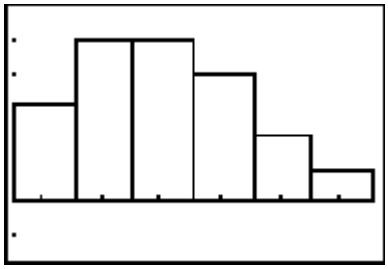
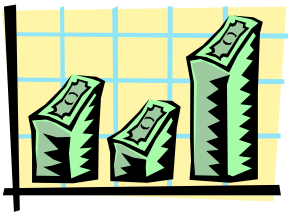
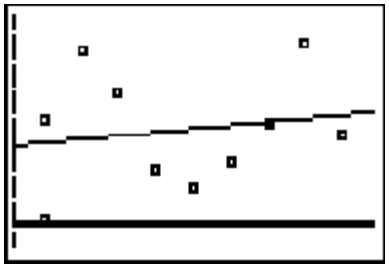


Statistics and Data Analysis

- A five-day unit plan by: Erin Olsen
- For grade level Nine
- Five 45 minute lessons
- Materials: View Screen, TI-83+, Computer, Motion detector, and Cbl





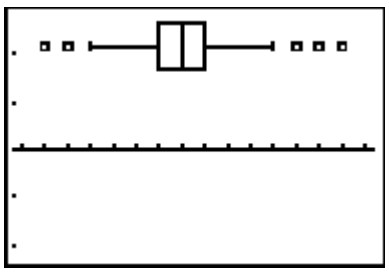


Table of Contents

Objectives and Standards

Students will be able to:

- ✓ Describe relations between measures of central tendency or measures of spread.
- ✓ Select and use appropriate statistical methods to analyze data.
- ✓ Use statistics to describe data sets or compare/contrast data sets.
- ✓ Select, create and use appropriate graphical representations of data, including histograms, box plots, scatter plots, and stem and leaf plots.
- ✓ Develop recognition and understanding of data representations.
- ✓ Discuss and understand the correspondence between data sets and their graphical representations.
- ✓ Develop and evaluate inferences and predictions that are based on data and their graphical representations.

Standards:

NCTM

- Algebra
  - Draw reasonable conclusions about a situation being modeled.
  - Approximate and interpret rates of change from graphical and numerical data.
- Measurement
  - Make decisions about units and scales that are appropriate for problem situations involving measurement.
- Data Analysis & Probability
  - Understand the differences among various kinds of studies and which types of inferences can legitimately be drawn from each.
  - Understand histograms, parallel box plots, and scatterplots and use them to display data.
  - For univariate measurement data, be able to display the distribution, describe its shape, and select and calculate summary statistics.
  - For bivariate measurement data, be able to display a scatterplot, describe its shape, and determine regression coefficients, regression equations, and correlation coefficients using technological tools.
  - Identify trends in bivariate data that find functions that model the data or transform the data so that they can be modeled.
- Communication
  - Communicate their mathematical thinking coherently and clearly to peers, teachers, and others.
  - Analyze and evaluate the mathematical thinking and strategies of others.

    - Use the language of mathematics to express mathematical ideas precisely.
- Connections
    - Recognize and use connections among mathematical ideas.
    - Understand how mathematical ideas interconnect and build on one another to produce a coherent whole.
- Representation
    - Create and use representations to organize, record, and communicate mathematical ideas.
    - Select, apply, and translate among mathematical representations to solve problems.
    - Use representations to model and interpret physical, social, and mathematical phenomena.

NYS

- Modeling/Multiple Representations
    - 4E: Systems of linear equations.
- Measurement
    - 5D: Collecting and organizing data, measures of central tendency, and quartiles and percentiles.
    - 5G: Equation of a line.
- Uncertainty
    - 6C: Probability distribution.
- Patterns/Functions
    - 7D: Determine and model real-life situations with appropriate functions.

Resources

Internet

1. The Shodor Education Foundation, Inc.
    - Robert M. Panoff
    - http://www.shodor.org
2. Illuminations
    - Helene Silverman
    - http://illuminations.nctm.org
3. Regents Exam Prep Center
    - Dr. Kenneth Eastwood
    - http://regentsprep.org
4. Exploring Data
    - Education Queensland
    - http://exploringdata.cqu.edu.au
5. Cartoons
    - Mark Evans
    - http://website.lineone.net/~evans.mw/cartoon.html

Textbook

1. Rubenstein, R. (1992). The University of Chicago School Mathematics Project. Glenview, IL: Scott, Foreman and Company

   *Chapters 1 and 2; pages 1-97

Unit Plan Five-Day Outline

Day One: Defining: Measures of Central Tendency

Students will be exposed to the measures of central tendency including, mean, median, and mode as well as range and frequency. Through group exploration of "pitching" index cards, students will calculate measures of central tendency and spread. By doing this exploration, the students will have the opportunity to make the new knowledge their own. The TI-83+ will be utilized to find the mean and median of a set of data. Students will be able to describe relations between measures of central tendency and where they could be used in real world situations.

Day Two: Stem and Leaf Plots

Students will explore stem-and-leaf plots using real-world data. As a class, the students will discuss the overall shape, distribution, the presence of outliers, and the clustering of data in their plots. In groups, students will explore using a virtual stem-and-leaf plotter. Students will be able to analyze a stem-and-leaf plot and discuss its advantages over other types of plots.

Day Three: Bar Graphs and Histograms

Students will be introduced to histograms, bar graphs, and the concept of class interval. Through a variety of activities including working with a virtual manipulative on the computer, students will learn how data can be graphically represented. The class will take part in a group discussion about the manipulation of the class interval and vertical scales that will deepen and formalize the students' understanding of histograms. Students will be able to distinguish between bar graphs and histograms and know the circumstances when each should be used.

Day Four: Scatter Plots, Correlation, and Linear Regression

Students will graph a variety of data sets by hand and on the TI-83+. Through this, the different types of correlations that exist in a scatter plot graph will be explored. The students will work in pairs to complete a scatter plot by hand and then on the graphing calculator. Together they will find the "best fit" line by hand and then on the TI-83+. The answers the students receive using the two methods will be discussed in groups and then as a class. Students will be able to explain linear regression and how it relates to finding the average rate of change.

Day Five: Box & Whisker Plots

Students will explore the five-number summary and will be able to find each value through a variety of explorations. As a class, the students will take measurements of their heights using a motion detector, CBL, and graphing calculator. Through this exploration, the students will learn how to manipulate collected data on the TI-83+, reveal the five-number summary using 1-Var Stats and by the creation of a box and whisker plot. Students will be able to analyze the box and whisker plot including information about the quartiles, and outliers.

Materials and Equipment

Day One: Defining: Measures of Central Tendency

TI-83+ graphing calculator, overhead view screen, yard sticks, and index cards.

Day Two: Stem-and-Leaf Plots

TI-83+ graphing calculator, computer with Internet access to the website-

http://www.shodor.org, and the weather section of the newspaper

Day Three: Bar Graphs and Histograms

TI-83+ graphing calculator, computer with Internet access to the website-

The National Library of Virtual Manipulatives

Day Four: Scatterplots, Correlation, and Linear Regression

TI-83+ graphing calculator, overhead view screen, spaghetti, and wall charts for measuring heights and arm spans.

Day Five: Box and Whisker Plots

Motion Detector, CBL, TI-83+ graphing calculator, and DATAMATE software

Day One:
Defining: Measures of Central Tendency

Lesson Summary

Students will learn the about the measures of central tendency and spread. They will learn how to calculate each one and then participate in a hands-on activity that will allow the students to incorporate their new knowledge.

Materials

TI-83+ graphing calculator, overhead view screen, yard sticks, and index cards.

Objectives

The student will be able to:

- ✓ Identify the measures of central tendency.
- ✓ Calculate each measure given a set of data.
- ✓ Record data collected through group exploration.
- ✓ Compare relationships between measures of central tendency and spread.
- ✓ Develop individual sets of data and identify the measures of central tendency.
- ✓ Decide where the measures of central tendency can be used in a real world situation.

Standards

- *NCTM*
  Data Analysis & Probability, Communications, and Connections
- *NYS*
  Key Idea: Measurement: 5D

Anticipatory Set

For many students this lesson is a review of the measures of central tendency. Therefore activating their prior knowledge is important. Each student will be given a worksheet containing groups of related words. The students will choose the correct term from the word box.

Instructional Procedure

A concept elaboration strategy will be conducted in order to introduce new vocabulary where each student will:

1. Write the book's definition of the term.
2. Examples that will help each student remember the key points.
3. Restate the definition using their own words.

*The students will do this for the following words: mean, median, mode, range, and frequency.

Modeled Practice

1. A set of forty math test scores out of ten points will be presented on the overhead projector.
2. As a class we will prepare a frequency table of the scores.
3. Then, using the frequency table the mean, median, mode, and range will be calculated.
4. A class discussion of the results will be discussed and interpreted.
5. Following the discussion, the students will be shown how to calculate the mean and median on their graphing calculator.

Guided Practice

1. Students will get into groups of three and be given ten index cards labeled 1-10.
2. The students will then find a space in the classroom or hallway where they can “pitch” or throw the cards one at a time.
3. Another student will measure the distance in inches from the starting point to the spot where the card fell.
4. The third student will record the information on his/her worksheet.
5. Each student will take turns at each station.
6. In groups, the students will determine the mean, median, mode, and range for the collected data.
7. While the students are collecting their data, I will be walking around the room to facilitate them in their learning and provide my assistance when it is needed.
8. Students will hand in their paper containing their answers and collected data so that I will be able to check their understanding of the information.

Closure

Each student will write a paragraph on where they could use measures of central tendency in a real world situation.

Independent Practice/Homework

Pages 27-28, problems 1-10, 14,16, and 17.

Measures of Central Tendency and More

Fill in the blank with the appropriate key term that describes the other three.

1. Mean_____
   Arithmetic Average
   Around

2. Median_____
   Middle Value
   Central Point
   Mid Point

3. Range_____
   Span
   Greatest # minus Lowest #
   Spread

4. Mode_____
   Most Common
   Most Often
   Appears the most

5. Frequency_____
   Occurrence
   Happening
   Indicated with a tally mark

Word Box:	**Mode**	**Mean**	**Range**
	Frequency	**Median**	

Concept Elaboration Strategy

Define: Mean-The sum of a list of numbers, divided by the total number of numbers in the list.

Give an example: 20+50+34+13 = 117

117/4 = 29.25

Restate the definition: Mean is the sum of the numbers divided by how many numbers there are.

Define: Median- "Middle value" of a list. If the list has an odd number of entries, the median is the middle entry in the list after sorting the list into increasing order. If the list has an even number of entries, the median is equal to the sum of the two middle (after sorting) numbers divided by two.

Give an example: 13, 20, 34, and 50

Median = 27

13, 20, 34, 50, 67

Median = 34

Restate the definition: Median is the number in the middle.

Define: **Mode-** The most common (frequent) value. A list can have more than one mode.

Give an example:

20, 15, 19, 20, 17

20 is the mode because it appears the most.

Restate the definition: The number that appears the most, there can be none, and there can be more than one.

Define: **Frequency-** The number of items occurring in a given category.

Give an example:

20, 15, 19, 20, 17

Number	Frequency
15	|
17	|
19	|
20	||

Restate the definition: How many times each number appears.

Define: **Range-**The range of a set of numbers is the largest value in the set minus the smallest value in the set.

Give an example: 15, 17, 19, 20, and 20

20-15 = 5

5 is the range.

Restate the definition: The greatest number in the set minus the smallest number in the same set.

Transparency

Forty students took a math test marked out of ten points
Their results are as follows:

9, 10, 7, 8, 9, 6, 5, 9, 4, 7, 1, 7, 2, 7, 8, 5, 4, 3, 10, 7, 3, 7, 8, 6, 9, 7, 4, 2, 3, 9, 4, 3, 7, 5, 5, 2, 7, 9, 7, 1

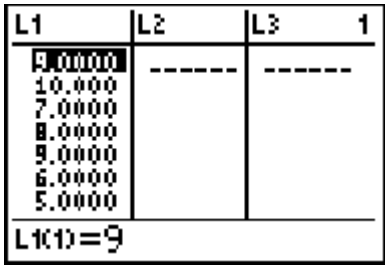
a.) Prepare a frequency table of the scores
b.) Using the frequency table, calculate the mean, median, mode, and range
c.) Interpret these results

Score on test	Frequency
1	||
2	|||
3	||||
4	||||
5	||||
6	||
7	|||| ||||
8	|||
9	|||| |
10	||

Mean: 5.9
Median: 7
Mode: 7
Range: 9

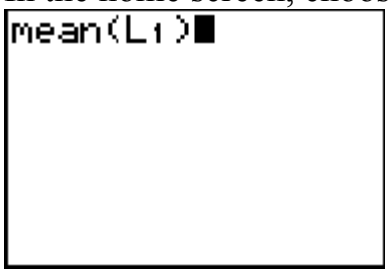
Using the TI-83+ to find mean and median

Enter data in L1

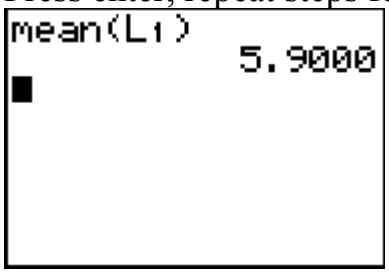


Choose math and then mean under the list menu

In the home screen, choose L1



Press enter, repeat steps for median



Worksheet

Throw each card, measure each distance in inches, record your data, and calculate the measures of central tendency.

Card Number	Distance in inches
1	42.5
2	56
3	20
4	54.2
5	65.6
6	33.4
7	67.8
8	42
9	38.5
10	42.5

Mean: 46.25 inches__

Median: 42.5 inches__

Mode: 42.5 inches___

Range: 47.8 inches___

Day Two
Stem and Leaf Plots

Lesson Summary

Students will explore stem-and-leaf plots using real-world data. As a class, the students will discuss the overall shape, distribution, the presence of outliers, and the clustering of data in their plots. In groups, students will explore using a virtual stem-and-leaf plotter.

Materials

TI-83+ graphing calculator, computer with Internet access to the Website www.shodor.org, and the weather section of the newspaper.

Objectives

The student will be able to:

- ✓ Show how the stem and leaf plot works.
- ✓ Construct a stem and leaf plot given a set of data and calculate the measures of central tendency.
- ✓ Use a virtual stem-and-leaf plotter to plot a set of chosen data and calculate the measures of central tendency.
- ✓ Examine the graph and identify specific outliers and peaks.
- ✓ Determine the general shape of the distribution using their collected data.
- ✓ Judge whether the use of the stem and leaf plot is advantageous over other types of plots.

Standards

- *NCTM*
  Algebra, Data Analysis & Probability, Communications, and Representations
- *NYS*
  Key Idea 5: Measurement: 5D

Homework Check

The answers to last night's homework will be placed on the overhead. Students will be allowed five minutes to check their answers and discuss with their neighbor any problems they had. The teacher will be walking around the room to see if each student made a conscious effort to complete the assignment.

Anticipatory Set

The students will be asked to list as many ways data can be organized. The student with the longest list will be rewarded. Hopefully he/she will mention a stem-and-leaf plot and then be able to explain what a stem-and-leaf plot is to the class. If not, I will prompt the students to look in their books to see what plot they missed on their list.

Instructional Procedure

1. Explain that a stem-and-leaf plot is another way a set of data can be organized and displayed.
2. This type of plot shows the shape and distribution of the data.
3. Each data value is split into a "stem" and "leaf".

4. The “leaf” is the last digit of the number and the other digits form the “stem”.

Modeled Practice

1. Each student will be given a copy of the weather section from the newspaper.
2. Pick ten cities; record the city and the high temperature for that corresponding city.
3. Creating a stem-and-leaf plot will be demonstrated on the overhead
4. Each student will follow along and create a stem-and-leaf plot that incorporates the temperatures of their chosen cities.
5. Each student will then calculate the mean, median, and mode for their set of data.
6. The students will then pair up and describe their stem-and-leaf plot with their partner discussing the location, overall shape, outliers, and clustering of each.

Guided Practice

1. In a room with computers that have Internet access, students work in groups of two.
2. Using a virtual stem-and leaf-plotter located at www.shodor.org the students will create a plot using any numbers and as many numbers as the want.
3. After exploring the virtual plotter, the students will chose 20 numbers and calculate the mean, median, and mode.
4. The students will record their data as well as their answers and hand their papers in before they leave.

Closure

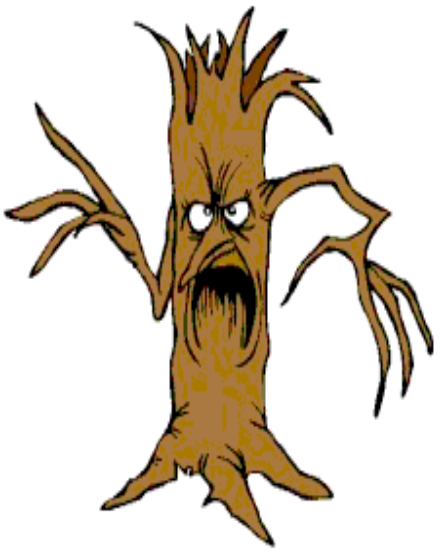
As an exit slip write one reason why you would use a stem-and-leaf plot over another type of plot.

Independent Practice/Homework

Complete the stem-and-leaf plot worksheet

Weather Data

City	High Temperature
Atlanta	57
Boston	52
Cincinnati	44
Detroit	38
Kansas City	49
Las Vegas	60
Miami	84
San Juan	88
Seattle	54
Tampa	83



*Construct a stem-and-leaf plot in the space below and find the mean, median, and mode.

3	8
4	4 9
5	2 4 7
6	0
7	
8	3 4 8

Mean=60.9
Median=55.5
Mode=N/A

Name: _____

Stem-and-Leaf Plot

Please answer the following questions using a Stem-and-Leaf Plot:

1. Your class just took your last math test of the year. These are the scores: 97, 99, 81, 78, 73, 95, 33, 97, 64, 100, 85, 83, 85, 88, 79, 81, 93, 86, 83, and 71.
    a. What are the mean, median, and mode of the data?
    b. Describe the shape of the data set. Is it symmetric or are there extreme values in the low or high numbers?
    c. Given this information would you say that the students did well on the exam? Explain.

    Please show your work:

2. You decide to get odd jobs for the summer and you keep a log of your daily earnings. In one month you earned the following amounts: $12, $9, $15, $17, $20, $12, $9, $6, $13, $15, $11, $17, $18, $14, $15, $20, $25, $13, $12, $15.
    a. What are the mean, median, and mode of the data?
    b. Describe the shape of the data set. Is it symmetric or are there extreme values in the low or high numbers?
    c. What was your average daily income? If you worked approximately 4 hours a day, what was your average hourly income?

    Please show your work:

3. During the summer you also play in a recreational soccer league. Your coach asks you to quickly determine what was the most common score for the season. She was asking you to find the _____ (mean, median, or mode) of the list of scores. Your team's scores were: 9, 5, 2, 10, 7, 6, 7, 2, 5, 7, 9, 10, 9, and 6.
    a. What are the mean, median, and mode of the data?
    b. Describe the shape of the data set. Is it symmetric or are there extreme values in the low or high numbers?
    c. Draw the appropriate stem-and-leaf plot and write your answer for the most common score.

4. An extra challenge: You visit the grocery store with your dad and he is using a whole lot of coupons in order to save money. You decide to find the mean, median, and mode of the coupons and you want to use a stem-and-leaf plot, but then you realize that the coupons contain decimals. How can you use the stem-and-leaf plot?

Here are the values of the coupons: $0.50, $0.75, $0.30, $1.00, $0.45, $0.30, $0.75, $0.25, $0.25, $1.50, $0.75, $0.45, $0.30, and $0.50. Use the method that you described to find the mean, median, and mode of these values!

NAME: __ANSWER KEY_____

1.

3	3
4	
5	
6	4
7	1 3 8 9
8	1 1 3 3 5 5 6 8
9	3 5 7 7 9
10	0

a. Mean= 82.55 Median= 84 Mode=81, 83, 85, 97

b. The presence of an outlier (33) makes the set asymmetrical.

c. The students did well on the exam because the average of all the test scores is 82.55. If you take out the 33, the average increases to 85.16.

2.

0	6 9 9
1	1 2 2 2 3 3 4 5 5 5 5 7 7 8
2	0 0 5

a. Mean = 14.4 Median = 14.5 Mode = 15
b. The shape of the data is symmetric.
c. The average daily income is $14.40. If you worked 4 hours a day, the average hourly income would be $3.60.

3.

0	2 2 5 5 6 6 7 7 7 9 9 9
1	0 0

a. Mean = 6.714 Median = 7 Mode = 7, 9
b. The data is not symmetric, but there are no extreme high or low values.
c. MODE

4.

.2	5 5
.3	0 0 0
.4	5 5
.5	0 0
.6	
.7	5 5 5
.8	
.9	
1.0	0
1.1	
1.2	
1.3	
1.4	
1.5	0

a. Mean = $0.575 Median = $0.475 Mode = $0.30, $0.75

Day Three
Bar Graphs and Histograms

Lesson Summary

The goal of this lesson is to introduce the students to histograms and the concept of class interval. Through a variety of activities including using a virtual manipulative, the students will learn how data can be graphically represented. Students will be able to distinguish between bar graphs and histograms and determine the context in which each should be used.

Materials

TI-83+ graphing calculator, computer with Internet access to the website- The National Library of Virtual Manipulatives

Objectives

The student will be able to:

- ✓ Identify the differences between bar graphs and histograms.
- ✓ Construct a histogram given a set of data.
- ✓ Read and understand the overall shape of the data, symmetry, location, spread, and the existence of outliers.
- ✓ Interpret information in a histogram and match it to its associated summary statistics.
- ✓ Deduce the most accurate histogram through exploration of different intervals and scales.
- ✓ Determine when advantages and disadvantages exist of using the histogram over other types of plots.

Standards

- *NCTM*
  Algebra, Measurement, Data Analysis & Probability, Communication, Connections, and Representations
- *NYS*
  Key Idea 5: Measurement: 5D
  Key Idea 6: Uncertainty: 6C

Homework Check

Students will be allowed five minutes to discuss their homework in groups of four. The teacher will be walking around the room to check to see if the students made a conscious effort to complete the assignment. Also the teacher will answer questions if the students are having trouble.

Anticipatory Set

Students should already be familiar with the bar graph because it is frequently used in other content areas such as social studies and science to represent a set of data. Also, bar graphs are frequently used in magazines and newspapers. A variety of articles from newspapers/magazines that utilize the bar graph to represent data will be presented to the class. Students will brainstorm how and what types of information is presented in a bar graph.

Instructional Procedure

1. The concept of the histogram will be introduced to the class.

2. Students will be shown how their stem-and-leaf plot from the previous day resembles a histogram when it is rotated onto its side.
3. The differences between the bar graph and histogram will be explained and in what situations you use each one.
4. A concept elaboration strategy will be conducted in which each student will:
    - Write the book's definition of the term
    - Examples that will help each student remember the key points
    - Restate the definition using their own words.

*The students will do this for the following words: bar graph, histogram, and class interval.

Modeled Practice

1. As a class, we will create a histogram for a set of test scores by hand.
2. We will explore the histogram using different vertical scales.
3. We will explore the histogram using different class intervals.
4. We will discuss which graph represents the data the best.
5. The measures of central tendency will be calculated.
6. Exploration using the TI-83+ graphing calculator to manipulate the intervals of a histogram will be discussed.

Guided Practice

1. In the computer lab, the students will be able to use a virtual manipulative to explore histograms.
2. The source of the manipulative is *The National Library of Virtual Manipulatives-Histograms*
3. Students will complete the Analysis of Histogram worksheet that corresponds to the computer activity.

Closure

Discuss and review as a class what the differences are between a histogram and bar graph. Write a paragraph in response to if you think the histogram has advantages/disadvantages over the other graphical representations we have learned about. Why or why not?

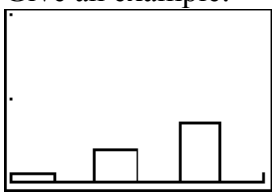
Independent Practice/Homework

Pages 41-43, Problems 1-10, and 12 in addition to completing the matching histograms and summary statistics worksheet.

Concept Elaboration Strategy

Define: **Bar Graph-** a diagram showing a system of connections or interrelations between two or more things by using bars

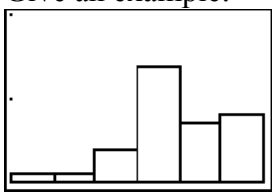
Give an example:



Restate the definition: A histogram with gaps between the bars.

Define: **Histogram-** a bar graph such that the area over each class interval is proportional to the relative frequency of data within this interval

Give an example:



Restate the definition: A bar graph with no gaps.

Define: **Class Interval-** dividing the range of all values into non-overlapping intervals, called class intervals, in such a way that every piece of data is contained in some class interval

Give an example:
0-10, 11-20, 21-30, 31-40, and 41-50

Restate the definition: dividing the data into equal intervals such that every piece of data falls into one and only one interval.

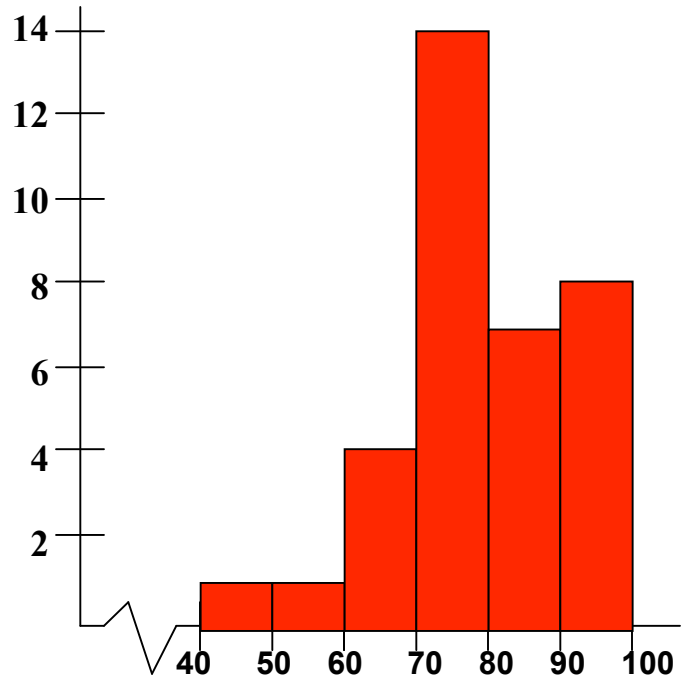
Transparency

Ms. Chair's math class's test scores

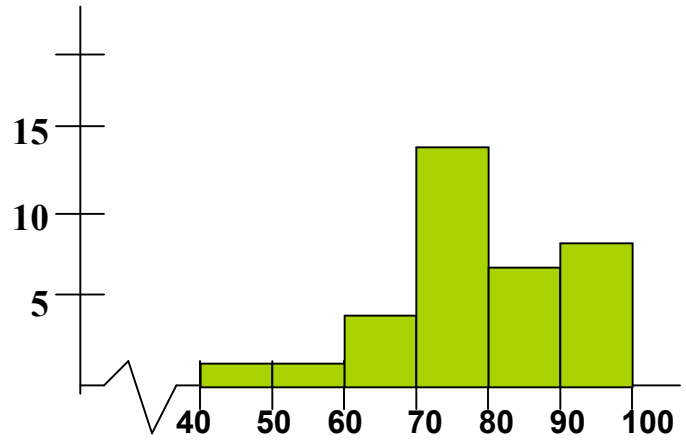
43 68 73 78 80 88 92
52 70 74 78 82 89 93
65 70 75 78 85 90 94
66 71 75 78 87 90 94
67 72 76 79 87 90 98

a. Construct a histogram for the above test scores.
    1. Use a vertical scale of 2 and a class interval of 10.
    2. Use a vertical scale of 5 and a class interval of 10.
    3. Use a vertical scale of 2 and a class interval of 20.
    4. Use a vertical scale of 5 and a class interval of 20.
b. Decide which graph you think best fits the data and why.
c. Calculate the measures of central tendency.

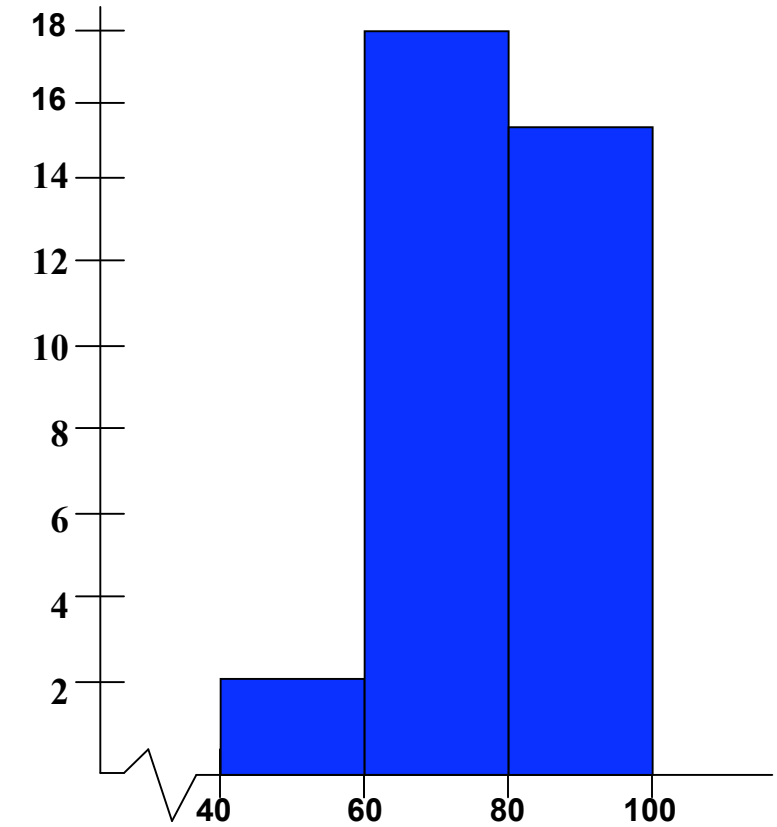
A-1



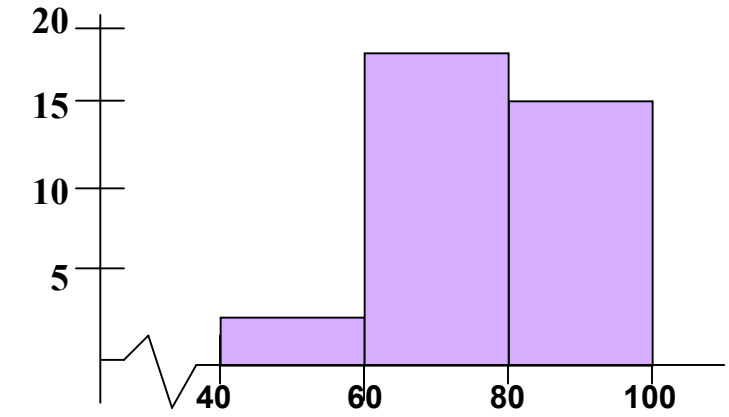
A-2



A-3



A-4



Mean-78.4857
Median- 78
Mode-78
Range-55

Analysis of Histograms

Choose a set of data and list it here:

Then answer the questions pertaining to the set of data you chose.

1. Change the vertical and horizontal scales for the set of data and observe changes in the histogram.

2. Is data increasing or decreasing along the horizontal axis?

3. Does the answer depend on the horizontal scale? Vertical scale?

   No

4. Can you make the data look more stable by changing the scales? Can you make it look unstable?

   Yes

5. Find a place where a slight change in the horizontal scale leads to a big change in the appearance of the histogram. How and why does this happen?

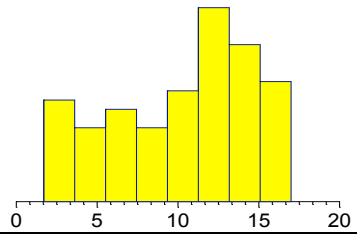
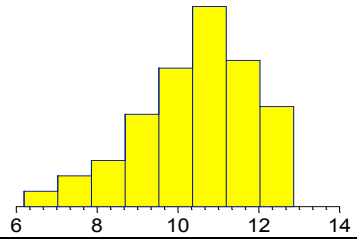
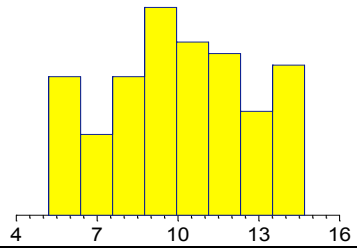
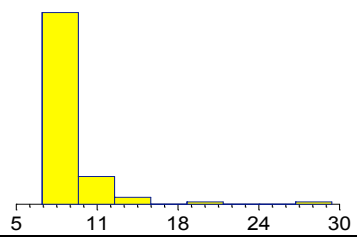
   If for example there are values for a set of data in a certain interval, changing the size of the interval will affect the appearance of the histogram.

6. A histogram can sometimes be misleading. State one situation where this is true.

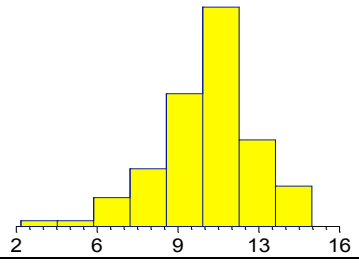
   For example, if a data set contained a value that was extremely high (or low) you could arrange the intervals to include that value so that the histogram appears to not have any outliers.

Matching Histograms and Summary Statistics

Match each histogram with a set of summary statistics, by writing the letter in the space provided.

1. _____D_____ 	A. Mean 10.5 Median 10.7 IQR 2.0
2. _____A_____ 	B. Mean 10.1 Median 10.1 IQR 4.2
3. _____B_____ 	C. Mean 10.2 Median 10.5 IQR 2.5
4. _____E_____ 	D. Mean 10.2 Median 11.9 IQR 6.8

5. C



E.

Mean	8.8
Median	8.0
IQR	1.9

Day Four
Scatter Plots, Correlation, and Linear Regression

Lesson Summary

Students will learn how to make a scatter plot and find a "best fit" line. Students will work in pairs to gather two sets of data and record them on a class chart. The whole class will work to design a scatter plot of the data. Class discussion will focus on interpreting the meaning of individual coordinates, the overall graph, and the "best fit" line.

Materials

TI-83+ graphing calculator, overhead view screen, spaghetti, and wall charts for measuring heights and arm spans.

Objectives

The student will be able to:

- ✓ Plot points on a given set of coordinates.
- ✓ Construct scatter plots of two variable data and interpret individual data points.
- ✓ Classify certain data points as outliers and examine the measures of central tendency.
- ✓ Construct a "best fit" line by hand and with the TI-83+ graphing calculator.
- ✓ Determine the equation of the "best fit" line and draw conclusions about the trends in data.

Standards

- *NCTM*
  Algebra, Data Analysis & Probability, Problem Solving, Communication, and Representations
- *NYS*
  Key Idea 4: Modeling/Multiple Representation: 4E
  Key Idea 5: Measurement: 5D, 5G
  Key Idea 7: Patterns/Functions: 7D

Homework Check

The answers to last night's homework will be placed on the overhead. Students will be allowed five minutes to check their answers and discuss with their neighbor any problems they had. The teacher will be walking around the room to see if each student made a conscious effort to complete the assignment.

Anticipatory Set

The students will be presented with a mathematical cartoon about statistics and outliers. The cartoon is intended for the students to activate their prior knowledge about statistics including mean, median, and outliers. The students will be asked for their interpretations of the cartoon. This discussion will lead into the introduction of scatter plots and how outliers are identified.

Instructional Procedure

2. A concept elaboration strategy will be conducted in which each student will:
    - Write the book's definition of the term
    - Examples that will help each student remember the key points
    - Restate the definition using their own words.

*The students will do this for the following words: Scatter plot, correlation (positive, negative, and none), and outliers

Modeled Practice

1. A set of data will be presented on the overhead projector relating the cost of engine repairs per year and the number of oil changes per year.
2. As a class, the points will be plotted on a coordinate graph.
3. Discussion about the points and their correlation will take place.
4. The teacher will pass out spaghetti to each student and they will be asked if they can visualize a "best fit" line that would represent the summary of the data
5. The student will be instructed to lay their piece of spaghetti on that imaginary line.
6. The teacher will then demonstrate on the overhead view screen how to accurately find the "best fit" line by doing linear regression.
    - Enter data in L1 and L2.
    - Choose the scatter plot option under STAT PLOT.
    - Graph the plot and explain to the students that this should look like the work that they just did by hand.
    - Press the STAT button, scroll over to the CALC menu, choose LinReg(ax+b).
    - This should paste LinReg(ax+b) in the home screen.
    - Next, paste L1, L2, such that your screen looks like LinReg(ax+b) L1,L2,
    - Next press the VARS button, scroll over to Y-Vars, choose function, and then choose Y1. This step pastes the equation of the line into Y1.
    - Press ENTER to complete the calculation.
    - Finally press GRAPH to see how your "best fit" line looks on your scatter plot.
7. Conduct a class discussion to see how close the students were to finding the "best fit" line with their spaghetti.

Guided Practice

1. The class will conduct an experiment to determine if there exists a relationship between a person's height and their arm span.
2. Students will work in pairs to collect the data of their height and arm span.
3. After each student has their data, they will then record it on a class chart.

4. Each student pair will work together to plot the points in order to construct a scatter plot on the TI-83+ calculator.
5. Then they will determine the “best fit” line and be able to perform linear regression.

Closure

Individually, each student will write a paragraph discussing the appearance and interpretation of the scatter plot and its “best fit” line. They will also be asked to include the measurement of a person’s arm span that is 7 foot tall according to their “best fit” line.

Independent Practice/Homework

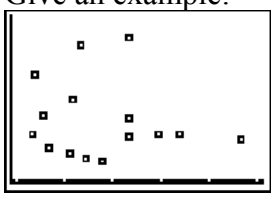
Page 22, problem 14
Page 97, problem 10

Cartoon (transparency)

Concept Elaboration Strategy

Define: **Scatter plot-**A graph on a set of coordinates that displays plotted points of a set of data that has two variables.

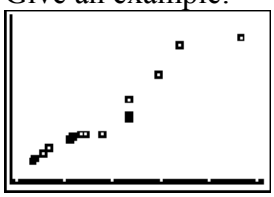
Give an example:



Restate the definition: Points on a graph that represent data.

Define: **Positive correlation-**A set of data that has a "best fit" line whose slope is positive.

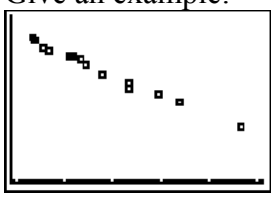
Give an example:



Restate the definition: When the points on a scatter plot resemble the equation y=x.

Define: **Negative correlation-**A set of data that has a "best fit" line whose slope is negative.

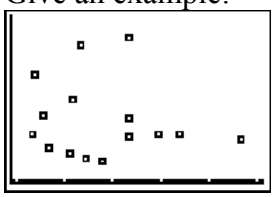
Give an example:



Restate the definition: When the points on a scatter plot go downhill.

Define: **No correlation-**A set of data that has a "best fit" line whose slope is close to zero.

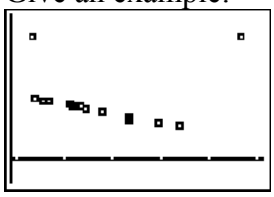
Give an example:



Restate the definition: When the points are all over the place and you can't tell whether they are going up or down.

Define: **Outliers-**A data point (or points) that lie far outside most of the rest of the points in the data set.

Give an example:



Restate the definition: Any point that lies outside the area where the majority of points are plotted.

Transparency

Cost of repair	# of oil change
300	3
300	5
500	2
400	3
700	1
400	4
100	6
250	4
450	3
650	2
600	0
0	10
150	7

1. The table gives data relating the number of oil changes per year to the cost of car repairs. Plot the data on the grid provided, with the number of oil changes on the horizontal axis.

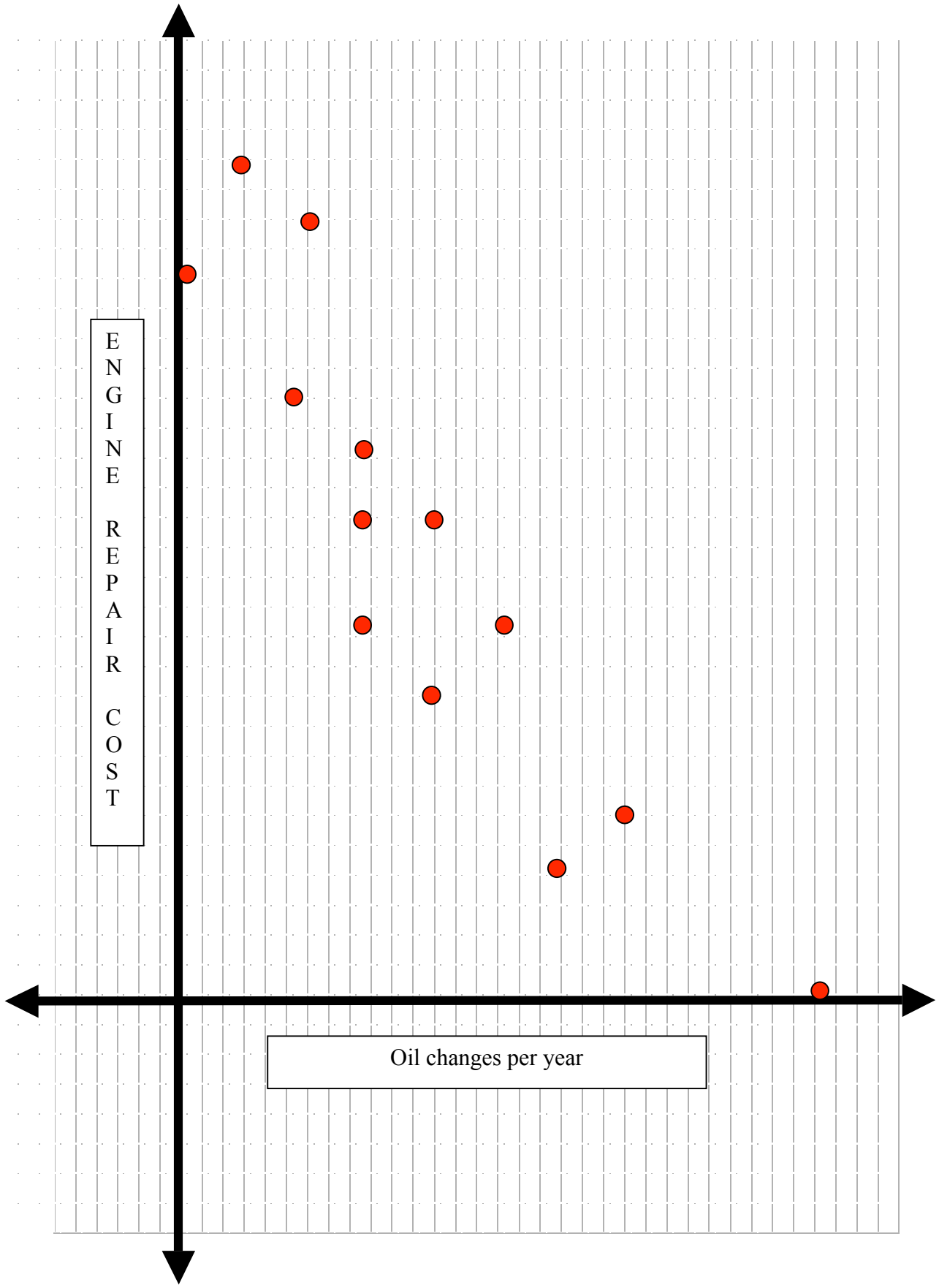
2. Is the data linear? If so, draw a best-fit line.
   Yes

3. Find the slope of the line. Describe in words what the slope represents.
   Slope = -100

4. Find the x- and y-intercepts. Explain in terms of oil changes and engine repairs what each represents.
   X-intercept = 7.5 (oil changes), y-intercept = 750 (dollars)

5. Write the equation of the line.
   Y= -100x + 750

6. Use the equation to predict the cost of engine repairs if the car had four oil changes. How accurate do you think your prediction is? Explain your answer.
   $350
7. Explain the relationship between the number of oil changes and the cost of repairs.
   The more oil changes the car has results in fewer car repairs.


ENGINE REPAIR COST
Oil changes per year

Day Five
Box & Whisker Plots

Lesson Summary

The goal of this lesson is for the students to explore the five-number summary as well as the concept of outliers. Through discussion and two activities students will learn how to create a box and whisker plot by hand and by using the TI-83+. Students will be able to determine the advantages and disadvantages of using a box and whisker plot and are able to compare it to a histogram.

Materials

TI-83+ graphing calculator, CBL, motion detector, DATAMATE software, pencils, worksheets, overhead, overhead view screen, strips of number paper.

Objectives

The student will be able to:

- ✓ Name the key concepts in the five-number summary
- ✓ Construct a box and whisker plot using the TI-83+ calculator and identify the five-number summary given a set of data.
- ✓ Make use of a motion detector, CBL, and TI-83+ to collect data.
- ✓ Analyze the data collected using statistical concepts.
- ✓ Construct a variety of plots on the TI-83+ using the collected data.
- ✓ Determine the advantages and disadvantages of using the box and whisker plot to describe the collected data.
- ✓ Interpret information given in a box and whiskers plot and match the distribution to that of its associated histogram.

Standards

- *NCTM*
  Algebra, Data Analysis & Probability, Communication, Connections, and Representations
- *NYS*
  Key Idea 5: Measurement: 5D
  Key Idea 6: Uncertainty: 6C

Homework Check

The teacher will choose students to present one of their answers on the board from last night's homework so that every problem is displayed on the board. Students will check their answers while each student is presenting his or her answer. Homework will be collected and checked.

Anticipatory Set

The students will be presented with a problem that relates to a basketball game and the score at the end of each quarter. It is easy to find the score after each period, but what if you had a set of data and you wanted to find the number that lies at the 25^{th} percentile, 50^{th} percentile, and the 75^{th} percentile. The students will already know how to find the number at the

50th percentile because it is the median. The students will use their prior knowledge to brainstorm how to find the lower quartile as well as the upper quartile.

Instructional Procedure

A concept elaboration strategy will be conducted in which each student will:

1. Write the book's definition of the term
2. Examples that will help each student remember the key points
3. Restate the definition using their own words.

*The students will do this for the following words: box & whisker plot, lower quartile, upper quartile, maximum, and minimum.

Modeled Practice

1. Students will be given rectangular strips of paper containing a set of data that is equally spaced in ascending order.
2. The student will be instructed to fold the paper into four equal parts and then unfold the strip.
3. Once the paper is unfolded, the class will discuss what each fold means, if the fold does not pass directly through a data value, discuss how you would find this value.
4. After discussion, show students how to create a box and whisker plot and also a modified one containing outliers.
5. Have students enter the data from the strip into the TI-83+, choose the box and whisker plot option and graph the plot.
6. Have students find the values of the numbers in the five-number summary, compare their drawing with the calculator and discuss if they match the folds in the strips of paper.

Guided Practice

1. Measure the heights of students using the CBL and motion detector.
    - Position the motion detector about 2.5 meters from the floor, making sure there are no obstructions beneath the detector.
    - Plug the motion detector into the CBL and the CBL into the TI-83+.
    - Turn the TI-83+ on and start the DATAMATE software.
    - The screen will show the distance to the floor in meters.
    - Have the motion detector record the distance to the head of each student and not record the distance to the floor in between the student measurements; set up DATAMATE to record only selected events.
        - Select SETUP from the main screen.
        - Press up to move to the mode selection.
        - Press ENTER to move to the SELECT MODE menu.
        - Select SELECTED EVENTS from the SELECT MODE menu.
        - Select OK to return to the main screen.

        - Select START.
    - Have the first student stand directly under the motion detector. Press ENTER to make a measurement.
    - Repeat for all students.
    - When finished, press the STORE button to end data collection.
    - Press ENTER to see a plot of distance readings. These need to be converted into heights.
        - Press ENTER to return to the graph selection screen.
        - Select MAIN SCREEN to return to the main screen.
        - Select QUIT to leave DATAMATE.
        - Follow directions on calculator screen to return to the home screen.
        - Then enter the distance to the floor in meters and subtract the list that contains the heights and then store the expression to the same list.
        - For example, 2.5 – L1 STO_ L1.
        - Press ENTER to complete the operation.
- To calculate a number of statistics all at once, use the 1-Var Stats function.
- Press STAT and scroll over to open the CALC menu. Press ENTER to paste the 1-Var Stats function to the home screen.
- Press 2^{nd} [L1]
- Press ENTER to complete the calculation.
- Scroll down to reveal the values of the mean, minimum, median and maximum as well as the lower quartile and upper quartile.

2. Have students create a box & whisker plot to view the graph of the data distribution of their collected data.

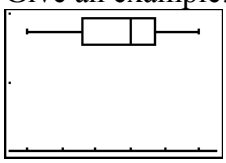
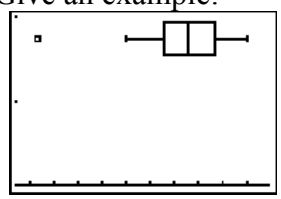
Closure

Complete the Data Collection and Analysis worksheet with a partner. This activity will allow the students to think critically and assess their knowledge of what was learned though the exploration.

Independent Practice/Homework

Pages 35-36, Problems 1-10, 12, and 13 in addition to the matching box plots and histogram worksheet.

Concept Elaboration Strategy

Define: **Box & Whisker Plot-** is a visual representation of the five-number summary of a data set. The ends of the box are the upper and lower quartiles; the whiskers extend to the highest and lowest values and are no longer than 1.5 times the interquartile range. Give an example:  Restate the definition: a graphical representation of the five-number summary.	Define: **Lower Quartile-**Is the median of the numbers below the location of the median. Give an example: The 25th percentile Restate the definition: The median of the lower part of the data.
Define: **Upper Quartile-**Is the median of the numbers above that location. Give an example: The 75th percentile Restate the definition: The median of the upper part of the data.	Define: **Minimum-**The smallest value in a set of data. Give an example: 12, 24, and 36 Minimum = 12 Restate the definition: Smallest number
Define: **Maximum-** The largest value in a set of data Give an example: 12, 24, and 36 Maximum = 36 Restate the definition: Largest number	Define: **Outlier-** Individual points that lie beyond 1.5 times the interquartile range or the whisker. Give an example:  Restate the definition: Extreme high and

Strips of paper for quartile exercise

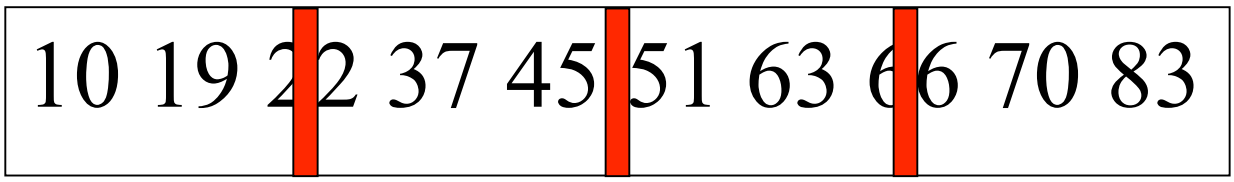
10 19 22 37 45 51 63 66 70 83

10 19 22 37 45 51 63 66 70 83

10 19 22 37 45 51 63 66 70 83

10 19 22 37 45 51 63 66 70 83

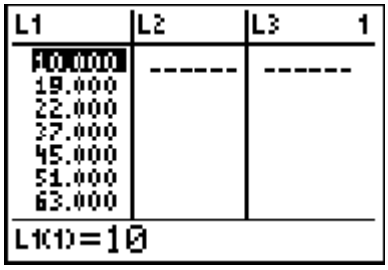
The red lines indicate the folds in the strip of paper



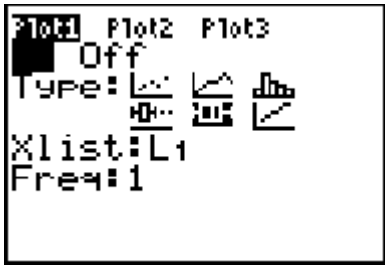
Lower Quartile Median Upper Quartile

Using the TI-83 for Box & Whisker Exploration

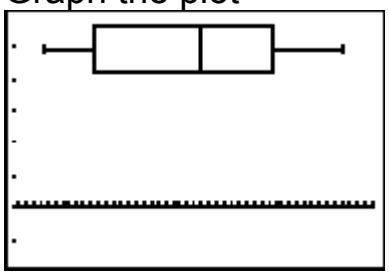
Enter data in list 1



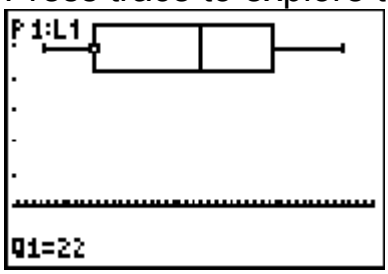
Choose the box & whisker plot



Graph the plot



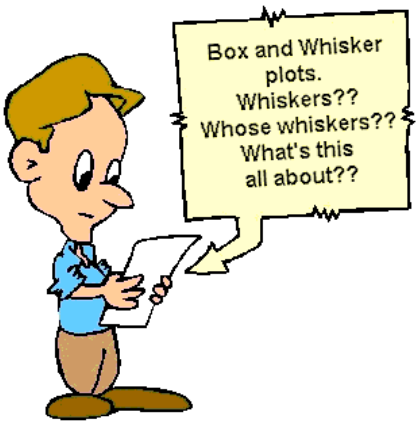
Press trace to explore the box plot



Name: _____ Partner: _____

Box & Whisker Plot
DATA COLLECTION and ANALYSIS

Data Table



Number of Students=30	From Box & Whisker Plot
Minimum	4'8''
Median	5'5"
Mean	5'52"
Maximum	6'4"
Lower Quartile	5'3"
Upper Quartile	5'8"

Questions

1. Notice that the box part of the plot represents the middle portion of the data, while the whiskers stretch to the lowest and highest numbers in the data set. The size and location of the box tells you about the distribution of values in the data set. A large box indicates that the data is spread out, while the smaller box means the data is clustered closely. Discuss the size and location of the box part of the plot; describe how it relates to the heights of the members of the class.

   The box is somewhat small compared to the range of the whiskers, implying that the class heights are somewhat bunched to the middle values.

2. Is the median located near the center of the box? What does the location of the median line in the box tell you about the distribution of heights for the middle half of the data?

The median is 5’5”, which is to the left of center in the box. This means that there is a smaller difference between the minimum and median then there is between the median and maximum.

3. The length of the whiskers on the box plot gives a hint as to the distribution of the data. If one whisker is significantly longer than the other, we say the data is skewed in the direction of the shorter whisker. This usually means that the values are bunched together near the shorter whisker. Describe the whiskers on the plot. What do the whisker lengths tell you about the heights for the class?

   The whiskers are about the same lengths on the plot, which tells you that the shortest person is about as much shorter than average as the tallest person is taller than average.

4. The presence of an outlier might cause you to think that the data is skewed in one direction or another when it really is not. Can you identify any outliers in the data set?

   There are no clear outliers, so the answer to the previous question does not change.

5. Describe how the box plot would be affected if one of the members of your class were exceptionally tall at 7’5”.

   The right-hand whisker would extend considerably, the upper quartile value would increase slightly, but the other values would remain the same.

6. Suppose that the values for the tallest and shortest members of the class were removed from the data set. How would the median, lower quartile, and upper quartile values change, if at all?

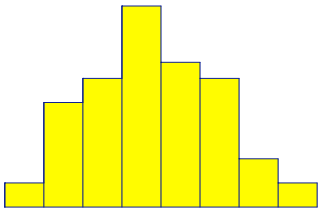
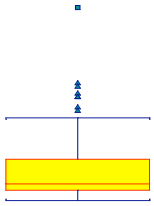
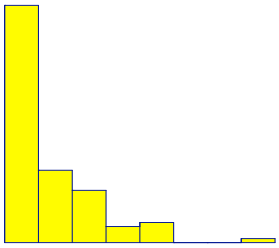
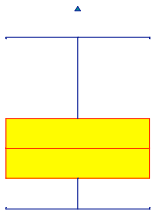
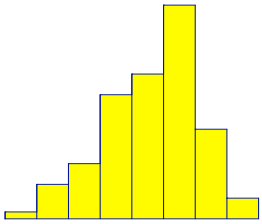
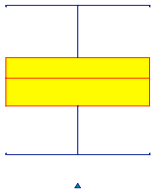
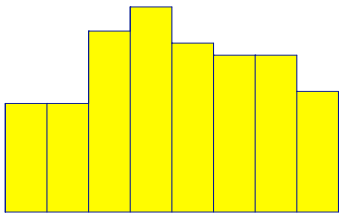
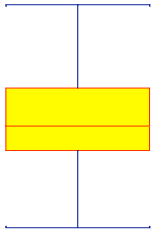
The median value would not change, while the lower quartile value would increase slightly. The upper quartile value would decrease slightly.

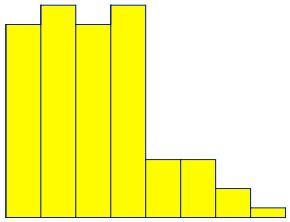
7. Determine the advantages and disadvantages of using the box and whisker plot over other types of plots including histograms and stem and leaf plots.

   The advantages of the box plot is that this plot shows the lower quartile, median, and upper quartile where in the histogram and stem-leaf plot these measures are not easily seen. The disadvantage of the box plot is that it does not display the frequency of each number, nor does it show each individual number as does the stem-leaf plot does.

Matching Histograms and Box plots

Match each histogram with its box plot, by writing the letter of the box plot in the space provided.

1. __D_____ 	A. 
2. ____A___ 	B. 
3. ____C___ 	C. 
4. ___E____ 	D. 

5. B 	E.