

Ghost Kitchen Marketing

Gregory Franks



Table of Contents



01

**Problem
Statement**



02

**Data and
Background**



03

Goals



04

Modeling



05

Recommendations



06

**Future
Development**



01

Problem Statement

Develop models that will enable targeted marketing campaigns and tailored menu offerings, presenting the top locations with recommendations for top menu items and keywords to use for paid search advertisement.



02

Data and Background



Data Sources

Food Desert

Data.world

Reddit

/finedining + /budgetfood

CDC NPAO

Data.gov

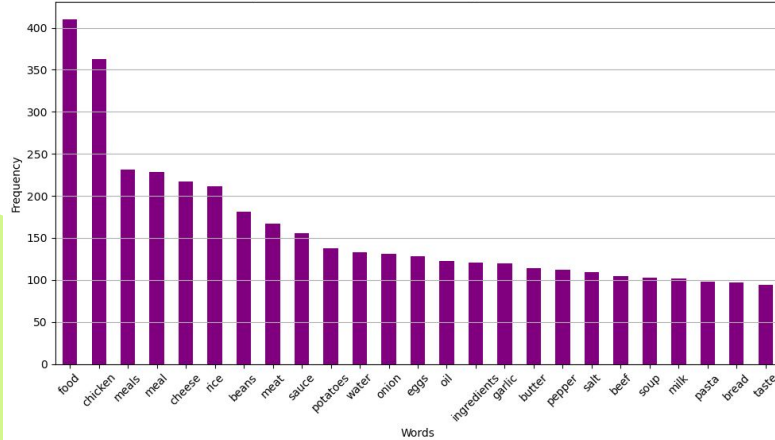


Reddit Data

Rice, Beans,
Soup, etc.



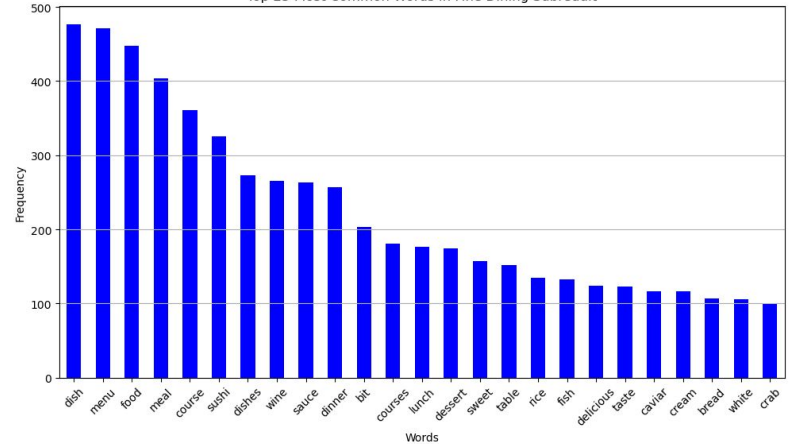
Top 25 Most Common Words in Budget Food Subreddit



Caviar, Crab,
Truffles, etc.



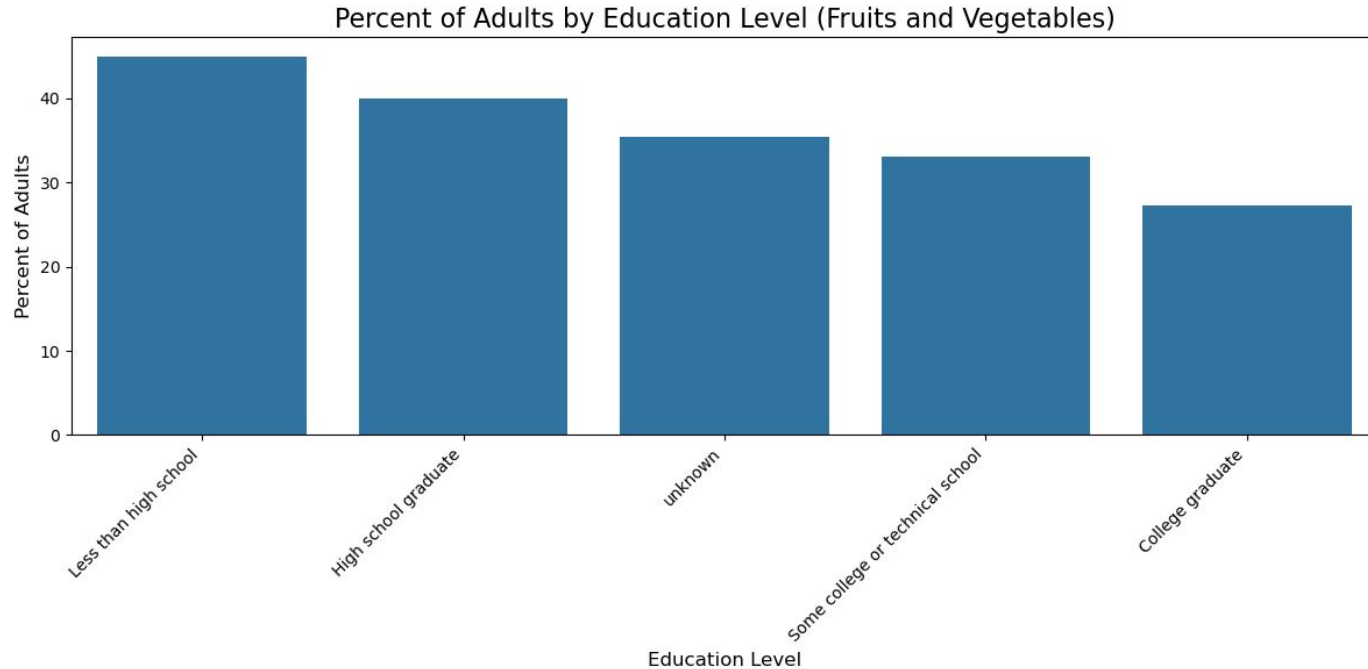
Top 25 Most Common Words in Fine Dining Subreddit



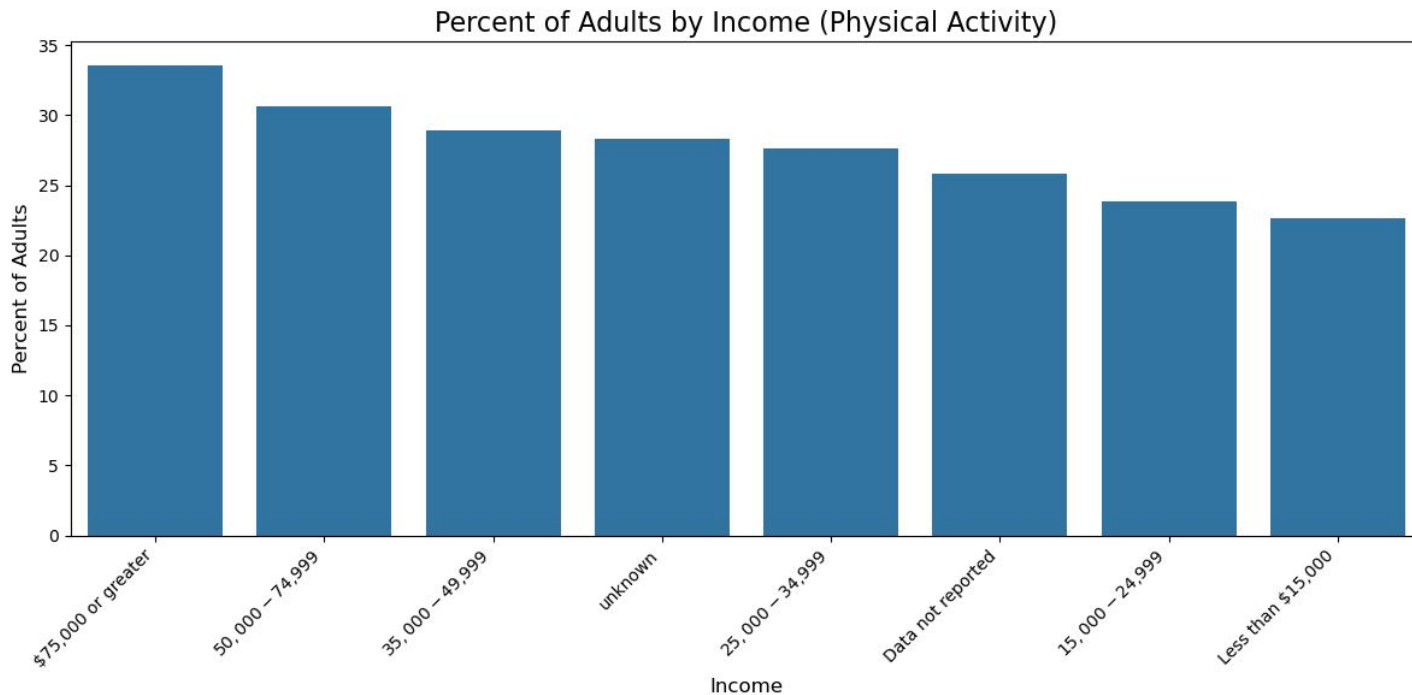
Census Tracts in Alabama



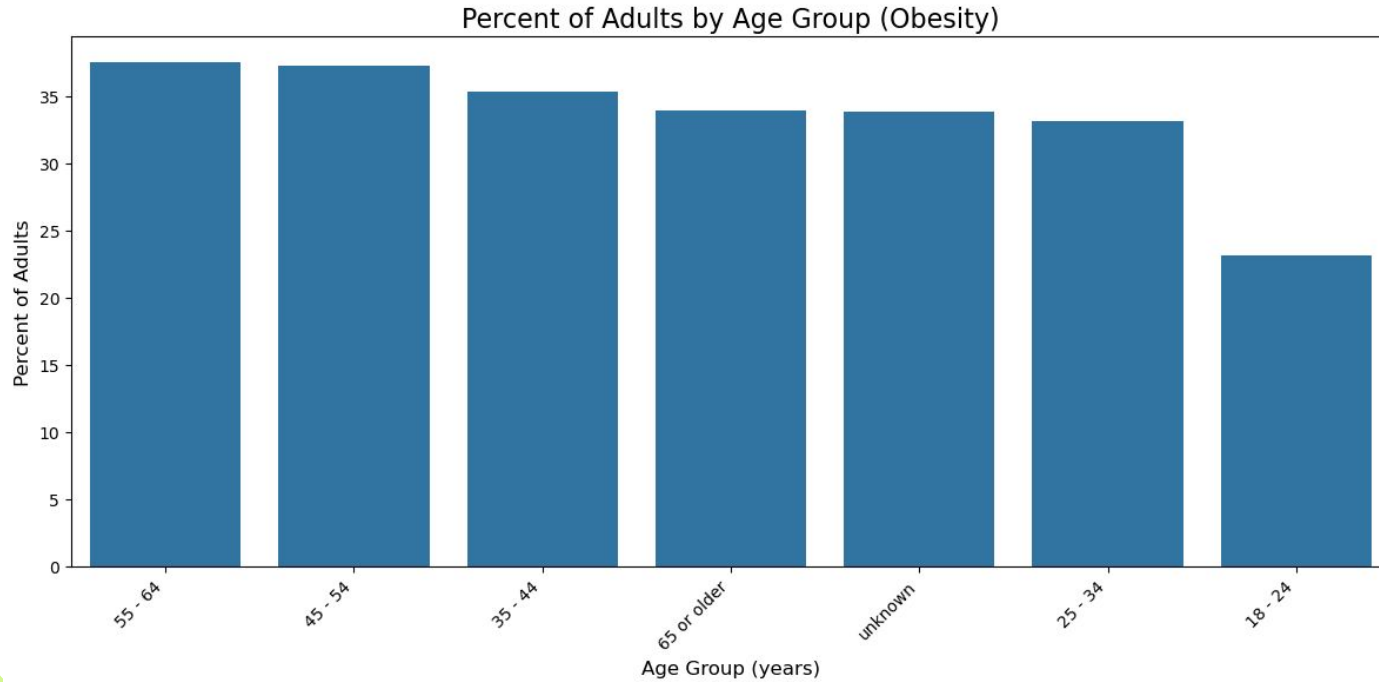
CDC: Fruits + Vegetables



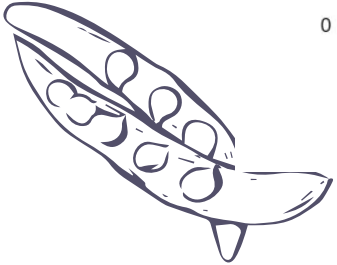
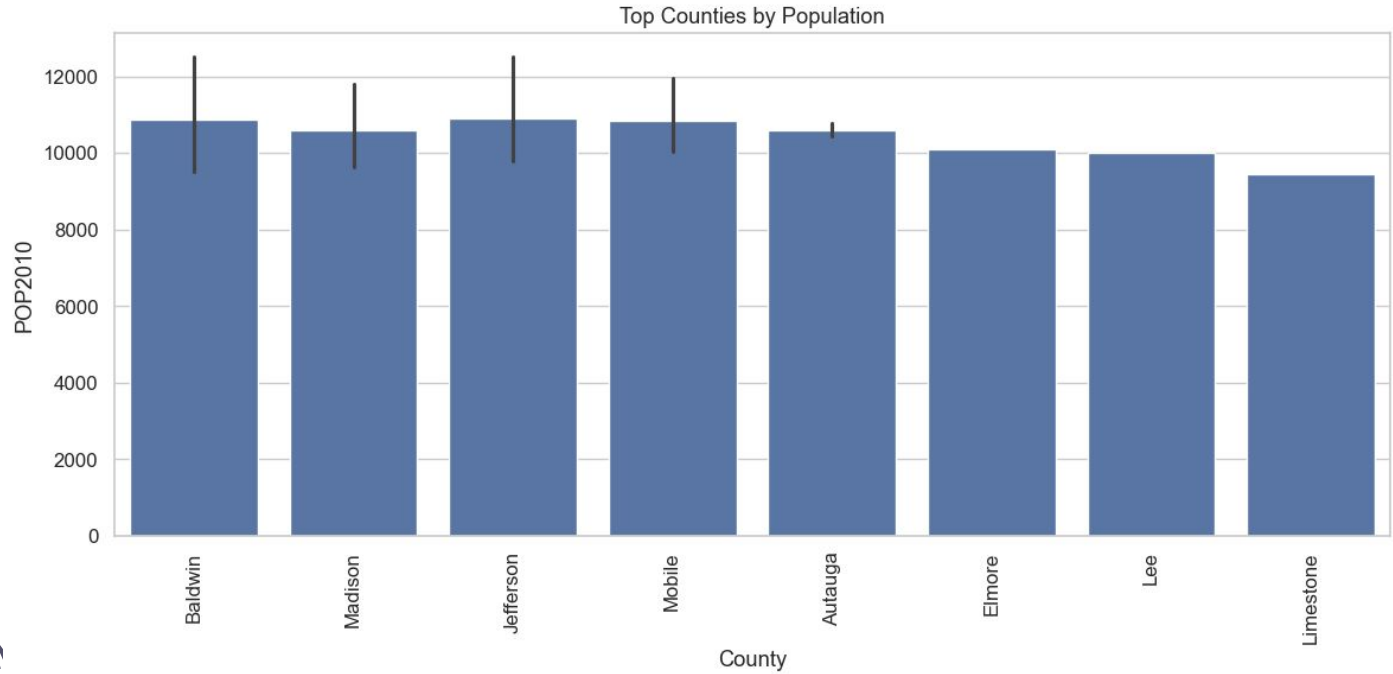
CDC: Physical Activity



CDC: Obesity



Food Deserts



A stylized illustration of an avocado, split in half to show the pit. It is drawn with dark blue outlines and some internal shading. The avocado is positioned to the left of the text.

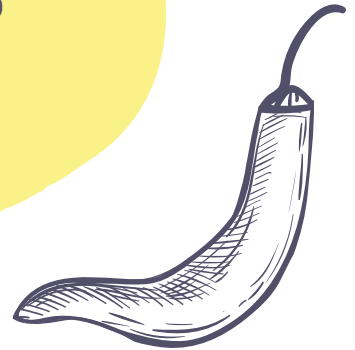
03

Goals

The background features three large, abstract, organic shapes: a light green one in the top left, a yellow one in the bottom left, and a tan one in the bottom right.

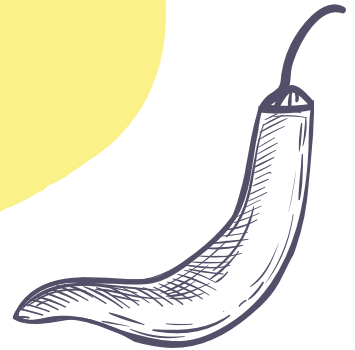
Location Analysis

- Avoid top income area - higher cost preferences will not be met
- Avoid lowest income area - not enough money to spend on our products
- Avoid top population area - need to have room to grow and expand, don't want to fail and have nothing but smaller areas to consider
- Avoid low population area - need to have enough people to advertise and sell to to make a reasonable profit



Keyword + Menu Analysis

- Focus on foods that are low cost to maximize net revenue; increase net of high volume purchases
- Focus on meals and ingredients that are popular with lower income groups to make sure there is enough volume of search terms
- Avoid foods that are high cost because they will not be on the menu and their search terms should be used as negative keywords





04

Modeling

Random Forest

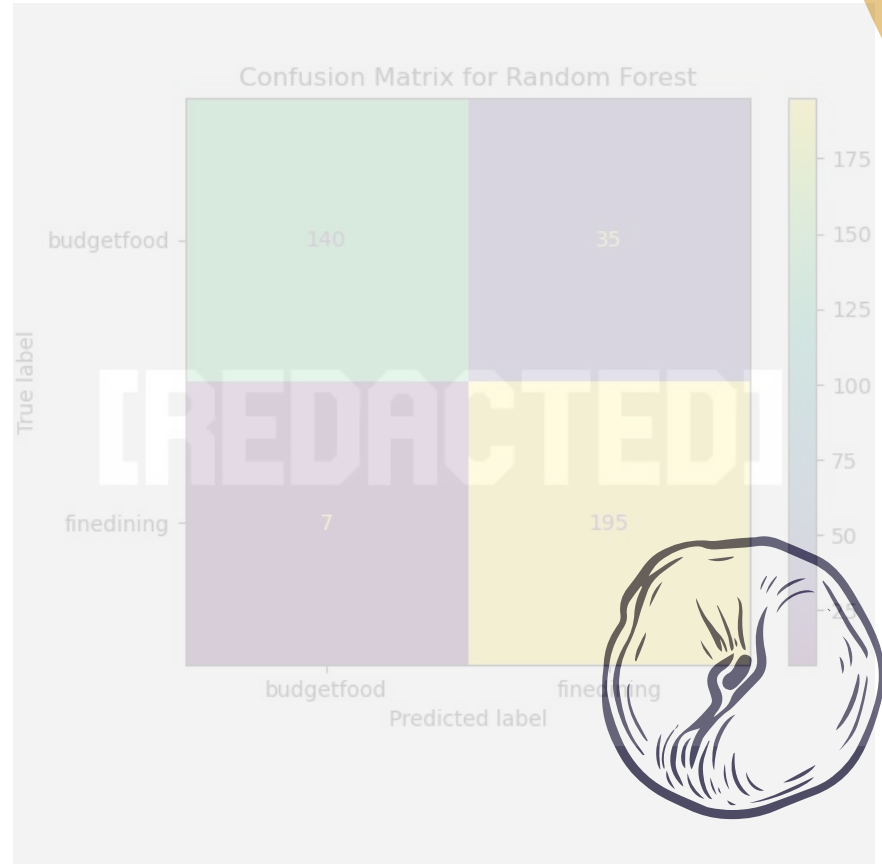
Performance:

Accuracy: 0.8886 (88.86% correct predictions)

Precision: 0.8964 (89.64% of positive predictions were correct)

Recall: 0.8886 (88.86% of actual positive instances were identified)

F1 Score: 0.8874 (harmonic mean of precision and recall)



Logistic Regression

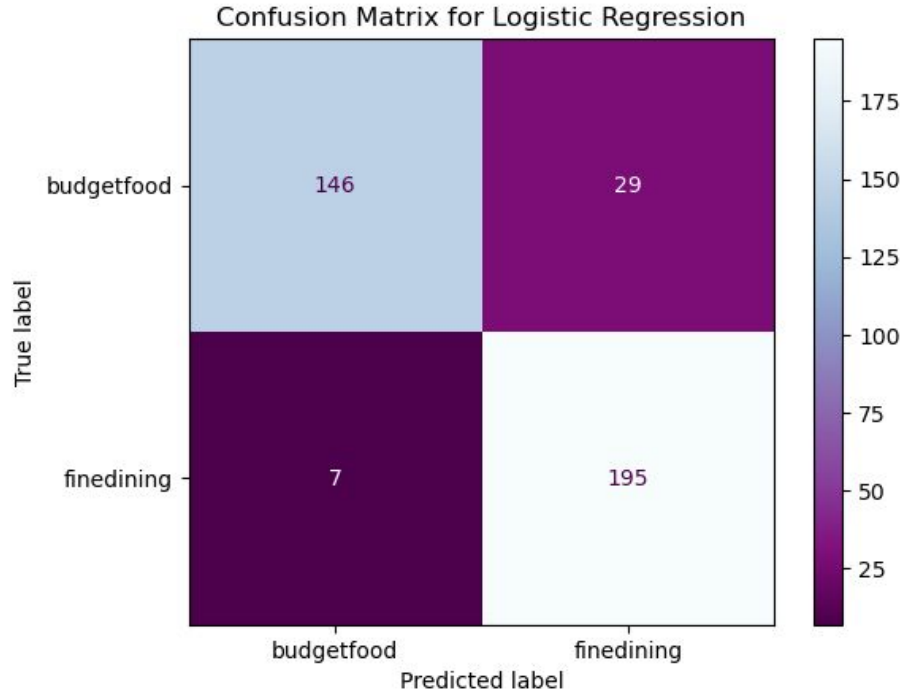
Performance:

Accuracy: 0.9045 (90.45% correct predictions)

Precision: 0.9094 (90.94% of positive predictions were correct)

Recall: 0.9045 (90.45% of actual positive instances were identified)

F1 Score: 0.9038 (harmonic mean of precision and recall)



Kmeans

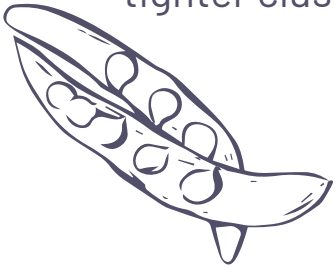
Performance:

Silhouette Score: 0.443

Inertia: 15436534561.065

Silhouette score indicates moderate cluster separation and cohesion

Lower inertia suggests tighter clusters



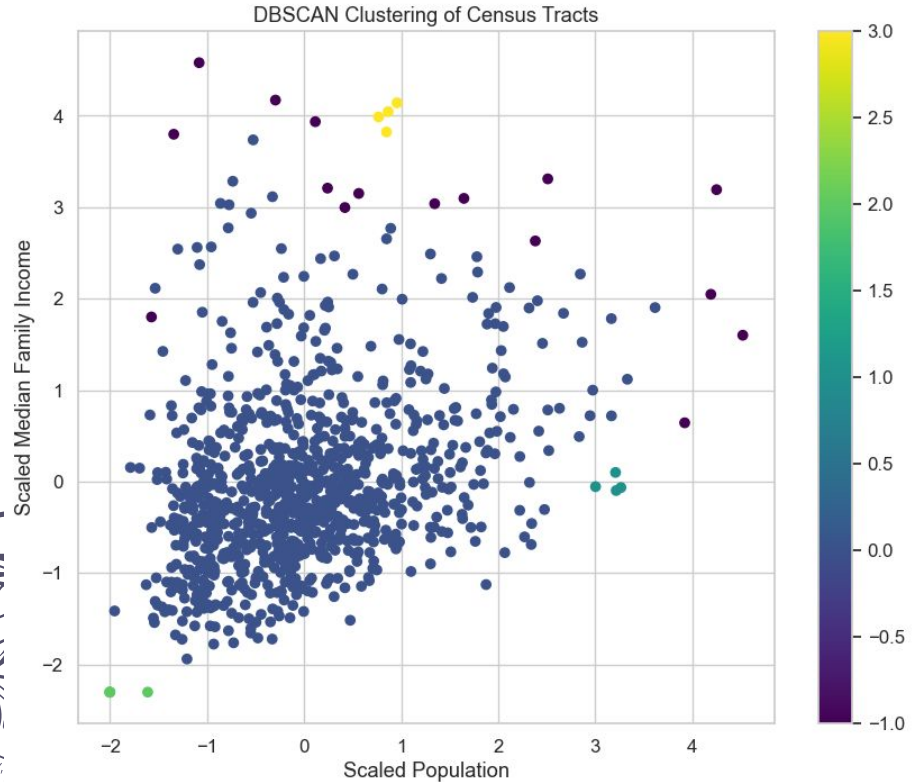
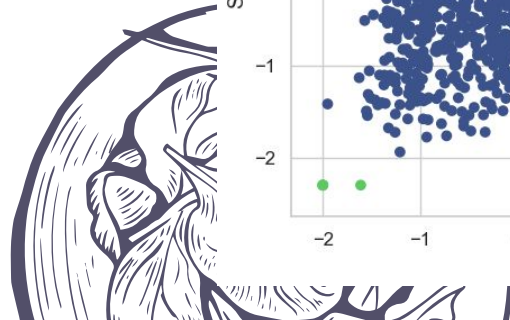
DBSCAN

Performance:
Silhouette Score: 0.288

Inertia: 2645004622287.102

Higher silhouette score
indicates better-defined
clusters than K-means
(possibly)

(Higher inertia is not directly
comparable due to different
clustering approaches)






05

Recommendations



Recommendations

- Establish initial operations and test in Baldwin, Elmore, and Mobile counties
 - Develop a menu centered around rice, soup, chicken, potatoes, and beans, etc.
 - Concentrate paid search advertising on keywords related to these core ingredients
 - Use high-cost items like truffles, sushi, caviar, and crab, etc. as negative keywords in advertising campaigns
- 

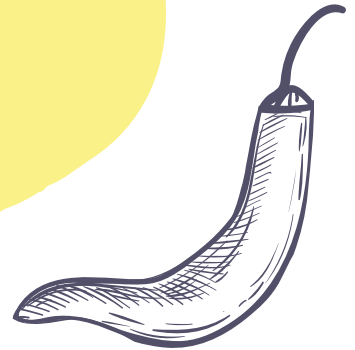
06

Future Development



Future Development

- Feature Engineering
- Ensemble Methods
- Geospatial Considerations
- Improve Data Acquisition





Questions?