integrated learning

Statistical Measurements

for the

GMAT MATH SECTION

Image that a politician in your state is up for re-election. Trying to impress the public with the work she has done, the politician announces at a press conference that at the end of her term the average yearly salary per person in the state is $80,000.

Would you support her based on that number? It sounds good – it *sounds* like everyone in her state is making $80,000. But can we be sure? A reporter wants to know if her statistic is valid, and he picks three groups of five people randomly from the state. Each group has the same average, $80,000. Let's look at the different groups now.

	Group 1 (avg. = $80,000)	**Group 2 (avg. = $80,000)**	**Group 3 (avg. = $80,000)**
Person 1	80,000	71,000	15,000
Person 2	80,000	74,000	30,000
Person 3	80,000	82,500	30,000
Person 4	80,000	82,500	100,000
Person 5	80,000	90,000	225,000

What can we learn from this chart? Group 1 represents what the politician wants her state to believe: that everyone earns $80,000. Group 2 represents what most people would see as an acceptable situation: some people earn more, some earn less, but they all earn about the same and together they average $80,000. And Group 3 represents what is probably most likely: that most people don't earn very much at all but a few people earn so much as to sway the average in their favor.

Statistical measurements are used to condense the behavior of multiple actors into a single number that we can work with and comprehend. Unfortunately, the average is a particularly poor statistical measure, as the same average can represent an unlimited number of situations.

How do we avoid this pitfall then? There are several other important statistical measurements that help us understand the average and paint a more complete picture of what is happening. Those measurements are *Range, Median, Mode, and Standard Deviation.*

Statistical Measurements
Range **Median** **Mode** **Standard Deviation**

By understanding these four measurements on top of the average, you will be able to evaluate a situation more sophisticatedly than ever before. You'll also do well on the GMAT. We will start with *Range, Median, and Mode*, the continue with *Standard Deviation*.

Range

The range of any list of numbers is the difference of the highest and lowest numbers.

Example 1

Find the range of the lists below:

I.	8, 8, 8, 8, 8
II.	1, 5, 3, 14, 17
III.	6, 8, -14, 7, -2

Answer

I.	Range = 8 – 8 = 0
II.	Range = 17 – 1 = 16
III.	Range = 8 – (-14) = 22

Median

The median of any list of numbers is the number in the middle when the list is arranged in numeric order. If there is no middle number, the median is the average of the two middle numbers.

Example 2

Find the median of the lists below:

I.	3, 4, 5, 6, 7
II.	8, 14, 5, 6, 2, 21, 21
III.	8, 14, 5, 6, 2, 21

Answer

I.	Median = 5
II.	First arrange the list in order: 2, 5, 6, 8, 14, 21, 21. Median = 8
III.	First arrange the list in order: 2, 5, 6, 8, 14, 21. The median is the average of the two in the middle, 6 and 8, or 7.

Mode

The mode of any list of numbers is the number that appears the most often. If two or more numbers appear the same number of times, they are all considered the mode.

Example 3

Find the mode of the lists below:

I.	3, 5, 4, 5, 7, 5, 2
II.	4, 2, 9, 4, 2

Answer

I.	Mode = 5
II.	Mode = 2 and 4

So, now that we understand Range, Median, and Mode, let's see them applied to our politician's claim from above:

	Group 1 (avg. = $80,000)	Group 2 (avg. = $80,000)	Group 3 (avg. = $80,000)
Person 1	80,000	71,000	15,000
Person 2	80,000	74,000	30,000
Person 3	80,000	82,500	30,000
Person 4	80,000	82,500	100,000
Person 5	80,000	90,000	225,000
Range	**$0**	**$19,000**	**$205,000**
Median	**$80,000**	**$82,500**	**$30,000**
Mode	**$80,000**	**$82,500**	**$30,000**

As you can see, it is possible to use these statistics to help understand what an average is really all about. By pressing for more information, our reporter will be able to further understand what an average of $80,000 means, without having to study the state's entire population list. An average of 80,000 with a range of 205,000 and a median and mode of 30,000 is clearly a very uneven list. But no statistic describes an average with the clarity that Standard Deviation does.

Standard Deviation

A. Introduction

As we have seen, every list of numbers has an average. Also as we have seen, the numbers of the list do not need to actually be anywhere near the actual average. In fact, most numbers in the list differ, or *deviate* from the average by some degree.

The *standard deviation* tells us generally how the numbers will deviate from the average. The greater the deviation, the higher the standard deviation.

Integrated Learning Tip: While there is a complicated and intricate Standard Deviation formula, the GMAT does not require you to know it. Your goal in learning Standard Deviation should be to understand what it is, how it works, and what it represents, rather than how to calculate its value.

Example 4
Which of the following lists has the highest standard deviation?

I. 8, 8, 8, 8, 8
II. 6, 7, 8, 9, 10
III. 4, 6, 8, 10, 12

Answer: List III
You can probably see quickly that all three of the above lists have the same average, 8. So which list has the most deviation from 8? In List I, all the numbers are the same as the average. There is no deviation at all. It is List III that shows the greatest deviation.

Example 5
Which of the following lists has the highest standard deviation?

I. 6, 7, 8, 9, 10
II. 4, 6, 8, 10, 12
III. 18, 20, 22, 24, 26

Answer: II and III have the same Standard Deviation
This is a trick question. While the average of List III is higher than List II, the numbers of List III are the same distance from its average of 22 as the numbers of List II are from its average of 8. They are both lists of even numbers, and as such, have the same dispersion. Therefore, they have the same standard deviation.

B. Counting by Units of Standard Deviation
Though we don't need to know how to calculate standard deviation, we do need to know how to use it. The standard deviation of a certain list is always expressed along with the average of the same list.

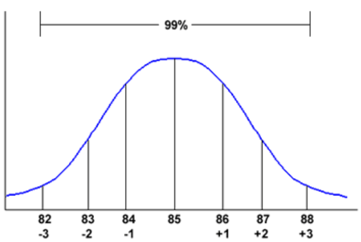
For example, one might say that the average gambler's winning in Las Vegas is $200, with a standard deviation of $30. It is then possible to count in units of standard deviation. One standard deviation would be $30, two standard deviations would be 2x30=$60, and three standard deviations would be 3x30=$90.

Example 6
Farmers in the United States grew an average (arithmetic mean) of 80 tons of corn each, with a standard deviation of 10 tons. What value is two standard deviations away from the mean?

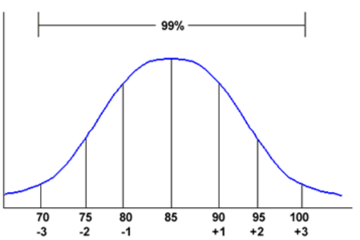
A) 90
B) 82
C) 78
D) 75
E) 60

Answer: 60
Since the standard deviation is 10, two standard deviations is 2x10=20. Standard deviations can be counted above and below the average, which, in this case, is 80. So two standard deviations away from 80 would be 20 tons less or more. 80 – 20 = 60.

C. Standard Deviation in Practical Use
Imagine a situation in which a college professor is bragging to his colleagues about a recent exam he gave his class. The exam was difficult, he tells them, but the average grade was still 85%. He interprets this to mean that he is a good teacher – his students have learned well. Hearing this, his suspicious colleague asks him what the *standard deviation* of the grades was. Let's look at and graph two possible scenarios.



Average Grade: 85
Standard Deviation: 1



Average Grade: 85
Standard Deviation: 5

It is a statistical fact that 99% of all the numbers of a given list will lie between –3 and +3 standard deviations away from the mean. Therefore, we can see in the graphs above that the range of test scores in the case where the standard deviation is 1 is from 82-88, or 6 points, while the range of scores when the standard deviation is 5 is from 70-100, or 30 points.

Which one corroborates the professor's claim that his students have learned the material well? The graph with the lower standard deviation, the first one, tells us that his students all generally scored in the same range, while the second one, the one with higher standard deviation, tells us that his students' scores were more erratic. Yes, some students scored very well, but others scored very poorly. The first one supports the professor, while the second proves his colleague's suspicions correct.

Now let's take a look at our original example, the one with the politician. How erratic are her numbers, the ones that so neatly averaged out to $80,000? Let's see what the standard deviation is of our three scenarios.

	Group 1 (avg. = $80,000)	Group 2 (avg. = $80,000)	Group 3 (avg. = $80,000)
Person 1	80,000	71,000	15,000
Person 2	80,000	74,000	30,000
Person 3	80,000	82,500	30,000
Person 4	80,000	82,500	100,000
Person 5	80,000	90,000	225,000
Standard Deviation	**$0**	**$6,774.95**	**$78,294.32**

So we can see that when there is no deviation at all, as in Group 1, the standard deviation is necessarily 0. When there is slight deviation, as in Group 2, the standard deviation is also slight. When the numbers vary wildly from one another, as in Group 3, the standard deviation is very high.

The politician is caught. When pressed about her statistic, she tells the audience that the average salary in her state is $80,000 with a standard deviation of $78,294. The reporter sees through her empty average, and slams her in the papers.

D. Standard Deviation GMAT Practice Questions

1. An Olympic diver received the following scores: 6.0, 5.5, 7.0, 6.5, and 5.0. The standard deviation of her scores is in which of the following ranges?

A) 0 – 1.9
B) 2 – 3.9
C) 4 – 6.9
D) 7 – 7.9
E) 8 – 9.9

2. Is the standard deviation of Roberta's 8 test scores higher than the standard deviation of Melissa's scores on the same 8 tests?

1) Melissa's average score was 78 and Roberta's average score was 85.

2) The range of Melissa's test scores was 0 and the range of Roberta's test scores was 10.

3. What was the standard deviation of Company R's earnings per day in January and February of this year?

1) The standard deviation of Company R's earnings per day in January of this year was $2.3 million.

2) The standard deviation of Company R's earnings per day in February of this year was $2.3 million.

Answers

1. A; 2. B; 3. E

Explanations

1. Ans: A

Concepts tested:

1. Standard Deviation can never be larger than the range of the numbers. 2. Average of consecutive numbers is the number in the middle.

Explanation:

DO NOT calculate the standard deviation of these five numbers. That's what they want you to do, and it will take A LOT of time. Rather, it's important for the GMAT that you understand the *meaning* of standard deviation. It will save you a lot of time on the test.

Standard Deviation describes the way numbers are organized around their average. The greater the standard deviation, the greater the range of numbers, and vice versa. But most specifically, the standard deviation actually tells us the range of the numbers: If you were given an average of 200, and a standard deviation of 2, and the numbers would fall in the range of 198-202, or 196-204, depending on the number of standard deviations. Compare that with an average of 200 and a standard deviation of 15. It's the same average, but now the numbers are organized much farther way, between 185-215, or 170-230.

Look at this list – the average is easily calculated at 6.0 (another GMAT trick – since they're consecutive, it's the number in the middle!) and the range is from 5 to 7. That means the greatest distance any number is from the average is 1. In that case, the standard deviation could never be more than 1! Even if there were 200 numbers between 5 and 7, the SD wouldn't be larger than 1. So the answer has to be A.

2. Ans: B

Concepts tested:

1. The standard deviation of a list who's range is 0 is 0. 2. Standard deviation can only be derived from the numbers in the list, and has nothing to do with the average.

Explanation:

The first statement tells us that Roberta's average score was higher, but that cannot be enough information because it does not tell us about the dispersion of each of the girls' scores around the average. Remember that any average number comes from a list of numbers and the standard deviation describes how those numbers are organized in the list. If you were given an average of 200, and a standard deviation of 2, and the numbers would fall in the range of 198-202, or 196-204, depending on the number of standard deviations. Compare that with an average of 200 and a standard deviation of 15. It's the same average, but now the numbers are organized much farther way, between 185-215, or 170-230. Based on that, the averages alone cannot describe the list of numbers they come from, and this statement cannot answer the question.

The second statement tells us that the range of Melissa's test scores was 0. That means there is no dispersion, and thus, her standard deviation must be 0 as well. Since Roberta has a range of 10, there must be some dispersion, so she must have a higher standard deviation than Melissa.

3. Ans: E
Concept tested:
Standard deviation can only be derived from the numbers in the list, or, standard deviation tests dispersion around the average.

Explanation:
Think about what you have learned about Standard Deviation: it is the way the numbers in a list are organized around the average. For example, if you're given an average of 10 and a standard deviation of 2, you know that most numbers are organized in the range of 8-12, or 6-14, depending on how many standard deviations away from the average we're talking about. If you were given an average of 200, you could still have a standard deviation of 2, and the numbers would fall in the range of 198-202, or 196-204, also depending on the number of standard deviations. Compare that with an average of 200 and a standard deviation of 15. It's the same average, but now the numbers are organized much farther way, between 185-215, or 170-230.

From this you should see that standard deviation refers to the *dispersion* of the numbers, not the actual average, and so given *only* the standard deviation, we know nothing at all about the average or the numbers used to make that average.

As you learned in the lesson, the formula for standard deviation involves using all the numbers in the given list of numbers. In the case of this problem, the numbers would have to be the company's earnings each day in January and each day in February.

Statement 1 tells us only what the standard deviation in January was, but tells us nothing about the actual numbers being used. We know that the standard deviation was $2.3 million. But what was the average? It could be $10 million or $200 million, for example, and still have the same standard deviation. It's not enough to describe what happens in February.

Statement 2 presents the same issue.

We are left with C or E. Since each month has the same standard deviation, does that mean that they will have the same standard deviation together?

That’s what they want you to think, but it’s the other way around completely. Standard deviation is strict – in order to find it, you MUST have the numbers in the list to figure out the average first, and then the standard deviation. Since we don’t have the daily company earnings for this company in January and February, we cannot find the standard deviation for that time.