From Phillips, Brian (Ed.), (1997) Papers on Statistical Education presented at ICME-8 (International Congress on Mathematics Education-8) Seville, Spain, July 14-21, 1996. Hawthorn, Australlia: Swinburne Press. ISBN 0 85590 753 3

STATISTICS: TOWARDS THE 21ST CENTURY

Kay Lipson & Peter Jones
Swinburne University of Technology
Australia

Statistics education in the 21st century will be concerned with the statistical process, and not with computation, which has been the focus of statistics teaching and learning for most of the 20th century. What do we mean by the statistical process?. The main stages are described in figure 1(after Graham, 1987):

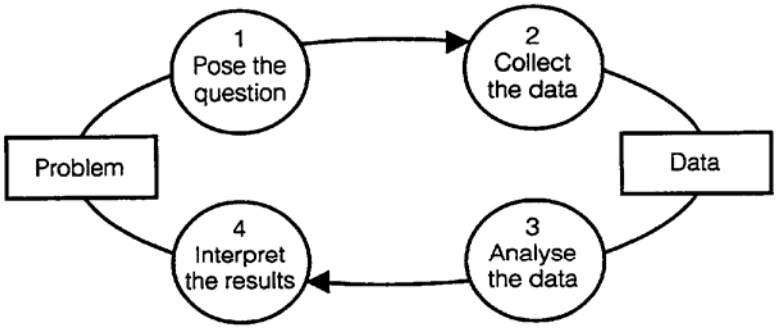

Figure 1: The Statistical Process

Each of the four steps proposed in figure 1 are important in their own right. But it is only when they are encountered as an integrated unit that the philosophy and the logic of the statistical process are encountered. Exposing students to the cycle of the statistical process gives them an opportunity to build an understanding of the place and the purpose of the various statistical techniques which have been so highly valued in the past. The challenge for the statistical educator as we proceed into the twenty-first century is to develop teaching strategies embedded in courses which give students the opportunity to experience first hand the statistical process in its entirety. The difficulty with this objective is that some aspects of the process, such as data collection, may be extremely time consuming, whilst others, such as generalising from the data analysis in answering the question, may rely upon sophisticated statistical techniques which are beyond the scope of most school level courses.

One way in which students may be able to experience more of the statistical process first hand is through the power of modern technology. In the remainder of this paper we will investigate some of the roles the technology might play in the statistics classroom of the future.

Using technology as a source of questions

One of the most remarkable advances in technology in the past few years has been the potential for students to connect to the World Wide Web. Communication via the net is quick, easy and extremely effective. Exploring the Web gives us access to information regarding topical issues from all over the world. As well as the ability to go direct to the source of the information, some specialist sites which support statistical education have been developed. One such site is the CHANCE course (see figure 2), whose pedagogical approach is based on a philosophy of discussing statistics arising from published articles. By reading an appropriately chosen newspaper or magazine article, which is both relevant and topical to a particular group of students, the class may begin a discussion on an issue which readily lends itself to statistical analysis.

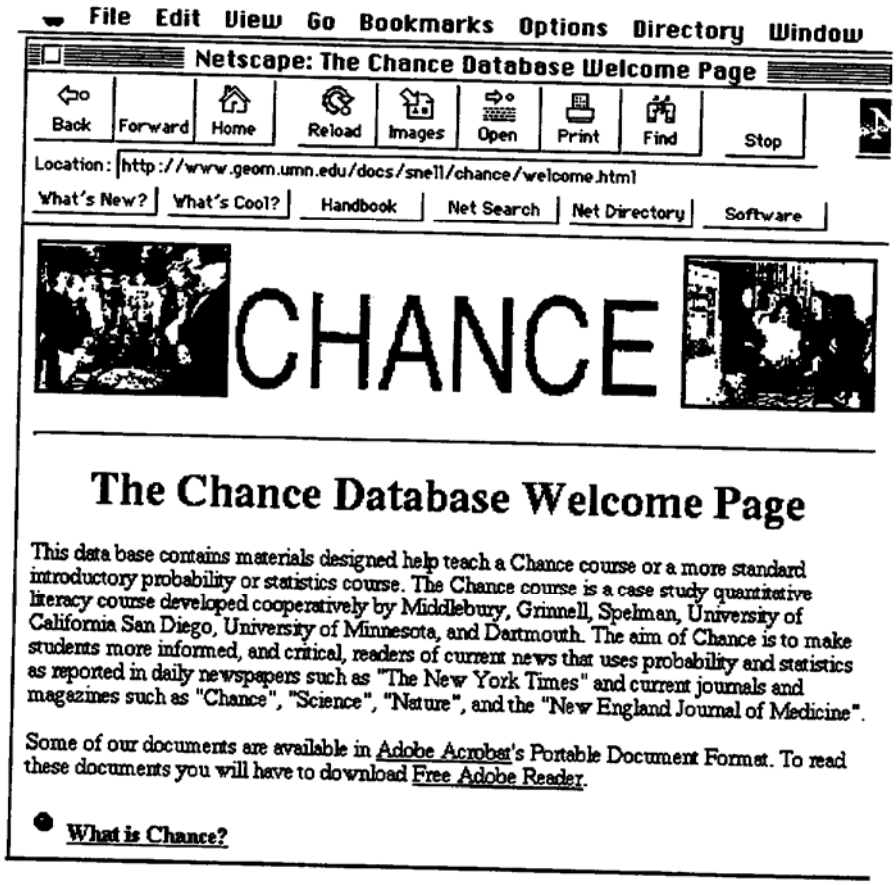

Figure 2: Home page for the CHANCE course

Using technology as a source of data

Once a question has been formulated, data are required in pursuit of a statistical answer. Sometimes the data can be collected first hand (primary data), but more often than not secondary data is more appropriate – that is, data already collected by another. The Web is a valuable source of secondary data. In this year of the Olympic games, it might be interesting to examine a sporting question, such as "Who is Sergei Bubka, and why he is he considered such a star in track and field events?". So now we have a question and some data, how do we proceed? A quick search of the net will reveal many interesting data sets relevant to this question. We have chosen to look at the world's all time top 20 results, shown in figure 3.

The data is readily converted to a useful form. We chose to copy the data and paste it into a word processing package, where is could be easily edited and prepared for subsequent analysis.

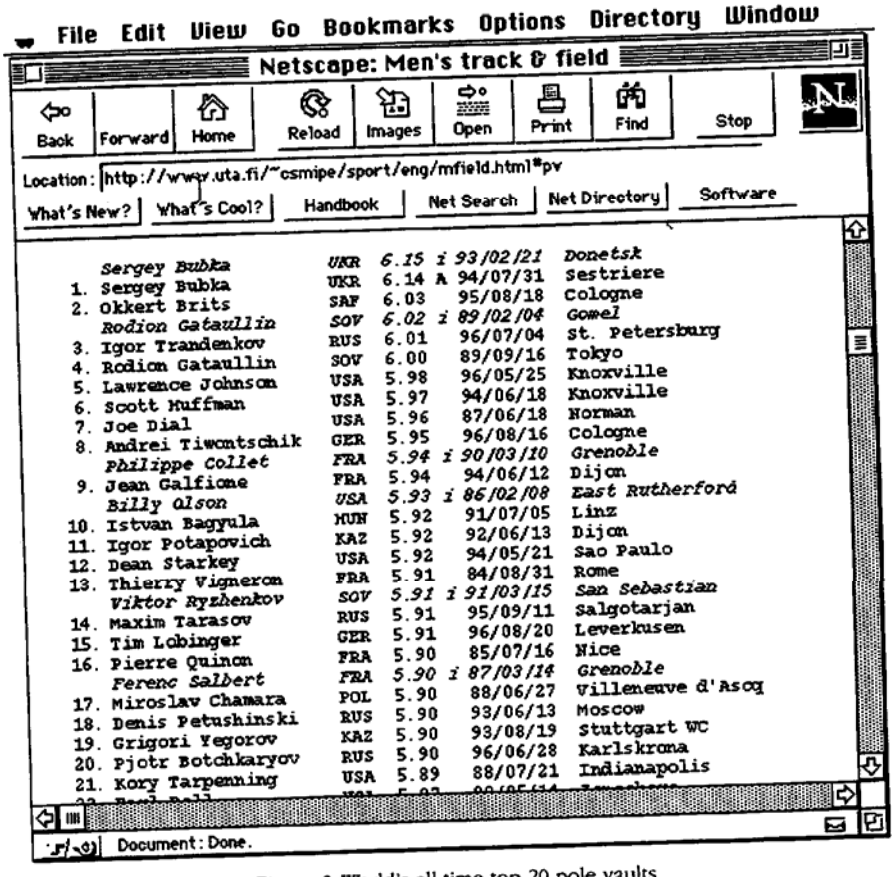

Figure 3: World's all time top 20 pole vaults

Using technology to analyse data

How should we now proceed about analysing the data. Of course, having the data already in the computer suggests we should use the computer based statistical package of our choice to proceed. However, there are other paths which may be followed in using the technology to analyse the data. The one which will be followed here uses the TI-83, a graphics calculator which offers most of the features of a computer based statistical package that are needed at the introductory level with the convenience of a calculator. Using the Graph Link software and cable, the data can be loaded directly into the calculator from the computer where it was downloaded, avoiding any necessity to enter the data by hand. Some of the interim calculator screens are shown in figure 4.

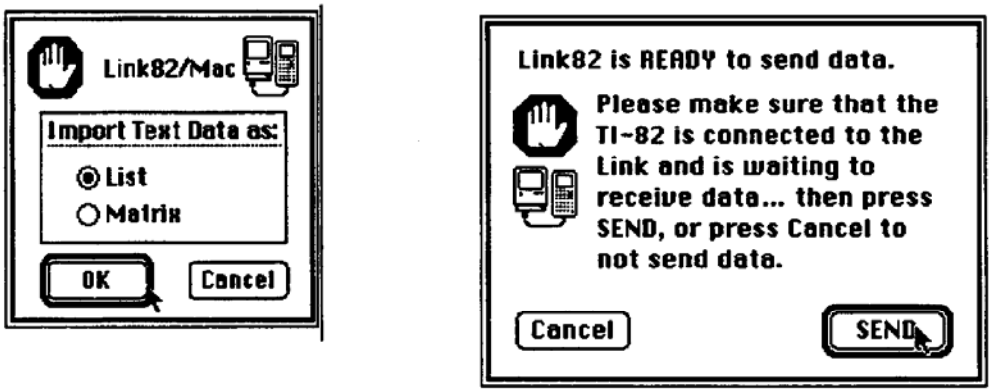

Figure 4: Using the Link 82 software the data can be transferred directly into the graphics calculator ready for further analysis.

Most of the desired statistical plots, such as histograms and boxplots are readily available. A boxplot of the pole vault distances readily shows that their distribution is positively skewed, and that Sergei Bubka stands well out from his peers, as shown in figure 5.

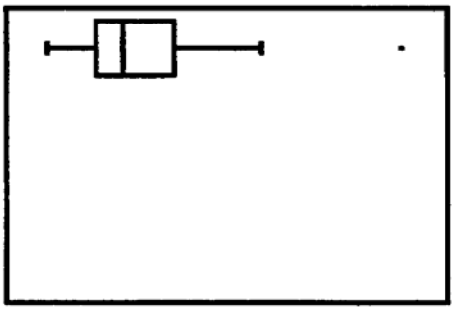

Figure 5: A boxplot of the world's all time best pole vault distances for males showing that Sergei Bubka is indeed an outlier amongst his peers.

The graphics calculator also has the additional feature that plots may be overlaid, with the result that multiple representations of the data can be created. For example, a histogram of the pole vault distances can be constructed on the same pot as the boxplot, as shown in figure 6. As well

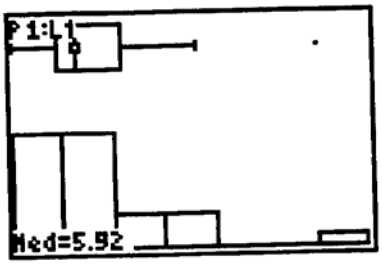

Figure 6: A histogram and a boxplot of the pole vault distances for males, with the median located.

as the multiple plots, by putting a cursor onto the screen using the TRACE key, both the boxplot and histogram may be interrogated interactively. In figure 6, the cursor is on the median of the boxplot, and the value of the median shown. The student may use this multiple plot to locate by eye the position of the median on the histogram. Such a plot not only gives students insights into the actual data set being analysed, but also into features of the statistical tools being used for the analysis.

Before proceeding any further with the analyses it might be interesting to compare the male and female achievements in the pole vault. Downloading the data for females gives rise to the plot shown in figure 7, indicating that female pole vault distances are also positively skew, and also has its super athlete, Emma George from Australia!

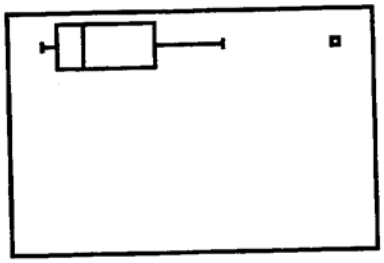

Figure 7: A boxplot of the world's all time best pole vault distances for females showing Emma George as an outlier.

Of course, the next interesting question might be to compare male and female distances. This can be readily investigated visually on the graphics calculator, as shown in figure 8. From this plot it may be seen quite clearly that there is quite a discrepancy between the male and female distances for the elite athletes.

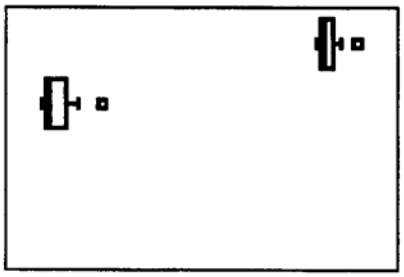

Figure 8: Boxplot of the world's all time best pole vault distances for both males and females.

Revisiting the statistical process

Of course, as we proceed with the analysis we find more and more questions keep arising. Is this set of data the most appropriate for answering the original question? Should we be comparing Bubka to his peers, or finding some way of comparing him to elite athletes in other sports, such as number of world records, number of world championships, length of time ranked number one in the world, and so on? The list of possible, and available, data seems endless. Perhaps we should try to find some summary statistics which puts together much of this information, and allows us to rank athletes from many different areas. With access to the WWW, students are able to keep refining the question, collect the appropriate data, and analyse it quickly and efficiently. The technology is allowing them to creatively deal with a real world statistical problem.

Discussion and conclusion

Clearly these are fascinating and challenging times for statistics educators. As we approach the twenty first century, the advances in hardware and software are immense and there is no reason to believe this rate of change will not continue. Some teachers of statistics have been dismayed by the changes, as they challenge the value of much of what we have taught in the past. But there is every reason to believe that by embracing the technology in appropriate ways and exploring what it has to offer, the teaching and learning of statistics has a great potential to be improved, and the cause of promoting understanding in statistics advanced.

Finally, as shown in figure 9, technology plays a critical role in facilitating and enhancing the statistical process. The challenge to statistical educators is to look for ways in which the technology allows us to explore the statistical process with our students in ways we simply couldn't do before, rather than use the technology to replicate old, and maybe no longer relevant, skills.

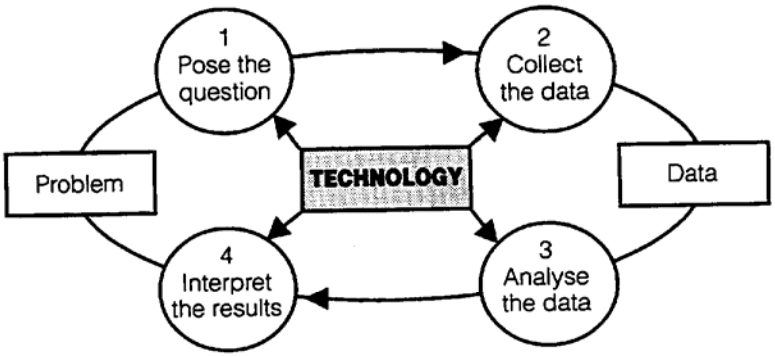

Figure 9: The Statistical Process - Underpinned by technology

References

Graham, A. (1987). *Statistical investigations in the secondary school.* London: Cambridge.