Brandon Weaver

Joseph Shifman

Ethan Nguyen

Yafet Mekonnen

7 November 2021

Combating Food Insecurity Technical Details

All code for this project was written in R and ran in RStudio. R libraries used include readxl, the Tidyverse package, usdata, and geofacet. We used R's built-in correlation function using the Pearson correlation coefficient to find features that highly correlate with a tract having the "LA1and10" flag.

We included data from the National Household Food Acquisition and Purchase Survey(3). Primary grocery store travel mode was categorized into four buckets: drives own car, drives someone else's car, someone else drives me, and other forms of transportation. The final factor, "other forms of transportation", includes walking, biking, and using public or paid transportation. To compare census tracts and vehicle access, we turned the poverty rate of the tract into a factor dividing on the quarter percentiles.

In order to explore how education levels correlate with the likelihood of low food access, a linear model was made based on the original challenge dataset and an educational dataset published by USDA ERS. The educational dataset is organized by county so we had to group the original dataset by county and find the total number of tracts within that county with the "LA1and10" flag. The two datasets were joined on a column representing concatenation of the county name and the state name. The data was split 60:20:20 for training, validation and testing. The final statistical results were that the model had a p-value of about zero which means that we were able to reject the null hypothesis. The adjusted r^2 value of the model was 0.7899, meaning there is about a 79% chance that the different education levels are correlated with the number of LA tracts in a county.

To explore the distribution of ages, tracts were grouped by the "LA1and10" flag to identify potential differences in the distribution of low access tracts and tracts without that classification. Running two two-sample t-tests (on mean proportion of seniors and mean proportion of kids), at an alpha of 0.05, resulted in both rejecting the null hypothesis. Though the result was statistically significant, the matching confidence interval was relatively small; we are 95% confident that the true population difference between low access tracts and non-low access tracts proportion of seniors is between 0.657% - 0.886%, and the difference in the proportion of kids (at that same confidence level) is between 0.984% - 01.176%.

To examine the race distribution we looked at the seven ethnic groups that have low access to a supermarket. We took the sum of each ethnic group per state and also the total count of a specific race and a percentage was calculated. For example, for the white race the formula would be percent_White= ((White / Total_White) * 100) where White represents sum(as.integer(lawhite10)) and Total_White represents sum(TractWhite). We then used the pivot_longer function to get the result for each state, with seven rows containing each race. Geofacet was used to create a horizontal bar chart and display the geographical distribution according to state.

Works Cited

1.“County-Level Data Sets Download Data.” *Educational attainment for the U.S., States, and counties, 1970-2019*, Economic Research Service U.S. DEPARTMENT OF AGRICULTURE, 24 Feb. 2021, https://www.ers.usda.gov/data-products/county-level-data-sets/download-data/.

2.“Definitions of Food Security.” *Www.ers.usda.gov*, Economic Research Service U.S. DEPARTMENT OF AGRICULTURE, 8 Sept. 2021, https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/definitions-of-food-security.aspx.

3.“FoodAPS National Household Food Acquisition and Purchase Survey.” *USDA ERS - FoodAPS National Household Food Acquisition and Purchase Survey*, Economic Research Service U.S. DEPARTMENT OF AGRICULTURE, 28 Apr. 2021, https://www.ers.usda.gov/data-products/foodaps-national-household-food-acquisition-and-purchase-survey/.

4.Haskell, Scott. “Food Insecurity and Food Deserts: How Are They Related?” *Institute for Food Laws and Regulations*, Michigan State University, 11 Feb. 2021, https://www.canr.msu.edu/news/food-insecurity-and-food-deserts-how-are-they-related.

5. “This Is Statistics Fall Data Challenge 2021 Dataset.” *Notion*, 13 Oct. 2021, https://zzlalo.notion.site/This-is-Statistics-Fall-Data-Challenge-2021-Dataset-2b24d244f1fd4b678b4718416e94ddb5.