The Statistics Teacher Network

Number 47 ASA/NCTM Joint Committee on the Curriculum in Statistics and Probability Winter 1998

Statistics in the Classroom ...

Investigating Centers— "Thinking Art"

Henry Kranendonk
Rufus King High School
Milwaukee, Wisconsin

I started my geometry class last semester by asking students to define geometry based on their previous work with this subject. Most students, as expected, included references to areas and perimeters, circles and triangles, or other traditional topics. One young man, however, defined geometry as "thinking art." This definition struck me as particularly intriguing. Beyond the formulas and structure, geometry often investigates the characteristics of objects, shapes, and ideas that might be more appropriately described as art. I used his definition to introduce students to an investigation of *centers.*

The investigations of centers were developed by a team of statisticians and secondary teachers in a project entitled "Data-Driven Mathematics." The theme suggested by our project's title was not only to teach statistics, but to make connections to all aspects of secondary mathematics using a data related curriculum. Jeffrey Witmer (a statistician from Oberlin College) and I specifically developed the lessons involving connections to geometry. We entitled our work "Exploring Centers."

Raisin Geometry: Exploring the Balance Point

Several lessons in the first part of the investigation involve manipulating cut out geometric shapes from poster paper and raisins. Included in the cut outs are various triangles, hexagons, concave and convex quadrilaterals, and a convex pentagon. The students are given a small box of raisins and instructed to select several raisins of similar size. A raisin is taped to the vertex of each shape. The resulting model is then balanced on the eraser end of a new pencil. The approximate location of a balance point **B** is marked on the cut out shape. This point is discussed as a center representing a *balance* of the distances *and* weights of the raisins. The cut out is then placed on a coordinate grid. The coordinate values for each vertex and the observed balance point **B** are recorded on a data table as illustrated below.

Point	x-value	y-value
P_1	7	5
P_2	-6	5
P_3	-2	-4
Balance point B	-0.5	2

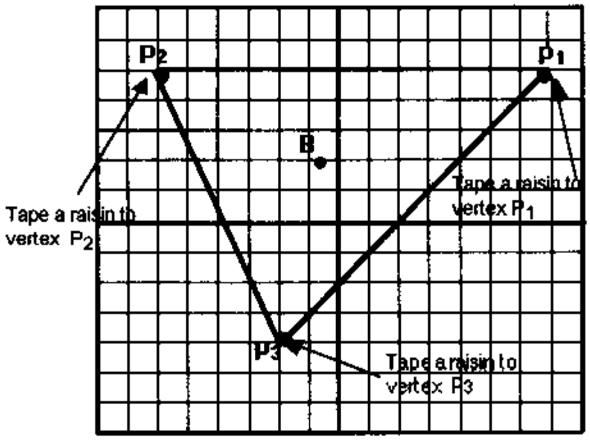
The centroid location, *C*, is determined as the means of the *x* and *y* values of the vertices.

$$\bar{x} = \frac{(x_1 + x_2 + x_3)}{3} = \frac{(7 + -6 + -2)}{3} = -.333$$

$$\bar{y} = \frac{(y_1 + y_2 + y_3)}{3} = \frac{(5 + 5 + -4)}{3} = 2$$

The coordinate location of the resulting centroid **C** is compared to the location of the balance point **B**. Students discover the two locations are very close. The centroid is viewed as an "artistic" average of the balance point.

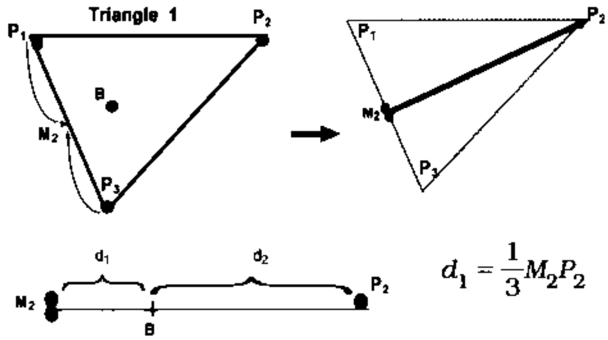
Also In This Issue...



A process called "collapsing the raisins" was also developed to reinforce several special features of the cut out geometric shapes. The intent of this process was to relate the point-mass activities to a student's previous work in geometry. Basically this phrase describes a process in which the raisins are "collapsed" to a new point. The new shape outlined by including this collapsed point is "balanced" by the same point as the original arrangement of the raisins. To illustrate this process, consider *Triangle 1* with one raisin taped to each of the vertices. The raisins at point $\mathbf{P_1}$ and point $\mathbf{P_3}$ are combined ("squished" on top of each other) at the midpoint of the segment $\overline{\mathbf{P_1P_3}}$. Label this new point as $\mathbf{M_2}$. The resulting arrangement of the 3 raisins (or, the triangle *collapsing to* a line segment) contains the "balance point" of the original arrangement of the 3 raisins. Similarly the 2 raisins at $\mathbf{M_2}$ and the one raisin at $\mathbf{P_2}$ can be combined (squished together) at the point one-third of the distance along $\overline{\mathbf{M_2P_2}}$. (Why one-third?) The balance of the new arrangement (a line segment *collapsing to* a single point) *is* the balance point **B** of the original arrangement of the raisins.

The process of "collapsing the raisins" demonstrates the intersection of the medians of a triangle meet in one point—*interestingly* this point is also the centroid of the raisins forming the original triangle.

A similar set of collapsing steps guide discoveries of other geometric figures, including the following: the balance point of a parallelogram is the intersection of the diagonals; the balance point of a regular hexagon (or any regular polygon) is the center of the circumscribed circle; and the balance point of a concave quadrilateral might lie outside the interior region (a figure we termed the *boomerang* quadrilateral). Although not a "proof" of any of



$$d_1 = \frac{1}{3} M_2 P_2$$

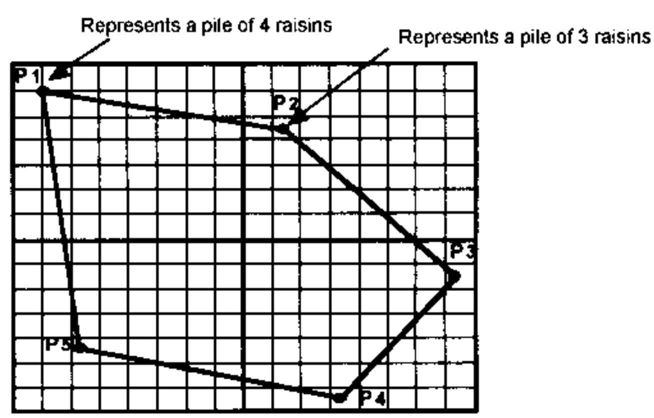
these properties, the activities reinforced the summaries using a concrete, student developed model.

The Weighted Means

One of the investigated shapes, a convex pentagon, is used for a special lesson. In this lesson the following problem is presented to the students:

> Vertices of a pentagon will be weighted by ***an unequal number of raisins.*** Create a total pile of four raisins (one on top of each other) and tape this stack as close as possible to vertex $\mathbf{P_1}$. Similarly, tape a total stack of three raisins at $\mathbf{P_2}$, one raisin at $\mathbf{P_3}$, one raisin at $\mathbf{P_4}$, and finally, one raisin at $\mathbf{P_5}$. Although the vertices trace out a pentagon, this new model will have a different balance point than the model investigated with just one raisin taped at each vertex. Why?

Both the balance point and the centroid of this figure are investigated. The balance point is initially estimated by balancing the model on a pencil acting as a *fulcrum*. The centroid is calculated by placing the model in a coordinate grid as illustrated. The varying raisin counts at each vertex represent a new feature of this lesson.



The *weighted means* of the x and y coordinates describe the calculated centroid of this figure. The centroid is developed as:

$$\bar{x} = \frac{\sum_{i=1}^{5} W_i x_i}{\sum_{i=1}^{5} W_i} \quad \text{and} \quad \bar{y} = \frac{\sum_{i=1}^{5} W_i y_i}{\sum_{i=1}^{5} W_i}$$

Here again, the location of this calculated centroid and the observed balance point are compared. (And again, the locations are very close!) Essentially, the students are guided to discover this center is the balance of the *weights* and distances. What is the significance of this point? Are there any realistic applications demonstrating the importance of this center beyond the geometric models? This is where the following lesson is introduced.

Population Centers

Consider the following problem:

> ***Pretend*** you were an important political person back in the 1980s. You were appointed by the Governor of Wisconsin at that time to head a committee to determine the location of an important job service agency in Wisconsin. Your committee was responsible for helping the people located in the communities of Milwaukee, Waukesha, Port Washington, Belgium, Beloit, and Racine. A map of this particular area of Wisconsin is provided. Also included in this sketch are the 1980 population statistics. The recommendations of your committee were expected to service the people of this area for at least the next 20 years. A review of your decision will be made in 1990 and 2000.

The development in this lesson is directed at the goal of locating a center that balances population and distance. Students develop the specifications for this location around the idea of replacing the number of raisins with the number of people. The weighted mean resulting from this analogous situation describes a "center" balancing the locations and populations of each of the towns and cities.

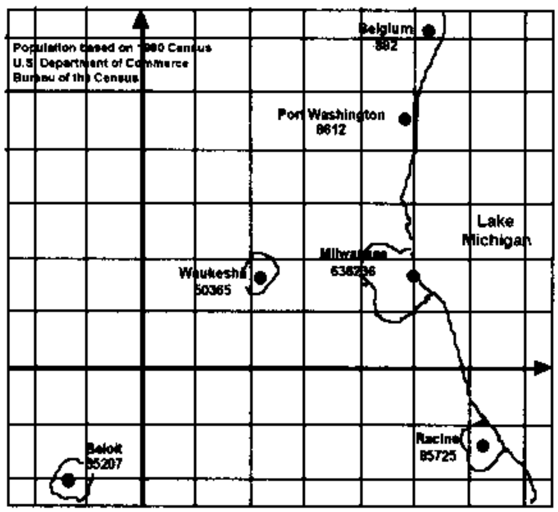
The development of this lesson is aided by a spreadsheet. Clearly estimating a balance point using raisins is no longer possible! (Imagine placing over 600000 raisins on top of each other to represent Milwaukee!) A recent implementation of this lesson was extensively conducted within our computer lab. Calculations were developed with a spreadsheet, locations graphed with a drawing program, and essay questions answered with a word processor.

Belgium Problem!

The spreadsheet implementation allows for students to explore several "what if ...?" questions. One of them explores the location of the population center if Belgium was removed from this problem. This particular question caused concern among the students as they found the location of the new population center did not visibly change if Belgium was removed from the calculation. Is this "fair"? Why did this happen? Although Belgium represents a node in the network of communities, the magnitudes of the city/town populations developed a sense of estimating a location of the population center. Students generally contrasted Belgium's minimal influence to the major shift observed if Milwaukee was removed from the model. The question related to the fairness of the population center generated intriguing responses. Most students felt this method of determining a center was "unfair" to Belgium. As each of the methods identify a similar location for this "center," a summary of the importance of this location is developed.

The Larger Picture

This lesson set-ups a major project of this module—namely, estimating the population center of the United States. This project primarily adds a bigger picture of the same activity suggested with the population lesson. A spreadsheet approach to this problem involves the following: placing an xy-coordinate grid over a map of the 48 states, using the coordinate locations of each of the state capitals as an estimate of a "node" (or vertex) for each state, and using the state's population as the



weight of that state. (How to deal with Alaska and Hawaii is a major new piece in this lesson!) The resulting center (centroid) is very close to the location the U.S. Bureau of the Census published as the population center for the United States using the 1990 census data (i.e., Crawford County, Missouri). Interestingly, this center can also be approximated by simply taping 10 raisins on a cut out poster paper map of the United States to the locations of the state capitals of the 10 most populous states. Balance this map on the top of a pencil and observe the balance point! (Remember the Belgium problem.)

(Editor's Note: Henry submitted this article at the request of the Editor. Four of the ten modules in the Data Driven Mathematics project are currently available through Dale Seymour/Cuisenaire.)

Software Review ...

ActivStats 2.0

Paul F. Velleman (1998)

ISBN 0-201-31068-6 $44.25 (student) ISBN 0-201-31070-8 $90.50 (institution); Addison Wesley Longman; http://hepg.awl.com http://www.datadesk.com

ActivStats is a complete introductory statistics textbook on a CD. It is the most exciting development in statistics education I have ever experienced.

ActivStats has all the features of a traditional statistics textbook: coverage of statistical concepts with text, graphics, and exercises (plus a table of contents, index, and glossary). What's different is that *ActivStats* provides powerful features that are only available on a computer: interactive lessons, videos, and (amazingly) a built-in data analysis package (the award-winning *Data Desk*).

ActivStats emphasizes data visualization and data analysis. Less emphasis is placed on the mathematical mechanics of statistics, which can now be mostly handled by the computer, as in real data analysis. The sequence of topics is as follows: data and variables, univariate distributions, relationships between variables (correlation, regression, surveys, experiments), randomness, probability, random variables, sampling distributions, and statistical inference.

The sequence of topics in *ActivStats* exemplifies the currently most popular sequence for the introductory course. I recommend a substantially different emphasis and sequence of topics for the introductory course [1996], but I shall not discuss that here.

ActivStats consists of twenty-four "lessons," each of which is between two and six "pages" long. Each page in a lesson appears on the computer screen as a scrollable block of text and graphics. Each page gives a high-level summary of a sub-topic of the lesson topic.

Most pages contain between one and five icons, each representing an "activity." When a student clicks on an icon, it expands on the screen into a window containing a presentation. Activities come in four flavors:

1. *Narrated Videos* (average length around two minutes). These describe real-world problems that motivate the statistical topic under discussion on the page.

2. *Narrated Mini-Lectures* (average length around five minutes).

These are on a single statistical topic and are carefully illustrated with innovative moving graphics. Many are punctuated with "hands-on" sessions in which students interact with the computer to gain experience with the topic.

3. *Data Analysis Exercises* (average length around eight minutes). These guide students through the use of *Data Desk* to display or analyze data to reinforce the statistical topic under discussion.

4. *Concept-Review Exercises* (average length under two minutes).

These contain a set of sentences describing the main ideas presented earlier on the page. Students fill blanks in the sentences with supplied words, and the computer indicates the correctness of the answers. These help students to assess their understanding of the ideas.

Each lesson is accompanied by a set of homework questions (roughly nine per lesson), many of which are intended to be answered using *Data Desk* (with supplied data). Each lesson also has a set of projects (roughly two per lesson), which are similar to the homework questions except that students must actually collect data as part of each project.

Each *ActivStats* lesson has cross-references to eleven currently popular introductory statistics textbooks. Thus a teacher can teach a course with *ActivStats* in conjunction with a standard introductory text. One of these texts is by David Moore, a leading statistics educator and 1998 president of the American Statistical Association. His book, *The Active Practice of Statistics* (1997), is specifically tailored to be a companion textbook to *ActivStats* and employs the same philosophy and contains the same twenty-four lesson topics as *ActivStats*.

Of special note are the innovative "visualization tools" that accompany many of the mini-lectures. Each tool is a graphical display of data (e.g., a bar chart, scatterplot, or density plot) together with one or more ways the student can easily change properties of the data or display. (Many tools have "sliders" that allow the student to change the value of a property through a continuous range of values.) When a student changes the value of a property, the display immediately and (if necessary) continuously updates to reflect the new value(s) of the property. These fascinating tools are revealing windows for students into both data and statistical concepts. A controllable moving picture is worth ten thousand words.

Also of note is the tight integration between *ActivStats* and *Data Desk*. If a student clicks on an activity or exercise that uses *Data Desk*, the *Data Desk* program is automatically launched and loaded with (a) the step-by-step instructions for the student to follow to do the work and (b) the dataset to be analyzed.

ActivStats contains several other nice touches, including careful discussion of many stumbling blocks of statistics and superb graphics. These features all reflect the great care and tremendous effort Paul Velleman and his development team have put into this remarkable computer-based statistics "textbook."

ActivStats is the most notable of several new products that reflect a major turning point in statistics education: Students now have the opportunity to *actively* learn statistics, using their media of most frequent choice—computers and television. Students who actively participate in their learning learn substantially more than students who passively attend lectures or engage in passive reading of a textbook (Garfield 1995).

Velleman has produced a first-rate computer-based statistics "textbook"—a significant contribution. However, he has also made a much more important contribution: he has given us a rich new paradigm for teaching statistics.

An important feature of the paradigm is that it is extendible—teachers can add their own material to *ActivStats*. If Velleman provides good standards, openness, and easy-to-use procedures, it is likely that many statistics teachers will develop activities that run in the *ActivStats* environment for teaching many statistical topics, possibly at all levels. This will enable a sharing and natural selection of activities by statistics teachers, which will help us to zero in on optimal approaches to teaching statistical concepts.

Because products like *ActivStats* deliver substantially more functionality than standard (paper-based) textbooks, it is clear that these products will, in time, completely replace standard textbooks. However, the following problems must be solved before computer-based textbooks will be the norm:

1. The low "pixel count" of current computer screens detracts from computer-based statistics textbooks. A computer screen is a dense matrix of tiny dots called "pixels" (picture elements). The computer generates images on the screen by controlling the color and brightness of each pixel. A typical screen may be 800 pixels wide and 600 pixels high, giving roughly a half-million pixels on the screen. The printed page in a typical standard textbook (ignoring margins) is equivalent to 5400 x 9300 pixels, giving roughly fifty million pixels on the page—a one-hundred-fold increase over a typical computer screen. The pixel count of the printed page of current textbooks evolved over the centuries as optimal for displaying "printed" information. Thus, as many readers will know from experience, the printed page is substantially more effective than the computer screen for displaying (static) information. (*ActivStats* makes the best of this unfortunate situation by generally using only small blocks of static text, and by using motion to help focus attention.)

2. Some students may be tempted to "surf" computer-based textbooks, jumping somewhat randomly from topic to topic, like surfing on the web. This approach leads to minimal learning since the later topics in a statistics course invariably require knowledge of earlier topics. Similarly, there is a sense among many students that when using a computer one must move quickly, without pausing to reflect. But statistics students must learn to pause and reflect on what they are learning, and sometimes even backtrack for review. Thus the challenge in designing a computer-based textbook is to keep students focused on the appropriate material until it is properly learned. *ActivStats* uses several techniques to help students stay focused, but more techniques may be needed.

3. Computer-based textbooks can only be used in courses if students have adequate access to computers. For *ActivStats*, I suspect the average college or university student will need five or more hours of readily available access in some weeks.

As a practicing statistician, I am attracted to careful empirical research. Therefore, I look forward to research by Velleman and others as they compare courses using *ActivStats* with other introductory statistics courses. I believe this research will demonstrate unequivocally that carefully designed "multimedia" teaching (supplemented with some lecturing) is superior to other approaches in two critical areas: (1) careful multimedia teaching gives students substantially better understanding of the use of statistics in empirical (scientific) research, and (2) careful multimedia teaching gives students substantially more respect for our field.

With this last thought, I heartily recommend that all introductory statistics teachers evaluate this extraordinary teaching tool with a view to using it in their courses.

References

Garfield, J. (1995). "How Students Learn Statistics," *International Statistical Review*, 63, 25-34.

Macnaughton, D. B. (1996) "The Introductory Statistics Course: A New Approach." Available at http://www.matstat.com/teach

Moore, D. S. (1997), *The Active Practice of Statistics*, New York: Freeman.

Technical Information

ActivStats runs on Windows and Macintosh computers. Minimum requirements under Windows are Windows 95, 486/66, 12 Mb RAM, 640 x 480 x 256 color monitor, double-speed CD-ROM drive, and an 8- bit sound card that is QuickTime compatible. Minimum requirements on the Mac are System 7.5, 68040/25 or Power Macintosh, 8 Mb free RAM, 640 x 480 x 256 color monitor, and a double-speed CD-ROM drive.

Reviewed by
Donald Macnaughton
MatStat Research Consulting Inc.
donmac@matstat.com

The PRE-STAT project is seeking participants for Summer '98 Faculty Development Workshops on statistics education to be held at Appalachian State University July 19-25 and Montana State University July 26-August 2. Applicants should be college faculty who are involved with the preservice education of middle or secondary school mathematics teachers. PRE-STAT is supported by NSF's Undergraduate Faculty Enhancement program. You will find further information at WWW.PRESTAT.APPSTATE.EDU . Or contact Mike Perry, project director, Dept. Math. Sciences, Appalachian State University, Boone, NC 28608, tel (704) 262-2362, E-mail PERRYLM@APPSTATE.EDU.

The 1998 American Statistics Poster & Project Competitions

Recognizing that there is a need to encourage and promote the understanding and application of statistics, the American Statistical Association, through its Center for Statistical Education, is continuing to offer two annual statistical competitions. Both competitions are supervised by the NCTM/ASA Joint Committee on Curriculum in Probability and Statistics, Grades K-12. The project competition was started in 1987, the poster in 1990. Both encourage students to experience the scientific method in formulating and solving problems and to discover the usefulness of statistics in the analysis process. Data for the question(s) of investigation could be collected through survey or an experiment.

The Project Competition:

There are three categories in the project competition: grades 4-6, 7-9, and 10-12. Entries must be from teams of two to six students. Projects from a single individual are not accepted.

A statistical project is a written report that describes the process, and presents the results, of selecting a question for investigation, collecting and analyzing the data, interpreting the results to answer the question, and reflecting on the study. The methods of analysis depend on the level of the students. In the 4-6 grade category, graphs and some summary statistics such as means, medians, and modes, are often presented. Parallel boxplots may be used to compare two groups. In the grades 7-9 group, formal tests of inference are possibly seen, but there should be evidence of more attention paid to, and a higher understanding of, the scientific method, as well as more creative and interesting questions to investigate. The 10-12 level should reflect much more involved questions with inferential procedures used in the analysis.

There are six components emphasized in judging: question of interest, research design and data collection, analysis of data, conclusions, reflection on the process, and final presentation. The most critical aspects of judging focus on whether the data have been collected in a manner that will permit the question to be answered, whether the analysis is appropriate for the design, and whether the conclusions are consistent with the analysis and can answer the question.

The current chair of the project competition committee is Linda Young (biom025 @unlvm.unl.edu).

The Poster Competition:

There are four categories in the poster competition: grades K-3, 4-6, 7-9, and 10-12. Entries may be from individuals or from teams of up to four students. Note that as activities are often done at the primary level in large groups, in the grades K-3 category only, a team may consist of the entire class.

A statistical poster is a visual display containing two or more related graphics (plots, charts, maps, etc.) that summarize a set of data, that look at the data from different points of view, and that answer some specific questions about the data. Statistical graphics are exceptionally powerful tools for data analysis and presentation.

Entries are judged on the basis of the overall impact of the display, the clarity of the message, the appropriateness of the graphics for the data, and the creativity shown in the individual graphics. For the primary grades it is helpful to think that the purpose of a statistical poster is to visually tell a story, from the data, about some phenomena, revealing to the viewer the conclusions that can be drawn. For the higher grades it is reasonable to expect that the poster be used not only for data presentation, but also as a graphical problem-solving tool. Indeed, at its best, the poster should demonstrate that the scientific method of problem solving has been used.

A poster has one serious drawback - it must stand alone. There is neither narrator to tell the story nor an accompanying report to fill in the details. Thus the overall message must be clear. Not only must the viewers be able to understand the individual graphics; they must be able to see the relationships between the graphics and how the graphics address the question being studied.

The current chair of the poster competition committee is Linda Quinn (lmq2@po. cwru.edu).

Structure of the Competitions:

There are no entry fees for either competition, and there is no restriction on the number of entries per school. In both competitions, with increased support from several ASA Sections, this year's prize structure has been

Conference Announcements...

Here are two opportunities for new and experienced AP Statistics teachers who are attending the NCTM Convention in Washington DC on April 3.

CWAC: Conference-Within-A-Conference (requires ticket) Experiencing AP Statistics: Designing and Implementing a Course Involving Activities-Based Components

Friday April 3, 8:30-4:30 Session 455
Salon A (Marriott).

Presenters:

James Bohan, Manheim Township School District, Lancaster, PA Al Coons, Buckingham Browne & Nichols School, Cambridge, MA Cathy Lincoln, Groton School, Groton, MA Paul Myers, Woodward Academy, College Park, GA Allan Rossman, Dickinson College, Carlisle, PA Dan Yates, Electronic Classroom, Richmond, VA

Statistics comes alive through the study of real data. This conference focuses on the needs of first time AP Statistics teachers and those interested in activities-based approaches. By including activity-based exercises, teachers will deal with technology, the need to approach statistics at a high conceptual level, and open-ended assessment. Participants in this pragmatic workshop will experience different types of individual and group activities. Curriculum design, texts and resources, project implementation, and assessment will be discussed. Appropriate uses of technologies including calculators, and computers.

Third Annual AP Statistics Teachers Meeting

Friday April 3, 1998, Tentative time: 4:30 PM to 6:30 PM. When time and location are confirmed, they will be posted at www.bbns.org/us/ap_stats . Flyers about the meeting will also be available at related vendors booths at the NCTM.

This two hour informal meetings bring together experienced and new AP Statistics teachers, authors, publishers, and technology vendors. The main focus of this meeting is to share information. Some vendors and publishers will have materials available to review. Appetizers will be provided. The discussion opens with the gathering of some statistics from the AP stats audience about what books and technologies are being used, student populations, etc. Open discussion follows by questions raised by those in attendance.

Al Coons, the organizer of each the sessions, can be contacted at alcoons@aol.com and there is more information at www.bbns.org/us/math/ap_stats .

improved to three prizes with awards up to $200, $100, and $50 in each category. The winners' schools receive plaques. The winning projects and posters are displayed each year at the annual NCTM meeting. The posters are also shown at the annual professional Joint Statistical Meetings.

The deadline for entries is April 15, 1998. Judges are teachers of statistics and statisticians. Their decisions are final. Submitted entries become the property of ASA and cannot be returned.

For application forms or further information, contact Sue Kulesher by e-mail at the ASA site sue@amstat.org or by phone at 703-684-1221 x150. Rules and entry forms may also be obtained from the ASA web site at www.amstat.org/education/ql-projects.html.

Jerry Moreno
STN Editor

[Editor's Note: The editor thanks Linda Young and Linda Quinn, the current chairs of the project and poster competition committees, resp., and John Schollenberger, past chair of the poster competition for information from which this article was written.]

Keep Us Informed...

The Statistics Teacher Network is a newsletter published three times a year by the American Statistical Association—National Council of Teachers of Mathematics Joint Committee on the Curriculum in Statistics and Probability.

We need your letters, announcements, articles, and information about what is happening in statistics education! Please send hard copy, and, if possible, a disk written in standard ASCII text to the editor:

Jerry Moreno, Dept. of Mathematics, John Carroll University, University Heights, OH 44118

or **moreno@jcvaxa.jcu.edu**

or **fax: (216) 397-3033**

Production Editor: *Derek Lawlor, American Statistical Association*

To be added to the mailing list or make an address change, please send your name and address to: *Statistics Teacher Network,* c/o American Statistical Association, 1429 Duke St., Alexandria, VA 22314-3415; (703) 684-1221; fax: (703) 684-2036; E-mail: chris@amstat.org

Printed in the U.S.A

Statistics Teacher Network
c/o American Statistical Association
1429 Duke Street
Alexandria, VA 22314-3415
USA

Non-Profit
Organization
U.S. Postage
PAID
Permit No. 361
Alexandria, VA

Let us know your Zip +4!