Hi! This is team EcoView. Our team members include

- Violet Huang: yuanninghuang@uchicago.edu
- Hantao Xiao: hantaox@uchicago.edu
- Anmin Yang: anmin@uchicago.edu

1. Project Description

Background

Climate change has become a universal challenge that needs to be addressed urgently. According to the United Nation's report, global greenhouse gas emissions need to reach net zero by 2050 to successfully halt climate change. Despite the urgency of such task, there is still a divide in the recognition of the urgency of climate change among partisan lines in the U.S.. In a survey conducted by Pew Research Center in 2020, 75% of Democrats think climate change should be a top priority for the president and the Congress, while only 21% of Republicans think the same.

Project's goal

In accord with United Nation's goal, we aim to explore the relationship between the sustainability each state has pulled in the U.S. and the political leaning of each state from 2018 to 2021 to understand whether states committed more to sustainability differ from states who don't in terms of political ideology.

Research question

From 2018 to 2021, what is the relationship between sustainability and political leaning at the state level?

Hypothesis

From 2018 to 2021, states that are more sustainable are also more democratic.

Significance

This project is significant in two regards. Firstly, while previous studies such as the Pew Research Center survey have examined the perception of climate change at the individual level, our project shifts the focus to the state level. This is important because states play a crucial role in combatting climate change and have a greater impact than individuals. Secondly, our project assesses the actions taken by each state to quantify their efforts in addressing climate change, as opposed to perception. By doing so, we can accurately measure and compare the commitment of each state towards combating climate change.

2. Methodology

Conceptualization

Political leaning is defined as the extent of a state being Democratic or Republican.
We conceptualize sustainability as the efforts and effects of a state in conserving environment and using renewable resources. To further elaborate on the concept of sustainability, we break it down into four aspects: clean energy jobs, clean energy share, legislative efforts, and environmental quality. We define clean energy jobs as the amount of people working in the clean energy sector. Clean energy share
measures the proportion of energy generated using clean energy source. Legislative efforts captures the extent of a state proposing and installing environmental protection and energy conservation laws and policies at the state and federal level. Finally, Environmental quality measures the amount of pollutants a state emits and the cleanliness of environment.

Data Collection

We have five concepts: political leaning, clean energy jobs, clean energy share, legislative efforts, and environmental quality. Each concept may have one or more variables to capture the concept. For each variable, we denote its definition and source

1. Political Leaning
    a. Variable: Political leaning score
        i. Definition: This variable calculates the average margin difference between how a state or district votes and how the country votes overall as the indicator for political leaning.
        ii. Source: Partisan Lean Index
2. Sustainability - Clean energy jobs
    a. Clean Energy Job Employment (scraped)
        i. Definition: This variable represents the amount of labor working in the clean Energy industry by state
        ii. Source: E2 Clean Energy Annual Reports 2018 - 2022
    b. Clean Energy Job Growth
        i. Definition: This variable measures the percentage change in clean energy employment compared to the previous year by state
        ii. Source: Computed from Clean energy job employment
    c. Clean Energy Job Share
        i. Definition: This variable measures the ratio of total workforce working in the clean energy industry
        ii. Source:E2 Clean Energy Annual Reports 2021 - 2022, Bureau of Labor Statistics Employment and Unemployment Reports 2018 - 2019
3. Sustainability - Clean energy share
    a. Green Ratio

        i. Definition: This variable measures the proportion of green source in electricity generation
        ii. Source: US Energy Information Administration
4. Sustainability - Legislative Efforts
    a. Policy Score
        i. Definition: This score indicates the implementation level energy efficiency policies in a state. It is a metric developed by The American Council for an Energy-Efficient Economy
        ii. Source: The American Council for an Energy-Efficient Economy
    b. Number of Bill (Scraped)
        i. Definition: This variable measures the number of environmental protection bills proposed in a year by each state
        ii. Source: GovTrack Website
5. Sustainability - Environmental Quality
    a. Good Day Ratio
        i. Definition: This variable measures the proportion of days with good air quality in a year
        ii. Source: United States Environmental Protection Agency
    b. Emission Per Capita
        i. Definition: This variable measures the amount of emission of harmful particles per capita
        ii. Source: United States Environmental Protection Agency

Data Cleaning and Wrangling

1. Political Leaning Score: the data is missing for year 2019 and year 2021. We use the average of nearby years to approximate the missing value
2. Clean Energy Job Employment: this data only includes the top 10 states in clean energy employment for year 2017 and 2018. We use clean energy employment data from 2019 to 2021 to approximate the 2017 and 2018 data for the rest of the 40 states using linear regression
3. Clean Energy Job Growth: we use the below formula to calculate the clean energy job growth for each year after we have each year's clean energy job employment data by state:

$$Clean\ Energy\ Job\ Growth\ = \frac{Clean\ energy\ job\ employment\ current\ year - Clean\ energy\ job\ employment\ previous\ year}{Clean\ energy\ job\ employment\ previous\ year}$$

4. Clean Energy Job Share: this data is recorded for year 2020 and 2021 from the E2 annual reports. For year 2018 and 2019, we use December employment data from BLS to calculate the variable by the formula below

$$Clean\ Energy\ Job\ Share\ = \frac{Clean\ Energy\ Job\ Employment}{Total\ Workforce\ in\ December}$$

5. Emission Per Capita: The original data set does not consider the effect of population on emission, therefore we divide the original data by each state's population to obtain emission per capita
6. After we have a master data sheet for all the variables we wanted, we normalized the data for each year on a scale of 0 to 1. For each state, we also calculated the average data for all four years

Data Analysis Procedure

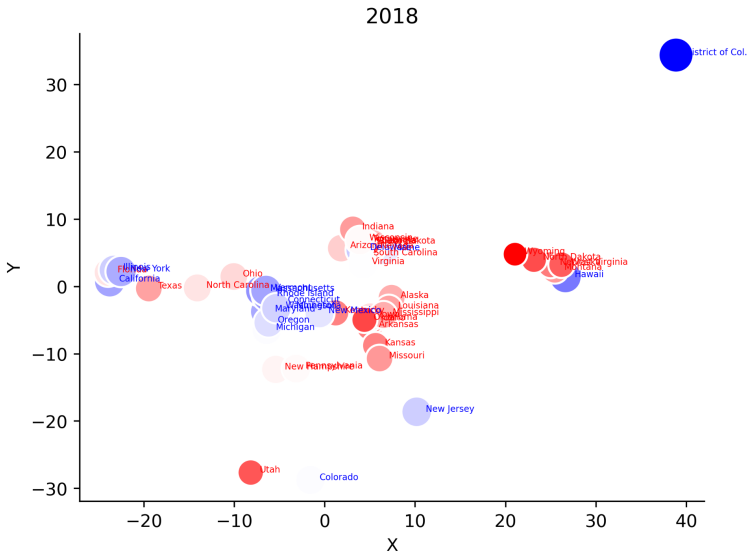
We first reduced the dimension of our independent variables and made a plot to see if blue states and red states will form two separate clusters. Then, we plotted a heat map for each year and for the averaged data. We used OLS regression to find the relationship between these sustainability variables and political leaning. We also conducted panel data regression on these data to explore the effect of time on the results

3. Results

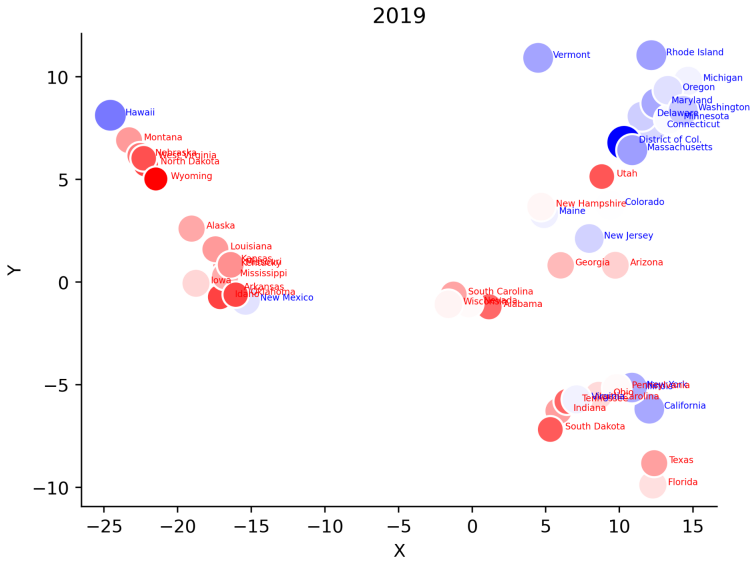
Dimension Reduction

We can see that blue states and red states both tend to form their own clusters. This shows that our hypothesis is likely to be correct and gives us confidence for later analysis.

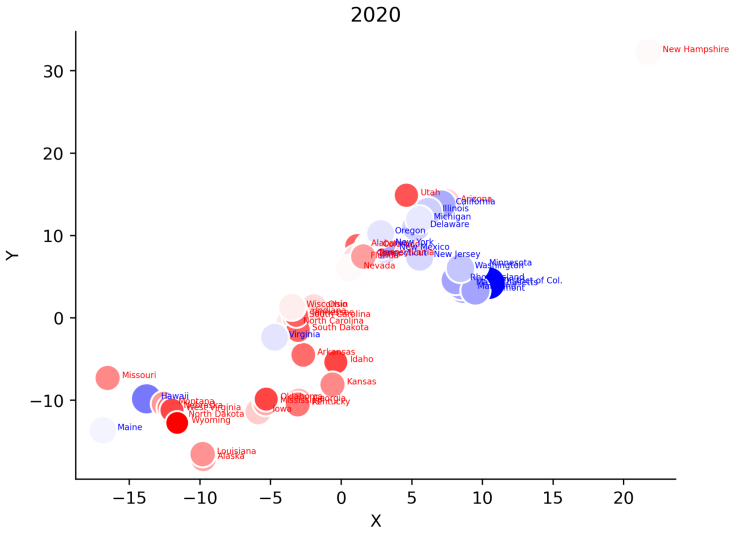
1. 2018



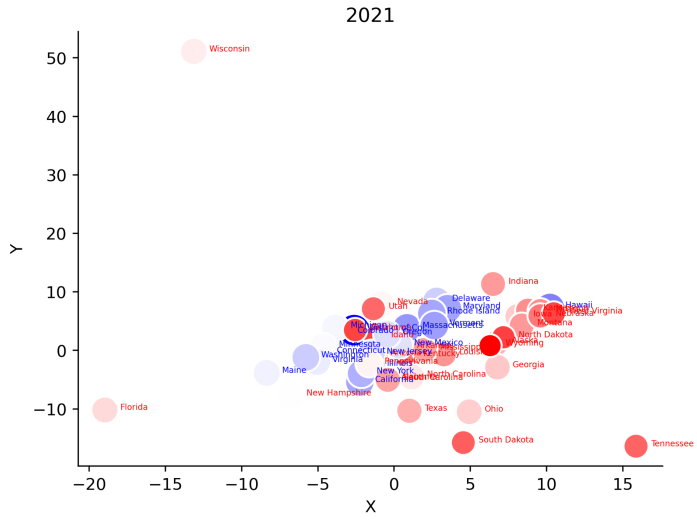
2. 2019


2019
Vermont
Rhode Island
Michigan
Oregon
Maryland
Washington
Minnesota
Connecticut
District of Col.
Massachusetts
Utah
New Hampshire
Maine
Colorado
New Jersey
Georgia
Arizona
Hawaii
Montana
Nebraska
West Virginia
North Dakota
Wyoming
Alaska
Louisiana
Kansas
Kentucky
Mississippi
Iowa
Arkansas
Oklahoma
New Mexico
South Carolina
Wisconsin
Alabama
Indiana
California
South Dakota
Texas
Florida
Y
X

3. 2020


2020
New Hampshire
Utah
Arizona
California
Illinois
Michigan
Delaware
Oregon
New Mexico
New Jersey
Nevada
Minnesota
District of Col.
Wisconsin
North Carolina
South Dakota
Virginia
Arkansas
Idaho
Missouri
Kansas
Hawaii
West Virginia
North Dakota
Wyoming
Maine
Louisiana
Alaska
Y
X

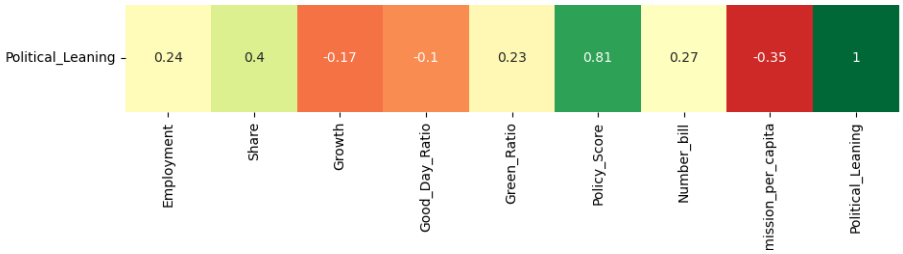
4. 2021



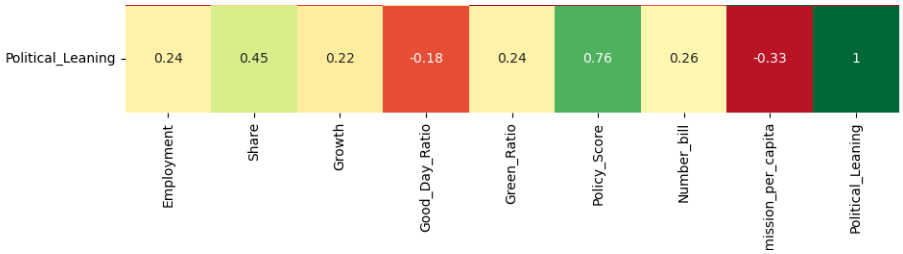
Heat Map

To focus on the effects of sustainability variables on political leaning, we cropped the heat maps to only show the correlation of the last row. By comparing the averaged heat map and the heat maps for 2018 to 2021, we can see that despite minimal changes in the correlation of most variables, the correlation between Clean Energy Jobs Growth and Political Leaning experienced a significant change from 0.22 to -0.35. Therefore, we need to run our OLS regression for each year separately to analyze this effect. From the heat map, we can see that there is a strong correlation between Policy Score and Political Leaning across all years. We can expect that this relationship is significant in the following regression.

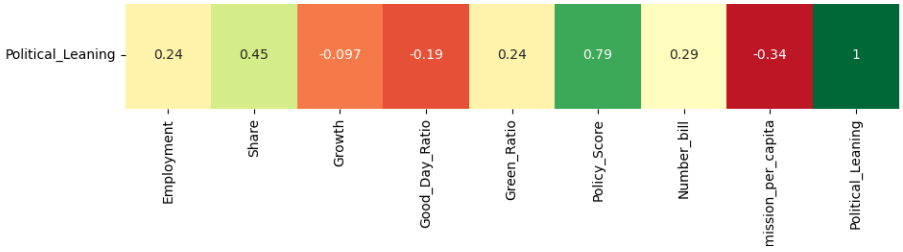
1. Average



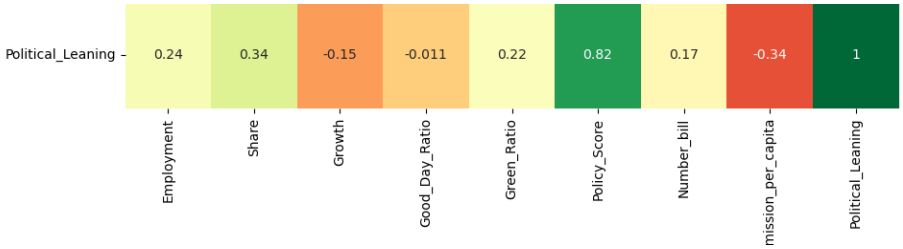
2. 2018



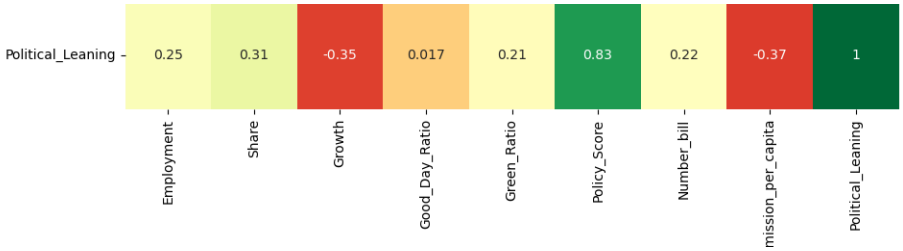
3. 2019



4. 2020



5. 2021



OLS Regression

The regression results coincide with the results from heat maps. We can see that policy score and political leaning have a strong, positive, and statistically significant relationship across four years. It's interesting to note that albeit the small correlation between Clean Energy Jobs Growth and Political Leaning, this relationship is significant in the OLS regression in and only in 2019.

1. 2018

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.2790	0.113	2.472	0.018	0.051	0.507
Employment	0.0998	0.175	0.571	0.571	−0.253	0.453
Share	−0.1541	0.122	−1.265	0.213	−0.400	0.092
Growth	−0.0894	0.079	−1.129	0.265	−0.249	0.070
Good_Day_Ratio	−0.0043	0.097	−0.044	0.965	−0.200	0.192
Green_Ratio	−0.0013	0.085	−0.015	0.988	−0.173	0.170
Policy_Score	0.6003	0.084	7.146	0.000	0.431	0.770
Number_bill	−0.2219	0.165	−1.347	0.185	−0.554	0.110
emission_per_capita	−0.0742	0.097	−0.763	0.449	−0.270	0.122

2. 2019

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.2146	0.137	1.569	0.124	−0.061	0.491
Employment	−0.0068	0.132	−0.052	0.959	−0.273	0.259
Share	−0.0869	0.119	−0.730	0.469	−0.327	0.153
Growth	−0.1877	0.083	−2.248	0.030	−0.356	−0.019
Good_Day_Ratio	0.0579	0.114	0.508	0.614	−0.172	0.288
Green_Ratio	−0.0342	0.081	−0.424	0.674	−0.197	0.129
Policy_Score	0.6459	0.081	7.938	0.000	0.482	0.810
Number_bill	−0.0941	0.110	−0.857	0.396	−0.316	0.127
emission_per_capita	−0.0715	0.088	−0.811	0.422	−0.249	0.106

3. 2020

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.3263	0.117	2.796	0.008	0.091	0.562
Employment	−0.0551	0.239	−0.230	0.819	−0.537	0.427
Share	0.1056	0.119	0.890	0.379	−0.134	0.345
Growth	−0.1401	0.097	−1.439	0.157	−0.337	0.056
Good_Day_Ratio	−0.1059	0.098	−1.076	0.288	−0.304	0.093
Green_Ratio	−0.0341	0.090	−0.379	0.707	−0.216	0.148
Policy_Score	0.5029	0.098	5.111	0.000	0.304	0.702
Number_bill	0.0080	0.217	0.037	0.971	−0.430	0.446
emission_per_capita	−0.1672	0.113	−1.484	0.145	−0.395	0.060

4. 2021

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.1847	0.117	1.580	0.122	−0.051	0.421
Employment	−0.2970	0.217	−1.369	0.178	−0.735	0.141
Share	0.1579	0.126	1.251	0.218	−0.097	0.413
Growth	0.1240	0.154	0.806	0.425	−0.187	0.435
Good_Day_Ratio	−0.0496	0.097	−0.513	0.611	−0.245	0.146
Green_Ratio	−0.0073	0.087	−0.083	0.934	−0.183	0.169
Policy_Score	0.4702	0.098	4.801	0.000	0.273	0.668
Number_bill	0.0637	0.102	0.626	0.535	−0.142	0.269
emission_per_capita	−0.1599	0.109	−1.469	0.149	−0.380	0.060

Panel Regression

The OLS regression model only captures the impact of sustainability variables on political leaning for the current year. However, political leaning may exhibit time-lagged effects that are not immediately apparent until subsequent years. To address this limitation, we utilize panel regression to complement the OLS model.

Our findings suggest that, irrespective of policy score, emissions per capita have a negative association with political leaning. This implies that an increase in emissions per capita is likely to decrease the probability of a state being blue in the next one to two years. We also observe different patterns in the ratio of clean energy employment to total employment between the pre- and post-COVID periods. Our hypothesis is that the

impact of the pandemic on the clean energy sector in red and blue states may be inconsistent.

1. Panel Regression 2018 - 2019

Political_Leaning	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Employment	.1410461	.1354429	1.04	0.298	-.1244171	.4065093
Share	.1775976	.0689454	2.58	0.010	.0424671	.3127281
Growth	-.013589	.0098067	-1.39	0.166	-.0328097	.0056318
Good_Day_Ratio	.0205491	.0121284	1.69	0.090	-.0032221	.0443204
Green_Ratio	-.0433591	.0358197	-1.21	0.226	-.1135645	.0268463
Policy_Score	.1063594	.0411994	2.58	0.010	.02561	.1871088
Number_bill	-.0232947	.0118588	-1.96	0.049	-.0465376	-.0000519
emission_per_capita	-.2372662	.1033147	-2.30	0.022	-.4397592	-.0347731
_cons	.3288159	.0500718	6.57	0.000	.2306769	.4269548

2. Panel Regression 2020 - 2021

Political_Leaning	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Employment	.1207923	.1368558	0.88	0.377	-.1474402	.3890247
Share	.1811628	.1033486	1.75	0.080	-.0213968	.3837224
Growth	.0039501	.0101636	0.39	0.698	-.0159702	.0238705
Good_Day_Ratio	-.0540522	.0220723	-2.45	0.014	-.0973131	-.0107913
Green_Ratio	-.0262758	.0676173	-0.39	0.698	-.1588032	.1062517
Policy_Score	.1011032	.0429086	2.36	0.018	.017004	.1852024
Number_bill	-.0009207	.0232574	-0.04	0.968	-.0465043	.044663
emission_per_capita	-.2271984	.0875414	-2.60	0.009	-.3987764	-.0556204
_cons	.3574344	.0673223	5.31	0.000	.2254852	.4893837

3. Panel Data Analysis Lag 1 Period

L.Political_Leaning	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Employment	.1391418	.1155526	1.20	0.229	-.0873371	.3656206
Share	.0594866	.0273876	2.17	0.030	.0058078	.1131654
Growth	.003701	.0065378	0.57	0.571	-.0091127	.0165148
Good_Day_Ratio	-.0006	.0122531	-0.05	0.961	-.0246155	.0234156
Green_Ratio	.0325906	.0446282	0.73	0.465	-.0548791	.1200602
Policy_Score	.109572	.0228621	4.79	0.000	.0647631	.1543808
Number_bill	-.015708	.0150408	-1.04	0.296	-.0451875	.0137715
emission_per_capita	-.1581511	.0612365	-2.58	0.010	-.2781724	-.0381298
_cons	.2870352	.0435954	6.58	0.000	.2015898	.3724807

4. Panel Data Analysis Lag 2 Period

L2. Political_Leaning	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Employment	.1479558	.1308255	1.13	0.258	-.1084576	.4043691
Share	.1372112	.0809876	1.69	0.090	-.0215215	.295944
Growth	-.0009115	.0065027	-0.14	0.889	-.0136566	.0118336
Good_Day_Ratio	.0148112	.0142512	1.04	0.299	-.0131207	.0427431
Green_Ratio	.0656976	.0503283	1.31	0.192	-.032944	.1643392
Policy_Score	.0671552	.0286891	2.34	0.019	.0109256	.1233847
Number_bill	-.00211	.0147499	-0.14	0.886	-.0310193	.0267993
emission_per_capita	-.1755526	.0732076	-2.40	0.016	-.3190369	-.0320684
_cons	.2525065	.0514742	4.91	0.000	.151619	.353394

4. Finding

Dimension reduction analysis provides insights into the potential effectiveness of our selected metrics in predicting political status. Subsequently, through correlation analysis, we quantified the strength of the relationship between each metric and political leaning. All features, except the good day ratio that quantifies air quality, were found to be correlated with political leaning.

Our OLS regression analysis of each year identified sustainable political score as a robust indicator of political leaning. States that are efficient in environmental policy tend to lean blue.

We also used panel regression analysis to identify emission per capita as a time-lagged predictor of political leaning. Specifically, we found that higher emission per capita this year may increase the possibility of a state turning red in the following one to two years.