

Algebra 2

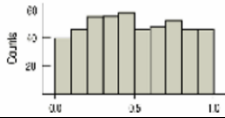
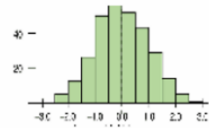
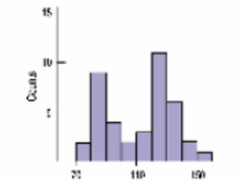
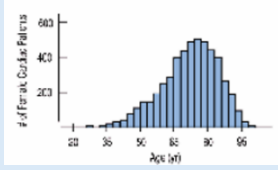
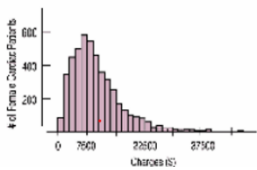
Chapter 9 Notes

Name: _____

Day	Date	Assignment (Due the next class meeting)
		9.1 Worksheet Measures of Center, Box Plots, and Outliers
		9.2 Worksheet Data Distributions
		9.3 Worksheet Normal Curves and Histograms
		9.4 Worksheet Normal Curves and z-scores
		Ch 9 Practice Test
		Ch. 9 Test

- * Be prepared for daily quizzes.
- * Every student is expected to do every assignment for the entire unit.
- * Try www.khanacademy.org if you need help outside of school hours.
- * Need help with the material? Check out Earl's website: www.mathguy.us
- * Student who complete 100% of their homework second semester will receive a pizza party and 2% bonus to their grade!

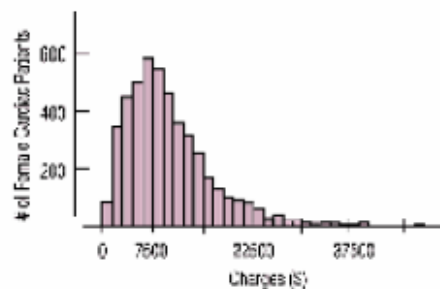
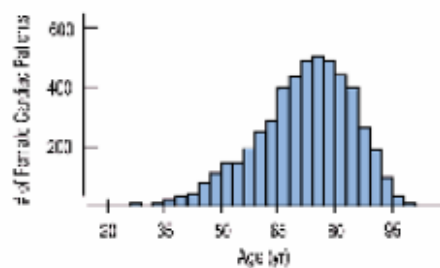
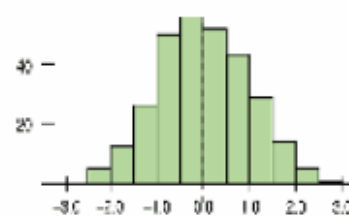
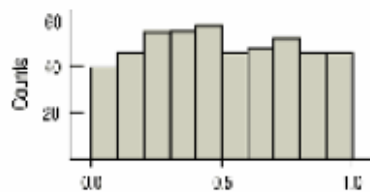
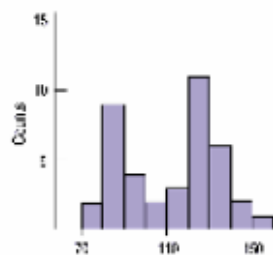
Chapter 9 Key Terms

Uniform Distribution: Same Height 	Normal Distribution: “Bell” Curve 
Bimodal Distribution: 2 Modes (highest points) 	Skewed Left: “Tail” on the left side of the graph. 
Skewed Right: “Tail” on the right side of the graph. 	<u>Categorical Data vs. Quantitative Data:</u> • Categorical: categories (not #'s) • Quantitative: numerical data you can measure
Mean ($\bar{x}$): “average” • To find the mean you add all of the data and then divide by the number of terms. •	Median: “middle” # when data is listed in order from least to greatest.
Mode: “most”, the value that occurs the most often in the data set.	Range: measures the “spread” of the data. Highest value – Lowest value
Quartile: a statistical term that divides a set of data into 4 equal parts or quarters. • Q_1 is the median of the 1st half of data • Q_3 is the median of the 2nd half of data •	Interquartile Range (IQR): a measure of spread of the data. $Q_3 - Q_1$
Standard Deviation (σ): is a measure of variation from the mean typically used with normal distributions.	Outliers: an outlier is any value MORE than 2 standard deviations (σ) from the mean ($\bar{x}$).
Percentile Ranking: a score has a certain percentile based on the percentage of scores below that score.	Probability: the likelihood that a favorable outcome will occur.
Z-score: tells us how many standard deviations a value is from the mean. $Z - score = \frac{value - mean}{standard deviation} = \frac{x - \bar{x}}{\sigma}$	

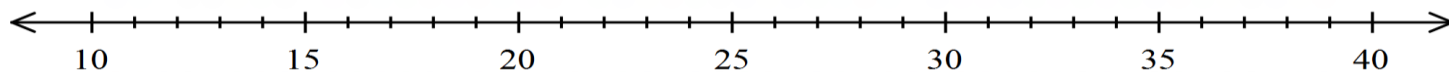
Algebra 2

9.1 Notes Shapes

Chapter 9 Notes



Students: come up to the board and create a dot plot of your ages:



Which shape does this most closely resemble?

If your teacher decides to add his or her age, how does it change the shape?

When is data normal?

What shape do you think these data would have?

1. The test scores for an Algebra 2 test:
2. The # of Instagram followers for kids in Alg 2 at Damonte:
3. The # of pushups the boys in the room can do:
4. The # on a 6-sided die if tossed 100 times:

Types of Data:

What type of data?

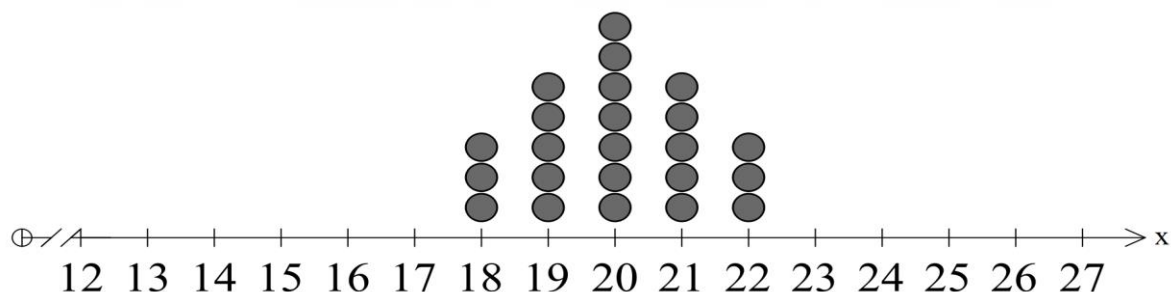
Categorical:

Type of pet
GPA
Gender
Level of approval
Age
Phone Number

Quantitative:

Center

Here are the ages of students in a college course. Mark the median, mean, and mode.



Let's say a 12-year old child genius is added to the course. Now mark the median, mean, and mode. What has happened to these values?

What is the new shape of the data?

Measures of Center

- Mean:
- Median:
- Mode:

5-number summary

Create a 5-number summary from the data on the right, which represents the number of minutes people waited for their doctor when arriving on time for an appointment:

Min:	6	28
Q1:	5	0468
Median:	4	0346
Q3:	3	334446678
Max:	2	289
Find the range:	1	8

Find the Interquartile Range:

Now create a box plot from the 5 number-summary:

If everyone is asked to fill out an additional piece of paperwork and their wait times all increase by 5 minutes, what happens to the:

IQR:

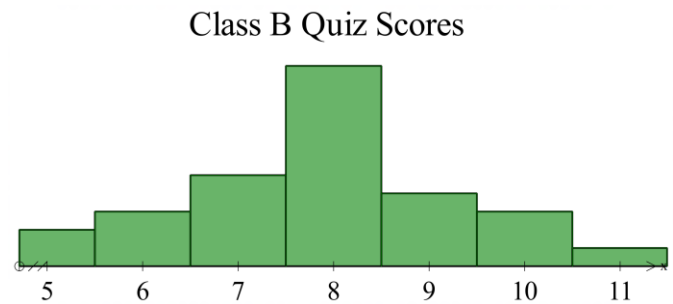
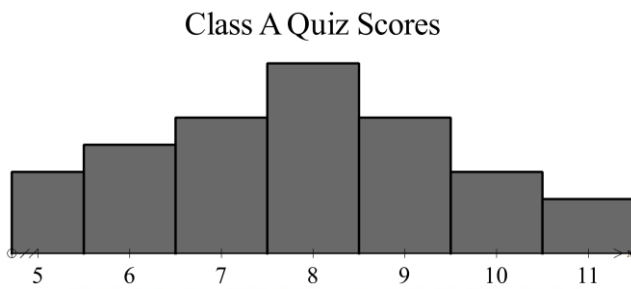
Median:

If all of the wait times are doubled, what happens to the:

IQR:

Median:

Spread



Compare the center and the range of the two classes:

Which class would have a larger IQR?

Which class shows less variability?

Measures of Spread

- Standard Deviation:
- Range:
- IQR:

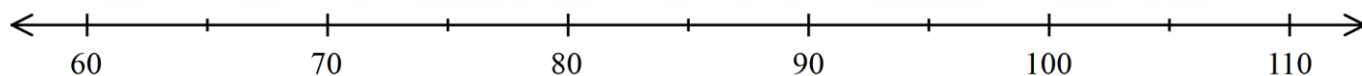
Outliers

If data looks **normal** (symmetric/unimodal): $mean \pm 2 \cdot standard\ deviation$.
(We will focus more on this in 9.3.)

If data doesn't look *normal*:

Here are some Algebra 2 test scores. Find the five-number summary, identify any outliers, and create a box plot. *Denote the outliers with asterisks.*

63, 72, 75, 78, 80, 81, 85, 86, 86, 86, 87, 87, 88, 90, 95, 96, 98, 108



What is the shape of this data?

Would you expect the mean or median to be larger?

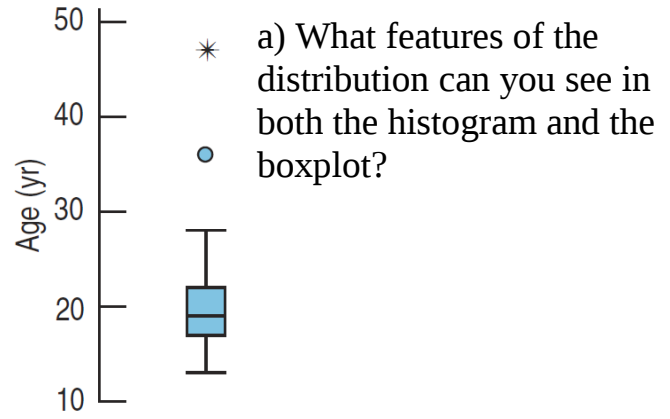
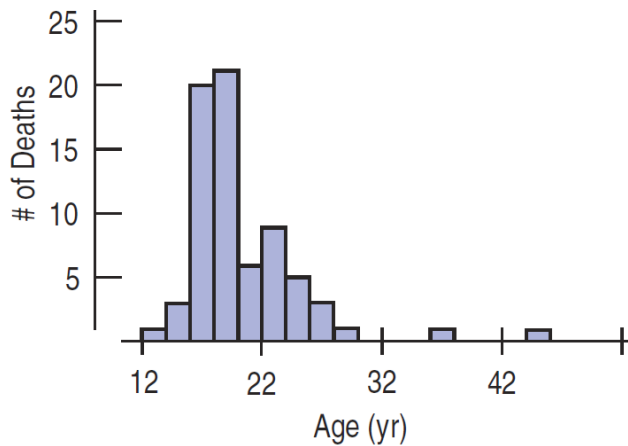
What is the IQR?

What is the range?

What percentage of the data is between 80 and 108?

Between 63 and 80?

Example: Crowd Management Strategies monitors accidents at rock concerts. In their database, they list the names and other variables of victims whose deaths were attributed to “crowd crush” at rock concerts. Here are the histogram and boxplot of the victims’ ages for data from 1999 to 2000:



- a) What features of the distribution can you see in both the histogram and the boxplot?
- b) What features of the distribution can you see in the histogram that you could not see in the boxplot?
- c) What summary statistic would you choose to summarize the center of this distribution? Why?
- d) What summary statistic would you choose to summarize the spread of this distribution? Why?

9.2 Notes: Statistical Studies and Sampling Techniques

Population – all the members of a group being studied

Parameter – a measurement of the *population*

Sample – a subset of the population that's being studied

Statistic – a measurement of the *sample*

1. The principal wants to know how many students at her school have jobs. She samples 100 of the students and found that 25% of them were employed.

Population:

Sample:

Parameter:

Statistic:

2. A biology teacher wants to know which dissection his students liked best this semester, so he randomly selects eight students from each of his five classes to participate in a survey.

What is the population?

What is the sample?

3. At the end of the month, a restaurant manager used the total food sales for the month and the total number of customers for the month to calculate that the mean amount spent per customer during that month was \$12.59.

What's the population?

Is this value a statistic or a parameter?

Sample survey – asks every member of sample the same set of questions and records the answers

Observational Study – measures or observes members of a sample in such a way that they are not affected by the study.

Experiment – involving applying a treatment to some group or groups and measures the effects of the treatments

Identify if this is a survey, observational study, or experiment:

4. Lions are watched in the wild and the amount of time they sleep is recorded.
5. Ask randomly chosen students in the school their opinion and record the results.
6. Doctors track children throughout their childhood to see if cancer is ever diagnosed.
7. Split plants into groups randomly and give each group different conditions such as extra water, light, or different fertilizer. Leave one group under the current conditions to see if the new treatments show improvement.
8. A gym asks their customers if they would prefer the gym to be open earlier in the morning.
9. A gym counts the number of customers who work out before 8am.

Bias – a systematic error in a study that results in a sample misrepresenting a population

Variability – spread or range of the data. Within most data there is natural variation.

- *Do you think there could be bias in these samples and why?*

10. A soft drink company calls 500 people at random and asks, “Isn’t it true that our product is better than our rival’s product?” and 75% of the people respond, “Yes.”

11. A PTA has a community meeting and they ask everyone who attends if they agree with the budget and 70% agree.

12. A teacher wants to know if kids at her school enjoy her class, so sends 5 randomly selected students from each of her classes to a computer lab to take a survey.

13. A teacher wants to know if his students find it easy to turn in homework, so he gives a survey to all of his students and asks them to turn it in when they’re done filling it out. He figures since he has 150 students, he will have enough.

14. A scientist wants to know how happy people are in the morning at 9:00. She samples 100 people in a subway car on their way to work and her statistic was 35%. Her colleague sampled 100 people at the dog park and his statistic was 65%. Are their differences due to variability or bias?

15. A radio station wants to know if its listeners liked the last song, so they tell them to go to their facebook page and complete the survey.

Experiment – involving applying a treatment to some group or groups and measures the effects of the treatments.

Treatments – the treatments must be **randomly assigned** to more than one group. If there are no treatments *assigned* and/or they are not *randomly* assigned, there is no experiment.

Control Group – A control group is a group that receives the treatment of *nothing* (can be a placebo). This way the **experimental group** can be compared with a baseline. There doesn't need to be a control group in an experiment, but it is very common.

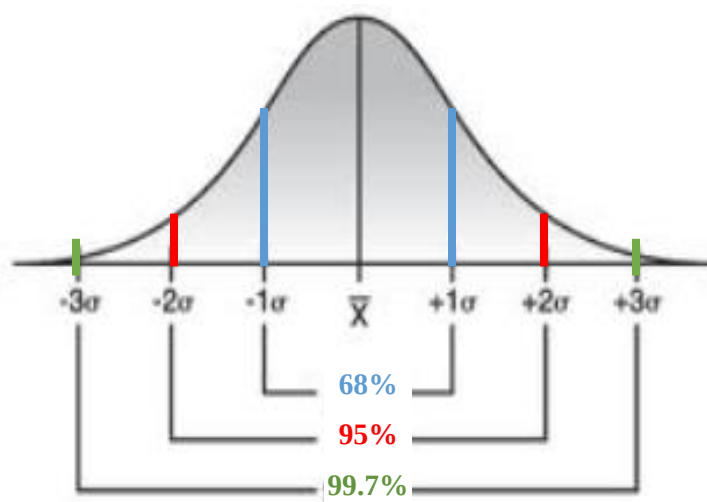
16. Describe a design for a controlled experiment to test a new drug for treating the flu on mice. How does randomization apply to your design?

17. Design a controlled experiment to test whether drinking coffee improves memory. How would you choose the experimental and control groups?

9.3 Notes: Normal Curves and Histograms

Given a distribution with a normal shape that is made out of a curve, the distribution can be separated into sections by adding and subtracting _____ from the mean.

The percentages of the area under the curve for each section will be the same for every normal curve, as shown below.



The percentages under the curve can be remembered by using the

_____ - _____ - _____ rule.

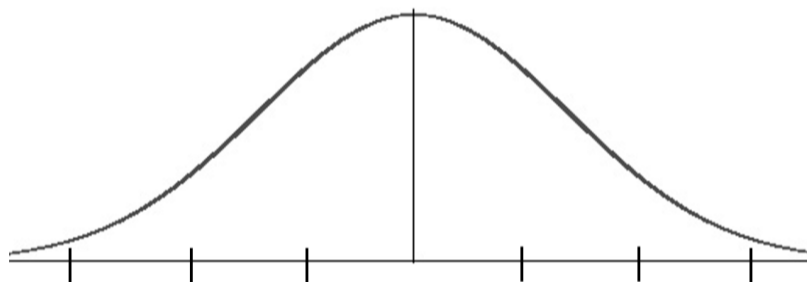
Thus, the area under the curve that is within _____ standard deviation of the mean is _____% of the total area under the normal curve.

Also, the area under the curve that is within _____ standard deviations of the mean is _____% of the total area under the normal curve.

Finally, the area under the curve that is within _____ standard deviations of the mean is _____% of the total area under the normal curve.

Ex 1: The height of adult American males is normally distributed with a mean of 70 in and a standard deviation of 3 in.

- a) Draw the normal curve to represent the heights of adult American males.



- b) What percentage of adult American males are between 64 inches and 76 inches tall?

- c) What percentage of adult American males are greater than 70 inches tall?

- d) What percentage of adult American males are less than 67 inches tall?

- e) What percentage of adult American males are greater than 76 inches tall?

- f) What percentage of adult American males are between 67 and 79 inches tall?

- g) True or false? Approximately the same number of adult American males are taller than 79 inches as those who are shorter than 61 inches.

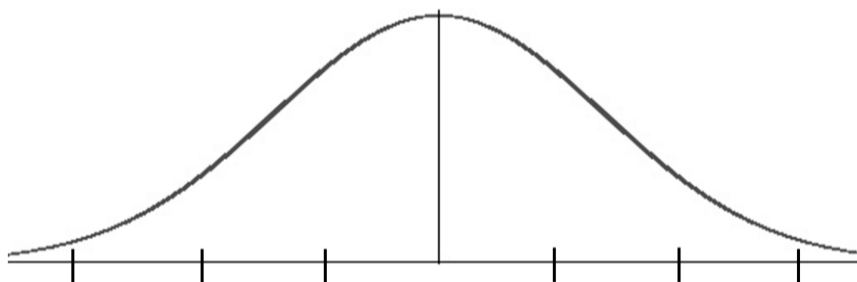
- h) As of the US Census taken in 2010, there are approximately 113 million adult men in the United States. *How many* of them are between 67 and 73 inches tall?

Outliers: A score is an outlier if it is more than _____ standard deviations away from the mean.

Percentile ranking: A score has a certain percentile based on the percentage of scores _____ than that score.

Example 2: In 2011, SAT scores are have a normal distribution with the mean score of the math portion of the SAT is 518, with a standard deviation of 115.

a) Draw the normal curve.



b) What is the **probability** that a student scored less than 403 on the math portion of the SAT?

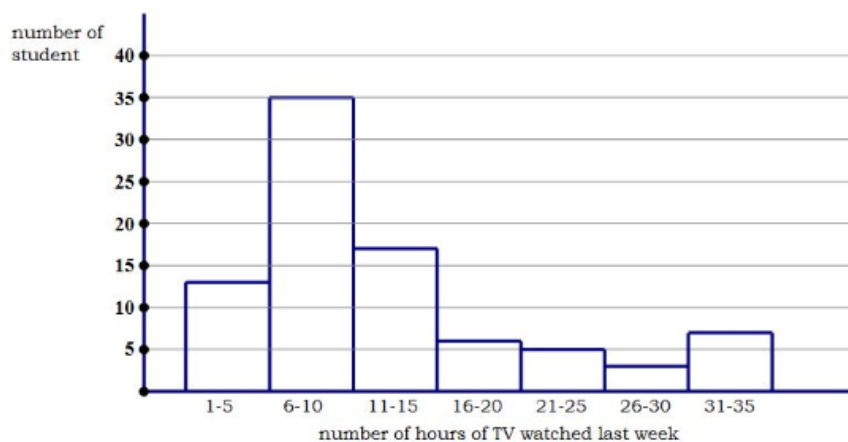
c) What is the **probability** that a student scored between 288 and 518 on the math portion of the SAT?

d) What is the **probability** that a student scored between 288 and 633 **or** more than 748 on the math portion of the SAT?

e) Terrance scored 750 on the math portion of the SAT. Is his score an outlier? Justify your conclusion.

f) Sara scored 633 on the math portion of the SAT. The college that she wants to attend requires that students score in the 85th percentile or higher on that math portion of the SAT. Is Sara's score high enough for her to be considered by this college?

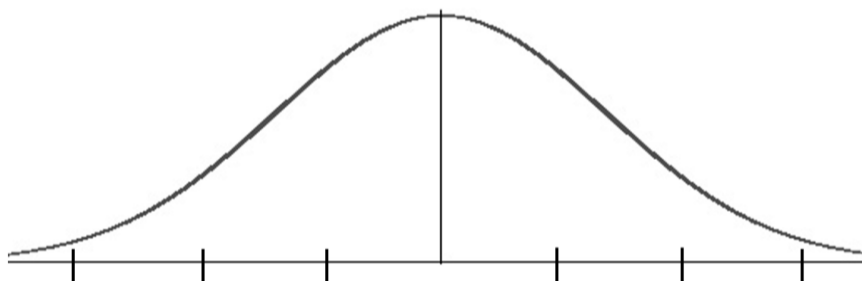
Example 3: In a survey, high school students were randomly selected and asked how many hours of television they had watched in the previous week. The histogram to the right displays their answers.



- a) Approximately how many students participated in the survey?
- b) Describe the shape of the distribution.
- c) Approximately how many students watched 10 hours or less of TV last week?
- d) Approximately how many students watched between 16 and 30 hours of TV last week?
- e) In which category is the median of the data? Explain.
- A) 6 – 10 hours B) 11 – 15 hours C) 16 – 20 hours D) 21 – 25 hours
- f) Is it possible to calculate the *mean* of the data from a histogram? Explain.
- g) Which has a higher value, the median of the data or the mean of the data? Explain.

9.4 Notes: Normal Curves and z-scores

Ex 1: The heights of adult American males is normally distributed with a mean of 70 in and a standard deviation of 3 in. Draw the normal curve to represent the heights of adult American males.



b) Suppose you want to find the probability that an adult American male is less than 75 inches tall. Why is this problem hard to do?

We want to find out *how many* standard deviations 75 is away from our mean. To do this, we can calculate a **z-score**.

z-score formula:
$$z = \frac{\text{value} - \text{mean}}{\text{standard deviation}} = \frac{x - \bar{x}}{\sigma}$$

A z-score tells us _____ standard deviations we are away from the mean.

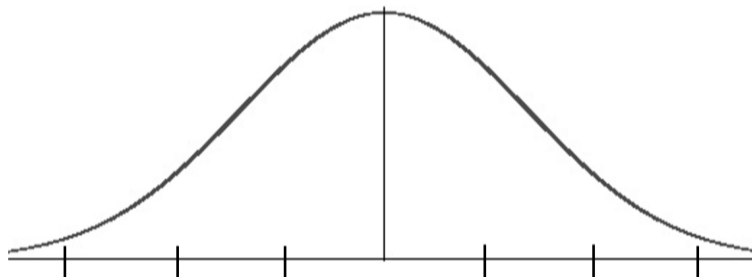
c) Calculate the z-score for 75 inches.

d) We can use a table of values (which will be given to you) in order to find the probability that we randomly choose someone **less than** our given score. Use the table below to find the probability that an adult American male is less than 75 inches tall.

z	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
-2	.0228	.0179	.0139	.0107	.0082	.0062	.0047	.0035	.0026	.0019
-1	.1587	.1357	.1151	.0968	.0808	.0668	.0548	.0446	.0359	.0287
-0	.5000	.4602	.4207	.3821	.3446	.3085	.2743	.2420	.2119	.1841
0	.5000	.5398	.5793	.6179	.6554	.6915	.7257	.7580	.7881	.8159
1	.8413	.8643	.8849	.9032	.9192	.9332	.9452	.9554	.9641	.9713
2	.9772	.9821	.9861	.9893	.9918	.9938	.9953	.9965	.9974	.9981

Example 2: In 2011, SAT scores have a normal distribution with the mean score of the math portion of the SAT being 518, with a standard deviation of 115. What is the **probability** that a student scored less than 450 on the math portion of the SAT?

Step 1: Draw a sketch of the normal curve.



Step 2: Calculate the z-score.
Why is this a negative value?

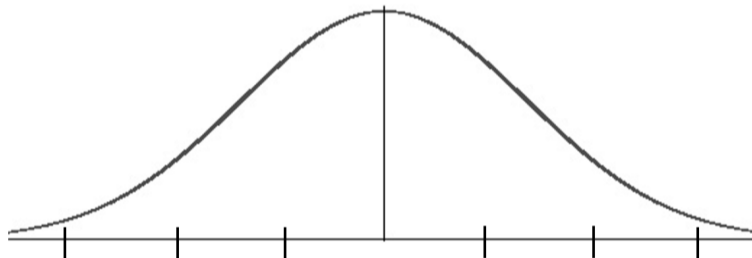
Step 3: Use the table below to answer the question.

<i>z</i>	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
−2	.0228	.0179	.0139	.0107	.0082	.0062	.0047	.0035	.0026	.0019
−1	.1587	.1357	.1151	.0968	.0808	.0668	.0548	.0446	.0359	.0287
−0	.5000	.4602	.4207	.3821	.3446	.3085	.2743	.2420	.2119	.1841
0	.5000	.5398	.5793	.6179	.6554	.6915	.7257	.7580	.7881	.8159
1	.8413	.8643	.8849	.9032	.9192	.9332	.9452	.9554	.9641	.9713
2	.9772	.9821	.9861	.9893	.9918	.9938	.9953	.9965	.9974	.9981

Example 3: In 2011, the mean score of the math portion of the SAT is 518, with a standard deviation of 115. What is the **probability** that a student scored **more than** 450 on the math portion of the SAT?

Example 4: In 2011, SAT scores have a normal distribution with the mean score of the math portion of the SAT being 518, with a standard deviation of 115. What is the **probability** that a student scored **between** 450 and 600 on the math portion of the SAT?

Step 1: Draw a sketch of the normal curve.



Step 2: Find the z-score for 600.

Step 3: Find the probability for **scoring less** than 600.

<i>z</i>	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
−2	.0228	.0179	.0139	.0107	.0082	.0062	.0047	.0035	.0026	.0019
−1	.1587	.1357	.1151	.0968	.0808	.0668	.0548	.0446	.0359	.0287
−0	.5000	.4602	.4207	.3821	.3446	.3085	.2743	.2420	.2119	.1841
0	.5000	.5398	.5793	.6179	.6554	.6915	.7257	.7580	.7881	.8159
1	.8413	.8643	.8849	.9032	.9192	.9332	.9452	.9554	.9641	.9713
2	.9772	.9821	.9861	.9893	.9918	.9938	.9953	.9965	.9974	.9981

Step 4: Reminder: You already know the probability for scoring less than 450, which is _____. Now, you can _____ the probabilities for 600 and 450 to find the probability of scoring **between** those values.

Why does this strategy work?