# Environmental Justice and Climate Change Sentiment

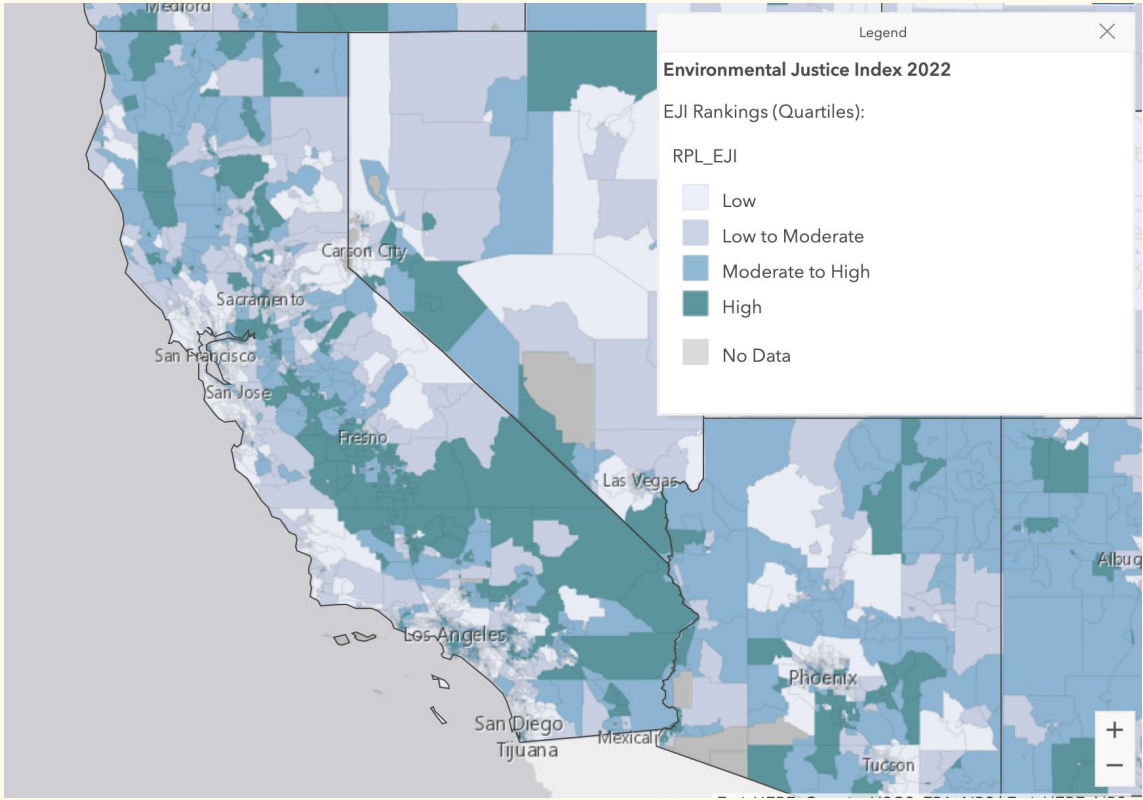
Bootcamp Final Project

Bella Mullen

# Dataset 1: Environmental Justice Index

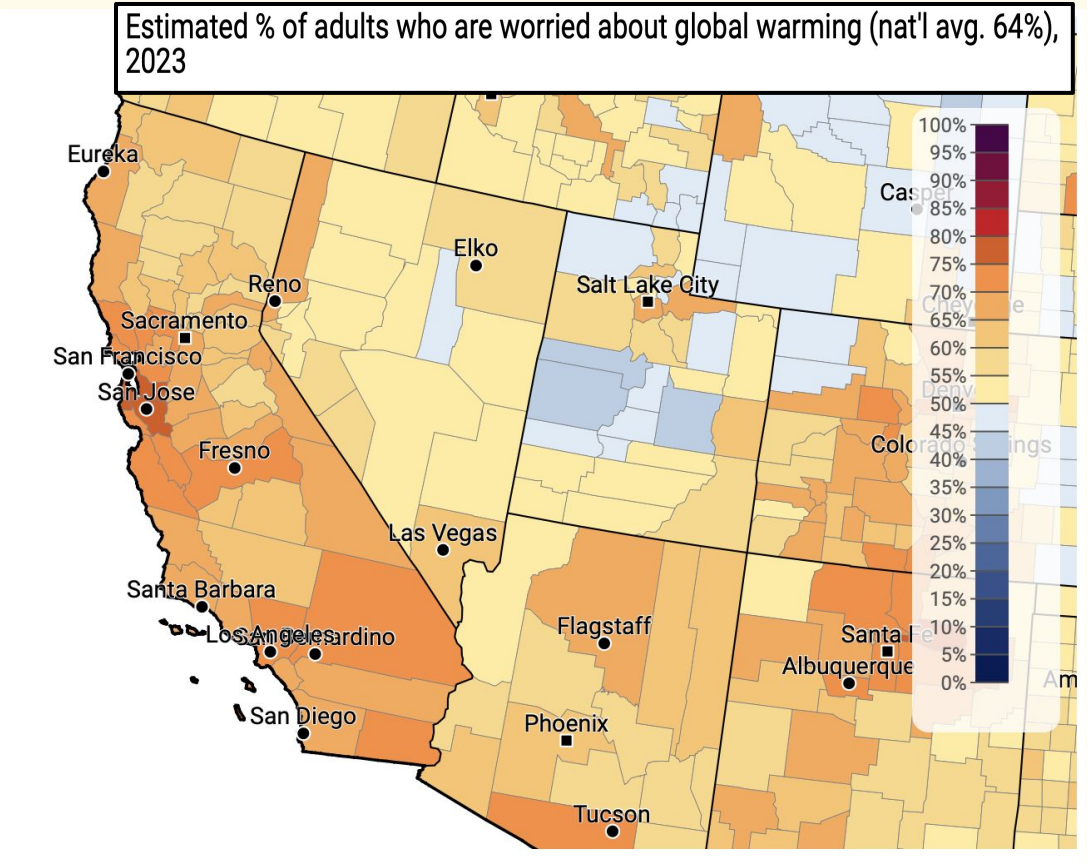
A tool designed to measure the cumulative impacts of environmental burden through the lens of human health and health equity

- This goal is for everyone to enjoy the same degree of protection from environmental and health hazards, and to live, learn, and work in a healthy environment
- Data from:
  - U.S. Census Bureau
  - U.S. Environmental Protection Agency
  - U.S. Mine Safety and Health Administration
  - U.S. Centers for Disease Control and Prevention



# Dataset 2: Yale Climate Opinion Maps

- Data through fall 2023
- Provides estimates of U.S. climate change beliefs, risk perceptions, and policy preferences at the state and local levels



```
# Selecting relevant features (independent variables) from the dataset 'eji_co_merge' These
X = eji_co_merge[['E_OZONE', 'E_PM', 'E_DSLPM','E_TOTCR', 'E_NPL', 'E_TRI', 'E_TSD', 'E_RMP',
y = eji_co_merge[['citizens_combined']]

# Splitting the data into training and testing sets with a 80-20 split
# 'random_state=42' ensures reproducibility of the results
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

```
# Initializing a Lasso regression model and setting up a parameter grid for alpha values to
# Alpha controls the strength of the regularization (higher alpha -> more regularization)

lasso = Lasso()
parameters = {'alpha': np.arange(0.01, 1, 0.05)}
```

```
# Performing grid search cross-validation to find the best alpha value for the Lasso model
reg = GridSearchCV(lasso, parameters)
reg.fit(X_train, y_train)

# Extracting the best Lasso model with the optimal alpha value
## Using the best parameters
model = reg.best_estimator_
```

```python
# Creating a DataFrame to display the Lasso coefficients corresponding to each feature
lasso_coefficients = pd.DataFrame({
    'Feature': column_names,
    'Coefficient': model.coef_
})

print(lasso_coefficients)
```

| | Feature | Coefficient |
|---|---|---|
| 0 | Days above O3 standard | -0.000000 |
| 1 | Days above PM2.5 standard | -0.000000 |
| 2 | Ambient concentrations of diesel | 0.000000 |
| 3 | Probability of contracting cancer | -0.000000 |
| 4 | Proximity to EPA National Priority List site | 0.000000 |
| 5 | Proximity to EPA Toxic Release Inventory site | 0.000000 |
| 6 | Proximity to EPA Treatment, Storage,and Dispos... | 0.000000 |
| 7 | Proximity to EPA risk management plan site | -0.000000 |
| 8 | Proximity to coal mines | 0.000000 |
| 9 | Proximity to lead mines | 0.000000 |
| 10 | Proximity to green space | 0.188540 |
| 11 | Houses built pre-1980 | 0.037684 |
| 12 | Walkability | 0.000000 |
| 13 | Proximity to railroad | 0.000000 |
| 14 | Proximity to high-volume road or highway | 0.000000 |
| 15 | Proximity to airport | -0.000000 |
| 16 | Impacted watershed | 0.000000 |

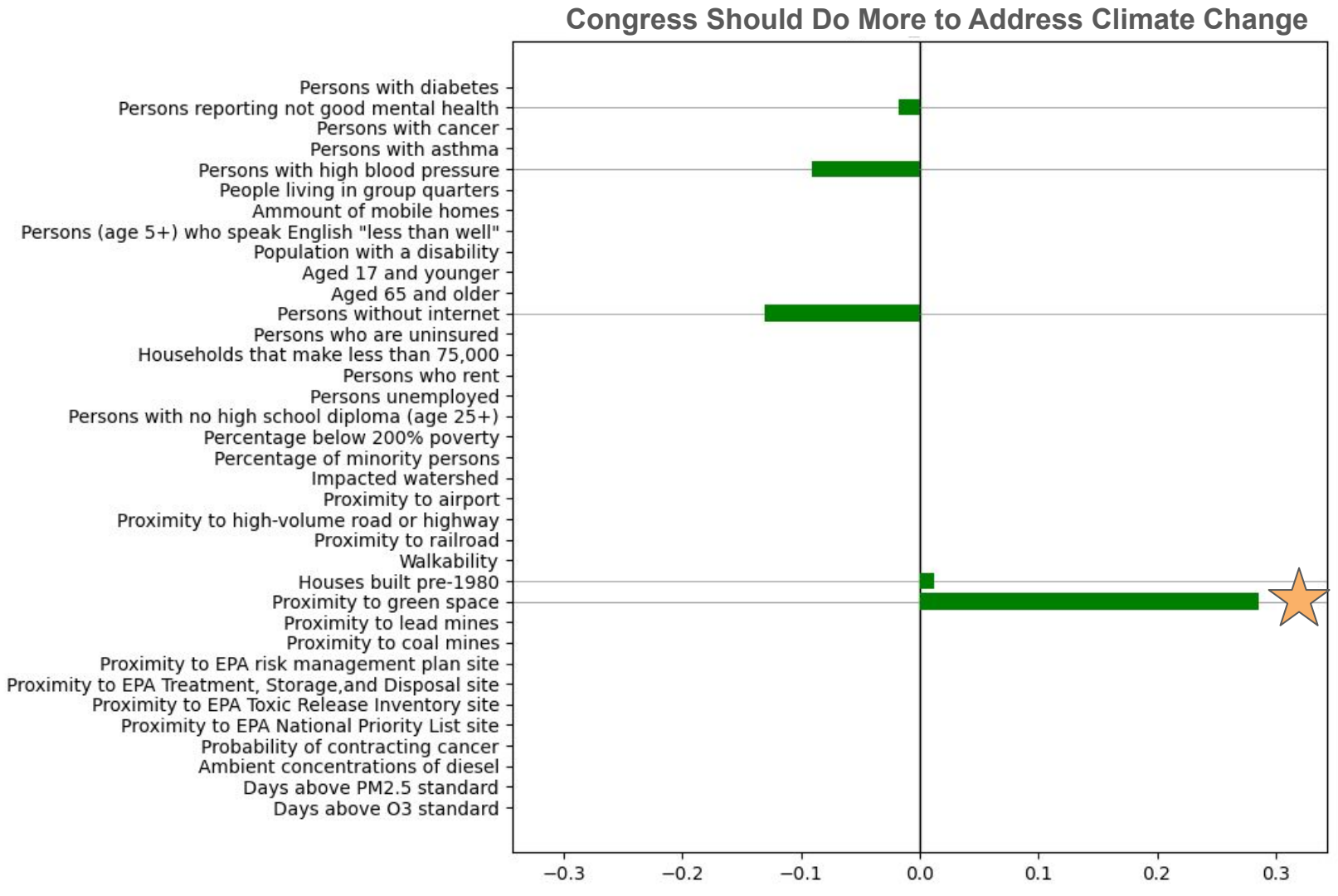
# Who should act?

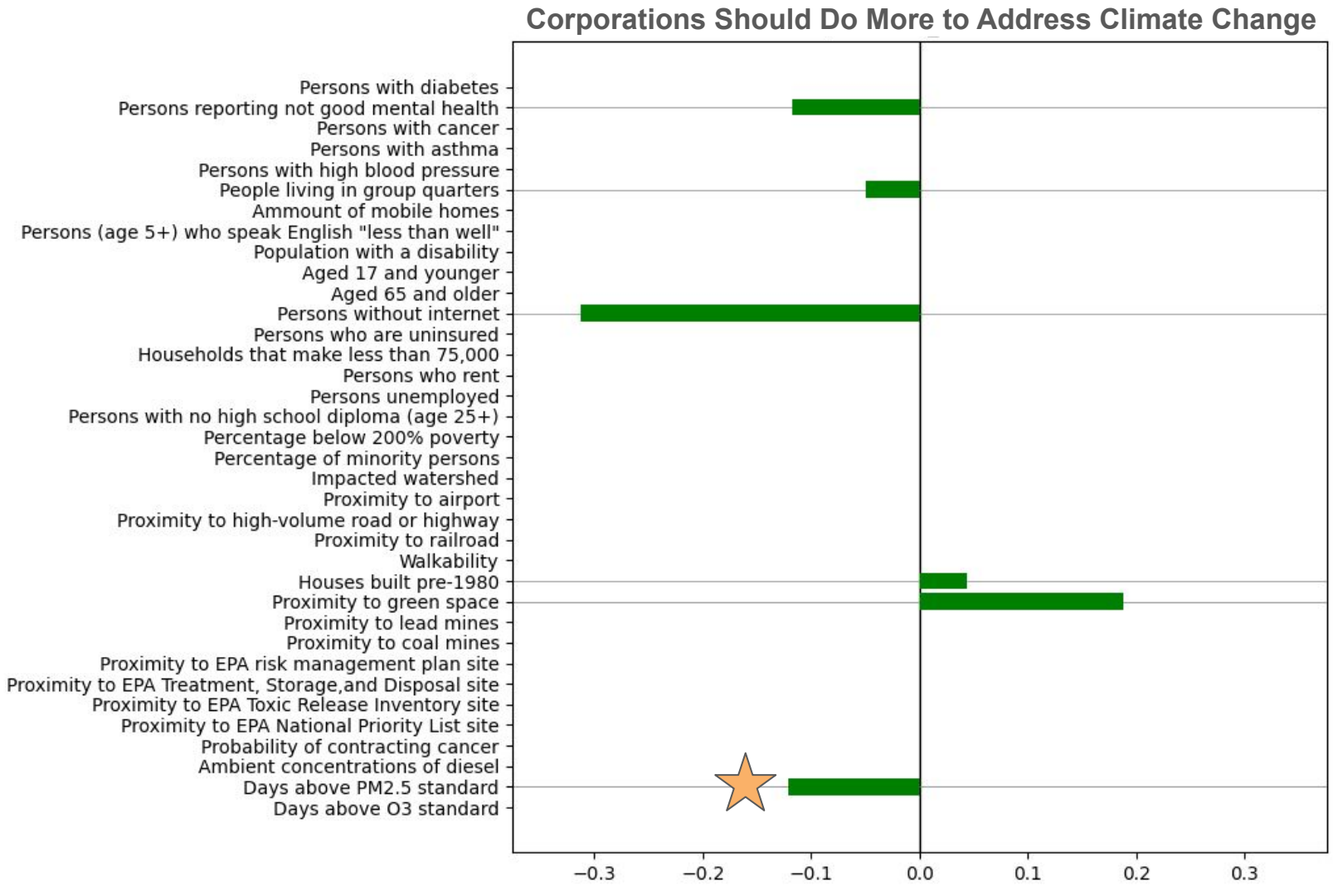
- **Citizens_combined:**
  - Estimated percentage who think citizens themselves should be doing more or less to address global warming.
- **Congress_combined:**
  - Estimated percentage who think Congress should be doing more or less to address global warming
- **Corporations_combined:**
  - Estimated percentage who think corporations and industry should be doing more or less to address global warming

| Citizens | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.188540 |
| | Persons without internet | -0.258834 |

| Congress | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.285941 |
| | Persons without internet | -0.130169 |

| Corporations | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.187994 |
| | Persons without internet | -0.313111 |


Congress Should Do More to Address Climate Change
Persons with diabetes
Persons reporting not good mental health
Persons with cancer
Persons with asthma
Persons with high blood pressure
People living in group quarters
Ammount of mobile homes
Persons (age 5+) who speak English "less than well"
Population with a disability
Aged 17 and younger
Aged 65 and older
Persons without internet
Persons who are uninsured
Households that make less than 75,000
Persons who rent
Persons unemployed
Persons with no high school diploma (age 25+)
Percentage below 200% poverty
Percentage of minority persons
Impacted watershed
Proximity to airport
Proximity to high-volume road or highway
Proximity to railroad
Walkability
Houses built pre-1980
Proximity to green space
Proximity to lead mines
Proximity to coal mines
Proximity to EPA risk management plan site
Proximity to EPA Treatment, Storage,and Disposal site
Proximity to EPA Toxic Release Inventory site
Proximity to EPA National Priority List site
Probability of contracting cancer
Ambient concentrations of diesel
Days above PM2.5 standard
Days above O3 standard
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3


Corporations Should Do More to Address Climate Change
Persons with diabetes
Persons reporting not good mental health
Persons with cancer
Persons with asthma
Persons with high blood pressure
People living in group quarters
Ammount of mobile homes
Persons (age 5+) who speak English "less than well"
Population with a disability
Aged 17 and younger
Aged 65 and older
Persons without internet
Persons who are uninsured
Households that make less than 75,000
Persons who rent
Persons unemployed
Persons with no high school diploma (age 25+)
Percentage below 200% poverty
Percentage of minority persons
Impacted watershed
Proximity to airport
Proximity to high-volume road or highway
Proximity to railroad
Walkability
Houses built pre-1980
Proximity to green space
Proximity to lead mines
Proximity to coal mines
Proximity to EPA risk management plan site
Proximity to EPA Treatment, Storage,and Disposal site
Proximity to EPA Toxic Release Inventory site
Proximity to EPA National Priority List site
Probability of contracting cancer
Ambient concentrations of diesel
Days above PM2.5 standard
Days above O3 standard
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3

# Who should act?

- **Proximity to green spaces**
  - Might reflect a higher environmental awareness or value placed on nature in these areas
  - Might improve mental health overall
- **Lack of internet access**
  - Could be due to reduced access to information, awareness, and mobilization tools
  - Could be attributed to a lack of other basic needs, which might take priority over climate change action

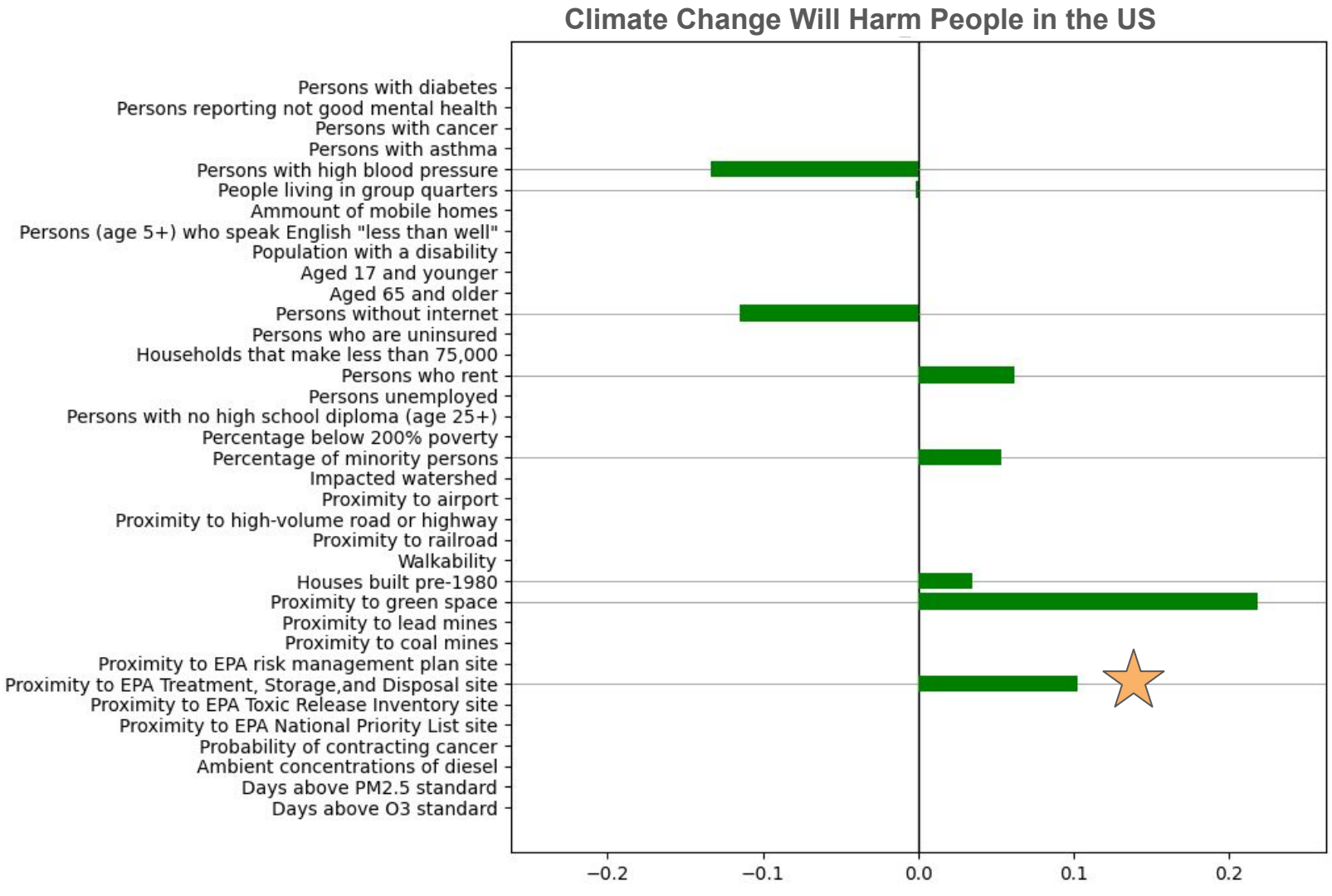
# Who is harmed?

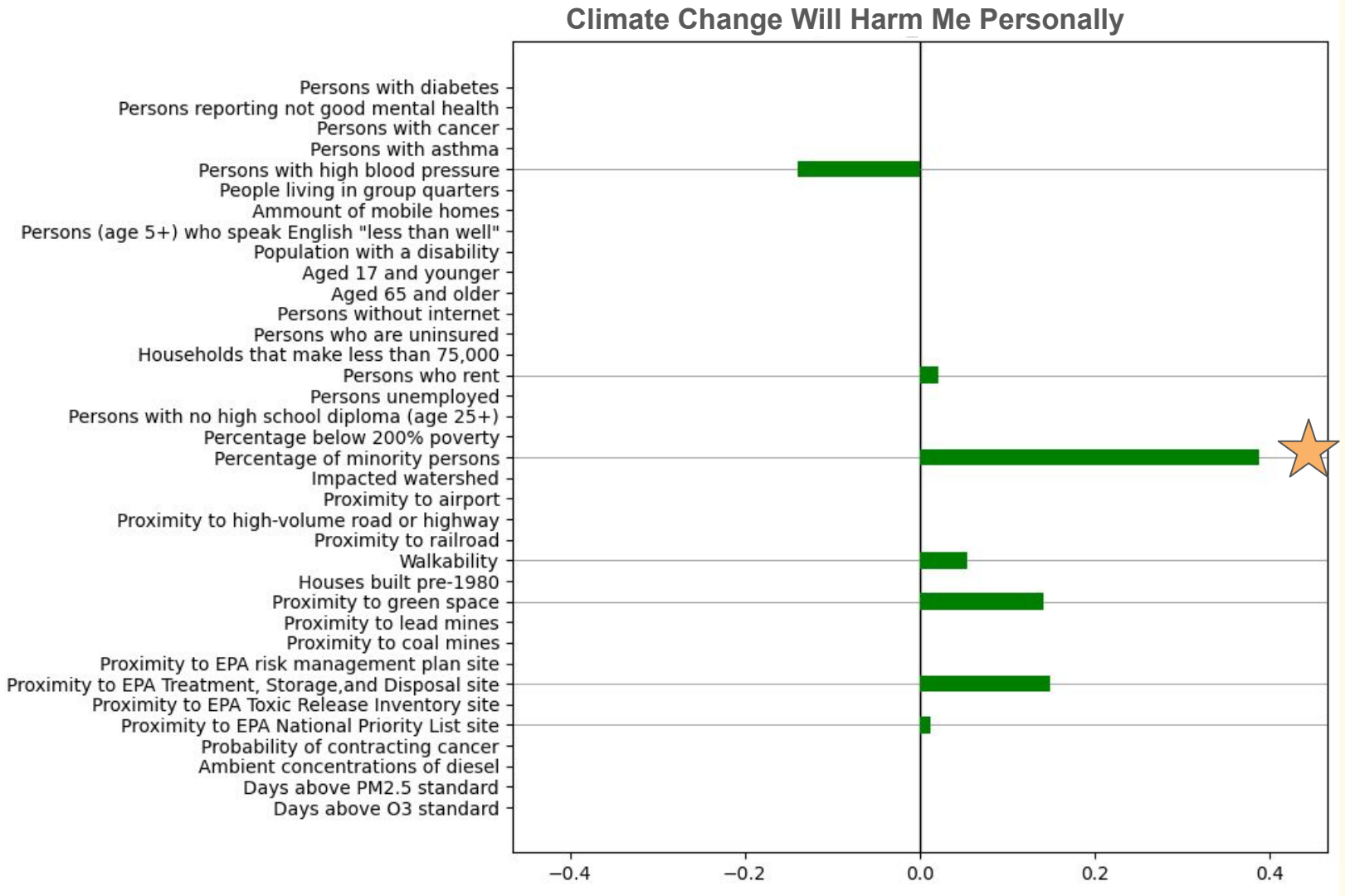
- **Futuregen_combined:**
  - Estimated percentage who think global warming will harm future generations
- **Harmus_combined:**
  - Estimated percentage who think global warming will harm people in the US
- **Personal_combined:**
  - Estimated percentage who think global warming will harm them personally a moderate amount/a great deal

| Future Gen | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.265506 |
| | Persons without internet | -0.153935 |

| Harm US | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.218492 |
| | Persons with high blood pressure | -0.133470 |

| Personal | Feature | Coefficient |
|---|---|---|
| | Percentage of minority persons | 0.388241 |
| | Persons with high blood pressure | -0.139409 |


Climate Change Will Harm People in the US
Persons with diabetes
Persons reporting not good mental health
Persons with cancer
Persons with asthma
Persons with high blood pressure
People living in group quarters
Ammount of mobile homes
Persons (age 5+) who speak English "less than well"
Population with a disability
Aged 17 and younger
Aged 65 and older
Persons without internet
Persons who are uninsured
Households that make less than 75,000
Persons who rent
Persons unemployed
Persons with no high school diploma (age 25+)
Percentage below 200% poverty
Percentage of minority persons
Impacted watershed
Proximity to airport
Proximity to high-volume road or highway
Proximity to railroad
Walkability
Houses built pre-1980
Proximity to green space
Proximity to lead mines
Proximity to coal mines
Proximity to EPA risk management plan site
Proximity to EPA Treatment, Storage,and Disposal site
Proximity to EPA Toxic Release Inventory site
Proximity to EPA National Priority List site
Probability of contracting cancer
Ambient concentrations of diesel
Days above PM2.5 standard
Days above O3 standard
−0.2
−0.1
0.0
0.1
0.2


Climate Change Will Harm Me Personally
Persons with diabetes
Persons reporting not good mental health
Persons with cancer
Persons with asthma
Persons with high blood pressure
People living in group quarters
Ammount of mobile homes
Persons (age 5+) who speak English "less than well"
Population with a disability
Aged 17 and younger
Aged 65 and older
Persons without internet
Persons who are uninsured
Households that make less than 75,000
Persons who rent
Persons unemployed
Persons with no high school diploma (age 25+)
Percentage below 200% poverty
Percentage of minority persons
Impacted watershed
Proximity to airport
Proximity to high-volume road or highway
Proximity to railroad
Walkability
Houses built pre-1980
Proximity to green space
Proximity to lead mines
Proximity to coal mines
Proximity to EPA risk management plan site
Proximity to EPA Treatment, Storage,and Disposal site
Proximity to EPA Toxic Release Inventory site
Proximity to EPA National Priority List site
Probability of contracting cancer
Ambient concentrations of diesel
Days above PM2.5 standard
Days above O3 standard
−0.4
−0.2
0.0
0.2
0.4

# Who is harmed?

- **High blood pressure**
  - People with poor health may deprioritize global or future environmental risks in favor of personal health concerns
- **Minority populations**
  - Might be more likely to believe in personal harm from global warming, possibly due to increased vulnerability to environmental risks or cultural connections to the earth

# Key Takeaways:

**People's immediate surroundings (proximity to green space and internet access) and personal health challenges heavily impact climate change sentiment in the US.**

- **Proximity to Green Space Increases Climate Awareness**
    - People living near green spaces are consistently more likely to believe in the reality of global warming, feel they have experienced its effects, discuss it more, and expect greater action from citizens, Congress, and corporations to address climate change.
- **Limited Internet Access Reduces Climate Engagement**
    - Areas with higher percentages of people lacking internet access show reduced belief in global warming, less personal concern, and lower expectations for action from both individuals and institutions, likely due to limited access to information on climate issues.
- **Health Issues Lower Climate Concern**
    - Populations with higher rates of health problems, such as mental health challenges, high blood pressure, and diabetes, tend to engage less with climate change, likely because immediate personal health concerns take priority over broader environmental worries.

# Data Cleaning Process:

- Merged data by county
- The EJI data was by census tract so I calculated a weighted average for each county
- Weighted by population
- Could have lost neighborhood-level variation if compared by county

```
# aggregate census tract data into county-level data by calculating weighted averages f

#Calculate the weighted average for each theme using E_TOTPOP (population variable) as
def weighted_avg(group, avg_name, weight_name):
    """Calculate weighted average."""
    d = group[avg_name]
    w = group[weight_name]
    return (d * w).sum() / w.sum()

# Group by COUNTY and apply weighted average or sum where appropriate
eji_clean_grouped = eji_clean.groupby('COUNTY').apply(
    lambda x: pd.Series({
        'E_TOTPOP': x['E_TOTPOP'].sum(),
        'E_OZONE': weighted_avg(x, 'E_OZONE', 'E_TOTPOP'),
        'E_PM': weighted_avg(x, 'E_PM', 'E_TOTPOP'),
        'E_DSLPM': weighted_avg(x, 'E_DSLPM', 'E_TOTPOP'),
        'E_TOTCR': weighted_avg(x, 'E_TOTCR', 'E_TOTPOP'),
        'E_NPL': weighted_avg(x, 'E_NPL', 'E_TOTPOP'),
        'E_TRI': weighted_avg(x, 'E_TRI', 'E_TOTPOP'),
        'E_TSD': weighted_avg(x, 'E_TSD', 'E_TOTPOP'),
        'E_RMP': weighted_avg(x, 'E_RMP', 'E_TOTPOP'),
        'E_COAL': weighted_avg(x, 'E_COAL', 'E_TOTPOP'),
        'E_LEAD': weighted_avg(x, 'E_LEAD', 'E_TOTPOP'),
        'E_PARK': weighted_avg(x, 'E_PARK', 'E_TOTPOP'),
        'E_HOUAGE': weighted_avg(x, 'E_HOUAGE', 'E_TOTPOP'),
        'E_WLKIND': weighted_avg(x, 'E_WLKIND', 'E_TOTPOP'),
        'E_RAIL': weighted_avg(x, 'E_RAIL', 'E_TOTPOP'),
        'E_ROAD': weighted_avg(x, 'E_ROAD', 'E_TOTPOP'),
        'E_AIRPRT': weighted_avg(x, 'E_AIRPRT', 'E_TOTPOP'),
        'E_IMPWTR': weighted_avg(x, 'E_IMPWTR', 'E_TOTPOP'),
        'EP_MINRTY': weighted_avg(x, 'EP_MINRTY', 'E_TOTPOP'),
        'EP_POV200': weighted_avg(x, 'EP_POV200', 'E_TOTPOP'),
        'EP_NOHSDP': weighted_avg(x, 'EP_NOHSDP', 'E_TOTPOP'),
        'EP_UNEMP': weighted_avg(x, 'EP_UNEMP', 'E_TOTPOP'),
        'EP_RENTER': weighted_avg(x, 'EP_RENTER', 'E_TOTPOP'),
        'EP_HOUBDN': weighted_avg(x, 'EP_HOUBDN', 'E_TOTPOP'),
        'EP_UNINSUR': weighted_avg(x, 'EP_UNINSUR', 'E_TOTPOP'),
        'EP_NOINT': weighted_avg(x, 'EP_NOINT', 'E_TOTPOP'),
        'EP_AGE65': weighted_avg(x, 'EP_AGE65', 'E_TOTPOP'),
        'EP_AGE17': weighted_avg(x, 'EP_AGE17', 'E_TOTPOP'),
        'EP_DISABL': weighted_avg(x, 'EP_DISABL', 'E_TOTPOP'),
        'EP_LIMENG': weighted_avg(x, 'EP_LIMENG', 'E_TOTPOP'),
        'EP_MOBILE': weighted_avg(x, 'EP_MOBILE', 'E_TOTPOP'),
        'EP_GROUPQ': weighted_avg(x, 'EP_GROUPQ', 'E_TOTPOP'),
        'EP_BPHIGH': weighted_avg(x, 'EP_BPHIGH', 'E_TOTPOP'),
        'EP_ASTHMA': weighted_avg(x, 'EP_ASTHMA', 'E_TOTPOP'),
        'EP_CANCER': weighted_avg(x, 'EP_CANCER', 'E_TOTPOP'),
        'EP_MHLTH': weighted_avg(x, 'EP_MHLTH', 'E_TOTPOP'),
        'EP_DIABETES': weighted_avg(x, 'EP_DIABETES', 'E_TOTPOP')
    })
).reset_index()

# Add the GEOID column, ensuring it matches the county
eji_clean_grouped['GEOID'] = [eji_clean[eji_clean['COUNTY'] == county]['GEOID'].iloc[0]

#save back into eji_clean
eji_clean = eji_clean_grouped
```

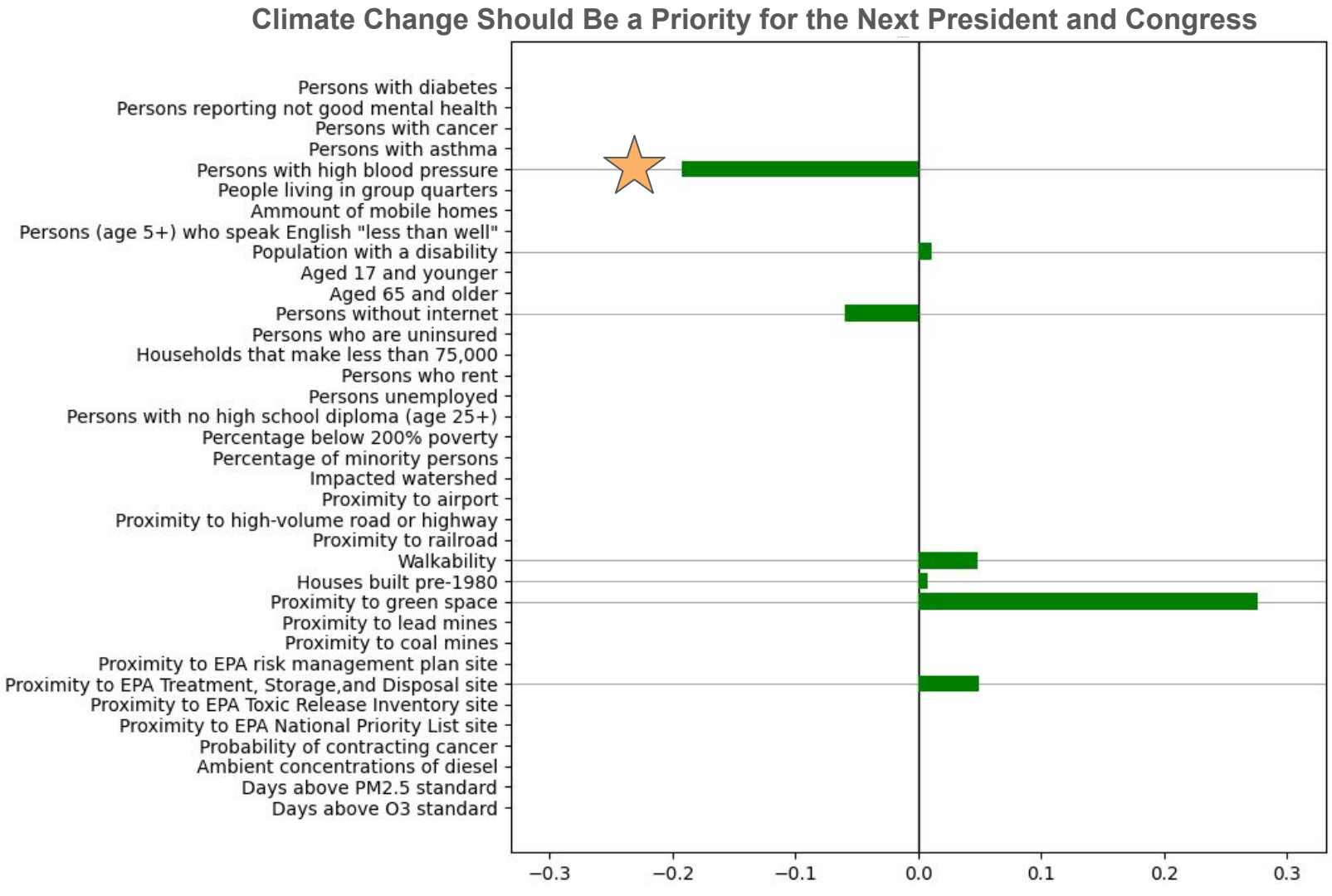
# Voting/ legislation:

- **Gwvoteimp_combined:**
  - Estimated percentage who say a candidate's views on GW are important to their vote
- **Priority_combined:**
  - Estimated percentage who say global warming should be a high priority for the next president and Congress
- **Regulate_combined:**
  - Estimated percentage who somewhat/strongly support regulating CO2 as a pollutant

| Gwvoteimp | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.328951 |
| | Persons without internet | -0.096976 |

| Priority | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.276327 |
| | Persons with high blood pressure | -0.192180 |

| Regulate | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.173203 |
| | Persons without internet | -0.388985 |


Climate Change Should Be a Priority for the Next President and Congress
Persons with diabetes
Persons reporting not good mental health
Persons with cancer
Persons with asthma
Persons with high blood pressure
People living in group quarters
Ammount of mobile homes
Persons (age 5+) who speak English "less than well"
Population with a disability
Aged 17 and younger
Aged 65 and older
Persons without internet
Persons who are uninsured
Households that make less than 75,000
Persons who rent
Persons unemployed
Persons with no high school diploma (age 25+)
Percentage below 200% poverty
Percentage of minority persons
Impacted watershed
Proximity to airport
Proximity to high-volume road or highway
Proximity to railroad
Walkability
Houses built pre-1980
Proximity to green space
Proximity to lead mines
Proximity to coal mines
Proximity to EPA risk management plan site
Proximity to EPA Treatment, Storage,and Disposal site
Proximity to EPA Toxic Release Inventory site
Proximity to EPA National Priority List site
Probability of contracting cancer
Ambient concentrations of diesel
Days above PM2.5 standard
Days above O3 standard
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3

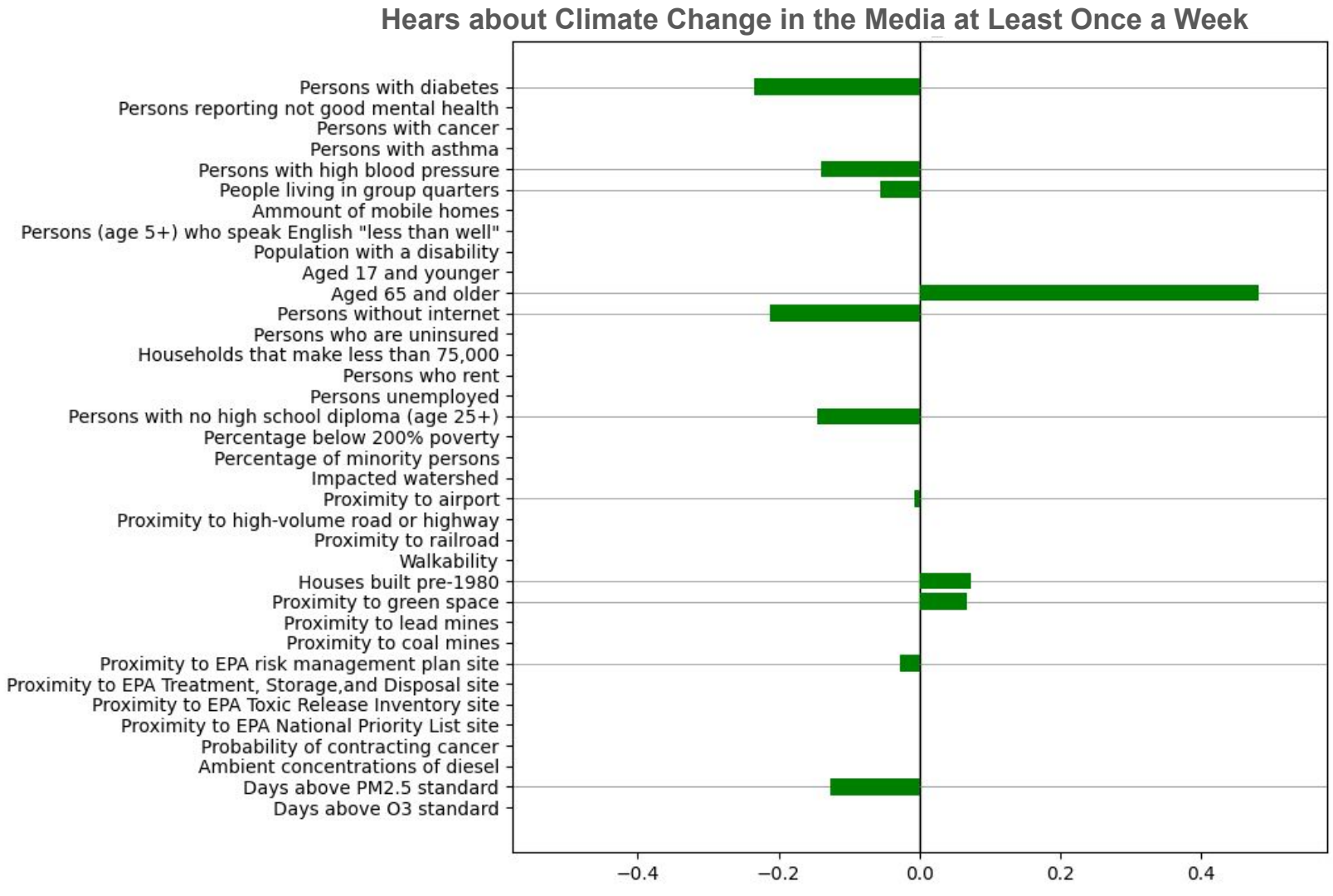
# Climate change discussion:

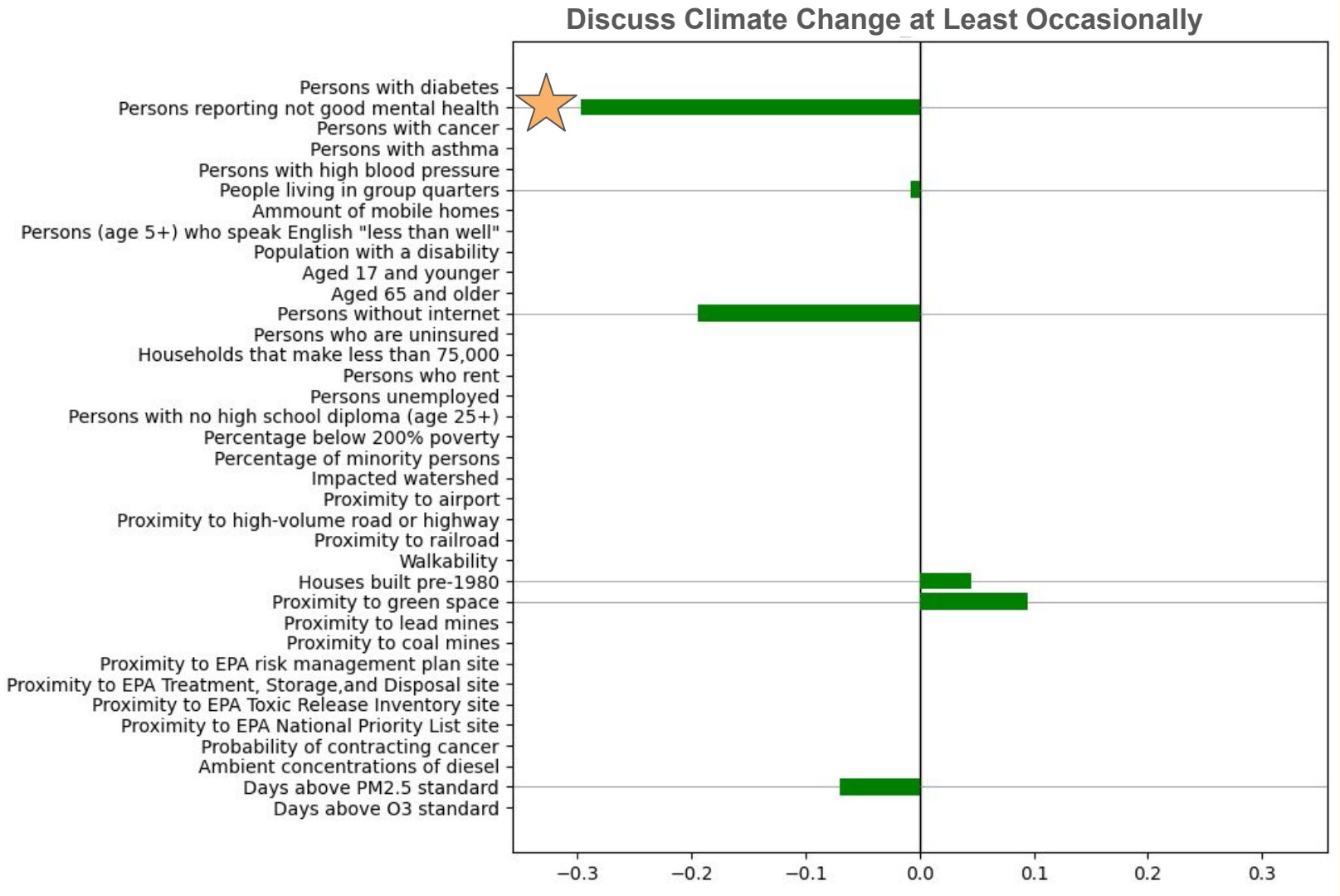
- **Discuss_combined:**
  - Estimated percentage who discuss global warming with friends and family
- **Mediaweekly_combined:**
  - Estimated percentage who hear about global warming in the media
- **Human_combined:**
  - Estimated percentage who think that global warming is caused mostly by human activities

| Discuss | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.094367 |
| | Persons with not good mental health | -0.297540 |

| Mediaweekly | Feature | Coefficient |
|---|---|---|
| | Aged 65 and older | 0.481906 |
| | Persons with diabetes | -0.235376 |

| Human | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.314892 |
| | Persons with asthma | -0.098549 |


Hears about Climate Change in the Media at Least Once a Week
Persons with diabetes
Persons reporting not good mental health
Persons with cancer
Persons with asthma
Persons with high blood pressure
People living in group quarters
Ammount of mobile homes
Persons (age 5+) who speak English "less than well"
Population with a disability
Aged 17 and younger
Aged 65 and older
Persons without internet
Persons who are uninsured
Households that make less than 75,000
Persons who rent
Persons unemployed
Persons with no high school diploma (age 25+)
Percentage below 200% poverty
Percentage of minority persons
Impacted watershed
Proximity to airport
Proximity to high-volume road or highway
Proximity to railroad
Walkability
Houses built pre-1980
Proximity to green space
Proximity to lead mines
Proximity to coal mines
Proximity to EPA risk management plan site
Proximity to EPA Treatment, Storage,and Disposal site
Proximity to EPA Toxic Release Inventory site
Proximity to EPA National Priority List site
Probability of contracting cancer
Ambient concentrations of diesel
Days above PM2.5 standard
Days above O3 standard
−0.4
−0.2
0.0
0.2
0.4


Discuss Climate Change at Least Occasionally
Persons with diabetes
Persons reporting not good mental health
Persons with cancer
Persons with asthma
Persons with high blood pressure
People living in group quarters
Ammount of mobile homes
Persons (age 5+) who speak English "less than well"
Population with a disability
Aged 17 and younger
Aged 65 and older
Persons without internet
Persons who are uninsured
Households that make less than 75,000
Persons who rent
Persons unemployed
Persons with no high school diploma (age 25+)
Percentage below 200% poverty
Percentage of minority persons
Impacted watershed
Proximity to airport
Proximity to high-volume road or highway
Proximity to railroad
Walkability
Houses built pre-1980
Proximity to green space
Proximity to lead mines
Proximity to coal mines
Proximity to EPA risk management plan site
Proximity to EPA Treatment, Storage,and Disposal site
Proximity to EPA Toxic Release Inventory site
Proximity to EPA National Priority List site
Probability of contracting cancer
Ambient concentrations of diesel
Days above PM2.5 standard
Days above O3 standard
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3

# Climate change discussion:

- **Older adults (aged 65 and older)**
    - Could be caused by higher traditional media consumption rates and exposure to climate-related content

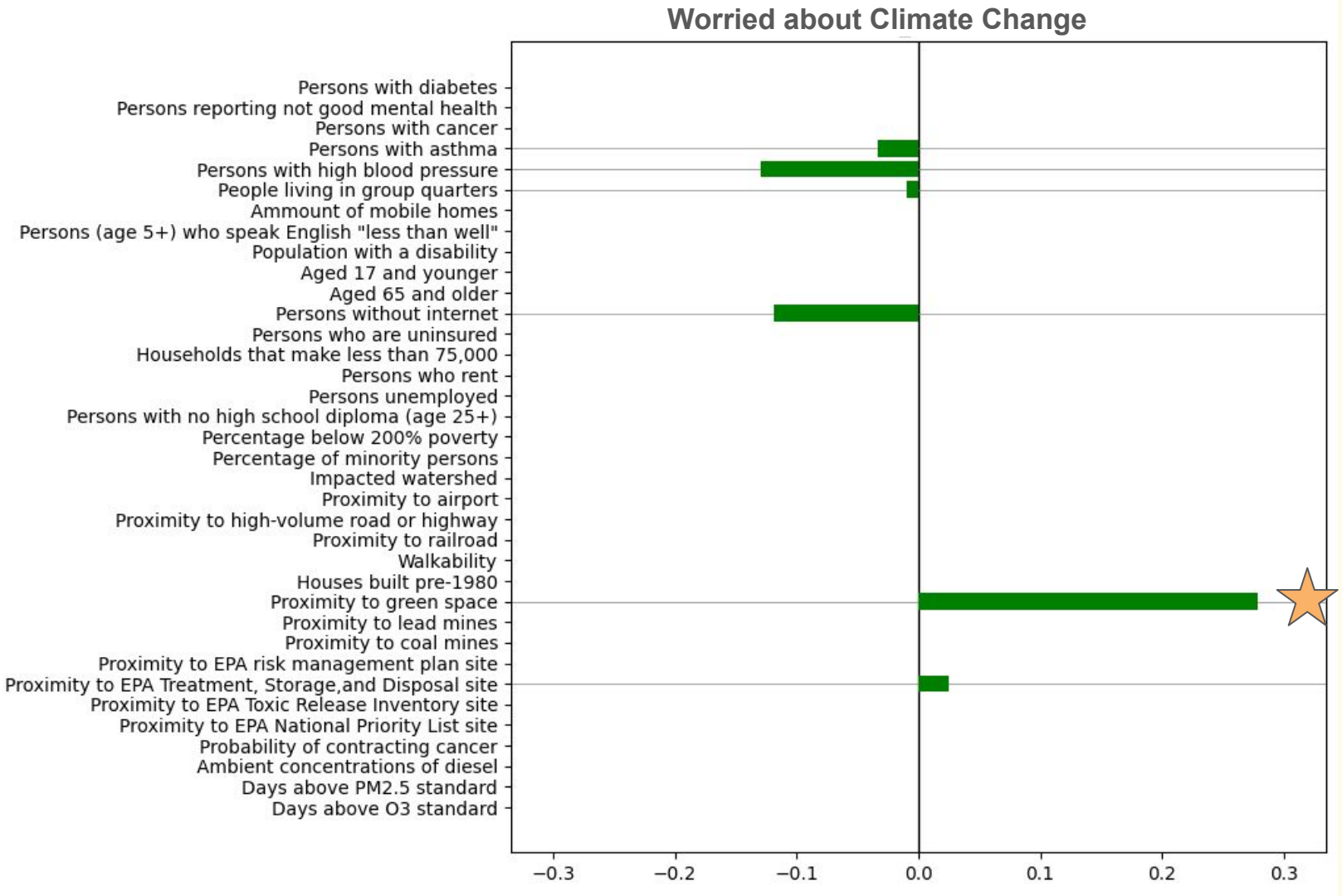
# Climate change concern:

- **Exp_combined:**
  - Estimated percentage who responded that they have personally experienced the effects of global warming
- **Happening_combined:**
  - Estimated percentage who think that global warming is happening
- **Worried_combined:**
  - Estimated percentage who are not very/not at all worried about global warming

| Exp | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.241683 |
| | Persons without internet | -0.192961 |

| Happening | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.259983 |
| | Persons without internet | -0.220575 |

| Worried | Feature | Coefficient |
|---|---|---|
| | Proximity to green space | 0.278782 |
| | Persons with high blood pressure | -0.129625 |


Worried about Climate Change
Persons with diabetes
Persons reporting not good mental health
Persons with cancer
Persons with asthma
Persons with high blood pressure
People living in group quarters
Ammount of mobile homes
Persons (age 5+) who speak English "less than well"
Population with a disability
Aged 17 and younger
Aged 65 and older
Persons without internet
Persons who are uninsured
Households that make less than 75,000
Persons who rent
Persons unemployed
Persons with no high school diploma (age 25+)
Percentage below 200% poverty
Percentage of minority persons
Impacted watershed
Proximity to airport
Proximity to high-volume road or highway
Proximity to railroad
Walkability
Houses built pre-1980
Proximity to green space
Proximity to lead mines
Proximity to coal mines
Proximity to EPA risk management plan site
Proximity to EPA Treatment, Storage,and Disposal site
Proximity to EPA Toxic Release Inventory site
Proximity to EPA National Priority List site
Probability of contracting cancer
Ambient concentrations of diesel
Days above PM2.5 standard
Days above O3 standard
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3