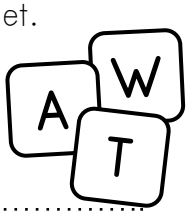
Unit: Data & Statistics
Student Handout 1

Name _____

Date _____ Pd _____

POPULATIONS AND SAMPLES

A board game uses various letter tiles to form words. Each letter of the alphabet has at least one tile. Brandon groups the letters A, E, I, O, and U and states that they are a subset of the set. What set is Brandon referring to? How do you know?



SAMPLES	• A sample is a _____ of an entire group. • It refers to a portion of the group with specific characteristics. • Data collected from the sample is called a _____ _____ .
POPULATIONS	• When collecting data, the _____ references the entire group. • Data collected from the entire population is called a _____ _____ .

The following statements are populations. Determine a possible sample that is part of a population.

POPULATION	SAMPLE
People living in a specific neighborhood	
Families with children	
Seniors at East High School	

The following statements are samples. Determine a possible population that each could be a part of.

SAMPLE	POPULATION
Children who play football for the YMCA	
Students who were on the B honor roll for the first nine weeks	
Households that have cable in the Riverside Apartment complex	

RANDOM SAMPLES

- Everyone in the population must have an _____ _____ of being selected in order for a sample to be _____ .

Ex: A survey is taken of favorite sports in seventh grade, but only girls are surveyed. This is _____ a random sample.

Read the situation below. Then, determine if the various methods would produce a random sample. Justify your thinking in the space below.

1. The senior class president would like to host an after-graduation lock-in. There will be games, gifts, and food! In order to decide if classmates would be willing to buy tickets and attend, she has decided to conduct a survey. She will select 25 students to meet and hear input on the prizes, food, and ticket price. There are 500 members of the senior class.

a. She selects every tenth student as they get off the bus one morning.

b. She asks her choir director to allow her to survey the choir.

c. She meets with the counselor, who selects every twentieth student on the school roster.

d. She assigns each senior a number and randomly selects 25 numbers.

e. She makes an announcement that people interested should meet in the library after school.

f. A roster of the senior class is printed, and each student's name is put into a hat. The principal selects 25 names.

Use your understanding of random sampling to answer the question below.

2. The public library is considering changing their hours to better support their patrons' needs. They decide to survey their patrons asking when they typically visit the library. Describe how you could select a random sample of 50 library patrons.

Summarize today's lesson:

Unit: Data & Statistics
Homework 1

Name _____
Date _____ Pd _____

POPULATIONS AND SAMPLES

Determine the population and sample in each problem below.

1. A survey of 2,541 American households discovered that 64% of the households own one car.

Population: _____

Sample: _____

2. The average height of every fifth member of the varsity football team was 5'11".

Population: _____

Sample: _____

3. A recent study found that 40% of New York City residents earn more than the median salary.

Population: _____

Sample: _____

4. In a news poll, 73% of readers indicated that they were satisfied with their job.

Population: _____

Sample: _____

Use your understanding of random sampling to answer the questions below.

5. A school board member surveys parents to learn how they feel about the new school boundaries. Mrs. Donovan shares an example of a random sample and Mr. Giles shares an example of a sampling method that is not random. What could they have shared?

MRS. DONOVAN

MR. GILES

6. The mayor is hoping to develop a new outdoor park area. He decides to survey the town residents by surveying YMCA members on a Saturday morning. Is this a random sample? Justify your answer.

Unit: Data & Statistics
Student Handout 2

Name _____

Date _____Pd _____

POPULATION INFERENCES

POPULATION INFERENCES

- Once a sample has been collected from a population, an _____ about the entire population can be made by setting up a proportion.
- These inferences can also be used to compare different populations and make _____ .

Use the information below to answer the questions.

1. Eastside Middle School conducted a survey of randomly selected 6th graders to determine which elective they were most likely to participate in. The results are shown in the table.

	SPANISH	THEATER	ART
6TH GRADE	14	20	6

a. How many students were surveyed?

b. What percentage of students surveyed selected Spanish as their elective of choice?

c. There are 350 students in the 6th grade at Eastside Middle School. Using this sample, about how many students can be expected to sign up for art?

d. Of the 350 students in the 6th grade, how many students can be expected to sign up for theater?

e. The school decides to drop any elective that has less than 10% of the students' interest. Miguel says that since only 6 students would like to sign up for art, it should be dropped. Explain why this is or is not correct.

_____

_____

_____

f. Mark each of the following statements as true or false based on the data.

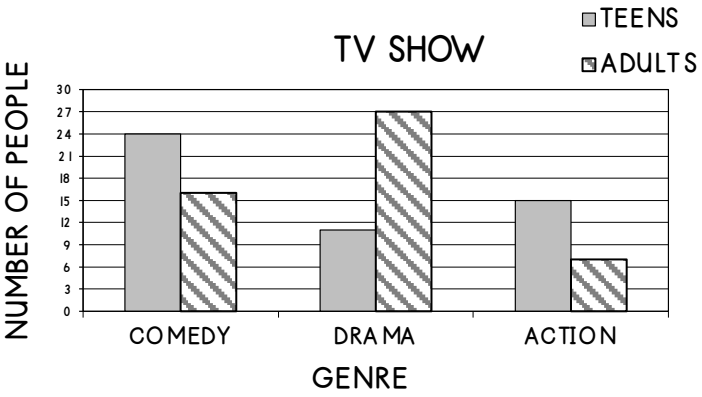
_____ Over half of the students at Eastside Middle School will select theater for their elective.

_____ Combined, the Spanish and art classes account for 50% of the elective results.

_____ Out of the 350 students at Eastside Middle School, at least 120 will sign up for Spanish.

2. A video company randomly selected teen and adult subscribers and asked them their preferred genre of TV shows.

Use the survey results in the graph at right to answer the questions below.



a. How many total subscribers were surveyed?

b. What percentage of subscribers surveyed selected a drama as their TV show of choice?

c. The video company estimates that there are actually 2,000 adult subscribers. Using this sample, how many adult subscribers can be estimated to watch action shows?

d. What percentage of subscribers that preferred comedies were adults?

e. Mark each of the following statements as true or false based on the data.

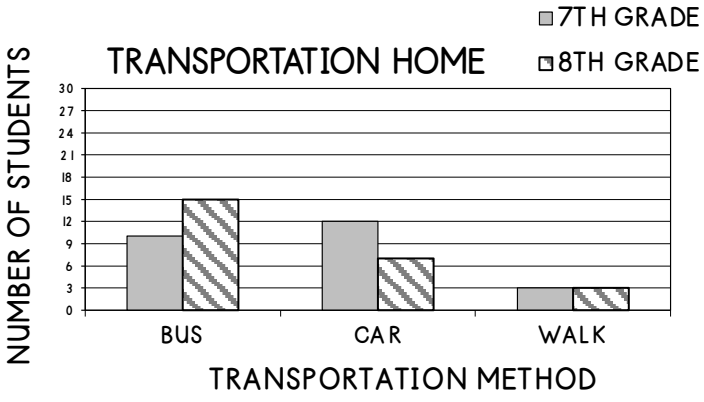
_____ More teens were surveyed than adults.

_____ Adults liked dramas more than comedies and action shows combined.

_____ Half as many teens chose action over comedy.

Use your knowledge of population inferences to answer the question below.

3. The school counselor randomly selected a group of students at Ford Junior High to survey ways they get home from school. Based on the results below, which of the following does **not** represent the data?



A. A total of 50 students were surveyed.

B. Sixty percent of the 8th graders rode the bus home.

C. The number of car riders and walkers in 8th grade is equal to the number of car riders in 7th grade.

D. Exactly half of the students surveyed rode the bus.

Summarize today's lesson:

Unit: Data & Statistics
Homework 2

Name _____

Date _____ Pd _____

POPULATION INFERENCES

The Mitchell Junior High newspaper staff conducted a survey on school start times with two samples of 40 randomly selected students to represent the entire 600-member student population. The results are shown below.

	7:45 AM START TIME	8:30 AM START TIME	9:00 AM START TIME
SAMPLE #1	12	18	10
SAMPLE #2	14	19	7

1. Read each headline and explain if the statement is supported by the data.

HEADLINE A

OVER 300 STUDENTS WANT TO START AFTER 8:00 AM

HEADLINE B

8:30 AM START TIME MOST POPULAR WITH STUDENTS

HEADLINE C

SMALL PERCENTAGE OF STUDENTS WANT TO START AT 7:45 AM

HEADLINE D

LESS THAN 100 STUDENTS WANT TO START AT 9:00 AM

Headline A: _____

_____

Headline B: _____

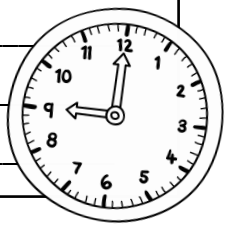
_____

Headline C: _____

_____

Headline D: _____

_____



Use your understanding of population inferences to answer the questions below.

2. A school board randomly samples 80 students to determine their opinion on requiring school uniforms for the next school year. The table shows the results of the survey.

OPPOSED	UNDECIDED	IN FAVOR
11	24	45

a. If 1,200 students are in the district, how many students can be expected to oppose school uniforms?

b. Bernice says that based on the survey, a student is more likely to be undecided or opposed than in favor. Do you agree or disagree? Why or why not?

Unit: Data & Statistics
Student Handout 3

Name _____

Date _____Pd_____

MEASURES OF CENTER

MEASURES OF CENTER

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

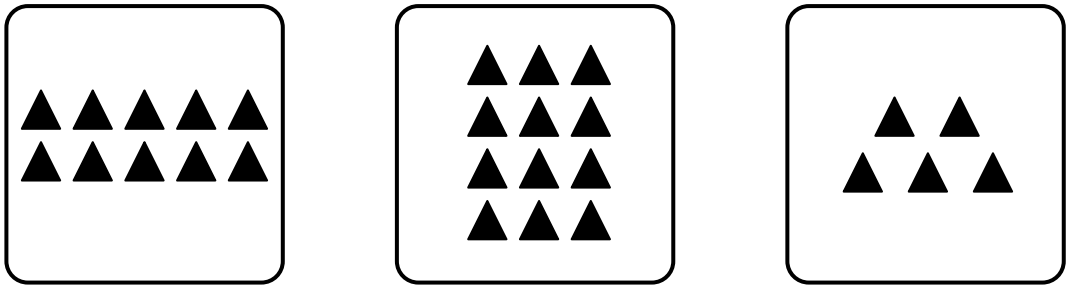
MEAN

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

1. Finding the _____ of all the data points
2. _____ by the number of data points

- The mean can be thought of as determining the _____ _____.

Mrs. Hawkins asked students to divide the blocks into three different buckets. How many blocks would be in each bucket if they were shared fairly? Discuss your process for finding the fair share.



Use your understanding of mean to answer the questions below.

1. The list below shows the number of points the Timberwolves basketball team scored.

43, 72, 62, 57, 60, 49, 57, 49, 63, 55

Sum: _____

Mean: _____

2. The following data set represents the height in inches of people in line to ride a roller coaster.

48.5, 36, 42, 37.5, 50.2, 42.6

Sum: _____

Mean: _____

3. The following data set represents the speed of several cars as they pass over the Widow Falls Bridge.

SPEED	55	57	58	60	62
NUMBER OF CARS	2	0	2	3	1

List: _____

Sum: _____

Mean: _____

MEDIAN

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

Determine the median of the data sets below.

4. The following data set represents the number of menu items at seven local restaurants.

35, 24, 50, 22, 32, 28, 46

Median: _____

5. The following data set represents the number of leaves that have fallen off several trees in Mr. Smith's neighborhood.

59, 112, 68, 37, 62, 107, 85, 94

Median: _____

Use your understanding of mean and median to answer the questions below.

6. The number of items purchased by the last five customers at Gregory's Groceries is shown in the table below.

NAME	# OF ITEMS
Adrian	42
Mallory	15
Luis	23
Gerardo	56
Bett	22

a. Mean: _____

b. Median: _____

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?

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

18	13	12	14	11	?

_____

Summarize today's lesson:

Unit: Data & Statistics
Homework 3

Name _____

Date _____ Pd_____

MEASURES OF CENTER

Solve the problem on card A. Draw a line from the arrow on card A to its solution in the top corner of a card in column #2. Then solve the corresponding problem. Continue to draw lines showing the path from each card to its solution in the opposite column until you end at the solution on card A.

COLUMN #1

A

8.25

The table represents the number of minutes of commercials during hour-long TV shows. Find the median.

# OF MINS	12	13.5	14	14.5
FREQUENCY	2	2	3	1

C

131.5

The data set below has a mean of 20.5. What would be the new mean if 17 was added to the list?

24, 18, 20, 13, 25, 23

E

24

The cards below have a mean of 10.6. What is the missing number?

12.6	10	?	11.4

G

11.3

Calculate the mean of the data set below.

116, 108, 125, 208, 132

COLUMN #2

13.75

B

The data set below has a median of 23. What would be the new median if 25 was added to the list?

22, 24, 17, 26, 21, 18, 29, 27

20

D

Calculate the median of the data set below.

11.17, 12.33, 11.82, 10.83, 10.97, 11.43

8.4

F

Calculate the median of the data set below.

108, 110, 168, 137, 152, 126

137.8

H

The table represents the number of elective classes offered at different middle schools. Find the mean.

# OF CLASSES	7	8	9	10
FREQUENCY	2	3	2	1

Unit: Data & Statistics
Student Handout 4

Name _____
Date _____ Pd _____

MEASURES OF VARIABILITY

MEASURES OF VARIABILITY

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

IQR [INTERQUARTILE RANGE]

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

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

1. The data set below represents the number of animals in different exhibits at a zoo.

48, 86, 15, 27, 18, 52, 103

a. Write the data from least to greatest.

b. What is the minimum number of animals?

c. What is the maximum number of animals?

d. What is the median number of animals?

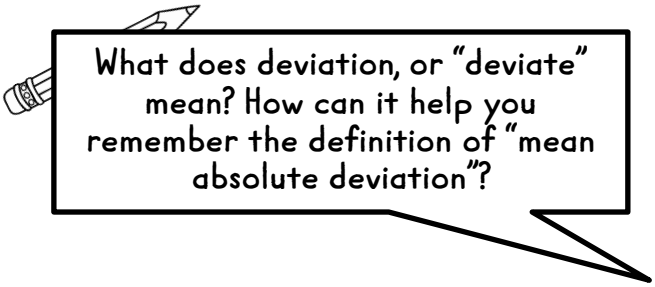
e. What is the median of the first half of the data? (first quartile)

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

g. What is the interquartile range?

MEAN ABSOLUTE DEVIATION

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



STEPS TO CALCULATE M.A.D.

1. Find the _____ of the data set

2. Find the _____ from each data point and the mean

3. Find the _____ of those distances

Use the steps above to complete the table and find the mean absolute deviation in the situation below.

2. Mr. Breedlove recorded the number of minutes he spent traveling to work for the last seven days.

48, 40, 62, 52, 60, 42, 39

# OF MINUTES	DISTANCE FROM MEAN
48	
40	
62	
52	
60	
42	
39	
TOTAL DEVIATION	

a. Find the mean of the data.

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.

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

Use measures of variability to answer the questions below.

3. Mrs. Johnson recorded the number of pages five of her students read that day. Complete the blanks below in order to make each statement true.

OF PAGES
20
22
18
19
21

- The minimum number of pages read is _____ and the maximum number of pages read is _____.
- The interquartile range of _____ means that there is a _____ amount of variability in the number of pages read.
- The average distance that each number of pages read is from the mean is _____.
- Since the mean absolute deviation is a _____ number, this means that the number of pages read are _____.

Unit: Data & Statistics
Homework 4

Name _____

Date _____ Pd _____

MEASURES OF VARIABILITY

Use your understanding of variability to answer the questions below.

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

108, 53, 47, 22, 56, 62

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

12.2, 6.8, 3.4, 8.8

3. The average daily rainfall in Seattle for the month of June has a mean absolute deviation of 1.3 inches. What conclusion can you make about the average daily rainfall in Seattle for the month of June?

4. If the mean absolute deviation is a really large number, then what does that mean about the data set?

5. The number of students in 10 kindergarten classes is shown below.

20, 18, 21, 18, 22, 24, 18, 23, 23, 24

a. What is the median number of students in the kindergarten classes? _____

b. What number of students describes the first quartile? _____

c. What number of students describes the third quartile? _____

d. What number of students describes the interquartile range? _____

6. Bryan thinks that a low interquartile range for average daily temperatures means that the temperatures were really low. Do you agree or disagree? Why or why not?

Unit: Data & Statistics
Student Handout 5

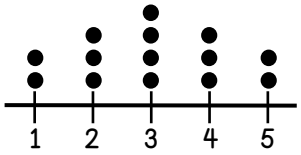
Name _____

Date _____ Pd _____

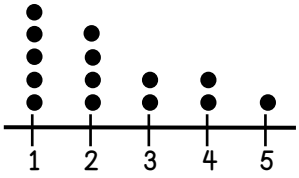
COMPARING DOT PLOTS

SHAPE

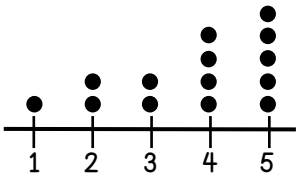
- Data can take on three different shapes:



SYMMETRIC



SKEWED RIGHT



SKEWED LEFT

SPREAD

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

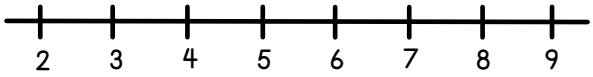
CENTER

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

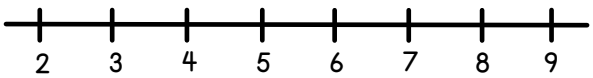
Use your understanding of dot plots to answer the questions below.

1. Two different classes were asked how many minutes each student traveled to school in the mornings.

ROOM 434: 5, 7, 8, 8, 4, 4, 7, 5, 4, 7
ROOM 425: 4, 5, 5, 6, 3, 6, 4, 7, 5, 5



MINUTES TRAVELED TO SCHOOL IN ROOM 434



MINUTES TRAVELED TO SCHOOL IN ROOM 425

2. Describe the shape of the two dot plots above. Which group appears to have the higher number of minutes traveled to school?

3. What is the median of each data set?

a. Room 434: _____

b. Room 425: _____

4. How does the median number of minutes in Room 434 compare to that of Room 425?

Use the information from the dot plots on the previous page to answer the following questions.

5. Mark each of the following statements as true or false based on the information.

_____ a. The distribution for Room 434 is symmetrical while the distribution for Room 425 is skewed left.

_____ b. The median number of minutes in Room 425 is greater than the median number of minutes in Room 434.

_____ c. Room 434 has a smaller range than Room 425.

_____ d. The distribution for Room 425 is symmetrical while the distribution for Room 434 has no pattern.

6. Set up a table to calculate the mean absolute deviation of each classroom.

Room 434: 5, 7, 8, 8, 4, 4, 7, 5, 4, 7

# OF MINUTES	DISTANCE FROM MEAN
5	
7	
8	
8	
4	
4	
7	
5	
4	
7	
TOTAL DEVIATION	

MAD: _____

Room 425: 4, 5, 5, 6, 3, 6, 4, 7, 5, 5

# OF MINUTES	DISTANCE FROM MEAN
4	
5	
5	
6	
3	
6	
4	
7	
5	
5	
TOTAL DEVIATION	

MAD: _____

a. What observations can you make about the mean absolute deviation of each classroom?

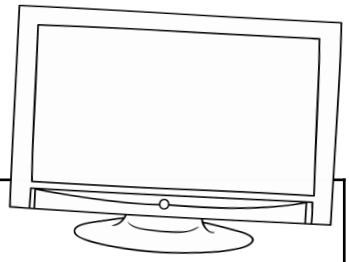
b. Room 434 has a mean average deviation almost _____ that of Room 425.

Unit: Data & Statistics
Homework 5

Name _____

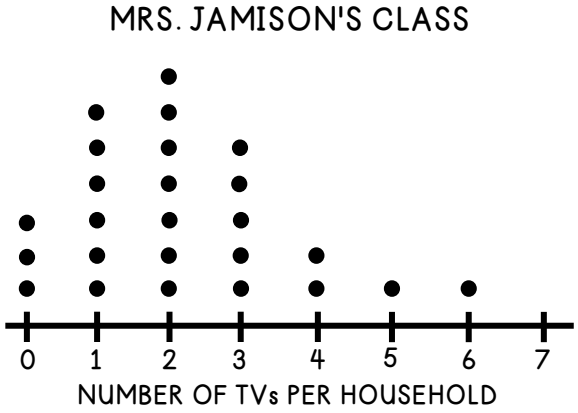
Date _____ Pd _____

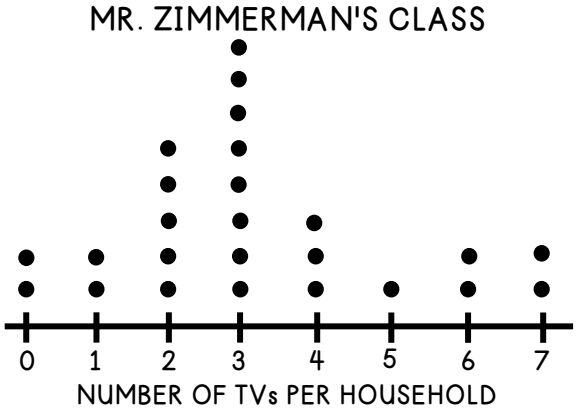
COMPARING DOT PLOTS



Use the dot plots to answer the questions below.

The number of TVs per household for Mrs. Jamison's class and Mr. Zimmerman's class is shown below.





1. What is the median number of TVs in Mrs. Jamison's Class?

2. What is the median number of TVs in Mr. Zimmerman's Class?

3. The mean absolute deviation of Mrs. Jamison's class is 0.24. What does this mean?

4. The mean absolute deviation of Mr. Zimmerman's class is 0.46. What does this mean?

5. Mark each statement below as true or false.

_____ The range of the number of TVs per household in Mrs. Jamison's class is less than the range of the number of TVs per household in Mr. Zimmerman's class.

_____ The median number of TVs per household is equal in both Mrs. Jamison's class and Mr. Zimmerman's class.

_____ The number of TVs per household with the highest frequency for both Mrs. Jamison's class and Mr. Zimmerman's class is 3.

6. When Asher looks at the data, he says that both Mrs. Jamison's class and Mr. Zimmerman's class has data that is skewed right. Do you agree or disagree? Why or why not?

_____

_____

_____

Unit: Data & Statistics
Student Handout 6

Name _____
Date _____ Pd _____

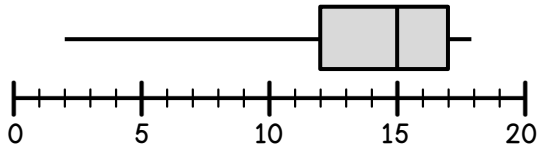
COMPARING BOX PLOTS

BOX PLOTS

- A box plot displays a data distribution using _____ key numbers.

  The difference between the first and third quartile is called the _____, which is a measure of variability.

  - _____: the smallest piece of data
  - _____: the median of the lower half of data
  - _____: the median (midpoint) of the data
  - _____: the median of the upper half of data
  - _____: the largest piece of data

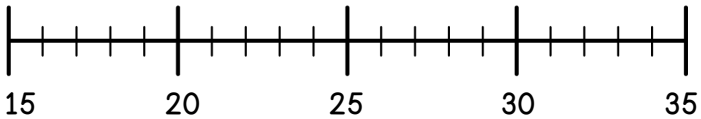


Use the data below to create a five-number summary and sketch a box plot.

1. The following data set represents the number of pages in a book read each day.

34, 32, 19, 35, 33, 30, 34

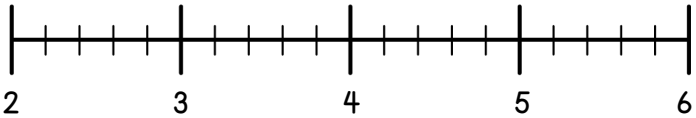
Min: _____
Q1: _____
Med: _____
Q3: _____
Max: _____



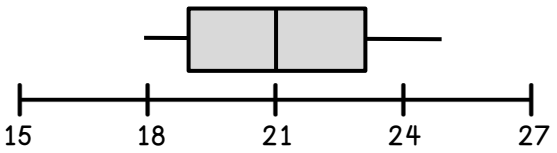
2. The following data set represents the number of miles Monica walked each day.

4.2, 3.8, 4.7, 5.8, 3.2, 4.1, 5

Min: _____
Q1: _____
Med: _____
Q3: _____
Max: _____



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

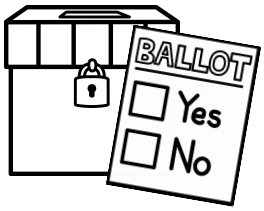


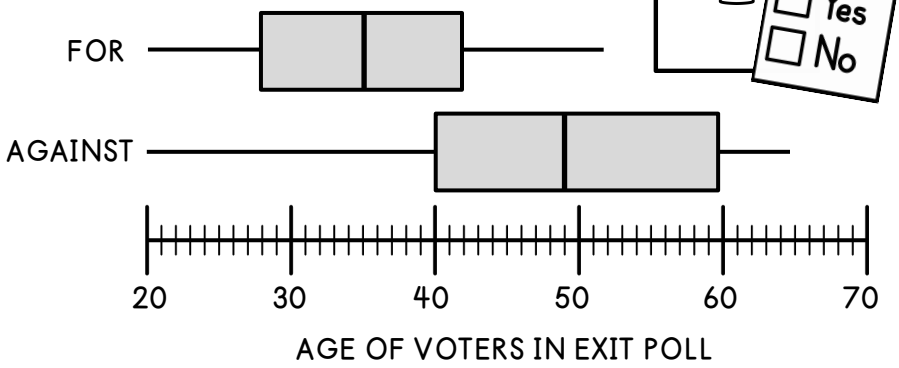
Min: _____ Q1: _____ Med: _____ Q3: _____ Max: _____

Use the box plots below to answer questions 4-6.

In a local election, a school bond is proposed to fund a natatorium.

Alex conducts an exit poll in which he asks voters to share their age and whether they were for or against the school bond. He then displays the data in a box plot.





4. What is the difference in the measure of center of the voters who are for and against the natatorium?

5. Determine the IQR of voters who voted against the natatorium. Determine the IQR of voters who voted for the natatorium.

6. Complete the statements below:

   a. The range in age of voters for the natatorium was _____ the range in age of voters against the natatorium.

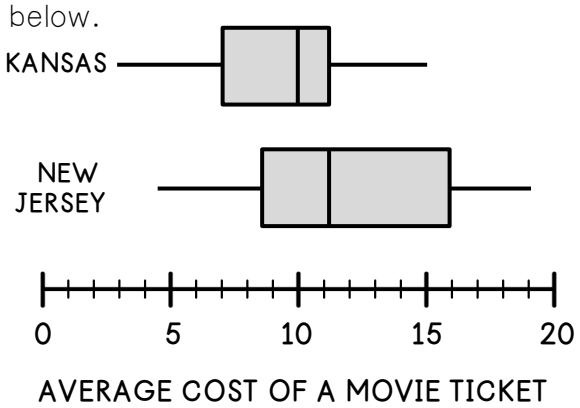
   b. The IQR of those who voted against the natatorium was almost _____ times the IQR of those who voted for the natatorium.

In question 7, compare the two box plots below. Justify your solutions.

7. A survey was conducted to determine the average price of a movie ticket in Kansas and New Jersey. The data was displayed using the box plots shown below.

a. What was the median cost of a movie ticket in each state?

b. Which state has a wider range in variability? Why?



c. A movie goer spends $14 on a movie ticket in New Jersey. What can you conclude about the price the customer paid?

Unit: Data & Statistics
Homework 6

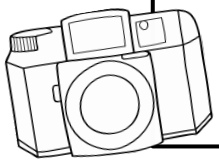
Name _____

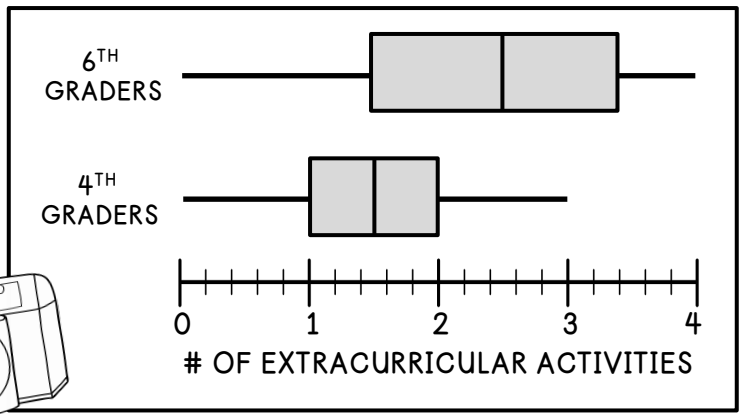
Date _____ Pd_____

COMPARING BOX PLOTS

Use the box plots at the right to answer questions 1-4, and then apply your understanding of box plots in question 5.

The 4th and 6th grade students at Tree Hill Elementary completed a survey on how many extracurricular activities they participate in. The data is shown in the box plots at the right.





1. What is the median number of extracurricular activities for each grade?

2. What is the difference in the spread of the data between the two grades?

3. Determine the variability of the number of extracurricular activities for each grade.

4. Read each of the following statements and mark them as true or false.

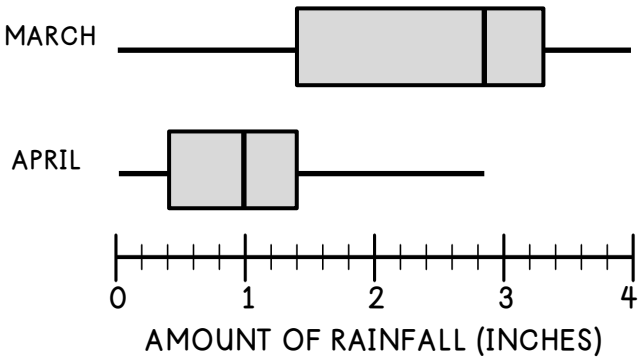
_____ The third quartile of the data for 4th graders is less than the third quartile of the data for 6th graders.

_____ The first quartile of the data for 6th graders is equal to the median of the data for 4th graders.

_____ The maximum number of extracurricular activities for 4th graders is greater than the maximum number of extracurricular activities for 6th graders.

_____ Both grades have at least one student who does not participate in any extracurricular activities.

5. The boxplots show the amount of rainfall for the months of March and April. Which statement is represented by the data?



A. March has half the variability of April.

B. The third quartile of the data in April is greater than the first quartile of the data in March.

C. The maximum amount of rainfall in April is equal to the median amount of rainfall in March.

D. The difference in the centers of March and April is greater than 2.

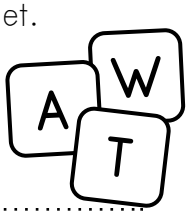
Unit: Data & Statistics
Student Handout 1

Name _____
Date _____ Pd _____

POPULATIONS AND SAMPLES

A board game uses various letter tiles to form words. Each letter of the alphabet has at least one tile. Brandon groups the letters A, E, I, O, and U and states that they are a subset of the set. What set is Brandon referring to? How do you know?

Ex: Vowels are a subset (or a small group) within all the letter of the alphabet



SAMPLES	• A sample is a _____subset_____ of an entire group. • It refers to a portion of the group with specific characteristics. • Data collected from the sample is called a _____sample_____ _____statistic_____ .
POPULATIONS	• When collecting data, the _____population_____ references the entire group. • Data collected from the entire population is called a _____population_____ _____statistic_____ .

The following statements are populations. Determine a possible sample that is part of a population.

POPULATION	SAMPLE
People living in a specific neighborhood	Ex: the elderly people within the neighborhood
Families with children	Ex: families with only one child
Seniors at East High School	Ex: seniors who drive their car to East High School

The following statements are samples. Determine a possible population that each could be a part of.

SAMPLE	POPULATION
Children who play football for the YMCA	Ex: people who participate in YMCA programs
Students who were on the B honor roll for the first nine weeks	Ex: all students enrolled during the first nine weeks
Households that have cable in the Riverside Apartment complex	Ex: all households in the Riverside Apartment complex

RANDOM SAMPLES

- Everyone in the population must have an equal chance of being selected in order for a sample to be random.

Ex: A survey is taken of favorite sports in seventh grade, but only girls are surveyed. This is NOT a random sample.

Read the situation below. Then, determine if the various methods would produce a random sample. Justify your thinking in the space below.

1. The senior class president would like to host an after-graduation lock-in. There will be games, gifts, and food! In order to decide if classmates would be willing to buy tickets and attend, she has decided to conduct a survey. She will select 25 students to meet and hear input on the prizes, food, and ticket price. There are 500 members of the senior class.

a. She selects every tenth student as they get off the bus one morning.

This is not a random sample because it only includes students who ride the bus.

b. She asks her choir director to allow her to survey the choir.

This is not a random sample because it only includes students who are members of the choir.

c. She meets with the counselor, who selects every twentieth student on the school roster.

This is a random sample.

d. She assigns each senior a number and randomly selects 25 numbers.

This is a random sample.

e. She makes an announcement that people interested should meet in the library after school.

This is not a random sample because it only includes students who volunteer to participate.

f. A roster of the senior class is printed, and each student's name is put into a hat. The principal selects 25 names.

This is a random sample.

Use your understanding of random sampling to answer the question below.

2. The public library is considering changing their hours to better support their patrons' needs. They decide to survey their patrons asking when they typically visit the library. Describe how you could select a random sample of 50 library patrons.

Ex: Use a computer to generate a list of library patrons, then randomly select 50 to survey.

Summarize today's lesson:

Note: This is a strategy for moving information from short-to long-term memory. Have students write 2-3 sentences.

Unit: Data & Statistics
Homework 1

Name _____
Date _____ Pd _____

POPULATIONS AND SAMPLES

Determine the population and sample in each problem below.

1. A survey of 2,541 American households discovered that 64% of the households own one car.

Population: American households

Sample: American households that own a car

2. The average height of every fifth member of the varsity football team was 5'11".

Population: the varsity football team members

Sample: every fifth member of the team

3. A recent study found that 40% of New York City residents earn more than the median salary.

Population: New York City residents

Sample: New York City residents that earn more than the median salary

4. In a news poll, 73% of readers indicated that they were satisfied with their job.

Population: readers who took the news poll

Sample: readers that were satisfied with their job

Use your understanding of random sampling to answer the questions below.

5. A school board member surveys parents to learn how they feel about the new school boundaries. Mrs. Donovan shares an example of a random sample and Mr. Giles shares an example of a sampling method that is not random. What could they have shared?

MRS. DONOVAN: Ex: "We could go through the district roster and poll every 100th parent."

MR. GILES: Ex: "We could survey the parents at the school board meeting."

6. The mayor is hoping to develop a new outdoor park area. He decides to survey the town residents by surveying YMCA members on a Saturday morning. Is this a random sample? Justify your answer.

Ex: No, residents who are involved at the YMCA may be residents who are more connected to the community or more active outdoors and may skew his data.

Unit: Data & Statistics
Student Handout 2

Name _____
Date _____ Pd _____

POPULATION INFERENCES

POPULATION INFERENCES

- Once a sample has been collected from a population, an inference about the entire population can be made by setting up a proportion.
- These inferences can also be used to compare different populations and make predictions.

Use the information below to answer the questions.

1. Eastside Middle School conducted a survey of randomly selected 6th graders to determine which elective they were most likely to participate in. The results are shown in the table.

	SPANISH	THEATER	ART
6TH GRADE	14	20	6

a. How many students were surveyed?

40 students

b. What percentage of students surveyed selected Spanish as their elective of choice?

35% selected Spanish

c. There are 350 students in the 6th grade at Eastside Middle School. Using this sample, about how many students can be expected to sign up for art?

about 53 students

d. Of the 350 students in the 6th grade, how many students can be expected to sign up for theater?

175 students

e. The school decides to drop any elective that has less than 10% of the students' interest. Miguel says that since only 6 students would like to sign up for art, it should be dropped. Explain why this is or is not correct.

Miguel is incorrect. Six students out of 40 signed up for art, which is 15% of the sample. We can infer that based on the sample, the population will have about 15% of students sign up for art.

f. Mark each of the following statements as true or false based on the data.

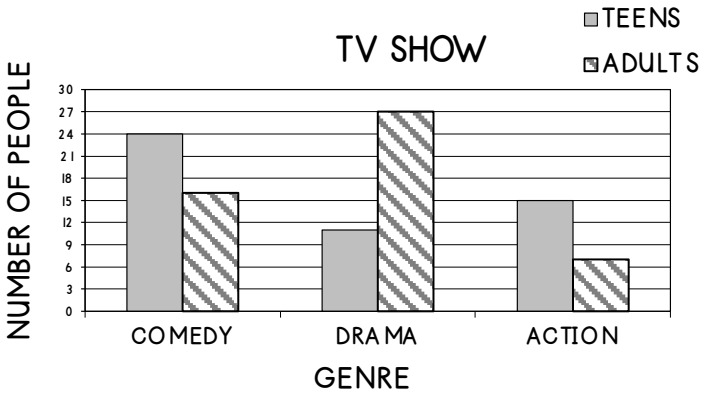
false Over half of the students at Eastside Middle School will select theater for their elective.

true Combined, the Spanish and art classes account for 50% of the elective results.

true Out of the 350 students at Eastside Middle School, at least 120 will sign up for Spanish.

2. A video company randomly selected teen and adult subscribers and asked them their preferred genre of TV shows.

Use the survey results in the graph at right to answer the questions below.



a. How many total subscribers were surveyed?

100 subscribers

b. What percentage of subscribers surveyed selected a drama as their TV show of choice?

38% selected drama

c. The video company estimates that there are actually 2,000 adult subscribers. Using this sample, how many adult subscribers can be estimated to watch action shows?

Approximately 280 adults

d. What percentage of subscribers that preferred comedies were adults?

40%

e. Mark each of the following statements as true or false based on the data.

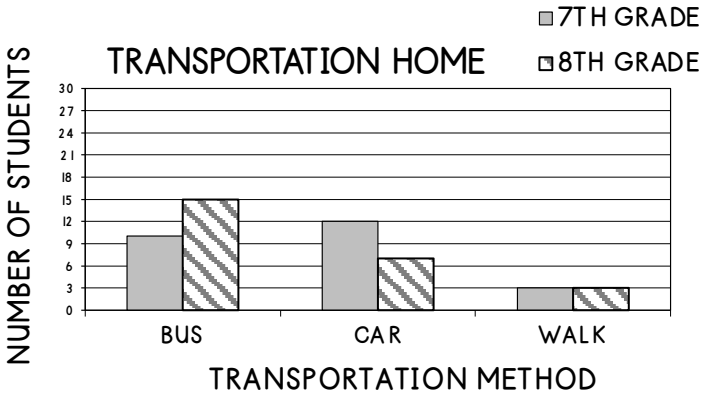
false More teens were surveyed than adults.

true Adults liked dramas more than comedies and action shows combined.

false Half as many teens chose action over comedy.

Use your knowledge of population inferences to answer the question below.

3. The school counselor randomly selected a group of students at Ford Junior High to survey ways they get home from school. Based on the results below, which of the following does **not** represent the data?



A. A total of 50 students were surveyed.

B. Sixty percent of the 8th graders rode the bus home.

C. The number of car riders and walkers in 8th grade is equal to the number of car riders in 7th grade.

D. Exactly half of the students surveyed rode the bus.

Summarize today's lesson:

Unit: Data & Statistics
Homework 2

Name _____
Date _____ Pd _____

POPULATION INFERENCES

The Mitchell Junior High newspaper staff conducted a survey on school start times with two samples of 40 randomly selected students to represent the entire 600-member student population. The results are shown below.

	7:45 AM START TIME	8:30 AM START TIME	9:00 AM START TIME
SAMPLE #1	12	18	10
SAMPLE #2	14	19	7

1. Read each headline and explain if the statement is supported by the data.

HEADLINE A
OVER 300 STUDENTS WANT TO START AFTER 8:00 AM

HEADLINE B
8:30 AM START TIME MOST POPULAR WITH STUDENTS

HEADLINE C
SMALL PERCENTAGE OF STUDENTS WANT TO START AT 7:45 AM

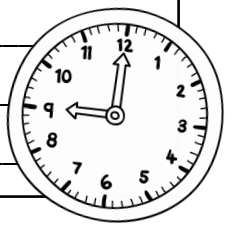
HEADLINE D
LESS THAN 100 STUDENTS WANT TO START AT 9:00 AM

Headline A: True; in sample #1, 28 out of 40 people wanted to start after 8:00 AM. Based on that data, about 420 students wanted to start after 8:00 AM. (Or 390 students using sample #2)

Headline B: True; in both sample #1 and sample #2, the greatest number of people chose 8:30 AM.

Headline C: False; in sample #1, 30% of the students wanted to start at 7:45 AM, and in sample #2, 35% of the students wanted to start at 7:45 AM.

Headline D: False; using sample #1, 150 of the students wanted to start at 9:00 AM, and using sample #2, 105 of the students wanted to start at 9:00 AM.



Use your understanding of population inferences to answer the questions below.

2. A school board randomly samples 80 students to determine their opinion on requiring school uniforms for the next school year. The table shows the results of the survey.

OPPOSED	UNDECIDED	IN FAVOR
11	24	45

a. If 1,200 students are in the district, how many students can be expected to oppose school uniforms?

165 students

b. Bernice says that based on the survey, a student is more likely to be undecided or opposed than in favor. Do you agree or disagree? Why or why not?

Ex: I disagree because in the sample over half of the students were in favor of school uniforms.

Unit: Data & Statistics
Student Handout 3

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.

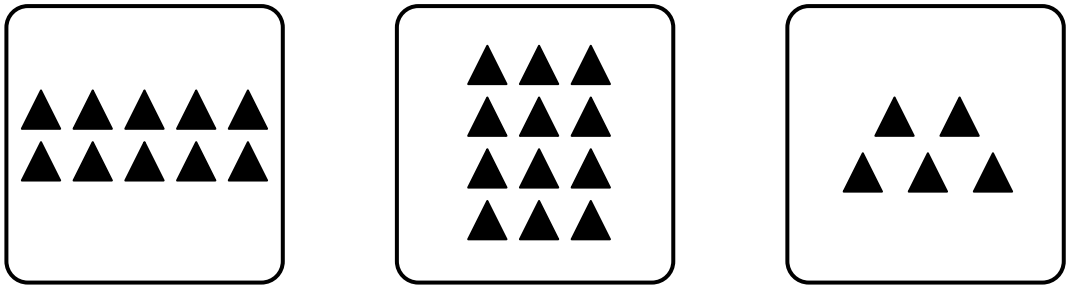
MEAN

- 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___.

Mrs. Hawkins asked students to divide the blocks into three different buckets. How many blocks would be in each bucket if they were shared fairly? Discuss your process for finding the fair share. Each bucket would need 9 blocks in order to be shared fairly.



Use your understanding of mean to answer the questions below.

1. The list below shows the number of points the Timberwolves basketball team scored.

43, 72, 62, 57, 60, 49, 57, 49, 63, 55

Sum: 567

Mean: 56.7

2. The following data set represents the height in inches of people in line to ride a roller coaster.

48.5, 36, 42, 37.5, 50.2, 42.6

Sum: 256.8

Mean: 42.8

3. The following data set represents the speed of several cars as they pass over the Widow Falls Bridge.

SPEED	55	57	58	60	62
NUMBER OF CARS	2	0	2	3	1

List: 55, 55, 58, 58, 60, 60, 60, 62

Sum: 468

Mean: 58.5

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 the data sets below.

4. The following data set represents the number of menu items at seven local restaurants.

35, 24, 50, 22, 32, 28, 46

Median: 32

5. The following data set represents the number of leaves that have fallen off several trees in Mr. Smith's neighborhood.

59, 112, 68, 37, 62, 107, 85, 94

Median: 76.5

Use your understanding of mean and median to answer the questions below.

6. The number of items purchased by the last five customers at Gregory's Groceries is shown in the table below.

NAME	# OF ITEMS
Adrian	42
Mallory	15
Luis	23
Gerardo	56
Bett	22

a. Mean: 31.6

b. Median: 23

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 not very closely related. There is quite a bit of spread between the data points (the data varies greatly).

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

18	13	12	14	11	?

4

Summarize today's lesson:

Unit: Data & Statistics
Homework 3

Name _____

Date _____ Pd_____

MEASURES OF CENTER

Solve the problem on card A. Draw a line from the arrow on card A to its solution in the top corner of a card in column #2. Then solve the corresponding problem. Continue to draw lines showing the path from each card to its solution in the opposite column until you end at the solution on card A.

COLUMN #1

A

8.25

The table represents the number of minutes of commercials during hour-long TV shows. Find the median.

# OF MINS	12	13.5	14	14.5
FREQUENCY	2	2	3	1

C

131.5

The data set below has a mean of 20.5. What would be the new mean if 17 was added to the list?

24, 18, 20, 13, 25, 23

E

24

The cards below have a mean of 10.6. What is the missing number?

12.6 | 10 | ? | 11.4

G

11.3

Calculate the mean of the data set below.

116, 108, 125, 208, 132

COLUMN #2

B

13.75

The data set below has a median of 23. What would be the new median if 25 was added to the list?

22, 24, 17, 26, 21, 18, 29, 27

D

20

Calculate the median of the data set below.

11.17, 12.33, 11.82, 10.83, 10.97, 11.43

F

8.4

Calculate the median of the data set below.

108, 110, 168, 137, 152, 126

H

137.8

The table represents the number of elective classes offered at different middle schools. Find the mean.

# OF CLASSES	7	8	9	10
FREQUENCY	2	3	2	1

Unit: Data & Statistics
Student Handout 4

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.

IQR [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. The IQR 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.

1. The data set below represents the number of animals in different exhibits at a zoo.

48, 86, 15, 27, 18, 52, 103

a. Write the data from least to greatest. 15, 18, 27, 48, 52, 86, 103

b. What is the minimum number of animals? 15

c. What is the maximum number of animals? 103

d. What is the median number of animals? 48

e. What is the median of the first half of the data? (first quartile) 18

f. What is the median of the second half of the data? (third quartile) 86

g. What is the interquartile range? 68

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 M.A.D is the average distance that points are "away" from the mean.

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

Use the steps above to complete the table and find the mean absolute deviation in the situation below.

2. Mr. Breedlove recorded the number of minutes he spent traveling to work for the last seven days.

48, 40, 62, 52, 60, 42, 39

# OF MINUTES	DISTANCE FROM MEAN
48	1
40	9
62	13
52	3
60	11
42	7
39	10
TOTAL DEVIATION	54

a. Find the mean of the data.

49

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.

7.7

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

The average distance that each number of minutes is from the mean of 49 is 7.7.

Use measures of variability to answer the questions below.

3. Mrs. Johnson recorded the number of pages five of her students read that day. Complete the blanks below in order to make each statement true.

OF PAGES
20
22
18
19
21

- The minimum number of pages read is ___18___ and the maximum number of pages read is ___22___.
- The interquartile range of ___3___ means that there is a ___small___ amount of variability in the number of pages read.
- The average distance that each number of pages read is from the mean is ___1.2___.
- Since the mean absolute deviation is a ___small___ number, this means that the number of pages read are ___close together___.

Unit: Data & Statistics
Homework 4

Name _____

Date _____ Pd _____

MEASURES OF VARIABILITY

Use your understanding of variability to answer the questions below.

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

108, 53, 47, 22, 56, 62

18

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

12.2, 6.8, 3.4, 8.8

2.7

3. The average daily rainfall in Seattle for the month of June has a mean absolute deviation of 1.3 inches. What conclusion can you make about the average daily rainfall in Seattle for the month of June?

Ex: The average daily amounts of rainfall in Seattle are close to the mean. There is very little spread in the data.

4. If the mean absolute deviation is a really large number, then what does that mean about the data set?

Ex: A large mean absolute deviation means that the data set is not very close to the mean. The data varies greatly from the mean.

5. The number of students in 10 kindergarten classes is shown below.

20, 18, 21, 18, 22, 24, 18, 23, 23, 24

a. What is the median number of students in the kindergarten classes? 21.5

b. What number of students describes the first quartile? 18

c. What number of students describes the third quartile? 23

d. What number of students describes the interquartile range? 5

6. Bryan thinks that a low interquartile range for average daily temperatures means that the temperatures were really low. Do you agree or disagree? Why or why not?

Ex: Disagree. A low interquartile range means that the middle 50% of the data has a small amount of variability in the temperatures, but not necessarily low temperatures.

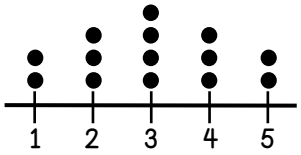
Unit: Data & Statistics
Student Handout 5

Name _____
Date _____ Pd _____

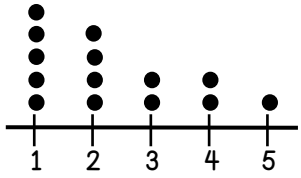
COMPARING DOT PLOTS

SHAPE

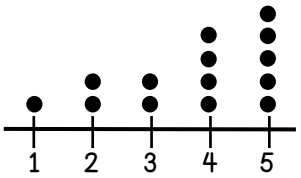
- Data can take on three different shapes:



SYMMETRIC



SKEWED RIGHT



SKEWED LEFT

SPREAD

- The variability, or spread, in the data points describes how far apart the data is from one another. This can also be represented by the _____range_____ or _____interquartile range_____.

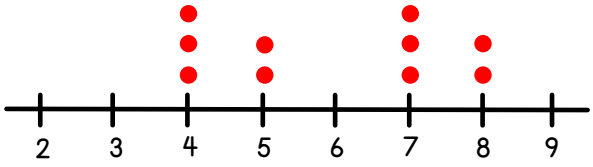
CENTER

- 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.

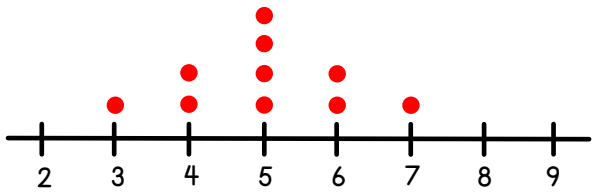
Use your understanding of dot plots to answer the questions below.

1. Two different classes were asked how many minutes each student traveled to school in the mornings.

ROOM 434: 5, 7, 8, 8, 4, 4, 7, 5, 4, 7
ROOM 425: 4, 5, 5, 6, 3, 6, 4, 7, 5, 5



MINUTES TRAVELED TO SCHOOL IN ROOM 434



MINUTES TRAVELED TO SCHOOL IN ROOM 425

2. Describe the shape of the two dot plots above. Which group appears to have the higher number of minutes traveled to school?

Room 434 doesn't have a pattern, while in Room 425 the data is symmetric. Room 434 has a higher number of minutes traveled to school.

3. What is the median of each data set?

a. Room 434: 6 minutes

b. Room 425: 5 minutes

4. How does the median number of minutes in Room 434 compare to that of Room 425?

The median number of minutes traveled to school in Room 434 is 1 minute longer.

Use the information from the dot plots on the previous page to answer the following questions.

5. Mark each of the following statements as true or false based on the information.

false a. The distribution for Room 434 is symmetrical while the distribution for Room 425 is skewed left.

false b. The median number of minutes in Room 425 is greater than the median number of minutes in Room 434.

false c. Room 434 has a smaller range than Room 425.

true d. The distribution for Room 425 is symmetrical while the distribution for Room 434 has no pattern.

6. Set up a table to calculate the mean absolute deviation of each classroom.

Room 434: 5, 7, 8, 8, 4, 4, 7, 5, 4, 7

# OF MINUTES	DISTANCE FROM MEAN
5	0.9
7	1.1
8	2.1
8	2.1
4	1.9
4	1.9
7	1.1
5	0.9
4	1.9
7	1.1
TOTAL DEVIATION	15

MAD: 1.5

Room 425: 4, 5, 5, 6, 3, 6, 4, 7, 5, 5

# OF MINUTES	DISTANCE FROM MEAN
4	1
5	0
5	0
6	1
3	2
6	1
4	1
7	2
5	0
5	0
TOTAL DEVIATION	8

MAD: 0.8

a. What observations can you make about the mean absolute deviation of each classroom?

The MAD of room 434 is 1.5, while the MAD of room 425 is 0.8. Room 425 has data that is closer to the mean. There is slightly more variability in Room 434.

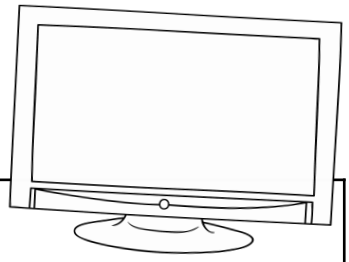
b. Room 434 has a mean average deviation almost twice that of Room 425.

Unit: Data & Statistics
Homework 5

Name _____

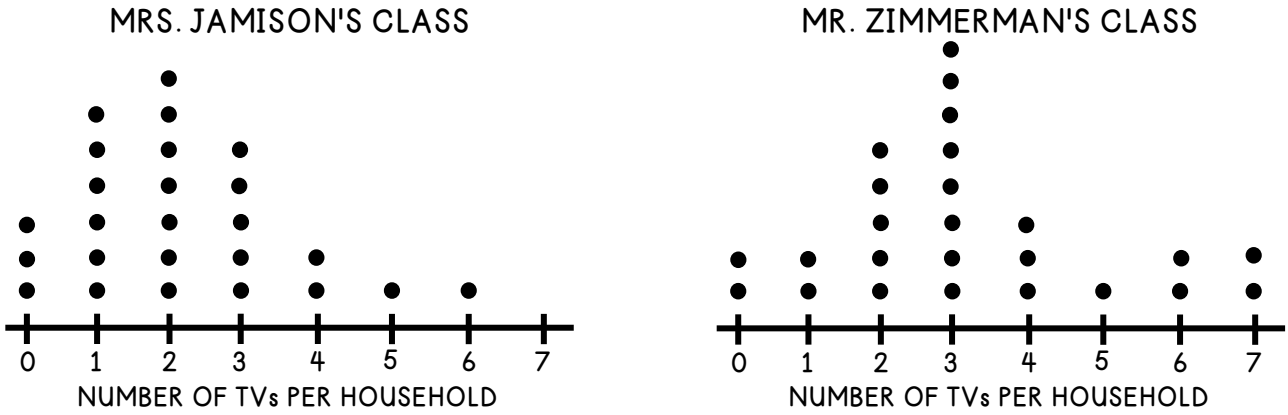
Date _____ Pd _____

COMPARING DOT PLOTS



Use the dot plots to answer the questions below.

The number of TVs per household for Mrs. Jamison's class and Mr. Zimmerman's class is shown below.



1. What is the median number of TVs in Mrs. Jamison's Class?

2 TVs

2. What is the median number of TVs in Mr. Zimmerman's Class?

3 TVs

3. The mean absolute deviation of Mrs. Jamison's class is 0.24. What does this mean?

The data is very close to the center.

4. The mean absolute deviation of Mr. Zimmerman's class is 0.46. What does this mean?

The data from Mr. Zimmerman's class is slightly more spread out than Mrs. Jamison's.

5. Mark each statement below as true or false.

true The range of the number of TVs per household in Mrs. Jamison's class is less than the range of the number of TVs per household in Mr. Zimmerman's class.

false The median number of TVs per household is equal in both Mrs. Jamison's class and Mr. Zimmerman's class.

false The number of TVs per household with the highest frequency for both Mrs. Jamison's class and Mr. Zimmerman's class is 3.

6. When Asher looks at the data, he says that both Mrs. Jamison's class and Mr. Zimmerman's class has data that is skewed right. Do you agree or disagree? Why or why not?

I partially agree with Asher. The data in Mrs. Jamison's class is skewed right, while the data in Mr. Zimmerman's class is more symmetric.

Unit: Data & Statistics
Student Handout 6

Name _____
Date _____ Pd _____

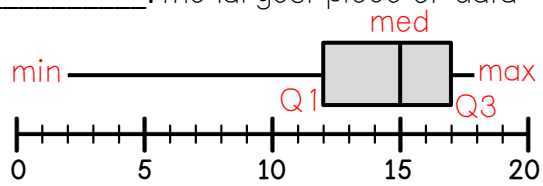
COMPARING BOX PLOTS

BOX PLOTS

- A box plot displays a data distribution using ___five___ key numbers.

  The difference between the first and third quartile is called the ___interquartile range, or IQR___, which is a measure of variability.

  - ___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

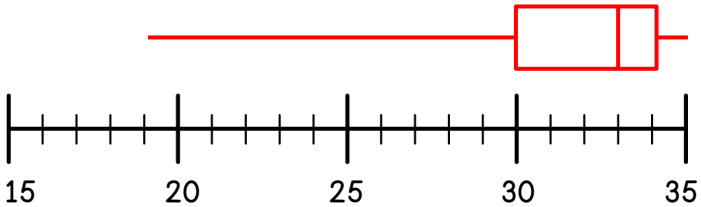


Use the data below to create a five-number summary and sketch a box plot.

1. The following data set represents the number of pages in a book read each day.

34, 32, 19, 35, 33, 30, 34

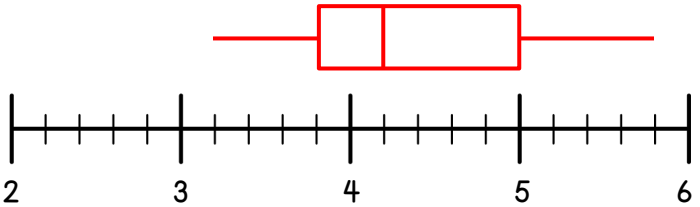
Min: 19
Q1: 30
Med: 33
Q3: 34
Max: 35



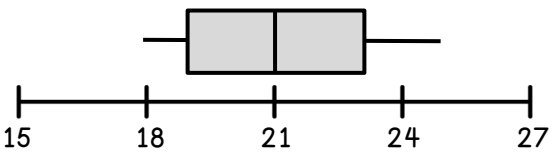
2. The following data set represents the number of miles Monica walked each day.

4.2, 3.8, 4.7, 5.8, 3.2, 4.1, 5

Min: 3.2
Q1: 3.8
Med: 4.2
Q3: 5
Max: 5.8



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

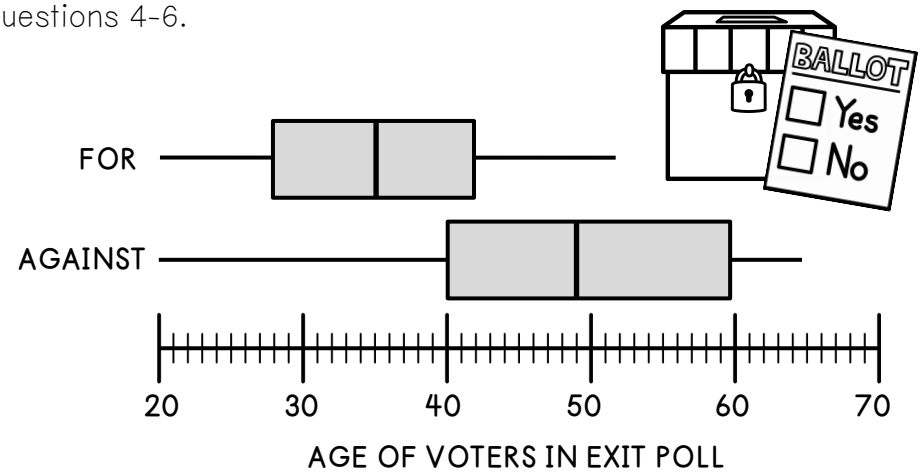


Min: 18 Q1: 19 Med: 21 Q3: 23 Max: 25

Use the box plots below to answer questions 4-6.

In a local election, a school bond is proposed to fund a natatorium.

Alex conducts an exit poll in which he asks voters to share their age and whether they were for or against the school bond. He then displays the data in a box plot.



4. What is the difference in the measure of center of the voters who are for and against the natatorium?

The difference in measure of center is approximately 14 years.

5. Determine the IQR of voters who voted against the natatorium. Determine the IQR of voters who voted for the natatorium.

The IQR of the voters who were against the natatorium is 20 years, while the IQR of the voters who voted for the natatorium is 14 years.

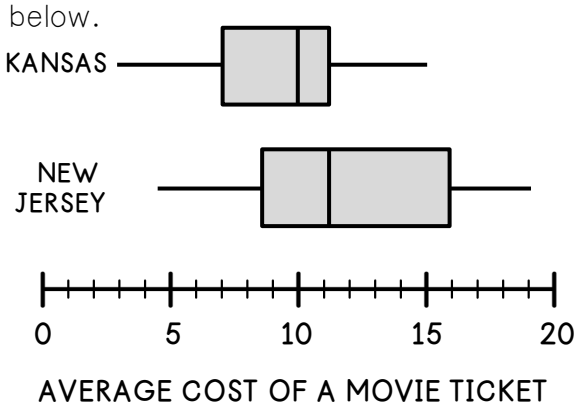
6. Complete the statements below:

a. The range in age of voters for the natatorium was less than the range in age of voters against the natatorium.

b. The IQR of those who voted against the natatorium was almost 1.5 times the IQR of those who voted for the natatorium.

In question 7, compare the two box plots below. Justify your solutions.

7. A survey was conducted to determine the average price of a movie ticket in Kansas and New Jersey. The data was displayed using the box plots shown below.



a. What was the median cost of a movie ticket in each state?

Kansas: $10
New Jersey: $11

b. Which state has a wider range in variability? Why?

New Jersey appears to have a wider range in variability because it has a wider spread of data.

c. A movie goer spends $14 on a movie ticket in New Jersey. What can you conclude about the price the customer paid?

The movie goer spent a few dollars more than the median cost for a ticket.

Unit: Data & Statistics
Homework 6

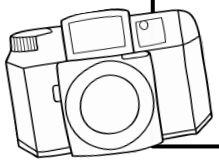
Name _____

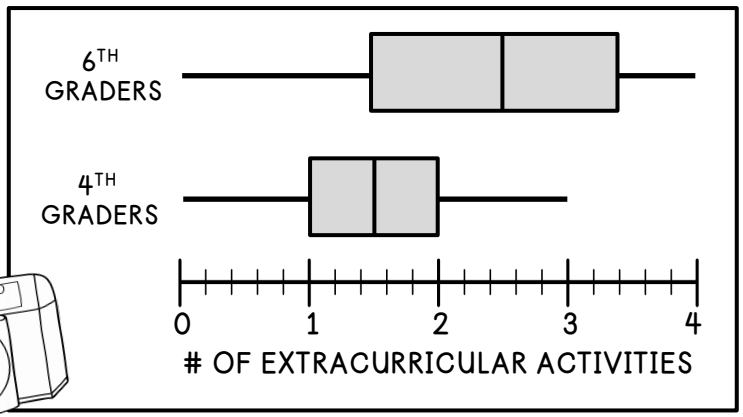
Date _____ Pd_____

COMPARING BOX PLOTS

Use the box plots at the right to answer questions 1-4, and then apply your understanding of box plots in question 5.

The 4th and 6th grade students at Tree Hill Elementary completed a survey on how many extracurricular activities they participate in. The data is shown in the box plots at the right.





1. What is the median number of extracurricular activities for each grade?

6th graders – 2.5
4th graders – 1.5

2. What is the difference in the spread of the data between the two grades?

1 extracurricular activity

3. Determine the variability of the number of extracurricular activities for each grade.

6th graders – 1.9
4th graders – 1

4. Read each of the following statements and mark them as true or false.

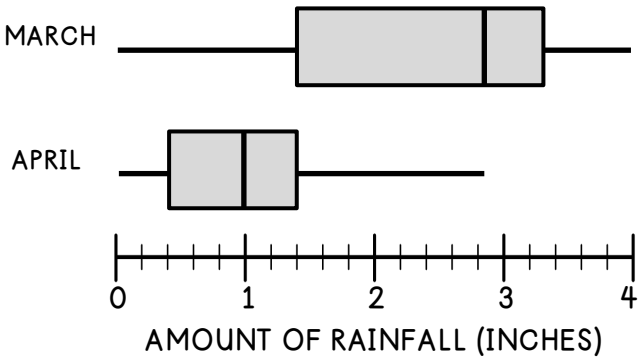
True The third quartile of the data for 4th graders is less than the third quartile of the data for 6th graders.

True The first quartile of the data for 6th graders is equal to the median of the data for 4th graders.

False The maximum number of extracurricular activities for 4th graders is greater than the maximum number of extracurricular activities for 6th graders.

True Both grades have at least one student who does not participate in any extracurricular activities.

5. The boxplots show the amount of rainfall for the months of March and April. Which statement is represented by the data?



A. March has half the variability of April.

B. The third quartile of the data in April is greater than the first quartile of the data in March.

C. The maximum amount of rainfall in April is equal to the median amount of rainfall in March.

D. The difference in the centers of March and April is greater than 2.