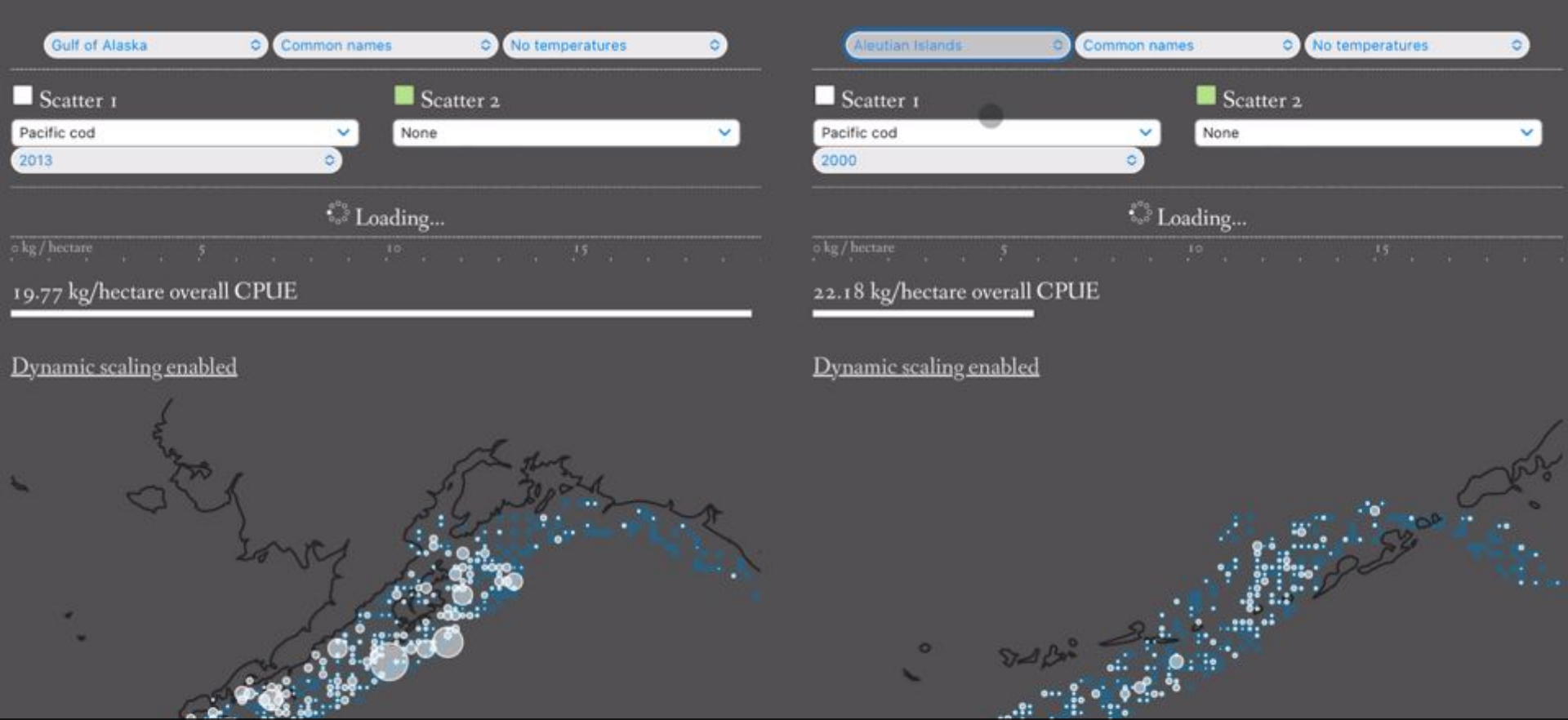
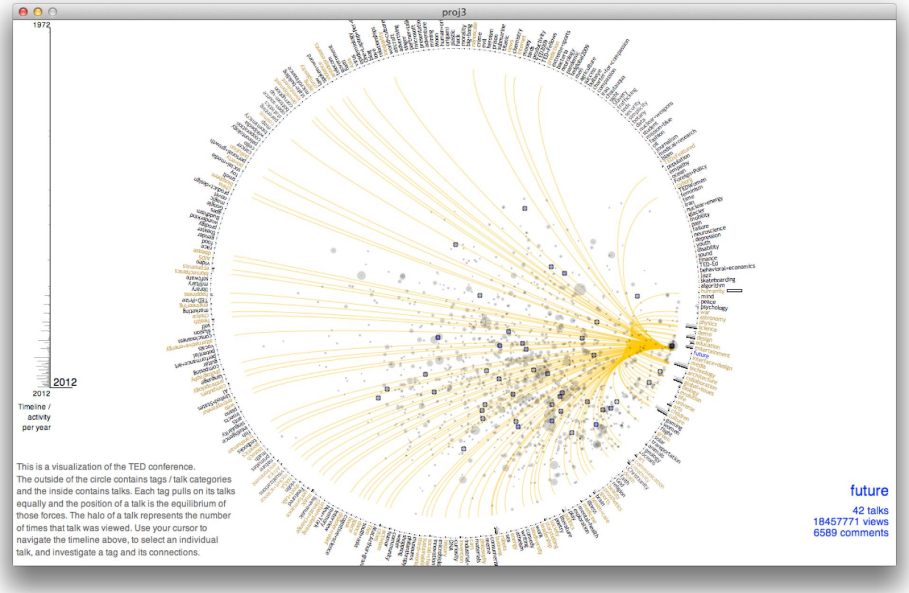


Lecture 1

A Samuel Pottinger
Jan 22, 2025
Stat 198: Interactive Data Science and Visualization

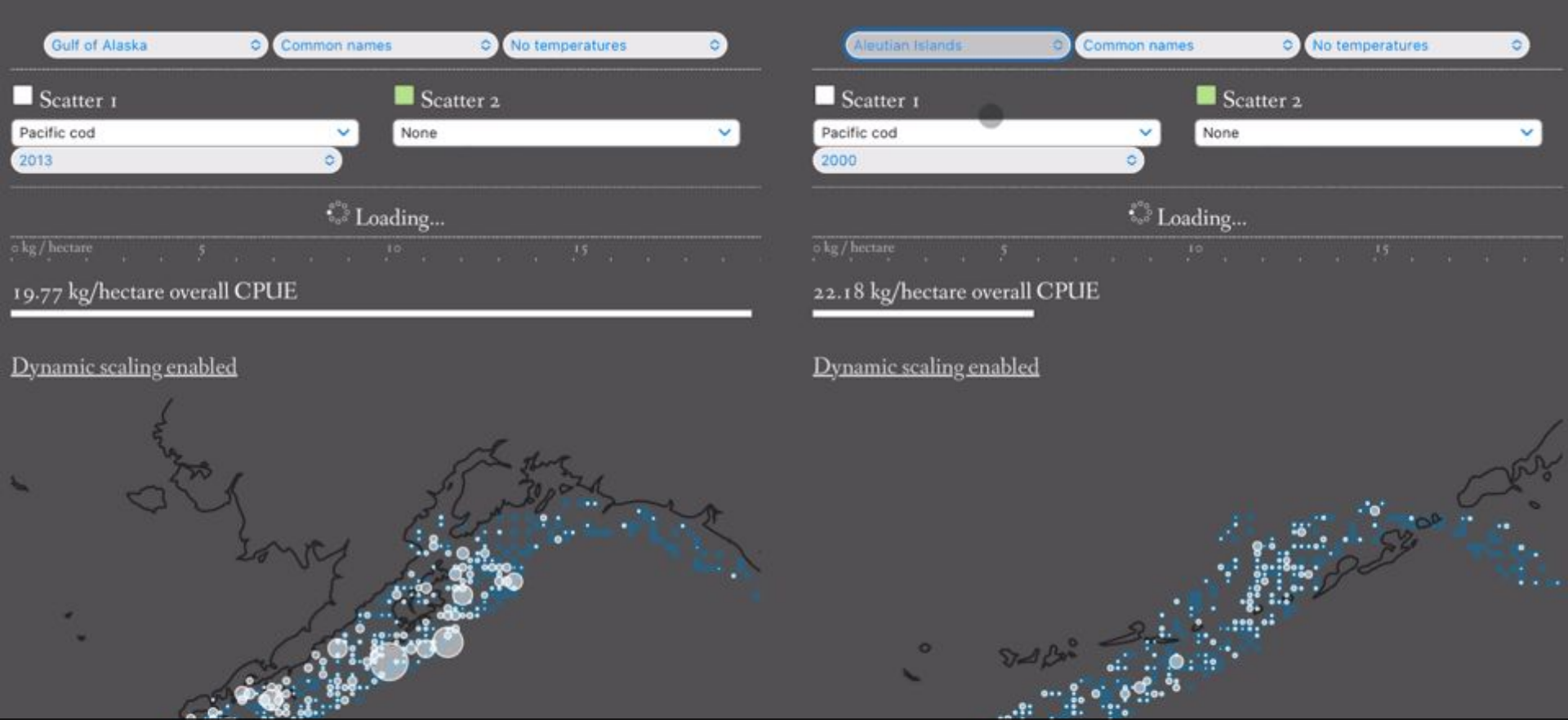


Today

About the course

Overview of concepts

Some logistics



About the Course

Hello! I'm a data scientist, software engineer, and information designer.

Sam Pottinger
A more human-centered AI/ML
https://gleap.org

```
UC Berkeley | Data + Environment
      EVERY | Data + Synthetic Biology
       IDEO | Data + Design
     Plenty | Data + Indoor Agriculture
      Apple | Data + Engineering
     Google | Data + Visualization
    LabJack | Data + Hardware

 Processing | Data + Love in Java
Sketchingpy | Data + Love in Python
```

I’d love to learn more about you and for you to meet some of your fellow students.

Our first homework assignment will test drive the Zulip to do introductions. Keep an eye on #learner-graded and #learner-audit for more information.

What you will be able to do after the course.

- Build data visualizations and other interactive experiences to **share your findings** with others.
- Tell **impactful stories** that engage your readers emotionally through data.
- Invite your audience in as **co-creators** to build new meaning alongside you in your work and collaborate with AI / ML to design solutions and make decisions.
- **Craft tools to explore data-heavy questions** and uncover insights.
- Incorporate **ethics and accessibility** into your data visualization work.

What we will cover together

Section	Concepts	Tech
Hello	Overview of data visualization.	(Creative) Python
Primitives	Perception / cognitive science for viz.	Sketchingpy, Matplotlib
Combination	Data viz within human-centered design.	Geospatial and graph data
Conversation	Game design.	Alternative user inputs
Context	Accessibility and ethics.	Adaptive technologies
Skills	Iterative process.	JavaScript, D3, P5.js

What we will do together

Weekly Exercises

Weekly Reading

Interactive Experience

Final Project + Present

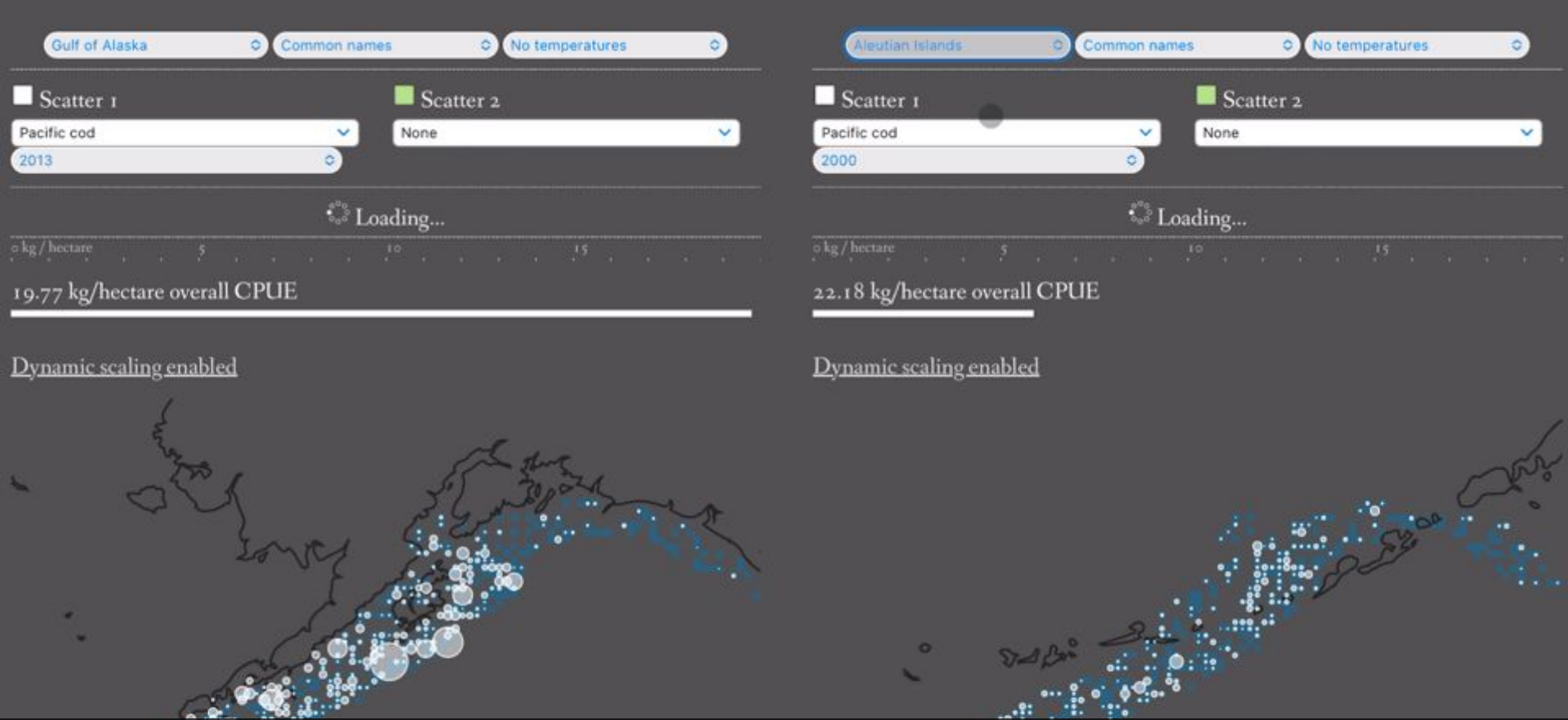
What we will do together

Weekly Exercises	Weekly Reading
Interactive Experience	Final Project + Present

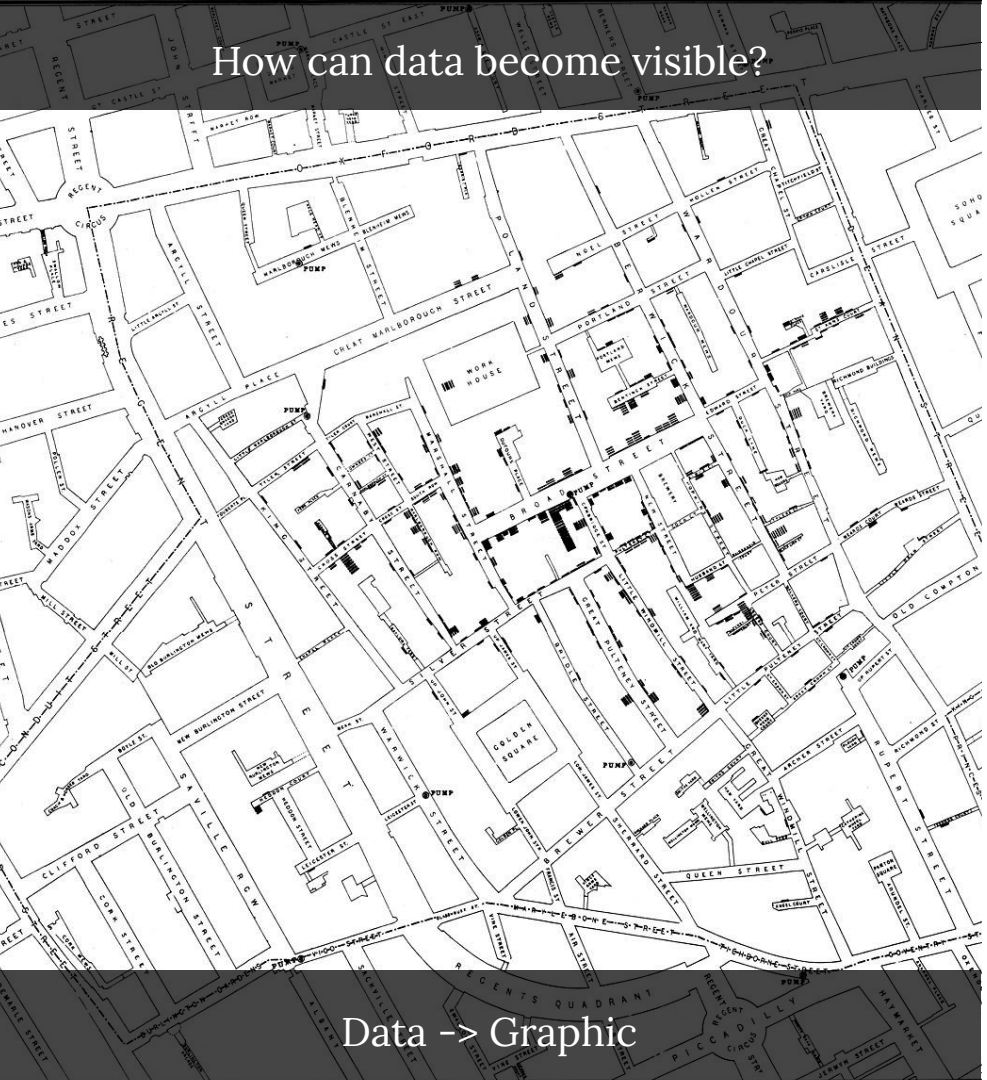
Check out the course manual for more information.

Some things we won't do

- A deep investigation of data manipulation / cleaning.
- Full treatment of evaluative methods.
- Server-side engineering.
- Tableau, PowerBI, ...



4 Perspectives



Data Visualization in 4 Acts

| As representation

As task

As message

As dialogue

Why and how different groups do data visualization.
How you can think about it in your work.

Year	Number of Wolves	Number of Moose
1980	50	664
1982	14	700
1984	24	811
1986	20	1025
1988	12	1653
1990	15	1216
1992	12	1600
1994	15	1800
1996	22	1200
1998	14	700
2000	29	850
2002	17	1000
2004	29	750
2006	30	385
2008	23	650
2010	19	510
2012	9	750
2014	9	1050
2016	2	1300
2018	2	1500

Premise: The human visual system is good at spotting patterns.

What is the relationship between wolves and moose in Isle Royale?

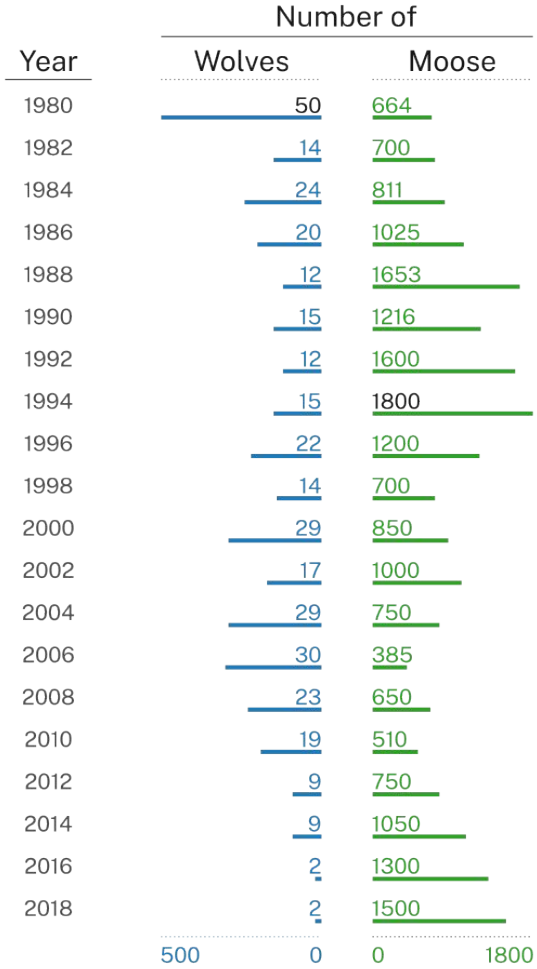
What year saw the most moose?

Year	Number of Wolves	Number of Moose
1980	50	664
1982	14	700
1984	24	811
1986	20	1025
1988	12	1653
1990	15	1216
1992	12	1600
1994	15	1800
1996	22	1200
1998	14	700
2000	29	850
2002	17	1000
2004	29	750
2006	30	385
2008	23	650
2010	19	510
2012	9	750
2014	9	1050
2016	2	1300
2018	2	1500

Premise: The human visual system is good at spotting patterns.

What is the relationship between wolves and moose in Isle Royale?

What year saw the most moose?



Year	Number of Wolves	Number of Moose
1980	50	664
1982	14	700
1984	24	811
1986	20	1025
1988	12	1653
1990	15	1216
1992	12	1600
1994	15	1800
1996	22	1200
1998	14	700
2000	29	850
2002	17	1000
2004	29	750
2006	30	385
2008	23	650
2010	19	510
2012	9	750
2014	9	1050
2016	2	1300
2018	2	1500
	500 0	0 1800

Premise: The human visual system is good at spotting patterns.

What is the relationship between wolves and moose in Isle Royale?

What year saw the most moose?

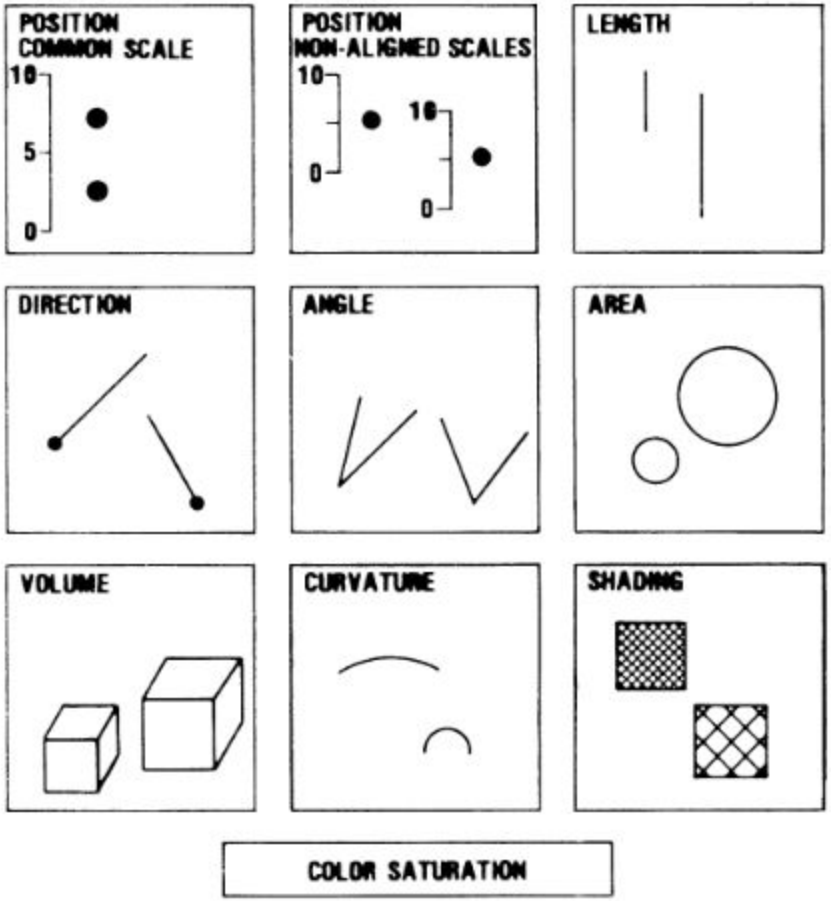


Figure 1. Elementary perceptual tasks.

Example: This first way of thinking about data visualization focuses on encoding.

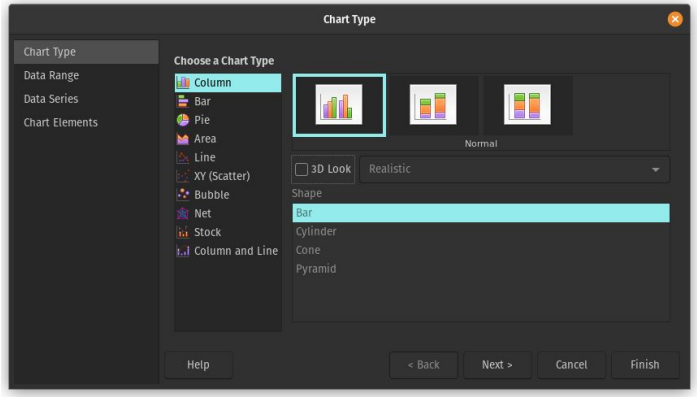
How do we “map” attributes of data to visual attributes?

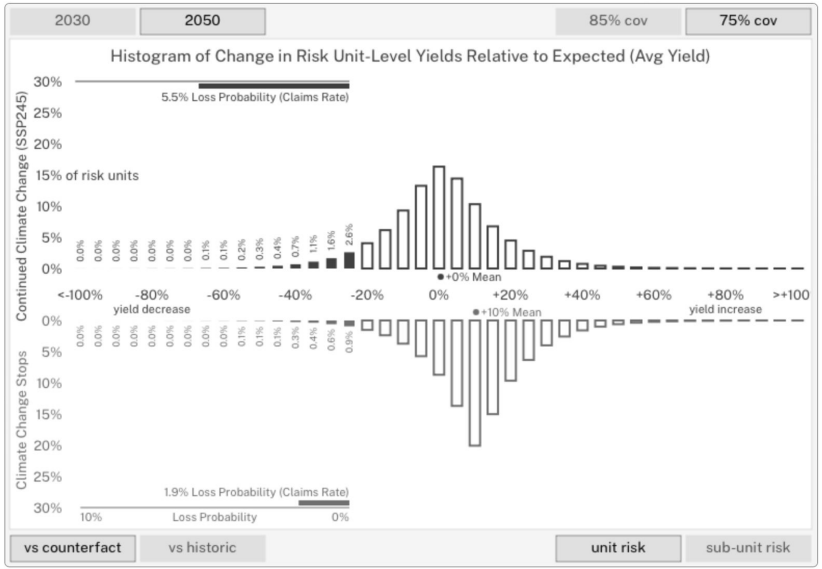
What visual encodings are better than others?

How do we make visualizations accessible?

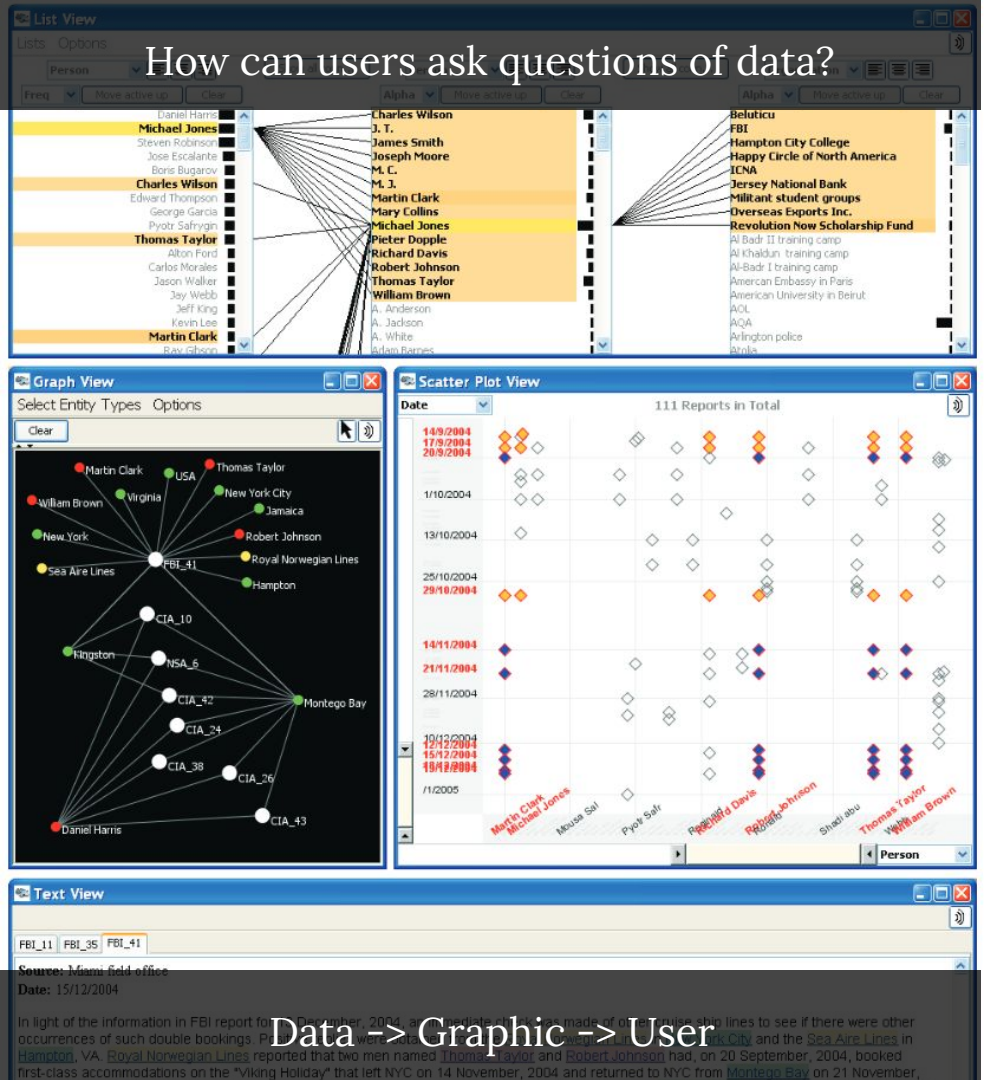
Offers: Flexibility beyond the chart wizard but principles to guide us.

Gives us the basic building blocks for how humans process visual information but lets us use that understanding in many different ways.





Where this will come back: Learn more about human visual system to understand when to use different chart types or encoding devices.



Data Visualization in 4 Acts

As representation

| As task

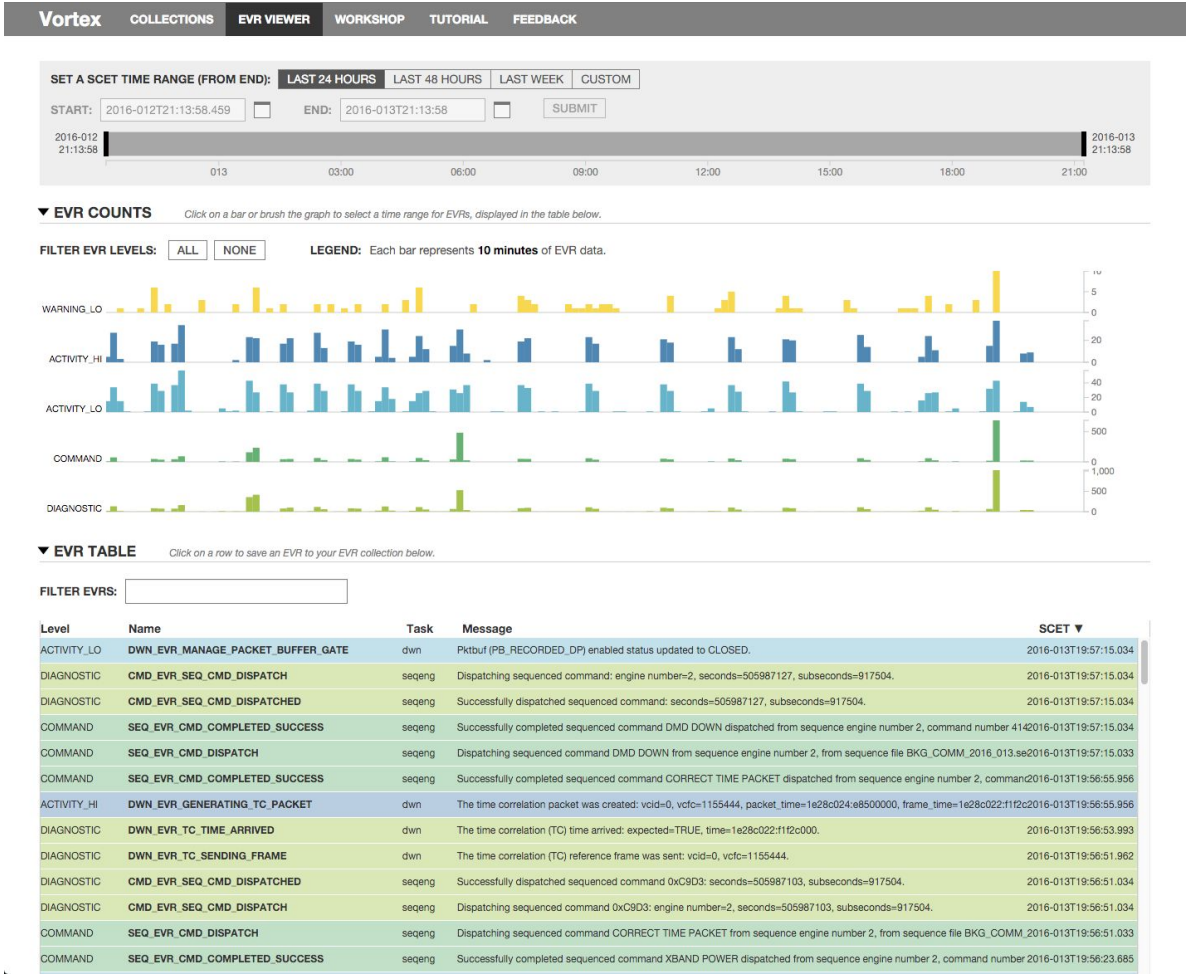
As message

As dialogue

Why and how different groups do data visualization.
How you can think about it in your work.

Premise: Visualizations are part of a broader user journey.

A structured way to think about the user in the context of data visualization.



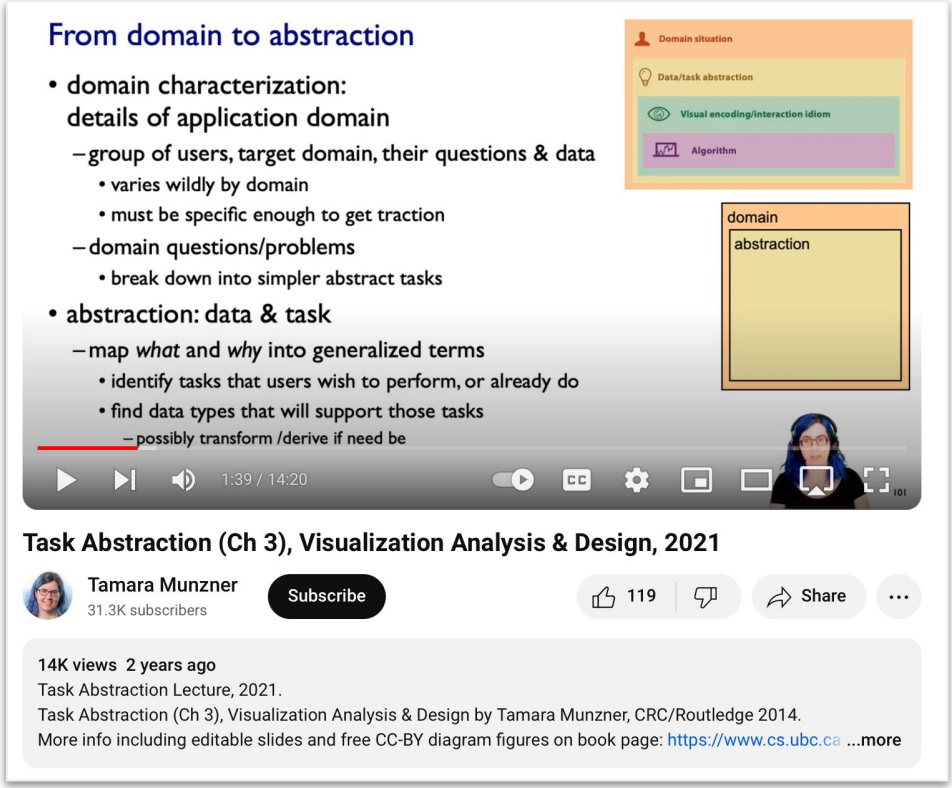
Example: Rachel Binx at NASA.

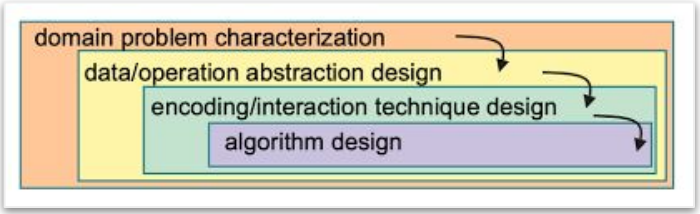
Looking at “event records” sent from spacecraft to NASA.

Interviewed a bunch of users to figure out how they worked with these data previously (log files).

Binx talks about how people had never seen their data before visually and the periodicity of events was revelatory for example.

Boils down into “tasks” the user executes and build user experiences to support those tasks.





Offers: Structured evidence-based understanding of the user to support them in their tasks.

Oritents around domains, tasks, questions, and data.

Fits within a broader modern user experience design dialogue.

Where this comes back: Discussion of how to use more traditional design concepts including those employed in other forms of product and UX design as part of data visualization and interactive data experiences.

Data Visualization in 4 Acts

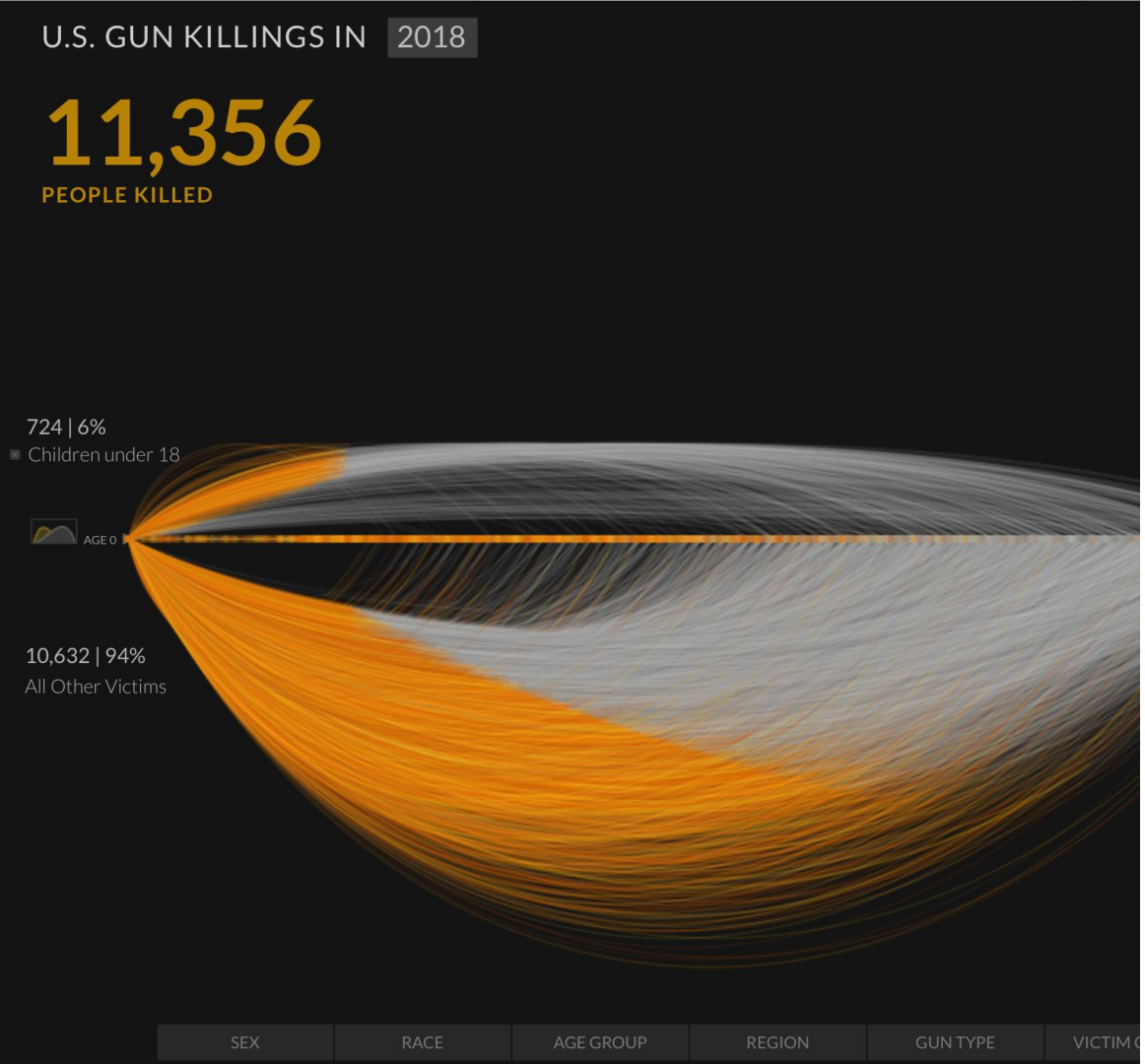
As representation

As task

| As message

As dialogue

Why and how different groups do data visualization.
How you can think about it in your work.



Premise: Forms given to data enable authors to convey a message to a reader.

How does the reader feel when going through a visualization?

Where is efficiency helpful but where does it conflict with the message of the piece?

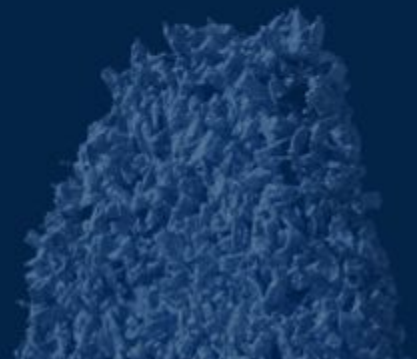
How might we defy reader expectations or have them confront prior held beliefs?

guns.periscopic.com

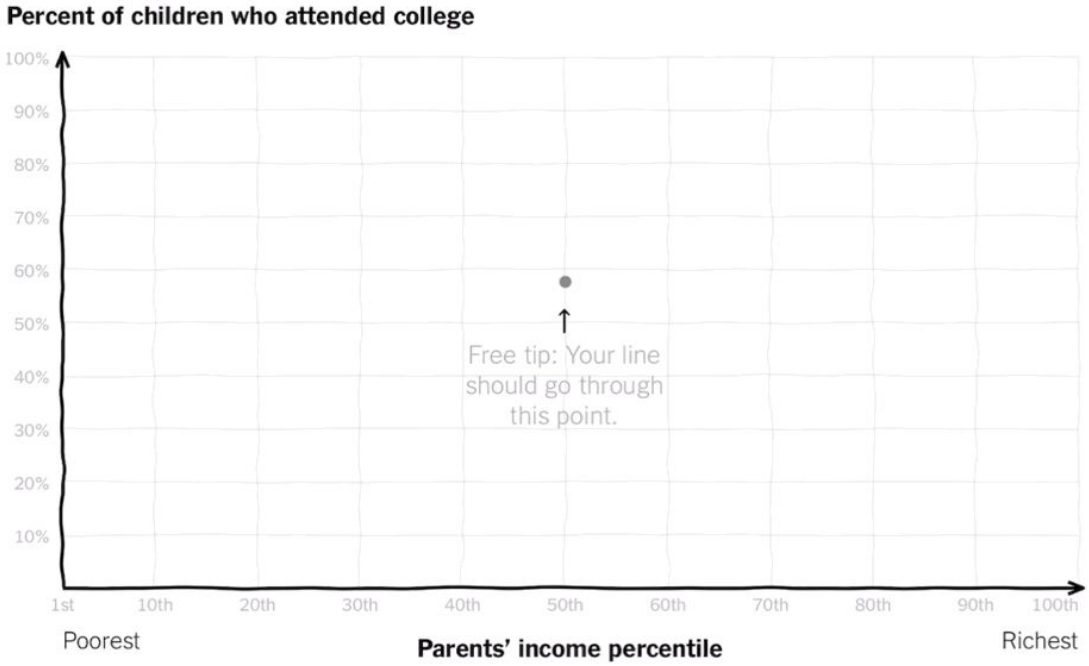
A Treaty To End Plastic Pollution. Forever.

World leaders must take steps to drastically limit the impact of plastics on the environment and human health

CHOOSE OUR FUTURE



Draw your line on the chart below



I'm done

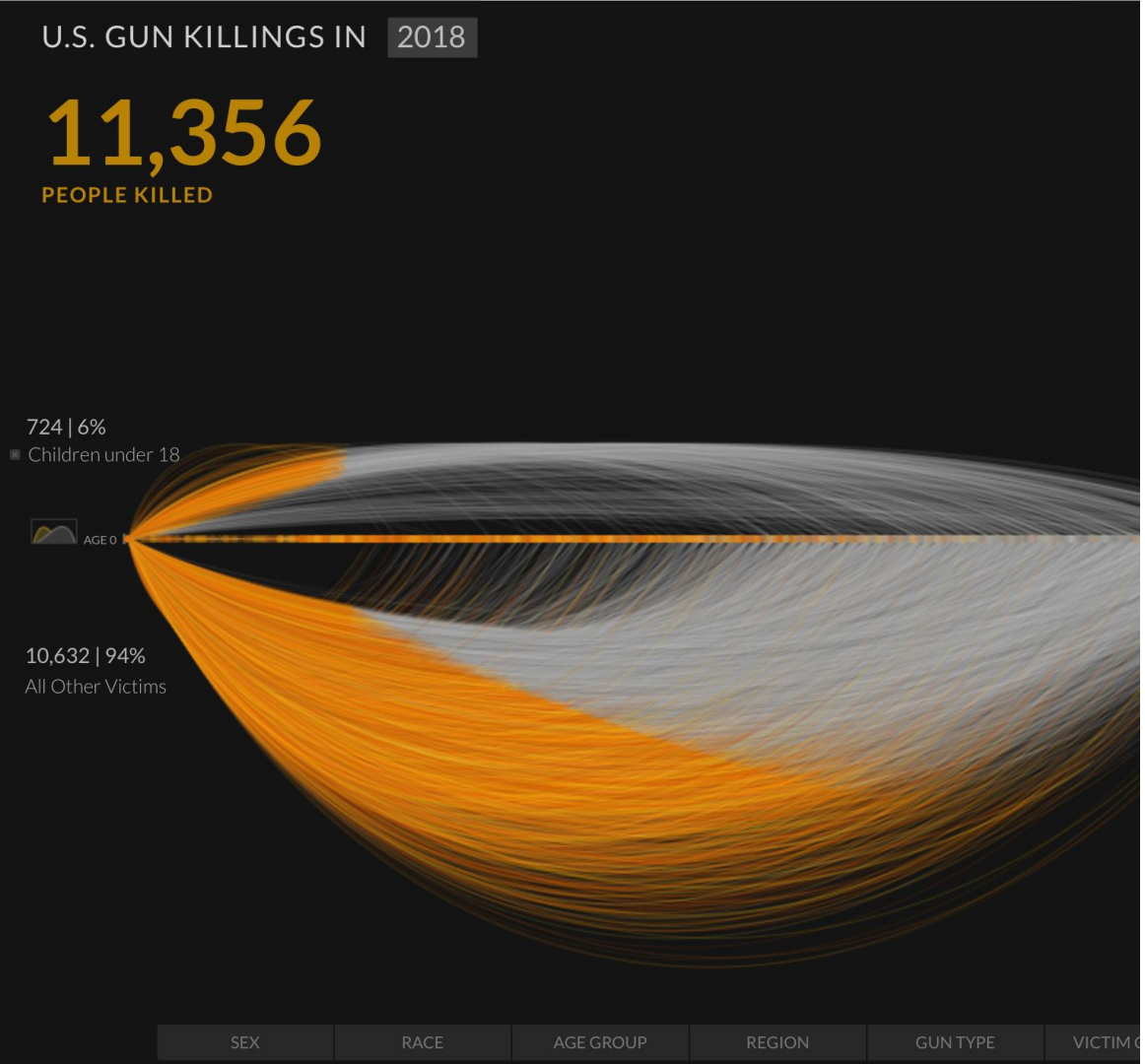
Start over

Offers: A way to convey messages with logos and pathos.

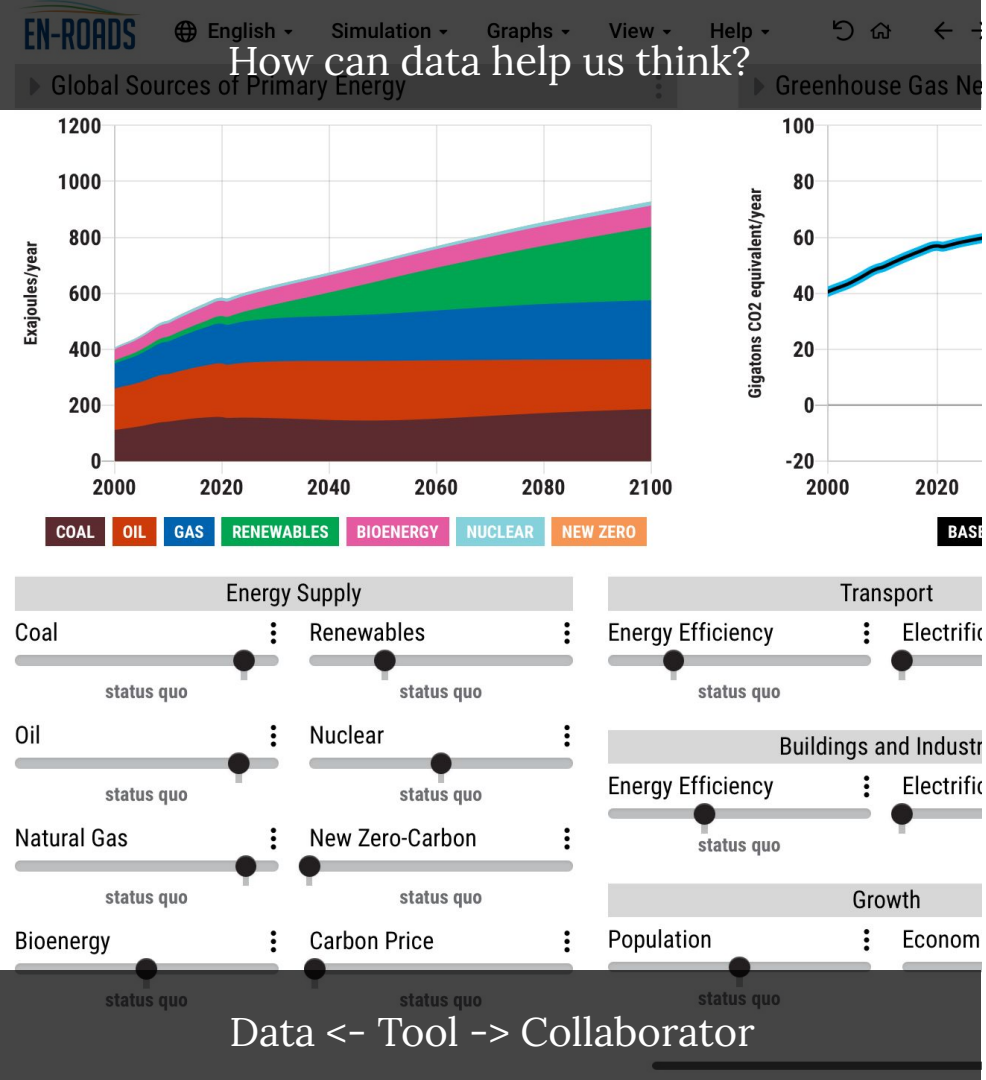
How to invoke emotional response.

How to challenge reader assumptions.

How to understand the process by which messages and meaning are interpreted.



Where this will come back: Techniques we can borrow from art and design to guide and evoke an emotional response.



Data Visualization in 4 Acts

As representation

As task

As message

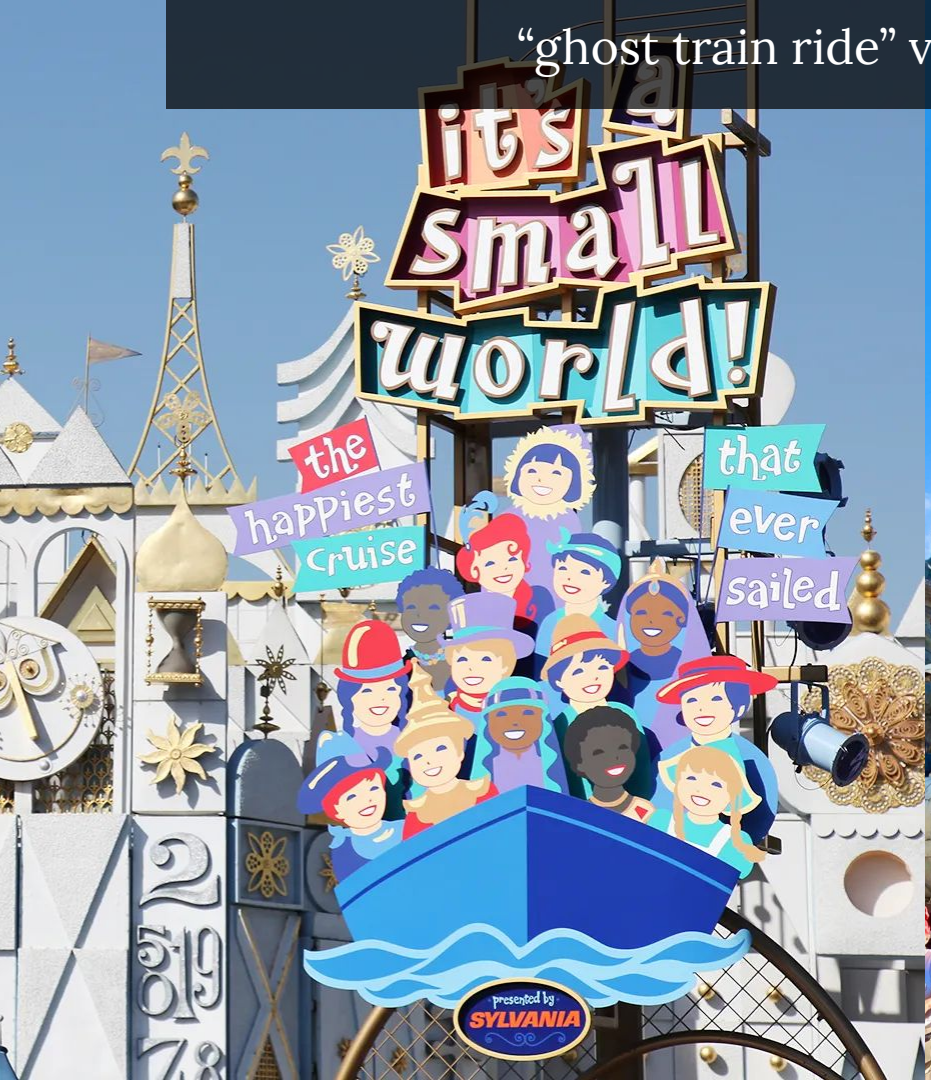
As dialogue

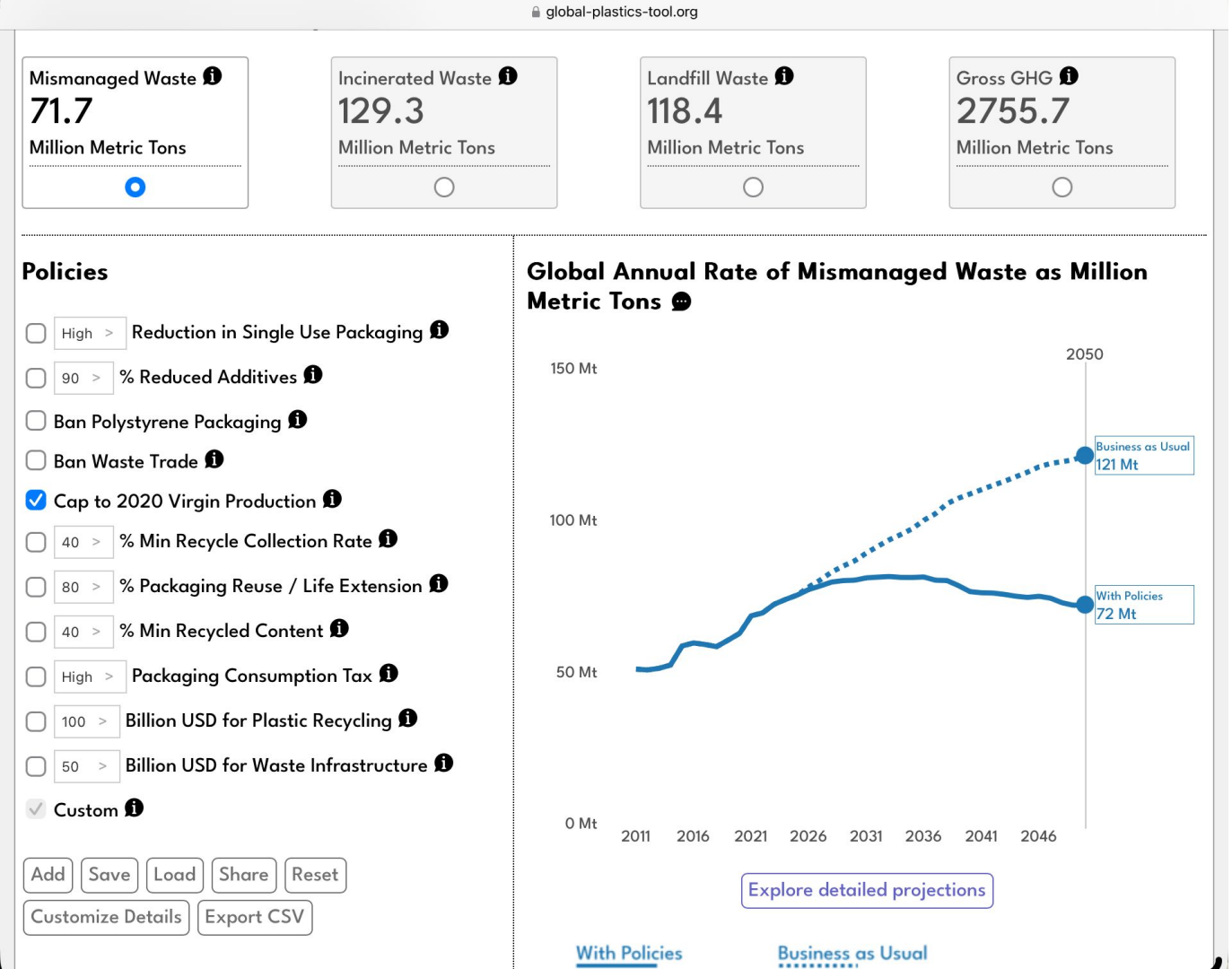
Why and how different groups do data visualization.
How you can think about it in your work.

Premise: Data as humane dynamic media.

The designer creates media for thought, elevating the reader to an author of tools and co-creator of meaning.

“ghost train ride” vs “open world”



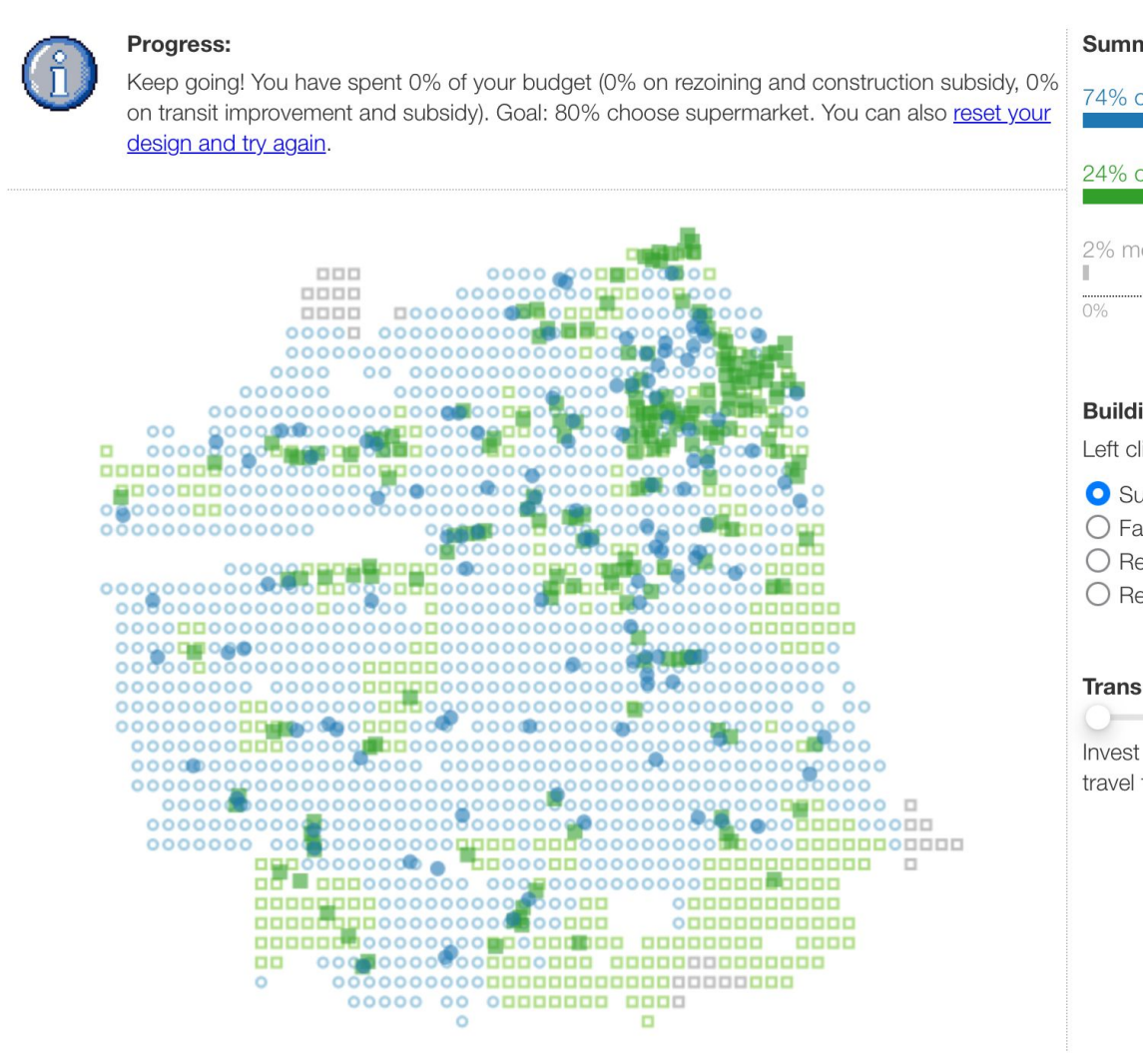


Example: Finding a solution to the plastics crisis.

A layered experience in which the user can simulate different policies.

An invitation to build outside the original designer's intention.

https://global-plastics-tool.org



Offers: Co-creation and user agency.

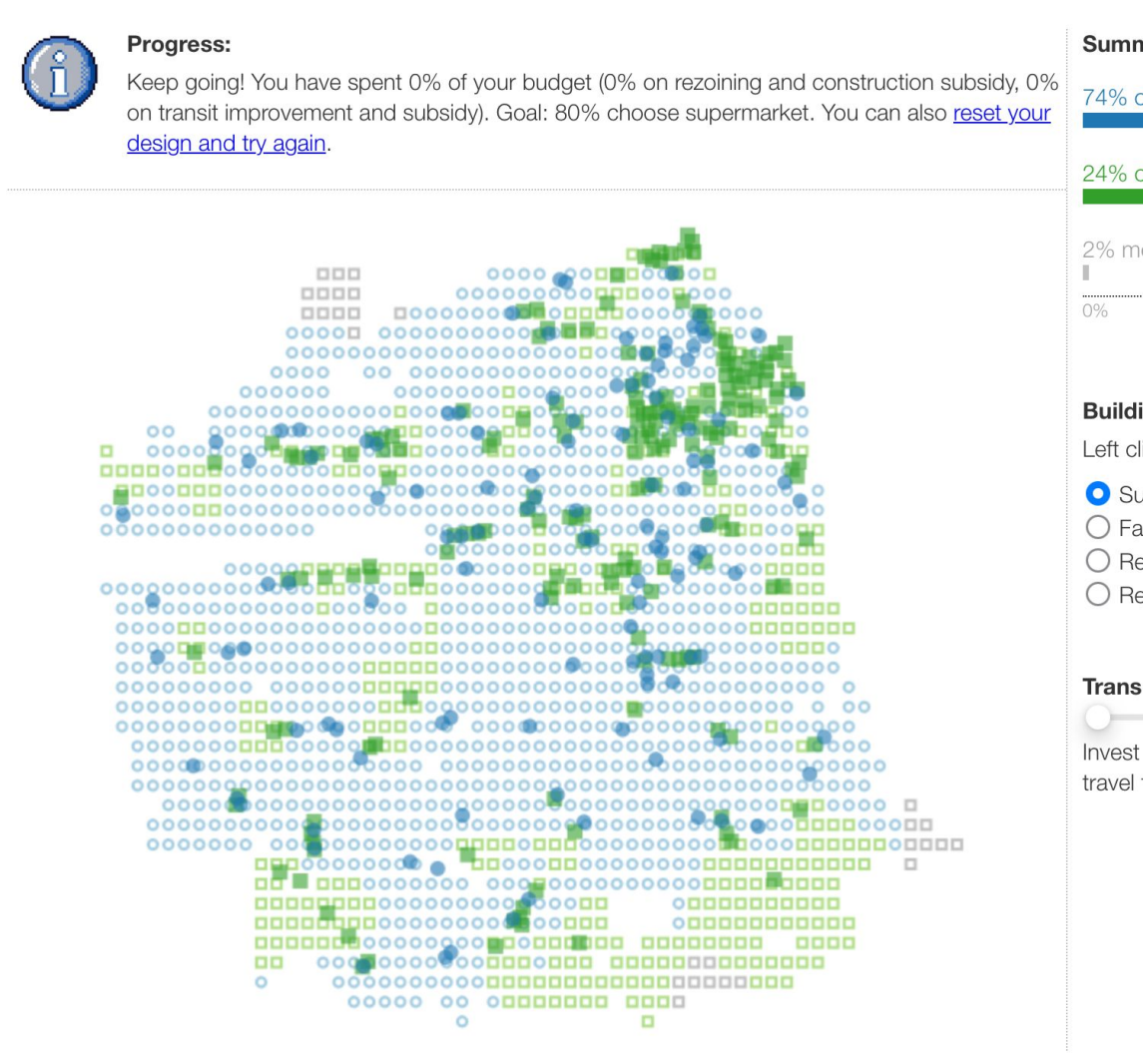
Often leaning on game design concepts.

How to teach with/without tutorializing.

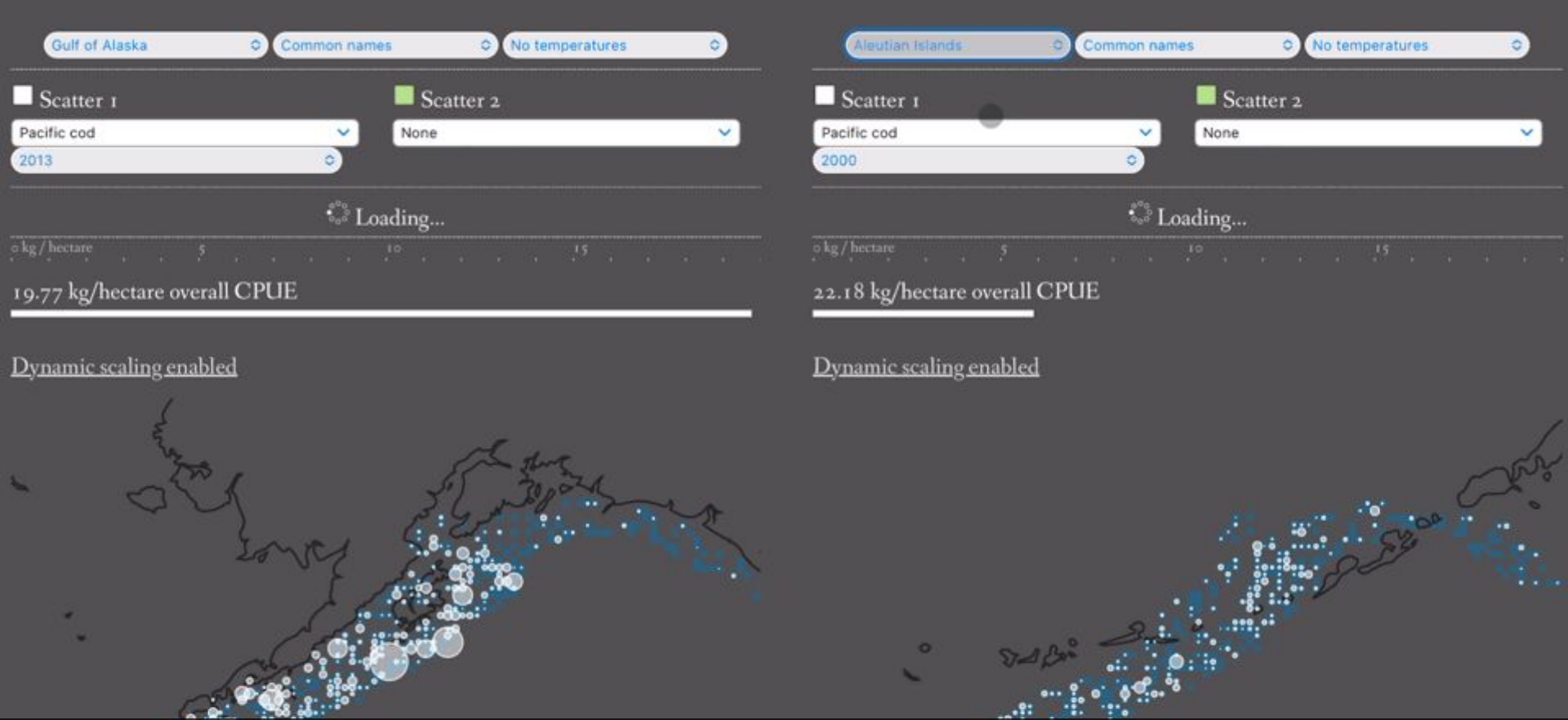
How to create spaces to interrogate assumptions.

How to build media to be repurposed.

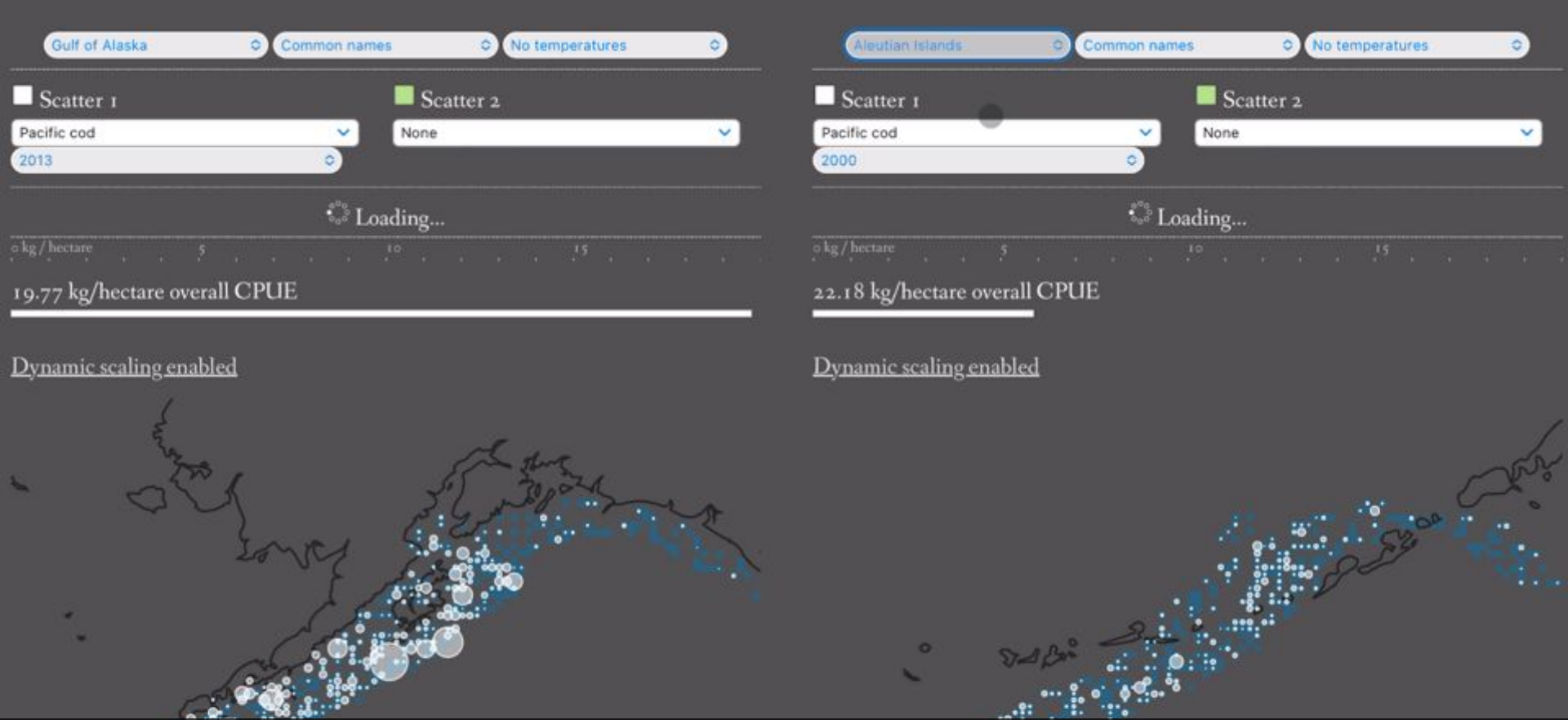
How to design experiences where the user is co-author.



Where this comes back: How to employ interaction and game design to create digital spaces where users can explore data more freely and go beyond your own narrative.


Gulf of Alaska
Common names
No temperatures
Scatter 1
Pacific cod
2013
Scatter 2
None
Loading...
0 kg / hectare
5
10
15
19.77 kg/hectare overall CPUE
Dynamic scaling enabled
Aleutian Islands
Common names
No temperatures
Scatter 1
Pacific cod
2000
Scatter 2
None
Loading...
0 kg / hectare
5
10
15
22.18 kg/hectare overall CPUE
Dynamic scaling enabled

Logistics


Gulf of Alaska
Common names
No temperatures
Scatter 1
Pacific cod
2013
Scatter 2
None
Loading...
0 kg / hectare
5
10
15
19.77 kg/hectare overall CPUE
Dynamic scaling enabled
Aleutian Islands
Common names
No temperatures
Scatter 1
Pacific cod
2000
Scatter 2
None
Loading...
0 kg / hectare
5
10
15
22.18 kg/hectare overall CPUE
Dynamic scaling enabled

Works Cited

[1]
G. Aisch, A. Cox, and K. Quealy, “You Draw It: How Family Income Predicts Children’s College Chances.” The New York Times Company, May 28, 2015. [Online]. Available: https://www.nytimes.com/interactive/2015/05/28/upshot/you-draw-it-how-family-income-affects-childrens-college-chances.html

[2]
R. Binx, “Vortex.” 2015. [Online]. Available: https://rachelbinx.com/data-visualization/vortex

[3]
R. Binx, “Designing for Realtime Spacecraft Operations.” BocoupLLC, Apr. 2016. [Online]. Available: https://www.youtube.com/watch?v=HuYKhSHcRSQ

[4]
Book Books Book, “We Feel Fine.” YouTube, Dec. 2009. [Online]. Available: https://www.youtube.com/watch?v=vi8WrnWNSzU

[5]
W. S. Cleveland and R. McGill, “Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods,” *Journal of the American Statistical Association*, vol. 79, no. 387, pp. 531–554, Sep. 1984, doi: 10.1080/01621459.1984.10478080.

[6]
Flipflops, “I feel… everything.” Flipflops.org, Oct. 2007. [Online]. Available: https://www.flipflops.org/category/thoughtful/page/2/

[7]
V. Hart and N. Case, “Parable of the Polygons.” Nicky Case, 2022. [Online]. Available: https://ncase.me/polygons/

[8]
Isle Royale National Park MIchigan, “Wolf & Moose Populations.” National Parks Service, Mar. 29, 2024. [Online]. Available: https://www.nps.gov/isro/learn/nature/wolf-moose-populations.htm

[9]
S. D. Kamvar and J. Harris, “We feel fine and searching the emotional web,” in *Proceedings of the fourth ACM international conference on Web search and data mining*, Hong Kong China: ACM, Feb. 2011, pp. 117–126. doi: 10.1145/1935826.1935854.

[10]
A. KIRK, *DATA VISUALISATION: a handbook for data driven design*. S.l.: SAGE PUBLICATIONS, 2024.

[11]
A. Kirk, “Visualizing Data.” Visualising Data Ltd, 2024. [Online]. Available: https://visualisingdata.com/

[12]
T. Munzner, “A Nested Model for Visualization Design and Validation,” *IEEE Trans. Visual. Comput. Graphics*, vol. 15, no. 6, pp. 921–928, Nov. 2009, doi: 10.1109/TVCG.2009.111.

[13]
T. Munzner, “Visualization Analysis and Design.” AK Peters Visualization Series, 2014. [Online]. Available: https://books.apple.com/us/book/visualization-analysis-and-design/id1567434451

[14]
T. Munzner, *Visualization analysis and design*. in A.K. Peters visualization series. Boca Raton: CRC Press, Taylor & Francis Group, CRC Press is an imprint of the Taylor & Francis Group, an informa business, 2015.

[15]
T. Munzner, “Task Abstraction (Ch 3), Visualization Analysis & Design, 2021.” YouTube, 2021. [Online]. Available: https://www.youtube.com/watch?v=pHljd-cgICY

[16]
M. Nix, *Visual simplexity: die Darstellung großer Datenmengen*. Frankfurt am Main: entwickler.press, 2013.

[17]
Periscopic, “U.S. Gun Deaths.” Periscopic, 2018. [Online]. Available: https://guns.periscopic.com/

[18]
A. Pottinger, “FoodSim: San Francisco.” 2023. [Online]. Available: https://foodsimsf.com/

[19]

[19]
A. Pottinger, “Income Gaps.” 2023. [Online]. Available: https://incomegaps.com/

[20]
A. Pottinger, “Interactive Data Science.” 2024. [Online]. Available: https://interactivedatascience.courses/

[21]
A. S. Pottinger *et al.*, “Combining Game Design and Data Visualization to Inform Plastics Policy: Fostering Collaboration between Science, Decision-Makers, and Artificial Intelligence,” 2023, *arXiv*. doi: 10.48550/ARXIV.2312.11359.

[22]
A. S. Pottinger, L. Connor, B. Guzder-Williams, M. Weltman-Fahs, and T. Bowles, “Climate-Driven Doubling of Maize Loss Probability in U.S. Crop Insurance: Spatiotemporal Prediction and Possible Policy Responses,” 2024, *arXiv*. doi: 10.48550/ARXIV.2408.02217.

[23]
A. S. Pottinger and G. Zarpellon, “Pyafscgap.org: Open source multi-modal Python-basedtools for NOAA AFSC RACE GAP,” *JOSS*, vol. 8, no. 86, p. 5593, Jun. 2023, doi: 10.21105/joss.05593.

[24]
J. N. Rooney-Varga, F. Kapmeier, J. D. Sterman, A. P. Jones, M. Putko, and K. Rath, “The Climate Action Simulation,” *Simulation & Gaming*, vol. 51, no. 2, pp. 114–140, Apr. 2020, doi: 10.1177/1046878119890643.

[25]
J. Schell, *The art of game design: a book of lenses*, Third edition. Boca Raton: CRC Press/Taylor & Francis Group, 2020.

[26]
J. Snow, *On the mode of communication of cholera*. London: John Churchill, 1855. [Online]. Available: https://archive.org/details/b28985266/page/n57/mode/2up

[27]
J. Stasko, C. Gorg, Z. Liu, and K. Singhal, “Jigsaw: Supporting Investigative Analysis through Interactive Visualization,” in *2007 IEEE Symposium on Visual Analytics Science and Technology*, Sacramento, CA, USA: IEEE, Oct. 2007, pp. 131–138. doi: 10.1109/VAST.2007.4389006.

[28]
The Document Foundation, *LibreOffice*. (2024). The Document Foundation.

[29]
ThoughtLab, The Wendy and Eric Schmidt Center for Data Science and Environment, and Benioff Ocean Science Laboratory, “A Treaty to End Plastic Pollution. Forever.” University of California, 2023. [Online]. Available: https://plasticstreaty.berkeley.edu/

[30]
B. Victor, “Inventing on Principle.” CUSEC, 2012. [Online]. Available: https://www.youtube.com/watch?v=PUv66718DII

[31]
B. Victor, “Media for Thinking the Unthinkable.” MIT Media Lab, Apr. 04, 2013. [Online]. Available: https://vimeo.com/67076984

[32]
Visual Computing BLOG, “Tamara Munzner discussed quantification in terms of a nested model of visualization design and evaluation.” Transregional Collaborative Research Center. [Online]. Available: https://www.visual-computing.org/2018/10/17/computerscienceconferenceweek/201810_conferenceweek_munzner-2/

[33]
C. Ware, “Colin Ware | The Data Visualzation Research Lab.” University of New Hampshire. [Online]. Available: https://ccom.unh.edu/vislab/people/colin_ware/

[34]
C. Ware, *Information visualization: perception for design*, Fourth edition. Cambridge, MA: Morgan Kaufmann, 2021.

[35]
Wikipedia Contributors, “Snow-cholera-map-1.jpg.” Wikimedia Foundation, Inc., 2020. [Online]. Available: https://en.wikipedia.org/wiki/File:Snow-cholera-map-1.jpg

[36]

[36] Wikipedia Contributors, “Bret Victor.” Wikimedia Foundation, Inc., Jun. 22, 2023. [Online]. Available: https://en.wikipedia.org/wiki/Bret_Victor

[37] Wikipedia Contributors, “Star Wars: Galaxy’s Edge.” Wikimedia Foundation, Inc., Sep. 21, 2024. [Online]. Available: https://en.wikipedia.org/wiki/Star_Wars:_Galaxy%27s_Edge

[38] Wikipedia Contributors, “It’s a Small World.” Wikimedia Foundation, Inc., Sep. 24, 2024. [Online]. Available: https://en.wikipedia.org/wiki/It%27s_a_Small_World

[39] N. Yee, “Motivations for Play in Online Games,” *CyberPsychology & Behavior*, vol. 9, no. 6, pp. 772–775, Dec. 2006, doi: 10.1089/cpb.2006.9.772.

Thanks to https://unsplash.com/photos/DHl49oyrn7Y


CC BY-NC-SA 4.0