22.1.2021

The Divided Counties of America

An Analysis of the Impact of the Venue-Composition of Counties on the Election-Results in the 2020 Presidential Election

Dario Nališ, M.Sc.

Table of Contents

In the 2020 election, as well as in most recent elections, large cities in the USA overwhelmingly voted for the Democrat candidate, while smaller counties voted for the Republican candidate. Looking at an election map usually shows a largely red map, with the occasional blue dots in the large cities. However, as with all data, there is still a lot of variation to be explained. Many studies of course already analyzed the relationships between socioeconomic-status, education, ethnicity, etc., and voting behavior. With Foursquare, we have the opportunity of a totally different kind of analysis. Instead of analyzing the impact of population characteristics on voting behavior alone, we can explore if voting behavior is related to the kinds of venues different areas are made up of. We are going to use the Foursquare API to collect information on venues within the counties. We are going to cluster the counties by their venue-composition and use machine learning to predict how those clusters voted in the 2020 presidential election, above and beyond the usual sociodemographic variables.

Background and Problem Description

When watching the election coverage in the year 2020, there were quite a few interesting facts about voting behavior to observe. Besides the obvious red states and blue states which stay the same every election cycle with only small variations, we could again see the large difference between rural areas and metropolitan areas - the rural areas overwhelmingly voting for Trump and the metropolitan areas overwhelmingly voting for Biden. Because vote count reporting does not happen all at once or consistently distributed over the country, but happens in large chunks at a time, we observed very dramatical changes throughout election night and the week afterward. For example, Biden started with a large lead in the state of Texas early election night, which shocked many Trump-voters. When we looked at the counties that reported results, we could see that the lead was because almost only Dallas - a large city - was reporting results. Later Texas went to Trump quite decisively. A similar thing happened in the state of Ohio - a must-win for Trump - where Biden had a lead for a long period of time before the rural areas were reporting the results and Trump again won quite considerably. The most dramatic shifts happened in the night from December 3rd to December 4th. Trump went into the night with massive leads in Pennsylvania and Michigan, and a considerable lead in Wisconsin and already declared himself the winner, only to lose those states after large amounts of absentee ballots in the largest cities like Philadelphia, Detroit and Milwaukee were counted.

In my analysis, I want to look more deeply at the differences of features of counties and their influence on the election-results. The question to be answered is this: **Are there any features in the composition of counties when it comes to their venues, which can predict how many voters voted for Joe Biden and how many voted for Donald Trump above and beyond demographic variables like socio-economic status, education or ethnicity composition?** I want to understand more deeply what kind of counties favored Joe Biden and what kind of counties favored Donald Trump.

Who could be interested in this research? There is a famous Netflix Documentary called "The Great Hack", which depicts the successful efforts by Cambridge-Analytica to apply data science to influence the 2016 presidential election. Cambridge-Analytica was hired by the Trump-campaign to help them optimize their campaign strategy for optimal targeting of voters. They primarily focused on individual voter variables, to target those populations which could successfully be persuaded. As we can recognize from this fact, the candidates and political parties themselves could be stakeholders in our project. Elections are often close races and getting the extra-edge can put a candidate on top in the end. To be more specific, the Republican Party might be interested in understanding more thoroughly why they are so unsuccessful in metropolitan areas. A better understanding of the features within cities could help them understand their potential voters so they could adapt their message and campaign strategies to address this particularly difficult audience.

Data

I used three types of data for our analysis. The first were the election results by county. It consisted of 3152 counties with the relative and total votes President Trump and now President Biden received, the name of the state the county belongs to, the county's name and the county's Federal Information Processing System (FIPS)-code which is a numerical identifier of each county. The data was provided on a public github repository by Tony McGovern[1]. My second type of data were sociodemographic data of each county. I obtained data on unemployment rates and median household income of counties from the Economic Research Service of the United States Department of Agriculture[2]. The latest data were from 2019. The same source provided me with data from 2018 on the highest level of educational attainment. The data identified the proportion of the adult population (25 years or older) who completed a bachelor's degree, who started some college without graduating, high-school-graduates and proportion of population who did not graduate from high-school. I obtained data on ethnicities and population-sizes from the United States Census Bureau[3]. I simplified the data by using proportions of the population which identify as White-only, African American or mixed, Asian or mixed and Hispanic or mixed. The agencies did not provide data for all 3152 counties. I only used counties on which I had data on all variables which resulted in a final dataset of 3112 counties.

The final type of data I used were venue data from Foursquare Places[4]. Foursquare is a platform on which users mark and rate venues they visited and thus Foursquare created a large database of venues within regions of the whole world. Each venue belongs to a specific

[1] https://github.com/tonmcg/US_County_Level_Election_Results_08-20/blob/master/2020_US_County_Level_Presidential_Results.csv
[2] https://www.ers.usda.gov/data-products/county-level-data-sets/
[3] https://www.census.gov/content/census/en/data/datasets/time-series/demo/popest/2010s-counties-detail.html
[4] https://developer.foursquare.com/places

venue-category like coffee-shop, gas-station or park. I downloaded the 100 most popular venues within each of the 3112 counties using Foursquare-API. However, I did not obtain 100 venues for each county, because some of them are quite small, so no 100 venues were identified in all counties. I then calculated the relative frequency of a venue-category within a county and received a unique profile of venue-categories for each county.

Methodology

In the following section I am going to describe which type of analyses I performed along with a more in-depth description of some of the methods.

Exploratory Analyses of the 2020 Presidential Election

First, I did some basic analyses of the 2020 elections. I calculated the total and relative votes Trump and Biden received in the entire USA, as well as the average percentages per county. In the second step, I did a cluster analysis of the venue-data obtained from Foursquare and mapped the counties by cluster to get a visual representation of where the clusters can be found within the USA. I then investigated the relationships the clusters had with the demographic data, to get a better understanding of the clusters. To complete the exploratory analyzes I investigated the relationships of all the independent variables I was going to use in the inferential statistical analyses.

Clustering Counties by Venues

Foursquare provided a total of 595 venue-categories. Examples for categories are: 'Art Gallery', 'Bakery', 'Bus Station', or 'Medical Supply Store'. For each of the categories I calculated the relative frequency in which the category appeared within a county. For example, 9% of all venues in Los Angeles County belonged to the category 'Beach'. Each county had a unique composition of relative frequencies of venue-categories. I performed k-means clustering with scikit-learn[5] to receive a cluster each county would fall into. K-means clustering analyzes the similarities of data-point's features, generates central-points called centroids by calculating the smallest mean-distance of data-points to the centroids and puts each data point into categories called clusters by virtue of their nearest centroid, so that each data-point in the end belongs to one cluster. Clustering is a way of analyzing similarities of data-points, so we get an understanding how counties are similar or dissimilar with regards to their venue-composition. The practical result is that we reduce the thousands of counties to a small number of clusters, so we can later use the clusters as predictors for election outcome. This serves two purposes: First, we can interpret which kinds of counties preferred which candidate more easily – imagine having to interpret the influence of 595 venue categories. Secondly, we prevent the later models from overfitting – a problem which arises when we have too many features to predict an outcome where the model fits the noise in the data more than the actual population.

[5] https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html

A hard problem of k-means clustering is deciding how many clusters are generated to divide the datapoints into. There is no one correct solution for deciding k – or the number of clusters. A rule of thumb is to continuously apply k-means with increasing k, until the next k does not contribute to much information gain any more. The way I am going to estimate which k is best is by looking at the decrease in inertia with each added cluster. Inertia is the mean-distance to the centroid a data-point belongs to. It basically measures how well centroids represent the data within a cluster. The larger the inertia, the higher the average dissimilarity of a data-point with their cluster. In our case a low inertia would mean that the counties are very similar to other counties within their cluster. If we add a cluster and inertia decreases by a lot, we know that the cluster represents the data-points well and is a good source of information. If inertia does not decrease by a lot when another cluster I added, it means that the cluster does not contribute much information about our data. In that case adding another cluster would only complicate our understanding. To decide on the number of clusters, I used the elbow-method. The elbow-method is a graphical method to visualize how much the inertia decreased by increasing k. The elbow point is called such, because the graph shows a sharp angle at the point at which increasing k does not decrease inertia to a large degree. The problem with this method is, that there is often not a clear elbow point. To circumvent this problem, I looked at how many counties fell into the smallest cluster – which is an additional method of understanding the value of an additional cluster – if only a few counties fall into a single cluster, there is not much to be gained by including that cluster.

After obtaining the clusters, I wanted to understand which venues the clusters are comprised of. I ordered the venue-categories according to the highest difference in frequency between the clusters, so I got a good picture what types of venues each cluster was characteristically comprised of. In the next step I used the python package folium[6] to visualize the clusters on a United States map, so I could get an understanding how the clusters are distributed spatially across the country.

In summary, I applied k-means clustering to help us understand the similarities and dissimilarities of counties with regards to their venue-composition. Each county was ascribed to a cluster. I generated a map which depicts the cluster each county belongs to on a map of the USA. In the later analyses I used the cluster as an independent variable to predict election outcome.

Exploratory Analyses of Independent Variables

I used bivariate correlations to analyze the relationships between all the independent variables: Median household income, unemployment rate, proportion of levels of education (no high-school, high-school, some college, college graduates), county population, mean-age, proportions of ethnicity (White-only, African American or mixed, Asian or mixed, Hispanic

[6] https://python-visualization.github.io/folium/

or mixed), and clusters. I focused more deeply on the demographic differences between the clusters.

Inferential Statistics: Predicting Election Outcome

My strategy to estimate the impact the venue-composition of counties had on the election results consisted of two major steps. First, I created a linear regression model with demographic variables as predictors and relative votes received by Trump as outcome. I calculated the explained variance, which shows us how much of the variation could be explained by demographic variables alone. In the next step I included the clusters as predictor and calculated the increase in explained variance. It is to be expected that the clusters are correlated with some of the demographic variables. I am particularly interested in the added explanation power the venue-data can provide in predicting election-results. Additionally, I looked at the regression coefficients of each predictor variable to estimate the magnitude of impact it had on the election results.

Predicting Election Outcome with Demographic Variables

I used a multiple linear regression model with demographic variables as independent variables and relative votes Trump received in a county as the dependent variable to estimate the total and relative effects of demographic variables on the election-results. The goal of this study was academic, which means I wanted to get a theoretical understanding of the magnitude of the effects variables had on election-results, instead of having the most precise prediction, as you would want, if you had to base an action or decision on your results. My independent variables needed to be preprocessed first to accomplish this goal. First, I standardized all the independent variables. After that I wanted to create one variable for education and one variable for socio-economic status. Up to this point I still had 4 different variables for education. I performed a principal component analysis to reduce the four dimensions to one and received on variable, which had a normal distribution, so it could be easily included in the regression model. To obtain one single variable for socio-economic status I calculated the mean of the standardized unemployment rates and median household incomes. First, I multiplied unemployment rates with -1, because unemployment is a negative indicator of socioeconomic status, while median household income is a positive indicator.

Before performing any machine-learning, I first calculated the bivariate correlations of demographic variables with relative Trump-votes. In the next step, I randomly split the dataset into a train-set consisting of 85% of counties and a test-set consisting of 15% of counties, using the train-test-split function from scikit-learn[7]. I used scikit-learn's [8]linear regression method to train and test the model. I then calculated the explained variance using R^2, a statistic depicting the total amount of variance in the dependent variable explained by all the dependent

[7] https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.train_test_split.html
[8] https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html

variables combined. I repeated the calculation of R^2 on the test-data. The second type of result I received where the regression coefficients for each independent variable. The regression coefficient can be interpreted as the change in relative Trump votes in a county, when the value of the predictor changes by one standard deviation, given that all other variables remain constant.

Estimating the Impact of Venue-Category Clusters on Election Outcome

In the next step, I included the cluster-variable as predictor in the regression model. I repeated the same analyses as mentioned above with the demographic variables. First I compared the differences in election-results between the clusters, calculating the average and total votes each candidate received in each cluster. To get a precise estimate of the added impact clusters had on election outcome, I used the same train- and test-split as before. The increase in R^2 is to be interpreted as the isolated explained variance of relative Trump-votes in a county by cluster alone. Again, I calculated the regression weights. However, this time, I used the entire dataset to train the linear regression model, to obtain the most precise estimates for regression-coefficients. The cluster regression coefficient can be interpreted as the increase or decrease in relative Trump votes in a county belonging to a cluster, compared to other clusters, while all other variables remain constant.

Results

Exploratory Analyses of the 2020 Presidential Election

In the popular vote President Trump received 74.0 million (47.7%) votes and President Biden received 81.1 million (52.3%) votes (see figure 1.). When we look at the results within counties (see figure 2), President Trump received 65.0% of the votes on average, while President Biden received an average of 33.3% of the votes. The large disparities between the total popular vote and the average popular vote per county are due to Trump having received more votes than Biden in 82.7% of the counties. However, most of the counties are very small and they more than got balanced by the large dominance President Biden had in the large cities.

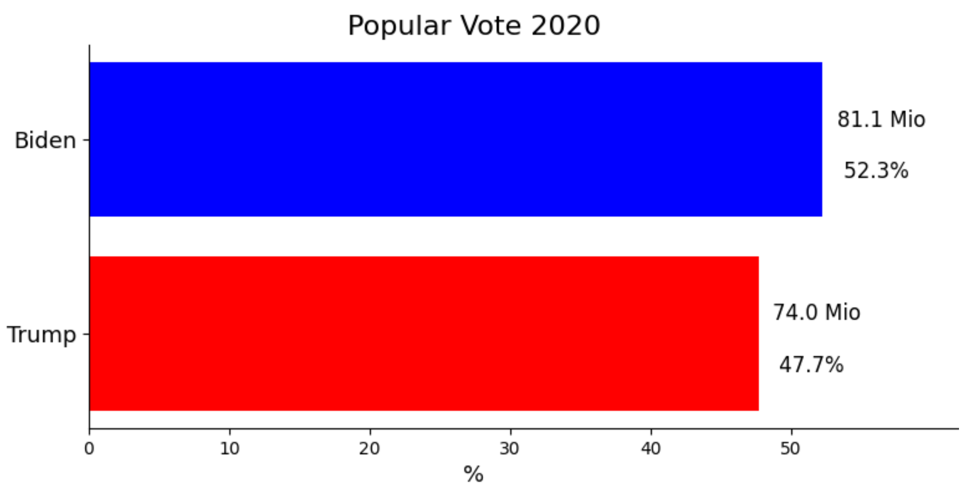

Figure 1.

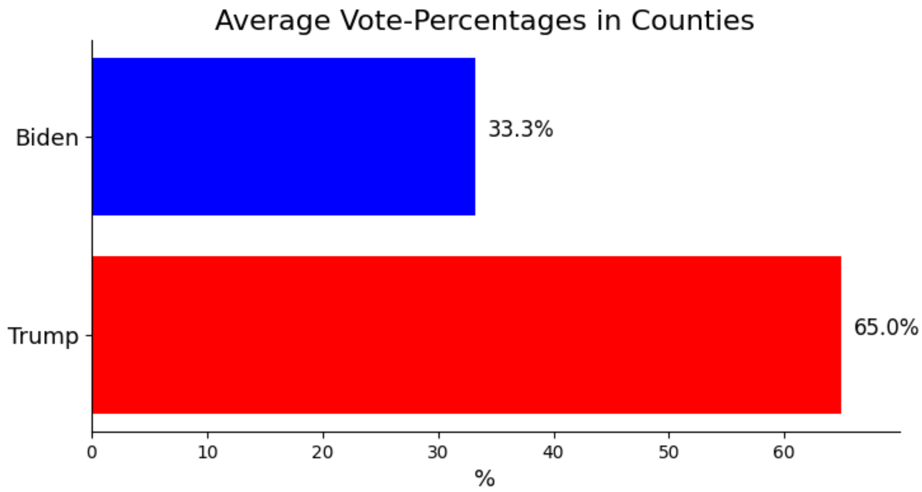

Figure 2. The percentages show the average relative votes Trump and Biden received in a county. The numbers might appear surprising; however, they only express the overwhelming dominance of Trump is small counties and Biden in big cities.

Clustering Counties by Venues

I successively fit 10 k-means clustering models with increasing k. Figure 3 shows the decrease in inertia with increasing k. There is only one elbow point at k=2, which means that the decrease in inertia is substantial between k=1 and k=2 and low with any additional added cluster. However, the angle at k=2 is not very sharp, so I decided to apply a second method to determine the number of clusters. Figure 4 depicts the number of counties which fall into the smallest clusters, depending on the total number of clusters. The smallest cluster rapidly declines in size when a third cluster is added, with only approximately 350 counties in the third cluster. One could very well add a third cluster to increase prediction accuracies, however I decided to use only the two large clusters for further analysis, because the third cluster would be a very specific small one, compared to the two large clusters. In the end, Cluster 0 entailed 1364 counties and cluster 1 entailed 1748 counties.

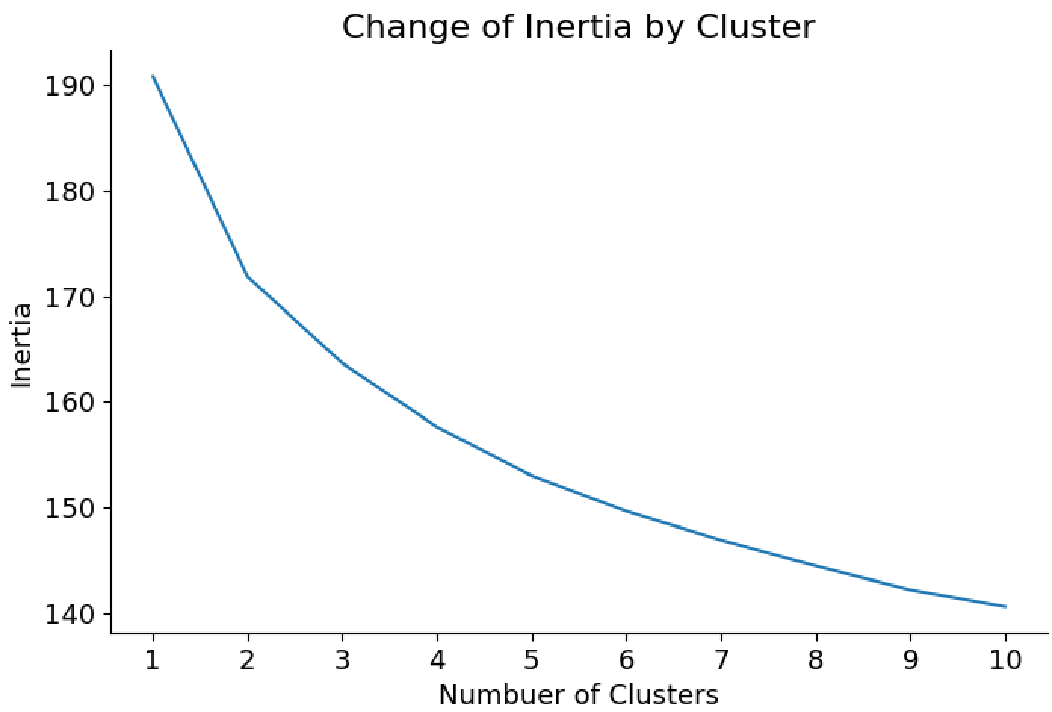

Figure 3

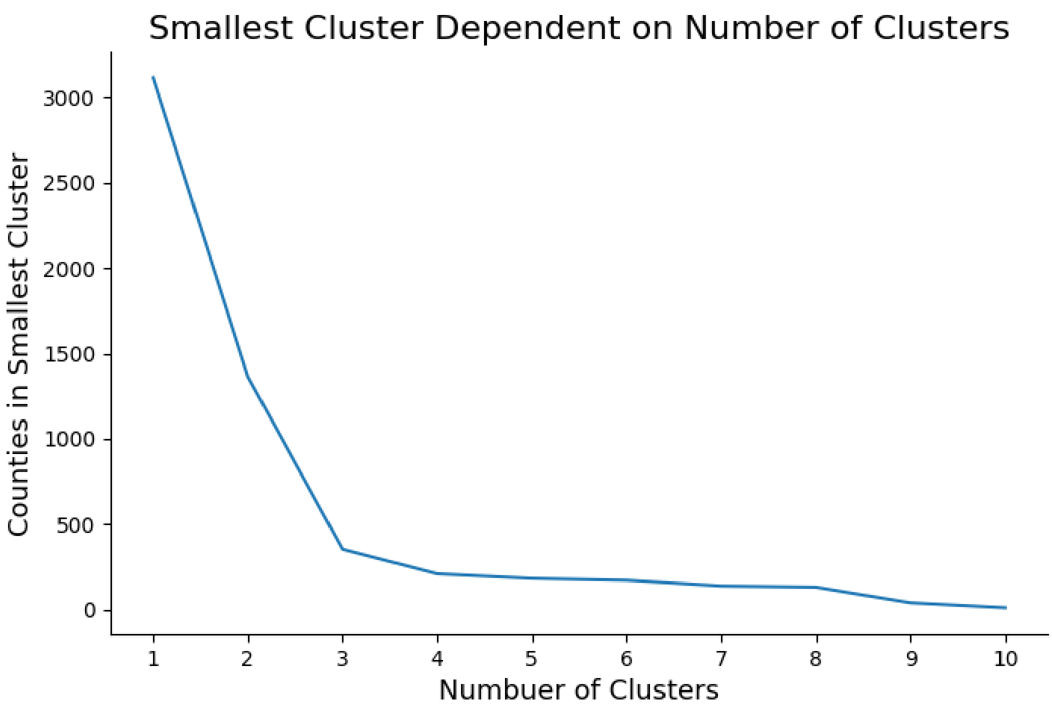

Figure 4

Table 1 shows the 20 venue categories with the largest differences in relative frequency between the clusters. Cluster 0 is characterized by higher relative frequencies of fast-food restaurants, discount stores, sandwich places and gas stations, while cluster 1 is characterized by higher relative frequencies in coffee shops, bars, American restaurants and breweries.

Table 1. Relative Frequency of Venue Categories in Clusters. These are the Top 20 venue categories with the highest difference in relative frequency between the clusters. The first 10 are more frequent in Cluster 0 and the last 10 more frequent in Cluster 1.

Venue Category	% in Cluster 0	% in Cluster 1	Difference in %
Fast Food Restaurant	12.7	2.7	10.0
Discount Store	11.0	2.1	8.9
Sandwich Place	7.5	3.5	4.0
Gas Station	5.7	2.0	3.7
Pizza Place	7.3	4.3	3.0
Pharmacy	2.4	1.0	1.4
Fried Chicken Joint	1.8	0.4	1.4
Big Box Store	1.7	0.3	1.4
Convenience Store	3.5	2.6	0.9
Video Store	0.6	0.3	0.3
Coffee Shop	0.8	2.9	2.1
Bar	0.6	2.5	1.9
American Restaurant	4.3	6.1	1.8
Brewery	0.2	1.6	1.4
Café	0.7	1.7	1.0
Park	0.7	1.6	0.9
Bakery	0.3	1.1	0.8
Italian Restaurant	0.6	1.4	0.8
Burger Joint	0.8	1.5	0.7
Trail	0.2	0.9	0.7

Figure 5 depicts a map of the United States. Each circle represents a county. The colors express the cluster a county belongs to. The green circles are counties belonging to cluster 0 and the yellow circles are counties belonging to cluster 1. Cluster 0 has a high density in the middle and eastern parts of the country, while cluster 1 has a high density in the western and northern parts, as well as at the coasts. Almost all counties at the eastern and western coasts fall into cluster 1.

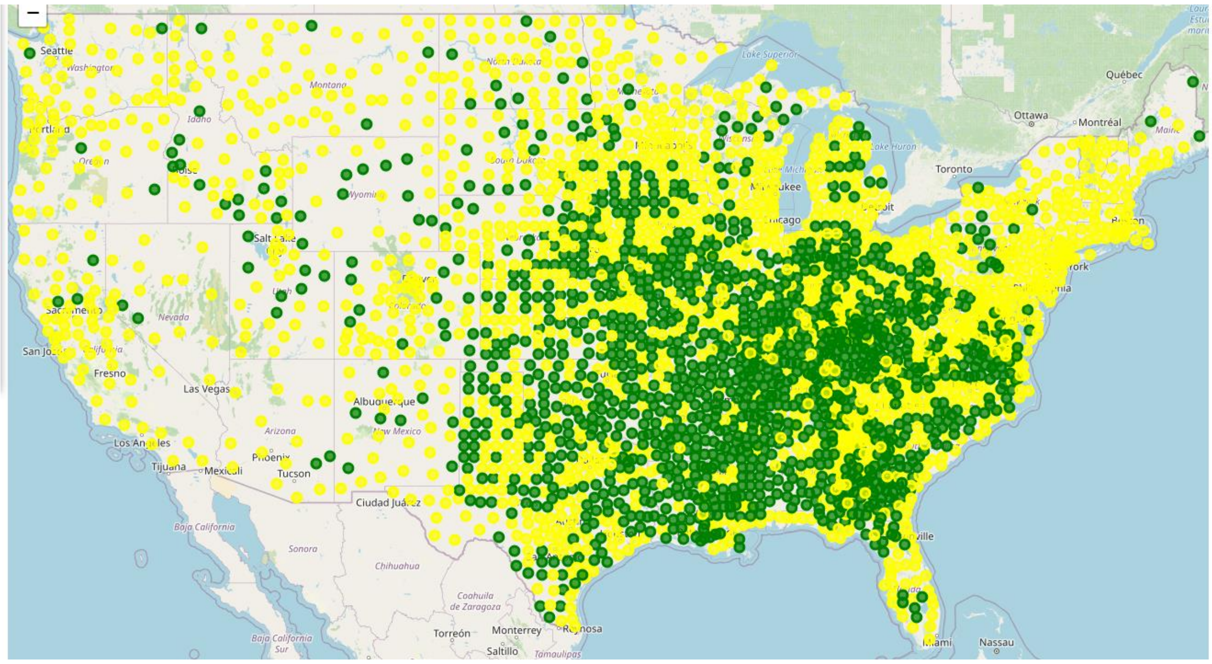

Figure 5. Counties mapped by clusters. Green: Cluster 0, Yellow: Cluster 1.

Exploratory Analyses of Independent Variables

Table 2 shows all bivariate Pearson correlations between the independent variables. Again, the unit of measurement is the county and not an individual voter, which means that for example age does not necessarily correlate negatively with having Asian descent, but that more people of Asian descent live in counties with a younger average population. The last row of the table shows us the correlation of cluster with demographic variables. A positive correlation means that cluster 1 is characterized by having larger values in the associated variable. The largest differences between the clusters could be found in education levels. Cluster 1 shows very significantly higher levels of education than cluster 0. Cluster 1 has larger populations than cluster 0. There is hardly any difference in the average age of the citizens. Cluster 1 also has a higher socio-economic status as indicated by a positive correlation of cluster with median income and a negative correlation with unemployment rate. When it comes to ethnicities, the clusters are not highly differentiated. The strongest relationship can be found in the proportion of people with Asian descent – there are more Asians in cluster 1 than in cluster 0. Cluster 0 has a higher proportion of African Americans and cluster 1 a larger proportion of Hispanics Cluster 1 has a marginally higher proportion of Caucasians (Whites) than cluster 0.

Table 2. Correlations of all independent variables.

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. unemployment rate												
2.median income	-.44											
3. no high school*	.37	-.53										
4. high school*	.25	-.54	.25									
5. some college*	-.15	.09	-.48	-.28								
6. bachelor or higher*	-.36	.72	-.60	-.77	-.02							
7. population	-.05	.28	-.05	-.32	-.10	.33						
8. age	.08	-.18	-.19	.27	.11	-.14	-.18					
9. Caucasian	-.27	.17	-.27	.12	.19	-.007	-.16	.28				
10. African American	.24	-.23	.32	-.008	-.24	-.07	.09	-.17	-.87			
11. Asian	-.14	.48	-.16	-.46	-.10	.51	.47	-.23	-.24	.04		
12. Hispanic	-.003	.04	.42	-.30	-.08	-.007	.20	-.32	.04	-.09	.17	
13. cluster	-.16	.39	-.37	-.43	.16	.49	.21	.03	.07	-.13	.26	.10

*Note: The variables indicating degree of education within a country show the proportion of inhabitants above the age of 25 with the maximum education as indicated by the name of the variable. In the later inferential analyses, the four variables where combined into one indicator of education by applying a principal component analysis on them.

Inferential Statistics: Predicting Election Outcome

Predicting Election Outcome with Demographic Variables

The first regression model was performed with all demographic variables, excluding the cluster variable. R^2 within the train-set reached .59 and .56 in the test-set, indicating that 56% of the variance of relative Trump votes between counties can be explained by all the demographic variables combined. Column model 1 in Table 3 depicts the regression coefficients. Education level had the strongest impact on the election results showing that a change of 1SD in education level was associated with a decrease of 5.4% in votes Trump received. The second largest impact was the proportion of counties' Caucasian population. A change in 1SD resulted in an increase of 4.9%-points for Trump. Increases in non-Caucasian population resulted in decrease in Trump votes between 2.7%-points (Hispanic) and 3.8%-points (African American). Socioeconomic status resulted in a 3.3%-points increase in Trump votes per SD. Compared to the correlation shown above, where income was negatively correlated with Trump-votes, socioeconomic status has a positive effect on Trump votes, when all the other variables remain constant. Population size had a marginal effect of a 1.2%-point decrease in Trump votes. The average age within a county hardly had any effect on the election outcome.

Estimating the Impact of Venue-Category Clusters on Election Outcome

In the second step, cluster was included as independent variable. The R^2 within the train-set was .60, which means that cluster explained 1.1% additional variance within the train-set. R^2 was .58 in the test-set, indicating 2.4% additional variance explained. This means that 2.4% of all the variation in Trump votes could be explained by the cluster a county belongs to, independently of all demographic variables. Column model 2 shows us all regression weights in the combined model (all variables including the cluster variables). The impact of most variables decreased slightly, which was to be expected, because cluster is correlated with most variables. Education level and proportion of Caucasian population were still the strongest predictors of election outcome. The magnitude of the cluster regression coefficient cannot be properly compared with the others because the scale is different. However, we can see that President Trump had a 4.3%-points higher expected relative vote in cluster 0 compared to cluster 1, when all other variables remain constant.

Table 3. Regression coefficients and explained variance of multiple linear regression.

	regression coefficient	
variable	model 1: demographic only	model 2: including cluster
socioeconomic status	3.3	3.0
education level	-5.4	-4.7
population	-1.2	-1.3
age	-0.3	-0.2
Caucasian	4.9	4.7
African American	-3.8	-4.1
Asian	-3.1	-2.6
Hispanic	-2.7	-2.3
cluster		-4.3
R^{2*}	0.561	0.584

The independent variables have all been standardized, except for the cluster variable. The dependent variable is the relative votes Trump received in a county in %. The regression coefficients of demographic variables can be interpreted as the expected increase in Trump votes (in %) if the independent variable changes by 1 SD, while all other variables remain constant. The regression coefficient of the cluster variable can be interpreted as the expected increase (decrease!) in Trump votes in cluster 1 compared to cluster 0, given that all other variables remain constant.
*Note: The value of R^2 was calculated as the explained variance of relative Trump votes within the test-set.

Discussion

Counties, Venues, and Election Results

The results show us that the composition of venues within a county had a unique capacity to predict the 2020 election results. After keeping all demographic variables, it was shown that the cluster increased or decreased the percentage points a candidate received by 4.3%-points. This is not as low a number as it might appear at first sight. First, we know that swing states are decided within single digit percentage points. Secondly, a 4.3%-point increase for one candidate means a 4.3%-point decrease for the other candidate, which leads to a total difference of 8.6%-points. This is an exceptionally large difference which changes election outcomes very significantly. Additionally, it must be noted that my work was more academic in the sense of wanting to understand the impact of venue-composition on election-outcomes substantially, instead of simply wanting to maximize predictive capability of my model. A cluster analysis of course decreases the amount of data used to predict an outcome, so the actual impact of venue-composition on election-results is most likely even higher.

All counties in the United States can be roughly divided into two categories by their venues. One category is characterized by fast-food restaurants, discount stores, sandwich places, and gas stations and the other category is characterized by coffee shops, bars, American restaurants, and breweries. I can only roughly speculate what the common characteristics of these venues are. The first cluster appears to be simpler, more traditional American, maybe working class and probably more rural, while the second cluster appears to be more urban, hip, maybe touristy and socially progressive. We get a more in-depth understanding about the clusters, when we look at how they are distributed across America and how they are made up when it comes to their demographic composition. The first cluster has a high density in middle eastern and southern parts of the nation – parts which do fall into the category of traditional America. The second cluster has a large density at the coastal regions, in New England and western parts of the country. Especially the coastal regions are known to be more progressive, less traditional and intuitively we know they are Democrat strongholds. When we look at the sociodemographic composition of the two clusters, we find that education is what differentiates the clusters the most. The first cluster, which I characterized as more traditional-American has way lower education-levels than the second cluster, which I characterized as more progressive. Of course, higher education-levels are associated with higher income, so it is hardly surprising that the second cluster is wealthier than the first. Also as is probably widely known, traditional America can be found more in the rural areas, with lower population sizes, which can be seen in our clusters as well. The first cluster has lower population than the second cluster. When it comes to ethnicity, especially the differences in Asian population stand out, while the differences in African American or Caucasian population are not as large as could be expected. When we explore the relationships of counties with substantial Asian populations, we find that those counties also tend to be richer, larger and have a higher education standard.

The Impact of Education and Importance of Culture.

Even though it was not the planned topic of this study, nevertheless I was astonished by the immense impact education-levels had on the election results. Education is the strongest predictor of all. Its impact is so large, as to even reverse the otherwise positive effect income would have on voting for Trump. As we can observe when comparing correlations and regression coefficients, wealthier counties tended to vote for Biden, but when the effect of other variables (and the strongest is education) is controlled for, counties with higher socioeconomic status voted for Trump. If we believe that voting for Trump or the Republicans can be considered beneficial for the wealthy, education reverses the effects so that now wealthier counties vote for Biden, maybe even against their own economic self-interest. Depending on which candidate or party you prefer, this might be a major accomplishment of Democrats and Progressives to promote their political views, or as a large failure of Republicans and Conservatives to reach the educated with their message. When we take the differences between the clusters seriously and apply my interpretation that one cluster represents traditional America and the other progressive America, we can hypothesize, that the culture of the Nation might be systematically changing through academia, which is in accord with what the famous Professor of Psychology, Jordan B. Peterson highlighted. It is even now a well-known fact, that most professors – at least in the social sciences, vote Democrat and promote progressive views. From a conservative viewpoint it must be alarming that the so-called culture war has not been properly fought and might be the reason for present-day and future losses in elections.

Recommendations

The data shows us that there is a true division going through the counties of the United States. There appear to be two different cultures, which not only show themselves in the demographics, but also in the venue-composition of the counties. One culture can be termed traditional or conservative and the other progressive. When we look at the impact education has on the election outcome, my suggestion to Republicans for the future is to engage in the culture. Try acquiring prominent positions in colleges and universities, or even develop an alternative conservative academic infrastructure. Also engage in the "culture-war" through producing high quality entertainment with conservative messages. Try conquering back the urban highly educated counties by a suited intelligent message, which can compete with progressive ideas. To Democrats I recommend to truly engage in policies that are fitted to strengthening the working-class, which do not necessarily have a college degree. I think the Democratic party somewhat abandoned its traditional focus group of the working-class and too much engaged with intellectual progressive ideas. Social justice does no longer appear to mean equality of opportunity and care for the less fortunate, but an ideology of groupthink and victim-culture.

Conclusion

President Biden in his inaugural speech, on January 20th, 2021, stressed the importance of unity and his ambition of uniting the country. When we look at the data, our analyses provided us with, the country does indeed appear to be divided into two nations. Even when we look at the differences in venues that are characteristic for a county, theses differences show to have a significant political impact, even as to influence election results to a large enough extent as to be decisive. The differences appear to not be racial, or economic, as a Democrat's worldview might suggest, but seem to be deeply cultural. I think unity can only be achieved through actual understanding, and this would mean to understand traditional America for a Democrat and to understand culturally progressive America as a Republican.