Data for Environmental Justice (d4ej)

Tiered modules for introducing environmentally and socially conscious data science

March, 2023

Table of Contents

# About These Modules

This is a collection of modules intended to teach students about environmental justice through the [EPA’s EJ Screen tool](https://www.epa.gov/ejscreen).

These modules use a three-tiered system to educate across experience levels. Modules are also broken up by the intended audience: data science and/or STEM majors, social science majors, and those participating in socially engaged arts learning.



# Audience and Tiered Structure

These modules are made to provide opportunities for students, especially **underserved and minority students**, at higher education facilities to explore and critically think about environmental justice data using EJScreen and other tools.

We intend for these modules to be discipline agnostic, so that students of multiple disciplines and majors can participate in environmental data science. We have also designed these modules with under-resourced instructors in mind, such as those teaching at Minority Serving Institutions, community colleges, primarily undergraduate universities, and historically women’s colleges.

In order to provide for a range of experience levels, we have created **three tiers** of educational content. The tier 1 module targets the lower division level undergraduate students, across a variety of disciplines (e.g., biology, social science).

Tiers 2 and 3 will target more advanced students, potentially junior and senior undergraduate students, graduate students or students completing a research capstone. Tiers 2 and 3 will also assume some background knowledge in data science. Finally, we have included additional follow-on learning objectives for students interested in social science disciplines or socially engaged art. We encourage instructors to gauge their students’ interests and mix objectives from the different tiers and disciplines.

# 1 Guide for Instructors

## 1.1 Tier 1

Why are we doing this? Value of:

* Env Justice exploration
* Increasing relevance of learning material for students
* Data science skills for career potential (even the most basic of exposure)

Student-centered teaching approaches

How to empathetically approach environmental justice issues

Data Science ethics (and a slide for them to use in their courses?)

Guide for assessing student learning

## 1.2 Tier 2

Video tutorial with associated R code

Also a pdf (R Markdown) guide to follow

Guide for assessing student learning

## 1.3 Tier 3

Best-practices for guiding/managing student independent projects and assessing students equitably

Guide for assessing student learning

# 2 Learning Outcomes by Subject

When students complete this module, they will be able to:

## 2.1 Science/Data Science

### 2.1.1 Overarching LOs - to be applied at all tiers

* Understand how data science can be used to create environmental solutions for communities
* Place data science questions in context (ecological, environmental, community solution, etc)
* Understand complexities and limitations of data
* Evaluate drawbacks/benefits of tools like EJScreen
* Interpret results in context (ecological, environmental, community solution, etc)

### 2.1.2 Tier 1 (Intro level)

*Prerequisite Knowledge: None!*

* Explain how environmental indices can affect their community
* Evaluate the differences in the tools (EJScreen vs CEJST vs state-based(?) tools)
* And the benefits/drawbacks of the tools and how underlying data influences results (e.g., EJScreen uses census data - that is biased)
* Evaluate the positives and negatives of abstracting a place to one number
* Understand how weighting can impact results
* Question policy-makers and land managers on environmental justice issues
* Collaboratively develop action plans to move forward from their findings (wording of this sentence?)

### 2.1.3 Tier 2 (Mid level)

*Prerequisite Knowledge: Basic introduction to data science and statistical analyses, e.g.*

* Access data through R
* Execute pre-written example code and interpret the results
* Construct and modify R code to test hypotheses
* Choose a place and tell a story about why it is identified as an EJ place. What is missing? Is there a place that you thought would show up in EJ screen but does not? What data gap makes that happen?

### 2.1.4 Tier 3 (Upper Division)

*Prerequisite Knowledge:*

* Student-driven project initiatives (SMART principles)
* Formulate a testable question
* Justify why this question is interesting with appropriate background information
* Create a justified hypothesis
* Obtain data from public sources (like EJ screen)
* Process raw data into usable formats
* Analyze data with appropriate statistical methods to answer the question
* Visualize data
* Contextualize results in broader context ((ecological, environmental, community solution, etc)
* Communicate results through - e.g. a paper, poster, flash talk, other format
* quantitative models to address scientific questions?
* Testable question
* Placed in the context
* Obtaining, cleaning, transforming, and processing raw data into usable formats?
* Apply a range of statistical methods for inference and prediction…
* Build data science products that can be used by a broad audience - or can be transferable to other broader contexts

## 2.2 Social Science:

Geared towards students who Never have made a map before

### 2.2.1 Tier 1:

* Explain how environmental data science tools reflect our understandings of race and can both perpetuate and challenge racism
* Interpret maps
* Expand understanding of maps (through resources like this counter mapping project and memory maps)
* Navigate the EJScreen tool and/or other similar tools to answer relevant, student-generated research questions about environmental (in)justice
* Understand how these can benefit their own community and neighborhood

### 2.2.2 Tier 2:

* Involve in ethnographic studies
* Be able to infer data with a broad socio-economic context
* Visualization of data using programming languages such as R
* Maybe tie-up with different environmental law firms to get a hands-on learning experience by interning/volunteering!
* Be able to come up with concept maps to project a boarder relationship with different interactions
* Think of gathering qualitative data through interviews and surveys that are based around ethics

### 2.2.3 Tier 3:

* Placement opportunities for students interested in continuing this field of science
* Introduce public health implications of the data and research?
* Discuss data ethics?

## 2.3 Socially Engaged Art:

Geared towards students who

* Are interested in creatively expressing and communicating their data analysis
* Are interested in connecting and engaging in reciprocal story sharing with local community members about pertinent environmental justice issues

### 2.3.1 Tier 1: communication (2-3 weeks)

*Tier objective* : introduce students to science communication, socially engaged art, and research translation with hands-on activities between students and with the general public

* Students read foundational literature on the history of socially engaged art practices, and how science, art, and agency are tied together.
* Students create a representation of the results from the Data Science and Social Science subjects that can be shared with classmates and the broader community
* Representations could take the form of ArcGIS StoryMaps, collage, art installation, composition, art/dance/theater performance, a poster, presentation, etc.
* Do a site visit with students and teachers to see the EJ community first-hand and learn from locals (example: Dakota Bdote tour) Class creates a gallery show and/or hosts an event to share creative works with each other and community members.

### 2.3.2 Tier 2: storytelling (2 weeks)

*Tier objective* : Students and community members come to a more holistic understanding of the different experiences and perspectives related to environmental justice, for example of how personal experiences are part of shared experiences or a larger picture

* Organize a gathering of students and residents with different breadths of traditional and ancestral knowledge like teachers and Indigenous leaders.
* Storytelling preparation:
  + Hold a reflection session (individually or in groups) and a writing workshop to be able to put ideas and thoughts into words more effectively
  + Run a storytelling workshop for students to practice telling and listening to stories
* Facilitate an organic sharing and listening of stories between students and community members related to environmental justice from embodied experiences, research, data analysis

### 2.3.3 Tier 3: co-creation of knowledge (longterm, multi-year)

*Tier objective*: Build and sustain healthy relationships between students, local stakeholders, and Indigenous leaders. Over time, co-create a collective understanding of the root causes of environmental justice issues in the local community, brainstorm ways to sustainably address these issues, and empower the community to tackle these issues.

* Enable communication channels for continued support between students and residents, knowledge exchange and future collaborations.
* Engage students and community members in regular meetups and activities to develop a community of students engaged in environmental justice
* Support participants (students and community members) through funding sources
* Secure funding for a competition for participants to propose a new or an extension to an existing EJ project that the winning team can work on for a year.

Multi-lingual: <https://www.enlightenment.org/develop/legacy/program_guide/multilingual_pg>

# 3 Assessment

## 3.1 Student Assessment:

### 3.1.1 Tier 1

* Exam questions
* Give a benefit and drawback of using one environmental index to describe an area. Possible answers: Benefits: relatively low-cost, easy, utilized pre-existing resources. Drawbacks: May or may not be an accurate representation of the health, no way to check it. A lot depends on the weighting, quality of the data going in, etc.
* Suppose you look at two different indices and get very different numbers (e.g., one assigns an area 50/100, the other assigns 75/100). Give two possible reasons for these differences. Possible answers: Includes different raw data, different weighting, includes different variables.
* Guided questions for any pre reading
* Guided worksheet for students throughout module
* Low-pressure quiz questions
* Art project-draw a scene from your neighborhood. Include the three biggest things you see as impacting the environmental health of your neighborhood. Are - all three on the index? If not, how hard would it be to collect data to include them? Could it be collected nationwide?

### 3.1.2 Tier 2

* Individual or collaborative project
* Group projects encouraging working together

### 3.1.3 Tier 3

* Build your own case study and come up with data analysis to present to the class
* Several principles or programming functions demonstrated successfully (e.g., data wrangling / specific R functions)
* Apply to your own research question (if graduate student)

## 3.2 Assessment & evolution of the module itself

* Assessment & evolution of the module itself (Nate, Ellen, ):
* Survey of students pre- and post- course (self efficacy, excitement for data science, data science is relevant to me, belonging indices, etc)
* Survey of faculty/instructors that are actually teaching the course (self-efficacy)
* Incorporate feedback into further development of the module

**Online repository for student submissions, so students can see what others are working on across the country**

# 4 Introduction

**Note fantastic slides from Lourdes Vera here:** [**https://docs.google.com/presentation/d/1hGCkesvIgccjunUQzVRuHEp1aZNIDyMW9BdAonpN40M/edit#slide=id.p**](https://docs.google.com/presentation/d/1hGCkesvIgccjunUQzVRuHEp1aZNIDyMW9BdAonpN40M/edit#slide=id.p)

When organizations/governments need to make decisions that impact communities, they have to take a lot of information and condense it. It is not practical to go out to every site and measure every environmental variable in every location that they want to make a decision about. It costs money and takes time, both of which are always limited. The government/organizations will often look to easily-accessible data to make decisions about whether/how to do potentially damaging actions that impact the environment.

We are going to show you one tool that people use to make decisions that can impact your community, EJ Screen. EJ Screen is an environmental justice screening and mapping tool.

First, let’s take a look at the [EJ Screen](https://ejscreen.epa.gov/) website. Spend a few minutes poking around the site: What data are available? How are they represented?



# 5 Introduction

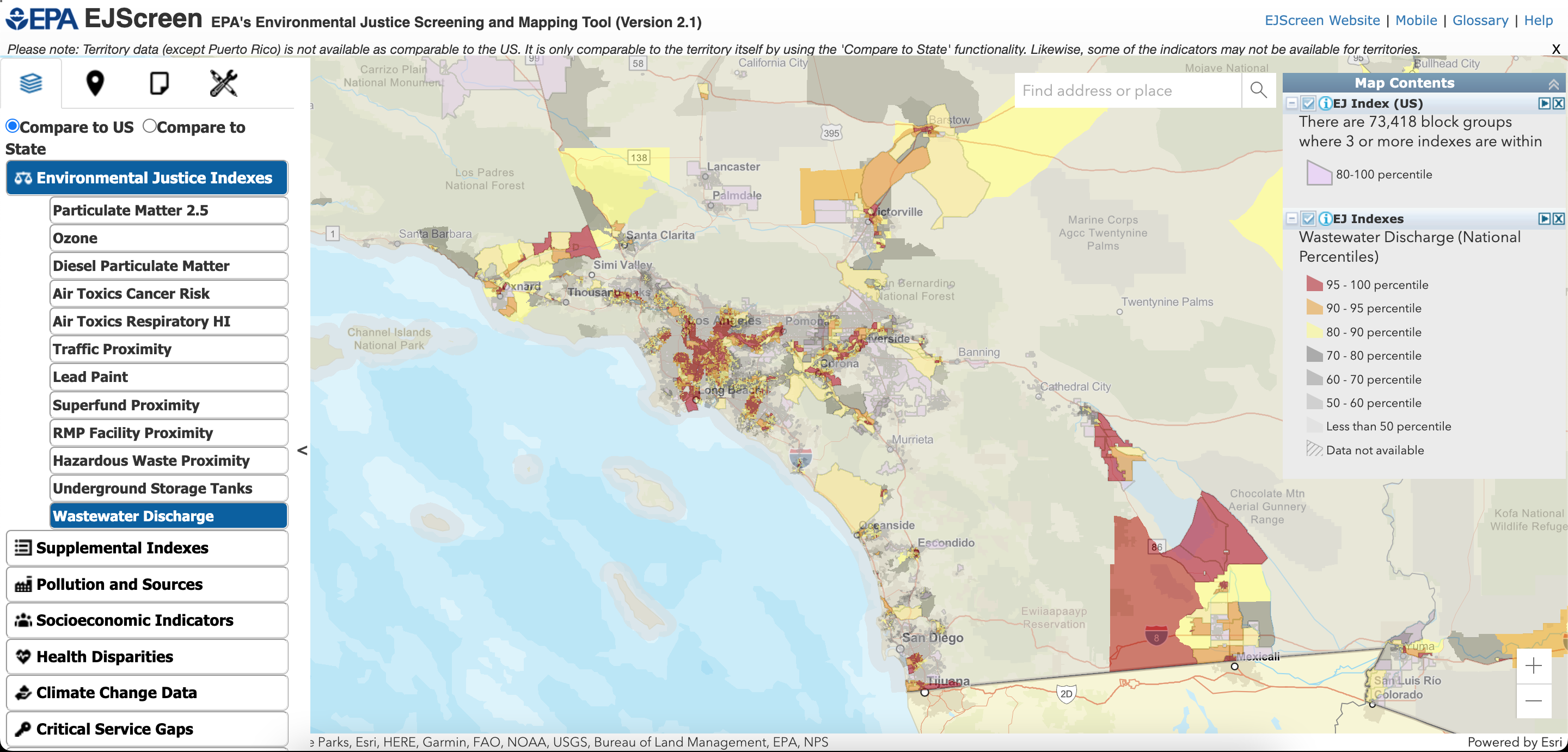
EJ Screen incorporates many different data sources. Let’s dig into one of the data sources that is part of EJ screen so we can understand it better.

We will import the data source used to make the wastewater index that is part of the EJ screen tool.

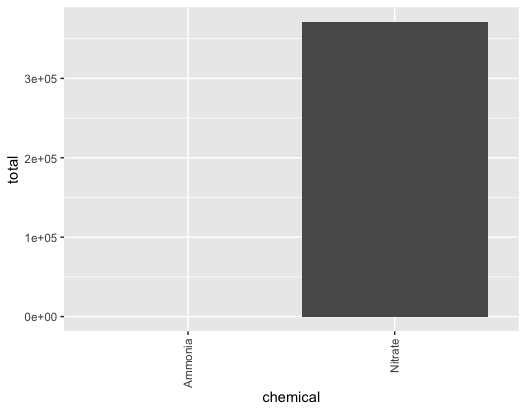
# 6 Activity

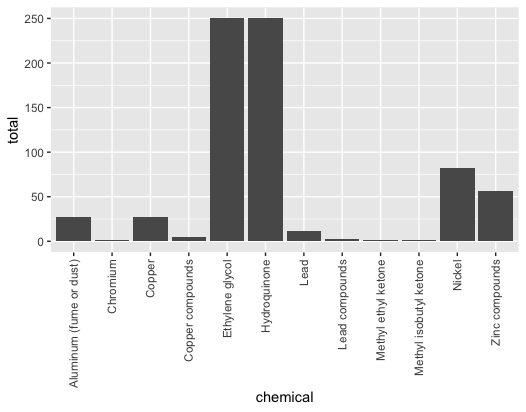
Looking at the EJ Screen website, wastewater discharge is an EJ factor in many Southern California regions.

We can dive into the data and look at the content of the wastewater in two different rural counties: Riverside and Imperial



We find that the chemical composition is very different! These are also counties with large amounts of agricultural run-off, but these are not accounted for in this database. The nitrate wastewater runoff in Imperial county is from the US Navy.





library(ggplot2)  
data <- read.csv(file = ‘TRI\_table\_CA2.csv’)  
county\_name = “IMPERIAL”  
county = data[data$COUNTY\_NAME == county\_name,]  
## What do the columns mean?  
# TOTAL\_PRODUCTION\_RELATED\_WASTE. = sum of all reports  
# TOTAL\_PRODUCTION\_RELATED\_WASTE..1 = average of all reports  
# TOTAL\_PRODUCTION\_RELATED\_WASTE..2 = count of reports  
# county$TOTAL\_PRODUCTION\_RELATED\_WASTE..5 = std of all reports  
# county$TOTAL\_PRODUCTION\_RELATED\_WASTE..6 = variance of all reports  
## Plot total by facility  
county1 = aggregate(x = county$TOTAL\_PRODUCTION\_RELATED\_WASTE., # Specify data column  
 by = list(county$FACILITY\_NAME), # Specify group indicator  
 FUN = sum)  
county1 <- county1[order(county1$x),]  
p<-ggplot(data=county1, aes(x=Group.1, y=x)) +  
 geom\_bar(stat = ‘identity’)  
p + theme(axis.text.x = element\_text(angle = 90, vjust = 0.5, hjust=1))  
## Plot total by chemical  
chemical = aggregate(x = county$TOTAL\_PRODUCTION\_RELATED\_WASTE., # Specify data column  
 by = list(county$CAS\_CHEM\_NAME), # Specify group indicator  
 FUN = sum)  
p<-ggplot(data=chemical, aes(x=Group.1, y=x)) +  
 geom\_bar(stat = ‘identity’)  
p + theme(axis.text.x = element\_text(angle = 90, vjust = 0.5, hjust=1))  
## Plot chemicals that are released into the water  
county\_water <- county[county$WATER\_TOTAL\_RELEASE > 0,]  
chemical = aggregate(x = county\_water$WATER\_TOTAL\_RELEASE, # Specify data column  
 by = list(county\_water$CAS\_CHEM\_NAME), # Specify group indicator  
 FUN = sum)  
p<-ggplot(data=chemical, aes(x=Group.1, y=x)) +  
 geom\_bar(stat = ‘identity’)  
p + theme(axis.text.x = element\_text(angle = 90, vjust = 0.5, hjust=1))

# About the Authors

Coming soon!

# NEXT STEPS

## 6.1 Short term:

* Organize target faculty to get their input on module
* Input from experts/practitioners familiar with EJScreen
* Dissemination of the modules and resources
  + ESA discovery (Alison)
  + CourseSource

## 6.2 Long term:

* Apply for a grant to work with teachers to develop modules and pilot the modules in their courses
* Select a team interested in working on submitting funding applications
* Identify and select grant opportunities for developing educational resources or funding a working group on environmental data science education
* Submit grant applications!
* Potential Funding Sources:
  + Research Experience for Teachers (NSF) <https://www.nsf.gov/pubs/2018/nsf18089/nsf18089.jsp>
  + Existing Data Science centers?: e.g. ESIIL: <https://esiil.org/>
  + Curriculum Development Grants (e.g. Cal State <https://www.calstate.edu/impact-of-the-csu/research/csuperb/Documents/2023CurriculumDevelopmentRFP-final.pdf>)
  + Educational grants: <https://vcresearch.berkeley.edu/brdo/funding-education-research-programs-and-training>
  + Institute of Education Sciences grants: <https://ies.ed.gov/funding/ncer_progs.asp>
  + Sloan Center For Systemic Change (Graduate Level): <https://sloan.org/programs/higher-education/diversity-equity-inclusion/scsc-call>
  + HHMI - no open relevant calls but has had relevant grants in the past
  + NEA Foundation: <https://www.neafoundation.org/educator-grants-and-fellowships/>
  + <https://www.spencer.org/grant_types/racial-equity-special-research-grants>
  + <https://www.spencer.org/grant_types/small-research-grant>
  + <https://mccartheydressman.org/teacher-development-grants/>
  + NSF division of undergraduate education: <https://beta.nsf.gov/funding/opportunities?sort_bef_combine=nsf_funding_upcoming_due_dates_DESC&f%5B0%5D=division%3A215>
* Translate/develop modules into different languages ? (maybe geared towards a different audience though)
* Make a data nugget? (no environmental justice focused ones already??) <https://datanuggets.org/search-current-data-nuggets/>
* FIX THE PACKAGE AND/OR Contact the maintainer

# Resources

[ejscreen R package](http://ejanalysis.github.io/ejscreen/)

Google slides for instructor to give background: <https://docs.google.com/presentation/d/1hGCkesvIgccjunUQzVRuHEp1aZNIDyMW9BdAonpN40M/edit#slide=id.p>

Background readings (and/or intro lecture?) to frame: What is EJ and why is it relevant? What is EJScreen? Why was this tool created and how is it used (and by who)? What data was used/is used for the tool? What are the general outputs of the tool? - general overview, diving into issues and understanding later

Google slides for lecture for class to give students background Guided worksheet for readings to help students break them down (and key with example answers) Do we want to revise the existing EJscreen pdf document that describes how to use it? Clearly commented code that underlies the module with potential variations

Definitions and government resources: AY Environmental and Climate Justice Definitions and History WebQuest

Global Environmental Justice Atlas <https://ejatlas.org/>

There is page an EJScreen Office Hours & Training

HackMD document–Project Overview: <https://hackmd.io/o2GWqTwHSjCDtKB3zV-QRg>

HackMD document–Tier 1: <https://hackmd.io/VWg4OEozTNi_DYO-0QZToA>

This is a very helpful breakdown on EJ Screen

One activity could be to analyze EJ screen through algorithmic ecology: <https://stoplapdspying.medium.com/the-algorithmic-ecology-an-abolitionist-tool-for-organizing-against-algorithms-14fcbd0e64d0>

Example of citizen science x socially engaged art practices (listening to climate stories, documenting phenology:) <https://phenology.umn.edu/>

Pollinator / environmental awareness x art internship for youth <https://plainsartbuzzlab.wixsite.com/buzzlab/about-us>

Integrating indigenous science to learning promoting inclusive science <https://futurumcareers.com/what-if-indigenous-science-were-part-of-the-science-curriculum> Environmental Justice Dashboards <https://pubhealthgis.uic.edu/environmental-justice/>

# References