

Visualizing Data at the USDA Using Animations, Small Multiples and Interactivity

Irwin Anolik USDA - National Agricultural Statistics Service
1400 Independence Avenue, SW Washington, DC 20250-2001

Abstract

Plans are underway at the USDA's National Agricultural Statistics Service (NASS) to enhance data customers' ability to view, analyze and interact with NASS historical data using any web browser or mobile device.

We discuss concepts and technologies implemented to enhance the ability of NASS's data customers to find historical data of interest, make comparisons, and see important trends across variables.

The visualization methods presented include 1) animated maps which depict changes over time, 2) sparklines, invented by Edward Tufte and described as "word-sized graphics", employed in small multiples which allow the display of large amounts of historical data on a single web page or mobile device screen, and 3) interactive charts to enable our data customers to drill down and focus on more detailed views of these data displays. Each of these methods facilitates the display of large volumes of data and allows data customers to extract information from our statistics that is difficult or impossible to obtain from traditional static charts or tabular displays of data.

Introduction

Well-designed data graphics and maps can help someone understand the underlying data much better than a simple table of numbers. Until recently, historical data offerings on the NASS website were generally limited to static charts and maps, and text and tables via text files, spreadsheets and PDF documents.

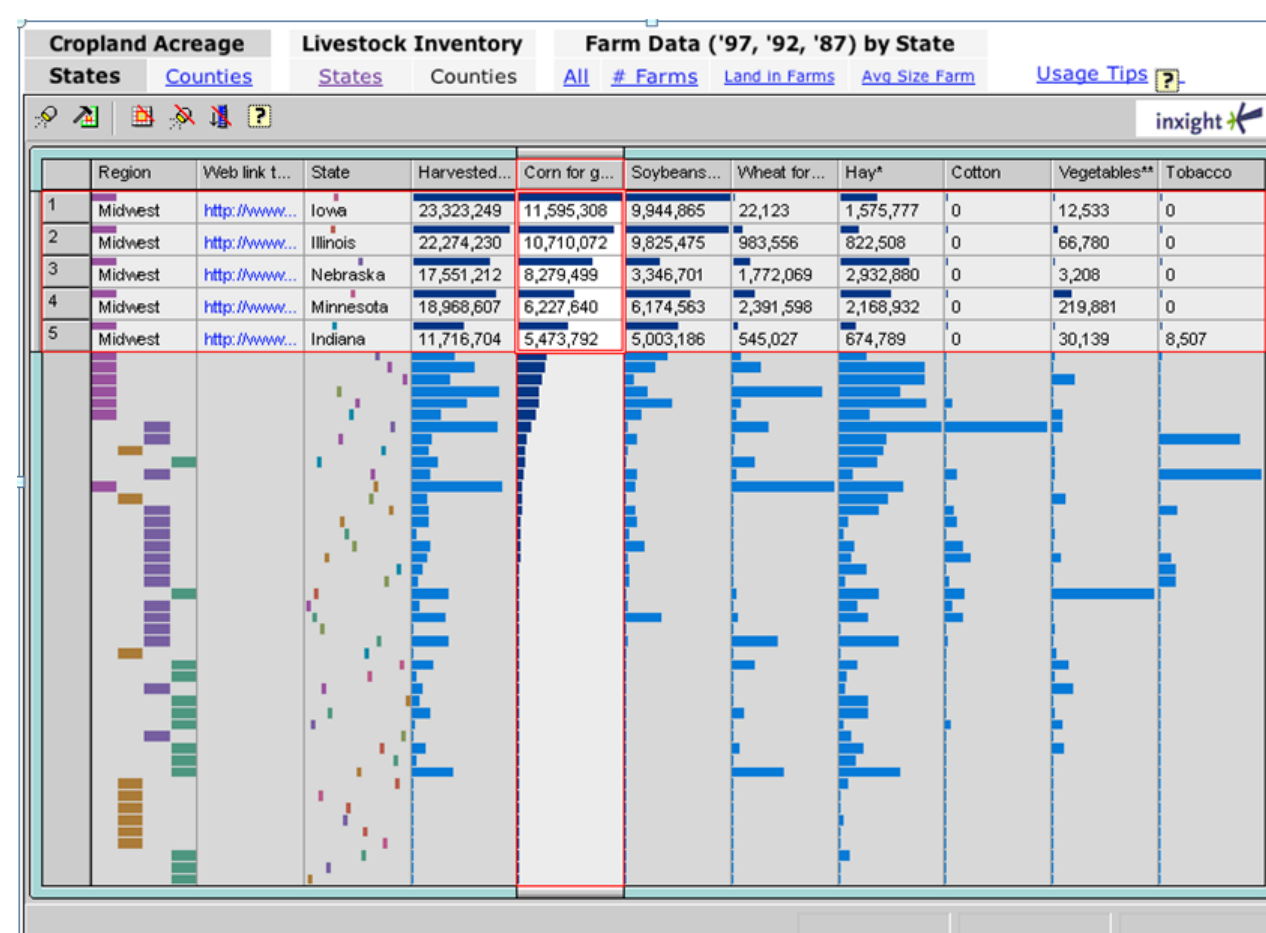
Modern computers and the internet have made it much easier to make graphics out of tabular data and give these graphics the qualities of animation and interactivity. However, a computer alone doesn't make a good graphic. In this presentation, we explore some principles of good data graphics, which we apply to NASS historical data series on the NASS website. We also share some novel techniques that are available to create effective ways of visualizing these data. Given the rich and comprehensive nature of NASS historical data, it is important to effectively and efficiently present this information. Charts can quickly become cluttered with too much data, but good graphical design can reveal new insights from these data.

In his seminal book entitled *The Visual Display of Quantitative Information*, Edward Tufte writes that "Graphical elegance is often found in simplicity of design and complexity of data." He then discusses ways to improve displays of statistical information starting with the proper choice of a format and design. He recommends to:

- 1) use words, numbers, and drawings together,
- 2) reflect a balance, a proportion, a sense of relevant scale,
- 3) often have a narrative quality, a story to tell about the data,
- 4) create charts in a professional manner, with the technical details of production done with care, and
- 5) avoid using "chartjunk" or those visual elements in charts and graphs that are not necessary to comprehend the information represented on the graph, or that distract the viewer from this information.

"Table Lens"

The "Table Lens" is a visualization method to dynamically explore large amounts of tabular data. Using the spreadsheet model, the Table Lens displays regional data values in columns and rows without scrollbars and without obscuring any data and filling the cells with scaled horizontal small bars. Each row in the Table Lens represents a single region (e.g. state or county) and the columns represent a specific indicator.



Plans at USDA-NASS include use of an open source tool known as "datacomb" recently developed by Chris Polis which adds functionality and compatibility with modern browsers and devices.

The "Table Lens" has the powerful combined abilities to sort records, and then, focus or "zoom in" on areas of interest in the data (to reveal exact numerical information). The advantage is obtained by using a so-called "focus + context" or "fisheye" technique. These techniques allow interaction with large information structures by dynamically distorting the spatial layout of the structure according to varying interest levels. The relationship between focus and context supports searching for patterns in the big picture and dynamically investigating interesting details without losing framing context.

A significant feature of a table is the regularity of its content with information along rows or columns being consistent. This is reflected in the fact that regional tables usually have the name of the region at row and column edges that identify associated indicator data values. The Table Lens adjusts the layout of a table without bending any rows or columns. A Table Lens can explore patterns in regional tables and investigate various explanatory statistics using a multivariate dataset. The Table Lens should enable sorting by clicking on the header of the column that you want to sort by. Clicking again on the same header should reverse the order of the records. The graphic above shows an example of what a Table Lens could look like using a NASS dataset.

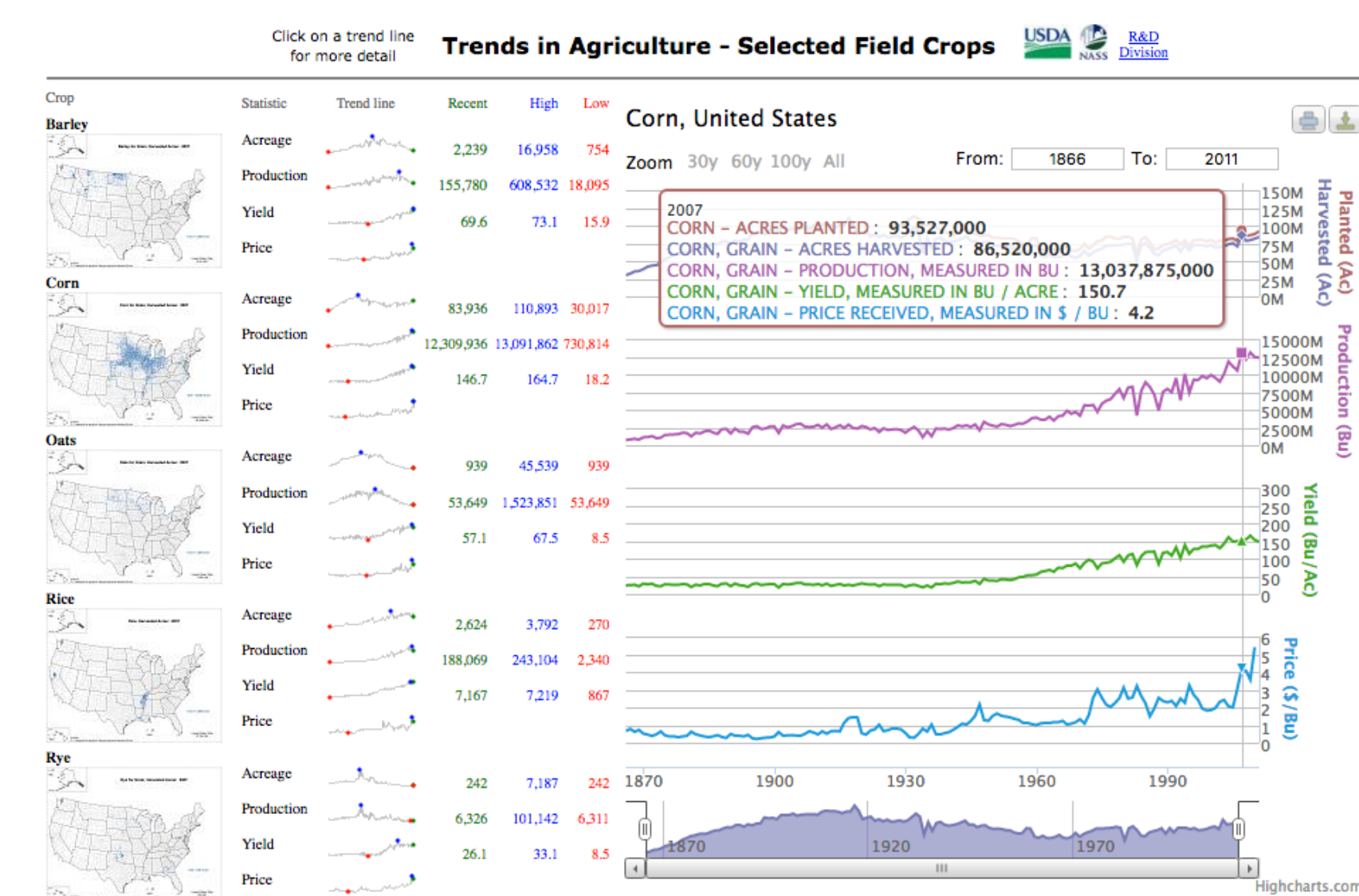
The zoom capabilities in the Table Lens work by clicking and dragging. To select an area to zoom in on, click and hold the left mouse button on the record that you want the area to start with. Then drag the mouse downwards until the regions of interest is covered. The selected regions will be enlarged to reveal exact numerical values as well as the horizontal bar representation. An example of the zoom function is shown in the graphic above.

Sparklines and Interactive Charts

A sparkline, as defined by Tufte, is "a small intense, simple, word-sized graphic with typographic resolution."

Sparklines ... can be everywhere a word or number can be: embedded in a sentence, table, headline, map, spreadsheet, graphic."

On the NASS website, sparklines will allow us to show long-term trends for multiple crops and multiple variables for each crop. Most of the sparklines represent 150 data points (one for each year), allowing us to visualize several thousand data points on one page in a way that is easy to understand and derive meaning from. A sparkline is a very small line chart, typically drawn without axes or coordinates. Here, they represent the general shape of the variation over time for acreage, production, yield, and price for major field crops, in a simple and highly condensed way. Note how these sparklines are small enough to be embedded in a table, and several sparklines are grouped together as elements of a small multiple. A small multiple is a series or grid of small similar graphics or charts, allowing them to be easily compared.



Sparklines have obvious applications for economic data, by tracking changes over time, showing overall trend as well as local detail. Note that sparklines are to be depicted on the NASS website as part of a data table, along with relevant color coded minimum, maximum, and most recent values and corresponding dots on each sparkline.

Interacting with the data graphic

Clicking on one of the sparklines for a given crop enables you to zoom in on variables of interest for that crop, and also enable interaction with the trend lines in a way that will display individual data values and zoom in on a desired period of time.

These tools are similar to what one sees on financial websites, such as those provided by Google and Yahoo.

Note that small multiples of maps are used in the appropriate position to provide spatial context alongside the series of sparklines.

The tools being developed by NASS will generate sparklines using the jQuery JavaScript library, and the accompanying interactive charts are created using the Highcharts charting library. Importantly, this enables data customers to use any mobile device (e.g. running iOS or Android) to access sparklines and interactive charts on the NASS website.

Conclusions

The charts and maps being developed for the NASS website and presented here were produced using a variety of products and programming languages. These programs and products are all useful technologies. They are most useful when information designers think critically about how best to display the data of interest.

When looking at NASS's historical data, researching and applying best principles and appropriate technologies to create well-designed graphics can help data customers find the information they want, and learn more about the big picture, and the story behind the data.



Animations

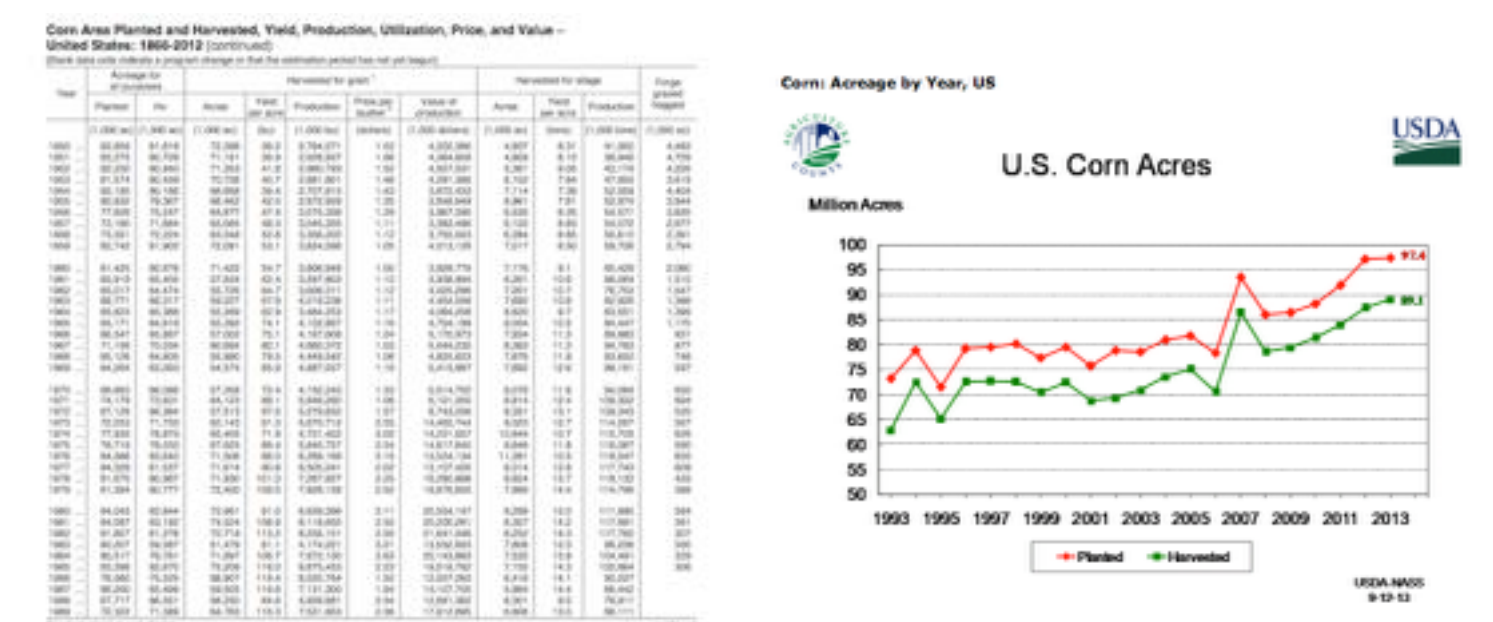
Animating a map or a graphic gives data customers a very effective way to visualize change over time. It can also provide the ability to control the playback of an animation with a pause-play button and scroll bar.

The use of the internet has allowed animated maps to become interactive. The user can witness representations of changes over time, while manipulating the direction of view, and the pace or the parameters of the map displayed (MacEachren 1998).

Map animations to be made available on the NASS website will help data customers visualize changes over time: namely, crop acreage animations that incorporate statistical graphics.

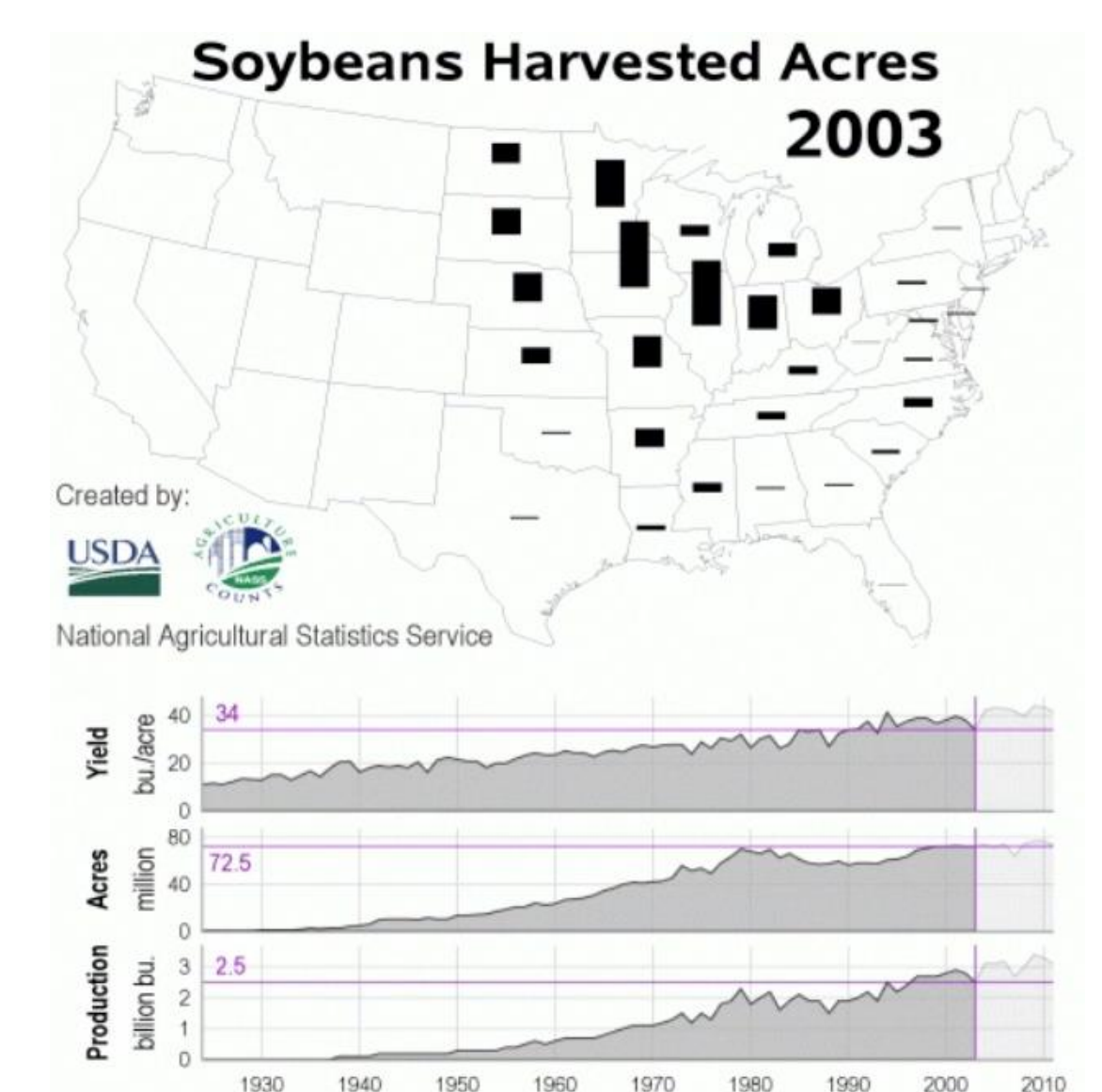
Maps can be enhanced by integrating charts and graphs. This was done commonly in the past when people had to draw maps and charts by hand. Many examples of this can be found at the David Rumsey Map Collection website: <http://www.davidrumsey.com>

Until recently, the NASS web site presented data such as crop acreage as part of a comprehensive PDF or text file or as a static graph showing the last 20 years of data. Examples of these are shown here:



The crop acreage animations for major crops to appear on the NASS website integrate time series charts that animate along with the map. As the animation progresses, the relevant values for production, acreage and yield appear on the plots. The harvested acreage for each state for a given crop, as represented by the bars over each state, changes year by year thus telling a more comprehensive long term story for the given crop.

One can then clearly see the trends for the US totals of production, acreage, and yield over time as well as how harvested acres changes geographically (by state) over the same time period.



The crop acreage maps and charts, were created using SAS. The animations were initially created using Apple's QuickTime player technology, and plans are underway to update them taking advantage of modern web standards and available JavaScript libraries.

References

- David Rumsey Map Collection website: <http://www.davidrumsey.com>
- MacEachren, Alan M. (1998). Cartography, GIS and the World Wide Web. *Progress in Human Geography* 22(4): 575-585.
- Tufte, E. (1990). *Envisioning Information*. Cheshire, CT: Graphics Press.
- Tufte, E. (2001). *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.
- Tufte, E. (2006). *Beautiful Evidence*. Cheshire, CT: Graphics Press.

Acknowledgments

Special thanks to Carol Crawford, Lee Ebinger, Bob Seffrin, and Eric Wilson for their expertise and encouragement in helping develop and implement the changes discussed and presented here.