

# **Equity Data Handbook**

CA Water Board's Office of Information Management and Analysis (OIMA)

2024-07-01

# Table of contents

<b>Welcome!</b>	<b>5</b>
<b>1 Background</b>	<b>7</b>
<b>I Getting Started</b>	<b>9</b>
<b>2 Getting Started</b>	<b>10</b>
<b>3 Best Practices</b>	<b>13</b>
3.1 Plan & Prepare . . . . .	13
3.2 Collect & Process . . . . .	13
3.3 Assure & Analyze . . . . .	13
3.4 Preserve & Store . . . . .	13
3.5 Publish & Share . . . . .	13
3.6 Discover & Integrate . . . . .	13
3.7 Describe . . . . .	14
<b>4 Establishing Common Language</b>	<b>16</b>
<b>II Plan &amp; Prepare</b>	<b>19</b>
<b>5 Planning</b>	<b>20</b>
<b>6 Data Preparation</b>	<b>21</b>
<b>III Collect &amp; Process</b>	<b>22</b>
<b>7 Data Sources</b>	<b>23</b>
<b>8 Surveys</b>	<b>24</b>
8.1 Picking a Survey Software . . . . .	25
<b>9 Data Limitations</b>	<b>26</b>

<b>IV Assure &amp; Analyze</b>	<b>28</b>
10 Data Analysis	29
11 Data Visualization	30
11.1 Data Visualization Tools . . . . .	30
12 Demographics Data	31
12.1 Demographics Data Needs Context . . . . .	31
12.2 Commonly used Racial Equity Data Sources . . . . .	32
12.3 For a growing list of Water Boards Developed Racial Equity Data Tools: . . . . .	32
12.4 Collecting Data . . . . .	32
12.5 Point Data . . . . .	32
12.6 Lines . . . . .	32
12.7 Polygons . . . . .	33
13 Working with CalEnviroScreen	34
13.1 Missing Values for CalEnviroScreen Scores . . . . .	34
13.2 Inconsistent Census Tract Boundaries in CalEnviroScreen 4.0 Shapefile	34
14 Turning Data into Information is Challenging - Even More so with Racial Equity Data Work	36
15 Demographics Data Needs Context	37
<b>V Preserve &amp; Store</b>	<b>38</b>
16 Preservation & Storage	39
<b>VI Publish &amp; Share</b>	<b>40</b>
17 Data Sharing	41
<b>VII Discover &amp; Integrate</b>	<b>42</b>
18 Preservation & Storage	43
<b>VIII Describe</b>	<b>44</b>
19 Documentation	45

<b>IX Use Cases</b>	<b>46</b>
<b>20 Use Case 1</b>	<b>47</b>
<b>21 Demographics Use Case</b>	<b>48</b>
<b>22 Resources</b>	<b>49</b>
22.1 Other Equity & Data Handbooks / Toolkits / Guides . . . . .	49
22.2 Presentations . . . . .	49
22.3 Relevant Government Documents . . . . .	49
22.3.1 <u>California Water Boards</u> . . . . .	49
22.3.2 <u>California Office of Data and Innovation</u> . . . . .	50
22.3.3 <u>The White House</u> . . . . .	50
22.4 Relevant Literature . . . . .	50
22.5 Websites . . . . .	50
<b>23 Inspiration</b>	<b>51</b>

# Welcome!

This Equity Data Handbook is an online resource written by the The State Water Resources Control Board (State Water Board) and the nine Regional Water Quality Control Boards (Regional Water Boards), collectively known as the California Water Boards (Water Boards).

Content in this Handbook includes best practices and guidance for Water Boards staff on incorporating racial equity concepts into their data-related work. Specifically, this Handbook will provide guidance and resources to help Water Boards staff conduct each phase of the data life cycle through a racial equity lens - from the planning and design of a project to data collection methods, visualization development (e.g., maps, fact sheets, etc.) and more!

A depiction of the Data Life Cycle is provided below for context. We must have meaningful engagement and partnership with our diverse communities during each phase of the Data Life Cycle, especially those that have been historically underserved namely Black, Indigenous, and other People of Color (BIPOC). If we focus on uplifting those most highly impacted we will inevitably improve the the data products and services we develop and, most importantly, the experiences and outcomes for all communities.

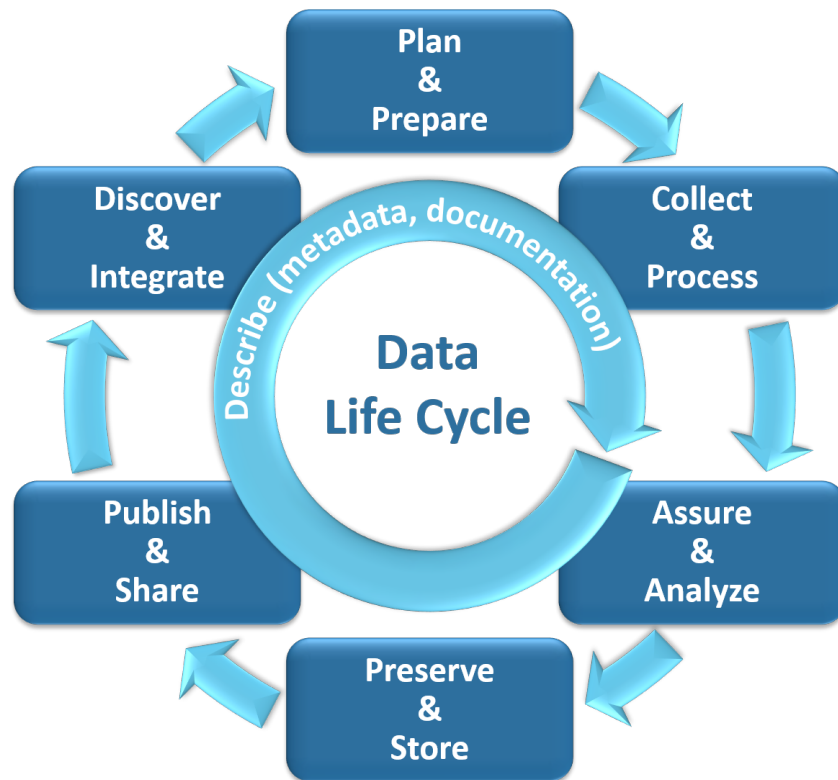


Figure 1: Graphic representing each phase of the Data Life Cycle.

This [Quarto book](#) is an open, living, and continuously iterating resource. If you have suggestions for additions or revisions you think should be incorporated into this book, please email [equitydatahelp@waterboards.ca.gov](mailto:equitydatahelp@waterboards.ca.gov).

# 1 Background

During its August 18, 2020 meeting, the State Water Resources Control Board (State Water Board) publicly acknowledged that the historical effects of institutional racism must be confronted throughout government, and it directed staff to develop a priority plan of action (Aug 18, 2020 Meeting [Agenda](#), [Recording](#)). The State Water Board's Racial Equity Team held public and employee listening sessions to help develop a draft resolution. After a public comment period on the draft resolution in spring 2021, the Racial Equity Team made significant updates to the resolution. On November 16, 2021, the State Water Board adopted [Resolution No. 2021-0050](#), "Condemning Racism, Xenophobia, Bigotry, and Racial Injustice and Strengthening Commitment to Racial Equity, Diversity, Inclusion, Access, and Anti-Racism" (Racial Equity Resolution) which affirms the State Water Board's commitment to racial equity and directs staff to undertake a variety of actions to achieve racial equity throughout all State Water Board programs and activities (Nov 16, 2021 Meeting [Agenda](#), [Recording](#)).

The Racial Equity Resolution is one milestone on our ongoing journey to operationalize equity throughout our organization. The next step is to implement the [Racial Equity Action Plan](#), which includes specific actions the State Water Board will take to address racial inequities, as well as metrics to measure our progress. With this Action Plan, we envision a sustainable California where race no longer predicts where clean water is available or who has access to it.

## Note

It's important to note that racial equity and equity in general is an outcome and there is no such thing as a "racial equity data-set" or "racial equity data". Instead we should think of data as a tool to help us achieve the overall outcome of equity - and how we view, use, and act on data and related tools is what will determine whether we operationalize equity or perpetuate injustice.

Development of the Racial Equity Action Plan began in Spring 2022 and involved public and employee engagement and tribal consultations. The Water Boards' Racial Equity Team presented the Racial Equity Action Plan (2023-2025) to the State Water Board as an informational item at the Board Meeting on January 18, 2023 ([Agenda](#), [Slides](#), [Recording](#)).

Goal 1a of the Racial Equity Action Plan is to ensure Water Boards data are accessible, equitable, and culturally relevant. One action captured under that goal is the development of a Racial Equity Data Action Plan which must:

1. Develop training and best practices guidance for Water Boards staff on incorporating racial equity concepts into the planning and design of data collection methods and visualizations (e.g., maps, factsheets, etc.) projects.
2. Identify and expand existing opportunities for public participation in science and community data gathering programs to develop new data collection methods, support existing programs, and incorporate community datasets into the database.
3. Create a publicly accessible data catalog tool / interface that includes existing demographic data, Water Boards program data, and other available data (such as heat maps or flood hazard maps) to inform the implementation of the Racial Equity Action Plan.

The Racial Equity Data Action Plan is being developed by the Water Board's Racial Equity Data Task Force which is a group of volunteer staff from across the Water Board led by the [Office of Information Management and Analysis](#) and the [Office of Public Participation](#). This Handbook is intended to address Item 1 above by providing staff a best practices guide for incorporating racial equity concepts into the Water Board programs using data and information. Item 2 above will be iterative and grow as staff and programs begin to utilize the guidance and tools found within this handbook. To fulfill Item 3 above staff have created a [California Water Boards Racial Equity Data Resource Hub](#) which will grow as more programs create and publish racial equity based tools and visualizations.



# **Part I**

## **Getting Started**

## 2 Getting Started

Program staff that are beginning a racial equity based data project should meet with the Racial Equity Data Team to discuss key aspects of the project including who the audience is and what questions or information you are trying to convey using the available racial equity data. This engagement should occur prior to beginning any analysis or data collection. To initiate a consultation with the Racial Equity Data Team, please send an email to [equitydatahelp@waterboards.ca.gov](mailto:equitydatahelp@waterboards.ca.gov).

This Best Practices and Guidance handbook is a blend of emerging, comprehensive guidance on:

- How to break down the management questions where racial equity information is being posed against administrative data; and
- How to apply best practices on data use, data visualization and storytelling to begin an iterative process into Advancing Racial Equity.

This resource is specifically structured to support Water Boards staff on incorporating racial equity concepts into the planning and design of data collection methods and visualizations (e.g., maps, factsheets, etc.) projects. The guidance and best practices provided serves as a strategic guide emphasizing the importance of collection, analysis and utilization of racial equity data.

These users could include:

- Program Staff
- Program managers
- Executives
- Agency Partners
- Tribal Governments
- The Public

Furthermore, we strongly suggest staff review the following [framework outlined by the Government Alliance for Racial Equity \(GARE\)](#) to normalize, organize, and operationalize racial equity throughout data integration (see image below). In addition to why it is important to [lead with race](#):

“with the recognition that the creation and perpetuation of racial inequities has been baked into government, and that racial inequities across all indicators for success are deep and pervasive. We also know that other groups of people are still marginalized, including based on gender, sexual orientation, ability and age, to name but a few. Focusing on racial equity provides the opportunity to introduce a framework, tools and resources that can also be applied to other areas of marginalization. It is critical to address all areas of marginalization, and an institutional approach is necessary across the board. As the local and regional government deepens its ability to eliminate racial inequity, it will be better equipped to transform systems and institutions impacting other marginalized groups.”

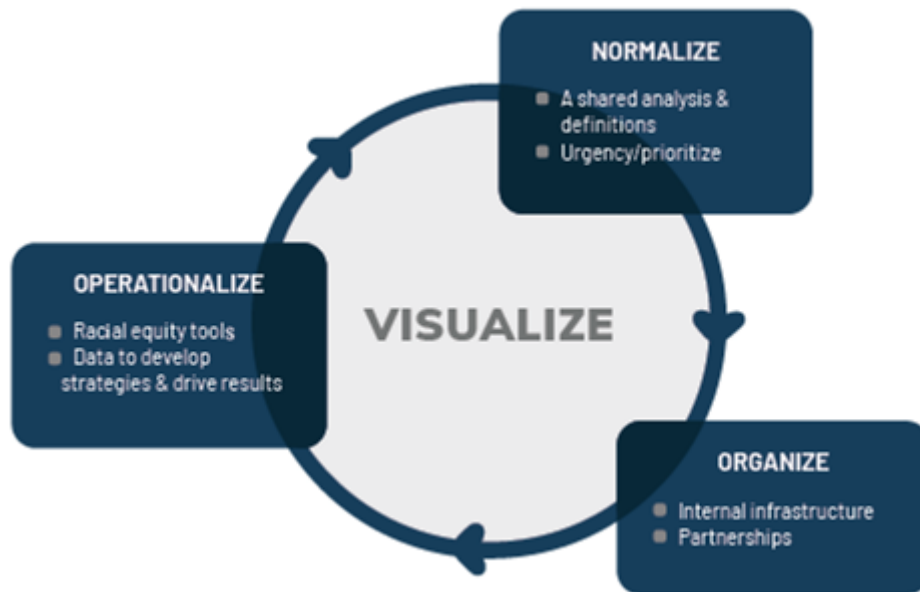


Figure 2.1: GARE model of change. Source: GARE Communications Guide, May 2018

Furthermore, we recommend staff take the Advancing Racial Equity training series offered by the [Water Boards Training Academy](#) to foster a consistent baseline knowledge of racial equity work and the importance of applying a racial equity lens to our work.

**In summary, before embarking on your racial equity analysis or data project you should:**

1. Consult [Racial Equity Data Subcommittee](#) by emailing [equitydatahelp@waterboards.ca.gov](mailto:equitydatahelp@waterboards.ca.gov) to help identify audience and potential questions or information your teams' analysis may answer and provide guidance.
2. Review the [Racial Equity Data Project Guidance](#) and answer the questions found within the Project Scoping section by filling out this form: <https://forms.office.com/g/>

0HuJqQkiJs

3. Review [GARE Framework: Normalize, Organize, and Operationalize](#)
4. If possible, take the Advancing Racial Equity training series offered by the [Water Boards Training Academy](#)

A running list of Water Board specific racial equity resources are available on [Environmental Justice Roundtable SharePoint](#)

## **3 Best Practices**

We strongly encourage staff to read this Handbook in its entirety. However, for those who want to jump to best practices, or who just need a refresher, we have provided them below, with links to pages that have more details, guidance, and resources.

### **3.1 Plan & Prepare**

Remember - this phase is all about.... Best practices for this phase include:

### **3.2 Collect & Process**

Remember - this phase is all about.... Best practices for this phase include:

### **3.3 Assure & Analyze**

Remember - this phase is all about.... Best practices for this phase include:

### **3.4 Preserve & Store**

Remember - this phase is all about.... Best practices for this phase include:

### **3.5 Publish & Share**

Remember - this phase is all about.... Best practices for this phase include:

### **3.6 Discover & Integrate**

Remember - this phase is all about.... Best practices for this phase include:

## 3.7 Describe

Remember - this phase is all about.... Best practices for this phase include:

- Use the best available science while relying on current, generally accepted Agency procedures for conducting risk assessment and economic analysis.
- Use existing frameworks and data from other parts of the regulatory analysis, supplemented as appropriate.
- Be consistent with the basic assumptions underlying other parts of the analysis, such as using the same baseline and option scenarios.
- Use the highest quality and most recent data available. Discuss the overall quality and main limitations of the data (e.g., completeness, accuracy, validation).
- Discuss available evidence of factors that may make population groups of concern more vulnerable to adverse effects (e.g., unique pathways; cumulative exposure from multiple stressors; and behavioral, biological, or environmental factors that increase susceptibility).
- Identify unique considerations for subsistence populations when relevant.
- Carefully select and justify the choice of a comparison population group.
- Carefully select and justify the choice of the geographic unit of analysis and discuss any particular challenges or aggregation issues related to the choice of spatial scale.
- Analyze and compare effects in baseline and across policy scenarios to show differences in effects.
- Present summary metrics for relevant population groups of concern as well as the comparison population group.
- When data allow, characterize the distribution of risks, exposures, or outcomes within each population group, instead of presenting only average effects.
- Disaggregate data to reveal important spatial differences (e.g., demographic information for each facility/place) when feasible and appropriate.
- Discuss the severity and nature of the health consequences for which differences between population groups have been analyzed.
- Clearly describe data sources, assumptions, analytic techniques, and results.
- Discuss key sources of uncertainty or potential biases in the data (e.g., sample size, using proximity as a surrogate for exposure) and how they may influence results.
- When possible, conduct sensitivity analysis for key assumptions or parameters that may affect findings.

- Make elements of Environmental Justice (EJ) assessments as straightforward and easy for the public to understand as possible.

## 4 Establishing Common Language

When working with racial equity data we suggest establishing a common language and definitions to cultivate a collective understanding of underlying concepts and historical context. Creating and agreeing upon a common language can help foster transparency, challenge assumptions, and center the voices of marginalized communities. Yet the efficiency of these efforts hinges on a shared language that facilitates understanding and collaboration. By grounding discussions in a common language, we can build trust and empower our team and community.

Establishing a common language and definitions are critical to creating a shared understanding, however we acknowledge that language can be used deliberately to engage and support community anti-racism coalitions and initiatives, or to inflame and divide them. It is important to note that although the language in this resource may be commonly used, it is not the sole definition and some may disagree with the definitions and their use. More specifically, in this resource we intentionally use the acronym BIPoC (Black, Indigenous, people of color) as a term that seeks to recognize the unique experience of Black and Indigenous People within the United States. We recognize that naming is power, and we remain committed to using language that supports pro-Blackness and Native visibility, while dismantling white supremacy.

List of definitions:

Short list of important definitions (i.e. data definitions and we can address BIPOC here)

[Glossary](#) | [Racial Equity Tools](#) - include glossary of definitions?

Below are a set of definitions provided by the Water Board Racial Equity Team in the development of the Racial Equity Resolution and Racial Equity Action Plan and are those that we adhere to in this document. (citations can be found [here](#))

- **Equality** describes circumstances in which each individual or group is given the same or equal treatment, including the same resources, opportunities, and support. However, because different individuals or groups have different histories, needs, and circumstances, they do not have equal positions in society or starting points. Providing the same resources, support, or treatment does not guarantee that everyone will have fair or equal outcomes.



- **Ethnicity** is a term used to describe subgroups of a population that share characteristics such as language, values, behavioral patterns, history, and ancestral geographical base. Social scientists often use the terms ethnicity and ethnic group to avoid the perception of biological significance associated with race.
- **Institutional racism** describes the ways in which policies and practices perpetuated by institutions, including governments and private groups, produce different outcomes for different racial groups in a manner that benefits the dominant group. In the United States, institutional racism includes policies that may not mention race but still result in benefiting white people over people of color.
- **Race** is a social construct used to categorize humans into groups based on combinations of shared physical traits such as skin color, hair texture, nose shape, eye shape, or head shape. Although most scientists agree that such groupings lack biological meaning, racial groups continue to have a strong influence over contemporary social relations. Historically in the United States, race has frequently been used to concentrate power with white people and legitimize dominance over non-white people.
- **Racial equity** means Race can no longer be used to predict life outcomes and outcomes for all groups are improved. For example, when we hold income constant, there are still large inequities based on race across multiple indicators for success, including the environment, education, jobs, incarceration, health and housing.
- **Racism** is any prejudice against someone because of their race when systems of power reinforce those views.
- **Structural racism** is the normalization and legitimization of an array of historical, cultural, institutional, and interpersonal dynamics that routinely advantage whites while producing cumulative and chronic adverse outcomes for people of color. Structural racism encompasses the entire system of white domination, diffused, and infused in all aspects of society, including its history, culture, politics, economics, and whole social fabric. Structural racism is more difficult to locate in a particular institution because it involves the reinforcing effects of multiple institutions and cultural norms, past and present, continually reproducing old and producing new forms of racism. Structural racism is the most profound and pervasive form of racism; all other forms of racism emerge from structural racism.

**Systemic racism** can be said to encompass both institutional and structural racism. Glenn Harris, president of Race Forward, defines systemic racism as “the complex interaction of culture, policy and institutions that holds in place the outcomes we see in our lives.” The legacy of systemic racism can be seen in a variety of outcomes affecting people of color, such as housing insecurity, a ten-fold wealth gap between white and Black or Latinx households, a dramatic over-representation of people of color in prison, and disparities in education, health, and exposure to environmental pollution.

**Non-Inclusive Terms to Avoid:**

- “Brown Bag” - The term “brown bag” has a historical connotation with creating an exclusive gathering that required attendees to have a lighter skin tone than a brown paper bag to participate and gain access.
  - Alternatives: lunch in, lunch and learn
- “Chief” - this term is used throughout the Water Boards to indicate positions and job titles. This term is appropriated from the Indigenous Peoples of North America and should be avoided wherever possible.
  - Alternatives: manager, lead, head
- “Grandfathered in” - The American South created absurd voting requirements that targeted Black people and made it almost impossible to vote. The name for these requirements is the “Grandfather Clause.” They wrote the Amendment in a way to imply the practice was not discriminatory. They created stringent new voter requirements such as literacy tests. These requirements did not apply to people who had voted before 1867. Slaves did not know they were free until June 19, 1865. However, slavery was abolished on January 1, 1863, making it nearly impossible for a person formally kept in captivity to be legally allowed to vote.
  - Alternatives: legacied, exempted, preapproved
- “Master \_\_\_\_\_” - using the term Master to describe something that is the main or centralized source of information like a PowerPoint presentation or slide deck is inappropriate due to the connotations associated with slavery.
  - Alternatives: primary, main
- “White Paper” - while this term is widely used to describe an authoritative document, the term has historical implications that evoke negative associations especially with Tribes.
  - Alternatives: Issue paper, briefing document, prospectus

**Part II**

**Plan & Prepare**

## 5 Planning

## 6 Data Preparation

# **Part III**

## **Collect & Process**

## 7 Data Sources

## 8 Surveys

A good comprehensive guide for survey design can be found here: <https://files.eric.ed.gov/fulltext/ED619797.pdf>

Creating surveys that yield actionable insights is about details. And writing effective survey questions is the first step. You do not have to be an expert to build and distribute an effective online survey, but by checking your survey against tried-and-tested benchmarks, you can help ensure you are collecting the best data possible.

Tips for Building an Effective Survey:

1. Make Sure That Every Question Is Necessary.
2. Keep it Short and Simple.
3. Ask Direct Questions.
4. Ask One Question at a Time.
5. Avoid Leading and Biased Questions.
6. Speak Your Respondent's Language.
7. Use Response Scales Whenever Possible.
8. Avoid Using Grids or Matrices for Responses.
9. Rephrase Yes/No Questions if Possible
10. Take Your Survey for a Test Drive

<https://www.qualtrics.com/blog/10-tips-for-building-effective-surveys/>



## 8.1 Picking a Survey Software

Most Water Board staff will use Microsoft Forms which is available to all staff through the Microsoft 365 suite of applications. Microsoft Forms has a lot of advantages because of its integration with other Microsoft tools like Excel and PowerBi which allow for the survey results to be analyzed and visualized. Here is video on how to make that connection between Forms and PowerBi via Sharepoint that allows for consistent updating of results: <https://youtu.be/XBFVDedwLiY?si=O161oYja-FBhG1W7>

One issue with Microsoft Forms and other free software like Google Forms is that they produce wide data in Excel or Google Sheets. This type of data is more difficult to transform.

## 9 Data Limitations

In most cases, the Water Boards programs were not developed or designed to collect demographic data as a matter of process which means that most will rely on a few key data sources.



Figure 9.1: Data Benefit vs Risk

Incorporating a racial equity lens during data analysis includes incorporating individual, community, political, and historical contexts of race to inform analysis, conclusions, and recommendations. Solely relying on statistical outputs will not necessarily lead to insights without careful consideration during the analytic process, such as ensuring data quality is sufficient and determining appropriate statistical power. Disaggregation of data is also a series of tradeoffs. Without disaggregating data by subgroup, analysis can unintentionally gloss over inequity and lead to invisible experiences. On the other hand, when analysts create a subgroup, they may be shifting the focus of analysis to a specific population that is likely already over-surveilled. ([Centering Racial Equity Throughout Data Integration](#))

Centering racial equity means paying attention to which data are highlighted and how they are framed, as well as the readability and accessibility of the communication method. This involves strategic consideration of the audience and the mode of dissemination that most effectively conveys the information. There are many ways to communicate information. These

include briefs, interactive documents, websites, dashboards, social media content, data walks, posters, briefs, and infographics. Regardless of the form, content geared toward the public should avoid jargon that may be otherwise appropriate for internal program staff or academic audiences, while also using person-centered language and translating materials into languages most applicable to your community context. ([Centering Racial Equity Throughout Data Integration](#))

Furthermore, good quality data regarding marginalized communities is often lacking, but it is still important to discuss impacts to BIPOC communities. It may be appropriate in some cases to still present or analyze this data and also present caveats for the data limitations. In other cases, it may be more appropriate to rely only on qualitative discussion based on information derived from background research and feedback from affected communities.

**Part IV**

**Assure & Analyze**

## 10 Data Analysis

# **11 Data Visualization**

## **11.1 Data Visualization Tools**

## 12 Demographics Data

Adding demographics data to your data project can help increase understanding of potential correlations or relationships between your data and demographic and socioeconomic characteristics of locations of interest.

Depending on what demographics data sources you decide to use, the methods needed to combine, overlay, or compare with the data you are interested in may vary. Below we outline methods of comparing demographics data to point, line, and polygon data types.

### 12.1 Demographics Data Needs Context

It's important to remember that you can always benefit from setting context before trying to communicate demographic or specific racial equity answers to questions posed by our Board, a member of the public, or a partner in this work. What your program is about, what does it do, how well does it do these things (aka Performance Report), etc. This may take a few different visualizations to help frame the context of the program mission but will help viewers understand how you are approaching the racial equity data work within the scope of your program.

Often the go-to resources for making inferences to demographic and socioeconomic characteristics is the National Census dataset and the associated American Community Survey dataset. While we are fortunate to have just updated this dataset in 2020 there are limitations and potential inaccuracies associated with relying solely on census data to enumerate demographic characteristics within a given census tract. This [tool from the Department of Finance exists to measure this limitation](#).

A detailed example of using R programming to estimate demographics and other characteristics with U.S. census data to be used for custom spatial features is available and can be tailored to programs with the help of a data scientist proficient in R and staff familiar with the program. [https://daltare.github.io/example-census-race-ethnicity-calculation/example\\_census\\_race\\_ethnicity\\_calculation.html](https://daltare.github.io/example-census-race-ethnicity-calculation/example_census_race_ethnicity_calculation.html)

## 12.2 Commonly used Racial Equity Data Sources

1. [CalEnviroScreen 4.0](#)
2. [United States Census Data](#)
3. [American Community Survey Data](#)
4. Internal Administrative Data (i.e. Human Resources data)
5. Local and Regionally Collected Data

## 12.3 For a growing list of Water Boards Developed Racial Equity Data Tools:

[California Water Boards Racial Equity Data Resource Hub](#)

More general data literacy resources on our website and in our [toolkit](#).

## 12.4 Collecting Data

## 12.5 Point Data

```
library(tidyverse)
library(palmerpenguins)
penguins |>
  mutate(
    bill_ratio = bill_depth_mm / bill_length_mm,
    bill_area  = bill_depth_mm * bill_length_mm
  )
```

①  
②

- ① Take `penguins`, and then,
- ② add new columns for the bill ratio and bill area.

## 12.6 Lines



## 12.7 Polygons

## 13 Working with CalEnviroScreen

[CalEnviroScreen](#) can be a helpful tool in creating visualizations and performing analysis as it provides a number of index, as well as a “rolled-up” score that combines environmental and demographic data together. However, there can be things to consider, a couple of which are discussed below.

### 13.1 Missing Values for CalEnviroScreen Scores

Users conducting an analysis with the [CalEnviroScreen](#) (CES) 4.0 dataset should be aware that it contains missing values, both for individual indicators and overall CES scores. These missing values are distinct from zeros, which are also in the CES dataset. For more information about the missing (and zero) values, see the data dictionary ([calenviroscreen40resultsdatadictionary\\_F\\_2021.pdf](#)) that accompanies the CalEnviroScreen 4.0 results Excel workbook, available for download as a zip file [here](#).

In the CES 4.0 data (for the version available as of April 2023), the shapefile containing CES 4.0 scores (available [here](#)) encodes these missing values as negative numbers (-999 for most variables, and -1998 for one variable). The Excel workbook containing CES 4.0 scores (available [here](#)) encodes these missing values as NA. Users should account for these missing values – and their different encodings – as needed when doing any analysis using CES data. Also, note that the CalEnviroScreen 3 shapefile (June 2018 update version) encoded missing values as 0, so users should be aware of this change if/when updating an analysis from CES 3 to CES 4.0 data.

### 13.2 Inconsistent Census Tract Boundaries in CalEnviroScreen 4.0 Shapefile

In the CES 4.0 data (for the version available as of April 2023), the shapefile containing CES 4.0 scores (available [here](#)) uses a simplified version of the polygons that represent 2010 census tracts. The boundaries of the census tracts defined by these simplified polygons do not always align with the boundaries of neighboring census tracts, resulting in slight gaps or overlaps between some neighboring census tracts. These inconsistencies are not likely to have a significant impact on most uses of the CES data, but they could impact some types of

analysis based on CES data. For example, when assessing sites or facilities based on the CES score of the census tract they are located in, sites located near a census tract boundary could be associated with more than one census tract (and more than one CES score) in areas where there are overlapping census tract polygons, or not associated with any census tract (and no CES score) in areas where there are gaps between census tract polygons. This issue may be addressed in a future release of the CES dataset; in the meantime, a possible workaround is to use the official 2010 census tract boundaries from the US Census Bureau for any calculations, then use census tract IDs to tie this information to the associated CES score for each tract.

## 14 Turning Data into Information is Challenging - Even More so with Racial Equity Data Work

Turning data into information in the context of racial equity involved navigating complex ethical considerations. The process requires an understanding of the potential impact on Black, Indigenous, and other People of Color (BIPOC) communities and the responsibility to mitigate perpetuating or reinforcing biases. Upholding ethical standards requires a commitment to maintaining privacy, accessibility, and fostering transparency throughout the data transformation process. Additionally, acknowledging the limitations of the data and being transparent about potential biases is essential for maintaining the integrity of the data and information generated and shared. The transformation of racial equity data into meaningful information requires a thoughtful and intentional approach which we will highlight in the next sections.

For example, many programs will rely on demographic and socioeconomic data, like those collected from the [U.S. Census](#) and the [American Community Survey](#) (ACS). Because the ACS is based on a sample, rather than all housing units and people, ACS estimates have a degree of uncertainty associated with them, called sampling error. In general, the larger the sample, the smaller the level of sampling error. To help users understand the impact of sampling error on data reliability, the Census Bureau provides a “margin of error” for each published ACS estimate. The margin of error, combined with the ACS estimate, gives users a range of values within which the actual “real-world” value is likely to fall. It is important to acknowledge this uncertainty up front to be transparent with your audience about the data and conclusions you are drawing. For more information on using [ACS data please see the Understanding and Using American Community Survey Data: What All Data Users Need to Know Handbook](#). Also for an interactive tool that measures the potential inaccuracies associated with relying on census data to enumerate demographic and socioeconomic characteristics in California please explore: <https://cacensus.maps.arcgis.com/apps/webappviewer/index.html?id=48be59de0ba94a3dacff1c9116df8b37>.

Another example of how to tell your data story in a transparent way can be found on the Office of Environmental Health and Hazard Assessment CalEnviroScreen 4.0 Race and Equity Analysis. <https://storymaps.arcgis.com/stories/f555670d30a942e4b46b18293e2795a7>

## 15 Demographics Data Needs Context

It's important to remember that you can always benefit from setting context before trying to communicate demographic or specific racial equity answers to questions posed by our Board, a member of the public, or a partner in this work. What your program is about, what does it do, how well does it do these things (aka Performance Report), etc. This may take a few different visualizations to help frame the context of the program mission but will help viewers understand how you are approaching the racial equity data work within the scope of your program.

Often the go-to resources for making inferences to demographic and socioeconomic characteristics is the National Census dataset and the associated American Community Survey dataset. While we are fortunate to have just updated this dataset in 2020 there are limitations and potential inaccuracies associated with relying solely on census data to enumerate demographic characteristics within a given census tract. A tool from the Department of Finance exists to measure this limitation and is available here <https://cacensus.maps.arcgis.com/apps/webappviewer/index.html?id=48be59de0ba94a3dacff1c9116df8b37>.

A detailed example of using R programming to estimate demographics and other characteristics with U.S. census data to be used for custom spatial features is available and can be tailored to programs with the help of a data scientist proficient in R and staff familiar with the program. [https://daltare.github.io/example-census-race-ethnicity-calculation/example\\_census\\_race\\_ethnicity\\_calculation.html](https://daltare.github.io/example-census-race-ethnicity-calculation/example_census_race_ethnicity_calculation.html).

**Part V**

**Preserve & Store**

## **16 Preservation & Storage**

**Part VI**

**Publish & Share**



## 17 Data Sharing

add guidance on open/sharing practices & locations - GitHub, Open Data Portals, etc.

## **Part VII**

# **Discover & Integrate**

## **18 Preservation & Storage**

# **Part VIII**

## **Describe**

## 19 Documentation

Add guidance on documentation, reproducibility & transparency, oh my!

This is intended to be a living document and providing documentation is extremely important.

# **Part IX**

## **Use Cases**

## 20 Use Case 1

change page title to short description of use case

## **21 Demographics Use Case**



## 22 Resources

Here you will find a curated list of presentations, webpages and other resources related to the development, implementation and scaling of the principles and practices outlined in this Data Equity Handbook.

All Water Boards authors are **bolded** below.

### 22.1 Other Equity & Data Handbooks / Toolkits / Guides

- [Water Boards Developed Racial Equity Data Tools](#)
- [GARE Racial Equity Toolkit: An Opportunity to Operationalize Equity](#)
- [Beyond Compliance Network Advocacy Toolkit](#)
- [Academic Data Science Alliance Data Science Ethos](#)
- [CA Water Boards College of Water Informatics Data Toolkit](#)
- [500 Women Scientists Guide for Inclusive Scientific Meetings](#)

### 22.2 Presentations

[Analyzing Water Boards and Demographic Data for Equity](#). Jun 2024. Hannah Cushman Garland. State Water Board Racial Equity Data Subcommittee Webinar. [Recording](#) | [Download and Use the Code](#) | [View Code](#)

### 22.3 Relevant Government Documents

#### 22.3.1 California Water Boards

- Racial Equity Resolution (2021) - [English](#), [Español](#)
- Racial Equity Action Plan (2023) - [English](#), [Español](#)
- [Racial Equity Resolution Annotated References](#) (2021)

### **22.3.2 California Office of Data and Innovation**

- [California Data Strategy Report](#) (2024)
- [California Data Strategy](#) (2020)

### **22.3.3 The White House**

- [The Environmental Justice Science, Data, and Research Plan](#) (2024)
- [Year of Open Science Fact Sheet](#) (2023)

## **22.4 Relevant Literature**

O'Brien, M., Duerr, R., Taitingfong, R., Martinez, A., Vera, L., Jennings, L.L., Downs, R.R., Antognoli, E., ten Brink, T., Halmai, N.B., David-Chavez, D., Carroll, S.R., Hudson, M. and Buttigieg, P.L. (2024) '[Earth Science Data Repositories: Implementing the CARE Principles](#)', *Data Science Journal*, 23(1), p. 37.

## **22.5 Websites**

- [Racial Equity at the Water Boards](#)
- [Openscapes at the Water Boards](#)
- [California Office of Data and Innovation](#)
- [Government Alliance on Race and Equity \(GARE\)](#)

## 23 Inspiration

The impetus for developing this Data Equity Handbook began in August 2020 when the State Water Resources Control Board (State Water Board) publicly acknowledged that the historical effects of institutional racism must be confronted throughout government, and it directed staff to develop a priority plan of action.

Since then, the State Water Board, its [Office of Information Management and Analysis](#) (OIMA) and OIMA's many internal and external partners have been developing and compiling material, and adding to this Data Equity Handbook as time and bandwidth allow.

This Data Equity Handbook is inspired by many sources, including:

- [Water Boards Racial Equity - Resolution and Related Actions](#)
- [Government Alliance on Race and Equity](#) (GARE)
- [Openscapes](#) and their [Approach Guide](#)
- [NOAA Fisheries](#) (NMFS) [Open Science Resource Book](#)