Statistics with State Names

Overview:

The students will determine the number of times that each letter of the alphabet is used when writing the names of all 50 states and understand how various representations, including stem-and-leaf plots, box-and-whisker plots, and histograms, can be used to organize the data.

3-4th Grades

Data Analysis & Probability

A. Gather and organize data from surveys and classroom experiments including data collected from a period of time.

C. Construct charts, tables, and graphs to represent data.

E. Describe data using mode, median, and range.

5-7th Grades

Data Analysis & Probability

A. Read, create and use line graphs, histograms, circle graphs, box-and-whisker plots, stem-and-leaf plots and other representations when appropriate.

E. Collect, organize, display and interpret data for a specific purpose or need.

F. Determine and use the range, mean, median, and mode to analyze and compare data, and explain what each indicates about the data.

8-10th Grades

Data Analysis & Probability

A. Create, interpret and use graphical displays and statistical measures to describe data; e.g., box-and-whiskers plots, histograms, scatterplots, measures of center and variability.

C. Compare the characteristics of the mean, median and mode for a given set of data, and explain which measure of center best represents the data.

D. Find, use and interpret measures of center and spread, such as mean and quartiles, and use these measures to compare and draw conclusions about sets of data.

Materials

Students will need worksheets.

Procedure:

1) We are going to write the names of the 50 states.
2) Ask questions to start students thinking about the data to be generated:
- Which letter will you use most? Which letter will you use least?
- Will you use every letter of the alphabet?
- Which state name has the most letters?
- Can you answer these questions just by looking at a map? (Need to establish a reason for organizing data in some format)
3) Use the States Names activity sheet to identify the frequency of each letter.
4) Talk about the need for a systematic process for organizing the data.
5) Have students find the range of the data.
6) Create a stem-and-leaf plot of the data.
7) Which of the questions from 2 could we now answer easily using the stem-and-leaf plot?

8) Create a box-and-whiskers plot of the data. Help students find the five numbers need for this plot. Talk about the median of a set of data and how to find it.
9) Have students find the mean and mode of the data.
10) Have students compare the different graphs and how the data is represented in each.

Extensions:

- Have students create a histogram of the data using an appropriate class size.
- Have students look at other state-related data such as population or the length of names of the state capitals.
- Have students do a similar project with the names of the students in the class.

Assessment:

- Have students write a paragraph about the process they used to organize the data and why they need to have a systematic way to organize rather than trying to use the raw data.
- Have students write a paragraph comparing the way the data is represented in the different graphs and which graph makes it easiest to answer the questions. Why don't any of the graphs help answer the question about the state with the most letters.

STATE NAMES

The names of all 50 states are listed below. Record the number of times that each letter of the alphabet is used.

Alabama Alaska Arizona Arkansas California Colorado Connecticut Delaware Florida Georgia Hawaii Idaho Illinois Indiana Iowa Kansas Kentucky Louisiana Maine Maryland Massachusetts Michigan Minnesota Mississippi Missouri Montana Nebraska Nevada New Hampshire New Jersey New Mexico New York North Carolina North Dakota Ohio Oklahoma Oregon Pennsylvania Rhode Island South Carolina South Dakota Tennessee Texas Utah Vermont Virginia Washington West Virginia Wisconsin Wyoming

LETTER	FREQUENCY	LETTER	FREQUENCY
A		**N**	
B		**O**	
C		**P**	
D		**Q**	
E		**R**	
F		**S**	
G		**T**	
H		**U**	
I		**V**	
J		**W**	
K		**X**	
L		**Y**	
M		**Z**	

1. Which letter is used most often?

2. Are there any letters that are not used at all?

3. What state requires the most letters to spell?

State Name Statistics

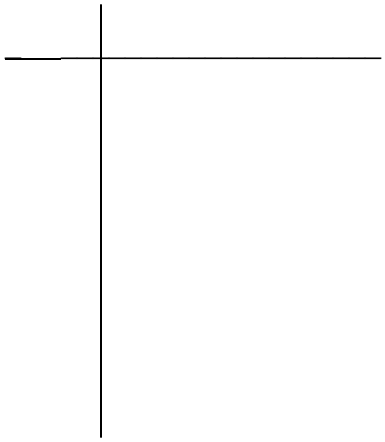
List the numbers of times the letters are used from highest to least from your tally chart.

What is the range (largest – smallest numbers) in the set? _____

Find the median (middle of the data set) of the set. _____

Find the mode (most frequently occurring number) for the data set. _____

Create a stem-and-leaf plot of the numbers:



Don’t forget your legend to tell readers how to read the stem-and-leaf plot

Look at the data listed from largest number to smallest number of times a letter occurs.

Find the median and draw a line to show where it is located.

Look at the top half of the data – the data that is larger than the median. Find the median of this set of data. This number is called the upper quartile. Draw a line in your data list to show where the upper quartile is located.

Look at the bottom half of the data – the data that is smaller than the median. Find the median of this set of data. This number is called the lower quartile. Draw a line in your data list to show where the lower quartile is located.

Complete the box-and-whiskers plot below by filling in the indicated numbers.