Visualizing Historical Agricultural Data: The Current State of the Art

Irwin Anolik (USDA – National Agricultural Statistics Service)

Abstract

This paper reports on methods implemented at the National Agricultural Statistics Service (NASS) which enhance data users' 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 that are not possible with static tables or charts.

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 users 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 users to extract information from our statistics that is difficult or impossible to obtain from traditional static charts or tabular displays of data.

Keywords: (dissemination, visualization, animation, sparklines, small multiples)

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 paper, 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 this 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.

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).

This paper discusses 2 sets of map animations available on the NASS website to help data users visualize changes over time: Crop Progress and Condition Animations, and Crop Acreage Animations (incorporating statistical graphics)

Crop Progress and Condition Animations

Crop Progress Percents relate to acres and indicate the progress of field activities or crop development. Generally, an acre should be considered at or beyond a certain stage of development when 50 percent or more of the plants in that acre are at or beyond that stage.

Crop condition generally relates to yield prospects based on the condition at a given time, and the maps display percent good + excellent where good and excellent are defined as:

- Good - Yield prospects are normal. Moisture levels are adequate and disease, insect damage, and weed pressures are minor. Pastures are providing adequate feed supplies for the current time of year.

- Excellent - Yield prospects are above normal. Crops are experiencing little or no stress. Disease, insect damage, and weed pressures are insignificant. Pastures are supplying feed in excess of what is normally expected at the current time of year.

The crop progress and condition data is published weekly throughout the growing season as a series of dozens of tables for each crop progress and condition combination. Here is a recent example of one of those tables

Cotton Harvested – Selected States

[These 15 States harvested 98% of the 2012 cotton acreage]

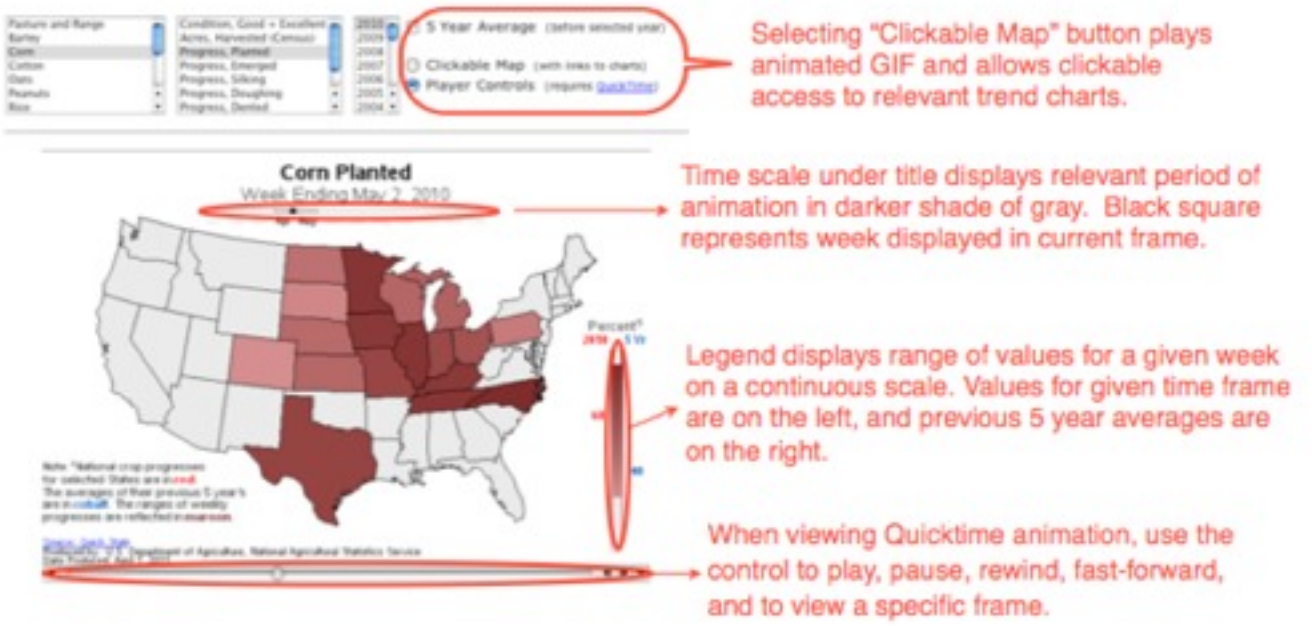
State	Week ending			2008-2012 Average
	September 29, 2012	September 22, 2013	September 29, 2013	
	(percent)	(percent)	(percent)	(percent)
Alabama	2	-	-	8
Arizona	19	5	7	19
Arkansas	20	2	5	16
California	-	-	1	1
Georgia	8	1	2	7
Kansas	1	-	-	-
Louisiana	57	15	30	48
Mississippi	23	1	3	23
Missouri	22	-	-	16
North Carolina	2	-	1	5
Oklahoma	2	-	-	2
South Carolina	3	-	2	7
Tennessee	18	2	4	14
Texas	15	10	11	16
Virginia	-	-	-	5
15 States	13	5	7	14

- Represents zero.

As stated on the NASS Research and Science website, “Animated U.S. choropleth maps of crop progress and condition state-level estimates greatly enhance current NASS offerings regarding crop progress and condition by more effectively showing the crop story to data users. Condition maps' sequential hues portray percent "good + excellent" for the crop by state, with weekly animated change. Their legend includes national condition. Progress maps' monochromatic shades portray percent progress for the crop data item by state, and their legend also includes the national 5-year average progress.”

Much effort went into selecting appropriate color schemes for the generated maps, using the “colorbrewer” schemes (developed by Cynthia Brewer at Penn State) as a guide. Colorbrewer is an online tool designed and developed to help select high quality effective color schemes for maps and other graphics.

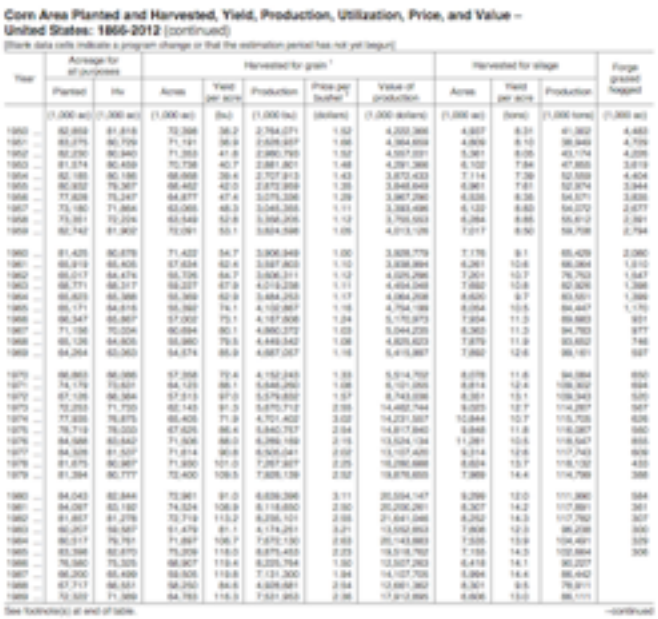
Here is a quickstart guide that shows what the screens look like and how to access the animations:

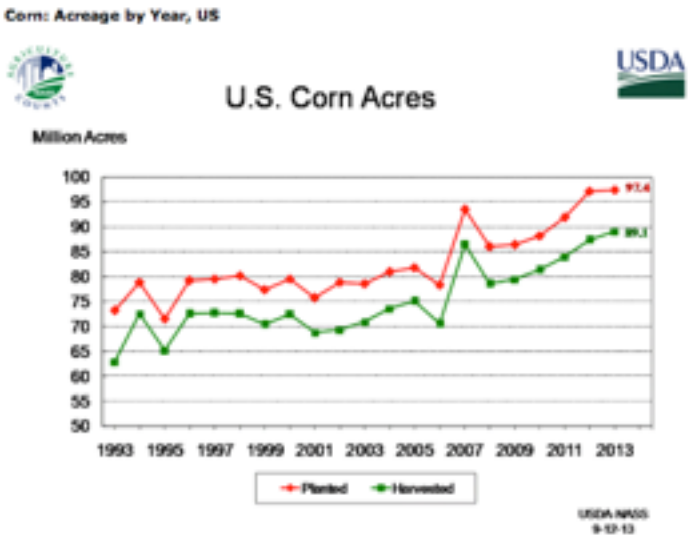


Crop Acreage Animations

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)

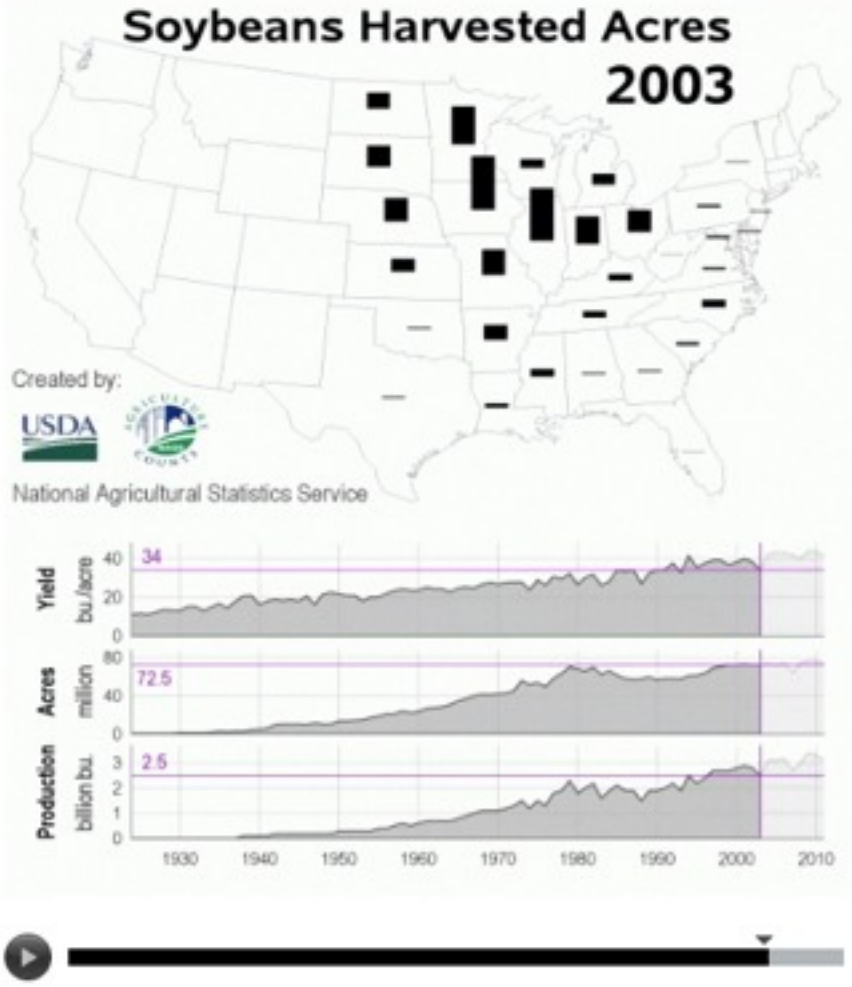
Until recently, the NASS web site presented data such as crop acreage as part of a comprehensive PDF or text file or as a static graph showing the last 20 years of data. Examples of these are shown here:


Corn Area Planted and Harvested, Yield, Production, Utilization, Price, and Value – United States: 1866-2012 (continued)



The crop acreage animations for major crops that now 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.



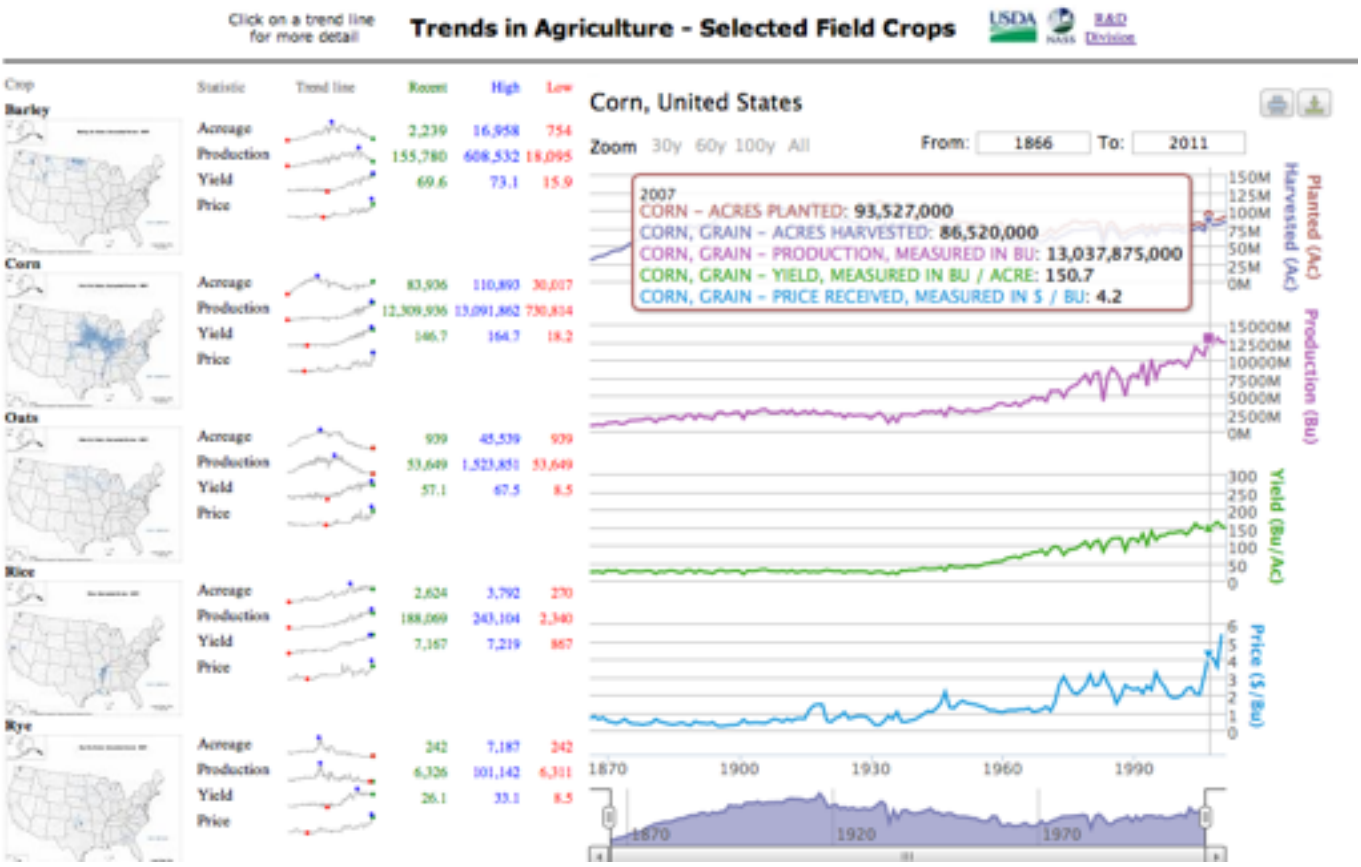
The crop progress and condition maps mentioned previously, as well as the crop acreage maps and charts, were created using SAS. The animations were created using Apple's QuickTime player technology.

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 allow us to show long-term trends for multiple crops and multiple variables for each crop. Most of the sparklines represents 145 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. On the NASS web site, 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 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.

Allow the ability to interact 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 very similar to what is available on the most famous 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.

Unlike Google and Yahoo, which use Flash technology to provide interactivity for the above mentioned websites, the tools developed by NASS generate sparklines using the jQuery JavaScript library, and the accompanying interactive charts were created using the Highcharts charting library. Importantly, this enables data customers to use their mobile devices (iOS and Android) to access sparklines and interactive charts on the NASS website.

Conclusion

The charts and maps developed for the NASS website and presented in this paper were produced using a variety of products and programming languages. These programs and products are all useful technologies, but 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-users find the information they want, and learn more about the big picture, and the story behind the data.

References

ColorBrewer Color Advice for Maps (http://colorbrewer2.org)

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.

NASS Research, Science, and Technology website (http://www.nass.usda.gov/Research_and_Science/)

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.