VIDEO: “Histograms with Continuous Data” (2:11)

(00:00)
Unfortunately, most scientific data does not take on discrete values. Rather it might take on any value in an interval or range. For example, suppose a researcher measures the height of 100 corn plants. The data might look like this- 3.01 centimeters, 5.2 centimeters and so on. There are some issues supporting our previous support because it might be hard figuring out what our unique value are. There is a difference between 3.02 and 3.01 with an experimental error. The solution is to group or pin the data and count how many values fall in each group. We can group any type of data but let’s see how to do it for continuous data.

(00:38)
Binning means is dividing the data up into groups and counting how many values are in each group. We will start with the highest- 3.01 cm, 5.2cm and so on. Suppose the smallest value in the data set in zero meaning the plant didn’t come up and the largest value 8 cm. The range is 8 which is the difference between the max value and the minimum value. The first data value 3.01 goes into pin 4 while the second data value 5.2 goes into pin 6 and so on. After we process the data, we might have a frequency table that looks like this. There were 30 values between 0 and 1 inclusive and only 20 values that were greater than 1 and less than or equal to 2.

(1:31)
A histogram displays a frequency table and shows the same information in graphical form but the frequency table on the left the histogram looks like this. It is 8 bars corresponding to 8 groups or rows of the frequency table the heights of the bars show the counts columns of the frequency table.

(01:51)
One point of confusing shows the difference between the histogram and a bar chart. Let’s compare the two- in a histogram the heights of the bar represent the counts. In a bar chart they represent the actual data values. The x axis plots something else in this case plant id.