



# VISIBILITY GUIDE

An Examination  
of Aggregation  
and Suppression  
Practices in Race  
and Ethnicity Data

March 2025



TheRPGroup

Research, Planning & Professional Development  
for California Community Colleges

SM

# Table of Contents

Introduction.....3

How Data Aggregation Hides Voices.....4

    A False Sense of Homogeneity.....5

How Data Suppression Silences Voices.....6

    The Asterisk Nation.....7

Guiding Principles for Visibility.....8

Conclusion.....11

Additional Resources.....12

About The RP Group.....13



# Introduction

The RP Group is committed to fostering social responsibility, collaboration, equity, diversity, and inclusion. These values are fundamental to our [mission, vision, and values](#), [Code of Ethics](#), [2024-2028 Strategic Plan](#), and broader [commitment to advancing equity](#).

We believe that data can be a powerful tool for improving education, especially for historically marginalized students. However, it is crucial to use data responsibly and ethically. Therefore, we must scrutinize our data practices and how they shape narratives about student experiences.

This resource is designed to help our organization, and the broader California community college institutional research, planning, and effectiveness (IRPE) community adopt more equitable data practices. We will examine research practices that may unintentionally perpetuate systemic racism and oppression, providing concrete examples and discussion points to promote equity.

**A key focus of this guide is data aggregation and suppression.** These common practices, while often used to protect privacy and maintain statistical integrity, can inadvertently hide or silence the voices and experiences of marginalized students. Data suppression and aggregation can have several negative consequences:

- ❖ **Silencing historically marginalized, minoritized, or excluded voices:** By obscuring the experiences of specific student populations, these practices hinder our ability to understand their unique needs and challenges.
- ❖ **Masking diversity within groups:** Aggregating data can erase the diverse experiences within student groups.
- ❖ **Creating oversimplified realities:** These practices can lead to inaccurate and/or incomplete depictions of complex student experiences.

These practices can limit our understanding of student experiences, making it difficult to address inequities effectively.

This guide will first examine how aggregation can hide student voices and provide an example of a racial group adversely affected by this practice. Then, we will discuss how suppression silences student voices and provide an example of how a specific ethnic group is impacted by this practice. Finally, we will provide guidelines for data practitioners to consider when collecting, categorizing, and reporting student data.

# How Data Aggregation *Hides* Voices

## The Illusion of a Monolithic Experience

Aggregation is a categorization practice that involves combining two or more groups. Most categorization practices in higher education, like race and ethnicity aggregation, adhere to federal guidelines, such as those outlined in the [Integrated Postsecondary Education Data System](#) (IPEDS).

For example, IPEDS requires that institutions report students who identify as “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” under the broader Asian category. Additionally, any student who may identify with more than one racial group is to be reported as multiracial.

These guidelines were designed to enforce civil rights laws by requiring institutions to disaggregate data and identify potential inequities. However, the level of disaggregation mandated by these minimum reporting standards is insufficient.

The level of aggregation mandated by these minimum reporting standards can create a false sense of homogeneity within groups, masking differences in students’ experiences. In turn, this aggregation **creates the illusion of a monolithic experience**—that all students within a particular group are living through a similar experience because of their association with the broader group. Consequently, this practice can hinder the IRPE community’s ability to address disparities for students whose experiences are being hidden.



# EXAMPLE

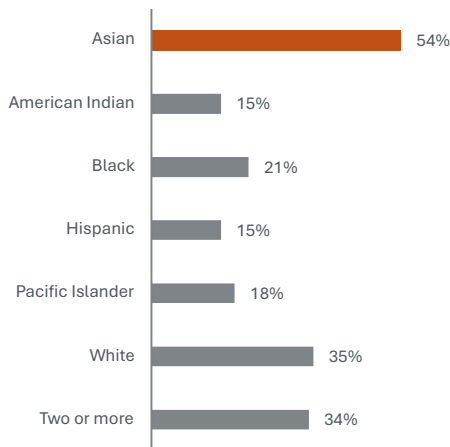
## A False Sense of Homogeneity

A prime example of this illusion of monolithic experiences revolves around the practice of grouping race/ethnic categories such as Asian American, Native Hawaiian, and Pacific Islander (AANHPI) groups. The AANHPI category has at least 21 distinct groups, all of which have their own unique languages and cultures. However, because these distinct groups are combined and categorized under one category (“Asian”), it creates the illusion of a monolithic experience among AANHPI groups and hides disparities that may be present among them.

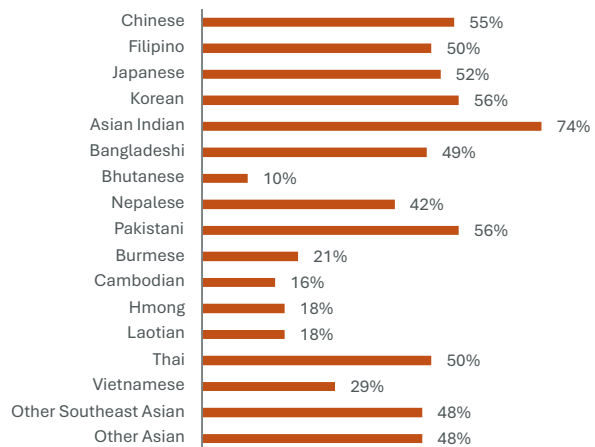
For example, the two graphs below display different views of degree attainment data: one that disaggregates the data using broader ethnic/racial groups (Figure 1) and another that disaggregates the Asian racial groups further into 17 distinct categories (Figure 2).

Figure 1 oversimplifies the Asian population and suggests that they are uniformly successful in achieving awards. Any disparities that may exist among AANHPI groups within this Asian category are not apparent. Figure 2, in contrast, reveals that several Asian groups—such as the Bhutanese, Cambodian, Hmong, and Laotian groups—experienced the lowest bachelor’s degree attainment rates. This finding would have been hidden without further disaggregation.

**Figure 1.** % of Adults Who Completed a Bachelor’s or Higher - 2016



**Figure 2.** % of Asian Adults Who Completed a Bachelor’s or Higher - 2016



# How Data Suppression *Silences* Voices

## A Story Never Told

Data suppression is a privacy practice that involves removing personally identifiable information from datasets to protect people's identities. This practice is often employed when there is a small number of people in the group, which could make them more easily identifiable. While intended to protect participant confidentiality and data reliability, this practice can inadvertently obscure the experiences of specific, marginalized, and minoritized groups.

For instance, The RP Group's guide, [Effective Practices for Data Suppression](#), advocates suppressing data points associated with groups of 10 or fewer individuals. This practice is common in fields like education but can lead to the systematic silencing of smaller, often marginalized and minoritized, groups based on, for example, race/ethnicity, sexual orientation, or gender identity.

While privacy protection is crucial, the harm comes when data from the same groups are repeatedly suppressed. This continual data suppression prevents the narratives of specific groups from ever being told, limiting further understanding of their unique challenges and experiences. This practice can also implicitly devalue these groups, suggesting there are too few of them to be worth considering.

For example, Native American, Alaska Native, Hawaiian Native, and Pacific Islander students often make up the smallest proportions of student populations. According to the California [Community Colleges Chancellor's Office Management Information Systems Data Mart](#), in the 2023-2024 academic year, each of these groups comprised less than 1% of the total student headcount. Subsequently, due to their smaller numbers, these students are often marked as statistically insignificant and excluded from further analysis, exacerbating their marginalization and hindering efforts to address their specific needs.

## EXAMPLE

# The Asterisk Nation

Data practitioners employ suppression rules to suppress or conceal data of groups deemed “too small” for sufficient analysis by lumping them into an “other” category or marking them with an asterisk (\*).

This practice has affected American Indian and Alaska Native populations to such an extent that the [National Congress of American Indians](#) has created a label for this phenomenon—**The Asterisk Nation**—because an asterisk, instead of a data point, is used in data displays time and time again.



Consequently, Native populations’ experiences repeatedly become invisible in data analyses. This lack of reported data leads to a lack of understanding of and knowledge about the Native community, which in turn hinders advocacy efforts for Native communities.


If Native students continue to be invisible in the data, they will be absent in policymaking as well. The underrepresentation of Indigenous communities in educational data hinders the allocation of resources to address their specific needs, negatively impacting their educational outcomes and success.

# Guiding Principles for Visibility

## Illuminating Student Voices

Blindly adhering to typical data protocols can inadvertently silence students' voices. Data stewards have a responsibility to align data practices and methods with students' experiences and voices. It is crucial to avoid a one-size-fits-all approach and be intentional in your methodology each time. Ergo, data methods should be tailored to the specific student population being served to ensure the most inclusive storytelling.

The RP Group created the three principles below to offer guidance on how data practitioners can adopt more equitable data practices to center students and amplify their voices. We also provide examples of how these principles can be applied to racial and ethnic data.



**Honor**

Honor the information we collect from people who take the time to share it.



**Center**

Center the data practices and methods to the experiences and voices of the target population.



**Balance**

Balance data privacy and security with visibility and utility.

## Honor the Information

We often collect a wealth of student data but fail to utilize it effectively. This lack of action can erode students' trust in our data initiatives and undermine our commitment to using their information meaningfully. To respect the time and effort people invest in sharing their information, we must carefully consider what data we collect and how we will utilize it.

Below are two examples of how we can honor the demographic information students report to us:

- ❖ **Leverage the Expanded Ethnicity Categories:** The California Community Colleges IRPE community has access to expanded ethnicity and race categories through CCCApply. These expanded categories offer a more nuanced understanding of student populations beyond the limitations of IPEDS categories. This detailed demographic data can be utilized for more targeted interventions and equitable resource allocation depending on the makeup of your student population.
- ❖ **Utilize an “If Any” Protocol:** To enhance the granularity of our racial and ethnic data, we can employ an “if any” rule to identify students who select multiple races or ethnicities. This approach can help mitigate the potential underrepresentation of smaller multiethnic groups and increase statistical power for analysis. For example, if a student selects both American Indian or Hawaiian Native and African American/Black, they can be counted in both categories and in the multiracial category.

## Center Student Voices

To ensure an accurate and equitable representation of all student groups, it is crucial to center your data practices and methods on the experiences and voices of the student population you serve. In determining the appropriate data methods, the following should be considered:

- ❖ **Provide multiple views of the data:** To maximize the insights gained from the data, consider presenting it in various ways to illuminate as many stories as possible. For instance, if the goal is to report on the experiences of African American/Black students, one could report the outcomes for these students based on how they self-identify: solely as African American/Black or as African American/Black and one or more other races/ethnicities.
- ❖ **Explore different reference groups.** Relying solely on specific reference groups (e.g., White or Asian) can inadvertently perpetuate existing inequities. To challenge the status quo and uncover more nuanced findings, explore alternative reference groups, analytical approaches, or strategic considerations to enhance statistical power. For instance, in regression models, you could recode your race/ethnicity categories whereby the reference group is African American/Black, centering the averages based on outcomes of African American/Black students and using that as a launchpad for discussion of the findings.



## Balance Privacy with Visibility

When reporting data findings, it is essential to prioritize privacy while ensuring students' voices are still heard. To balance these competing needs, consider the following strategies:

- ❖ **Combine Data:** By combining data from multiple years or disparate sources, you can significantly increase your sample size for groups who have historically had their experiences and outcomes suppressed or masked. This strategy offers several advantages, including increased statistical power and enhanced generalizability, while protecting privacy. For instance, if a single year's data is insufficient to reveal meaningful trends, combining data from the preceding 3-5 years can significantly boost the sample size. Alternatively, instead of relying solely on data from a specific college, leveraging data from multiple colleges can provide a broader perspective.
- ❖ **Consider the Audience:** Before sharing data from small sample sizes, it is crucial to carefully assess the reader's specific needs and role. More specifically, while sharing such data with the general public might not be appropriate, it may be suitable for a more targeted audience, such as a group of college personnel who already have access to this information. For instance, student support services staff serving special populations (e.g., veterans, foster youth) may need access to disaggregated outcomes to provide focused support and resources to these students depending on their intersectional identities.

# Conclusion

To advance social justice and equity, we must critically examine our existing data practices to ensure they promote student equity and success without perpetuating systemic racism and oppression.

This guide highlighted aggregation and suppression practices and how their indiscriminate application can hide and silence our students' voices. Specifically, we covered how aggregating Asian ethnicities into one broad group hides the experiences of smaller ethnic groups within this larger racial grouping. We also highlighted the impact data suppression has had on the Native American community, rendering them The Asterisk Nation, perpetually symbolized with an asterisk for representing smaller sample sizes.

While this guide only reviewed the impact of these two practices on racial and ethnic groups, it is important to remember that these practices can also affect other student groups, such as the LGBTQIA+ community, gender groups (e.g., nonbinary), or smaller affinity groups (e.g., women in leadership).

Moreover, it is essential to acknowledge that other data methodologies and practices may inadvertently result in adverse consequences. We strongly recommend reviewing the additional resources at the end of this document for further elucidation.

To combat the potential harm of data practices, practitioners should prioritize their intentionality by adhering to The RP Group's three **Guiding Principles for Visibility: Honor, Center, and Balance**. While implementing these intentional methods may require extra time and resources, it is a worthwhile investment in our students' success.

# Additional Resources

The following resources offer additional information on the topics explored in this guide. While not comprehensive, this list provides a foundation for further exploration.

- ❖ [A Toolkit for Centering Racial Equity Throughout Data Integration](#)
- ❖ [Asterisk Nation: One Tribe's Challenge to Find Data About Its Population](#)
- ❖ [Counting a Diverse Nation: Disaggregating Data on Race and Ethnicity to Advance a Culture of Health](#)
- ❖ [Decolonize Data: Accurate Data Tells Accurate Stories](#)
- ❖ [Do No Harm Project](#)
- ❖ [How to "QuantCrit": Practices and Questions for Education Data Researchers and Users](#)
- ❖ [Layers of Identity: Rethinking American Indian and Alaska Native Data Collection in Higher Education](#)
- ❖ [Twice Invisible: Understanding Rural Native America](#)
- ❖ [Understanding the Culture of Health for Asian American, Native Hawaiian and Pacific Islanders \(AANHPIs\)](#)
- ❖ [Video - Reclaiming Native Truth Intro](#)

# About The RP Group

The Research and Planning Group for California Community Colleges (The RP Group) is a nonprofit, nonpartisan team made up of a dedicated Board of Directors and staff, working together with a shared and resolute goal: to increase the success of California community colleges and beyond.

## Our Mission

The RP Group is a leader that supports equitable outcomes for minoritized and marginalized students through race-conscious, equity-minded research, planning, and professional development. We uplift student voices and empower researchers and planners to improve institutional effectiveness by dismantling systemic barriers and injustices.

## Our Vision


The RP Group envisions itself to be a change agent at local, state, and national levels that addresses longstanding injustices to ensure that students' demographics do not predict their postsecondary success.

## Our Values

- ❖ Championing Equity and Antiracism
- ❖ Centering Student Voice
- ❖ Leading Transformational Change
- ❖ Working Collaboratively and Building Community
- ❖ Demonstrating Integrity and Quality

Visit [www.rpgroup.org](http://www.rpgroup.org) for more information.



 2025. This work is openly licensed via [CC BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/).