FORGOTTEN AMERICA

How a Conflation of Social, Economic, and Health Challenges are Crippling Eastern Kentucky

Authors:
Samuel C. Arneson
Brandon M. Lin
Jelena Mitrovic
Landon L. Wilson

October 2, 2021

Submitted to:
2021 Carolina Data Challenge

Table of Contents

Introduction

By many measures, Kentucky is one of the hardest places to live in the United States. 54% of the counties in Kentucky are considered to be in persistent poverty (tied for second worst in the US), with over 20% of their respective populations having lived below the poverty line for the last 30 years. The economic challenges residents of Kentucky face are only exacerbated by the social vulnerabilities that plague their communities, especially in the eastern half of the state. Using social vulnerability data from the Center for Disease Control, MSA TEAM 2 looked at the overall social vulnerability and resilience scores of census tracts to identify particular tracts of Kentucky that are harder hit than other regions of the state. Further research into access to education, food, and healthcare as well as prevalence of agriculture and infrastructure throughout Kentucky reinforced a troubling pattern: eastern Kentucky lags behind the rest of the state in almost every measure of social and economic vulnerability.

Methodology & Analysis

Census Tract Clustering

MSA TEAM 2 used k-means clustering to identify tracts within Kentucky that had similar social vulnerabilities. Table 1 lists the features that were used to create the clusters.

Table 1: Social Vulnerability Variables used for Census Tract Clustering

Variables	
Percent of Population Living in Poverty	Percentage Minority
Percent of Population 25+ Years of Age without a High School Diploma	Percent of Population over 5 Years Old who Speak English “Less than Well”
Percent of Population under 17 Years Old	Percent of Population over 65 Years Old
Percentage Unemployment	Percent of Mobile Homes
Percent of Housing Structures with Ten or More Units	Percent of Housing Units with More People than Rooms
Per Capita Income	Percent of Households with No Vehicles
Percent of Single Parent Households with Children under 18	Percent of Population in Institutionalized Group Quarters
Percent of Population with Disability	-

Logistic Regression

In order to get a better understanding of how social vulnerability was associated with other challenges, MSA TEAM 2 created a logistic regression model to predict the probability that a given census tract would be considered a low-access tract at ten miles (LA-10 tract). According to the United States Department of

Agriculture, an LA-10 tract is a rural census tract in which at least 500 residents (or at least 33% of the population) live greater than 10 miles from the nearest supermarket or grocery store.

The social vulnerability factors that were ultimately significant in predicting a census tract's status as an LA-10 tract are listed in Table 2, ordered by their significance.

Table 2: Social Vulnerability Variables Significant in Predicting a Low-Access Tract at Ten Miles

Variable	**Significance**
Percent of Housing Structures with Ten or More Units	4.84E-124
Percent of Mobile Homes	6.74E-81
Per Capita Income	1.18E-21
Percentage Minority	2.21E-21
Percent of Single Parent Households with Children under 18	2.14E-12
Percent of Housing Units with More People than Rooms	1.39E-05
Percent of Population over 65 Years Old	2.76E-04

The two most significant variables in Table 2 relate to housing infrastructure and give an indication of the degree to which a given census tract may be considered rural. The remaining variables provide color to the demographics of the population, including their relative wealth, the racial diversity of the population, and the age of the population.

Results

Census Tract Clustering

The characteristics of the clusters produced by the k-means clustering algorithm are listed in Table 3. Figure 1 overlays each cluster onto Kentucky's census tracts.

Table 3: Characteristics of Census Tract Clusters

Cluster	**Location**	**Affluence**	**Vulnerability**	**Disability**
Yellow	Urban	High	Low	Low
Red	Urban	Moderate	Low	Low
White	Suburban	High	Low	Low
Light Blue	Rural	Moderate	Moderate	Moderate
Dark Blue	Rural	Low	High	High

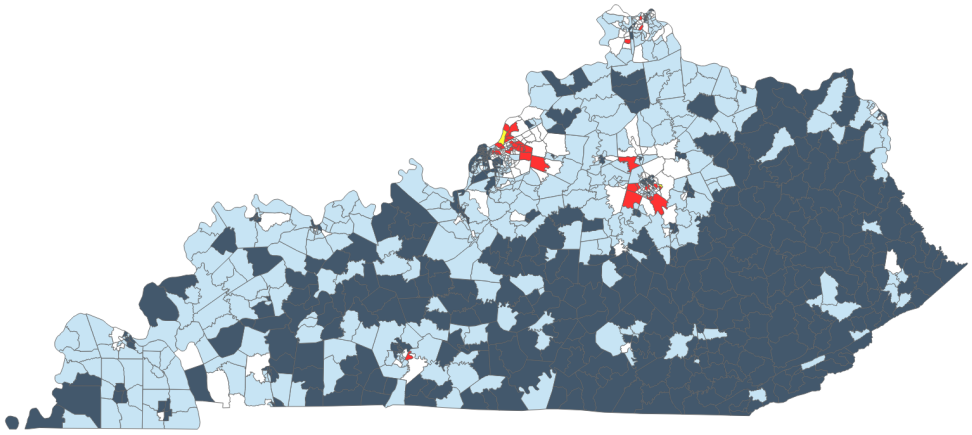

Figure 1: Kentucky Census Tracts Colored by Social Vulnerability Cluster

Maps

In addition to comparing social vulnerability characteristics across census tracts, MSA TEAM 2 analyzed the relative resiliency of each region as measured by National Risk Index data from the Federal Emergency Management Agency (FEMA). Resiliency, a measurement of a community's ability to prepare for, adapt to, withstand, and recover from natural disasters, is illustrated for Kentucky census tracts in Figure 2.

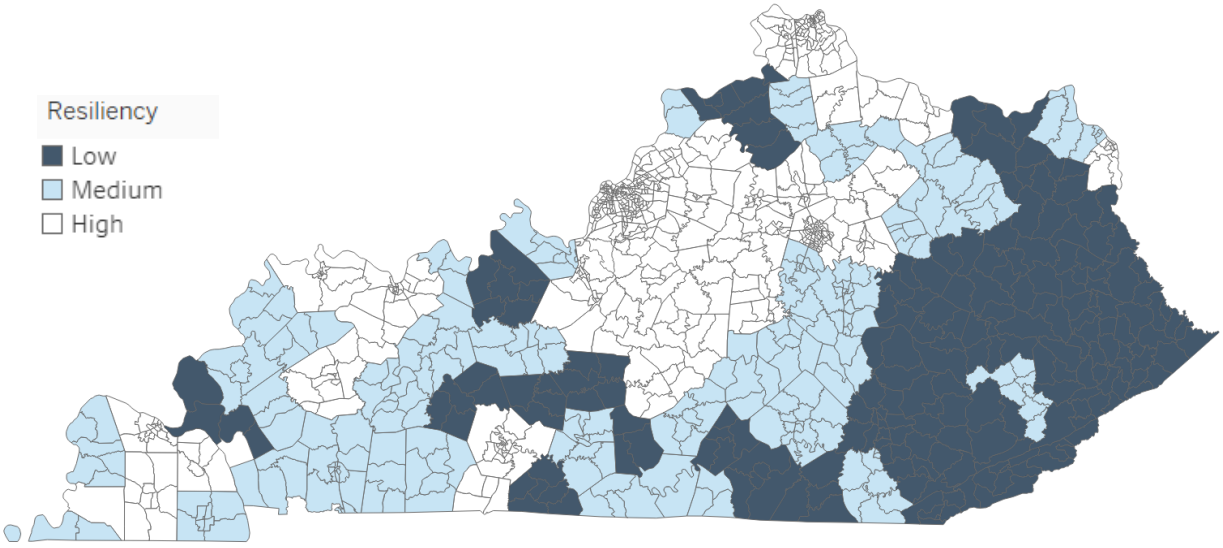

Figure 2: Census Tract Resiliency in Kentucky

Figure 2 demonstrates that a large proportion of communities in eastern Kentucky, already plagued by social and economic vulnerability, also have minimal resilience in the face of natural hazards.

MSA TEAM 2 also evaluated the value of agriculture and building infrastructure across communities in Kentucky as proxies of potential economic output. Figure 3 shows the bivariate relationship between

agriculture and building value in each census tract, as measured by FEMA's National Risk Index data. Regions shaded in dark blue suffer from a combination of low agricultural output and low building value.

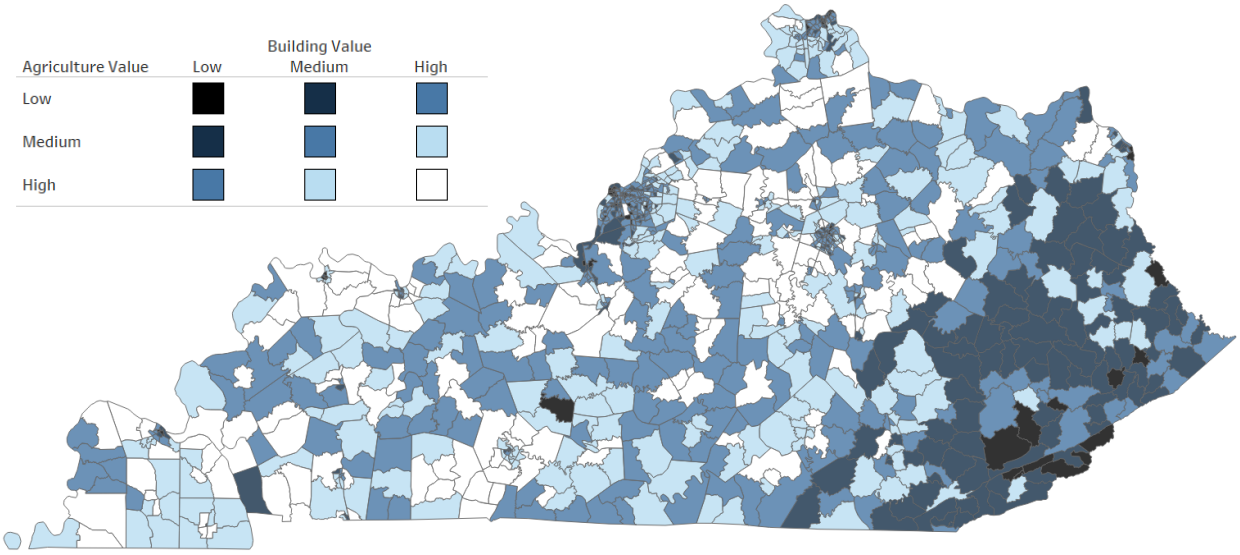

Figure 3: Relationship between Agriculture and Building Value of Kentucky Census Tracts

In Figure 3, the volume of census tracts in eastern Kentucky that have both low agriculture and building value suggests that these communities have minimal opportunities for economic advancement. Without valuable infrastructure or a means of agricultural production, it is no surprise that many of these communities are considered to be in persistent poverty.

Logistic Regression Predictions

The logistic regression model built to predict LA-10 tracts based on social vulnerability scores had a concordance percentage of around 88.29%, correctly ranking true LA-10 tracts above non-LA-10 tracts 88% of the time. Table 4 displays the odds ratio percentages of selected explanatory variables from the logistic regression model.

Table 4: Odds Ratio Percentages of Select Variables from Logistic Regression Model

Variable	**Odds Ratio Percent**
Per Capita Income, between $30,000 and $75,000	-61.76
Per Capita Income, above $75,000	-94.10
Percentage Minority, between 20% and 50%	-54.61
Percentage Minority, between 50% and 70%	-67.61
Percentage Minority, above 70%	-73.12
Percent of Housing Units with More Rooms than People	5.58

Percent of Population over 65 Years Old, above 25%	89.59
Percent of Mobile Homes, between 10% and 25%	377.19
Percent of Mobile Homes, between 25% and 57%	605.54

The odds ratio percentages in Table 4 paint a vivid picture of the communities that lack access to grocery stores and supermarkets. Areas with per capita income above $75,000 are 94% less likely to be a LA-10 tract than communities with a per capita income below $20,000. Rural, poor, white communities with families packed in mobile homes and a high percentage of the population above 65 years old comprise the tracts that are mostly likely to be classified as having low access to reliable food sources within 10 miles. Figure 4 shows the predicted probability of a region in Kentucky being an LA-10 tract based on the conflation of these attributes.

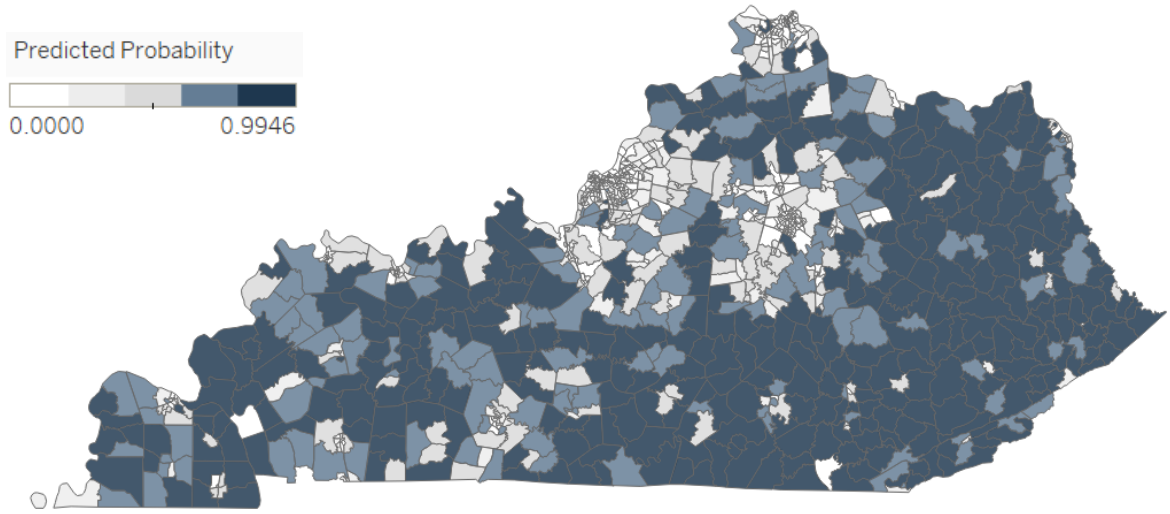

Figure 4: Predicted Probability of being an LA-10 Tract Based on Social Vulnerability

Figure 4 demonstrates a concentration of regions with a high probability of being an LA-10 tract throughout eastern Kentucky. While some of these communities may not currently qualify as LA-10 tracts, their social vulnerability attributes suggest they may be likely to become one soon. In total, over 1.5 million Kentucky residents live in census tracts that have a predicted probability of being an LA-10 tract greater than or equal to 80%, colored in Figure 4 is dark blue.

Recommendations

While the patterns of social and economic vulnerability in eastern Kentucky paint a troubling picture, further research into the causes and solutions is needed. Specifically, MSA TEAM 2 recommends expanding analysis to the following:

- Historical, political, and institutional factors that have contributed to the high rates of vulnerability throughout eastern Kentucky
- Federal and state financial support given to Kentucky census tracts over the past several decades in comparison with funds given to similar tracts in other states
- Allocation of funds spent on infrastructure, healthcare, food, and education throughout Kentucky
- Census tracts in the same clusters as those in eastern Kentucky that have been able to make economic and social improvements for their populations

Conclusion

Through clustering, MSA TEAM 2 identified the communities with the highest risks of social vulnerability and lowest resilience in the face of environmental and economic disasters. The logistic regression illustrated how these social vulnerabilities correlate with a community's likelihood to have limited access to a reliable food source. Lastly, visualizing access to education, healthcare, agriculture, and valuable infrastructure reinforced the multitude of issues that residents of Kentucky face.

Ultimately, the patterns of social vulnerability in communities that are also beset by limited access to healthcare, food, and education creates a disturbing outlook for hundreds of thousands of eastern Kentucky residents. Understanding the social and economic forces that plague these communities is a key first step to alleviating the problems they face.