Building Inclusive Data Visualizations

Negeen Aghassibake, MSIS

Data Visualization Librarian

University of Washington Libraries

W UNIVERSITY LIBRARIES

hot topics in practice

Northwest Center FOR PUBLIC HEALTH PRACTICE

> " **One of the ways that there is a continuing genocide against American Indians/Alaska Natives is through data. When we are invisible in the data, we no longer exist. When I see an asterisk that says "not statistically significant," or they lump us together with Pacific Islanders and Asian Americans – you can't lump racial groups together. That is bad data practice.**

— Abigail Echo-Hawk,
Urban Indian Health Institute

Question for the Viewers

?

Have you ever seen a chart that seemed problematic or made you question its accuracy in relation to your understanding of the topic?

A. Yes

B. No

C. I'm not sure

NOTE: If you answered "**Yes**," please share details in the chat box.

Understanding Inclusive Visualization

Task: Take a brief moment to open your web browser and complete an image search for the term "CEO."

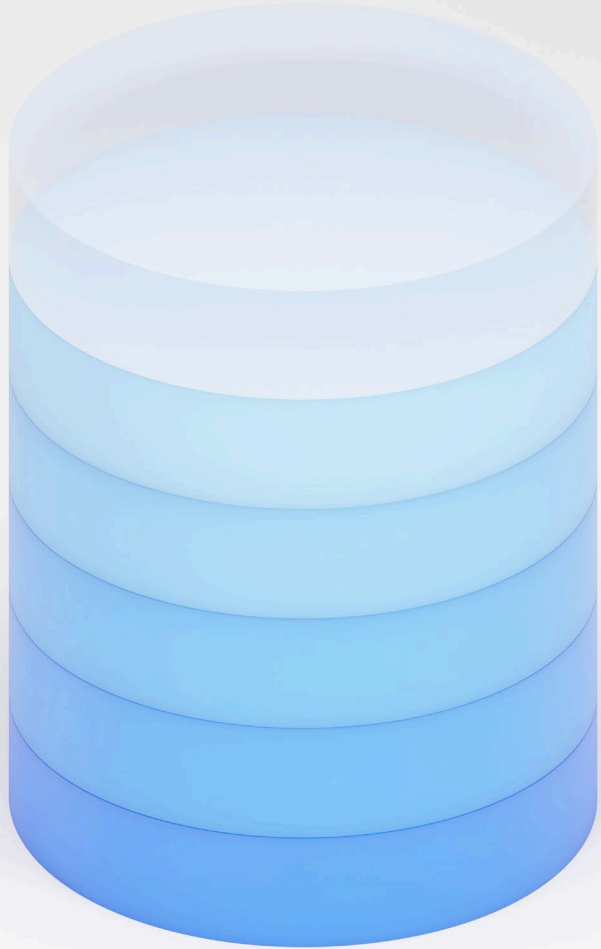We'll return to your results later in the presentation.

CEO

Data is Real Life

> " **Data represents real life. It is a snapshot of the world, in the same way that a picture catches a small moment in time...**

— Georgia Lupi, "Data Humanism, the Revolution will be Visualized"

Data try to represent

but are not real people,

places, or things.

Clean

Normalize

Wrangle

Tidy

Prepare

Survey Design Example: United States

How is race defined in the US?

Consider someone who is Middle Eastern.

About

The U.S. Census Bureau must adhere to the 1997 Office of Management and Budget (OMB) standards on race and ethnicity which guide the Census Bureau in classifying written responses to the race question:

White – A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

Black or African American – A person having origins in any of the Black racial groups of Africa.

American Indian or Alaska Native – A person having origins in any of the original peoples of North and South America (including Central America) and who maintains tribal affiliation or community attachment.

Asian – A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.

Native Hawaiian or Other Pacific Islander – A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

The 1997 OMB standards permit the reporting of more than one race. An individual's response to the race question is based upon self-identification.
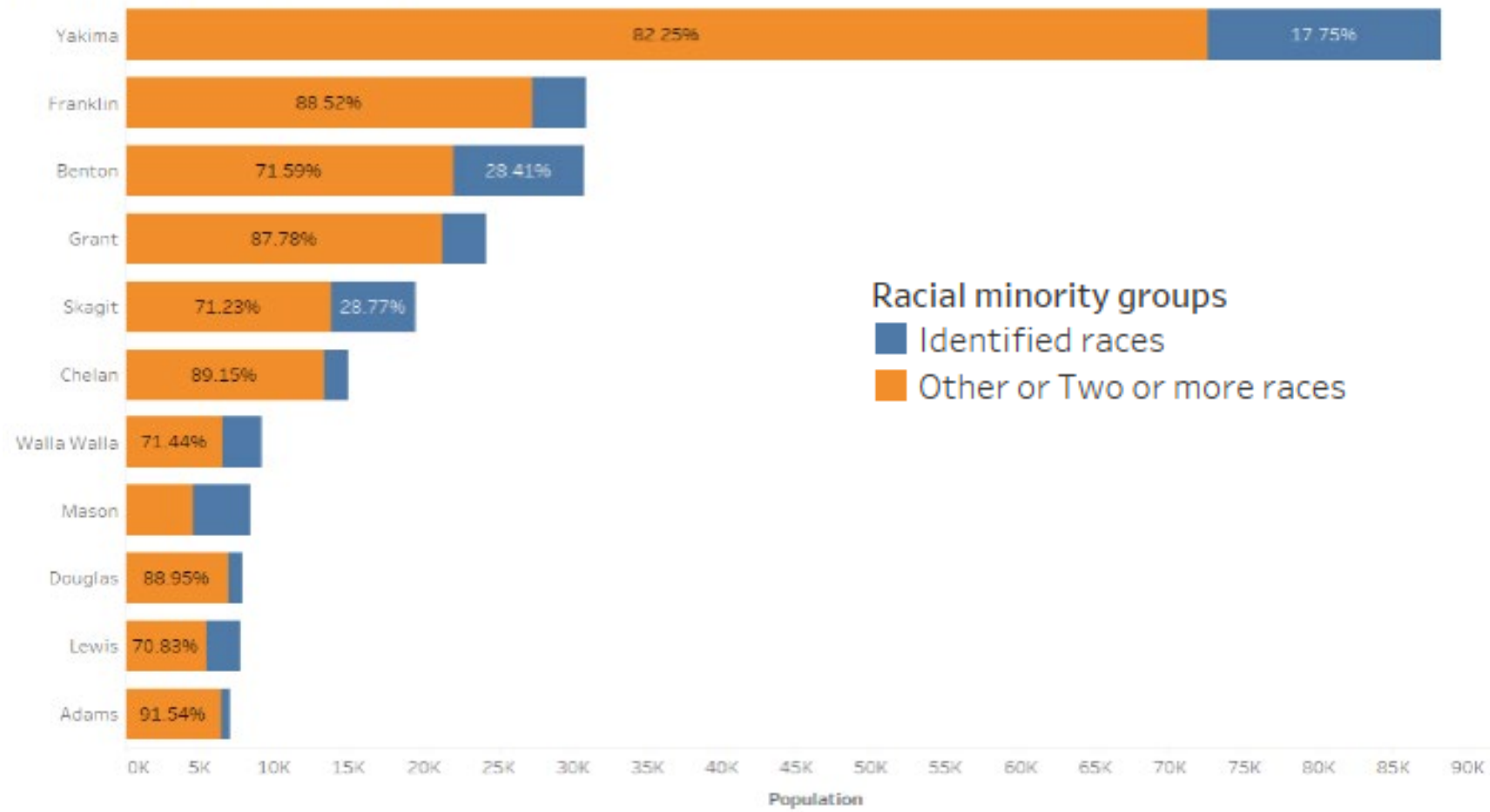
Source: U.S. Census Bureau, "About," https://www.census.gov/topics/population/race/about.html

> " There is a long history of excluding categories like 'Arab' or 'Middle Eastern and North African' from government surveys...See, if you exclude those groups from the US Census, a lot of Arabs might feel like there's no choice but to tick the 'white' box, helping to artificially increase those numbers for political gain.

— Mona Chalabi

Source: "Choices" Instagram post, October 13, 2020, https://www.instagram.com/p/CGSUEFSF2q8/

Identifying "Other": Racial Minorities in Washington

Racial minorities in WA counties

County	Other or Two or more races	Identified races
Yakima	82.25%	17.75%
Franklin	88.52%	
Benton	71.59%	28.41%
Grant	87.78%	
Skagit	71.23%	28.77%
Chelan	89.15%	
Walla Walla	71.44%	
Mason		
Douglas	88.95%	
Lewis	70.83%	
Adams	91.54%	

Racial minority groups
- Identified races
- Other or Two or more races

Population

> **"** *...non-binary gender and data represents complicated terrain for computational applications for numerous reasons. But we have an ethical and empirical imperative to tackle this complexity.*
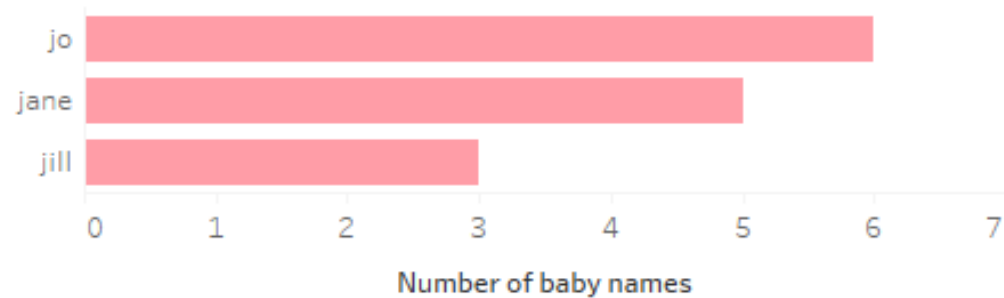
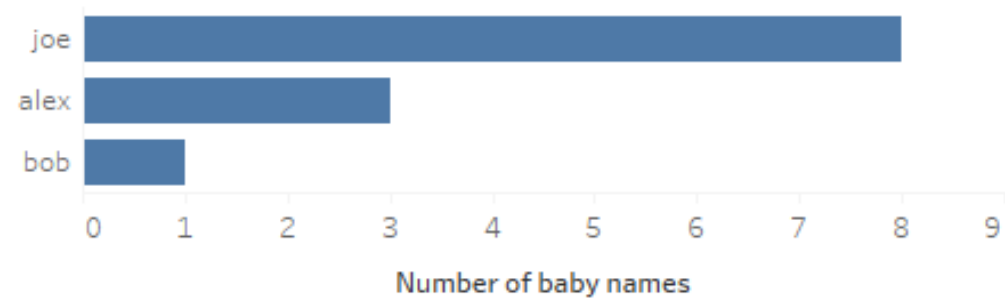— Catherine D'Ignazio

Binary Approaches

Baby names
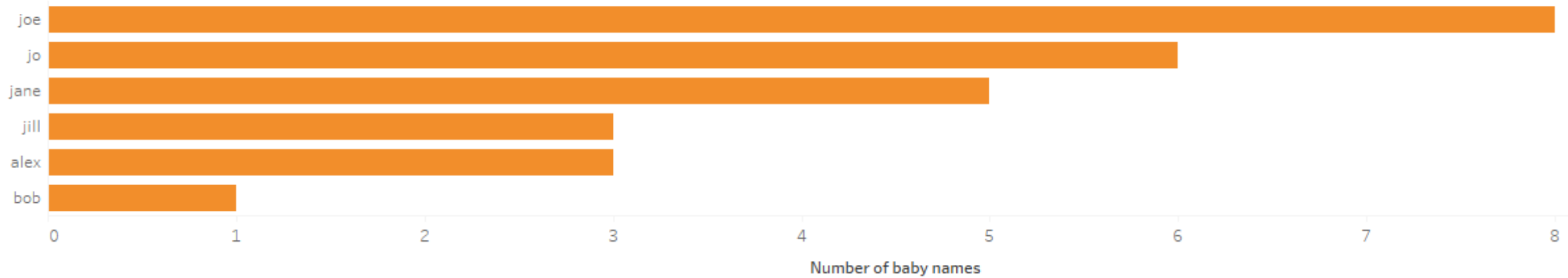Note: These charts were created using mock data.

Source: Mock data, chart created using Tableau

Non-Binary Gender and Data

Baby names

Note: This chart was created using mock data.

Source: Mock data, chart created using Tableau

Identifiable Data

When making data decisions, find a balance between being inclusive but also not identifying individuals.
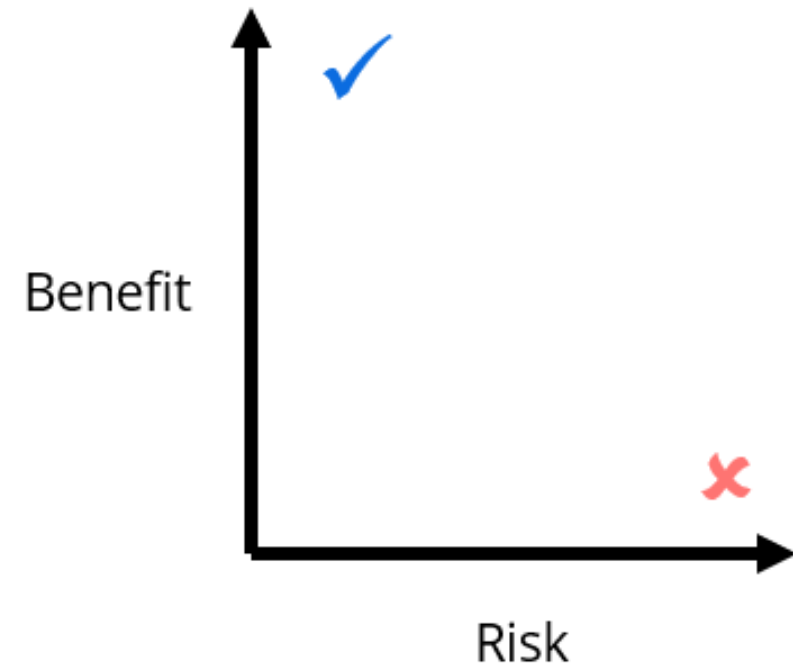
We Are Always Making Choices

- Collaborate with others to include multiple perspectives

- Be transparent in the decisions that you make when you collect, analyze, and visualize data

- Put care into how you speak about and represent the people who have contributed to your data

- Share data with the community that you are working with along the way

- Develop a custom dataset that pulls in non-Western, BIPOC-centered data sources, primary and secondary data, etc.

- Create a guide that explains the limitations of the data and what other sources to consider

- Legal standards

- Funders' standards

- Ethical standards →

- ✓ Visualizing responsibly and empathetically

- ✓ Understanding outcomes vs. intentions

- ✓ Remembering that data is not neutral

Selecting a topic

- Why are you selecting it?

- Who benefits from your work on this topic? Who may be harmed?

- Are you the right person to be doing this work?

Getting the data

- Whose labor went into gathering the data?

- What biases impacted the collection process?

- Who has control over the data? Does this conflict with who the data belongs to?

Cleaning the data

- Who cleaned the data?

- How did they clean the data?

- Who is made invisible, erased, or prioritized?

" One struggle in how we talk about data visualization is how we understand it: that goals, abstraction, and numbers intersect to make a whole. It's not just for the mathy types, but for anyone who seeks to understand their world through data. It's an artifact, one curated, logged, manipulated via calculations and comparisons, and finally sent out to the world as a message.

— Bridget Cogley

Source: Data Viz Philosophy: Better than Bar Charts, TableauFit Blog, December 1, 2020

Inclusive Practices

Question for the Viewers

?

*Think back to your image search for the term "CEO."
Do the images in the search results reinforce stereotypes?*

A. Yes

B. No

C. I'm not sure

NOTE: If you answered "**Yes**," please share your thoughts in the chat.

Inclusive Design Choices: Images

- Images and other visual representations of people are not immune from inequities in other parts of society

- They can reinforce harmful stereotypes, so if you search for an image or use an image that isn't yours, think critically about where and how it is used
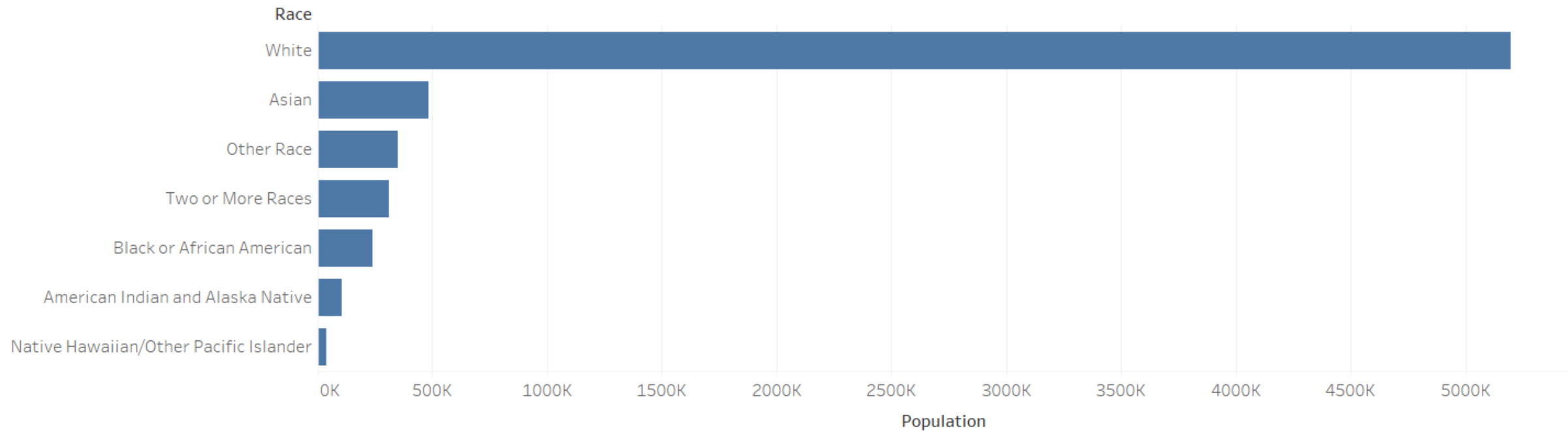
Inclusive Design Choices: Language

- Accessible and inclusive language applies to titles, labels you use

  - Example: CNN's 2020 election coverage

- Word order also matters

  - What the viewer sees can set the benchmark for what they compare to other categories

Inclusive Design Choices: Order

- What are your readers/viewers seeing first?

- What they see first can set the benchmark upon which they compare other categories

- Alternatives to categorizing by size: alphabetically, by stakeholder

- Who isn't represented? Who is misrepresented? Who is over or underrepresented?
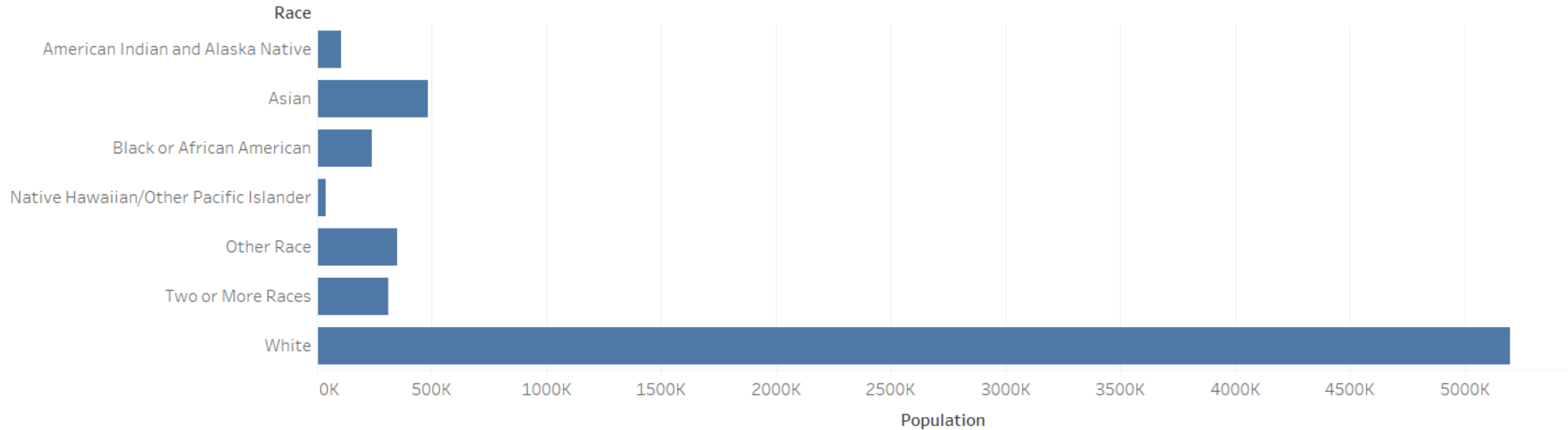
Example: Population Size

Population by race in WA

Race

White	
Asian	
Other Race	
Two or More Races	
Black or African American	
American Indian and Alaska Native	
Native Hawaiian/Other Pacific Islander	

0K   500K   1000K   1500K   2000K   2500K   3000K   3500K   4000K   4500K   5000K

Population

Example: Alphabetical

Population by race in WA

Race

Race	
American Indian and Alaska Native	
Asian	
Black or African American	
Native Hawaiian/Other Pacific Islander	
Other Race	
Two or More Races	
White	

Population: 0K, 500K, 1000K, 1500K, 2000K, 2500K, 3000K, 3500K, 4000K, 4500K, 5000K

Example: Stakeholder Group

Population by race in WA
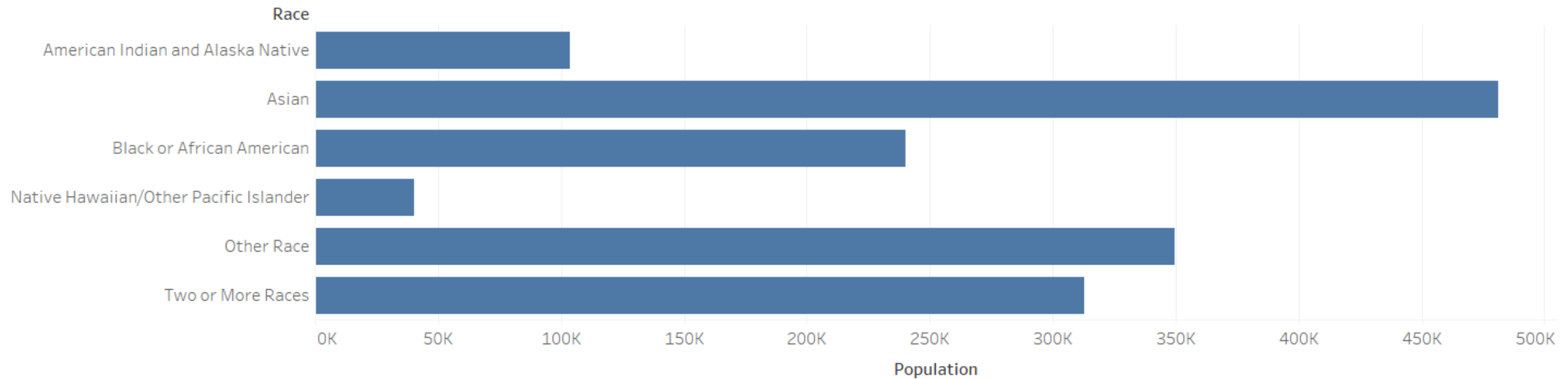
Race

	Population
American Indian and Alaska Native	~105K
Asian	~480K
Black or African American	~240K
Native Hawaiian/Other Pacific Islander	~40K
Other Race	~350K
Two or More Races	~310K

Population axis: 0K – 500K

In "Homegoing Stories," COVID Black uses the following inclusive design principles:
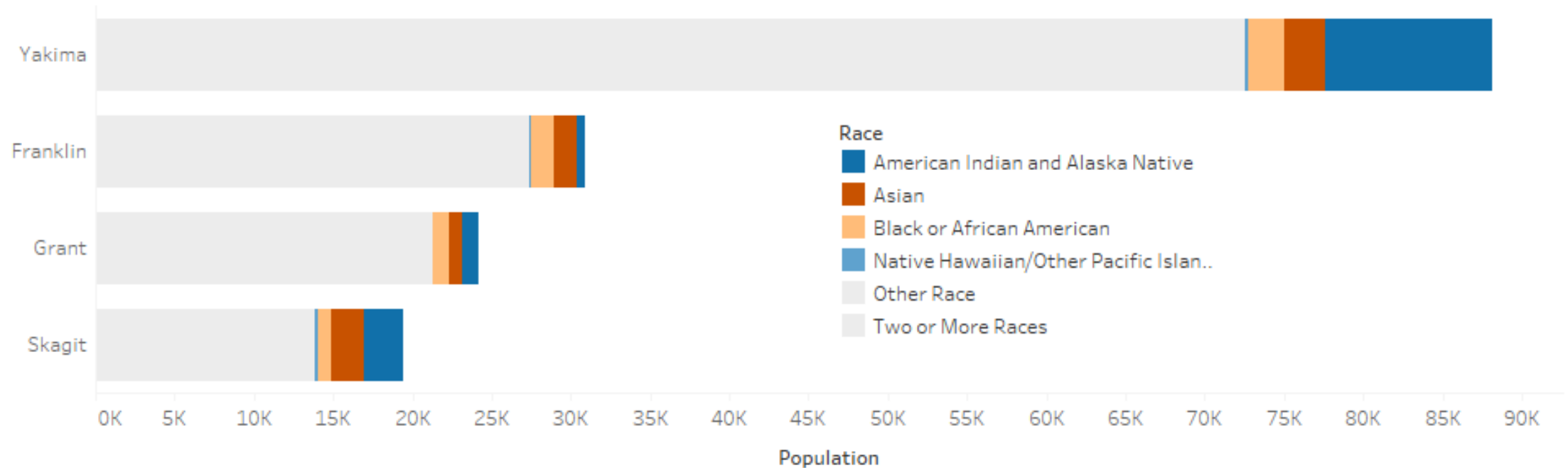
- Focuses on stakeholder group (Black Americans)

- Sorts names alphabetically

- Focuses on human elements

Inclusive Design Choices: Space and Color

- Are you creating visual divides in a way that excludes or prioritizes one group over another?

- How much space something takes up and its color in a visualization can correlate to perceived importance or value

    - Example: Using gray to represent "other" devalues those in that group

Example: Impact of Color

Population by race in WA (select counties)

Race
- American Indian and Alaska Native
- Asian
- Black or African American
- Native Hawaiian/Other Pacific Islan..
- Other Race
- Two or More Races

Mapping, History, and Political Issues

- Who set the borders — both historically and in the design of the map?

- Who is missing in the labels?

- Who is made invisible?

Mapping, History, and Political Issues

"...I looked up multiple maps of India and have not yet had two sources show the same map. This is most likely in relation to the Kashmir conflict which has been ongoing for nearly seventy years now. I do not work for Tableau but I am not sure they as an organisation are going to be willing to take sides in this dispute. The maps that Tableau use come from OpenStreetMap..."

Accessibility

- Using "the least amount of ink on the page" can make it more difficult for screen readers

- Alternative text, color choices

- Designing around bandwidth

  - Example: Visualizations that require high bandwidth internet when your key population, community, or stakeholders do not have access to high bandwidth can make it difficult for them to access it

Accessibility: Color

How many times did you snooze your alarm in January?
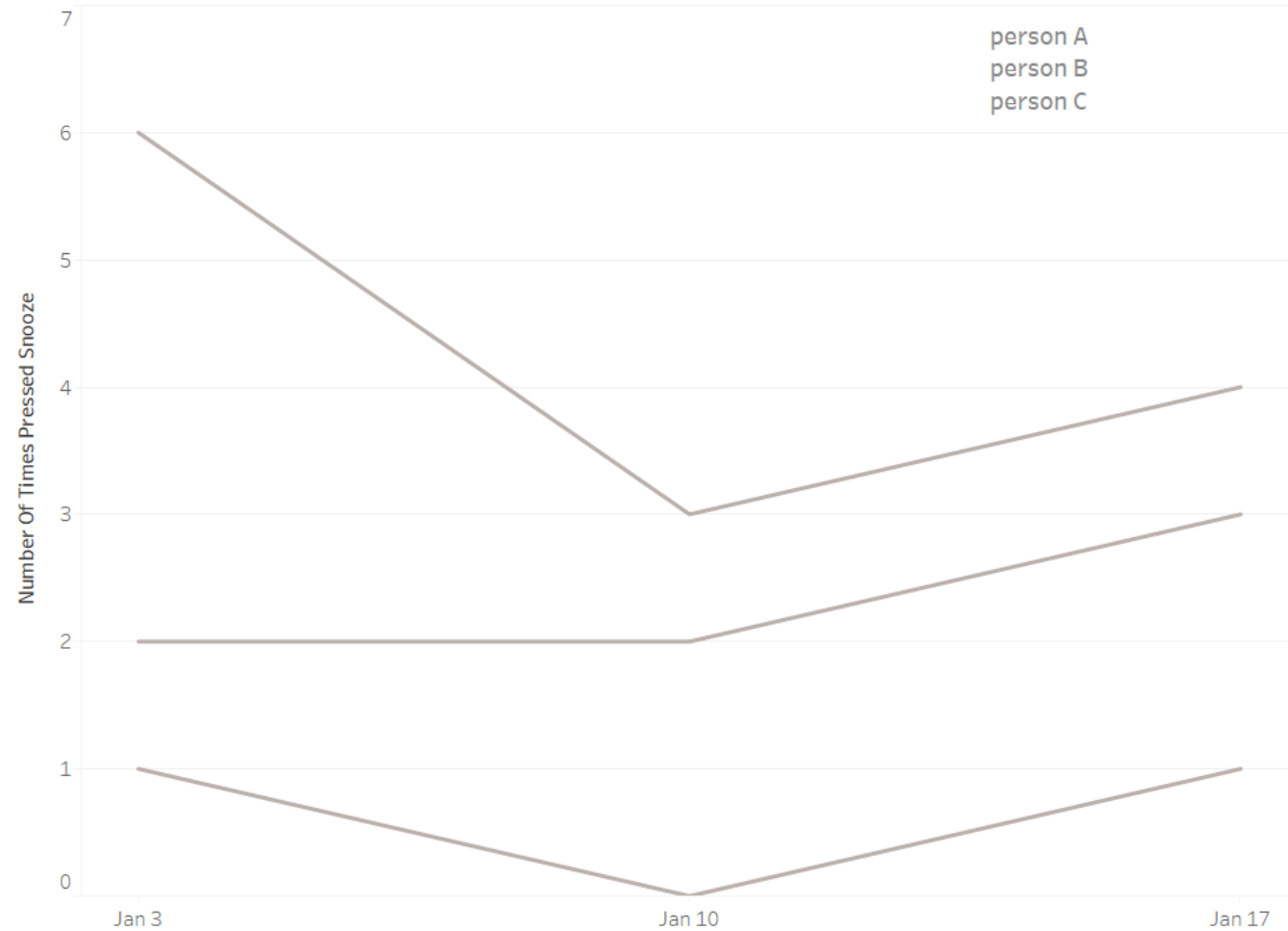**Note:** This chart was created using mock data.
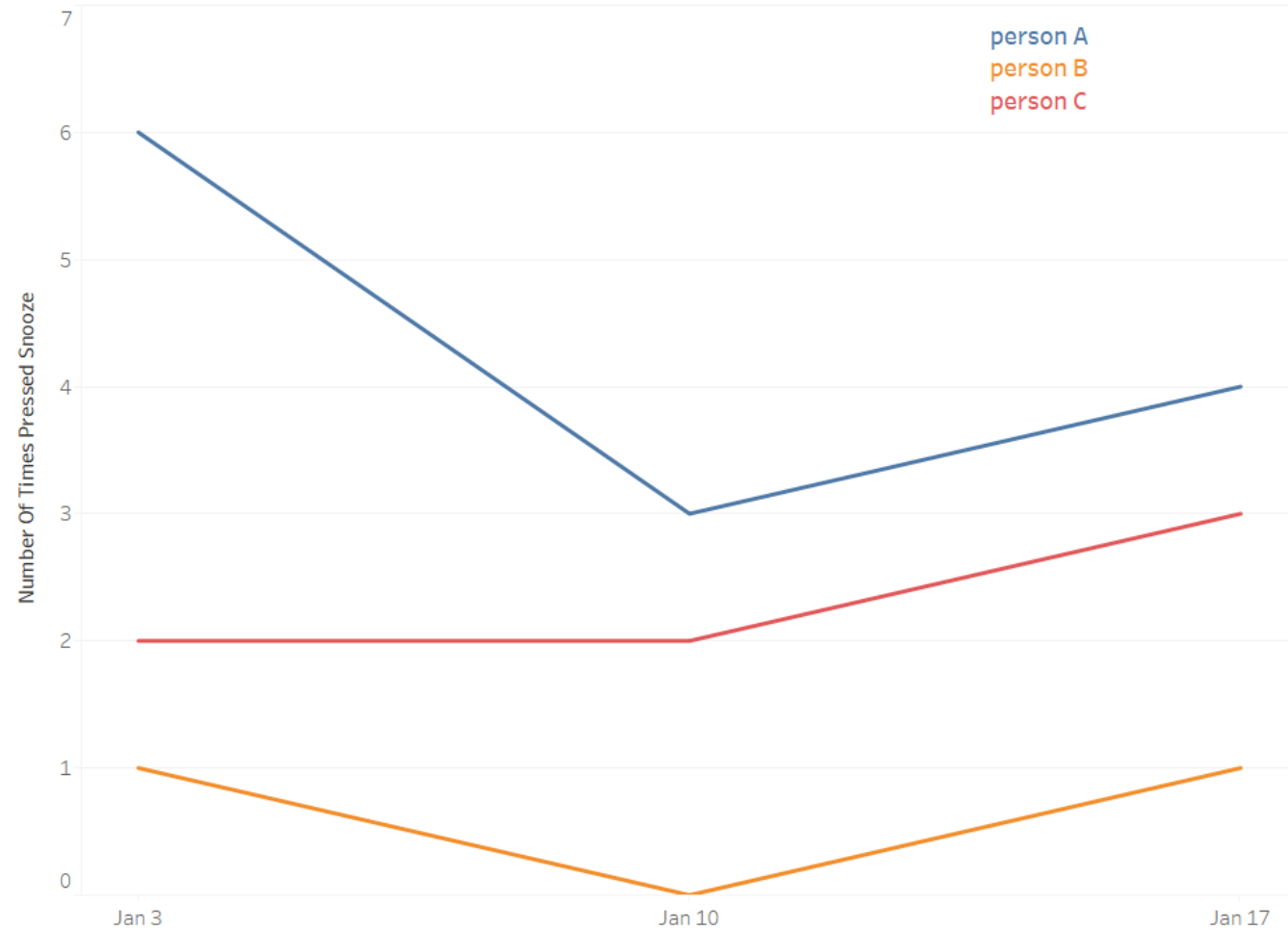
Source: Mock data, chart created using Tableau

Accessibility: Only Using Color

How many times did you snooze your alarm in January?
Note: This chart was created using mock data.

person A
person B
person C

Number Of Times Pressed Snooze

Jan 3   Jan 10   Jan 17

Source: Mock data, chart created using Tableau

Accessibility: Color and Shape Changes

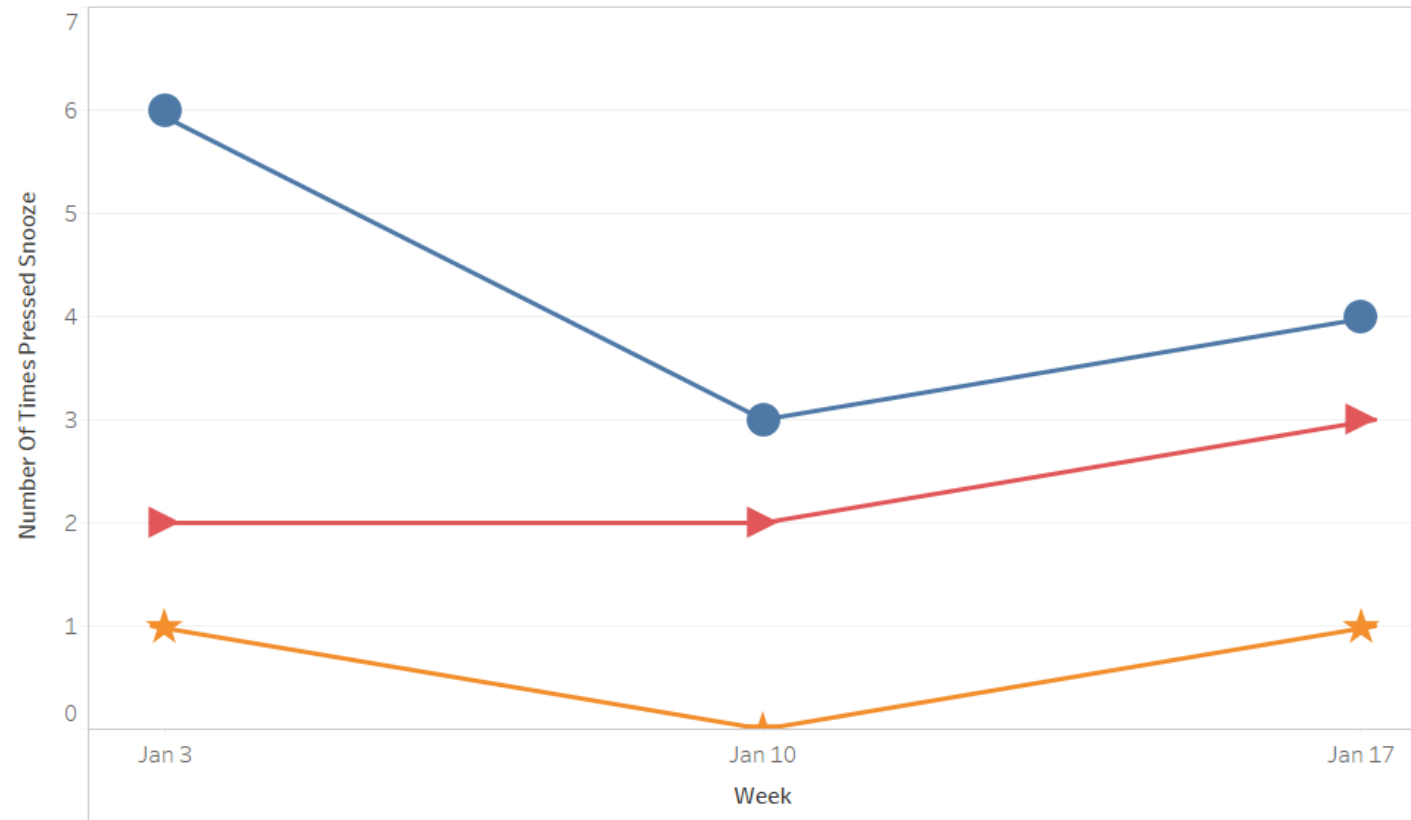How many times did you snooze your alarm in January?
Note: This chart was created using mock data.

Person
● person A
★ person B
▶ person C

Line chart that shows the number of times person A, person B, and person C snoozed their alarms over the course of a few weeks in January. The chart indicates that alarms were snoozed at higher rates immediately following the holidays.

Source: Mock data, chart created using Tableau

Key Takeaways

✓ Data and data processes do not always capture complexity and are not always inclusive.

✓ Whenever possible, work *with* your community on data collection and sharing processes.

✓ Mark any areas of uncertainty or data quality issues on a chart or in an accompanying note/guide.

✓ Consider the colors and language you use when creating charts. Are they inclusive? Do they prioritize or minimize any groups?

✓ Create more accessible charts by using multiple elements (such as shape *and* color) to communicate the information.

QUESTIONS?

To ask a question, please click the [Q&A] icon in the Zoom toolbar to open your Q&A Pod.

Resources

Abigail Echo-Hawk on the art and science of 'decolonizing data'
Crosscut.com article by Manola Secaira
https://crosscut.com/2019/05/abigail-echo-hawk-art-and-science-decolonizing-data

Chartability (chart accessibility)
Frank Elavsky and Fizz Studio
https://chartability.fizz.studio/

COVID Black's Homegoing Geographic and Category Data
Tableau Dataset by COVID Black
https://public.tableau.com/app/profile/covid.black/viz/HomegoingGeographicandCategoryData/HomegoingStories

Data Feminism
Catherine D'Ignazio and Lauren F. Klein – The MIT Press
https://data-feminism.mitpress.mit.edu/

Data Humanism, The Revolution will be Visualized
Data Humanism – A Visual Manifesto by Giorgia Lupi
http://giorgialupi.com/data-humanism-my-manifesto-for-a-new-data-wold

Indigenous Identity: More Than "Something Else"
Native Americans in Philanthropy
https://nativephilanthropy.org/2020/11/24/indigenous-identity-more-than-something-else/