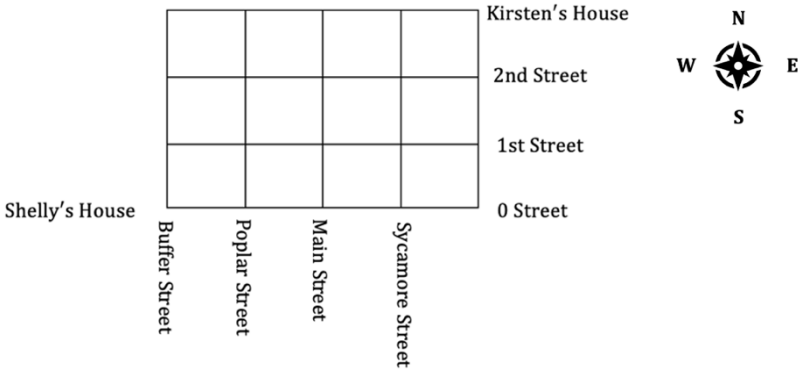
Module 4 Working with Statistics

Section 4.1 Data Types

Practice Problems 4.1

For Problem 1-5, solve the word problem.

1. Shelby wants to go to Kirsten's house to visit her. She will be traveling north and east. How many different routes can Shelly take?



2. In Example 3 from the Lesson Notes, is the data categorical or numerical?

3. In Example 4 from the Lesson Notes, is the data categorical or numerical?

4. For the types of data below, write *C* next to *categorial data* and *N* next to *numerical data*.

a) Birth month

b) Language spoken most often in a country

c) Travel time to church

d) Travel method to church

e) City where you live

f) Ages in a group study

g) Amount of money earned for mowing lawns

h) Hours slept each night

i) Favorite sport

j) Favorite song

5. For the types of data below, write *D* next to *discrete data* and *C* next to *continuous data*.

a) Birth dates of students in a Sunday school class

b) Death dates of World War I Veterans at Arlington Cemetery

c) Travel time on vacation

d) Travel distance on vacation

e) Height of water as a pool is filled

f) Temperatures during the month of June

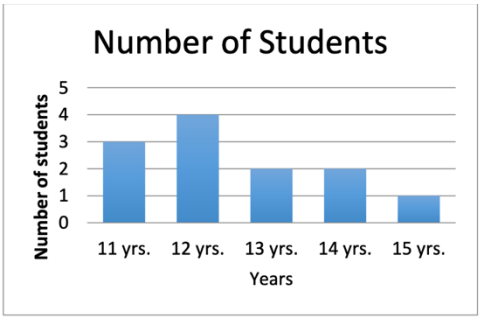
g) Amount of money in your savings account

h) Amount of money in your savings account on January 1st of each year

i) Temperature at 6:00 A.M. each day of the week

j) Wingspan of a butterfly

For Problem 6-10, use the chart below to the left to solve the problem.



6. Is the data in the chart discrete or continuous?

7. Is the data in the chart categorical or numerical?

8. How many categories are there in the chart?

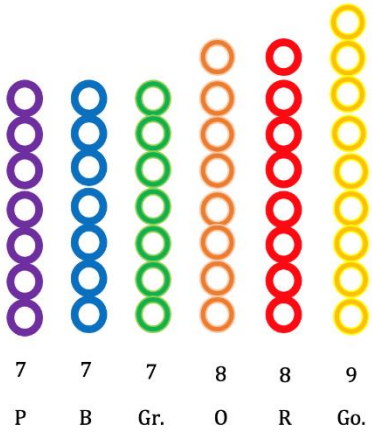
9. How many students were categorized by age?

10. How many more 12-year-olds are there than 15-year-olds?

Section 4.2 Univariate Data

Practice Problems 4.2

For Problem 1-9, use the Fruity Loop example from the Lesson Notes to solve the problem.



1. If 23 is the median number in $\frac{1}{4}$ cup of Fruity Loops (when you line them all up), what is the total number of Fruity Loops in the $\frac{1}{4}$ cup? Is this the median of our example?

2. If 23.5 is the median number in $\frac{1}{4}$ cup of Fruity Loops (when you line them all up), what is the total number of Fruity Loops in $\frac{1}{4}$ cup? Is this the median from our example?

3. Is it possible to find the median categorically when items are sorted out by color?

4. If you pick one Fruity Loop from the $\frac{1}{4}$ cup of Fruity Loops, what is the probability of picking each color?

$P(P) =$ $P(B) =$

$P(Gr.) =$ $P(Go.) =$

$P(O) =$ $P(R) =$

5. Which color has the highest probability of being picked? What does this mean?

6. Why does the mode of the colors have the highest probability of being picked?

7. What is the probability of picking purple (P) or blue (B) or green (Gr.) if you reach in and pick one Fruity Loop?

8. What is the probability of not picking a gold (Go.) if you reach in and pick one Fruity Loop?

9. Use a different method for finding the probability of not picking a gold when you pick one Fruity Loop.

For Problem 10, use the given information to solve the problem.

Probability problems often involve selections (picking things) based on the likelihood of an event occurring. This challenge problem involves one pick that has 100% certainty of getting the desired outcome.

Suppose you are a mail clerk. There are three bags in the mailroom. They are marked "Letters," "Secrets," and "Letters & Secrets." Your manager labeled each bag incorrectly to confuse anyone wanting to steal his secrets and go into business for themselves.

10. Suppose you were curious and had only one minute to draw one document from one bag. Which bag would you choose from so that you might discover which bag is full of secrets?

To solve this problem, set up a simulation. Take three bags and set them out. Label one "Letters," one "Secrets," and one "Letters & Secrets." Mark three cards with different words or symbols and place one each in of the bags. Draw one card from one bag and see if you can determine where the secrets are. Write down what you pick. Keep working the simulation until you figure out the answer. You might want to try the cards in different bags.

Section 4.3 Mean Distribution

Practice Problems 4.3

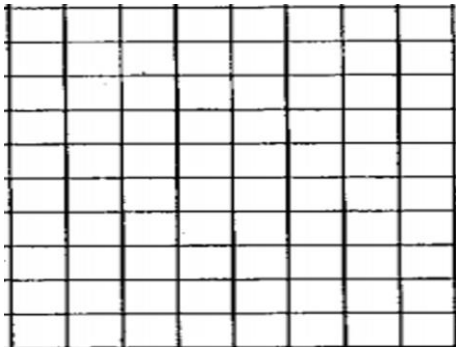
For Problem 1-4, use the given information and table to solve the problem.

Language Spoken	Number of Students
Russian	5
English	8
Portuguese	3
Spanish	5
Swahili	2
Tagalog	2

At a missionary school in Brazil, there are students from all over the world. The number of students and the primary language they speak is represented in the table.

1. What is the mode of the data and what does it represent?

2. Make a bar graph of the data.



3. Find the median, range, and mean of the data of the numerical data? What information do they give you? What is the mode of the categorical data?

For Problem 5-10, use the information given to solve the problem.

5. The following are the ages of the students in a tour group at the zoo: 12, 12, 11, 13, 14, 14, 11, 7, 8, 14, and 14. Find the median, mode, range, and mean of the group; what information do they give you?

6. If two parents' ages are 35 and 36 and they join the group from Problem 5 at the zoo, how does this affect the median, mode, range, and mean of the group? How does this affect the interpretation of the data?

7. Multiple Choice: There are three packages ready to be mailed. They weigh 3.8 pounds, 2.6 pounds, and 4.4 pounds. If a fourth package was mailed with them that made the average weight of the packages 4 pounds, how much does it weigh?

a) 3.4 pounds b) 2.2 pounds c) 5.0 pounds d) 5.2 pounds

8. Barbara scored 60 out of 100 points on her first pre-algebra test. She plans to study for the second test which is also 100 points and will be averaged with the first test to get a final grade. What score does Barbara need on the second test to pass the class (a passing grade is 70 points)?

Chelsea and Kameron run the 100-meter dash at five track events. Below are their times:

Chelsea's 100-meter Dash Times					
Meets	1	2	3	4	5
Time (seconds)	12.8	13.3	13.3	12.9	13.5

Kameron's 100-meter Dash Times					
Meets	1	2	3	4	5
Time (seconds)	13.2	13.0	12.8	12.8	13.5

Statistic	Chelsea	Kameron
Median		
Mode		
Range		
Mean		
Maximum		
Minimum		

9. Each girl, Chelsea and Kameron, ran in five meets, but not at the same time, and placed first twice, second twice, and third once. The district meet is coming up. Fill out the table for each runner and then explain which one the coach should choose to run in the district meet and why.

10. Isaiah plays a game at the mall in which gophers pop up out of holes and the player tries to tap their heads with a mallet as quickly as possible. Isaiah plays the game five times. Below are his statistics. Which values give Isaiah the most information about his reaction time? Which values give him the least information about his reaction time?

Statistic	Reaction Time (seconds)
Median	**1.4**
Mode	**1.02**
Range	**1.36**
Mean	**1.5**
Maximum	**2.2**
Minimum	**0.84**

Section 4.4 Bar Graphs or Histograms
Practice Problems 4.4

For Problem 1, use the given information to solve the problem.

In order to determine a scale for data in a frequency table, we look at the range of numbers and include all the numbers for each range of numbers. We want to decide on an interval that divides the numbers into equal parts but would not include too many intervals. We can make the range larger so the interval divides evenly. For example, an approximate scale for 3% through 12% would be 0% through 15% with intervals of 3%.

1. Find an appropriate scale and interval for each range of numbers below.

a) 3% through 99%

b) 15 through 25

c) 6 through 39

d) 10% through 83%

For Problem 2-11, solve the problem given.

2. Bar graphs generally have spaces between the bars but not necessarily. Why is this not important?

3. Histograms generally do not have spaces between the bars. Why is this important?

4. Write *Bar Graph* or *Histogram* under the term given and explain why you did so.

a) Categorical Data

b) Numerical Data

c) Discrete Data

d) Continuous Data

5. Write *Bar Graph* or *Histogram* under the statement given and explain why you did so.

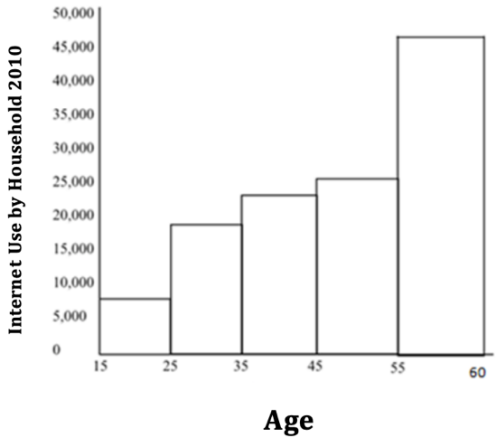
a) Hours of battery life by brands of computers

b) Household income by city

c) Annual high and low temperatures in countries around the world

d) Time it takes students to get ready for school in the morning

6. The histogram below shows the number of households using the Internet (in thousands).



a) What is the scale and interval for households?

b) What is the scale and interval by age?

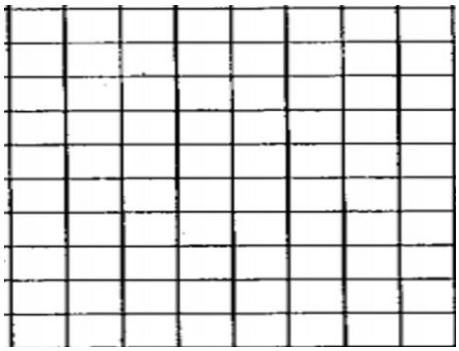
c) Is a 25-year-old included in the 15-25-year-old interval or 25-35-year-old interval?

d) There are four regions from which this data was taken: Northeast, Midwest, South, and West. Does geological location in this instance represent categorical data or numerical data?

e) If we made a graph based on data by region, would we use a bar graph or a histogram?

7. Each member of a family picked their favorite dessert for a family reunion. The results were as follows: 21 chose chocolate pie; 13 chose chocolate cake; 15 chose brownies; 17 chose chocolate mousse. What is an appropriate scale for the y-axis and interval of this data? Why is a bar graph preferrable to a histogram in this instance?

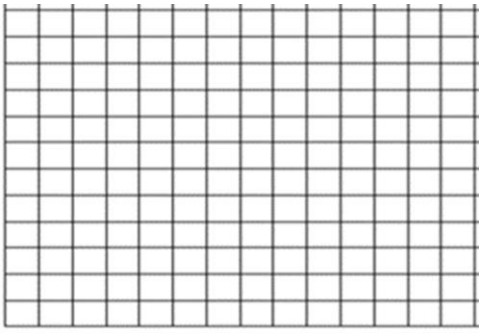
8. Make a bar graph for the results of Problem 7.



9. Below is a frequency distribution for the religious affiliation of the first forty-three presidents of the United States.

Religious Affiliation of President	Frequency of Occurrence
Episcopalian	𝍸 𝍸 I
Unitarian	IIII
Deist	I
Presbyterian	𝍸 III
Reformed Dutch	II
Methodist	IIII
Liberal	I
Disciples of Christ	III
Baptist	III
Congregationalist	I
Quaker	II
Roman Catholic	I
Southern Baptist	I
United Christ of Church	I

a) Make a bar graph of the frequency table. Let the category along the bottom be the religious affiliation of the president. If you use graph paper, let each square represent one president.

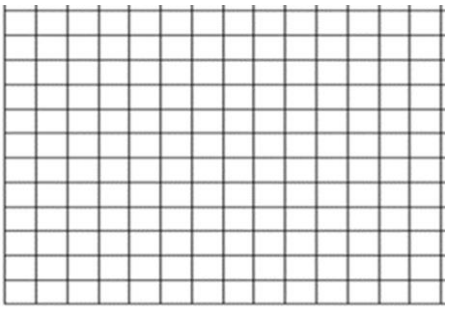


b) What is the mode of religious affiliation of these presidents?

10. Below is a frequency table for the ages of death of the first forty-three presidents of the United States.

Age of Death of President	Frequency of Occurrence
45-49	I
50-54	II
55-59	IIII
60-64	𝍸 I
65-69	𝍸 I
70-74	𝍸 II
75-79	IIII
80-84	III
85-89	II
90-94	III

a) Make a histogram of the frequency table. The earliest death of a president is age 46. Because that is the youngest and there are no deaths between ages 0-45, you can start with 45 on the bottom. A 5-year interval would be ages 45-49. The next 5-year interval would be ages 50-54.



b) Why did we make a histogram for this data rather than a bar graph?

11. Why is finding the range of the data not as helpful in Problem 6 and Problem 9 as in Problem 10?

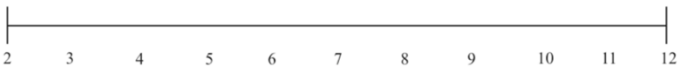
Section 4.5 Normal Distribution

Practice Problems 4.5

For Problem 1-3, use the given information to solve the problem.

1. Roll two dice fifty times. Find the sum of each roll of the two dice and put an X on the line plot above the number you get. What does the line plot look like when you are finished?

a) Why does it have this shape?



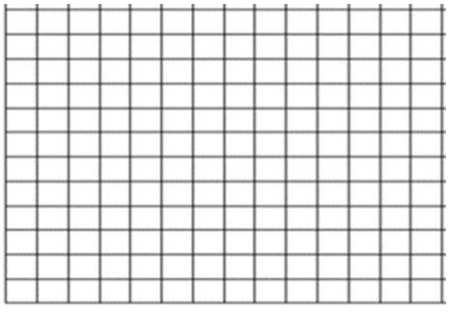
b) Which number on the line plot did most of your X's land on? Explain why.

c) Which numbers on the line plot did the least amount of your X's land on? Explain why.

2. Complete the frequency table below for your data. What represents frequencies on the line plot? What is the category?

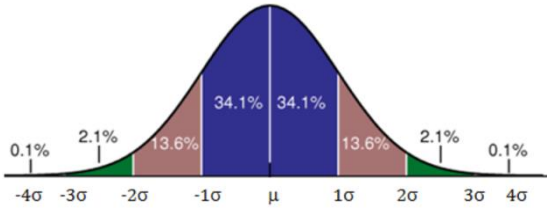
Sum of Two Dice	Frequencies

3. Make a bar graph of your data.



For Problem 4-7, use the given information and IQ test results to solve the problem.

IQ (Intelligence Quotient) is derived from a standard test score and is used to test intelligence. The mean in an age group is 100 and one standard deviation is 15 points (this used to be 10).



4.

a) If μ is 100, find the IQ scores for 1σ, 2σ, 3σ, 4σ, -1σ, -2σ, -3σ, and -4σ.

b) If a person has an IQ of 94, what standard deviation is this from the mean?

c) If a person has an IQ of 113, what percent of the population falls within this interval?

d) What is the possible range of an IQ score if it falls two standard deviations below the mean?

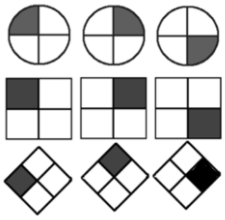
5. In the normal distribution of data, why are the mean, median, and mode all the same?

6. What percent of data falls 4 standard deviations from the mean?

7. What is the IQ score -2σ from the mean?

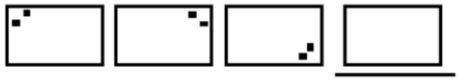
For Problem 8-11, solve the IQ test sample problem given.

8. Find the next shading in the diagnostic series of shapes below.



9. Draw what comes next in the sequence given.

a)



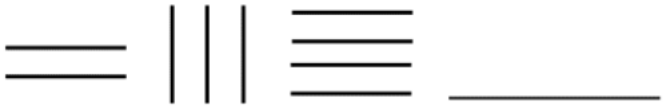
b)

○□, □□○○, ○○○□□□, □□□□○○○○, _____

c)

d)

e)



10. Tell what number comes next in the sequence given.

a) 1, 3, 5, _____

b) 13, 11, 9, _____

c) 1, 1.5, 2, 2.5, 3, _____

d) 1, 3, 6, 10, 15, _____

11. Tell what symbol comes next in the sequence given.

a) MI, MIMI, MIMIMI, _____

b)

_____

c)

_____

Section 4.6 Pie Charts

Practice Problems 4.6

For Problem 1-10, solve the problem given.

1. Circle the following samples of data that model a normal bell-shaped curve:

a) Heights among a given age group of men

b) Weights among a given age group of women

c) Scores on an ACT or SAT college preparatory exam

d) Number of siblings of primary school students

e) Internet usage percentage at home or at work

f) Favorite sports among middle-aged South American men

g) Average income in countries in Europe

h) Scores on a pre-algebra test for 7th graders

For Problem 2-5, use the given data to solve the problem.

Below are the test scores in a 7th grade pre-algebra class:

92%	84%	99%	76%
64%	59%	72%	71%
81%	82%	65%	66%
	93%	84%	

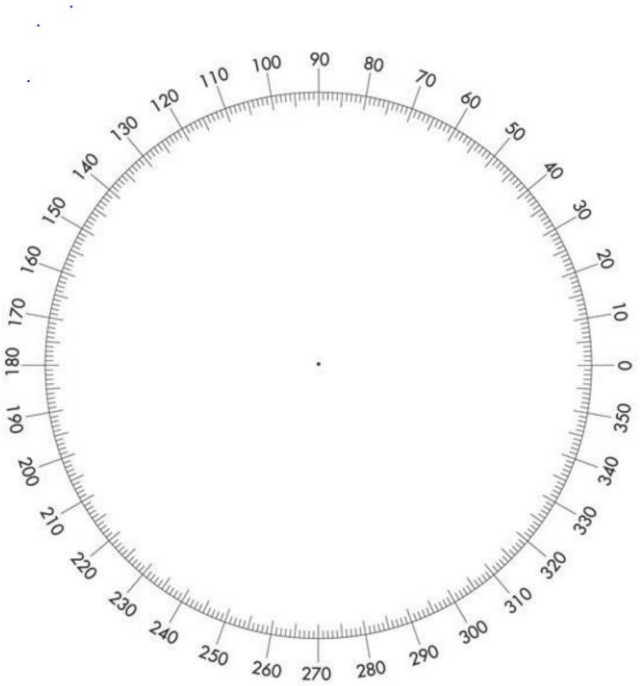
Grade	Frequency
A: 90-100%	
B: 80-89%	
C: 70-79%	
D: 60-69%	
F: 0-59%	

2. a) What is the class average for the test?

b) Complete the frequency table above for the class.

c) Make a pie chart for the percentages of As, Bs, Cs, Ds, and Fs given the grading scale is as follows:

Grade
A: 90-100%
B: 80-89%
C: 70-79%
D: 60-69%
F: 0-59%

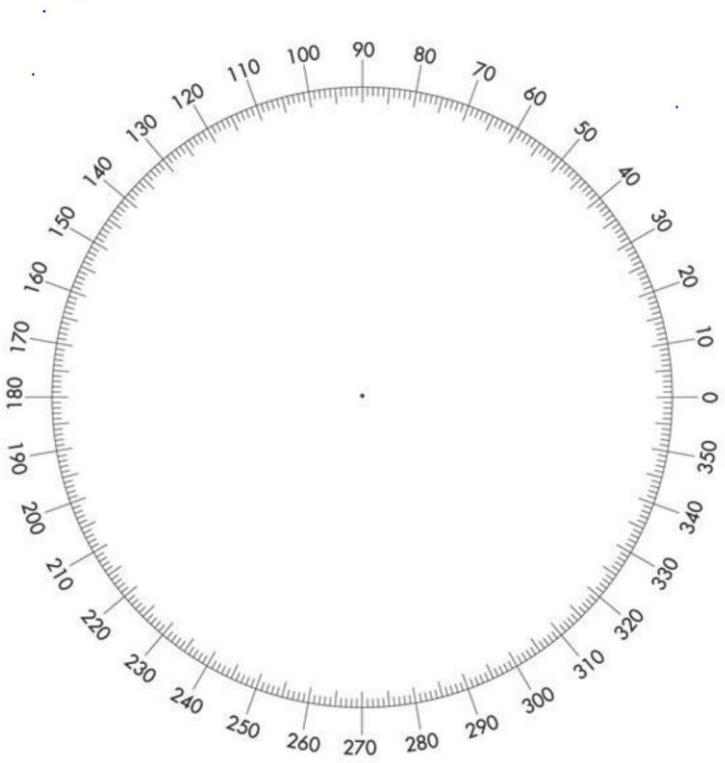


d) Does this grading scale seem reasonable?

e) Should this test be administered again using a different scale? Explain why.

3. How does the pie chart from Problem 2 change if the grading scale is as follows:

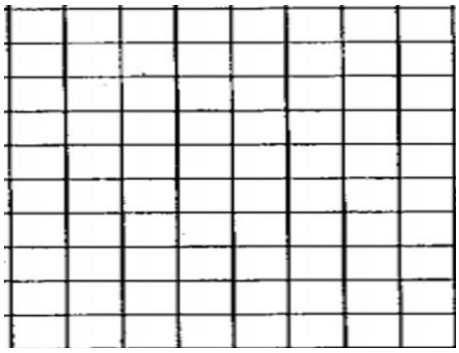
Grade	Frequency
A: 93-100%	
B: 86-92%	
C: 79-85%	
D: 71-78%	
F: 0-70%	



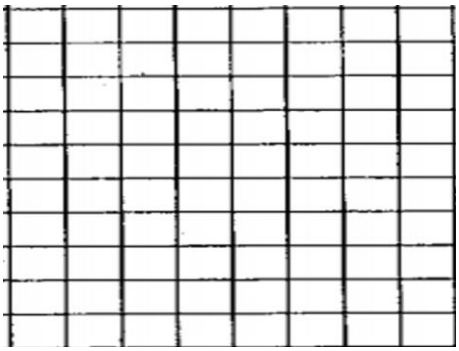
a) Does this grading scale seem reasonable?

b) Should this test be administered again using a different grading scale? Explain why.

4. Make a histogram of the data from Problem 2. Is this preferrable to a pie chart? Explain why or why not.



5. Make a histogram of the data from Problem 3. Is this preferrable to a pie chart? Explain why or why not.



For Problem 6-10, solve the word problem given.

6. Describe each type of graph/chart and give an example of their uses: line graph; bar graph; histogram; pie chart.

7. What is the best graph/chart to use to show the growth of a plant over three weeks?

8. What is the best graph/chart to use to show the types of computers used by different businesses?

9. What is the best graph/chart to use to show the number of each type of medal awarded to United States' Olympic Champions?

10. Does a pie chart display discrete data or continuous data?

Section 4.7 Stem-and-Leaf Plots

Practice Problems

For Problem 1, use the given information and table to solve the problem.

In the table below are the test scores from Example 1 of the Lesson Notes.

STEM	LEAF
6	1 2 4
7	1 2 4
8	3 3 4 4
9	2 7 9

1. How many students took the test? List their scores in order from least to greatest.

For Problem 2-10, use the information given to solve the problem.

2. There are fourteen test scores from the 7^{th} grade pre-algebra class from Problem 2 of the previous section. Make a Stem-and-Leaf plot of all their grades.

3. Does changing the grading scale for the test from Problem 2 change the Stem-and-Leaf plot? Explain why or why not.

4. Why would a Stem-and-Leaf plot not be useful for the number of students in the Mountain Home School District by grade level?

5. Would a Stem-and-Leaf plot be useful if we knew the ages of each of the students in the 9^{th}-12^{th} grade of the 131 students in the Mountain Home School District from Problem 4?

6. Below is a Stem-and-Leaf plot of the Gopher Game task completion time:

STEM	LEAF
4	7 8 9
5	2 2 3 4 5
6	1 2 4

The STEM is the ones units and the LEAF is the tenths units.

a) What was the fastest time?

b) What was the slowest time?

c) How many tasks were timed?

d) Does this appear to be a normal distribution?

e) What was the average time?

f) Do you think the fastest times were at the beginning of the tasks or the end of the tasks?

g) Complete the frequency table for the Gopher Game task completion time.

Time (seconds)	Frequency
4.0-4.9	
5.0-5.9	
6.0-6.9	

h) Make a histogram and see if it appears to be a normal distribution.

7. Below is the number of shoppers at a local grocery store for each day of the week. Make a Stem-and-Leaf plot of the data. Let the tens place be the Stem and the ones place be the Leaf.

64, 49, 53, 57, 71, 68, 65

8. How can you quickly determine if the numbers in a Stem-and-Leaf plot represent all the data?

9. Is it better to use a Stem-and-Leaf plot or a histogram for the number of favorite desserts among patrons of a four-star restaurant?

10. Make a Stem-and-Leaf plot to represent the ages of family members at a cousins' get-together shown below. What age group do most of the members fall in? What is an appropriate key for the data?

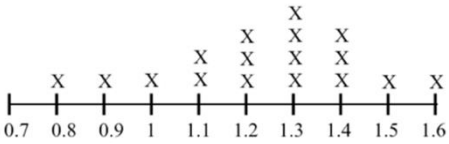
64, 62, 61, 57, 40, 43, 16, 18, 41, 66, 78, 77, 61, 79, 60

Section 4.8 Box-and-Whisker Plots

Practice Problems 4.8

For Problem 1-3, use the given information to solve the problem.

Leticia invites Missy, Patty, Cynthia, Janine, Sophie, Jade, Jackie, Christina, and Angela to a party. She also invites Carl, Jason, Steven, Asher, Regan, August, and Gary. Leticia challenges her guests to the Gopher Game but in this case, it is played on a computer; therefore, rather than tapping gophers on the head in real-life, the players tap a button on the keyboard to virtually tap the gophers on the head. To the right are the reaction times (in seconds) of the party guests when using their dominant hand.



1. a) What is the slowest reaction time? b) What is the fastest reaction time?

c) What is the median of the data? d) What is the mean of the data?

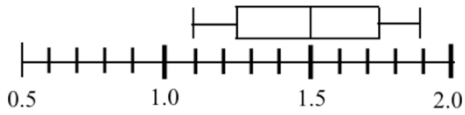
e) If this data were made into a Box-and-Whisker plot, which would have a longer box: the lower quartile or the upper quartile? Explain why.

2. What would you say is the typical fastest reaction time for the data: the median or the mean? Explain why.

3. Make a Box-and-Whisker plot of the data.

For Problem 4 and 5, use the given information to solve the problem.

To the right is a Box-and-Whisker plot of the computer Gopher Game for seventeen children when using their non-dominant hand.



4. a) What is the slowest reaction time? b) What is the fastest reaction time?

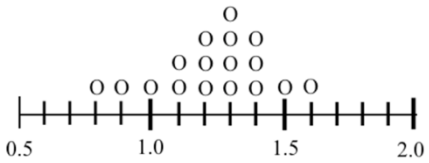
c) What is the median of the data? d) What is the mean of the data?

d) Because the lower rectangular box is longer, does this mean more times fall in the lower quartile?

5. Make a line plot of the data.

For Problem 6-10, use the given information to solve the problem.

To the right is a graph of Gopher Game data that appeared on a computer screen. Jennifer says it is for the slowest reaction times. Cameron says it is for the fastest reaction times. (The assumption is that the fastest reaction times occurred when the player used their dominant hand and the slowest reaction times occurred when the player used their non-dominant hand.)



6. What do you say the graph represents?

7. Could this be the data for the Box-and-Whisker plot from Problem 4 when the children played the Gopher Game with their non-dominant hand?

1.1, 1.2, 1.2, 1.2, 1.3, 1.3, 1.3, 1.4, 1.5, 1.7, 1.7, 1.7, 1.7, 1.8, 1.9, 1.9, 1.9

8. Could this also be the data for the Box-and-Whisker plot from Problem 4 when the children played the Gopher Game with their non-dominant hand?

1.1, 1.1, 1.2, 1.2, 1.3, 1.4, 1.4, 1.4, 1.5, 1.6, 1.6, 1.6, 1.7, 1.8, 1.8, 1.9, 1.9

9. Which data points would have to change to yield a different Box-and-Whisker plot from Problem 4?

10. Change five data values in Problem 4 so the data is still in order from least to greatest but the five-point summary and Box-and-Whisker plot do not change.

Section 4.9 Scatterplots
Practice Problems 4.9

For Problem 1 and 2, use the given information about the study to solve the problem.

A study indicates that eating blueberries before standardized tests results in higher scores. In this study, teenagers in the control group were given blueberries before a test while the other group were not. Over the course of the nineteen-month study, those who ate blueberries before testing showed significantly higher scores.

1. a) What is the independent variable in this study? b) What is the dependent variable in this study?

2. In the study, is the dependent variable qualitative (categorical) or quantitative (numerical)? Is the independent variable qualitative or quantitative?

For Problem 3, indicate the correct category for the term(s) given.

3. Tell whether the variable given is qualitative (categorical) or quantitative (numerical).

a) Eye Color b) Religion

c) Shoe Size d) Height

e) Gender f) Weight

For Problem 4 and 5, use the given information to solve the problem.

A car manufacturer wants to know how bright to make brake lights so when cars behind them see them they will have time to hit the brakes and not crash into the car in front of them.

4. a) Is the brightness of the brake lights the dependent or independent variable?

b) Is the time it takes to hit the brakes the dependent or independent variable?

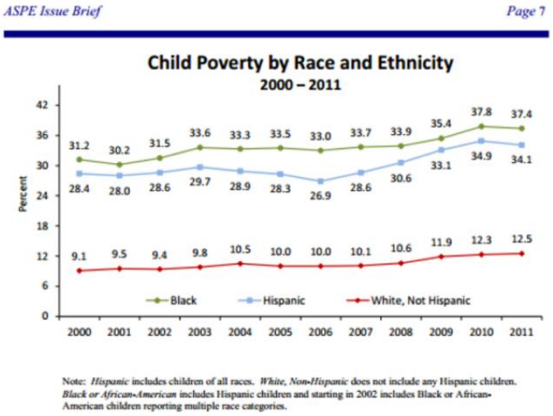
5. In Problem 4, is the dependent variable discrete or continuous? Is the independent variable discrete or continuous?

For Problem 6, use the given table to solve the problem.

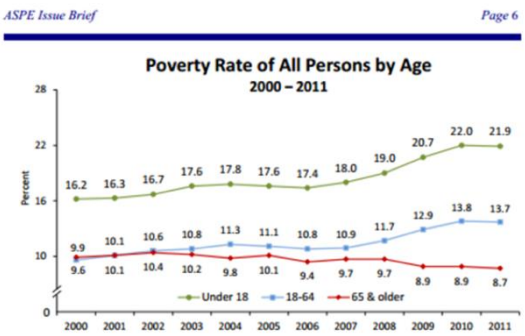
Shoe Name	Shoe Preference	Shoe Price
Defense II	1	180.00
Chromeless	2	90.00
KVJD	3	115.00
Elite X	4	140.00
Custom Cut	5	240.00
XEL	6	180.00
XCEED	7	180.00
Xception	8	240.00
HyperPro	9	66.99
Hypedunk	10	74.97
ProElite	11	100.00
ConAirPro	12	143.99
Kwanlq	13	64.96
SwisherIV	14	115.00
HoopsterII	15	143.96

6. Does there seem to be a correlation between shoe preference and shoe price?

For Problem 7-9, use the given table to solve the problem.



7. Looking at the table of Child Poverty by Race and Ethnicity for 2001-2011, what can you generalize about Black, Hispanic, and White childhood poverty? Is there a correlation between race and poverty?



8. Looking at the table, what do the line plots tell you about poverty and age? Is there a correlation between the two?

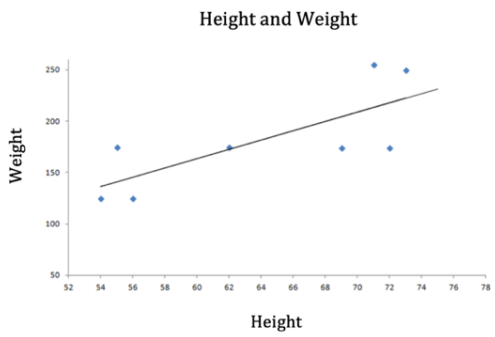
9. Birth Rate tends to be lower in richer countries. Therefore, as income increases, the birth rate decreases. Is this a positive or negative correlation?

10. GDP, which stands for Gross Domestic Product, is the products, goods, and services of a country and measures the size of the country's economy and how it is growing. When the GDP is growing steadily, employment is likely to be high as well. Is this a positive or negative correlation and what is its causation?

Section 4.10 Bivariate Data

Practice Problems 4.10

For Problem 1, use the scatterplot to solve the problem.



1. A trend line has been drawn near the data. Is there a strong (high) positive correlation, a weak (low) positive correlation, or no correlation at all between height and weight?

For Problem 2-6, solve the problem given.

2. In bivariate data, is one set of data always the cause of a second set of data?

3. Given an xy-coordinate graph and a strong correlation, is x dependent on y or is y dependent on x?

4. Interpolation is the process of filling in data points between the data you already have. Looking at Example 1 from the Lesson Notes and using interpolation, what would you predict someone who studied for 2.5 hours to score on the test?

5. Did the someone from Problem 4 who studied for 2.5 hours perform as expected on the test according to the line of best fit?

6. Extrapolation is the process of going outside the data set to make predictions. If you looked at the graph from Example 1 in the Lesson Notes and predicted what a student would score if they studied for 5 hours, are you using interpolation or extrapolation?

For Problem 7-9, solve the multiple-choice problem given.

7. Which graph connects the dots?

a) A line graph b) A scatterplot

8. Which graph is continuous?

a) A line graph b) A scatterplot

9. Which of the following examples describe situations that are not causal relationships?

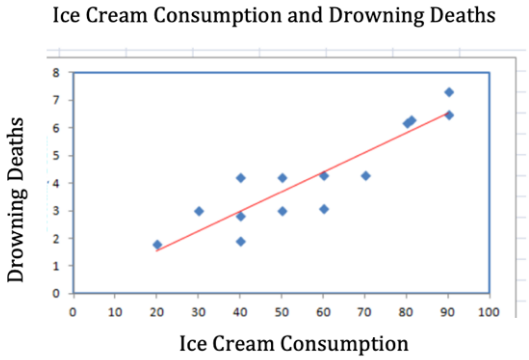
a) Your height and your arm span (wing span)

b) The farther you travel on vacation, the more miles you put on your odometer

c) The more miles you put on your odometer, the more gas you use

d) The more gas you use, the higher the price per gallon of gas

For Problem 10, use the given scatterplot to solve the problem.



10. a) Is this a strong correlation or a weak correlation?

b) Is this a positive correlation or a negative correlation?

c) Does consuming ice cream cause drowning?

d) Is there any other reason ice cream consumption and drowning deaths increase at the same time?

Section 4.11 Probability Sampling

Practice Problems 4.11

For Problem 1-6, use the given survey and table to solve the problem.

A survey of teenagers from ages thirteen to eighteen was taken at a theme park. Every fourth teenager who purchased a ticket at the park was asked to take the survey, which asked them how many hours they spend watching television daily. The table below shows the results.

Teenager	Gender	Hours Spent Watching TV Daily
01	M	3.6
02	F	2.5
03	M	2.1
04	M	5.2
05	F	0
06	M	4.6
07	M	3.9
08	M	3.6
09	M	4.2
10	F	2.1
11	M	2
12	F	4
13	M	4
14	F	3.3
15	M	5.2
16	M	5.5
17	F	4.5
18	F	5.0
19	F	5.5
20	M	5.5
21	M	2.5
22	M	3
23	F	2.5
24	M	6
25	F	3
26	M	6.5
27	F	4.1
28	M	6.5
29	M	5.5
30	M	4.2

1. a) How could you pick a random sample of the group and choose fifteen teenagers to take the survey?

b) Which method was used to choose the teenagers for the survey as they entered the theme park?

c) Make a dot plot of the results for every even-numbered teenager from the table.

d) Which method was used to choose the teenagers from c)?

e) Describe the variability in the data.

f) Can you make any conclusions about the television-watching habits of teenagers?

2. Make a frequency table and histogram for every odd-numbered teenager from the table.

a) Which method was used to choose the teenagers in this instance?

b) Describe the variability in the data. Let the intervals be 1-hour increments.

c) Do your conclusions agree with your conclusions from f) in Problem 1?

3. How is it possible for a teenager to spend 0 hours watching television?

4. Make a Box-and-Whisker plot of the data from the sample in Problem 1c. Can you make any conclusions about the television-watching habits of teenagers from this data? Is this a more accurate representation of the data than the others? Were there others that were more accurate than this?

5. Why is a dot plot not best for the data from Problem 1c?

6. In reflecting on the results, is a larger sample or smaller sample more accurate?

For Problem 7-10, use the information given to solve the problem.

7. Put thirty different numbers in a basket and select fifteen of them from it to participate in a study. Which sampling method was used to choose the fifteen participants?

8. Suppose students pick a card upon entering a classroom. The number they pick is the group they will work in for the day. What type of sampling is this?

9. Every fourth student that enters a classroom is the leader for their group. What type of sampling is this?

10. All the athletes at a school are chosen to take a test. What type of sampling is this?

Section 4.12 Non-Probability Sampling

Practice Problems 4.12

For Problem 1 and 2, choose your answer from the following sampling methods to best complete the problem:

Convenience Judgment Random Systematic

Voluntary-response Stratified Quota

1. Suppose you want three friends to come to your birthday party…

a) You choose the first three friends you see in the next few days to come to your birthday party.

b) You keep track of how many times each friend calls you in a week and then pull numbers from a hat to see if it matches the number of times each friend called you. You invite the matches to come to your birthday party.

c) You toss a coin and have your friends call *Heads* or *Tails*. You invite the first three that guess correctly to your birthday party.

d) You decide that all those with blonde hair can come to your birthday party because you have blonde hair.

e) You pick the three friends that have been your friends for the longest period of time to come to your birthday party because they probably know you the best.

2. Suppose you want to find out how reliable your public transit system is…

a) You get on a different bus at a different time each day and administer a survey to fellow passengers.

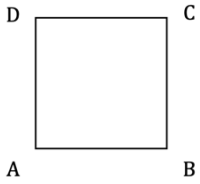
b) Your friend asks fellow passengers on your bus to grab a survey on their way out to fill out and mail back.

c) Your other friend stands outside your bus and surveys every third person who exits the bus.

d) The last friend who was at your birthday party says that each of you and your friends should ride a different bus and survey everyone on the bus.

For Problem 3 and 4, use the given information and diagram to solve the problem.

Geometric probabilities are based on area models. If you have a square, ABCD (shown to the right), and draw the diagonals from A to C and B to D, then you can label the point of intersection E.

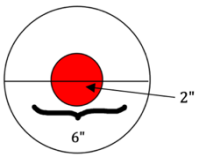


3. If you shade in Triangle AEB and spin a spinner centered at E, what is the probability it would land in the shaded region?

4. If you shade in Triangle CED as well and spin the spinner centered at E, what is the probability it will land in the non-shaded region?

For Problem 5 and 6, use the given diagram to solve the problem.

Suppose you have the dartboard shown to the right that has an inner circle with a diameter of two inches and a full diameter of six inches.



5. If you throw a dart at the board, what is the probability it will land in the red area?

6. If you throw a dart at the board, what is the probability it will land in the white area?

For Problem 7-10, solve the problem given.

7. Suppose a young lady decides to invite everyone that rides her bus to her birthday party. What type of sampling is this?

8. Why might judgment sampling be biased (favor certain outcomes)?

9. Why might convenience sampling be errored?

10. Do you think some of the sampling methods have more bias (favor certain outcomes) than others?

Section 4.13 Sampling Bias

Practice Problems 4.13

For Problem 1-4, identify the method or methods being used in the survey and identify any bias to solve the problem.

1. Suppose you want to conduct a survey about fast-food so you go to a fast-food restaurant and interview every fifth person that comes through the door between 11:00 AM and 2:00 PM. Which method or methods are you using? Can you identify any bias in your survey?

2. Suppose you are doing a survey about favorite fast-food restaurants so you send a survey to all of your friends and relatives asking them to fill it out. Which method or methods are you using? Can you identify any bias in your survey?

3. Telemarketers often call your home and ask you to take a five-minute survey. Which method or methods are they using? Can you identify any bias in their survey?

4. Suppose you shop online for some clothes and shoes. At the end of each purchase is an optional ten-question survey. If you fill it out, you are given a 20%-off coupon for your next purchase. Which method or methods are they using? Can you identify any bias in their survey?

For Problem 5-8, use the given information to solve the problem.

Just as bias can affect surveys, errors may affect problem-solving. Calculations may involve commonly-made mistakes.

5. Identify any common mistakes that could be made in the following calculations:

$3^5 = 243$ $3^4 = 81$ $3^3 = 27$ $3^2 = 9$

What is 3^1? What is 3^0?

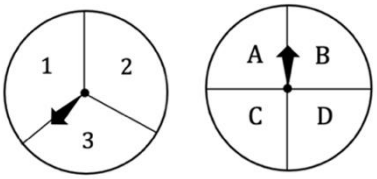
6. What is halfway between $\frac{3}{4}$ and $\frac{7}{16}$?

7. What is the next term in the following sequence?

1, 4, 9, 16, _____

8. Shana has taken four tests and scored a 75%, 89%, 94%, and 91% on them. Right now, she has an average of 87.25%. Shana wants to get an A in this class and has one more test left. What must she get on the fifth and final test to get an A in the class (an average of at least 90%)?

For Problem 9 and 10, use the spinners shown below to solve the problem.



First Spinner Second Spinner

9. What is the probability the first spinner will land on an even number *and* the second spinner will land on D?

10. What is the probability the first spinner will land on an even number *or* the second spinner will land on D?

Section 4.14 Module Review

For Problem 1-5, use the given information to solve the problem.

Below are the number of miles that a group of employees at a manufacturing plant commute daily to get to work. The new CEO wants to offer an incentive plan to encourage employees who live more than a certain number of miles away from the plant to be on time.

24, 14, 18, 22, 36, 41, 6, 44, 80, 13, 27, 43, 49

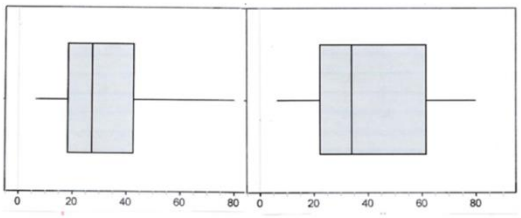
1. Find the range, median, mode, mean, maximum, and minimum of the data. As a consultant hired to solve the CEO's problem, you need to figure out the minimum number of miles an employee can commute to work and still deserve the incentive to be on time.

2. Is the data discrete or continuous?

3. Is the data numerical or categorical?

4. Which of the plots below represents the data? How do you know?

a) b)



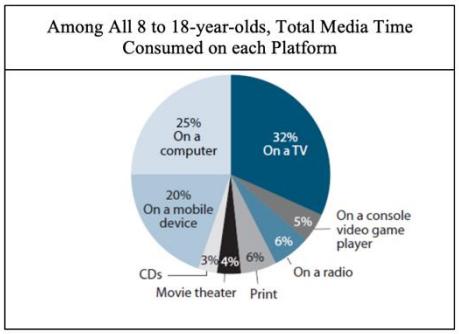
5. a) Are there any numbers that are outliers in the data? If so, what are they?

b) Is there more variability in the upper quartile or lower quartile of the data?

c) Which quartile includes those employees on the incentive plan?

For Problem 6, use the given pie chart to solve the problem.

6. A study was done by the Kaiser Family Foundation in 2011 called "Generation M^2: Media in the Lives of 8 to 18-year-olds." Below to the left is a pie chart of the results.



a) Which media is most often used by 8 to 18-year-olds?

b) Which media is least often used by 8 to 18-year-olds?

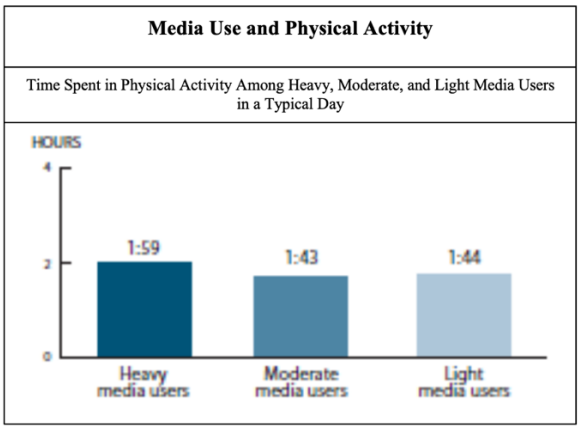
c) Why do you think CDs are used less than radio when so many teenagers listen to music?

d) What do all the percentages add up to?

e) What does the total percentage in d) represent?

For Problem 7, use the given bar graph to solve the problem.

7. According to the bar chart (below) from the Kaiser Family Foundation's "Generation M^2" survey, heavy media users spent more time doing physical activity than moderate media users or light media users.

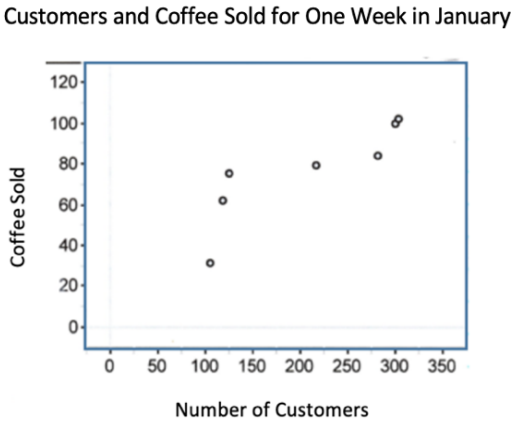


a) What might be the explanation for heavy media users spending more time in physical activity than moderate and light users?

b) About how much more time daily did heavy media users spend in physical activity than light media users?

c) How much less time did moderate media users spend in physical activity than heavy media users?

For Problem 8 and 9, use the given information and scatterplot to solve the problem.
Below to the left is a scatterplot of customers that come in to a restaurant and the amount of coffee sold at the restaurant in one week in January.



8. Which statement best describes the relationship between the number of customers and the coffee sold for this week in January?

a) As the number of customers decreases, the coffee sales decrease.

b) As the number of customers increases, the coffee sales increase.

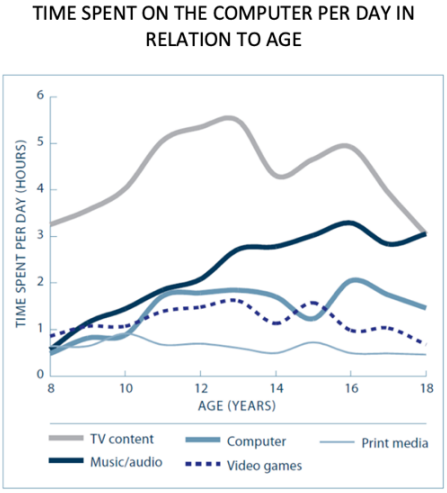
c) As the number of customers increases, the coffee sales decrease.

d) As the number of customers decreases, the coffee sales increase.

9. Which is the dependent variable, the number of customers or the amount of coffee sold?

For Problem 10, use the line graph to solve the problem.

10. The line graph below contains data from the "Generation M²" survey.



a) Which age group spends the most hours per day on the computer?

b) Which age group spends the least hours per day on the computer?

c) In general, which media do all age groups spend the least amount of time on of the five media platforms?

d) About how many hours do 14-year-olds spend playing video games daily?

For Problem 11 and 12, solve the problem given.

11. Suppose you want to conduct your own survey for media use among teenagers. You used the following methods:

Method 1: You call all your friends and ask them to come to your house at 3:00 PM to take your survey.

Method 2: You ask all the teenagers at your family reunion to take your survey.

Method 3: You give your survey to every third person that comes to Youth Group Bible Study on Wednesday night.

a) Which method uses convenience sampling?

b) Which method uses voluntary-response sampling?

c) Which method uses systematic sampling?

12. Cars go on sale in the spring. The more car-lots that have sales, the more cars they sell during the spring sale. Is there a strong-positive correlation, strong-negative correlation, weak-positive correlation, weak-negative correlation, or no correlation between spring sales and car sales?

Section 4.15 Module Test

For Problem 1-3, use the given information to solve the problem.

Sierra and Samantha want to try out for a bowling team. There are two spots open. The average scores of each of the eight bowlers currently on the team (including alternates) are listed below.

101, 114, 112, 100, 102, 103, 107, 108, 114

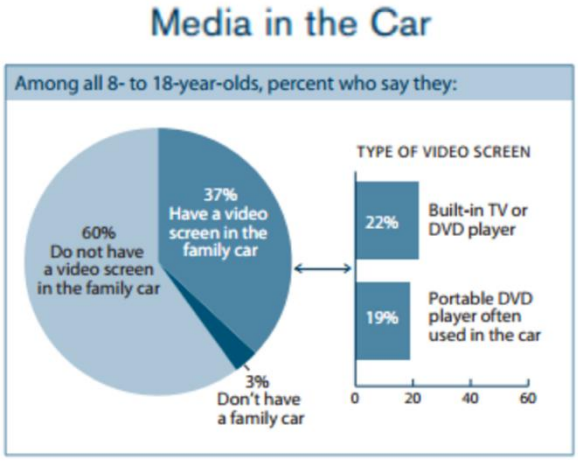
1. a) Find the range, median, mode, mean, maximum, and minimum of the data.

b) Does it seem likely the girls will make the bowling team if Sierra's average score is 106 and Samantha's average score is 102?

2. Make a Box-and-Whisker plot of the data. What is the median of the lower quartile? What is the median of the upper quartile?

3. Are there any outliers in the data? What can be said about the spread of data or variability in the scores?

For Problem 4 and 5, use the given pie chart to solve the problem.

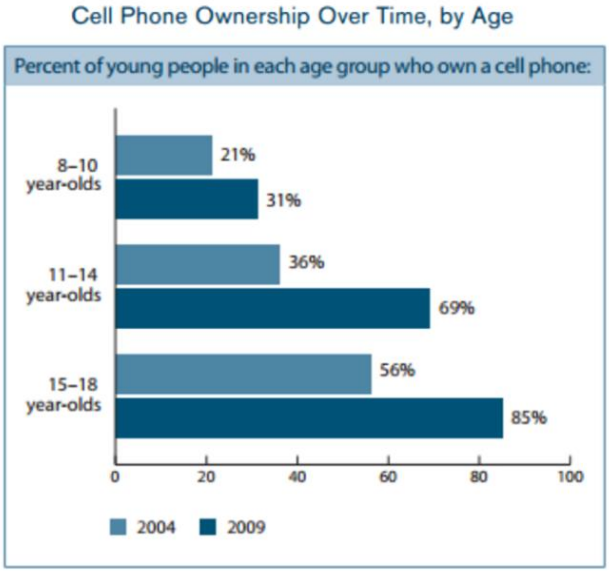


4. Based on the pie chart (from a survey conducted by the Kaiser Foundation in 2011), do most families have a video screen in their car? Why do the bars that represent type of video screen add up to less than 100%? How do you explain the 3% of teenagers that do not have a family car?

5. Is the media in the pie chart numerical or categorical?

For Problem 6, use the given bar graph to solve the problem.

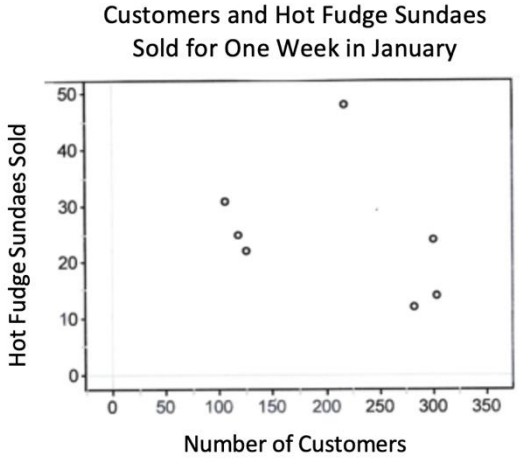
6.



a) Why do you think cell phone ownership increases with age?

b) Why are the dark blue bars always longer than the light blue bars?

For Problem 7 and 8, use the given scatterplot to solve the problem.



7. Which statement best describes the relationship between the number of customers and the hot fudge sundaes sold for this week in January?

a) There is a strong-positive correlation.

b) There is a weak-positive correlation.

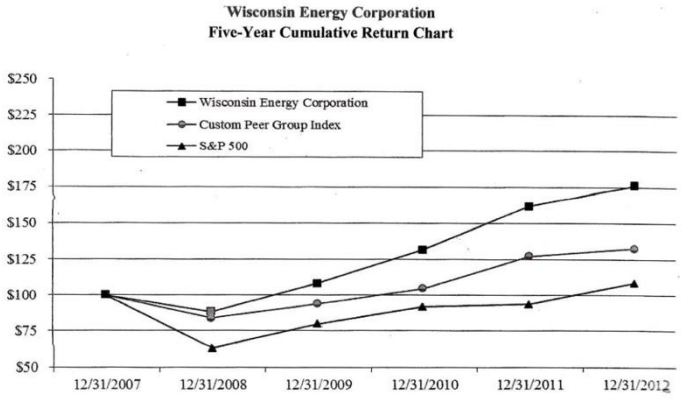
c) There is a strong-negative correlation.

d) There is a weak-negative correlation.

e) There is no correlation.

8. Which is the independent variable and dependent variable between the number of customers and the number of hot fudge sundaes sold?

For Problem 9, use the given line plot to solve the problem.



[Taken from WEC 2012 Annual Report Page 6]

a) Which would have been the best investment to make in 2008?

1) Wisconsin Energy Corporation

2) Custom Peer Group Index

3) S&P 500

b) How many years did S&P 500 fall below the initial investment of $100.00 per share?

c) Which stock has been most consistent in rising above the initial $100.00-per share investment?

10. Is the data in the Wisconsin Energy Corporation five-year summary discrete or continuous?

For Problem 11, classify each sampling method as probability sampling or non-probability sampling.

11.

a) Systematic sampling

b) Judgment sampling

c) Quota sampling

d) Random sampling

e) Convenience sampling

For Problem 12, use the data below of scores on a math test to solve the problem.

85%, 87%, 45%, 89%, 88%, 87%, 88%, 94%, 96%, 100%, 93%, 95%, 95%, 95%

a) Determine the 5-point summary for the data.

b) Make a Box-and-Whisker plot for the data.