

Data Visualization

Presentation for DOE Energy Innovator Fellows

Bentham Paulos, Affiliate, Berkeley Lab

June 5, 2025

This work was funded by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.



Introduction and Theory



What is data visualization?

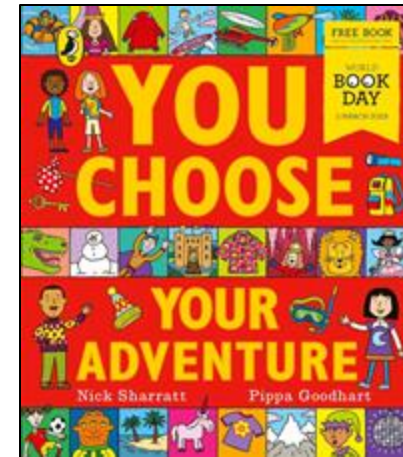
Data viz is storytelling.

What is my story? What do I want to say with the data?

Anticipating what the user wants to know.

Too much detail gets in the way of the story.

Or, data viz is “choose your adventure.”



Theory of data visualization

Who – what – how

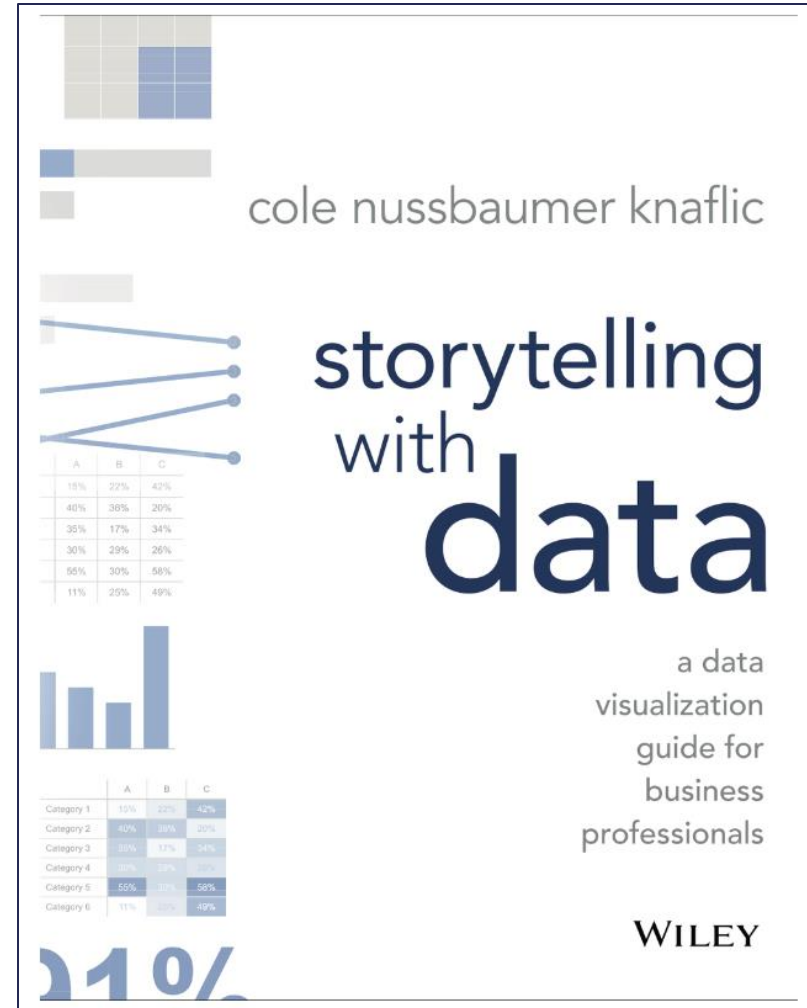
Who is your audience?

What do you want them to know?

How can you best reach them?

Explanatory vs. Exploratory

Do you want to explain specific findings, or let the audience interrogate your data with their own questions?



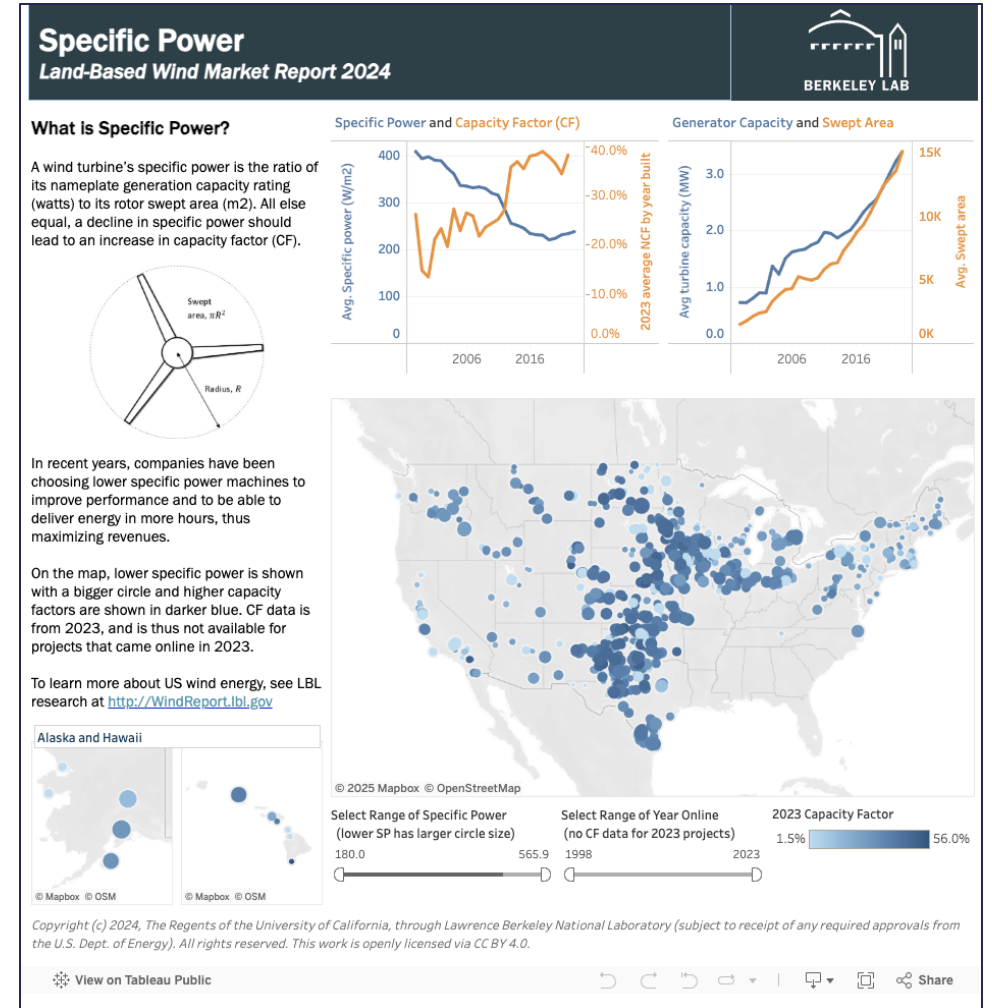
Telling a specific story (explanatory)

Berkeley Lab wind data on “[Specific Power](#)”

Explains a little-known technical concept

Uses text, illustration, map, and charts

Filters affect the maps



The user asks questions (exploratory)

EIA [Electricity Data Browser](#)

- Way too much information to tell one story – choose your own!
- Many choices and filters
- Formats: time series, columns, maps

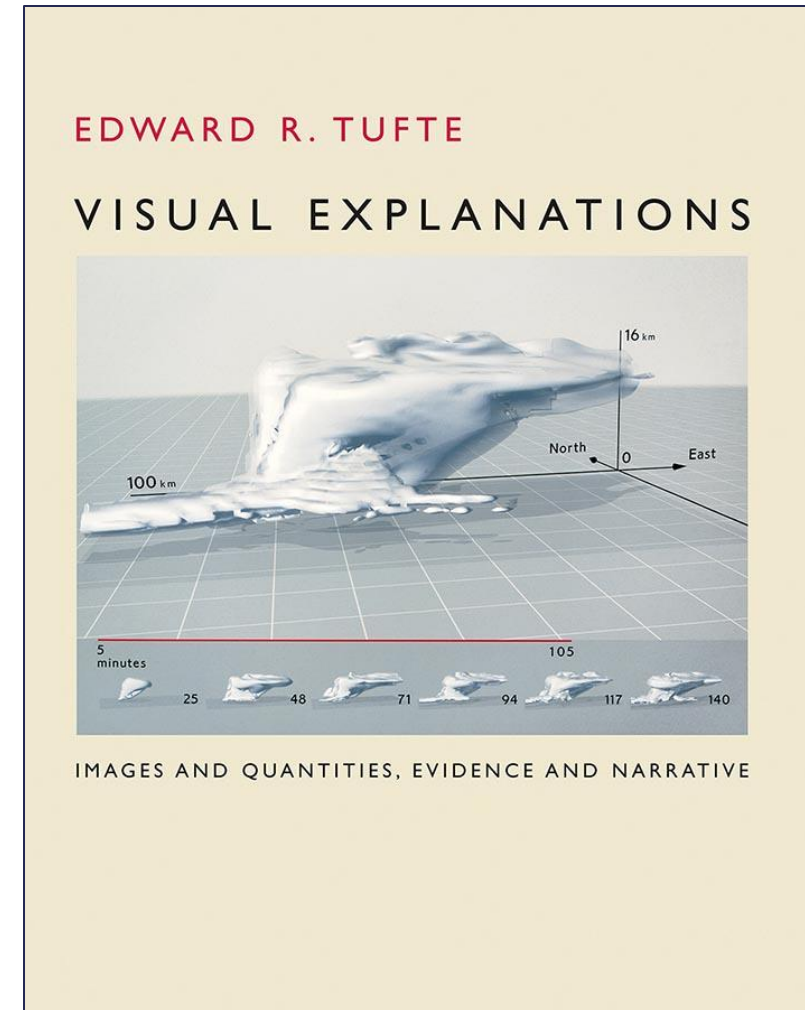
The screenshot shows the 'Electricity Data Browser' interface. At the top, there's a navigation bar with '< Electricity Data' and the title 'Electricity Data Browser'. Below this is a blue banner with the text 'Explore the new Beta version with expanded plant level data for water cooling and emissions.' The main content area is divided into two sections: 'Change data set' and 'View a pre-generated report'. The 'Change data set' section has a dropdown menu with 'Total consumption' selected. The 'View a pre-generated report' section has a dropdown menu with 'Choose a report' selected. Below these sections is a map of the United States with a legend on the right. The legend is titled 'Map controls & legend' and includes 'VALUE OPTIONS: Values Growth', a date selector for 'Feb 2022', 'Play' and 'Reset' buttons, and a color scale for 'thousand Mcf' ranging from 0 to 195,334. A 'Value is not available' legend is also present.

This screenshot shows the filter options in the Electricity Data Browser. It is divided into three columns: 'ORDER BY (choose one)', 'THEN BY (choose one or more)', and 'THEN BY (choose one or more)'. The 'ORDER BY' column has 'FUEL TYPE' selected, with options for 'Coal', 'Petroleum liquids', and 'Petroleum gas'. The first 'THEN BY' column has 'SECTOR' selected, with options for 'All sectors', 'Electric power', 'Electric utility', 'Independent power producers', 'Electric utility non-cogen', and 'Electric utility cogen'. The second 'THEN BY' column has 'GEOGRAPHY' selected, with options for 'United States', 'New England', 'Connecticut', 'Maine', 'Massachusetts', 'New Hampshire', 'Rhode Island', 'Vermont', 'Middle Atlantic', 'New Jersey', 'New York', 'Pennsylvania', 'East North Central', 'Illinois', 'Indiana', 'Michigan', and 'Ohio'. There are 'Check all' and 'Uncheck all' buttons for each column. At the top right of the filter section are 'SUBMIT', 'CANCEL', and 'X' buttons. Below the filter section is a navigation bar with 'FILTER/ORDER', 'ANALYSIS', 'Time-series', 'Column', and 'Map' tabs. There are also 'HELP' and 'DOWNLOAD' buttons. At the bottom of the filter section is a time series selector with 'Jan 2001' and 'Mar 2025' dates, and 'Annual', 'Quarterly', and 'Monthly' format options.

This screenshot shows the 'Change data set' and 'View a pre-generated report' sections of the Electricity Data Browser. The 'Change data set' section has a dropdown menu with 'Net generation' selected. The 'View a pre-generated report' section has a dropdown menu with 'Choose a report' selected. Below these sections is a navigation bar with 'FILTER/ORDER', 'ANALYSIS', 'Time-series', 'Column', and 'Map' tabs. There are also 'HELP' and 'DOWNLOAD' buttons. At the bottom of the navigation bar is a time series selector with 'Jan 2001' and 'Mar 2025' dates, and 'Annual', 'Quarterly', and 'Monthly' format options.

Basic rules of data viz (Edward Tufte)

- 🔗 Use standard units and scales within a figure
- 🔗 Be as simple as possible, “use ink for data not decoration,” in the least space
- 🔗 Put data in context
- 🔗 Use color as a narrative element
- 🔗 Use the right amount of data to tell your story
- 🔗 Don't distort data – “Graphics reveal data”



Making Effective Choices



Making effective choices

Charts & graphs

Maps

Dashboards



Charts & Graphs: Simple

Line charts

Bar charts

Scatter plots

Pie charts

Column charts

Treemap charts

Heatmap charts

Pareto charts

Geo charts

Waterfall charts

Donut chart

Funnel chart

Bubble chart

Histogram

Candlestick chart

Area chart

KPI charts

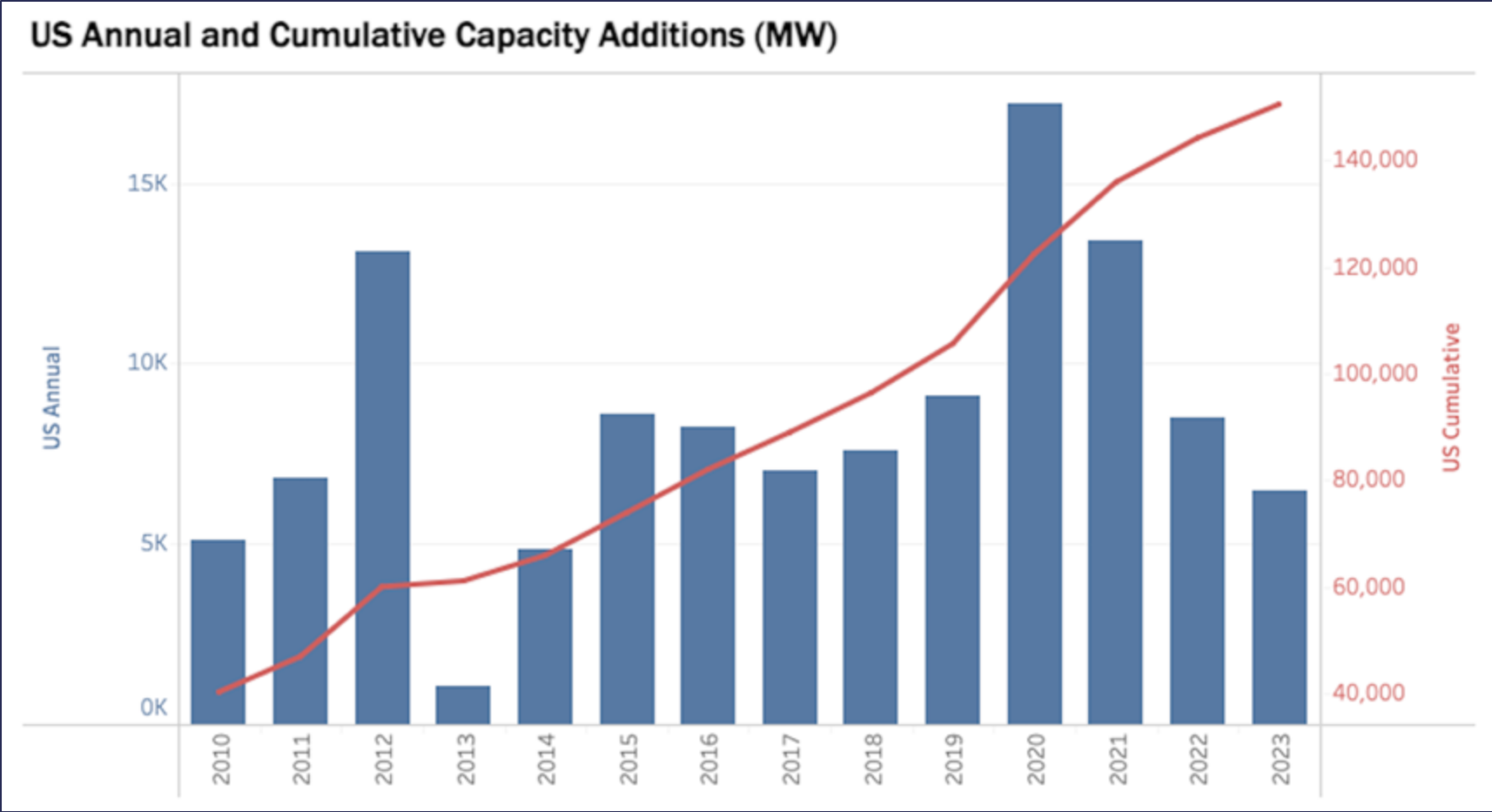
Scatter maps

Sankey charts

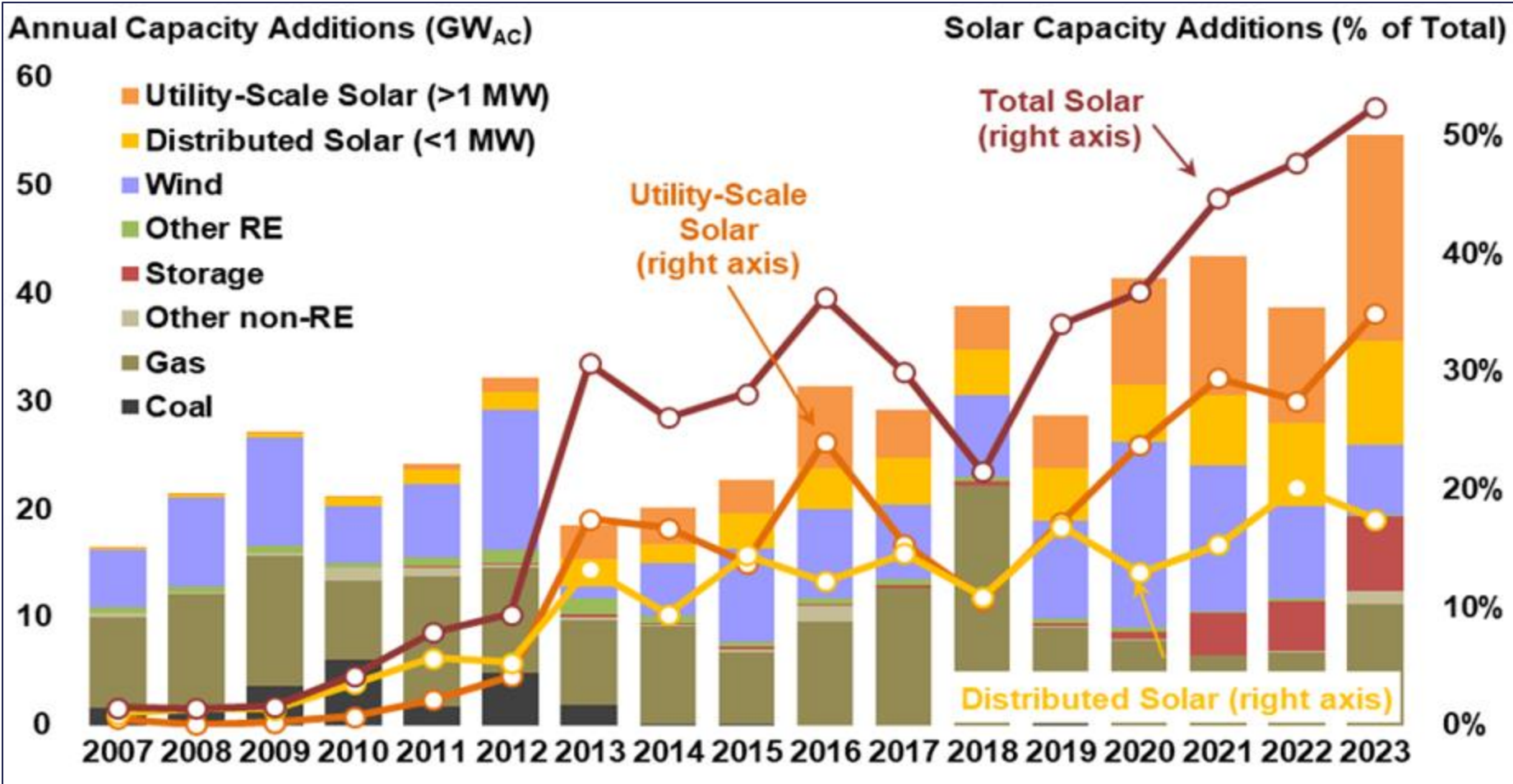
Radar charts



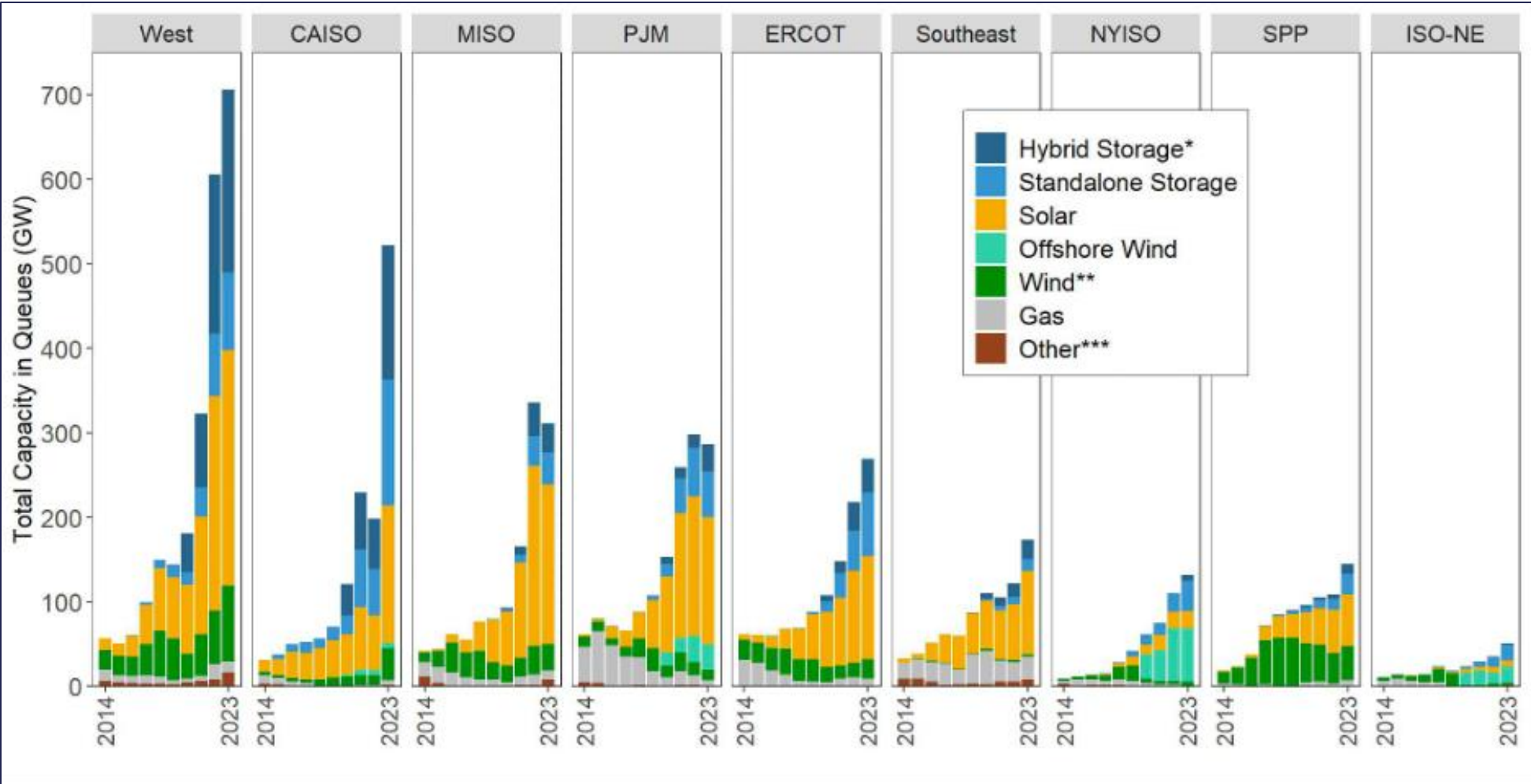
Charts & Graphs: Multidimensional



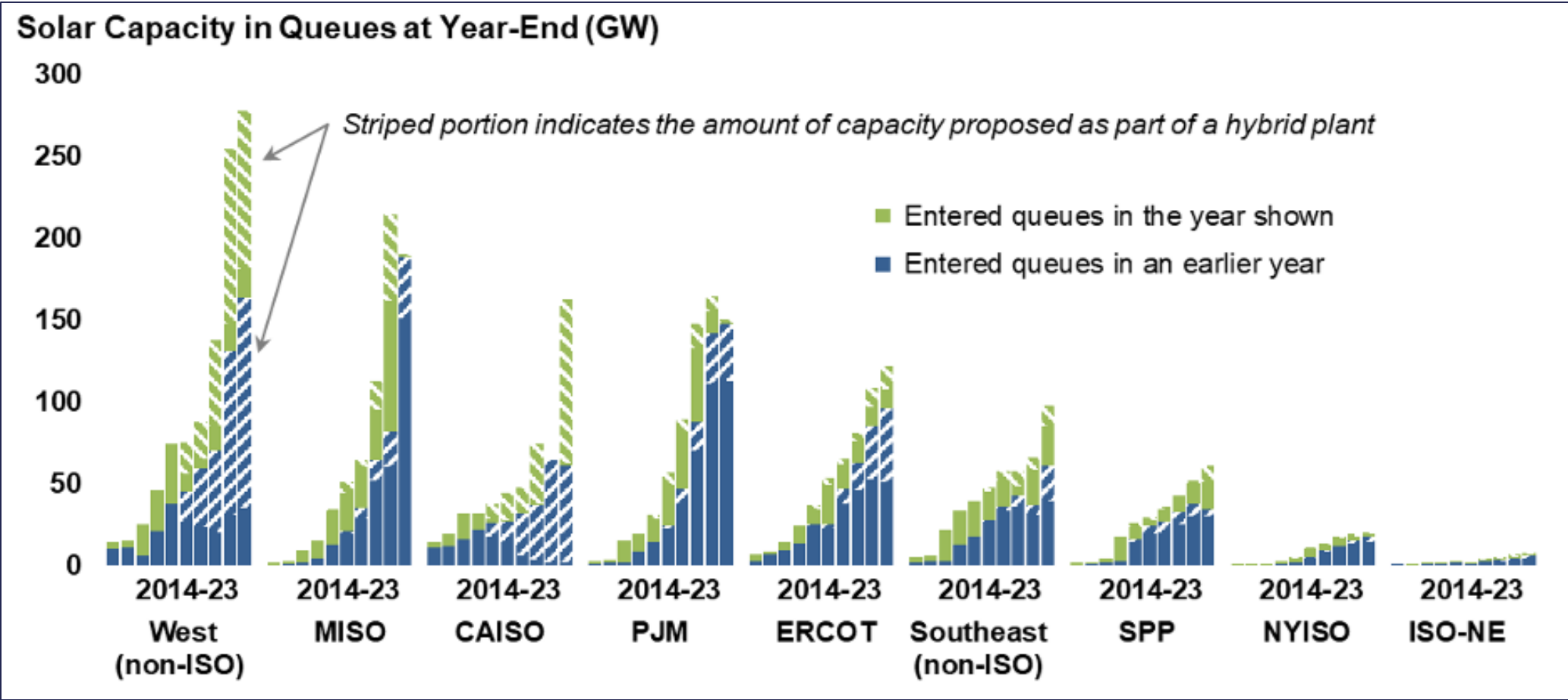
Charts & Graphs: Very Multidimensional



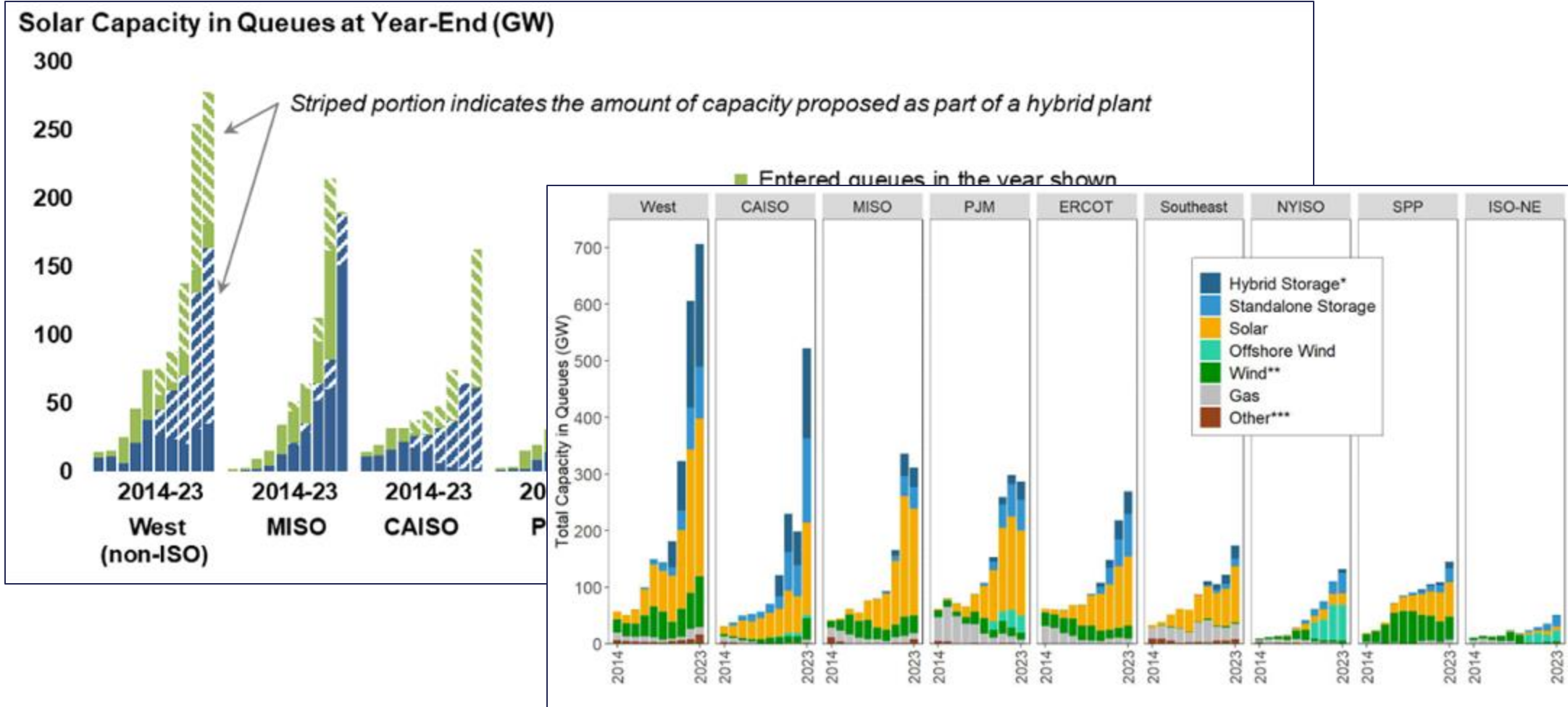
Use of color and texture



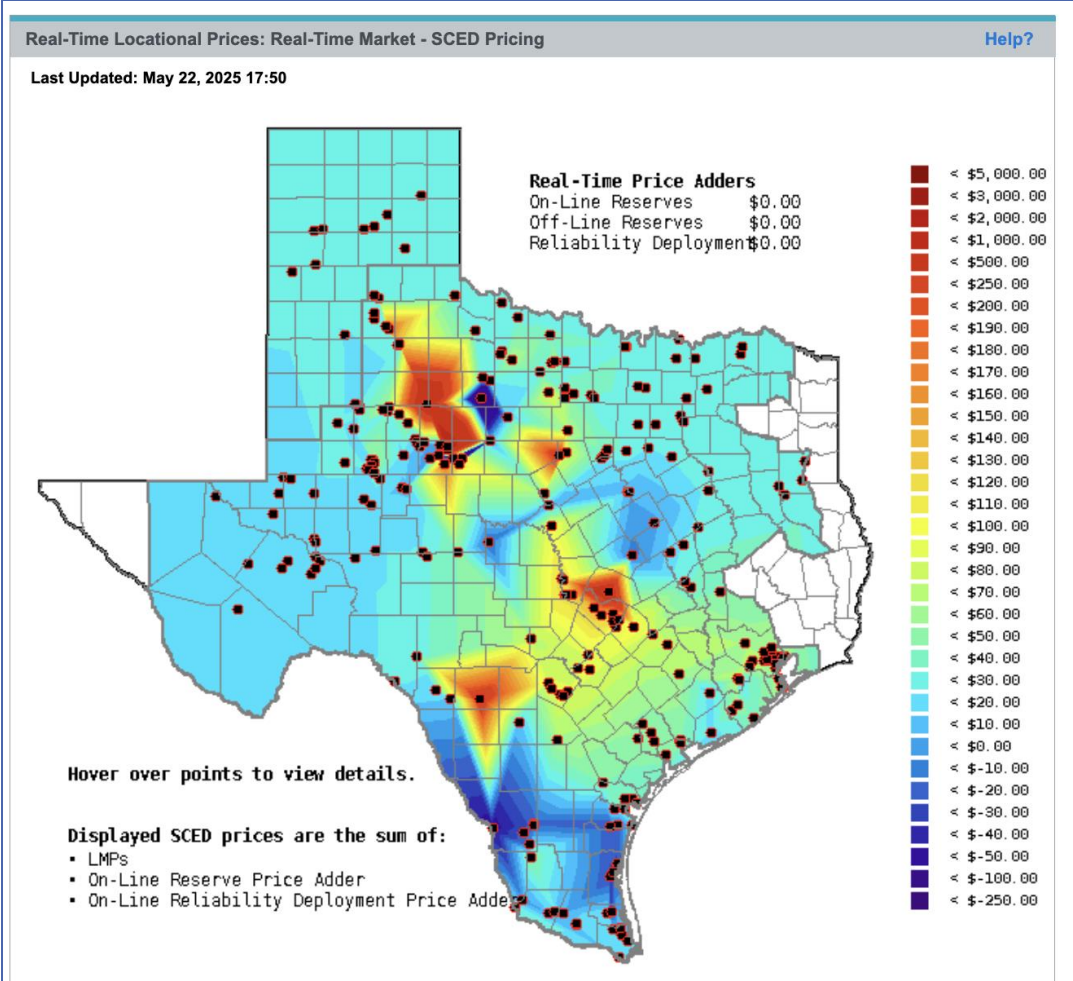
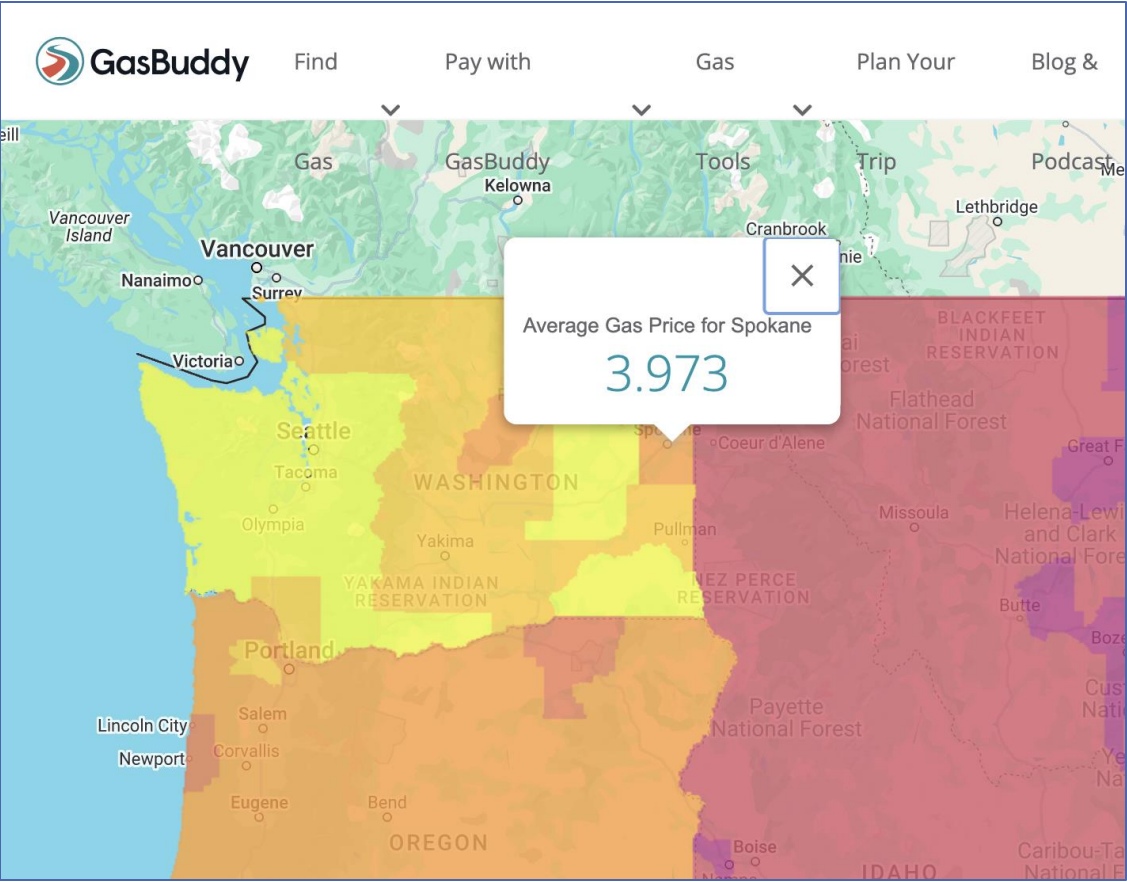
Use of color and texture



Use of color and texture

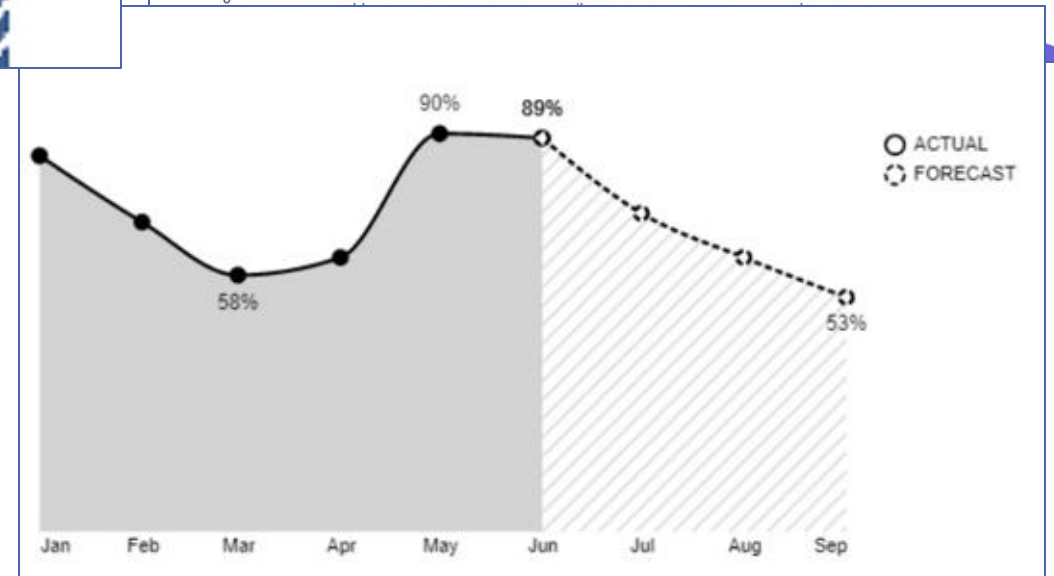
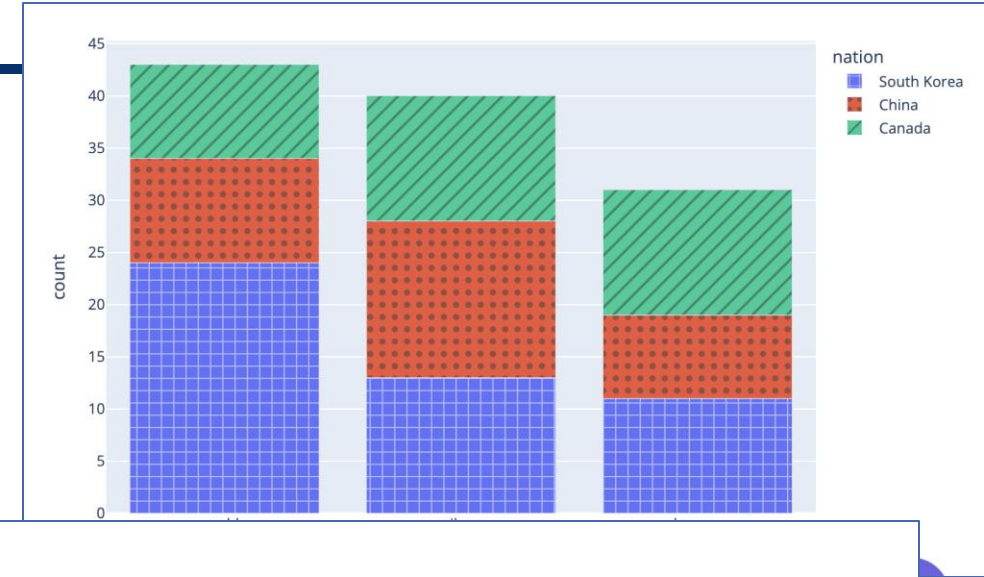
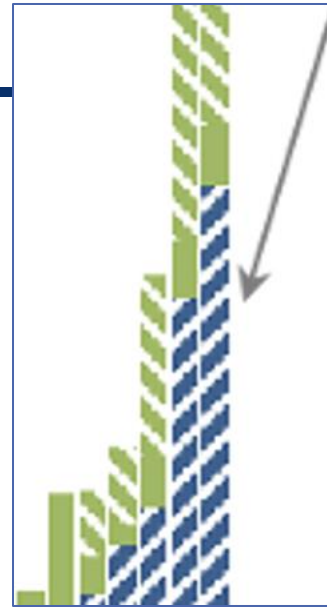


Many more options... Heatmaps

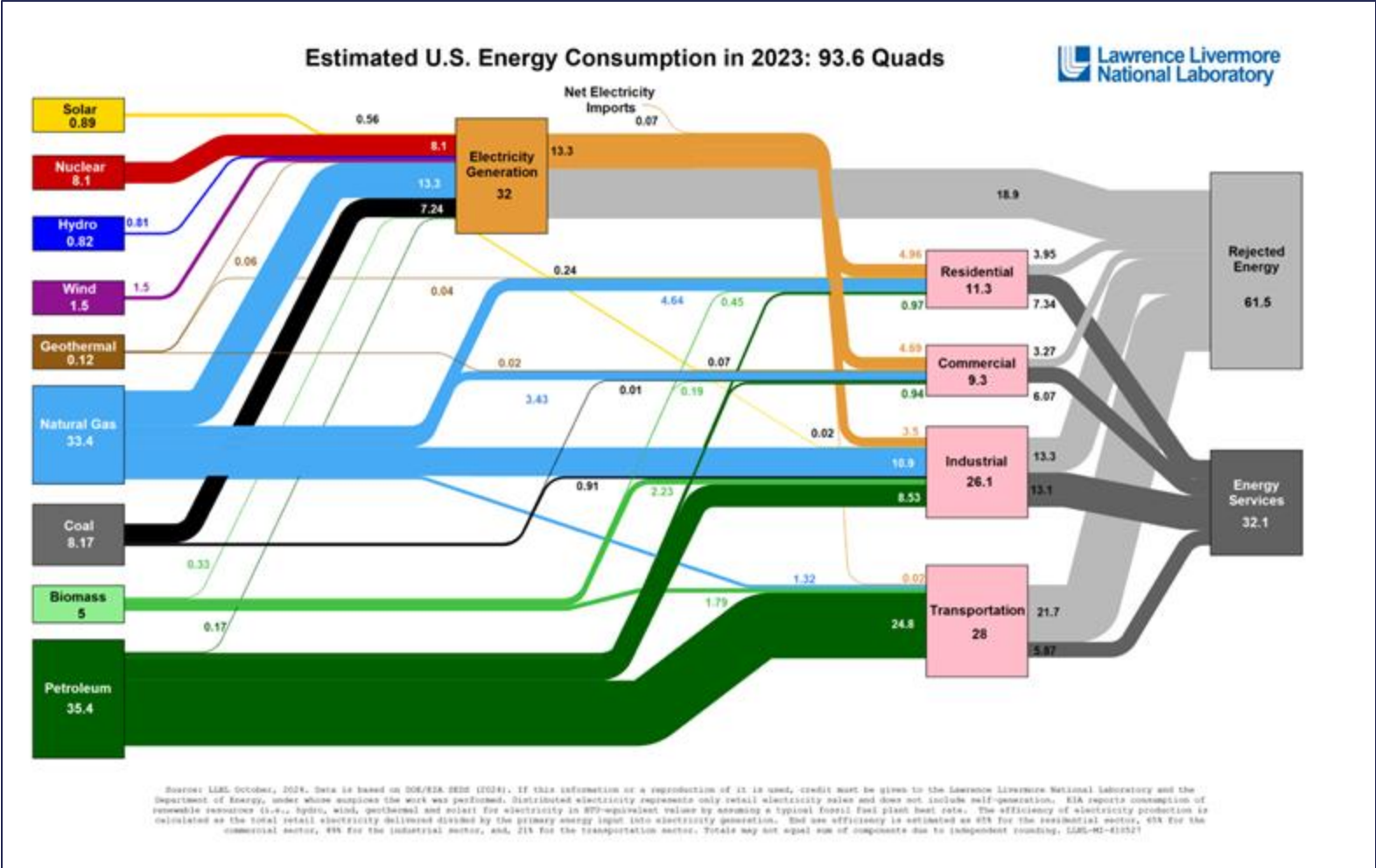


Be careful about texture

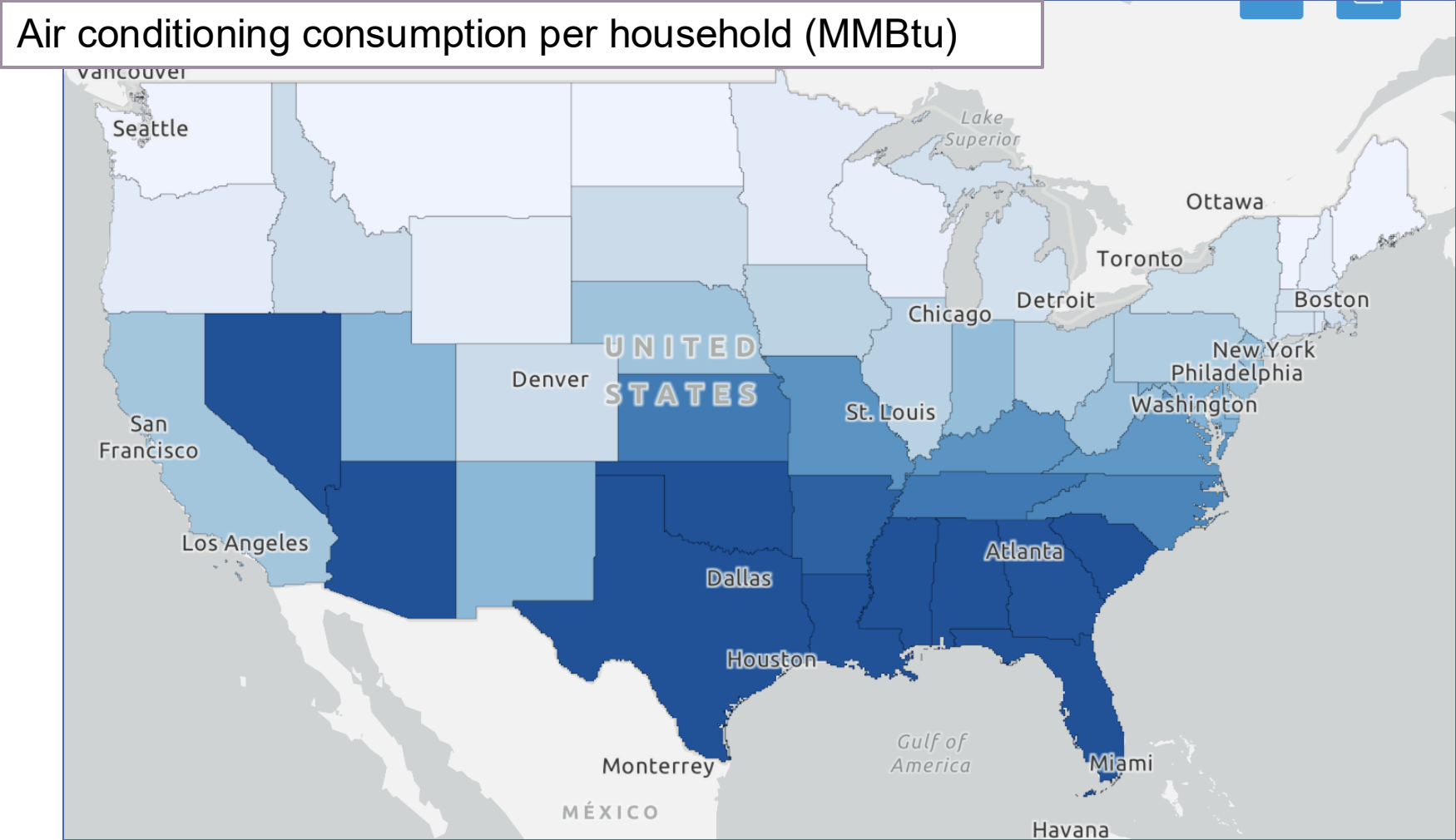
- Can be very messy, harder to read, create optical illusions
- Especially bad for small or crowded figures
- Color is almost always better, especially for on-screen display
- Some good uses
 - + Showing the difference between actual and forecast
 - + Binary things: data points that are in or out of a set



LLNL's Sankey diagram (flowcharts)

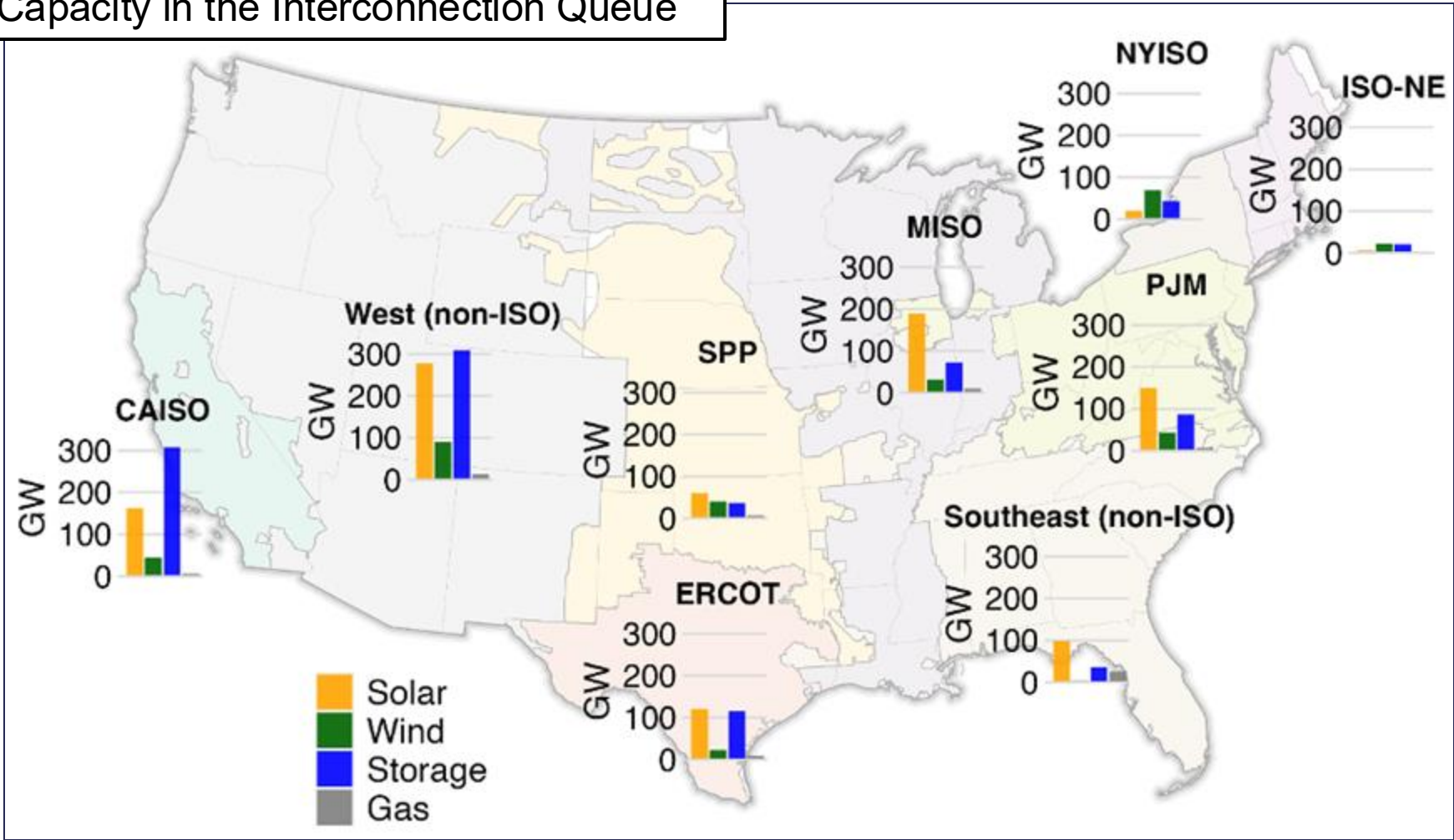


Maps: simple

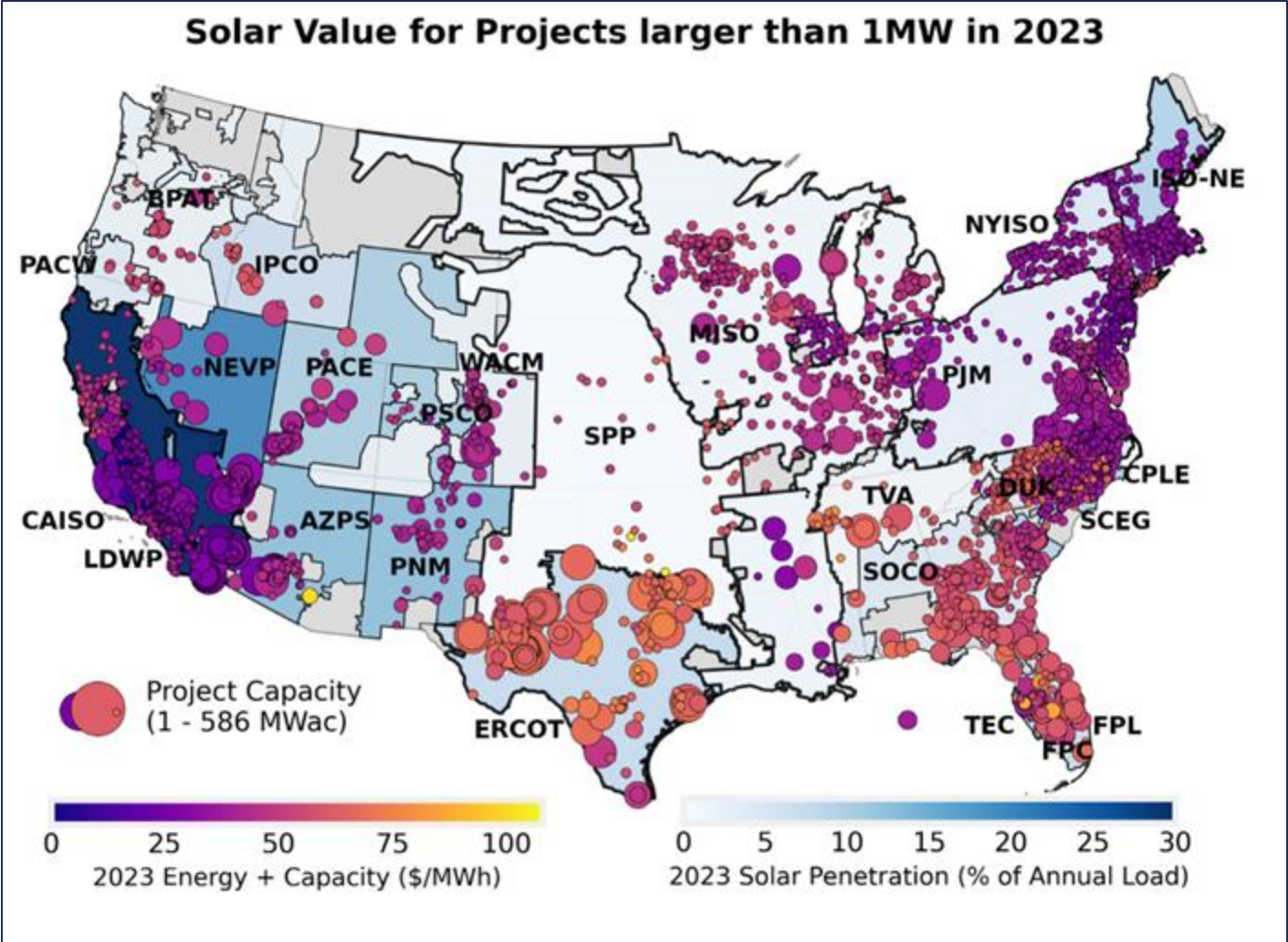


Maps: multi-dimensional

Capacity in the Interconnection Queue



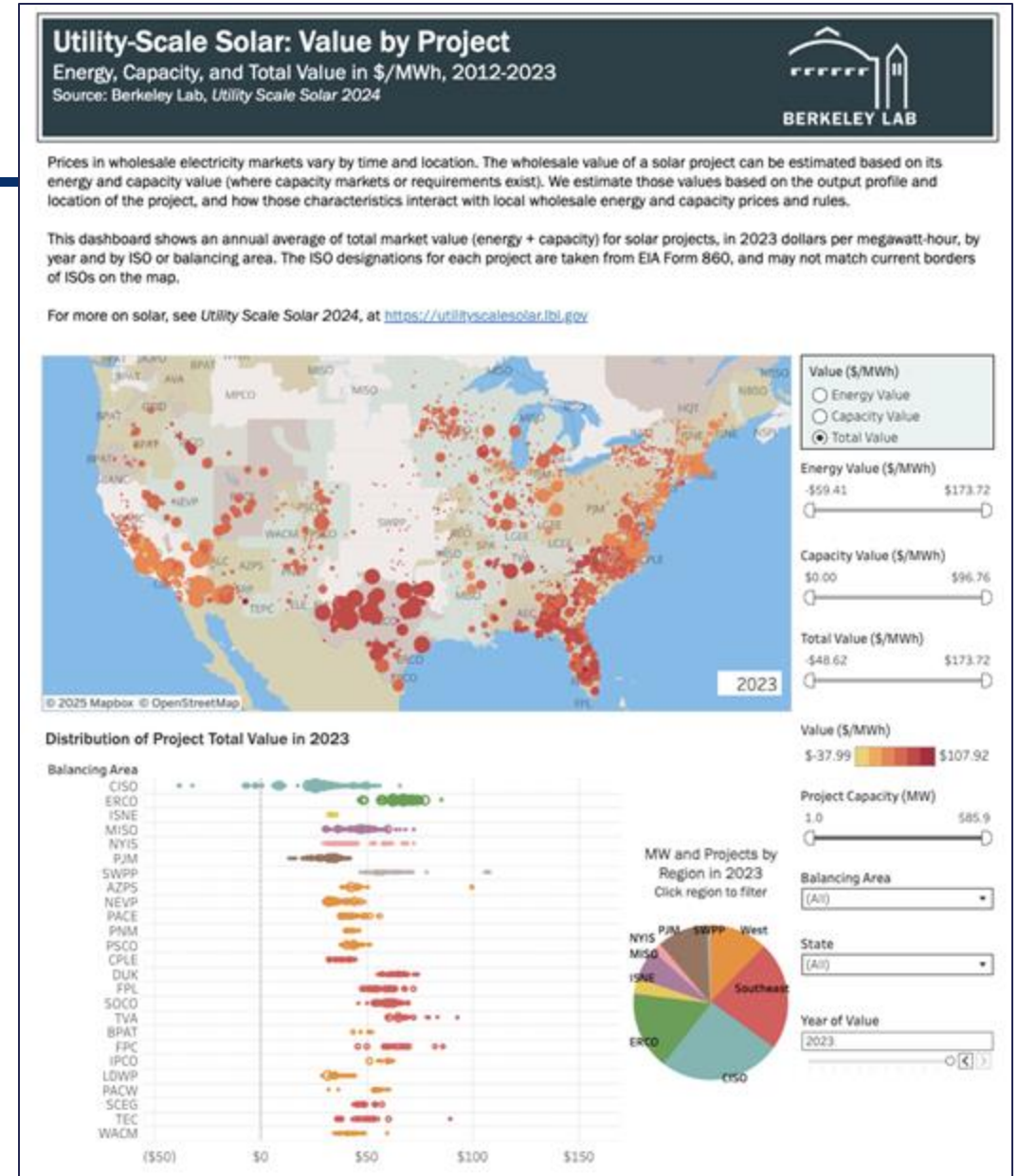
Maps: very multi-dimensional



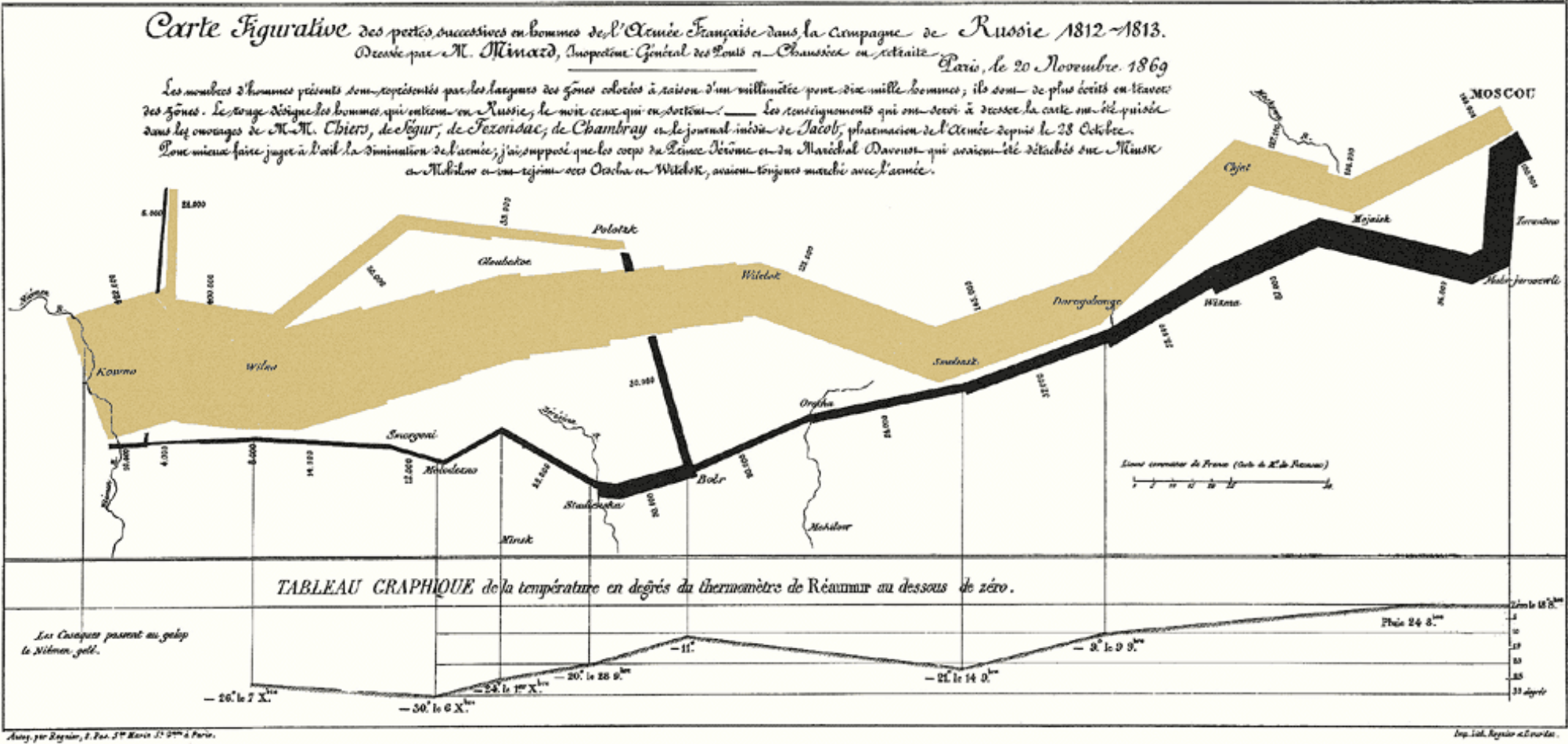
Combining Maps and Graphs

Berkeley Lab [solar value](#)

- 🔗 Filters affect both map and charts
- 🔗 Can select by energy, capacity, total value, project size, geographic area, and year
- 🔗 Can tell different stories to different users



The classic combination: Napoleon's march on Moscow



Dashboards

For tracking progress or showing current status

Often static, without filters

Often updated automatically through API (application programming interface)



Interactivity



Interactivity: when, why, and how

Exploratory - Let your user tell their own stories

Explanatory - Tell multiple stories using filters

Key features: Filters, mouseovers, data downloads

Not good for print. Okay for presentations. Best for online.

Can allow for greater complexity



Key aspects of interactivity

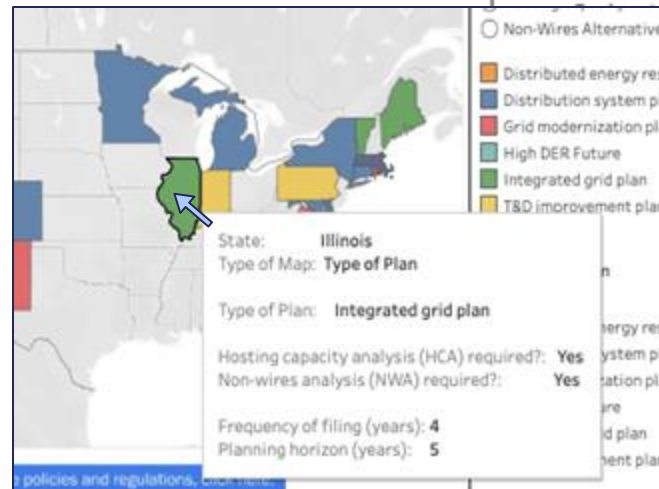
Filters – select a subset of data to tell a more specific story ([EV/gas tool](#))

Mouseovers – pop-up of information about that data point

Linked figures - Maps & charts change together

URL links - link to sources

Downloads – images, crosstabs, data

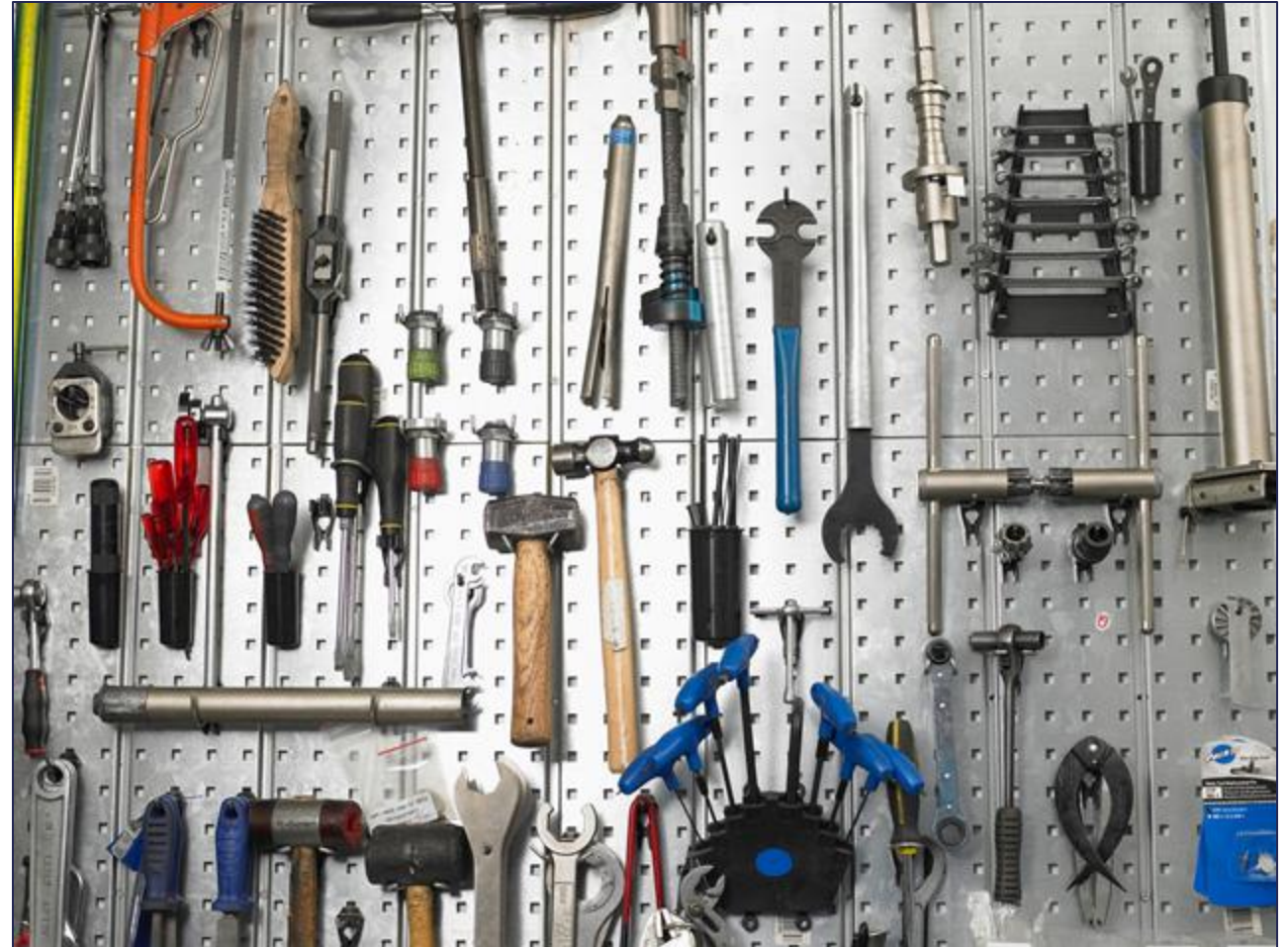


Overview of Tools



Tools

1. Data analysis
2. Data visualization
3. Maps
4. Graphics
5. Energy data
6. Energy photos



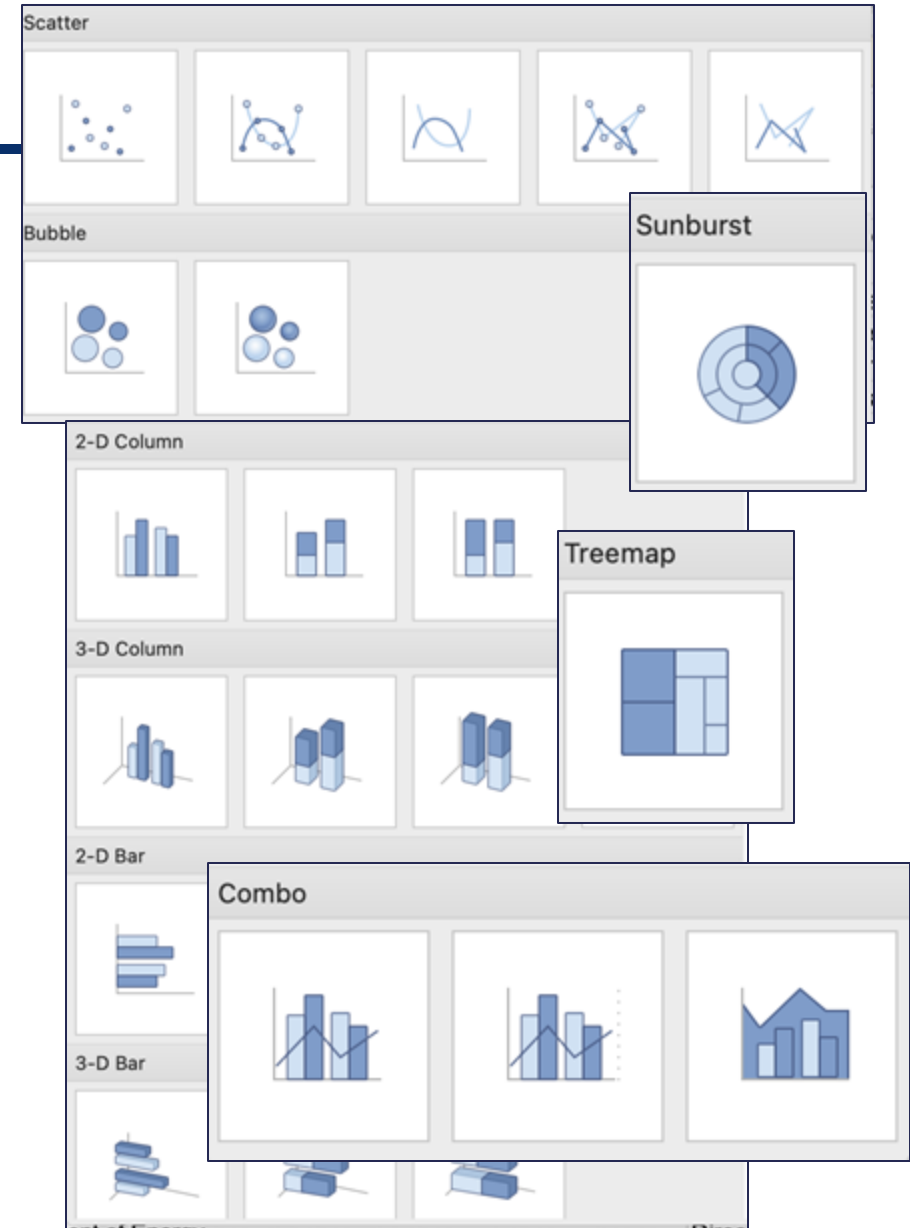
Tools: Data analysis

Spreadsheet software

- 🔗 MS Excel and Google Sheets
- 🔗 Endless tutorials, classes, how-to videos online, plus books, etc.

Coding languages & packages

- 🔗 R and Python
- 🔗 SAS, SPSS, Stata
- 🔗 GitHub

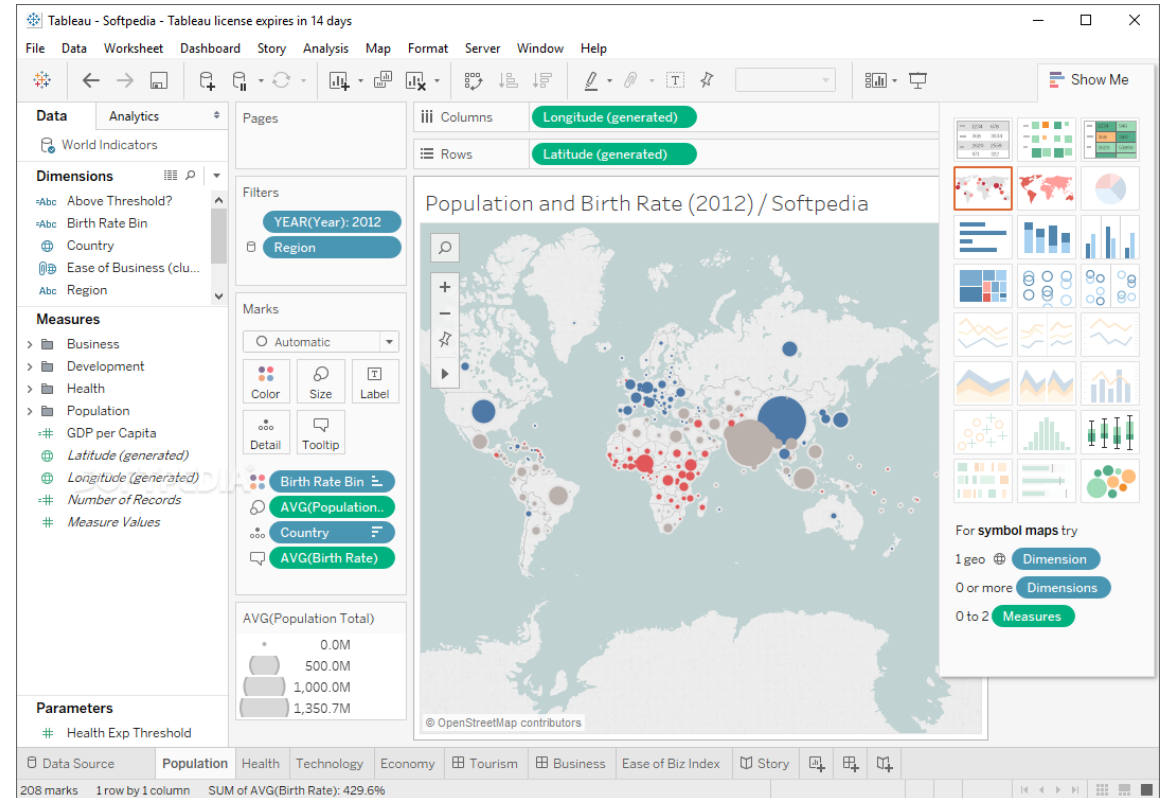


Tools: Data viz

Simple - [Datawrapper](#), [Infogram](#)

Medium - [Tableau](#) (Public version is free), can integrate programming languages and advanced maps. Also Microsoft [Power BI](#).

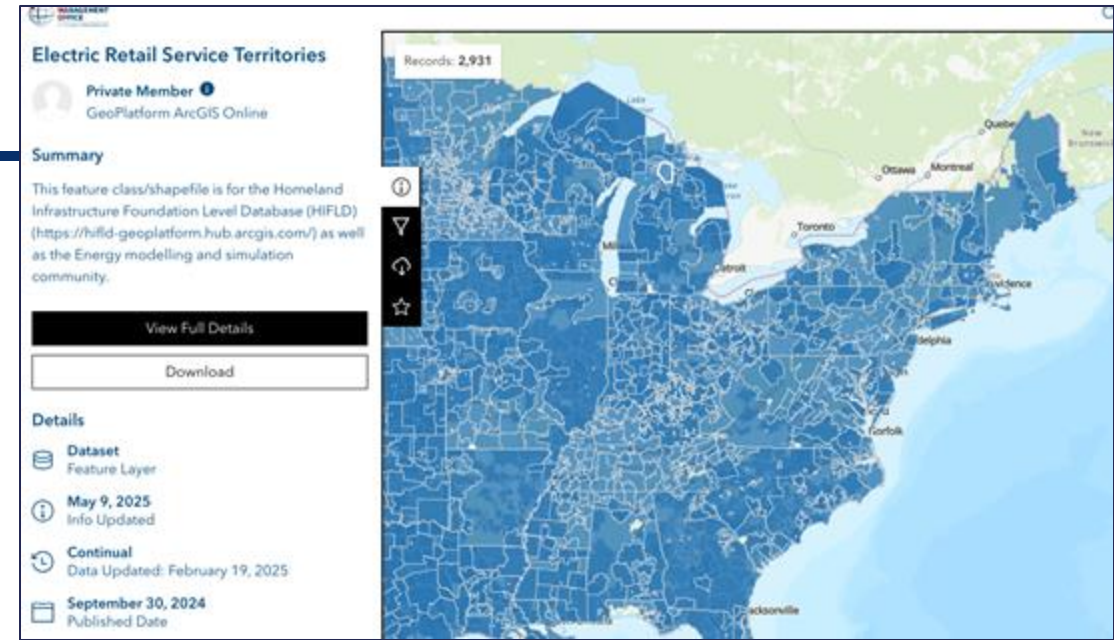
Complex - Packages for R (like [ggplot2](#)) and Python (like [Plotly](#)). Many are open source. Now with AI!



Tools: Mapping

Simple - [Google My Maps](#), Google Earth

Medium - [Tableau Public](#), [Felt](#), [ArcGIS Online](#), [Mapbox](#) (integrates with Tableau)



Complex - GIS software like QGIS (free), [ArcGIS Online](#) (freemium)

🔗 Make your own from raw data, or peruse and combine existing maps

Energy maps: [Homeland Infrastructure Foundation-Level Data](#) (33 energy layers, over 300 altogether). [EIA Energy Atlas](#)



Tools: Energy data sources

[EIA](#) - Electricity Data Browser and so much more

[EIA Open Data](#) - API portal

[OpenEI](#) - “hundreds of topics crowdsourced from industry and government agencies,” 2500 data sets! 3.9 Petabytes!

[IEA](#) - Data sets, explorer, chart library (6000!)

[Public Utility Data Liberation](#) (PUDL) Project

[Grid Status](#)

Spreadsheet Add-Ins



Access EIA data — right in your spreadsheet

You can download EIA energy data (and economic data from the St. Louis Federal Reserve (FRED)) directly to your spreadsheet for charts, graphs, and tables. Version 2.1.0 was released March 30, 2023.

[MICROSOFT EXCEL ADD-INS >](#)










Embed Graphs & Maps



Access EIA data — right on your website

Stay current and embed EIA charts and maps on your website that automatically update when the data is updated.

[WIDGETS >](#)

 Natural Gas > <i>Exploration and reserves, storage, imports and exports, production, prices, sales.</i>	 Petroleum & Other Liquids > <i>Crude oil, gasoline, heating oil, diesel, propane, and liquids including biofuels and natural gas liquids.</i>
 Electricity > <i>Sales, revenue and prices, power plants, fuel use, stocks, generation, trade, demand & emissions.</i>	 Coal > <i>Reserves, production, prices, employment and production distribution, stocks, imports and exports.</i>
 Nuclear & Uranium > <i>Uranium fuel, nuclear reactors, generation, spent fuel.</i>	 Renewable & Alternative Fuels > <i>Includes hydropower, solar, wind, geothermal, bioethanol.</i>
 Consumption & Efficiency > <i>Energy use in homes, commercial buildings, manufacturing, and transportation.</i>	 Total Energy > <i>Comprehensive data summaries, comparisons, and projections integrated across all energy sources.</i>
 Analysis & Projections > <i>Monthly and yearly energy forecasts, analysis of energy topics, financial analysis, congressional reports.</i>	



Tools: Energy photo and image libraries

“Creative Commons” (CC) licensing – free for use, may require attribution and limit editing

Free sources:

- [NREL Image Gallery](#)
- [Flickr](#) (search for CC photos; follow DOE, NREL, others)
- [Pixabay](#), [Unsplash](#), [WikiMedia Commons](#), and many more

Paid sources: iStock, Shutterstock, Getty, etc.

Aleeee! - AI image generators like [ChatGPT](#) and [Gemini](#)



Tools: Illustrations

Make your own:

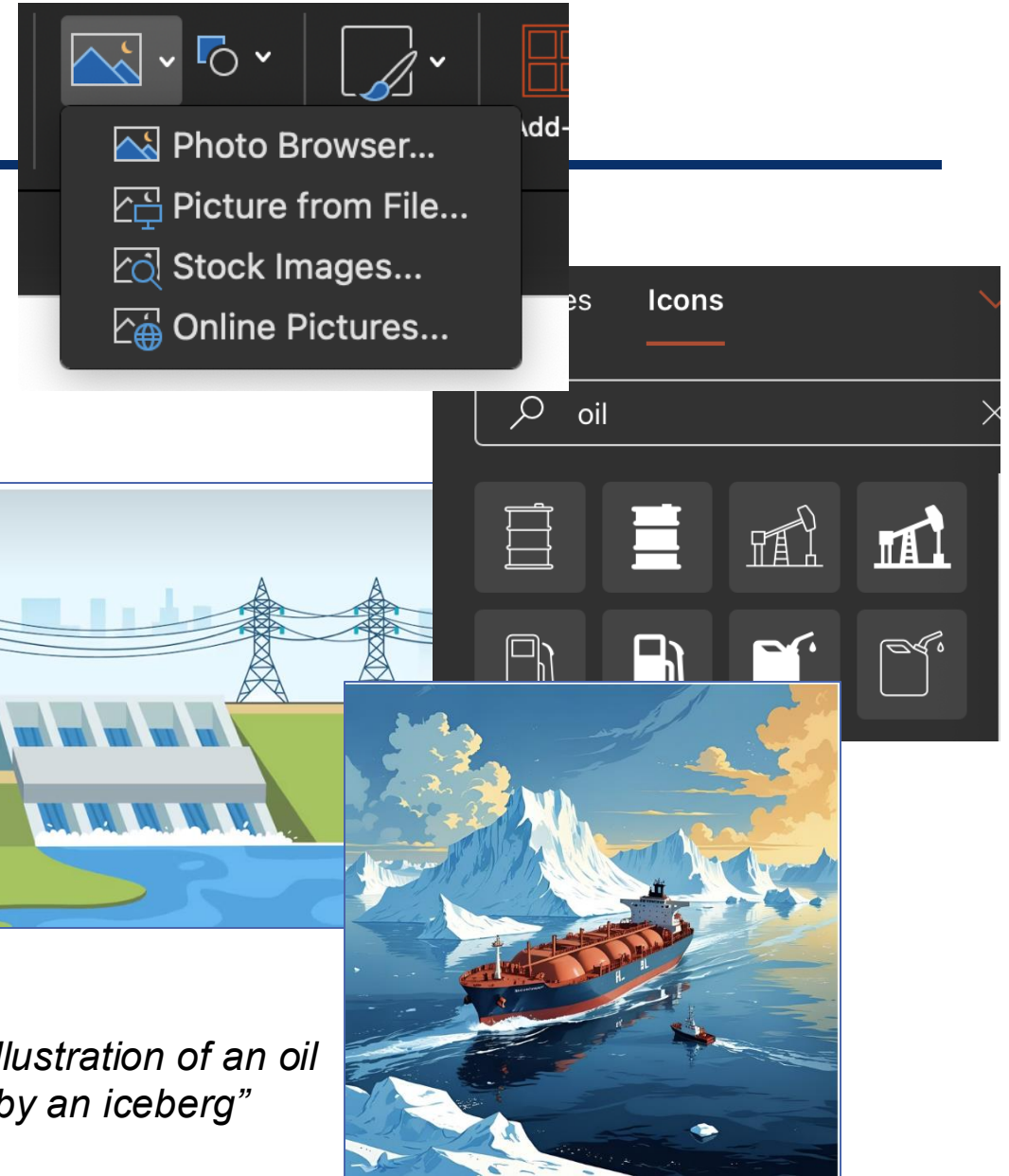
- 🔗 Powerpoint or Google Slides (now have libraries of graphics)
- 🔗 [Canva](#)
- 🔗 [Corel Vector](#)

Many free and freemium sources, like [Unsplash Illustrations](#)

Many AI image generators like [ChatGPT](#) and [Gemini](#)

Vector vs. raster

Prompt: “color illustration of an oil tanker passing by an iceberg”



Where to get inspired and learn

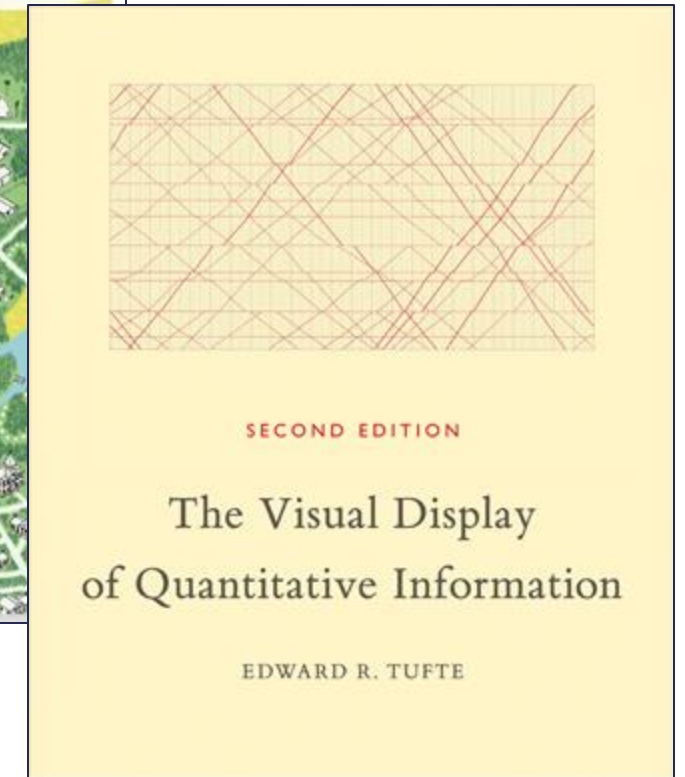
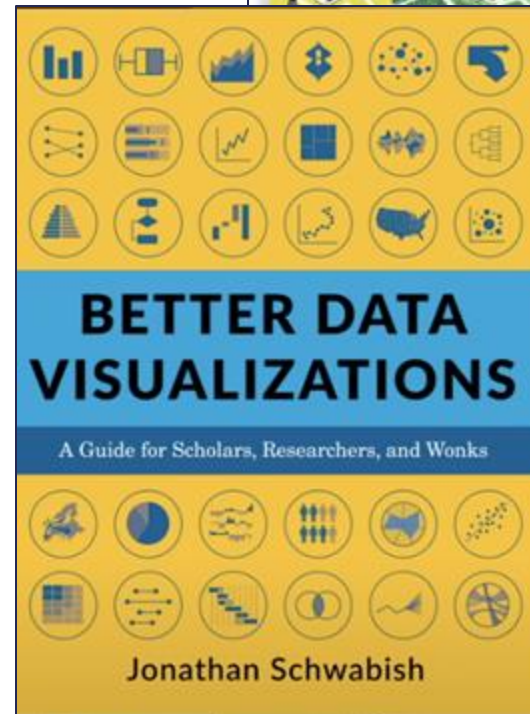
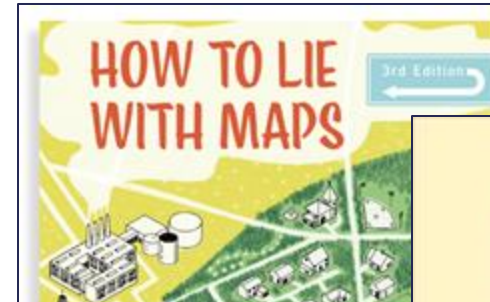
Online:

[Our World in Data](#)

[Gapminder](#), especially [Bubbles](#)

[One Chart at a Time](#) videos

Tableau Public [Viz of the Day](#)



The End

Ben Paulos

Affiliate, Berkeley Lab

benpaulos@lbl.gov

emp.lbl.gov/tools-data

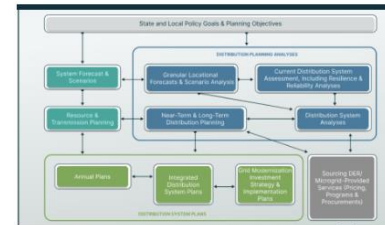


Tools & Data

Research Area Project



Chinese Coal Mine Methane (CMM) Dashboard



Interactive Decision Framework for Integrated Distribution System Planning



State Distribution Planning Requirements



US Photovoltaic Database



Wholesale Electricity Market Price Drivers



PRESTO



Disclaimer

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or The Regents of the University of California.

Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.

Copyright Notice

This manuscript has been authored by an author at Lawrence Berkeley National Laboratory under Contract No. DE-AC02-05CH11231 with the U.S. Department of Energy. The U.S. Government retains, and the publisher, by accepting the article for publication, acknowledges, that the U.S. Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes



Contacts

Bentham Paulos: benpaulos@lbl.gov

Lisa Schwartz: lschwartz@lbl.gov

For more information

Download publications from the Energy Markets & Policy: <https://emp.lbl.gov/publications>

Sign up for our email list: <https://emp.lbl.gov/mailling-list>

Follow the Energy Markets & Policy on Bluesky: [@BerkeleyLabEMP.bsky.social](https://bsky.app/profile/BerkeleyLabEMP.bsky.social)

Acknowledgements

This work was funded by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or The Regents of the University of California.

