

## **Functional Specification**

- **Background:** (The problem being addressed)
  1. This project will create a data set that allows researchers to explore whether a relationship exists between beliefs about climate change, demographic characteristics and transportation behaviors (for example walking, cycling, driving behaviors, use of transit types and number of cars owned) in the United States.
  2. The final product would allow the user to explore (through basic statistics and visualizations) questions such as:
    - Do race, ethnicity, socioeconomic status (education and income), and gender is related to climate change perceptions?
    - Do people's beliefs about climate change have a relationship with their transportation behavior (for example, do people who believe that climate change is happening and caused by humans drive less/ own less cars/ use more public transit/ bike more?).
    - How public opinion about global warming changed relative to 2014 and we want to see if this change is related to other data ( e.g., census)
- **User profile:** (Who uses the system. What they know about the domain and computing (e.g., can browse the web, can program in Python))
  1. The user for this product is an academic researcher who is able to code in Python. They should be comfortable with manipulating data sets in python, visualization tools and performing basic statistical calculations.
- **Data sources:** (What data you will use and how it is structured)
  1. Yale Climate opinion maps (2014 - 2018)
    - Measures of public opinion about different aspects of global warming. This data is structured as % 'Yes' responses to each question in the survey at state, congressional district, and county levels.
  2. 2015 United States Census (income)
    - provides information on a wide range of social, economic, demographic, and housing characteristics. Topics covered include
    - income, employment, health insurance, the age distribution, and education, among many others. The data is organized by states as well as counties.
  3. 2017 National Household Travel Survey
    - a nationally representative survey and the primary source of data on the travel behavior of the American public. It includes daily travel by all modes, including characteristics of the people traveling, their household, and their vehicles, allowing for the analysis of personal and household travel. The data for the 2001, 2009, and 2017 waves of this survey are publicly available through the Federal Highway Administration (FHWA). The data is organized in four related files: personal, household, trip, and vehicle, each with different unit of analysis. The common identifier across each file is the Household ID.

- **Use cases:** (Describing at least two use cases. For each, describe: (a) the objective of the user interaction (e.g., withdraw money from an ATM); and (b) the expected interactions between the user and your system)
  1. Reveal most impactful relationships between climate opinion and demography (e.g., race, income, age, education)
    - What information does user provide?
      - Not much?...maybe they can select what variables they want to control for? Thoughts?
    - What responses does the system provide?
      - Table of correlation/regression/significance values for various pairs of intervals suitable for use in future research/policy decisions.
      - Some visualization of this table using a grid/colors
      - Maps of correlation/regression/significance where analysis is performed individually for each state.
  2. Reveal relationship between a specific climate opinion and demographic variable (choose by the user)
    - What information does user provide?
      - Select opinion of interest and demographic variable of interest from dropdown boxes ideally.
      - Perhaps some options for how the analysis is done (e.g. control for certain variables?)
    - What responses does the system provide?
      - Some correlation/regression/significance values
      - Maps of correlation/regression/significance where analysis is performed individually for each state.
  3. *Possible*: Reveal most impactful relationships between climate opinion and transport habits. (Part of use case 1)
  4. *Possible*: Reveal relationship between a specific climate opinion and transport habit. (Part of use case 2)

\*\* Note: the user will be able to download raw data from specific variables as well as graphics and statistical outputs for each use case.