Measures of Central Tendency

Purpose:
Participants will determine the mean, median, and mode of a data set.

Overview:
In pairs, participants will input data into the lists of their graphing calculator. They will determine the mean, median, and mode of the data and provide possible interpretations for these measures of central tendency.

TExES Mathematics 4-8 Competencies. The beginning teacher:

IV.012.B Applies concepts of center, spread, shape, and skewness to describe a data distribution.

IV.012.D Demonstrates an understanding of measures of central tendency (e.g., mean, median, mode) and dispersion (e.g., range, interquartile range, variance, standard deviation).

TEKS Mathematics Objectives. The student is expected to:

5.13.B Describe characteristics of data presented in tables and graphs including the shape and spread of the data and the middle number.

6.10.B Use median, mode, and range to describe data.

7.12.A Describe a set of data using mean, median, mode, and range.

7.12.B Choose among mean, median, mode, or range to describe a set of data and justify the choice for a particular situation.

8.12.A Select the appropriate measure of central tendency to describe a set of data for a particular purpose.

Terms.
Data, mean, median, mode

Materials.
For instructor:

- Transparencies
- Overhead graphing calculator and LCD panel

For each participant:

- Data Sheet
- Activity Sheet
- Graphing calculator

Transparencies.

- *Measures of Central Tendency*

Activity Sheet(s).

- *Measures of Central Tendency*

Procedure:

Steps	Questions/Math Notes
1. Have participants enter the data either by hand or by linking two calculators and sending the lists from the calculator that has the data to the other calculator. Ask participants to work in pairs and to check one another's input of data. Ask participants for observations based on the data.	Let L1 = Region Number, L2 = Total Number of Students Enrolled in that Region during 1987-88, and L3 = Total Number of Students Enrolled in that Region during 1997-98. Do not enter the 4^{th} column of data.
2. Once the data for L1, L2, and L3 are entered into all calculators, have participants compute the total student enrollment change from 1987-88 to 1997-98 in L4. Once L4 has been calculated, have participants exchange calculators with their partner and check the L4 column for correctness. If L4 is correct, then most likely all the data is correct.	Calculating the total student enrollment change from 1987-88 to 1997-98 in L4 can be done by placing the cursor on L4 and typing L3 – L2. Then press ENTER.
3. Ask participants to predict the mean of L2. Ask participants to predict the mean of L3.	*Why is making this prediction difficult?* (The data are not organized in ascending or descending order, and the values are spread apart so it is difficult to get a sense of an average.) *What could the mean represent in this situation?* (The mean could represent the average number of students that would reside in one region if the state's student enrollment was evenly distributed among the regions.) *Do you expect the mean of L3 to be larger or smaller than the mean of L2? Why?* (We expect the mean of L3 to be larger than the mean of L2 because all numbers in L3 are larger than their corresponding partners in L2.)
4. Ask participants to predict the median of L2.	*Why is making this prediction difficult?* (The data are not organized in ascending or descending order so it is difficult to get a sense of the middle score.) *What could the median represent in this situation?* (The median could represent the number of students who are enrolled in the "middle"

Ask participants to predict the median of L3.	region.) *Do you expect the median of L3 to be larger or smaller than the median of L2? Why?* (We expect the median of L3 to be larger than the median of L2 because all numbers in L3 are larger than their corresponding partners in L2.)
5. Ask participants to predict the mode of L2. Ask participants to predict the mode of L3.	*Is predicting the mode difficult? Why or Why not?* (There is no mode for L2; all data entries are unique.) *Is predicting the mode difficult? Why or Why not?* (There is no mode for L3; all data entries are unique.)
6. Ask participants to calculate the mean of L2 and L3. Notes on how to get the calculator to compute the mean are given below.	*How will you calculate the mean of L2? L3?* (The mean is the average. The data for L2 must be summed together and that sum divided by 20. The same approach is used for L3.)
7. Ask participants to calculate the median of L2 and L3. Notes on how to get the calculator to compute the median are given below. Ask participants to determine the mode of the L2 and L3.	*How will you calculate the median of L2? L3?* (The data must be in ascending or descending order, and the middle data entry is the median.) *Why is the median so much smaller than the mean of the two data sets?* (The median is the middle data point. Since much of the data is less than 200,000 total students, the median is closer to 100,000. The large values at the end of the data set skew (or stretch) the mean to the right; thus, giving the mean a much larger value than the median.) *What is the mode of L2? L3?* (Neither L2 or L3 have a mode. Each data value is listed only once.)
8. Ask participants to calculate the 1st and 3rd quartiles of L2 and L3. Notes on how to get the calculator to compute the quartiles are given below.	*How will you calculate the 1st quartile of L2 or L3?* (Calculate the 25th percentile of the data set. The 25th percentile of the data set would be the mean of the 5th and 6th data entries or the mean of the first half of data values.) *How will you calculate the 3rd quartile of L2 or L3?* (Calculate the 75th percentile of the data set.

	The 75th percentile of the data set would be the mean of the 15th and 16th data entries or the mean of the second half of data values.) *What can you tell about the data by looking at the quartiles of L2 and L3?* (From the original data, we can see that every region grew in size over the ten-year span of time. The quartile values lead us to believe, however, that the data over time has become more dispersed. The Q1s of L2 and L3 are approximately 3,000 students apart. The Q3s of L2 and L3 are approximately 65,000 students apart. This would lead us to believe that the larger regions (at the end of the data set when it is listed in ascending order) are growing at a faster rate than the smaller regions.

Calculator Instructions.

Entering Data in Lists. We begin by entering the data into the calculator. On a TI-83 calculator, press STAT, 1: Edit, and enter data.
L1 = Texas Education Region Numbers
L2 = Total Students Enrolled in Region 1987-88
L3 = Total Students Enrolled in Region 1997-98
To check L2 and L3, highlight L4 and enter 2nd STAT, L3 – L2, then press enter. Check the data with the numbers listed in the handout.

Transferring Data from One Calculator to Another. Use the TI calculator cable to link two calculators together. For the calculator which contains the data, press 2nd LINK, SEND, 4: List, SELECT, and then highlight L1, L2, and L3.
For the calculator which is going to receive the data, press 2nd LINK, RECEIVE, ENTER.
Once the receiving calculator is ready to receive the data, press TRANSMIT on the sending calculator. If errors appear, check the cable for proper connection.

Computing the Mean. To compute the mean of L2, we sum the numbers in L2 and divide the sum by the number of regions (20). To do this, we press 2nd QUIT to get to the home screen. Then we press 2nd LIST, MATH, 5: sum(L2), ENTER. We get the sum (total number of students enrolled in Texas schools during the 1987-88 school year) to be 3,224,916. To determine the mean, we divide the sum by the number of data items (which in this case is 20). On average, there were 161,245.8 students for each region in 1987-88.

Another way to compute the mean is to press 2nd LIST, MATH, 3: mean(L2), ENTER.

Computing the Median. To compute the median of L2, we need the data in ascending or descending order. In order to preserve L2, we can copy the data set to L5 and then convert it to ascending order. To copy the data of L2 to L5, highlight L5, then press 2nd, 2 for L2, now press ENTER. The list for L2 will appear in L5. Copy L3 to L6 using the same procedure. To convert the lists to ascending order, press STAT, 2: SortA(L5), ENTER. (SortA sorts the data into ascending order; SortD sorts the data into descending order.) The middle data point is halfway between the 10th and 11th data values or 104,697 students.

Computing the Mode. There is no mode for L2 or L3.

Computing Quartiles. The 25th percentile of a data set is called the first quartile and the 75th percentile is the third quartile. The first quartile value of L5 is an average of the 5th and 6th data entries when the data are listed in ascending order. Hence, the first quartile value is the mean of 56,229 (Region III) and 77,765 (Region XVI) or 66,997 total students. The third quartile value of L5 is the average of the 15th and 16th data entries when the data are listed in ascending order. Hence, the third quartile value is the mean of 180,493 (Region XIII) and 222,668 (Region I) or 201,580.5 total students.

The first quartile value of L6 is an average of the 5th and 6th data entries when the data are listed in ascending order. Hence, the first quartile value is the mean of 57,730 (Region III) and 80,711 (Region XVI) or 69,220.5 total students. The third quartile value of L6 is an average of the 15th and 16th data entries when the data are listed in ascending order. Hence, the third quartile value is the mean of 247,989 (Region XIII) and 284,614 (Region I) or 266,301.5 total students.

L2 (1987-1988 data)	L3 (1997-1998 data)
Q1 = 66,997 students	Q1 = 69,220.5 students
Q3 = 201,580.5 students	Q3 = 266,301.5 students

Of course, 0.5 of a student isn't a reasonable answer.

We can also determine the first and third quartile values using the calculator. Press STAT, CALC, 1: 1-Var Stats L3, ENTER. This command provides the mean value, the sum of the x values, the sum of the x^2 values, the sample standard deviation, the population standard deviation, the value of n, minimum value of x, Q1, Median, Q3, and the maximum value of x.

Solutions to Activity Sheet:

1.

L2 (1987-1988 data)	L3 (1997-1998 data)
mean of L2 = 161,245.8 students	mean of L3 = 194,575.65 students
median of L2 = 104,697 students	median of L3 = 120,286 students
no mode for L2	no mode for L3

2.

L2 (1987-1988 data)	L3 (1997-1998 data)
Q1 = 66,997 students	Q1 = 69,220.5 students
Q3 = 201,580.5 students	Q3 = 266,301.5 students

Of course, 0.5 of a student isn't a reasonable answer.

3. The mean of L3 = 194,575.65 students could be interpreted in the following way:
   If the student population were distributed evenly throughout the state of Texas in 1997-98, then each region in Texas would have had approximately 194, 576 students.
   The two outliers (Regions IV and X) cause the mean to be enlarged and, hence, not the best representation of the overall data.

   Which region most closely represented the state's mean population in 1997-98?

   Was your Regional Service Center servicing more or less than the state's average student population in 1997-98? What % above or below the state average was enrolled in your region?

4. The median of L3 = 120,286 students could be interpreted in the following way:
   The middle enrollment in Texas' 20 regions in 1997-98 was 120,286 students.
   The median is a more appropriate interpretation of the data than the mean.

   Which regions most closely represent the median of the 1997-98 data?

Why is the mean of L3 so much larger than the median?

5. From the original data, we can see that every region grew in size over the ten-year span of time. The quartile values lead us to believe, however, that the data over time has become more dispersed. The Q1s of L2 and L3 are only approximately 3,000 students apart. The Q3s of L2 and L3 are approximately 65,000 students apart. This would lead us to believe that the larger regions (at the end of the data set when it is listed in ascending order) are growing at a faster rate than the smaller regions.