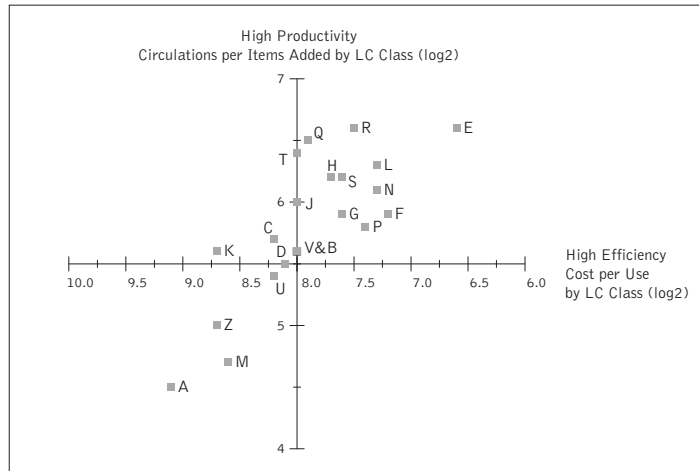


SIMPLIFY: Don't let graphical design intrude on the data

- Avoid piling up patterns and tick marks, overusing outlines, and keying variables to legends
- Label within the data frame rather than on a legend.
- Use color selectively; conserve color to communicate significant information.
- Juxtapose multiple graphs rather than superimpose elements on a single data frame *if* the superimposition clutters the picture
- Order data (for example, expenditures in labeled categories) to optimize visual perception

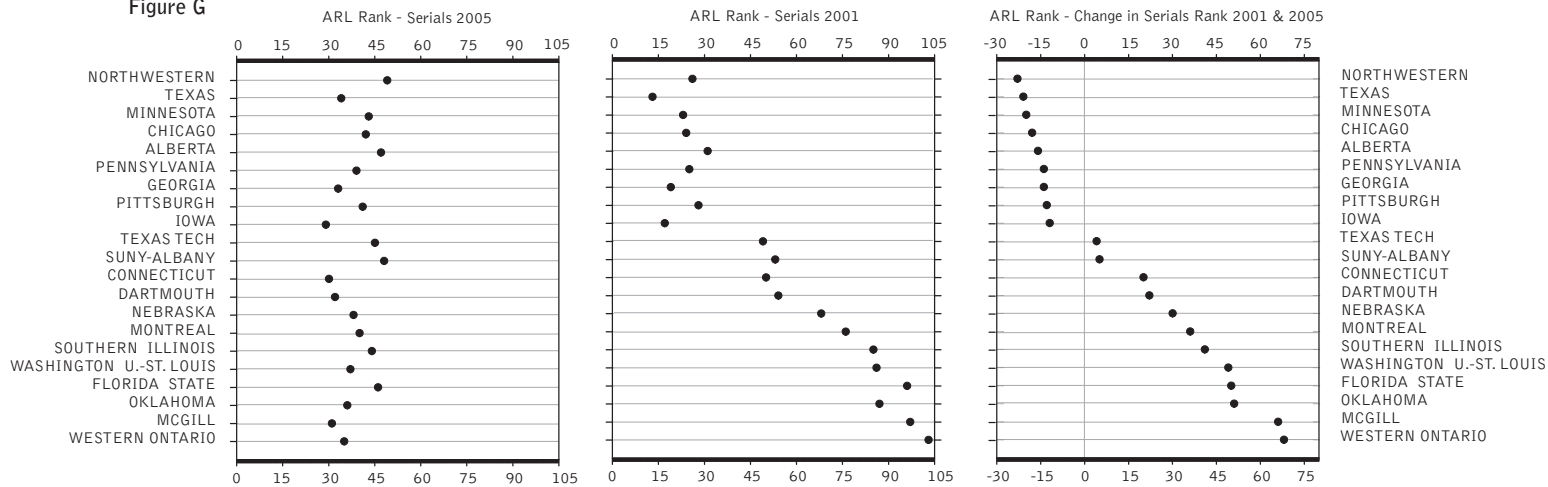
Figure F



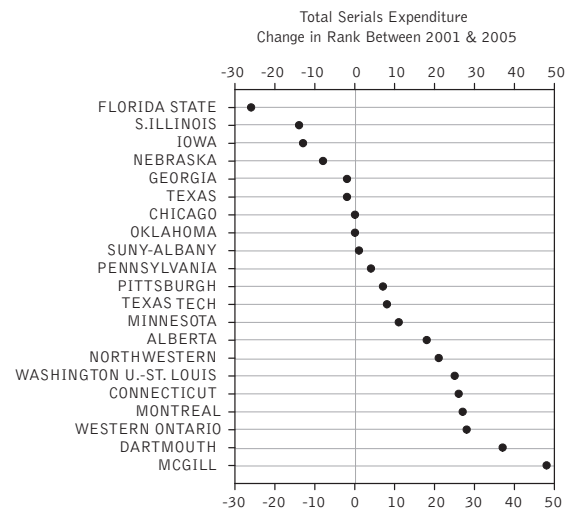
Design Simplicity|Data Complexity: Designs should strive for simplicity, but that principle doesn't necessarily hold for the data they encode. A well designed graph should summarize lots of data without creating visual puzzles for the eye and brain.

Figure F is an example of expanding complexity. It employs the same 22 observations used in Figure B, the ratio of circulations to items added within LC Classes, and sets them along side cost per use factors for each class. The graph condenses 44 data points that help to reveal the cost efficiency (measured by cost per use) and productivity (measured by use per items added) of Penn's acquisitions program in specific fields. So, for example, although the cost efficiency of materials purchased in U.S. History (E) is about twice that of Medicine (R), the medical literature is very slightly higher in productivity. A graph like this can stimulate discussion about collection priorities and how to frame them, even as it points to the possibilities of more detailed analysis.

Figure G



Picturing Data: The dot plot is a useful device for picturing data relationships. The plots in Figure G are mini spreadsheets that show an odd shifting of ranks within ARL. In the top row, the grids are snapshots of ARL Serials ranks for the ten libraries above and below Penn in 2005. The size of the gap between the 2005 and 2001 Serials rank determines the order. Northwestern suffered the largest negative shift—23 positions—so it's first. W. Ontario's spectacular 68-position rise placed it last. Panel 3 shows these +/- shifts. Between '01 and '05 we see a significant realignment of libraries in their serials holdings ranks. Anomalies appear when comparing the holdings rank to the ranks for serials spending (see the panel on the right). Compare, for example, Florida State's 26-position decline in spending to its 50-position improvement in holdings. The panels picture several layers of information, including trends, and "winners" and "losers" and raise interesting questions about how we measure and compare library inputs.



The graphs on these pages were produced using **Origin 7.5** from Origin Lab Corp. www.OriginLab.com and modified in Abobe Illustrator; CS2 for this handout.

Library Assessment Conference | Preconference Workshop

Data Analysis and Presentation, September 25 & 27, 2006

Joe Zucca | zucca@pobox.upenn.edu

Workshop Overview

This workshop looks at some problems of data collection and analytical presentation experienced by librarians. To frame discussion and help illustrate general concepts, I will draw on a few initiatives from the University of Pennsylvania Library. The underlying measurement issues will sound familiar; despite the idiosyncracies of Penn's approach. The purpose is to make concrete the challenges of

- developing metrics,
- performing measurement,
- collecting and managing data, and
- creating effective presentations.

At a deeper level, I hope the work we will discuss generates a few ideas about the fostering of an evidence-based management culture within research libraries.

Agenda

I. Getting started

- "Just the overwhelming scope of it all..." Review of the workshop survey inputs
- We must measure (right?)
- Contextualizing the discussion: the University of Pennsylvania

II. Framing issues, evaluating methods

Focus: Research and instructional services—some approaches to data collection, analysis and presentation

- Developing metrics
- Operationalizing data collection
- Dealing with analysis and presentation

Break

III. Framing issues and evaluating data:

Focus: Collection management and use

- Developing assessment tools
- Coping with data overload and computational problems
- Visualizing data

Wrap up

Along with this handout, you have a packet containing a variety of graphs and tables. I will use these "exhibits", a few Excel spreadsheets, and some web pages throughout our discussions. I hope you have time to rummage through the packet before we begin.