VIDEO: "Histogram features" (3:23)

(00:00)
Now we will go over the features you should look for using the Daphne island example. First, look at the general shape this is diagramed as bell shaped it is also symmetric meaning that you can flip the picture around the vertical line in the center and the histogram will look the same. Try symmetry line sometimes helps in deciding how symmetric the data is. By reading the x position in the left most and the right most bin, you can estimate the range of the data- I will guess roughly 6-14mm. When I calculate the min and the max directly, I find 5.8-13.9 mm. I guess I could have made a better estimate by reading off the bin positions more carefully. By estimate the highest values I look at which bin positions have the tallest bars. We are going to estimate that most of the values are between 8.5-11mm. When I actually calculate how many values are in this interval from the data, I find that 81% of the values are in this interval. We should also look for any outliners. Notice that there is a box out there way on the right that is just under 14mm, I guess it said there were a couple of birds in it and it turned out to be 2- just a luck guess on my part. Finally, by looking at the areas of the bins you can estimate the median of the data. The median is the vertical dividing line- half of the shaded area is on the left and half is on the right. I estimate about 9.5- the actual value was about 9.7mm. It is not very easy to estimate the data set average from the histogram. However, if your area is perfectly symmetric your two values would be the same. I calculated the mean of the data to be directly 9.6mm which is pretty close to 9.7 supporting the observation symmetry. However, be careful with this just because the mean and median are close together does not mean that distribution is symmetric. You actually have to look at the histogram for this.

(2:00)
There are a number of histogram shapes that you are likely to encounter in real data. The bell-shaped curve is the most familiar. However, in biological and other physical system you will often see a long tail shape it looks bells shaped but it is not symmetrical. Rather it has a long tail to one direction usually to the right and on the other side it may cut off. Another off shape is decaying it starts at a certain level and just drops off smoothly without any bumps. Lab shapes are also possible. Bimodal shapes have two bumps rather than one. Sometimes you see a bimodal shape when you use histogram grades for a course- there was some people who understood the material and some who did not. We will talk more about what these shapes tell you in the lesson of distribution and sampling.

(02:48)
This video gave general guidelines about interpreting histograms which we will summarize here. Start by estimating the standard features. Unusual values are unexpected shapes are starting points for investigation. Use rulers or draw lines to increase accuracy. Histograms with different bin sizes give you different information. It is easier to estimate data information using few bins than a small number of bins may not give an accurate shape. Try different bin sizes to improve your interpretation and finally verify your observations by calculation.