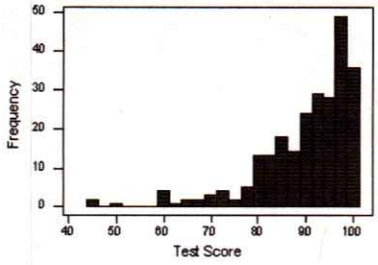
Statistical Reasoning
Chapter 2 Review: Describing Distributions of Data

Name answer key
Date _____ Day _____

Types of Displays

Bar Graph, Pie Graph, Dotplot, Stemplot, Histogram, Boxplot



1. How would you describe the overall shape of this distribution?

   skewed left

2. Where will the mean fall in respect to the median? (don't try to calculate it)

   The mean is less than the median (<)

3. What numerical measures would best describe it?
   (mean/standard deviation OR median/IQR)

   skewed → five # summary → median/IQR

4. Which display best represents categorical data?

   pie chart or bar graph

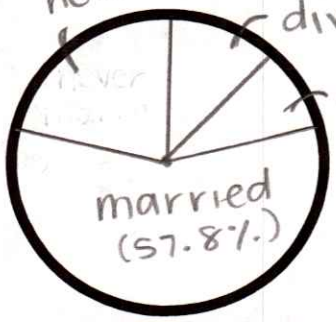
5. What's the main difference between the data bar graphs and pie chart can display?

   pie charts can't display data that totals over 100%.

6. Use the following data to create a pie chart:

Marital Status	Count (thousands)
Never Married	23,260
Married	62,250
Widowed	10,050
Divorced	12,100
Total	**107,660**

%
21.6% → 78°
57.8% → 208°
9.3% → 33°
11.2% → 40°



skewed right (tail is on the right)

120 is an outlier

7. Create a boxplot for the following data. Describe the shape. Describe any outliers.

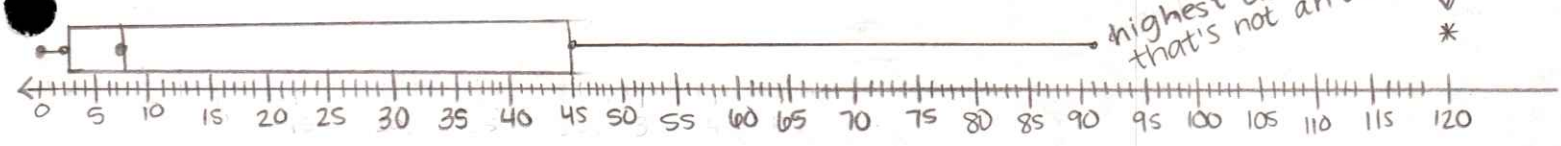
Number of text messages sent by teenagers in the last 24 hours:

0	8	2	28	21	7	4	1	23	88	7	1	22
8	120	82	0	91	42	17	2	3	48	7	52	

min = 0 Q1 = 2.5 med = 8 Q3 = 45 max = 120

LB = -61.25

UB = 108.75 → 120 is an outlier



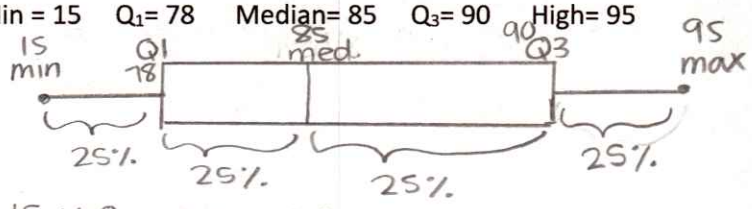
Statistical Reasoning
Chapter 2 Review: Describing Distributions of Data

Name _____
Date _____ Day _____

Vocabulary

Median, IQR, 5 numbers summary, Mean, Standard deviation, Outliers, Variability

8. A set of data has the following 5-number summary: Min = 15 Q_1= 78 Median= 85 Q_3= 90 High= 95



   a. What's the interquartile range?

   IQR = Q3 - Q1 = 12

   b. Is there an outlier? How do you know?

   LB = 78 - 1.5(12) = 60 → since 15 < 60, we know there is at least 1 outlier on the low end

   UB = 90 + 1.5(12) = 108 → since 95 < 108, we know there are no outliers on the high end

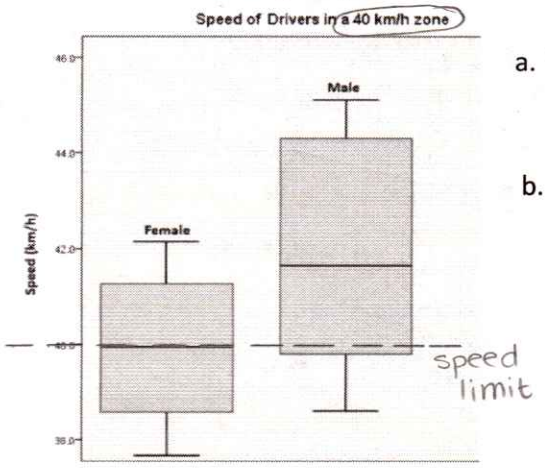
   c. What percentage of the data lies above 90?

   25%

   d. What percentage of the data lies below 85?

   50%

9. Use the following boxplot



   a. Which one shows more variability?

   male

   b. Roughly what percentage of males are going exactly the speed limit or speeding?

   40 km/hr or faster

   approximately 75%.

10. At a local university, the mean of the total number of classes taken by students earning their Bachelor's Degree is 35. What does this mean in this situation? The average # of classes is 35

   a. If all students graduating with a Bachelor's degree took an extra 3 classes, what would happen to the mean? What would happen to the standard deviation?

   mean: would go up by 3
   s.d.: no change

   b. Would anything happen to the median? What about the IQR?

   median: would go up by 3
   IQR: no change

Statistical Reasoning
Chapter 2 Review: Describing Distributions of Data

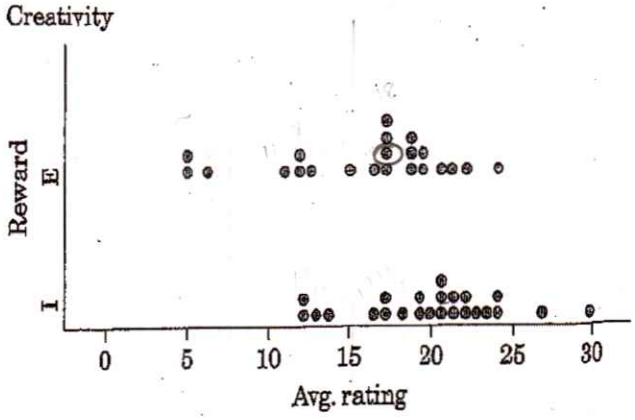
Name _____
Date _____ Day _____

Measures of center and spread

Median/IQR, Mean/Standard Deviation

11. Psychologists designed a study involving 47 experienced creative writers who were college students. Students were divided into two groups using a chance process (like drawings name from a hat). One group were given external reasons for writing (like public recognition/making money/pleasing their parents). The other group were given a list of statements about internal reasons (expressing yourself/enjoying working with words). Both groups were then instructed to write a poem about laughter. Each student's poem was rated separately by 12 different poets using a creativity scale.

Here's the data collected during this study:



a) What's the median of the students who were given external reasons? 18

b) What's the median of the students who were given internal reasons? 21

c) Without calculating it, where would you expect the mean to fall in respect to the median for the students given external reasons? mean < median since the data is skewed left

d) Which one has the most variability? How do you know?

the ones given internal reasons → they are more spread out

e) Did either one have one or more outliers? How do you know?

external reasons → there is a large gap between the minimum values and the other cluster of scores

12. . Mr. Jones and Mrs. Barber give the same history test and the mean result of each class is 85. Mr. Jones claims that even though they got the same average, his class did better overall because his standard deviation was 0 while Mrs. Barber's was 5. What do you think? Did his students do better? Why or why not?

Since Mr. Jones' had a smaller standard deviation, his scores were more consistent. Although Mrs. Barber may have had the students with the highest grades, she also had the ones with the lowest grades (her grades are more spread). I would say Mr. Jones' class did better overall.