Given one's access to greenspaces, political affiliation, and mental health, which is the most significant predictor of one's economic status?

An analysis of the relationship of greenspaces, political status, and mental health on economic status

TEAM 8: Arushi Sharma, Anastasiia Rudenko, Kelly Yang, Sashrika Pandey

Citadel & Citadel Securities Women's Datathon Spring 2023
February 4, 2023

CONTENT:

1. EXECUTIVE SUMMARY

The influence of green spaces on human happiness and life has already been thoroughly researched. But how big is this influence? They say: "you cannot buy happiness", – and yet, it doesn't mean we don't need money to pay our bills, right? How much of an impact do green spaces have on one's financial situation? How big is it compared to, say, politics or mental health diagnoses – other major forces in the individual's life? In this study, we attempted to answer these questions.

1.1 Background Summary

Over the course of the datathon, we attempted to predict the influence of green spaces, mental health, and political affiliation on the individual's economic status. We aimed to determine which of these three significant life factors has the greatest effect on an individual's financial situation. This analysis can provide a better insight into the relative importance of green spaces versus politics and mental health in society.

Let us look closely at the three aspects we're talking about.

A. POLITICAL AFFILIATION:

Political stress, intolerance, and bigotry are all tremendously stressful to witness through media, be a part of and even be a target of. According to a survey conducted by The Harris Poll on behalf of the American Psychological Association:

"Ahead of the most divisive election in decades, more than two-thirds of U.S. adults (68%) say that the 2020 U.S. presidential election is a significant source of stress in their life, a large increase from the 2016 presidential election when 52% said the same"

It is unfortunate to witness such a tendency, but we cannot deny that politics became a huge part of our life, our thoughts, and our concerns.

Political conflicts are particularly stressful because political views are so divisive, and these disagreements occur not only between acquaintances but between close friends and especially family members, which is when things become even tenser. On top of that, many changes are happening in the country that can be particularly stressful for those who are concerned about the direction things are taking. This brings even more potentially divisive discussions to the front.

How far does it influence our life? How deeply does it impact our economy as individuals?

B. GREENSPACES:

It's important to recognize that greenspaces provide countless benefits and a highly positive influence on human mental and physical health. The ScienceDirect studies show that green spaces have a great effect on improvements in memory, attentiveness, and learning ability, and a reduction in stress for small kids and young people. Not to mention their studies on increased levels of cardiovascular health, a lower risk of depression, and social isolation among the elderly.

Yet, on the other hand, despite being highly important, greenspaces do not bring equal benefits to everyone and are more beneficial for the well-being of certain socio-economic groups. An International Journal of Environmental Research and Public Health shows that people of lower socioeconomic status get greater benefits from urban green space than more privileged groups, especially in terms of reducing stress and improving mental health.

How is this inequality of benefits provided by greenspaces connected to the political affiliation of individuals? How much correlation does it have with the financial situation of each human being?

C. MENTAL HEALTH:

Mental health plays a crucial role in determining the quality of our lives. Studies have shown that prioritizing mental health is the most effective way to enhance overall well-being. It influences our thoughts, emotions, and behavior, as well as our ability to cope with stress, interact with others, and make healthy decisions. According to the World Health Organization (WHO), the number of individuals struggling with mental illness has dramatically increased in recent years, with the COVID-19 pandemic exacerbating the situation globally.
According to the Mayo Clinic, there's no sure way to prevent mental illness. However, while dealing with a mental illness, it can help to take steps to control stress, increase resilience, and boost low self-esteem. Does it influence our economical status and ability to cover our financial expenses? Certainly. But – to what extent?

D. CONNECTION OF THESE THREE:

Looking at all the questions stated above, another is raised – how is all this connected? Which one of these areas of human life is the most influential? Let's go on this journey together and see.

1.2 Summary of Findings

From the analysis that we did use several datasets, we concluded that the green coverage of an area was 67.903% accurate in predicting the economic status of an individual. This percentage indicates that there could be a correlation, if not a casual relationship, between the greenhouse coverage of a region and the resulting economic status. We concluded that mental health was 73.849% accurate in predicting the economic status of the individual. Thus, there is a correlation between mental health and economic status. While the political system model is still currently running, we can infer from previous data done from multiple studies that an individual's socioeconomic status may influence political attitudes and engagement. Thus, these findings answered our question by emphasizing that mental health was the most significant predictor of economic status, but green coverage was immensely close to that finding.

2. TECHNICAL APPROACH

2.1 General Approach

After looking over the provided databases, we realized we would have to find additional data to join between the existing databases. After finding connector databases that had information linking parameters such as zip code to the GEO_ID of a given area, we mapped the zip code or state (based on the deepest level of granularity we could extract from the data) to the feature of interest. The features we decided to explore were the percentage of green coverage in an area, the at-risk percentages of denizens of a state based on 2020 data, and their political affiliation. As our goal was to analyze which of these features was most correlated with the economic status of an individual, for each of the features we cleaned and joined the data with the zip code in the provided data sets. We then used a Support Vector Machine library to create Support Vector Classification models for each of the individual features. Based on the data that we had collected, we split the data into training, validation, and test samples. After training the models, we trained on the validation sample to identify the hyperparameter that yielded the greatest validation accuracy and decided on a training size that also led to the greatest validation accuracy. Following this step, we finally trained on the test data for each of the following features.

2.2. Data Collection and Cleaning

In addition to the datasets provided to us, we created four additional datasets of our own to add to our analysis. We began by analyzing Mental Health America's Mental Health Risk Assessments View of Depression dataset from 2022 to determine the percentage of people at risk for depression in the 50 states per year. Second, we analyzed the Internal Revenue Service (IRS)’s SOI Tax Stats - Individual Income Tax Statistics data for 2022 to determine the

economic status of individuals in different zip codes across the United States. In a third study, we examined the correlation between zip code and political standing to analyze individuals' political affiliations. Finally, we mapped a trek that stored the zipcodes for the tree and mental health dataset to connect all of our data in terms of zip code.

To clean the datasets, we used the Python Pandas library to parse and merge different datasets. While merging the datasets, we cross-checked different column values (e.g., zip code, percent park, median income, etc.) with each other. So if there was a certain zip code within one dataset but not within another, we did not include that zip code in our final DataFrame. For example, to create a data frame that mapped the percent park and median income to a particular zip code, we first used a dataset that contained the number of people within six different income brackets per zip code and found the median for each zip code, creating a DataFrame with zip code and median income as columns. Then, we used another dataset to map GEOID to a zip code to map a particular zip code to a percent park. Finally, we combined these two DataFrames to yield the final table that mapped the percent park and median income to a particular zip code. We followed a similar process for the rest of our datasets.

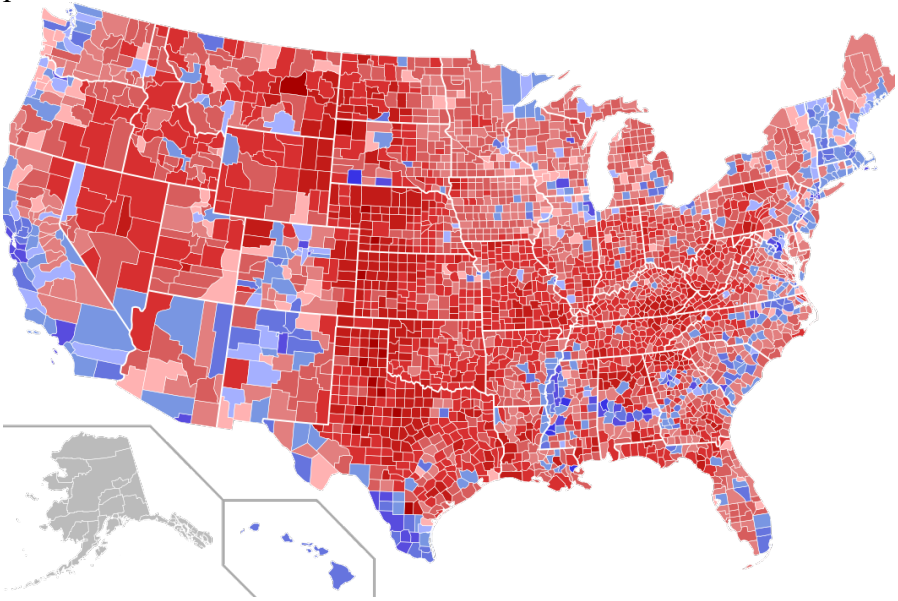
2.3. Conclusion

To conclude, it is clear that there is an indicative relationship between the impact of greenhouses on the economic status of individuals, based on our analysis. This allows us to conclude that the greenhouses that are created correlate with the economic status of individuals. One’s political leanings may also be indicative of economic status. However, the most significant correlation we found with one’s economic status was mental health, indicating that green coverage has the potential to have a greater correlation with economic status but isn’t the most indicative measure. However, greenspaces are something that we can control, by adding

more of them in different locations. Thus, this analysis can better assist greenspace creators as they create more greenspaces and the impact that they can have on individuals, with a potential impact financially.

2.4. United States Maps (Political, Economic, and Mental Health)

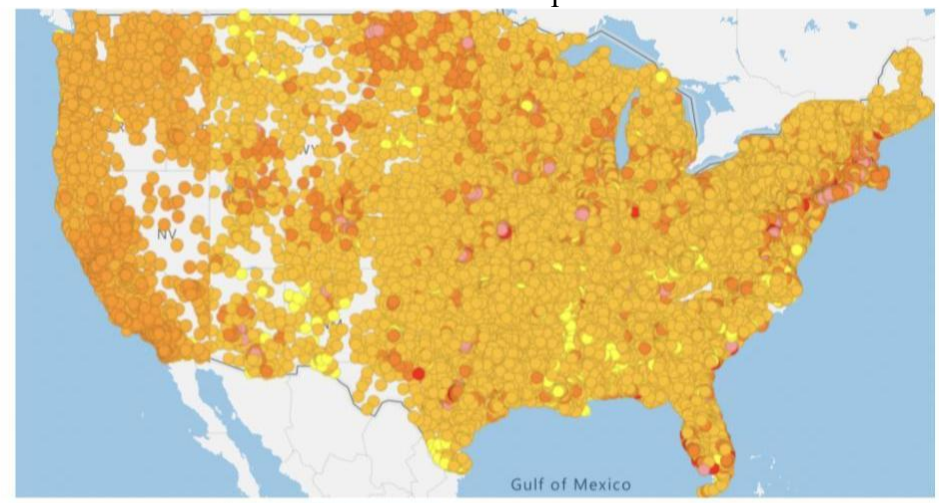
A. **Political Map** of the United States: This political map shows the two different political parties in the counties across the United States.



LEGEND: Red = Republican Party & Blue = Democratic Party

Source: The divide between us: Urban-rural political differences rooted in geography from Washington University in St. Louis

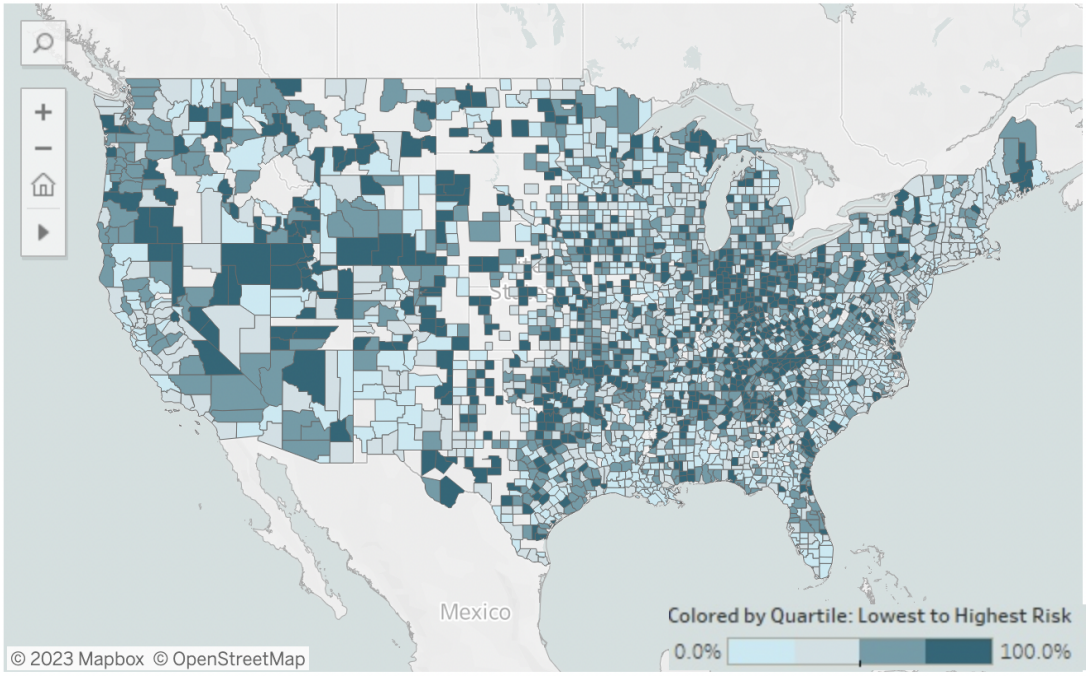
B. **Economic Map** of the United States: This economic map shows the different economic statuses across the United States based on zip code.



LEGEND:		
	Purple =	$1 under $25,000
	Yellow =	$25,000 under $50,000
	Light Orange =	$50,000 under $75,000
	Dark Orange =	$75,000 under $100,000
	Pink =	$100,000 under $200,000
	Red =	$200,000 or more
	Blue =	

Source: Individual Income Tax Statistics(IRS)'s SOI Tax Stats Data based on individual income tax returns filed with the IRS for 2022

C. **Mental Health Map** of the United States: This map shows the quarantine risk from lowest to highest risk of people with a depression diagnosis in 2021 based on the different counties in the United States.



Source: Mental Health America's Mental Health Risk Assessments View of Depression dataset from 2022 based on different quarters and the lowest to the highest risk

-page break-

3. APPENDIX

3.1 Roadblocks encountered

1. When looking for the political data, it was difficult to find the data of the party affiliation that correlated with a specific zip code that was free and easily accessible. We had to clean up the dataset that we found for the political data to get the information that we needed.

3.2 Caveats

1. The mental health data was not organized by zip code, but by county and state. We utilized a function to take the county and state information and analyze the data after parsing that instead.

3.3 Future research areas

1. **Analysis beyond the USA**: in the current analysis, we looked closely into the correlation between three aspects of the life of Americans – greenspace, political affiliation, and mental health – and how it affects our economic status as individuals. We can expand this research beyond the USA area and look more globally at the interrelation between the mental well-being, green spaces, and political involvement in places around the world (other than the USA), including Europe, Asia, the Middle East, Africa, etc.
2. **Expand analysis of greenspaces** onto blue spaces – such as reservoirs, oceans, seas, rivers, lakes, and others – and regions' weather conditions: as we saw over the research so far, greenspace is very important for an individual's health (both mental and physical). We highly conjecture that blue spaces are as beneficial as green ones. Same counts for the weather conditions – people tend to feel happier and mentally healthy with more

balanced and sunny weather over rainy or extreme. An interesting observation can be additionally made that blue spaces can influence weather to some extent – areas near reservoirs are known to have more smooth and balanced weather.

3. **Add dimension about the influence of world events** onto greenspaces and politics: one of the major world events that can ever happen and that influences all the aspects of the life of individuals is war. In addition to our current analysis, we can add a dimension on how war changes the number of greenspaces accessible worldwide and how it influences politics – and based on that, how it influences the financial situation.

THANK YOU!

We really enjoyed doing this project😄
