



# Climate Change Sentiment vs. Country CO<sub>2</sub> Emissions

In what ways do a given country's **CO<sub>2</sub> emissions** and **population's climate change opinion** reflect each other?



# Dataset

A custom dataset was produced from the following sources:

- **Climate Change Opinion Survey 2022** from Data For Good at Meta, accessed Feb 2025
- **Emissions by Country** from Zenodo via Kaggle, accessed Feb 2025

Emissions from 2021 have been compared against the climate change opinion survey in the year directly following (2022), and all emissions data should be considered to be from 2021. Countries for which either form of data was unavailable have been omitted.

# Emissions by country

The total CO2 emissions of a given country, in tons.

**Mean:** 155.73 tons

**Median:** 38.49 tons

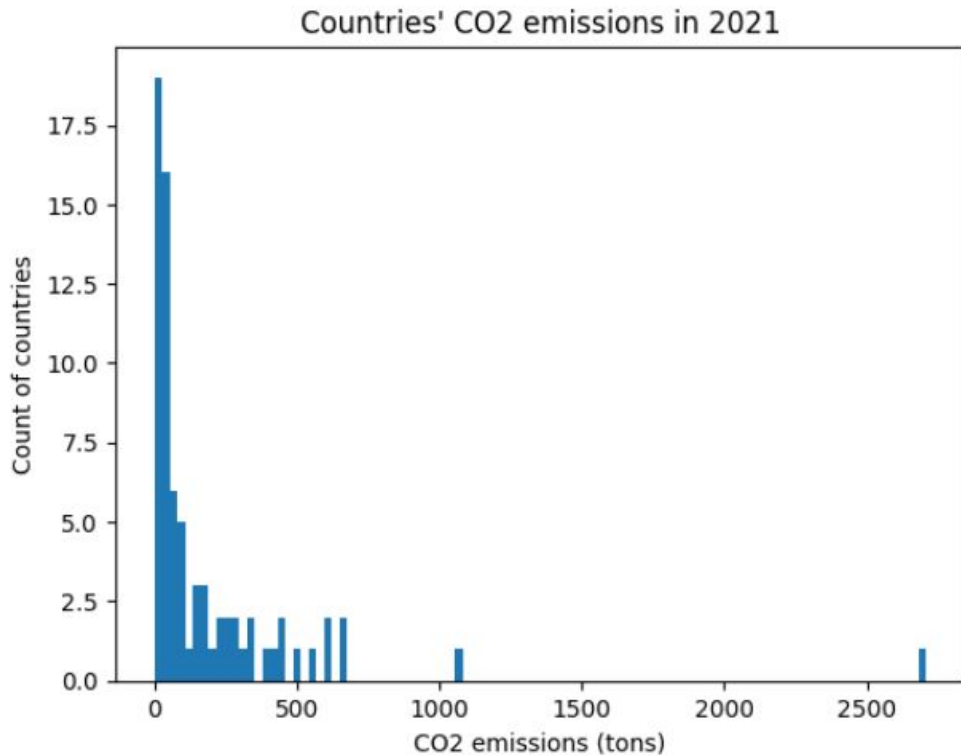
**Mode:** 0.0

**Spread:** Range of 2709.7

**Tails:** Steep positive tail away from mode of 0

There are significant high outliers in this data, which indicates that some countries produce vastly more CO2 emissions than others (>1000 tons). These should nevertheless be included as they will provide valuable comparative information.

It is important to note going forward that the *precise* mode is unlikely to be valuable information as precisely duplicated numbers in these datasets are unlikely. Therefore the mode will be derived from binned counts within the histogram.



# Belief in climate change

The percentage of a country's respondents who responded "Yes" to the question "Climate change refers to the idea that the world's average temperature has been increasing over the past 150 years, will increase more in the future, and that the world's climate will change as a result. What do you think: Do you think that climate change is happening?"

**Mean:** 85%

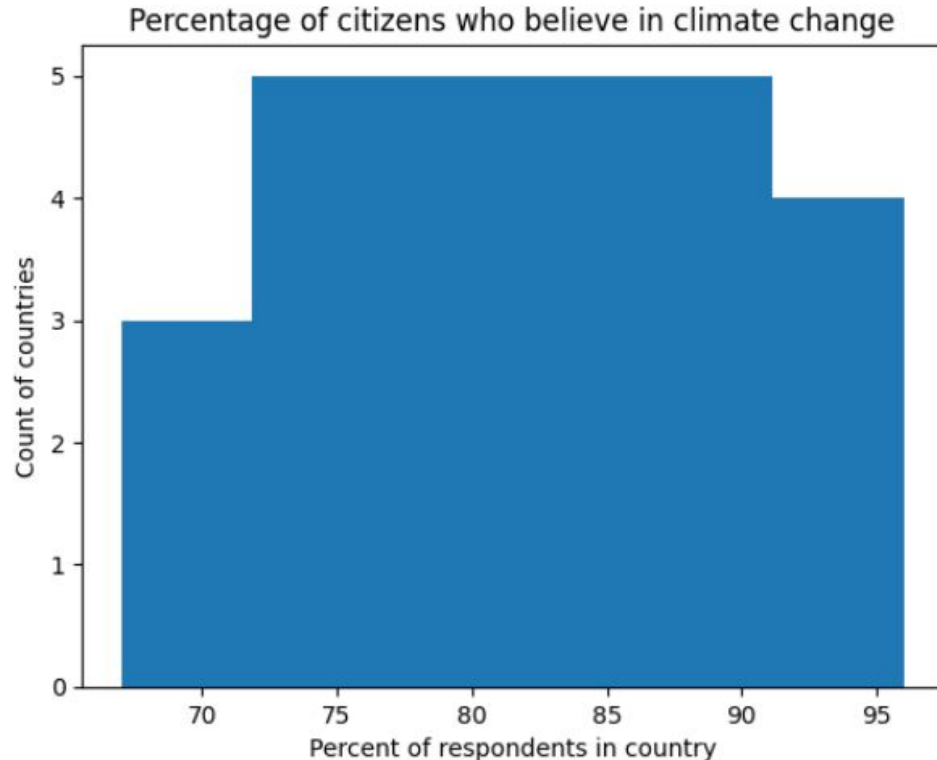
**Median:** 87%

**Mode:** wide range within bins, ~75% to ~90%, i.e. ~82.5%

**Spread:** Range of 28.6 percentage points

**Tails:** Values are mostly concentrated toward the mode

This (particularly the spread) indicates that there is not a lot of global variation in belief in climate change. The lowest belief rate was 67.4%, still a majority, with the highest at 96%.



# Education on climate change

The percentage of a country's respondents who responded "I know a moderate amount about it" or "I know a lot about it" to the question "How much do you know about climate change?"

**Mean:** 50%

**Median:** 44%

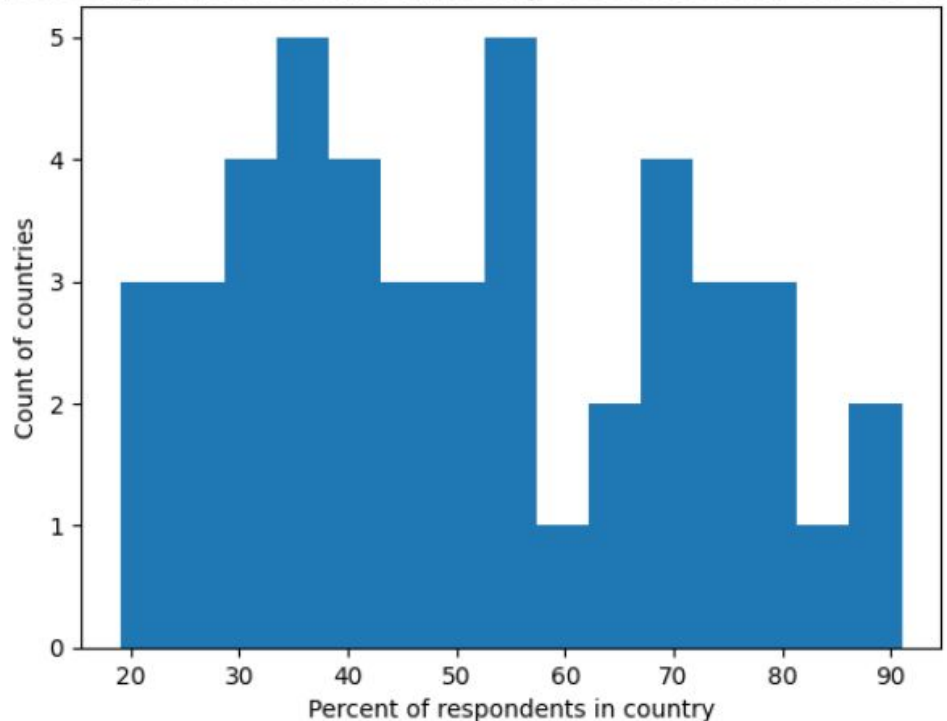
**Mode:** ~35% and ~55%, i.e. ~45%

**Spread:** Range of 72.4 percentage points

**Tails:** A slight positive tail, indicating that lower values (closer to the mode) are more common

This distribution indicates that there is a wide global spread of climate change knowledge (as self-reported). The breadth of percentages higher than the mode indicate that the higher percentages are likely pulling up the overall mean.

Percentage of citizens who think they are educated on climate change



# Sentiment on the importance of climate change

The percentage of a country's respondents who responded "Extremely important" or "Very important" to the question "How important is the issue of climate change to you personally?"

**Mean:** 60%

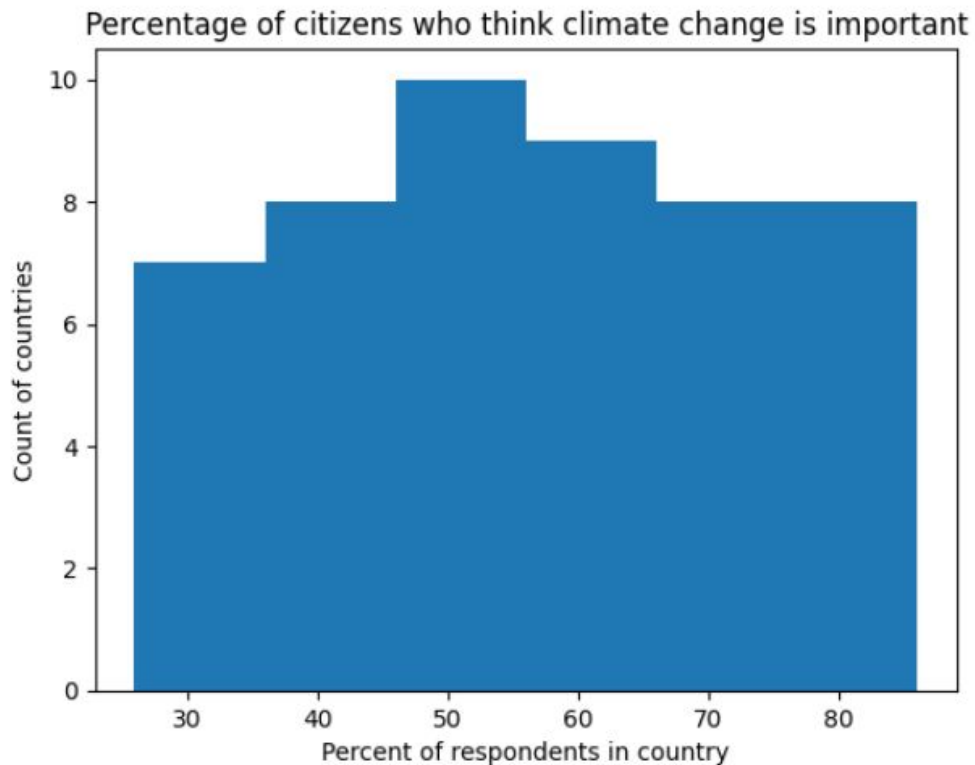
**Median:** 61%

**Mode:** ~50-55%, i.e. ~53%

**Spread:** Range of 59.9 percentage points

**Tails:** Approximately even distribution, skewed high

The mean and median are significantly higher than the mode, which is caused by the positive skew. In effect, it is more likely for the percentage to be higher than the mode than not. Still, there is a wide spread of percentages here, with a minimum of 26% and a maximum of 86%.



# Sentiment on climate change's threat

The percentage of a country's respondents who responded "Very serious threat" or "Somewhat serious threat" to the question "Do you think that climate change is a very serious threat, a somewhat serious threat, or not a threat at all to the people in the country (or territory) where you live, over the next 20 years?"

**Mean:** 78%

**Median:** 79%

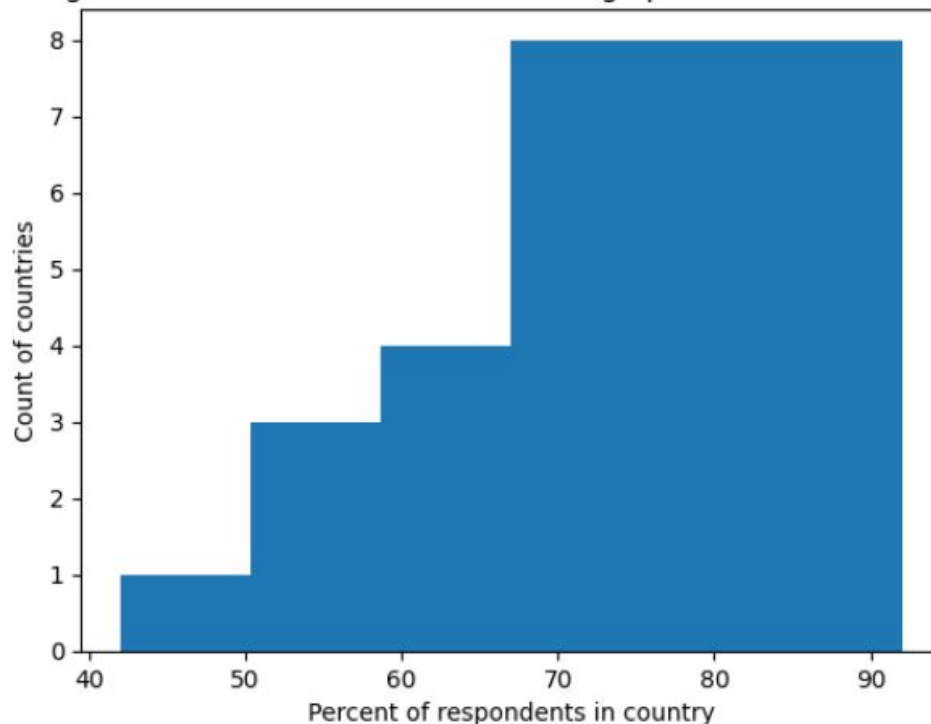
**Mode:** ~70-90%, i.e. ~80%

**Spread:** Range of 50.3 points

**Tails:** Steep negative tail from mode, indicating a higher likelihood of countries responding with a high level of concern

This spread indicates that it is more likely for a country's respondents to be concerned about the threat of climate change than not.

Percentage of citizens who think climate change poses a threat within 20 years





# Comparative PMFs

The dataset will be used to generate two subsets:

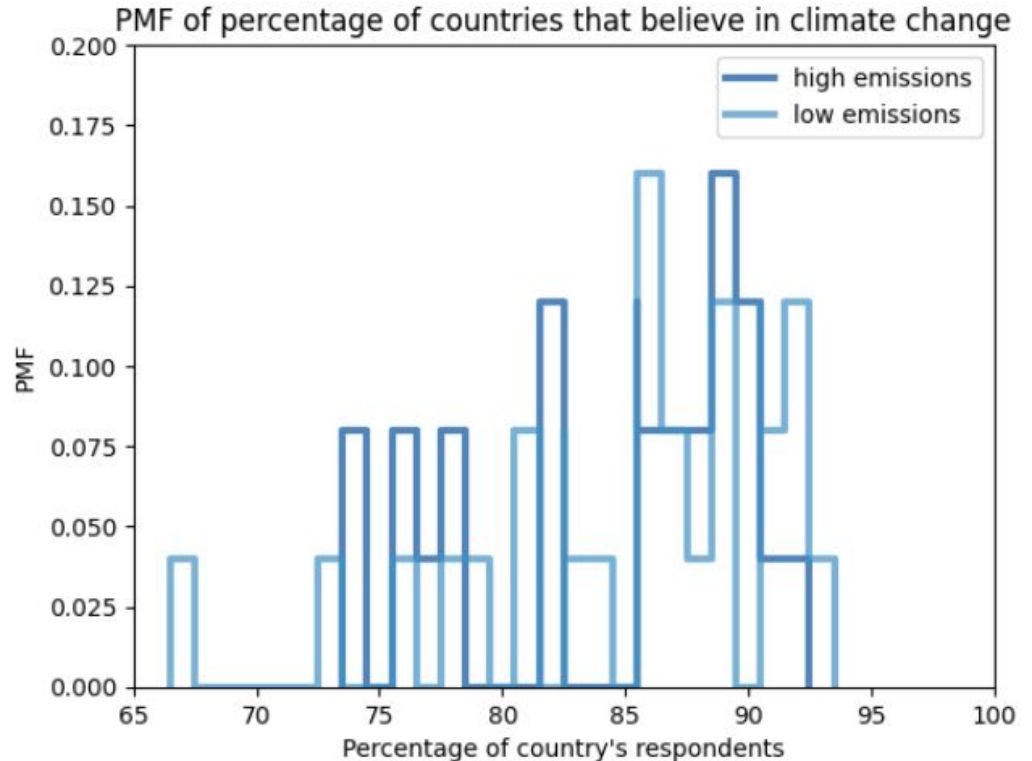
- the countries with the highest emissions ( $\geq$  **75th percentile**);
- the countries with the lowest emissions ( $\leq$  **25th percentile**).

The following slides compare these two scenarios and their relative climate sentiment survey results.



# PMF: Belief in Climate Change

On visual inspection, countries that have higher CO2 emissions may be slightly more likely to have a population that believes in climate change than a country that generates low emissions. However, the disparity is too inconsistent for a visual inspection to be conclusive.

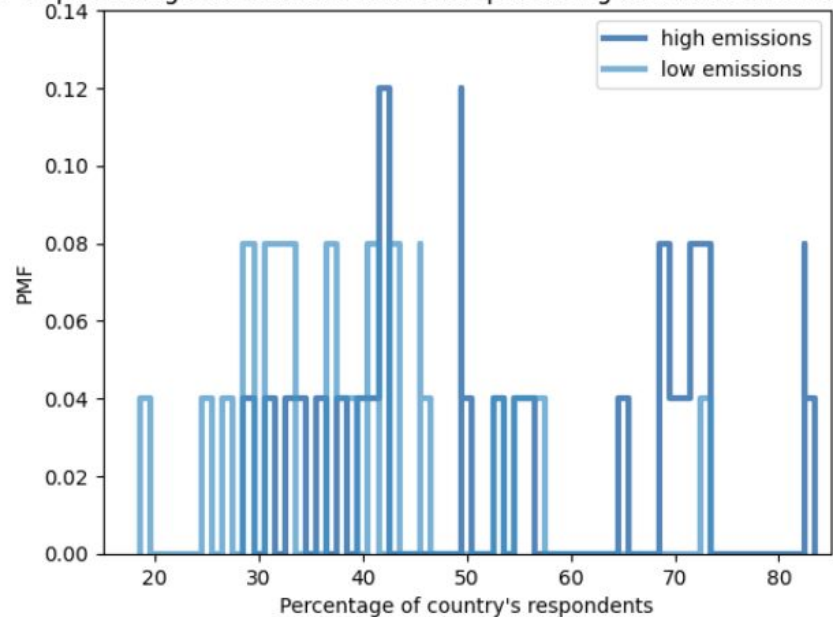


# PMF: Education on Climate Change

Contrary to the PMF on belief in climate change, countries that generate more emissions are more likely to have a majority (>50%) of the population that self-reports being educated on climate change.

While this doesn't reflect the actual knowledge of a country's citizens, it does likely reflect the availability of resources and climate change-related reporting.

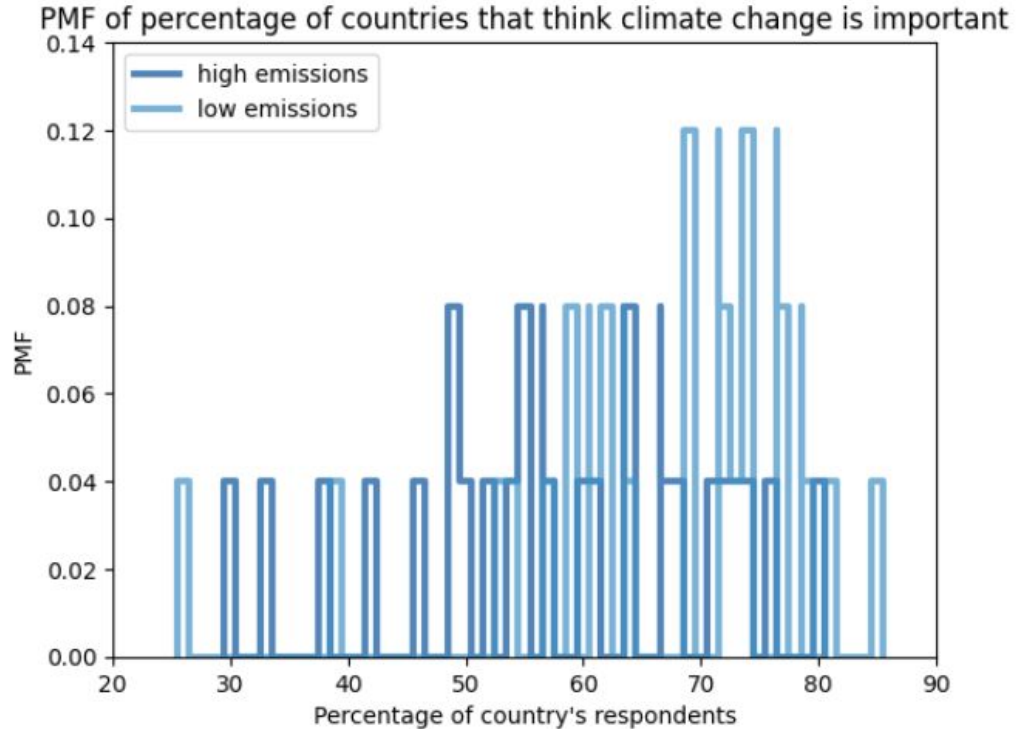
PMF of percentage of countries that self-report being educated on climate change



# PMF: Importance of Climate Change

Countries with low emissions have populations that are more likely to report that climate change is personally important to them.

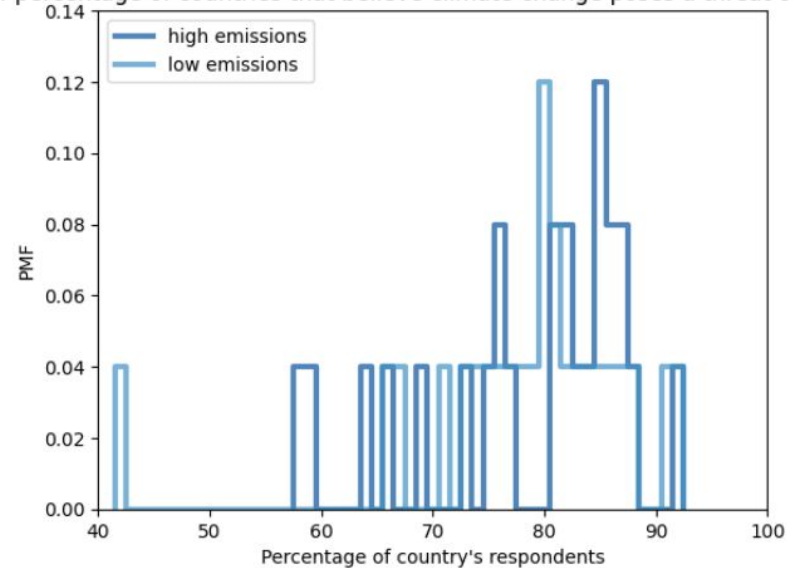
This may indicate that countries with low emissions consider themselves likely to be affected by countries with high emissions (e.g., countries they cannot affect). It may also be a reverse causality: Countries with low emissions may have low emissions *because* its citizens care about the topic.



# PMF: Perceived Threat of Climate Change

Countries with higher emissions may be marginally more likely to believe that climate change poses a threat over the next 20 years. However, these values are closely aligned and are likely not significantly different between the two groups.

PMF of percentage of countries that believe climate change poses a threat over 20 years





## Comparative CDFs

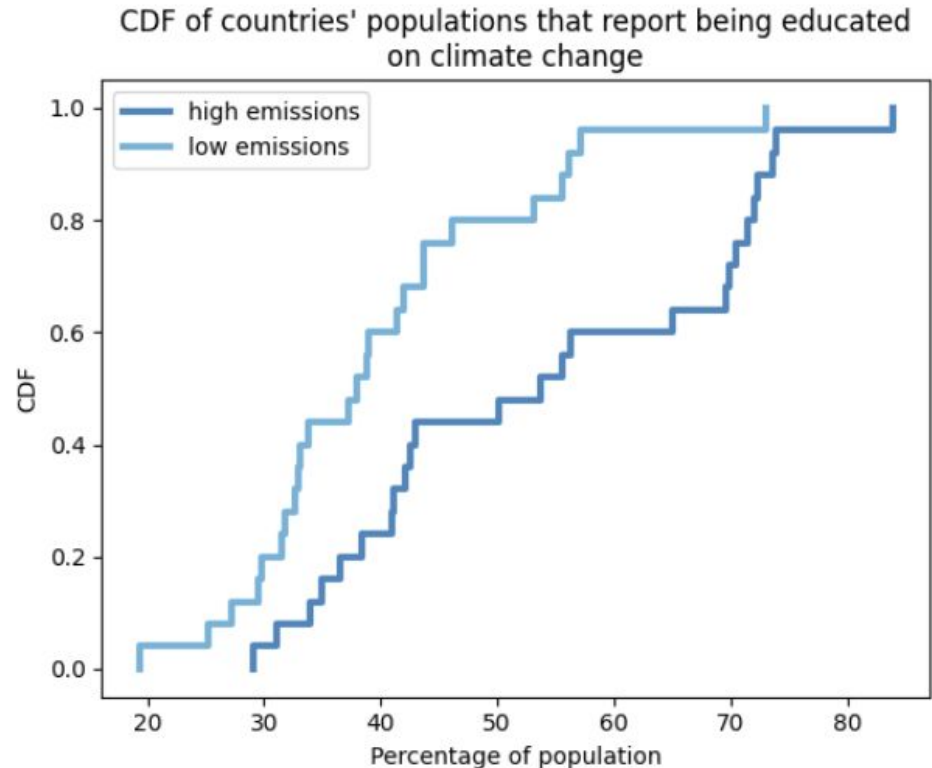
CDFs are a better analytical tool for this scenario because of the number of values. However, the PMFs provide likely avenues for comparison.

From the disparities between the high- and low-emission groups in the PMF comparisons, the **education** and **importance** variables show the most disparity and therefore are of the most interest, for now, as potential predictors.

# CDF: Education on Climate Change

The CDF shows much more clearly than the PMF that **higher-emission countries** are more likely to have populations that are educated on climate change.

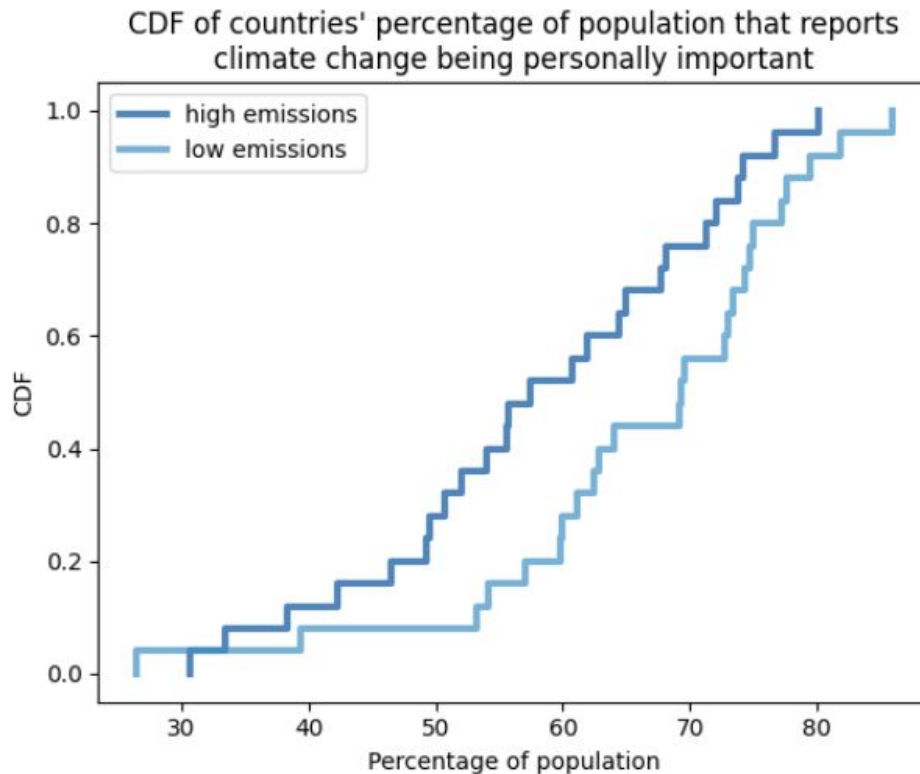
This could be because of public education surrounding efforts to reduce or affect emissions.



# CDF: Personal Importance of Climate Change

Similarly, the CDF shows that **high-emission countries' citizens consistently report being less personally concerned about climate change** than low-emission countries' citizens.

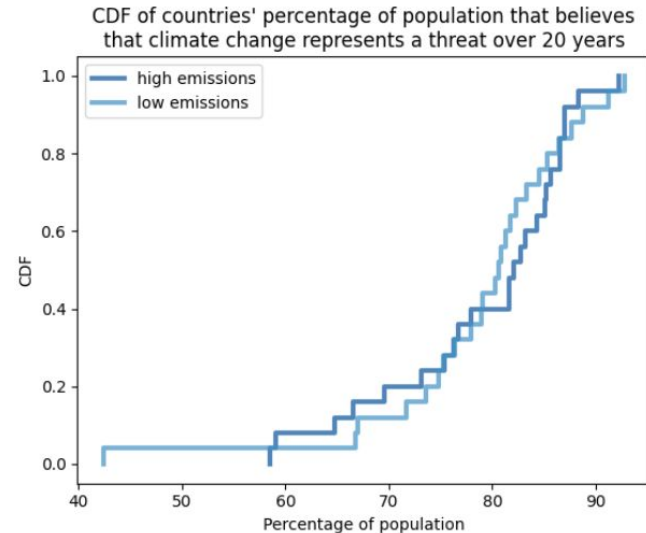
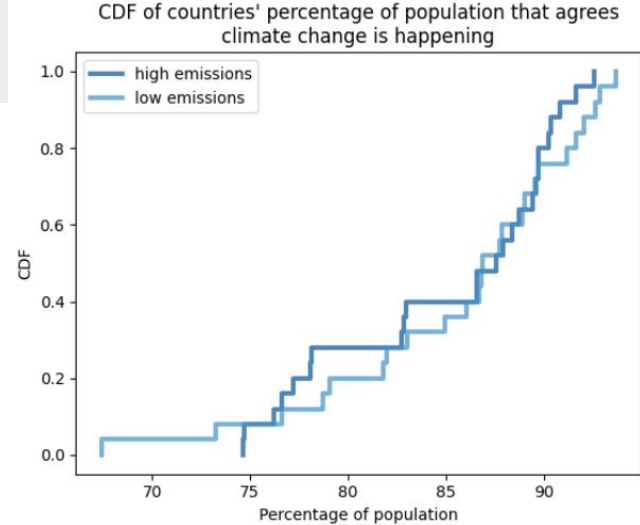
This could be for many reasons, e.g., perhaps high-emissions countries (which are more likely to have a high GDP) are wealthier and therefore more likely to be insulated from the personal effects of climate change, such as housing loss or the loss of agricultural jobs.



## CDF Analysis

Focusing on the sentiment variables of **climate education** and **personal importance of climate change**, there are clear disparities between high- and low-emission countries' citizens' survey responses and overall sentiment.

The two variables that were less promising in the PMF stage due to a lack of clear visual tendency have proven to be equally unclear in CDF analysis (see right) and are therefore less promising for further analysis.





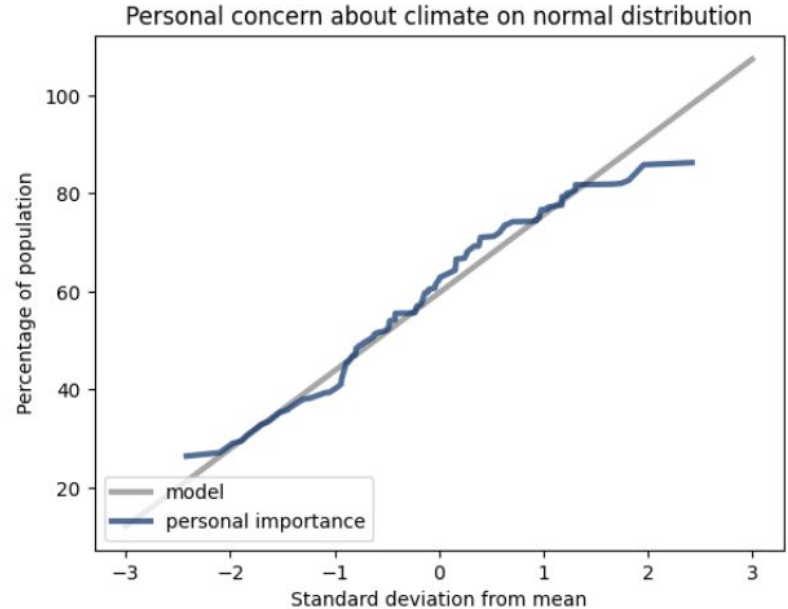


# Analytical Distribution Using Normal Distribution

While no distribution model fits the personal importance variable distribution perfectly, a simple normal distribution is likely the best fit.

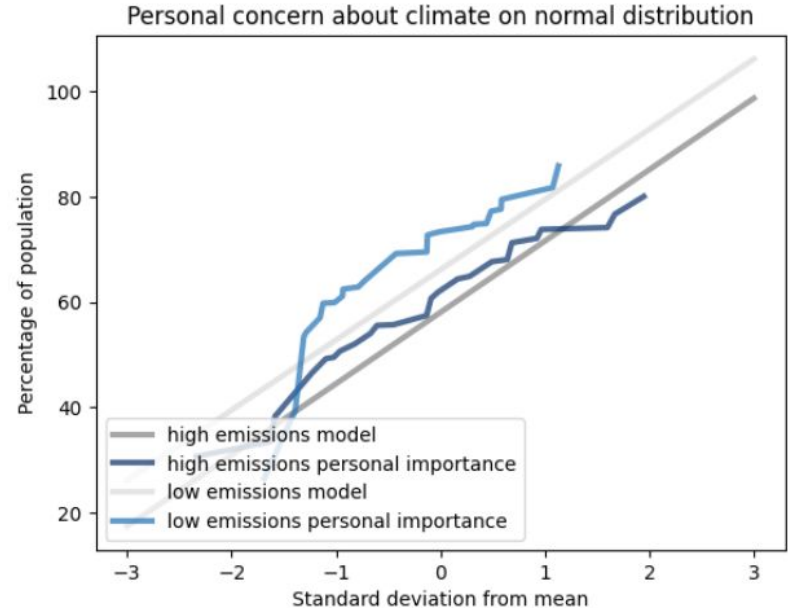
# Normal Distribution

The normal distribution best represents the personal concern metric (e.g., compared to a lognormal or Pareto distribution). The tails are slightly skewed, but not only is this expected behavior, the body of data is necessarily small (we cannot invent new countries), which likely creates the noise that causes deviations from the normal distribution.



# Normal Distribution

Normal distribution can be expected in this scenario, but the necessarily small sample size of countries causes enough noise to throw this model into doubt. However, modeling all countries (see previous slide) as opposed to the high- and low-emissions countries (see right) uses a larger body of data and therefore a more regular, “smooth” model. This indicates that a normal distribution is accurate, which can be expected from survey data.





# Scatter Plots

The best way to visually compare the relationship between climate sentiment and emissions is through scatter plots. Here, scatter plots will be used to compare:

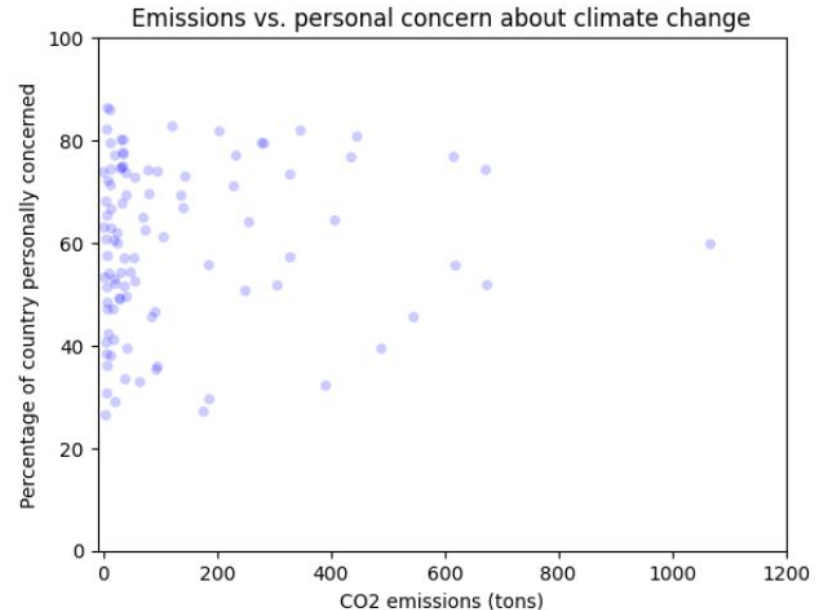
- **total CO2 emissions and countries' percentage of personal concern about climate change**
- **total CO2 emissions and countries' percentage of education on climate change**

## Emissions & Personal Concern

Contrary to the CDF analysis, scatter plots do not indicate a clear correlation between emissions and personal concern about climate change.

To the right, all countries are represented in a single scatter plot. The mean is 60%, which can be compared to the high- and low-emission breakouts (next slide).

*Scatter plot has had one (>2500 CO<sub>2</sub>et) outlier removed for readability.*

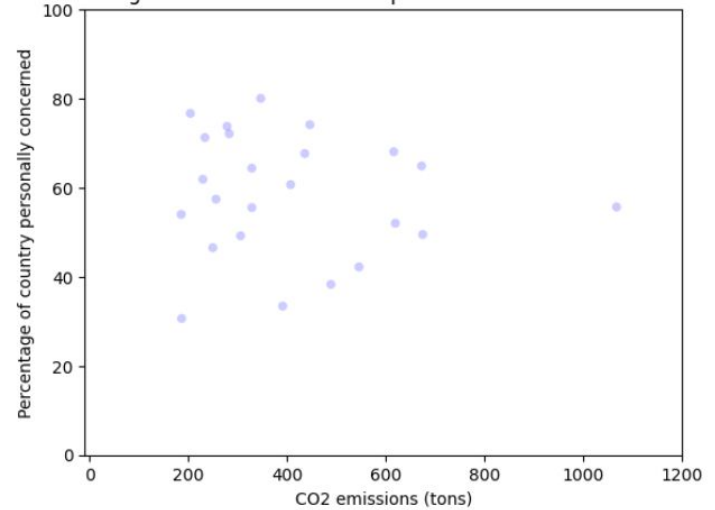


# Emissions & Personal Concern

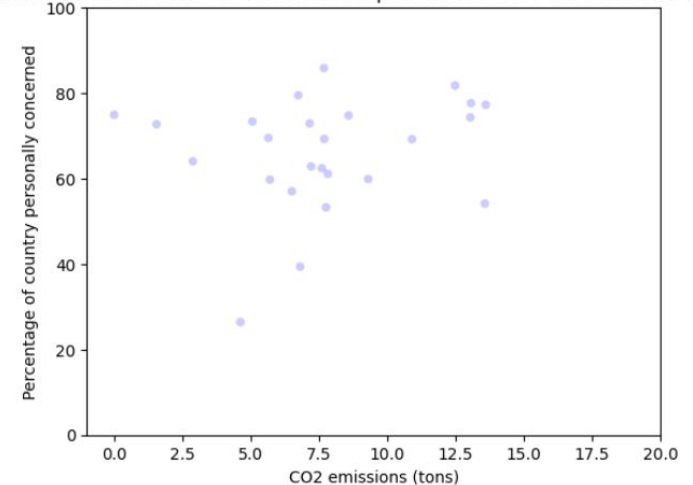
Subdividing the scatter plots by high-emission countries and low-emission countries, the large cluster of low-emission countries sits higher than that of the high-emission countries, reflecting their respective means of 66% (low emissions) and 58% (high), compared to the all-country mean of 60%.

*Scatter plot has had one (>2500 CO2et) outlier removed for readability.*

Emissions in high-emission countries vs. personal concern about climate change



Emissions in low-emission countries vs. personal concern about climate change





## Emissions vs. Concern: Correlation & Covariance

Testing for correlation between a country's emissions and the percentage of the population that is personally concerned about climate change, we find that:

- The covariance value is **122.4**;
- Pearson's correlation value is **0.024**;
- Spearman's rank correlation value is **0.102**.

For the purposes of analysis, correlation is more useful here, and Spearman's correlation more valuable because of the significant presence of outliers in the dataset. A positive correlation of ~0.1 indicates that emissions have a very minor predictive effect for personal concern, or vice versa.

Notably, high-emission countries, when isolated, report a negative correlation (-0.14 Pearson, -0.12 Spearman).



## Emissions vs. Concern: Causation

Though there is a minor correlation, it is difficult to determine the direction of potential causation with this data. It may be said that emissions have an effect on a given citizen's climate concern because of emissions' effect as a public concern. On the other hand, it may be said that climate concern has an effect on emissions because it drives political action and legislation.

**Further insight on a longer time frame** is necessary.





## Emissions vs. Concern: Hypothesis Testing

For hypothesis testing, a Spearman correlation is calculated for a random permutation (from *numpy.random.permutation*) and, from the hypothesis test, the p-value is determined.

Using Spearman's correlation, the p-value is **0.309**.

This indicates that there is a 31% probability that the distribution and correlation could be by random chance. Ideally, the p-value would be  $\leq 0.1$ , indicating a  $\leq 10\%$  likelihood. 31% is likely not statistically significant.

A larger dataset would be ideal for determining whether or not this correlation is truly significant.



# Emissions vs. Concern: Regression Analysis

A model was constructed using the method of least squares to determine the potential usefulness of predicting climate concern from a country's CO2 emissions.

As expected, the p-value is high enough to be a concern, at **0.358**. The intercept of **49** can be used to construct a plot of the model (see next slide).

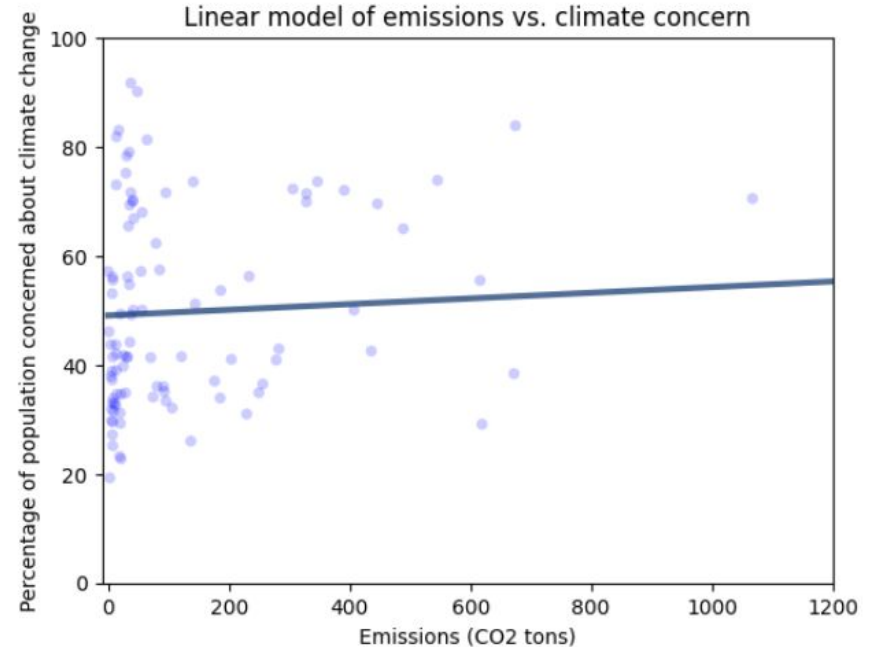
OLS Regression Results						
Dep. Variable:	climate_awareness			R-squared:	0.009	
Model:	OLS		Adj. R-squared:	-0.002		
Method:	Least Squares		F-statistic:	0.8526		
Date:	Thu, 27 Feb 2025		Prob (F-statistic):	0.358		
Time:	19:32:19		Log-Likelihood:	-425.33		
No. Observations:	99		AIC:	854.7		
Df Residuals:	97		BIC:	859.9		
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	49.0909	2.006	24.474	0.000	45.110	53.072
total_emissions	0.0052	0.006	0.923	0.358	-0.006	0.016
Omnibus:	12.424	Durbin-Watson:	1.970			
Prob(Omnibus):	0.002	Jarque-Bera (JB):	7.184			
Skew:	0.491	Prob(JB):	0.0275			
Kurtosis:	2.118	Cond. No.	396.			

# Emissions vs. Concern: Regression Analysis

While there is a marginally positive slope, as expected, this model shows that the p-value of the emissions variable is fairly high at 0.358.

Additionally, while the slope is centered on the scatter plot, the majority of the points do not align with the slope.

Therefore, total emissions likely have low predictive power against climate concern, and this model has minimal predictive value.





## Concluding Remarks: Challenges

One of the primary challenges of working with country data is the necessarily low number of data points available for analysis. It would be valuable to reattempt this analysis on data that is further divided into subregions. Having such a limited dataset presents problems while attempting to fit the data to a distribution model, and ultimately presented some very ill-fit data.

Additionally, further research comparing climate sentiment and wealth, e.g. through GDP, may be valuable as this is one of the most prominent divisions between high- and low-emission countries. Further analyses may also benefit from focusing on climate education efforts and political movements, such as the presence of climate action movements or petroleum lobbies.



## Concluding Remarks: Analysis

In this exploratory data analysis project, country emissions data and country-wide climate opinion survey data was used to attempt to determine whether a country's CO2 emissions have an effect upon climate opinion.

The exploratory analysis is inconclusive. The most promising variables that differ strongly between high- and low-emission countries are whether or not citizens were likely to report that they were well-educated on climate change, and whether citizens were likely to be personally concerned about climate change. High-emission countries' populations are more likely to have populations that report being educated on climate change, while low-emission countries' populations are more likely to report being personally concerned about its effects.

However, emissions appear to have minimal predictive power for either of these variables, or vice versa.