# Linking Climate Opinions to US Census Data

•••

CSE 583 Robin, Yakelyn, Hamid, Rebeca

## Background

What demographic and social factors help predict climate change opinions?

Are climate opinions predictors of behaviors that impact emissions (for example transportation behaviors)?

In this project, we have created a tool to help **visually explore statistical relationships** between demographic variables and opinions on climate change.

#### Data sources

We used three publicly available sources of data, using values at the **county level** 

Yale Climate
Opinions
2018
(+ metadata)

County land area data (Census)

United States
Census
2015

#### Use cases

User: A researcher who is interested in 1) predictors of climate opinions 2) climate opinions as predictors of climate related behaviors.

This is a tool to help the user quickly assess relationships worth exploring (not for causal inference).

Visualize statistical relationships between climate opinions and demographic variables Visualize statistical relationships between climate opinions and reported transportation habits.

## Demo.

Link to Demo

## Design

Raw data Processed data Module Output

Yale Climate Opinions 2018 (+ metadata)

County land area data

United States Census 2015 Data sources merged at county level

Prep data

Estimate statistics

Correlation Regression P-value

Plot heatmap

Interactive plots
In HTML

#### Design

Raw data Processed data Module Output

Yale Climate Opinions 2018 (+ metadata)

County land area data

United States Census 2015 Using pandas

Using linregress from scipy

Data sources merged at county level

Prep data

Estimate statistics

Using bokeh with javascript callbacks

Correlation Regression P-value

Plot heatmap

Interactive plots
In HTML

## Design

Raw data Processed data Output

Yale Climate Opinions 2018 (+ metadata)

County land area data

United States Census 2015 Using pandas

• •

Prep data

Data sources merged at county level

Using linregress from scipy

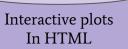
Estimate statistics

Correlation
Regression
P-value

Using bokeh with javascript callbacks

Plot heatmap





## **Project Structure**

#### Git hub repository

Data might end up in package directory

```
ClimOps
   climops
       calculate_statistics.py
      create_heatmap.py
       plot heatmap.py
      prepare_data.py
      heatmap.html
      scatter.html
     - tests
           _init__.py

    test calculate statistics.py

         - test_prepare_data.py
      - version.py
   data
      LND01.xls
      - YCOM 2018 Data.csv
      YCOM_2018_Metadata.csv
      acs2015 county data.csv
  Docs
      Component_specification.md
      Functional_specification.md
     climops.html
   Examples
   create_heatmap.ipynb
  images
   L logo.png
  LICENSE
  README.md
  requirements.txt
  environment.yml
   setup.py
```

#### **Lessons** learned

- Good practices for coding and software design
- Experience with Git and version control
- Proper package structure
- Divide tasks efficiently and set deadlines for specific work to be done
- Unit testing
- Things always take longer than expected...

#### **Future work**

We had several ideas which we weren't able to implement within the timeframe:

- Be able to control for certain variables (e.g. population density, or state)
- Perform PCA (principal component analysis) on the data.
- Bring in other data sets e.g. more detailed transport behaviour or natural hazard impacts.
- Maps to show if there is a spatial pattern to some of the relationships

## TEAM TO DO LIST (blue are not done yet)

- Final updates to functional and component specification
- Write installation instructions
- Write tests for stats module
- Finish tests for data cleaning
- Write examples of how to use our modules??
- Continuous Integration
- PEP8 and team code review
- Add this presentation and technology review to the documents
- Move data folder to climops directory