

# Disaggregating American Community Survey Data by Race and Ethnicity

Alicia VanOrman

September 20, 2022

[www.prb.org](http://www.prb.org)

# 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**

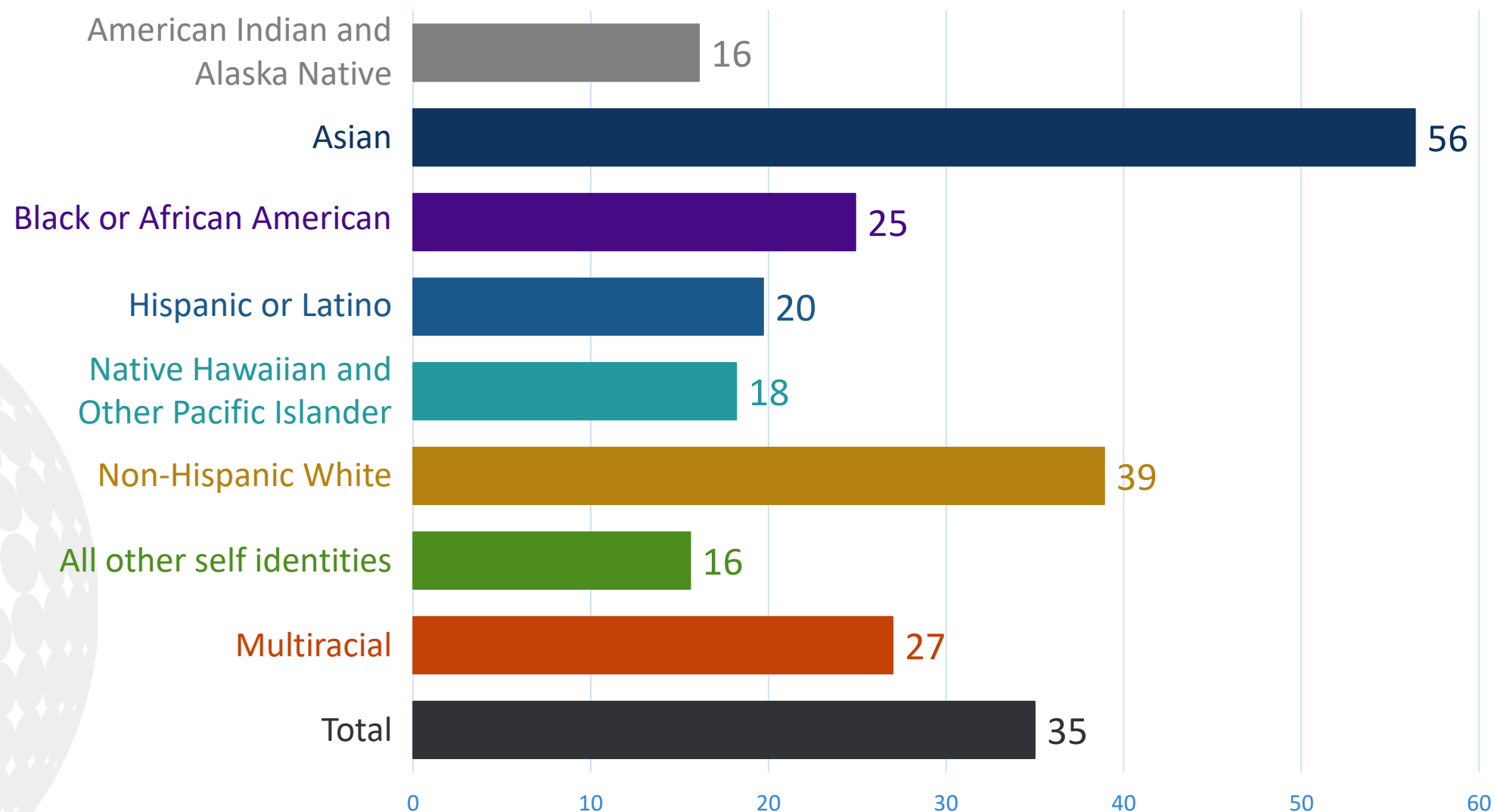
- Transform equity from a concept to a measurable objective

**Informs public policy**

**Improves program effectiveness**

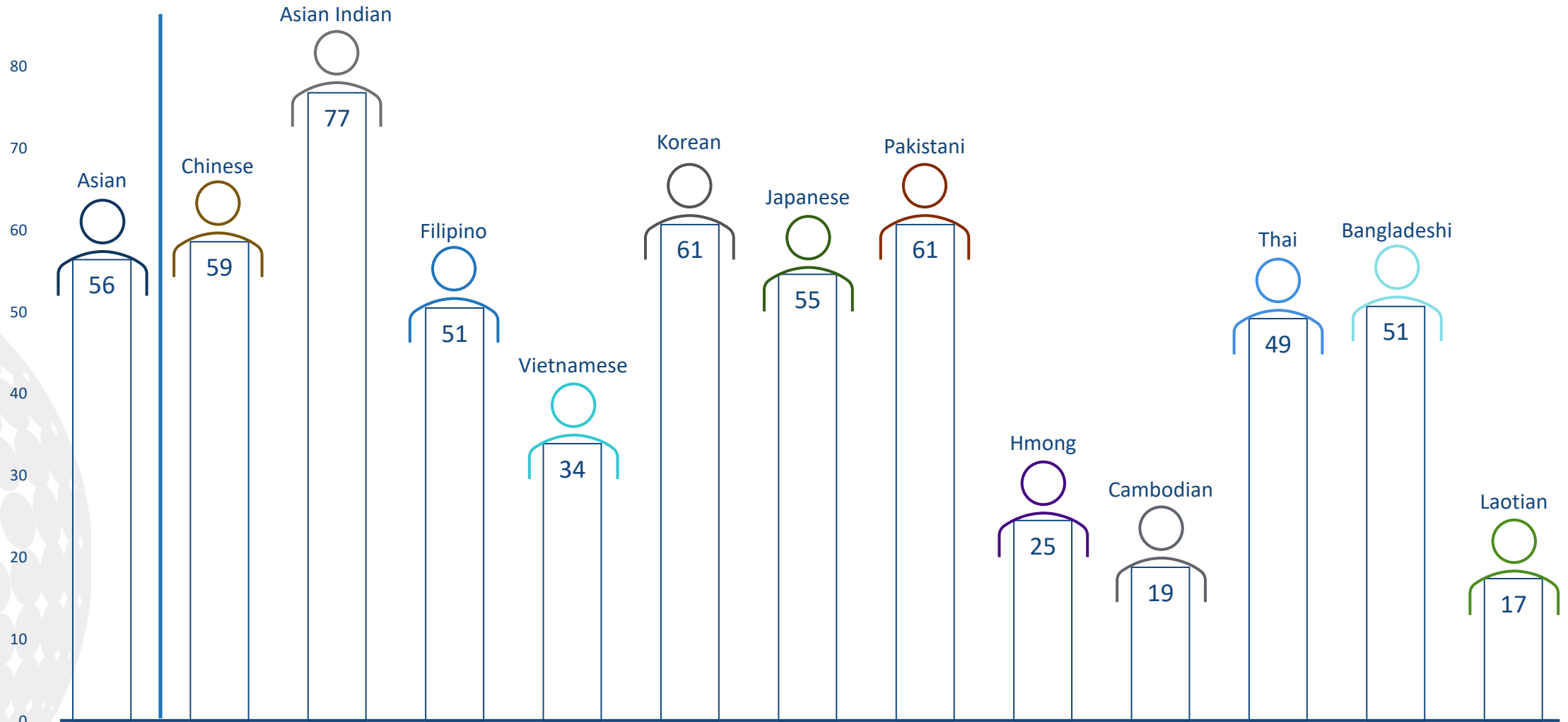


# Percent of People Ages 25+ with a Bachelor's Degree or more



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

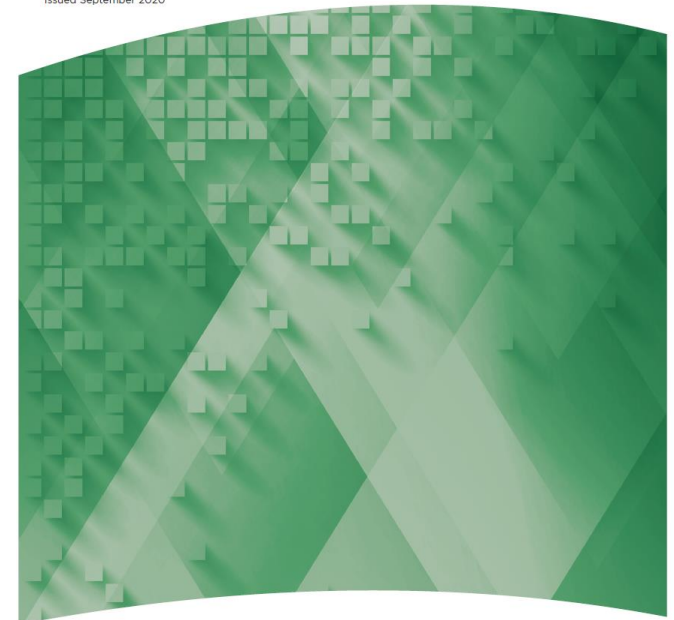
- Standard error, margin of error (MOE), confidence interval, coefficient of variation

## 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](#)

Understanding and Using  
American Community Survey Data  
*What All Data Users Need to Know*

Issued September 2020



United States<sup>®</sup>  
**Census**  
Bureau

U.S. Department of Commerce  
U.S. CENSUS BUREAU  
census.gov



# 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

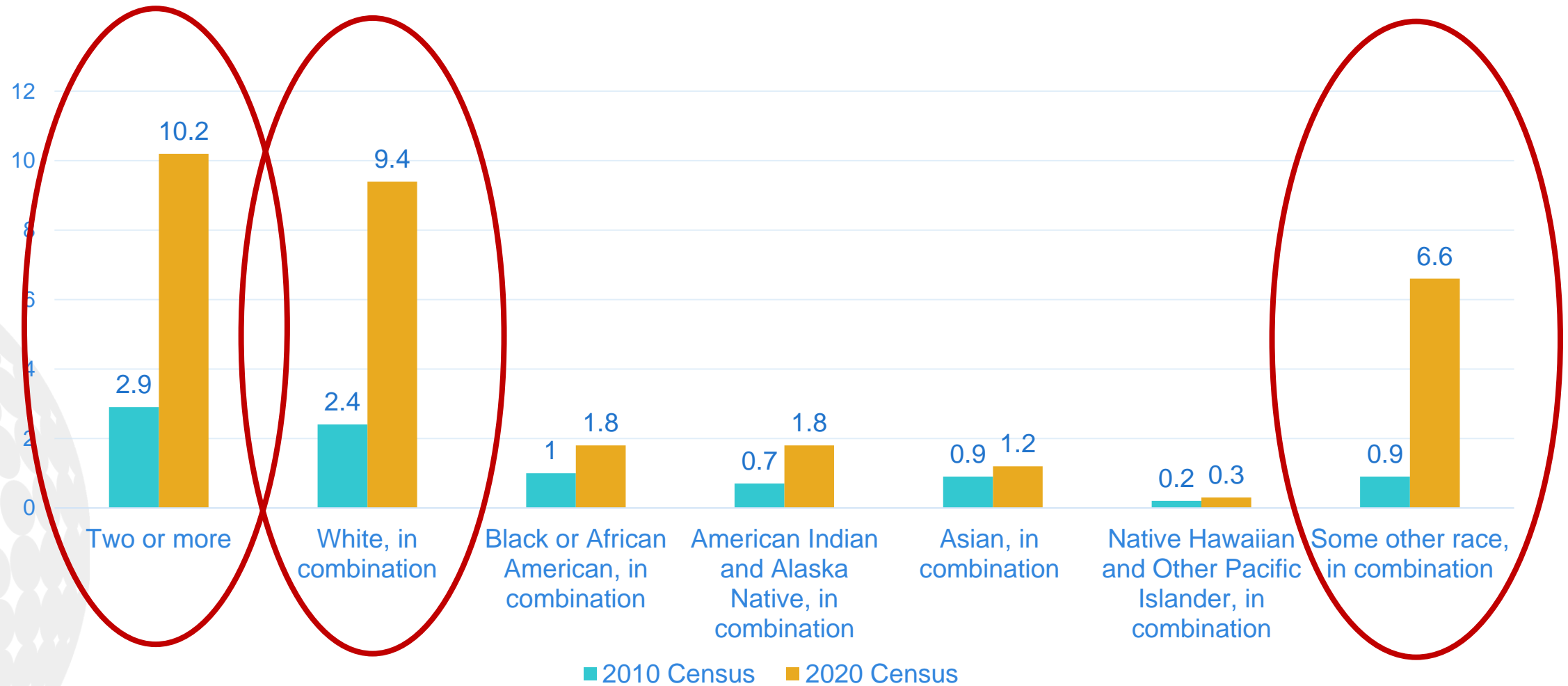
- 2020 Census vs 2010 Census, 2020 ACS vs prior years

## Demographic change

- Births, deaths, and migration



# Percent of Total Population by Multiracial Combinations



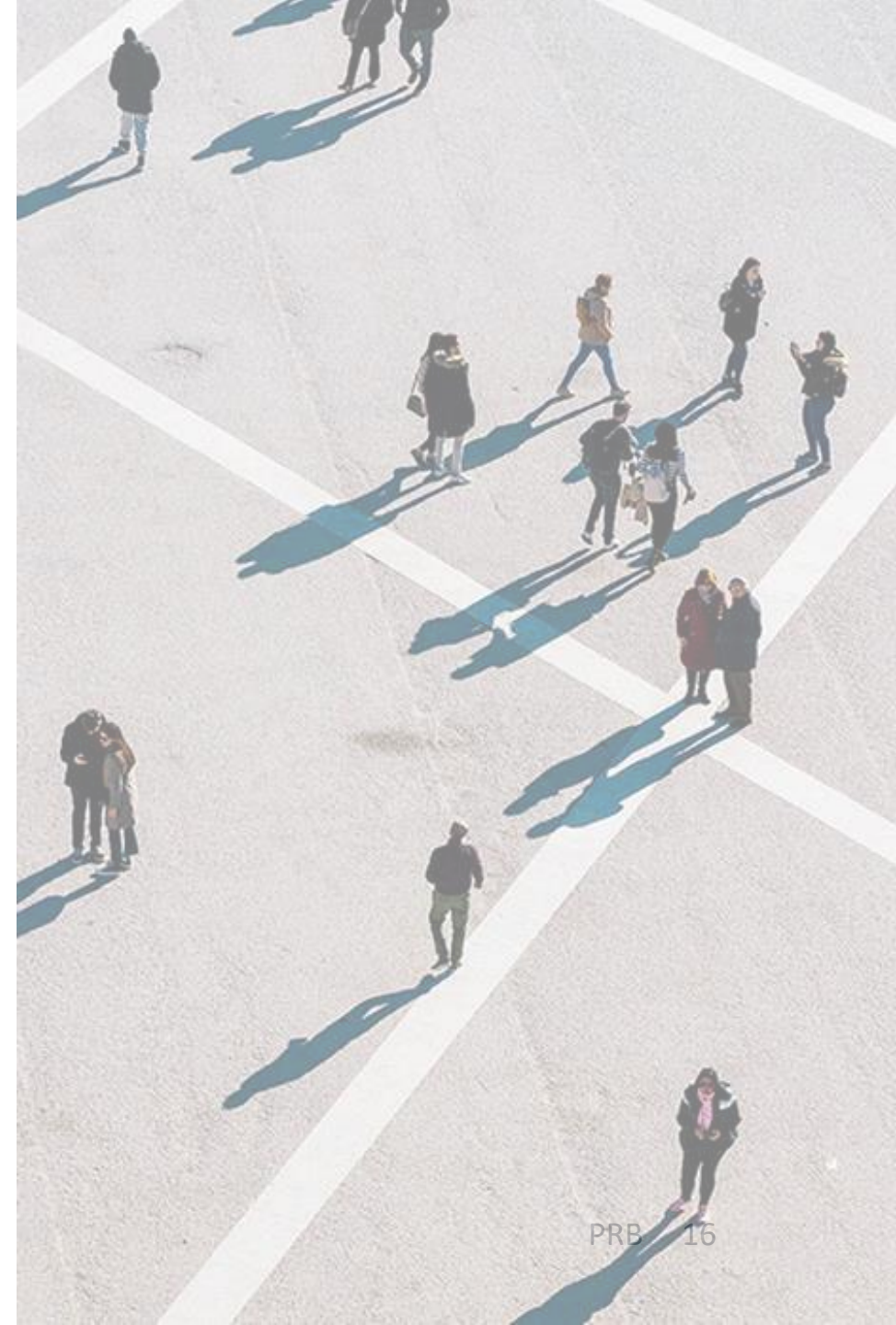
# Trend analysis

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

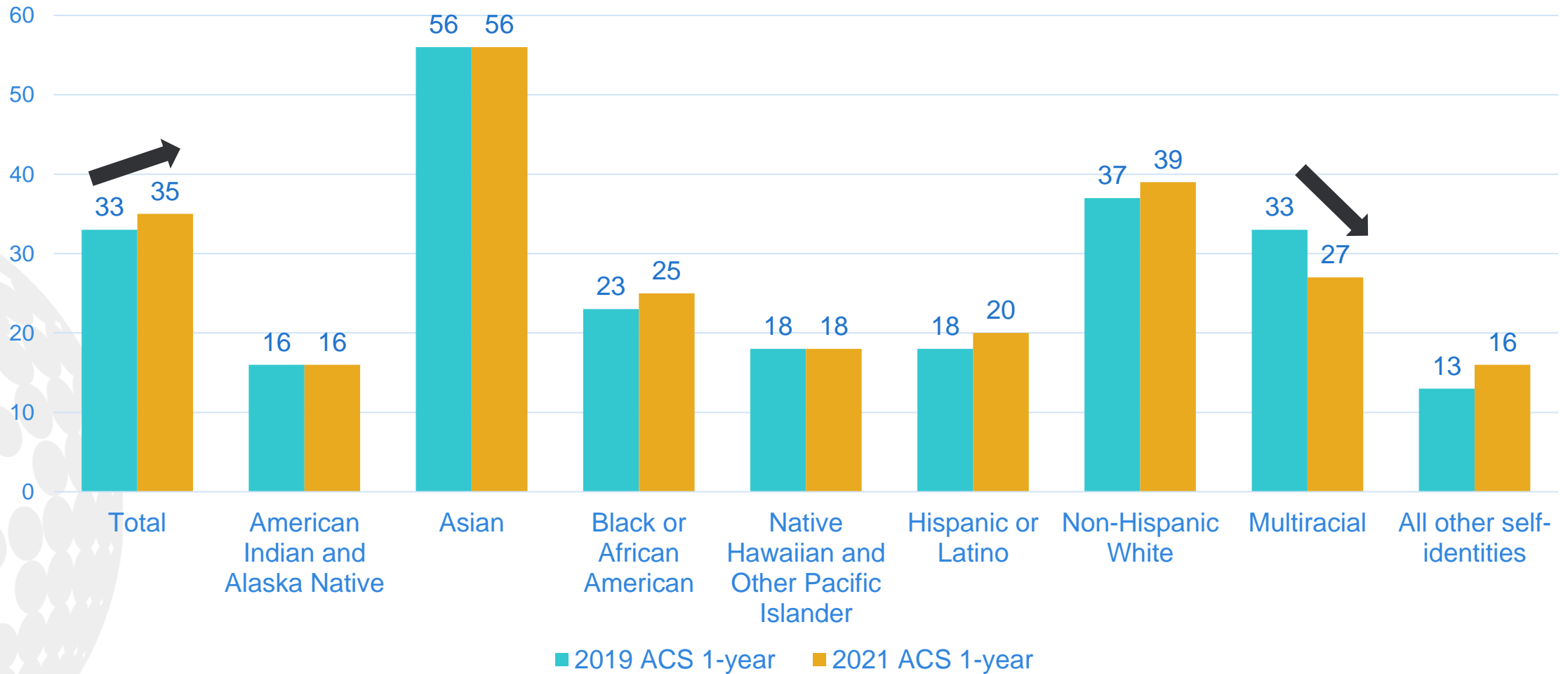
- U.S. Census Bureau recommends comparing with caution

## Identify change in racial and ethnic distribution within a broad group

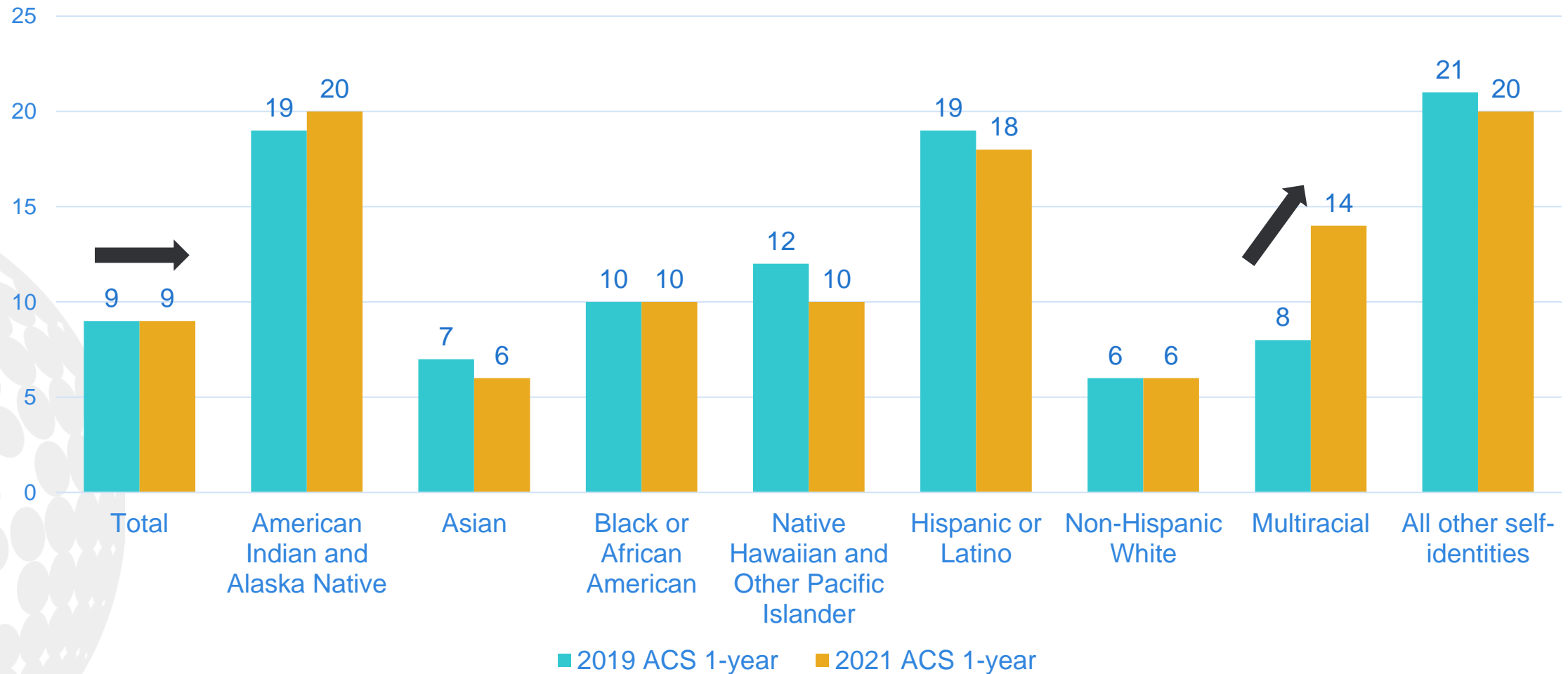
## Link changes in group characteristics to changes in rates



# 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



The screenshot shows the IPUMS USA website homepage. At the top, there is a dark blue header with the IPUMS USA logo on the left and navigation links (LOG IN, REGISTER, IPUMS.ORG) on the right. Below the header is a horizontal menu with links: HOME, SELECT DATA, MY DATA, and SUPPORT. A banner image below the menu shows various people in different settings. The main content area is divided into two columns. The left column contains a sidebar with links under the heading 'IPUMS USA', including 'ABOUT', 'REGISTER', 'DONATE TO IPUMS', 'DATA', 'SUPPLEMENTAL DATA', 'DOCUMENTATION', and 'SUPPORT'. The right column features the text 'U.S. CENSUS DATA FOR SOCIAL, ECONOMIC, AND HEALTH RESEARCH' followed by a paragraph about IPUMS USA's mission. Below this is a blue button labeled 'Get Data' and a light blue button labeled 'Analyze Data Online'. At the bottom, there is a section titled 'WHAT IS IPUMS?' with a paragraph explaining the organization's goals.

LOG IN | REGISTER | IPUMS.ORG

**IPUMS USA**

HOME | SELECT DATA | MY DATA | SUPPORT

U.S. CENSUS DATA FOR SOCIAL, ECONOMIC, AND HEALTH RESEARCH

IPUMS USA collects, preserves and harmonizes U.S. census microdata and provides easy access to this data with enhanced documentation. Data includes decennial censuses from 1790 to 2010 and American Community Surveys (ACS) from 2000 to the present.

USE IT FOR GOOD -- NEVER FOR EVIL

CREATE YOUR CUSTOM DATA SET

Get Data

ONLINE TOOL FOR ANALYSIS

Analyze Data Online

WHAT IS IPUMS?

IPUMS provides census and survey data from around the world integrated across time and space. IPUMS integration and documentation makes it easy to study change, conduct comparative research, merge information across data types, and analyze individuals within family and community context. Data and services available free of charge.



# Thank You!

Alicia VanOrman

[avanorman@prb.org](mailto:avanorman@prb.org)

[www.prb.org](http://www.prb.org)