has revenue information as well. Of course, you may not want to create two charts, you want to include it in the same chart in which case there are options in excel that allow you to add data, so we can say add and

we say this is the series values and the name of the series is revenue. So, instead of creating a separate chart, I have now added the revenue information in the same chart.

Column charts are used when we want to do comparisons over time. So here we have multiple dimensions. The data that we are going to use is sales across location across time, so you can see there are five different sales territories and we are looking at their revenues across four different quarters, then we will use a column chart. So we just highlight this data, go to insert and say, choose a column chart and this is what the column chart looks like, sales by location by quarter. So quarter 1 is blue, quarter 2 is red, quarter 3 is green & quarter 4 is purple, and then you have the same information repeated for different locations.

Pie charts. Remember, pie charts are used when we have information that is a proportion to 100%. So for example, if we go back to our keyword data, if I wanted to understand what proportion of my visits came from people who were searching for funky shoes, then I can create a pie chart. So this is the data of my visits. I want to understand what proportion came looking for funky shoes, so I go to insert and I click on pie chart. So here you can see that I created a pie chart but I don't know what these numbers are, so I am going to add data labels. The data labels are showing up as a number

of visits but I wanted as %, so I go to format data labels and I say I want %, not the value, so now we have %. Now this may not be very readable, so I can always format this to show text in say, white colour. So instead of black, I want to show white and now I have the percentages in white. So the answer to my question, what proportion of people came looking for funky shoes is 17%.

Finally, another very commonly used type of chart is a scatter plot or an X Y plot. Remember, X Y plots are created when we want to look at relationships between two variables. So for example, I want to understand; is there a relationship between hours spent in a mall and the total dollars spent by customers. I will use an X Y plot. So this is the hours spent in a mall and this is the dollar spent on average by customers. So I go here and I use this scatter plot. You can see that there does not really seem to be a clear relationship between hours spent in the mall and dollars spent by a customer.

So these are examples of very very commonly used charting options and more importantly, when to use what type of a charting option. There are other charting options as well that are a little bit more advanced. We will look at those and then, we will also look at how to format and label these charts to make it easy for an audience to understand what is being actually displayed in the chart.