

A silhouette of a person stands on a rocky path, looking out over a vast mountain range under a hazy, sunset-colored sky. The person is positioned in the center-left of the frame, facing away from the viewer. The mountains in the distance are layered and shrouded in a soft, golden light. The overall mood is contemplative and serene.

FOOD DESERT

By Sagar Jain
-Rutgers University (MSDS)

What is Food Desert?



Food Desert!!

- Area where healthy and affordable food is difficult to obtain.
- Mainly those areas with low income and less accessibility are affected.
- Food Desert Theory: There is a direct relationship between a population's access to fresh food and health.

Consequences:

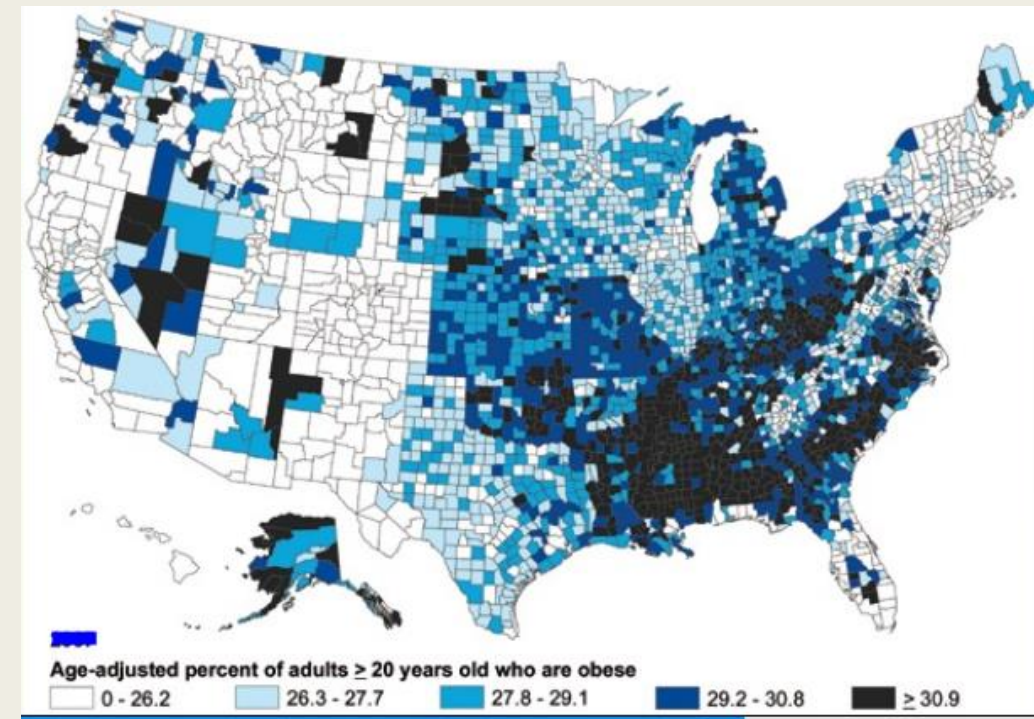
- Food Insecurities.
- Excessive Junk Food Consumption.
- Health problems:
 - **Diabetes**
 - **Cardiovascular, Neurological Problems**
 - **Kidney issues**
 - **Gout**
 - **Obesity**
- Child malnourishment.
- Origin of Food Swamps:
 - Stores high sodium content items
 - No Vegetables
 - Canned Food at inflated prices
- 16,000 residents of New Hanover County are affected by this.



Reasons:

- Mostly people living in these areas have low fixed income.
- Only 67% of the people owns the vehicle.
- Grocery stores and SuperMarkets are scarce as they fear generating less revenue.
- Community Farms are uncommon.
- Less transportation facilities available.





If we will compare the two maps, the obesity rates are higher in areas where there is no grocery stores for atleast a mile.

Solution:

- Fortunately we have a great community which is working together to overcome this problem.
- Upfront organizations like
 - CapeFearCollective,
 - Feast Down East(FDE),
 - NourishNC are working to address this issue from ground level.



Feast Down East:

- Bringing the mobile grocery Store to the people through their experience in food hop operation.
- Targeting the food desert communities.
- Working 4 days a week.
- Accepting Food Stamps through their dollar to dollar matching program which is proved to be very effective and instrumental. 1EBT = 2\$



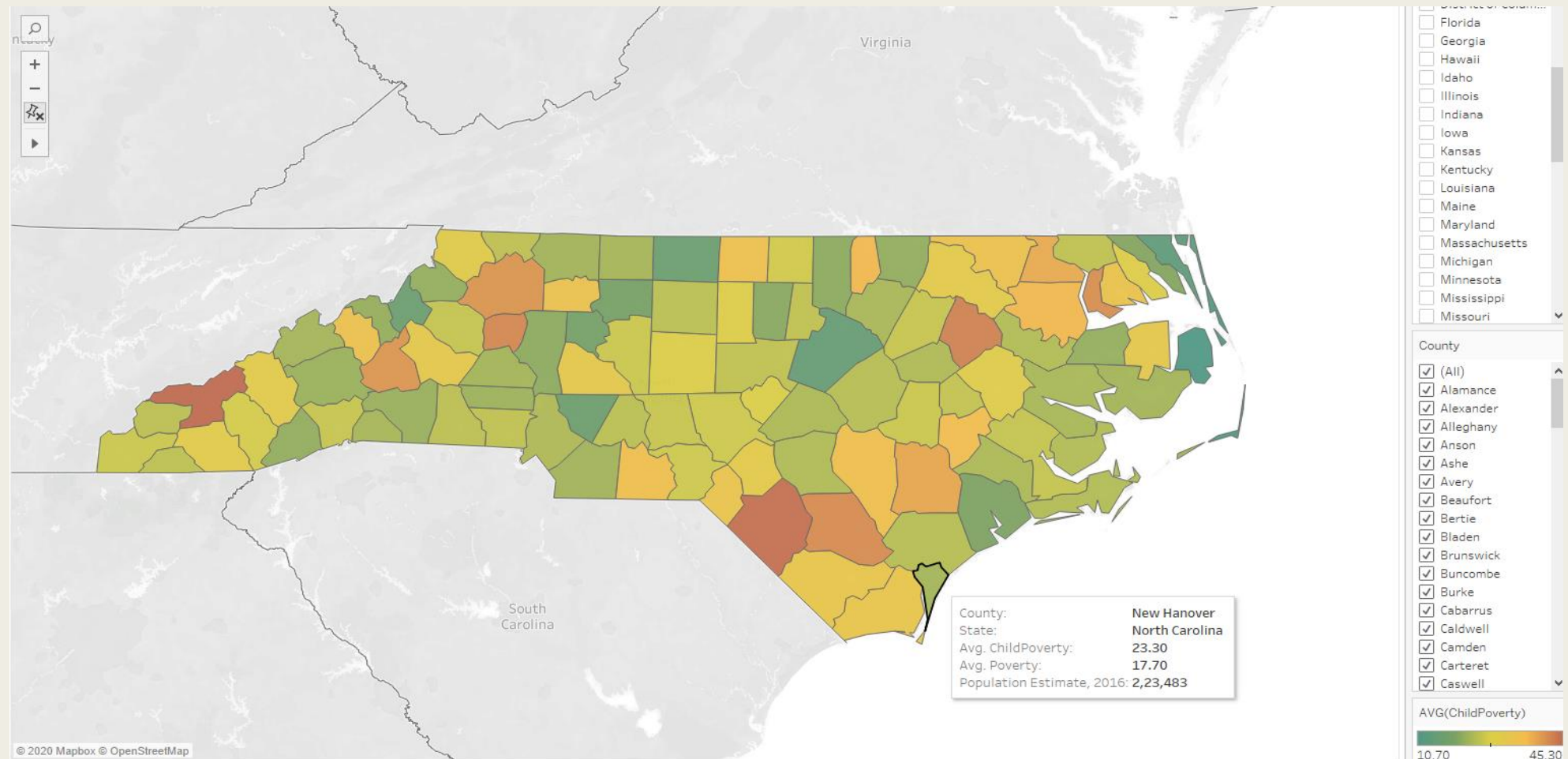
Nourish NC:

- Back Pack program for Kids.
- Over 1100 kids are being taken care by this program.
- Maintaining Balance with the mission to provide food with dignity.
- Open convenient stores for people suffering from hunger. No ID's required.
- Food pharmacy program.

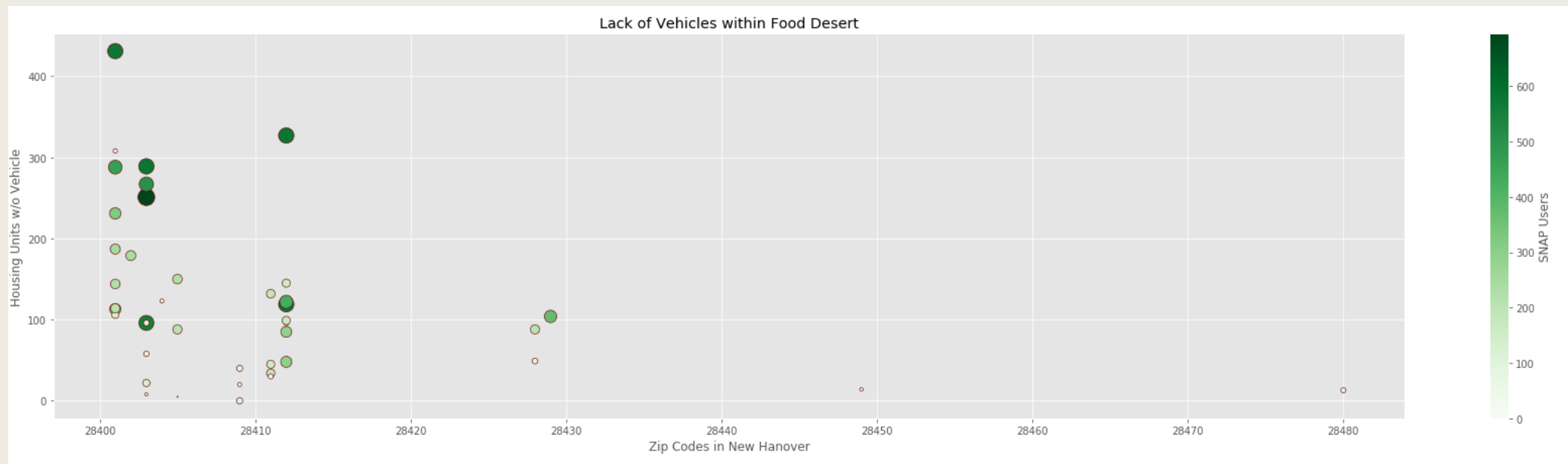


Analytics

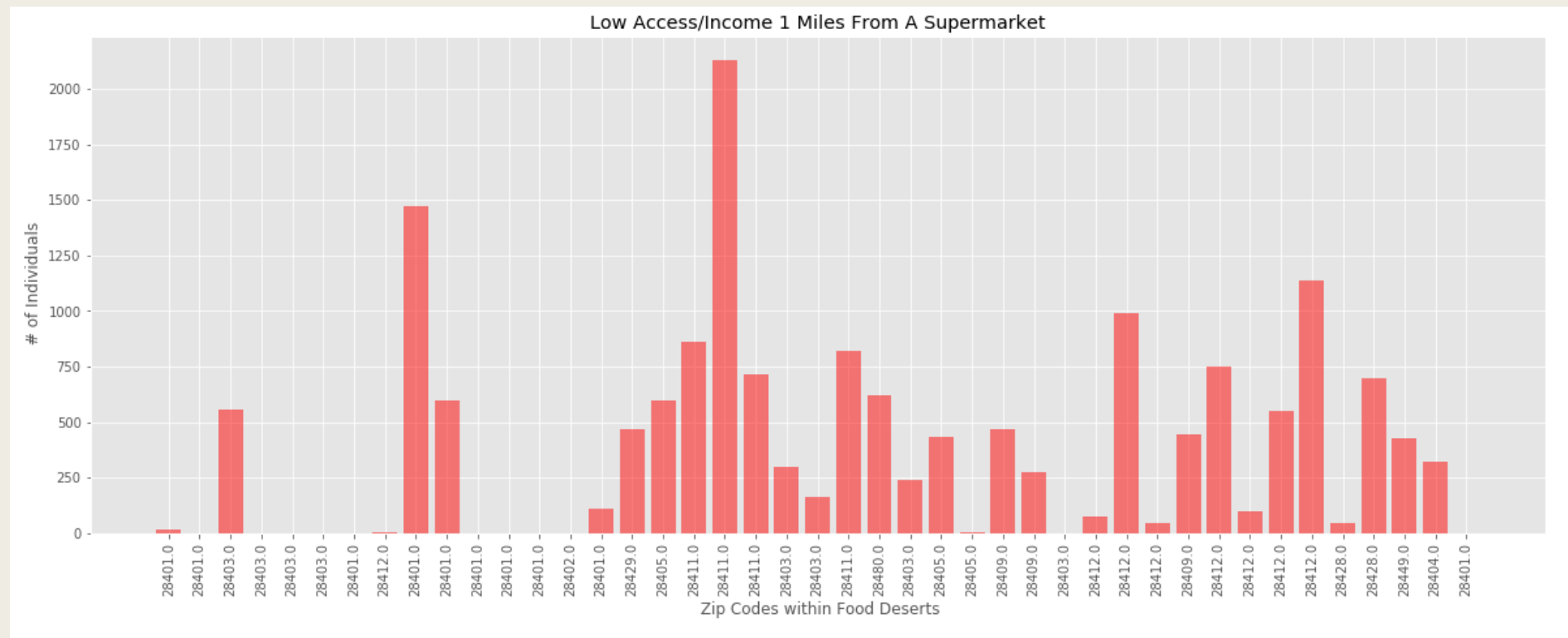
- Child Poverty distribution across different counties in North Carolina.



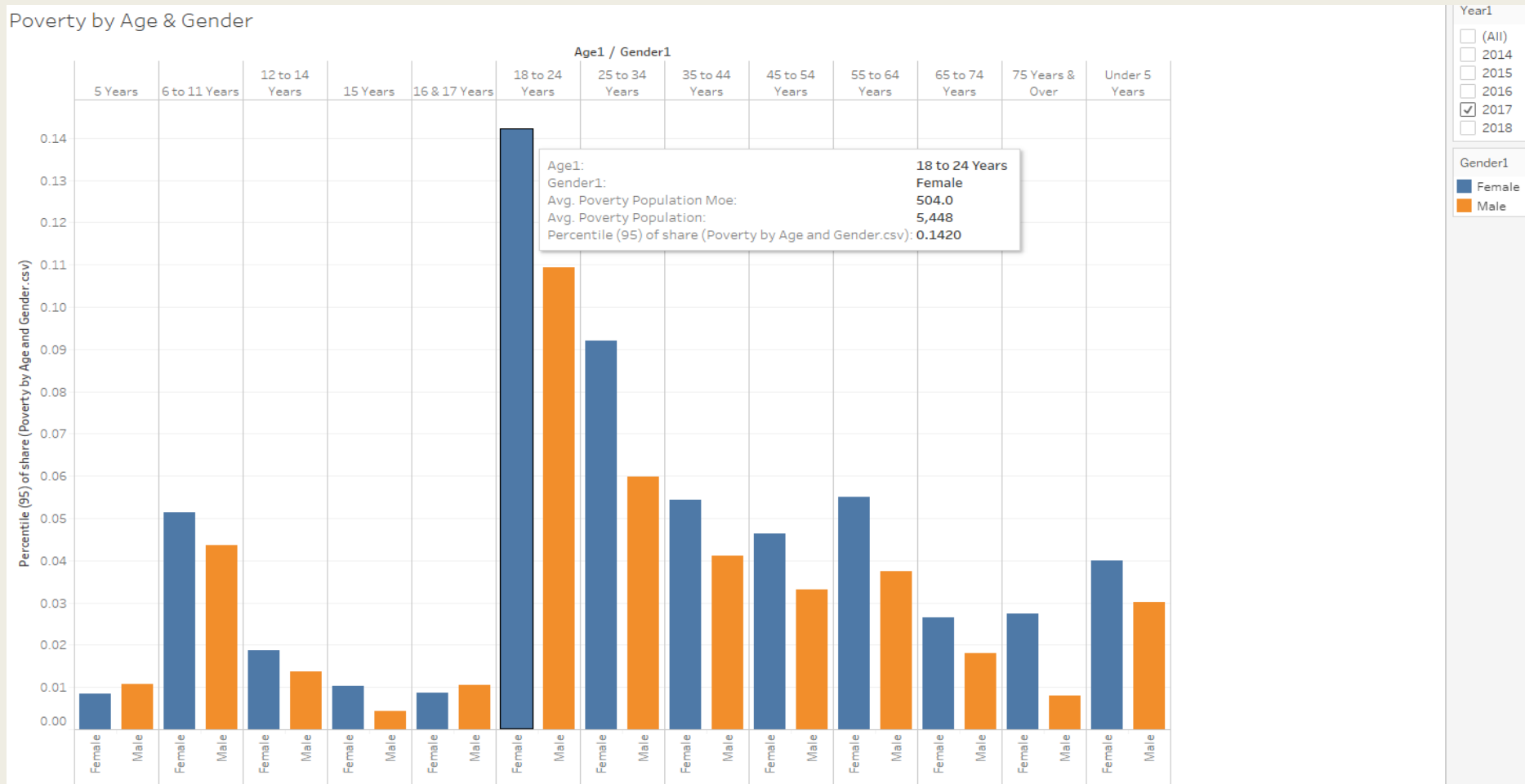
- Distribution of Housing units without vehicle across various ZIP Codes.
- The density of bubble signifies the number of SNAP users.



- Number of people with low access/income and 1 mile away from a Supermarket, living across various regions of New Hanover County.



- 18% of the population for whom poverty status is determined in New Hanover County, NC (38.4k out of 213k people) live below the poverty line, a number that is higher than the national average of 13.1%. The largest demographic living in poverty are Females 18 - 24, followed by Males 18 - 24 and then Females 25 - 34.





Data Science Approach

- **Food Desert Theory:**

There is a direct relationship between a population's access to fresh food and health.

In order to prove this theory, we need to answer 3 questions:

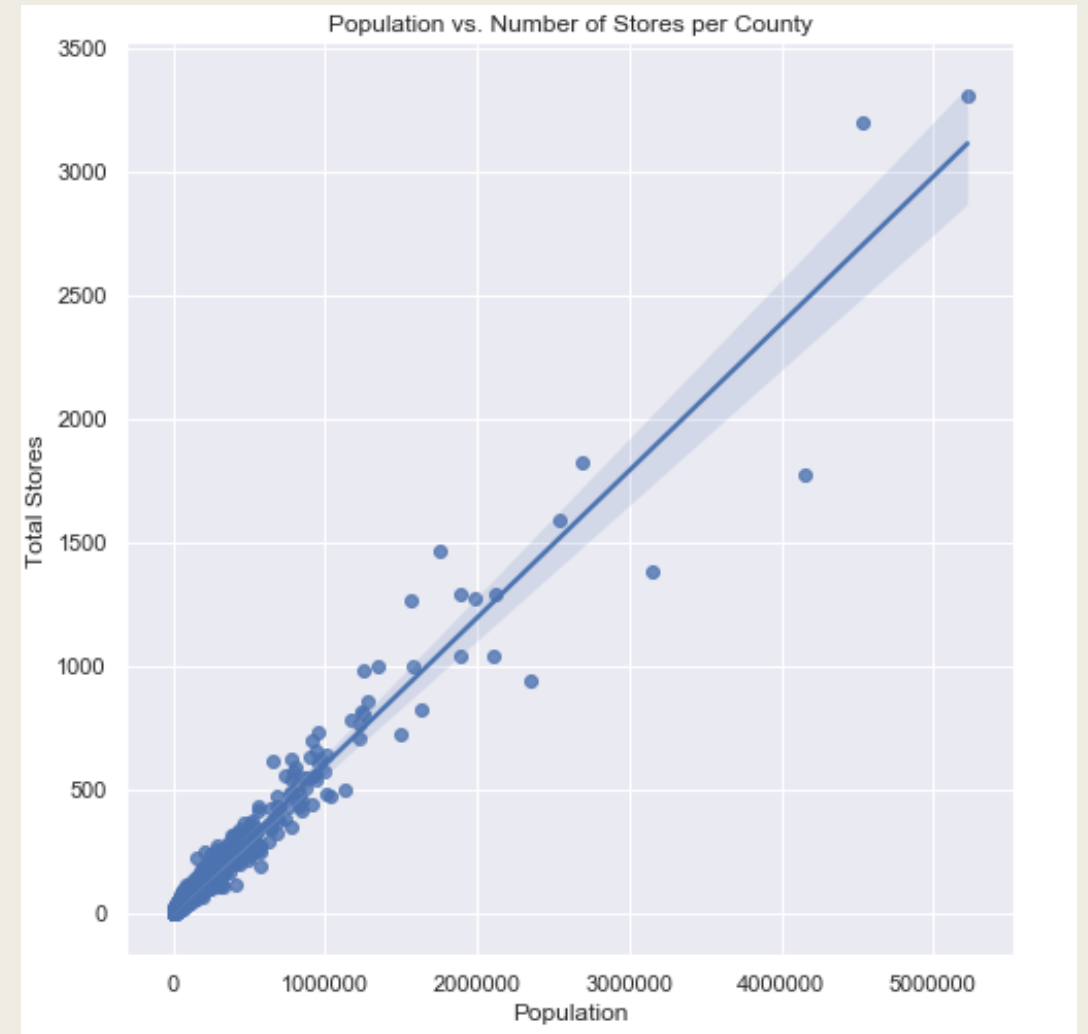
- Does a change in access to grocery stores show correlation to change in health over time?
- Are there outliers to this theory?
- Are there other factors that impact health beyond access such as socio economics or region or culture?

Approach

- I decided to focus on 4 aspects before collecting data and building the base model.
 1. Number of people living in every county across United States.
 2. Total number of employed/unemployed i.e. civilians labor force in each county.
This feature needs to be normalized by dividing it with the total population of the county.
 3. Total number of convenient store in each county, again needs to be normalized.
 4. Considering the demographics data such as income, poverty percentage, and unemployment rate along side percentage of each race in each of the counties.

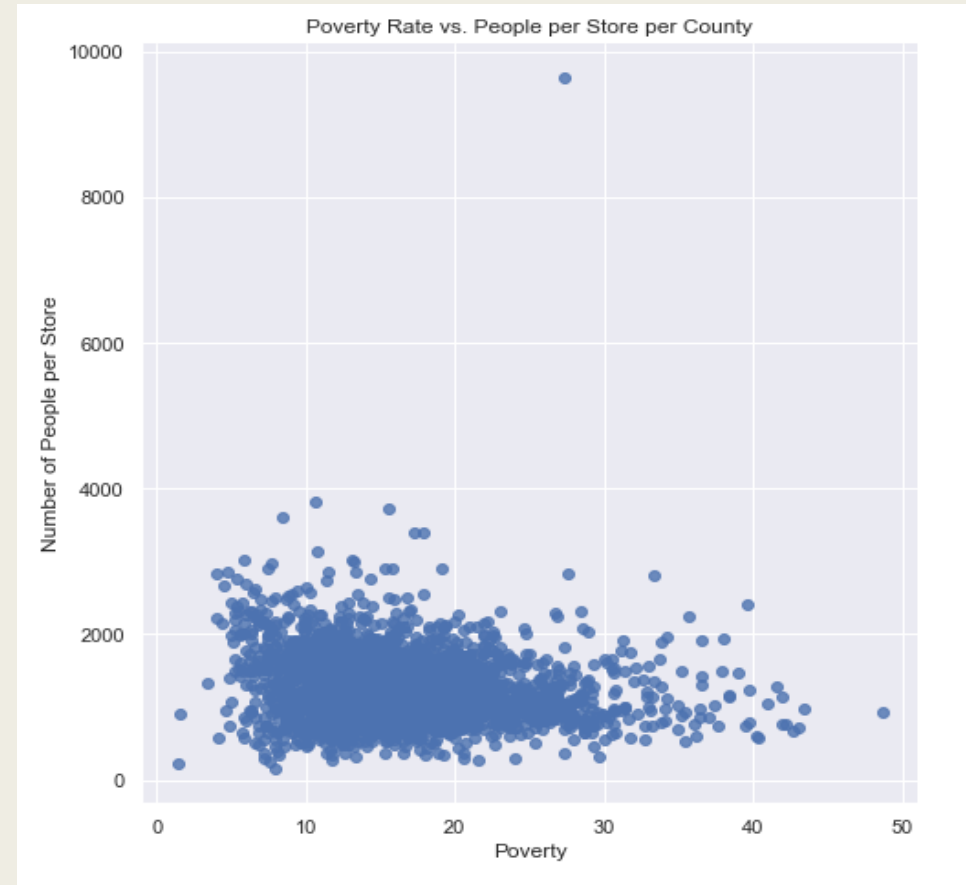
- Data was collected from 3 different sources and finally merged together to form one cleaned dataset.
- Target variable was Number Of People Per Store Per County in order to determine the store density in that particular region.
- **Correlation:**

There is a direct correlation between population size of the county and number of total stores.

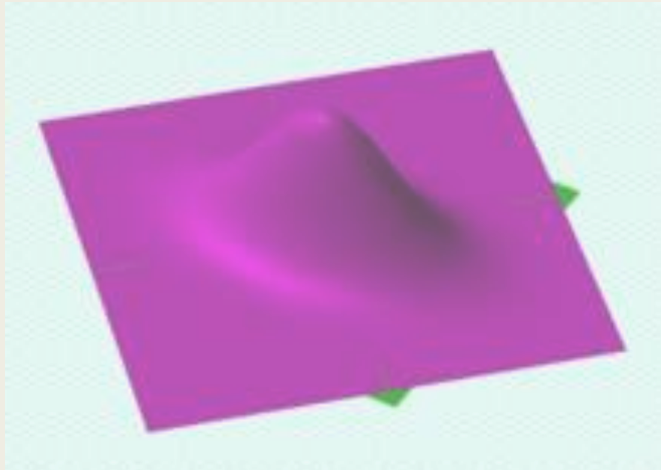


Correlation

- No Correlation between people per store and poverty.
- Similarly no Correlation was observed between people per store and unemployment.
- There is one outlier which might be interesting to investigate.



Covariance Matrix between
given features and target
variables.



```
1 # compute covariance between each feature and people per store
2 all_ = X.copy(deep=True)
3 all_['People per Store'] = y
4 all_.cov().loc[:, ['People per Store']]
```

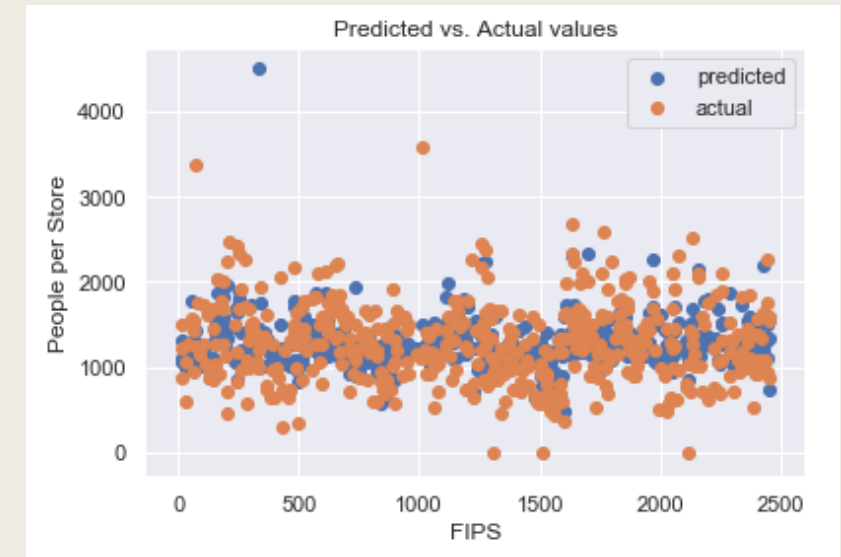
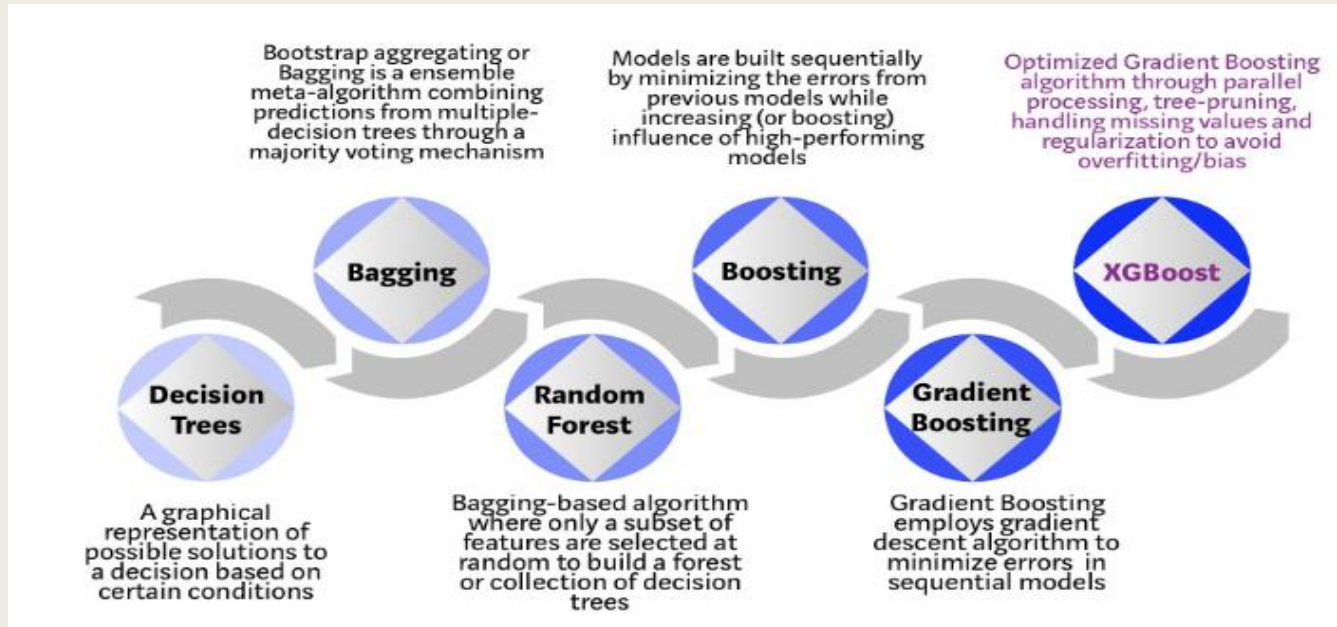
People per Store	
Hispanic	6.390389e+02
White	-4.104964e+02
Black	-5.703030e+02
Native	-5.006583e+01
Asian	2.882093e+02
Pacific	8.830140e+00
Income	2.081861e+06
Poverty	-5.272893e+02
Unemployment	3.680510e+01
Civilian Labor Force %	-2.627420e+01
People per Store	2.458045e+05

Final Model

Built an ensembled multiple regression Machine Learning Model.

- Gradient Boosting Regressor ML model:

Gradient boosting is a machine learning technique for regression and classification problems, which produces a prediction model in the form of an ensemble of weak prediction models, typically decision trees.



Conclusion

- The model has an r^2 (coefficient of determination) value of about 0.25.
- Means that about 25% of the variance of the y variable (people per store) can be accounted for by the set of features.
- There is not a strong correlation between the demographics of a county and the occurrence of stores.
- A better metric to use for adjustment could be the geographic size of the county.
- In addition, if the county is a rural, farming county, there are more likely to be fewer stores because more people would buy their groceries from farmers markets and local stores.
- Further analysis is needed to determine if these factors could better explain the variance in number of people per store.

Suggestions for the Community work and Future work

- FET can conduct various surveys while food hop operation.
 - Supply chain Management.
 - Sending text Message before arrival.
- We can identify food patterns based on the information collected through surveys.
- Association rule for items can also be implemented based on the demand of a particular item (similar to one used by wholesale giants like walmart).
- A pilot program with taxi services like Uber/Lyft can be very instrumental towards this which can offer \$2.50 rides to grocery stores.
- A Mobile App which allows seniors to order groceries online for delivery to neighborhood locations.

Data Source and References:

- <https://www.countyhealthrankings.org/app/north-carolina/2020/overview>
- https://www.huduser.gov/portal/datasets/usps_crosswalk.html#codebook
- Grocery store and population data: <https://www.ers.usda.gov/data-products/food-environment-atlas/data-access-and-documentation-downloads/>
- Income data: <https://www.ers.usda.gov/data-products/county-level-data-sets/download-data/>
- Demographics data: https://www.kaggle.com/muonneutrino/us-census-demographic-data#acs2015_county_data.csv
- <https://www.ers.usda.gov/data-products/food-access-research-atlas/download-the-data/>
- <https://www.ffiec.gov/census/report.aspx?year=2019&county=129&tract=ALL&state=37&report=demographic>

A person wearing a red coat is standing next to a vintage suitcase. The person's hand is visible, resting on the suitcase. The suitcase is light-colored with dark brown leather straps and buckles. The background is blurred, suggesting an outdoor setting. The text "THANK YOU" is overlaid in the center in a white, sans-serif font. There are white L-shaped brackets in the top-left and bottom-right corners of the image.

THANK YOU