



Graph Designs for Reviewing Transactions and the Changing Balance

Stephen Few

September 12, 2006

This article features Xan Gregg's (of SAS Institute) winning solution for the second scenario in the 2006 Data Visualization Competition.

This article is the second in a five-part series that features the winning solutions to our 2006 Data Visualization Competition. The second scenario in our competition asked participants to create a visual display that would enable people to review their checking account activity.

Here's the scenario as it was described to participants:

You have been asked to design a visualization that can be used by bank customers to track their checking account activities during a single month. You want to make it just as easy for people to see the individual deposit and withdrawal transactions as well as the resulting balance. You also want to make it easy for them to see the precise dollar amounts of the transactions and balance whenever they wish.

In addition to this description, participants were given the following data in an Excel file.

March 2006 account activity
Beginning balance is \$4,218.33
All quantitative values are expressed in U.S. dollars

Date	Transaction Description	Amount
03/05/06	Mortgage payment	-2,580.19
03/13/06	Car loan payment	-305.00
03/15/06	Salary check	3,155.84
03/25/06	Student loan payment	-1,010.17
03/28/06	Household expenses	-1,748.52
03/29/06	Entertainment expenses	-107.35

While judging the many solutions that were submitted for this scenario, in addition to clarity of communication and ease of use, I was looking for a display that would allow people to see the following:

- Both the checking account transactions that occurred during the month as well as the changing balance
- The varying amounts of time that passed between the transactions and resulting balances as well as the points in time when they occurred

- Precise values (deposits, withdrawals, and the resulting balance) without cluttering the visual display
- The overall change that occurred between the beginning and end of the month

The Winning Solution

Xan Gregg of SAS Institute submitted the winning solution for this scenario, which he created using JMP, one of SAS' many products. JMP provides a rich collection of statistical functions, visualizations, and data interaction methods for analyzing quantitative information.

Take a moment to look at Xan's solution (Figure 1) to determine for yourself how well he succeeded in communicating the month's checking account activity.

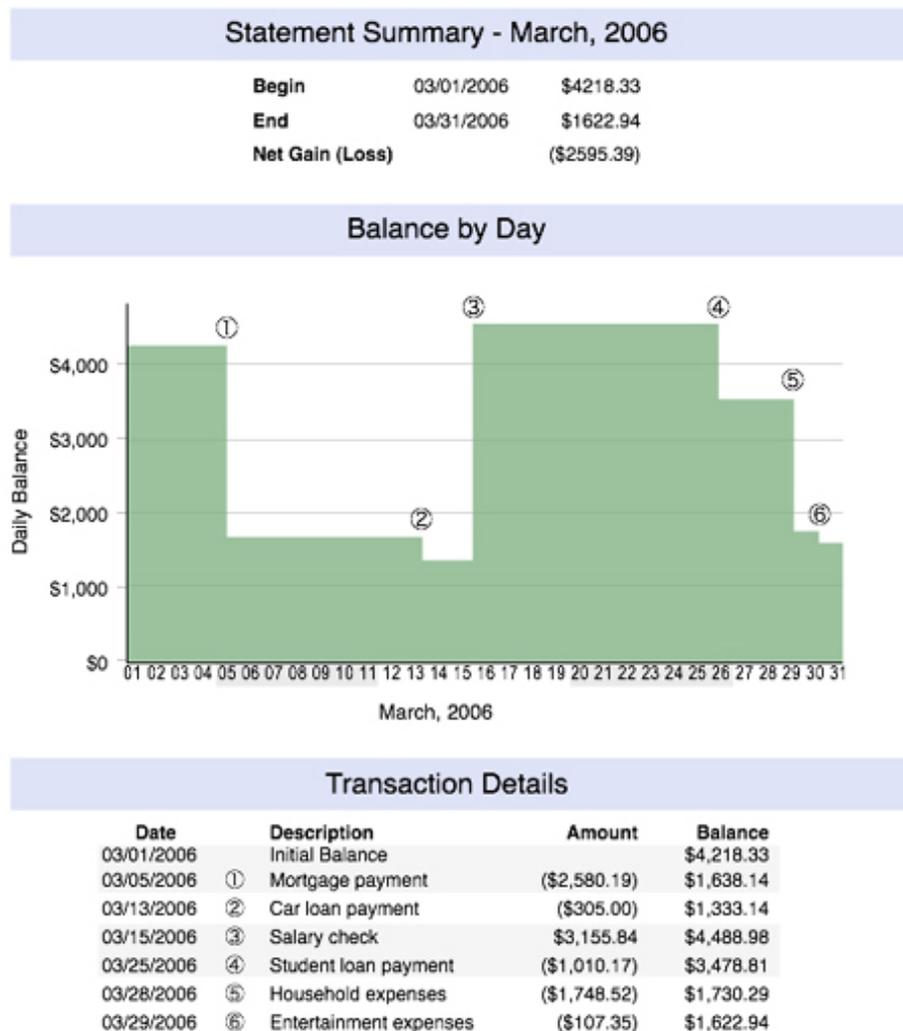


Figure 1: Xan Gregg of SAS Institute's winning solution.

Here are some of the highlights that made Xan's solution stand out:

- The first quality that I noticed was that this display is aesthetically pleasing. The soft colors and minimal visual means that he used to group the information and

assist the reader in scanning it (for example, the light gray background color in the table that assists the reader's eyes in scanning across the rows) work quite well.

- The organization of the information into summary, balance by day, and transaction details is simple, meaningful, and clear.
- It is very easy to track the account's balance as it goes up and down through the course of the month.
- It is also easy to see the duration of each balance amount and precisely when it changed.
- Precise transaction values, as well as the resulting balance, can be easily associated with the graphical display in a way that doesn't clutter or distract from the graphical pattern it reveals.
- The transaction table is nicely formatted (with one exception, which I'll mention below), making it easy to read and compare individual values. For example, by displaying a consistent number of digits for each part of the dates in the leftmost column (that is, always a two-digit month, two-digit day, and four-digit year), each part of a date (month, day, and year) is precisely aligned with the same parts of the other dates, which makes them easy to compare.
- Weeks during the course of March are subtly delineated in the graph by a light gray fill color behind the dates for every other week (Sunday through Saturday). This adds useful information without clutter.

In addition to learning from Xan's thoughtful design decisions, let's scrutinize his display to see if it could be improved. Given the purpose of this display, I believe a few design changes could make it even more effective.

- While it is extremely easy to see the changing balance, it is not equally easy to see the deposit and withdrawal transactions that caused these changes in the balance. More could be done to make it easier to focus on the individual transactions separately from the balance, and to clearly separate deposits from withdrawals. I'll show a design of my own that attempts to do this.
- There is an inconsistency in how numbers are displayed. Those that appear in the summary section do not include commas (for example, \$4218.33 rather than \$4,218.33), while those that appear in the transaction details section do. It is almost always useful to include commas to make numbers easy to read.
- In the transaction table, when parentheses are used to identify negative numbers, the entry is right-aligned to the parenthesis, rather than to the last digit of the number, which causes the reader's eyes to bounce back and forth slightly when scanning the numbers up or down through the column. This is a minor flaw, which could be easily corrected by right-aligning the numbers (excluding the right parenthesis) and allowing the right parenthesis to fall to the right of the number.

Other than the first possibility of improvement in this list, these items are quite minor. To illustrate how the deposit and withdrawal transactions might be featured with as much clarity

as the changing balance, I'll show you a design that I created prior to the competition (Figure 2). This is an example of a display that is sometimes called a *waterfall graph*, but one that I think does a better job than most waterfall graphs that I've seen.

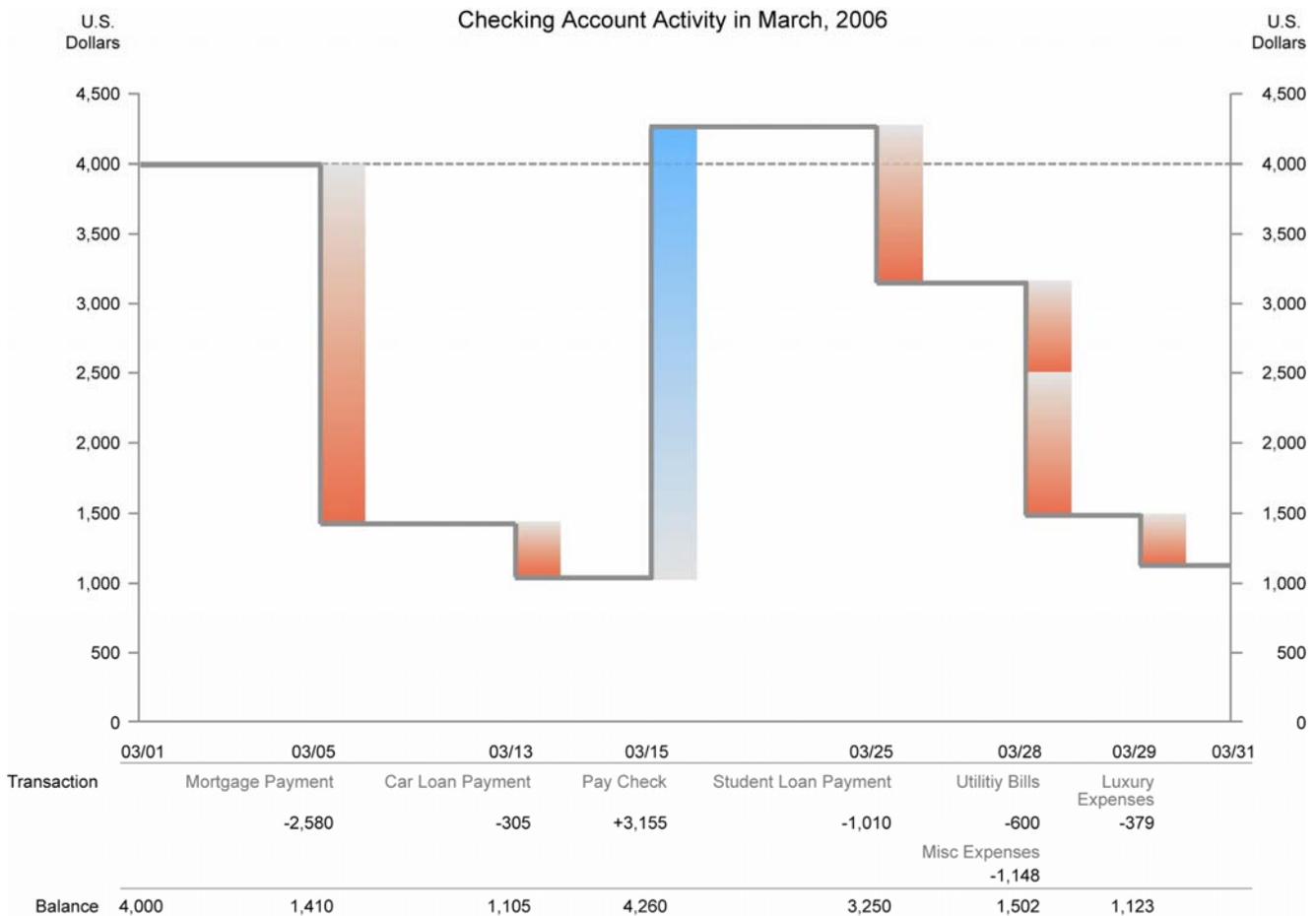


Figure 2: My version of a waterfall graph, designed to equally feature the transactions and the balance.

Notice how you can focus your attention on the balance, which is encoded as the thick gray line that moves up and down through time, but you can also shift your attention to the transactions, encoded as blue or red bars (blue for deposits and red for withdrawals). (I intentionally avoided the use of green and red bars, which are typically used to indicate positive and negative values, because green and red cannot be distinguished by those who are colorblind.) I advise against using changing gradients of color (from light to dark, lowly saturated to fully saturated) when they serve no useful purpose, but in this case, the gradients effectively draw the reader's eyes upward for deposits and downward for withdrawals, which is a meaningful application of this visual effect. One other feature of this display that might have added some value to Xan's solution is the dashed line, which makes it easy to see how activity throughout the month compared to the balance at the beginning of the month.

The way that the precise values in the table are closely associated with the graph works makes them easy to access without cluttering the graph. This method, however, could not be used for displays that are automatically generated by software from ongoing checking account activity, because when transactions are close to one another in time, their entries in

the table might overlap. An approach such as the one that Xan used would work much better for automated reports.

Solutions that Fell Short

Several of the solutions that were submitted exhibited particular problems that are worth noting so you'll know to avoid them. The first example (Figure 3) doesn't graph the values at all, so we aren't able to easily follow the ups and downs in checking account activity through the course of the month. Graphs excel in displaying trends, patterns, and exceptions in quantitative data. Tables don't provide this useful picture of the data. Figure 3 only uses one visual device to encode information, seen in its use of green and red colors to distinguish deposits from withdrawals at a glance. Unfortunately, the 10% of males and 1% of females who are colorblind, most of whom cannot distinguish red from green, would not be able to see this distinction. Notice also that gradients of color used to highlight each row serve only to give rows a pseudo-3-D appearance, causing them to jump out for no reason. The saturated reds and greens also make it harder to read the numbers in the table, as opposed to a light background such as white.

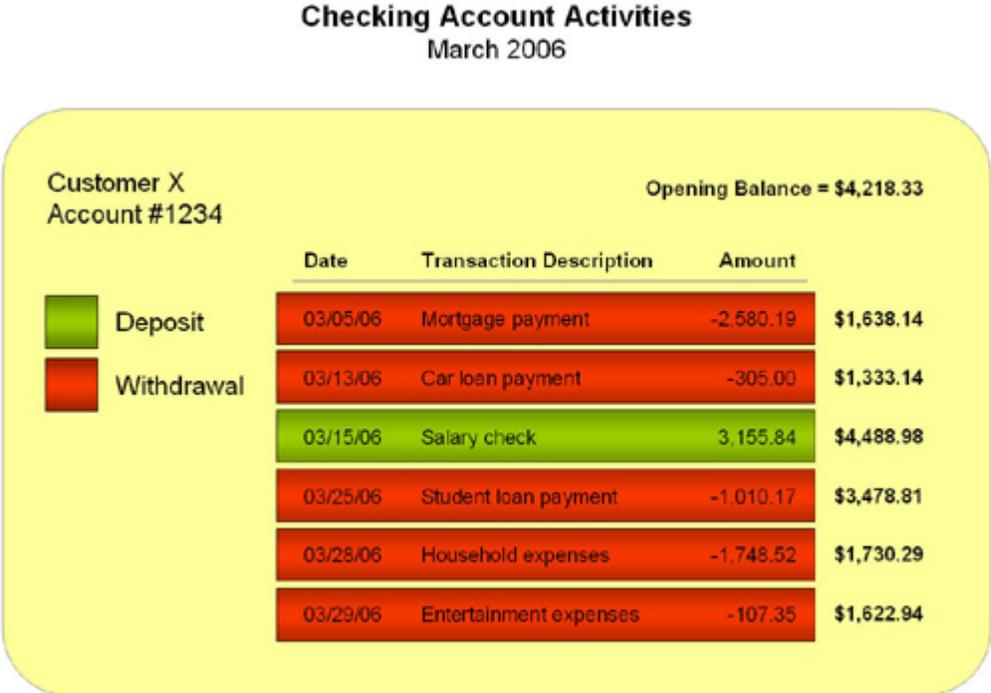


Figure 3: This display fails to visualize the changes in checking account values through time.

The next example (Figure 4) includes a graphical display, which the last one lacks, but it visually encodes the changing values in a way that suggests something that didn't really happen. Does anything about the graph bother you?

Checking Account Statement Ending March 31st, 2006

March 2006 Account Activity		
Beginning Balance Is		\$4,218.33
Date	Transaction Description	Amount
03/05/06	Mortgage Payment	-2,580.19
03/13/06	Car Loan Payment	-1,315.17
03/15/06	Salary Check	3,155.84
03/25/06	Student Loan Payment	-1,010.17
03/28/06	Household Expenses	-1,748.52
03/29/06	Entertainment Expenses	-107.35
Ending Balance Is		\$612.77

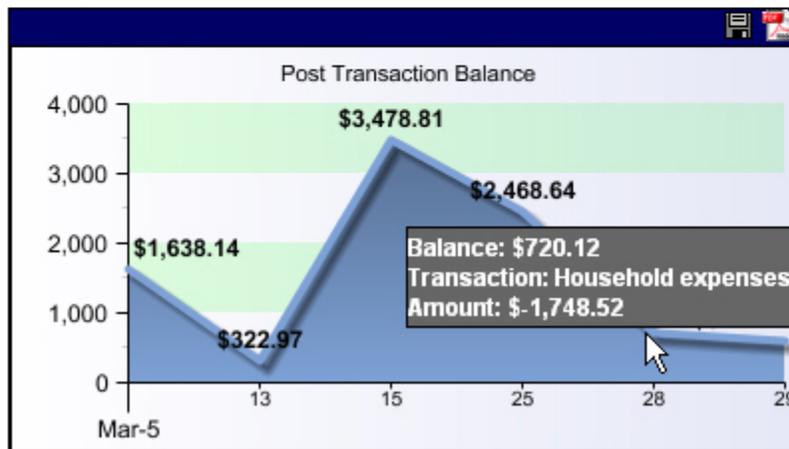


Figure 4: This example graphs the checking account balance in a way that fails to display the true nature of the changing balance.

Notice that this line graph, by directly connecting each balance values to the next in this way (that is, by drawing the line directly between the values to form a slope), suggests that the balance changed gradually over time from one value to the next. Line graphs usually directly connect one value to the next, and this usually works because the values are displayed for equal intervals of time, such as for each day in a case such as this when the balance should be shown for each day throughout the month. Xan's solution and my waterfall graph both display changes in the balance as an abrupt rise or fall from one value to the next on a given day, which is the true nature of this change, rather than a gradual change across several days. When changes between the values in a line graph are encoded as abrupt changes, the display is called a *step graph*.

I hope you've found this description of Xan's winning solution and my commentary useful, including the critique of design elements that failed to work effectively. Next month, I'll feature the winning solution to scenario #3 of our 2006 Data Visualization Competition, which takes on the challenging task of simultaneously displaying several variables related to house sales in a specific real estate market.

About the Author

Stephen Few has worked for over 20 years as an IT innovator, consultant, and teacher. Today, as Principal of the consultancy Perceptual Edge, Stephen focuses on data visualization for analyzing and communicating quantitative business information. He provides training and consulting services, writes the monthly *[Visual Business Intelligence Newsletter](#)*, speaks frequently at conferences, and teaches in the MBA program at the University of California, Berkeley. He is the author of two books: *Show Me the Numbers: Designing Tables and Graphs to Enlighten* and *Information Dashboard Design: The Effective Visual Communication of Data*. You can learn more about Stephen's work and access an entire [library](#) of articles at www.perceptualedge.com. Between articles, you can read Stephen's thoughts on the industry in his [blog](#).