

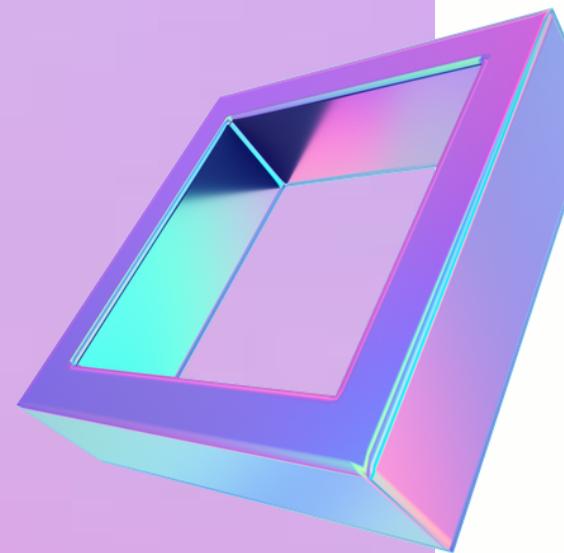
APPLIED DATA SCIENCE

Grocery Store Scarcity

Charlotte, North Carolina

Sakib Bhuiyan

Agenda



Part 1 Introduction

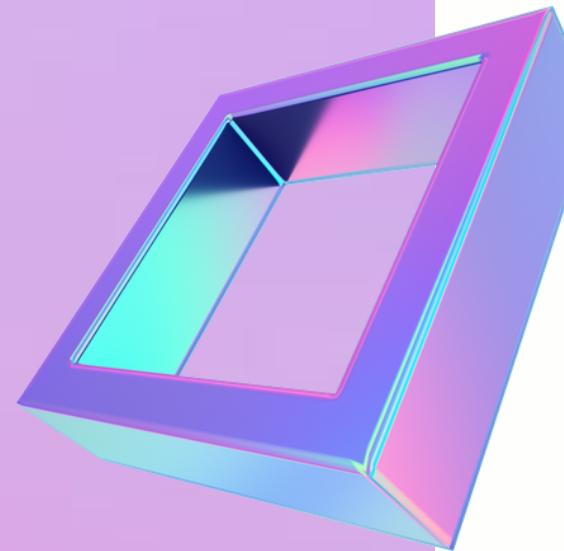
Part 2 Data

Part 3 Methodology

Part 4 Analysis

Part 5 Results and Discussion

Agenda



Part 1 Introduction

Part 2 Data

Part 3 Methodology

Part 4 Analysis

Part 5 Results and Discussion

Introduction

Background information

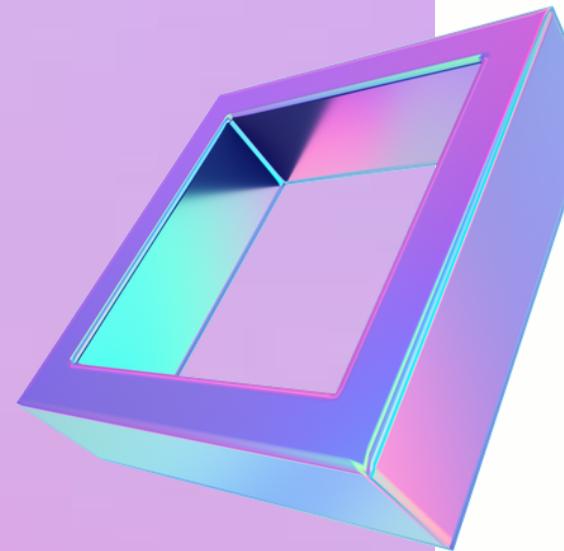
- **Food deserts:** geographic areas where access to affordable, healthy food options is limited because grocery stores are too far away.
 - **23.5 million Americans** live in food deserts, and nearly half of them are also low-income.
 - The lack of access to healthful foods can also contribute to diet-related conditions, such as **diabetes, high blood pressure, and cardiovascular disease**

Introduction

The business (and public health) problem

- The presence of food deserts has become both a public health and political issue in my hometown: **Charlotte, North Carolina**
- We will attempt to better understand which areas in Charlotte have limited access to grocery stores.
- These results can be used by city planners and grocery store retail chains to **prioritize grocery store openings in a way that makes healthy food options accessible to more Charlotteans.**

Agenda



Part 1 Introduction

Part 2 Data

Part 3 Methodology

Part 4 Analysis

Part 5 Results and Discussion

Data

Sources

- We were interested in the following data (**the source is provided in parentheses**):
 - Regions in Charlotte, NC. We decided to define regions by their ZIP code. (**Map of Charlotte, NC**)
 - Population data for Charlotte ZIP codes (**worldpopulationreview.com**)
 - Grocery stores that are within a specified radius of each Charlotte ZIP code (**Foursquare Places API**)
 - Location data for each ZIP code in Charlotte and each grocery store (**geopy geocoding services and Foursquare Places API**)

Data

Feature Description

Feature	Type	Description
ZIP Code	string	Each of the ZIP codes within the Charlotte city limit.
Latitude	float	Latitude coordinate for each location.
Longitude	float	Longitude coordinate for each location.
Population	int	Count of residents within each ZIP code.
Venue Count	int	Count of grocery stores within a specified radius of the ZIP code center.
Venue Scarcity	float	Calculated field that indicates "For every x number of people, there is 1 grocery store."
Grocery Store Relative Scarcity	float	Max absolute scale of <code>Grocery Store Scarcity</code> . Each <code>Grocery Store Scarcity</code> value is divided by the maximum value. Range is 0.0 to 1.0.

Data

Geocode Data Frame

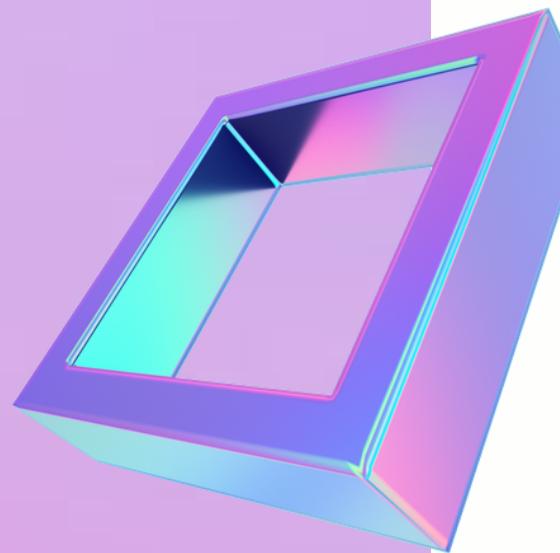
	ZIP Code	ZIP Code Latitude	ZIP Code Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	28202	35.227552	-80.842109	Harris Teeter	35.231231	-80.843956	Supermarket
1	28202	35.227552	-80.842109	Whole Foods Market	35.221301	-80.846211	Grocery Store
2	28202	35.227552	-80.842109	Trader Joe's	35.211608	-80.835468	Grocery Store
3	28202	35.227552	-80.842109	Target	35.214366	-80.834270	Big Box Store
4	28202	35.227552	-80.842109	Berrybrook Farm	35.202310	-80.844790	Grocery Store

Data

Scarcity Data Frame

	ZIP Code	Latitude	Longitude	Population	Venue Count	Venue Scarcity	Relative Scarcity
0	28202	35.2276	-80.8421	13498	14	964.142857	0.515089
1	28204	35.217	-80.8278	7199	20	359.950000	0.192302
2	28203	35.2116	-80.8568	16655	20	832.750000	0.444893
3	28206	35.2469	-80.829	12036	39	308.615385	0.164876
4	28205	35.2205	-80.7886	48798	50	975.960000	0.521402
5	28209	35.183	-80.8569	23533	38	619.289474	0.330852

Agenda



Part 1 Introduction

Part 2 Data

Part 3 Methodology

Part 4 Analysis

Part 5 Results and Discussion

Methodology

Steps

- First, we aggregated all of the data that was needed from the various sources.
 - We used Charlotte ZIP codes and retrieved population data for each ZIP code from **WorldPopulationReview.com**.
 - We used the **Foursquare Places API** to search for **grocery stores within a 1-mile radius**.
 - We used a map to visualize the **distribution of grocery stores** across Charlotte.

Methodology

Steps

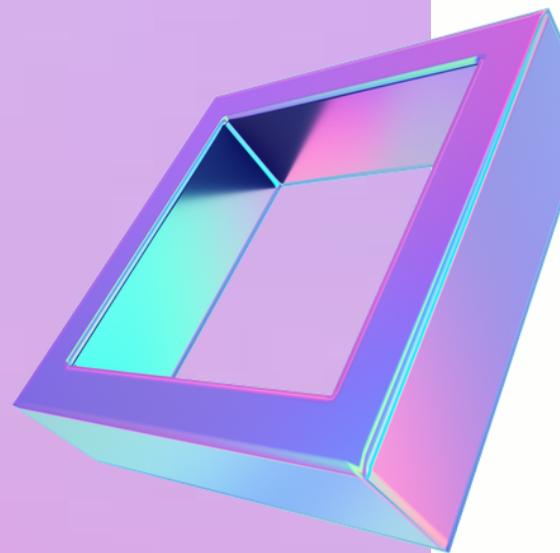
- Next, we created **new features** to understand the **grocery store density for each of the zip codes**.
 - Specifically, we created a new feature called **Grocery Store Scarcity**.
 - The value of this feature can be interpreted as “**Each grocery store within this ZIP code serves ‘x’ people.**”
 - The higher this grocery scarcity values, the more people that 1 grocery store is serving.
 - ZIP codes with high Grocery Store Scarcity values have a low density of grocery stores within the area, and there may not be enough grocery stores to adequately serve the population within the ZIP code.

Methodology

Steps

- We used these new features to identify the ZIP codes that have low grocery store density (high Grocery Store Scarcity) and high grocery store density (low Grocery Store Scarcity).
- Finally, we created visualizations to better understand the data.
 - **Boxplots:** understand the distribution of Population and Grocery Store Scarcity.
 - **Two-axis bar plot:** compare the Population of each ZIP code to its Grocery Store Scarcity.
 - **Map:** illustrate Population and Grocery Store Scarcity based on location.

Agenda



Part 1 Introduction

Part 2 Data

Part 3 Methodology

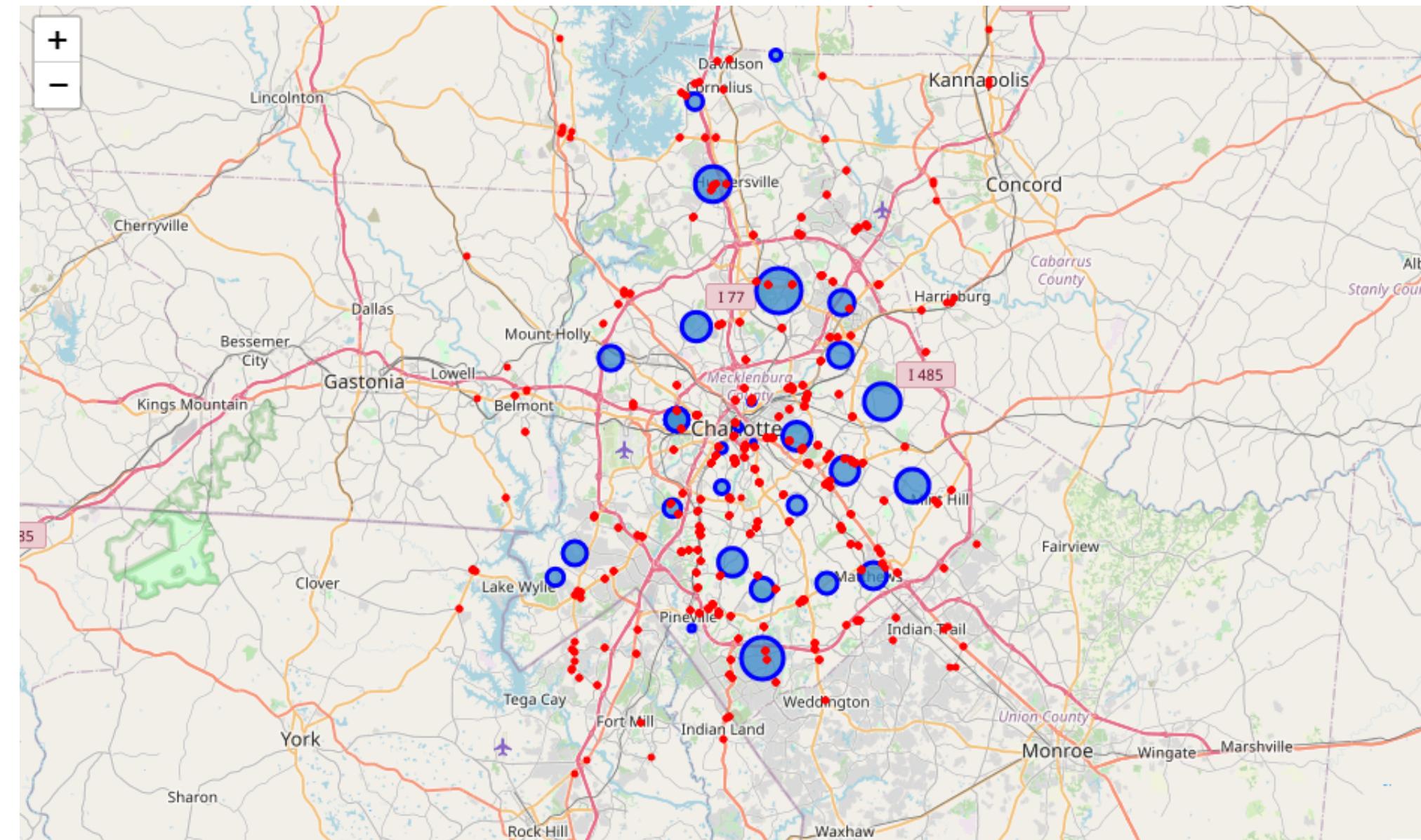
Part 4 Analysis

Part 5 Results and Discussion

Analysis

Map of Charlotte Grocery Stores

- **Blue markers:** ZIP code centers. Size is proportional to Population.
- **Red markers:** grocery stores within 1-mile radius of ZIP code center.



Analysis

Top 5/Bottom 5 Grocery Store Scarcity

Top 5

	ZIP Code	Population	Venue Scarcity
7	28031	28077	1871.800000
13	28262	43209	1800.375000
25	28277	72132	1717.428571
12	28269	77248	1544.960000
19	28215	62543	1250.860000

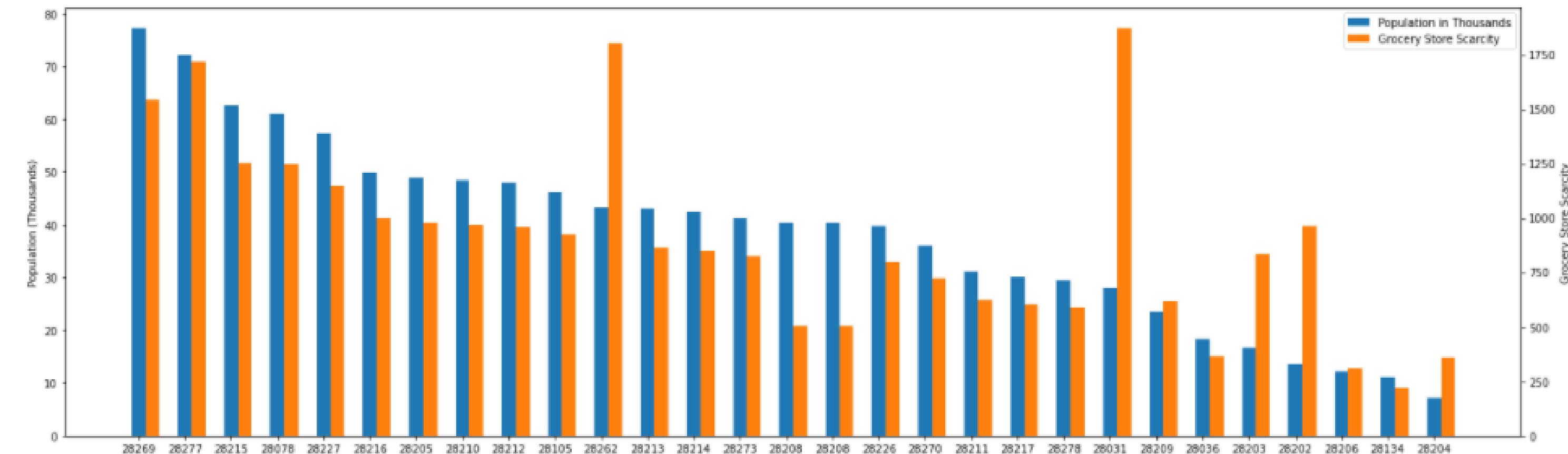
Bottom 5

	ZIP Code	Population	Venue Scarcity
6	28208	40284	503.550000
8	28036	18266	365.320000
1	28204	7199	359.950000
3	28206	12036	308.615385
20	28134	11037	220.740000

Analysis

Imbalance in Population and Grocery Store Scarcity

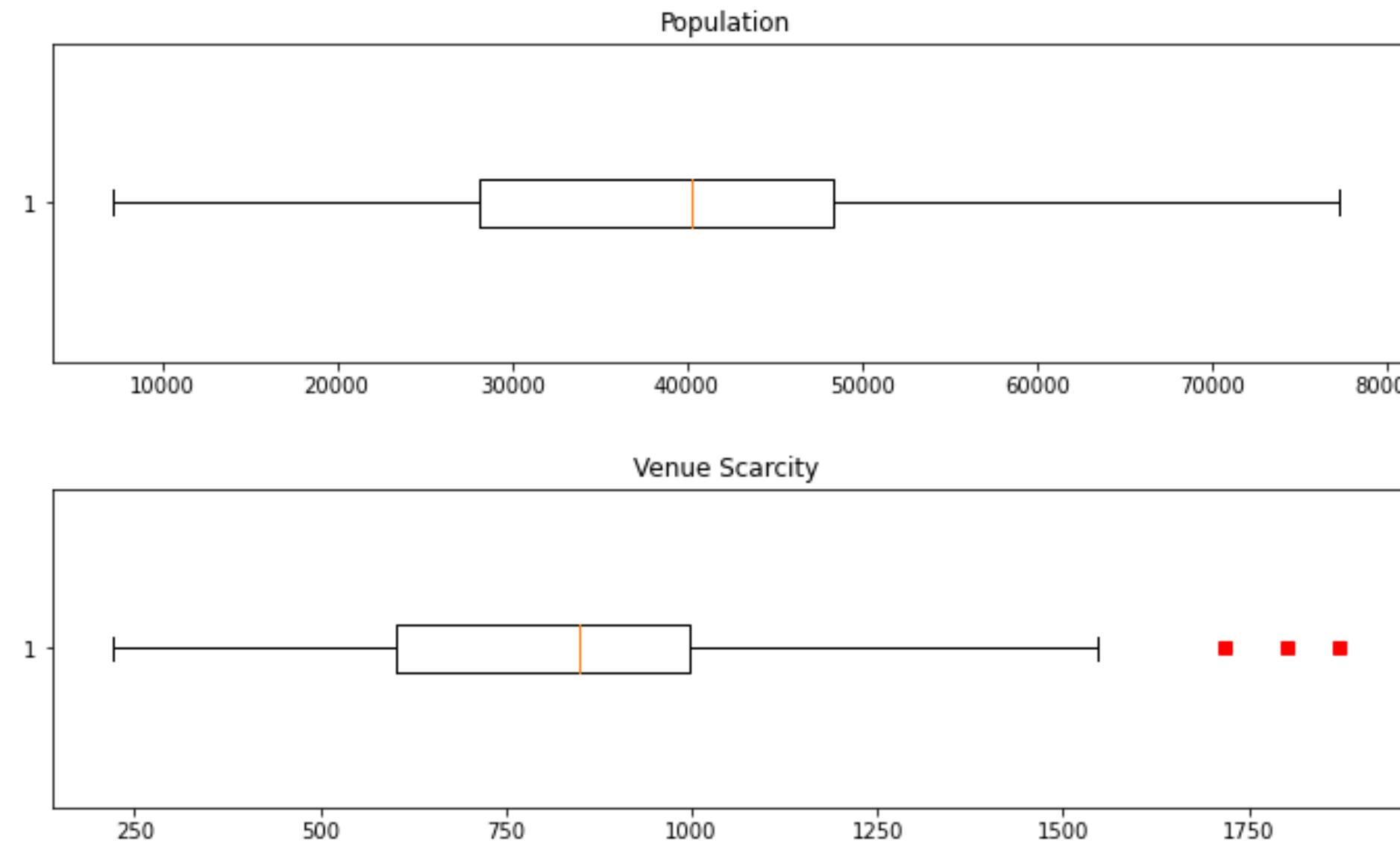
- **Blue Bars:** Population
- **Orange bars:** Grocery Store Scarcity
- Noticeable imbalance between the two bars for the following ZIP codes:
 - 28262
 - 28031
 - 28202



Analysis

Distribution of Population and Grocery Store Scarcity

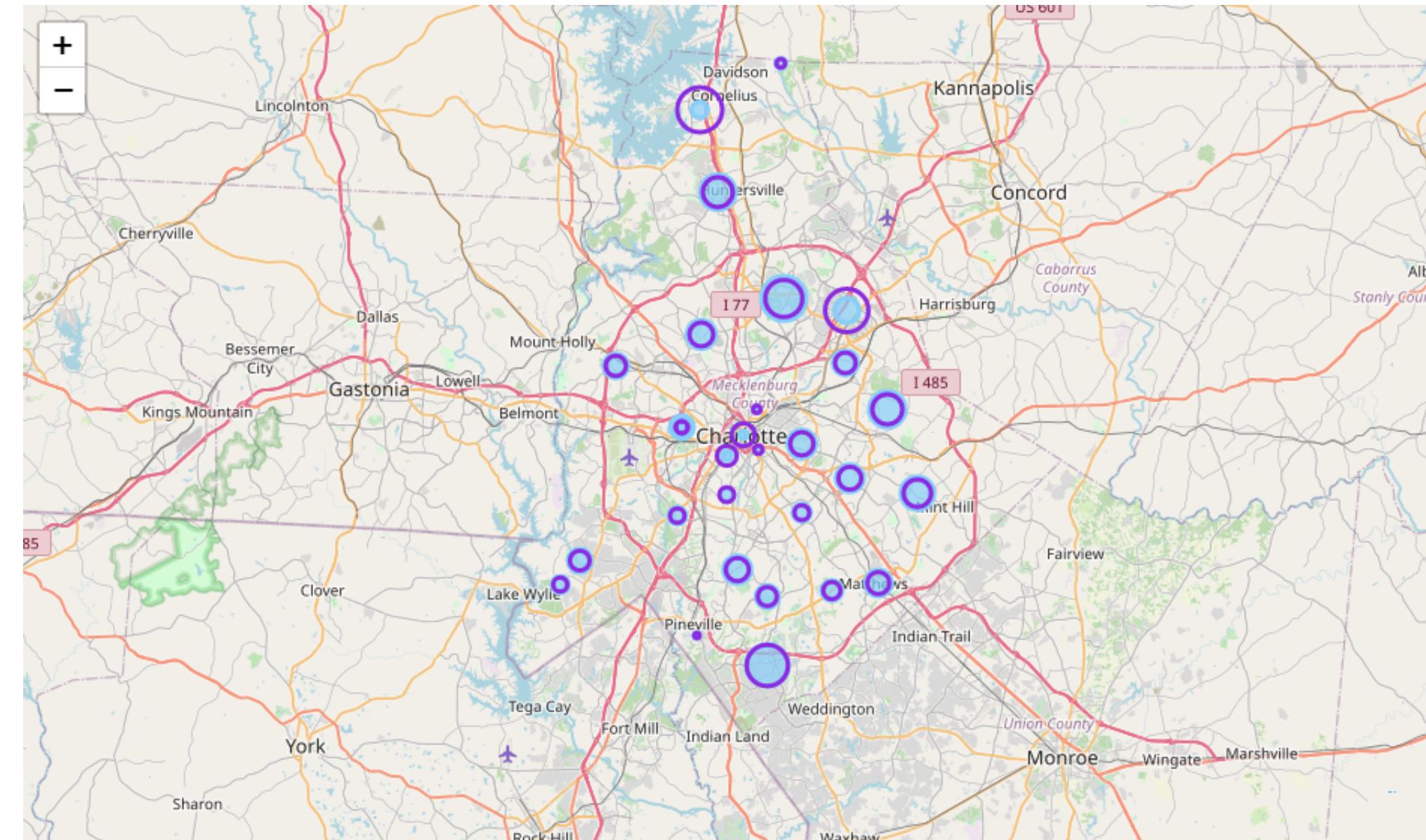
- **Population:** symmetric distribution
- **Grocery Store Scarcity:** 3 noticeable outliers (28031, 28262, 28277)



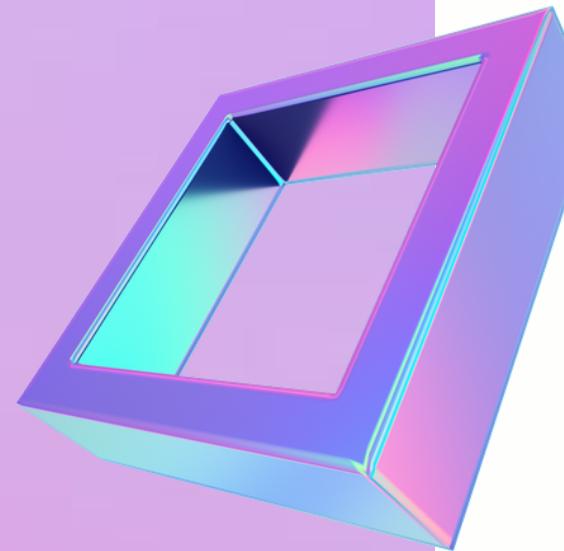
Analysis

Map of Population and Grocery Store Scarcity

- **Blue markers:** ZIP code centers. Size is proportional to Population.
- **Purple markers:** Size is proportional to Grocery Store Scarcity.
- Purple marker larger than the blue marker: there may be a lack of enough grocery stores in the area.



Agenda



Part 1 Introduction

Part 2 Data

Part 3 Methodology

Part 4 Analysis

Part 5 Results and Discussion

Results and Discussion

Main Takeaways

- There are, indeed, certain ZIP codes in Charlotte that have an imbalance between their population and the density of grocery stores.
 - These ZIP codes could potentially be **food deserts**,
 - Opening additional grocery stores in these specific areas could have a larger impact than opening additional grocery stores in other ZIP codes.
- The lack of grocery stores is **most prevalent in ZIP codes 28262 and 28031**.
 - These ZIP codes are found along the perimeter of the Charlotte city limits.

Results and Discussion

Future Expansion Potential

- Future work would require more data to add additional features for each zip code. A few suggestions for future work are below:
 - Include **income data** for each ZIP code. This could help explain whether grocery stores are more prevalent in affluent areas because retailers believe their revenue will be higher in these areas.
 - Include **demographics (race, age, education)** for each ZIP code. This could help explain whether food deserts disproportionately affect certain demographics, and whether retailers should be incentivized to open grocery stores in specific under-served areas.
 - Use **clustering with income and demographic data** to understand which ZIP codes are the most similar.

Thank you!

