2. A community consists of 580 members. In the summer, the number of members that participate in recreation sports is 325 and the number of members that participate in outdoor festivals is 182. If the probability that a member participates in either sports or festivals is $\frac{411}{580}$, what is the probability that a member participates in both sports and festivals?

3. Given $P(X) = 0.4, P(Y) = 0.57$, and $P(X \cup Y) = 0.288$, find the value of $P(X \cap Y)$.

4. Given $P(B) = 0.37, P(C) = 0.5$, and $P(B \text{ } and \text{ } C) = 0.315$, find the value of $P(B \text{ } or \text{ } C)$.

5. Given $P(N) = 0.5, P(M \cap N) = 0.2$, and $P(M \cup N) = 0.56$, find the value of $P(M)$.

Skill #39: Conditional Probability

- Please be able to know and apply the following formula. It must be memorized. $P(A|B) = \frac{P(A \cap B)}{P(B)}$

probability of A given B occurred

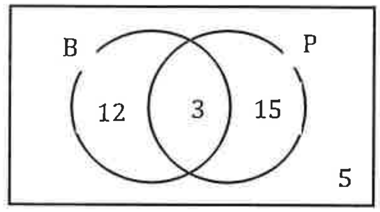
Find $P(male|cat\ owner)$.

	Cat	Dog	Total
Male	11	29	40
Female	22	30	52
Total	33	59	92

To find $P(M|C)$, first look at the C column.
That total becomes the denominator.
Then look at how many males are in that column.

So $P(M|C) = \frac{11}{33}$

Find $P(biology|physics)$.



To find $P(B|P)$, first look at the P circle.
The total in that circle becomes the denominator.
Then look at how many bio students are in that circle. So $P(B|P) = \frac{3}{18}$

A survey of a town in Syracuse revealed that the probability that a resident supports converting I-81 to a community grid is 0.6. The probability that a resident favors education funding <u>given</u> that they support the community grid is 0.62. Determine the probability that a randomly selected resident of this town favors education funding and supports the community grid.

The word "AND" tells us that we need to find the intersection.

$$P(E|G) = \frac{P(E \cap G)}{P(G)}$$
$$0.62 = \frac{P(G \cap E)}{0.6}$$
$$P(G \cap E) = 0.372$$

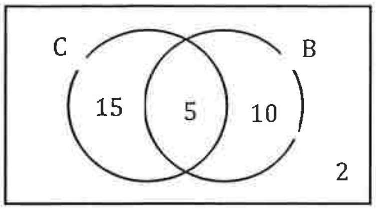
Using the information about conditional probability, compute the following probabilities.

1. Using the table below, determine the following probabilities.
   a. $P(Sandals|Hat)$

   b. $P(Sandals|No\ Hat)$

	Sandals	No Sandals	Total
Hat	25	15	40
No Hat	50	10	60
Total	75	25	100

2. Using the Venn Diagram below, determine the following probabilities.
   a. $P(city|beach)$



   b. $P(beach|city)$

3. Given $P(X) = 0.45, P(Y) = 0.7$, and $P(X \cap Y) = 0.375$, find the value of $P(X|Y)$.

4. Given $P(A) = 0.8, P(B) = 0.61$, and $P(A|B) = 0.8$, find $P(A \cap B)$.

5. In a class, 32.1% of students have a brother B, 57.1% of students have a sister S, and 14.3% have both a brother and a sister.
   a. Determine the probability of S given B, to the *nearest tenth of a percent.*

   b. The teacher wants to make a slide for her class that explains these results. What could he write?

6. In a math class, a test was given the same day that an assignment was due. The probability that a student passed the test was $\frac{17}{27}$ and the probability that a student completed the assignment was $\frac{2}{3}$. If the probability that a student passed the test and completed the assignment was $\frac{5}{9}$, what is the probability that a student completed the assignment given that they passed the test?

Skill #40: Test for Independence

- To show that two events are independent, you must show one of the following: (1) $P(A \cap B) = P(A) \cdot P(B)$
  (2) $P(A|B) = P(A)$

The two-way table shows the number of students in 10th and 11th grade who have seen Avengers: Endgame opening weekend.

	Saw Endgame	Did not see Endgame	Total
10th graders	33	99	132
11th graders	37	111	148
Total	70	210	280

Are "being a 10th grader" and "seeing Endgame" independent events?

$$P(sophomore|Endgame) = P(sophomore)$$
$$\frac{33}{70} = \frac{132}{280}$$
$$0.4714285714 = 0.4714285714$$

Yes – being a 10th grader and seeing Endgame are independent events.

Given the following information, answer the following questions.

1. Given that the events A and B are independent with $P(A) = 0.94$ and $P(B) = 0.4$, determine the value of $P(A|B)$.

2. Given that the events A and B are independent with $P(A) = 0.36$ and $P(B) = 0.7$, determine the value of $P(A \cap B)$.

3. Given events A and B such that $P(A) = 0.6, P(B) = 0.5$, and $P(A \cup B) = 0.8$, determine whether or not A and B are independent. Justify your answer. (Hint: First, determine $P(A \cap B)$.)

4. Students were asked whether they preferred Coke or Pepsi. The results are given in the table below.

	Coke	Pepsi
Male	10	15
Female	12	13

Is being female independent of liking Pepsi? Justify your answer using two methods.

5. Adults were asked whether they preferred baseball or football. The results are given in the table below.

	Male	Female
Baseball	30	15
Football	45	25

Is liking football independent being male? Justify your answer using two methods.

Skill #41: Statistics Vocabulary

- Just a bunch of definitions to memorize.

Types of Studies

- Controlled Experiment: Divide the sample into two groups called *treatment* and *control.* The *control* group serves as a baseline that receives no treatment (placebo). The groups must be *randomly assigned* to participants. Compare the results of treatment group to the control group.
- Observational Study: Measure or observe members of a sample in such a way that they are not affected by the study. *You do not "do" anything to them – you just observe.*
- Survey: Ask a sample of people a series of questions. This sample, if chosen randomly, can represent the entire population. (If a survey is given to everyone in a population, it is called a *census.*)

Types of Groups

- Population: Everyone in a group.
- Sample: A group that represents the population.

A survey of 100 randomly selected high school students showed that 90% of them were excited for summer.

- What is the sample and what is the population?
  The sample is the 100 randomly selected high school students.
  The population is all high school students.
- Given this information, what can you conclude about the *population* of all high school students?
  We can conclude that it is likely that approximately 90% of the population will be excited for summer.

Types of Samples

avoid bias by trying to create a random sample.

- Random Sample: All the members of a population of equally likely to be chosen.
  example: a survey is given to everyone who is selected by a random number generator
  RANDOM SAMPLES ARE UNBIASED!
- Systematic Sample: Order the population in some way. Then select it using regular intervals.
  example: a survey is given to every 5^{th} person who enters the mall
  SYSTEMATIC SAMPLES ARE UNBIASED! (They are a type of random sample.)
- Convenience Sample: Select any members of the population who are conveniently and readily available.
  example: a survey is given only to the friends and family of the researcher
  CONVENIENCE SAMPLES ARE BIASED!
- Volunteer or Self-Selected Sample: Select only members of the population who volunteer.
  example: a survey is left on a table in the mall to be done by those who volunteer.
  VOLUNTEER SAMPLES ARE BIASED!

Given the definitions above, answer the following questions.

1. A coach wants to test the effect of a new sports vitamin drink on mile time. She separates her sample into two groups, giving one group the new sports vitamin drink and giving the other group water. After a certain amount of time, she compares the results of a timed mile trial. Which type of study best describes this situation?

2. A researcher wants to determine if there is a link between attendance during the school year and summer school attendance. He gathers data from school districts and compares the results. Which type of study best describes this situation?

3. **Multiple Choice.** Which scenario might *not* lead to biased sampling?

   (1) Giving all those who answer a survey a free entry in a drawing to receive a new phone.
   (2) Having people call in to a company to give feedback about a truck driver's driving.
   (3) Distributing a survey to every 5^{th} person who enters the mall.
   (4) Giving a survey about baseball coaching to those people who you know are baseball coaches.

4. Mary is an interior designer and orders 500 white roses for an upcoming design. After they arrive, she randomly chooses 10 dozen roses from the truck and brings them inside.
   a. Identify the sample and the population in the given scenario.

   b. State one conclusion that could be made about the population if Mary finds that 2% of the sample was found to be unsatisfactory.

5. **Match the type of sampling to the scenario. Write the number on the line.**
   On Wednesday during homeroom, the four Student Senate officers (Kevin, Mary, Sarah, and Larry) are given a stack of surveys about the school construction and told that they are responsible for getting the surveys completed by their peers. They split up the stack between the four of them and then...

   _______ A. Kevin gets his surveys completed by every student in his homeroom.

   _______ B. Mary stands outside school on Thursday and has every third person who enters the school complete her survey.

   _______ C. Sarah leaves her surveys in the lunchroom with a sign on them that says, "Please take this survey!"

   _______ D. Larry walks to each homeroom and has two students from each complete his survey.

Options:				
	(1)	Random sample	(3)	Voluntary sample
	(2)	Systematic sample	(4)	Convenience sample

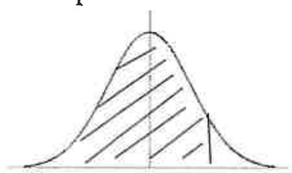
Skill #42: Mean and Standard Deviation (Calculator Skill)

- Combo move: 2^{nd} – VARS – 2:normalcdf – then input the lower bound (-10000 if you don't have one), the upper bound (10000 if you don't have one), the mean μ, and the standard deviation, σ.

A certain type of dog has a mean of 22 pounds and a standard deviation of 6.5 pounds.

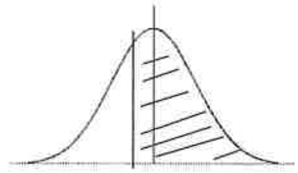
- What percentage of dogs weigh **less than** 33 pounds?

$$normalcdf(-10000, 33, 22, 6.5) = 0.954706368 = 95.5\%$$



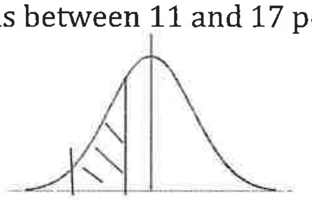
- What percentage of dogs weigh **more than** 21 pounds?

$$normalcdf(21, 10000, 22, 6.5) = 0.5611344957 = 56.1\%$$



- What is the probability, to the *nearest hundredth*, that a dog weighs between 11 and 17 pounds?

$$normalcdf(11, 17, 22, 6.5) = 0.1755844611 = 0.18$$



Out of 700 dogs, how many would weigh between 11 and 17 pounds?

$$700(0.18) = 126 \text{ dogs}$$

Given the following data, answer the questions.

1. If the amount of time that students spend doing extracurricular activities each week is normally distributed with a mean of 8 hours per week and a standard deviation of 3 hours,
   a. What is the probability that a student does between 4 and 9 hours of extracurricular activities each week?

   b. Given 700 students, how many students would complete between 4 and 9 hours of extracurricular activities?

2. A shoe manufacturer collected data regarding women's shoe size and found that the sizes were normally distributed with a mean size of 8.5 and a standard deviation of 1.5.
   a. Determine the probability that a woman's shoe size is less than 9.

   b. Given 400 women, how many would have a shoe size less than 9?

   c. Determine the probability that a woman's shoe size is greater than 8.

   d. Given 560 women, how many would have a shoe size greater than 8?

   e. Determine the probability that a woman's shoe size would be between 5 and 8.5.

   f. Given 335 women, how many would have a shoe size between 5 and 8.5?

3. Men's heights follow an approximate normal distribution with a mean of 70 inches and a standard deviation of 3 inches.
   a. Determine the percentage of males whose heights are above 74 inches.

   b. Determine the percentage of males whose heights are below 68 inches.

   c. Determine the percentage of males whose heights are between 67 and 74 inches.

Skill #43: Confidence Intervals

- To make a confidence interval: take your mean and add and subtract 2 standard deviations. Write your answers in parentheses with the lower one first – (_____ , _____). Usually you will round to the *nearest hundredth!*
- If your value is within your confidence interval, it is an ordinary (usual) value. If your value is outside your confidence interval, it is unusual!

Juan is a high school student with $4500 in his savings account. A study at his high school surveyed 98 students and found a mean savings of $2500 with a standard deviation of $75.50. Given this data, and using a 95% confidence level, is Juan's savings unusual? Justify your answer.

$$2500 \pm 2(75.50) \Rightarrow 2500 + 2(75.50) = \$2651$$
$$\Rightarrow 2500 - 2(75.50) = \$2349$$

The 95% confidence interval is $(2349, 2651)$. Juan's amount is unusual since it falls outside the confidence interval.

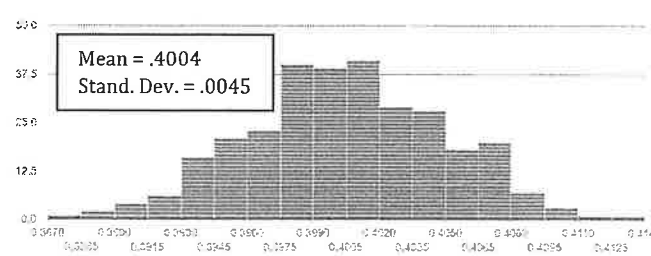
Given the information, answer the following questions.

1. A teacher gives a standardized test to his class. The mean score is 635 points and the standard deviation is 28.
   a. Using this data, create a 95% confidence interval and interpret this confidence interval in context.

   b. What is the margin of error?

2. A study of a local high school tried to determine the mean number of text messages that each student sent per day. The study surveyed a random sample of 111 students in the high school and found a mean of 199 messages sent per day with a standard deviation of 75 messages. Jake is a student at this school and sends 300 texts one day. Using a 95% confidence interval, is Jake's amount unusual? Justify your answer.

3. The CEO of a company claimed that 40% of his employees would like to participate in a community day. A simulation based on this claim was created. The chart shows the results of 100 surveys simulated 300 times. The CEO conducted a survey of a random sample of 100 employees and 35 said that they wanted to participate. Based on the results of the simulation, and assuming a 95% confidence level, is this a reasonable result? Explain.



Skill #44: P-Hat and Standard Error

- If you're not given enough information, you can find your estimated mean, p-hat $\hat{p}$, and your estimated standard deviation (error) using: $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
- Yes, you would need to memorize this formula. However, this is the last formula I would try to memorize. It has not been needed very often.

A survey was given to a random sample of 1250 residents of a city to determine whether or not they would support a community grid. Of those surveyed, 672 respondents said they were in favor of the plan. Determine a 95% confidence interval, to the *nearest hundredth,* for the proportion of people who favor the community grid. What is the margin of error?

$$\hat{p} = \frac{672}{1250} = 0.5376 \qquad SE = \sqrt{\frac{0.5376(1-0.5376)}{1250}} = 0.014$$

The margin of error is 2 standard deviations $= 2(0.014) = 0.028$

The 95% confidence interval is: $0.5376 \pm 2(0.014) \Rightarrow (0.5096, 0.5656) \Rightarrow (0.51, 0.57)$

Answer the following questions.

1. A survey was given to a random sample of 1600 voters in the United States to determine if they favored an increase in taxes to benefit education. 960 of the respondents stated that they would favor such an increase.
   a. Determine the sample proportion.

   b. Determine the standard error and margin of error. Round to the *nearest thousandth.*

   c. Determine a 95% confidence interval, to the *nearest thousandth.*

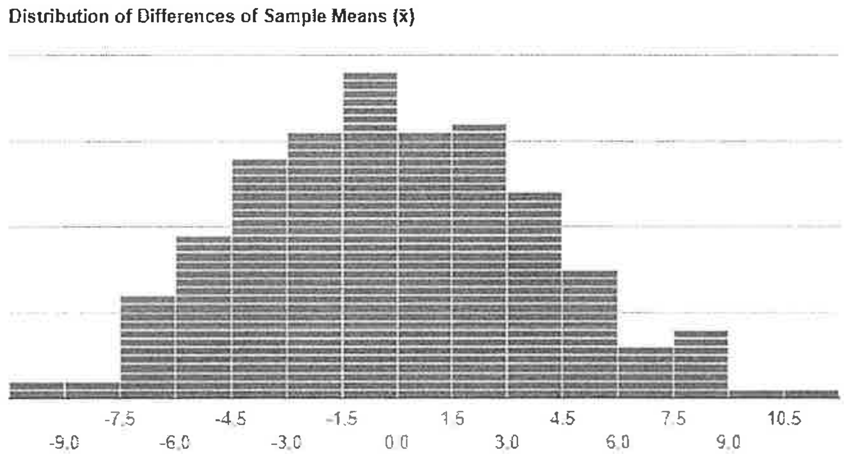
2. A survey of 50 students was given to determine how many favored an updated dress code policy. Out of the 50 surveyed, 15 said they wanted to update the policy. A simulation based on this claim was completed 500 times. Determine the sample proportion and margin of error.

Skill #45: Differences in Means

- Given two sets of data, find the two means then find the difference between the two (subtract them).
- If that mean difference happens less than 2.5% of time (or in the tail ends of the distribution), then "WOW! That's statistically significant!" If it falls in the 95%, then "BORING! Not statistically significant."

20 students were randomly split up into two equally sized groups. Each member of Group 1 watched videos for homework and each member of Group 2 were given traditional problem sets as homework. Both groups were given the same test at the end of the year and the results are shown in the table below. A simulation was conducted in which the subjects' scores were rerandomized into two groups 250 times. The results are shown in the graph below. What inference can be made about the results?

Group 1	Group 2
71	77
98	64
71	73
67	75
81	77
88	59
76	67
79	75
63	74
71	63
Mean: 76.5	Mean: 70.4



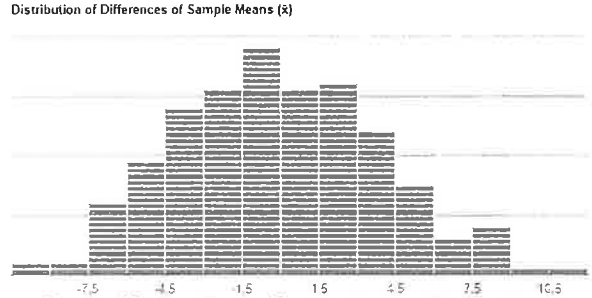
Our mean difference is: $76.5 - 70.4 = 6.1$

A mean difference of 6.1 or higher happens approximately $\frac{16}{250} = 0.064 = 6.4\% > 2.5\%$

Since this percentage is greater than 2.5%, the value of 6.1 does not fall in the tail ends of the distribution. Therefore, it falls within the 95% confidence interval. This means our mean difference is ordinary – it is likely to happen by chance.

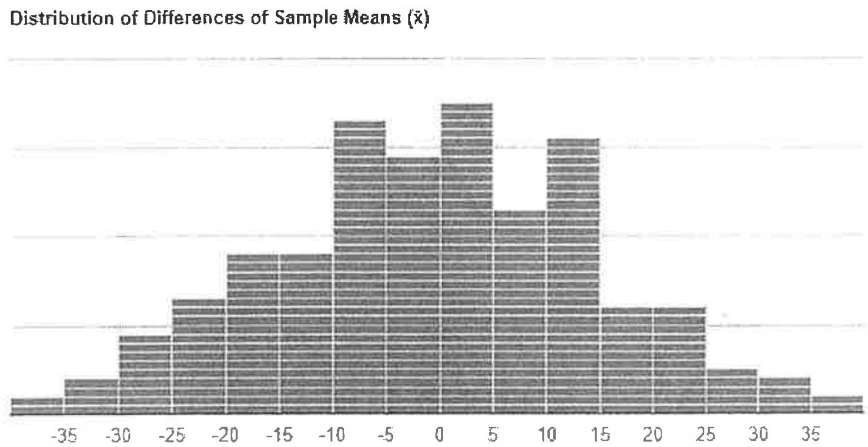
Given the data, answer the following questions.

1. Given the same scenario as above, Joelle found a mean difference between groups 1 and 2 to be -7.6. What inference can be made about this result. Justify your answer.



2. 20 adults were randomly split up into two equally sized groups. Each member of Group 1 exercised to lower cholesterol and each member of Group 2 was given medicine to lower cholesterol. Both groups were given the same blood test at the end of the year and the results are shown in the table below. A simulation was conducted in which the subjects' levels were rerandomized into two groups 250 times. The results are shown in the graph below. What inference can be made about the results?

Group 1	Group 2
195	180
200	192
220	205
235	240
242	170
282	150
275	195
260	250
221	220
200	215
Mean:	Mean:



3. 20 baseball pitchers were randomly split up into two equally sized groups. Each member of Group 1 did strength training to improve pitching speed and each member of Group 2 completed extra pitching exercises to improve pitching speed. The top speed of each pitcher in both groups were recorded at the end of the year and the results are shown in the table below. A simulation was conducted in which the subjects' speeds were rerandomized into two groups 300 times. The results are shown in the graph below. What inference can be made about the results?

Group 1	Group 2
80	94
98	81
85	89
92	92
96	95
88	98
91	88
95	93
94	95
99	85
Mean:	Mean: