

# Food Access Data Analysis



Alexis Burton, Yaekob Boru, Jamar Adams, John Miller

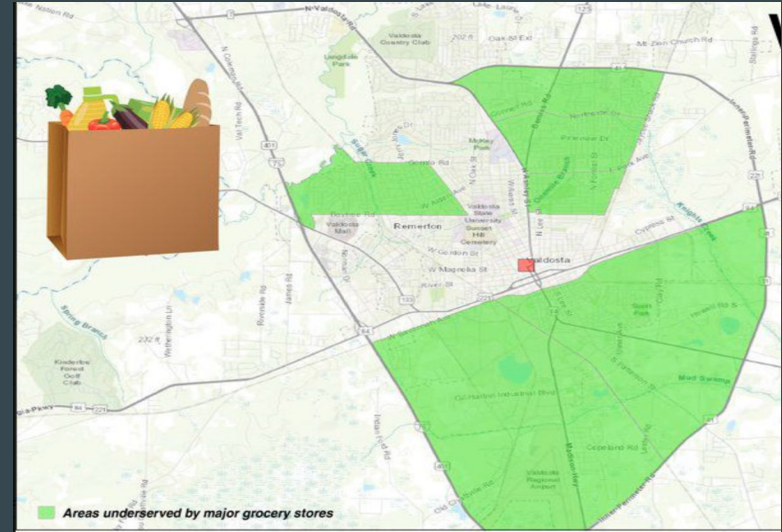
# Food Access Research

- Motivation to identify and analyze communities with limited access to healthy food.
- Urban neighborhoods and rural towns without access to fresh, affordable food are known as food deserts.
  - About 23.5 million Americans live in food deserts.
  - Urban population more than 1 mile from nearest supermarket
  - Rural Population more than 10 miles from nearest supermarket



# Inequities Associated with Food Deserts

- Dominated by fast food providers and corner stores with lack of fresh produce
- These populations are thus prone to health issues:
  - Obesity, cardiovascular diseases, diabetes.
  - Covid-19 Pandemic exacerbated these issues.



# Research Questions

- What socioeconomic factors are associated with communities with limited food access?
- Does race share a correlation with food access?
- Motivation: Understand and analyze underserved communities.

# Dataset Description

- Source: United States Department of Agriculture Food Access Research Atlas 2019
  - Contains census tract information for every county in United States
  - Census Tract: a subdivision of the county level
    - 3143 counties, over 73,000 census tracts
  - Key Variables: demographic data, median family income, population, poverty rate, low access 1 and 10.

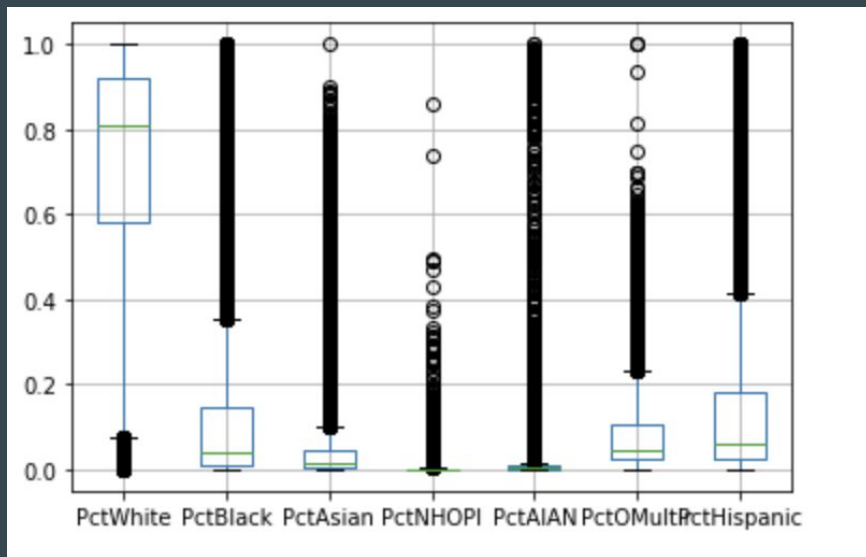
# Exploratory Data Analysis

- Originally looked at racial data:

```
Logistic Regression: 0.974  
Scaled Logistic Regression: 0.981  
Undersampled Logistic Regression: 0.982  
SVM Accuracy: 0.982
```

- Great accuracy, however, PCA indicated one variable had too much weight
- Thus, we looked further at the income data.

# Exploratory Data Analysis



Importance	
Urban	0.065373
Pop2010	0.205181
OHU2010	0.194710
PovertyRate	0.165930
MedianFamilyIncome	0.182706
PctLowI	0.186100

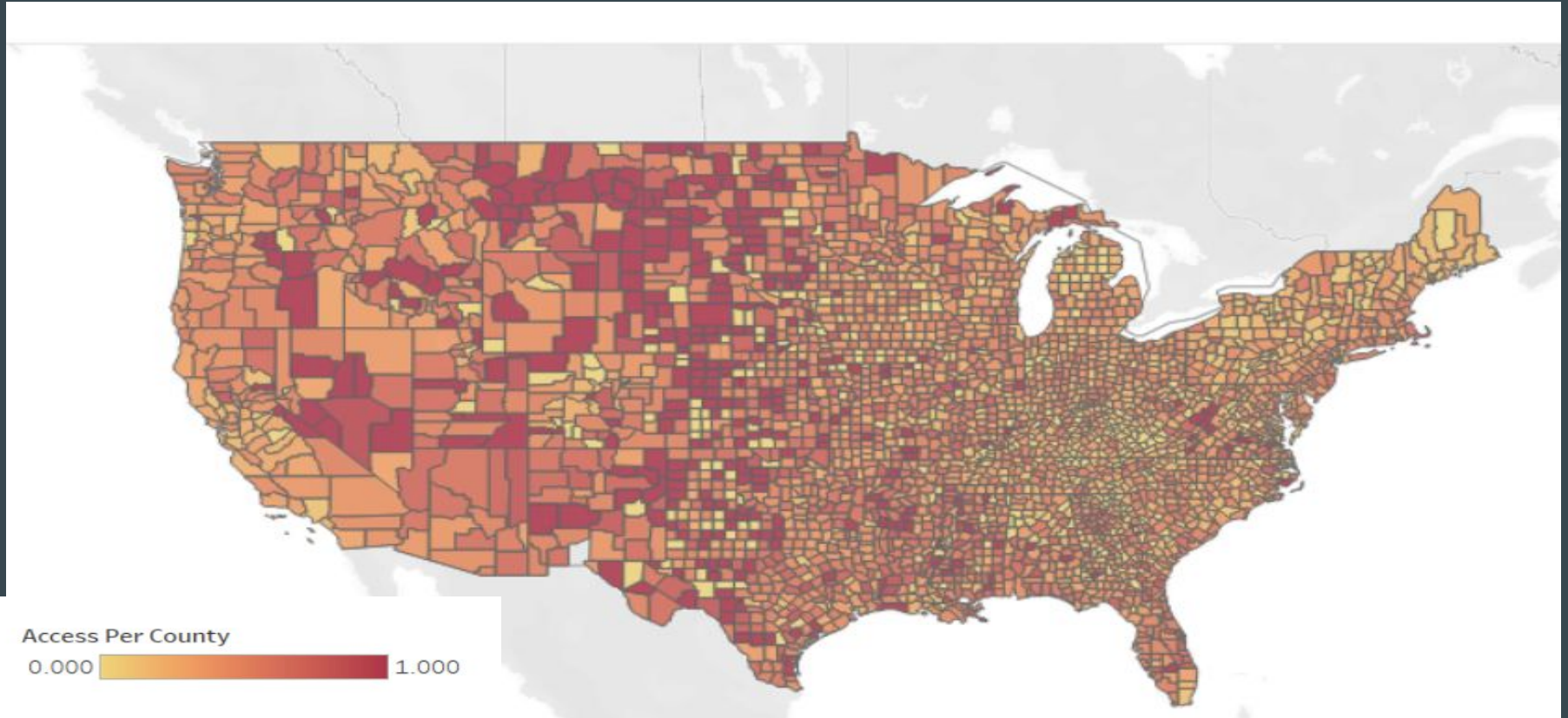
# Machine Learning Analysis

- Overall Accuracy: 69.72%
- Predicted food deserts with 72% accuracy
- Predicted non food deserts with 64% accuracy
- Used Supervised Machine Learning Model with low access 1 to 10 as target variable.

	Predicted 0	Predicted 1
Actual 0	9391	1756
Actual 1	3678	3121

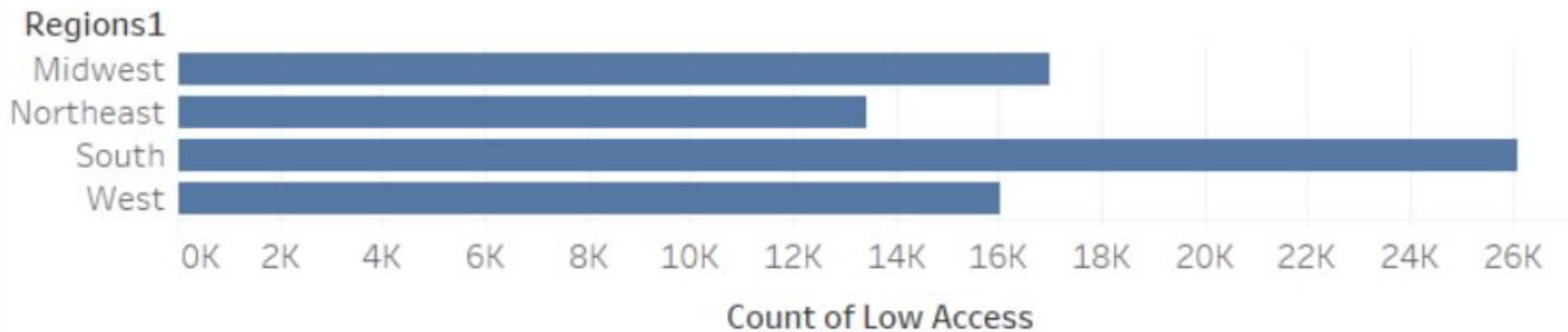


# Dashboard

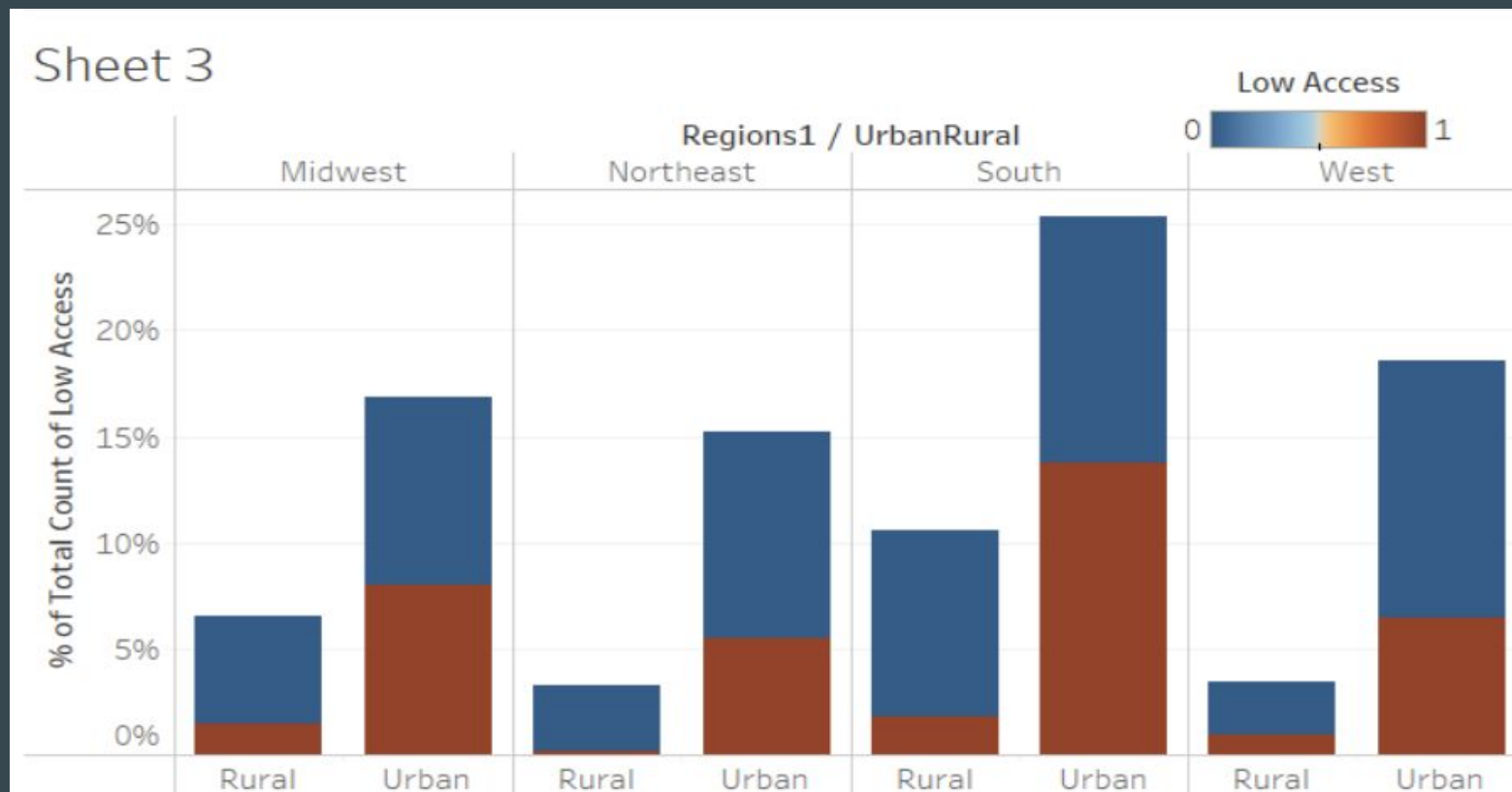


# Dashboard

## Sheet 1



# Dashboard



# Conclusion

- Racial data is significantly significant
- Income does not define food access for all populations.
  - This indicates other factors at play.