**Analysis of:**

**Supermarket Locations**

**Within United States Census Tracts**

Andrew Wang

CIS 320 Data Analytics

Professor Thomas

March 2, 2016

**Table of Contents**

|  |  |
| --- | --- |
| Introduction | 2 |
| Method | 2 |
| Hypothesis | 3 |
| Selected Data Fields | 3 |
| Uploading Data | 4 |
| Bar Graph | 5 |
| Rural Versus Urban | 6 |
| Conclusion | 7 |

**Introduction**

Inability to access healthy food is a serious problem in the United States. The task of finding healthy and affordable food is difficult and challenging, forcing many Americans away from a healthy diet. There are multiple reasons for people to be unable to reach healthy foods, such as: distance of stores or number of stores in an area, household income, or vehicle access and available transportation.

Food Access Research Atlas was created with ArcGIS Server, a geographic information system built with Environmental Systems Research Inc. to gather data on topographic and satellite maps. The Food Desert Locator looks for low income areas where supermarkets are categorized by distance. Urban locations consider half and 1 mile to be far away while rural locations will see 10 and 20 miles as far away. Vehicles are a common way to reach food sources in the United States, and a way to identify access to supermarkets is with availability to motor vehicles.

The data use in this atlas, <https://catalog.data.gov/dataset/food-access-research-atlas>, is based on information collected in 2010 from supermarkets, Decennial Census, and the American Community Survey. To be considered a food desert for this dataset, the census tracts only looks at locations with low income and low access.

**Methods**

The method techniques measured half kilometer square grids to determine census tracts where there were super markets (USDA ERS). Another important factor to look at is the difference between urban and rural locations. Urban locations measured sections that said 1 mile from supermarkets were close and rural locations measured 10 miles and said those were close by.

**Food Access Research**

**Hypothesis**

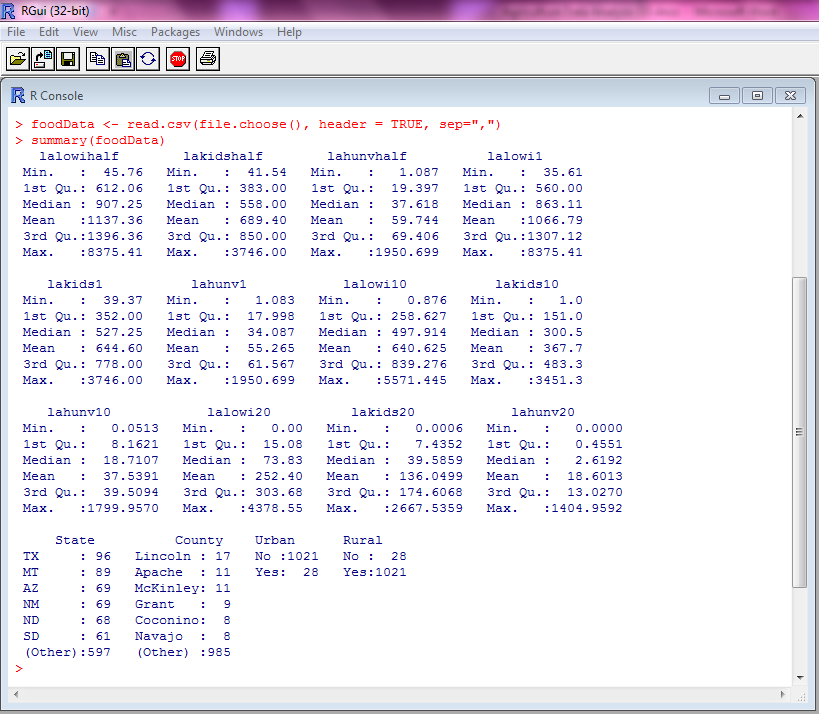
* Are there more people in low income households with tracts that have vehicle access?

**Selected Data Fields**

The original data set was too large and needed to be cut down to size before I began my analysis. I felt these entries would be able to compare to each other and provide insight to my question.

* State: What state in the US the tract is part of.
* County: What county the tract is part of.
* Urban: This field tells me if the county where the census took place was a city.
* Rural: This field tells me if the county where the census took place was a countryside.
* lakidshalf: Low access children within half a mile of a supermarket.
* lakids1: Low access children within 1 mile of a supermarket.
* lakids10: Low access children within 10 miles of a supermarket.
* lakids20: Low access children within 20 miles of a supermarket.
* lalowihalf: Low access people within half a mile of a supermarket
* lalowi1: Low access people within 1 mile of a supermarket.
* lalowi10: Low access people within 10 miles of a supermarket.
* lalowi20: Low access people within 20 miles of a supermarket.
* lahunvhalf: None and low access vehicle access at ½ mile.
* lahunv1 : None and low access vehicle access at 1 mile.
* lahunv10: None and low access vehicle access at 10 miles.
* lahunv20: None and low access vehicle access at 20 miles.

**Uploading Data**



Here is a summary of the data set that has been cleaned and uploaded into R.

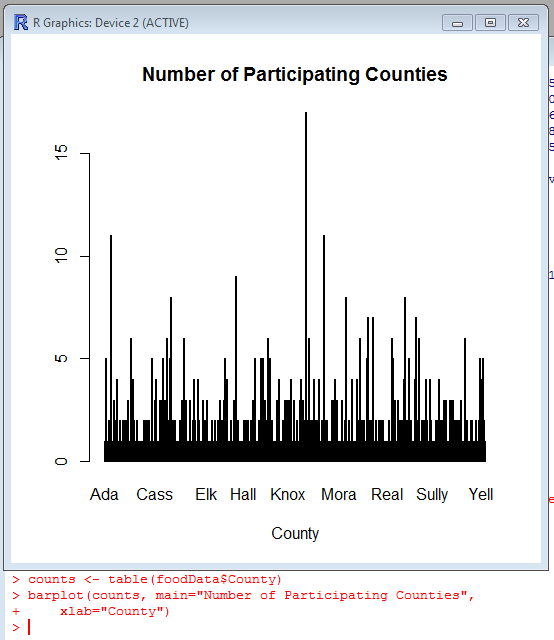
foodData <- read.csv(file.choose(), header = TRUE, sep=",")

This line of code was used to pull the file from my computer.

summary(foodData)

This line of code displays the data in terms of minimum, maximum, median, and quartiles.

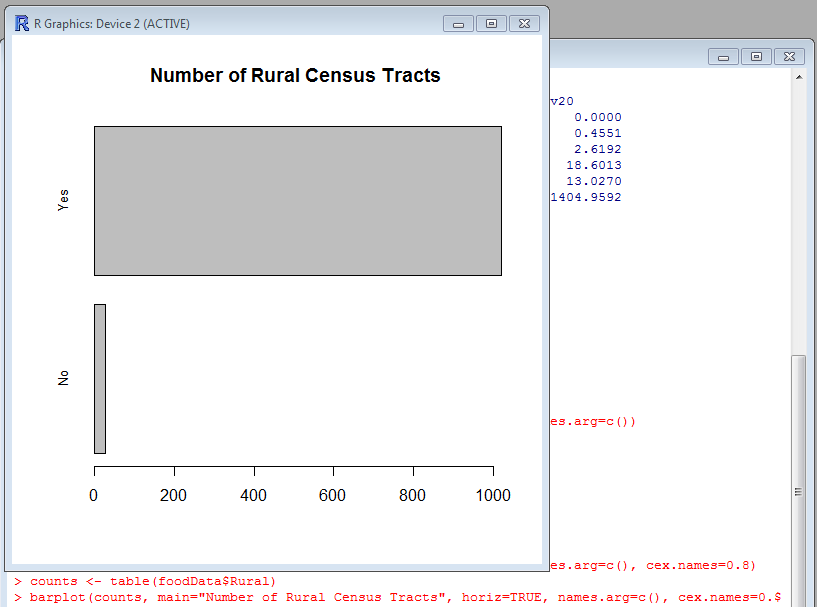
**Bar Graph**



Here is a bar graph showing the many different counties and how often they show up on the census. The number of times they show up tells us that there are more tracts within these counties that have low access or low income households.

There are far too many occurrences of counties to show all of their names, but we can tell that the highest reaching bar belongs to a county named Lincoln as the summary tells us there are 17 instances recorded.

**Rural Versus Urban**



This graph here tells us the amount of rural census tracts that were in the dataset. From our summary, we can tell that the exact number is 1021. The remaning 28 tracts belong to the urban locations. From this, we can tell that most of the low income and low access families are overwhelmingly likely to live in rural locations, far away from the influence of cities and their supermarkets. It might be possible that lack of immediate profit could be the reason for the few supermarkets in rural locations

**Conclusion**

From the data that we have gathered, are we able to answer the question: “Are there more people in low income households with tracts that have vehicle access?”

We can see that a majority of people reside in the Lincoln county and that most census tracts are in rural locations. Looking at the data summary, we can tell that as the miles from supermarkets increase, we see that the amount of vehicles in each location decreases. We can also see that as the distance from supermarkets increase, the number of people that live within those tracts decrease. From this, we can see that there is s correlation that low income families also have less access to vehicles, resulting in a greater difficulty to reach sources of healthy food.

**Bibiography**

Ploeg, Michele V., and Dutko Paula. "USDA ERS - Food Access Research Atlas: About the Atlas." *USDA ERS - Food Access Research Atlas: About the Atlas*. USDA ERS, 1 Mar. 2013. Web. 01 Mar. 2016.