Benchmarks Online

Research and Statistical Support
University of North Texas

RSS Matters

Link to the last RSS article here: [Stata 9: Introduction to Mata](#) - Ed.

Using Sparklines for Visual Representation, Part 1

By **Patrick McLeod**, ACS Research Consultant

Introduction

"A picture is worth a thousand words." Within the realm of science and social science, a picture, whether in the form of a graph or plot, can be the difference between confusing a reader and being able to provide your reader with a concise picture of your data, model and/or outcome.
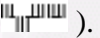
Even the most germane graph or plot is, by convention, divorced from the text in which is embedded. While the author or authors of an article, thesis or dissertation can attempt to frame visual representation with the most appropriate possible text, this can sometimes disrupt the flow of discussion and can create space problems for journal submissions. Do you exclude the graph or plot if you're operating under space constraints? Do you attempt to re-arrange an entire chapter of a thesis or dissertation in order to place a graph or plot in a more appropriate location?

For some categories of data, the solution lies with a tool called sparklines. Developed from concept through implementation by Dr. Edward Tufte (http://www.edwardtufte.com/tufte/), Professor Emeritus at Yale University, sparklines allow for the in-line representation of data in a document.

What Are Sparklines?

Tufte defines sparklines as "small, high-resolution graphics embedded in a context of words, numbers and images. Sparklines are data-intense, design-simple, word-sized graphics." Sparklines allow the researcher to display data in-line and, most importantly, in context.

Sparklines have different formats to match different types of data (from http://www.sparklines.org): Series data on the Federal deficit from the Office of

Management and Budget from 1983-2003 (▮▮▮▮▮▮▮▮▮▮▮▮ᵃ▮), the five-year daily closing stock price of Cisco (∿∿∿) or the postseason performance of the 2004 Boston Red Sox en route to their first World Series championship since 1918 (▮▮▮▮▮▮▮).

Implementing Sparklines

The Bundesliga (http://www.bundesliga.de/) is Germany's top-level professional football league. The Bundesliga season is 34 games long; while this seems like a small number of games given the length of the season (more like our MLB season than our NFL season), there are numerous other team and player commitments that run concurrently with the Bundesliga season: European competition for the more successful clubs (Champions League and UEFA Cup) and international competitions which halt domestic play since many players are absent from their clubs representing their respective nations.

As any fan of sports can attest, there are an overwhelming amount of numbers associated with your favorite player, your favorite team and the league in which your team plays. For example, if you wanted to know the record of every team in the Bundesliga from a particular season, you could simply represent it in a table like the following one (modified from http://www.bissantz.de/sparklines/demos/bundesliga/season_2004-2005.html):

P	Team	W	D	L	Goals	GD	TP
1.	Bayern MŸnchen	24	5	5	75:33	42	77
2.	FC Schalke 04	20	3	11	56:46	10	63
3.	Werder Bremen	18	5	11	68:37	31	59
4.	Hertha BSC Berlin	15	13	6	59:31	28	58
5.	VfB Stuttgart	17	7	10	54:40	14	58
6.	Bayer Leverkusen	16	9	9	65:44	21	57
7.	Borussia Dortmund	15	10	9	47:44	3	55
8.	Hamburger SV	16	3	15	55:50	5	51
9.	VfL Wolfsburg	15	3	16	49:51	-2	48
10.	Hannover 96	13	6	15	34:36	-2	45
11.	SV Mainz 05	12	7	15	50:55	-5	43

P=Place finished, Team=Team Name, W=Win, D=Draw, L=Loss, Goals (for:against), GD=Goal Differential, TP=Total Points

While this table provides us with the basic information we need to see how all the teams in the Bundesliga performed in the 2004-2005 season, wouldn't it be nice to be able to display a team's game by game performance in terms of wins and losses along with this information?

Thinking along conventional lines, displaying game by game performance in a tabular form would take up a ridiculous amount of room (34 games x 18 teams = 612 total games); in fact, so much room would be taken up that you will rarely find such a display outside of a sports almanac or season-in-review publication.

Implementing the same idea graphically using sparklines is possible in the space of a table. Note the in-line graphics under the column labeled "Course": A blue uptick is a win and a red downtick is a loss. A dash represents a draw. In a comparable amount of space to what it takes to present the team names, 612 data points of wins, losses and draws can be displayed in a meaningful graphic:

P	Team	W	D	L	Goals	GD	TP	Course
1.	Bayern MŸnchen	24	5	5	75:33	42	77	
2.	FC Schalke 04	20	3	11	56:46	10	63	
3.	Werder Bremen	18	5	11	68:37	31	59	
4.	Hertha BSC Berlin	15	13	6	59:31	28	58	
5.	VfB Stuttgart	17	7	10	54:40	14	58	
6.	Bayer Leverkusen	16	9	9	65:44	21	57	
7.	Borussia Dortmund	15	10	9	47:44	3	55	
8.	Hamburger SV	16	3	15	55:50	5	51	
9.	VfL Wolfsburg	15	3	16	49:51	-2	48	
10.	Hannover 96	13	6	15	34:36	-2	45	
11.	SV Mainz 05	12	7	15	50:55	-5	43	
12.	FC Kaiserslautern	12	6	16	43:52	-9	42	

In Part 2 of this article coming in the September 2005 edition of *Benchmarks Online*, I will cover the technical details of implementing sparklines in Microsoft Word, Microsoft Excel, LaTeX, HTML, PHP and Ruby. Below is a list of some links to information on Sparklines, including discussions on Edward Tufte's website and Sparklines.org.

Until next month, happy computing!

Links:

- Edward Tufte on Sparkline Implementation: http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0000Lk&topic_id=1

Edward Tufte on New Sparklines: http://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0001OR&topic_id=1

- Sparkline.Org: http://www.sparkline.org/

- Sparkline Wiki: http://sparkline.wikispaces.org/

Return to top