Robert_S2_L10

Wed, 1/12 12:25PM 10:25

SUMMARY KEYWORDS

squared, square root, standard deviation, deviations, kilometers, write, sum, calculate, number, divide, step, variability, auto theft, calculator, units, equal, subtract, squared deviations, values, compute

SPEAKERS

Robert McKeown

R Robert McKeown 00:04

Let's continue our analysis of standard deviation. So right here we have a recipe to follow. And the first step is to compete, compute the mean. And then we want to subtract the mean from each score. So I could write it like that, or I could write it with the I sometimes it's just easier to write it without the i. And then we want to square each individual deviation from the mean. And we want to add them all up. And then we want to divide by n. So we're gonna have like so. And then finally, we want us to take the square root of this whole thing, right there. And there's our formula. So we've got a example here, we've got auto theft cases reported in the city of Toronto, for the years, 2018 2019, and 2020. And we want to calculate the standard deviation of cases over these past three years. Well, with the app, we're going to calculate the standard deviation, and auto theft cases reported in the City of Toronto, we're also going to do that we need to first compute the mean. So what was the average number of reported auto theft cases in the City of Toronto, so if we want to know that, and we'll call this, we'll call it x bar. So x bar is going to be equal to 4946 plus 5461, plus 5820. Now we want to divide 16,227 by three, and that's going to give us the mean. Which is equal to 5409. Now that we have the mean, we want to subtract that mean from each of the observations. So we're going to have 4946 minus 5409, which is going to give us negative 463. Their next observation is this 5461 minus the mean of 5409. Which is going to give us 52. And 5820 minus 5409 is going to be equal to 411. So this is step one there step one, this is step two, we've now completed step two, so what about step three? While we need to square each of these deviations from the mean? And so we can do that. Negative 463 squared 52 squared, and 411 squared. And I'm going to use my calculator to make it easier. And there are the squares. And now we've completed step three, what about step four? While we want to sum all the squared deviations, so we want to add these up. Again, I'm gonna use a calculator, make it a little bit faster. And I get 385,994. And that is step four, maybe we could write the form a little bit closer. That's for now what about step five? We're good to just take 385,994 And we're going to divide it by three. That's going to give us 125 1664 and looks like two thirds, but I'll just write it as decimal six, seven. And now, step six, take the square root of 128 664 67. And we definitely want to use a calculator for this one, and I get a rounded answer 359. We'll round it to a whole number. And our standard deviation of auto theft cases in Toronto over the past three years, is on average 359. And that tells us what the average distance from the mean was for these auto theft cases that were reported. Here's an example that I like because it demonstrates a really useful property of standard deviations.

We've got distances from Calgary to Winnipeg, Calgary to Bismarck in the United States. So if I told you that the mean 1,223.3 kilometers, then you could calculate the deviations. Looking at our answers that I hope you're able to calculate on your own, you can see that we've got our mean here we've got our x bar, we've got our x's minus x bar in the deviation column, we've got X minus X bar squared, then when we sum up our x minus x bar, or deviations from the mean, we get 42,340.7. This value right here, there's our sum when we divide by three, right here, we've got 14,113.6. And this has special significance, it's actually called something called the variance. We'll discuss that a little bit later. When we take the square root of 14,113.6, we get the standard deviation 118.8. Earlier, I told you that I really liked this example, because it demonstrated a useful property of standard deviation. And I guess I should make sure that it's written right here. So we know what it is. Now, you might ask yourself, why do we what is where's this formula that we're using coming from? Why do we add up all the squared deviations, then take the square root of these things? And there's a good intuition why. The question here is why do we square each fat value? So we'll break that question into two smaller questions. Well, when we square the deviation from the mean, it removes negative values. So if we're 20 units below the mean, that's equivalent to being 20 units above the mean, squaring removes the negative sign, negative 20 times negative 20, is going to be 400. Just like 20 times 20 is equal to 400. That's useful, because remember, the sum of these deviations here is always equal to zero, the sum of the deviations is always equal to zero. That's a property of the mean, when we take deviations from the mean, the sum of those deviations from the mean is always going to be equal to zero. If we simply add up the deviations, we're not going to get anywhere, we're just going to end up with zero every single time, we're not going to have a deeper understanding of the variability in a series of numbers. So that's why we square the deviations from the mean. Now why do we take the square root of the sum of the squared values? So we had this number right here, and then we took the square root of it. Why did we do that? Well, importantly, we did that to return the values we're working with back into the units from which they started, this standard deviation here is and kilometers. This observation here is in squared. kilometers, and it's a lot harder for me anyway to think about squared kilometers and work in square kilometers than it is for me to work and just plain old regular kilometers. The reason why we square it so that we can get a better by a measure of the variability or better measure the variability. The reason why we take the square root is to return the numbers into the units of which we started here we're looking at kilometers. So when we take the square root, we return to kilometers