Goal 3: Data Analysis & Probability

Variability & Data Sets

OBJECTIVE	CORE LEARNING GOALS
• The student will determine the affect that variability in data has on the mean and median.	3.1.2 The students will use measures of central tendency and variability to solve problems and make informed decisions.
DRILL Calculate the mean and the median for each date set. Discuss why your values make sense for the given data. 1. 3, 4, 50, 60, 70 2. 30, 40, 50, 60, 70	MATERIALS Activities: "Box and Whisker Plots and Measures of Central Tendency" Additional Problems Answer Keys Graphing calculators

CALCULATOR SKILLS

Enter data into a list.
Create a box an whisker plot. and use the trace feature to determine the five-number summary.
Calculate the mean and median.

ACTIVITIES

1. **Drill**. Have the students work in pairs to complete this drill. Students should remember that the mean is the numerical average, and the median is simply the middle number. After the pairs have had a chance to discuss the questions let the students explain to the class why their answers make sense. Answers will vary, but you should expect students to notice that the second data set is symmetrical or has numbers equal distance from the middle number while the first data set has two values that are way to the left of the middle number.

2 **Exploration -** Have students work with their group or with a partner to complete "Box and whisker Plots and Measures of Central Tendency". Calculators should be available for the students to use. Some students may want to actually plot the boxplots to see these screens on their on calculator. While the students are working circulate to check on their progress, prompt their thinking through leading questions, and give them a chance to ask you questions.

3. **Class Discussion –** After most students have completed this activity have a class discussion. This gives you the opportunity to check to make sure that all groups understand the concept of the lesson. See the answer key for the major points that students should bring up during this discussion.

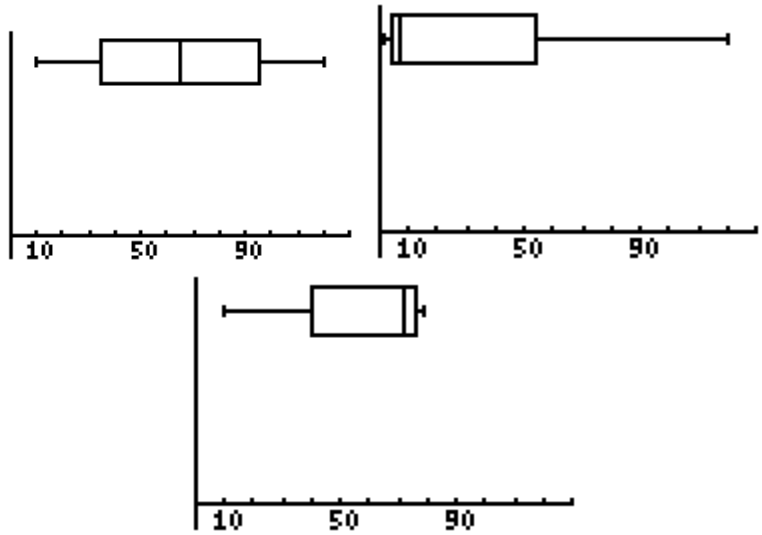
4. **Additional problems –** These problems are provided for you to use in any one of a variety of ways such as homework, an individual check for understanding, review, or quiz.

ASSESSMENT	HOMEWORK
Drill Class discussion Group work	

Box and Whisker Plots and Measures of Central Tendency

The box and whisker plots for three data sets are shown below.

Data Set I **Data Set II** **Data Set III**



1. Compare the shapes of the box and whisker plots.

2. What do these shapes tell you about the variation in the data sets.

3. The data sets for these box and whisker plots are given below.

Data Set I:										
10	20	30	40	50	60	70	80	90	100	110
120										

Data Set II:										
1	2	3	4	5	6	7	8	9	100	110
120										

Data Set III:										
10	10	11	70	71	72	73	74	75	76	78
79										

   a. Calculate the mean of each data set.

   b. Locate the mean of each data set on its box and whisker plot by drawing a vertical line that passes through the mean.

   c. How do the mean and median of each data set compare?

4. What does the shape of a box and whisker plot tell you about the relationship between the mean and the median of its data set?

Box and Whisker Plots and Measures of Central Tendency
Additional Problems

1. The chart below gives average precipitation for Baltimore, Maryland for a number of years starting in 1974.

Year	Precipitation (in inches)	Year	Precipitation (in inches)
1974	41	1985	41
1975	62	1986	38
1976	44	1987	39
1977	40	1988	36
1978	45	1990	43
1979	61	1991	30
1980	32	1992	39
1983	58	1993	45
1984	38	1994	41

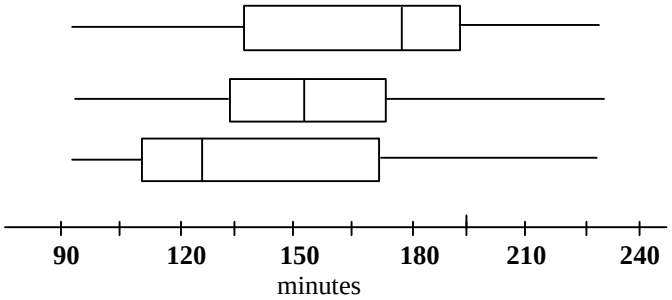
Source: http://www.cdc.noaa.gov/cgibin/USclimate/rankall.calc.pl

- **Create a box and whisker plot using this data.**

- Describe the relationship between the mean and the median based on your box and whisker plot. Use mathematics to justify your answer.

- State what this relationship tells you about the average precipitation for Baltimore. Use mathematics and the original data from the chart to justify your answer.

2. The following parallel box plots show the effective duration of three over-the-counter pain relievers.



Which of the following statements is true?

a. The medians for the pain relievers are different but the means are the same.

b. All of the pain relievers are equally effective because they all have the same range.

c. Data for the middle pain reliever is approximately symmetric.

d. The basic shape of the data is the same for all three pain relievers.

Box and Whisker Plots and Measures of Central Tendency
Answer Key

Notes: When students are making judgements about which measure of central tendency to use, they need to understand how variability in the data affects the values of the measures of central tendency. Generally, if you are looking for the best representative value when there is a skewed set of data, it is better to use the median rather than the mean. This is because the mean can be influenced by outliers and therefore give a false idea of the truth.

1 - 2. Students should mention that Data Set I is symmetrical. They can see that the box and whisker plot is symmetrical whereas the other two plots are not. Data Set II has half the data that is 10 or less and the other half between 10 and 120, while Data Set III has half the data falling between 70 and 80 with the other half falling between 10 and 70. Students should have a solid understanding of median and how it is represented in a box and whisker plot.

3. a – b.

	Mean	**Median**
Data Set I	65	65
Data Set II	31.25	6.5
Data Set III	58.25	72.5

c. If the data are symmetrical, then the mean and the median are equal. However, when the data are skewed to the right as in Data Set II, the mean is greater than the median. When the data are skewed to the left as is the case with Data Set III, the mean is less than the median. In other words, the mean is pulled in the direction of the skewedness or outliers while the median is not effected.

4. If the box and whisker plot is symmetrical, the mean and the median will be the same. If the data are skewed to the right (you can tell this by a long whisker on the right side), then the mean will be greater than the median. If the data are skewed to the left (shown by a long whisker to the left side), then the mean will be less than the median.