BLUE RIBBON
DATA ANALYSIS and PROBABILITY PROJECT

Karen Yoho
Marion County Technical Center

PRESENTATION TO PEERS

DATE: September 27, 2002

PLACE: Marion County Technical Center

PEER GROUP: Staff Members of the Marion County Technical Center

These staff members represented the areas of agriculture, auto mechanics, aviation, business, CISCO training, communication, collision repair, construction, counseling, distributive education, forestry, health science, machine technology, manufacturing, special education, transportation, welding, and work-based learning.

None of these instructors had used the Fathom Dynamic Statistics™ Software, and they had little or no experience with the TI-83 Plus graphing calculator. Yet they were eager to experience these new technologies.

GOALS:

(1) To give the teachers a background in statistics that would help them prepare their students for the SAT 9 Test in the spring.

(2) To give the teachers experience with Fathom Dynamic Statistics™ Software and the TI-83 Plus graphing calculator, technologies that they had not yet experienced.

TARGET GROUP OF STUDENTS:

Applied Math 1 and 2, Algebra I, Conceptual Math, Algebra II
(and all students who will be taking the SAT 9 Test)

PRELIMINARY ACTIVITY

As the teachers entered the room, they were asked to provide some data for a part of the presentation that would be discussed later.

(1) With fingers and thumb spread apart as far as possible, each person used a meter stick to measure his or her span (the distance from the tip of the thumb to the tip of the little finger) in centimeters.

(2) Using only one hand, each person grabbed as many "gems" as possible from one bowl, placed them in another bowl, and counted them. (The "gems" are small glass objects similar to marbles.)

(3) Besides recording the information from the first two steps, each person was asked to write the number of the day on which they were born. (For example, if they were born on July 4 of some year, they were supposed to write "4".)

(4) Each teacher then walked to the front of the room and entered this data in a chart, using the Fathom Dynamic Statistics™ Software program.

[Please do NOT write your name on this paper.]

STATISTICS DATA FORM

Length of your "span" _____cm

[Distance from tip of thumb to tip of little finger]

Number of gems _____

[Number transferred to the second bowl.]

The number of the <u>day</u> of your birth _____

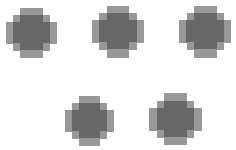
[Not the month or year]

Spread the fingers and thumb on one of your hands as far apart as possible.

Measure the distance (in centimeters) from the tip of your thumb to the tip of your little finger.

(Use decimals in the measurement if necessary.)

Record this information.



Put one hand behind your back.

Using the other hand, gather as many gems as possible in your hand.

Without turning that hand over, put these gems in another bowl.

Count the gems that were in your hand.

Record this information.

Return the gems to the original bowl (for the next person to use).

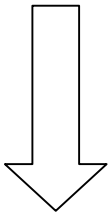
Write the number of the day you were born.
(For example, if your birthday is July 4, write “4”.)

Using the FATHOM SOFTWARE on the COMPUTER,

ENTER YOUR DATA IN THE CHART

and

PLACE YOUR COMPLETED FORM
HERE.



HOME-RUN LEADERS of the AMERICAN LEAGUE for the Years 1980-1990

Year	Player	Home Runs
1980	Reggie Jackson	41
1981	Eddie Murray	22
1982	Reggie Jackson	39
1983	Jim Rice	39
1984	Tony Armos	43
1985	Darrell Evans	40
1986	Jesse Barfield	40
1987	Mark McGwire	49
1988	Jose Canseco	42
1989	Fred McGriff	36
1990	Cecil Fielder	51

William M. Setek, Jr. (1992). *Fundamentals of Mathematics*, p.200.

MEAN = ____ MEDIAN = ____ MODE = ____ RANGE = ____

1. Determine the MEAN of the home run data.
   (Add the numbers in the "Home Runs" column and divide by the number of pieces of data.)

2. Determine the MEDIAN of the home run data.
   (Arrange the numbers in the "Home Runs" column in order from smallest to largest or from largest to smallest. The median is the number in the middle. If there are two numbers in the middle of the list, then average the two numbers by adding them and dividing by 2. If you have found the correct value of the median, then the number of pieces of data smaller than that value will be the same as the number of pieces of data larger than that value.)

   List the number of home runs in ascending, or descending, order below:

   ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____

3. Determine the MODE of the home run data.
   (Choose the number that occurs most often.)

4. Determine the RANGE of the home run data.
   (Subtract the smallest number of home runs from the largest number.)

Now use the TI-83 Plus graphing calculator to help you determine some of these values.

1. To enter the home run data in a list, press the **STAT** key and select **EDIT** (Since EDIT is highlighted, you can press ENTER or you can press "1".)

2. Be sure that the List L1 column is empty. (If there are numbers in the list, use the UP ARROW to highlight L1 and then press CLEAR and ENTER.)

3. Enter the data in the home run column in List L1, pressing **ENTER** or the **DOWN ARROW** after each number of home runs.

4. To have the numbers in List L1 arranged in ascending or descending order, press **STAT**. To have the list arranged in ascending order, press "**2**" for **SORTA(**. To have the list arranged in descending order, press "**3**" for **SORTD(**. You are now on the Home Screen.

5. To indicate which list is to be sorted, press **2nd L1** (L1 is above the "1" key), the right parenthesis "**)**", and **ENTER**.

6. To view List L1 in its new order, press **STAT** and **EDIT**. Notice that the numbers are now in order.

7. To determine the median, notice that 2 divides into 11 (the number of pieces of data) 5 whole times with a remainder. This means that there will be 5 pieces of data smaller than the median and 5 pieces of data larger than the median. Therefore, we are looking for the 6^{th} number in the list. Press the **DOWN ARROW** until **L1(6)** = appears in the lower left corner of the screen. This will tell you the 6^{th} number in List L1, which is the median for this list.

   Median = _____

8. To determine the mode with the graphing calculator, use the **UP** and **DOWN ARROWS** to search through List L1 for the number that occurs most often.

   Mode = _____

9. To determine the range, use the first and last numbers in the list; subtract the smaller number from the larger one.

Range _____

10. There's another way to determine the mean, median, and range with the TI-83 Plus.

Press **STAT** and then the **RIGHT ARROW** (to highlight **CALC**). Select **1-Var Stats** by pressing "**1**" or "**ENTER**" (since "1" is highlighted). You will now be on the Home Screen. Press **ENTER**. "$\bar{x}$" indicates the mean. So the mean is 40.18 (as determined in a previous step). Press the **DOWN ARROW** to locate "Med"; the median is 40 (as determined in Step 7). Since the minimum value, "minX", in the list of data is 22 and the maximum value, "maxX", is 51, subtract to find the range. (This should be the same value determined in Step 9.)

AGE AND HEIGHT
Children in Kalama

This data represents the average heights of a group of children in Kalama, an Egyptian village that is the site of a study of nutrition in developing countries. The data were obtained by measuring the heights of all 161 children in the village each month over several years.

The height of a child is not stable but increases over time. Since the pattern of growth varies from child to child, one way to understand the general growth pattern is by using the average of several children's heights.

AGE (in months)	HEIGHT (in centimeters)	GROWTH (in centimeters)
18	76.1	XXXXXXXXXX
19	77	0.9
20	78.1	
21	78.2	
22	78.8	
23	79.7	
24	79.9	
25	81.1	
26	81.2	
27	81.8	
28	82.8	
29	83.5	

David S. Moore and George P. McCabe (1989). *Introduction to the Practice of Statistics*, p. 118.

1. Complete the chart by determining the number of centimeters that the children grew each month. (For example, to decide how many centimeters the 19-month-olds grew, subtract the 18-month-old height from the 19-month-old height: 77 – 76.1 = 0.9 cm.)

2. What is the average number of centimeters that the children <u>grew</u> each month? (That is, determine the MEAN.)

_____

TI-83 PLUS GRAPHING CALCULATOR

3. Using the TI-83 Plus graphing calculator, enter the ages in List L1 and the heights in List L2. [Press **STAT** and **EDIT**; then enter the numbers in the appropriate list.]

4. To see a scatter plot of "Heights versus Ages", press **2nd** **y=** to get **STAT PLOT**. Press **ENTER** or "**1**" to get **Plot 1**. With **On** highlighted, press **ENTER** to turn the plot on.

5. Under "Type", highlight the first graph in the list and then press **ENTER**. For "Xlist" (the numbers that will appear on the x-axis), enter **L1** by pressing **2nd** "**2**". for "Ylist" (the numbers that will appear on the y-axis), enter **L2** by pressing **2nd** "**3**". Highlight the type of mark that you would like to have appear on your scatter plot, and then press **ENTER**.

6. Press **WINDOW**. Since the ages ranging from 18 to 29 months will be on the x-axis, type Xmin = **17**, Xmax = **31**, and Xscl = **2**. Since the heights ranging from 76.1 to 83.5 cm will be on the y-axis, type Ymin = **75**, Ymax = **85**, Yscl = **5**. ["Xscl" and "Yscl" indicate how far apart to put the "tick" marks on the x-axis and y-axis, respectfully.]

7. Press **GRAPH**.

8. Which of the following equations describes the line that best fits the data? [H = the height and A = the age.]
   (A) H = A (B) H = 80 (C) H = 0.63A (D) H = 0.63A + 65

   To check by using the TI-83 Plus graphing calculator, use y for H and x for A.

   Press "**y=**" and type "x" after "y1=". Press **GRAPH** to see how well the line y = x (that is, H = A) fits the data.

   Repeat this process for each of the other three equations until you can determine the equation of the line that best fits the data.

9. How many centimeters tall will a child in Kalama probably be at age 30 months?

   With the graph of the line of best fit displayed in the calculator window, press **TRACE**. Notice in the upper left corner of the screen that the trace cursor is on "P1:L1,L2", meaning "Plot 1 that uses the numbers in Lists L1 and L2". Press the **DOWN ARROW**. Now the trace cursor is on the line of best fit (y 1 = 0.63x + 65).

   Press the **RIGHT ARROW** several times and watch the cursor move up the line. Also, watch the changes in the x-values and y-values of the points (in the lower part of the screen). Continue pressing the **RIGHT ARROW** until you get as close to x = 30 as possible. What is the approximate value of y for x = 30?

10. To determine the equation for the line of best fit for a set of data on a scatter plot, press **STAT** and **RIGHT ARROW** (for **CALC**) and then press "**4**" [for "LinReg(ax+b)"]. You will be back on the Home Screen. Press **ENTER**. Notice that a = 0.63 and b = 64.9 or 65. Substitute these values in the equation y = ax + b to get y = 0.63x + 65, the line of best fit.

CORRELATION

Correlation is a numerical measure of the strength of a linear association. A correlation of "0" means that there is almost no relationship between two sets of data. Correlations of "1" and "-1" indicate strong relationships.

To determine the correlation coefficient on the TI-83 Plus, be sure that the "DiagnosticOn" feature is in effect. [After pressing **2nd CATALOG** and "**D**", use the **DOWN ARROW** to locate DiagnosticOn. Press **ENTER** twice.]

Press **STAT**, **CALC**, "**4**", and **ENTER**. The number after "r=" is the correlation coefficient.

1. Use the TI-83 Plus graphing calculator to graph the following data in a scatter plot with "Percent versus Age" (that is, "age" on the x-axis and "percent" on the y-axis).

BUDGET SPENT ON HEALTH CARE

AGE	PERCENT
20	26
30	35
40	42
50	45
60	59

Great Source Education Group (2000). *Algebra to Go* (p. 365)

Notice that the points in the scatter plot for "Budget Spent on Health Care" almost lie in a straight line and rise to the right. The fact that the series of points rises to the right indicates that the correlation is positive. The fact that the points are almost in a straight line that rises to the right indicates that the correlation is probably close to 1.

Use the TI-83 Plus to determine the correlation coefficient.

_____

2. Graph the following data in a scatter plot with "Percent versus Age".

AMUSEMENT PARK ATTENDANCE

AGE	PERCENT
20	76
30	70
40	68
50	52
60	40

Great Source Education Group (2000). *Algebra to Go* (p. 365)

Notice that the points in the scatter plot for "Amusement Park Attendance" descend to the right. This indicates a negative correlation. The fact that the points are almost in a straight line that descends to the right indicates that the correlation is probably close to –1.

Determine the correlation coefficient. _____

3. Graph the following data in a scatter plot with "Grade versus Height".

AVERAGE GRADES

HEIGHT (in inches)	GRADE
50	70
52	82
54	82
56	77
58	90
60	50
62	78
64	70
66	72
68	92
70	58
72	72
74	87
76	80

Great Source Education Group (2000). *Algebra to Go* (p. 365)

In the scatter plot of the data for "Average Grades" there does not seem to be a distinct pattern for the location of the points. The series of points are not close together and does not seem to rise or fall. This would indicate that a person's height is not closely related to the grade that he or she makes. Therefore, the correlation is probably near 0.

Determine the correlation coefficient. _____

4. Consider again the "Age and Height" data for the "Children in Kalama". Predict the value of the correlation coefficient and give reasons for your prediction. Then use the TI-83 Plus to determine the correlation coefficient and compare it to your prediction. [Graph the data in a scatter plot of "Height versus Age".]

_____

FATHOM DYNAMIC STATISTICS™ SOFTWARE

1. Consider again the "Budget Spent on Health Care", using Fathom.
   (A) Drag a case table from the shelf into the document.
   (B) Click once on **<new>**.
   (C) Type **Age** for the first attribute and press **Enter**.
   (D) Click once on **<new>**.
   (E) Type **Percent** for the second attribute and press **Enter**.
   (If necessary, drag on the right edge of the table to make the table wider.)
   (F) Double-click on the label **Collection 1** in the title of the table (or in the title of the collection).
   (G) Use this dialog box to rename the collection.
   Type **Budget Spent on Health Care** and click **OK**.
   (H) Enter the data in the two columns.
   (Drag the lower edge of the table to make it longer or shorter, if necessary.)
   (I) To make a scatter plot, drag the graph icon from the shelf to an empty area in the document (or choose **Graph** from the **Insert** menu).
   (J) Drag the word **Age** to the horizontal axis of the graph over the spot labeled **Drop an attribute here**.
   (K) Drag the word **Percent** to the vertical axis of the graph.
   (L) Adjust the size of the graph by dragging on the right edge and/or lower edge.
   (M) To determine the correlation, create a summary table.
      (1) Drag the summary table icon to the document or choose **Summary Table** from the **Insert** menu.
      (2) Drag the **Percent** attribute to the summary table and drop it on the down-pointing arrow.
      (3) With the summary table selected, choose **Add Formula** from the **Summary** menu.
      (4) Click on the + in front of **Functions**, **Statistical**, and **Two Attributes**. Then double-click on **correlation**.
      (5) Click on the + in front of **Attributes**.
      (6) Double-click on **Age**, type "," (a comma), double-click on **Percent**, and click **OK**.
      (7) Notice that the correlation appears in the summary box below the value of the mean.
2. Use Fathom to make a case table, scatter plot, and summary table that would display the correlation for (a) "Amusement Park Attendance", (b) "Average Grades", and (c) "Age and Height of Children in Kalama".

GEM EXPERIMENT DATA

Students

SPAN (in centimeters)	Number of GEMS	Number of DAY of BIRTH
24.3	65	21
20.2	56	9
22	45	12
24	51	8
23	57	18
16	33	26
19.8	51	17
19	38	4
20.5	49	24
21.7	49	9
19.5	39	24
23.2	46	6
22.8	44	21
22	42	21
20.3	38	18
20	37	18
22	56	17
21	43	2
21.5	32	23
21	51	13
22	48	26
19	27	6
20.2	29	18
23.1	54	23
24	10	12
23.1	42	15
21	76	3
23	47	10
24	62	19
20	43	22
23	55	16
23	34	27

1. What is the average size hand "span" of the students in the experiment?

2. What is the average number of gems that each student removed from the original bowl?

3. What is the median "span"?

4. What is the median number of gems that each student removed from the original bowl?

5. What is the mode for the birthdays?

6. What probably has a larger effect on the number of gems that are removed:

   the number of centimeters in a person's hand "span"
   OR
   the number of the day on which the person was born?

7. Make a scatter plot of "Gems versus Span". In other words, use the "Span" data for the x-axis and the "Gems" data for the y-axis. Predict the correlation coefficient and give reasons for your prediction. Then determine the correlation coefficient.

   Your Prediction: _____ Correlation Coefficient: _____

8. Make a scatter plot of "Gems versus Day of Birth". (Use the "Day of Birth" data for the x-axis and the "Gems" data for the y-axis.) Predict the correlation coefficient and give reasons for your prediction. Then use the TI-83 Plus to determine the correlation coefficient.

   Your Prediction: _____ Correlation Coefficient: _____

9. Write your reaction(s) to the results of this experiment. Were the results what you expected? If they were not, then give possible reasons for the discrepancy.

10. Repeat Steps 7 through 9 above for the data obtained from faculty members, as listed in the chart below.

GEM EXPERIMENT DATA

Faculty

SPAN (in centimeters)	Number of GEMS	Number of DAY of BIRTH
20	51	27
22	45	23
24	60	19
23	76	29
20	36	21
20.5	39	29
22.5	49	24
22	45	21
22	50	17
20	43	1
23	70	6
22	59	5
19.7	40	28

Correlation Coefficient for "Gems versus Span": _____

Correlation Coefficient for "Gems versus Day of Birth": _____

ANSWERS

Home-Run Leaders of the American League

1. 40.18 2. 40 3. 39 and 40 4. 29

7. 29 8. 39 and 40 9. 29

Age and Height (Children in Kalama)

1. Chart: 1.1, 0.1, 0.6, 0.9, 0.2, 1.2, 0.1, 0.6, 1.0, 0.7

2. 0.67 cm 8. D 9. 83.9 cm

Correlation

1. 0.98 2. –0.96 3. 0.24 4. 0.99

Gem Experiment

1. 21.5 cm 2. 45.3 or 45 gems 3. 21.85 cm

4. 45.5 or 46 gems 5. 18 6. no. of cm in the "span"

7. 0.28 8. –0.11 10. 0.74 and –0.19

ACTIVITY SOURCES

"The Data and Story Library". http://lib.stat.smu.edu/DASL

This online library consists of data files and stories that illustrate the use of basic statistics methods. A wide variety of topics help statistics teachers find real-world examples that will be interesting to their students.

"Key Crime & Justice Facts at a Glance". Bureau of Justice Statistics. U. S. Department of Justice. Page last revised on 9 Sept. 2002. http://www.ojp.usdoj.gov/bjs/glance.htm

Small versions of charts and brief statements of findings are presented with links to full size charts, additional information about the charts and findings, and the data that support the chart. Full size versions of selected trend charts that are suitable for overheads or handouts are also available.

"Correlation Coefficient".
http://noppa5.pc.helsinki.fi/koe/corr/cor7.html

A scatter plot is shown with a line of best fit and a correlation scale. As the student moves the cursor up and down the correlation scale from –1 to 1, the scatter plot changes accordingly.

"Guessing Correlations".
http://www.stat.uiuc.edu/~stat100/java/guess/GCApplet.html

Four scatter plots and four correlations appear each time the "New Plots" button is pressed. The student must match the plots with the correct correlation.

INSTRUCTIONAL GOALS AND OBJECTIVES

WV IGO AM1.11

Collect, organize, and interpret data using graphs, charts, and tables.

WV IGO AM1.15

Use appropriate software to practice and master Applied Mathematics I instructional objectives.

WV IGO AM1.17

Use graphic software to create graphs, charts, and tables from given data.

WV IGO AM2.10

Collect, organize, interpret data, and predict outcomes using the mean, mode, median, range, and standard deviation.

WV IGO AM2.14

Use appropriate software to practice and master Applied Mathematics II instructional objectives.

WV IGO AM2.18

Use graphing software to create graphs, charts, histograms, and tables of given data.

WV IGO AM2.19

Use graphing software to create process charts and histograms, run charts, scatter diagrams, and distribution curves.

WV IGO AI.20

Use appropriate software to practice and master Algebra I instructional objectives.

WV IGO CM.12

Create and interpret data using various methods of displaying numerical data, including frequency distributions, graphs, histograms, stem-and-leaf plots, and box-and-whiskers plots, using technology when appropriate.

WV IGO CM.18

Use appropriate software to practice and master Conceptual Mathematics instructional objectives.

WV IGO CM.19

Use a graphing calculator to graph linear equations.

WV IGO CM.20

Use graphing software to create graphs, charts, histograms, and tables of given data.

WV IGO A2.20

Use appropriate software to practice and master Algebra II instructional objectives.

WV IGO A2.22

Use a graphing calculator to graph linear equations.