## horizontal line



Project Proposal

DS 6040

**─**

Abigail Snyder

Elena Tsvetkova

Suraj Kunthu

# Data Set

The chosen data set is survey data from Yale Program on Climate Change Communication & George Mason University Center for Climate Change Communication.. The survey data catalogs American adults (age 18 and over) and their beliefs, attitudes, and policy opinions regarding global warming from 2008 through 2022. The data includes demographic data from the U.S. Census regarding gender, race and ethnicity, and level of education.

All of the data is in categorical form and stored as a .SAV file. There is a Survey Methods pdf which functions as a codebook for the data. The total weighted sample size of the entire dataset is (n=30,136). There are 30,136 observations and 54 features.

The citations for this data are below:

Yale Program on Climate Change Communication (YPCCC) & George Mason University Center for Climate Change Communication (Mason 4C). (2022). Climate Change in the American Mind: National survey data on public opinion (2008-2022) [Data file and codebook]. doi: 10.17605/OSF.IO/JW79P

Ballew, M. T., Leiserowitz, A., Roser-Renouf, C., Rosenthal, S. A., Kotcher, J. E., Marlon, J. R., Lyon, E., Goldberg, M. H., & Maibach, E. W. (2019). Climate Change in the American Mind: Data, tools, and trends. Environment: Science and Policy for Sustainable Development, 61(3), 4-18. doi: 10.1080/00139157.2019.1589300

# Questions

Some of the questions we are interested in are:

1. What is the probability that one’s political affiliation affects opinions on climate change?
2. What is the probability that demographic data, such as income and education determines opinions on climate change?

# Specifications

To answer this question, the variable *party* will be of importance, which asks survey respondents which political party they think of themselves as, with the options: -1. Refused, 1. Republican, 2. Democrat, 3. Independent, 4. Other; please specify:, 5. No party/not interested in politics. *ideology* may be the other variable that could be considered as a predictor variable, which asks respondents to classify themselves as: -1. Refused, 1. Very liberal, 2. Somewhat liberal, 3. Moderate, middle of the road, 4. Somewhat conservative, or 5. Very conservative.

In order to examine opinions on climate change, there are a number of variables that could be considered, including:

*happening:* Is global warming happening?

*harm\_future\_gen:* How much do you think global warming will harm future generations of people?

*worry:* How worried are you about global warming?

# Potential Solutions

## Bayesian Classification Models (LDA, QDA, etc.)

Can we use classification models to determine the probability that someone will be concerned / not concerned about climate change based on their demographic data?

# Roles

All three of us plan to equally split the load of the project. As such, there are no specific roles.