

► BUILDING CAPACITY



Darryl Booth, MBA

Building Capacity With a New Visualization: The Sankey Diagram

Editor's Note: A need exists within environmental health agencies to increase their capacity to perform in an environment of diminishing resources. With limited resources and increasing demands, we need to seek new approaches to the practice of environmental health. Acutely aware of these challenges, the *Journal* publishes the Building Capacity column to educate, reinforce, and build on successes within the profession using technology to improve efficiency and extend the impact of environmental health agencies.

The conclusions of this column are those of the author(s) and do not necessarily represent the official position of the National Environmental Health Association, nor does it imply endorsement of any products, services, or resources mentioned.

Darryl Booth has been monitoring regulatory and data tracking needs of environmental public health agencies for over 20 years. He is the manager of U.S. operations for Hedgerow Software.

We should use our data to tell compelling stories that enhance understanding and promote informed perspectives. It is very easy to produce pie charts, bar graphs, and even some fancy scatter plots in Excel. Add some eye-popping colors, shadows, and rounded edges, and you have a beautiful visualization. Anybody can do that. Environmental health professionals can do better!

One of my favorite and most communicative types of visualization is known as a Sankey diagram, and it is not among the standard Excel charting options. You have to go out of your way, but I think Sankey diagrams are beautiful. They so clearly convey

how data are categorized and recategorized. In fact, I bet the following example is immediately illustrative.

A Simple Budget Sankey Diagram

Figure 1 is an approachable example that shows a fictitious personal budget of \$1,750. We can see \$1,500 in wages plus a \$250 side hustle. We then see how that money is allocated to taxes, housing, savings, and other expenses. The data flow left-to-right, but you can reverse that order when creating the diagram.

In this month's column, I will show you how to do a Sankey diagram with environmental health data.

How to Use SankeyMATIC for Sankey Diagrams

I recommend a free tool called SankeyMATIC (<https://sankeymatic.com>). Once you have looked around the home page, click Try SankeyMATIC (<https://sankeymatic.com/build>) to get started. No login is ever required.

A Sankey diagram shows flows and how amounts move from one category to another. To tell SankeyMATIC what to draw, you type out instructions. You dictate where the flow starts (FROM), how much flows (AMOUNT), and where it goes (TO).

The format is as follows:

FROM [AMOUNT] TO

For the budgeting example in Figure 1, the formatted input is typed as follows:

Wages [1500] Budget

Other [250] Budget

Budget [450] Taxes

Budget [420] Housing

Budget [400] Food

Budget [295] Transportation

Budget [35] Other Necessities

Budget [150] Savings

Can you see how Wages (1500) and Other (250) are categorized as Budget? Then, you see how the Budget is redivided into Taxes (450), Housing (420), and so on. Spacing doesn't matter.

An Environmental Health Example

We are going to build a simple Sankey diagram that shows how food complaints enter your system (complaint mode) and how they

FIGURE 1

Example of a Sankey Diagram for a Personal Budget

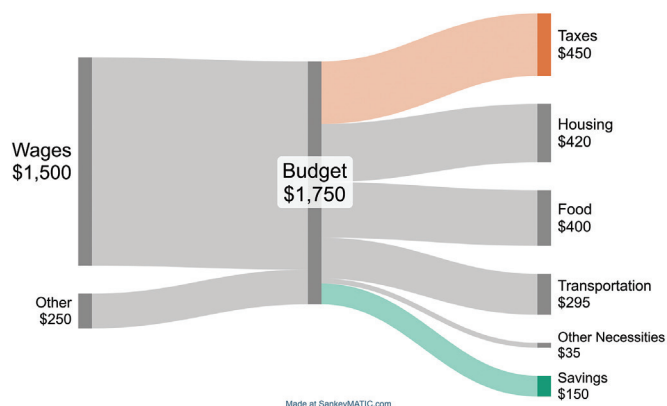


FIGURE 2

Example of a Sankey Diagram Showing How Food Complaints Enter a Health Department System

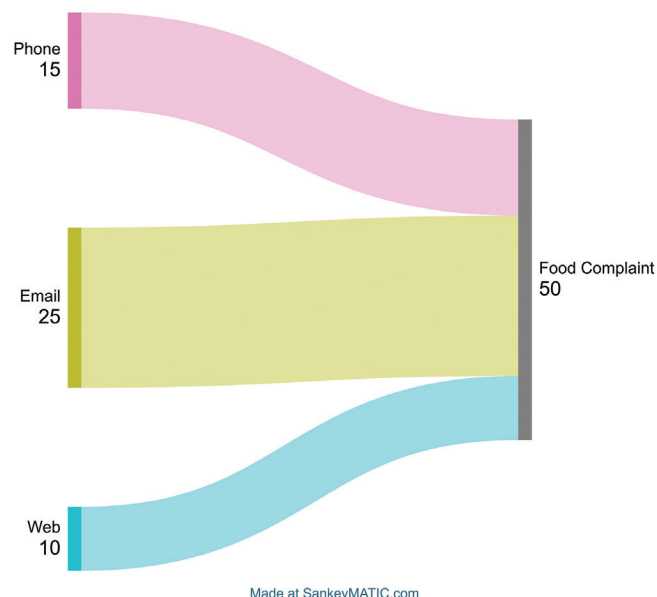


FIGURE 3

Example of a Sankey Diagram Showing How Food Complaints Enter a Health Department System and the Status of the Complaints

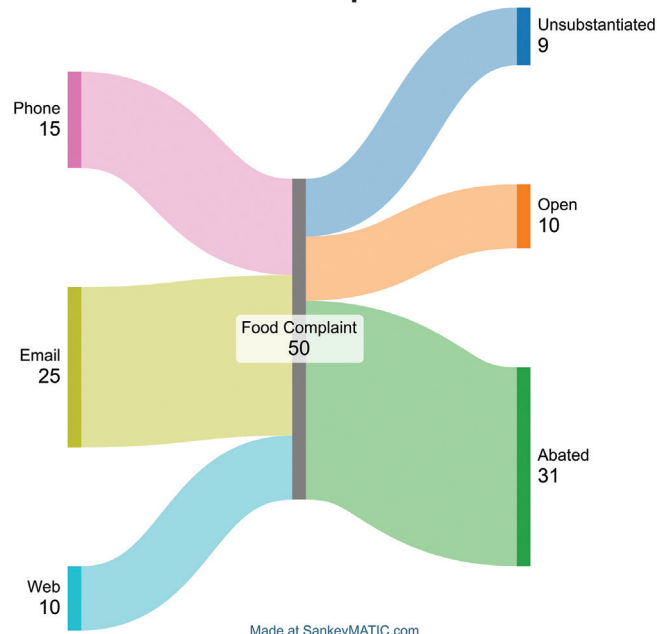
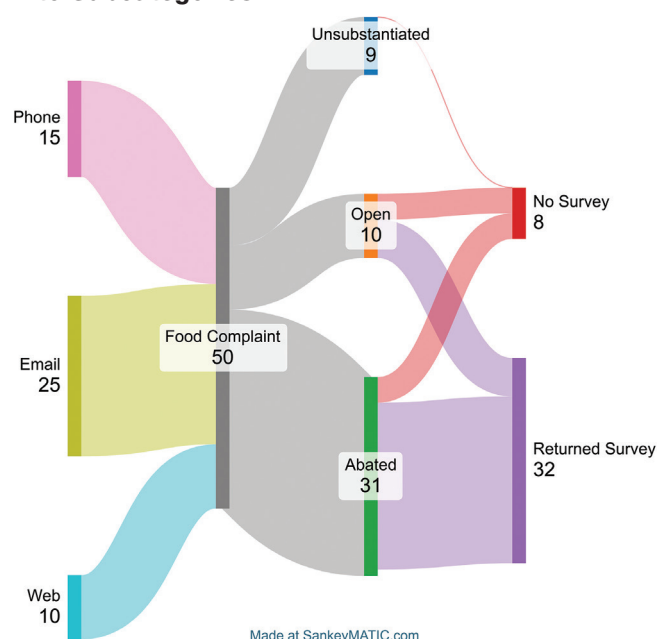


FIGURE 4

Example of a Sankey Diagram Showing How Food Complaints Enter a Health Department System and the Status of the Complaints Divided Into Subcategories



Use Sankey Diagrams in Your Work

Generate your own Sankey diagrams at no cost at:

- SankeyMATIC: <https://sankeymatic.com>

end up (disposition). I am going to start by pulling data from my data system, which could be a simple report or a filtered grid. In my system, I export complaints to CSV and open Excel, where filtering is very easy.

First, I filter for Telephone and Food Complaint, and I count 15. Next, I filter for Emailed Food Complaints, and I get 25. Lastly, I filter for Web Form Food Complaints, and I get just 10.

In the INPUTS section of the Sankey-MATIC form, I enter:

```
Phone [15] Food Complaint
Email [25] Food Complaint
Web [10] Food Complaint
```

With just this much data, I am already telling a story. Notice how the graph updates as you provide each row (Figure 2).

Now, to add the disposition, I return to my data system and filter/count again. First, I filter for Food Complaint, Unsubstantiated, and I count 9. Next, I filter for Food Complaints, Open, and get just 10. And then I filter for Food Complaints, Abated, and get 31.

Adding to what I keyed before, the entire INPUTS section now looks like the following:

```
Phone [15] Food Complaint
Email [25] Food Complaint
Web [10] Food Complaint

Food Complaint [9]
Unsubstantiated
Food Complaint [10] Open
Food Complaint [31] Abated
```

Figure 3 is now what our diagram looks like with the new set of data.

From here, I encourage you to play with the other settings, including the controls for color, labels, and much more. For numeric diagrams, find Labels > Number Format > Prefix to put a dollar sign before your numbers.

The result can be saved as an image when you hit Save as PNG (at the top). Choose Huge or SVG for the highest quality print versions. Your work can be saved with the Save My Work and Load From File buttons.

Where to Go From Here?

Well, it should not surprise you that the Sankey diagram lets you go further. You can further subdivide any category. Figure 4 is an example that includes whether the complainant returned a customer satisfaction survey. I just added five more statements to the bottom of our previous INPUTS section:

```
Unsubstantiated [0] No Survey

Open [6] Returned Survey
Open [4] No Survey

Abated [26] Returned Survey
Abated [4] No Survey
```

Try it and see how your chart changes.

I should include a few caveats. A Sankey diagram is not automatically the best for every type of data. To my thinking, if you are considering a pie chart, you can consider a Sankey diagram. Also, the numbers must balance. If you fail to resolve every number to a category, the website will show an out-of-balance error. But you can sharpen your pencil and fix that.

At the end of the day, this process is all about telling your story. The story can be about

money and budgets, violations, workload, or plan check outcomes. These areas each represent data trapped in your data system that are just waiting to tell a good story. 🌸

Postscript

For those individuals with an interest, this charting method was developed by Matthew Henry Phineas Sankey. He developed the technique to show how energy developed by a steam engine was put to use. Check out https://en.wikipedia.org/wiki/Sankey_diagram to learn more.

Corresponding Author: Darryl Booth, U.S. Operations Manager, Hedgerow Software. Email: darrylbooth@hedgerowsoftware.com

CALL FOR SUBMISSIONS

The *Journal* seeks guest authors for the Building Capacity column. Our goal is to provide a platform to share capacity building successes occurring across the country and within different sectors of the environmental health profession, including academia, private industry, and state, local, tribal, and territorial health agencies. Submissions will be reviewed by the NEHA technical advisors for data and technology and *Journal* staff for appropriate content, relevance, and adherence to submission guidelines. To learn more about the submission process and guidelines, please visit www.neha.org/contributors.

Did You Know?

The NEHA/AAS Scholarship is available to undergraduate and graduate students enrolled in an accredited college or university with a dedicated and recognized curriculum in environmental health sciences. The scholarships encourage an early commitment by students to pursue a career in environmental health. Further, the scholarships are meant to inspire past and present graduates to pursue postgraduate studies in environmental health science. Your donation empowers us to continuously award these scholarships to exceptional students who have a passion for environmental health. Every dollar donated supports the next generation of environmental health professionals. Learn more at www.neha.org/donate.