

Robert_S2_L05

Wed, 1/12 12:25PM 7:46

SUMMARY KEYWORDS

median, equal, values, central tendency, sum, observations, measure, deviations, midpoint, important, series, average, larger, smaller, average person, portfolio, outliers, variability, x minus, society

SPEAKERS

Robert McKeown

R

Robert McKeown 00:06

Hello, and welcome to our video comparing the median and the mean. Now, as we compare them, I want you to keep in mind that both the median and the mean are an average. So both of them represent an average, there are just two types of measuring the average. And both the median and the mean are a measure of central tendency, a measure of central tendency, so where do the values tend to center in a series now, let's take a look at a definition of the mean. The mean is the middle point, it is the middle point at which the value below and the values above the mean are equally weighted, that is the idea of a fulcrum of a seesaw. And also, the sum of the deviations. Remember, deviation was $X - \bar{X}$ was the observation each of the observations or each of the values within a series of values, take each of those values and subtract the mean from them. If we sum them all up, from i equals one to N using sigma notation, that sum is going to be equal to zero. And I'm going to demonstrate that to you in a moment with an example. The median is the middle point of a set of cases, the value that occurs at the midpoint of a series that is sorted from smallest to largest value. Importantly, the median is insensitive to extreme values, while the mean is sensitive to extreme values. For example, the median won't change if one of the if the second largest observation among 100 observations becomes larger, the median won't change, it won't change. While the mean if the second largest observation becomes even larger, that would influence the mean. In fact, if any of the values in a series change become larger or smaller, all else being equal, the mean is going to change while the median change, not necessarily. Now let's look at a very simple example. We've got X and within X , there are three values, two, seven, and nine. Now we can calculate the mean, the mean is just equal to two plus seven plus nine divided by three, which is going to give us 18 over three, and 18 divided by three is equal to six. So we've got a mean equal to six. Now let's take a look over here and let's calculate the deviations from the mean for each of these observations in X . So we could write this as $X - \bar{x}$. And that's going to be equal to you set notation to minus six, seven minus six, and nine minus six. So we're going to have a set of negative four. That's positive one, Seven minus six, and three. And notice that if we take the sum of our deviations from the mean, we're going to have negative four plus one plus three, which is just equal to zero. And this is always true for any series of values. If we calculate the mean for a series of values, and then we subtract the mean from each of the observations, the sum is always going to be equal to zero. So that's something important to keep in mind. It's good to be useful in the next section when we start Talking about variability.

It's an important understanding for many of the concepts and statistics, and as well as things like regression. Now we were talking about the mean here. What is the median? Well, the median is pretty easy to find out from X, it's the midpoint, we have a two, seven and a nine. The median here is going to be seven. So what can we say about that, if we've got the mean, is equal to six, the median is equal to seven. Well, we can see that if we let this if we let to become say, much, much smaller. So suppose we had X was equal to this set of negative 10. Seven, nine, notice that the median is still seven. While the mean would be negative 10 plus seven, plus nine over three, which is going to be equal to two. So looking at this very simple example here, and I apologize for writing it off, and kind of an awkward space on the slides. But hopefully, you can understand that the median here is not being influenced by the fact that two is quite a bit smaller than seven and nine. And so in the social sciences, business and economics, you want to ask yourself, do you want your measure of central tendency or average, to be impacted by these outliers? Sometimes the answer is yes, we do want to be impacted by outliers, it's important to consider that maybe you're looking at the history of a return on your portfolio, you probably want to keep in mind that you might lose 30% of the value for portfolio in a given year, like in 2008 with the financial crisis. But if you're looking at something like household income, and you've got a society with unequal wealth, where a few people own a lot of the wealth, maybe a better measure for the sort of average household, the average person in society maybe or excuse me, the median, maybe the median is going to give you a more accurate representation of how the average person lives.