GARY KADER AND MIKE PERRY

# Learning Statistics

IN ITS *CURRICULUM AND EVALUATION STANDARDS for School Mathematics* (1989), the National Council of Teachers of Mathematics recommends that the K–12 mathematics curriculum be broadened and designates statistics as an area deserving increased attention. The standards document promotes the concept that statistics be learned through the study of real problems with real data collected by the students. Rather than focus on developing formulas from which answers are simply computed, teachers should present statistics in a coherent fashion and develop the topic as a whole problem-solving process.

The *Standards* also encourages the use of appropriate technologies for learning mathematics. Appropriate technology allows us not only to expand what mathematics is taught but also to enhance how that mathematics is learned.

## Statistical Foundations

WHAT IS STATISTICS? STATISTICS IS A SCIENCE IN which data are used to answer a question. A statistical investigation begins with a problem, and the study of this problem typically involves the following four components suggested by Graham (1987): (1) pose the question; (2) collect the data; (3) analyze the data; and (4) interpret the results. This model furnishes structure for, and gives direction to, statistical problem solving. To give coherence to the *whole process* employed in statistical work, Perry and Kader (1992) suggest adding "communication of results" as the final product of a statistical investigation. The expanded Graham model is summarized in **figure 1.**

## Technology for Learning Statistics

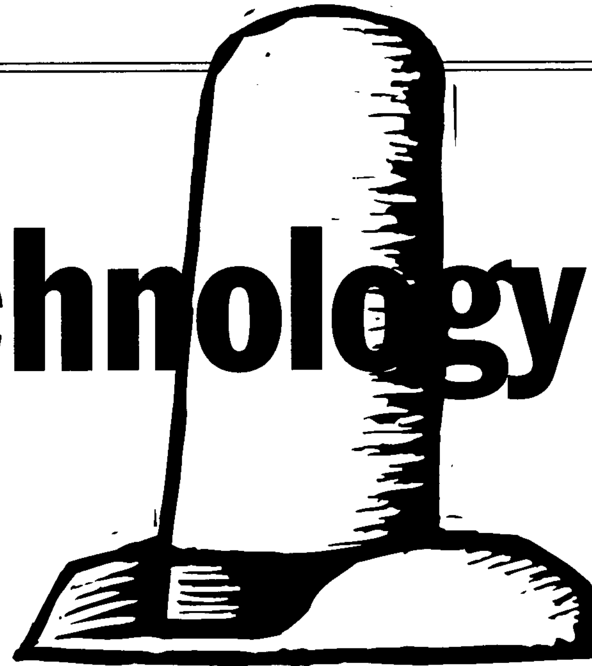
ONE DANGER OF INCLUDING STATISTICS IN THE MATHematics curriculum is that instruction may be focused too much on statistical techniques and calculations and not enough on statistical concepts. Our personal experience with mathematics teachers who teach statistics suggests that they have difficulty implementing statistics as a problem-solving strategy. The higher the grade level, the more difficult it is for teachers to break away from their reliance on the definition and computational aspects of statistics.

Statistical problem solving is often characterized as "detective" work. Appropriate technology for learning statistics should supply the tools necessary for the exploratory nature of this detective work as well as assist in developing the skills and processes required in statistical problem solving. Additionally, solving real statistics problems may involve large and complex data sets, which are not easily managed through hand calculations. Technology frees students from performing tedious calculations and allows them to experiment with a variety of data representations. The interactive and graphical capabilities of the microcomputer offer the opportunity to create a laboratory environment in which students experience statistical problem solving in a dynamic fashion. The appropriate integration of technology enhances the development of statistical concepts and methodologies, statistical thinking, and alternative problem-solving strategies.

In this article, we demonstrate and compare three commercially available statistical software products. Data Insights is designed for either Apple II–class or IBM-compatible microcomputers. Statistics Workshop, developed by BBN Laboratories, runs on Macintosh microcomputers. Data Insights and Statistics Workshop are promoted as being appropriate for use in grades 6–12. Data Desk was developed by Paul Velleman and runs on Macintosh microcomputers. The student version now available is college level. In its present state, Data Desk is not appropriate for use in middle schools; however, a school version is under development. This powerful software not only supplies the standard graphical and numerical statistical representations but also allows these representations to be *linked.* This linking allows the student to view multiple representations simultaneously and to detect underlying connections that may exist among several variables. Biehler (1993) gives a comprehensive review and critique of a variety of microcomputer software for learning probability and statistics in the schools.

*Prepared by* GARY KADER *and* MIKE PERRY, *Appalachian State University, Boone, NC 28608. Edited by* GEORGE BRIGHT, *University of North Carolina at Greensboro, Greensboro, NC 27412-5001.*

# with Technology



**Are some head sizes more common than others?**

The following problem was developed by the Statistics-Materials and Activities for Problem Solving project (STAT-MAPS). STAT-MAPS, is a three-year project (1991–94) funded by the National Science Foundation's Materials Research and Development Program (grant no. MDR-9150117). This problem can be studied on several levels. We have presented this problem at workshops for both middle-grade and secondary school teachers, and we have used it as an investigation in our introductory, college-level statistics courses. The problem is certainly appropriate for study by students in grades 6–12.

## The Problem—The Hat Shop

HATS ARE MADE IN A VARIETY OF STYLES AND SIZES. A merchant must decide what styles to keep in stock and how many hats of each size to order. Are some hat sizes more popular than others? If so, the merchant would want to order larger quantities of the more popular sizes. If not, then an equal number of each size would be ordered.

Hat size is clearly determined by head size. Several possible measurements of the human head might be used to describe head size. Mail-order catalogs ask you to measure your head circumference to determine your hat size. This description comes from the 1993 L. L. Bean summer catalog: *Measure around the largest part of the head with tape above the brow.* The catalog supplies a table from which you can determine your hat size from the circumference of your head.

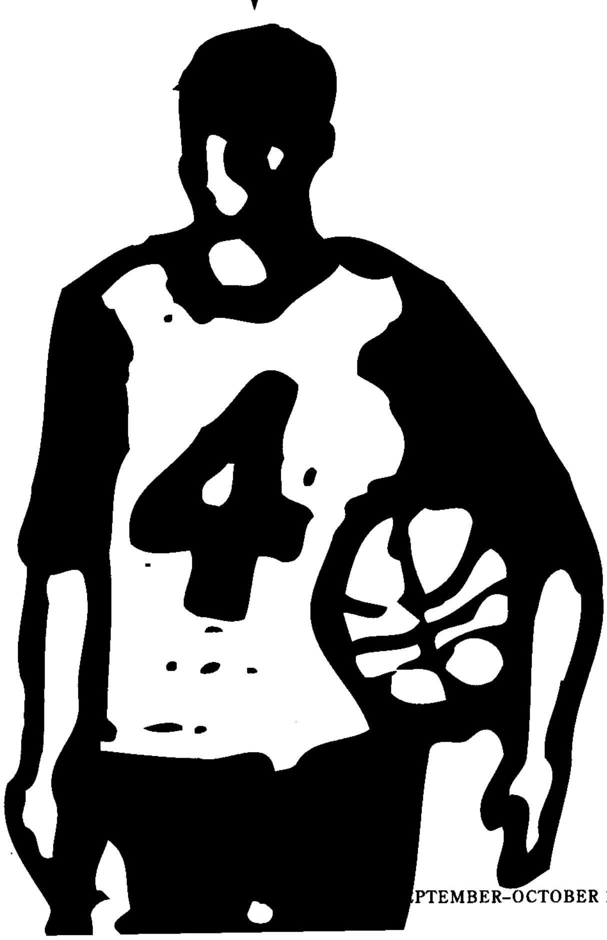
### The investigation

*Question*

Are some head sizes more common than others?

*Data*

An investigation of this question may initially involve measuring the students in your class. The head circumferences, in millimeters, of thirty students in an introductory statistics class at Appalachian State University are shown in **table 1**.



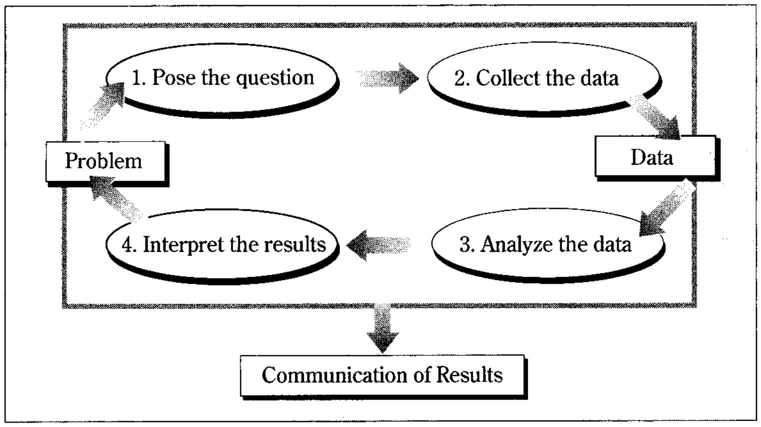



**Fig. 1 The process in statistical problem solving**

TABLE 1

**Head Circumference (in mm) of Thirty College Students**

| Student | Head Circumference | Student | Head Circumference |
|---|---|---|---|
| 1 | 588 | 16 | 555 |
| 2 | 578 | 17 | 555 |
| 3 | 569 | 18 | 579 |
| 4 | 575 | 19 | 575 |
| 5 | 573 | 20 | 579 |
| 6 | 554 | 21 | 565 |
| 7 | 568 | 22 | 594 |
| 8 | 570 | 23 | 578 |
| 9 | 563 | 24 | 578 |
| 10 | 561 | 25 | 546 |
| 11 | 538 | 26 | 541 |
| 12 | 568 | 27 | 572 |
| 13 | 550 | 28 | 562 |
| 14 | 542 | 29 | 591 |
| 15 | 588 | 30 | 561 |

*Analysis and interpretation*

The analysis component should consist of appropriate data representations that relate to the question being asked. Data representations should include numerical summaries and graphical displays of the data. However, we should keep in mind that the question about head sizes was asked to address the more general problem of ordering hats.

One appropriate graphical display for these data is called the *line plot*. The line plot is a one-dimensional graph that displays how often each value in this data set occurs. A line plot similar to the one generated by Data Insights for the thirty head circumferences is shown in **figure 2**.

An examination of this plot tells us something about the head sizes for these thirty students. We note immediately that all head circumferences are between 538 mm and 594 mm. Also, for each possible value between 538 and 594, some values do not occur at all, some occur exactly once, and some occur more than once. A concentration of head sizes falls between 560 mm and 580 mm, with eighteen out of thirty head sizes—more than half—in this narrow interval.

Of course, the original problem under study is determining how many hats of each size should be ordered, and hat sizes to the nearest millimeter are usually not available. In the conversion-to-hat-size table in the L. L. Bean catalog, for the most part the increases in hat size correspond to an increase in head circumference of three-eighths of an inch. In metric units, three-eighths of an inch converts to approximately 10 mm. Consequently, these data may be best examined by *grouping*. A convenient way for grouping data is to examine a *stem-and-leaf plot*. The default stem-and-leaf plot generated by Data Insights is similar to the one shown in **figure 3**.

We keep in mind that each stem in this plot represents an interval of values. For instance, stem 56 contains all circumferences in the interval 560 to 569. In this particular example, the interval length between stems is 10 mm, which happens to correspond to the increase in hat size. One advantage to using a stem-and-leaf plot is that the original measurements remain available.

Although the-stem-and-leaf plot has less detail than the line plot, it provides similar information about the head circumferences for these thirty students. First, observe that all circumferences measure between 538 mm and 594 mm. Also, note that not all stems have the same

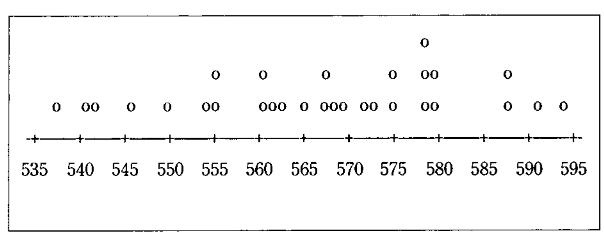



**Fig. 2 Line plot for thirty head circumferences**