Name _____

Date _____ Pd_____

INTRO TO STATISTICAL QUESTIONS AND DATA

DATA

- Data can be described as a collection of facts.
- For example: numbers, measurements, observations, and descriptions
  - **Categorical data:** this data represents characteristics and is often sorted into groups.
  - **Numerical data:** this data represents values that can be measured.

TASK 1 Read each statement to determine if it is referring to categorical or numerical data.

"I am collecting data to determine each student's favorite genre of book to read." LUKE

DEMARCUS
"I am collecting data to determine how many hours of sleep each student in our class gets."

JOSIE
"I am collecting data to determine how much money each student's family spends on groceries each week."

"I am collecting data on the color of cars that run the stoplight in front of our school." HEATHER

	IS THE DATA CATEGORICAL OR NUMERICAL? WHY?
LUKE	Categorical; the responses will be book genres
DEMARCUS	Numerical; the responses will be numbers of hours
JOSIE	Numerical; the responses will be amounts of money
HEATHER	Categorical; the responses will be colors of cars

Identify two numerical and two categorical data sets that could be observed about a standard deck of playing cards.

CATEGORICAL DATA: the different colored cards, the different suits of cards

NUMERICAL DATA: the number of face cards, the number of red or black cards

STATISTICAL QUESTIONS

- For a question to be considered statistical, it must meet two criteria:
  - the question must be able to be answered by data
  - the data must vary
- Ex: How many pets do the students in 6th grade at Jackson MS have?

TASK 2

Determine if each of the questions below are statistical or not. Rewrite any non-statistical questions to be statistical.

1. How many times a week do you go to football practice? Non-statistical Ex: How many times a week do the 6th grade students play a sport?	2. On average, how many hours of sleep do you get each night? Non-statistical Ex: On average, how many hours of sleep do the students in PE class get each night?	3. What is the maximum number of free throw shots students on the basketball team can make in a row? Statistical

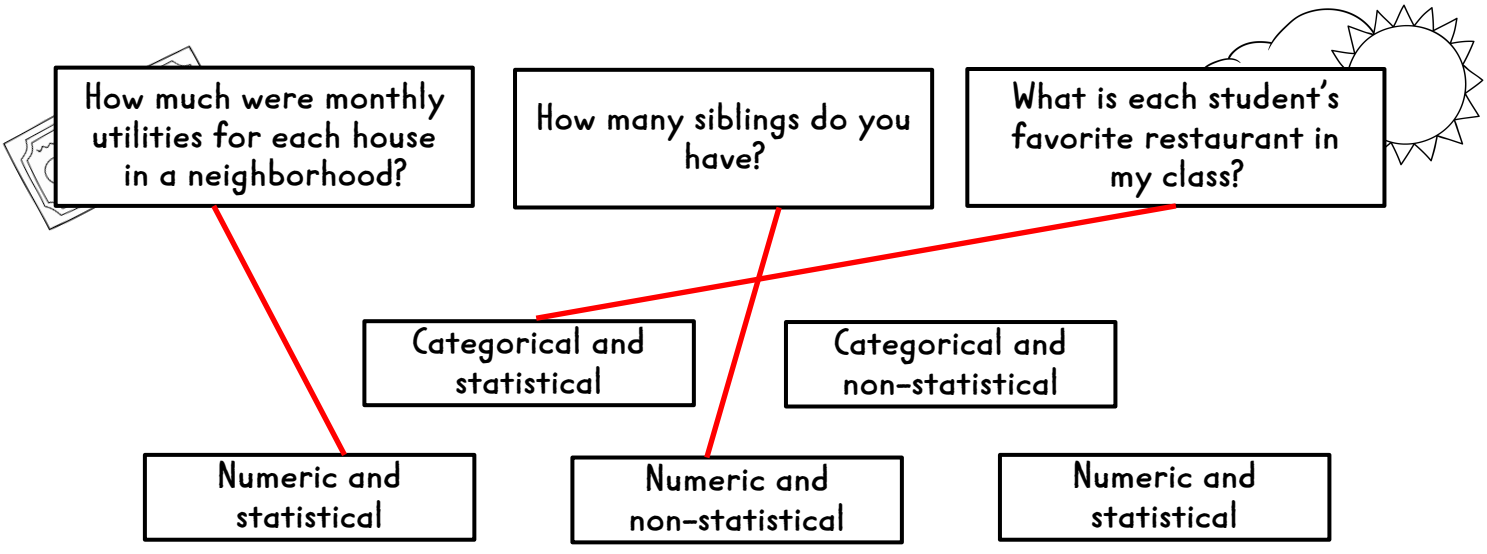
4. Create both a statistical and non-statistical question of your own below.

STATISTICAL: answers vary

NON-STATISTICAL: answers vary

TASK 3

Draw a line from each question to the best description of the question. Not all choices will be used.



Summarize today's lesson:

Name _____

Date _____ Pd _____

INTRO TO STATISTICAL QUESTIONS AND DATA

In 1-2, list whether the results of the survey would be categorical or numerical. Explain your reasoning.

SURVEY QUESTION	TYPE OF DATA AND EXPLANATION
1. How many pets do the students in my class have?	Numerical; the responses will be represented by a number of pets
2. What streaming service do the residents in my apartment complex subscribe to?	Categorical; the responses will be represented by the name of a streaming service

The students in Mr. Kurtz's class are asked to create nine different survey questions. List the letter of each question that is statistical in the table below. If it is not a statistical question, rewrite it to be statistical.

A How often do you play football?

B How many sit-ups can the students in PE class complete in 1 minute?

C What are the heart rates of the students in the 6th grade class?

D What are the types of sneakers students in our school wear?

E How many students tried out for the volleyball team?

F How many meters can the students on the swimming team swim?

G What is the max number of minutes that students on the track team can run?

H How many pairs of tennis shoes do the students in our school have?

I How many students in the 6th grade can do the long jump?

STATISTICAL	NON-STATISTICAL & REWRITE
B C D F G H	A; Example: How often do the 6th grade students play football?
	E; Example: How many students tried out for the different sports at our school?
	I; Example: How far can the 6th grade students jump in the long jump?

Name _____

Date _____ Pd_____

MEASURES OF CENTER

MEASURES OF CENTER

- A measure of center is a way of ___summarizing___ the data by providing a central point of data.

MEAN

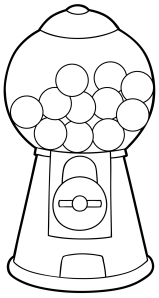
- The mean is a measure of center in which the ___average___ of the data is taken. It can be found by:

1. Finding the ___sum___ of all the data points
2. ___Dividing___ by the number of data points

- The mean can be thought of as determining the ___fair___ ___share___.

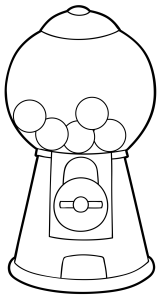
Determine the number of gumballs in each machine. Move the gumballs around to create an equal number of gumballs in each machine. What is the fair share?

MACHINE A



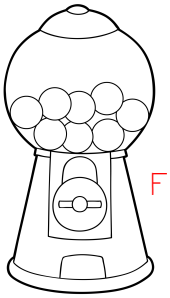
11

MACHINE B



5

MACHINE C



8

Fair share is 8

Use your understanding of mean to answer the questions below.

1. The list below shows the shoe size for eight students in Mr. Perdido's class.

13, 12, 9.5, 10.5, 12, 11, 12, 8

Sum: 88

Mean: 11

2. The following data set represents the age of people in line to vote.

27, 52, 60, 41, 33, 38, 42, 57, 72, 68

Sum: 490

Mean: 49

3. The following data set represents the number of times the phone rings when callers contact Linkhome Cable Customer Service.

NUMBER OF RINGS	4	5	6	7	8	9
FREQUENCY	2	1	2	0	2	1

List: 4, 4, 5, 6, 6, 8, 8, 9

Sum: 50

Mean: 6.25

MEDIAN

- The median is a measure of center in which the middle of the data is determined. It can be found by:
  1. Ordering the data from least to greatest
  2. Determining the middle of the data set
- If there is not one middle number, then find the average (mean) of the two middle numbers.

Determine the median of each of the data sets below.

4. The following data set represents the number of broken cookies in a test of nine boxes of Girl Scout Cookies.

2, 0, 2, 1, 1, 0, 3, 1, 1

Median: 1

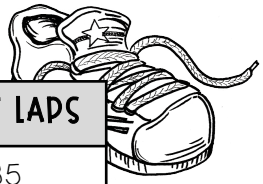
5. The following data set represents the number of inches of snowfall within a two-week period.

2, 1.5, 0, 0, 3, 1, 8, 6, 0, 0, 0, 1, 2, 3.5

Median: 1.25

Apply the mean and median of the data by answering the questions below.

6. Roanoke Middle School is holding a walk-a-thon, and the top five finishers are shown in the table below.



NAME	# OF LAPS
Brit	35
Margie	33
Jose	33
Amelia	32
Aldo	30

a. Mean: 32.6

b. Median: 33

c. Are the mean and median close to the same value or very different from one another? What does this tell you about the data?

Ex: The mean and median are very closely related because there is little variability between the data points.

7. Mrs. McConnell states that the numbers on the cards below have a mean of 12. What number is missing to make Mrs. McConnell correct?

7	16	12	10	11	?

16

Summarize today's lesson:

Name _____

Date _____Pd_____

MEASURES OF CENTER

Each of the cards on the left has the same solution as one of the cards on the right. Find the cards with matching solutions to complete the sentences below.

A

Calculate the mean of the data set below.

28, 40, 53, 39, 45

B

The data set below has a mean of 24. What would be the new mean if 17 was added to the list?

12, 30, 19, 27, 21, 35

C

The table represents the number of minutes it takes to walk to school. Find the mean.

# OF MINUTES	10	15	20	25
FREQUENCY	1	2	2	1

D

The cards below have a mean of 25. What is the missing number?

?	19	22	31

E

The table represents the number of miles to the nearest airport. Find the median.

# OF MILES	20	22	24	26
FREQUENCY	3	2	1	4

F

Calculate the median of the data set below.

15, 19, 24, 12, 9, 24, 19, 16, 21, 12

G

The data set below has a median of 39.5. What would be the new median if 43 was added to the list?

31, 41, 50, 28, 52, 38, 56, 27

H

Calculate the median of the data set below.

31, 20, 25, 34, 16, 33

1. Card A and Card G have the same solution of 41.
2. Card B and Card E have the same solution of 23.
3. Card C and Card F have the same solution of 17.5.
4. Card D and Card H have the same solution of 28.

Name _____

Date _____ Pd _____

MEASURES OF VARIABILITY

MEASURES OF VARIABILITY

- A measure of variability is a way of describing how spread out the data is. It can also be described as how much the data deviates from the center.

RANGE

- The range is a measure of variability that represents the spread in data. It can be found by:
  1. Determining the least and greatest values in the data set
  2. Subtracting the two values to determine the range in data

Use your understanding of range to answer the questions below.

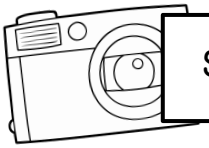
1. The list shows the number of pages in various novels: 286, 295, 307, 241, 396, 368 Range: 155	2. The Michaels family records their grocery bill each week. What is the range of the cost of their family grocery bill? $108.55, $86.20, $135.13, $176.97, $57.06 Range: $119.91
3. The Jonas family members read each evening and record their time. What is the range of the number of minutes read by the family? 120, 186, 62, 246, 98 Range: 184	4. The list shows the number of students in various college classes: 32, 19, 89, 102, 157, 25, 189, 36, 48 Range: 170

INTERQUARTILE RANGE

- The interquartile range (IQR) is a measure of variability that splits the data into four equal sized quartiles. It represents the middle 50% of the data. It can be found by:
  1. Ordering the data from least to greatest
  2. Finding the median
  3. Finding the median of the lower and upper quartiles
  4. Subtracting to find the IQR

Use your understanding of interquartile range to answer the questions below.

5. The data set below represents the different costs of cameras at an electronics store.



$28, $44, $108, $36, $59, $71, $66

a. Write the data from least to greatest.

28, 36, 44, 59, 66, 71, 108

b. What is the minimum price? 28

c. What is the maximum price? 108

d. What is the median price? 59

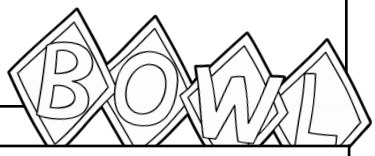
e. What is the median of the first half of the data? (first quartile) 36

f. What is the median of the second half of the data? (third quartile) 71

g. What is the interquartile range? 35

h. Is the IQR relatively small or large? What does this tell you about the spread of the data?

The IQR is $35 which means that the middle 50% of the data is within $35.



6. The data set below represents the scores Raven and her friends made while bowling.

105, 90, 110, 80, 115, 75, 90, 110, 100

a. Write the data from least to greatest.

75, 80, 90, 90, 100, 105, 110, 110, 115

b. What is the minimum score? 75

c. What is the maximum score? 115

d. What is the median score? 100

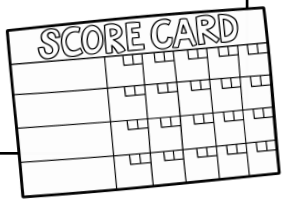
e. What is the median of the first half of the data (first quartile)? 85

f. What is the median of the second half of the data (third quartile)? 110

g. What is the interquartile range? 25

h. What does the interquartile range tell you about the variability of the data?

The IQR is 25 which means that the middle 50% of the data is spread out by 25 points.



Summarize today's lesson:

Name _____

Date _____ Pd _____

MEASURES OF VARIABILITY

Use your understanding of variability to answer the questions below.

1. Calculate the range of the data set.

6.5, 7.6, 9.1, 2.4, 8.8

6.7

2. Calculate the range of the data set.

33, 38, 35.5, 39.25, 31.75

7.5

3. Calculate the range of the data set.

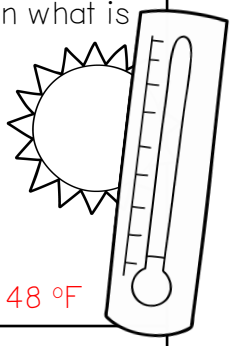
625, 638, 619, 677, 638, 659

58

4. When looking for a job, Micah is told that the monthly salary has a range of $876. He also knows that the starting salary is $1,793 per month. What is the maximum amount of money that Micah can earn in one month with this position?

$2,669

5. The range in temperature for the month of November in El Paso, Texas is 44°F. If the record high temperature is 92°F, then what is the record low temperature?



48 °F

6. The data set below represents the different costs of refrigerators at a local home improvement store.

$777, $498, $619, $379, $895, $1,256, $1,052

a. What is the median of the first half of the data? (first quartile) 498

b. What is the median of the second half of the data? (third quartile) 1,052

c. What is the interquartile range? 554

7. The data set represents the number of students in each elementary school in a large city. Louis found the interquartile range to be 175. Do you agree or disagree? Why or why not?

518, 579, 452, 537, 428, 603, 496

Louis is incorrect. Instead of finding the interquartile range, he found the range. The interquartile range is 127.

Note: In step 2 below, remind students that distance will always be positive.

MEAN ABSOLUTE DEVIATION

MEAN ABSOLUTE DEVIATION

- One way to describe the variability, or how spread out a set of data is, is by using mean absolute deviation.
- Mean absolute deviation is the average distance that the data points in a set of data are from the mean.
- A M.A.D. close to 0 means that the data values are close to the mean.

What does deviation, or "deviate" mean? How can it help you remember the definition of "mean absolute deviation"?

Sample: Deviate means to depart from, or go away. This can help me remember that mean absolute deviation is the average distance that points are "away" from the mean.

In order to find the mean absolute deviation of a set of data, follow the steps described in the table. Complete the table and apply the steps to find the mean absolute deviation in the situation below.

STEPS TO CALCULATE M.A.D.

1. Find the mean of the data set
2. Find the distance from each data point and the mean
3. Find the average of those distances

1. A flight attendant recorded the number of minutes it took to board the airplane for the last seven flights.

20, 19, 15, 23, 14, 17, 18

# OF MINUTES	DISTANCE FROM MEAN
20	2
19	1
15	3
23	5
14	4
17	1
18	0
TOTAL DEVIATION	16

a. Find the mean of the data.

18

b. List the distance between each data point and the mean in the table at the left.

c. Find the average of the distances, or the mean absolute deviation, to the nearest tenth.

2.3

d. Explain what the mean absolute deviation represents in the situation.

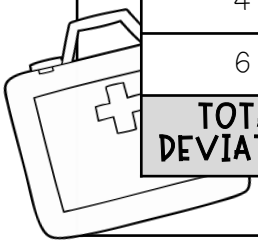
The average distance that each boarding time is from the mean of 18 is 2.3 minutes.

Practice finding mean absolute deviation, and round to the nearest tenth when necessary.

2. The number of people in hospital waiting rooms across the city are shown in the list at the right.

3, 5, 8, 10, 4, 6

NUMBER OF PEOPLE	DISTANCE FROM MEAN
3	3
5	1
8	2
10	4
4	2
6	0
TOTAL DEVIATION	12



a. Find the mean of the data.

6

b. List the distance between each data point and the mean in the table at the left.

c. Find the average of the distances, or the mean absolute deviation, to the nearest tenth.

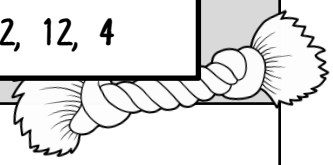
2

d. Explain what the mean absolute deviation represents in the situation.

The average distance that each waiting room number is from the mean of 6 is 2.

3. The numbers at the right represent the number of dogs at the dog park Mariella saw each day last week.

10, 7, 2, 12, 4



NUMBER OF DOGS	DISTANCE FROM MEAN
10	3
7	0
2	5
12	5
4	3
TOTAL DEVIATION	16

a. Find the mean of the data.

7

b. List the distance between each data point and the mean in the table at the left.

c. Find the average of the distances, or the mean absolute deviation, to the nearest tenth.

3.2

d. Explain what the mean absolute deviation represents in the situation.

The average distance that each number of dogs is from the mean of 7 is 3.2.

4. Coach Jackson found the mean absolute deviation for the number of minutes it took his PE class to run a mile. The mean absolute deviation was a very large value. Jaden assumes this means his class took a long time to run the mile. Do you agree with Jaden's conclusion? Explain why or why not.

No; mean absolute deviation only tells us how spread out the data is, not about the actual times themselves. Jaden's class had times that were very spread out.

Name _____

Date _____ Pd _____

MEAN ABSOLUTE DEVIATION

Use your understanding of variability to answer the questions below.

1. Find the mean absolute deviation for the set of values.

11, 16, 70, 66, 14, 15

24

2. Find the mean absolute deviation for the set of values.

68, 70, 64, 62

3

3. The speed of the last six semi-trucks is recorded in the list below. Use the information to answer the questions below.

61, 55, 66, 58, 50, 70

a. What is the mean?

60

b. What is the mean absolute deviation?

$5.\overline{6}$

SPEED	61	55	66	58	50	70
DEVIATION	1	5	6	2	10	10

c. Explain what the mean absolute deviation represents in the situation.

The average distance that each semi truck's speed is from the mean of 60 is $5.\overline{6}$.

4. The average monthly temperatures in New Orleans, LA has a mean absolute deviation of 43.5°F. What conclusion can you make about the average monthly temperatures in New Orleans, LA?

Ex: The average monthly temperatures in New Orleans, LA vary greatly from the mean.

5. If the mean absolute deviation is close to 0, then what does that mean about the data set?

Ex: A mean absolute deviation close to 0 means that the data set is very close to the mean. There is very little spread in the data.

6. The data set below represents the number of free throws the Tiger basketball team made in their last 8 games. Circle the names of the students that made a correct statement about the data.

8, 4, 16, 10, 5, 11, 15, 3

EVELYN: The mean of the data is 9.

NOVA: The total deviation of each data point from the mean is 32.

CARTER: The mean absolute deviation is 3.6.

Name _____

Date _____ Pd _____

DOT PLOTS

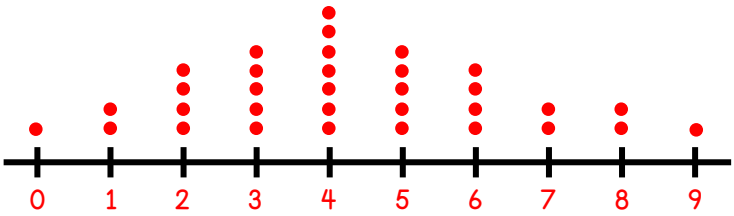
DOT PLOT

- A dot plot is a ___visual___ display of data using a number line and dots to represent each data point. The data that repeats itself most often is the mode. In a dot plot, the mode is the ___tallest___ value.

A survey in the 6th grade class asked students to record the number of devices they had in their home. The information is recorded in the table below.

# OF DEVICES	0	1	2	3	4	5	6	7	8	9
FREQUENCY	1	2	4	5	7	5	4	2	2	1

1. Use the data in the table to make a dot plot. Then answer the questions that follow.



a. What do you notice about the shape of the dot plot? List any other observations in the space below.

Ex: The data is evenly balanced.

b. Where would you say that most of the data lies on the dot plot? What is the median of the data?

Ex: Most of the data lies in the middle of the dot plot. The median number of devices is 4.

SPREAD

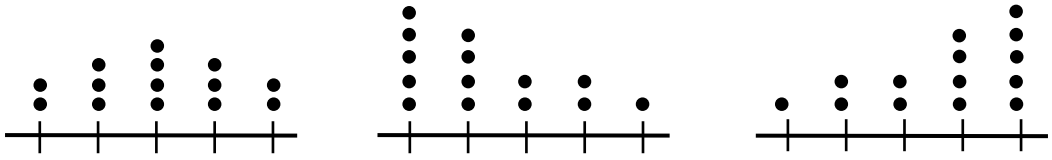
- The variability in the data points describes how far apart the data is from one another. This can also be represented by the ___range___ .

CENTER

- The median and the mean both represent the center of the data.
  - When the data is skewed, then the ___median___ is the best representation of the data.
  - When the data is symmetric, then the ___mean___ is the best representation of the data.

SHAPE

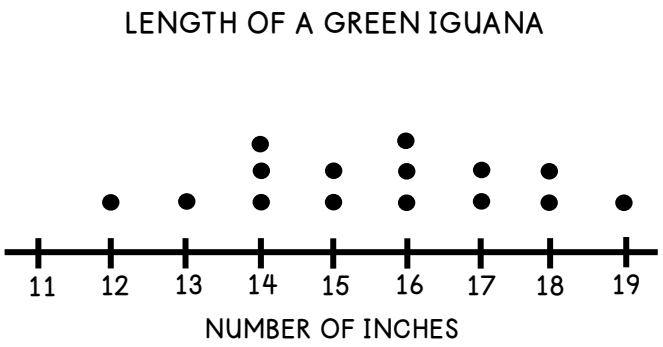
- Data can take on three different shapes:



symmetrical | skewed right | skewed left

Use your understanding of dot plots and statistical data to answer the questions below.

2. The ecological society sampled the green iguana population and made note of the length of each iguana sampled. The data is displayed in the dot plot below.



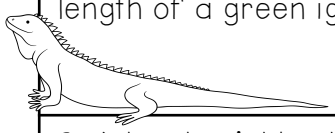
a. What is the mean? 15.6

b. What is the range? 7

c. What is the median? 16

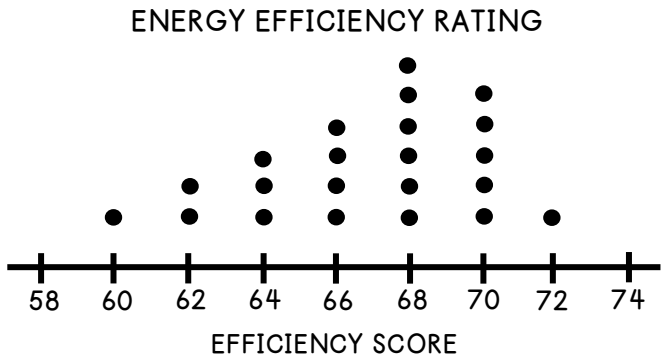
d. What is the interquartile range? 3

e. Based on the shape of the data, which measure of center is the best representation of the length of a green iguana?



Ex: Both the mean and the median are close, and the data is symmetrical. The mean is the best representation.

3. A local neighborhood is seeking to become certified in energy efficiency. The efficiency scores of each home are displayed in the dot plot below.



a. The peak score is 68.

b. The shape of the data distribution is skewed left.

c. The range of scores is 12.

d. Label each of the following statements as true or false.

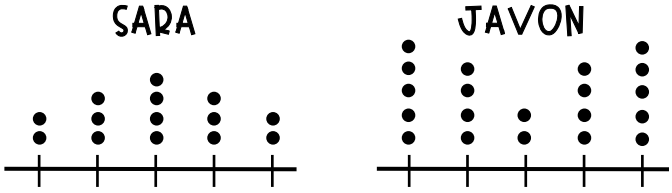
false A total of 20 homes were rated.

false Exactly half of the ratings were greater than 66.

false The most common home rating was 66.

4. Mr. Tips asked his students to sketch a dot plot that was symmetrical and had data that was close together. Which student(s) completed the task correctly?

a. Sara only
b. Javon only
c. Both Sara and Javon
d. Neither Sara nor Javon



Summarize today's lesson:

Name _____

Date _____Pd_____

DOT PLOTS

Four students were given dot plots to create. Sketch the dot plots in questions 1-4. Then use the clues in A-D to determine each student's data set.

A Sara's data set has a range of 5.

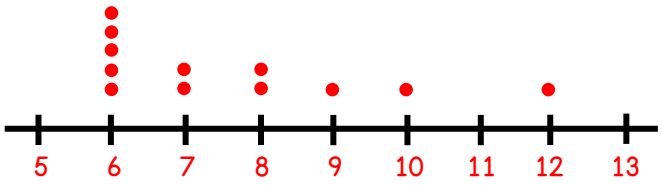
B Julius's data set is skewed right.

C Elisa's data set has a median of 10.5.

D Inez's data set is symmetrical.

1. The number of minutes a customer waits in line at the grocery curbside pickup:

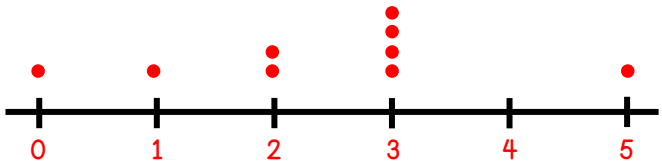
6, 9, 12, 7, 6, 10, 8, 8, 6, 6, 7, 6



Name: Julius

2. The number of miles Mrs. Estrada walks in a week:

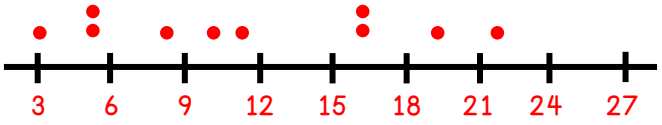
1, 2, 0, 5, 3, 3, 3



Name: Sara

3. The number of elementary schools in ten regional cities:

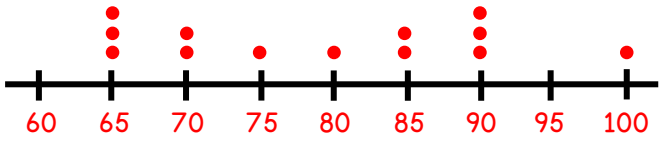
5, 8, 10, 22, 16, 16, 3, 11, 19, 5



Name: Elisa

4. The number of minutes spent on homework over a period of time:

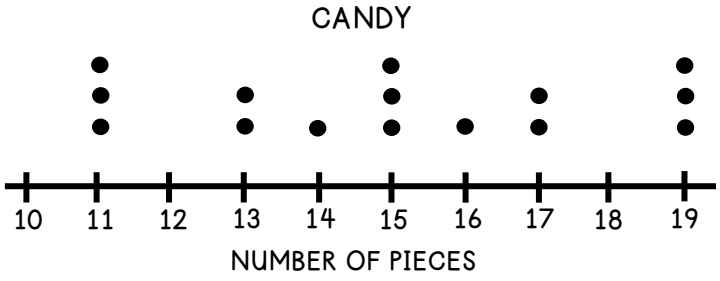
65, 80, 90, 65, 70, 75, 100, 90, 65, 70, 85, 85, 90



Name: Inez

Use your knowledge of dot plots to answer question 5.

5. A large bag of individual candy packs is opened. The amount of candy in each pack is counted and marked on the dot plot below. Which of the statements is not represented by the data?



A. Each bag of has 11-19 pieces of candy.

B. There is a total of 15 pieces of data.

C. The data is skewed right.

D. More than half of the bags had less than 16 pieces of candy.

49

Name _____

Date _____ Pd _____

QUIZ : DATA AND STATISTICS

Answer the questions below. Be sure to show your work.

Answers	
1.	A
2.	C
3.	A
4.	C
5.	D
6.	$9.25
7.	7.5
8.	B
9.	19
10.	C

1. Which of the following best describes the survey question,
"How many pairs of tennis shoes do the students in our school have?"

A. Statistical question, numeric data
B. Statistical question, categorical data
C. Non-statistical question, numeric data
D. Non-statistical question, categorical data

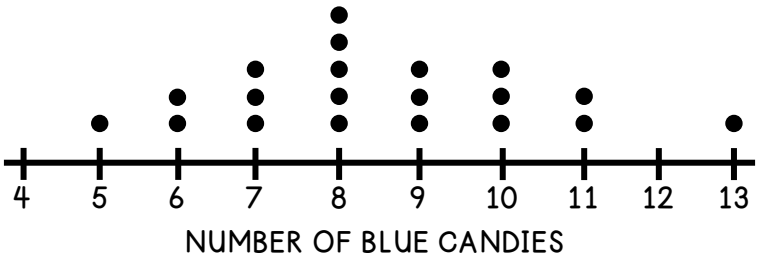
2. Which of the following best describes the survey question,
"How many days do you play football during the week?"

A. Statistical question, numeric data
B. Statistical question, categorical data
C. Non-statistical question, numeric data
D. Non-statistical question, categorical data

In questions 3-5, several small packages of colored candies were opened, and the number of blue candies were recorded on the dot plot below.

3. How many bags of candies were sampled?

A. 20
B. 22
C. 25
D. 30



4. Which of the following statements is a true statement about the number of blue candies?

A. Twenty-five packages of candies were sampled.
B. The data is skewed right.
C. The most common number of blue candies in a package is 8.
D. The median is 8.5.

5. Which of the following statements describes the general shape of the data?

A. The data is skewed left.
B. The data is skewed right.
C. The data is spread out.
D. The data is symmetrical.

Answer the questions below. Be sure to show work and justify your thinking.

6. The list below shows the allowance earned by eight of the students in Ms. Tomas's class. What is the mean allowance earned in Ms. Tomas's class?

$7, $12, $8.50, $10, $7, $8, $10.50, $11

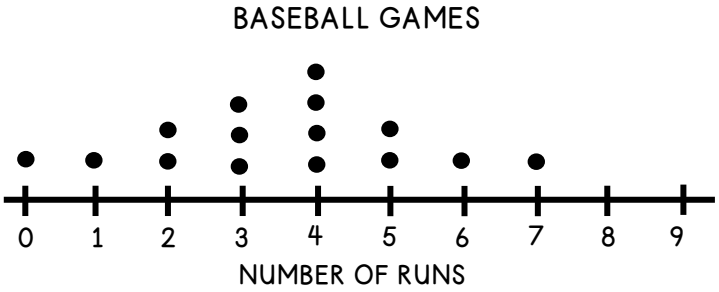
_____

7. What is the range of the data below?

33, 38, 35.5, 39.25, 31.75

_____

8. The data below represents the number of runs scored in the last 15 games. Which statement best describes the data?



A. The peak of the data is 3.

B. The data distribution is symmetrical.

C. The data distribution has several gaps.

D. The range of the data is 6.

9. The table represents the number of songs on various playlists. Find the median.

# OF SONGS	18	22	20	16
FREQUENCY	2	1	3	2

10. The mean absolute deviation of the height of Greyson and his five cousins is 0.8. Which of the following statements is true?

A. Greyson and his cousins must all be short, since the mean absolute deviation is such a small number.

B. The heights of Greyson and his cousins vary greatly from the mean.

C. Greyson and his cousins are all around the same height because the mean absolute deviation is a small number.

D. There is not enough information to tell whether or not Grayson and his cousins are similar in height or not.

Name _____

Date _____ Pd _____

HISTOGRAMS

HISTOGRAMS

- Histograms show data in ___intervals___ or ranges.
- They are similar to bar graphs, but the bars are not spaced apart. They are similar to dot plots because they represent the overall ___spread___ and ___shape___ of the data.

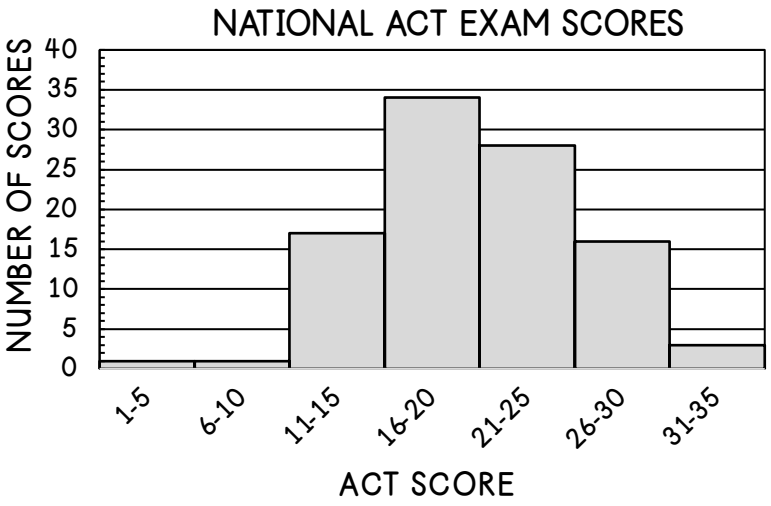
Ex: ___decades, age ranges___

Data from the National ACT Exam is shown in the histogram at the right.

a. How many scores were included in the data set? 100

b. List the number of scores for each range:

1-5:	1	21-25:	28
6-10:	1	26-30:	16
11-15:	17	31-35:	3
16-20:	34		



1. Which range of scores represents 17% of the data?

ACT scores 11-15

2. What percent of the scores are included in the 21-25 range?

28%

3. What two ACT score ranges make up more than 60% of the sample?

ACT score ranges 16-20 & 21-25

4. The histogram is symmetrical. What does this tell you about the median and the mean? In which interval would you predict the mean to be found?

The median and mean are close together. The mean is likely in the 16-20 range.

5. Label each of the statements below as true or false. Correct any false statements.

true a. The number of scores between 1-5 and 26-30 is equal to the number of scores between 11-15.

true b. The number of scores between 16-20 is double that of 11-15.

false c. There are an equal number of scores between 11-15 and 26-30. There are an equal number of scores between 1-5 and 6-10.

The local library surveys its patrons based on their age. Use the histogram below to answer questions 6-9.

6. Determine the total number of library patrons surveyed.

50 patrons

7. How many library patrons are younger than 30?

23 patrons

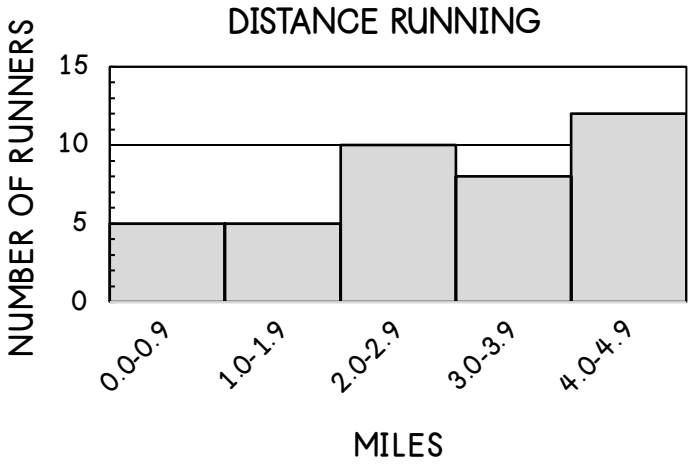
8. Based on the information in the graph, which two age ranges are responsible for exactly 40% of the patrons?

10-19 and 60-69

9. Based on the information in the graph, 30% of the library patrons are in which age range?

20-29

A group of runners chart their distance total for the week. Use the histogram below to answer questions 10-12.



10. How many runners run between 3.0-4.9 miles each week?

20 runners

11. How many runners run less than 4.0 miles each week?

28 runners

12. Which of the following statements best represents the data above?

A. Twenty-five percent of runners surveyed run 4.0-4.9 miles each week.

B. One-third of runners surveyed run 2.0-2.9 miles each week.

C. One-fifth of runners run 3.0-3.9 miles each week.

D. The number of runners surveyed who run less than 1.0 mile is equal to the number of runners surveyed who run 3.0-3.9 miles each week.

Summarize today's lesson:

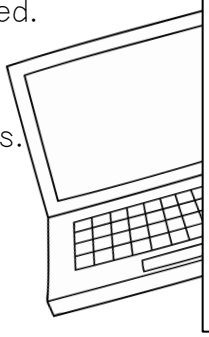
Name _____

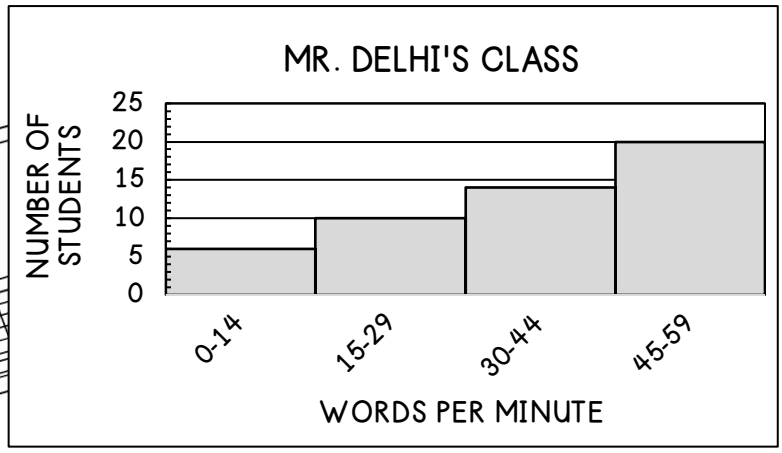
Date _____ Pd_____

HISTOGRAMS

In computer class, students take a typing test to determine their speed.

The histogram at the right shows the scores of the students in class.





1. Label each of the statements below as true or false. Correct any false statements.

false a. A total of 40 students can type from 0-59 words per minute.

A total of 50 students can type from 0-59 words per minute.

false b. The number of students who typed 45-59 words per minute is equal to the number of students who typed 0-29.

The number of students who typed 45-59 wpm was 20, while only 16 students typed 0-29 wpm.

true c. The number of students who typed 30-44 words per minute accounted for 28% of the students.

true d. Less than 15% of the students type 0-14 words per minute.

2. Based on the information above, what percent of students type less than 30 words per minute?

32%

3. Forty percent of the students were able to type at which speed interval?

45-59

4. Describe the distribution of the data. Estimate in which range the median of the data lies.

The data is skewed left. The mean and median will be towards the right of the middle part of the data.

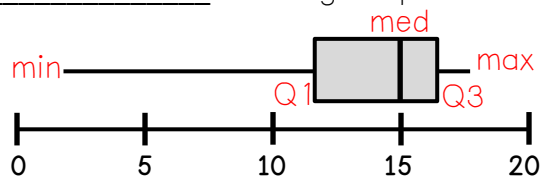
Name _____

Date _____ Pd _____

BOX PLOTS

BOX PLOTS

- A box plot displays data distribution using ___five___ key numbers. The difference between the first and third quartile is called the ___interquartile range___.
  - ___minimum___: the smallest piece of data
  - ___first quartile (Q1)___: the median of the lower half of data
  - ___median___: the median (midpoint) of the data
  - ___third quartile (Q3)___: the median of the upper half of data
  - ___maximum___: the largest piece of data

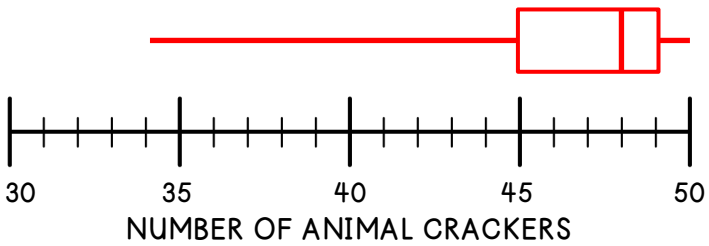


Use the data below to create a five-number summary and sketch a box plot.

1. The following data set represents the number of animal crackers in a snack-size box.

34, 50, 49, 47, 48, 45, 48

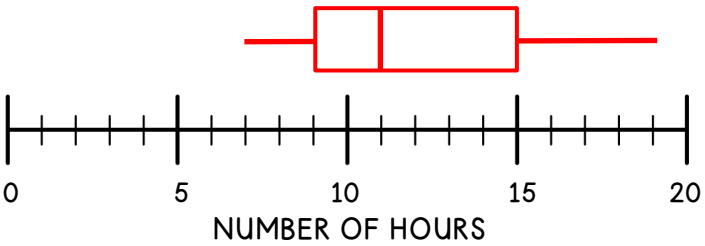
Min: 34
Q1: 45
Med: 48
Q3: 49
Max: 50



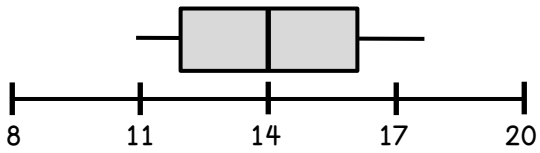
2. The following data set represents the number of hours a small candle will burn.

9, 7, 11, 16, 11, 19, 9, 10, 15, 14, 8, 12, 15

Min: 7
Q1: 9
Med: 11
Q3: 15
Max: 19



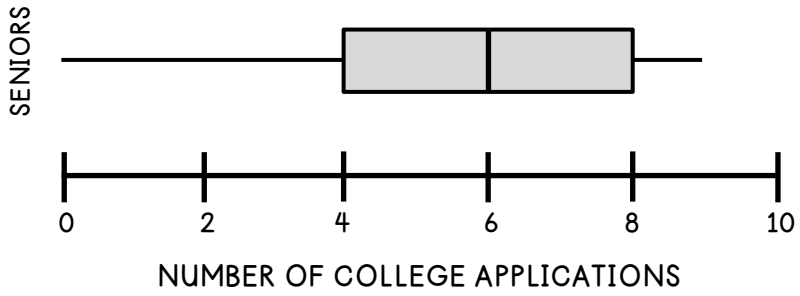
3. Use the given box plot to determine the five-number summary.



Min: 11 Q1: 12 Med: 14 Q3: 16 Max: 18

Use your knowledge of box plots to answer the questions below.

Christina conducts a poll to determine the number of colleges high school seniors applied to. She creates a box plot to represent the information.



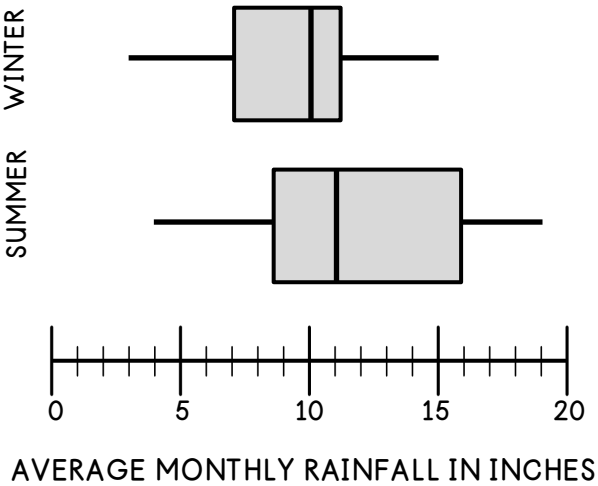
4. Determine the range of the number of colleges. 9 colleges	5. What is the median number of colleges applied to? 6 colleges
6. What is the interquartile range of the number of college applications? 4 colleges	7. Which quartile represents the greatest spread in data? first quartile

8. Describe why quartile 1 is larger than quartile 4. What do you observe about quartile 2 and 3?

Ex: Quartile 1 is larger because the lowest 25% of the data is more spread out than the highest 25% of the data. Quartiles 2 and 3 are equal in size, so the data points are equally distributed.

Use your knowledge of box plots to determine if the statements below are true or false.

9. The average monthly rainfall during the summer months and winter months are shown below.



true The range in rainfall in the winter was 12 inches.

false The median rainfall in the summer was 13 inches.

true For half of the months in the summer, the rainfall was greater than 11 inches.

true For half of the months in the winter, the rainfall was less than 10 inches.

Summarize today's lesson:

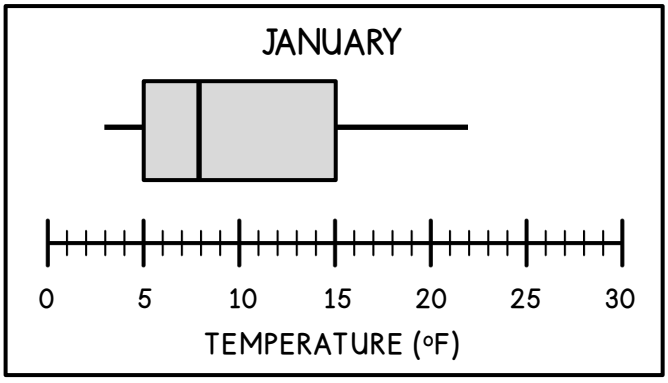
Name _____

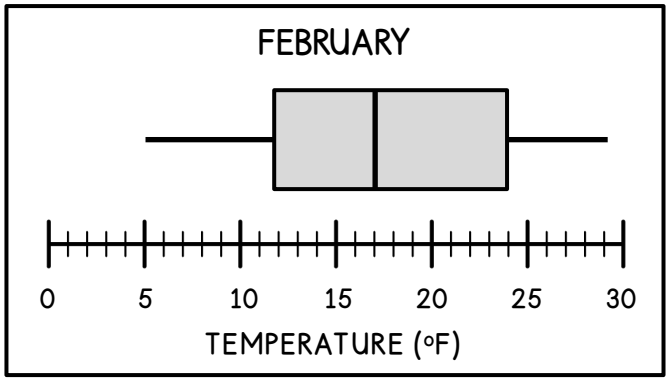
Date _____ Pd _____

BOX PLOTS

Use your understanding of box plots to answer the questions below.

1. The local meteorologist plots the daily low temperatures for the months of January and February on the box plots below. Use the box plots to mark each statement as true or false, and correct any false statements.





a. False The range of daily low temperatures in January was 15°F.
The range of daily low temperatures in January was 19°F

b. True In February, about 50% of the daily low temperatures were 17°F or higher.

c. True Less than 25% of the daily low temperatures in February were 25°F or higher.

d. False In February, the IQR was 20°F.
In February, the IQR was 12°F.

e. False The median daily low temperature in January was 15°F.
The median daily low temperature in January was 8°F.

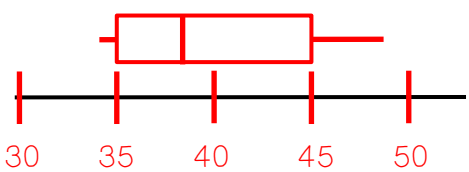
f. True Three-fourths of the daily low temperatures in January were above 5°F.

Using the data given below create a five-number summary and a box plot.

2. The following data set represents the number of fish crackers in a snack-size box.

34, 35, 37, 38, 42, 45, 49

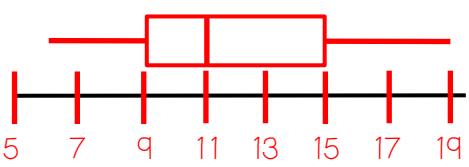
Min: 34
Q1: 35
Med: 38
Q3: 45
Max: 49



3. The following data set represents the number of hours various homes run their air conditioning in one day.

9, 6, 11, 16, 11, 19, 9, 10, 15, 14, 8, 12, 15

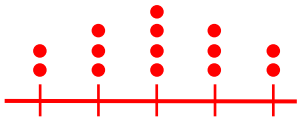
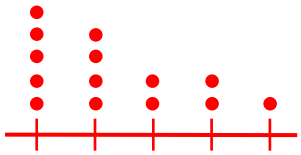
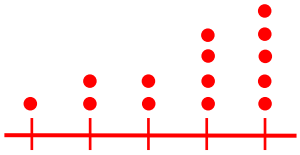
Min: 6
Q1: 9
Med: 11
Q3: 15
Max: 19



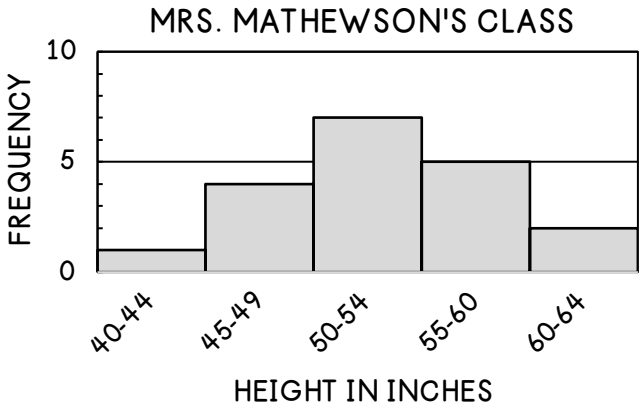
Name _____

Date _____ Pd _____

APPLYING CENTER AND VARIABILITY

	SYMMETRIC DATA	SKEWED DATA
SHAPE		 skewed right  skewed left
CENTER	mean - shows where the center of the data is, but not always a true representation of the data	median - midpoint of the data - might not be a data value - does not have to be in the middle of the range of values
VARIABILITY	MAD - a way to measure the spread of the data in relation to the mean - MAD close to 0 = data close to the mean	IQR - a way to measure the spread of data around the median - small IQR = data close to the median
GRAPH	histogram dot plot	box plot

1. The school nurse measured the height of each student in Mrs. Mathewson's class. She then created a histogram to display the students' height.



a. Describe the shape of the data.

symmetrical

b. What is the best measure of center?

mean

c. What is the best measure of variability?

mean absolute deviation

d. Javier says the mean of the data will fall between 50-54 inches. Do you agree or disagree?

Agree, since the data is symmetrical, the mean will most likely fall within the center of the data.

Practice comparing data using the information below.

2. Two studies were conducted at different banks. The summary measures how long a customer had to wait in line.

	MEAN	MAD
STATE BANK	8	6
NATIONAL BANK	8	1

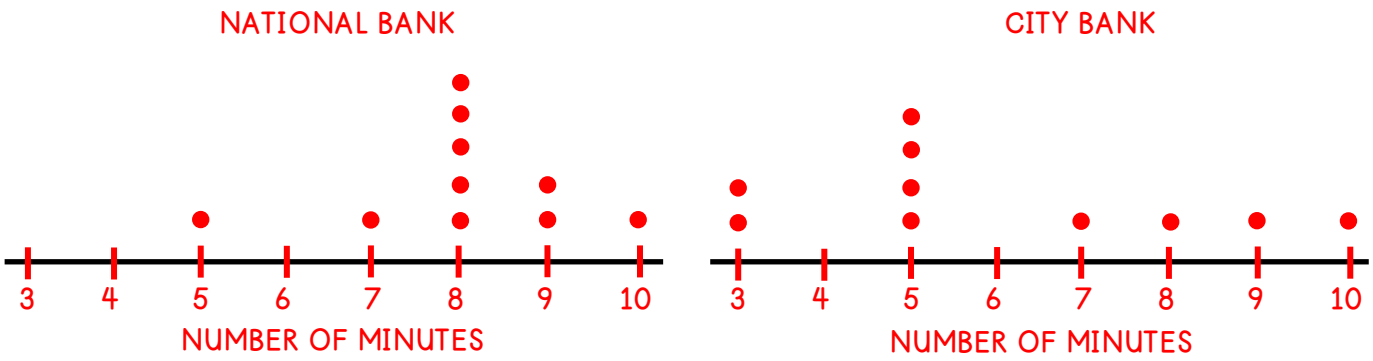
Describe the meaning of the data. At which bank would you rather stand in line? Why?

Ex: If I wanted consistency, I would rather stand in line at National Bank because I will likely wait between 7-9 minutes. At State Bank, the wait time varies from 2-14 minutes.

3. City Bank, a new bank in town, advertises that they have quick lines to satisfy busy customers. The data shown below compares the wait time for 10 different customers at both National Bank and City Bank.

	1	2	3	4	5	6	7	8	9	10
NATIONAL BANK	5	7	8	8	8	8	8	9	9	10
CITY BANK	3	3	5	5	5	5	7	8	9	10

a. Sketch a dot plot with the same scale to compare the wait time at both banks.



b. Calculate the mean number of minutes in line. Does City Bank live up to their advertisements?

The mean for National Bank is 8 minutes, while the mean for City Bank is 6 minutes.
Yes, on average City Bank customers wait 2 minutes less than National Bank customers.

c. If you wanted a consistent wait time, then which bank should you use? Why?

National Bank is more consistent with a MAD of 0.8 and a range of 5 minutes, while City Bank varies greatly with a MAD of 2 minutes and a range of 7 minutes.

Summarize today's lesson:

Name _____

Date _____ Pd _____

APPLYING CENTER AND VARIABILITY

An experiment is conducted in three different classes, where 10 students were given some fish crackers. The number of crackers that each student received is shown in the table below.

	1	2	3	4	5	6	7	8	9	10
CLASS 1	16	25	14	30	28	11	26	30	15	21
CLASS 2	24	26	27	24	25	24	26	26	27	24
CLASS 3	18	16	21	21	22	18	17	17	18	20

1. Determine the median of each class.

Class 1: 23

Class 2: 25.5

Class 3: 18

2. Determine the IQR of each class.

Class 1: 13

Class 2: 2

Class 3: 4

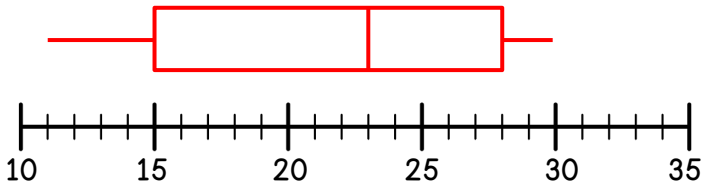
3. Describe how the results from Class 1, Class 2, and Class 3 differ.

Example: Classes 2 and 3 have results that are fairly close to the median. Class 1 has data that is very spread out with an IQR of 13.

4. If you were super hungry, then which class would be the best one to be in? Why?

Example: Class 2, because you would likely get around 25 crackers. Class 1 has a greater variability, so while you could get more crackers, you could also get far less.

5. Use the information from Class 1 to create a box plot representing the number of crackers that each student received.



Name _____

Date _____ Pd _____

DATA & STATISTICS UNIT STUDY GUIDE

Solve each of the problems below. These represent the types of questions on your test. Be sure to ask questions if you need more help with a topic.

I CAN RECOGNIZE A STATISTICAL QUESTION.

1. Determine if the survey question will result in categorical or numeric data. Then, determine if the question is statistical or not. If it is not, then rewrite the question to make it statistical.

a. What is each student's favorite restaurant in my school?

Categorical and statistical

b. What size shoes do you wear?

Numeric and non-statistic. What size shoes do the people in your family wear?

c. How many books have the students in my class read this month?

Numeric and statistical

d. How fast can the sixth graders run one mile?

Numeric and statistical

e. How many times do you go to PE each week?

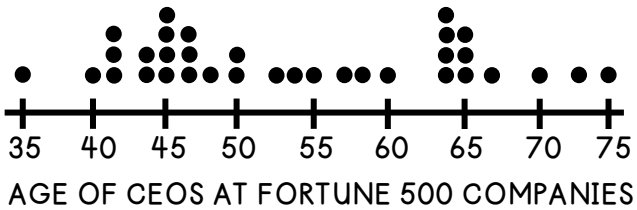
Numeric and non-statistical. How many times do students in different schools go to PE each week?

f. What is your favorite type of food?

Categorical and non-statistical. What are my friends' favorite types of food?

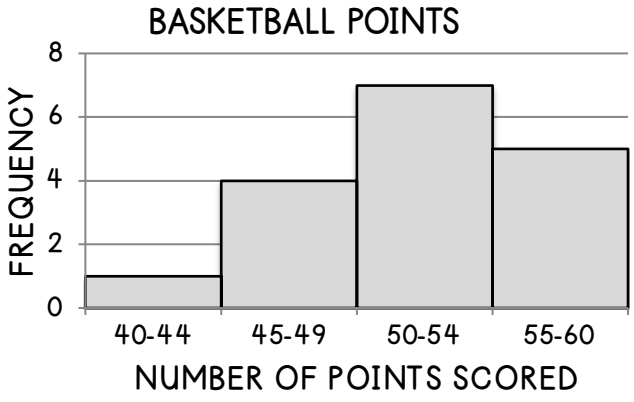
I CAN DESCRIBE A DATA DISTRIBUTION.

2. Describe the overall shape, spread, and center of the dot plot.



The data is not symmetrical and spread from 35 to 75. There are two centers around 45 and 65.

3. Describe the overall shape, spread, and center of the histogram.



The data is symmetrical and spread from 40-60. The center falls within 50-54.

I CAN FIND THE MEASURE OF CENTER.

4. The following data set represents the average high monthly temperature in Denver, CO.

44, 46, 54, 61, 71, 81, 88, 86, 77, 65, 52, 43

Mean: 64 Median: 63

5. The following data set represents the number of dollars 20 customers withdrew at an ATM.

NUMBER OF DOLLARS	20	40	60	80	100	120
FREQUENCY	10	5	2	1	2	0

Mean: 40 Median: 30

I CAN FIND THE MEASURE OF VARIABILITY.

6. The following data set represents the number of days Emma played soccer each week for the past 5 weeks.

5, 6, 5, 4, 5

NUMBER OF DAYS	DISTANCE FROM MEAN
5	
6	
5	
4	
5	
TOTAL DEVIATION	

a. Mean: 5 MAD: 0.4

b. What does the MAD represent in the context of the situation?

The average distance that each number of days is from the mean of 5 is 0.4.

7. The following data set represents the number of students who attended Monday swim classes.

8, 6, 4, 3, 8, 9, 6, 5, 2, 8

Q1: 4 Q3: 8 IQR: 4

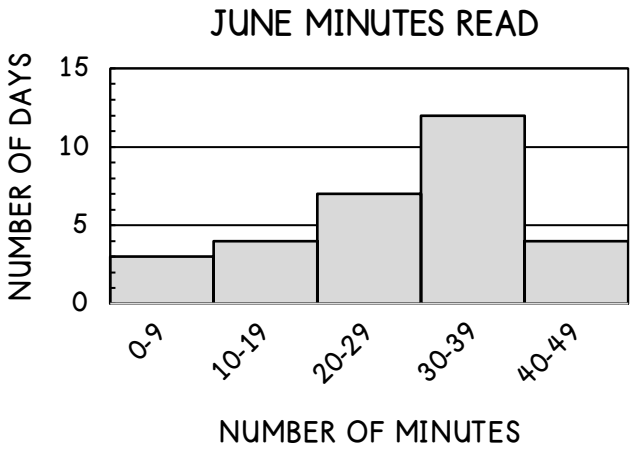
8. The following data set represents the cost of an order (rounded to the nearest dollar) in a fast-food drive thru.

10, 21, 34, 12, 6, 3, 14, 18, 22, 29, 9, 2, 15, 12, 18

Q1: 9 Q3: 21 IQR: 12

I CAN SUMMARIZE DATA SETS.

9. Mrs. Wilson has her son Benji chart the number of minutes he reads each day during the month of June. Use the graph below to answer the questions.



a. How many more days did Benji read 30-39 minutes than 0-9 minutes?

9 days

b. The sum of which two intervals is equal to the number of days that Benji read 20-29 minutes?

0-9 and 10-19
or 0-9 and 40-49

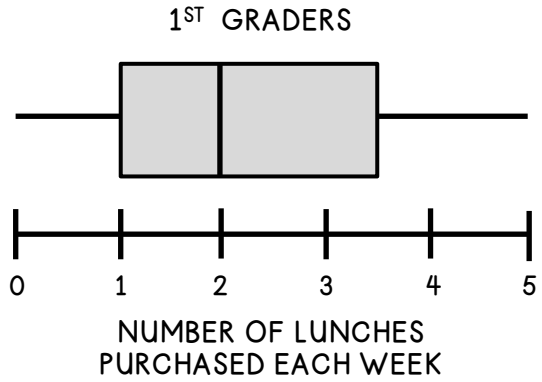
c. Which interval is likely to have the median number minutes read? Why?

30-39

d. Which two intervals account for approximately 60% of Benji's reading?

20-29 and 30-39

10. The school cafeteria takes note of the number of lunches purchased each week by students in first grade. Use the box plot to answer the questions.



a. What is the median number of lunches purchased?

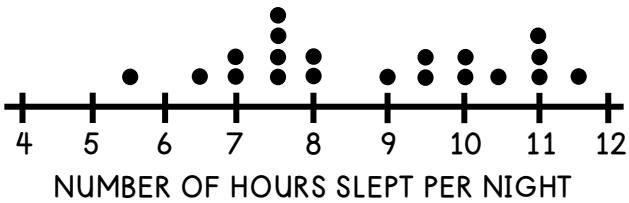
2 lunches

b. Half of the first graders purchased 2 or more lunches a week.

c. The interquartile range of the number of lunches purchased each week is 2.5.

d. The fourth quartile of the first graders purchased their lunch 3.5 times or more each week.

11. Use the data below to answer the questions.



a. Determine the mean number of hours slept each night.

8.75

b. Determine the median number of hours slept each night.

8.5

c. How do the mean and median differ?

The mean is the average number of hours slept, while the median represents the middle number of hours slept. They differ by 0.25 hours.

d. Which number of hours slept per night was most common?

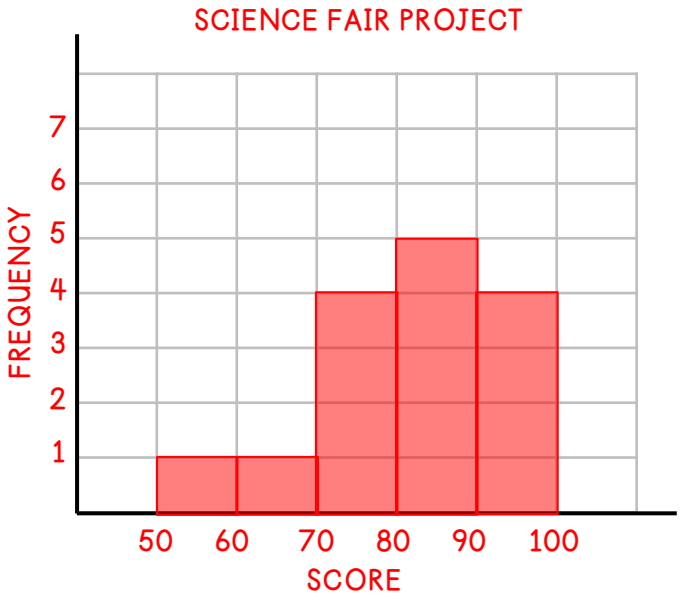
7.5

I CAN DISPLAY DATA ON DOT PLOTS, HISTOGRAMS, AND BOX PLOTS.

12. Use the list of scores below to complete a frequency table and a histogram displaying the grade students received on their science fair projects.

66, 89, 93, 81, 72, 78, 70, 55, 96, 84, 94, 92, 81, 86, 79

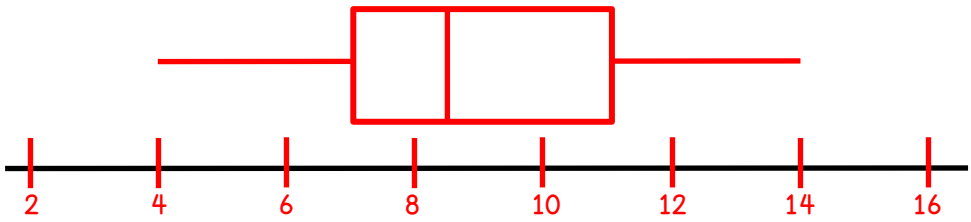
RANGE OF SCORES	TALLY	FREQUENCY
50-59	|	1
60-69	|	1
70-79	||||	4
80-89	|||| |	5
90-99	||||	4



13. The following data set represents the number of hours students spent on their science fair projects. Create a five-number summary and a box plot to display the data.

8, 4, 14, 12, 5, 11, 8, 9, 7, 9

Min: 4 Q1: 7 Med: 8.5 Q3: 11 Max: 14



I'VE GOT THIS:

I NEED TO STUDY:

Name _____

Date _____ Pd _____

DATA & STATISTICS UNIT TEST

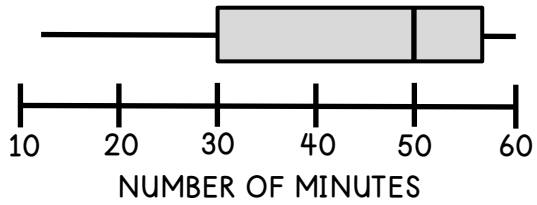
Solve the problems below. Be sure to show your thinking.

1. The number of points in the first five games of the basketball season are listed below. What is the mean number of points scored?

GAME 1	GAME 2	GAME 3	GAME 4	GAME 5
38	29	16	42	33

31.6

2. Students record the number of minutes they read at home. The box plot shows the summary of the results. Which statement best describes the data?



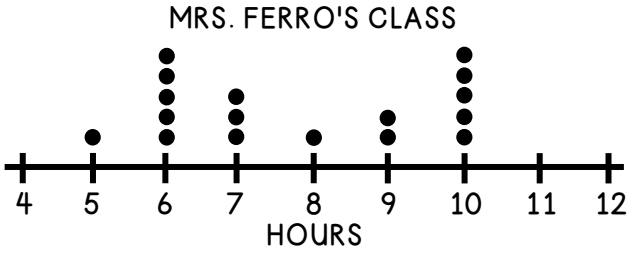
A. A quarter of the students read for 30-40 minutes.

B. The interquartile range is 30.

C. About half the students read for 50 minutes or more.

D. The average number of minutes read was 47.

3. The dot plot shows the number of hours the students in Mrs. Ferro's class volunteer each month. Which of the following statements is **not** correct?



A. There are a total of 17 students in Mrs. Ferro's class.

B. Exactly 7 students volunteered for less than 7 hours.

C. Less than half of the students volunteer for 9 hours or more.

D. The mode is 6 and 10.

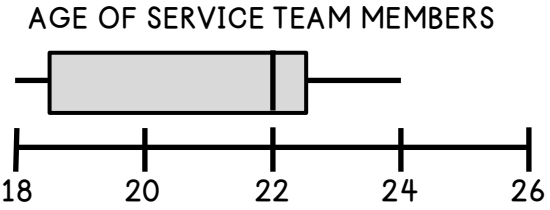
4. The number of wins in the 2015 season in the American League is shown below.

93, 87, 81, 80, 78, 95, 83, 81, 76, 74, 88, 86, 85, 76, 68

What is the range of the number of wins in the American League?

27

5. Based on the box plot below, which statement is **not** true?



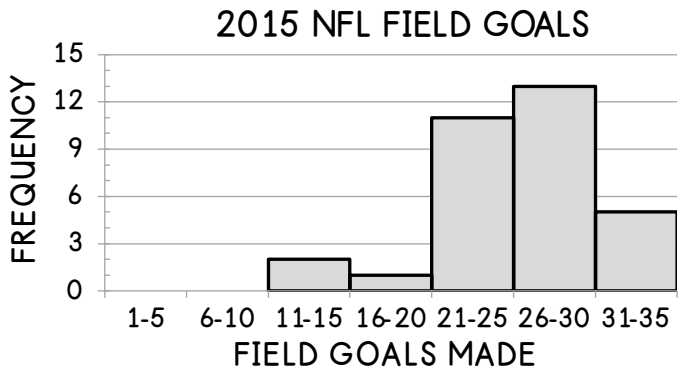
A. The oldest team member is 24.

B. The median age is 22.

C. Fifty percent of the participants are between 22 and 24.

D. Less than 25% of the participants are 20 or younger.

6. Which of the statements about the data in the histogram is true?



A. A total of 30 teams were included in the data.

B. Over half of the teams made 26-30 field goals in the season.

C. There were 13 teams that made between 21-25 field goals during the season.

D. A total of 18 teams made 26 or more field goals.

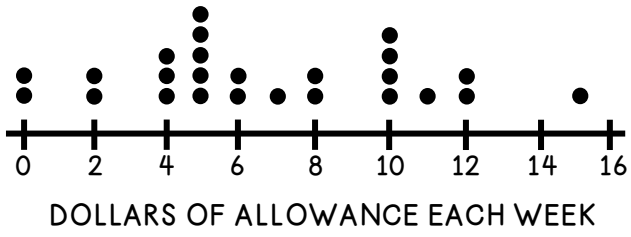
7. The data set below represents the different costs of monthly internet service.

$38, $56, $48, $72, $50, $66

a. What is the mean? $55

b. What is the mean absolute deviation? $9.67

8. The dot plot below represents the number of dollars in allowance that students receive each week in Mrs. Jimenez's class. What is the median amount of allowance?



A. 5 B. 6
C. 7.5 D. 8

9. Which of the following statements is the best description for the data in Question 8?

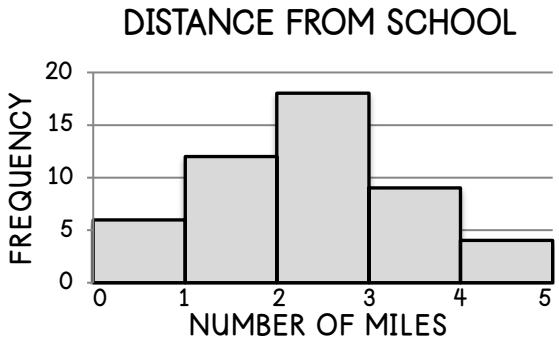
A. The data is spread out and skewed left.

B. Most of the data is far from the median and skewed right.

C. The range of the data is 13.

D. The data is spread out with $5 being the most common weekly allowance.

10. Which of the following statements best describes the data shown below?



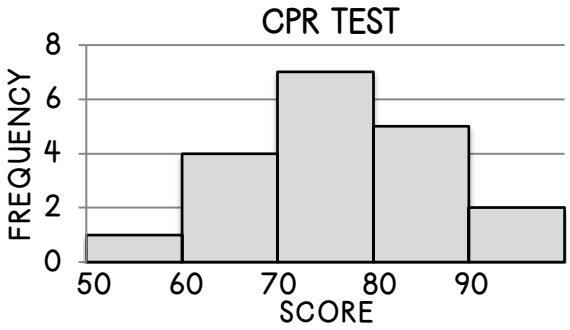
A. The distribution of the data is symmetrical, so the mean and median are likely within 2-3 miles.

B. The distribution of the data is skewed right, so the mean and median are likely within 1-2 miles.

C. The distribution of the data is skewed left, so the mean and median are likely within 1-2 miles.

D. The distribution of the data is skewed left, so the mean and median are likely within 3-4 miles.

11. Nineteen employees took a CPR test. The test scores are displayed below in a histogram.



Which of the following data sets could represent the scores shown in the histogram?

A. 53, 61, 68, 69, 71, 74, 75, 75, 77, 78, 78, 83, 84, 87, 88, 91, 91

B. 53, 68, 69, 70, 71, 74, 75, 75, 77, 78, 78, 81, 85, 86, 88, 91, 91

C. 57, 60, 63, 65, 66, 71, 72, 73, 77, 79, 79, 79, 80, 81, 81, 83, 89, 90, 91

12. Which of the following questions is not considered statistical?

A. What is the shoe size of the students in my class?

B. How much did the students in my class pay for a pair of shoes?

C. How long do you own a pair of shoes?

D. How many pairs of shoes do my classmates own?

13. Which of the following questions would be considered statistical?

A. What is the favorite subject of 5th graders at my school?

B. What is my favorite sport?

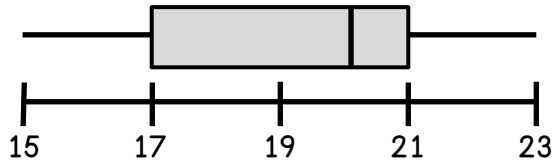
C. How old is the music teacher?

D. How far do you walk to school?

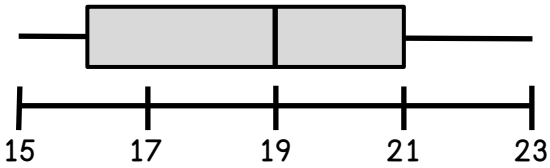
14. The number of people on various teams within a soccer league are shown below. Which box plot best represents the data?

15, 16, 16, 18, 19, 19, 20, 20, 21, 22, 23

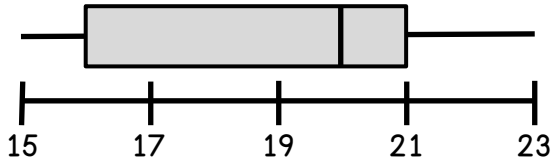
A.



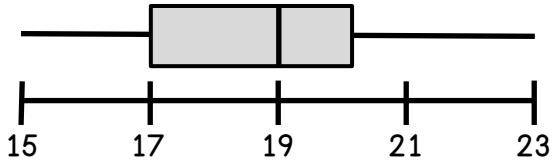
B.



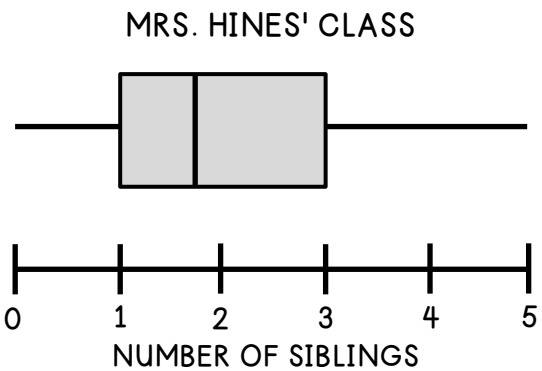
C.



D.



15. The box plot below shows the number of siblings each student in Mrs. Hines' class has. Which of the following statements is **not** true?



A. The interquartile range of the data is approximately 2.

B. The median number of siblings each student has is approximately 1.75.

C. There is at least one student who has no siblings.

D. Over half of the students in Mrs. Hines' class have two or more siblings.