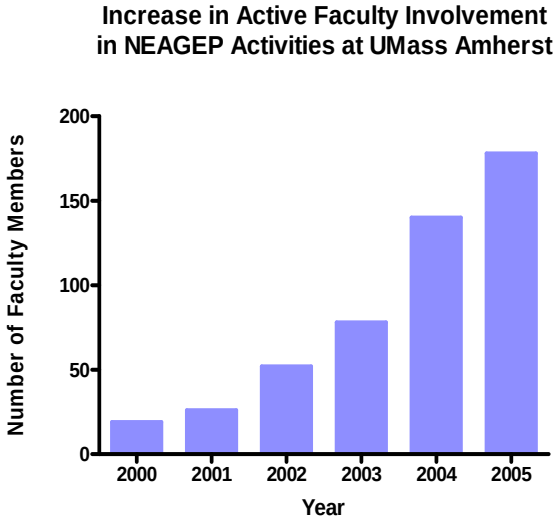
Collecting, Analyzing, and Displaying Data

Patricia B. Campbell, Rosa Carson, Campbell-Kibler Associates, Inc.,
and Shirley M. Malcom, AAAS

The following tips grew out of a review of a series of posters developed by universities participating in the National Science Foundation's Alliance for Graduate Education and the Professorate (AGEP) and presented at a 2005 AAAS-sponsored seminar to build institutional evaluation capacity. While the data from the posters may not be up-to-date, the lessons learned are timeless.

I. Make Your Message Clear

Data are displayed to get a message across to the reader. For the message to be received, it is important to be up-front and explicit about what message that you want people to take away and then use the data to support that message. In the following chart, the University of Massachusetts Amherst did just that. They wanted people to see changes in faculty involvement in the New England Alliance for Graduate Education and the Professorate (NEAGEP) over time, and that is exactly what this very clear and straightforward graph below does.



1 The works of Edward Tufte including *The Visual Display of Quantitative Information*, *Envisioning Information* and *Visual Explanations: Images and Quantities, Evidence and Narrative* http://www.edwardtufte.com/tufte/ are excellent resources on displaying data.

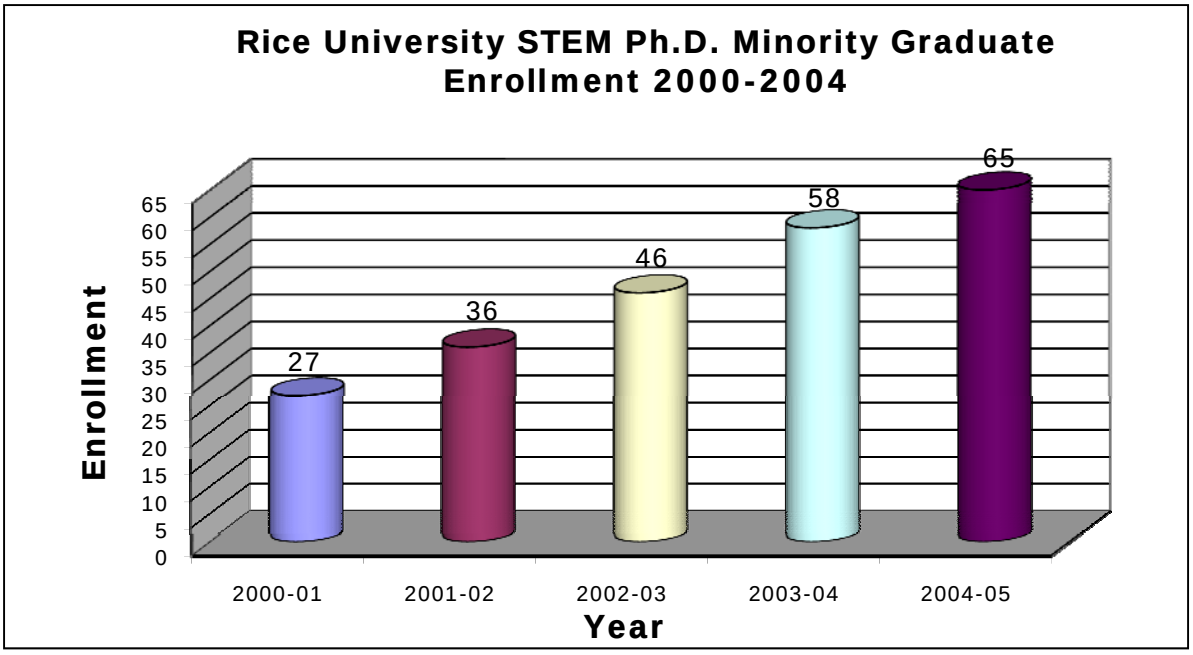
- Use font sizes large enough that people can easily read your work. Text in posters, power point presentations, and even transparencies needs to be at least an 18 point font, and print text needs to be at least a 10 point.
- Proofread and check your computations. Then do it again. If, for example, numbers in a table do not add up to the total, if numbers in the text do not reflect the numbers in the table, or if percentages are wrong, then your message will be lost.
- Use tables and graphs to present numbers and refer to them in the text as necessary. Do not use text to present a lot of data; it is too hard to read and compare numbers presented in text.

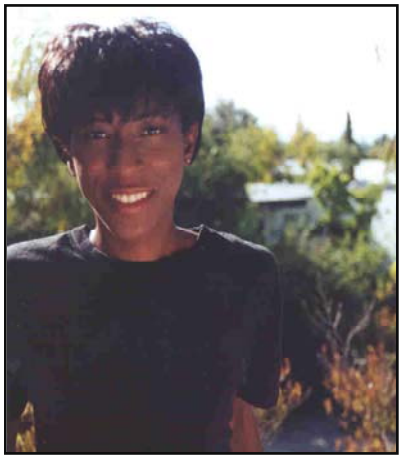
II. Use Pictures, Where Appropriate

Graphs and tables are useful tools to present data but pictures can provide an even more powerful way of getting your message out. This picture of SUNY Stony Brook AGEP students and staff with the University's Nobel Laureates presents evidence that, at Stony Brook, AGEP students are indeed interacting with top faculty members.

III. Use Statistics and Stories

Stories of the successes of individual students can be very powerful, but they say little about the overall success of a program. Statistics on the other hand can provide information about the program as a whole but tend not to attract the heart or, perhaps more importantly, the eye. Knowing this, the Rice AGEP provides both statistics and stories of the real people who make up those statistics.





Dr. Nancy Glenn
Ph.D., Statistics, 2002
Assistant Professor,
University of South Carolina,
Statistics Department (tenure track)

IV. Be Responsive to Your Audience

In displaying data, it is important to keep your audience in mind. Since different audiences have different levels of interest and expertise, to be most effective, the display of data should be targeted to the audience. This can be expensive and time consuming if it means that a number of different posters have to be designed and printed. The North Carolina AGEP has found a creative answer to the problem. Rather than having a whole poster, they have a number of separate independent pieces covering such areas as recruitment, retention, and evaluation. They then combine different pieces to create a poster targeted to a specific audience, pinning the individual pieces to the wall in the same way a full poster would be pinned. When presenting to a different group, the North Carolina AGEP goes back and reselects and recombines the individual pieces to create a different poster which reflects the needs and interests of that group.

V. Make Comparisons

"Compared to what?" is a typical and important question that readers ask when they are looking at data. In evaluation, as well as science, one data point or one set of data is not of great value. For example, the University of California did a study of doctoral completion rates after 10 years and found that 68% of underrepresented graduate students in the life sciences completed their PhDs in the ten year period. That piece of data on its own is not particularly useful. However, when comparisons are made across different subject areas and different racial/ethnic groups as Jim Litrownik of the University of California did in the following table, a much more comprehensive picture is presented; one that can be used to identify possible problem areas as well as areas of success.

Doctoral Completion Rates After Ten Years
(Fall 1988-90 Doctoral Entry Cohort)

	Under-represented Minority	Asian	White	All Students
Life Sciences	68%	74%	76%	75%
Engineering / CS	35%	41%	48%	45%
Physical Science & Math	57%	49%	65%	62%
Total Nat Sciences	55%	52%	65%	62%

The Florida AGEP also compared data across different subject areas, but in addition, they looked at it over time and in terms of Applicants, Admits, New Enrollees, Total Enrollees, and PhD Recipients. They then went the next step, discussing the results in the text and looking at how the data could be used in determining next steps and target areas.

Total Underrepresented Students	Engineering	Computer Engineering	Computer Science	Chemistry	Biological/ Agricultural Sciences
Applicants	94 (87)	10 (6)	7 (1)	19 (11)	68 (61)
New Admits	55 (45)	4 (4)	3 (1)	15 (0)	24 (18)
New Enrollees	24 (24)	3 (1)	2 (0)	3 (0)	15 (10)
Total Enrollees	153 (112)	14 (9)	2 (0)	19 (6)	83 (57)
PhD Recipients	20 (6)	1 (0)	0 (0)	11 (3)	20 (14)

An analysis of the data collected for each area can point to both areas of success as well as areas of opportunity. Computer Engineering and Computer Science have seen increases in applicants and have been very successful in enrolling those admitted to the program. The recent efforts to increase enrollments have not yet translated to increases in PhD recipients. Chemistry has had significant improvement in new admits, but there is an opportunity to improve the number of these admits who actually enroll as is the case in the Biological Sciences as well. The growth in total enrollees in all areas may be an indication that retention efforts are having an effect and that students are staying in the program. The UF AGEP has a retention rate of 86% to date.

At the graduate school level, comparison data can be found in several different places. Engineering data, including Masters and PhD enrollment and recipient data broken out by sex and by race/ethnicity for US students and by sex for foreign students are available from the Engineering Workforce Commission (http://www.ewc-online.org/). Data from all the subject areas can be found in the Survey of Earned Doctorates and Integrated Postsecondary Education Data System (IPEDS) using WebCASPAR (http://caspar.nsf.gov). The National Science Foundation's WebCASPAR database provides access to a large body of statistical data resources for science and engineering at U.S. academic institutions.

VI. Find Ways To Deal with Volatile Data

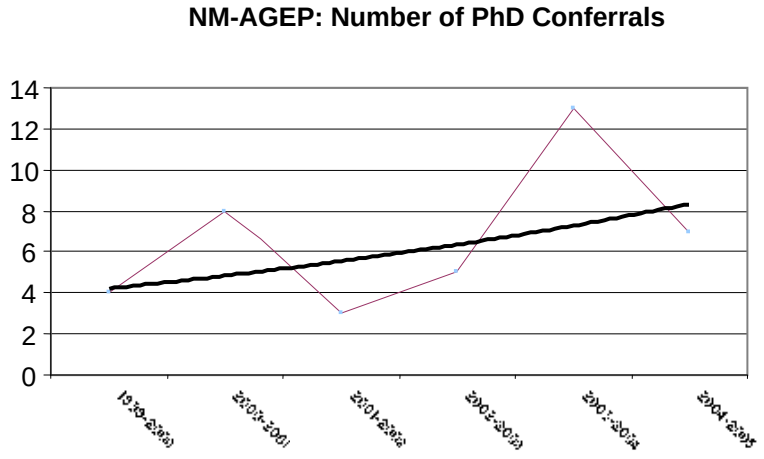
As the following example shows, data can vary greatly from year to year making it difficult to look for trends and patterns.

	1996-7	**1997-8**	**1998-9**	**1999-2000**	**2000-1**	**2001-2**	**2002-3**	**2003-4**	**2004-5**
All Under-represented Enrollees	5,059	5,262	5,722	5,039	5,207	6,119	5,600	6,596	6,305

One way to deal with volatility is to group data over several years and report the means rather than reporting data by year. That is what AAAS did in the following table, consolidating the data in the earlier table and making it clear that the number of all underrepresented enrollees in AGEP institutions is increasing.

	Pre AGEP (1996-7 to 1998-9)	**Early AGEP** (1999-2000 to 2001-2)	**Intermediate AGEP** (2002-3 to 2004-5)
All Underrepresented Enrollees	5,557	5,937	6,470

Another option is to, as David Johnson from the New Mexico AGEP did, create a graph with the yearly numbers and also with a trend line. That way the reader can see not only the yearly data but also see how overall the numbers are increasing.



Context can also be helpful in looking at volatile data. For example, if there were particular circumstances in a given year—either positive, like a new program, or negative, such as a large budget cut—including these as part of the graph can help readers not just see the trend but help them to understand what might be behind it.

VII. Use the Results

There is little value in collecting data without using them. To increase the probability that somewhat standardized data were collected and used, the Michigan AGEP Alliance has developed a series of templates to use in their process evaluation. The illustrative template below is used for Michigan AGEP Alliance planning and evaluation:

Goals(s)	Activity	Objectives	When	Target Population
Improve retention of current students in STEM fields	MI AGEP Fall Conference	Motivate students' interest in academic careers; provide information on networking strategies to improve career advancement	Oct 1, 2005	Currently enrolled graduate students at MAA institutions

A second template, like the one below, is used as a tool for questionnaire development.

Indicator	How measured (question stems)	Data Collection Method
Participant rating of helpfulness/ usefulness of sessions	What session did you find most inspiring or useful today? What was one important strategy or lesson you learned from that session? What session was the least useful? Why?	Self-administered written questionnaire
How well participants' experience matched session objectives: 1 Have you considered an Academic career?	Did the conference today help you to feel: (a) more positive; (b) more negative; (c) about the same about pursuing an academic career? If more positive, what did you learn that encouraged you? If more negative, what did you learn that discouraged you?	Self-administered written questionnaire

Process evaluation results are then shared across and within MAA institutions' programs.