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

- American Indian and Alaska Native: 16%
- Asian: 56%
- Black or African American: 25%
- Hispanic or Latino: 20%
- Native Hawaiian and Other Pacific Islander: 18%
- Non-Hispanic White: 39%
- All other self identities: 16%
- Multiracial: 27%
- Total: 35%

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

\[
\frac{\hat{P}_{y1} + \hat{P}_{y2} + \hat{P}_{y3}}{3} \neq \frac{\hat{N}_{y1} + \hat{N}_{y2} + \hat{N}_{y3}}{\hat{D}_{y1} + \hat{D}_{y2} + \hat{D}_{y3}}
\]

• 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

- Two or more: 10.2% (2010 Census), 9.4% (2020 Census)
- White, in combination: 2.9% (2010 Census), 2.4% (2020 Census)
- Black or African American, in combination: 1% (2010 Census), 1.8% (2020 Census)
- American Indian and Alaska Native, in combination: 0.7% (2010 Census), 1.8% (2020 Census)
- Asian, in combination: 0.9% (2010 Census), 1.2% (2020 Census)
- Native Hawaiian and Other Pacific Islander, in combination: 0.2% (2010 Census), 0.3% (2020 Census)
- Some other race, in combination: 0.9% (2010 Census), 6.6% (2020 Census)
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+

<table>
<thead>
<tr>
<th>Category</th>
<th>2019 ACS 1-year</th>
<th>2021 ACS 1-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>33</td>
<td>35</td>
</tr>
<tr>
<td>American Indian and Alaska Native</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>Asian</td>
<td>56</td>
<td>56</td>
</tr>
<tr>
<td>Black or African American</td>
<td>23</td>
<td>25</td>
</tr>
<tr>
<td>Native Hawaiian and Other Pacific Islander</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>37</td>
<td>39</td>
</tr>
<tr>
<td>Multiracial</td>
<td>33</td>
<td>27</td>
</tr>
<tr>
<td>All other self-identities</td>
<td>13</td>
<td>16</td>
</tr>
</tbody>
</table>

2019 ACS 1-year vs. 2021 ACS 1-year
Trends by Race and Ethnicity: % without Health Insurance

<table>
<thead>
<tr>
<th>Category</th>
<th>2019 ACS 1-year</th>
<th>2021 ACS 1-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>American Indian and Alaska Native</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>Asian</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Black or African American</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Native Hawaiian and Other Pacific Islander</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Multiracial</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>All other self-identities</td>
<td>21</td>
<td>20</td>
</tr>
</tbody>
</table>
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