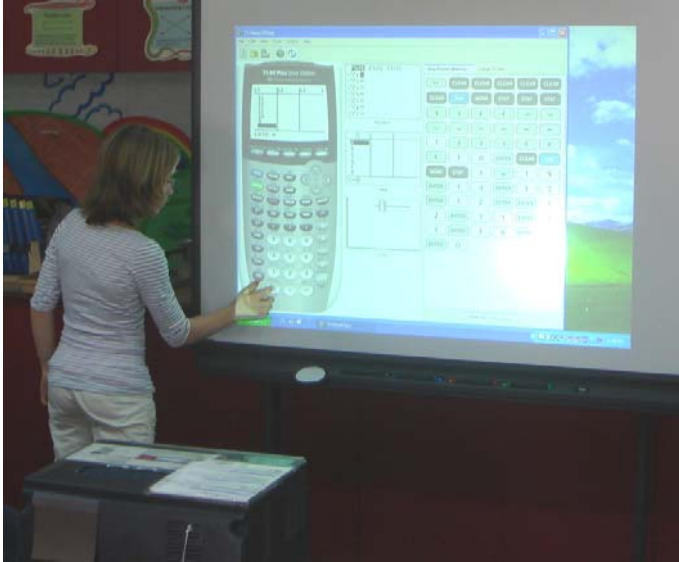


BOX AND WHISKER GRAPHS

A Hands-On Way to Teach Average, Minimum, Maximum, Median, Mode, and Range
With an Integrated Technology Component



Presentation by Tricia and Warren Apel
American Embassy School, New Delhi
NESA Teacher's Conference
April 2007 - Athens, Greece

How to make a human box-and-whisker graph:

Making a human box-and-whisker graph is a fun way to introduce your students to some new statistical and graphing techniques. This is a kinesthetic and hands-on lesson that makes complex math concepts come alive in your classroom. Additionally, it's a fun lesson at the beginning of the year, when students are still learning each other's names.

First, have students arrange themselves in line, ordered by name length. You'll need to establish some ground rules. Primary-age students can use their first name only, and should use whatever name or nickname they put on their nametag. Older students can calculate the length of their first and last names combined (you'll get a larger range of data if students use their full names, such as "Elizabeth" rather than the name they're usually called.)

A fun way to make this more challenging for middle-school students is to have them arrange themselves without speaking. They'll figure out ways of using fingers, sign language, and sticky notes to communicate.



Now that students are in order and know their name-length value, have them arrange themselves on a number line. You can download, print, and laminate our number line — the web address is in the appendix. Or you can simply make one from scratch paper. If your classroom isn't wide enough, you might need to use the hallway for this next step.



Have students stand behind the number that corresponds to their name length value. There may be three or four students lined up behind the more common values in the middle of the graph. Explain that the graph they've just made is a histogram. A histogram is a graph showing the frequency with which values occur. (Later in this packet, we'll explain the difference between a histogram and a simple bar graph).

Now that students have formed a histogram, you can review (or introduce) the concept of minimum, maximum, mode and range.

Minimum and maximum are easily visible

- The *minimum* is the student (or students) with the shortest name length
- The *maximum* is the student with the longest name length



The mode is more difficult to see. It is the most frequently occurring piece of data (the line with the most students.) Students should be able to look around and see that a name-length value, usually one towards the middle of the graph, has more students in it than others. There may be a “tie,” in which case this collection of data has more than one mode.

Now, have students calculate the median. The median is the value in the center of the data when the data is arranged in numerical order. To find the median, you essentially divide the class in half. First, have the students spread back out from the histogram into their numerically ordered line. Hand a ruler to the person at the minimum and the person at the maximum. Each time you say “pass,” have the students pass the ruler one person towards the center. The person who ends up with both rulers is the median. If there is an even number of students, the median will be the average of the two numbers.

Next, explain that for a box-and-whisker graph, the data has to be divided into fourths, or *quartiles*. You may want to use this teachable moment to reinforce that the prefixes “quadri” and “quart” mean four.

In order to further subdivide your students into quartiles, explain that they have to find the value in the center of the median and the maximum, as well as the value in the center of the median and the minimum.

The person at the median value should still have two rulers. Hand another ruler to the minimum and maximum. This time, each time you say “pass,” rulers will move away from the center and in from the ends. The rulers will meet at quartile one and quartile three.

Hand rulers to the minimum and maximum students again. Give the fifth ruler to the student in the median. Now use masking tape to create the box-and-whisker graph. Using the tape, connect the centers of the minimum and first quartile. Connect the centers of the maximum and third quartile. Connect the quartiles and median together at the ends.

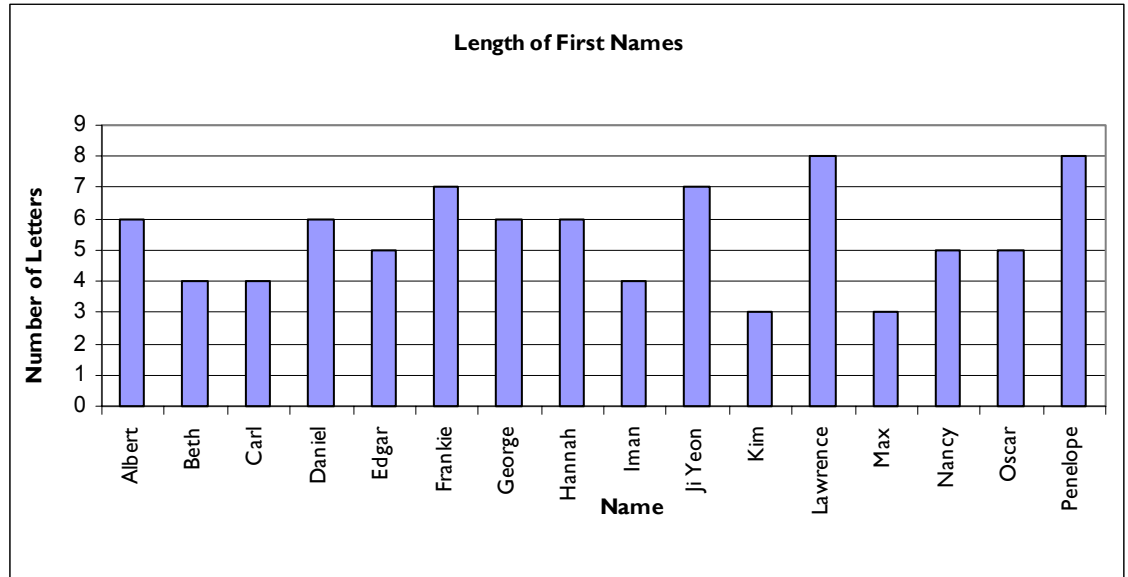
Have any students who are not holding rulers step outside of the graph briefly to observe the graph. Explain to them how the graph gives a very easy visual depiction of the four quartiles, the range, and the median. Quartiles that are physically small have the same number of people in them as quartiles that are longer in size, meaning that they have denser data in them. Being able to see the top half, bottom half, and middle half of data at a glance is very useful when trying to visualize a collection of data points.



What's the difference between a bar graph and a histogram?

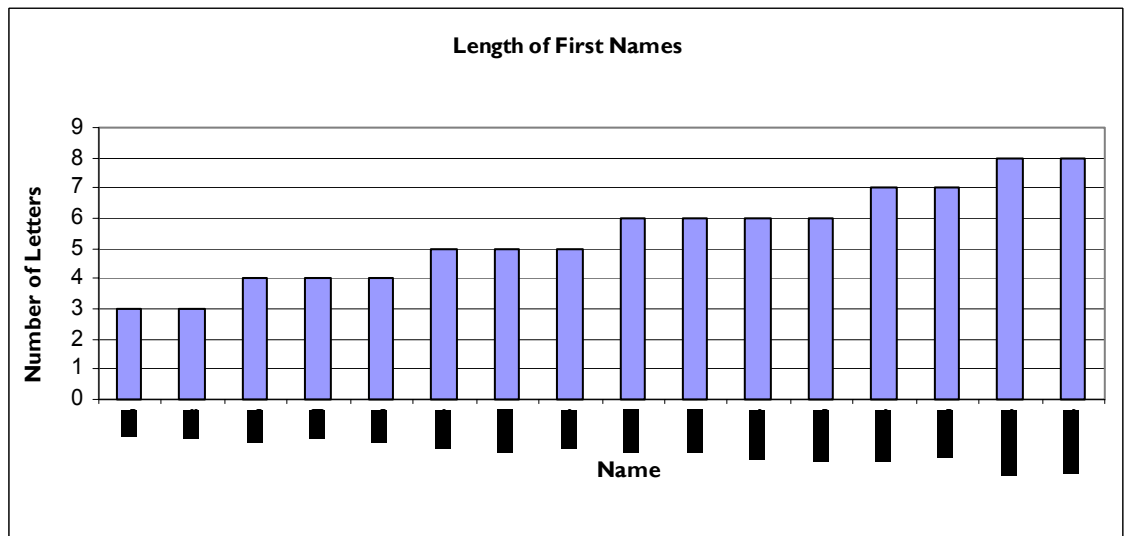
A box-and-whisker graph is a very powerful method to visually display a collection of data. It's easier to see the value by comparing the box-and-whisker graph to some other methods of data display. First, a simple bar graph is just a collection of bars, each one corresponding to a data point. Let's take an imaginary class of students who calculated the value of their first names. Since the bar (or "column") graph is the default choice in Excel, it's likely to be the graph many people choose to display their data.

	A	B
1	First Name	Length
2	Albert	6
3	Beth	4
4	Carl	4
5	Daniel	6
6	Edgar	5
7	Frankie	7
8	George	6
9	Hannah	6
10	Iman	4
11	Ji Yeon	7
12	Kim	3
13	Lawrence	8
14	Max	3
15	Nancy	5
16	Oscar	5
17	Penelope	8
18		



Aside from being able to find a student alphabetically, this graph doesn't really help us understand the data set as a whole. If we sort the class by name length, we can improve our data display a bit.

	A	B
1	First Name	Length
2	Kim	3
3	Max	3
4	Beth	4
5	Carl	4
6	Iman	4
7	Edgar	5
8	Nancy	5
9	Oscar	5
10	Albert	6
11	Daniel	6
12	George	6
13	Hannah	6
14	Frankie	7
15	Ji Yeon	7
16	Lawrence	8
17	Penelope	8



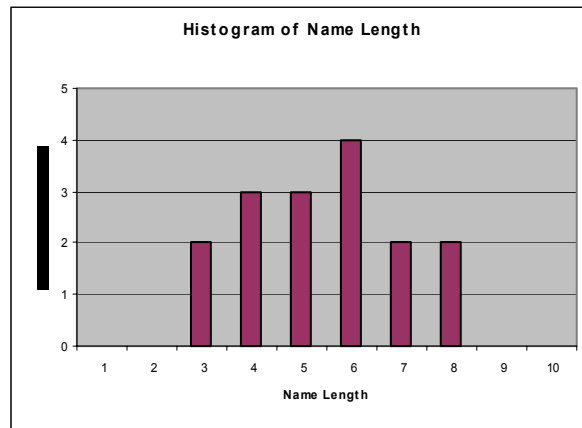
In this second graph, we can see that our mode is 6, since 4 different students have that length. But we still can't make any overall conclusions about the data set.

The next type of graph we could make is a histogram. We count the number of data points at each value. Students typically use tally marks for this. If you are familiar with Excel, you can use a COUNTIF formula to make it easier.

	A	B	C
1	First Name	Length	
2	Kim	3	
3	Max	3	
4	Beth	4	
5	Carl	4	
6	Iman	4	
7	Edgar	5	
8	Nancy	5	
9	Oscar	5	
10	Albert	6	
11	Daniel	6	
12	George	6	
13	Hannah	6	
14	Frankie	7	
15	Ji Yeon	7	
16	Lawrence	8	
17	Penelope	8	
18			
19	Name Length	Number of students	
20	1	0	
21	2	0	
22	3	2	
23	4	3	
24	5	3	
25	6	4	
26	7	2	
27	8	2	
28	9	0	
29	10	0	

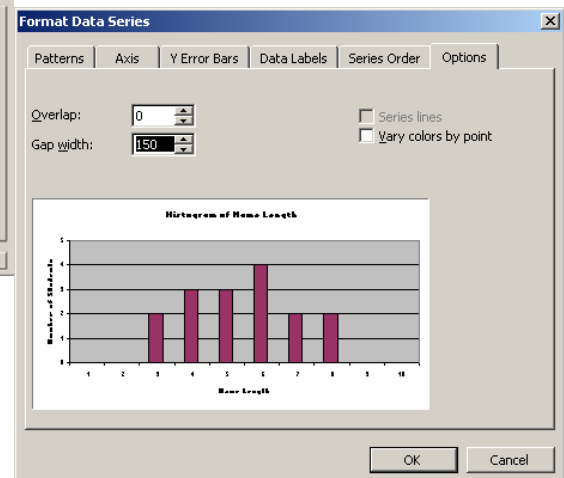
If we graph the tabulated data set, we can get a histogram.

Officially, a histogram must have adjacent bars. By default, Excel puts a gap between each bar, so this is actually just a bar graph of tabular data.

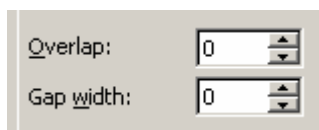


To make it a histogram, we would need to remove the gap. Here's how to do that.

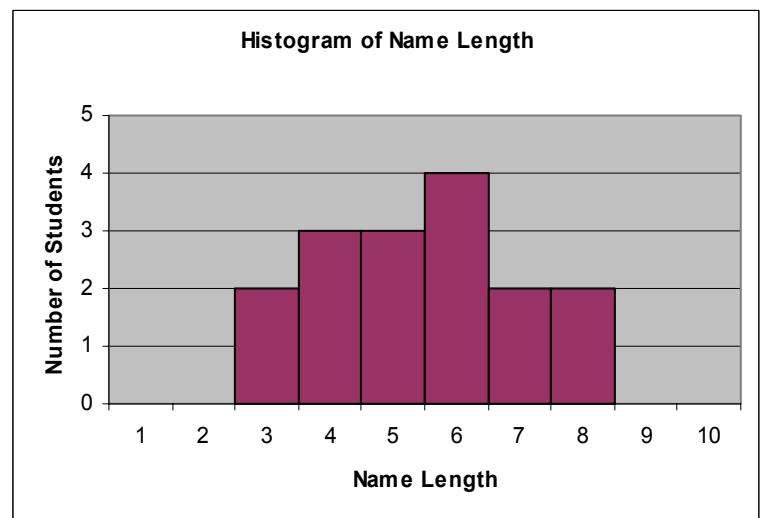
Double-click any of the colored bars to access the menu called "Format Data Series."



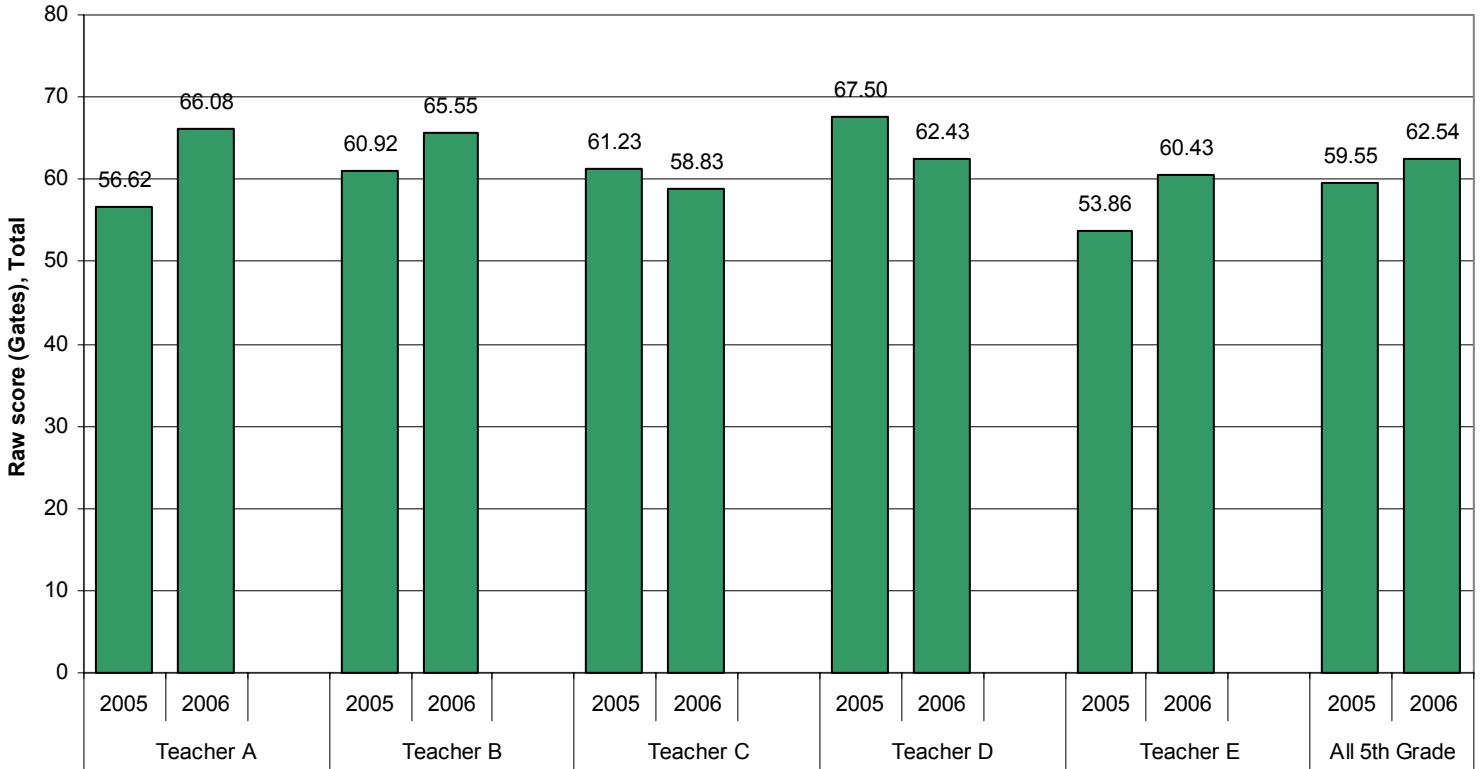
Click on the "Options" tab, and change the gap width from 150 to zero.



Press OK and you will have a real histogram! Notice how the shape of the graph begins to look like a "bell curve." We can really visualize the range, mode, and overall shape of our data set.

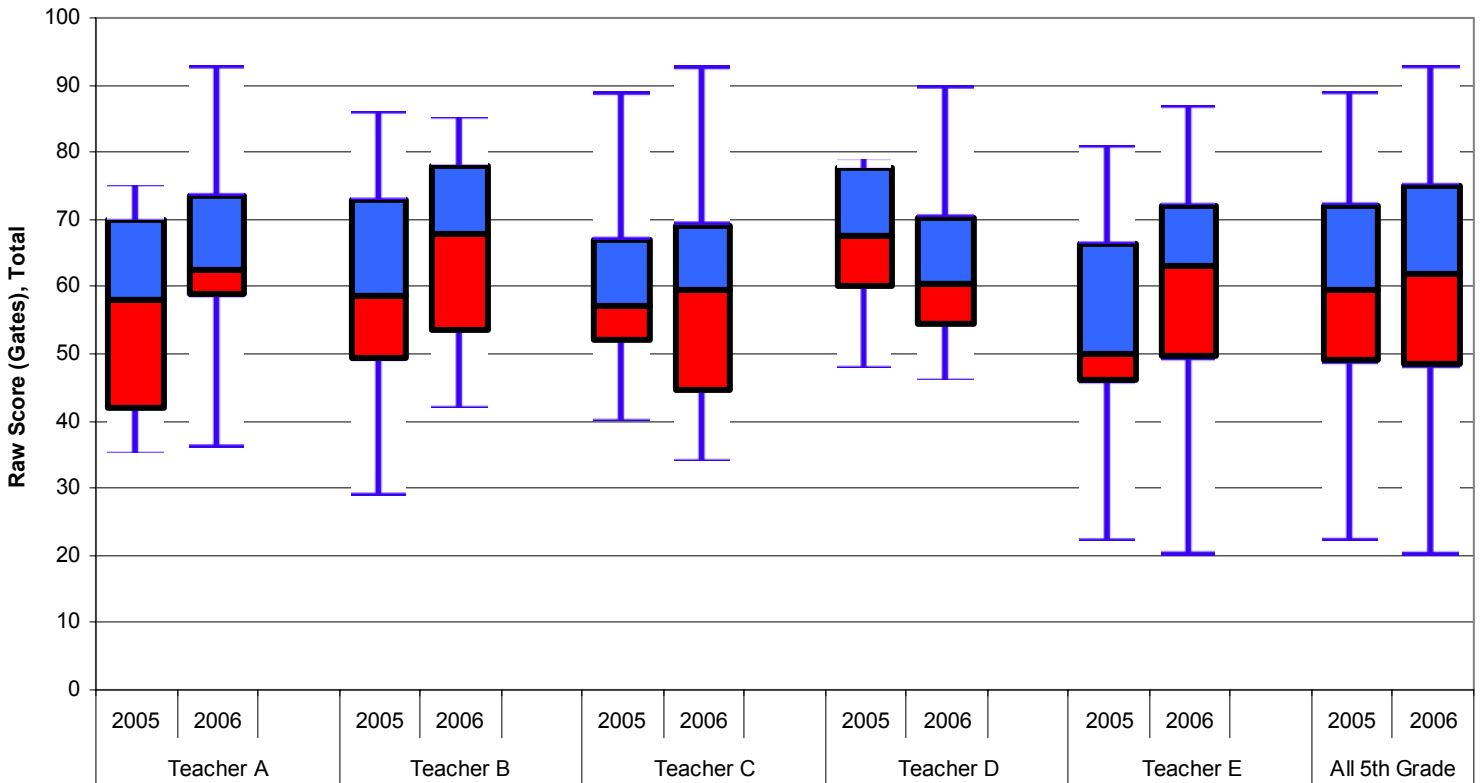


Average Student Performance on Beginning-of-Year Reading Assessment



Compare the results of this same assessment as displayed on an average-only bar graph (above), and a box-and-whisker graph (below.) See the written notes on the next page for commentary on these results.

Box-and-Whisker Analysis Beginning-of-year Reading Assessment



Commentary on the graphs on the previous page:

This is data from a fifth grade reading assessment. I've changed the teacher's names, but they give this assessment at the beginning of each year to see where their students are, and to help them differentiate their teaching for their new group of students each year.

Now you know as teachers that there's this phenomenon where we remember our students as they were at the end of the year, not the beginning. So in our minds, we always compare the new group of young 5th graders with the outgoing group of kids we sent off to sixth grade last May. It's not fair, but to us, every year it seems like the new group of students just aren't reading and cooperating on the same level as our great group of kids from last year. What our fifth grade team wanted to do was step back and compare their new crop of students with last years by using objective data. That way they could make a truly data-driven decision about how this year's group compares to last years', and they could design instructional strategies aimed specifically at moving this group of students forward.

The default type of graph in Excel is the bar or column graph. And the most commonly used statistical calculation among teachers is probably the average. The top graph is what we would have seen if we looked at the average performance from last year to this year.

Each of these sets of two bars is one of the five teachers, the average score of their students at the beginning of last year and this year. And this last set of bars is the overall 5th grade. We can see that in some classes, the average has gone up, in some it's gone down. The change isn't drastic, but it might give you an impression about what level these students are performing at. Now, to put this in perspective, this is a roughly 100-point standardized assessment, and our average scores are close to 60 or 65 across the board. The creators of this assessment say that's about seventh grade reading level, so we have nothing to worry about, but still, Teacher C thinks her students this year aren't as good at reading as last year's.

Now let's look at a box-and-whisker graph of the exact same assessment. Same students, same teachers, but now instead of looking at the average, we're going to divide the students up into quartiles.

If we assume that the class sizes are around 16, each of these quartiles is four students, arranged by score. Let's look at Teacher C, the one whose average score went down by 2 points. Her median increased. That means the top half of the class this year is scoring higher than the top half last year. But her whole class has become more diverse. Her top student is higher than last year's top student, and her lowest is lower than last year's as well. She's going to have to do more differentiating this year.

Let's flip back to the average and look at Teacher A. Her class average has improved. When we look at the box and whisker graph, we see that her lowest performing student is about the same level as last year's. But the bulk of the data – that middle half – is nearly the same as last year's top entire half. Look at this: the top quarter of the class are above any of the students from last year.

Teacher E is in a similar situation. His average went up, and so did his median. Just like Teacher A, three-fourths of his class is currently above 50 points, which is where only half his class was last year. But his lowest student is lower, and his top student is higher. Like Teacher C, he's going to have to differentiate more this year.

But to answer the big overall question – no, this year's students aren't lower than last year's. If anything, they're slightly higher, and slightly more diverse – a wider range of scores with a higher median and average.

Important buttons for building a box-and-whisker graph

WINDOW key—allows you to set the size of the graph window and scale of the graph axes

STATPLOT key—the 2ND function of the Y= key. Press 2ND then this button in order to plot a statistical graph

2nd key—acts like a “shift” to enable additional functions. For example, pressing “2nd” allows the 1 key to enter L1 instead.

STAT key—puts your calculator into statistics mode

1 key—if you press 2ND first, this button will access list one. Notice the small blue letters “L1” above it.

ON key—turns the calculator on. Pressing 2nd first will turn the calculator off.

TRACE key—lets you move quickly around a graph and determine values which have been plotted

GRAPH key—puts your calculator into graphing mode

ARROW keys—move your cursor around the screen

CLEAR key—lets you clear data, in this case from a list

ENTER key—press this after typing a number, and the calculator will enter that number into memory

PARENTHESIS keys—you need to close your parenthesis after entering the list you wish to sort



Directions for building a box-and-whisker graph using the TI graphing calculator:

- Locate a key on the calculator which refers to statistics: the **STAT** key
- To enter new data select **1: EDIT** under the **EDIT** menu option
- Use **L1** for list of one-variable statistics. To clear the **L1** column use the up arrow to highlight the **L1** and press **CLEAR** and **ENTER**.
- Enter data into the **L1** column.
- To sort data go back to the **STAT** key/function. Under the menu option of **EDIT** choose **SORT A** or **SORT D** and press **ENTER**.
- The calculator needs to know what you want sorted so enter **L1** (**2nd KEY** and the number **1 KEY**). Close your parenthesis **)** which is above the number **9** key and **ENTER** to sort.
- View List 1 by pressing **2ND** then **1**, and press **ENTER**
- Use the left and right arrow key to scroll across the data.

```

30000 CALC TESTS
1:Edit...
2:SortA(
3:SortD(
4:ClrList
5:SetUpEditor
    
```

L1	L2	L3	1
10	-----	-----	
12			
15			
9			

L1(5)=

```

SortA(L1)
Done
L1
(9 10 10 12 12 ...
    
```

- To calculate the average return to the **STAT** key and using the right arrow select the **CALC** menu and **ONE-VARIABLE STATS** under the menu option. Press **ENTER** to see the one-variable statistics.
- To make the box-and-whisker graph press **2nd Y=** for the **STAT PLOT** option.
- Turn off existing plots by pressing **PLOTS OFF** and **ENTER**.
- To return to the graphing menu **2nd Y=** and highlight option 1: **PLOT 1** and press **ENTER**.
- Turn on **PLOT 1** by selecting **ON** and pressing **ENTER**.
- Scroll down and select the box and whisker graph and press **ENTER**.
- Press the **WINDOW** button and set the $X_{min}=0$; $X_{max}=($ the

```

STAT PLOTS
1:Plot1...Off
  L1 L2
2:Plot2...Off
  L1 L2
3:Plot3...Off
  L1 L2
4:PlotsOff
    
```

maximum in your data set);
 $X_{scl}=1$; $Y_{min}=0$; $Y_{max}=5$; $Y_{scl}=1$

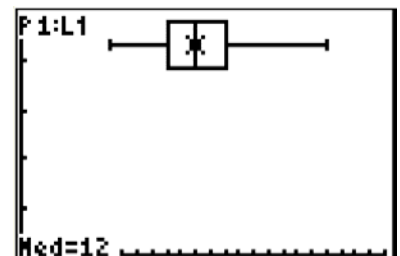
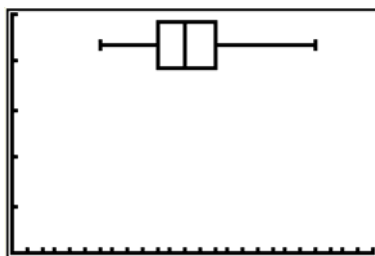
- To see your graph select the **GRAPH** key.
- To see the values on your graph select the **TRACE** key and move it right and left using the arrow keys.

```

20000 Plot2 Plot3
Off
Type: [ ] [ ] [ ]
Xlist: L1
Ylist: L2
Mark: [ ] + .
    
```

```

WINDOW
Xmin=0
Xmax=25
Xscl=1
Ymin=0
Ymax=5
Yscl=1
Xres=
    
```



Directions for using linker cubes for teaching/reinforcing statistics:

- Students build a cube column the length of their full first name (no nicknames) and last name.
- Using the cube columns, students organize themselves in **numerical order** either from smallest to biggest (*ascending order*) or from biggest to smallest (*descending order*).
- The shortest column (smallest) is the *minimum* and the longest (biggest) column is the *maximum*.
- Students can then group themselves by finding students with the exact same column length. I stress that even though we have collected data which repeats, each piece of data is important. Mathematicians and scientists record all the data even if the same number comes up multiple times. Looking at the groups you say that even though students have the same number of letters in their name does not mean that any of them do not matter in the *data set*.
- Students can visually see what group has the most students in it – the numerical data value that is most frequent in the data set is the **mode**.
- Have all students make a large circle and remind them that each person is a single piece of data and all of students in the circle make up the complete data set.
- Count how many students make up the data set.
- All students break apart their cube-column and put their pieces in the center of the circle. Discuss that all the cubes together make up all the letters in everyone's name from the data set. Compare this action with the mathematical operation of **addition**.
- Have students take one cube in turn until the cubes can not evenly be divided among the group. Discuss how the extra cubes in the center represent a part of the whole because they will not go around the whole group again. This would a decimal value. Then for young students and for the purpose of this activity ignore the left-over cubes. Compare the action of dividing the cubes evenly among the group/data set with the mathematical operation of **division**.
- Students rebuild their cube column with the cubes they took from the group pile. Discuss what they notice about the columns. The columns are all the same name length now. The data has been shared/divided evenly and now everyone in the room has the same name length – represented by their new cube-column. This evening-out of the data is **mean** or **average**.

This lesson is a powerful visual and kinesthetic activity for creating concrete understanding of the abstract concept of average.

NCTM Standards:

Data Analysis and Probability Standard Grades 3 ~ 8

- Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them
- Recognize the differences in representing categorical and numerical data
- Collect data using observations, surveys, and experiment
- Select, create and use appropriate graphical representations of data, including histograms, box plots, and scatterplots
- Select and use appropriate statistical methods to analyze data
- Use measures of center, focusing on the median, and understand what each does and does not indicate about the data set
- Find, use, and interpret measures of center and spread, including mean and interquartile range
- Compare different representations of the same data and evaluate how well each representation shows important aspects of the data
- Discuss and understand the correspondence between data sets and their graphical representations, especially histograms, stem-and-leaf plots, box plots and scatterplots
- Develop and evaluate inferences and predictions that are based on data
- Propose and justify conclusions and predictions that are based on data and design studies to further investigate the conclusions or predictions

Principles and Standards for School Mathematics, National Council of Teachers of Mathematics (NCTM), 2000.