SCENARIO:

- You are trying to decide if you want to take a class in school based on how the difficult the class is. You decide to use the grades of students who have taken the class previously as a measure of difficulty.
- What are some ways of looking at the data to make your decision?

Measures of Central Tendency

Median: The middle score in a rank-ordered distribution.

If the median score is 85%, would you consider this an easy class?

What if you found out that the grades were

42, 44, 50, 85, 85, 85, 85?

Is median a great measure of central tendency?

Measures of Central Tendency

Mode: The most frequently occurring score in a distribution.

If you find a class with a mode of 86 would this be an easy class?

Here are the grades:
14,25,32,45,50,60,86,86.

Is mode a great measure of central tendency?

Measures of Central Tendency

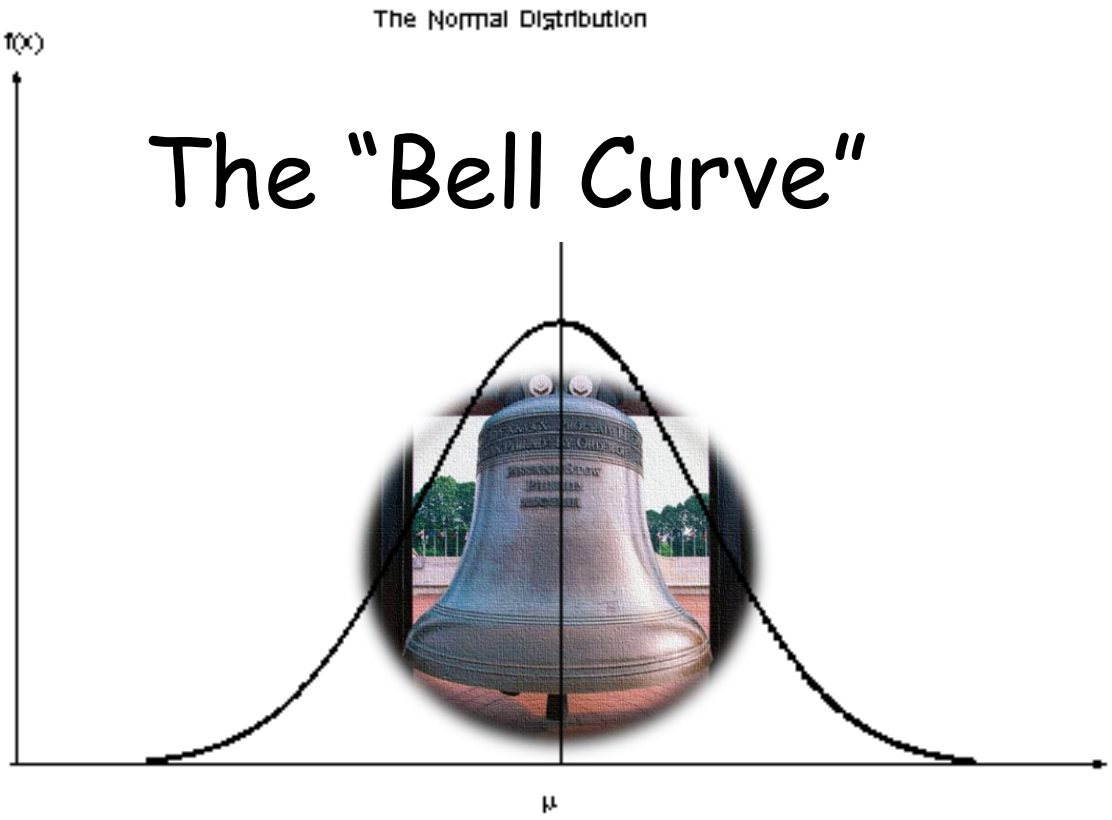
Mean: The arithmetic average of scores in a distribution obtained by adding the scores and then dividing by the number of scores that were added together.

You have found a class with a mean of 85 and have decided that this must be an easy class.

The grades were: 70,70,100,100. Would you feel confident that this an easy class?

Normal Distribution

- In a normal distribution, the mean, median and mode are all the same.

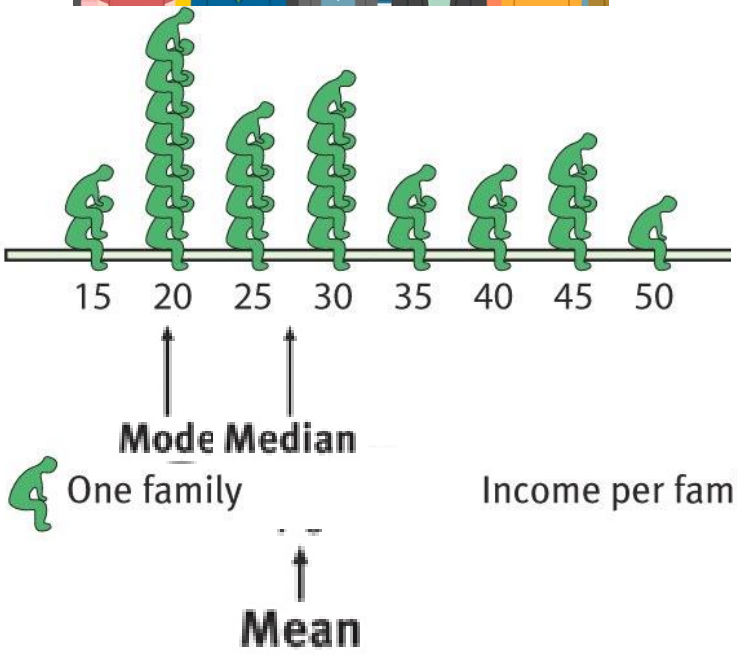


Measures of Central Tendency

- It is important to always note which measure of central tendency is being reported. If it is a mean, one must be especially alert to a few *atypical* scores.
- These scores could be distorting the data or causing a ***skewed distribution***.
- ***Skewed distribution***: When scores don't distribute themselves evenly around the center. (There are a few extremely high or low scores.)

Measures of Central Tendency

A Skewed Distribution

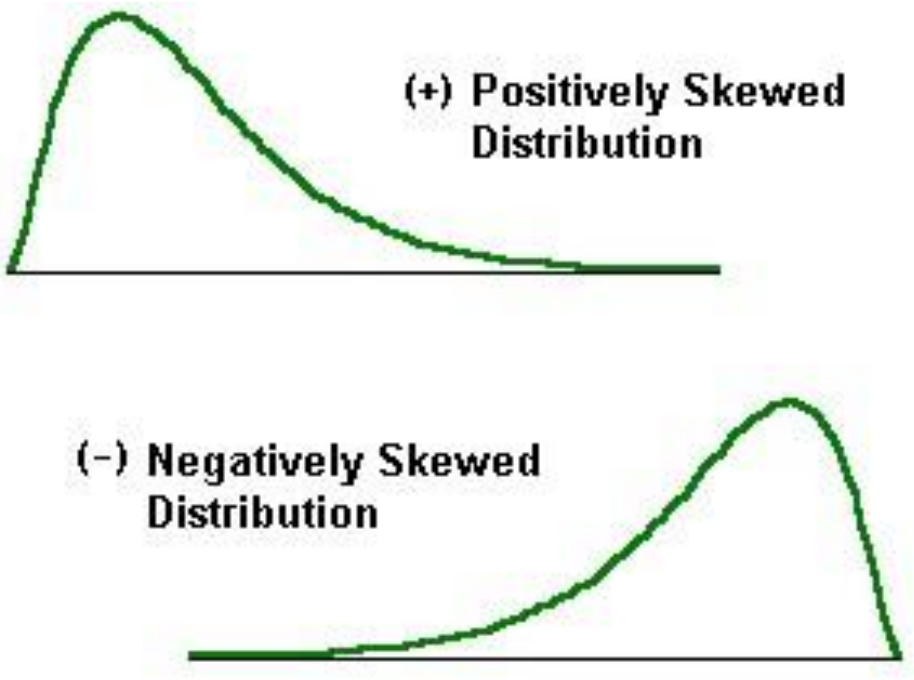


Panera BREAD®

Hey everyone! Is this seat taken?

Distributions

- **Outliers** skew distributions.
- If group has one high score, the curve has a positive skew (contains more low scores)
- If a group has a low outlier, the curve has a negative skew (contains more high scores)



Central Tendency

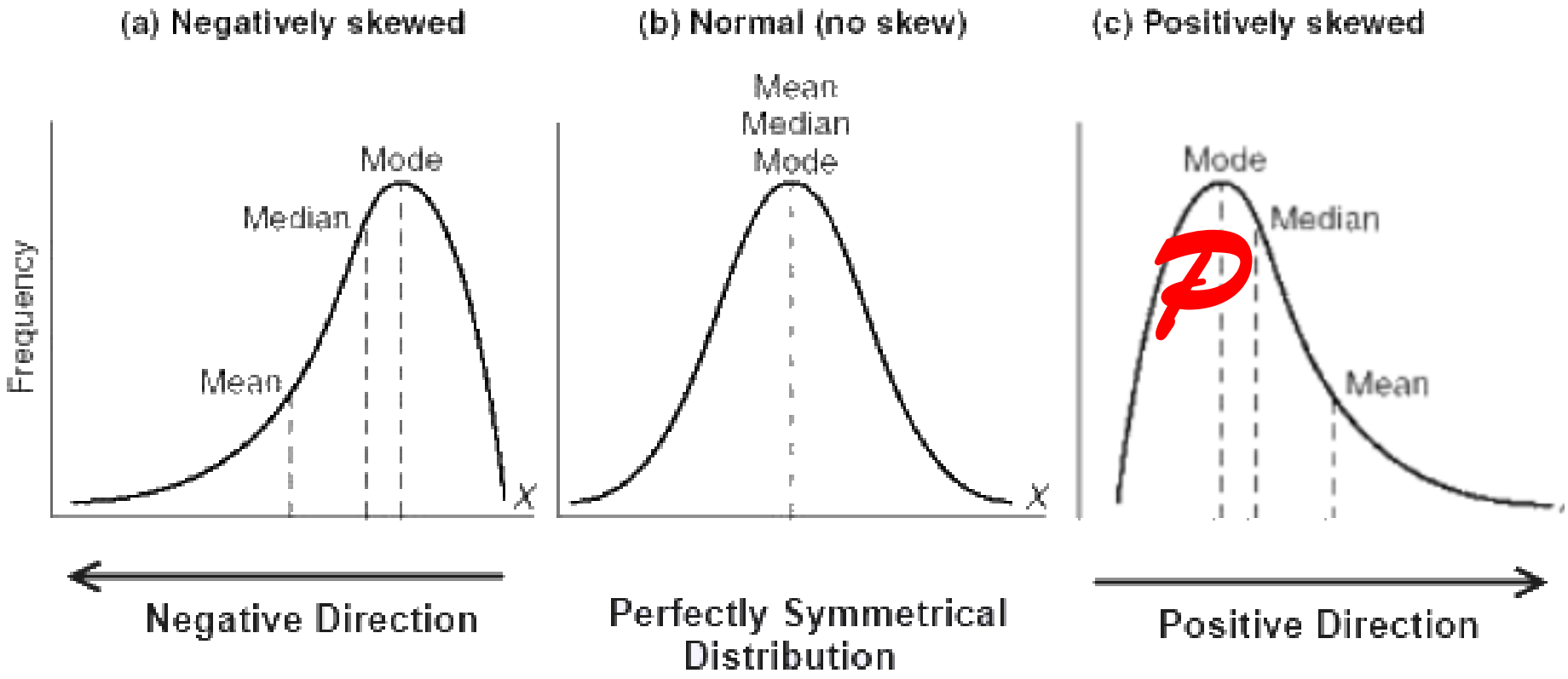
- Mean, Median and Mode.
- Watch out for extreme scores or outliers.

Let's look at the salaries of the employees at Dunder Mifflen Paper in Scranton:

$25,000-Pam
$25,000- Kevin
$25,000- Angela
$75,000- Andy
$75,000- Dwight
$75,000- Jim
$350,000- Michael

Measures of central tendency are Quick and easy, but outliers may distort the numbers.

Normal and Skewed Curves



Measures of variation

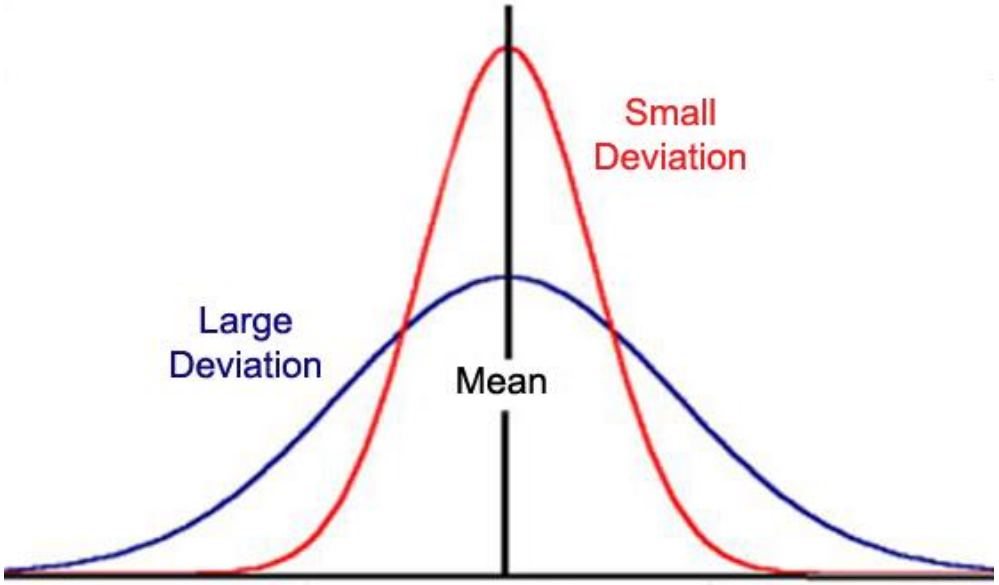
- Averages from scores with low variability are more reliable than those with high variability.
- **Range:** Difference between the highest and lowest scores in a distribution. Like with the mean, high and low scores could present a deceptively large range.

Hey diddle diddle,
the median's the middle;
YOU ADD AND DIVIDE FOR THE MEAN.
The mode is the one that appears the most,
and the range is the difference between.

Measures of Variation

Standard Deviation: A computed measure of how much scores vary around the mean.

Standard Deviation uses information from each score, so it better represents data.



Standard Deviation

SCORES	Score-Mean	(Score-Mean)²
18	-6	36
20	-4	16
24	0	0
25	1	1
33	9	81
		134

$$\frac{134}{5} = 26.8$$

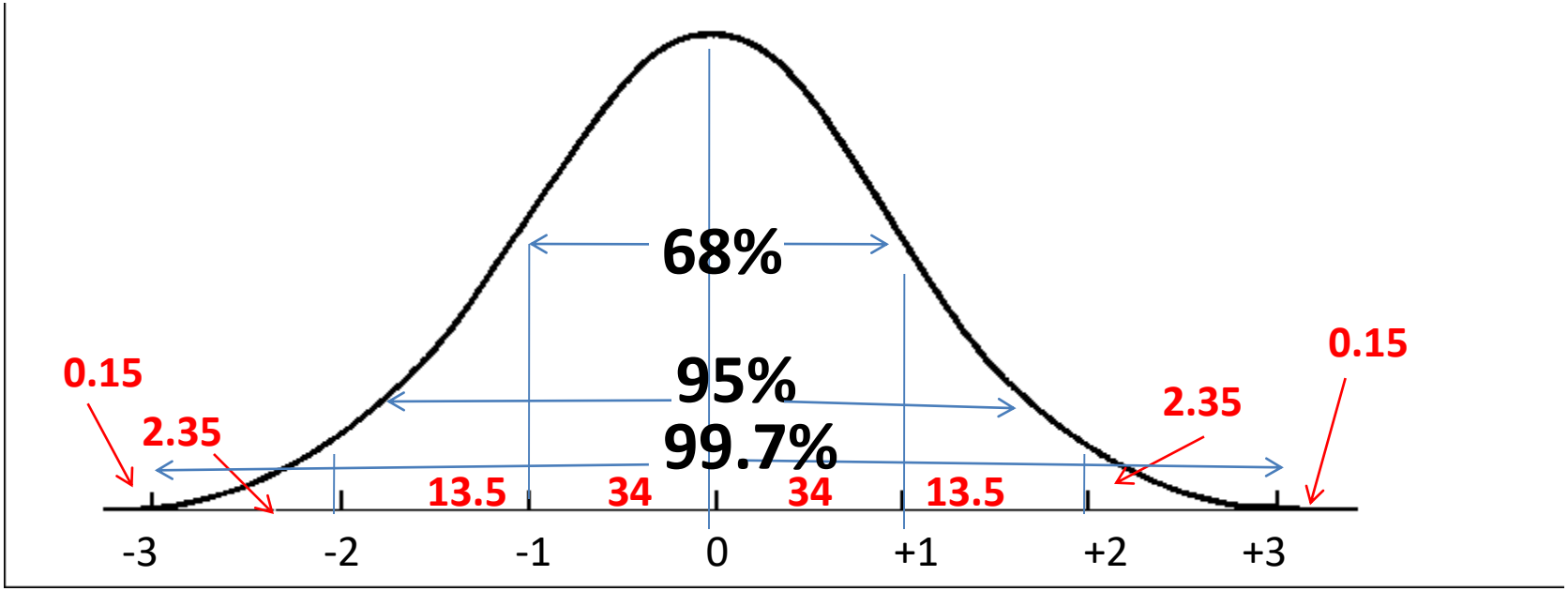
$$\sqrt{26.8}$$

MEAN: 24

Variance: how far a set of numbers are spread out from their average

- **26.8 is the "variance"**
- **Standard deviation is the "square root of the variance." (SD=5.17)**

Normal Curve



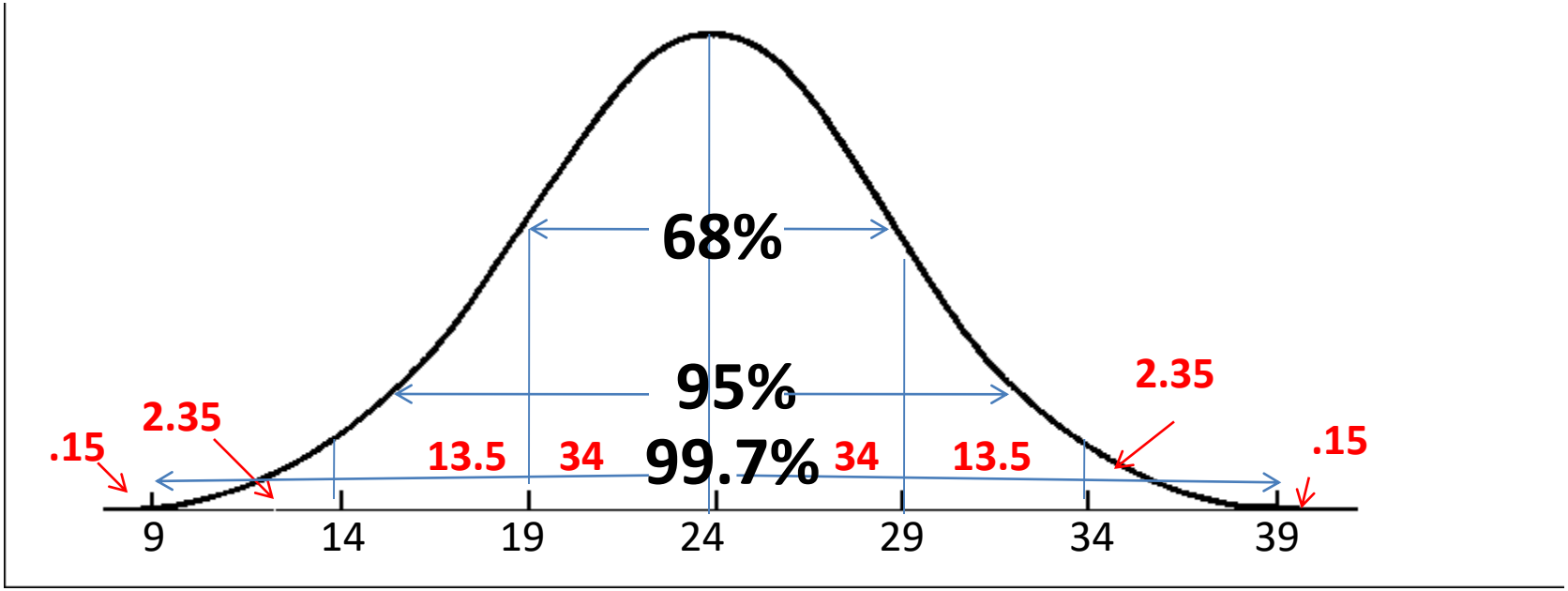
-Each mark represents one deviation away from the mean.

-Numbers in red are ***the percentage of people whose score falls within each standard deviation***.

-68% of people will fall within 1 standard deviation from the mean.

-95% of people will fall within 2 standard deviations from the mean.

Normal Curve



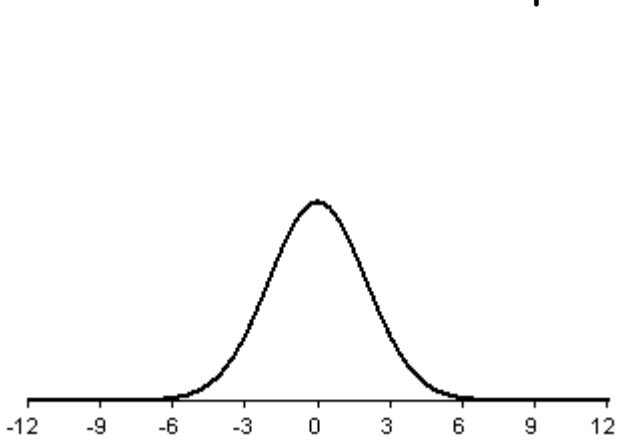
-Using our numbers from our standard deviation exercise, the normal curve would look like this.
68% would have scored within one standard deviation of the mean, or would have scored between 19 and 29.
95% would have scored within two standard deviations, or between 14 and 34.

ESTIMATING VARIANCE

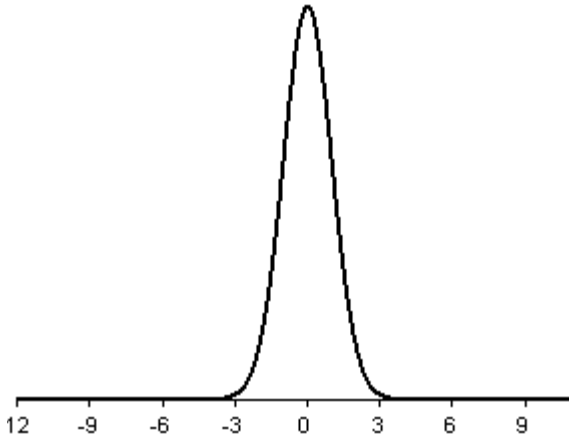
The three curves below represent standard deviations of 1, 2 and 3.

Which curve below would represent a standard deviation of 1? How do you know?

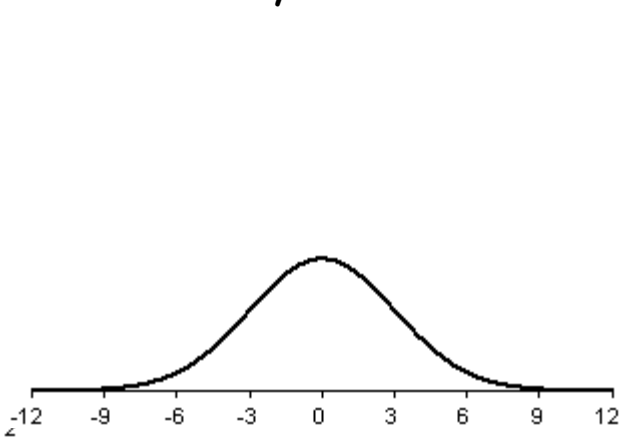
Which curve would represent a standard deviation of 3? How do you know?



A



B



C

THE GREATER THE VARIANCE IN RESULTS, THE GREATER THE STANDARD DEVIATION.

Standard deviation, the normal curve and baseball.