

---

This is your tutorial on scatterplots. Scatterplots are a particular type of graph. With a scatterplot, you have both an x and a y-axis. And on each axis is a piece of quantitative data or rather a variable that is quantitative.

So what you're doing is you are comparing two different variables on one graph. Now the coordinate point, so when you go over 4 and up 5 and place a dot, that point there is telling you what the values are for that one point. So it tells you what it is for variable 1 and for variable 2.

Now the information can be about people, places, things, whatever you've collected your information for, as long as it's quantitative. A great thing about the scatterplots is that you can easily show multiple data sets onto one plot. Now the way this is done is by using different symbols to represent the different data sets. We'll see an example of that. And we'll also go through an example of making a scatterplot.

So here, this information is telling us the age of a child and the height in inches. And we have age 1, 2, 3, 4, 5, and 8, and then the corresponding height for the children of those ages. Now when you have your scatterplot, you have an x-axis and a y-axis.

And your variables, one is going to go on each. A later tutorial will cover more about how to decide. But in this case, I'll just do it. We'll put age down here and have height on the y.

Now it's important to think about the fact that when we have our axis, we know that it's going to be age. But we have to decide how we're going to label the age. Now we have to do our spacing evenly.

So even though we skip from 5 to 8, when I'm labeling, I shouldn't do 1, 2, 3, 4, 5, 8. Because that's not even spacing. So if I'm going to do 1, 2, 3, 4, 5, I've skipped 1 each time. So I have to do 6 and then 7 and then 8.

Now what I'm trying to do to the height, I need to go from 25 up to 55. So I need to fit that span on this axis here. So you can do a couple of things. You can start from 0 and then count by 5s or 10s, or you can have a break in your graph.

Sometimes putting breaks in your graphs and starting counting from 20 can be misleading. So I choose not to do that. I'm going to count by 10s. So we have 10, 20, 30, 40, 50. Now when I'm placing my tick marks, I'm just estimating. But you should measure it out and make sure that it's properly and perfectly evenly spaced.

So now I have to plot my points. So for the first one, the age is 1, and the height is 25. So I'm going to go over to 1 and then up to about where 25 is. Then 2, 31, 3, 34, 4, 36. And see how I'm already up almost in line with the 40, that means I haven't done a good job estimating. So I'm going to try to get the 3, 34 a little bit lower. The 4, 36 a little bit lower, 5, 40. Now when you're making your graph, it is important to readjust like that if you notice you've been estimating poorly.

And on the last point, we're going to go over to 8. And then, oh, I only made my axis up to 50. But I need 55. So I'm going to add on 60. So that I can graph over 8 and up 55. Now you can adjust your graph as you're going if you realize you need more numbers, or if you have estimated poorly. So this is going to be our scatterplot for age versus height. You don't need to connect the lines. You just place on the points.

Now we'll see an example of multiple data sets on one scatterplot. So here, I know the points are a little bit small. But it's talking about Airbus and Boeing and comparing passengers and range. So we have range going across the bottom in thousands of nautical miles. And across the y-axis, they have the number of passengers. So it counts 1,000, 2,000, 3,000, 4,000, 5,000, and so forth.

Now what they've done is they've used these little red squares to indicate Boeing. And there's a key down the bottom that tells us that the red square is for Boeing. And then the other set of data, the blue diamonds, are for Airbus.

So when you have multiple data sets on one scatterplot, it's important to choose symbols that are different, choose symbols that are different enough that it's easy to see. So they did diamonds and squares. But they also changed the color.

And then you also need to give a key. So you need to tell us what this blue diamond means versus what this red square means. So the key is going to be very important. Otherwise, you can't read all the information from your multiple data sets.

And then both data sets are similar. They're both talking about airplanes. And they're both comparing passengers and range. So you can't put data sets together on one graph that really have nothing to do with each other because then you're not going to be able to draw any useful conclusions, and you might be misleading. This has been your tutorial on scatterplots.