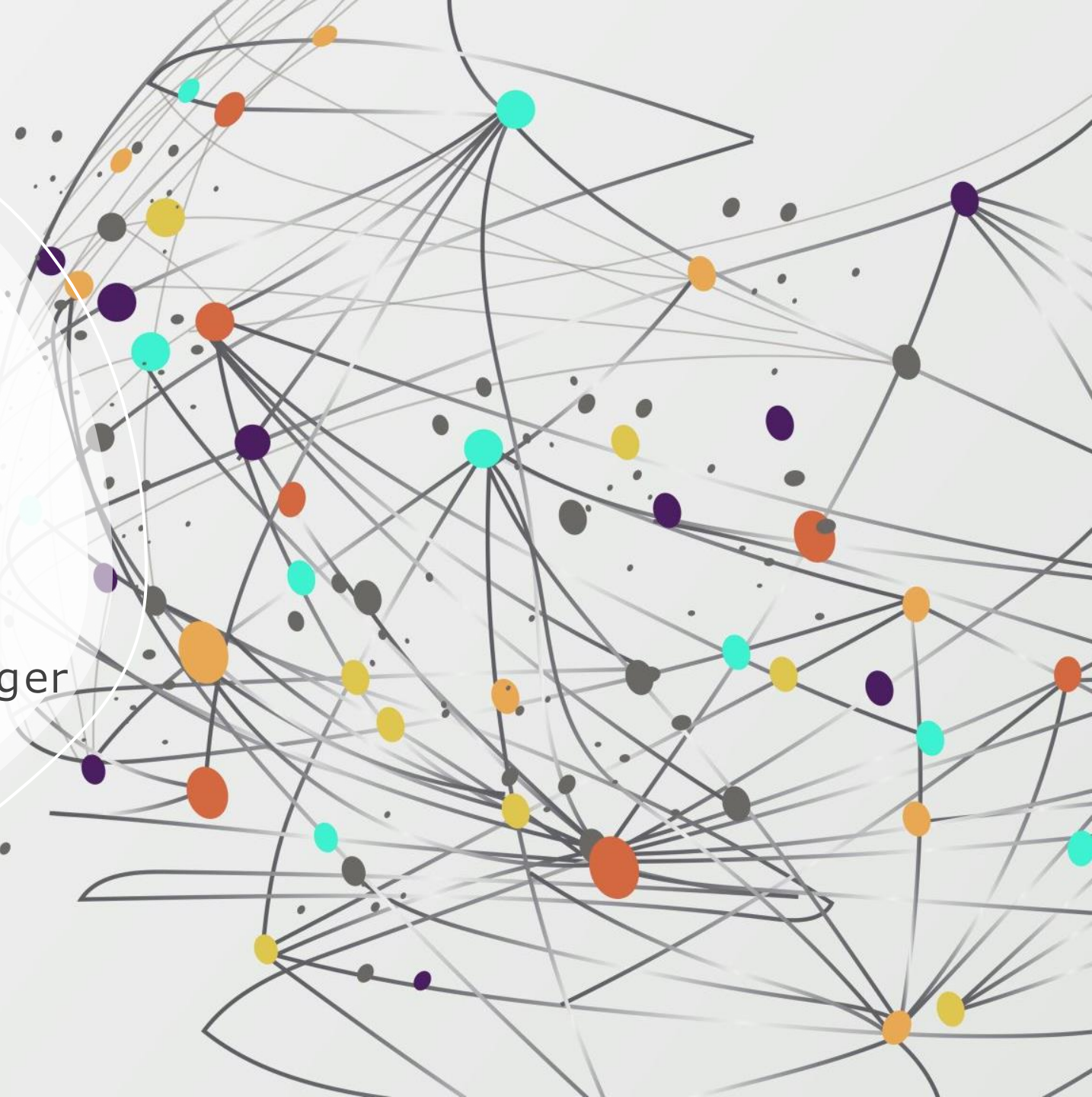


Considering Equity in Data Visualization

Silvia Canelón & Emily Seeburger

July 10, 2024



Agenda

We'll spend time today:

1. Introducing principles of equity in data analysis and visualization
2. Walk through some examples
3. Play a game!

Way to think about data



Summarize stats

Describing basic information about our data



Communicating impact

Highlighting outcomes related to a program, policy, or intervention



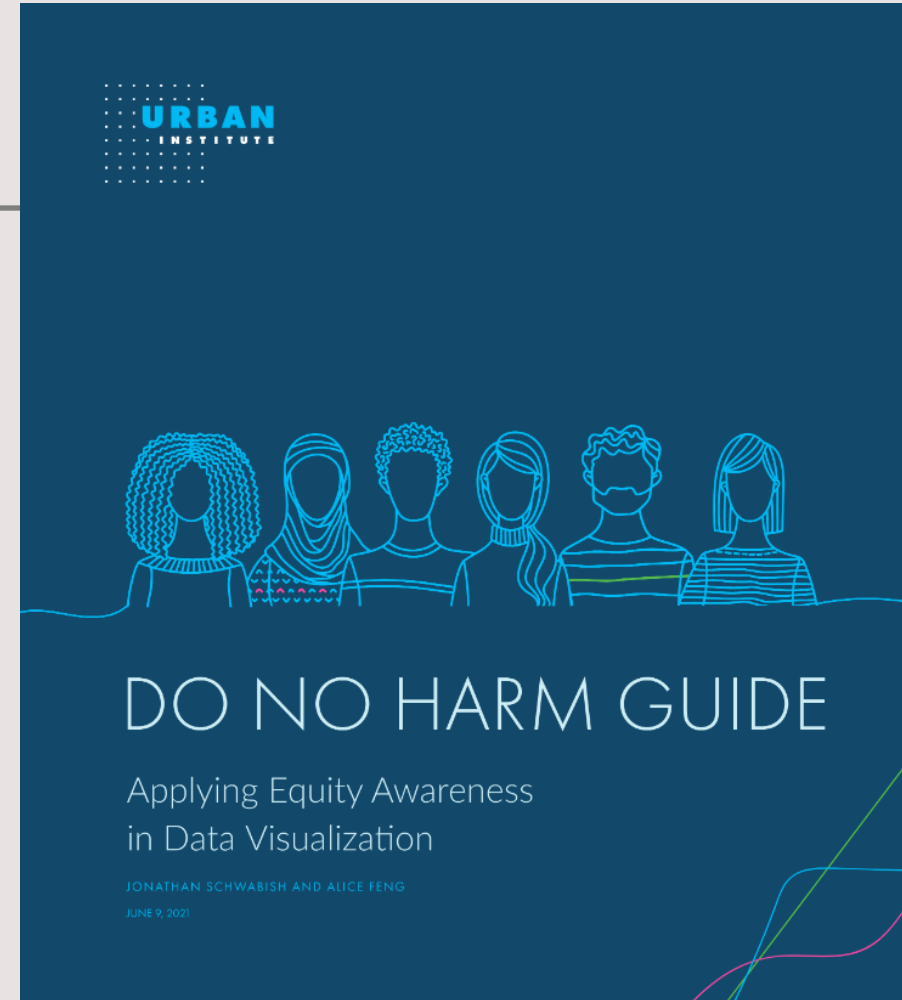
Community engagement & storytelling


Generating ideas and conversations

Urban Institute

The Urban Institute has put together a series of **Do No Harm Guides** on applying equity awareness in data visualization

We can lean on this work to help inform best practices for any research lab!





**Engaging or
reflecting lived
experiences**

Engaging & reflecting lived experience

- Engage community **early & often**
- **Be transparent** about our research goals and what we intend to do with the data
- Understand community needs and **how the data can be used to help**
- **Ask for community perspectives** on the data
- **Create a plan to disseminate findings** after the research is complete

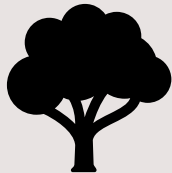
Deeply Rooted: Community-level data



Community health:
Heart Disease



Community violence:
Fatal & nonfatal shootings

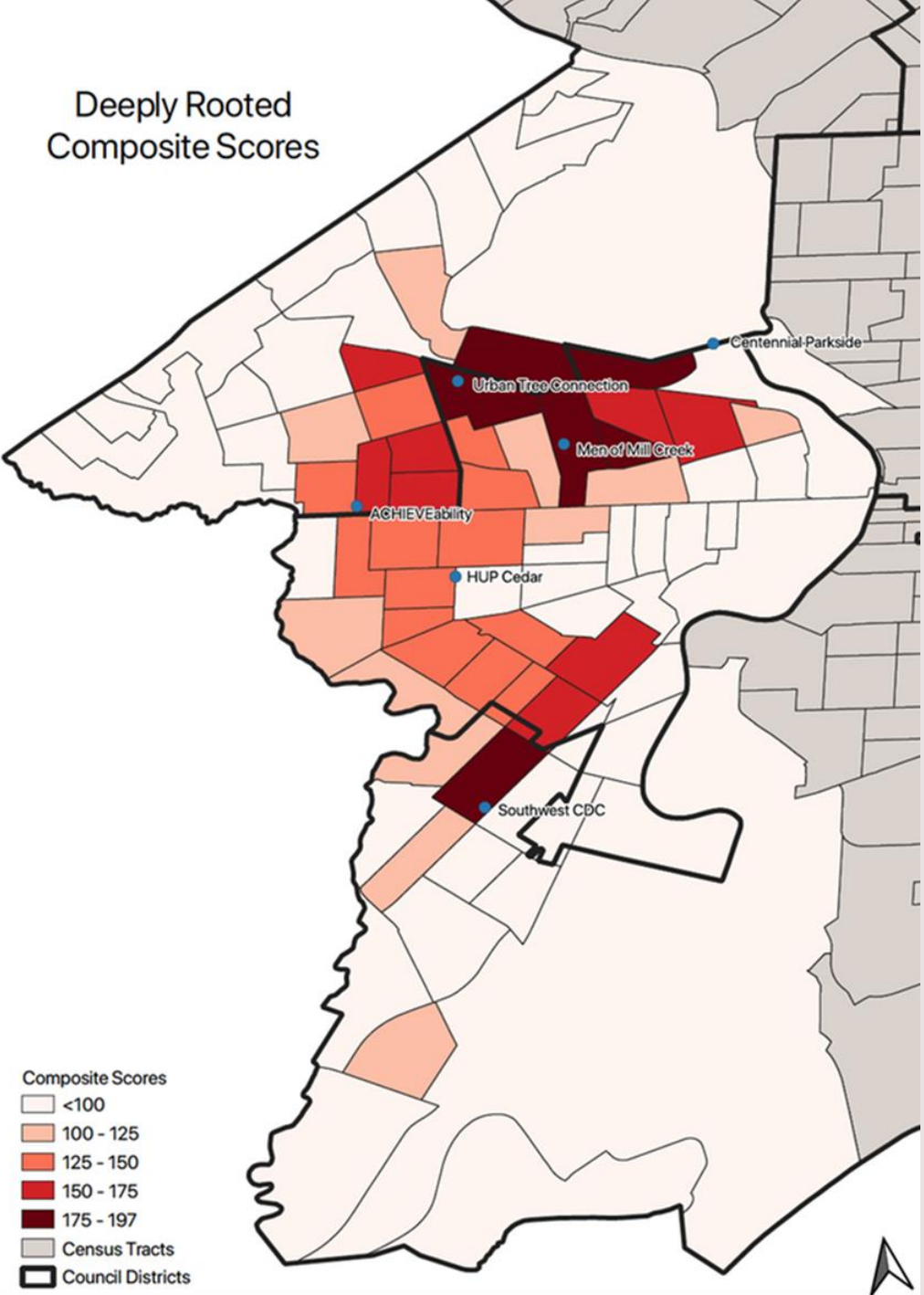


Urban nature:
Tree Equity Score



Vacancy:
Concentration of vacant lots

Deeply Rooted Composite Scores





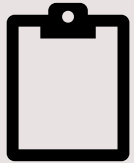
The background features a light gray gradient with several overlapping, rounded rectangular shapes. These shapes are defined by multiple thin, concentric lines in shades of gray and white, creating a layered, organic effect. The shapes are positioned in the upper and left portions of the frame, leaving the lower right area more open.

Language

Using language to describe people

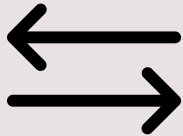
- Terms and phrases used to describe people are **always evolving**
- Use **people-first language** in labels, text, titles (e.g., people with disabilities, incarcerated people)
- Monitor current lexicon and **reflect the experiences of the people** we study and communicate with

Changing source data terminology



Surveys or datasets sometimes use **outdated** or **less preferred** terminology.

What can we do?



In publications, **change** to a new term and explain the decision and process within the text



In publications, **use the original** language and include an explanation of why this language is not preferred within the text

Ordering data with purpose

- Large demographic surveys in the US, like the census, order race starting with “White” and “Black” as the first two options
- “White” is often listed as the first option and coded as “1” in the data

9. What is Person 1's race?

Mark one or more boxes **AND** print origins.

- White – Print, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc. ↴

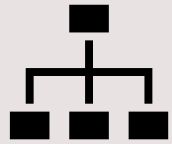
- Black or African Am. – Print, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc. ↴

- American Indian or Alaska Native – Print name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, etc. ↴

- | | | |
|---|-------------------------------------|--|
| <input type="checkbox"/> Chinese | <input type="checkbox"/> Vietnamese | <input type="checkbox"/> Native Hawaiian |
| <input type="checkbox"/> Filipino | <input type="checkbox"/> Korean | <input type="checkbox"/> Samoan |
| <input type="checkbox"/> Asian Indian | <input type="checkbox"/> Japanese | <input type="checkbox"/> Chamorro |
| <input type="checkbox"/> Other Asian –
Print, for example,
Pakistani, Cambodian,
Hmong, etc. ↴ | | <input type="checkbox"/> Other Pacific Islander –
Print, for example,
Tongan, Fijian,
Marshallese, etc. ↴ |

- Some other race – Print race or origin. ↴

Ordering data with purpose



Consider how readers will perceive the **relationship or hierarchy** between groups



If our study focuses on a **particular community**, that group should be presented first



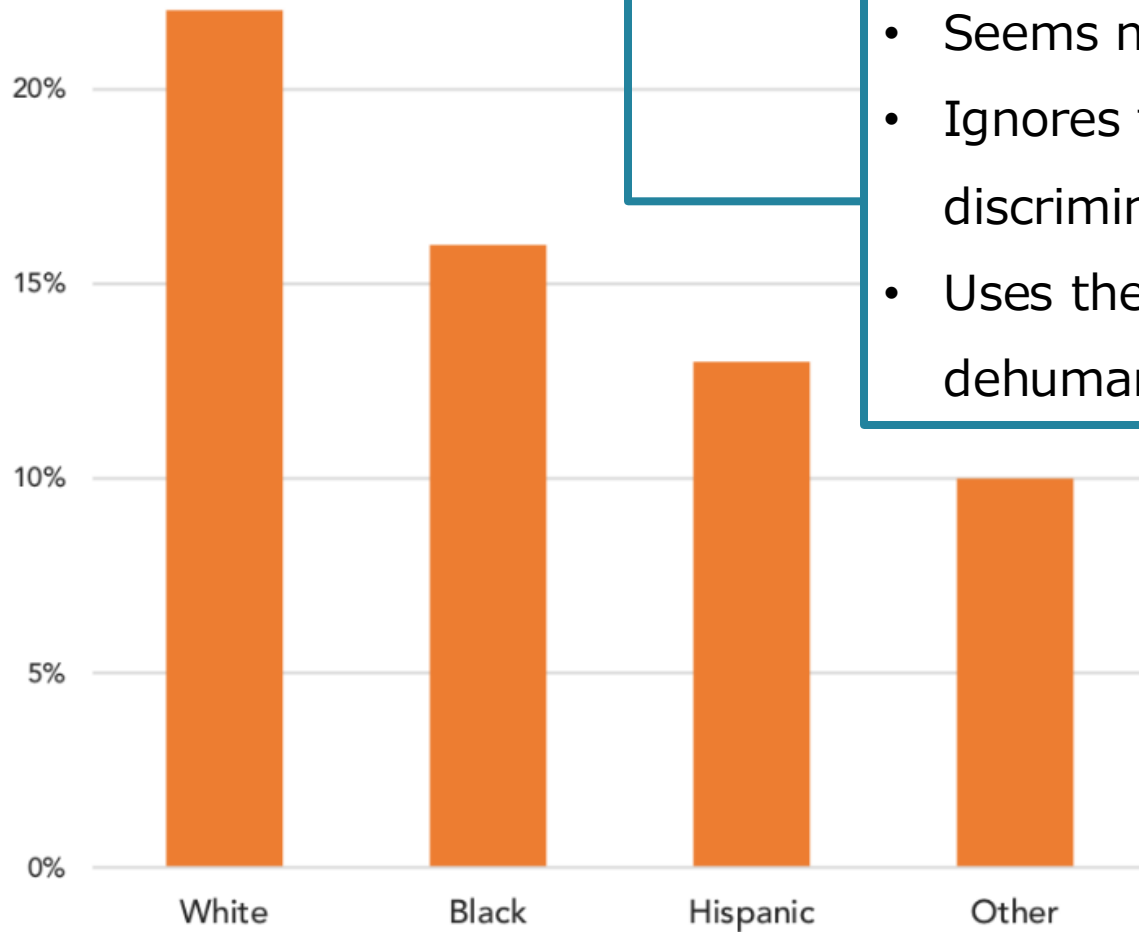
Consider whether the results could be presented according to a **quantitative relationship** (e.g., population size, magnitude of the results)

Writing titles that do some work

- Use concise, active titles that communicate **main findings** *and* **the role played by oppressive forces**
- It's important that these ideas be reflected **directly in the chart** rather than buried in surrounding text

Mental Health in Jail

Rate of mental health diagnosis of inmates

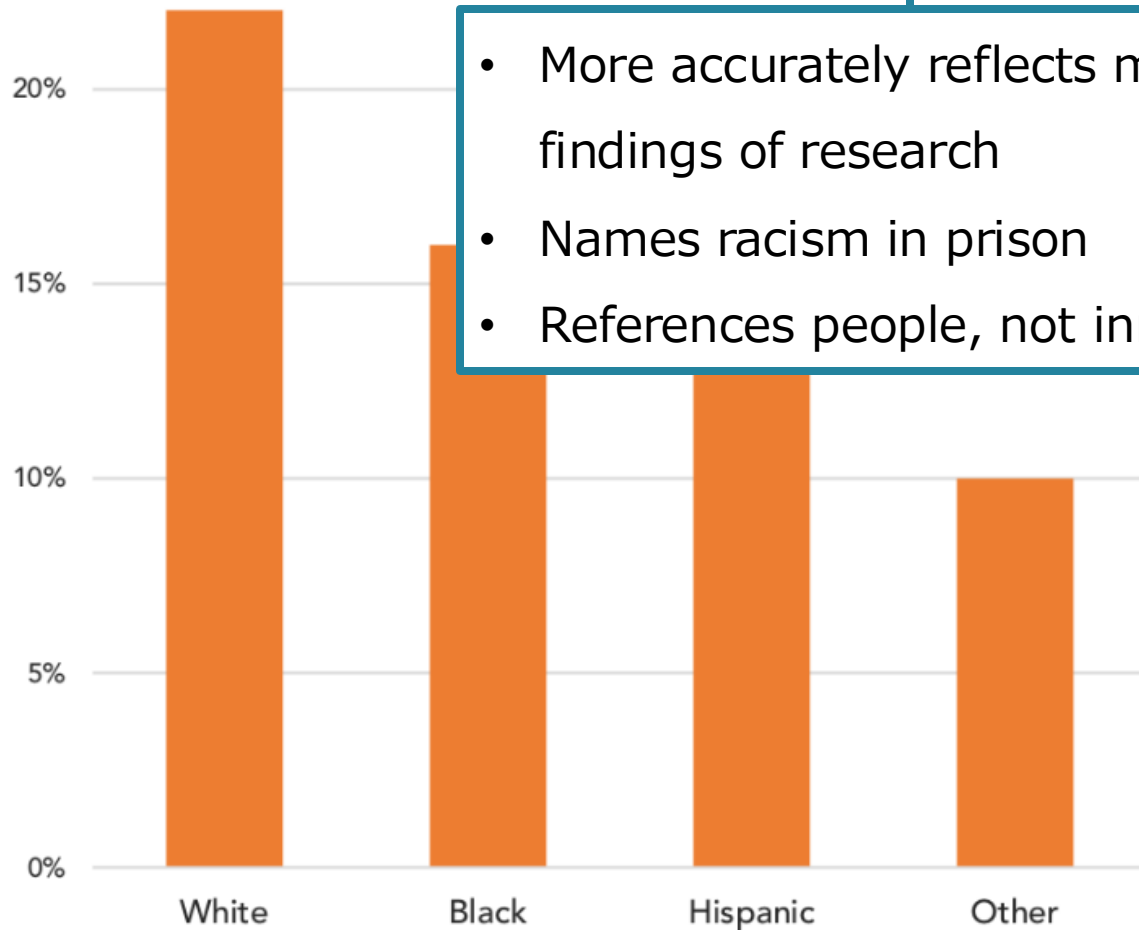


- Seems neutral and objective
- Ignores the role of racism and discrimination at play
- Uses the term “inmate” which is dehumanizing

Source: D'Ignazio and Klein (2020); Data from Fatos Kaba et al., “Disparities in Mental Health Referral and Diagnosis in the New York City Jail Mental Health Service,” *American Journal of Public Health* 105, no. 9 (2015): 1911–16.

Mental Health in Jail

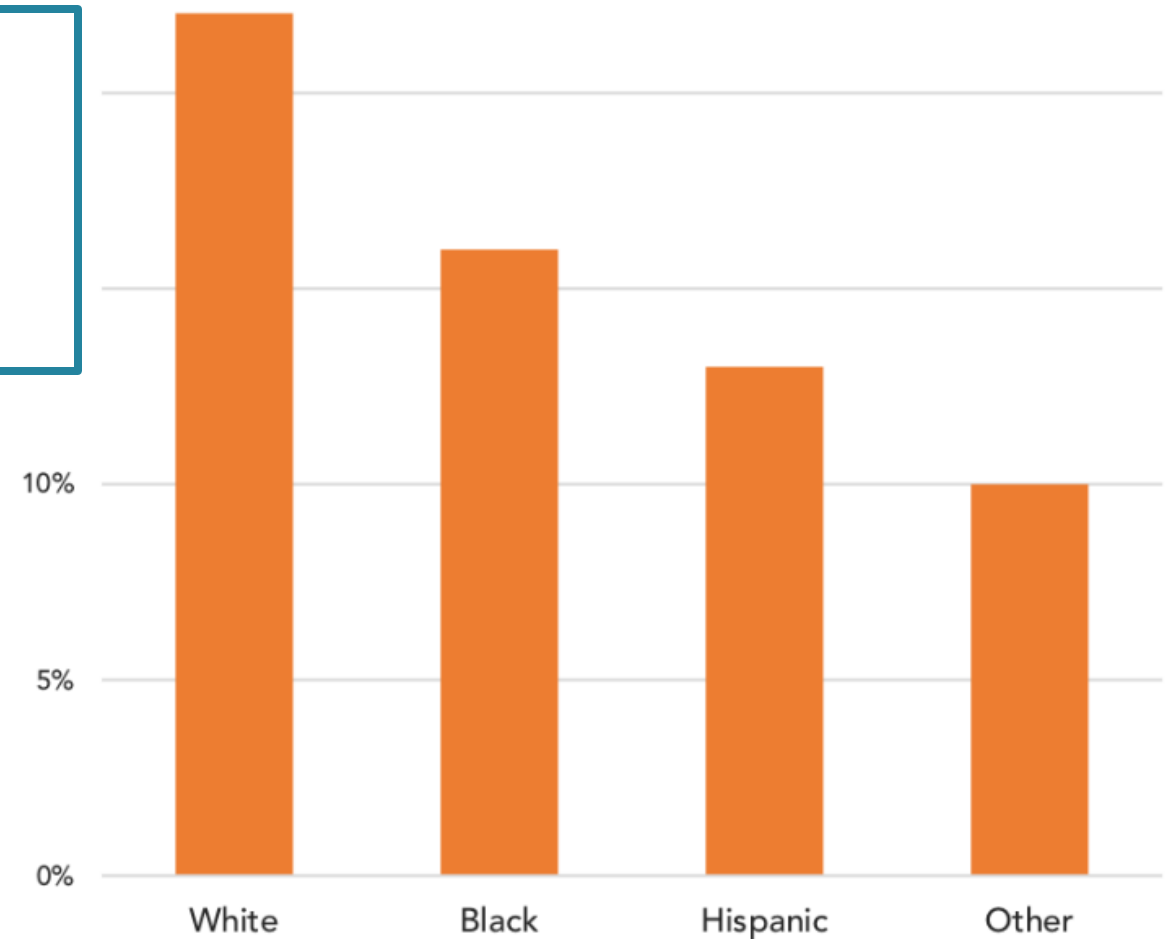
Rate of mental health diagnosis of inmates



- More accurately reflects main findings of research
- Names racism in prison
- References people, not inmates

Racism in Jail

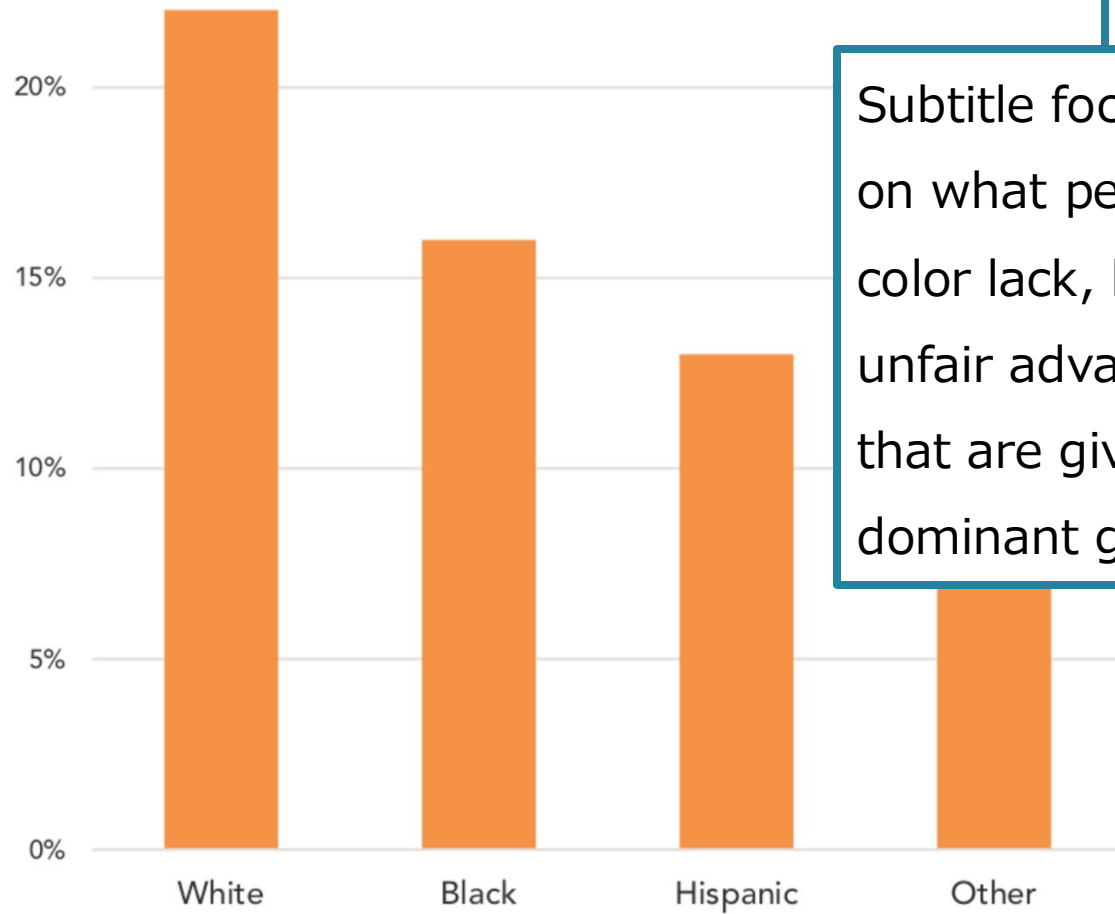
People of color less likely to get mental health diagnosis



Source: D'Ignazio and Klein (2020); Data from Fatos Kaba et al., "Disparities in Mental Health Referral and Diagnosis in the New York City Jail Mental Health Service," *American Journal of Public Health* 105, no. 9 (2015): 1911-16.

Racism in Jail

People of color less likely to get mental health diagnosis



Subtitle focuses not on what people of color lack, but on the unfair advantages that are given to the dominant group



**Colors, icons,
images & shapes**

Colors, Icons, Images & Shapes

- **Colors and images can add interest and depth to our data visualizations, but they can also limit accessibility and stigmatize our research participants**
- **This is especially true for work related to equity, substance use, gun violence, and poverty**

FEATURE

Trapped by the 'Walmart of Heroin'

A Philadelphia neighborhood is the largest open-air narcotics market for heroin on the East Coast. Addicts come from all over, and many never leave.



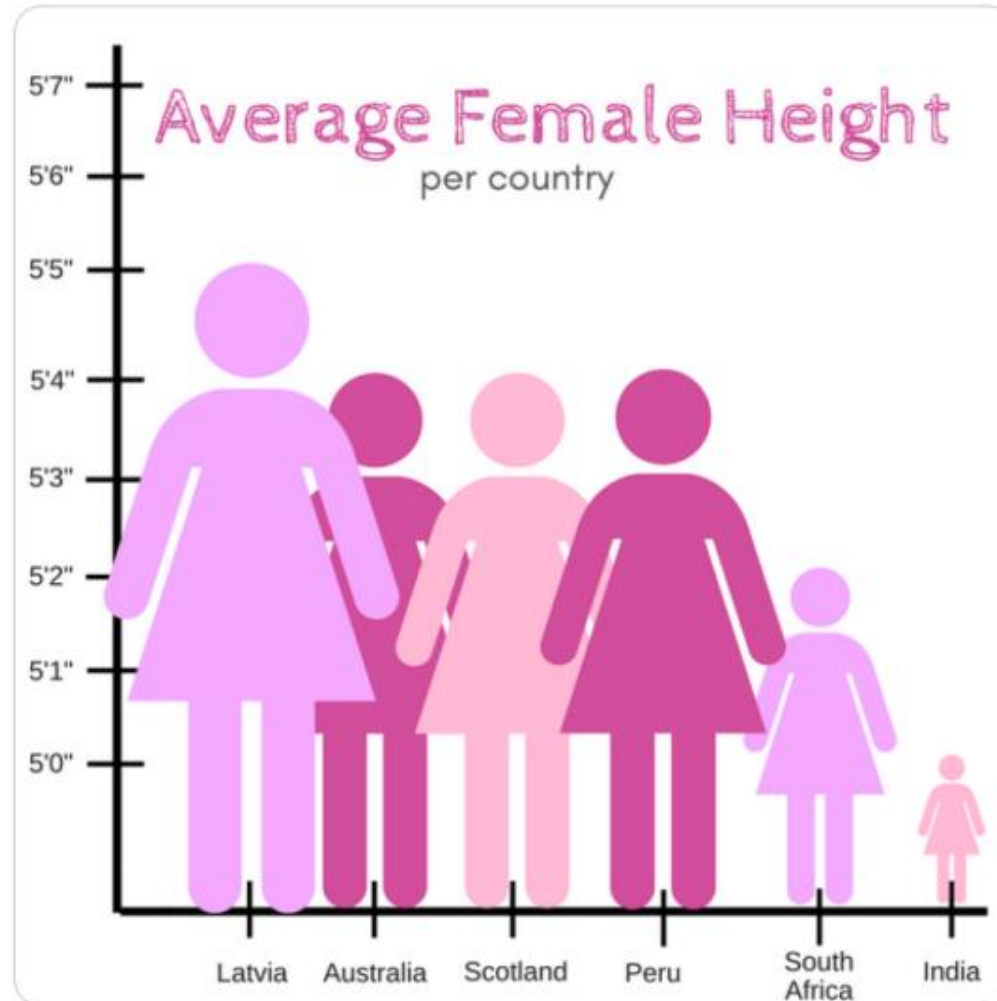


Sabah Ibrahim

@reina_sabah

...

As an Indian woman, I can confirm that too much of my time is spent hiding behind a rock praying the terrifying gang of international giant ladies and their Latvian general don't find me



Tools for color

- **Checking for color accessibility:**
<https://www.color-blindness.com>
- **Developing our own color palette:**
<https://colorbrewer2.org>

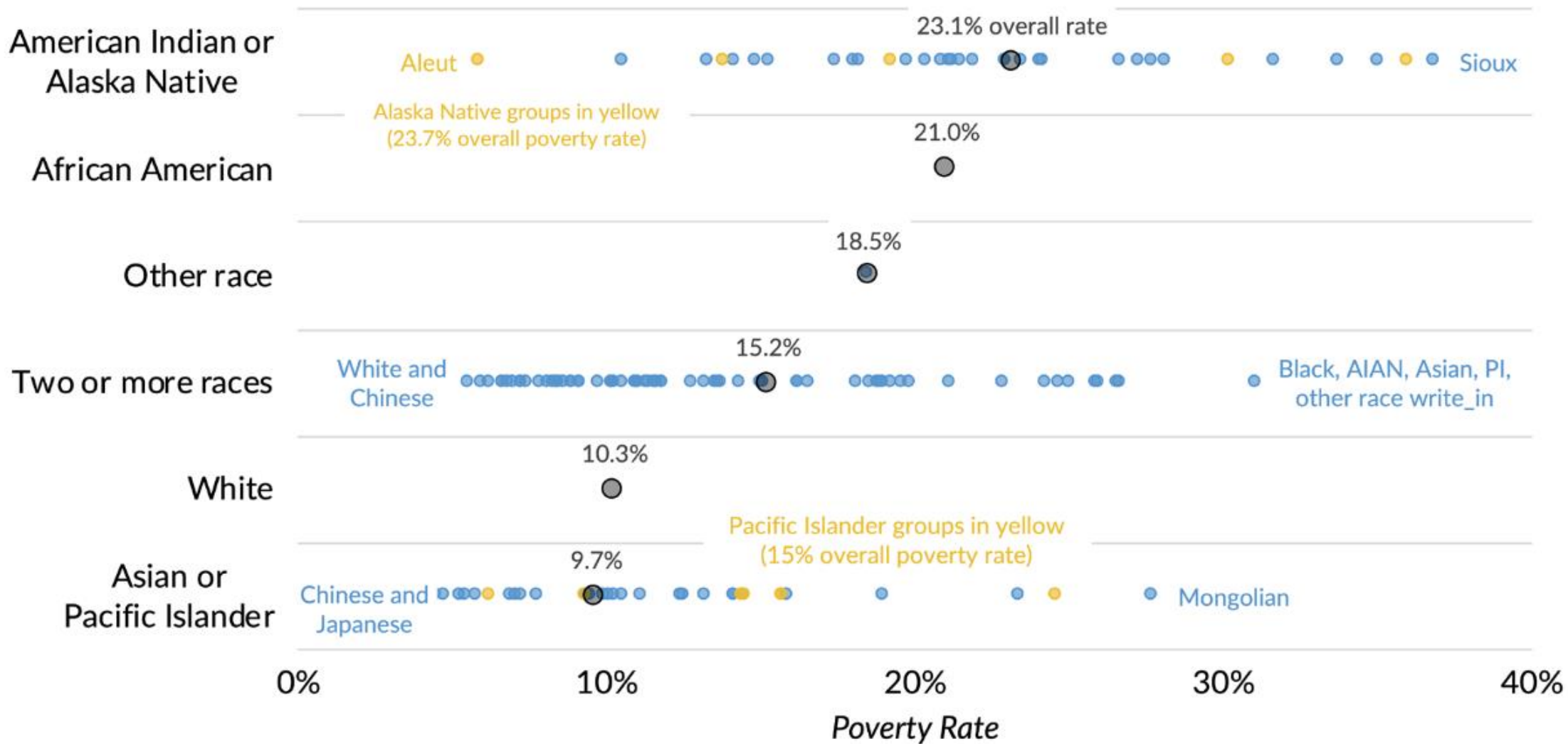


Considering missing groups

Lumping populations together

- Sometimes we lump together populations with **few observations** to make analysis more convenient
- Lumping together racial and ethnic groups can **mask important variation**

Variation in Poverty Rates



Source: Schwabish and Feng (2021).

Data: American Community Survey

Lumping populations together

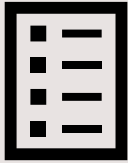


Reasons for not including more racial groups in analyses is often **sample size**



Sample sizes depend on actual size of the group, staffing, funding, as well as **active choices** made by the surveying organization, survey funders, and society

Being limited by category options



Major federal surveys, and electronic health records, are **limited in the options** they provide for gender, race, and other demographic variables

What can we do?



Make it clear to the reader **what information we have and do not have** in the data



Consider explicitly **calling out the absence** of data

* Gender:

✓
Female
Male
N/A
Unknown
Tax Entity

Does the Beneficiary reside a

* Contact Phone:

GENDER	Female
BIRTHDATE	Male
LANGUAGE	Female
	Human

Spouse's gender.*

Female

Male


Indeterminate

Othering the Other

- The “Other” category might include people with **varied characteristics and life experiences**
- Consider whether grouping populations this way is actually **meaningful**
- The term “**other**” **literally others individuals**, and emphasizes how they are different from the norm

Using alternatives to “other”

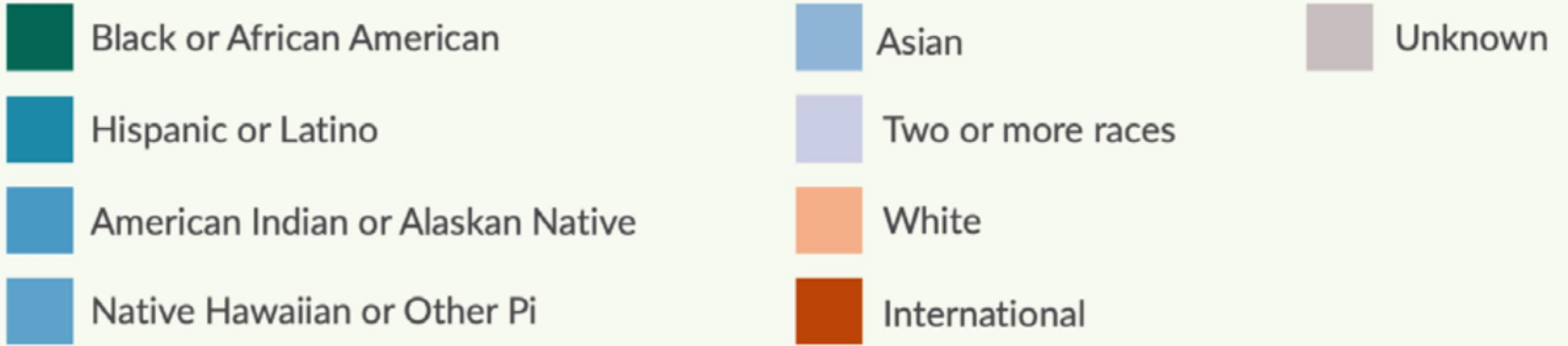
- Another race
- Additional groups
- All other self-descriptions
- People identifying as other or multiple races
- Identity not listed
- Identity not listed in the survey



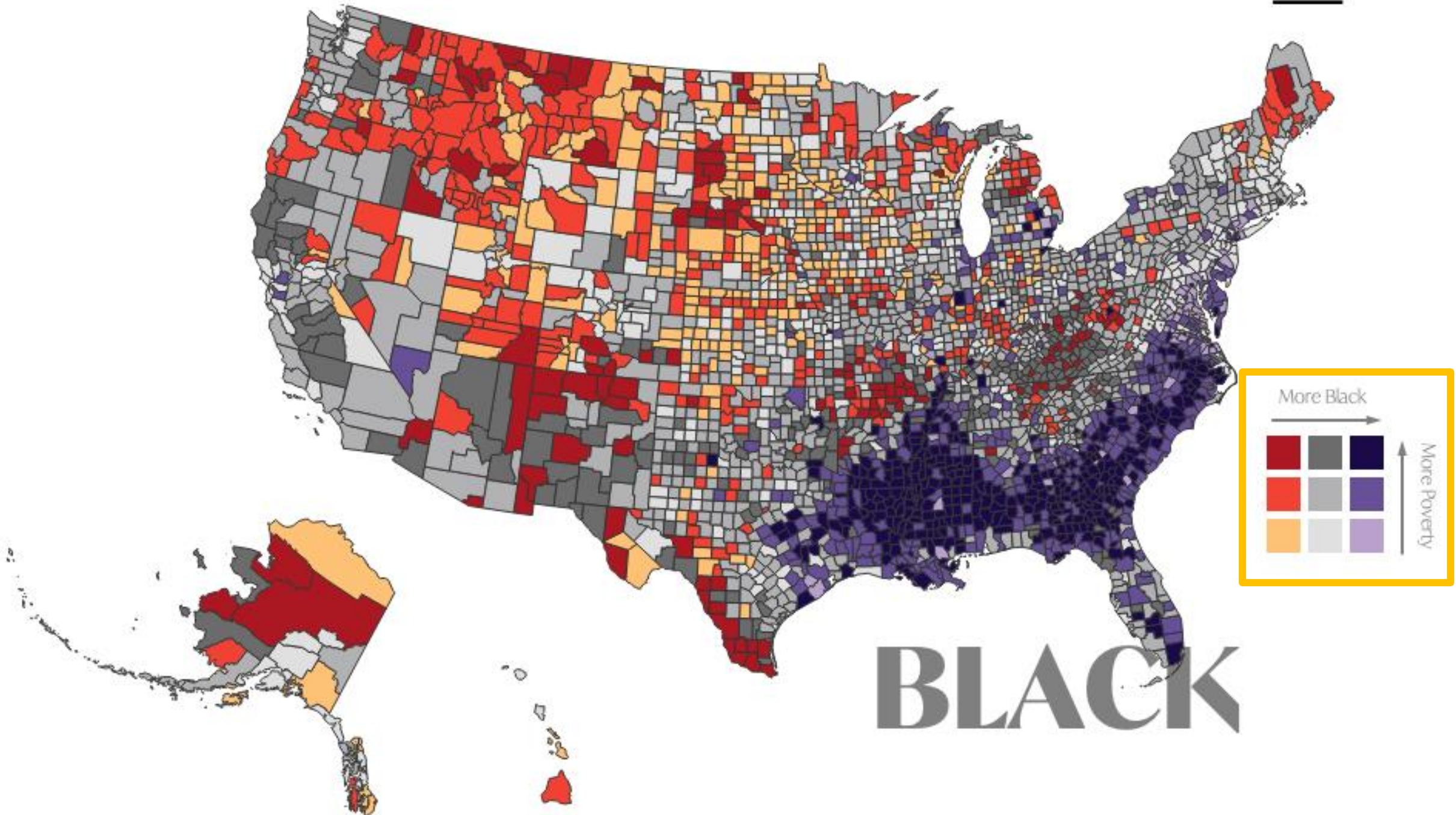
Game time!
Spot what's wrong



June 2020



July 2020

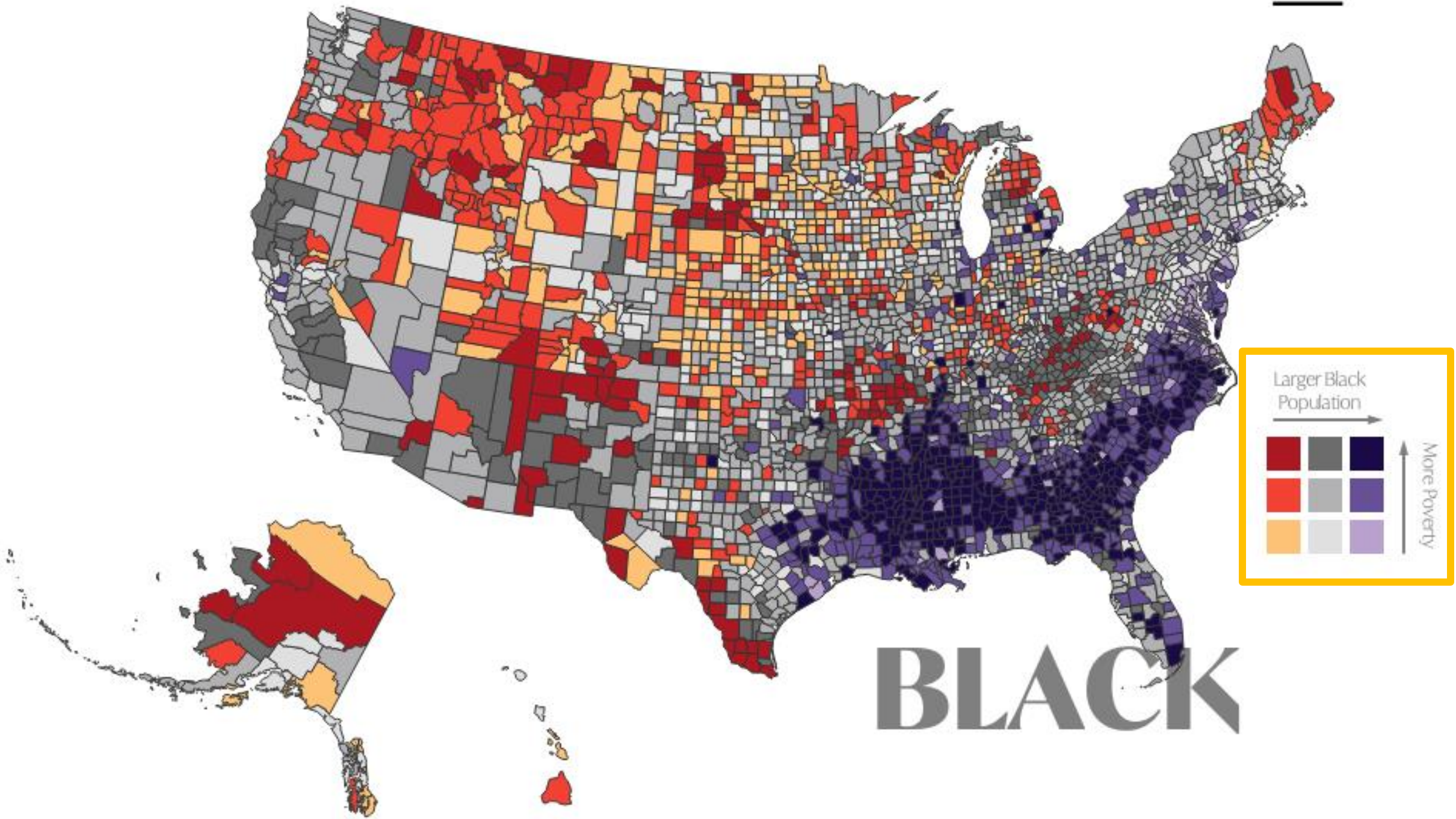


BLACK

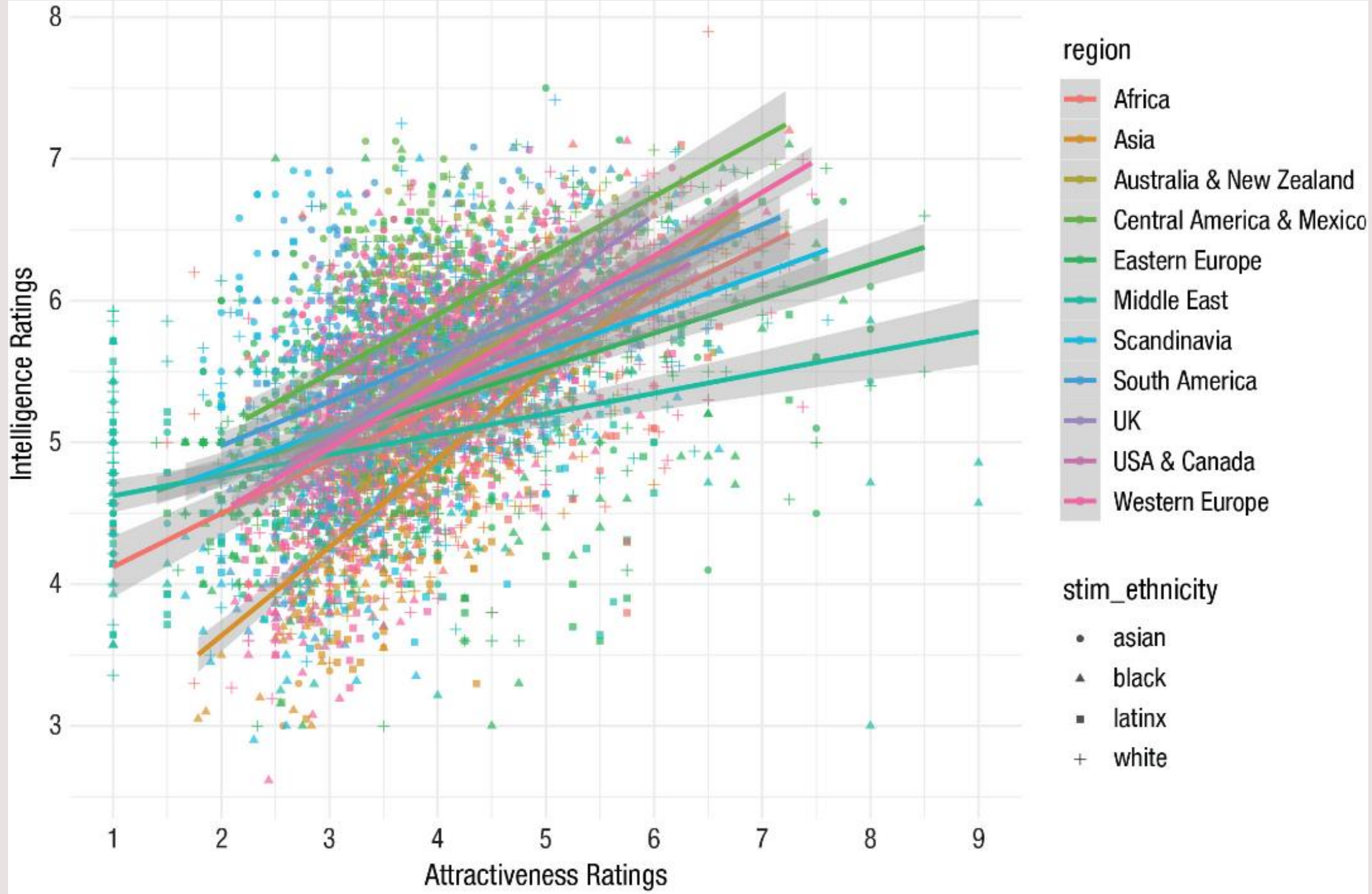
More Black



More Poverty



BLACK





Thank you!