**Scatter plot**

In this video, I’m going to show you how to make a scatterplot in SPSS. In order to make a scatterplot, with your x and your y axis variables should be interval or ratio in SPSS, this is defined as scale, and this is because for a scatterplot assuming there’s a meaningful order to both the x and the y axis variables. I’m also going to show you how to add a trend line.

So in our data file, we have employees’ age and we also have they’re total competency score out of 100, when they started their jobs so I want to look and see is there a relationship between how old someone is and how competent they are when they begin their job. So if we go to the graphs menu and ‘chart builder’. We want to choose scatterplot from the left, the first scatterplot option, so go ahead and double click or drag it into the gallery. And now we need to put our variables on the x and the y. Our independent variable, which is the variable that’s not going to change, so in other words my age will not change if I become more competent or less competent. That variable goes on the x axis. Your dependent variable, which is the variable that possibly could change, so for example my competency score may depend on my age, that variable goes on the y-axis. Go ahead and click ‘okay’, and here’s our scatterplot.

So what we’re looking for in a scatterplot is some kind of trend. So it looks to me like as people get older, they tend to be more competent when they start their job, and that’s probably because they have more experience and more education in their field. Now there are a few points here, a few individuals who look like they’re outside the normal, so these three people here are all different ages but they have a similar competency, so they’re not part of my general trend upwards. If you are doing a correlation or regression analysis and you want to add the trend line to your graph, double click on your graph until you open the Chart Editor, and then up here at the top you’ve got an icon, this is ‘Add Fit Line at Total’. So if you click on that, and your properties window, you have different fit methods. Now my relationship here between age and total competency is pretty clearly linear which means it’s going in a straight line upwards, but if you have a different shape to your trend, you had options here for quadratic or cubic etc. I’m going to leave it as linear. Close. Now when I do that I get an r-squared value and our squared value is coming from an r or correlation, and it’s also something which comes out in a regression model. Now we usually interpret our square by multiplying it by 100 to turn it into a percent, so that would be 5.2%, which is very, very low and what it tells us is that approximately 5.2% of the variation in competency between our employees can be explained by age, so not very much. That means there’s roughly 94/95% of total competency scores that cannot be explained by someone’s age.

We’re going to go ahead and close this. Now this one thing I’m going to show you, is a common mistake students make is to try to use ordinal variables when they make a scatterplot and they get something which looks a bit strange. So if we go to graphs, I’m going to show you what will happen. Hit the reset. Scatter in, and I’m going to choose two of my ordinal variables. So Organisational Commitment and, I’ll use Satisfaction. Click ‘okay’.

Now you can see it’s very difficult to notice a trend, and that’s because along the bottom I only have five possible categories, and along the side I only have five possible categories. So that means I only have a possibility of 25 different points that could be on this graph and sometimes you’ll get where all 25 points are filled in, so all of them have bubbles, so this is one reason why ordinal data is not suitable for a scatter graph because it’s very, very difficult to see any relationship or trend, so make sure that your data is scale when you try to do this type of graph.

END.