PRB
Informing A Smarter World

Disaggregating American Community Survey Data by Race and Ethnicity

Alicia VanOrman

September 20, 2022

Today's presentation

Importance of disaggregating data by race and ethnicity

Methodological and analytic challenges

Examples

Importance of disaggregating data by race and ethnicity

Total is more than a sum of its parts

Better understanding of the strengths, needs, and quality of life

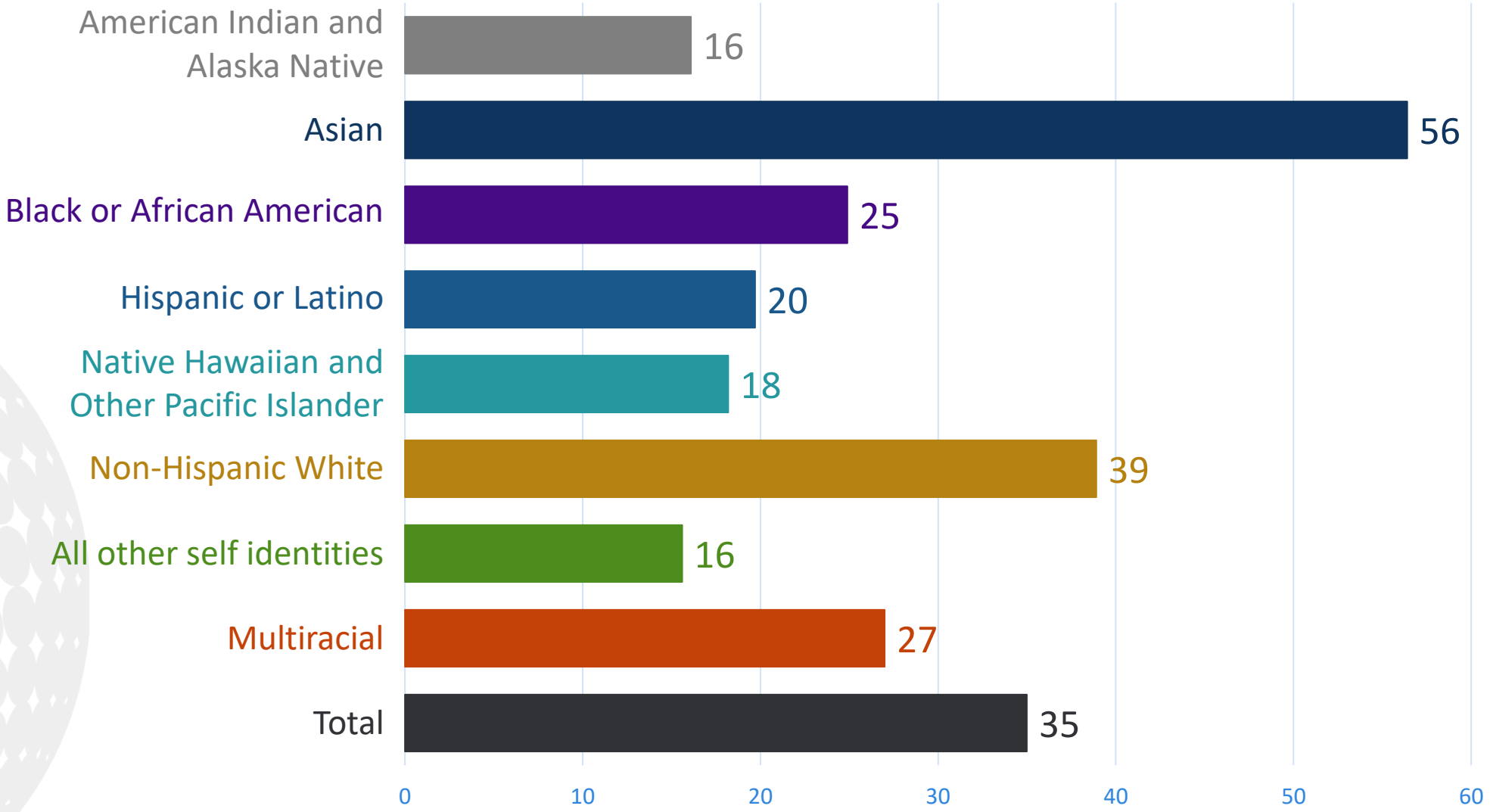
Identify, measure, and track racial equity gaps

- Transform equity from a concept to a measurable objective

Informs public policy

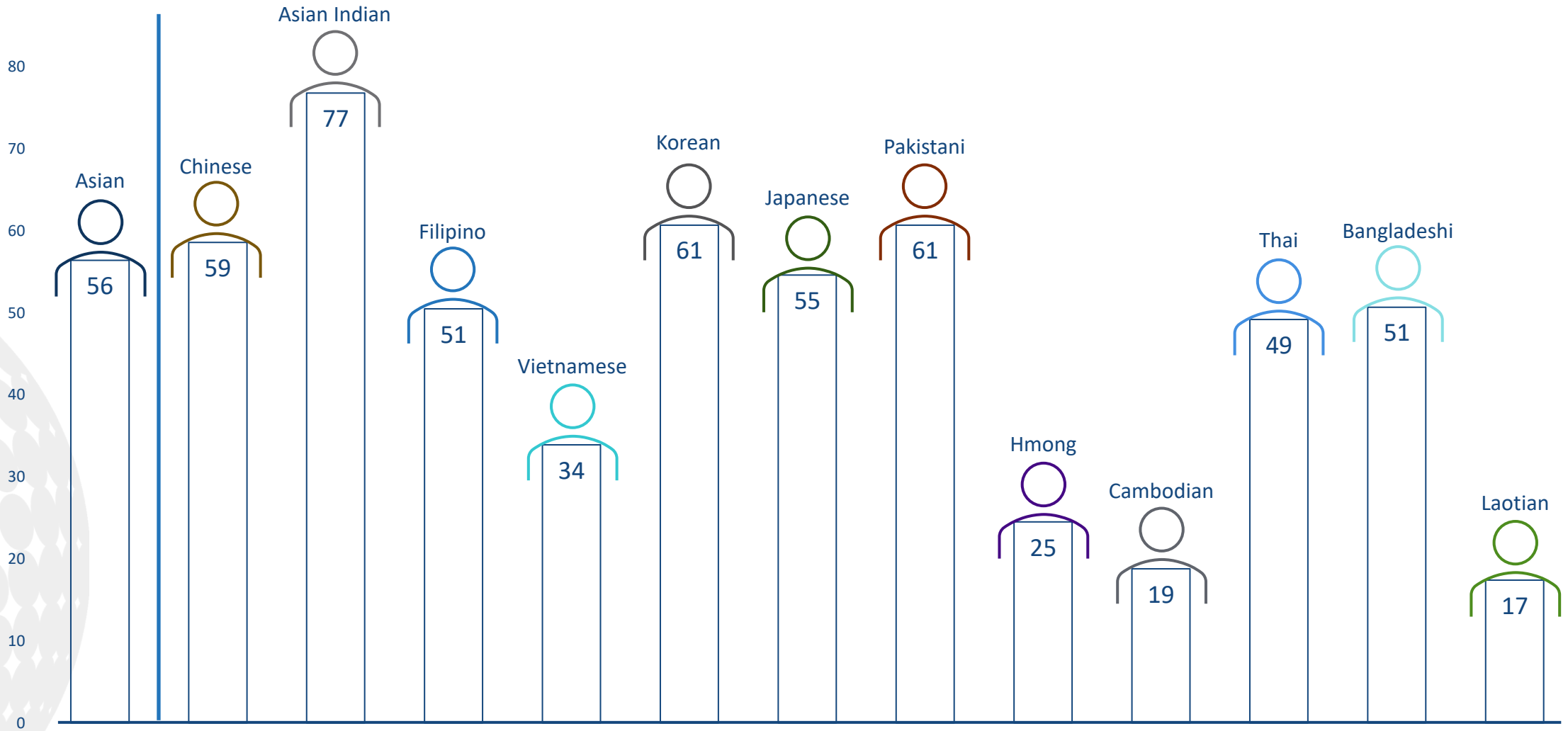
Improves program effectiveness

Percent of People Ages 25+ with a Bachelor's Degree or more



Source: U.S. Census Bureau, 2021 American Community Survey 1-year, Table S0201

Bachelor's Degree or More by Detailed Race



Source: U.S. Census Bureau, 2021 American Community Survey 1-year, Table S0201

Methodological and analytic challenges

Data can only be disaggregated by available categories

Small numbers and reliable estimates

Trend analysis

Working with small numbers

Maintaining confidentiality

- Smaller populations = easier to identify an individual

Providing reliable estimates

- Smaller populations = more sampling variability

Multiple options for assessing and dealing with these challenges

Assessing reliability

Calculate and evaluate measures of reliability

- Standard error, margin of error (MOE), confidence interval, coefficient of variation

ACS tables provide 90% MOEs for numbers and percents

- Formulas to calculate MOEs when combining categories and/or rates and percentages
- Understanding and Using American Community Survey Data: What All Data Users Need to Know

What is an acceptable amount of error?

No hard and fast rules; depends on the application

Confidence intervals for percents

- Less than 10 percentage points is one option

Coefficients of variation

- Smaller CVs (<15%) indicate greater reliability; larger CVs (>30%) indicate less reliable data
- Don’t use when proportion is close to zero

Options for presenting data

Reliable estimates

- Present the estimate with the reliability measure

Borderline reliable estimates

- Present the estimate with note to use caution
- Aggregate to increase sample size if applicable

Unreliable estimates

- Aggregate to increase sample size if applicable
- Suppress estimates
  - Highlight need for more or better data

Aggregating data

Expand the geographic area

- Combine smaller levels of geography into a larger group (e.g., combine counties to create county groups)

Combine multiple years of data

- 5-year ACS data
- Custom multiyear estimations: ACS Data User Group Webinar

Combine groups

- May not always be a good option
- Need to assess if combining groups of people is appropriate

Aggregating data

Conceptual considerations

- Tradeoff between data that is less current, has less geographic detail or less sub-group specificity

- Best option will vary based on goals, groups of interest
  - Sometimes no data is best
  - Sometimes limited data is best

Aggregating data

Technical considerations

- Microdata is more flexible
- Use aggregate numerators and aggregate denominators for rates

$$\frac{\widehat{P_{y1}} + \widehat{P_{y2}} + \widehat{P_{y3}}}{3} \neq \frac{\widehat{N_{y1}} + \widehat{N_{y2}} + \widehat{N_{y3}}}{\widehat{D_{y1}} + \widehat{D_{y2}} + \widehat{D_{y3}}}$$

- Standard errors/margins of error
  - ACS Accuracy of Data Documentation
  - ACS Data User Group Webinar
  - 5-year estimates: Variance Replicate Tables

Trend analysis

Racial and ethnic categories change across time

- Social construction of race and ethnicity

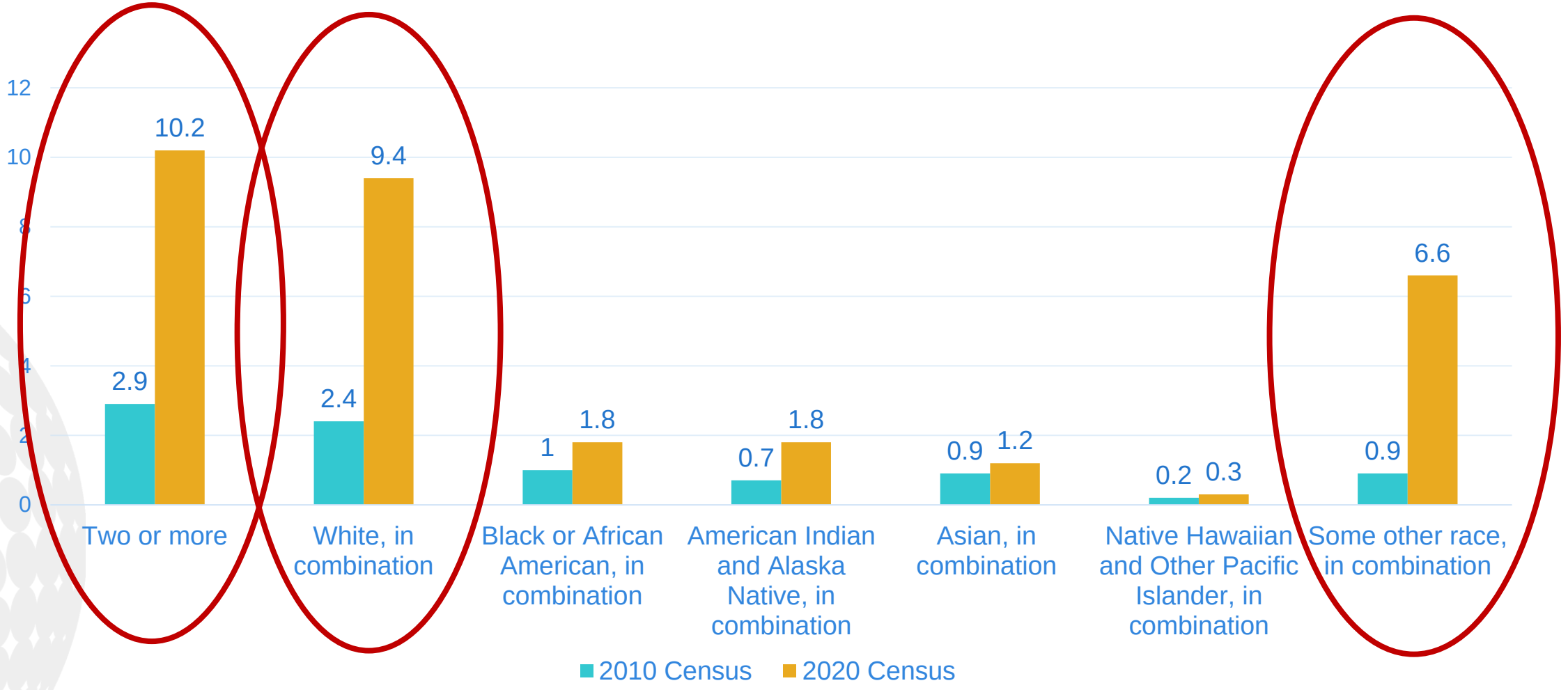
Questions and data processing changes across time

- 2020 Census vs 2010 Census, 2020 ACS vs prior years

Demographic change

- Births, deaths, and migration

Percent of Total Population by Multiracial Combinations



Trend analysis

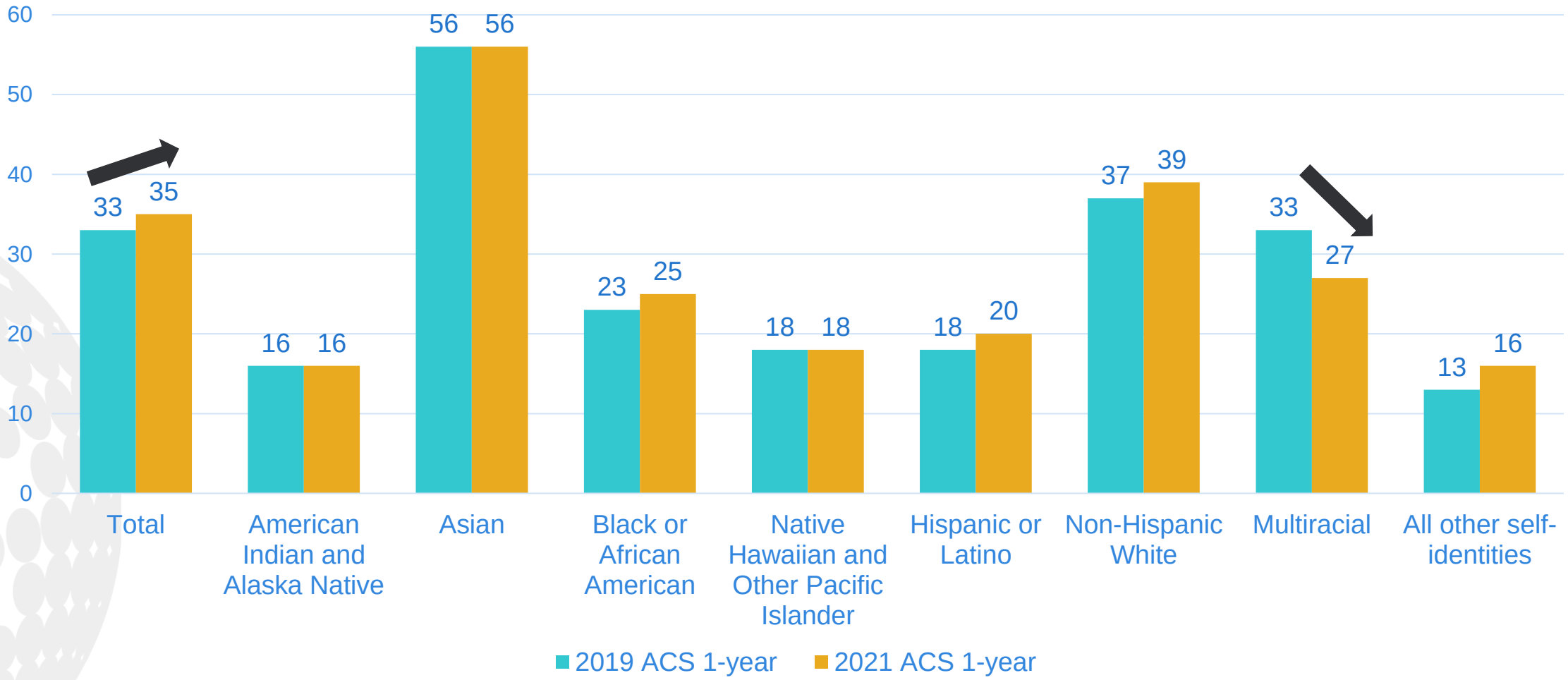
Comparing ACS data using new race/ethnicity design with data from 2019 and prior years

- U.S. Census Bureau recommends comparing with caution

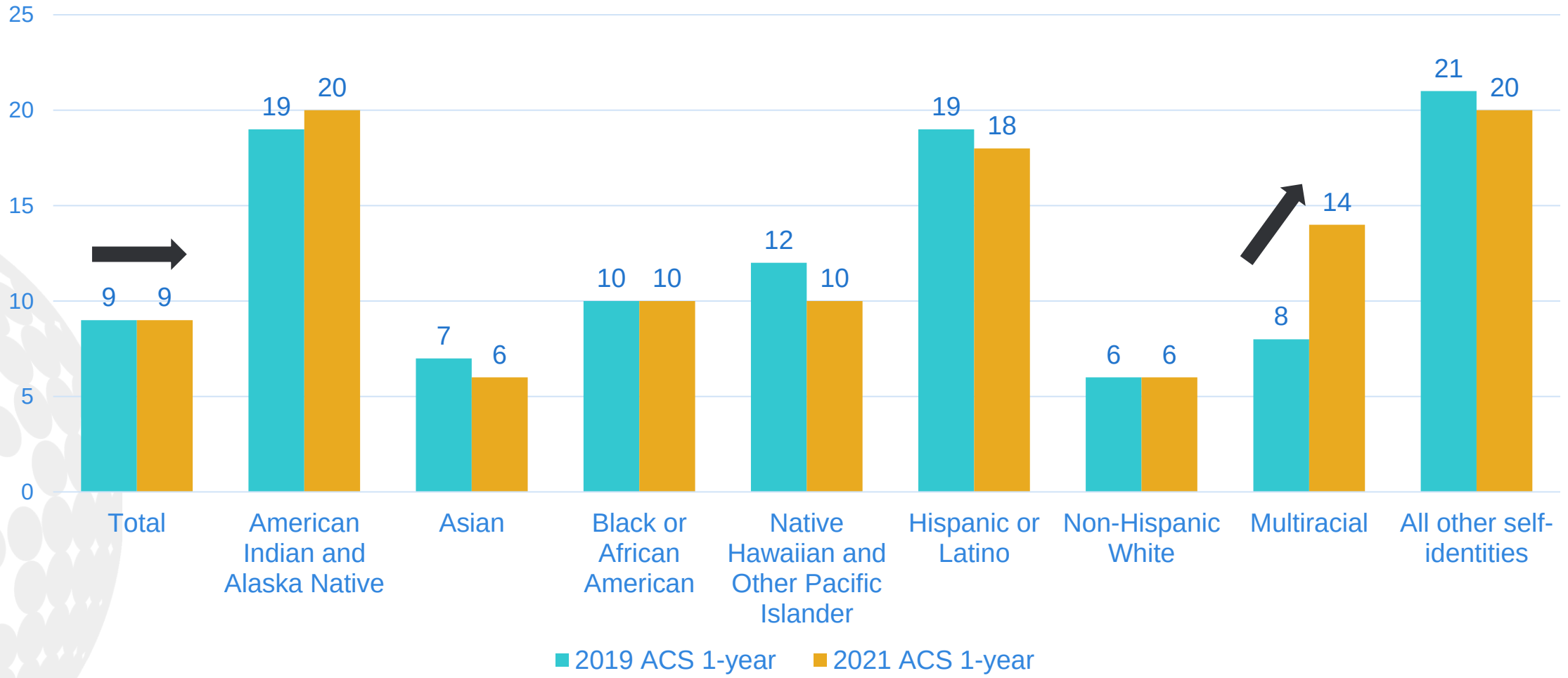
Identify change in racial and ethnic distribution within a broad group

Link changes in group characteristics to changes in rates

Trends by Race and Ethnicity: % with a BA+



Trends by Race and Ethnicity: % without Health Insurance



Accessing data by race and ethnicity:
Microdata

Public Use Micro Sample Data (PUMS)

- Download data files OR interactive MDAT tool
- Flexible, custom tabulations: More options to combine race and ethnicity
- Geographic data is limited
- MOEs are not provided, users must calculate MOES
  - See PUMS Accuracy of Data document
- Census Bureau webinar on using MDAT

Accessing data by race and ethnicity:
Microdata

PUMS via IPUMS

- Custom tabulations, 1-year and 5-year ACS data
  - Recode and create variables

- Download microdata
  - Need statistical software (e.g., SAS)
  - Multiple years of data at once, data are harmonized

Thank You!

Alicia VanOrman

avanorman@prb.org

www.prb.org