Project 4:
Global Warming Survey

Andy Deemer
Larry Jackelen
Saba Suhail
Vilo Avila

Problem Statement Background

Our nation is divided on whether or not global warming is happening; all while the planet burns.

The Paris Climate Agreement was met with global participation; however the withdrawal of the United States did not receive national condemnation, further showing this division.

The United States is now back in the Paris Climate Agreement, but it is still not politically costly to withdraw in the future. The country's view needs to change.

Problem Statement

We are meeting with key members of the staff of John Kerry - Special Presidential Envoy for Climate - to discuss the characteristics of the nation in the context of their belief in global warming

Leveraging their internal marketing and public service announcement capabilities, they want to know the types of people they should investigate further for reaching out to.

A major boon will be to convince persons that don't believe global warming is happening otherwise.

What types of characteristics and features have the largest effect on whether or not a person believes global warming is happening?

Approach

We set out to solve this problem through:

1. Data Collection through Survey Results

2. Data Exploration to identify key features and generate new features

3. Data cleaning of Survey Answers through ordinal encoding and dummification

4. Performing Feature Selection to refine our dataset before modeling

5. Capturing signal through Principal Component Analysis (PCA) of features not selected

6. Training multiple classification models

7. Interpreting results, both through model accuracy and feature importance

Data Collection

Survey

Climate change in the American Mind: National survey Data on Public Opinion (2008-2018)

The data is based on 19 waves of nationally representative surveys of US adults aged 18 and older. The surveys were conducted once in 2008 and then twice a year from 2010 to 2018. All questionnaires were self-administered by respondents in a web-based environment.

Samples were drawn from the "Ipsos Knowledge Panel", an online panel of members drawn using probability sampling methods. Prospective members were recruited using a combination of random digit dial and address based sampling techniques that cover virtually all resident phone numbers and addresses in the United States. The sample includes a representative cross-section of American adults.

The final dataset consisted of 22,416 data points.

A note on waves and years

As mentioned, these findings are a jumping off point–they are not generalizable directly to the population as a whole due to it being historical and not adjusted to census statistics.

However, these findings and models can point to where investigations, population mining, and outreach should start using the survey population as a test set.
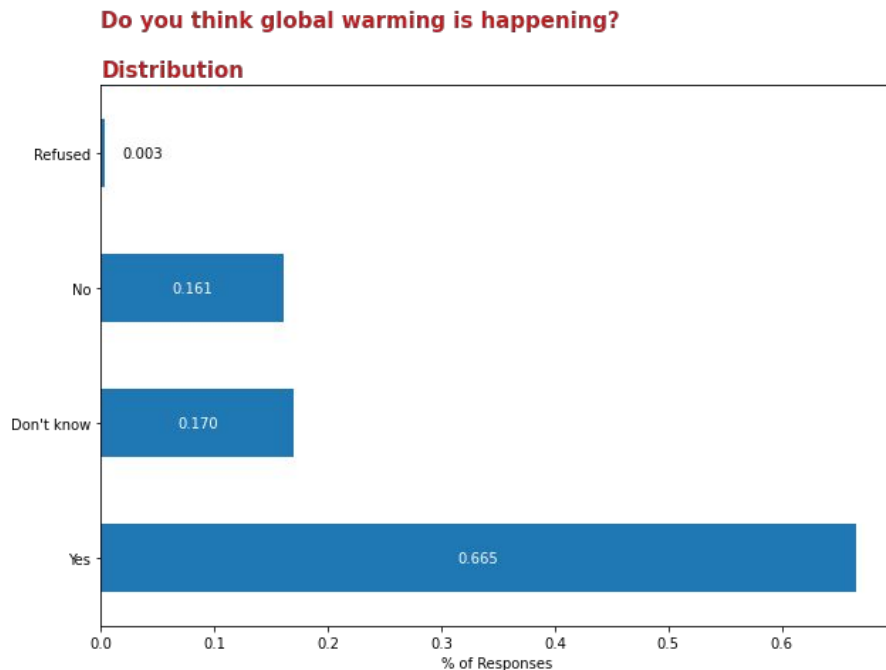
Data Exploration

Target Variable

"Refused" class made up 0.3% of dataset. Since this was our target variable, they were dropped.
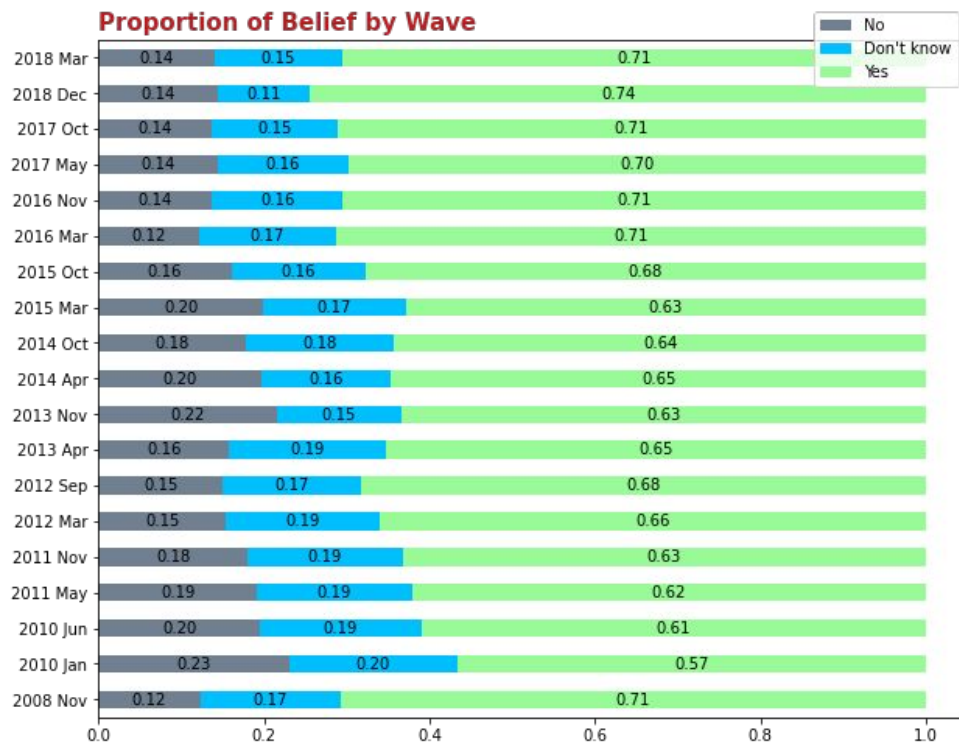
Baseline: 67%

Do you think global warming is happening?

Distribution

Response	%
Refused	0.003
No	0.161
Don't know	0.170
Yes	0.665

% of Responses

An interesting graph about our survey collection years

- 2008: 2,100 participants
- All subsequent waves average 1,100

- Trend of increasing proportion of "believers"

Goal: "Close the gap" from believers and non-believers. Those that don't know are most likely to be educated and persuaded into believing.

Proportion of Belief by Wave

Legend: No, Don't know, Yes

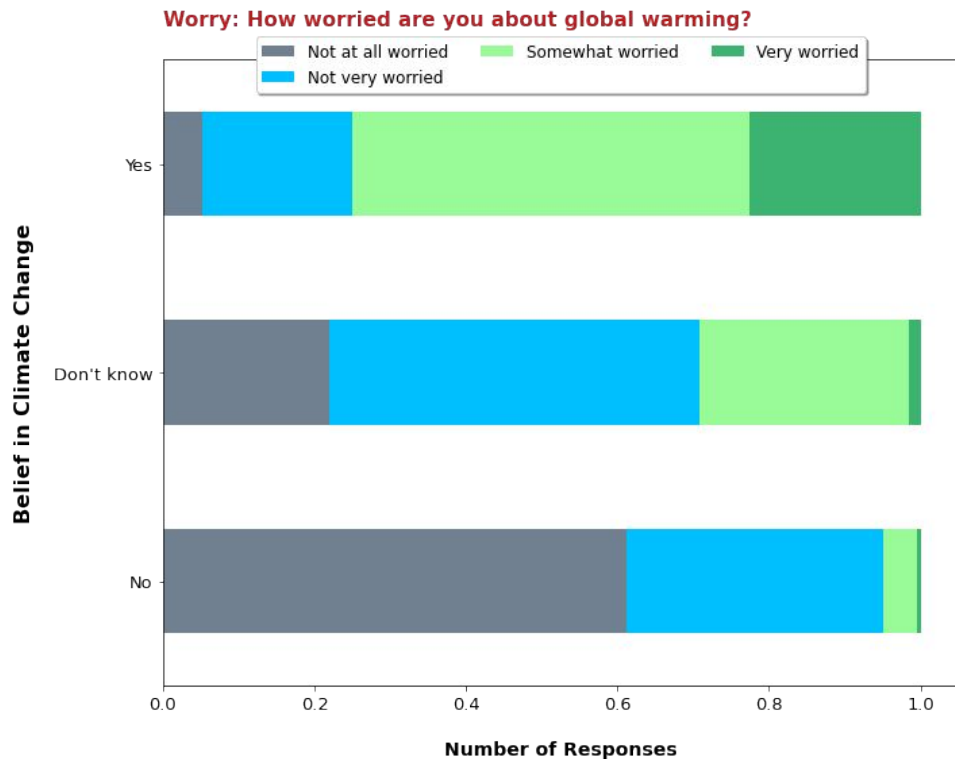| Wave	No	Don't know	Yes
2018 Mar	0.14	0.15	0.71
2018 Dec	0.14	0.11	0.74
2017 Oct	0.14	0.15	0.71
2017 May	0.14	0.16	0.70
2016 Nov	0.14	0.16	0.71
2016 Mar	0.12	0.17	0.71
2015 Oct	0.16	0.16	0.68
2015 Mar	0.20	0.17	0.63
2014 Oct	0.18	0.18	0.64
2014 Apr	0.20	0.16	0.65
2013 Nov	0.22	0.15	0.63
2013 Apr	0.16	0.19	0.65
2012 Sep	0.15	0.17	0.68
2012 Mar	0.15	0.19	0.66
2011 Nov	0.18	0.19	0.63
2011 May	0.19	0.19	0.62
2010 Jun	0.20	0.19	0.61
2010 Jan	0.23	0.20	0.57
2008 Nov	0.12	0.17	0.71

Most People have some level of worry

- Inconsistency in the level of worry between the believers and non-believers

- 40% of non believers are at least somewhat worried about climate change.

- 80% of people that don't know if climate change is happening are worried, suggesting they are likely to be persuaded.

THEME: when asked subsequent questions, the underlying feelings of individuals are revealed, often being inconsistent with the target question.
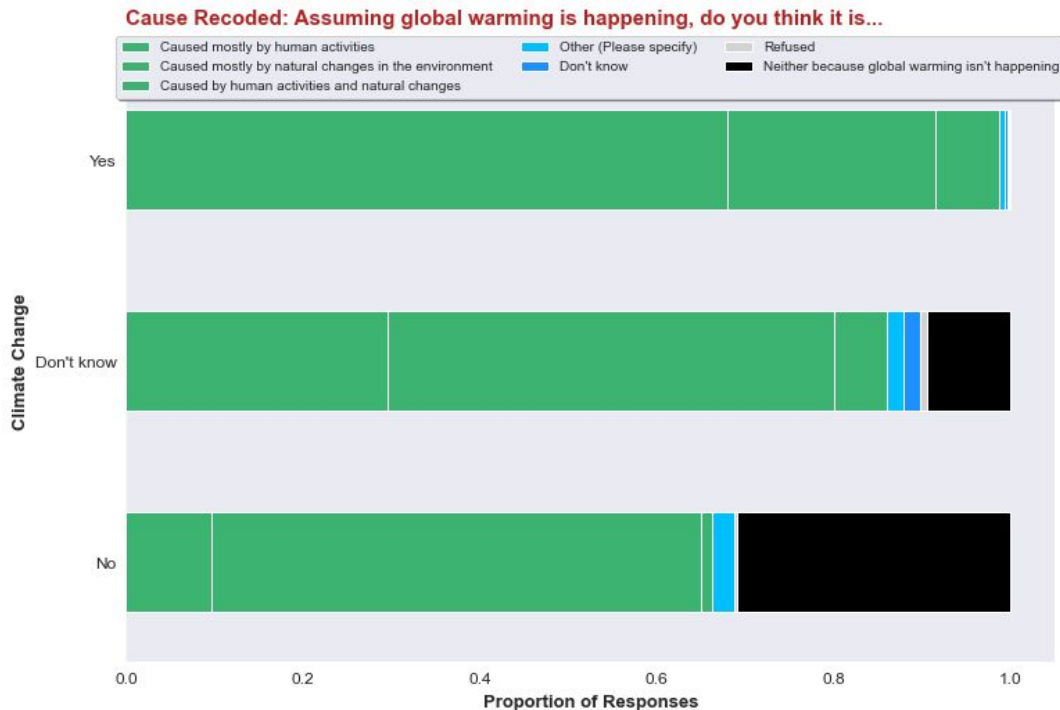
Recommend: capitalize on people's fear. They should be afraid. Discuss recent catastrophic events caused directly by climate change.

Worry: How worried are you about global warming?

Legend: Not at all worried, Somewhat worried, Very worried, Not very worried

Y-axis: **Belief in Climate Change** (Yes, Don't know, No)

X-axis: **Number of Responses** (0.0, 0.2, 0.4, 0.6, 0.8, 1.0)

Cause of Global Warming

- 2 Initial Questions
  - Cause of global warming
  - If 'other' specify
- -> cause_recoded
  - Surveyors deduced actual sentiments from these two initial question
  - Resulting in 7 final subgroups
- 3 of the answer subgroups suggest at least some belief in global warming.
- 85% of people who "don't know" answer 1 of these 3 subgroups
- 65% if non-believers did as well

Recommend: Clarify the distinction of weather global change is happening with weather it's caused by humans, naturally, or both. 99% of scientists agree that global warming is solely by humans

Cause Recoded: Assuming global warming is happening, do you think it is...

Legend:
- Caused mostly by human activities
- Caused mostly by natural changes in the environment
- Caused by human activities and natural changes
- Other (Please specify)
- Don't know
- Refused
- Neither because global warming isn't happening

Y-axis: Climate Change (Yes, Don't know, No)
X-axis: Proportion of Responses (0.0, 0.2, 0.4, 0.6, 0.8, 1.0)

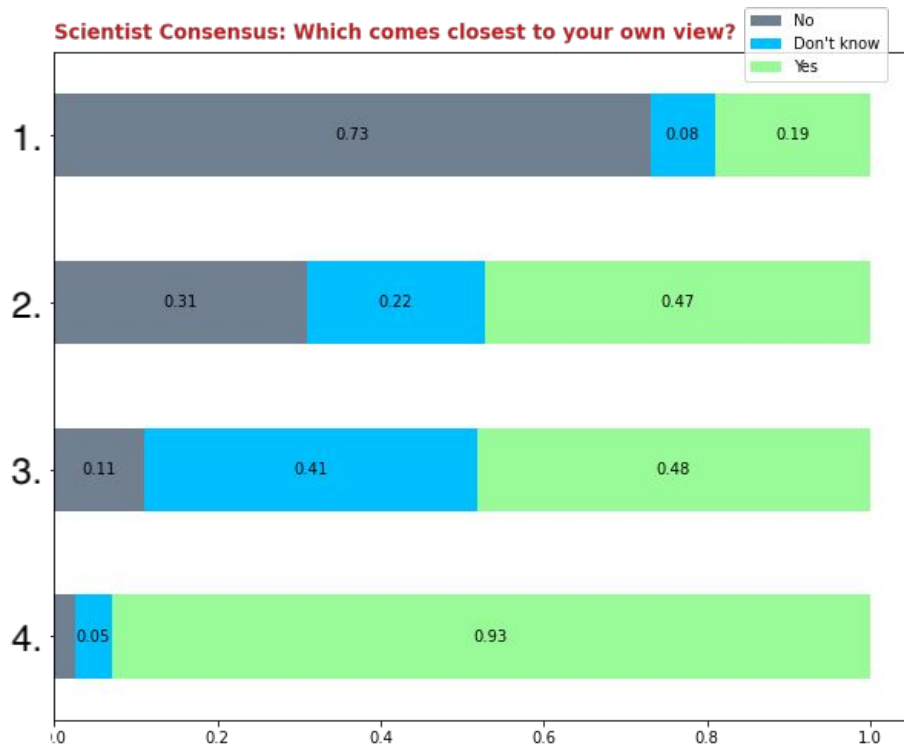# Clarify Actual Consensus in Scientific Community

Multiple choice answers:
1. Most scientists think global warming is not happening
2. There is a lot of disagreement among scientists about whether or not global warming is happening
3. Don't know enough to say
4. Most scientists think global warming is happening

- 50% of people who believe there is disagreement are on the fence or don't believe in climate change.

Recommend: Clarify true consensus in scientific community. This theory is no longer debated among the international community of scientists. Cite sources.

[More than 99.9% of studies agree: Humans caused climate change](#)

Scientist Consensus: Which comes closest to your own view?

Legend: No / Don't know / Yes

1. 0.73 | 0.08 | 0.19
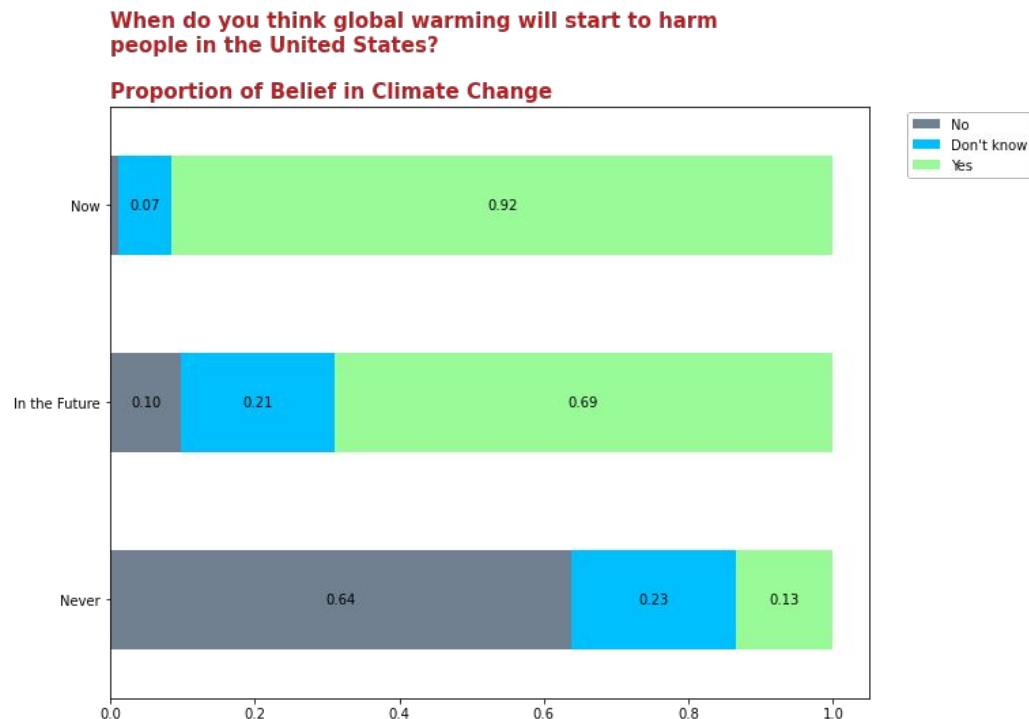2. 0.31 | 0.22 | 0.47
3. 0.11 | 0.41 | 0.48
4. 0.05 | 0.93

Harming people in the United States?

Procrastination for action is as harmful as inaction for global warming. It is happening now.

69% of respondents who mentioned that global warming is not **yet** harming people in the United States *believe* it is happening.

This is a great opportunity to push the urgency to act on people who already agree it is happening but perhaps not in their backyard.

When do you think global warming will start to harm people in the United States?

Proportion of Belief in Climate Change

Legend:
- No
- Don't know
- Yes

Now: 0.07, 0.92
In the Future: 0.10, 0.21, 0.69
Never: 0.64, 0.23, 0.13

The Harm Series

The survey makers included as part of their collection a series of five questions asked one after another about the **harm** of global warming.

These harms were related to different aspects. How much will global warming harm:

- You personally
- People in the United States
- People in developing countries
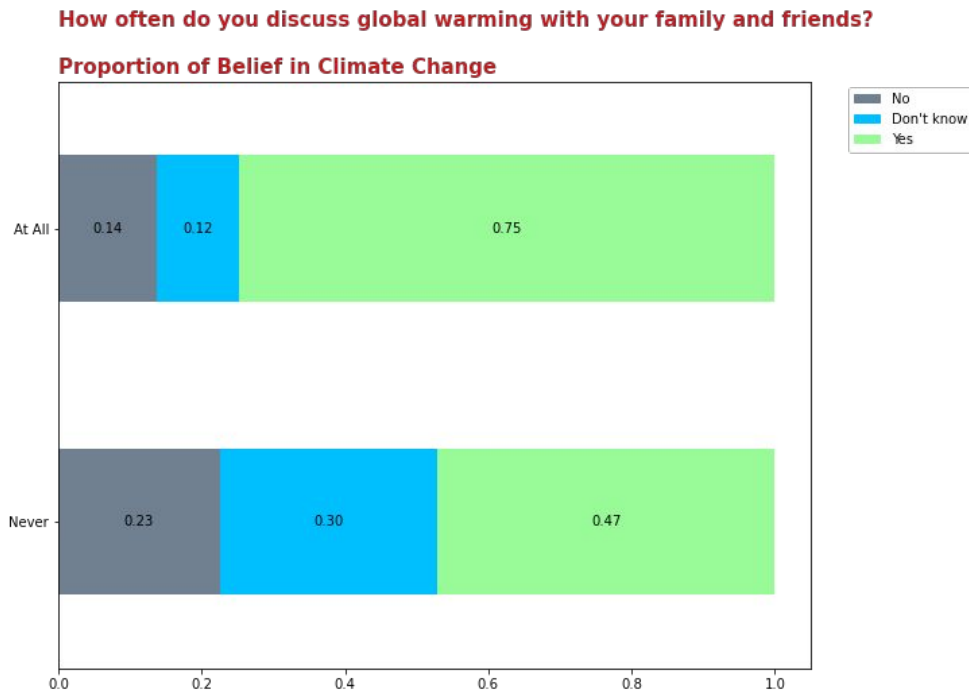- Future generations of People
- Plant and animal species

Each of these questions received above the baseline amount of agreement that global warming is happening. This shows some insight into how the survey population that does not think global warming is happening now, still has segments that expect it to happen…and for it to harm.

Discussion of Global Warming

72% of respondents indicated that they discuss global warming with family and friends.

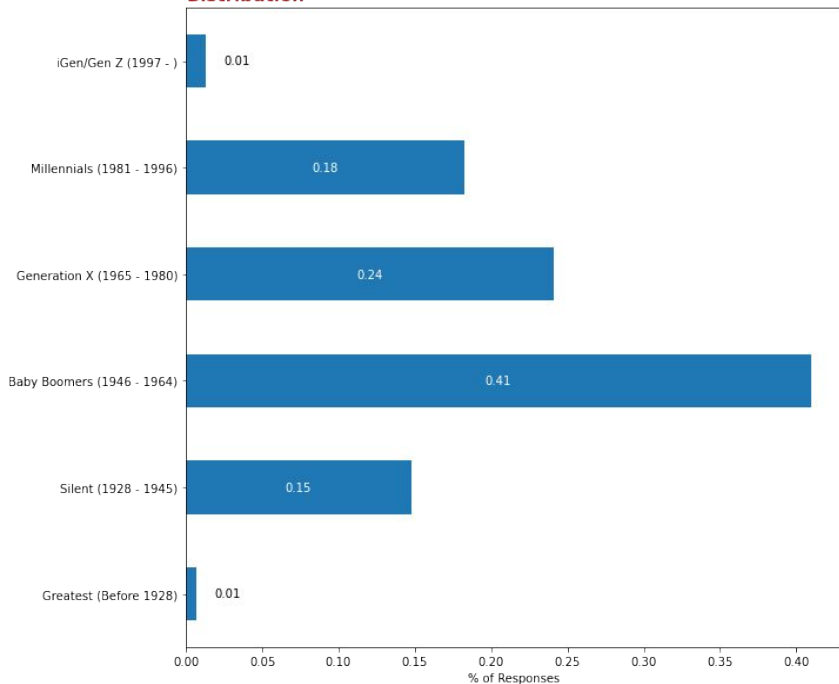These respondents believe in global warming at a **75%** rate!

Discussion is key to driving understanding and interest in a topic as critical and nuanced as global warming.
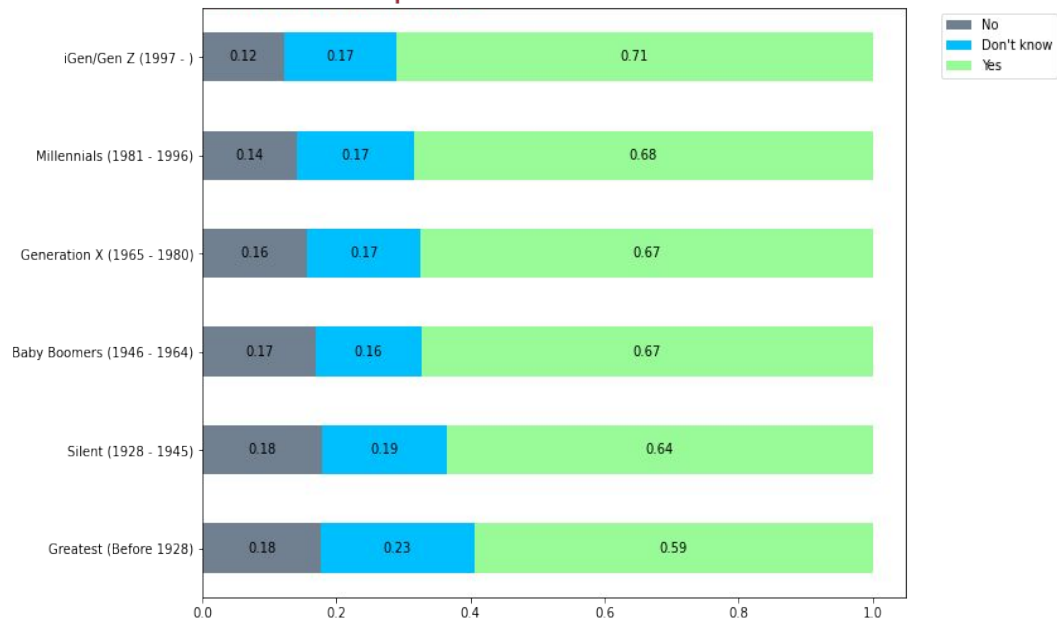

How often do you discuss global warming with your family and friends?

Proportion of Belief in Climate Change

Age Demographics
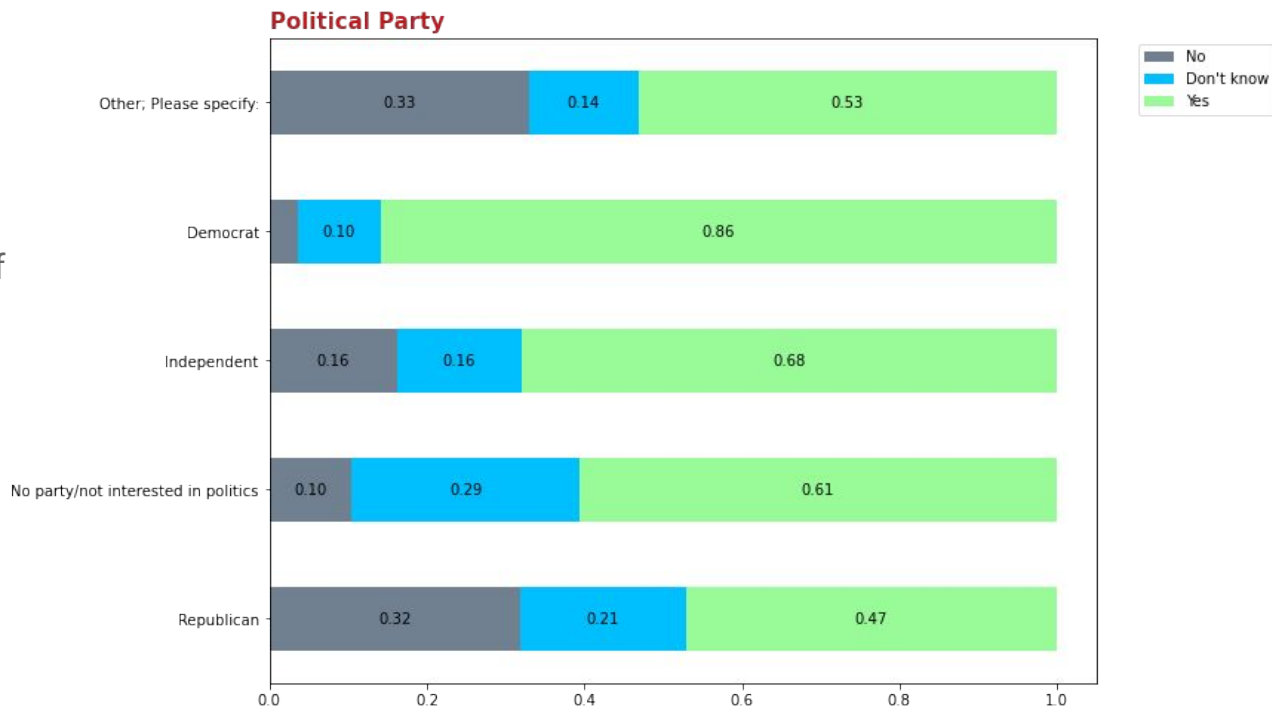
Education by Category

- The more educated, the higher the belief in climate change
- Those people with only high school degree or lower have more of a prevalence to not believe in climate change
- There is quite a jump in belief in climate change from having a high school degree or lower, to having a bachelor's degree
- We could start teaching climate change at a young age to reinforce the topic by the time they even reach high school age

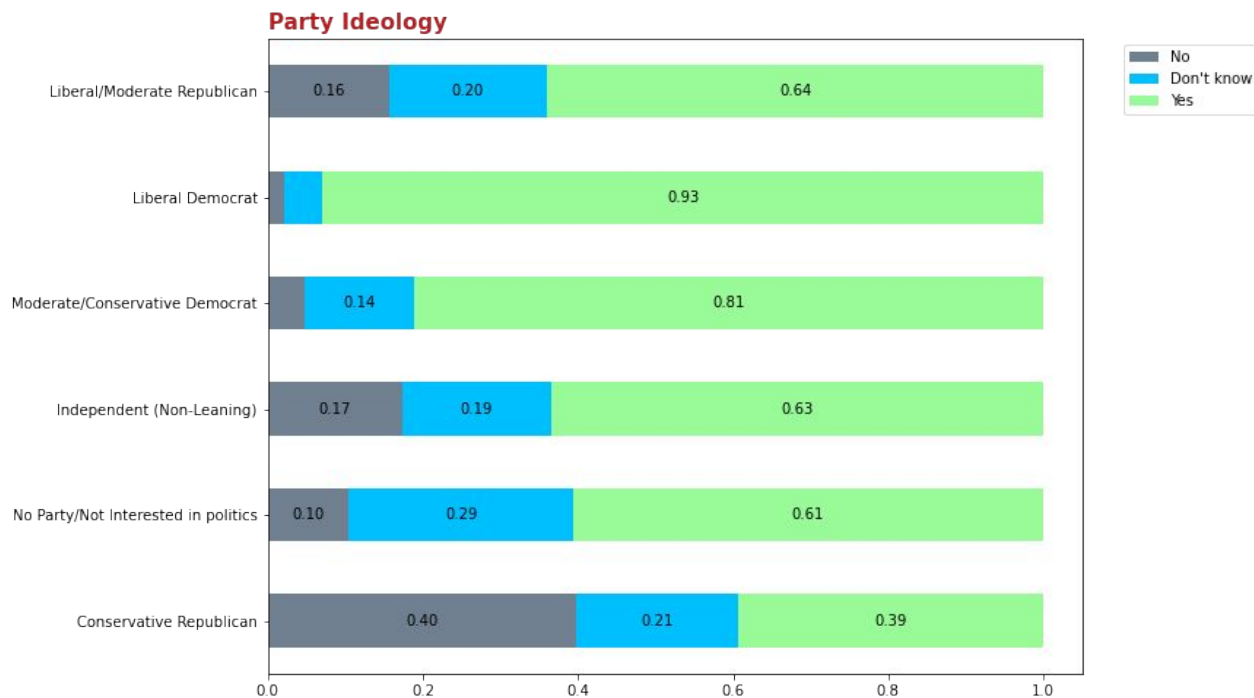
Education Category of Respondent
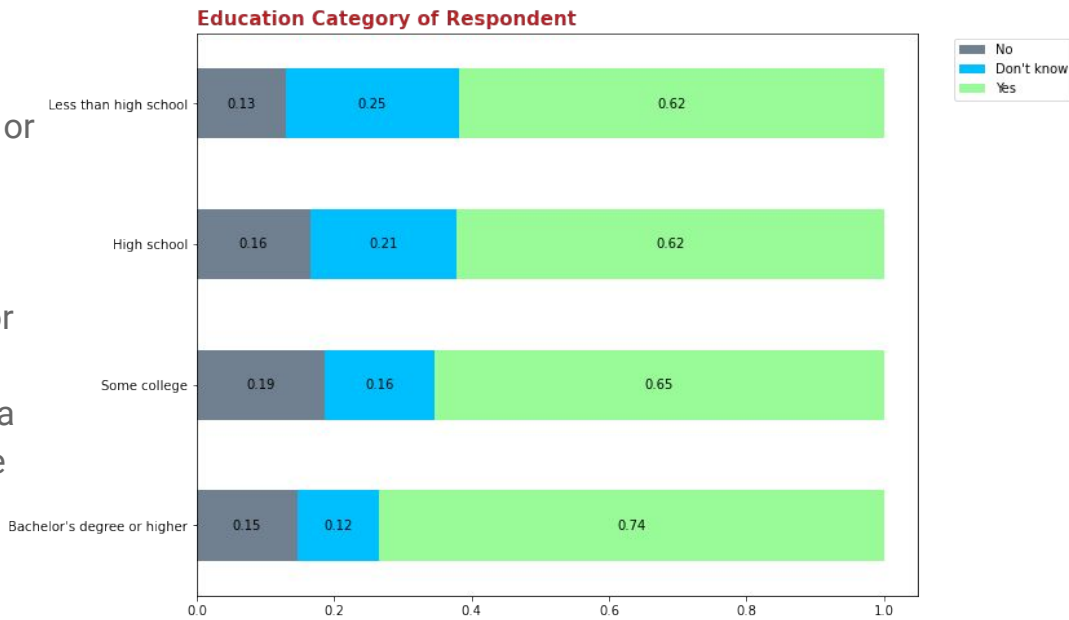
Political Affiliation

- Belief in climate change can be greatly affected by political views
- Democrats have the strongest belief in climate change
- Republicans have the least belief in climate change
- I would say it would be a good idea to target the republican party but we all know how touchy politics can be. The people who associate with no party, other, and independent may be better targets to change in opinion

Political Party

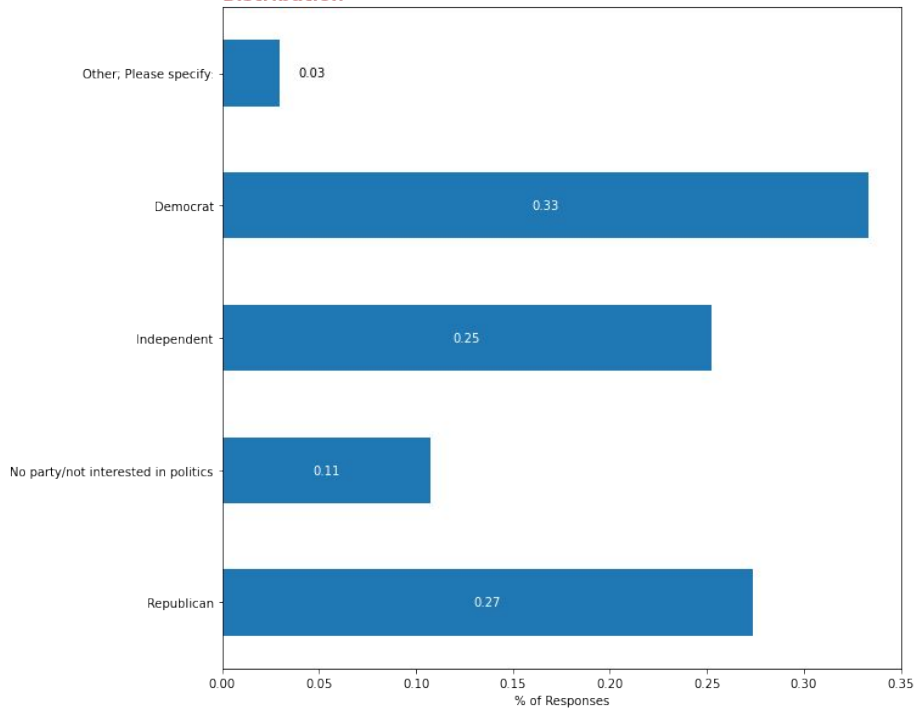|	No	Don't know	Yes
Other; Please specify:	0.33	0.14	0.53
Democrat		0.10	0.86
Independent	0.16	0.16	0.68
No party/not interested in politics	0.10	0.29	0.61
Republican	0.32	0.21	0.47

Political Ideology

- We see similar results with ideology within the respective political party with a much more pronounced extremes between liberal democrats and conservative republicans.
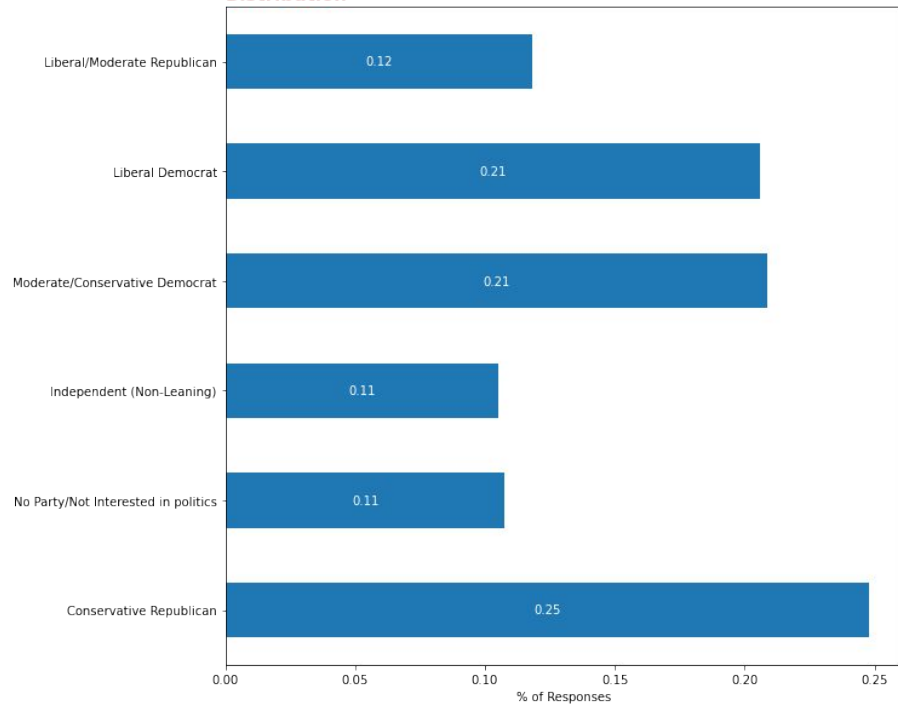- Again, it might be wise to target those who may be on the fence with their political views.

Party Ideology

Legend: No | Don't know | Yes
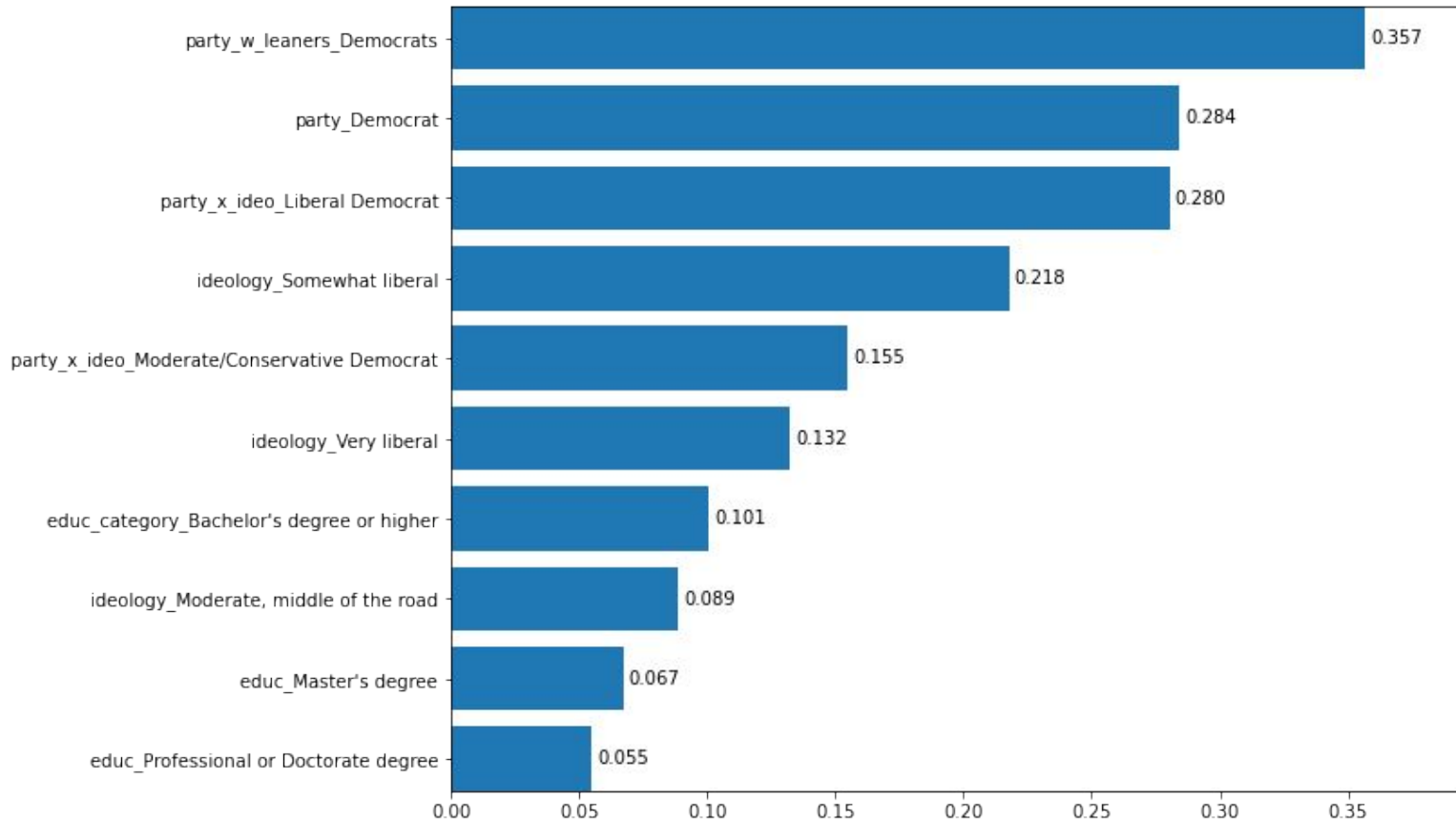
Category	No	Don't know	Yes
Liberal/Moderate Republican	0.16	0.20	0.64
Liberal Democrat			0.93
Moderate/Conservative Democrat		0.14	0.81
Independent (Non-Leaning)	0.17	0.19	0.63
No Party/Not Interested in politics	0.10	0.29	0.61
Conservative Republican	0.40	0.21	0.39

Political Distribution

Political Party Distribution

Category	% of Responses
Other, Please specify	0.03
Democrat	0.33
Independent	0.25
No party/not interested in politics	0.11
Republican	0.27

Political Ideology Distribution

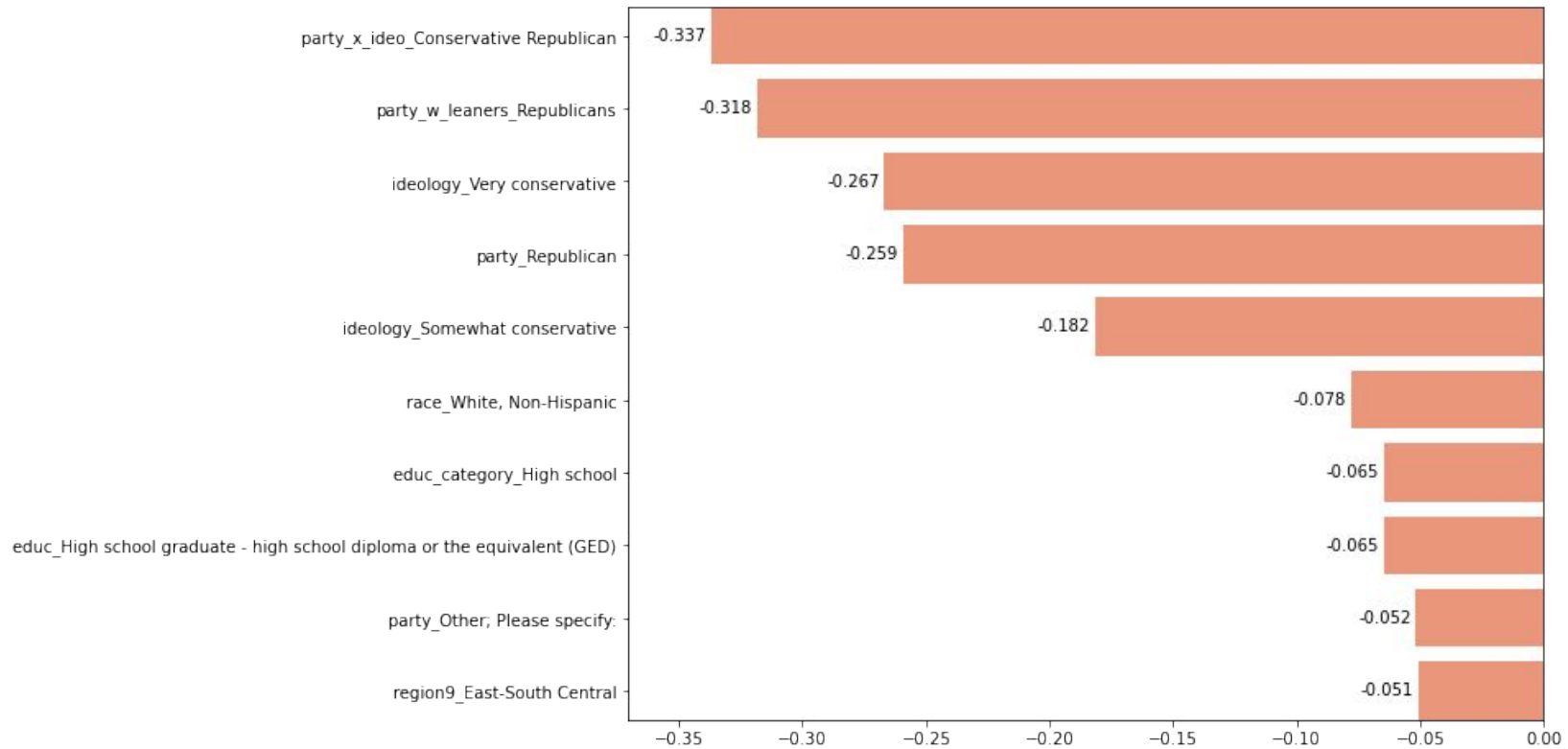| Category	% of Responses
Liberal/Moderate Republican	0.12
Liberal Democrat	0.21
Moderate/Conservative Democrat	0.21
Independent (Non-Leaning)	0.11
No Party/Not Interested in politics	0.11
Conservative Republican	0.25

Positive Feature Correlation to answering 'Yes'

Negative Feature Correlation to answering 'Yes'

Who knew that children were a big indicator? Here's a graph

It ain't.

Created a new variable children from the existing data.

Expected a positive correlation between having children and having a stake in earth's future.

Possibly people who believe in climate change have fewer or no children at all.

Household's number of children

Proportion of Belief in Climate Change

|	No	Don't know	Yes
more than 1	0.18	0.18	0.65
1	0.16	0.17	0.68
0	0.16	0.17	0.67

How religiosity relates to climate awareness?

It matters a lot!

People who go for service attendance for more than once a week and once a week have a higher proportion of people refusing to believe in climate change compared to all other groups.

We can try incorporating climate change awareness in service attendance.Possibly hang climate change awareness posters in churches/mosques/temples

Respondent Service Attendance

**Proportion of Belief in Climate Change**
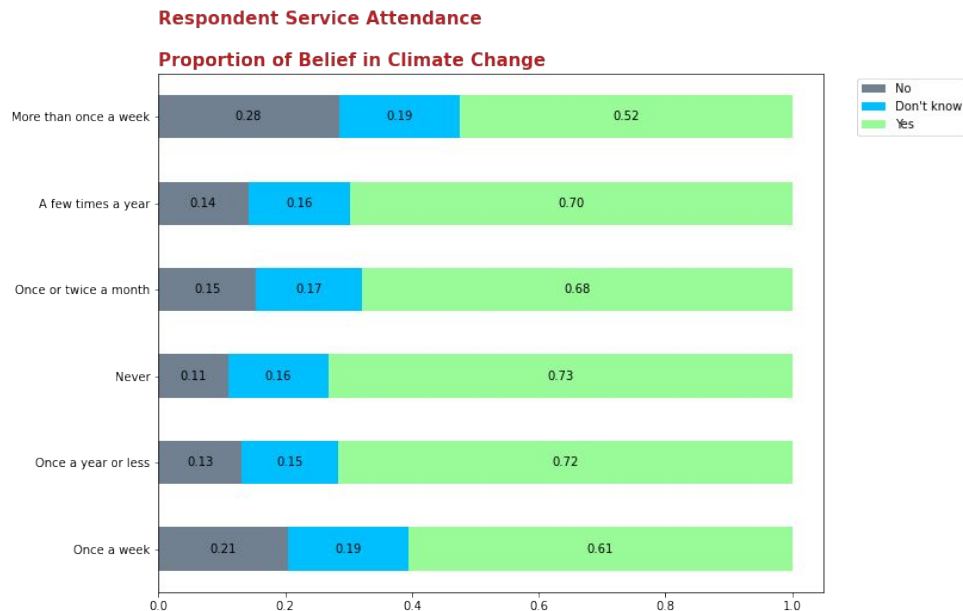
Employment Status:Generational Gap Visible

Retired people and self employed people have a slightly higher proportion of people not accepting climate change

Climate change awareness programme in Old age homes

Need to focus on these people to change thought process of coming generations too

Employment Status

Proportion of Belief in Climate Change
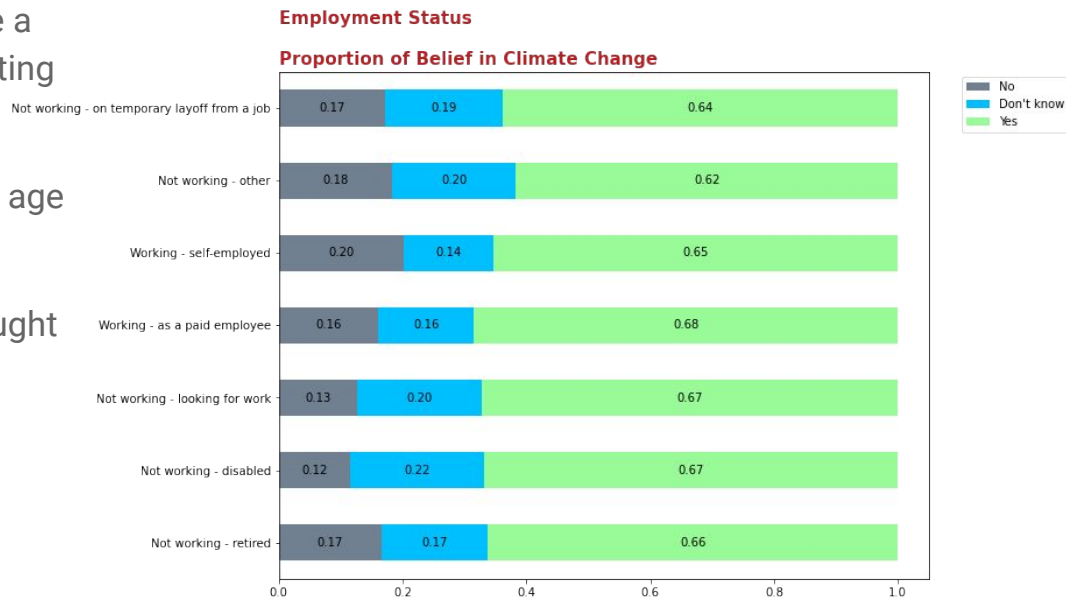
Data Cleaning

What to do with refused and nulls

The rows where people answered 5 or more responses with "refused" were treated as outliers and dropped. In total there were only 224 participants who refused to answer 5 or more questions and only accounted for 0.8% of the entire dataset.

The remaining refused answers were treated as null values and were iteratively imputed along with the rest of the null values for the dataset.

Feature Selection

- Survey had 24 multiple choice questions (nominal)
  - Several of these were reduced to binary OHE features
- Several features were used to manufacture features by binning subgroups. e.g. num_children
- Some features were "free-form" where participants were free to input any response in writing. These were helpful in analyzing the dataset, but ultimately disregarded when modeling.
- After EDA, feature selection and manufacturing, we used SelectKBest to select the top 20 features that capture most of the signal for our target.

PCA of non-selected features

- 60 features remained that were not chosen by SelectKBest
- PCA was used to reduce the non-chosen features down to 5 in order to capture some of the signal of the remaining features
- These 5 features were added to the 20 previously chose to give us our 25 features used for modeling

Modeling

Classifier Models

SVM

78% prediction accuracy

- 69% - No
- 66% - Don't know
- 81% - Yes
- Best model used the "rbf" kernel parameter which features cannot be extracted from

Logistic Regression
78% accurate in predicting class
Precision:
- 70% - No
- 52% - Don't know
- 82% - Yes

Top Features
- when_harm_US
- sci_consensus
- cause_recoded
- party_x_idea
- harm_plants_animals
- harm_personally
- worry

Random Forest
79% prediction accuracy
-74%-No
-55%-Don't Know
-83%-Yes
Top Features
- sci_consensus_ord
- when_harm_US_ord
- harm_US_ord
- harm_plants_animals_ord
- harm_future_gen_ord
- harm_dev_countries_ord
- harm_personally_ord

Conclusions and Findings

- For our best model we found that many of the questions that were asked relating to whether or not global warming may be harmful were most important for predicting our target. This holds true for our logistic regression model as it shared many of the same features
- This may come as no surprise but we found there to be a strong divide in belief in climate change based on political affiliation and even wider still depending on the ideology.
- Education level plays an important factor. The more educated a person is the more likely they are to believe in climate change.
- Age can be a factor however small. The older generations seem to not believe in climate change as much as the more recent ones.
- People who believe in climate change have fewer or no children at all.
- More religious people are more likely to not believe in climate change.
- This survey used a strategy of asking certain questions in a sequence in order to further pull out a participants actual underlying sentiment. This allows us and the surveyors to see how a person's sentiment is similar or different than the target class they're in.

Recommendations

- Recommend media that helps friends and family to learn about climate change more tangibly. It will make them worried. HBO should probably be telecasting "Our World" and "Before the Flood" more often.
- Speak to people in a language they understand- popular leaders,actors, famous scientists etc. Make climate awareness a compulsory subject in schools
- Highlight the heatwaves in US and Canada in social media
- Have climate change be taught in schools from an early age
- Convince people to change their political views (should be easy)
- Incorporate climate change programme in service attendance and old age homes

- Most people are worried about the effects of climate change, even if they say they don't believe in it or don't know. Based on this fact, we recommend focusing on showing the current natural disasters that are directly caused by global warming. These events directly impact the US population, more so in certain high risk areas like floodplains and hurricane prone environments.
- There's a large discrepancy in people's idea of why global warming is happening (by humans, natural, or both). Many people that don't believe in global warming or don't know still say it's caused by humans. The scientific community all believe it is caused by humans. We recommend addressing this discrepancy in lessons and public discussions.

Next Steps

The analysis can be further improved by:

More feature creation and consolidation, especially with the Harm series of questions

Imputing Nulls and Refused separately rather than together. This will capture signal separately for people not being served a question versus refusing to answer the question.

Leveraging the wave weight and weight aggregate in the analysis. This will allow the findings to be more generalizable toward the United States population as a whole, rather than just similar demographic populations

THANK YOU

Together, we can make a better tomorrow.

ILLUSTRATION BY FANATIC STUDIO/GETTY STUDIO