Video: “Comparing mean and median” (3:44 min)

Slide 1 “Comparing mean and median measures of central tendency” (00:00)

Measures of typical behavior help us characterize the data.

Slide 2 “Mean and median” (00:08)

The mean and the median are two such measures. They capture central tendency- or middle of the road behavior. In this lesson, we’ll show how to compute them and give a geometric interpretation. We’ll also discuss what the mean and the median can tell us about the data.

Slide 3 “Mean” (00:25)

The mean, which is sometimes called the average or arithmetic mean, is computed by adding up all the values and dividing by the number of values. In this example, the sum of the values is five and we divide by five to get one. You’ll often see the mean designated by an overbar over the variable name. You may also see this sigma notation indicating, “Adding up the values.”

Slide 4 “Geometric interpretation of the mean” (00:53)

Now let’s look at a geometric interpretation of the mean. Let’s start with the five values from the previous example, and we’ll put them on a number line arranged in their order. The mean can be thought of as the balance point or center of mass with equal weight on either side. The points are weighted by their distance from the mean. In the example, -2 is a distance 3 from the mean and 0 is a distance 1. The weight on the left hand side is four and this is balanced by a similar weight on the right hand side 4. When we move our fulcrum to the left, the line becomes unbalanced. The weight of the left is 2 which is too small so the mean must be to the right.

Slide 5 “Median” (1:43)

We compute the median by sorting the values and taking the one in the middle position. In our list here, after we’ve sorted, the middle value or median is one. In the case that the list has an even number of values; after we sort, we average the middle two values so this new list with six values has a median of 1.5.

Slide 6 “Geometric interpretation of median” (02:10)

The median also has a geometric interpretation in terms of center of mass, let’s look at an example. On the number line, we order our points and put our fulcrum at the median. The median is the balance point or center of mass if points each have a weight of one. In the example, there are two points to the left of one and two points to the right, so the line is balanced. As before, when we move the fulcrum to the left, the line becomes unbalanced.

Slide 7 "What happens for outliers?" (02:44)

At this point, you may be wondering why we need both mean and median if they give the same value. For most data sets they don't, and they tell us important things about the data. In particular, they give different information about the outliers. Let's look at an example. We replace the value 4 with 100, and the mean becomes 20.2, a significant difference from 1. The number line shows us why we need to move the fulcrum to 20.2. In contrast the outlier did not affect the median. The median is still 1. The line is balanced. When you look at a new data set, it's always good to compute both mean and median. If the mean and the median are close together, it is likely your data is symmetrically separated or clumped together. On the other hand, if the mean and the median are far apart, you should look for outliers, or perhaps your data is skewed in one direction.