Probability & Statistics Workbook

krista king
MATH

ONE-WAY DATA

1. Identify the variables in the following data description and classify the variables as categorical or quantitative. If the variable is quantitative, list the units.

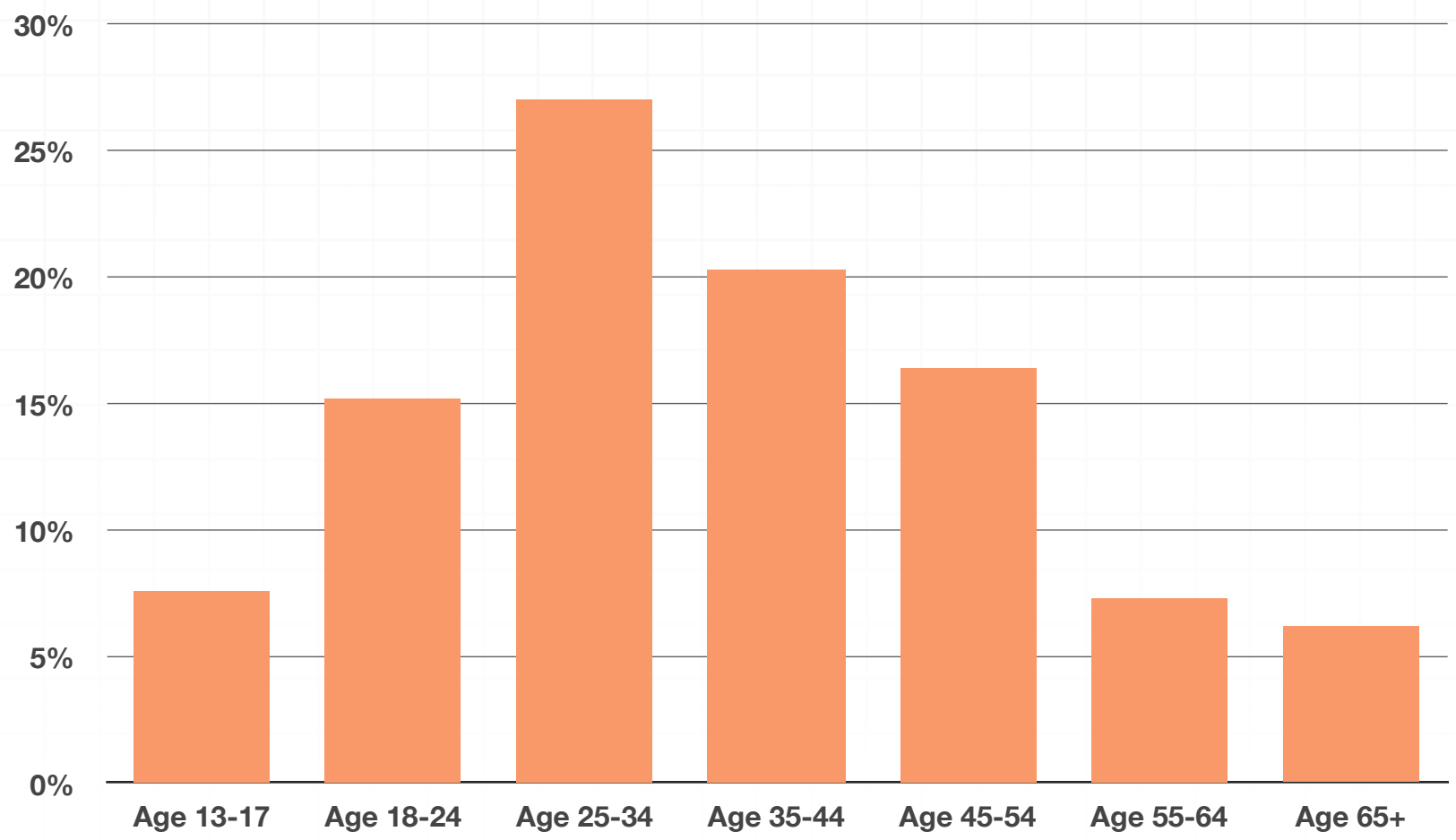
> "The Indianapolis 500 is a car race that's been taking place since 1911 and is often scheduled to take place over Memorial Day weekend. The race takes place at the Indianapolis Motor Speedway and a driver needs to complete 200 laps that cover a distance of 500 miles. Race results are reported by driver number, the driver's name, the type of car the driver uses, and the time to the nearest ten-thousandth of a second. If a driver doesn't finish the race, instead of the time to complete the race, their number of laps completed is recorded."

2. Casey is taking a survey of her senior class. She plans to ask the seniors this question:

> "In general do you think things have gotten better or worse for our students over the course of the year?"

Her survey has a checklist with these responses: Better, Worse, Stayed the same, and Don't know. Who are the individuals in the survey? What type of response variable is Casey looking for? Is it categorical or quantitative?

3. The graph below shows the age breakdown of Apple iPad owners in the United States in February, 2011. Who are the individuals in the data? What is the variable? Is it categorical or quantitative?



Source: www.statista.com

4. The table below shows the number of rejected products by worker and shift. Is the data below one-way data? Why or why not?

Worker ID	1st shift	2nd shift	3rd shift
1123	42	45	42
2256	45	74	32
6435	36	78	41

5. Why is this table an example of one-way data?

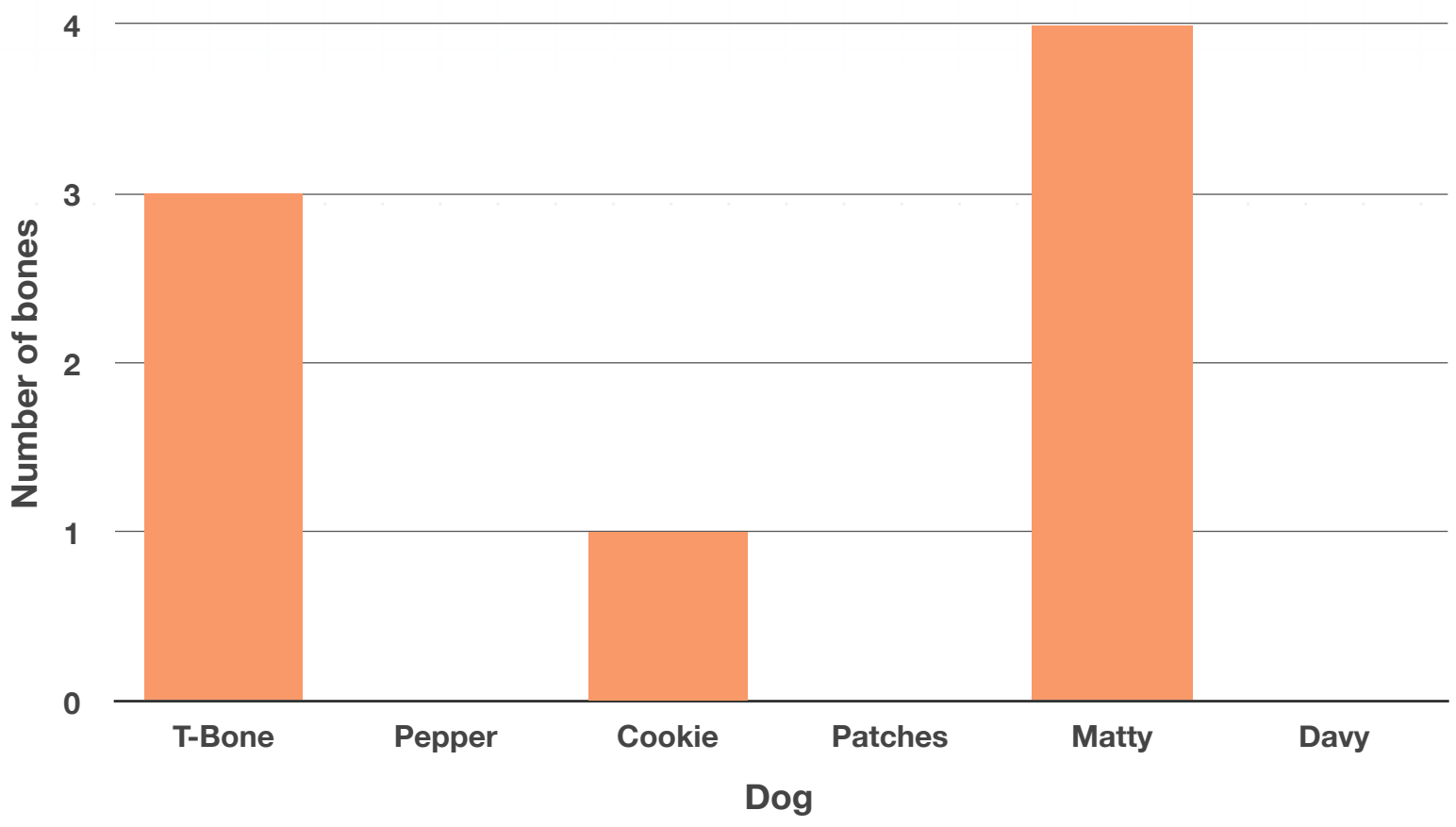
Flavor	Scoops sold	Contains chocolate?	Smooth or chunky?
Vanilla	300	No	Smooth
Chocolate	450	Yes	Smooth
Cookies & Cream	275	Yes	Chunky
Mint Chocolate Chip	315	Yes	Chunky
Fudge Brownie	375	Yes	Chunky
Rocky Road	250	Yes	Chunky

6. A botany student wants to test the claim of a diaper company that their product may be used in a compost pile. He creates 12 identical gardens and plants a random selection of 7 tomato plants in each one. He plans to have a fellow student use traditional compost on 6 of the garden plots and the compost from the diapers on the other 6. He does this so he doesn't know which plot is which. He plans to check the tomato plants for disease every two days for a month, and record the number of tomato plants with disease after each check. Would this experiment result in one-way data? Why or why not?

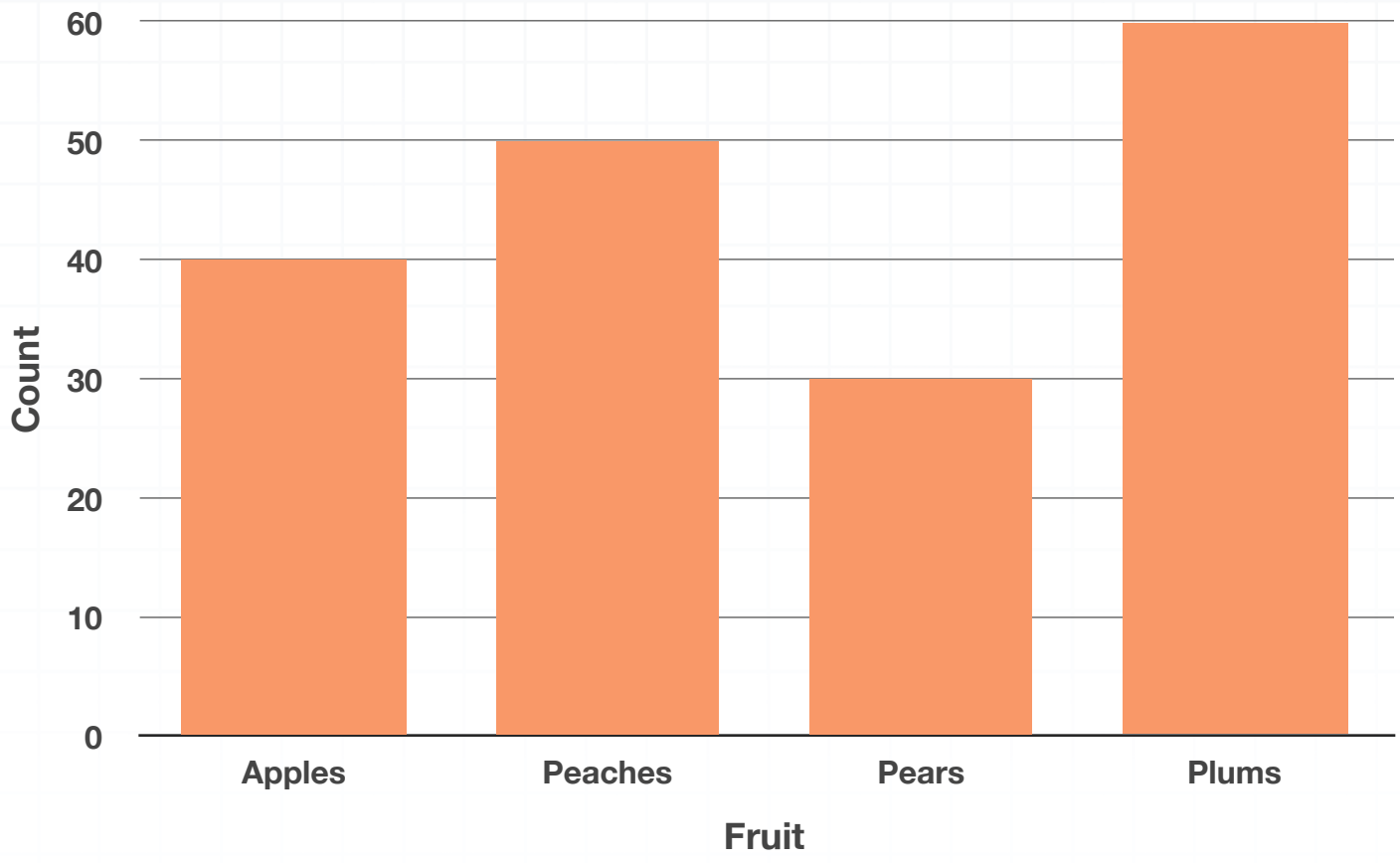
BAR GRAPHS AND PIE CHARTS

1. Both the bar graph and the table have missing information about the number of bones each dog consumed at doggie daycare. Use the graph and table together to fill in the missing pieces.

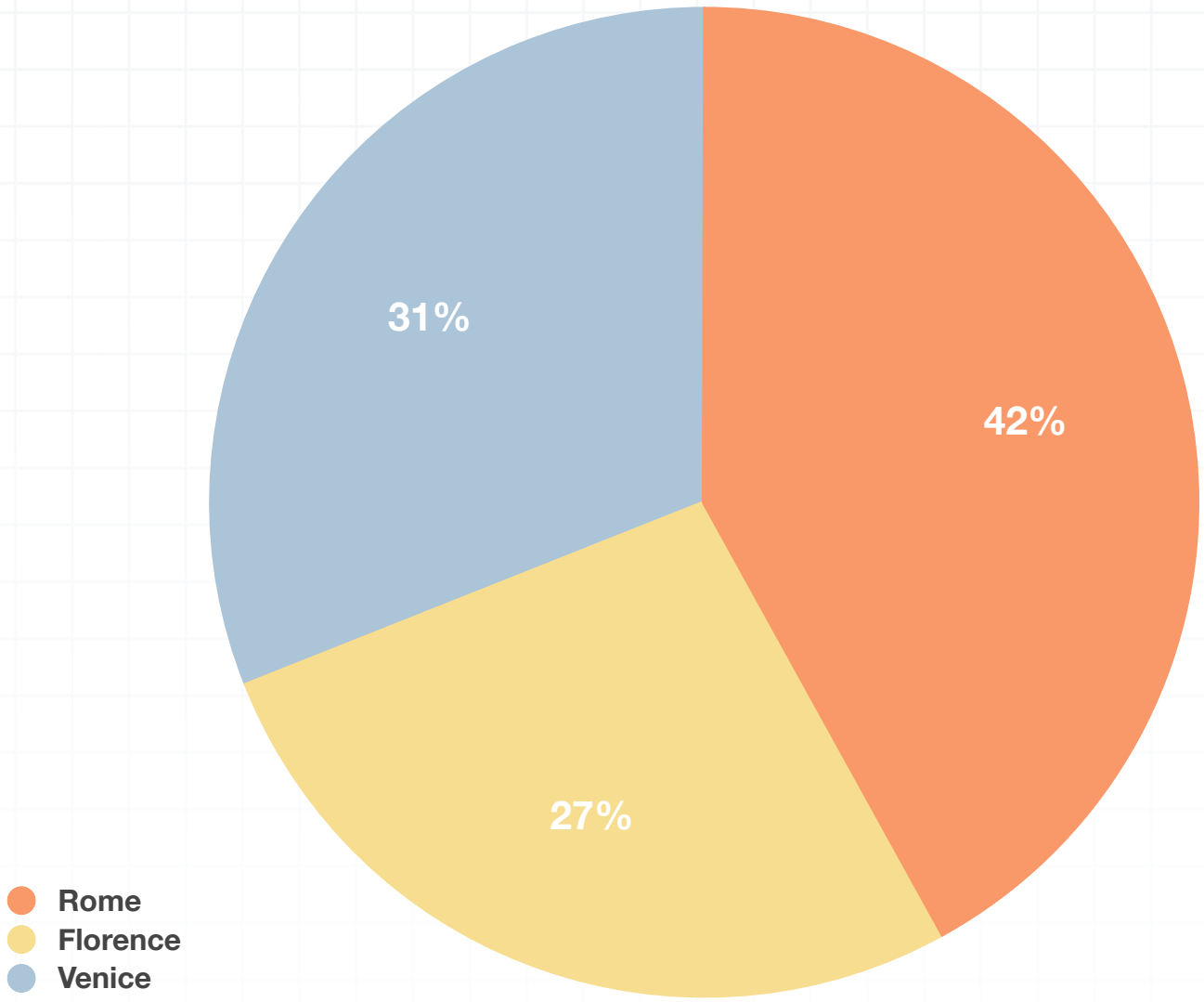
Dog	Number of bones
T-Bone	
Pepper	1
Cookie	
Patches	5
Matty	
Davy	2



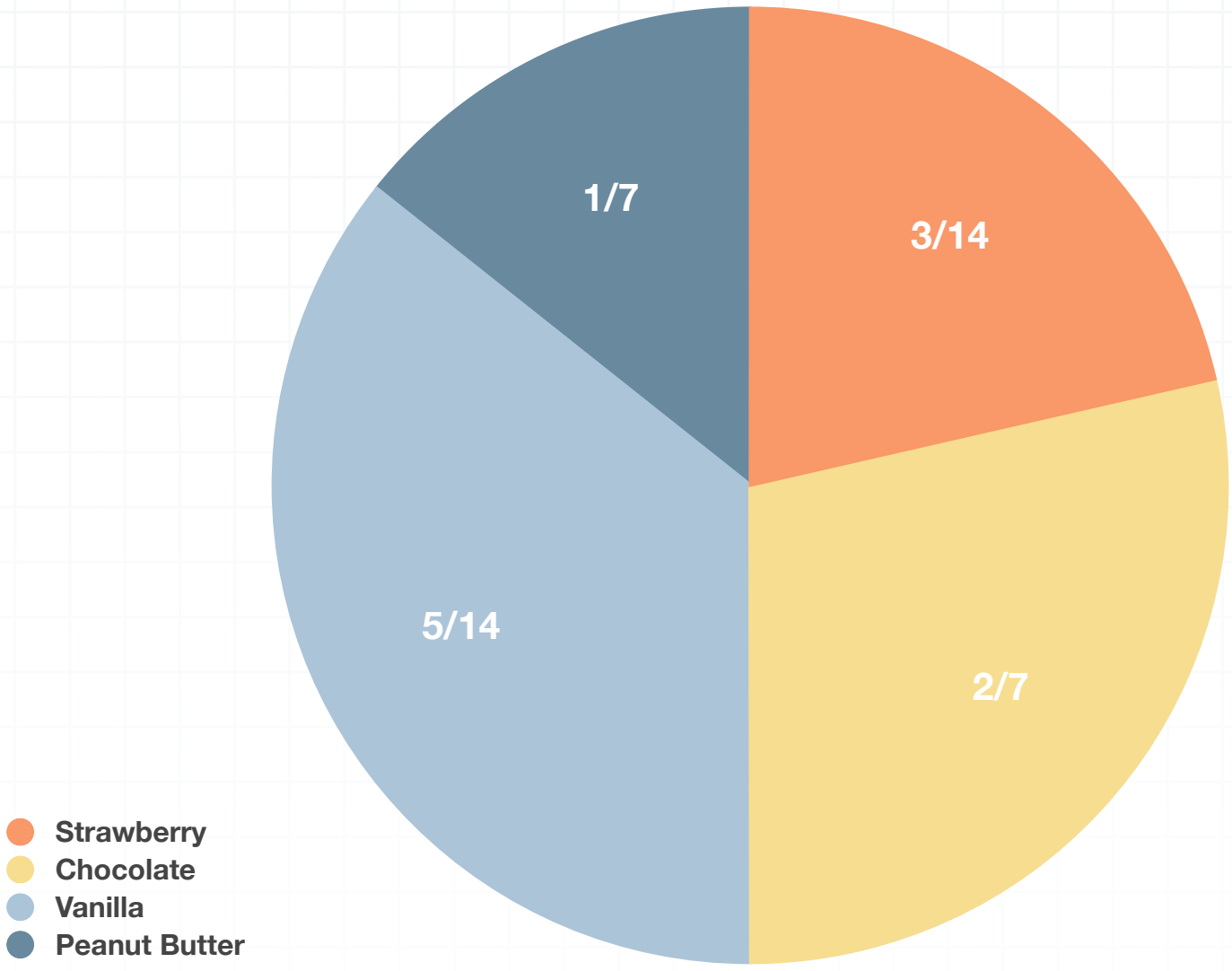
2. Eric's class went on a trip to an orchard. At the end of the trip they counted how many pieces of fruit came from each type of tree and graphed it in the bar graph shown below. Use the bar graph to create a pie chart of the data.



3. A tourist company took a survey of 600 clients and asked them which Italian city they were most interested in visiting. How many clients said they wanted to visit Rome?



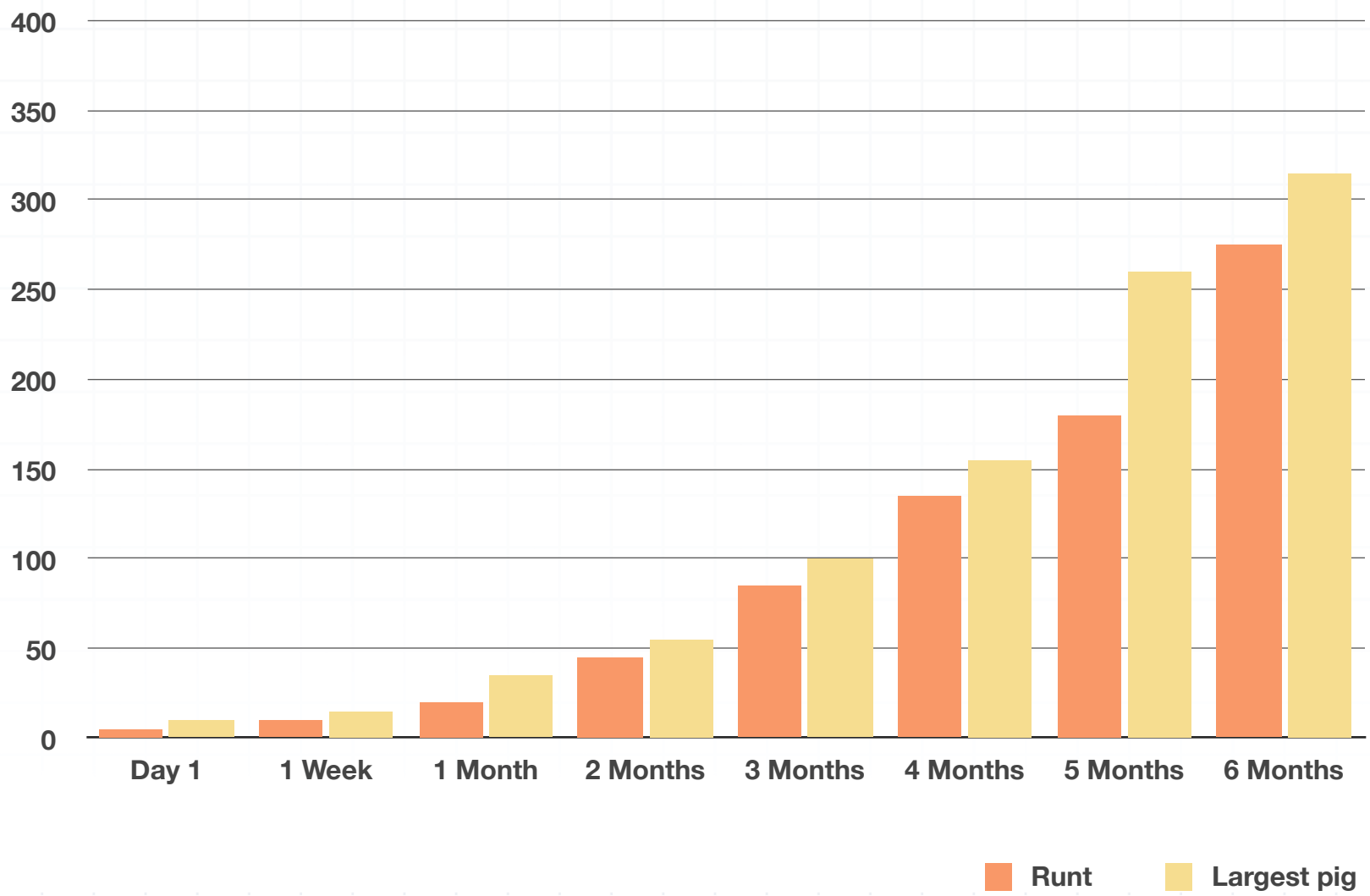
4. The pie chart shows how many ice cream cones of each flavor were sold. Assuming 280 total ice cream cones were sold in August, convert the pie chart to a bar graph.



5. A company is analyzing the results from a recent survey about why people left their employment. The results are shown in the data table below. In general, is a bar graph or a pie chart a better choice to display the data? Why?

Reasons for leaving job	
Reduced job duties	30%
Company restructuring	15%
Too much travel time	12%
Looking for more opportunity	11%
Need more personal time	9%
Poor expected company growth	8%

6. The comparison bar graph shows the growth of two pigs over their first 6 months of life. Which pig grew the most between 4 and 5 months?



LINE GRAPHS AND OGIVES

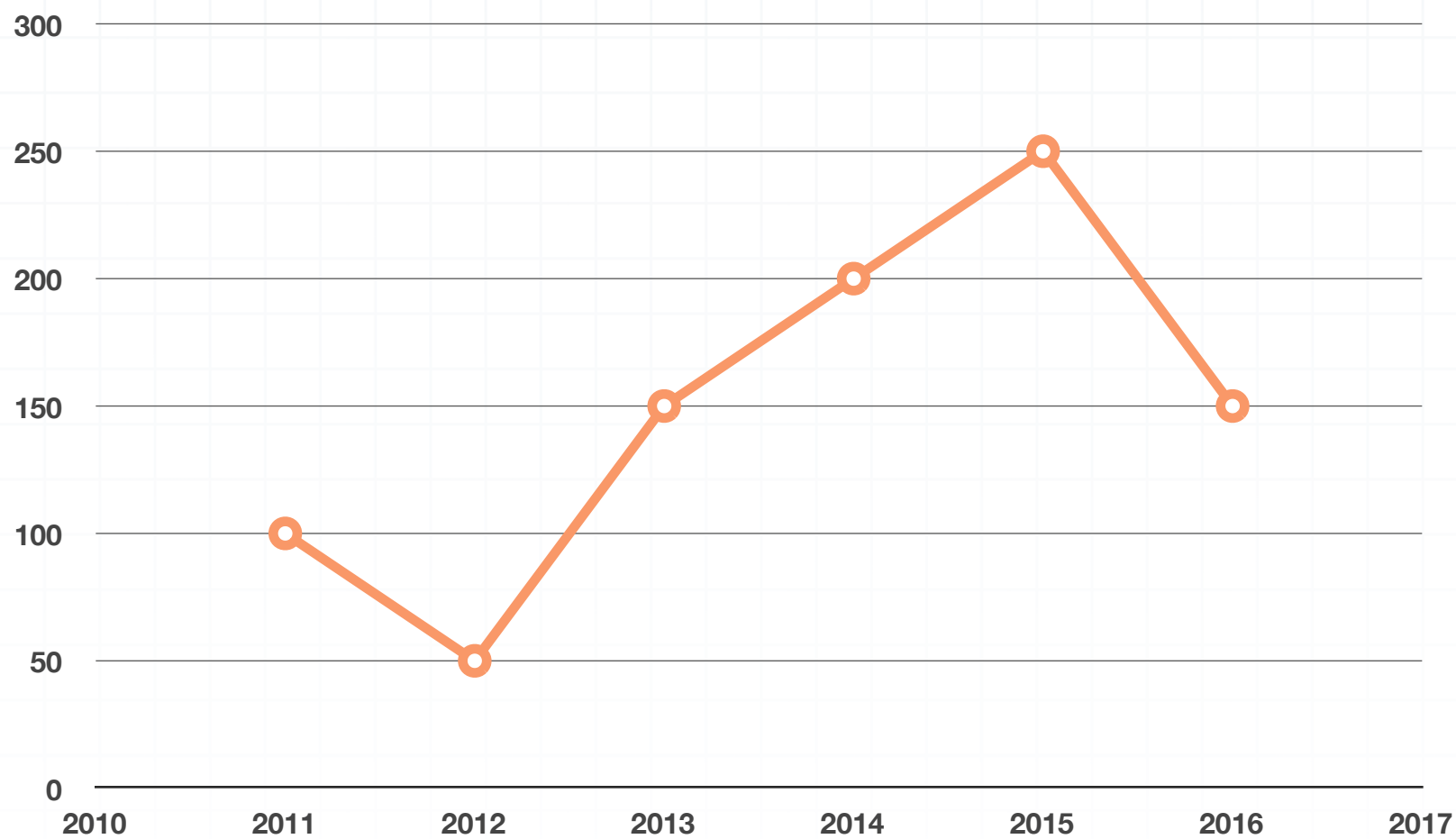
■ 1. Bethany started a sit-up program so that she can do 200 sit-ups in a day. At the end of week 6 she'll have completed $1{,}685$ sit-ups. Create an ogive of the data.

Week	Number of sit-ups
Week 1	350
Week 2	455
Week 3	600
Week 4	540
Week 5	1,275
Week 6	1,685

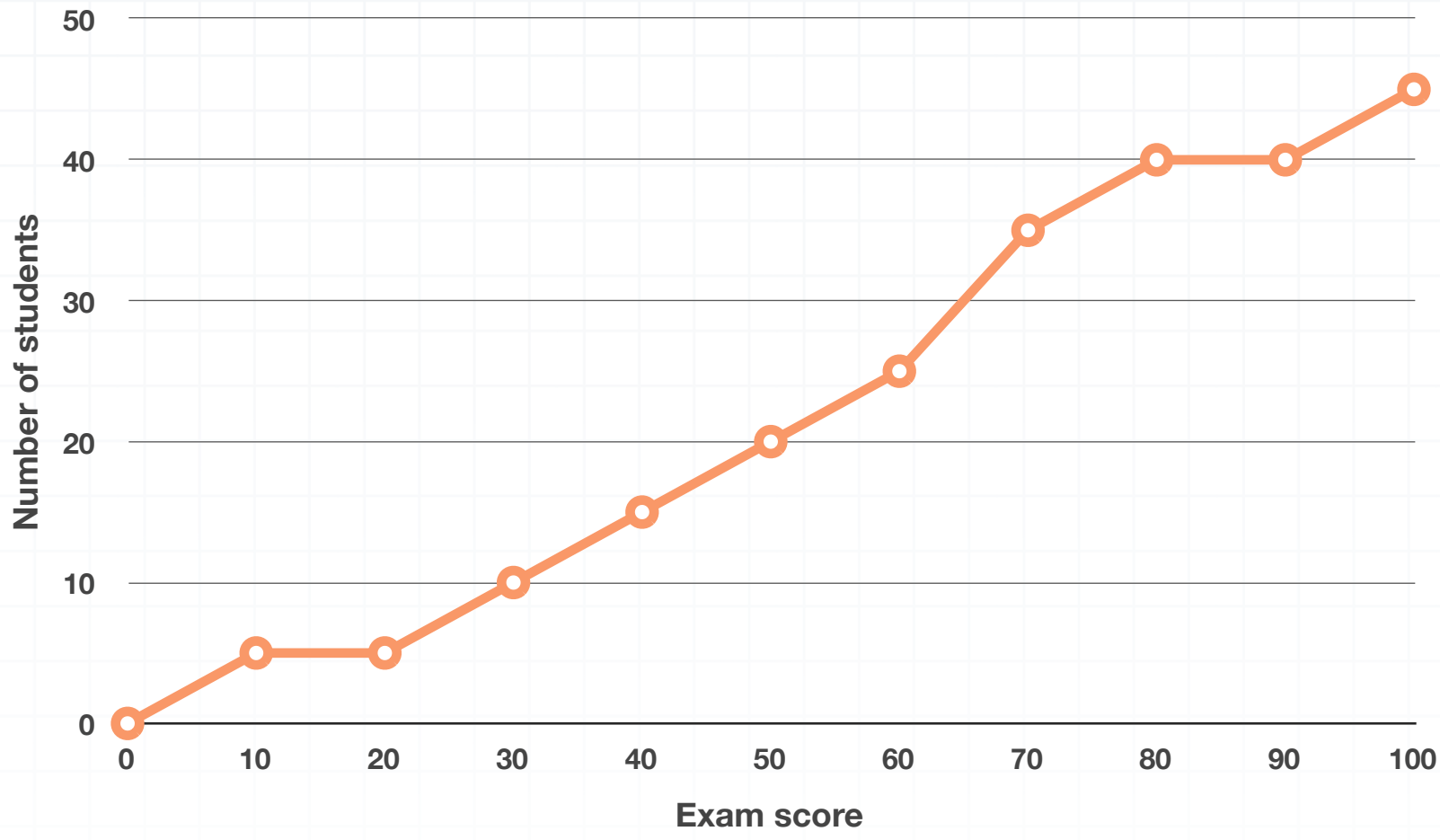
■ 2. The table shows passengers by year for Buster's Bus Service. Create a line graph of the data in the table.

Year	Passengers
2011	1,000
2012	500
2013	1,500
2014	2,000
2015	2,500
2016	1,500

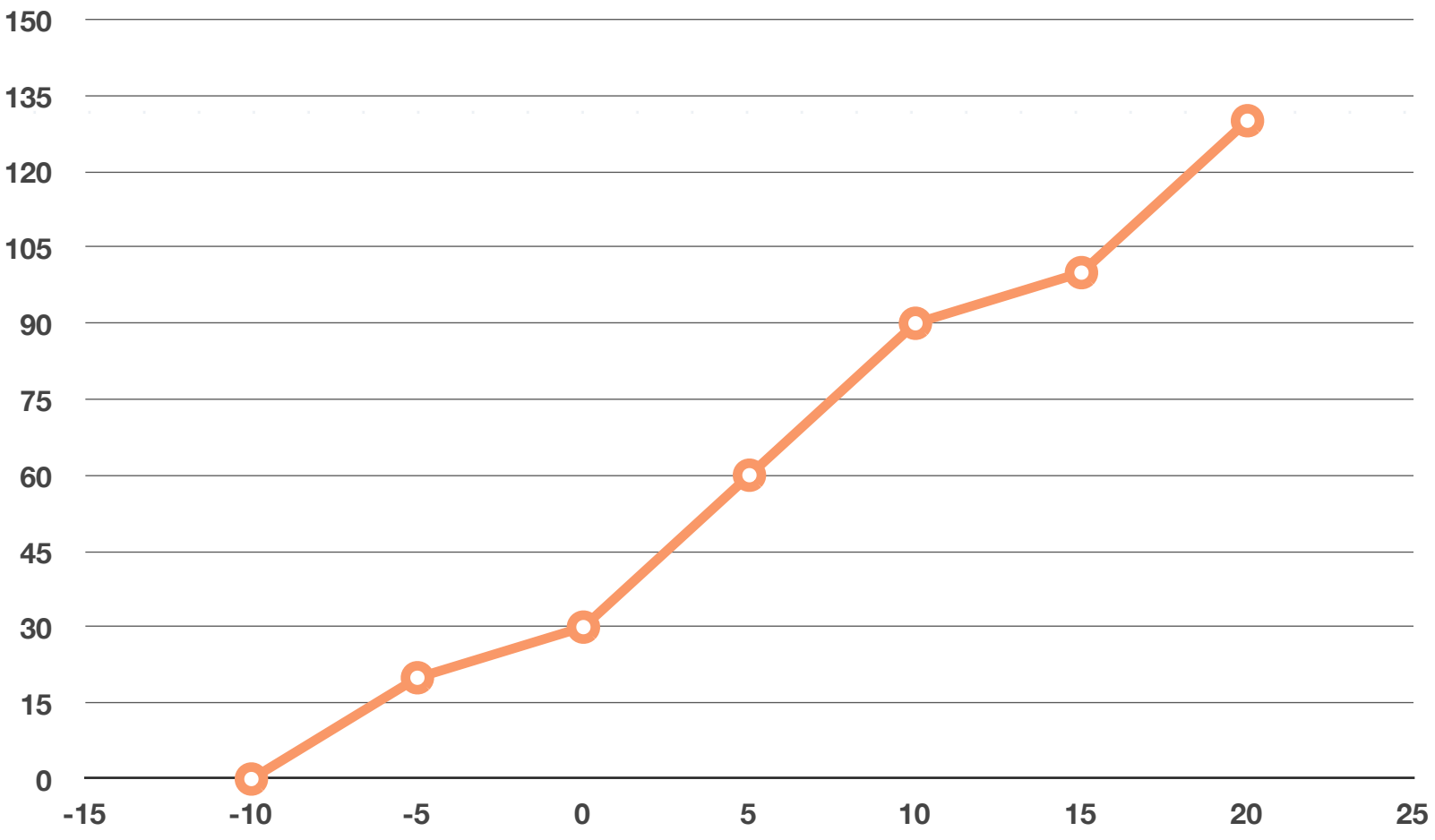
3. Between what two consecutive years was there the largest increase in car sales?



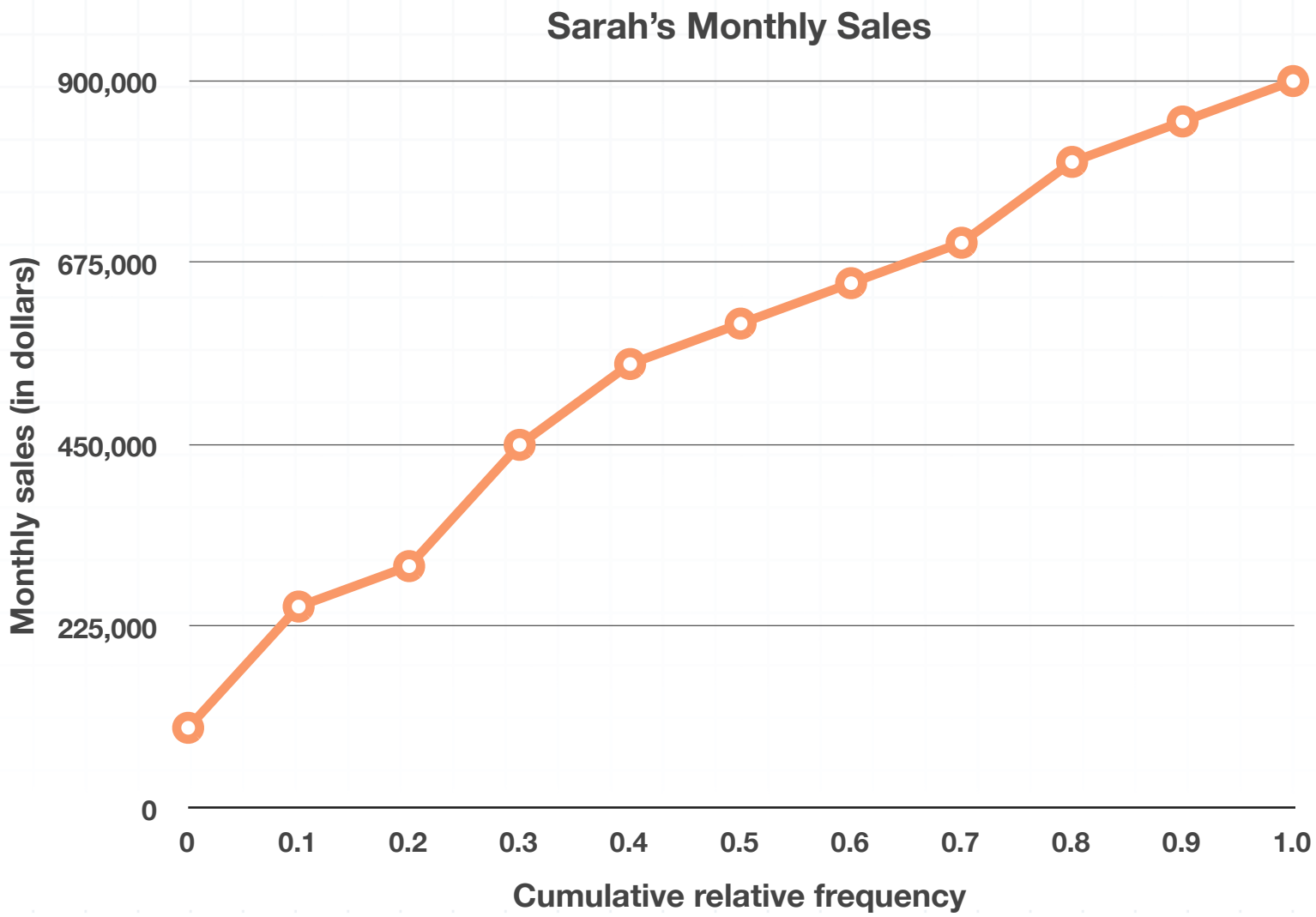
4. Mrs. Moore gave her students a midterm exam, then she created this ogive of the 45 exam scores. How many students got a score between $70\,\%$ and $90\,\%$?



5. Draw the line graph that corresponds to the ogive below.



6. Sarah's monthly sales to date are shown in the cumulative relative frequency plot below. What is the meaning of the circled point?



TWO-WAY DATA

■ 1. Create a comparison bar graph for the two-way table.

Favorite pet	Fish	Cat	Dog	Other
1st grade	8	15	7	9
2nd grade	13	10	12	5

■ 2. A pizza parlor wants to know if the age range of their customers affects pizza preferences. The pizza parlor asks each customer two questions:

1. Which type of pizza is your favorite: pepperoni, cheese, supreme or veggie?

2. What is your age range: Under 18, or 18 and over?

The results of the survey are as follows:

Of the 50 customers who prefer pepperoni pizza, 25 are under 18.

Of the 20 customers who prefer cheese pizza, 18 are under 18.

Of the 30 customers who prefer supreme pizza, 24 are over 18.

Of the 25 customers who prefer veggie pizza, 19 are over 18.

What type of data is the pizza parlor collecting, one-way or two-way? Create the best type of frequency table for the data.

■ 3. An elementary school creates the following two-way table. What is the best name for the row variable and what is the best name for the column variable?

	Walk	School bus	Day care vehicle	Carpool
Pre-school	1	10	20	26
First	5	12	14	19
Second	10	22	5	15
Third	8	33	3	10

■ 4. Which graph would be a better choice to display the data from the two-way table: a comparison bar graph or a comparison line graph? Create your chosen graph.

		Method of transportation			
		Walk	**School bus**	**Day care vehicle**	**Carpool**
Grade in school	**Pre-school**	1	10	20	26
	First	5	12	14	19
	Second	10	22	5	15
	Third	8	33	3	10

■ 5. Eric creates a survey asking students who ate a snack in the morning between classes if they felt sleepy or not. Here are his survey results:

Snack	Yes	Yes	No	No	No	No	Yes	No	Yes	No	Yes	Yes	No	Yes	No
Sleepy	Yes	Yes	Yes	No	No	No	No	Yes	Yes	No	Yes	Yes	No	No	Yes

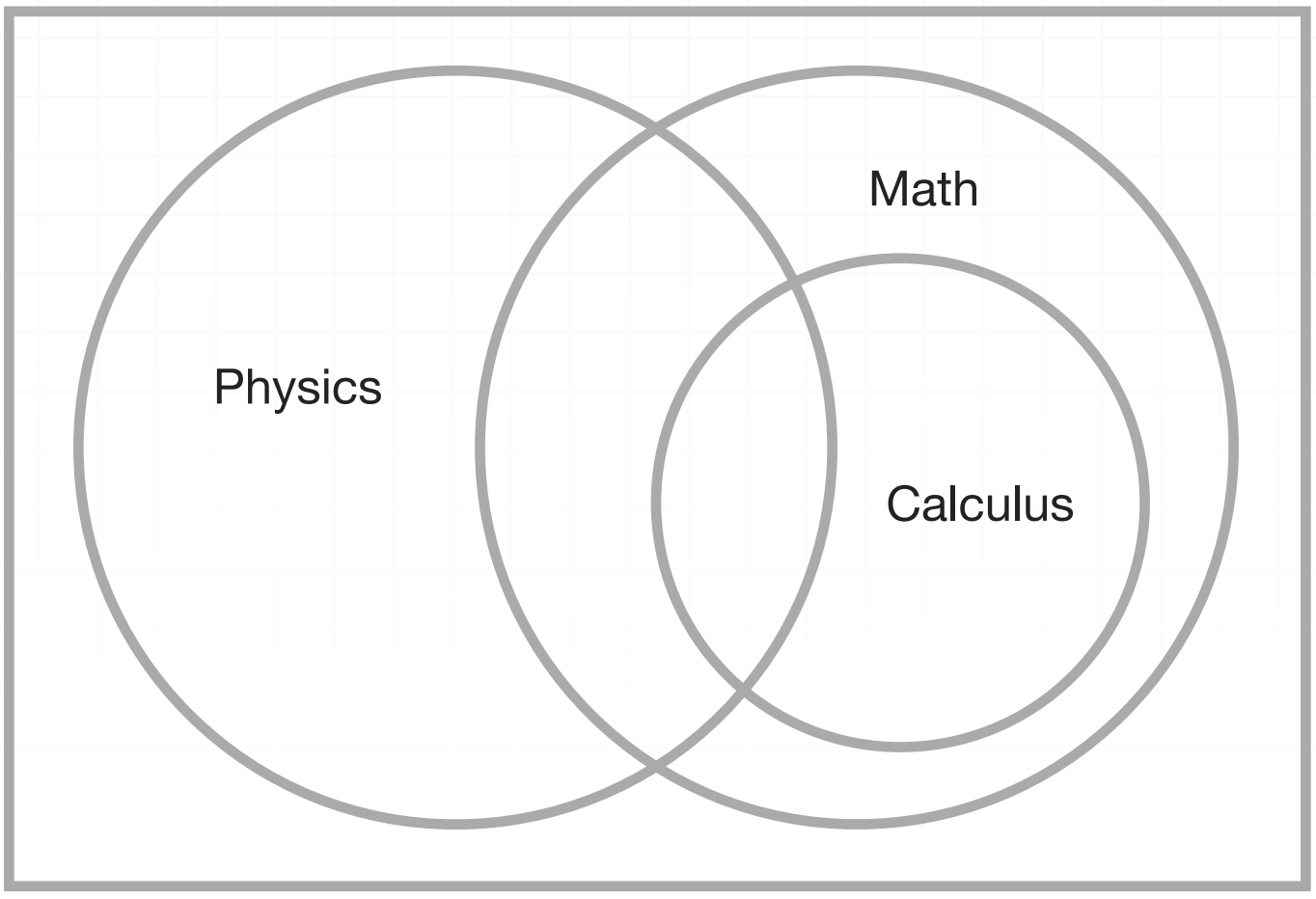
Create a two-way data table for Eric's survey.

■ 6. Is a comparison line graph an appropriate visual display for the data table, which shows monthly rainfall (in inches) for Dallas, Texas, January - August? Why or why not? If it's an appropriate display, create a comparison line graph. If it's not an appropriate display for the data, create a comparison bar graph.

	2015	2016	2017
January	3.62	1.04	4.39
February	2.96	2.20	2.33
March	2.53	2.67	1.06
April	5.56	4.60	3.38
May	16.96	6.25	0.70
June	3.95	3.60	8.44
July	0.92	3.89	4.12
August	0.46	4.42	4.24

VENN DIAGRAMS

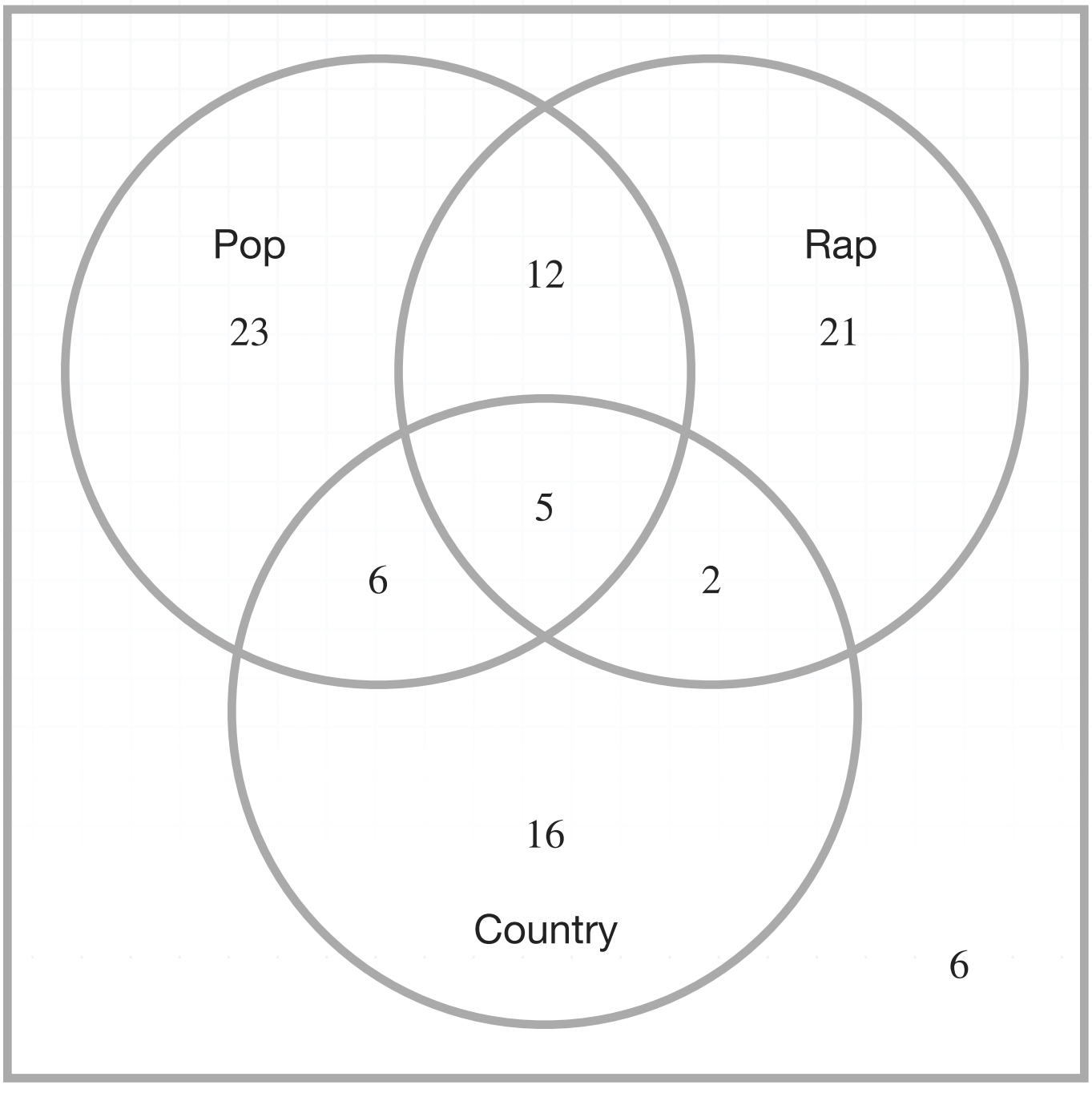
■ 1. What does the Venn diagram show about how Calculus is related to Physics and Mathematics?



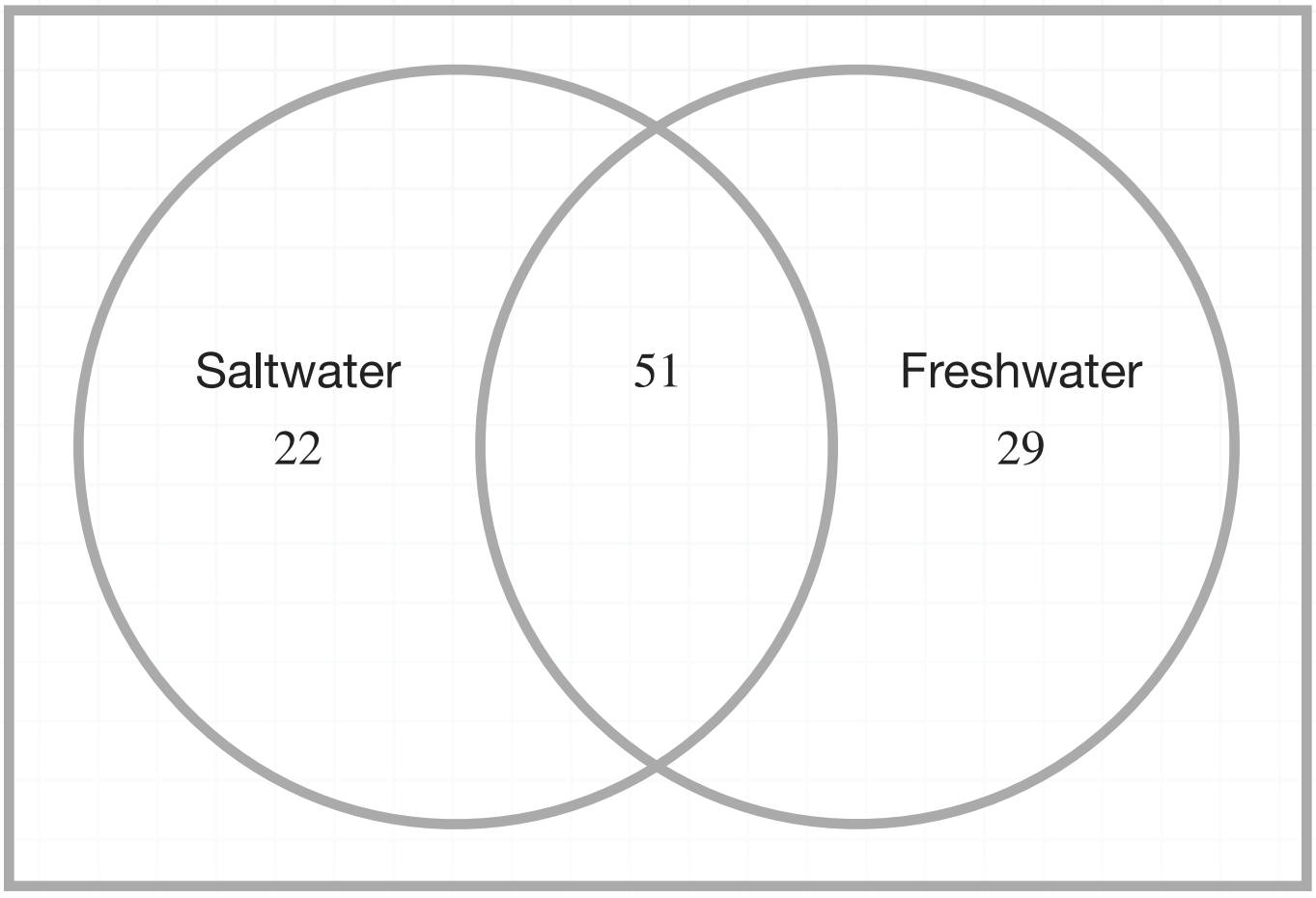
■ 2. Draw the Venn diagram for the number of humans in a room and the number of frogs in a room, if the room has 12 frogs and 15 humans.

■ 3. Students at Green Bow High School conducted a survey during lunch time to see what kind of music the students at the school liked. They recorded their results in a Venn diagram. How many students participated

in the survey? What percentage of the students who participated did not like Pop Music?



■ 4. A survey team is collecting data on a type of minnow that lives where a river meets the sea. They place nets in the river, where the river and sea meat and where there is only sea. They count the minnows caught in each net. What percent of the minnows were living in the brackish water? Brackish water is water that is a combination of fresh and saltwater.



5. Fill in the Venn diagram using the following information.

18 people's favorite exercise was swimming.

13 people's favorite exercise was running.

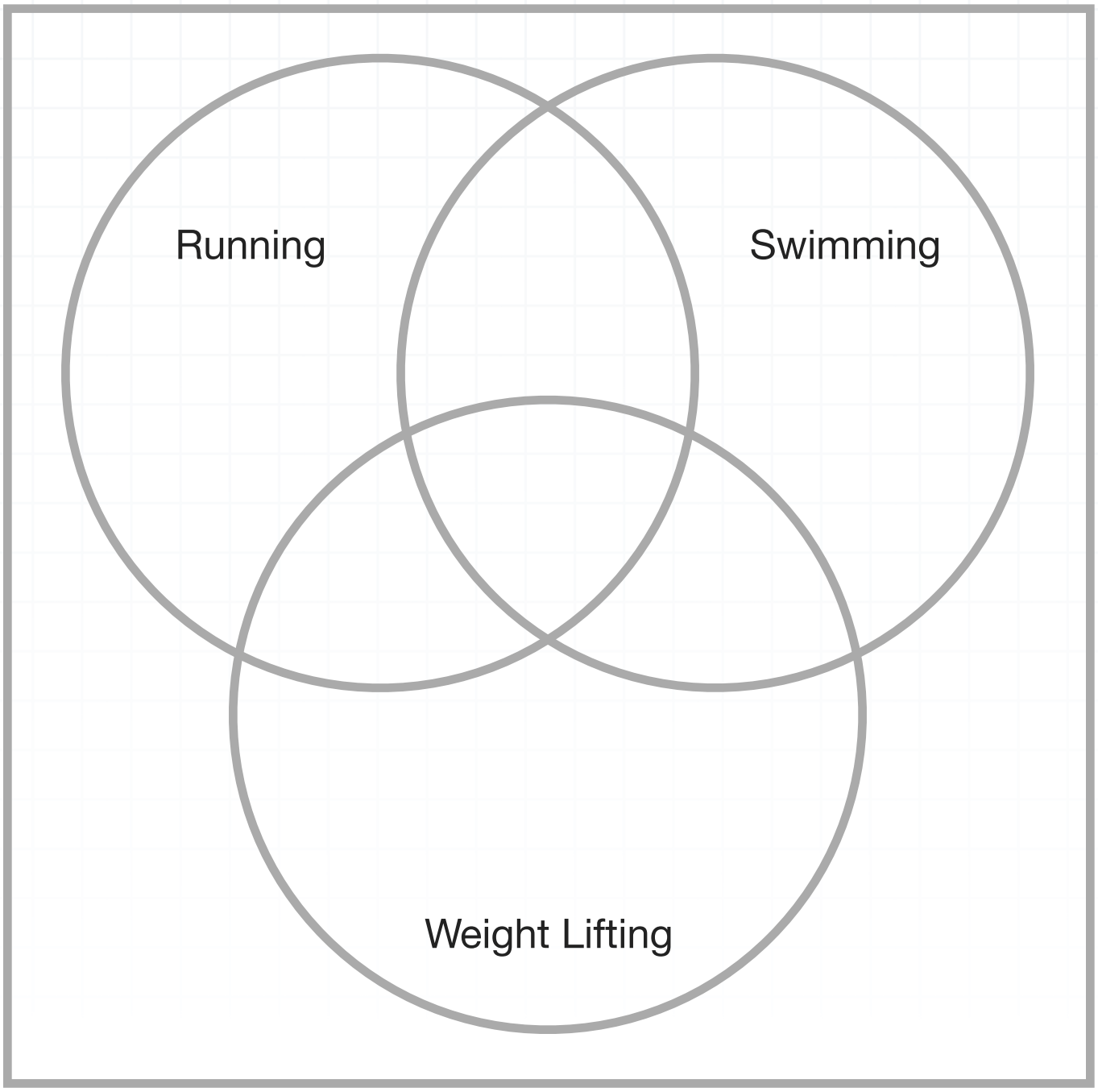
10 people only liked weight lifting.

5 people liked swimming and weight lifting equally.

4 people liked running and weight lifting equally, but not swimming.

5 people liked running and swimming equally, but not weight lifting.

2 people liked all three equally.



6. Eric creates a survey asking students who ate a snack in the morning between classes if they felt sleepy or not. He organizes his survey results into a two-way data table. Draw a Venn diagram for Eric's survey results.

		Do you feel sleepy?		
		Yes	**No**	**Total**
Did you eat a snack?	**Yes**	5	2	**7**
	No	3	5	**8**
	Total	**8**	**7**	**15**

RELATIVE FREQUENCY TABLES

1. Blake is surveying students in his class (made up of juniors and seniors) about whether or not they play video games on a daily basis. What type of relative frequency table is shown? Finish filling in the table.

	Play at least one video game daily	Don't play any video games daily	Total
Junior	23%		**75%**
Senior		14%	
Total			**100%**

2. Create the row-relative frequency table for the frequency table below displaying 9th grade students who participate in an after school activity, and then answer the question: What percent of female 9th grade students do not participate in an after school activity?

	Participate	Don't participate
Male	62	40
Female	57	38

3. Create the column-relative frequency table for this data table and then answer the question: What percentage of those who participate in an after school activity are male?

	Participate	Don't participate
Male	62	40
Female	57	38

4. Create the total-relative frequency table for the data, and then answer this question: Carl is in charge of creating an activity for the students in his college dorm. If Carl wants the highest possible turnout, which activity should he choose? Why?

	Movie	Bowling	Pizza Party
Male	20	40	55
Female	35	50	62

5. A city hall is looking into a dangerous intersection that has caused many bicycle accidents over the past month, due to rerouted traffic. They have counted the number of bicycle accidents and put them into a frequency table like the one below. Create the relative frequency table for the data and answer the following question: What day had the highest percentage of bicycle accidents?

Day of the week	Number of crashes
Sunday	13
Monday	10
Tuesday	8
Wednesday	6
Thursday	2
Friday	11
Saturday	14

6. Addie took a poll of the children in her neighborhood. She found that 15 of them watch 2 hours or more of cartoons per day. Out of the 15 that watch 2 hours or more, 10 watched the cartoons on a device other than the television. There were also 12 children who watched less than 2 hours of cartoons per day. For those 12 children, 2 of them watched cartoons on a device other than a television. Construct a two-way table to summarize the data and then construct a total-relative frequency table for the data.

JOINT DISTRIBUTIONS

1. To study the relationship between votes for a new park and people who have children, a community group surveyed voters. What percentage of those surveyed had children? Is this part of the joint, conditional, or marginal distribution?

	For	Against	No opinion
Children	125	50	30
No children	40	150	60

2. To study the relationship between votes for a new park and people who have children, a community group surveyed voters. What percentage of those surveyed were for the park and had children? Is this part of the joint, conditional, or marginal distribution?

	For	Against	No opinion
Children	125	50	30
No children	40	150	60

3. To study the relationship between votes for a new park and people who have children, a community group surveyed voters. What percentage of those with no children had no opinion? Is this part of the joint, conditional, or marginal distribution?

	For	Against	No opinion
Children	125	50	30
No children	40	150	60

4. Carl is in charge of creating an activity for the students in his college dorm, and he records their preferences by activity and gender. What percentage of the female students prefer pizza? To answer the question, did you use a marginal, joint, or conditional distribution?

	Movie	Bowling	Pizza Party
Male	20	40	55
Female	35	50	62

5. A pharmaceutical company is testing heart burn as a side effect of its new pain reliever. What conclusions can you draw from the marginal distributions of the study?

	Pain reliever	Placebo	Total
Minor heartburn	4	171	**175**
Major heartburn	102	25	**127**
No heartburn	10,568	10,478	**21,046**
Total	**10,674**	**10,674**	**10,674**

6. Consider the same data as the previous question. What do the conditional distributions (given the participant experienced minor heartburn, major heartburn, or no heartburn) tell us about the study?

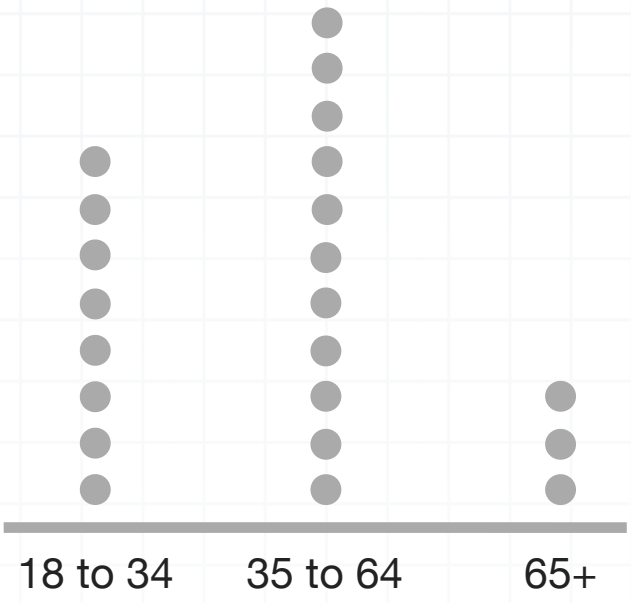
	Pain reliever	Placebo	Total
Minor heartburn	4	171	**175**
Major heartburn	102	25	**127**
No heartburn	10,568	10,478	**21,046**
Total	**10,674**	**10,674**	**10,674**

FREQUENCY TABLES AND DOT PLOTS

1. The frequency table shows the number of seed packets sold by each child during a pre-school fundraiser. Create a dot plot from the frequency table.

Name	Packets sold
Ivan	5
Stacy	6
Vanessa	3
Josh	8
Jamie	5
Kelly	7
Billy	10
Cassie	5
Tim	7
Kate	3

2. The dot plot shows the age of people who bought a bag of kale at a grocery store. Create a frequency table from the dot plot.



■ 3. The following data shows the number of red marbles drawn in a class lottery. Create a frequency table for the data.

0, 0, 0, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 5, 5, 5, 7, 7

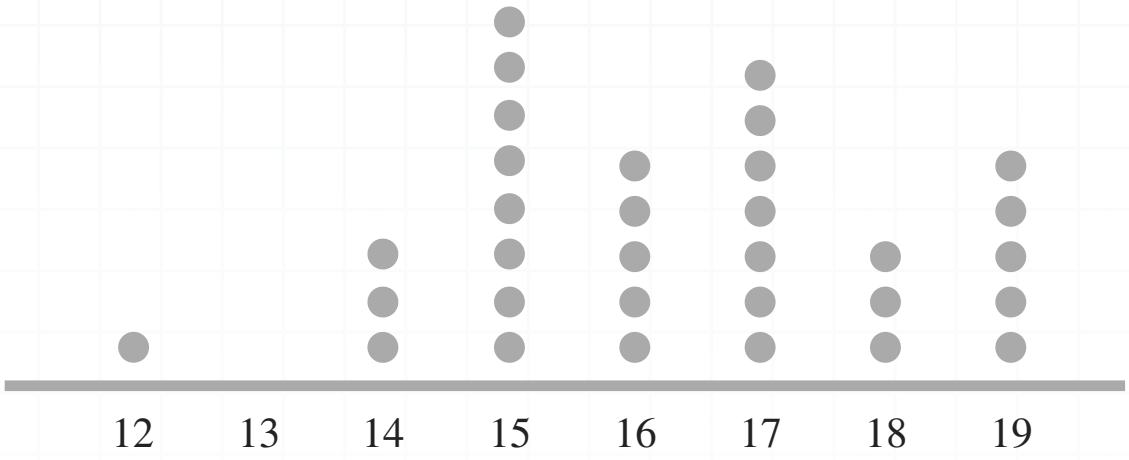
■ 4. The following data shows the favorite color of the students in Sebastian's kindergarten class. Create a frequency table for the data.

pink, pink, pink, pink, purple, purple, blue, blue, blue, blue, blue, red, red, red, yellow, orange, orange, green, green, green, black

■ 5. Kevin watches birds from his window and records what kind he sees. Create a dot plot from the data.

chickadee, redbird, redbird, redbird, chickadee, sparrow, sparrow, sparrow, sparrow, blue jay, crow, crow, redbird, chickadee, sparrow, sparrow, blue jay

6. The dot plot shows the ages of people in a lifeguard class at the local recreation center. How many people are enrolled in the class who are either 16, 17, or 18 years old?



HISTOGRAMS AND STEM-AND-LEAF PLOTS

■ 1. A doctor recorded the weight of all the babies that visited her clinic last week. How many babies weighed no more than 24 pounds?

1	5 5 7 8
2	2 4 6
3	5 6
4	
5	2 6
6	0

$1|5 = 15$

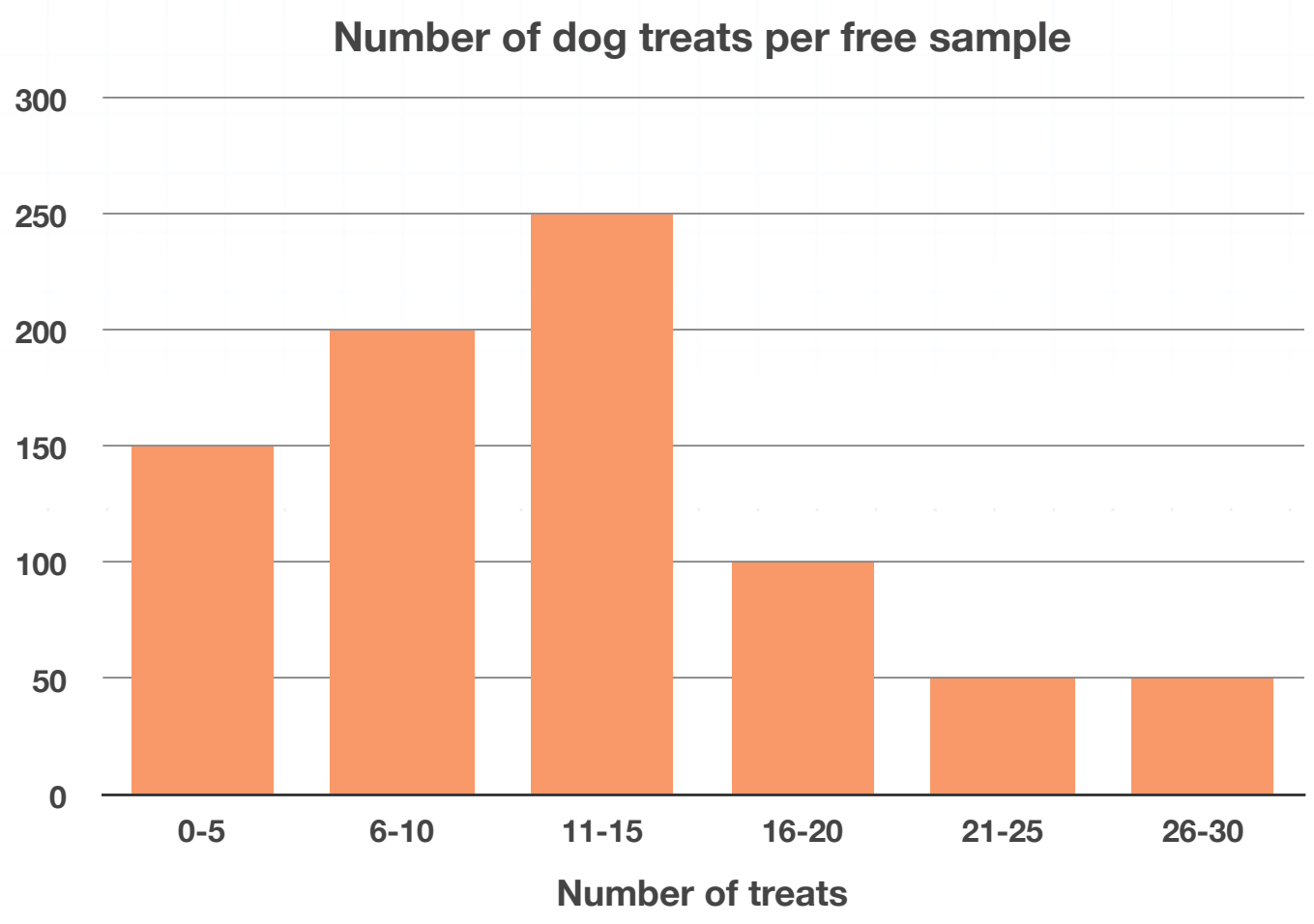
■ 2. The stem plot shows the number of clothing pieces on each rack at a clothing store. Create a histogram from the steam plot, and use buckets of size 10.

1	0 1 2 8
2	8 8 8
3	2 6 8 9
4	4 4 4
5	2 6
6	0

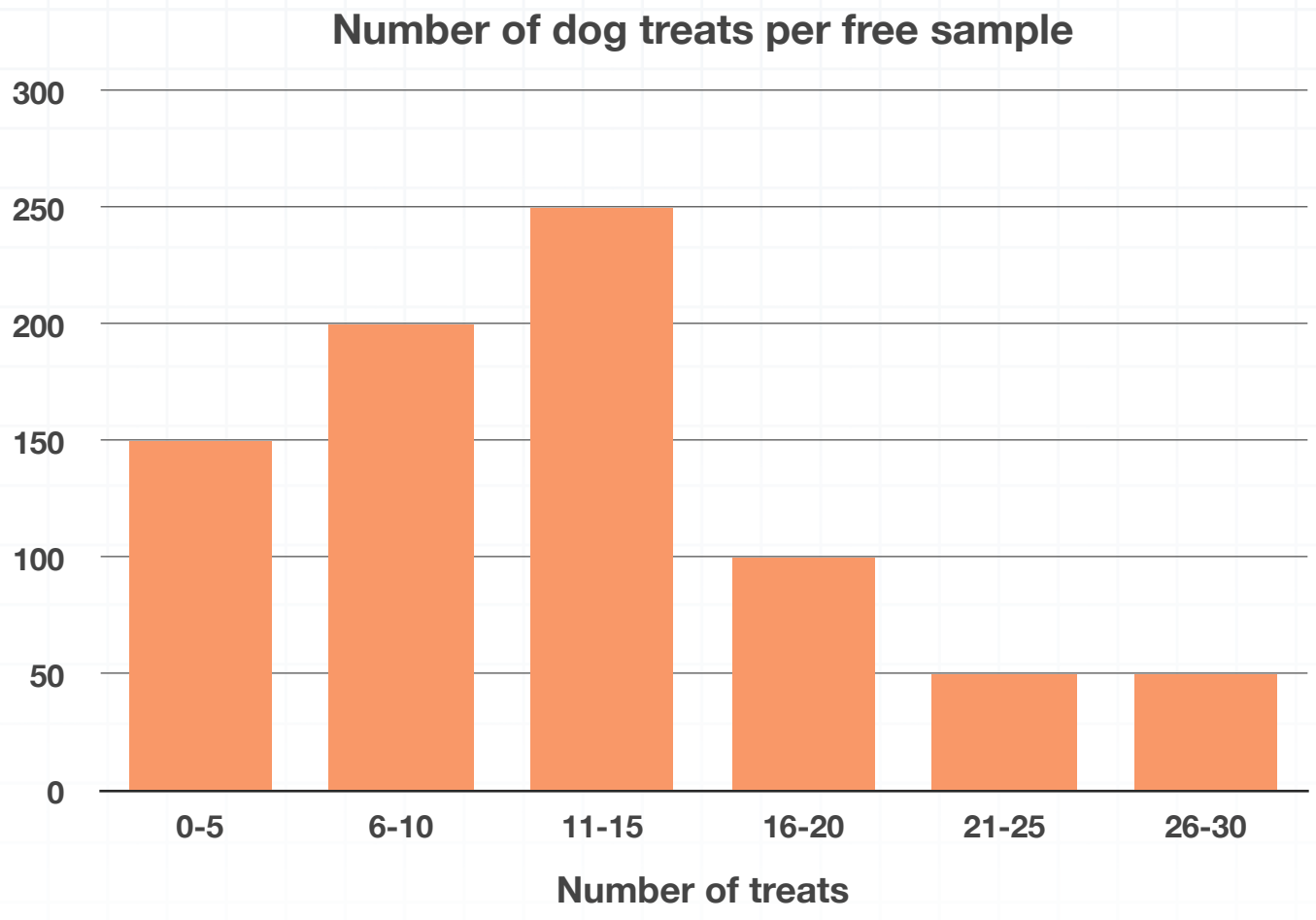
$1\,|\,0 = 10$

■ 3. Is it possible to create a stem-and-leaf plot from a histogram? Why or why not?

■ 4. A company mails out packets of dog treat samples based on a consumer's previous dog food purchases. How many times did the company mail a packet of $11 - 15$ treats?



■ 5. A company mails out packets of dog treat samples based on a consumer's previous dog food purchases. How many packets of dog treat samples did the company give out?



6. Create a stem-and-leaf chart from the list of student test scores.

60, 65, 80, 80, 81, 82, 88, 89, 90, 97, 98, 100, 100

CENTRAL TENDENCY: MEAN, MEDIAN, AND MODE

■ 1. What is the mean of the data set?

105, 250, 358, 422

■ 2. What is the median of the data set?

62, 64, 69, 70, 70, 71, 73, 74, 75, 77

■ 3. What is the mode of the data set?

1	3 7 8
2	1 4 6
3	5 5
4	
5	2 6

$1|3 = 13$

■ 4. What number could you add to the data set that would give you a median of 15?

1, 2, 8, 13, 20, 30, 31

5. A teacher lost Samantha's test after it was graded, but she knows the statistics for the rest of the class.

Class mean (including Samantha's test): $\mu = 85$

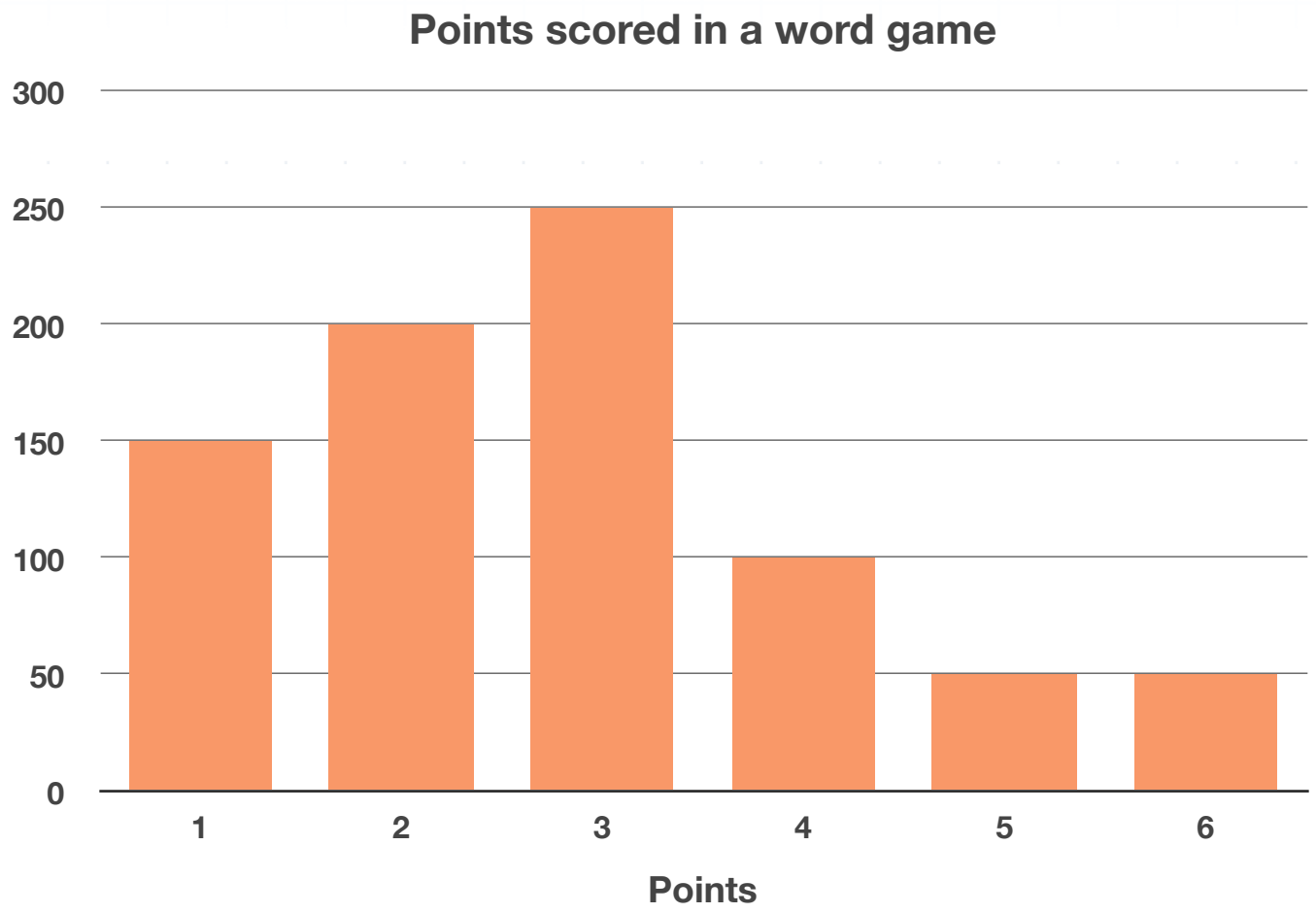
Total number of students who took the test: 20

Class test scores for everyone but Samantha were:

$$75, 75, 75, 80, 80, 80, 80, 80, 82, 82, 82, 82, 95, 95, 95, 95, 98$$

What did Samantha score on her test?

6. What is the mode of the data set?



SPREAD: RANGE AND IQR

■ 1. Sarah is visiting dairy farms as part of a research project and counting the number of red cows at each farm she visits. Here is her data:

$$0,\ 1,\ 1,\ 1,\ 2,\ 5,\ 5,\ 7,\ 7,\ 18,\ 24,\ 24$$

Calculate the IQR and range of the data set.

■ 2. A dog boarding company kept track of the number of dogs staying overnight and the frequency. What is the range of the data?

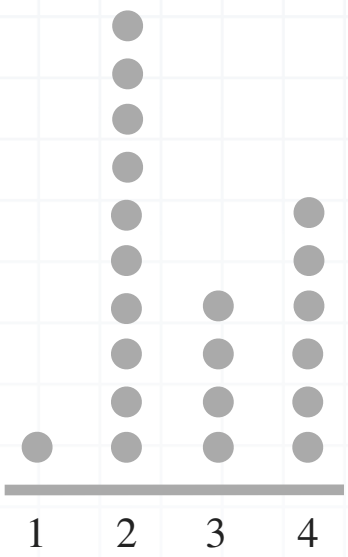
Number of dogs	Frequency
20	2
25	3
32	1
38	1
39	2
40	3
43	2

■ 3. Catherine counted the number of lizards she saw in her garden each week and recorded the data in a table. What is the interquartile range of the data?

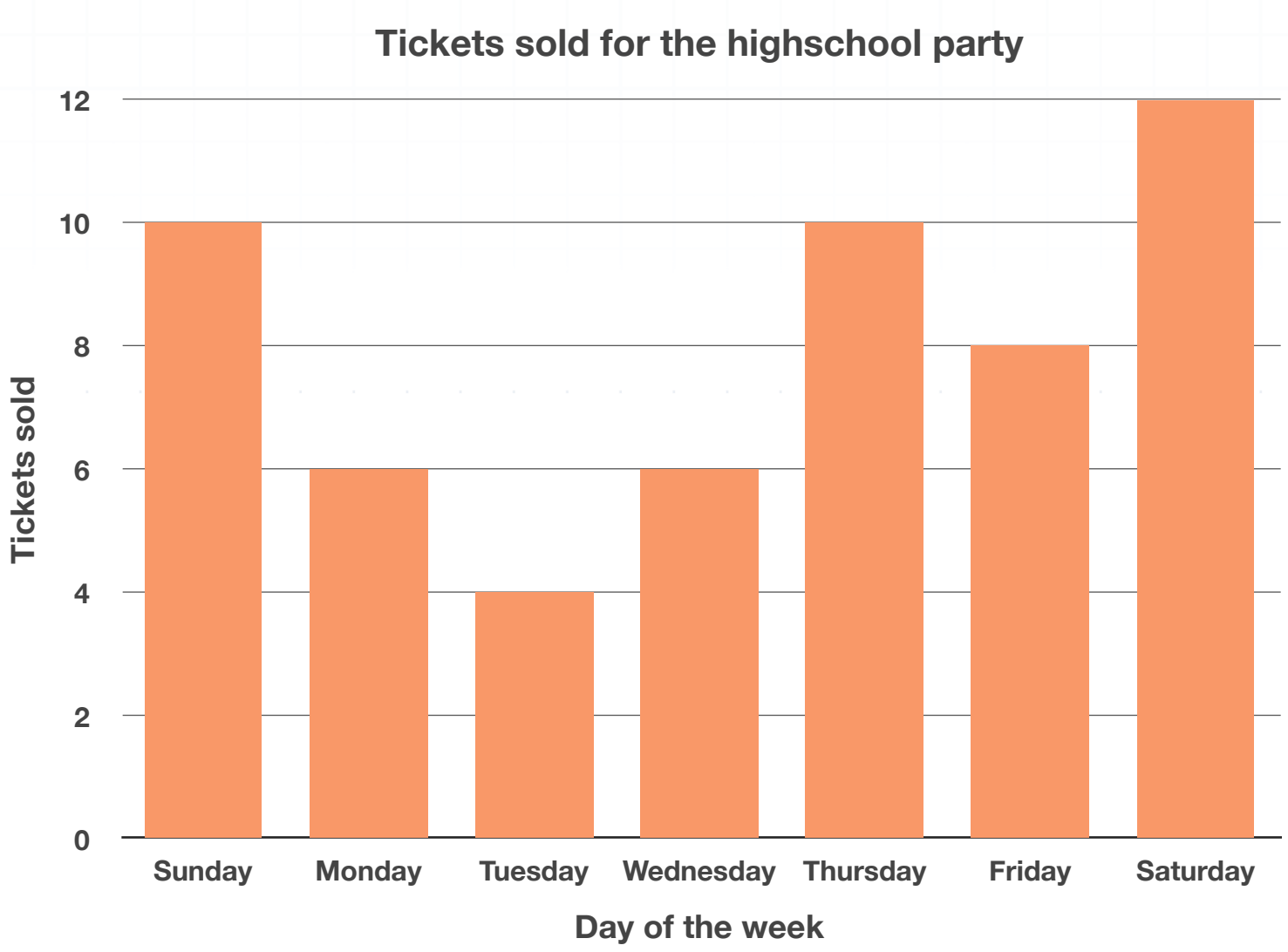
Number of lizards	Frequency
2	5
5	2
8	1
12	2
13	2
15	3
21	1

■ 4. The median of the lower-half of a data set is 98. The interquartile range is 2. If the data set has 9 numbers, what can you say about the median of the entire data set?

■ 5. The dot plot shows the number of trips to the science museum for a class of 4th graders. What is the range of the data set?



6. The bar graph shows the number of tickets sold for the high school party each day. What is the interquartile range of the data set?

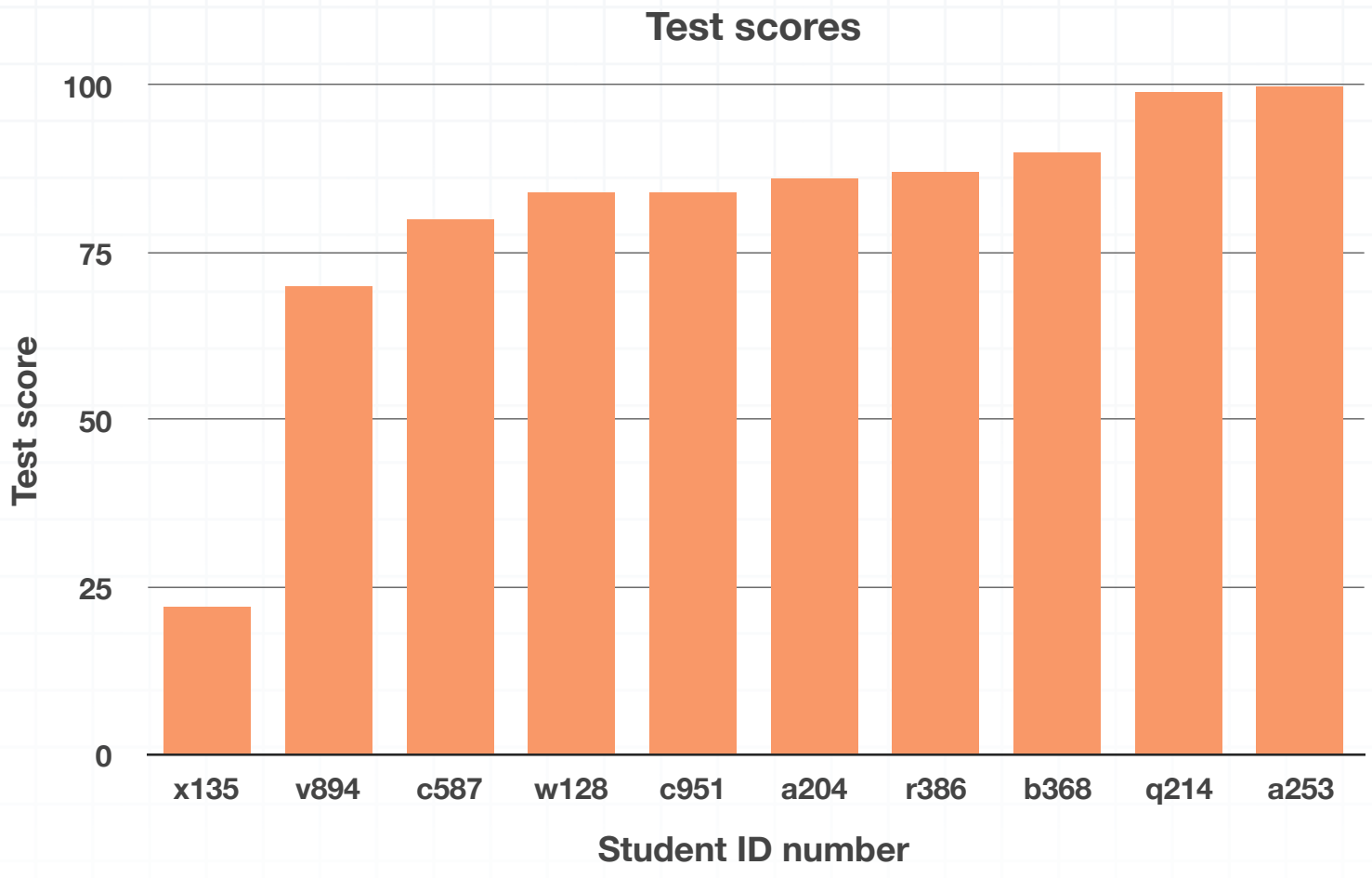


CHANGING THE DATA AND OUTLIERS

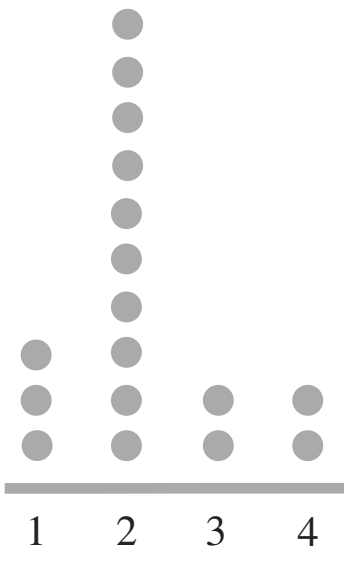
■ 1. The students in an English class ended up with a mean score on their recent exam of 65 points. The range of exam scores was 25 points. If each score is increased by $10\,\%$, what are the new mean and range?

■ 2. Spencer asked students at his high school what percentage of the school budget they thought was spent on extracurricular activities. The mean response was $8\,\%$ and the median response was $5\,\%$. There was one outlier in the responses. What do the mean and median tell you about the outlier?

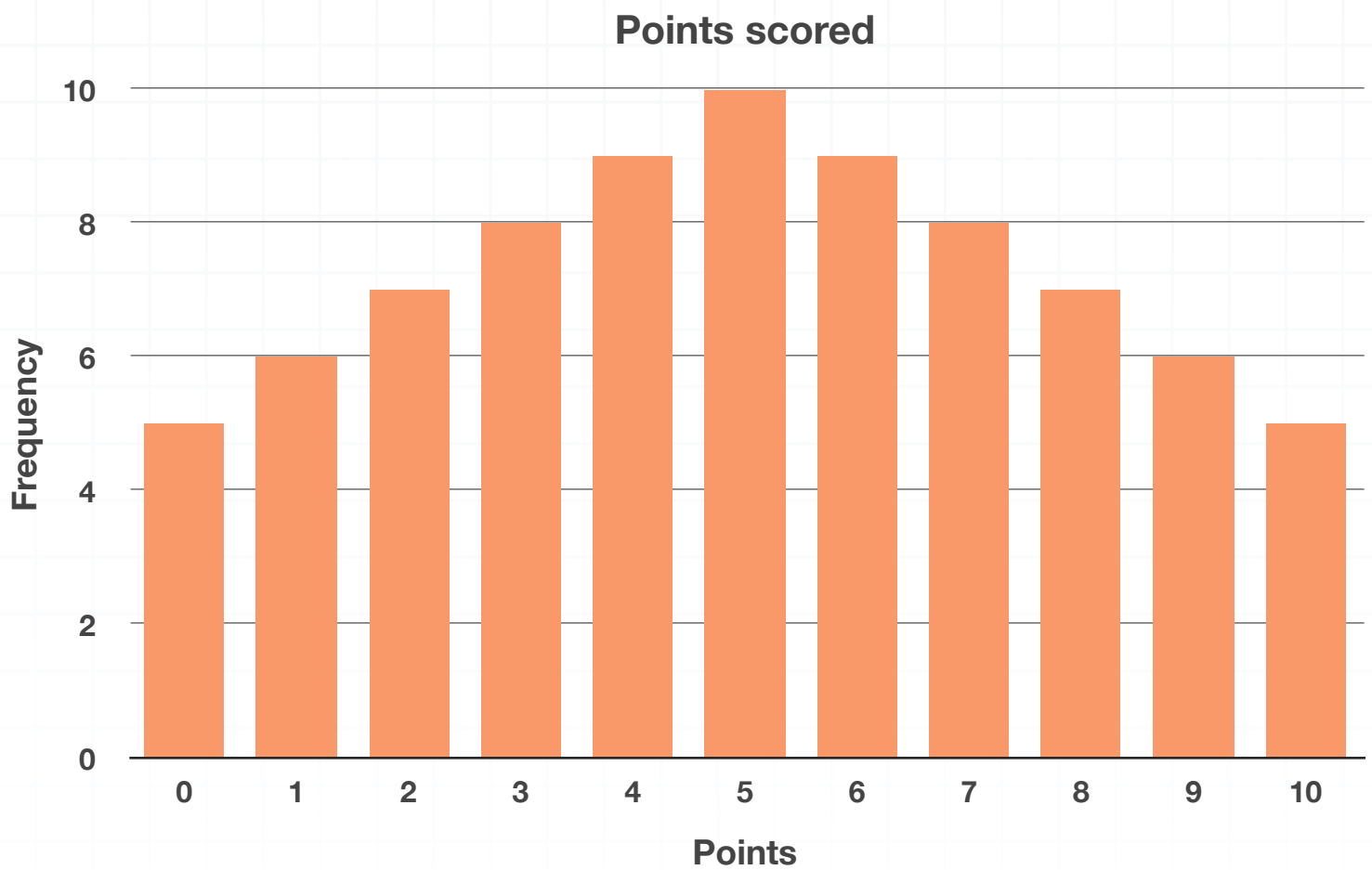
■ 3. How does the mean compare to the median in the data from the bar graph?



4. The dot plot shows the number of trips to the science museum for a class of 4th graders. How does the mean compare to the median in the data set below, and what does it tell you about the potential outliers in the data set?



5. What does the shape of this histogram tell you about the mean and median of the data?



6. An experiment is done in degrees Celsius. The original data had the following:

Mean: 102° Celsius

Median: 101° Celsius

Mode: 99° Celsius

Range: 7° Celsius

IQR: 4° Celsius

The formula to convert to degrees Fahrenheit is $F = (9/5)C + 32$. After the conversion to Fahrenheit, what are the new reported measures of the data set?

BOX-AND-WHISKER PLOTS

1. What is the range and interquartile range of the data set?

Median: 617,594

Minimum: 216,290

Maximum: 845,300

First quartile: 324,528

Third quartile: 790,390

2. These are average lifespans in years of various mammals:

35, 10, 40, 40, 20, 10, 15, 14, 18, 35

Find the five-number summary for the data.

3. Create a box plot based on the following information about a data set.

Mode: 300

Minimum: 100

First Quartile: 300

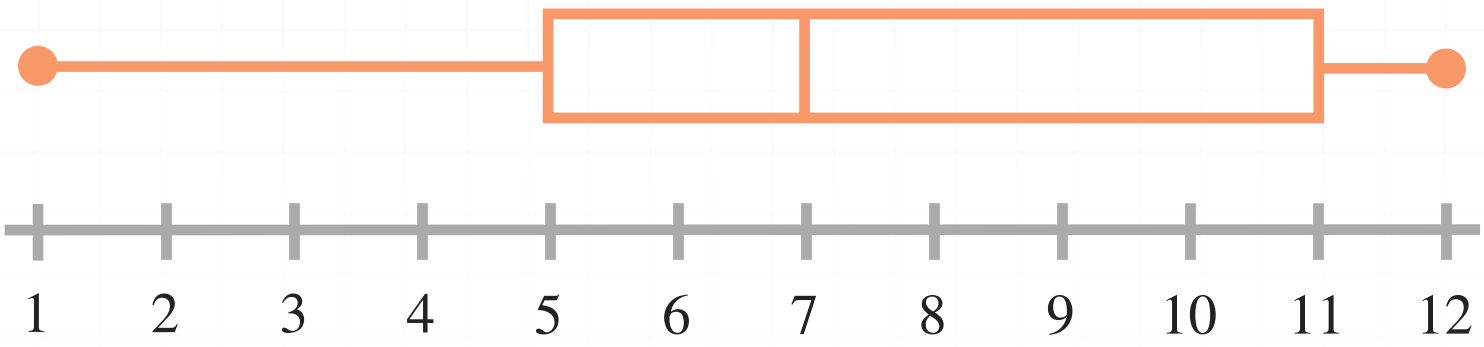
Median: 2,000

Mean: 1,887.5

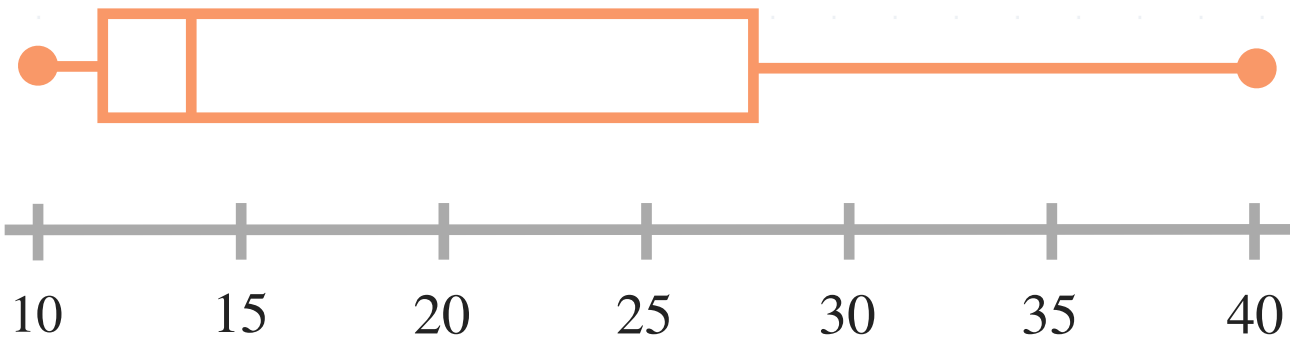
Third Quartile: 3,050

Maximum: 4,800

■ 4. How does the amount of data between 1 and 5 compare to the amount of data between 11 and 12?



■ 5. In which quartile of the data is the number 23 located?



■ 6. Create the box-and-whisker plot for the book ratings given in the stem and leaf plot.

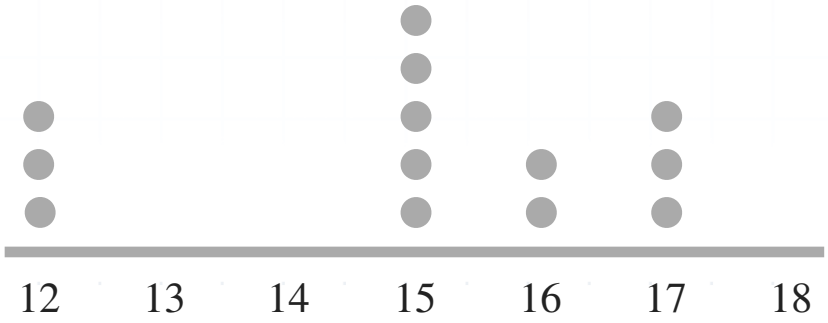
Stem	Leaf
1	3 7 8
2	1 4 6
3	5 5
4	
5	2 6

Key: $1 \mid 3 = 13$

MEAN, VARIANCE, AND STANDARD DEVIATION

■ 1. Mrs. Bayer's students take a test on Friday. She grades their tests over the weekend and notes that the average test score is 68 points with a population standard deviation of 5 points. She decided to add 10 points to all of the tests. What are the new mean and population standard deviation?

■ 2. What is the sample variance of the data set to the nearest hundredth? Use the sample mean rounded to the nearest hundredth for your calculation.



■ 3. Sometimes it can be helpful to calculate the standard deviation by using a table. Use the data to fill in the rest of the table and then use the table to calculate the sample standard deviation.

Data value	Data value - Mean	Squared difference
97		
110		
112		
121		
110		
98		
Total		

■ 4. The sum of the squared differences from the population mean for a data set is 212. If the data set has 25 items, what is the population standard deviation?

■ 5. For the data set $40, 44, 47, 55, 60, 60, 65, 80$, find

$$\sum_{i=1}^{n} (x_i - \bar{x})$$

for the data set. What does this say about why we square the $(x_i - \bar{x})$ in the variance and standard deviation formulas?

■ 6. Give an example of a situation where $\$5$ could represent a large standard deviation and another where $\$5$ could represent a small standard deviation.

FREQUENCY HISTOGRAMS AND POLYGONS, AND DENSITY CURVES

■ 1. A dog walking company keeps track of how many times each dog receives a walk. $40\,\%$ of all the dogs walked by the company received between 25 and 40 walks, and no dogs received more than 40 walks. How many dogs received between 0 and 25 walks, if the company walks 400 dogs?

■ 2. The number of crayons in each student's pencil box is

$$4,\ 1,\ 5,\ 5,\ 9,\ 11,\ 15,\ 13,\ 15,\ 14,\ 16,\ 17,\ 20,\ 16,\ 16,\ 17$$

Complete the frequency and relative frequency tables for the data and use it to create a relative frequency histogram.

Crayons	Frequency	Relative Frequency
1-5		
6-10		
11-15		
16-20		
Totals:		**100%**

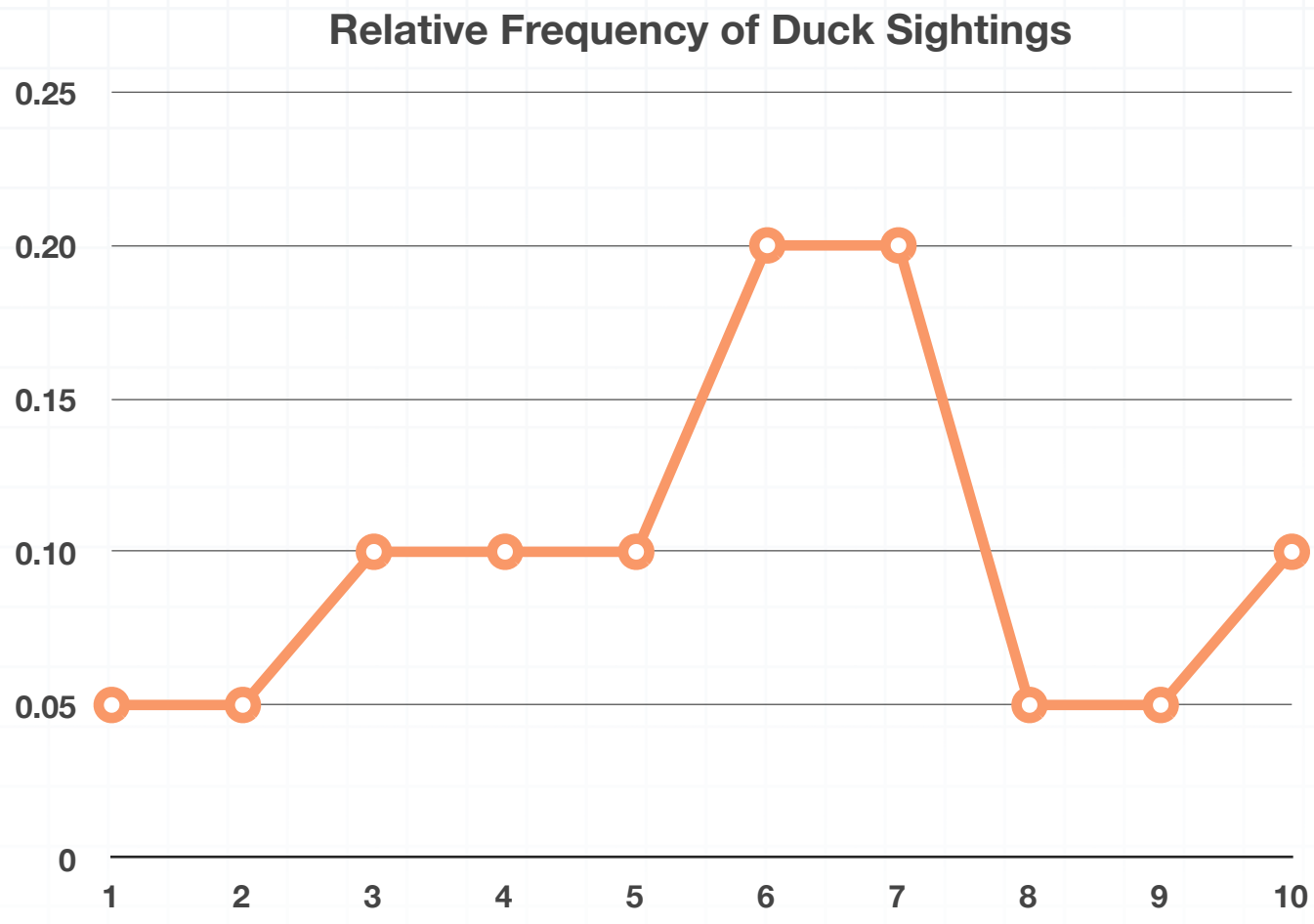
■ 3. The table shows the scores on the last history exam in Mr. Ru's class.

40	32	40	83
95	33	87	59
32	81	46	78
91	61	55	88
40	61	82	99
72	47	83	91
101	77	65	87

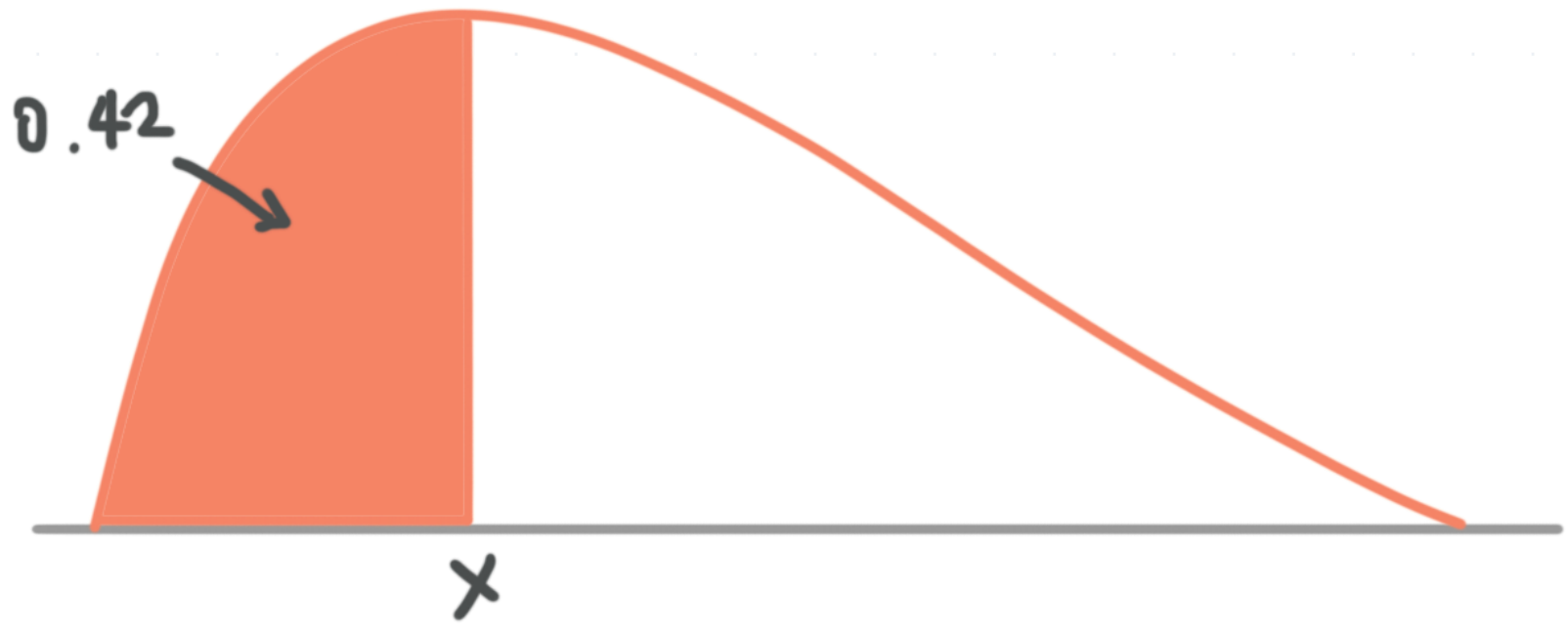
Complete the relative frequency table and create a frequency polygon for the data.

Score	Frequency	Relative Frequency
30-39		
40-49		
50-59		
60-69		
70-79		
80-89		
90-99		
100-109		
Totals:		

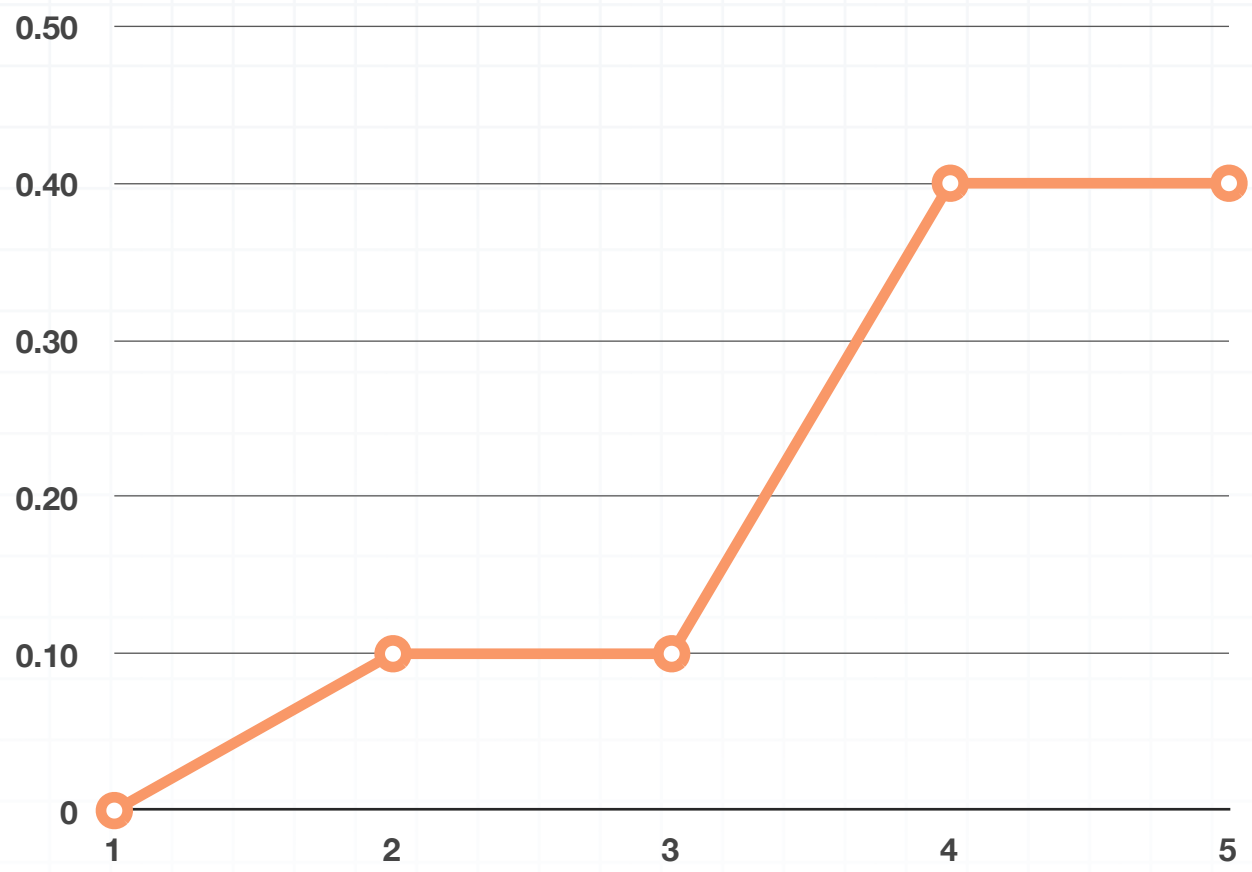
■ 4. Becky kept track of the number of ducks she saw at her neighborhood pond at $6:30$ a.m. every morning for 365 days. On how many days did Becky see more than 5 ducks?



5. What percentage of the population is greater than x for the density curve?

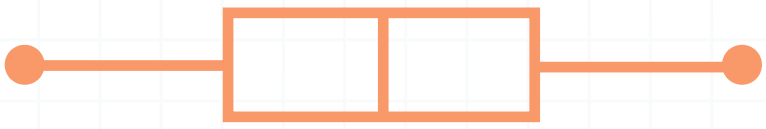


6. What percentage of the area in the density curve is between 3 and 5?

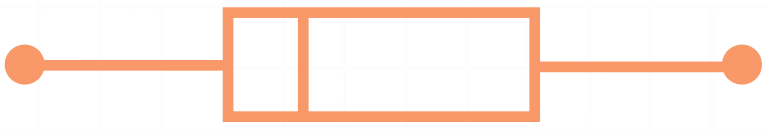


SYMMETRIC AND SKEWED DISTRIBUTIONS AND OUTLIERS

■ 1. Which type of distribution is modeled in the box plot (symmetric, negatively skewed, or positively skewed)?



■ 2. Which type of distribution is modeled in the box plot (symmetric, negatively skewed, or positively skewed)?



■ 3. The ages (in months) that babies spoke for the first time are

6, 8, 9, 10, 10, 11, 11, 12, 12, 13, 15, 15, 18, 19, 20, 20, 21

Are there outliers in the data set? If so, state what they are. What is the best measure of central tendency for the data? What is the best measure of spread?

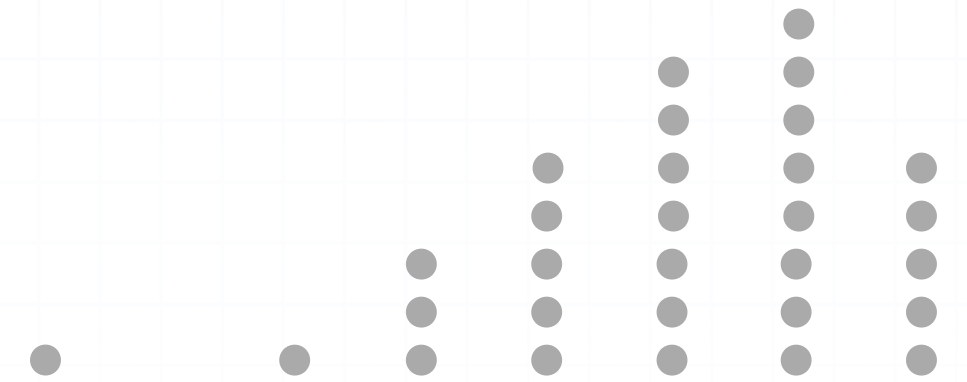
■ 4. The number of text messages sent each day by Lucy's mom is

0, 18, 19, 20, 20, 20, 21, 23, 23, 23, 24, 24,

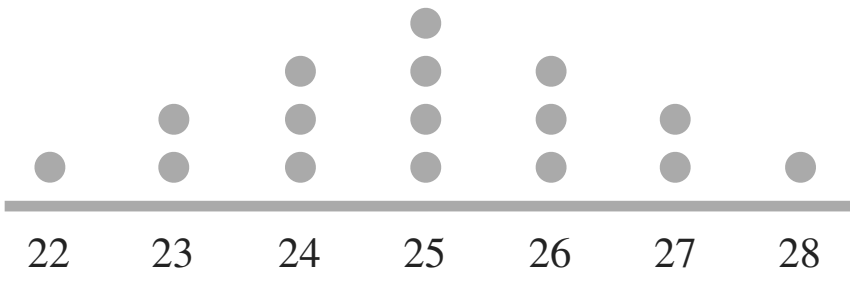
24, 24, 24, 25, 25, 25, 25, 25, 25, 30, 30, 31

Are there outliers in the data set? If so, state what they are. What is the best measure of central tendency for the data? What is the best measure of spread?

■ 5. Describe the shape, center, and spread of the data. State if there are outliers and what they are if they exist.



■ 6. Describe the shape, center and spread of the data. State if there are outliers and what they are if they exist.



NORMAL DISTRIBUTIONS AND Z-SCORES

■ 1. A population has a mean of 62 and a standard deviation of 5. What is the z-score for a value of 50?

■ 2. What percentile is a z-score of -1.68?

■ 3. A population has a mean of 170 centimeters and a standard deviation of 8 centimeters. What percentage of the population has a value less than 154 centimeters?

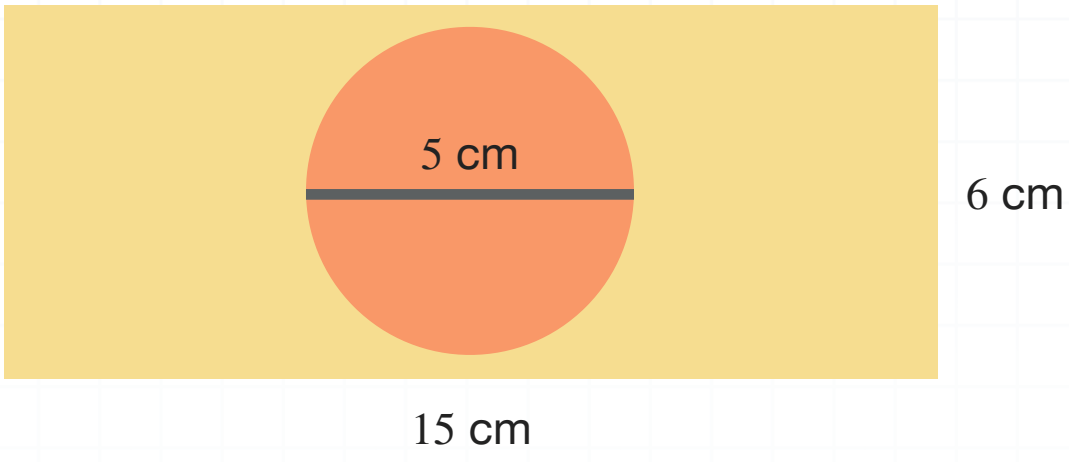
■ 4. The mean diameter of a North American Native Pine tree is $18''$ with a standard deviation of $4''$. What is the approximate diameter for a tree in the 21st percentile for this distribution? Assume an approximately normal distribution.

■ 5. The mean diameter of a North American Native Pine tree is $18''$ with a standard deviation of $4''$. According to the empirical rule, $68\,\%$ of North American Native Pines have a diameter between which two values? Assume an approximately normal distribution.

6. IQ scores are normally distributed with a mean of 100 and a standard deviation of 16. What percentage of the population has an IQ score between and 120 and 140?

SIMPLE PROBABILITY

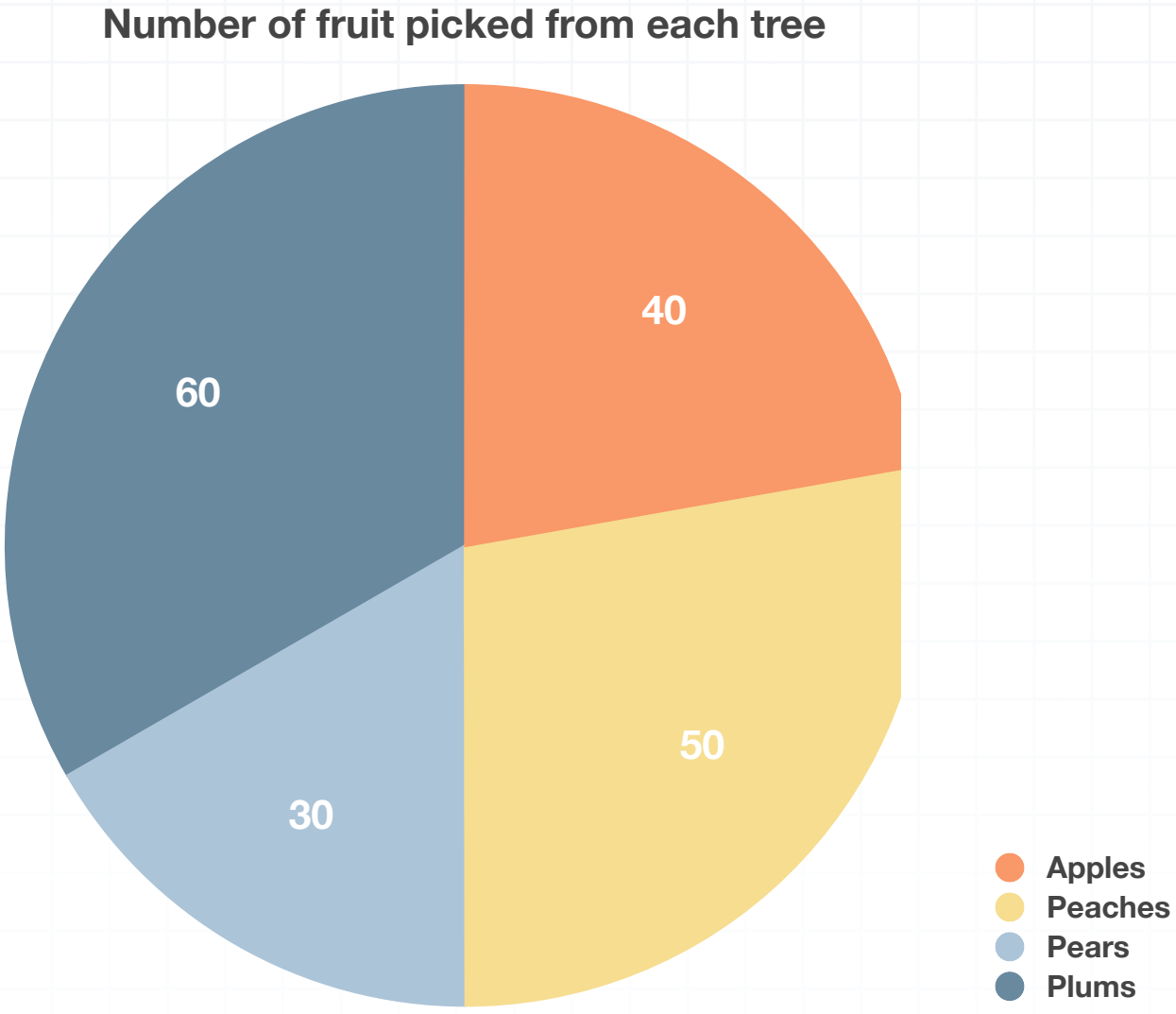
1. A child drops a marble onto a board. Suppose that it is equally likely for it to fall anywhere on the board. What is the probability, to the nearest percent, that it lands on the red circle?



2. A 12-sided number cube is rolled 60 times. Use the table to calculate $P(\text{rolling an } 11)$. Is this theoretical or experimental probability? Why?

Number rolled	1	2	3	4	5	6	7	8	9	10	11	12
Frequency	5	8	2	0	10	1	6	5	2	8	12	1

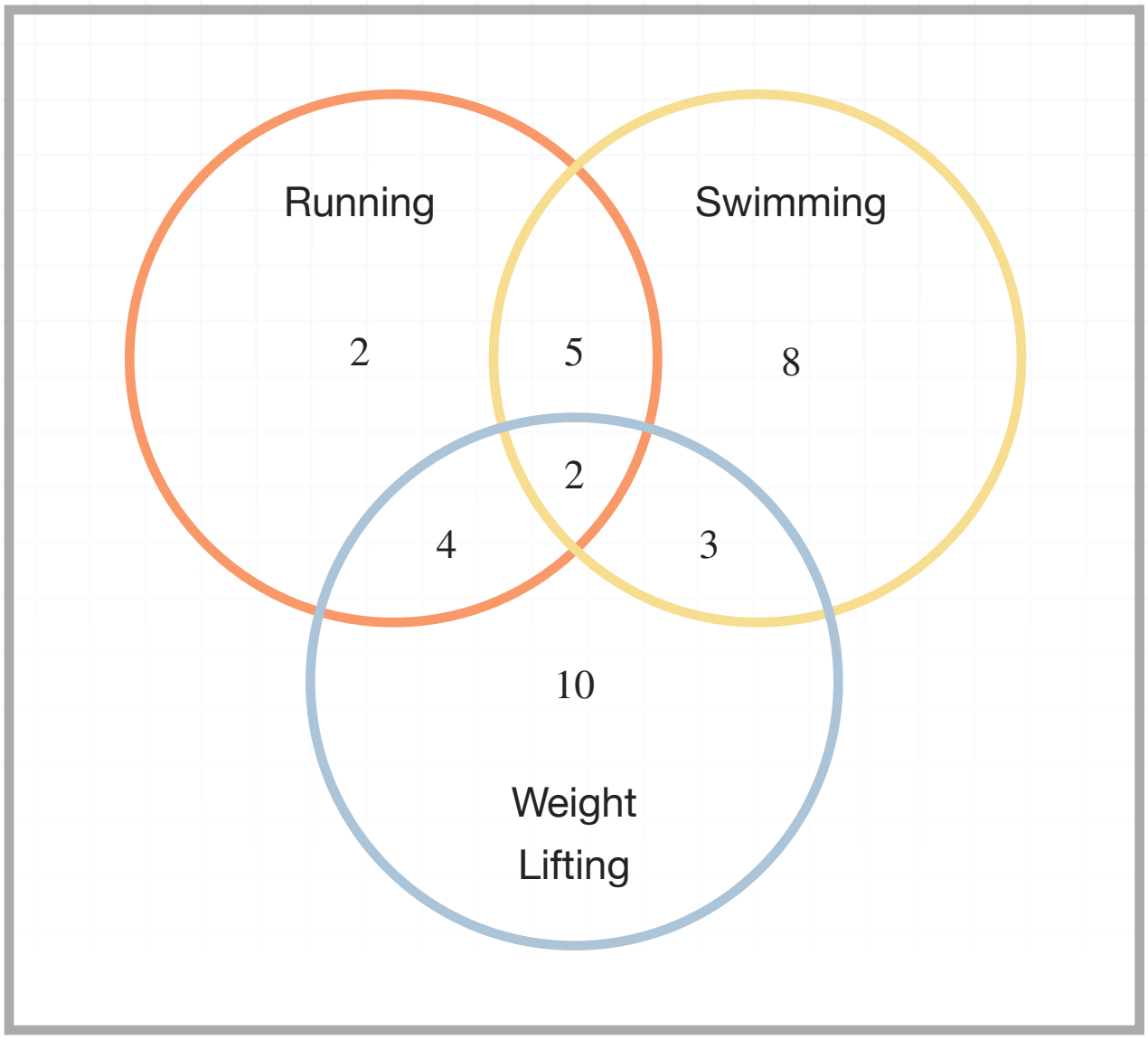
3. Monica's class went on a trip to an orchard. At the end of the trip they put all of the fruit they picked into one big basket. The chance of picking any fruit from the basket is equally likely. Monica's teacher picks out a fruit for her to eat at random. What is the probability that it's a plum (Monica's favorite)? Is this an experimental or theoretical probability? Why?



4. Jamal surveyed the people at his local park about their favorite hobby and recorded his results in a table. Based on the survey, what's the probability that someone who visits the park will choose Art as their favorite hobby? Is this a theoretical or experimental probability? Why?

Hobby	Count
Reading	14
Sports	28
Art	15
Total	**57**

■ 5. What is the probability that someone's favorite exercise was weight lifting only?

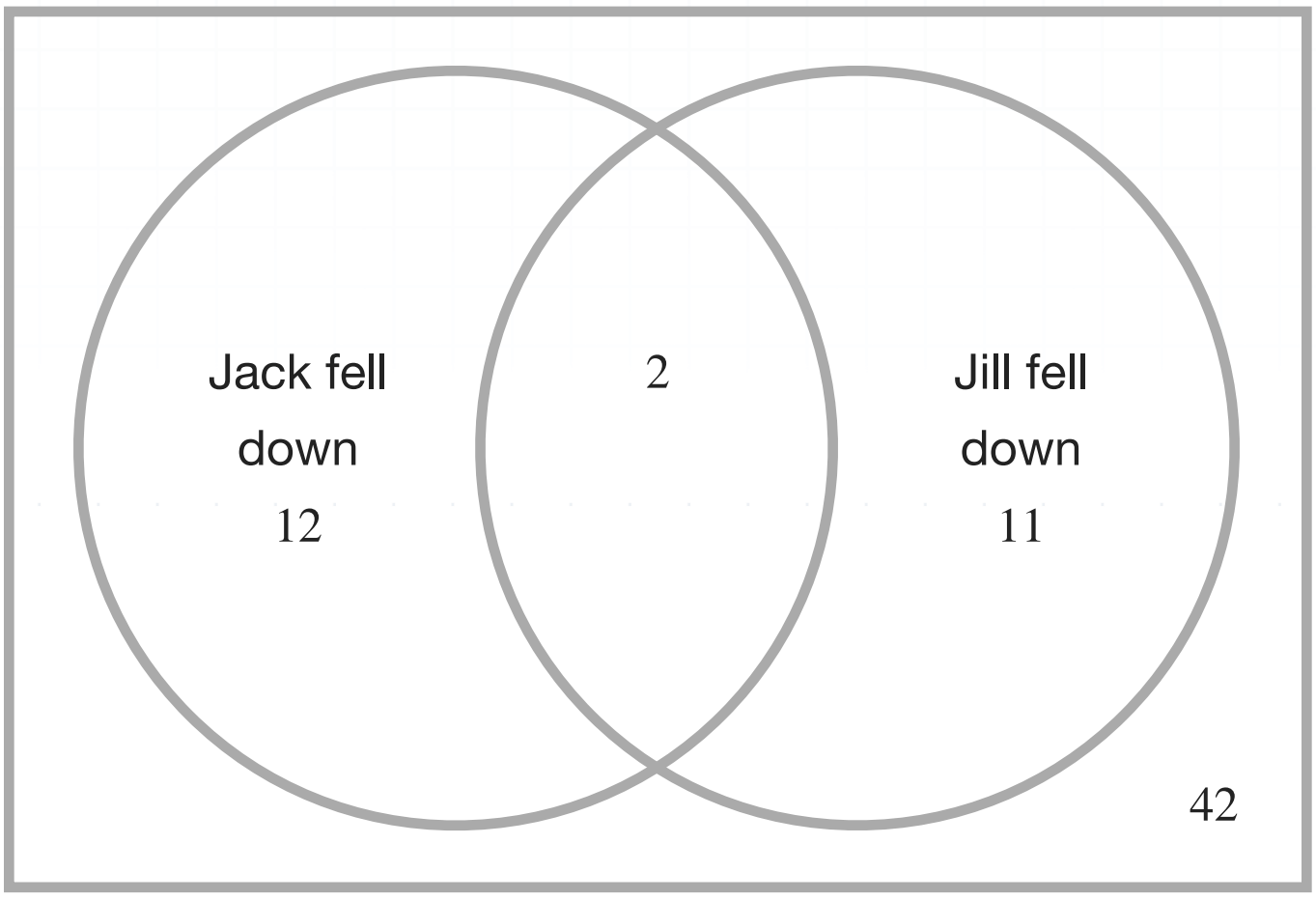


■ 6. What is the sample space for rolling two six-sided dice (the list of all possible outcomes)? What's the probability that the sum of the two dice is an odd number? Is this a theoretical or experimental probability? Why?

THE ADDITION RULE, AND UNION VS. INTERSECTION

■ 1. Given the probabilities $P(A) = 0.3$, $P(B) = 0.6$ and $P(A \cap B) = 0.05$, what is $P(A \cup B)$? Are A and B mutually exclusive events? Why or why not?

■ 2. The Venn diagram shows the number of times Jack and Jill fell when going up the hill. What is the probability that Jack fell down and Jill fell down? What is the probability that Jack fell down or Jill fell down?

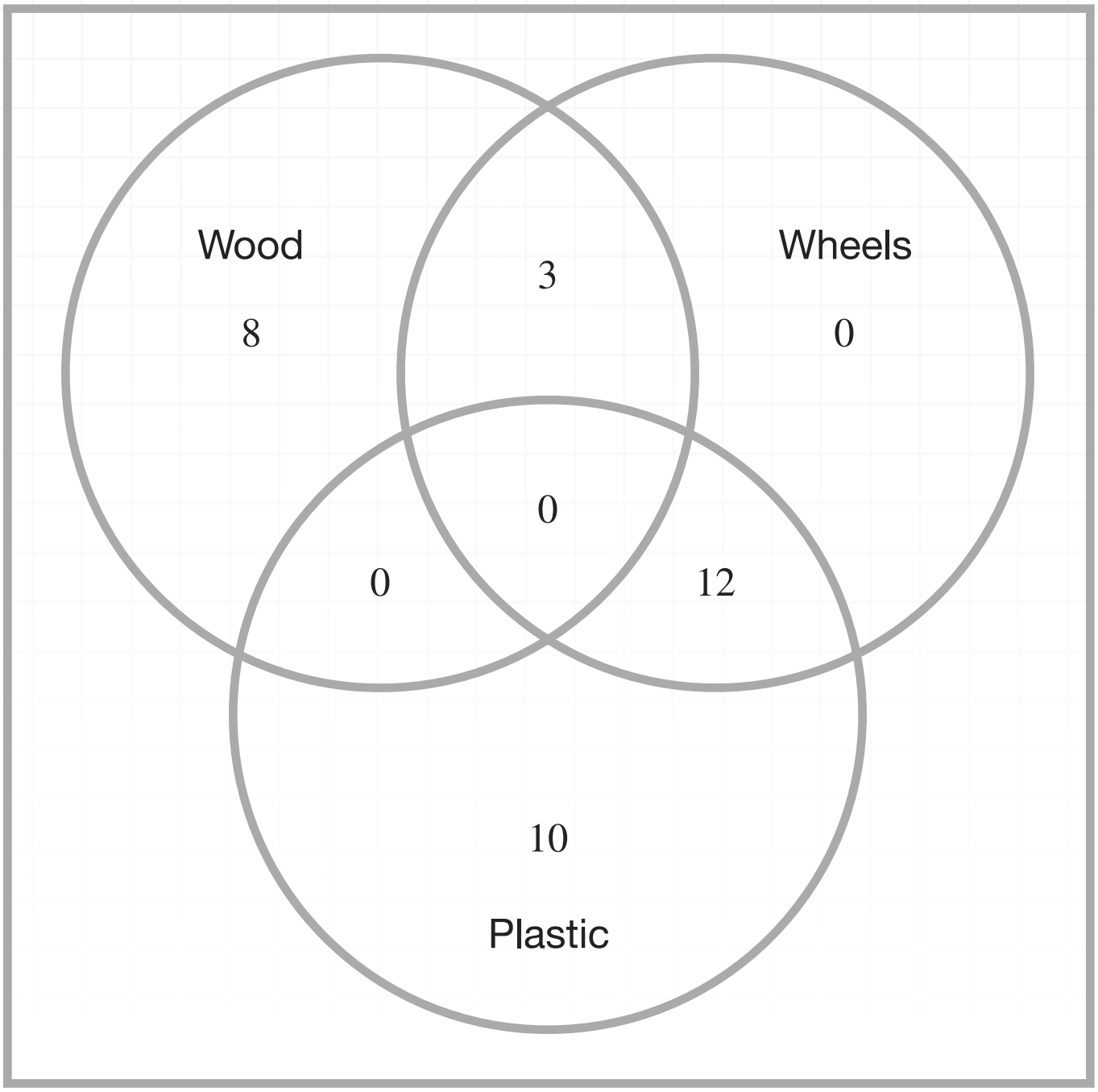


■ 3. When people buy a fish at a pet store the cashier can check off the color of the fish as mostly red, mostly orange or mostly yellow. Currently the probability of buying a red fish is 0.31, the probability of buying an

orange fish is 0.23, and the probability of buying a mostly yellow fish is 0.13 (there are colors of fish other than red, orange, and yellow).

Are the events buying a mostly red fish and buying a mostly orange fish mutually exclusive? Find the probability that the purchase of a randomly selected fish is either mostly red or mostly orange.

■ 4. The Venn diagram shows Mason's toy car collection. Are the events "plastic" and "wood" mutually exclusive? What is the probability that a vehicle is made from plastic or wood? Are the events "wood" and "wheels" mutually exclusive? What is the probability that a vehicle is made from wood and has wheels?



■ 5. Every student at a certain high school needs to choose exactly one fine arts elective. The frequency table shows the enrollment of electives for all students. Are the events “junior” and “architecture” mutually exclusive? What is the probability that a student taking architecture is a junior? What is the probability that a student is a junior or is taking architecture?

		Extracurricular activities			
		Art	**Architecture**	**Music**	**Total**
Grade	**Freshmen**	40	25	55	**120**
	Sophomore	52	12	71	**135**
	Junior	56	45	54	**155**
	Senior	30	60	20	**110**
	Total	**178**	**142**	**200**	**230**

6. James tosses a coin and rolls a six-sided die. What is the sample space for this situation? What is the probability the coin lands on heads and the die lands on a 2 or a 3?

INDEPENDENT AND DEPENDENT EVENTS AND CONDITIONAL PROBABILITY

■ 1. What is the probability of getting four heads in a row when you flip a fair coin four times?

■ 2. An old dog finds and eats $60\,\%$ of food that's dropped on the floor. A toddler wanders through the house and drops 10 pieces of cereal. What's the probability the dog finds and eats all 10 pieces?

■ 3. Amelia is choosing some pretty stones from the gift shop at the museum. The gift shop has a grab bag that contains 5 amethyst stones, 6 fluorite stones, 2 pink opals, and 7 yellow calcite stones. Amelia looks into the bag and takes out two stones, one at a time, at random. What is the probability that she gets an amethyst first and then a pink opal?

■ 4. Emily counted the shape and type of blocks that her little sister owns and organized the information into a frequency table.

		Block Shape		
		Cube	**Rectangular Prism**	**Total**
Block Color	**Red**	5	9	**14**
	Blue	4	10	**14**
	Total	**9**	**19**	**28**

Are events A and B dependent or independent events? Use the formula to explain your answer.

Event A is that the block is a cube.

Event B is that block is red.

Let $P(A)$ be the probability that a block drawn at random is a cube.

Let $P(B)$ be the probability that a block drawn at random is red.

■ 5. A bag has 4 cinnamon candies, 6 peppermint candies, and 12 cherry candies. Sasha draws 3 candies at random from the bag one at a time without replacement. Does the situation describe dependent or independent events? What is the probability of drawing a cinnamon first, then a cherry, and then a peppermint?

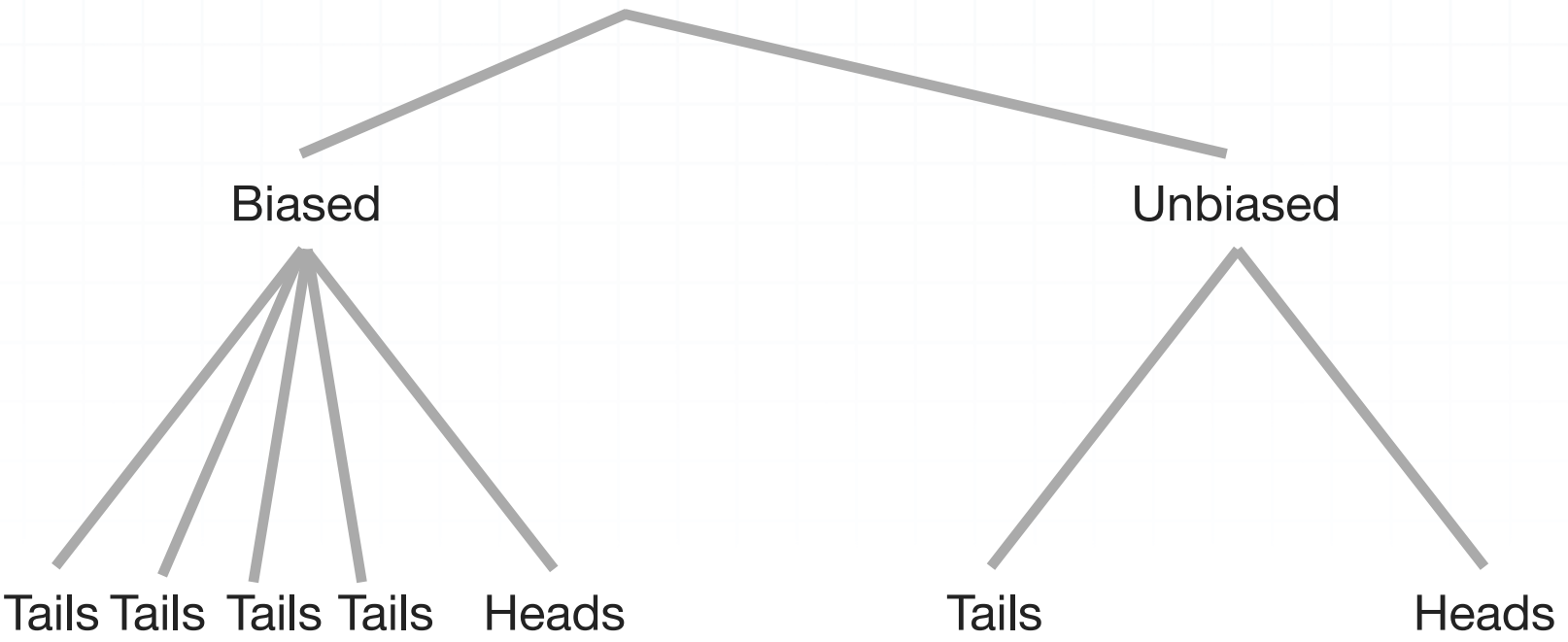
■ 6. Nyla has 12 stuffed animals, 7 of which are elephants (4 of the elephants play music and light up) and 5 of which are bears (2 of the bears play music and light up). Her mother randomly selects an animal to bring

with them on vacation. Let A be the event that she selects an elephant and B be the event that she selects an animal that plays music and lights up.

Find $P(A)$, $P(B)$, $P(A\,|\,B)$, and $P(B\,|\,A)$. State if events A and B are dependent or independent events, then find $P(A \text{ and } B)$.

BAYES' THEOREM

■ 1. You have two coins. One is fair and the other one is weighted to land on tails $4/5$ of the time. Without knowing which coin you're choosing, you pick one at random, toss the coin and get tails. What is the probability you flipped the biased coin? Complete the tree diagram to answer the question.



■ 2. You have two dice. One is fair and the other is biased. The biased die is weighted to land on 6 every 1 out of 36 rolls. There's an equal probability for all of the other five faces on the biased die. Without knowing which one you're choosing, you pick one of the dice, roll it, and get a 6.

Calculate the following and use them to answer the question: What is the probability that you rolled the fair die?

$$P(6 \mid \text{fair})$$

$P(\text{fair})$

$P(6)$

3. Charlie knows that, at his school,

$P(\text{senior}) = 0.40$

$P(\text{playing soccer}) = 0.15$

$P(\text{soccer and senior}) = 0.05$

Solve for the probability $P(\text{senior} \mid \text{soccer})$, then state whether or not Bayes' Theorem can be used to solve the problem.

4. You have two coins. One is fair and the other is weighted to land on tails $3/4$ of the time. Without knowing which coin you're choosing, you pick one at random, toss the coin, and get tails. What's the probability you flipped the biased coin?

5. A company is giving a drug test to all of its employees. The test is $90\,\%$ accurate, given that a person is using drugs, and $85\,\%$ accurate, given that the person is not using drugs. It's also known that $10\,\%$ of the general population of employees uses drugs. What is the probability that an employee tests positive due to an inaccurate result (a false positive)?

Let P represent a positive test for an individual.

Let N represent a negative test for an individual.

Let D represent the event that an employee is a drug user.

6. Two factories A and B produce heaters for car seats. A customer received a defective car seat heater and the manager at factory B would like to know if it came from her factory. Use the table below to determine the probability that the heater came from factory B.

Factory	% of production	Probability of defective heaters
A	0.55	0.020 P(DIA)
B	0.45	0.014 P(DIB)

DISCRETE PROBABILITY

■ 1. Let X be a discrete random variable with the following probability distribution. Find $P(X \geq 3)$.

X	1	2	3	4	5
P(X)	0.35	0.25	0.20	0.15	?

■ 2. Let B be a discrete random variable with the following probability distribution. Find μ_B and σ_B.

B	0	5	10	15
P(B)	1/5	1/5	2/5	1/5

■ 3. The table shows the distribution of size of households in the U.S. for 2016. Suppose we select a household of size at least 2 at random. What is the probability that this household has a size of at least 4?

Size of household	1	2	3	4	5	6	7+
P(size)	0.281	0.340	?	0.129	0.060	0.023	0.013

■ 4. A standard deck of cards is shuffled, and two cards are selected without replacement. Let R be the number of red cards selected. Construct a probability distribution for R.

■ 5. A local restaurant features a wheel you can spin before you pay your bill. The wheel is split into 8 equal size pieces. One of the sections gives customers a $\$10$ discount on their bill, two sections give a $\$5$ discount, three sections give a $\$2$ discount, and the rest of the sections give no discount. Find the expected value for the discount given by the wheel.

■ 6. John stops at the local gas station and decides to buy lottery tickets. Each ticket has a $20\,\%$ chance of being a winner. He will buy a lottery ticket and check to see if it's a winner. If it's a winner, he'll collect his money and be done. If it's not a winner, he'll buy another. He'll repeat this until he gets a winning ticket. But if he hasn't won by his fifth ticket, he won't buy any more tickets. Let L be the number of lottery tickets John will buy, then find $E(L)$.

TRANSFORMING RANDOM VARIABLES

1. We use the formula

$$^\circ F = \frac{9}{5}{}^\circ C + 32$$

to convert from Celsius to Fahrenheit. August is the hottest month in Hawaii with a mean temperature of $27^\circ C$. What is the mean temperature in Hawaii in $^\circ F$.

2. Let Z be a random variable with $\sigma_Z^2 = 49$. Let $W = (1/2)Z - 10$. Find σ_W.

3. The students in each 8th period classroom were asked to donate money for a school fundraiser. The class who raises the most money is awarded a pizza party. The school secretary records the amount raised by each class and makes a five-number summary for the data.

Min	Q1	Median	Q3	Max
4.50	15.25	22.00	38.75	95.50

Suppose the school has 45 8th period classrooms with 20 students per classroom. What was the median amount donated per student? With what IQR?

■ 4. The number of items sold at a concession stand is normally distributed with $\mu = 323$ and $\sigma = 30$. The average price per item sold is $\$1.25$. Different student clubs volunteer to work the concession stand throughout the year and get to keep half of their sales to go towards their club's activities. What is the probability that a club will get to keep more than $\$220$ in sales?

■ 5. The average length of a full-term new born baby is 20 inches with variance 0.81 inches. What are the mean and standard deviation of the length of a full-term new born, expressed in centimeters? Use 1 in $= 2.54$ cm.

■ 6. The weights of full-term new born babies are normally distributed with $\mu = 120$ ounces and $\sigma = 20$ ounces. Describe the shape, center, and spread for the weights of full-term new born babies as measured in pounds. Use 1 pound $= 16$ ounces.

COMBINATIONS OF RANDOM VARIABLES

■ 1. X and Y are independent random variables with $E(X) = 48$, $E(Y) = 54$, $SD(X) = 3$ and $SD(Y) = 5$. Find $E(X - Y)$ and $SD(X - Y)$.

■ 2. A and B are independent random variables with $E(A) = 6.5$, $E(B) = 4.4$, $SD(A) = 1.6$, and $SD(B) = 2.1$. Find $E(4A + 2B)$ and $SD(4A + 2B)$.

■ 3. The time it takes students to complete multiple choice questions on an AP Statistics Exam has a mean of 55 seconds with a standard deviation of 12 seconds. If the exam consists of 40 multiple choice questions, find the mean total time to finish the exam. Then find the standard deviation in the total time. What assumption must be made?

■ 4. Let M represent the height of a male over 21 years of age and let W represent the height of a female over 21 years of age. Let D represent the difference between their heights ($D = M - W$). Let $E(M) = 70$ inches, $\sigma_M = 2.8$ inches, $E(W) = 64.5$ inches and $\sigma_W = 2.4$ inches.

What is the mean and standard deviation of the difference between the two heights?

■ 5. The Ironman is a challenge in which a competitor swims 2.4 miles, then bikes 112 miles, and finally runs 26.2 miles. Suppose the times for each of the legs are normally distributed with the given means and standard deviations.

Swim: $\mu_S = 76$ minutes and $\sigma_S = 18$ minutes

Bike: $\mu_B = 385$ minutes and $\sigma_B = 32$ minutes

Run: $\mu_R = 294$ minutes and $\sigma_R = 25$ minutes

What percent of the competitors finish the Ironman in under 710 minutes?

■ 6. You buy a scratch off lottery ticker for $\$1$ at the local gas station. If you get three hearts in a row on your scratch off, the state will pay you $\$500$. Let X be the amount the state pays you and let X have the following probability distribution.

X	$0	$500
P(X)	0.999	0.001

Suppose you buy one of these scratch off tickets every day for a week (7 days). Find the expected value and standard deviation of your total winnings.

PERMUTATIONS AND COMBINATIONS

1. Calculate the binomial coefficient.

$$\binom{12}{7}$$

2. Calculate ${}_{10}P_3$.

3. How much greater is ${}_5P_2$ than ${}_5C_2$?

4. The high school girls' basketball team has 8 players, 5 of whom are seniors. They need to figure out which senior will be captain and which senior will be co-captain. To make it fair, they choose two players out of a hat. The first drawn will be captain and the second will be co-captain. How many different captain/co-captain pairs are possible?

5. How many different ways can the letters in the word "SUCCESS" be rearranged?

6. Mrs. B's kindergarten class has 14 students and Mr. G's kindergarten class has 16 students. Three students will be selected at random from each of these classrooms to ride on a float in the school parade coming up next week. How many different groups of 6 can be chosen to ride the float?

BINOMIAL RANDOM VARIABLES

■ 1. You toss a fair coin 15 times and record the number of tails.

Is this experiment modeled by a binomial random variable? If it isn't, explain why. If it is, determine its parameters n and p and express the binomial random variable as $X \sim B(n, p)$.

■ 2. You randomly select students from your school until you find a student in the school band. Assume there are 900 students in the school and 80 participate in the school band.

Is this experiment modeled by a binomial random variable? If it isn't, explain why. If it is, determine its parameters n and p and express the binomial random variable as $X \sim B(n, p)$.

■ 3. Let $X \sim B(n, p)$ be a binomial random variable with $n = 12$ and $p = 0.08$. Find $P(X = 4)$.

■ 4. Let Y be the number of times you roll a 1 on a fair 6-sided die if you do 10 trials. Fill in the following probability distribution for Y, rounding each probability to 4 decimal places.

Y	0	1	2	3	4	5	6	7	8	9	10
P(Y)											

5. For each binomial random variable, determine whether the shape of the probability distribution will be skewed right, skewed left, or symmetrical.

1. $X \sim B(n,p)$ with $n = 10$ and $p = 0.15$
2. $Y \sim B(n,p)$ with $n = 10$ and $p = 0.75$
3. $Z \sim B(n,p)$ with $n = 10$ and $p = 0.50$

6. Suppose an environmental biologist is studying juvenile sunfish mortality. He finds that only $30\,\%$ of juvenile sunfish survive in a certain lake. Out of 8 randomly selected juvenile sunfish, what is the probability that exactly 3 will survive?

"AT LEAST" AND "AT MOST," AND MEAN, VARIANCE, AND STANDARD DEVIATION

■ 1. Assume X is a binomial random variable. Let $X \sim B(n, p)$ with $n = 15$ and $p = 0.45$. Find $P(X > 7)$.

■ 2. According to a 2017-2018 survey, $68\,\%$ of U.S. households own a pet. Suppose we select 12 households at random. What is the probability that fewer than 8 of them own a pet?

■ 3. According to a 2017-2018 survey, $68\,\%$ of U.S. households own a pet. Suppose 200 households are selected at random. Find the expected value and standard deviation for the number of households that own a pet.

■ 4. $3\,\%$ of runners in the Boston Marathon do not finish. Suppose we select a SRS of 140 Boston Marathon runners. How many do we expect to finish the race?

■ 5. You roll a fair die 6 times. What is the probability you'll observe an even number in at most 3 of the rolls?

6. You roll two fair 6-sided die 10 times and observe the sum. What is the probability of rolling a sum of 7 on at least six of the rolls?

BERNOULLI RANDOM VARIABLES

■ 1. A game at the local county fair involves spinning a circular spinner that's divided into 8 congruent sections, only two of which are "winners." You buy 5 spins for $\$3.00$. If you land on "winner" on any of your 5 spins, you get to choose a stuffed animal.

Is this an example of Bernoulli trials?

■ 2. A game at the local county fair involves spinning a circular spinner that's divided into 8 congruent sections, only two of which are "winners." You buy 5 spins for $\$3.00$. If you land on "winner" on any of your 5 spins, you get to choose a stuffed animal.

Find the mean and standard deviation for each trial.

■ 3. A game at the local county fair involves spinning a circular spinner that's divided into 8 congruent sections, only two of which are "winners." You buy 5 spins for $\$3.00$. If you land on "winner" on any of your 5 spins, you get to choose a stuffed animal.

Find the mean and standard deviation for the number of winners expected in a set of 5 spins.

4. A game at the local county fair involves spinning a circular spinner that's divided into 8 congruent sections, only two of which are "winners." You buy 5 spins for $\$3.00$. If you land on "winner" on any of your 5 spins, you get to choose a stuffed animal.

Find the probability of observing no winners in a set of 5 spins.

5. A game at the local county fair involves spinning a circular spinner that's divided into 8 congruent sections, only two of which are "winners." You buy 5 spins for $\$3.00$. If you land on "winner" on any of your 5 spins, you get to choose a stuffed animal.

What is the probability of observing at least 1 winner in a set of 5 spins?

6. Your goal is to learn about the percentage of students with high ACT scores. You randomly select high school seniors and record their highest ACT score.

Explain why these aren't Bernoulli trials. Then design a way to conduct the experiment differently so that they can be considered Bernoulli trials.

GEOMETRIC RANDOM VARIABLES

■ 1. You toss a coin until you get "tails." Does this experiment represent a geometric random variable? If it doesn't, explain why. If it does, determine its parameter p and express the variable as $X \sim \text{Geom}(p)$.

■ 2. You randomly select students from your school until you find a student in the school band. Assume there are 900 students in the school and 80 participate in the school band. Does this experiment represent a geometric random variable? If it doesn't, explain why. If it does, determine its parameter p and express the variable as $X \sim \text{Geom}(p)$.

■ 3. Let $X \sim \text{Geom}(p)$ with $p = 0.25$. Find $P(X = 5)$.

■ 4. Suppose we roll a 6-sided fair die until we observe a 2. What is the probability that a 2 will be observed within the first 5 trials?

■ 5. Suppose we roll a 6-sided fair die until we observe a 2. What is the probability that a 2 won't be observed until at least the 6th trial?

6. According to a 2017-2018 survey, $68\,\%$ of U.S. households own a pet. Suppose we start randomly surveying households and asking whether they are pet owners. How many do we expect we will need to survey to find our first household that owns a pet?

TYPES OF STUDIES

■ 1. The following table shows the age and shoe size of six children. Does the data have a positive correlation, negative correlation, or no correlation?

Age	Shoe size
3	7
3	6
5	9
6	12
6	11
7	13

■ 2. A class conducts a survey and finds that $75\,\%$ of the school spends 2 or more hours on social media each day. Is the data one-way or two-way data? Is the study observational or experimental?

■ 3. The following table shows the number of classes from which students were absent and their final grade in the class. Does the data have a positive correlation, negative correlation, or no correlation?

Number of absences	0	0	1	2	3	3	3	5	5	6	7	10
Final grade	95%	97%	90%	86%	80%	74%	70%	65%	64%	58%	55%	45%

■ 4. The table below shows the favorite winter activities for 50 adults. Is the data one-way data? Why or why not?

	Skiing	Snowboarding	Ice Skating
Men	9	13	6
Women	8	7	7

■ 5. Is the following experiment an example of a double-blind experiment? If not, what could be changed to make it a double-blind experiment?

> "A soda company has developed a new flavor and wants to know how it compares in taste to competitor sodas. An employee of the soda company conducts a survey where participants are asked which soda tastes the best. The sodas are given to participants in unmarked plastic cups by the employee."

■ 6. A new cancer drug is being used to treat cancer in children and adults. The hospital conducts a study to measure the effectiveness of the new drug. Cancer patients are placed into groups according to their age and each age range is split into two groups. One group is given traditional treatment of the cancer and the other group is given the new drug. Is the data one-way or two-way data? Is the study observational or experimental?

SAMPLING AND BIAS

■ 1. The zoo conducts a survey on why patrons enjoy coming to the zoo. They ask families with children about why they like to visit the zoo as they're leaving. Give a reason why the sampling method may be biased.

■ 2. The owner of a restaurant gives a survey to each customer. Included in the survey is the question "Have you ever not tipped your waiter or waitress?" Give a reason why the sampling method may be biased.

■ 3. A health club wants to purchase a new machine and would like to know which machine members would most like to have. It creates a survey where members can rate the different machines that the health club is considering purchasing, and posts it at the reception desk for members to fill out if they choose to do so. Does the sample contain a bias? If so, what kind?

■ 4. A biologist wants to study a group of prairie dogs for parasites, but cannot examine the entire population. Which sampling method would be better in this case, a stratified random sample or a clustered random sample?

5. A hospital is studying the health effects of obesity. They group patients into different groups according to a specific weight range and study a variety of biometrics. What type of sampling is this?

6. A museum wants to find out the demographics of its patrons. They set up a survey and ask every 5th customer about their age, ethnicity, and gender. What type of sampling is this?

SAMPLING DISTRIBUTION OF THE SAMPLE PROPORTION

■ 1. The state representatives want to know how their constituents feel about the new tax to fund road improvements, so they send out a survey. Of the 5 million who reside in the state, $150{,}000$ people respond. $40\,\%$ disapprove of the new tax and $60\,\%$ are in favor of the new tax because of the improvements they've seen to the roads. Does this sample satisfy normality?

■ 2. An ice cream shop states that only $5\,\%$ of their $1{,}200$ customers order a sugar cone. You want to verify this claim, so you randomly select 120 customers to see if they order a sugar cone. Is this a normal sampling?

■ 3. The zoo conducts a study about the demographics of its patrons. Every 10th customer or group is recorded as a family, and defined as a group with children under 12 or not. They find that 45 families are recorded and only 20 are not part of a family with children under 12. That day there were 650 visitors or groups. What is the standard deviation for the sample?

■ 4. A pizza shop finds that $80\,\%$ of the 75 randomly selected pizzas ordered during the week have pepperoni. What is the standard deviation

for the sample if the pizza shop has a total of $1{,}000$ pizzas ordered during the week?

■ 5. A hospital conducts a survey on a particular day and finds that 10 patients of 30 randomly selected have high blood pressure. There were 325 patients in the hospital that day. What is the standard deviation for the sample?

■ 6. A study claims that first-born children are more likely to become leaders. The study finds that $72\,\%$ of $2{,}000$ first-born children are currently in or have held leadership roles in their careers. Another group of scientists wants to verify the claim, but can't survey all $2{,}000$ people, so they randomly sample 175 of the participants first-born children. What is the probability that their results are within $2\,\%$ of the first study's claim?

SAMPLING DISTRIBUTION OF THE SAMPLE MEAN

■ 1. The population of 32 year-old women in the United States have an average salary of $42{,}000$, but the distribution of their salaries is not normally distributed. A random sample of 24 women is taken. Does the sample meet the criteria to use the central limit theorem?

■ 2. There are 130 dogs at a dog show who weigh an average of 11 pounds with a standard deviation of 3 pounds. A sample of 9 dogs is taken. What is the standard deviation of the sampling distribution?

■ 3. A large university population has an average student age of 30 years old with a standard deviation of 5 years, and student age is normally distributed. A sample of 80 students is randomly taken. What is the probability that the mean of their ages will be less than 29?

■ 4. A cereal company packages cereal in 12.5-ounce boxes with a standard deviation of 0.5 ounces. The amount of cereal put into each box is normally distributed. The company randomly selects 100 boxes to check their weight. What is the probability that the mean weight will be greater than 12.6 ounces?

5. A hospital finds that the average body temperature of their patients is $98.4°$, with a standard deviation of $0.6°$, and we'll assume that body temperature is normally distributed. The hospital randomly selects 30 patients to check their temperature. What is the probability that the mean temperature of these patients $\bar{x}$ is within $0.2°$ of the population mean?

6. A company produces volleyballs in a factory. Individual volleyballs are filled to an approximate pressure of 7.9 PSI (pounds per square inch), with a standard deviation of 0.2 PSI. Air pressure in the volleyballs is normally distributed. The company randomly selects 50 volleyballs to check their pressure. What is the probability that the mean amount of pressure in these balls $\bar{x}$ is within 0.05 PSI of the population mean?

ALTERNATIVE AND NULL HYPOTHESES

■ 1. A current pain reliever has an $85\,\%$ success rate of treating pain. A company develops a new pain reliever and wants to show that its success rate of treating pain is better than the current option.

Decide if the hypothesis statement would require a population proportion or a population mean, then set up the statistical hypothesis statements for the situation. State if the test is a one- or two-tailed test.

■ 2. A research study on people who quit smoking wants to show that the average number of attempts to quit before a smoker is successful is less than 3.5 attempts. They set up their hypothesis statements as

$$H_0 : \mu = 3.5$$

$$H_a : \mu < 3.5$$

What are the Type I and Type II errors in this study? What could be the consequences of each type of error in a published report of the study?

■ 3. A factory creates a small metal cylindrical part that later becomes part of a car engine. Because of variations in the process of manufacturing, the diameters are not always identical. The machine was calibrated to create parts wires with an average diameter of $1/16$ of an inch. During a periodic inspection, it became clear that further investigation was needed to

determine whether or not the machine responsible for making the part needed recalibration.

Decide if the hypothesis statement would require a population proportion or a population mean, then set up the statistical hypothesis statements for the situation. State if the test is a one- or two-tailed test.

■ 4. A marketing study for a clothing company concluded that the mean percentage increase in sales could potentially be over $17\,\%$ for creating a clothing line that focused on lime green and polka dots. The clothing company used its sample data to test the hypothesis statements.

$$H_0 : \mu = 17$$

$$H_a : \mu > 17$$

Did the clothing company reject or fail to reject the null hypothesis? If their conclusion is incorrect, what type of error are they making (Type I or Type II) and what are the consequences of making that error?

■ 5. A factory creates a small metal cylindrical part that later becomes part of a car engine. Because of variations in the process of manufacturing, the diameters are not always identical. The machine was calibrated to create parts wires with an average diameter of $1/16$ of an inch. During a periodic inspection, it became clear that further investigation was needed to determine whether or not the machine responsible for making the part needed recalibration.

In the context of this situation, describe the Type I and Type II errors and the consequences of each. Based on the consequences, should you choose an α-level of 0.10 or 0.01?

■ 6. A new heartworm test is being developed to help improve the accuracy of detecting heartworms in dogs. The heartworm test was given to 60 dogs known to have heartworms, and it correctly identified 59 as positive for heartworms.

The test was also given to 64 dogs already known to be heartworm-free. Out of the 64 heartworm-free dogs, the test correctly identified 58 as heartworm-free.

The hypothesis statements for the study were:

H_0 : the dog has heartworms

H_a : the dog doesn't have heartworms

Define the Type I and Type II errors. Estimate the probabilities of each error, α and β, based on the study.

ONE- AND TWO-TAILED TESTS

1. A local high school states that its students perform much better than average on a state exam. The average score for all high school students in the state is 106 points. A sample of 256 high schoolers had an average test score of 129 points with a sample standard deviation of 26.8. Choose and calculate the appropriate test statistic for the data. Choose and calculate the appropriate test statistic for the data.

2. A dietician is looking into the claim at a local restaurant that the number of calories in its portion sizes is lower than the national average. The national average is $1{,}500$ calories per meal. She samples 35 meals at the restaurant and finds they contain an average of $1{,}250$ calories per meal with a sample standard deviation of 350.2.

3. In a recent survey, 567 out of a 768 randomly selected dog owners said they used a kennel that was run by their veterinary office for their dogs while they were away on vacation. The study would like to make a conclusion that the majority of dog owners use a kennel run by their veterinary office when the owners go on vacation. Choose and calculate the appropriate test statistic for the data.

■ 4. A school board wants to support opening a new day care center. They look at local random sample of 500 households with children under preschool age. 243 of the households were using a family member to care for their children that were under preschool age. The school board wants to determine if less than half of the households are now using a family member to care for their children at a statistically significant level.

1. Set up the hypothesis statements.
2. Check that the conditions for normality are met.
3. State the type of test: upper-tailed, lower-tailed, or two-tailed.
4. Calculate the test statistic using the appropriate formula.

■ 5. The highest allowable amount of bromate in drinking water is 0.0100 mg/L^2. A survey of a city's water quality took 50 water samples in random locations around the city and found an average of 0.0102 mg/L^2 of bromate with a sample standard deviation of 0.0025. The survey committee is interested in testing if the amount of bromate found in the water samples is higher than the allowable amount at a statistically significant level.

1. Set up the hypothesis statements.
2. Check that the conditions for normality are met.
3. State the type of test: upper-tailed, lower-tailed, or two-tailed.
4. Calculate the test statistic using the appropriate formula.

■ 6. A farmer reads a study that states: The average weight of a day-old chick upon hatching is 38.60 grams with a population standard deviation of 5.7 grams. The farmer wants to see if her day-old chicks are within the average quoted in the study. She takes a simple random sample of 60 of her day-old chicks and finds their average weight is 39.1 grams.

1. Set up the hypothesis statements.
2. Check that the conditions for normality are met.
3. State the type of test: upper-tailed, lower-tailed, or two-tailed.
4. Calculate the test statistic using the appropriate formula.

P-VALUES

■ 1. A private university is conducting a statistical test to determine whether or not the percentage of students who live on its campus is above the national average of $64\,\%$. They've calculated the test statistic as equal to 1.40. Set up the hypothesis statements and determine the type of test, then find the p-value.

■ 2. The national average length of pregnancy is 283.6 days with a population standard deviation of 10.5 days. A hospital wants to know if the average length of a pregnancy at their hospital deviates from the national average. They use the sample of the $9{,}411$ births at the hospital to calculate a test statistic of -1.6. Set up the hypothesis statements and determine the type of test, then find the p-value.

■ 3. The highest allowable amount of bromate in drinking water is $0.0100\ (\text{mg/L})^2$. A survey of a city's water quality took 61 water samples in random locations around the city and used the data to calculate a test statistic of 0.57. The city wants to know if the amount of bromate in their drinking water is too high. Set up the hypothesis statements and determine the type of test, then find the p-value.

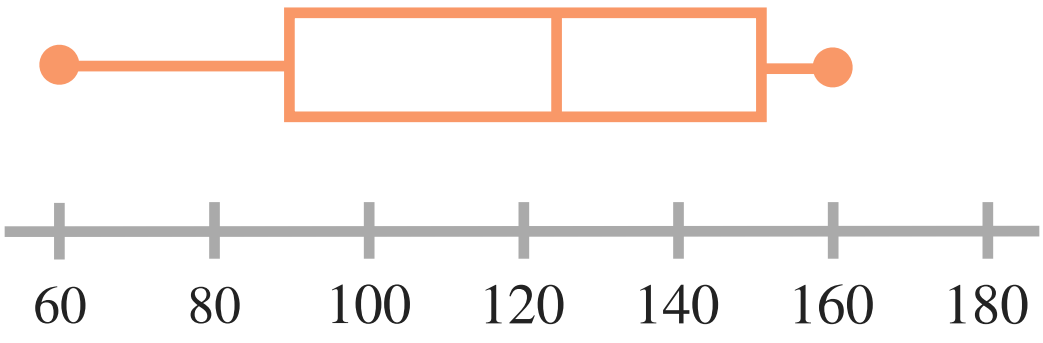
■ 4. A company produces red glow in the dark paint with an advertised glow time of 15 min. A painter is interested in finding out if the product behaves worse than advertised. She sets up her hypothesis statements as

$$H_0 : \mu = 15$$

$$H_a : \mu < 15$$

and using a sample of 14 observations from a random sample, she calculates a test statistic of -3.2. She assumes the distribution is approximately normal. What would be the conclusions of her hypothesis test at significance levels of $\alpha = .05$, $\alpha = .01$, and $\alpha = .001$?

■ 5. A manager reads an article that reports the average wasted time by an employee is 125 minutes every day. She takes a small random sample of 16 employees and monitors their wasted time. She calculates the following based on her observations: the average wasted time for her employees is 122 minutes with a standard deviation of 28.7. She wants to know if 122 minutes is below average at a significance level of $\alpha = .05$. She creates a box plot of her data for "wasted time at work" to check for normality:



1. State the population characteristic and whether you should use a t-test or z-test statistic.

2. Check that the conditions for performing the statistical test are met.
3. Set up the hypothesis statements.
4. State the type of test: upper-tailed, lower-tailed or two-tailed.
5. Calculate the test statistic using the appropriate formula.
6. Calculate the p-value.
7. Compare the p-value to the significance level and draw a conclusion.

6. A recent study reported that the $15.3\,\%$ of patients who are admitted to the hospital with a heart attack die within 30 days of admission. The same study reported that $16.7\,\%$ of the $3{,}153$ patients who went to the hospital with a heart attack died within 30 days of admission when the leading cardiologists were attending an academic conference.

Is there enough evidence to conclude that the percentage of patients who die when the lead cardiologists are away is any different than when they are at the hospital? Use a significance level of $\alpha = .05$ and also $\alpha = .01$.

1. State the population characteristic and whether you should use a t-test or z-test statistic.
2. Check that the conditions for performing the statistical test are met.

3. Set up the hypothesis statements.
4. State the type of test: upper-tailed, lower-tailed or two-tailed.
5. Calculate the test statistic using the appropriate formula.
6. Calculate the p-value.
7. Compare the p-value to the significance level and draw a conclusion.

CONFIDENCE INTERVALS OF A POPULATION MEAN

■ 1. A confidence interval for a study is $(11.5,18.5)$. What was the value of the sample mean?

■ 2. A student wanted to know how many chocolates were in the small bags of chocolate candies her school was selling for a fundraiser. She took a simple random sample of small bags of chocolate candy. From the sample she found an average of 17 pieces of candy per bag with a standard deviation of 2.030.

A box-plot of the data from the sample showed the distribution to be approximately normal. Compute and interpret a $95\,\%$ confidence interval for the mean amount of chocolate candy per bag.

■ 3. Consider the formula for a confidence interval for a population mean with an unknown sample standard deviation, how does doubling the sample size affect the confidence interval?

$$(a,b) = \bar{x} \pm t^* \cdot \frac{s}{\sqrt{n}}$$

■ 4. A magazine took a random sample of 540 people and reported the average spending on an Easter basket this year to be $\$44.78$ per basket

with a sample standard deviation of 18.10. Construct and interpret a $98\,\%$ confidence interval for the data.

■ 5. The national average of calories served in a restaurant meal is $1{,}500$ calories per meal from a sample of 31 randomly selected samples. For illustrative purposes, say that the a population standard deviation of the calories in a restaurant meal is 350.2. Construct and interpret a $90\,\%$ confidence interval for the mean number of calories in a restaurant meal.

■ 6. A bus that travels from Kansas City to Denver had the following travel times, in hours, on 11 randomly selected bus trips:

11.7	12.0	11.75	11.5
12.25	11.5	12.0	11.5
11.25	11.25	11.75	

Construct and interpret a $95\,\%$ confidence interval for the mean bus trip time in hours from Kansas City to Denver.

CONFIDENCE INTERVALS OF A POPULATION PROPORTION

■ 1. A court case questioning the use of a drug dog named Bentley, due to his low success rate of correctly alerting to drugs, was in the U.S. Court of Appeals (U.S. vs. Bentley). In Bentley's time on the job, it's estimated that he correctly identified drugs 59% of the time. How many different trials should they put Bentley through to show that this is his actual success rate at a 95% confidence level with a margin of error of $.05$?

■ 2. Sarah is conducting a class survey to determine if the percentage of juniors in favor of having the next dance at a local bowling alley is 65%. How many juniors should she survey to have a 90% confidence level with a margin of error of $.08$?

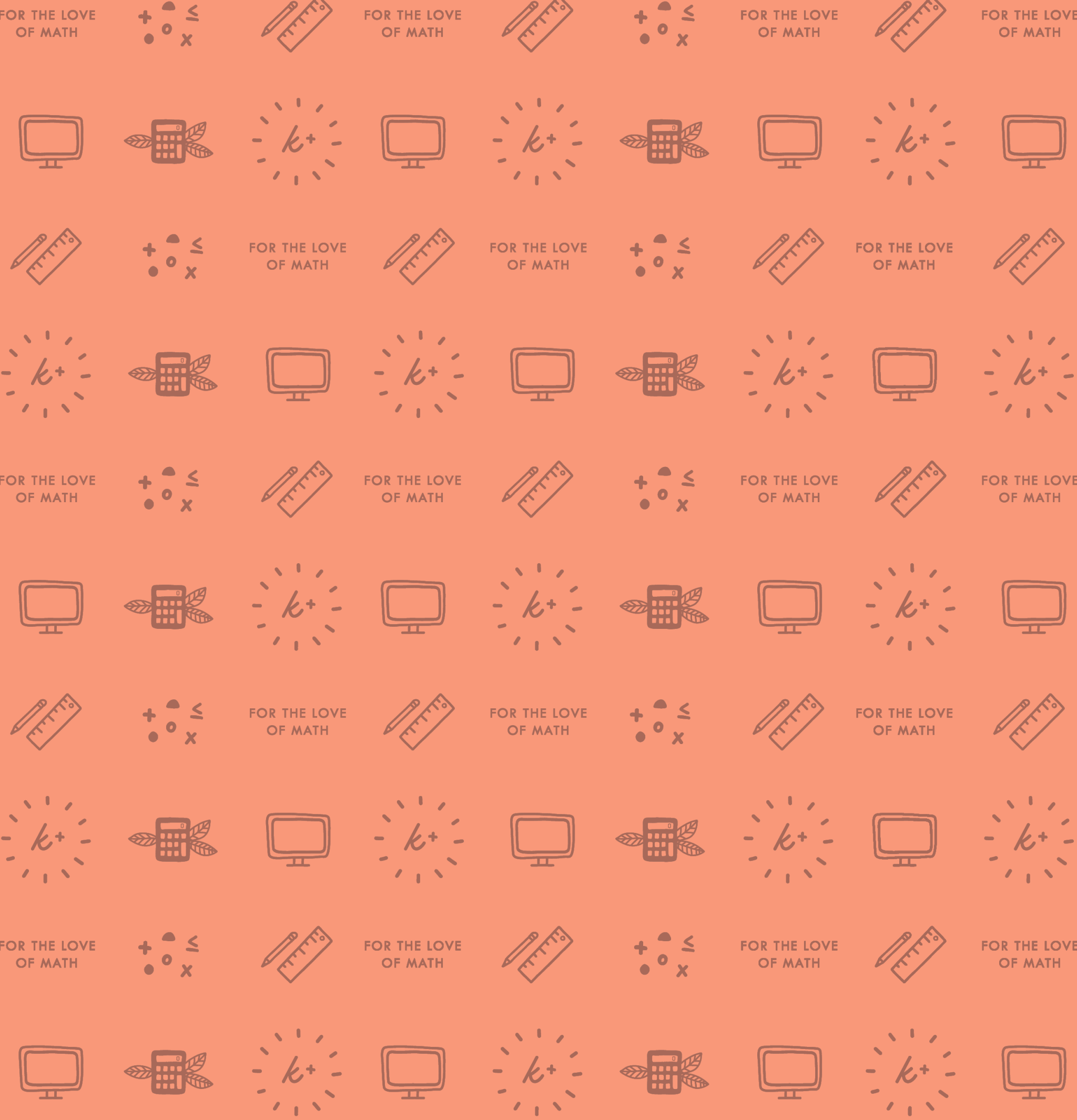
■ 3. A study suggests that 10% of practicing physicians are cognitively impaired. What random sample of practicing physicians is needed to confirm this finding at a confidence level of 95% with a margin of error of $.05$?

■ 4. A study shows that 78% of patients who try a new medication for migraines feel better within 30 minutes of taking the medicine. If the study involved 120 patients, construct and interpret a 95% confidence interval for

the proportion of patients who feel better within 30 minutes of taking the medicine.

■ 5. A study shows that 243 out of 500 randomly selected households were using a family member to care for their children who were under preschool age. Construct and interpret a $90\,\%$ confidence interval for the proportion of households using a family member to care for children under preschool age.

■ 6. According to a recent poll, $47\,\%$ of the 648 Americans surveyed make weekend plans based on the weather. Construct and interpret a $99\,\%$ confidence interval for the percent of Americans who make weekend plans based on the weather.



WWW.KRISTAKINGMATH.COM