Stem-and-Leaf Plots

Purpose:
Participants will create a back-to-back stem-and-leaf plot using two data sets.

Overview:
In order to generate a stem-and-leaf plot, pairs of participants will decide what place value to use for rounding and how many places to truncate the given data. They will also determine which numbers will be used for the stem. Then, they will create a back-to-back stem-and-leaf plot using the rounded data sets. Next, participants will compare the distributions of the two data sets. Finally, they will determine the mean, median, and 1st and 3rd quartile values for the two rounded data sets and compare these values to the actual ones that were generated in previous activities. Participants will discuss the value and uses of stem-and-leaf plots.

TExES Mathematics 4-8 Competencies. The beginning teacher:

IV.012.A Organizes and displays data in a variety of formats (e.g., tables, frequency distributions, stem-and-leaf plots, box-and-whisker plots, histograms, pie charts).

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).

IV.012.E Analyzes connections among concepts of center and spread, data clusters and gaps, data outliers, and measures of central tendency and dispersion.

IV.012.F Calculates and interprets percentiles and quartiles.

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.

5.13.C Graph a given set of data using an appropriate graphical representation such as a picture or line.

6.10.A Draw and compare different graphical representations of the same data.

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

7.11.B Make inferences and convincing arguments based on an analysis of given or collected 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.

8.13.B Recognize misuses of graphical or numerical information and evaluate predictions and conclusions based on data analysis.

Terms.
Data, mean, median, mode, quartiles, stem-and-leaf plot

Materials.

For instructor:

- Transparencies
- Overhead graphing calculator and LCD panel

For each participant:

- Data Sheet
- Activity Sheet
- Graphing calculator

Transparencies.

- *Stem-and-Leaf Plots*

Activity Sheet(s).

- *Stem-and-Leaf Plots*

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 which has the data to the other calculator. Ask participants to work in pairs and to check one another's input of 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 4th 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 decide what values are reasonable to use for the stem of a back-to-back stem-and-leaf plot for the data sets listed in L2 and L3. Truncating Decimals on the Calculator: When rounding to the hundreds place and truncating the ones and tens places of the original list, change the decimal mode from floating to fixed and choose 0 digits to be displayed to the right of the decimal. Then divide the numbers in the list by 100. The new list will be rounded to the nearest hundreds place and will have the ones and tens places of the original list eliminated. All of the above instructions can be accomplished by following these keystrokes: STAT, 1: Edit, highlight L4, MATH, NUM, 2: round(L3 ÷ 100, 0), ENTER. So, the entry line should read: L4 = round (L3 ÷ 100, 0). These keystrokes tell the calculator to divide L3 by 100, round to the nearest resulting whole number, and place the outcomes in L4.	*Would it be most useful to round the data to the hundreds, thousands, ten-thousands, or hundred-thousands place?* *If the data in L2 and L3 were rounded to the hundreds place and the ones and tens places were truncated, what numbers would serve as the stem?* (The stem numbers would range from 40 to 828. This is not practical.) *If the data in L2 and L3 were rounded to the thousands place and the ones, tens, and hundreds places were truncated, what numbers would serve as the stem?* (The stem numbers would range from 4 to 82. This is not practical.) *If the data were rounded to the ten-thousands place and the ones, tens, hundreds, and thousands places were truncated, what numbers would serve as the stem?* (The stem numbers would range from 0 to 8. This is a practical range.) *If the data were rounded to the hundred-thousands place and the ones, tens, hundreds, thousands, and ten-thousands places were truncated, what numbers would serve as the stem?*

	(The stem number would be 0. This is not reasonable.)
4. Ask participants to generate the stem-and-leaf plot for L2. List the leaves for L2 on the left side of the stem. Remind participants to sort the data in ascending order via the calculator to assist with creating the stem-and-leaf plot. Have participants make a histogram of the data that uses similar intervals to the stem-and-leaf plot.	*What window would you use for the histogram?* (The window should be: Xmin = 0 Xmax = 700000 Xscl = 100000 Ymin = -3 Ymax = 10 Yscl = 0) *Are there similarities between the stem-and-leaf plot and the histogram?* (When using the Trace function on the graphing calculator, the frequencies are the same for the histogram and the stem-and-leaf plot. Also, if the stem-and-leaf plot were rotated 90° clockwise and reflected vertically, then the two graphs would have the same shape.)
5. Ask participants to generate the stem-and-leaf plot for L3. List the leaves for L3 on the right side of the stem. Have participants make a histogram of the L3 data that uses similar intervals to the stem-and-leaf plot. Ask participants to rotate the stem-and-leaf plot 90° counter-clockwise.	*What similarities are there between the L2 and L3 portions of the stem-and-leaf plot?* (Both portions of the graph have most of their data stacked along the 0 and 1 part of the stem.) *What differences are there?* (The L3 data is more dispersed than the L2 data.) *What window would you use for the histogram?* (The window should be: Xmin = 0 Xmax = 900000 Xscl = 100000 Ymin = -3 Ymax = 10 Yscl = 0) *How does the stem-and-leaf plot compare to the histogram that was created in the Measures of Central Tendency Activity?* (The shapes of the data are the same. The frequencies are the same.)

Solutions to Activity Sheet:

1. Rounding to the nearest ten-thousands place makes the most sense. The stem values will range from 0 to 6 (the numbers in the hundred-thousands place).

2. The leaves will be the numbers in the thousands place; they will range from 0 to 9.

   Stem-and-Leaf Plot of L2

Frequency	Leaves | Stem
9	988865554 | 0
6	853110 | 1
3	982 | 2
0	| 3
1	3 | 4
0	| 5
1	6 | 6

3. Stem-and-Leaf Plot of L2 and L3

Frequency (L2)	Leaves | Stem | Leaves	Frequency (L3)
9	988865554 | 0 | 455668889	9
6	853110 | 1 | 13356	5
3	982 | 2 | 58	2
0	| 3 | 28	2
1	3 | 4 |	0
0	| 5 | 5	1
1	6 | 6 |	0
0	| 7 |	0
0	| 8 | 3	1

4. The stem-and-leaf plot tells us that almost one-half of the regions have a student population that is less than 100,000 and approximately 75% of the regions have a student population that is between 40,000 and 180,000. Also, the populations of the regions have become more dispersed over the ten-year span. The two largest regions (Regions X and IV) have increased more in size than the other districts.

5.

L2		L3	
mean	= 16.2 students	mean	= 19.4 students
Q1	= 7 students	Q1	= 7 students
median	= 10.5 students	median	= 12 students
Q3	= 20 students	Q3	= 26.5 students

   The numbers above (if multiplied by 100,000) are similar to the ones found in the Measures of Central Tendency Activity. Listed below are the results from the previous activity.

L2		L3	
mean	= 161,245.8 students	mean	= 194,575.65 students
Q1	= 66,997 students	Q1	= 69,220.5 students
median	= 104,697 students	median	= 120,386 students
Q3	= 201,580.5 students	Q3	= 266,301.5 students