



Math Virtual Learning

Probability and Statistics

April 15, 2020



Probability and Statistics

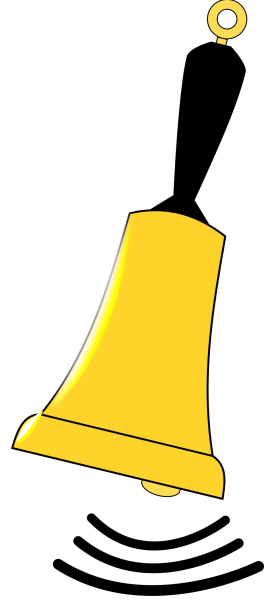
Lesson: April 15, 2020

Objective/Learning Target:

Students will be able to test whether a set of data is normal or not

Below are the standard deviations of 4 data sets

- Which data set is the most spread out?
- Which data set is the least spread out?



Standard Deviation of Set #1: 2.165

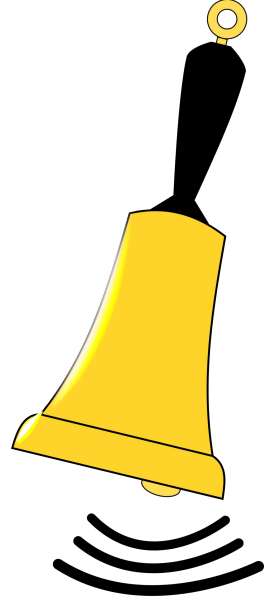
Standard Deviation of Set #2: 21.65

Standard Deviation of Set #1: 0.2165

Standard Deviation of Set #1: 216.5

Below are the standard deviations of 4 data sets

- Which data set is the most spread out?
- Which data set is the least spread out?



Standard Deviation of Set #1: 2.165

Standard Deviation of Set #2: 21.65

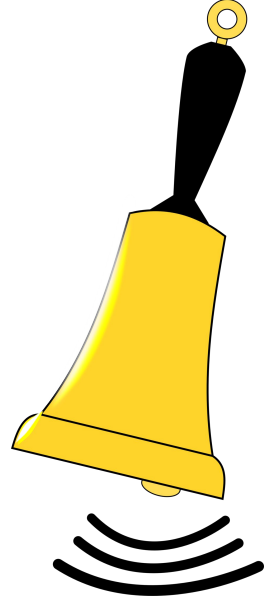
Standard Deviation of Set #1: 0.2165

Standard Deviation of Set #1: 216.5 --

The BIGGER the S.D. the further the numbers are to the mean...so the most spread out

Below are the standard deviations of 4 data sets

- Which data set is the most spread out?
- Which data set is the least spread out?



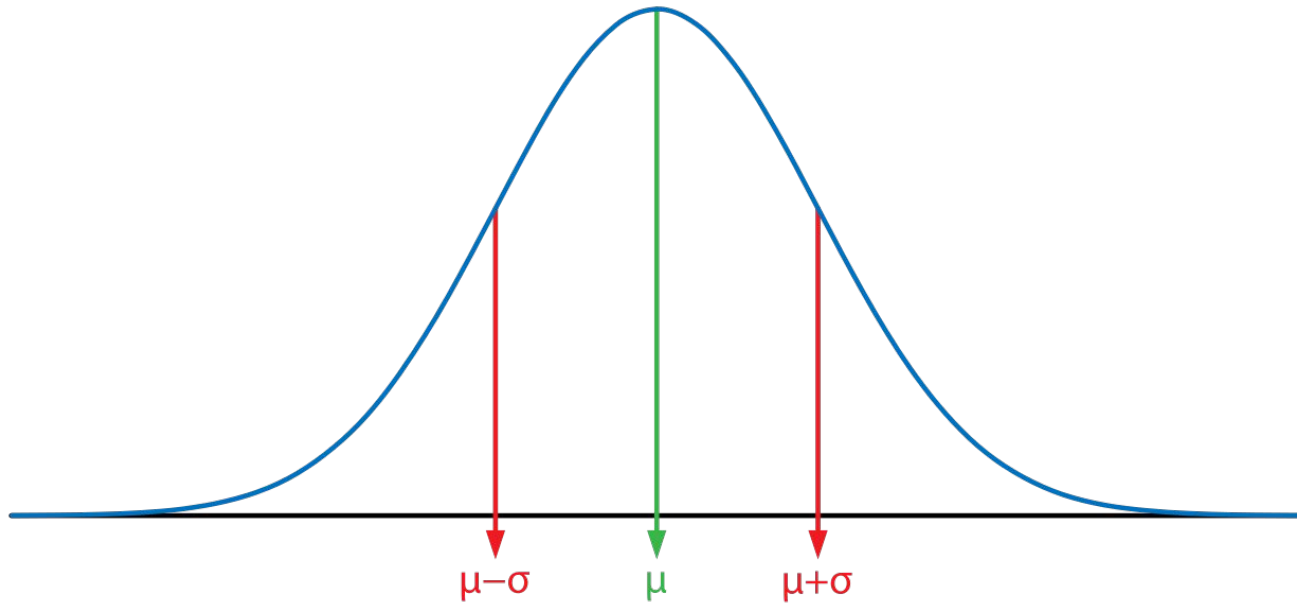
Standard Deviation of Set #1: 2.165

Standard Deviation of Set #2: 21.65

Standard Deviation of Set #1: 0.2165 --

Standard Deviation of Set #1: 216.5

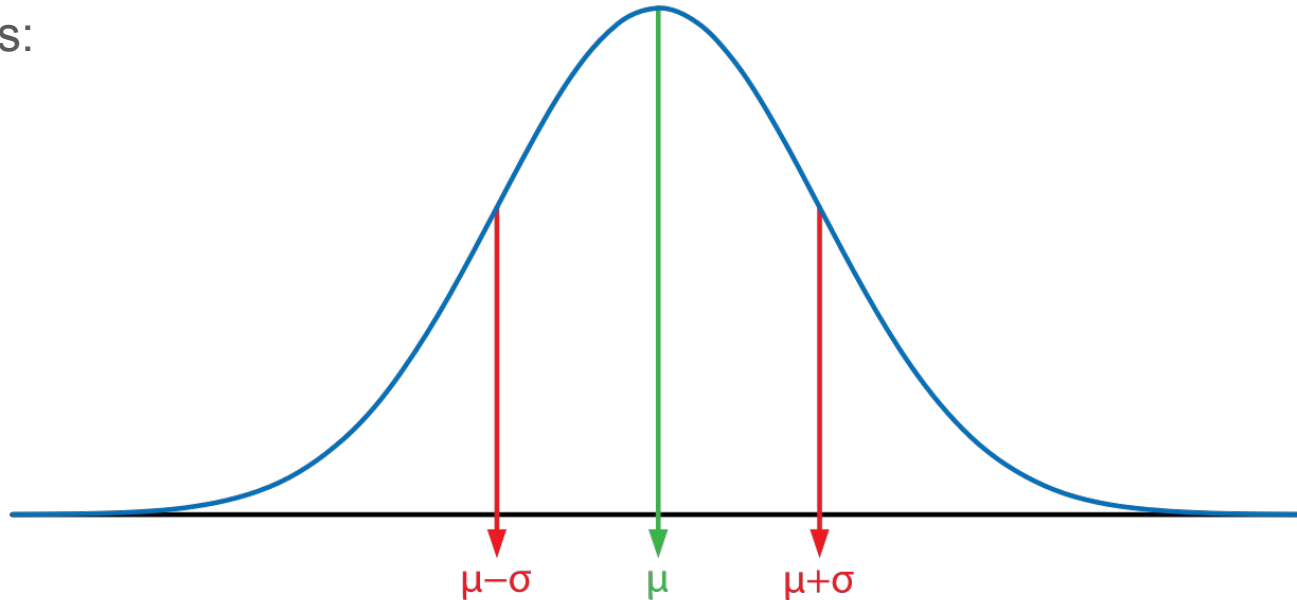
The **SMALLER** the S.D. the closer the numbers are to the mean...so the least spread out



Is it NORMAL???
What does that mean???
Is what normal???

Is what normal? Your data points

In order to imply some statistical formulas and rules, we must first know if our data set is “normal.” That means that IF we were to graph all of the points in our data set (which we won’t, but IF we did), then we would get a bell curve that looks similar to this:



So how do we check without graphing all the points?

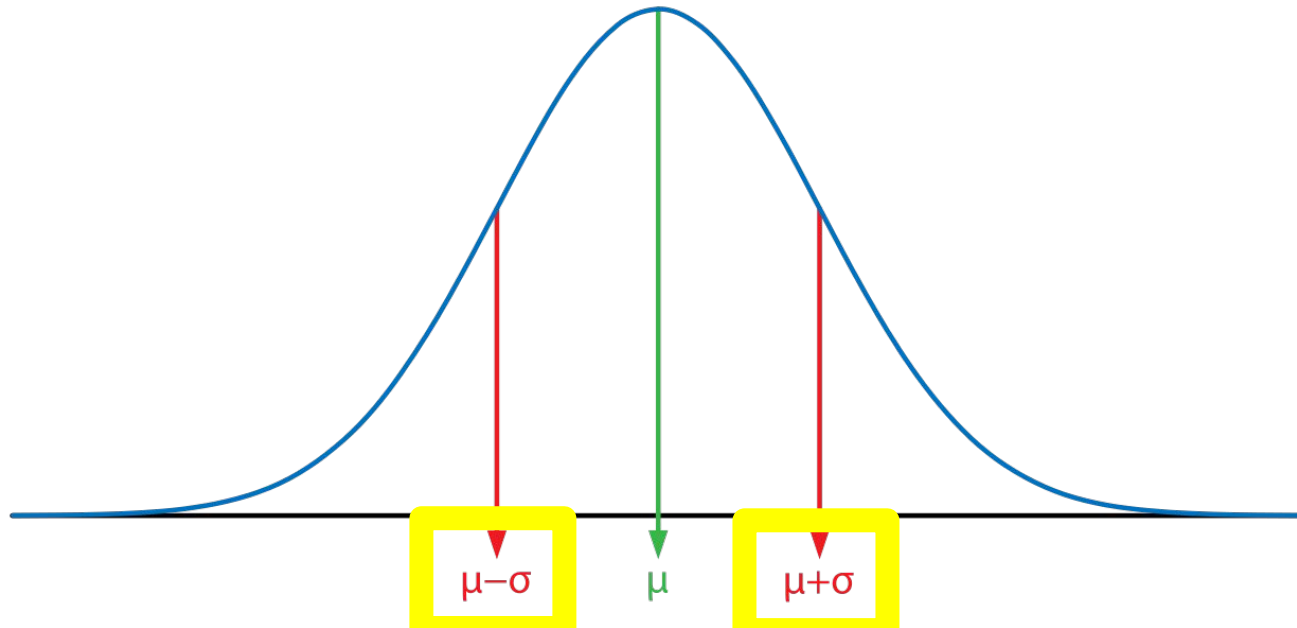
We use STANDARD DEVIATION

HINT:

If you did not look at the lessons for 4/13 and 4/14, then you will want to do that now because you will need to know how to calculate Standard Deviation in order to find if a data set is Normal.

What % is considered normal?

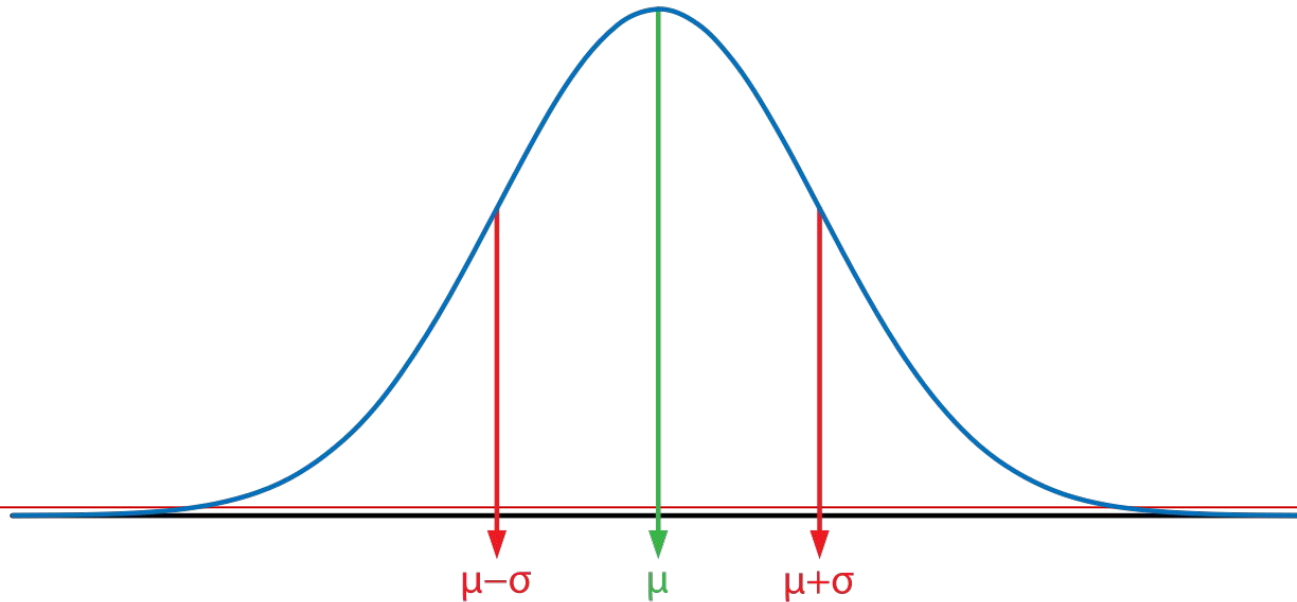
The rule says that when we look at our data, if 68% or more of our data points fall between the range of 1 standard deviation on either side of the mean, then it is considered “normal.”



What do the bars mean?

The **green bar** represents the mean

The **red bars** represent one standard deviation on each side of the mean



Let's try one to get a better understanding...

It's always easier to understand when doing an example, so let's start with a set of data:

The following numbers represent the hourly pay for 20 employees at a factory

\$29.17	\$34.39	\$22.47
\$35.32	\$37.02	\$25.53
\$29.72	\$27.50	\$32.19
\$27.17	\$36.59	\$31.32
\$25.82	\$25.67	\$27.35
\$23.47	\$33.36	\$20.25
\$34.72	\$29.68	

Let's try one to get a better understanding...

The following numbers represent the hourly pay for 20 employees at a factory

\$29.17	\$34.39	\$22.47
\$35.32	\$37.02	\$25.53
\$29.72	\$27.50	\$32.19
\$27.17	\$36.59	\$31.32
\$25.82	\$25.67	\$27.35
\$23.47	\$33.36	\$20.25
\$34.72	\$29.68	

Step 1: Enter data points into your calculator and pull out the Mean and Sample Standard Deviation

Mean: 29.4355 \approx **29.44**

Sample Standard Deviation (Sx):
4.8491 \approx **4.85**

Let's try one to get a better understanding...

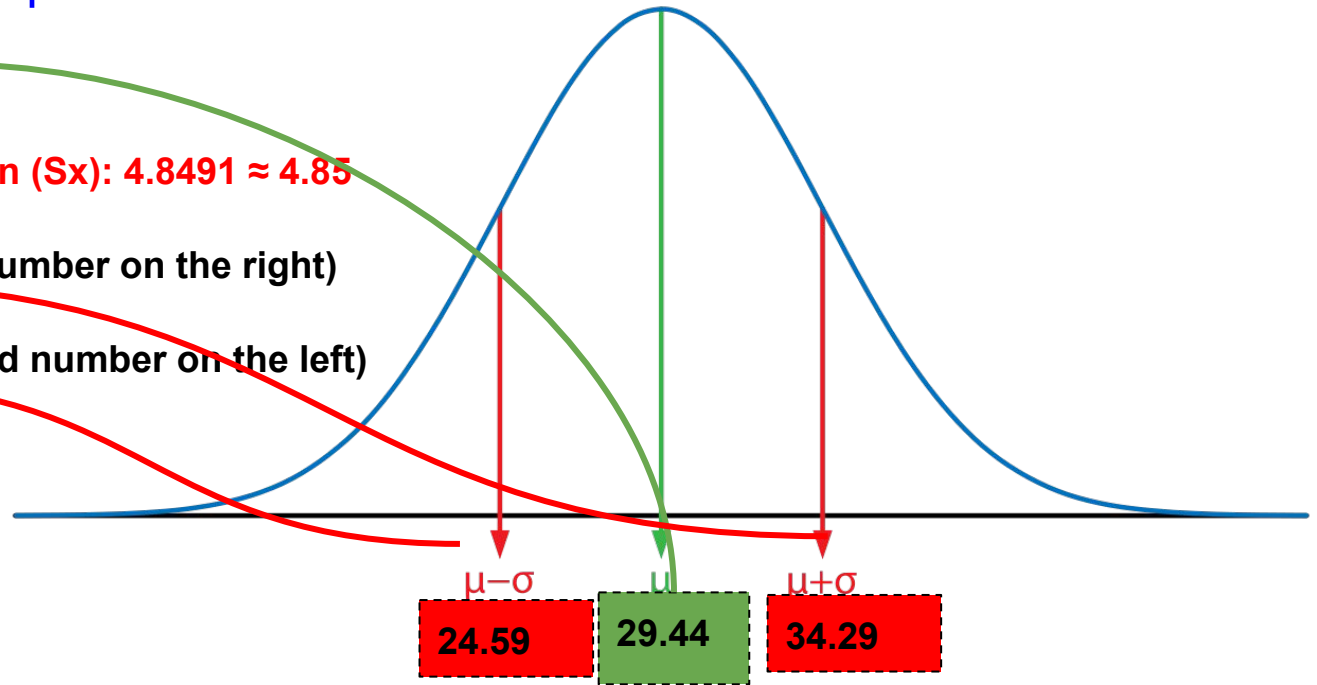
Step 2: Draw the graph and label it

Mean: $29.4355 \approx 29.44$

Sample Standard Deviation (S_x): $4.8491 \approx 4.85$

$29.44 + 4.85 = 34.29$ (red number on the right)

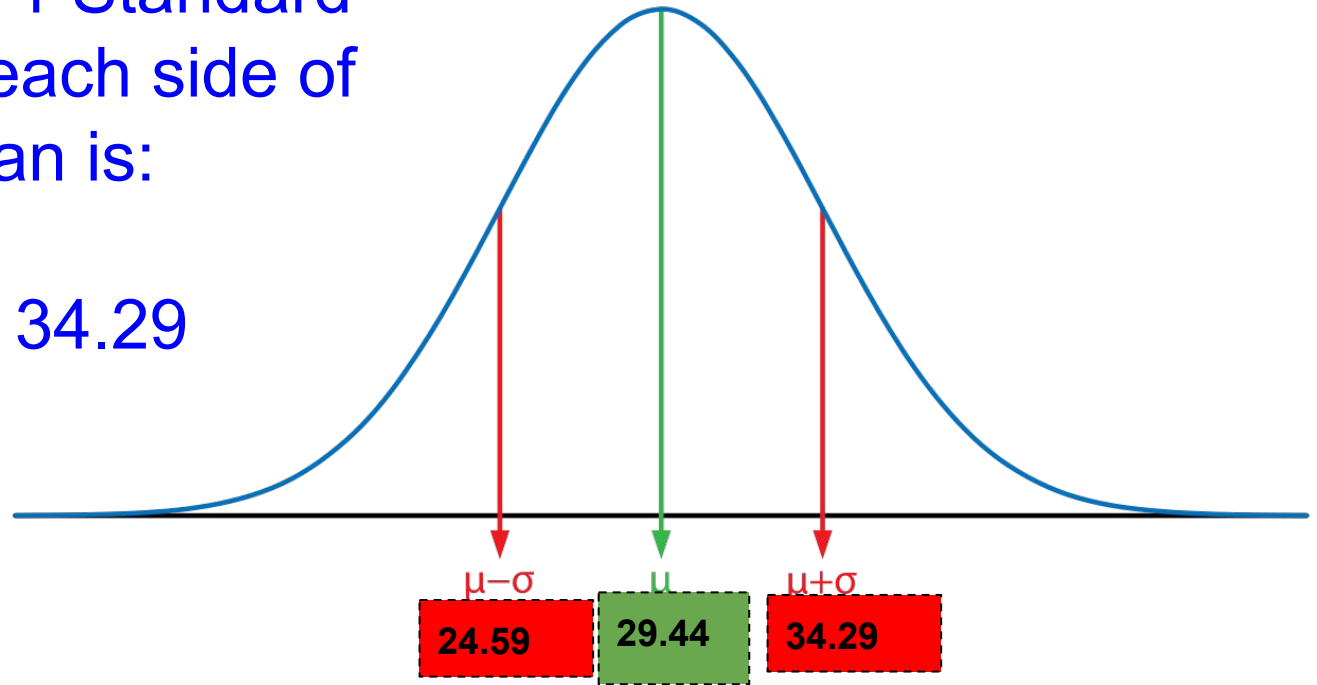
$29.44 - 4.85 = 24.59$ (the red number on the left)



Let's try one to get a better understanding...

Your range of 1 Standard Deviation on each side of the Mean is:

24.59 to 34.29



Let's try one to get a better understanding...

It's always easier to understand when doing an example, so let's start with a set of data:

The following numbers represent the hourly pay for 20 employees at a factory

\$29.17 ✓	\$34.39	\$22.47
\$35.32	\$37.02	\$25.53 ✓
\$29.72 ✓	\$27.50 ✓	\$32.19 ✓
\$27.17 ✓	\$36.59	\$31.32 ✓
\$25.82 ✓	\$25.67 ✓	\$27.35 ✓
\$23.47	\$33.36 ✓	\$20.25
\$34.72	\$29.68 ✓	

Step 3: Count how many of your original data points fall between that range of

24.59 to 34.29

There are 12 salaries that fall within our range of 1 standard deviation on each side of the mean.

Step 4: Calculate the %

$$12 \div 20 = .60 = 60\%$$

Let's try one to get a better understanding...

It's always easier to understand when doing an example, so let's start with a set of data:

The following numbers represent the hourly pay for 20 employees at a factory

\$29.17	\$34.39	\$22.47
\$35.32	\$37.02	\$25.53
\$29.72	\$27.50	\$32.19
\$27.17	\$36.59	\$31.32
\$25.82	\$25.67	\$27.35
\$23.47	\$33.36	\$20.25
\$34.72	\$29.68	

In order to be “Normal” we needed 68% or higher and we only got 60%

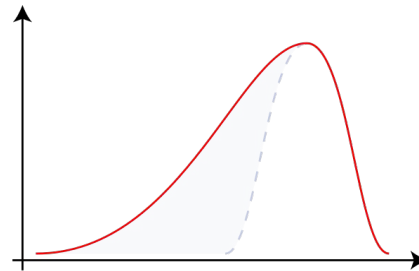
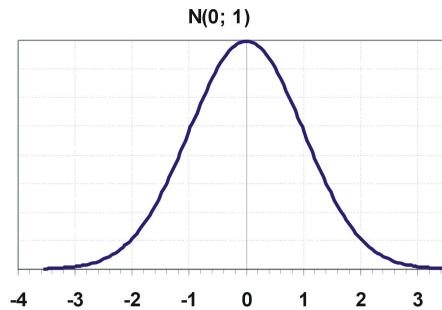
That means that if we graph all of these data points they will NOT make a symmetrical bell curve.

So is the Data “Normal” ???

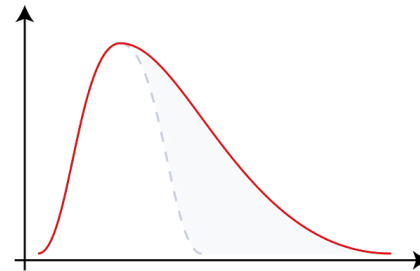
In this particular example the answer is NO the data is not “normal.”

Sometimes you will find that it is normal and sometimes you will find that it is not normal....and that is ok!

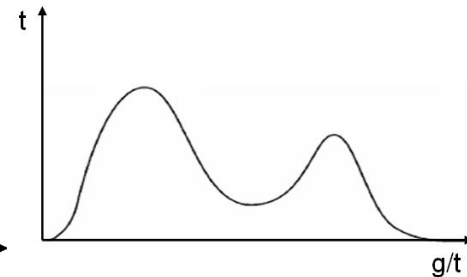
Not all data for real world problems will always make an approximately symmetrical bell curve shape. (sometimes it will be skewed, or bi-modal, etc)



Negative Skew



Positive Skew



Let's do another one...

The following data represents the travel time, in minutes, for 20 of your teachers driving to and from work each day (round trip):

26, 33, 65, 28, 34, 55, 25, 44, 50,
36, 26, 37, 43, 62, 35, 38 45, 32,
28, 34

Step 1: Enter data points into your calculator and pull out the Mean and Sample Standard Deviation

Mean: 38.8

Sample Standard Deviation (Sx): 11.7

Let's do another one...

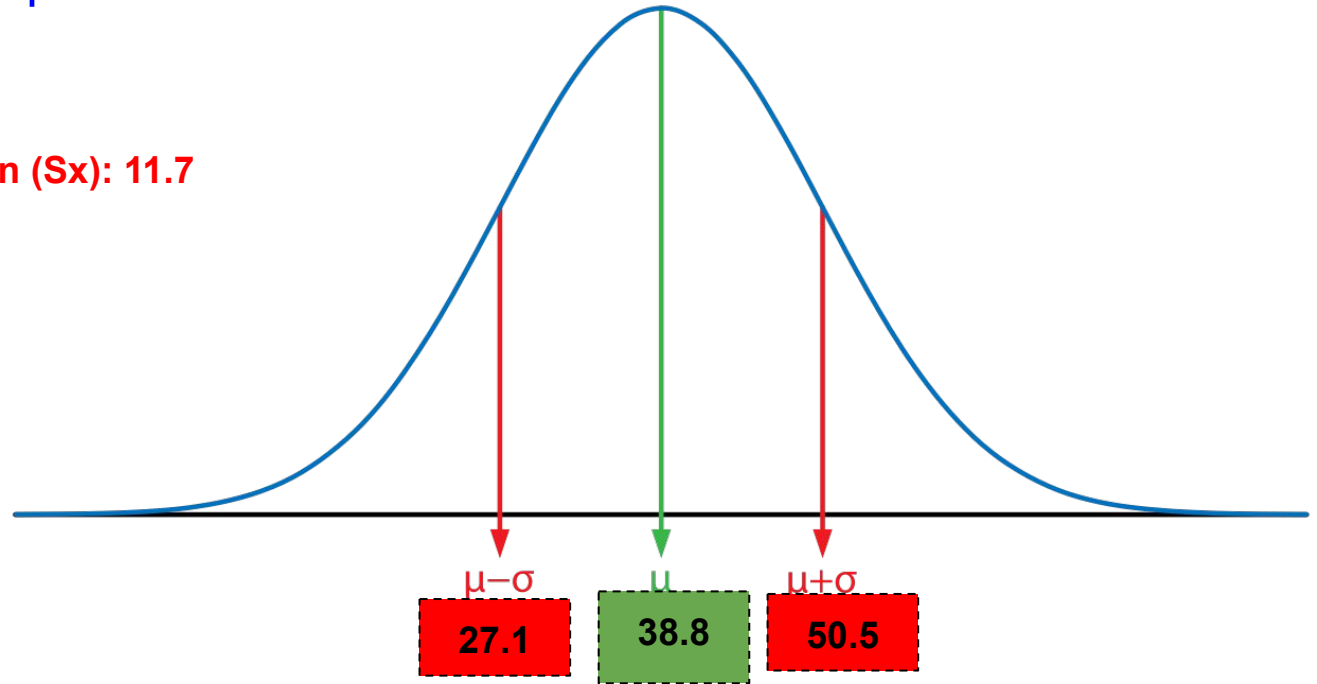
Step 2: Draw the graph and label it

Mean: 38.8

Sample Standard Deviation (Sx): 11.7

$$38.8 + 11.7 = 50.5$$

$$38.8 - 11.7 = 27.1$$



Let's do another one...

The following data represents the travel time, in minutes, for 20 of your teachers driving to and from work each day (round trip):

26, 33, 65, 28, 34, 55, 25, 44, 50,
36, 26, 37, 43, 62, 35, 38, 45, 32,
28, 34

Step 3: Count how many of your original data points fall between that range of

27.1 and 50.5

There are 14 drive times that fall within the range of 1 standard deviation on each side of the mean.

Step 4: Calculate the %

$$14 \div 20 = .7 = 70\%$$

Let's do another one...

The following data represents the travel time, in minutes, for 20 of your teachers driving to and from work each day (round trip):

26, 33, 65, 28, 34, 55, 25, 44, 50,
36, 26, 37, 43, 62, 35, 38, 45, 32,
28, 34

In order to be “Normal” we needed 68% or higher and we got 70%

That means that if we graph all of these data points they ARE NORMAL and WILL make a symmetrical bell curve.