Common Core Math 1A

Fall 2016	UNIT 1: ONE VARIABLE STATISTICS	HW
8/29 Monday	Overview of unit 1, categorical vs quantitative data	
8/30 Tuesday	Data collecting activity, 4 ways to describe data, dot plot, frequency tables, histograms	
8/31 Wednesday	Interpreting histograms Histograms on the Calculator.	
9/1 Thursday	Quiz Measures of center (mean and median): compare mean and median with outliers	
9/2 Friday	Measures of spread: standard deviation (sigma x) IQR	
9/5	LABOR DAY HOLIDAY!!	
9/6 Tuesday	Determining upper and lower bounds using the $1.5 \cdot IQR$ rule.	
9/7 Wednesday	Practice calculations, how does changing the data effect our calculations.	
9/8 Thursday	Quiz Box and whisker plots, side by side comparisons	
9/9 Friday	Early Release Work on Projects	
9/12 Monday	Work on Projects	
9/13 Tuesday	Review for Test	
9/14 Wednesday	Unit 1 Stats Test	
9/15 Thursday		

UNIVARIATE STATISTICS

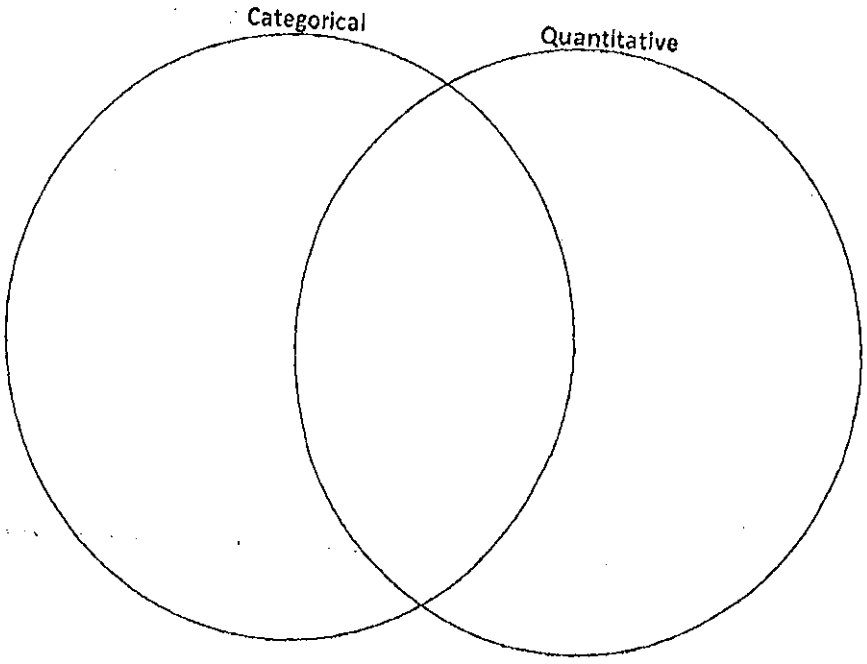
There are two types of data that we tend to collect in the world: Categorical (Qualitative) and Quantitative

Categorical _____

_____

Quantitative _____

_____



Write the letter of the data within the section of the Venn diagram where it is most applicable. Be prepared to justify your choices.

A. Average blood pressure	B. Who received high school diplomas	C. Right-handed or left-handed
D. Favorite TV show	E. Hours spent outdoors each day	F. Time of day you go to bed
G. Age group	H. Whether or not people smoke	I. Favorite food
J. Common majors in college	K. Average calories consumed in a day	L. Gender
M. Population	N. Number of brothers and sisters	O. Height
P. Shoe size	Q. Hours spent doing HW each night	R. Miles from home to school

Common Core Math 1

Name _____

Univariate Data
Categorical vs Quantitative Data

Date _____

Determine whether the following variables are **categorical** or **quantitative** by WRITING THE WORD for each one. If both, write "both."

1. Brand of vehicle purchased by a customer

2. Price of a CD

3. Number of students in a class of 30 who prefer peanut M&Ms over plain M&Ms

4. Phone number of all the students enrolled in school.

5. The height of a 1 year old child.

6. Number of students in a class of 35 who turn in a term paper before the due date.

7. Gender of the next baby born at a particular hospital.

8. Amount of fluid (oz) dispensed by a machine used to fill bottles with soda.

9. Thickness of the gelatin coating of a Vitamin C capsule

10. Brand of computer purchased by a customer

11. State of birth for someone born in the United States.

12. Price of a textbook

13. The zip codes of everyone in this class.

14. Actual weight of coffee in a one pound can.

15. The length of a rattlesnake.

Common Core Math 1

One-Variable Data

Homework: Introduction to Statistics

Name _____

Date _____

1. Determine whether the following data is **categorical** or **quantitative**

   a. The candidate a survey respondent will support in an upcoming election.

   b. The length of time of people's drive to work.

   c. The number of televisions in a household.

   d. The distance kickers for a football team can kick a football.

   e. The number of pages copied in the copy room each day.

   f. The kind of tree in each person's front yard in a neighborhood.

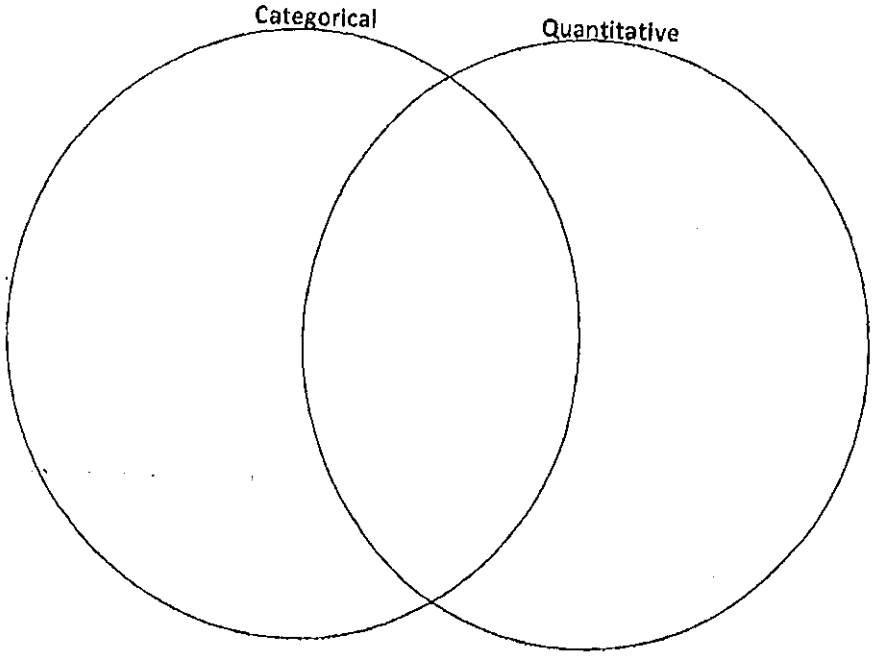
   g. The type of blood a person has.

   h. The jersey numbers of the football team.

   i. The heights of the tallest buildings in the world.

   j. The language spoken by 2000 people coming in to JFK Airport.

2. Come up with three examples of data that fits into each section of the Venn diagram that has not already been given. Be prepared to justify.



Notes

Dot Plot (or Line Plot) _____

_____

Frequency Distribution _____

_____

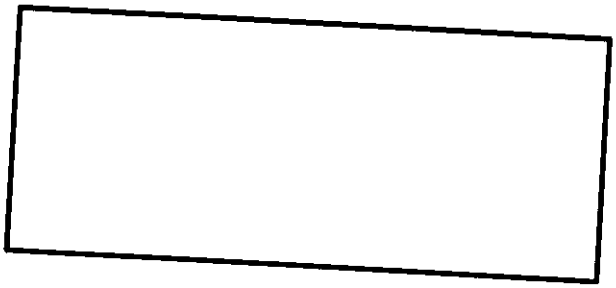
Histogram _____

Skew _____

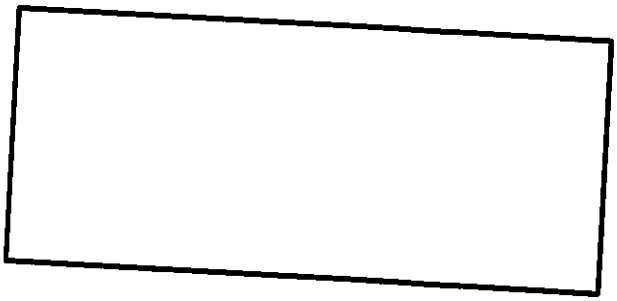
There are _____ characteristics used to analyze a distribution set. They are

_____, _____, _____, and _____.

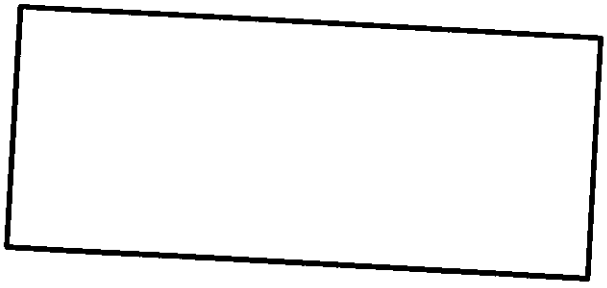
There are _____ shapes.



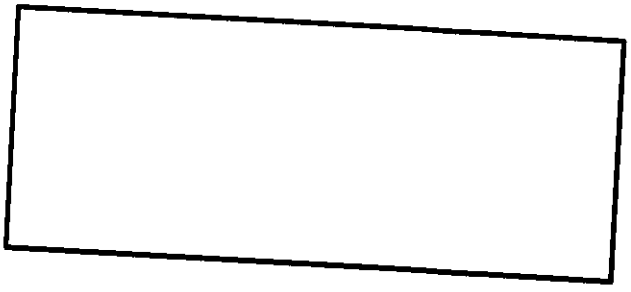
_____



_____



_____



_____

Name _____ Class _____ Date _____

Practice: Frequency Tables, Line Plots, and Histograms

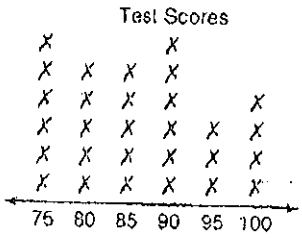
Make a line plot for the data.

1. boxes of juice sold per day:

   26 21 26 24 27 23 24 22

   26 21 23 26 24 26 23

Ms. Makita made a line plot to show the scores her students got on a test. At the right is Ms. Makita's line plot.



2. What does each data item or ✗ represent?

   _____

3. How many more students scored 75 than scored 95?

   _____

4. How many students scored over 85? _____

5. What scores did the same number of students get?

   _____

Nathan asked 24 classmates to estimate the total number of hours (to the nearest quarter hour) they spend doing homework Monday through Thursday. The frequency table below shows their responses.

6. Can you tell from the table how many students do homework for two hours or less? Explain. _____

   _____

7. How many more students do homework for at least 5 h than do homework for less than 4 h? _____

8. Make a histogram for the data. Use the intervals in the table.

Hours Spent Doing Homework

Number of Hours	Frequency
1 – 1.75	1
2 – 2.75	1
3 – 3.75	2
4 – 4.75	6
5 – 5.75	8
6 – 6.75	3
7 – 7.75	2
8 – 8.75	1

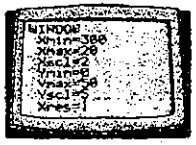
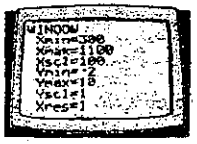
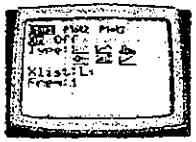
Technology Tip
TI-83 and TI-84

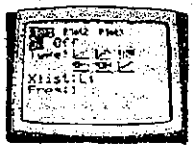
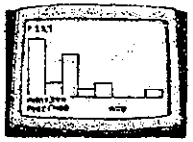
Displaying a Histogram

To display a histogram of data in a list, you need to specify the viewing window for the graph by selecting values for Xmin, Xmax, Xscl, Ymin, Ymax, and Yscl.

- Xmin is the number where you would like the horizontal axis to begin.
- Xmax is the number where you would like the horizontal axis to end.
- Choosing Xscl = 100 sets the width of each bar of the histogram to 100.
- Choosing Ymin = −2 leaves room at the bottom of the screen for the text.
- Choosing Ymax = 10 leaves room for a bar 10 units high.
- Selecting Yscl = 1 produces tick marks on the vertical axis that are 1 unit apart.

The following instructions display a histogram of total calories for Problem 8 on page 82.

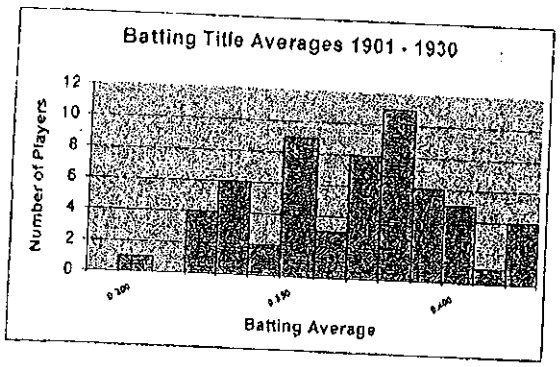
Calculator Commands	Expected Display
To enter data: WINDOW 3 0 0 ENTER	
1 1 0 0 ENTER 1 0 0 ENTER ▷ 2 ENTER 10 ENTER 1 ENTER	
To choose a histogram: 2nd Y= 1 (If the plot is OFF, press ENTER to select ON. Be sure other plots are OFF.)	

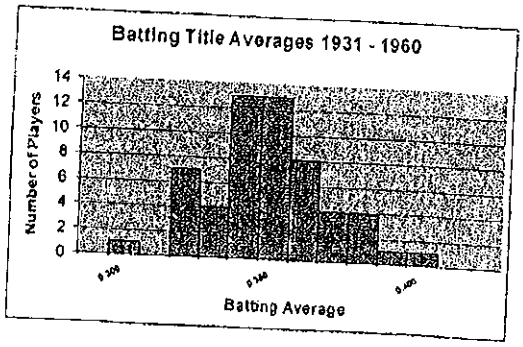
Calculator Commands	Expected Display
Arrow to the histogram next to Type: ENTER	
GRAPH TRACE	

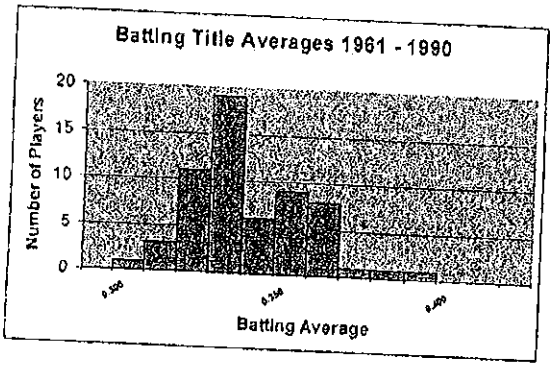
Note: If no vertical axis appears on your graph, press 2nd ZOOM to get the FORMAT menu. Then select AxesOn.

Histogram Worksheet

The 3 histograms below show the batting averages of the winners of the batting title in the major league baseball (for both the American & National leagues) for certain years in the 1900s. Batting average shows the percent (written as a decimal) of the time a certain player gets a hit. A player who has a batting average of 0.405 has gotten a hit in 40.5 % of the times that they were at bat. The batting title is an award given to the player with the highest batting average for a given season. Refer to the histograms as you answer questions 1 – 4.







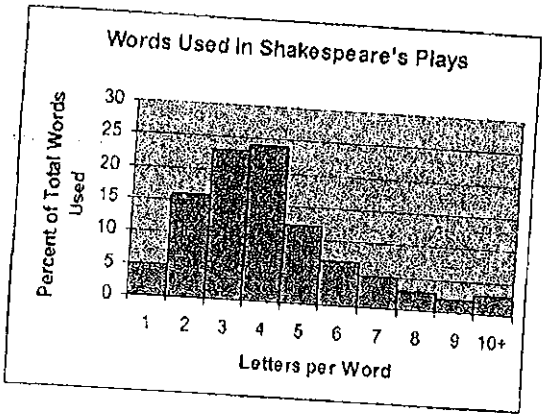
1. How many batting titles were won with a batting average of between 0.300 – 0.350 from 1901 to 1930?

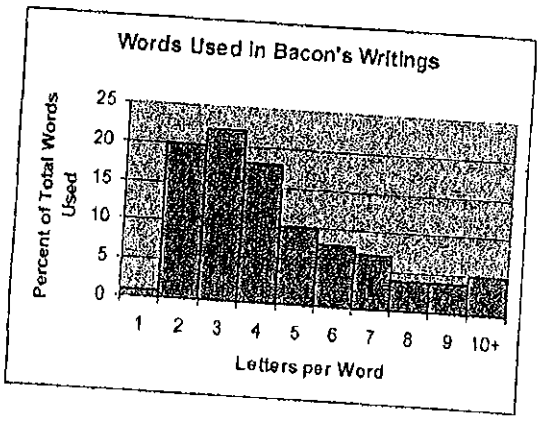
2. How many batting titles were won with a batting average of between 0.300 – 0.350 from 1931 to 1960?

3. How many batting titles were won with a batting average of between 0.300 – 0.350 from 1961 to 1990?

4. As the century progressed, what in general happened to the batting averages of the batting title winners? Justify your response from the data in the histograms.

For questions 5-10, refer to the following 2 histograms. These histograms were made in an attempt to determine if William Shakespeare was really just a pen name for Sir Francis Bacon. (A pen name is a fake name used by another person when writing). A few scholars have had this idea and in order to determine if this was true, a researcher had to count the letters in every word of Shakespeare's plays & Bacon's writing (and you thought you had a lot of homework). Their results are recorded in the histograms below.





5. How would you describe the shape of the histograms?

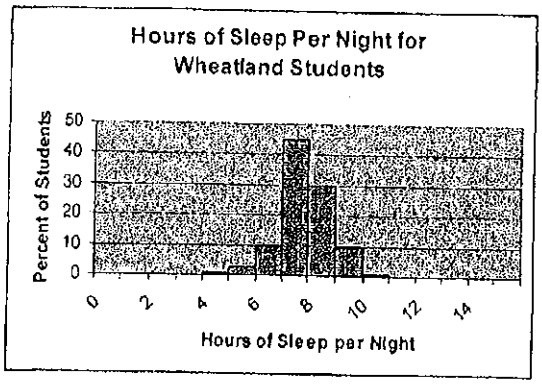
6. What percent of all Shakespeare's words are 4 letters long?

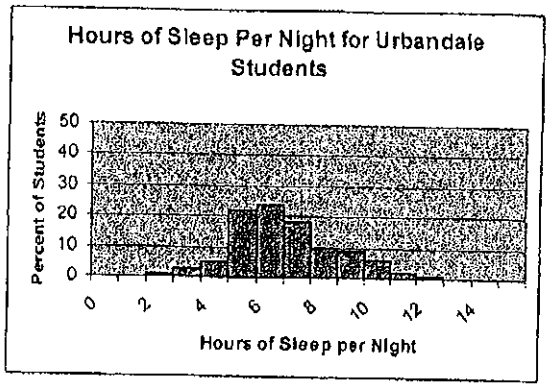
7. What percent of all Bacon's words are 4 letters long?

8. What percent of all Shakespeare's words are more than 5 letters long?

9. What percent of all Bacon's words are more than 5 letters long?

10. Based on these histograms, do you think that William Shakespeare was really just a pen name for Sir Francis Bacon? Explain.





Suppose that the two histograms above show the sleeping habits of the teens at two different high schools. Wheatland High School is a small rural school consisting of 100 students while Urbandale High School is located in a large city and has 3,500 students.

11. About what percent of the students at Wheatland get at least 8 hours of sleep per night?

12. About what percent of the students at Urbandale get at least 8 hours of sleep per night?

13. Which high school has a higher percentage of students that sleep between 9 – 10 hours per night?

14. Which high school has more students that sleep between 9 – 10 hours per night?

15. Wheatland's percent of students who sleep between 8-9 hours a night is _____ % more than Urbandale's percent of students who sleep between 8-9 hours per night.

Year	Age of Winning Actress	Age of Winning Actor
1928	22	42
1929	36	40
1930	28	62
1931	62	53
1932	32	35
1933	24	34
1934	29	33
1935	27	52
1936	27	41
1937	28	37
1938	30	38
1939	26	34
1940	29	32
1941	24	40
1942	34	43

Year	Age of Winning Actress	Age of Winning Actor
1943	24	49
1944	29	41
1945	37	40
1946	30	49
1947	34	56
1948	34	41
1949	33	38
1950	28	38
1951	38	52
1952	45	51
1953	24	35
1954	26	30
1955	47	38
1956	41	41
1957	27	43

16. The charts above show the age of the actress & actor who won the Oscar for best actress or actor during the first 30 years of the Academy Awards. Use the charts to make two histograms on the calculator (one for winning actresses ages & one for winning actors ages) displaying this information. Use bin widths of ten years (0-9; 10-19; 20-29 etc.) Sketch the histograms below.

17. Write a short paragraph discussing what your two histograms reveal.

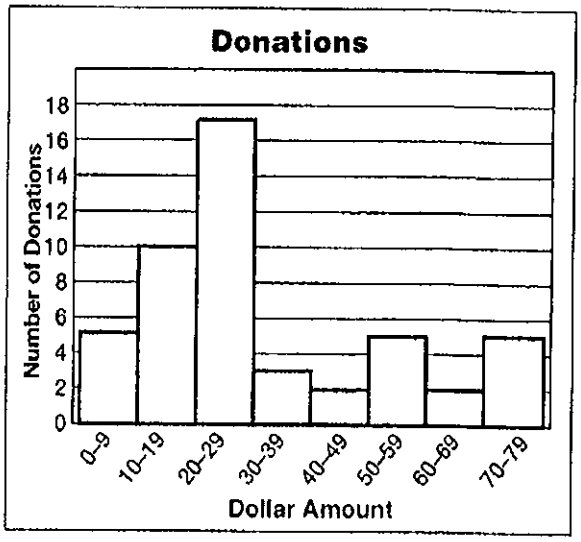
Name _____ Date _____

Practice 7-2

Histograms

Use the histogram for Problems 1–4.

The histogram shows the amount of money in dollars that people donated to a fund to help save the whales.



1. How many people donated between $50 and $59?

_____

2. For which interval did the most people donate money to the fund?

_____

3. How many people donated $50 or more to the fund?

_____

4. Looking at the histogram, do you think that people are more likely to donate large or small amounts of money? Explain.

_____

5. Make a histogram for the following information about the years of birth of members of the O'Brien family.

Years of Birth	Number of family members
1960–1969	|
1970–1979	||||
1980–1989	𝍸 |
1990–1999	𝍸 ||
2000–	||

Test Prep

6. If the histogram above were redone using the intervals 0–19, 20–39, 40–59, and 60–79, which interval would have the most donations?

A 0–19 C 40–59

B 20–39 D 60–79

7. How can a histogram be useful?

_____

_____

_____

Use with text pages 176–177.

Describing Data Numerically: Measures of Center

Median and Mean

Vocab.

Mean - numerical measure of center that is the _____ of the data

Median – numerical measure of center that describes the _____ value of a data set

Measures of Center:

What is a typical value?

Mean: sum of a set of data divided by the number of data (Do not round unless told to do so.)

Symbol for mean used in statistics _____

Median: middle value or mean of the two middle numbers when data is arranged in numerical order

Remember:

What are the four ways to describe distribution of data in context?

1.

2. 3.

4.

Median and mean are measures of _____ tendency.

How do I know which measure of central tendency to use?

	Mean	Median
Use when	To describe middle of a set of data when it DOES NOT have an outlier	To describe middle of a set of data when it DOES have an outlier
Advantages	• most popular in fields of business, engineering and computer science • Unique- only one answer • Useful when comparing data sets	• Extreme values or outliers do not affect the median as strongly as they do the mean
Disadvantages	• Affected by extreme values or outliers	• not as popular as mean

What happens to the mean of a data set when you add one or more data values that are outliers?

What happens to the mean of a data set when you add values that cluster near one end of the original data set? Explain.

Investigation: Experimenting with the Median

We can use the median of a set of data to describe what is typical about the distribution. Let's use this *measure of center* to describe the distribution of names of comediens. Below are twelve names, some stage names. Count the number of letters in each name and write that number in the column labeled "Number of Letters". Do not count spaces. Then, rank the names in order of shortest to longest. The shortest name is rank 1 and the longest name is rank 12.

Name	Number of Letters	Rank
Eddie Murphy		
Gabriel Iglesias		
Zach Galifianakis		
Dave Chappelle		
Jeff Dunham		
Kat Williams		
Margaret Cho		
Will Ferrell		
Carlos Mencia		
Daniel Tosh		
Kevin Hart		
Anjelah Johnson		

Order the cards from shortest name length to longest name length, and identify the median of the data. What is the median?

1. Remove two names from the original data set so that:

   a) the median stays the same. What names did you remove?

   b) the median increases. What names did you remove?

   c) the median decreases. What names did you remove?

2. Now, add two names to the original data set so that:

   a) the median stays the same. How long are the names?

   b) the median increases. What names did you add?

   c) the median decreases. What names did you add?

3. How does the median of the original data set change if

   a) a name with 50 letters is the longest name?

   b) a name with 2 letters is the shortest name?

   c) Would outliers affect the median?

4. Now calculate the mean of the original data set. Mean or $\bar{x}$ =

5. Now calculate the mean if the longest name was 50 letters long. Mean or $\bar{x}$ =

6. Compare the means from #4 and #5. Why do you think the result took place?

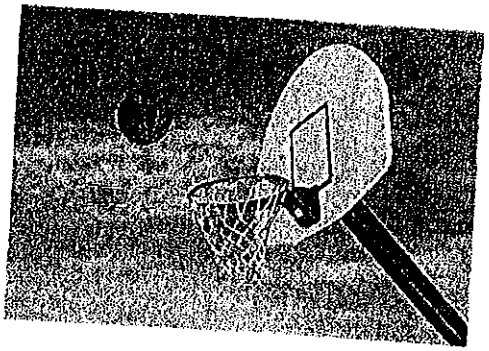
7. Make a conjecture on the effect of outliers on the mean.

8. If the longest name was 50 letters long, which measure of center, mean or median, would be the most appropriate to use to describe the data?

Discuss the following with your partner or group. Write your answers on your own paper. Be prepared to share your answers with the class.

The heights of Washington High School's basketball players are: 5 ft 9in, 5 ft 4in, 5 ft 7 in, 5ft 6 in, 5 ft 5 in, 5 ft 3 In, and 5 ft 7 in. A student transfers to Washington High and joins the basketball team. Her height is 6 ft 10in.

1) What is the mean height of the team before the new player transfers in? What is the median height?
2) What is the mean height after the new player transfers? What is the median height?
3) What effect does her height have on the team's height distribution and stats (center and spread)?
4) How many players are taller than the new mean team height? How many players are taller than the new median team height?
5) Which measure of center more accurately describes the team's typical height? Explain.



Choose 3 of these questions and explain how you answered them

1. The mean of 3 test scores is 80. What is the sum of these test scores?
2. The mean of 5 hourly wages is $5.95. What is the sum of these wages?
3. The mean of a set of numbers is 54. The sum of the numbers is 1.56. How many numbers are in the set?
4. The mean of a set of numbers is 0.39. The sum of the numbers is 1.56. How many numbers are in the set?
5. The mean of a set of numbers is 123. The sum of the numbers is 2,214. How many numbers are in the set?
6. Gini's test scores are 95, 82, 76, and 88. What score must she get on the fifth test in order to achieve an average of 84 on all five tests?
7. Laura's math test grades are 87, 93, 89, and 85. What grade must she get on the fifth test in order to get a mean of 90 for the term?
8. The Lachance family must drive an average of 250 miles per day to complete their vacation on time. On the first five days, they travel 220 miles, 300 miles, 210 miles, 275 miles, and 240 miles. How many miles must they travel on the sixth day in order to finish their vacation on time?
9. A small company has a limited budget for salaries. It can afford to pay an averages of $35,000 a year to its employees. If the first 5 employees are paid $37,000, $38,000, $33,000, $39,000, and $29,000, then how much money can they pay the sixth employee without exceeding their budget?

Standard Deviation

Standard deviation shows _____

_____

- Large standard deviations mean _____

  _____
- Small standard deviations mean _____

  _____
- Often symbolized by the lowercase greek letter sigma, which is denoted as σ.

92, 88, 80, 68, 52
Mean, $\bar{x}$ -
Std. Dev. (σ_x) –

92, 92, 92, 52, 52
Mean, $\bar{x}$ –
Std. Dev. (σ_x) –

What does the standard deviation tell you about the data sets?

Class C: 77, 76, 76, 76, 75
Estimate the standard deviation for Class C.

a) Standard deviation will be less than 17.4.
b) Standard deviation will be greater than 21.9.
c) Standard deviation will be between 17.4 and 21.9
d) Can not make an estimate of the standard deviation.

1. Calculate the mean and standard deviation of

(a) 14 15 18 20 23 18

(b) 41 45 34 45 46 47 50

2. The costs of a can of diet coke in 6 different shops are

47¢ 49¢ 50¢ 44¢ 48¢ 44¢

Calculate the mean and standard deviation of these costs.

3. The prices of a bag of sugar in 6 different shops are

86¢ 88¢ 84¢ 79¢ 81¢ 86¢

(a) Calculate the mean and standard deviation of these prices.

(b) In a different set of 6 the same bag of sugar has a mean price of 87¢ and a standard deviation of 5.2¢.

Make two comparisons between the prices in the two sets of shops.

4. (a) The marks of 7 pupils in an advanced higher math exam were

77 67 43 90 66 93 75

Calculate the mean and standard deviation of these marks.

(b) Another group of 7 pupils who sat the same exam had a mean of 78 and a standard deviation of 3.2.

Make two comparisons of the marks of the two groups.

6. The number of points scored by a football team over 7 games were

34 26 20 23 21 18 26

Calculate the mean and standard deviation of these scores.

7. (a) The number of pupils in 7 third year classes in a secondary school are

25 24 28 22 24 30 22

Calculate the mean and standard deviation of the class sizes.

(b) In the same school the mean and standard deviation of the number of pupils in 7 fourth year classes are 22 and 4.4 respectively.

Make two comparisons between the class sizes in third year and in fourth year.

8. Scientists are studying the differences between crocodiles and alligators.
(a) The lengths of 6 crocodiles are recorded in feet. The results are shown below.

18.2 23.0 17.3 22.0 20.8 18.1

Calculate the mean and standard deviation of these lengths.

(b) The lengths of 6 alligators are recorded. The results give a mean of 16.8 feet and a standard deviation of 1.85 feet.

Make two valid comparisons between the lengths of the crocodiles and the alligators.

Name: _____

Date: _____

Mean and Standard Deviation

Remember the mean of a set of data is a measure of center. It is a value that represents an equal share between individuals in the population. The standard deviation is a way of averaging the amount of variance from the mean. **Answer each prompt in a complete sentence.**

1. The following data represents the grades earned on a recent test for 7 students. Calculate the mean and standard deviation using your calculator.

Test 1:	**77**	**78**	**50**	**72**	**76**	**72**	**75**

$\bar{x}$:_____ σ_x :_____

2. The same group of students took another test and scored the following. Calculate the mean and standard deviation.

Test 2:	**75**	**76**	**60**	**70**	**75**	**71**	**73**

$\bar{x}$:_____ σ_x :_____

a. Compare Tests 1 and 2 by the information you collected. Also, speak on which class was more consistent with it's scores.

3. Based off what you found explain the differences in the standard deviations, and what that tells us in terms of the consistency of the test scores.

4. If the same students scored the following scores, WITHOUT using a calculator, what would the mean and standard deviation be?

Test 3:	**80**	**80**	**80**	**80**	**80**	**80**	**80**

$\bar{x}$:_____ σ_x :_____

Box plots, IQR, and Outliers

Vocabulary :

Box plot: A box plot is graph of quantitative data built using the _____number summary of the data. A box plot is also called a _____ and _____ plot. It is made by making _____ lines above three numbers on a number line- the first _____, the median, and the _____ quartile and then connecting the lines to form a _____. Minimum and maximum _____are represented by dots and are connected to the box. These lines are called the _____.

Interquartile Range: measure of spread; $Q_3 - Q_1$; sensitive to outliers

***Modified Box plot:**

- box plot that indicates ***outliers*** (if any) by representing them as dots ***not connected*** to the box
- whiskers connect the box to the lowest and/or highest data values that are *not* outliers

Outlier: A data value that does _____ the overall pattern of the data distribution;
a value that is more than $Q_3 + 1.5(IQR)$ or less than $Q_1 - 1.5\,(IQR)$.

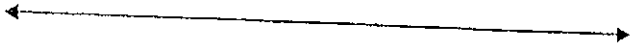
Example 1: Data: -3, 4, 6, 7, 9, 13, 14 (Med = 7; Q_1 = 4; Q_3 = 13)

13 + 1.5(13-4) or 4 - 1.5(13-4)

13 + 1.5(9) 4 - 1.5(9)

more than 26.5 less than -9.5

Boxplots
(no outliers)



Example 2: Data: 4, 8, 8, 8, 9, 10, 10, 14 (Med = 8; Q_1 = 8; Q_3 = 10)

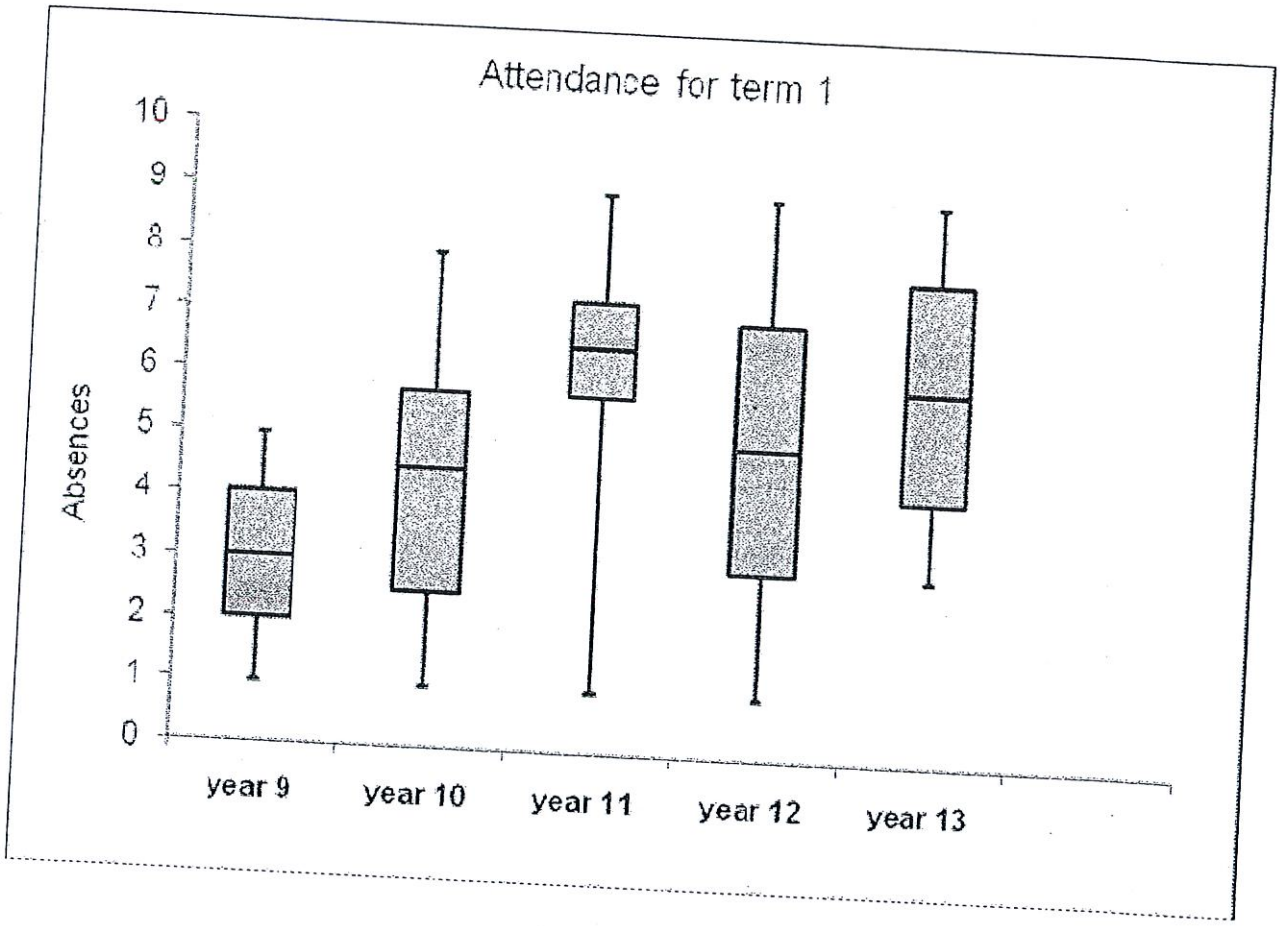
Boxplots
(with outliers)

Are there any outliers?

Find lower limit (lower bound):
Q1-1.5(IQR)

Find upper limit (upper bound):
Q3+1.5(IQR)

Understanding Box Plots



1. What percent of the absences in year 10 was less than 9? _____
2. Which year had the largest interquartile range? _____
3. What was the median number of absences in year 13? _____
4. Which year had a range of seven absences? _____
5. Which year had data that was skewed left? _____
6. What percent of the absences in year 9 was at least two? _____

Outliers and the $1.5 \times IQR$ Rule

1. Ms. Sanchez gave a test to one of her algebra classes. Here are the scores:

83, 86, 91, 78, 80, 33, 75, 91, 72, 82, 88, 84, 93, 99, 74, 79, 90

a. Look at the list of scores. Is there one score in the list that looks very different from the others? Which one? Explain.

b. One student in class was absent for the four days before the test was given. Which score most likely belonged to this student? Explain.

c. Sort the data and record the five-number summary in the chart below (You can use the 1-Var Stats command on your calculator to find these statistics).

	minimum	Q_1	median	Q_3	maximum
Algebra Test Scores					

d. Find the **interquartile range** (IQR). The interquartile range equals $Q_3 - Q_1$.

e. Statisticians have observed that in most data sets, almost all data values lie between a "lower bound" and an "upper bound." The bounds are at the first quartile minus 1.5 times the IQR, and the third quartile plus 1.5 times the IQR. Calculate the bounds according to the formulas given below:

Lower bound = Q1 - 1.5 × IQR = _____

Upper bound = Q3 + 1.5 × IQR = _____

f. Any data point that does not lie within the two bounds is called an **outlier**. Which test score is an outlier?

Switching Teams

2. When LeBron James played for the Cleveland Cavaliers during the 2009-2010 season, was he an outlier? Here are the statistics for all players on the team that year.

Rank	Player	Points per Game
1	LeBron James	29.7
2	Antawn Jamison	15.8
3	Mo Williams	15.8
4	Shaquille ONeal	12.0
5	Sebastian Telfair	9.8
6	Delonte West	8.8
7	Anderson Varejao	8.6
8	J.J. Hickson	8.5
9	Zydrunasllgauskas	7.4
10	Anthony Parker	7.3
11	Daniel Gibson	6.3
12	Jamario Moon	4.9
13	Jawad Williams	4.1
14	Leon Powe	4.0
15	Danny Green	2.0
16	Darnell Jackson	0.8
17	Cedric Jackson	0.2
18	Coby Kari	0.0

a. Record the five-number summary of the points per game in the chart below:

	minimum	Q_1	median	Q_3	maximum
Points Per Game					

b. Find the IQR.

c. Find the lower and upper bounds:

$Lower\ bound = Q1 - 1.5 \times IQR =$ _____

$Upper\ bound = Q3 + 1.5 \times IQR =$ _____

d. Any data value outside the bounds is an outlier. Are there any outliers on this team? If so, who?

3. Now apply the same analysis to the Miami Heat for the 2010-1011 season. Here are the statistics.

Rank	Player	Points per Game
1	LeBron James	26.7
2	Dwyane Wade	25.5
3	Chris Bosh	18.7
4	UdonisHaslem	8.0
5	Mike Bibby	7.3
6	Eddie House	6.5
7	Mario Chalmers	6.4
8	James Jones	5.9
9	Carlos Arroyo	5.6
10	Mike Miller	5.6
11	ZydrunasIlgauskas	5.0
12	Erick Dampier	2.5
13	Juwan Howard	2.4
14	Joel Anthony	2.0
15	Jamaal Magloire	1.9
16	Jerry Stackhouse	1.7
17	Dexter Pittman	1.0

a. Record the five-number summary of the points per game in the chart below:

	minimum	Q_1	median	Q_3	maximum
Points Per Game					

b. Find the IQR.

c. Find the bounds:

$Lower\ bounds = Q1 - 1.5 \times IQR =$ _____

$Upper\ bounds = Q3 + 1.5 \times IQR =$ _____

d. Any data value outside the bounds is an outlier. Are there any outliers on this team? If so, who?

Practice Calculating Values

must show all calculations

1. Find the five number summary: 23, 27, 18, 29, 24, 26, 30, 25, 28

2. Determine the mean: −4, −5, 7, 1, 3, −6, 4, −8, 11, 12

3. Determine the value of the 3rd (upper) quartile: 21, 29, 22, 28, 24, 26, 24, 21, 30

4. Determine the interquartile range (IQR): 3, 2, 2, 1, 3, 1, 2, 1, 1, 3, 2, 7, 4, 2, 3

5. Determine the lower bound: 15, 12, 19, 15, 17, 16, 16, 18, 17, 19

6. Determine the upper bound: 25, 25, 30, 26, 29, 26, 25, 22, 20, 21, 29

7. Determine the measures of center: 13, 20, 14, 19, 15, 13, 18, 14, 17, 14

8. Determine outliers (if any): 8, 6, 8, 9, 7, 8, 6, 8, 9, 9, 9, 8, 8, 8, 9

How does changing values affect the measures of center/spread?

Explain how each situation affects each measure of center and measure of spread. You might want to create a small data set to help you explain and determine your responses

1) Add 3 to every value.

2) Remove the largest three values.

3) Multiply the data set by 3.

4) Remove the smallest two numbers.

5) Insert a number smaller than the minimum and insert a number larger than the maximum.

Investigation
Effects on Center and Spread

Name: _____
Date: _____

We are going to investigate what happens to the center and spread when your original set of data is changed. Here is your list of data: 34 26 20 23 21 18 26

1. Calculate the following:

   Mean ($\bar{x}$): _____ Std. Dev (σ_x): _____ Med: _____ IQR: _____

2. Now add 5 to every value in the <u>original</u> list and recalculate the stats.

   Mean ($\bar{x}$): _____ Std. Dev (σ_x): _____ Med: _____ IQR: _____

3. What do you notice about the stats now?

4. Now double every value from the <u>original</u> set. Recalculate the stats.

   Mean ($\bar{x}$): _____ Std. Dev (σ_x): _____ Med: _____ IQR: _____

5. What do you notice about the stats?

Without using your calculator answer the next two questions.

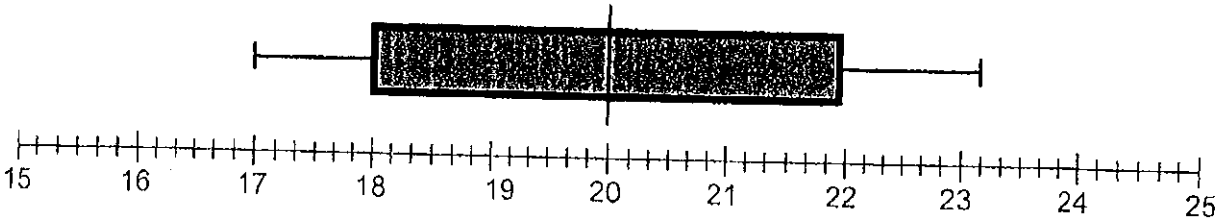
6. If the <u>original</u> data was increased by 8 what would happen?

7. If the <u>original</u> data was tripled, what would happen?

Name _____ Date _____

Making and Understanding Box and Whisker Plots- Independent Practice Worksheet

Complete all the problems.



1. What is the median?

2. What is the lower quartile ?

3. What is the Upper quartile ?

4. What is the minimum value?

5. What is the maximum value?

6. The above five values are called the _____.

7. What is the range of the data?

8. What percentage of data is below upper quartile?

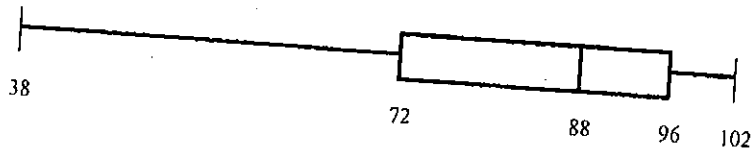
9. What percentage of data is located between the lower quartile and the median?

10. What percentage of data is above the median?

Box & Whisker Worksheet

For questions 1 – 6, refer to the box & whisker graph below which shows the test results of a math class.

Test Scores (as %) for 6th Period



_____ 1. What was the high score on the test?

_____ 2. What percent of the class scored above a 72?

_____ 3. What was the median score on the test?

_____ 4. What percent of the class scored between 88 & 96?

5. Do you think that this test was too hard for the students? Explain.

_____

_____

6. Would you expect the mean to be above or below the median? Explain.

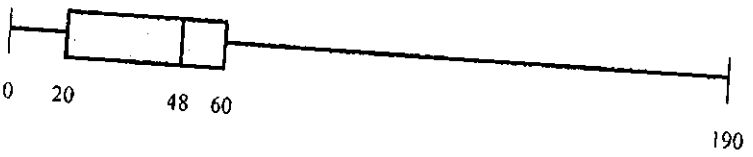
_____

_____

_____

For questions 7 – 11 refer to the box & whisker graph below that shows how much time was spent per night on homework for sophomore class at a certain high school during September.

Average Minutes Per Night Spent On Homework



_____ 7. What percent of the sophomores spend more than 60 minutes on homework per night?

_____ 8. What is the range of times that the middle 50% of the sophomores spend on homework per night?

_____ 9. How many sophomores do not do homework?

_____ 10. What percent of the sophomores spend less than 20 minutes per night on homework?

11. Would you expect the mean number of minutes per night to be higher or lower than the median? Explain.

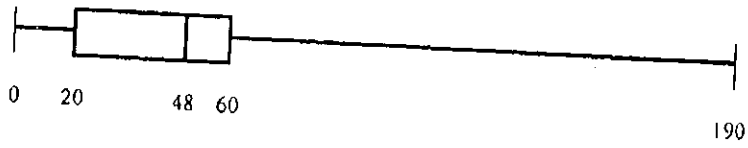
_____

_____

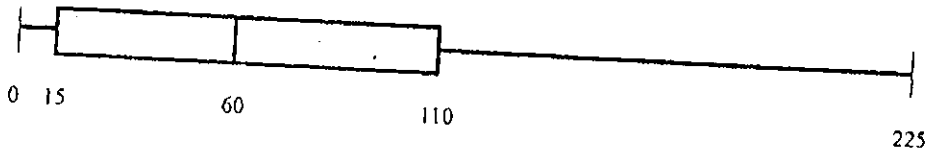
_____

For questions 12 – 23, refer to the box & whisker graphs below that compare homework time per night with TV time per night for the same group of sophomores.

TV & Homework Minutes per Night



Homework Time



TV Time

_____ 12. What percent of the sophomores watch TV for at least 15 minutes per night?

_____ 13. What is the 3rd quartile for the TV time data?

14. Is it more common for a sophomore at this high school to spend more than 1 hour on homework or more than 1 hour watching TV? Explain.

_____

_____

_____

For questions 15 – 23, identify if each statement is true, false, or cannot be determined.

_____ 15. Some sophomores didn't watch TV that month.

_____ 16. The TV box & whisker graph contains more data than the homework graph.

_____ 17. 25% of the sophomores spend between 48 & 60 minutes per night on homework.

_____ 18. 15% of the sophomores didn't watch TV that month.

_____ 19. In general, these sophomores spend more time watching TV than doing homework.

_____ 20. The TV data is more varied than the homework data.

_____ 21. The ratio of sophomores who spend more than 110 minutes per night watching TV to those who spend less is about 2:1.

_____ 22. 225 sophomores watch TV.

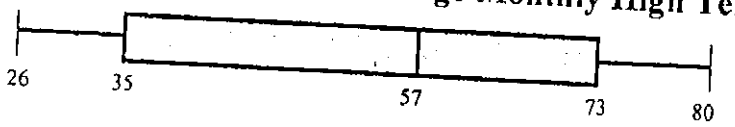
_____ 23. Twice as many sophomores watch TV for more than 1 hour than do homework for more than 1 hour.

24. Suppose that one family kept track of how many DVDs they rented each month for a two year period. The numbers for each month are shown in the table below. Make a box & whisker graph from this data.

J	F	M	A	M	J	J	A	S	O	N	D	J	F	M	A	M	J	J	A	S	O	N	D
3	5	2	8	1	5	0	3	6	4	9	15	3	6	4	1	10	3	8	7	2	9	0	11

For question 25, refer to the box & whisker graphs below that show the average monthly high temperatures for Milwaukee, Wisconsin & Honolulu, Hawaii.

Average Monthly High Temperatures



Milwaukee

Honolulu

25. Write a short paragraph comparing the temperatures in both cities.

Common Core Math 1
Univariate Data
Comparing and Contrasting

Name _____
Date _____

The miles per gallon for city travel of ten cars and ten SUVs are given below:

Cars	MPG	SUVs	MPG
Geo Metro	46	Jeep Grand Cherokee	14
Honda Civic CX	42	Ford Explorer	21
Hyundai Excel GS	29	Chevy Silverado	21
Mazda 323	29	Toyota Tacoma	22
Plymouth Sundance	26	Nissan Frontier	16
Saturn SL	28	Chevy Suburban	12
Eagle Summit	31	GMC Yukon	12
Nissan Sentra E	29	Ford F150	13
Ford Festiva GL	35	Jeep Wrangler	18

For each category, find the following information

1) The minimum MPG:
2) The maximum MPG:
3) The range:
4) The mean:
5) The median:
6) The lower quartile:
7) The upper quartile:
8) The interquartile range:
9) The standard deviation:
10) Outliers:

	Cars	SUV's
1)		
2)		
3)		
4)		
5)		
6)		
7)		
8)		
9)		
10)		

Compare and contrast the data

1. How are they different, how are they the same?

2. Compare and contrast the similarities and differences between the measures of center and spread between the two data sets.

3. Compare and contrast outliers or the absence of outliers between the two data sets.

4. Compare and contrast the distribution of the two data sets. Are they both skewed, are they symmetrical, are they complete opposite of each other, etc.?

5. Explain how your data would change if you were to combine your two categories.

Common Core Math I

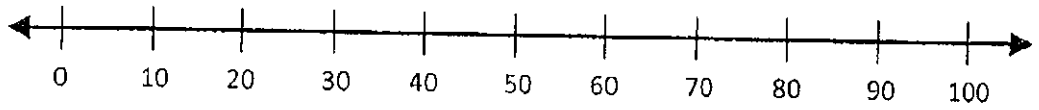
Name_____

Unit 1: One-Variable Data
Comparing Data Sets Practice

Listed below are the quality rating values of all natural peanut butters:
34, 40, 52, 57, 57, 60, 60, 63, 67, 69, 69, 69, 71, 89

The data values for all regular peanut butter are as follows:
11, 23, 23, 26, 29, 31, 31, 33, 34, 34, 35, 40, 40, 43, 45, 46, 49, 54, 54, 60, 76, 83, 83

1) Construct side-by-side boxplots for the two types of peanut butters. Make a sketch below.



2) Record the five-number summary for each in the table below. Calculate the IQR for each set of data and record in the table.

Statistic	Natural Peanut Butter	Regular Peanut Butter
Min		
Q_1		
M		
Q_3		
Max		
IQR		

3) Calculate the mean and standard deviation for each set of data and record below.

Statistic	Natural Peanut Butter	Regular Peanut Butter
$\bar{x}$		
σ		

4) Which measure of center and spread would be most appropriate to use to describe these two sets of data? Explain.

5) Compare the two data sets in context. Be sure to address shape, center, spread, and outliers. Which type of peanut butter is better?

Unit 1 Review

1. List the measures of center.

2. List the measures of spread.

3. List the 4 ways to describe a data distribution.

4. Sarah has been clothes shopping each Saturday for the past 6 weeks. She spent the following amounts: $109, $72, $99, $15, $99, and $89.

   a. Calculate the mean and median of Sarah's purchases.

   Mean _____ Median _____

   b. Which measure of central tendency would Sarah tell her parents to convince them that she is not spending too much money on clothes? Explain.

   c. Which value would Sarah tell her parents to convince them that she needs an increase in her allowance? Explain.

5. Find the interquartile range (IQR) of the data: 61, 65, 65 ,66, 72, 75, 77, 79, 81, 85, 89, 92, 99

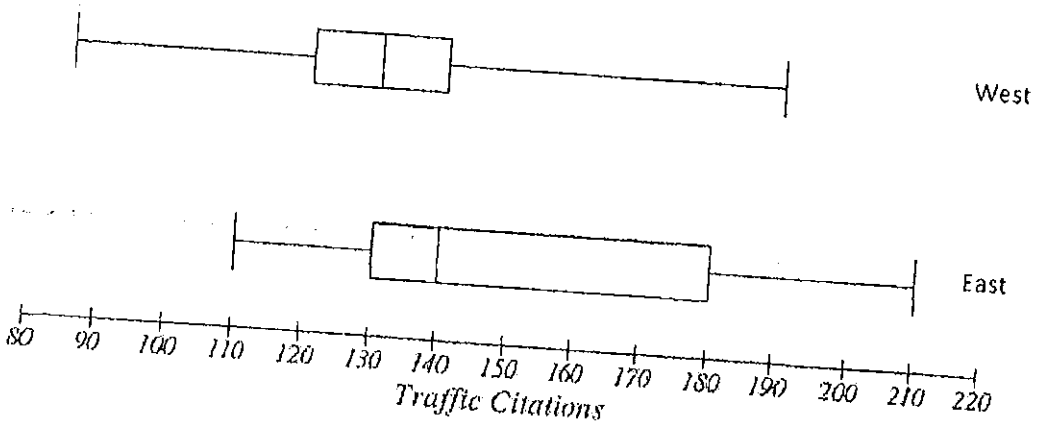
   A. 77 B. 87 C. 65.5 D. 21.5

6. On last week's math test, Mrs. Smith's class had an average of 83 points with a standard deviation of 8 points. Mr. Tucker's class had an average of 78 points with a standard deviation of 4 points.

   a) Which class was more consistent with their test scores? _____

   b) How do you know? _____

Use for 7-9. The number of traffic citations from two police stations (West and East) were recorded over a two-week period. Using the box plots below, choose the letter of the best answer.



7. What is the best estimate of the difference in the greatest number of citations given by each department in one day?

A. 10 B. 20 C. 30 D. 35

8. What is the difference in the median number of citations between the two stations?

A. about 10 B. about 15 C. about 22 D. about 40

9. Which statement is NOT true?

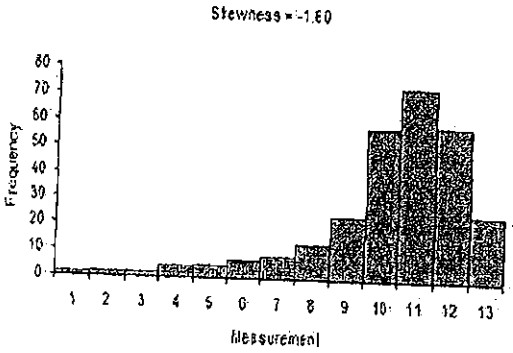
A. East gave the greatest number of citations in one day.
B. East gave the least number of citations in one day.
C. East has a greater IQR than West.
D. East has the greater median number of citations in one day.

10. Create a frequency distribution table and histogram for the following set of data. Make sure to LABEL YOUR GRAPH.

Age (in months) of First Steps			
13	9	12	11
10	8.5	14	9
12.5	10	13.5	9.5
6	7.5	10.5	9
8	11.5	10	12
10.5	11	13	12.5

Interval (age)	Tally	Frequency
6-6.9		
7-7.9		

11. Describe the data distribution below. Be sure to comment on shape, possible best measure of center, possible best measure of spread, and outliers.



Use the following table of data for 12-15

25	44	13
20	45	2
37	25	34
41	19	63

12. Find the interquartile range (IQR) of the data.

13. Find Q1 of the data.

14. Find the standard deviation of the data.

15. What is the range of the data?