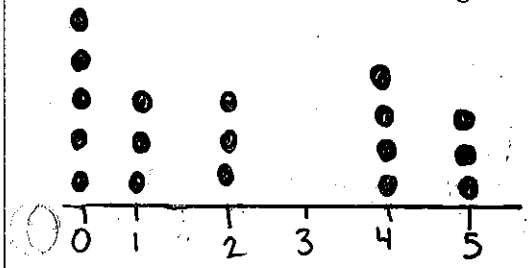
Name: _____ Date: _____

Graphical Displays for Data Homework

Kirsten plays softball in the spring. Each game, she records the number of times she reaches first base without being called out. Use the data in the table to solve problems 1 -5.

Game	Number of times at first	Game	Number of times at first
1	5 0	10	~~0~~ 2
2	~~1~~ 0	11	~~1~~ 2
3	~~2~~ 0	12	1 4
4	~~0~~ 0	13	~~0~~ 4
5	~~2~~ 0	14	5 4
6	~~1~~ 1	15	5 4
7	~~4~~ 1	16	4 5
8	~~4~~ 1	17	~~0~~ 5
9	~~0~~ 2	18	4 5

1. Create a dot plot showing the number of times Kirsten reached first base in each game.



2. Find the minimum, maximum, first quartile, and third quartile of the data set.

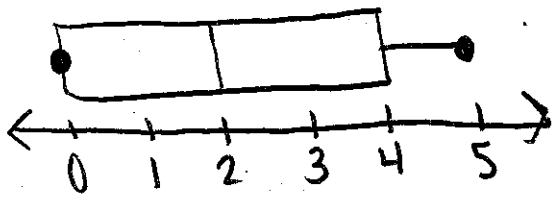
a. Minimum: 0

b. Maximum: 5

c. First Quartile: 0

d. Third Quartile: 4

3. Create a box plot showing the number of times Kirsten reached first base.



4. Find the interquartile range of the data. Are there any outliers?

IQR = 4 - 0 = 4

Outliers =

0 - (1.5)(4) = -6

4 + (1.5)(4) = 10

NONE!

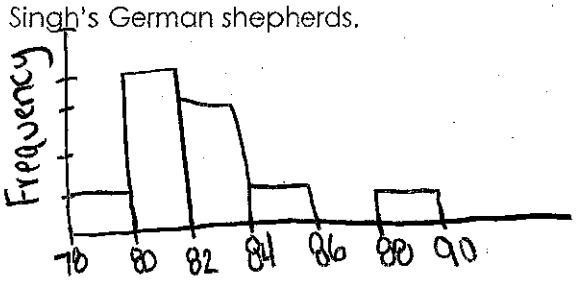
5. Kirsten wants to analyze her performance using this data. She wants to understand the range of her data and the frequency of different results. Which graph, the dot plot or the box plot, will be most useful to Kirsten? Explain.

Dot Plot

Dr. Singh is a veterinarian. He records the weights of each pet. The weights of 10 German shepherds, all 4-year-old males, are in the table below, rounded to the nearest pound. Use this information to solve problems 6-10.

Weight in pounds
~~80~~ 78
~~78~~ 80
~~82~~ 81
~~84~~ 81
~~81~~ 81
89 82
~~83~~ 82
~~81~~ 83
~~81~~ 84
~~82~~ 89

6. Create a histogram showing the weights of Dr. Singh's German shepherds.



7. Find the minimum, maximum, first quartile, and third quartile of the data set.

$Q_2 = 81.5$

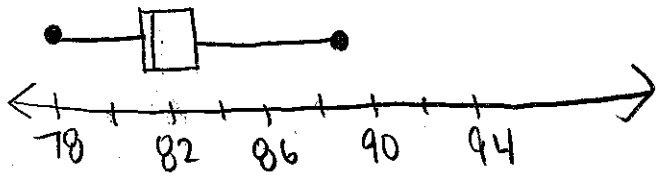
a. Minimum: 78

b. Maximum: 89

c. First Quartile: 81

d. Third Quartile: 83

8. Create a box plot showing the weights of the German shepherds.



9. Find the interquartile range of the data. Are there any outliers?

$IQR = 83 - 81 = 2$

$Q_1 - (1.5)(IQR) = 81 - 3 = 78$

$Q_3 + (1.5)(IQR) = 83 + 3 = 86$

89 is an outlier

10. Dr. Singh wants to analyze the weights of the German shepherds. He wants to understand the center and spread of his data, so that he has a better idea of an expected weight for a 4-year-old male German shepherd. Which graph would be most useful to Dr. Singh? Explain.

Box plot

Name: _____ Date: _____

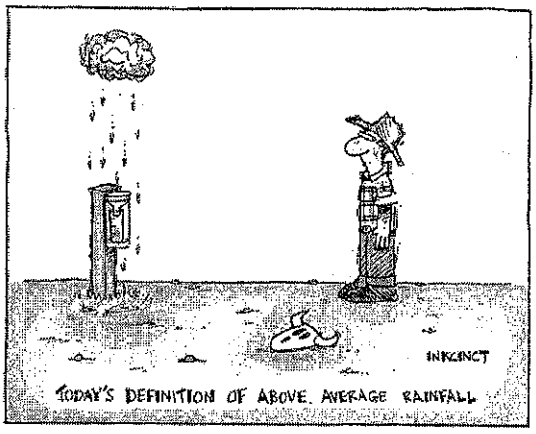
Measures of Central Tendency

1. Some people use "average" interchangeably for both mean and median. Consider this statement:

   "Just think of how stupid the average person is, and then realize half of them are even stupider?" George Carlin

   What type of "average" is George Carlin referring to, mean or median? Is it possible to have more than half of a population above this kind of average?

   Median



2. What is the difference between mean and median?

3. Give an example of data when the mean and median might have the same value.

   Give an example when the mean and the median do NOT have the same value.

4. Can the following statement be true? Why or why not?
   "Welcome to Lake Wobegon, where all the women are strong, all the men are good-looking, and all the children are above average." Garrison Keillor

5. Is it possible to have more than half of data values above (or below) the mean?

6. Find the mean, median, and mode for this set of data.

   ~~X~~ 11, 16, ~~XXX~~ 15, ~~X, X, X~~

   ~~4 5 6 7 7 8 11 11 15 16~~

   Mean = 9
   Median = 7.5
   Mode = 11 & 7

7. Kara had 85, 83, 92, 88, and 69 on her first five math tests. She knows that she needs an average of 85 to get a B. What score must she get on her last test to get a B?

$$\frac{85+83+92+88+69+n}{6}=85$$

$$\frac{417+n}{6}=85$$

$$417+n=510$$

$$n=93$$

Adapted from: Mathematics Vision Project

Measures of Spread – Range, IQR, and Mean Absolute Deviation

Man Absolute Deviation of a numerical data set is the average positive deviations of the data from the mean.

$$\textit{Mean Absolute Deviation} = \frac{\left|x_1 - \bar{x}\right| + \left|x_2 - \bar{x}\right| + \ldots + \left|x_n - \bar{x}\right|}{n}$$

A measure of distribution is a measure of how spread out data is, or how the data is distributed from its smallest values to its largest values. Suppose, for instance, that Joe has test scores of 60, 68, 69, 78, 90, 95, and 100. Sammy scores 78, 78, 79, 79, 82, 82, and 82.

8. Calculate Joe's mean test score. Then calculate Sam's mean test score. What do you notice about Joe's scores compared to Sammy's?

Measuring the mean will not tell you much about the characteristics of the test takers. A measure of distribution, or spread, will help you see that Sam consistently scores near 80, while Joe's scores are spread out, or distributed, over a much larger range.

9. To examine the distribution of test scores, find the mean absolute deviation. Follow the steps below to find the mean absolute deviation of Sam's test scores (Joe's example is given).

Steps	Joe	Sam
a) Calculate the mean, symbolically, $\bar{x}$, of the data.	Mean: $\bar{x} = \frac{60+68+69+78+90+95+100}{7}$, $\bar{x} = 80$	$\bar{X} = 80$
b) Find the deviation, or distance from the mean, for each piece of data.	Deviation: $60-80=-20$ $68-80=-12$ $69-80=-11$ $78-80=-2$ $90-80=10$ $95-80=15$ $100-80=20$	$78-80=-2$ $78-80=-2$ $79-80=-1$ $79-80=-1$ $82-80=2$ $82-80=2$ $82-80=2$
c) Find the absolute value of the mean deviations.	20 12 11 2 10 15 20	2 2 2 2 1 2 1
d) Find the average of the positive deviations found in part c.	MAD: $\frac{20+12+11+2+10+15+20}{7} = 12.86$	$\frac{12}{7} = 1.71$

10. Why do you think the MAD of Joe's test scores is higher than the MAD of Sammy's test scores?

Adapted from: Mathematics Vision Project

Name: _____ Date: _____

Central Tendency and Spread Homework

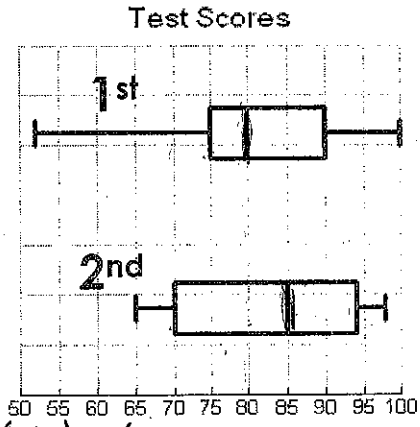
1. The table shows the scores from the top 10 players of our Homecoming basketball game.

Which player scored more than the upper quartile of the data?

A. Matt
B. Michael
C. Jim
D. Bobby

Player	Points	Player	Points
Michael	~~12~~ 3	Dave	~~9~~ 12
Brendan	~~6~~ 5	Heath	~~15~~ 14
Andrew	~~21~~ 6	Jack	~~3~~ 15
Jim	~~14~~ 9	Bobby	~~10~~ 16
Andre	~~5~~ 10	Matt	~~18~~ 21

For #2-3, use the graph to the right.



2. Fill in the blanks:
   - The median for 1st period is 80
   - The median for 2nd period is 85
   - The lowest score for 1st period is 52
   - The lower quartile for 2nd period is 70
   - The spread of the middle 50% for 2nd period is 70-94

2 ANSWERS!

3. Which statement below is NOT true?
   A. 2nd period had the highest score on the test X
   B. The median for 2nd period is 5 less than the median for ~~3rd~~ 1st X
   C. The LQ for 2nd period is 5 less than LQ for ~~3rd~~ 1st period
   D. The UQ for 2nd period is 94

Sample A: 2, 4, 4, 4, 8, 8, 10, 12, 12, 14 mean = 7.8

Sample B: 0, 1, 4, 7, 9, 9, 10, 12, 12, 15 mean = 7.9

4. Which statement accurately compares the two samples?
   A. The mean for Sample A is 1 greater than the mean of Sample B.
   B. The mean for Sample B is 1 greater than the mean of Sample A.
   C. The mean for Sample A is 0.1 greater than the mean of Sample B.
   D. The mean for Sample B is 0.1 greater than the mean of Sample A.

5. Your scores on the first 4 tests in Algebra were 85, 80, 90, and 93. What do you need to make on the 5th test to have a 90 average in the class?

$$\frac{85+80+90+93+n}{5} = 90 \qquad 348+n = 450$$

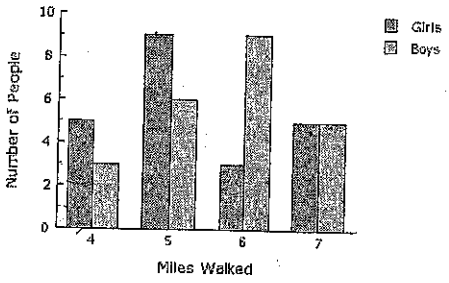
n = 102%

Good luck w/ that ☺

9

6. Which measure of central tendency is MOST EASILY affected by outliers? Mean

7. Forty-five people were asked about how many miles they walked in one week. The results are shown in the graph. How does the median number of miles walked for boys compare with the median number of miles walked for girls?



Girls → 4 4 4 4 5 5 5 5 5 5 5 5 6 6 6 7 7 7 7 7

Boys → 4 4 4 5 5 5 5 5 5 6 6 6 6 6 6 6 7 7 7 7

Girls → 5 Boys → 6

8. The table below shows the running times for science-fiction movies. Find the Mean Absolute Deviation of the data.

Running Times for Movies (min)					
98	87	93	88	126	108

$\bar{X}$ (mean) = 100

98 - 100
87 - 100
93 - 100
88 - 100
126 - 100
108 - 100

$2 + 13 + 7 + 12 + 26 + 8 = \frac{68}{6} = 11.33$

9. The summary statistics for all of the workers at a steel factory are shown. Three sample groups were taken from each of the three shifts. For which sample group is the mean deviation greater than that of the population? SKIP

Steel Factory Workers Ages

Mean Deviation: 11.23

Shift 1		Shift 2	Shift 3
23	17.4	19	21
19	21.4	22	23
50	9.6	24	25
49	8.6	40	40
67	26.6	45	35
34	6.4	29	19
30	10.4	33	70
59	18.6	29	40
40	0.4	39	22
33	7.4	59	23

$\bar{X} = 40.4$

MAD = 12.68

Adapted from: Mathematics Vision Project

50 | Mean Absolute Deviation

The mean absolute deviation or M.A.D. measures the spread of a set of data just like the interquartile range (IQR). Unlike IQR, however, M.A.D. uses every data point. Because they both use every data point, the mean (center) and M.A.D. (spread) tend to be used together to describe a set of data. These two measures of center and spread are appropriate for symmetric distributions.

Finding the mean absolute deviation

1. Find the mean of the data
2. Find the difference of every data point from the mean (called deviation)
3. Make every difference positive (absolute value)
4. Find the mean of the absolute differences

Find the M.A.D. for the following data sets. Round the answers to the nearest tenth.

1. 68, 70, 72, 73, 74, 75 Mean = 72 MAD = 2
2. 72, 75, 73, 99, 68, 79, 48, 60, 52, 59 Mean = 68.5 MAD = 11.1
3. 250, 300, 200, 400, 650, 225, 760, 1215 Mean = 500 MAD = 281.25
4. 22, 31, 57, 29, 62, 24 Mean = 37.5 MAD = 14.6
5. 1, 1, 3, 3, 6, 6, 5, 5, 10, 12

Mean = 5.2 MAD = 2.64

Name: _____ Date: _____

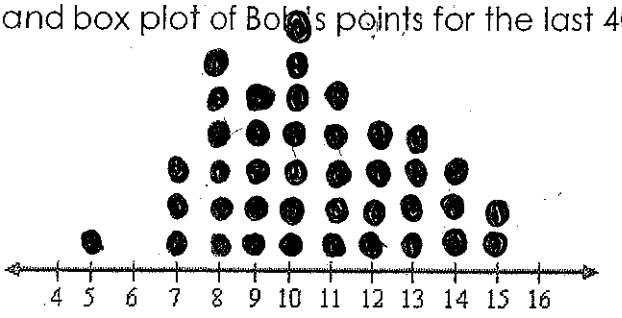
Performance Task: The Basketball Star
Is Bob or Alan a Basketball Star?

MCC9-12.S.ID.1 Represent data with plots on the real number line (dot plots, histograms, and box plots). Choose appropriate graphs to be consistent with numerical data: dot plots, histograms, and box plots.
MCC9-12.S.ID.2 Use statistics appropriate to the shape of the data distribution to compare center (median, mean) and spread (interquartile range, standard deviation- Advanced Algebra) of two or more different data sets. Include review of Mean Absolute Deviation as a measure of variation.
MCC9-12.S.ID.3 Interpret differences in shape, center, and spread in the context of the data sets, accounting for possible effects of extreme data points (outliers). Students will examine graphical representations to determine if data are symmetric, skewed left, or skewed right and how the shape of the data affects descriptive statistics.

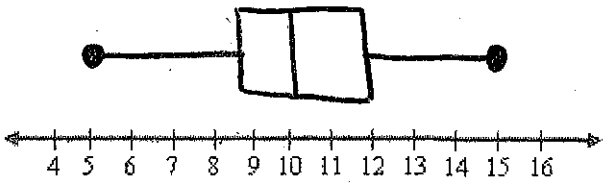
Bob's Points per Game

5 7 7 7 8 8 8 8 8 9 9 9 9 9 10 10 10 10 10 10 10 11 11 11 11 11 12 12 12 12 13 13 13 13 14 14 14 15 15 15

1. Bob believes he is a basketball star and so does his friend Alan. Create a dot plot and box plot of Bob's points for the last 40 games.



Bob's Points



Bob's Points

2. Describe Bob's data in terms of center, spread, and shape.

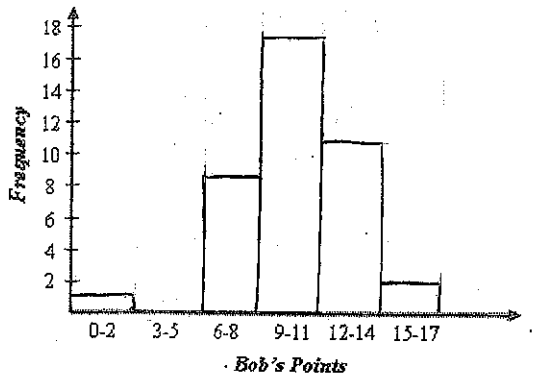
≈ 10 (center) · less spread · symmetric, unimodel, bell shape

Bob's friend Alan has the following points:

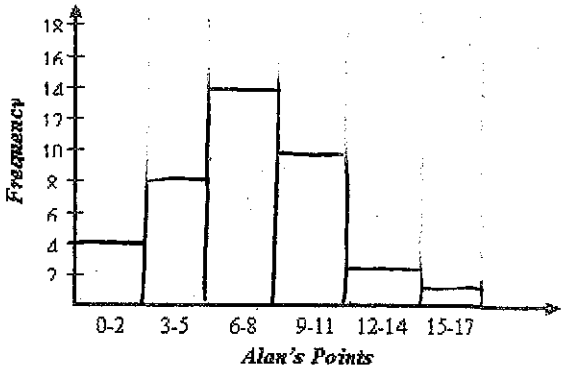
Alan's Points per Game

1, 2, 2, 2, 4, 5, 7, 7, 8, 10, 4, 4, 8, 9, 5, 6, 6, 6, 8, 10, 11, 11, 10, 12, 12, 6, 6, 9, 10, 10, 12, 11, 11, 7, 8, 9

3. Create a histogram of both Bob's and Alan's data.



Bob's Points



Alan's Points

4. Describe the shape of the two histograms from problem #3.

5. Use summary statistics to compare Bob and Alan's points per game.

	Min	Quartile 1 (Q1)	Median (Q2)	Quartile 3 (Q3)	Max	Mean	Range	IQR	MAD
Bob	5	8.5	10	12	15	10.4	10	3.5	
Alan	0	5	7	10	15	7.25	15	5	

Mean for Bob = 416 ÷ 40 = 10.4

Mean for Alan = 290 ÷ 40 = 7.25

6. Which graphical representation best displayed Bob's and Alan's data?

7. Based on the summary statistics is either friend a basketball star? Justify your answer.

Bob

1.	5
2.	7
3.	7
4.	7
5.	8
6.	8
7.	8
8.	8
9.	8
10.	8
11.	9
12.	9
13.	9
14.	9
15.	9
16.	10
17.	10
18.	10
19.	10
20.	10
21.	10
22.	10
23.	11
24.	11
25.	11
26.	11
27.	11
28.	12
29.	12
30.	12
31.	12
32.	13
33.	13
34.	13
35.	13
36.	14
37.	14
38.	14
39.	15
40.	15

$Q_1 = 8.5$

$Q_2 = 10$

$Q_3 = 12$

$Q_4 = 15$

Alan

1.	0
2.	1
3.	2
4.	2
5.	3
6.	3
7.	4
8.	4
9.	4
10.	5
11.	5
12.	5
13.	6
14.	6
15.	6
16.	6
17.	6
18.	7
19.	7
20.	7
21.	7
22.	8
23.	8
24.	8
25.	8
26.	8
27.	9
28.	10
29.	10
30.	10
31.	10
32.	10
33.	11
34.	11
35.	11
36.	11
37.	12
38.	12
39.	12
40.	15

$Q_1 = 5$

$Q_2 = 7$

$Q_3 = 10$

$Q_4 = 15$