The
Statistics
Teacher
Network

www.bio.ri.ccf.org/ASA/stn.html

Number 56 ASA/NCTM Joint Committee on the Curriculum in Statistics and Probability Winter 2001

Software Review...

Fathom Dynamic Statistics™ Software

http://www.keypress.com/fathom/

Given the limited budget constraints under which most math departments operate, choosing a software package is no easy task. I was faced with that decision last year when looking for software to use in my AP Statistics class. To justify spending the money on a site license for a software package, some of the questions I needed to answer were:

1. What will my students learn from using this program that they wouldn't be able to learn without it?
2. How easy is the software to use? How much time will be needed to learn the software before my students would be able to learn from the software?
3. What does this program do that can't be done with technology already available, such as the graphing calculator?
4. Can the software be used in more than one course?

Also In This Issue...

When Key Curriculum Press announced ***Fathom***, a dynamic statistics software package, I couldn't wait to try it. I am a long-time user of *The Geometer's Sketchpad*, another software program from Key, that revolutionized the way students learn geometry. So, needless to say, I had high hopes for *Fathom*. I am happy to report that it not only met my expectations, but exceeded them.

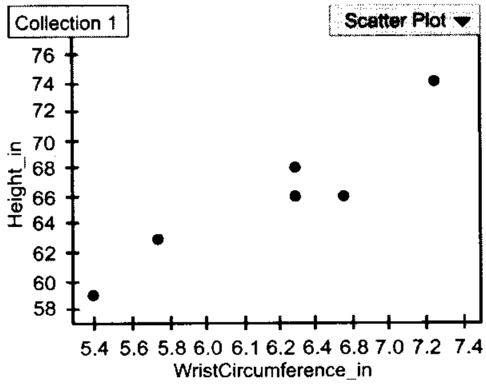
Fathom is a general-purpose tool for doing mathematics and mathematical modeling. The dynamic nature of the software allows the user to ask lots of "what if...?" questions, to make lots of good conjectures, and then to easily test and to see their conjectures in action. The following description comes from the *Fathom* website.

"*Fathom* does for data-driven mathematics and statistics what The Geometer's Sketchpad did for a geometric approach to mathematical modeling. With Fathom you can explore data, plot functions, and create animated simulations."

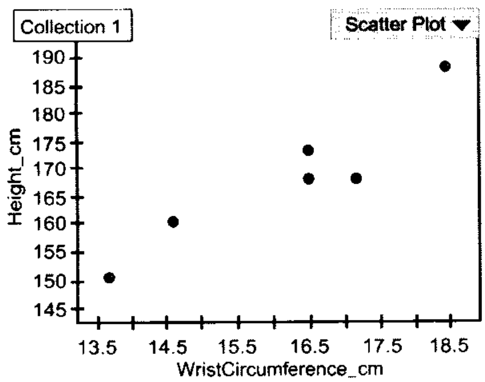
This is a very apt description. My students experienced the same "Aha!" while using *Fathom* that they had experienced in previous classes using Sketchpad. *Fathom's* dynamic capability allows students to manipulate data and to see the changes instantaneously. As students interact with the data, statistical concepts become clear. *Fathom* differs from traditional statistical software packages such as Minitab, in that it is designed to help students "learn" statistics versus "do" statistics. *Fathom* also has a complete statistical software package, so additional software is not needed to do statistical analysis.

To give you a better sense of *Fathom's* capabilities, I have included some examples of lessons I have done with my class, in addition to reviews from two other AP Statistics teachers, Susan Peters and Kristen Klegg.

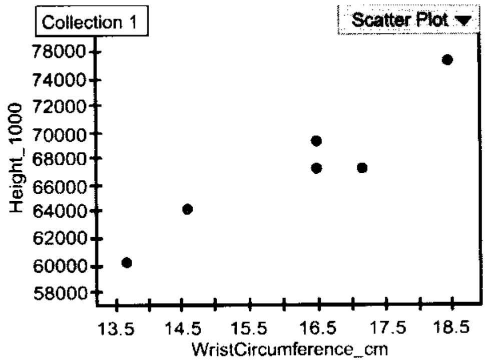
Below is data I used with my class on using a person's wrist circumference to predict his/her height. We were exploring the properties of the correlation coefficient, r. The original data was measured in inches. Below is a graph of the scatterplot of Height on Wrist Circumference.



We discussed how changing the units from inches to centimeters would affect the correlation. The class agreed that the correlation would remain the same. We defined two new attributes in *Fathom*, Wristcircumference_cm and Height_cm and plotted the new graph next to the original one.

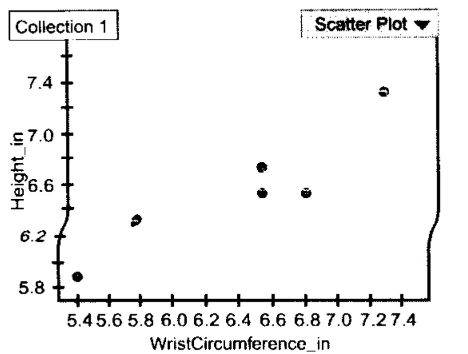


One of my students, Matt, made the comment that "r" didn't change because we multiplied both the height and the wrist circumference by the same number when we changed units. Matt conjectured that changing only one of the attributes would affect the correlation. I told him that I would plot Height in inches v. Wrist circumference in centimeters to explore his conjecture. Well, Matt's a bit of a skeptic, and he figured the only reason I offered to make this revised plot was because I knew it would work. He suggested instead that I plot Height in inches *1000 v. Wrist Circumference in centimeters. Matt's plot appears below with the correlation unchanged. Needless to say, after plotting the new graph, I now had Matt's full attention. "How could that be possible?" he wanted to know. This launched a discussion on the use of standardized values to calculate the correlation coefficient.



What did *Fathom* allow us to do that the graphing calculator could not? For one thing, it allowed us to display all three graphs at one time and to easily define new attributes. It also generated good discussion, a key ingredient in learning mathematics. Students could ask, "What if..." questions such as the one posed by Matt and have them answered almost immediately.

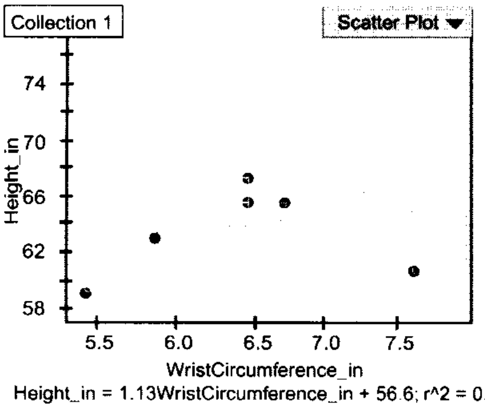
Susan Peters, a teacher at Twin Valley High School in Pennsylvania, made this comment. "I felt the most effective lessons with *Fathom* dealt with regression. Showing the squares with both the movable line and the least squares line really got across the point about what the least squares line really is. That was a lesson that just couldn't be taught on the graphing calculator. I also felt that *Fathom* was particularly effective in looking at outliers and influential observations and their effect on the least squares line."



Height_in = 7.00WristCircumference_in + 21.5; r^2 = 0.89;

In Matt's class, the Least Squares Regression Line (LSRL) was fitted to the Height v. Wrist Circumference data. Students can see the squares of the residuals and can also explore dynamically how changing a data point affects the LSRL. In the graph below, the point in the upper right (7.25, 74) was changed. The resulting graph and LSRL appear below.

Students can see first hand the effect one data value can have on the Least Squares Regression Line. Changing the point (7.25, 74) to (7.58, 60.95) changed the slope of the LSRL

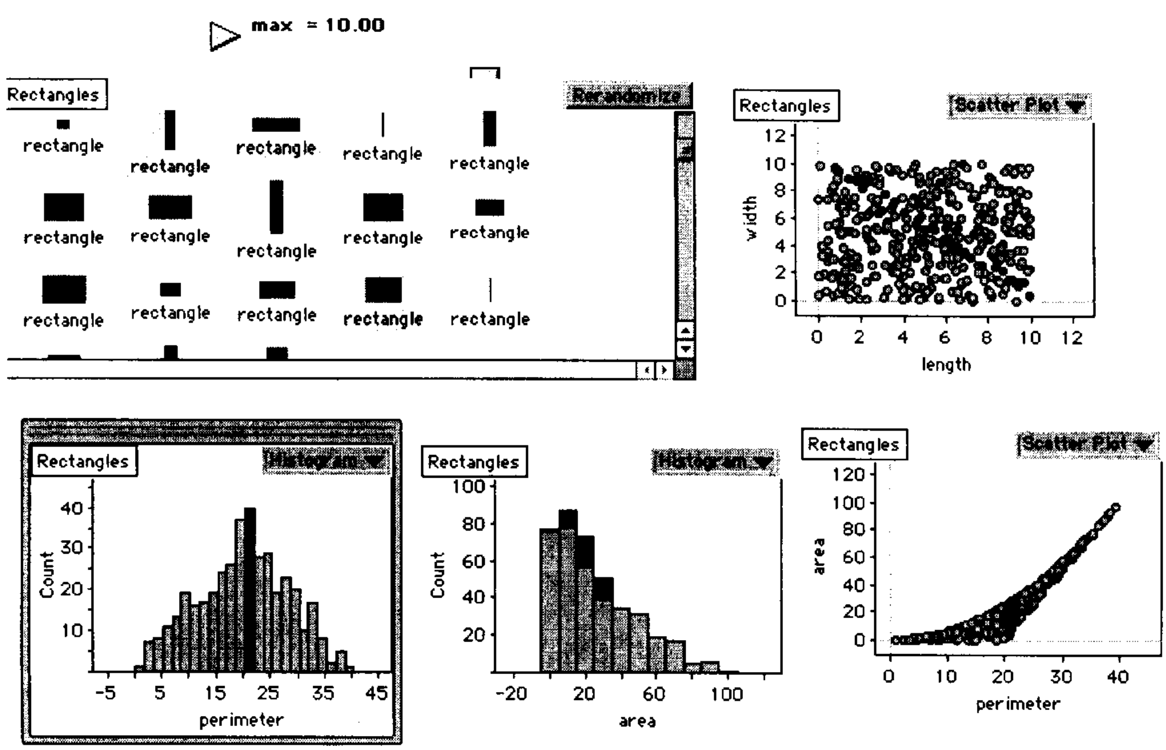


from 7.00 to 1.13 and r^2 from 0.89 to 0.064.

Kristen Klegg, a teacher at Chambourg High School in Chicago, had this to say about *Fathom*. "*Fathom's* ability to adjust statistical measurements, as data points are being dragged around, has been incredible for building my students' intuitive understanding of endless concepts. For example, when we were looking at scatterplots, my students were able to drag influential points and outliers and watch how much and in what ways the least squares regression line changes. Before *Fathom*, I would have the students use the calculator to find regression lines with and without outliers or influential points and compare these lines. This involved a lot of jumping back and forth and we were only able to compare a limited amount of lines on one screen before confusion would set in. For this reason, the students struggled with the idea that the influential points have a more drastic effect on the line than the outliers. But this year they saw it happening right before their eyes and were able to explain why, on their own, by just experimenting with endless changes. This dynamic power brings mathematical concepts to life."

So let's reconsider those original criteria for purchasing a software package. I think the above examples more than sufficiently answer questions one and three on my list. On to question two, "How easy is the software to use? How much time will be needed to learn the software before my students would be able to learn from the software?" *Fathom* is extremely intuitive and easy to use. My students were up and running in the first lab period. In fact, they rarely asked me for help, preferring to figure things out on their own. Sue Peters had this to say, "Overall, I appreciate the visual power of *Fathom*. My hope is that my students will have several of the key concepts of statistics clear in their mind just through their images from *Fathom*. *Fathom* has proven to be relatively easy for the students to use, and I haven't had to review functions used earlier in the year."

This leaves only the last question to be addressed, "Can the software be used in more than one course?" *Fathom* is packed with useful features to support the learning of mathematics from algebra to statistics. I will try to



give you an idea of some of the useful features.

Fathom contains dozens of graph types such as scatter plots, box plots, histograms, percentile plots, ntigrams (an equal-area histogram), bar charts, and ribboncharts, aligned along either axis and split by any categorical attribute. All graphs are dynamically linked, so when you select data in one representation, it is automatically selected in all representations. The representations of the data stay in sync, so when you add or drag a point, everything that depends on it updates.

On the bottom of the previous page is an activity from *Data in Depth*, a book of curriculum materials with over 300 pages of activities spanning mathematics courses from grades 8–12. It is included with the *Fathom Dynamic Statistics*™ software. In the "Area and Perimeter" activity, students use the simulation capabilities of *Fathom* to create rectangles with length and width between 0 and some maximum value (we used 10). The guided activity leads students through conjecturing what the different graphs will look like. In each case, students make a conjecture before using *Fathom* to plot the graph. We graphed scatterplots of length v. width, and perimeter v. area, and histograms of the perimeters of the rectangles and the areas of the rectangles generated by our simulation. The really neat part comes when you look at the relationship between the graphs. We selected the bar in the perimeter histogram that corresponded to Perimeters ≥ 20.32 units and Perimeters < 22 units. The appropriate rectangles were highlighted in each of the graphs. My geometry students were driven to investigate the relationships the saw; no prodding was required on my part.

Fathom is also extremely useful for exploring "Families of Functions." For example, students can explore, via the use of sliders, the effect that changing a, b, c, or d has on the graph of y=asin(bx+c)+d. Below is the graph of y=sin x and y=asin(bx+c)+d, where the values of a, b, c, and d can be changed by dragging the appropriate slider.

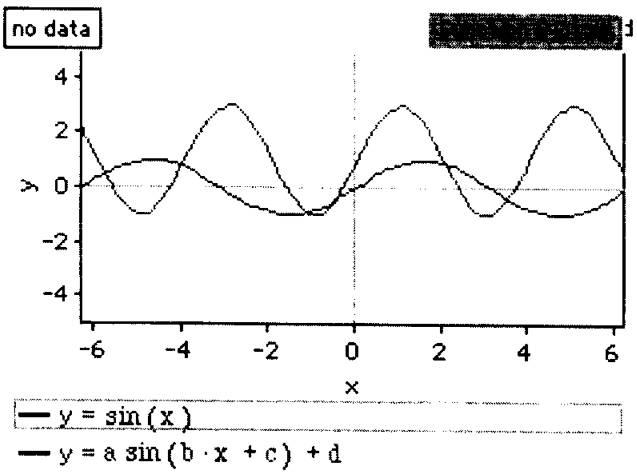
Most teachers have witnessed a heightened interest from their students when dealing with real data sets. But finding data sets and formatting them to use in a software program or with the graphing calculator is often a time consuming task. *Fathom* allows you to drag a URL from a web browser into a blank *Fathom* document and almost instantly, the data is formatted and ready to use. I have used this feature numerous times to obtain data sets for class, tests, worksheets and projects. One of my favorite sites for obtaining data is **DASL (Data and Story Library)**, an online library of data sets and stories listed by method, topic and subject *http://lib.stat.cmu.edu/DASL/*.

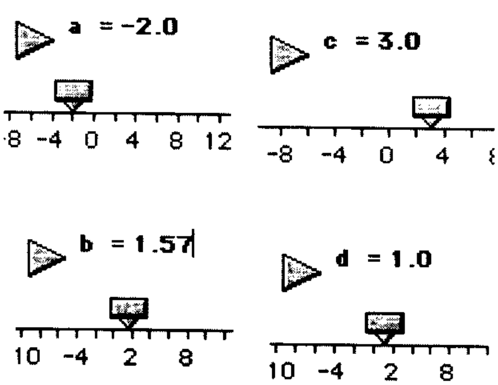
The **Data Analysis and Probability** strand of the *Principles and Standards for School Mathematics* states, "Instructional programs from prekindergarten through grade 12 should enable all students to:

- Formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them
- Select and use appropriate statistical methods to analyze data
- Develop and evaluate inferences and predictions that are based on data
- Understand and apply basic concepts of probability

Fathom is the perfect tool for motivating mathematics through the use of real-world data. It addresses each of the standards above, in addition to many of the standards in algebra, geometry and problem solving. If you only have funds in your budget to purchase one software program, I believe *Fathom* gives you the greatest "bang for your buck."

Ruth Carver
Germantown Academy
Fort Washington, PA
ruth@forum.swarthmore.edu





Statistics in the Classroom

A STATISTICAL ANAlYSIS OF CALIFORNIA'S SuperLOTTO Plus Game

by

Sanderson Smith, Cate School

http://www.cate.org/sms99/hmpsms.htm

sanderson_smith@cate.org

SuperLOTTO Plus was introduced by the California Lottery on June 4, 2000. There are nine ways to win, and the jackpot starts at seven million dollars.

The rules for *SuperLOTTO Plus* are relatively simple. There are two number sets:

Set A: The integers from 1 to 47 inclusive.
Set B: The integers from 1 to 27 inclusive.

A player spends one dollar for the privilege of selecting five numbers from Set A, and one number, called the MEGA number, from Set B.

Event	Symbolic Probability	Numerical Probability	Odds: 1 in ---
Match FIVE and MEGA	(5C5)(42C0)(1/27) / 47C5	2.41451E-08	41,416,353
Match FIVE and no MEGA	(5C5)(42C0)(26/27) / 47C5	6.27771E-07	1,592,937
Match FOUR and MEGA	(5C4)(42C1)(1/27) / 47C5	5.07046E-06	197,221
Match FOUR and no MEGA	(5C4)(42C1)(26/27) / 47C5	.000131832	7,585
Match THREE and MEGA	(5C3)(42C2)(1/27) / 47C5	.000207889	4,810
Match THREE and no MEGA	(5C3)(42C2)(26/27) / 47C5	.005405111	185
Match TWO and MEGA	(5C2)(42C3)(1/27) / 47C5	.002771852	361
Match ONE and MEGA	(5C1)(42C4)(1/27) / 47C5	.013512778	74
Match NONE and MEGA	(5C0)(42C5)(1/27) / 47C5	.020539423	49
	Total Probability of Winning Something =	.042574608	23

All prize payout amounts are pari-mutual. That means that the amount of money varies from week to week depending upon the sales levels and the number of winners. The jackpot prize, obtained when one matches all five numbers and the MEGA number, has a stated minimum value of seven million dollars.

There are nine ways to win in *SuperLOTTO Plus*. Probabilities and odds associated with the game are displayed in the table below. Notes and comments related to the table:

- When playing, two events are involved. One involves choosing five numbers from Set A, and the second consists of picking one number from Set B. The two events are statistically independent, so related probabilities can be multiplied, as is done in the table.

- The probability that a player matches the MEGA number is 1/27, and the probability that the MEGA number is not matched is 26/27.

- The symbolism nCr represents the number of different sets of r objects that can be selected from a set of n objects. nCr = (n)(n-1)...(n-r+1)/r!. For instance, consider the set {red, white, blue, green}. If two colors are to be chosen (without replacement and without regard to order), there are 4C2 = 6 ways that this can be done. The six possible sets are {red, white}, {red, blue}, {red, green}, {white, blue}, {white, green}, and {blue, green}. Modern calculators (the TI=83, for instance) have menus that include nCr.

- A number such as 2.41451E-08 is 2.41451 times 10 to the -8 power, or .000000024145.

To verify the "odds" column, as printed by the state of California, it is often helpful to represent very small probabilities in the form 1/x using the algebraic identity x = 1/(1/x). For instance, the probability of hitting the jackpot is .0000000241451. Using the algebraic identity, this can be written as .000000024145 = 1/(1/.000000024145) = 1/41,416,353. In other words, the **odds** of winning the jackpot are 1 **in** 41,416,353.

As the table indicates, the probability of hitting the jackpot is quite small. To use a time perspective, suppose that one could play the game once every second. In order to expect to hit the jackpot once, it would take over 15 months! Moreover, over 41 million dollars would have been spent in order to get a return that could be as small as 7 million dollars.

But then, while they may be few and far between, there are those who spend only $1 and end up as big winners.

A note for the Purist: Technically, the odds of winning E are a **to** b if P(E) = a/(a+b). So the odds of winning the jackpot are 1 **to** 41,416,352.

"Oh, many a shaft at random sent
Finds mark the archer little meant."
Sir Walter Scott: The Lord of the Isles

Using Technology to Reinforce Data Display and Analysis

by Vicky Jackson and J.D. Smith
Rolla Middle School Seventh Grade Mathematics Teachers

In a seven-day unit introducing the field of statistics, seventh grade middle school students research peer opinions to generate, organize, analyze and display data. Student learning is enhanced through the implementation of technology to achieve curriculum goals relating to research, statistics, and data analysis. Prerequisite skills include measuring angles using a protractor, determining the scales and intervals to use to construct a number line graph of a given set of data, knowing the meaning of percent, and being able to find equivalent fractions, decimals, and percents.

On **Day One**, students begin by discussing surveys and polls--their purpose and design, as well as bias, including information about marketing strategies and the psychological impact of manipulation by the use of vocabulary and the wording of a question. Each student has five to ten minutes to make up three questions that might be used to survey seventh graders. The importance of asking questions that are interesting and relevant to seventh grade students is discussed. Each question should offer at least three (but not more than eight) choices of answers.

Students are assigned to five different groups to discuss their questions. The groups have ten minutes to read their three questions to each other and discuss their ideas. Each group then submits one survey question to the class by sending one student from their group to the board to write the survey question, including the answer choices. The class discusses each question's merits, pointing out strengths and weaknesses in design and making any necessary improvements. The class then votes on one survey question to be included on a set of survey questions to administer to the entire seventh grade.

Day Two begins with a caution for students not to write their names on the surveys. After each student takes the survey, the surveys are collected and redistributed so that students do not receive their own surveys. This eliminates/minimizes the peer pressure to vote a certain way. Students make a frequency table for each survey question. The class counts the results by having students raise their hands or stand up to show the responses on the paper they are holding (not their own paper). Results are then recorded on frequency tables.

On **Day Three**, students record the totals of all results in frequency tables and discuss the results including: How were the results affected by the way the questions were asked? Were the questions designed well? How could they have been improved? Was the way each question asked unbiased?

Day Four begins with students making bar graphs. They are provided graph paper. Students review the required parts of a bar graph and how to construct a bar graph. Using the frequency table of the totals for survey question #1, students make a bar graph together in each class period. They then construct bar graphs for each of the rest of the frequency tables (totals).

Circle graphs are made on **Day Five**. Compasses, protractors and calculators are needed for this activity. Students use the data from the results of question #1 and the class constructs a circle graph together using conversions from fractions to decimals to percents and finding a percent of the whole (360 degrees). They construct two more circle graphs from two of the other survey questions of their choice, using the data in their frequency tables (totals).

On **Day Six**, students go to the computer lab and use Microsoft Excel and Word to make tables and graphs. They are guided through a handout of step-by-step instructions to use Microsoft Excel and Microsoft Word to produce a frequency table of the survey data, using the summation function to find the total of the data, and produce a bar and a circle graph of that data. Students compare the results to their hand drawn graphs from **Day Four** and **Day Five**.

A paragraph is written on **Day Seven** describing the "typical" seventh grader at the Rolla Middle School based upon the survey results. This activity supports writing across the curriculum. All frequency tables, bar graphs, and circle graphs are turned in at this time.

The Assessment Rubric is Class Participation in Discussions = 20 points, Survey Questions = 10 points, Frequency Tables = 20 points, Bar Graphs = 20 points, Circle Graphs = 20 points, Computer Pie Graph = 10 points. Up to 5 bonus points may be earned by creating a picture characterizing the surveyed students based upon the data collected in the survey. Students may draw a picture, generate a computer picture, cut and paste magazine pictures onto poster board, or in some other creative way portray that seventh grader visually.

Accommodations for special needs students include a reduction in the number of tables and graphs required (half). Students work with a buddy student. Students are graded on the guided practice, group work and resulting "products" rather than having a unit test over the concepts covered in this unit.

Follow-up activities include the use of graphing calculators to investigate the analysis of data. After students study the mean, median, and mode of sets of data and learn how to make a box-and-whisker plot, the students use graphing calculators to study these concepts. A classroom set of graphing calculators and an overhead LCD panel with calculator are needed for this activity. Students learn the basic functions of the graphing calculator through guided exploration. They then learn how to enter numerical data into a list, and calculate the mean, median, and mode of that data as well as Students learn to make box and whisker plots on the graphing calculator.

Handout: Making Bar and Circle Graphs Using the Computer

Microsoft Excel

1. Go to *Start/Programs/Office2000/Microsoft Excel.* Click Microsoft Excel.
2. Click in cell A1 and type the words Writing Utensil.
3. Click on the bar between the A and the B heading and drag the vertical bar to the right.
4. Click in cell B1 and type the word Total.
5. Type the writing utensils in the Writing Utensil column, using the down arrow to move to the next choice.
6. Type the total numbers in the total column using the down arrow.
7. Click the word *Total* (in cell B1) then go to the menu bar and click *Insert.* Click *Chart.*
8. See what each of the styles of charts look like, especially the Column, Bar, and Pie types.
9. Select the *Pie Graph,* then click *Next,* and then *Next* again.
10. In the Chart title box, name the chart Writing Utensils.
11. Click on the *Legend* tab and be sure the "Show legend" is checked. Then choose the placement.
12. Click on the *Data Labels* tab, and select "Show label and percent," and click on "Show leader lines."
13. Click "*Next.*"
14. Click "*Place chart*" and "*As new sheet.*"
15. Click *Finish.*
16. Click *Insert/Picture/Word Art.* Type your name. Adjust the size on the page. Move it to a better location.
17. Click *Print.* Click *OK.*

Microsoft Word

1. Go to *Start/Programs/Office2000/Microsoft Word.* Click Microsoft Word.
2. Click *Insert/Picture/Chart.*
3. Type the information into the chart cells as you did above in the Excel program.
4. Click *Chart/Chart Options.* Type a title. Label the axes (x and z).
5. Click *Chart/Chart Type.* Explore these categories.