	Zip Code	Land Area	Population	Trader Joe's Target Audience	Median Household Income	Foot Traffic	Trader Joe's	Competitors	Score
1	11215	5641350	73419	31107	156930	2789	0	1	0.521000
2	11234	19547800	94254	27707	86185	1231	0	1	0.520000
3	10016	1378330	56758	33257	136431	8572	1	2	0.514000
4	11226	3339500	99776	42717	66173	2075	0	1	0.509000
5	10001	1609400	26966	13812	101409	12938	0	2	0.508000

     I would advise Trader Joe's to open their next store in the zip code area **11215**. The rationale behind this choice lies in identifying metrics that are most relevant for launching a new Trader Joe's store. To arrive at a conclusion, I evaluated every zip code area in New York City based on the **land area, population, target audience, median household income, foot traffic, number of existing Trader Joe's stores, and number of competitors' stores**. Once the necessary data was gathered, I designed a model that assigns weights to every metric and outputs a composite score, summing up the product of weights and values for every criterion. The visualization above contains the top-5 recommendations for the next Trader Joe's location. The zip code area 11215 is the top choice of the model, which is on par with the average metrics from zip code areas in New York City, where Trader Joe's stores already exist. Moreover, the chosen zip code area also follows the pattern of greater values across all metrics of interest compared to the average data from all New York City neighborhoods.

Average data from zip code areas that already have at least one Trader Joe's

Land Area	Population	Trader Joe's Target Audience	Median Household Income	Foot Traffic	Trader Joe's	Competitors
4789968	59020	24211	112210	3368	1	1

Average data from all New York City zip codes

Land Area	Population	Trader Joe's Target Audience	Median Household Income	Foot Traffic	Trader Joe's	Competitors
4201310	48354	17986	86136	2174	0	0

     The recommended neighborhood possesses multiple advantages over the next options in the model's choice due to the balance between the criterion values. The neighborhood area is not too small, has many residents, and has a good ratio of Trader Joe's target audience to the total number of residents. The median household income is the highest among the top-5 choices, and there is potential for capturing market share due to the absence of other Trader Joe's stores nearby and the low level of competition from chains like Target and Whole Foods.

Preliminary Research

I utilized a top-down approach to achieve the desired outcomes. I started by researching Trader Joe's target audience and factors differentiating the company from its competitors. Then I derived a set of metrics for assessing the desirability of each New York City zip code area to Trader Joe's executives. Finally, I designed an algorithmic solution that breaks down the analysis of each metric into multiple subparts and then combines the insights to provide a data-driven recommendation.

Data gathered from research about Trader Joe's:

- Unique products with a mix of flavors/ingredients that are not found elsewhere (Terzigni, 2023).
- 80% of the products are store brands (Trader Joe's buys directly from suppliers to avoid middleman markup) (Terzigni, 2023).
- Big selection for customers with food allergies and special dietary preferences (Terzigni, 2023).

I defined Trader Joe's **target audience** as educated individuals with income above the median (Fink, 2021) who are progressive (e.g., support sustainable practices (Trader Joe's, 2022)) and are eager to try new products. I based the target audience on age and considered individuals between 20 and 45 years old as most likely to shop at Trader Joe's (Zippia, 2023).

I evaluated every zip code area in New York City upon the following metrics and *assumptions* connected to them:

- *Land area* — the larger the neighborhood size, the easier it would be to find a building that Trader Joe's can lease or an empty lot where the company can build a store.
- *Population* — choosing a location with a small number of residents might not be profitable.
- *Target audience* — despite a high population, the store might not get enough visitors if the subset of residents who belong to the store's target audience is not big enough.
- *Median household income* — Trader Joe's needs to ensure that there is a customer base that can afford its products. This is also important for maintaining the current pricing strategy.
- *Foot traffic* — locations with high foot traffic result in greater visibility, more new customers, and a higher level of accessibility due to the presence of transportation links.
- *Number of existing Trader Joe's stores* — if there at least one Trader Joe's already exists in the neighborhood and its area is not big, there might be no room for gaining substantial profit.
- *Number of competitors' stores* — if there are no competitors in the neighborhood, there is a risk of it being not suitable for the target audience. It is important to note that not all grocery stores should be considered competitors. I selected Target and Whole Foods as competitors of Trader Joe's for this scenario.

Recommendation Workflow Steps

1. Scrape all zip codes of New York City and their land areas (Name Census, n.d.).
2. Use Census Bureau APIs (U.S. Census Bureau, 2021) to retrieve data about population, Trader Joe's target population (combined 3 age ranges: 20-24, 25-34, 35-44), and median household income for every zip code area. Filter out the areas that do not contain any data about population (e.g., JFK airport) or median household income.
3. Visualize the distributions of the metrics obtained so far.
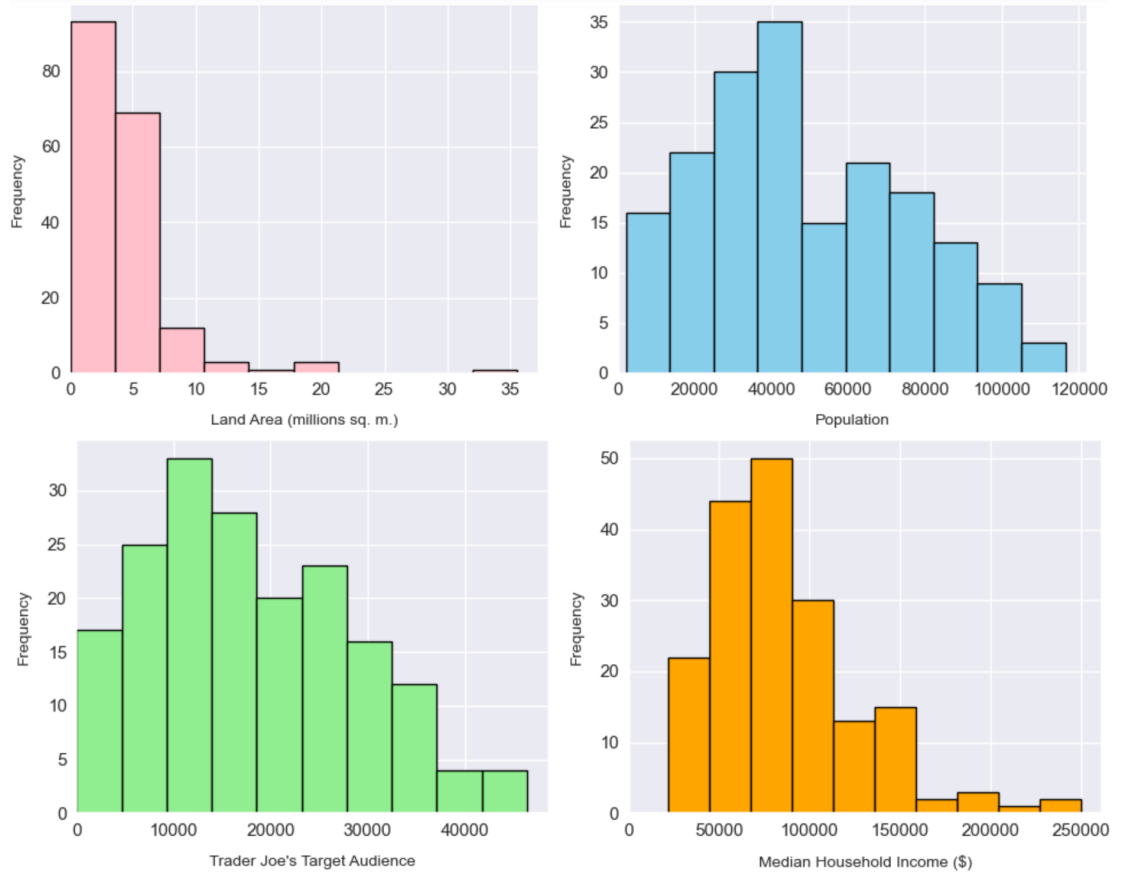
4. Gather data about the centroid coordinates of each zip code area (LOJIC Open Data Portal, 2020) and measure the distance between the zip code area and the rest of the zip code areas in New York City. Sort the neighbors of every zip code area by distance.
5. Gather data about foot traffic in New York City. Foot traffic is defined as the average number of people who pass a specific location point in the morning/evening of a weekday and the afternoon of a weekend (NYC Open Data, 2022). Since the foot traffic is tracked at a specific point, convert its coordinates with an appropriate zip code using OpenCage geocoding APIs (OpenCage, n.d.). The *assumption* made here is that one point represents the average foot traffic value in the whole neighborhood.
6. The foot traffic data is only available for some zip code areas. To fill the missing gaps, find 5 closest neighbors (zip code areas) that are present in the foot traffic dataset and apply inverse distance weighting (Wikimedia Foundation, 2023) to obtain foot traffic value for the given neighborhood.
7. Visualize the distribution of foot traffic.
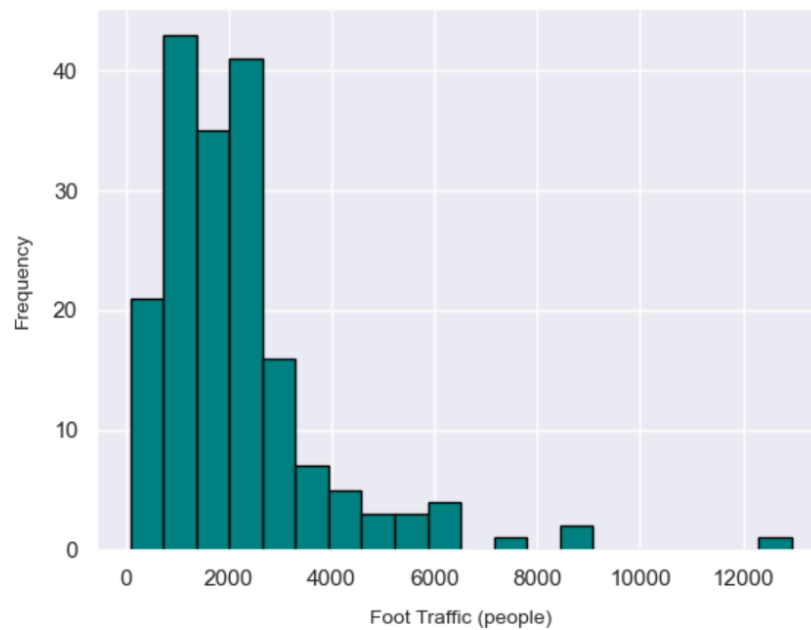
8. Find zip codes of existing Trader Joe's locations and competitor locations in New York City (Trader Joe's, n.d.). Calculate the number of Trader Joe's stores and the number of competitors' stores per zip code area.
9. Visualize the store count bar charts.
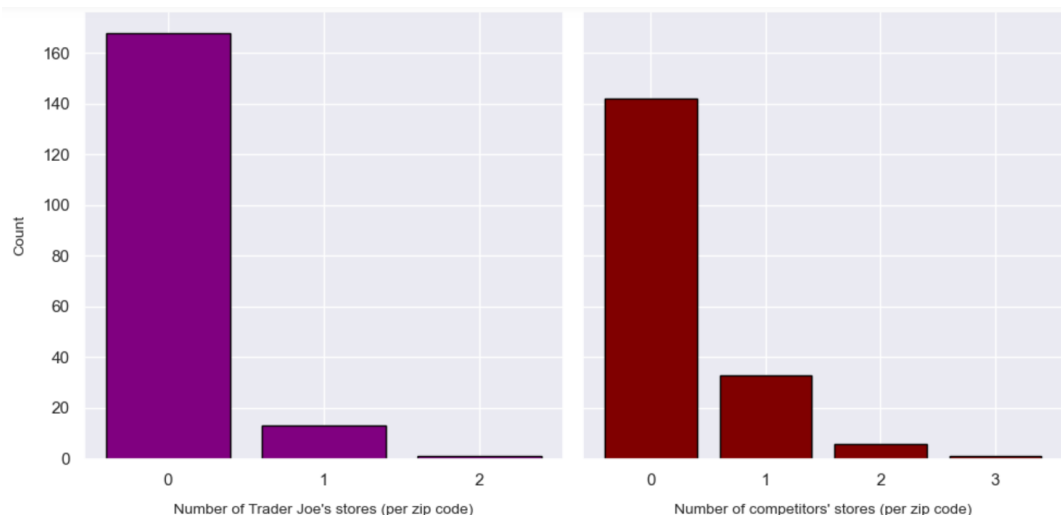
10. Since all metrics are represented on different scales, normalize them so that all metrics besides zip code have a value between 0 and 1.
11. Assign weights to each metric. Assuming all metrics should be weighted equally, their weights will equal ~0.14285, given that there are 7 variables. However, I have modified the importance of some metrics because *I think* that median household income and target audience should be more valuable for choosing the next store location. They are followed by the number of Trader Joe's stores (we want to "punish" the neighborhoods that already have a Trader Joe's store) and the number of competitor stores. Finally, land area, population, and foot traffic are marked as the least important out of the 7 metrics. Despite such allocation of weights, they are still roughly equal because the difference between the minimum and maximum weight values is 0.03.
12. Calculate the score for each neighborhood by multiplying the associated weights with metrics.

13. Sort the data by score and select the top-5 recommended locations. Manually examine the metric values to ensure that the first choice is optimal.
14. Compare the output to the metric trends present in locations with at least 1 Trader Joe's store and all locations across New York City.

Further Investigations and Improvements

If I was given more time and permitted access to proprietary and/or third-party data sources obtained for a fee, I would find foot traffic data of higher quality. Foot traffic is a very important variable in this decision-making process. My data is currently based on one geolocation point in some neighborhoods and a "synthetic" value obtained through inverse distance weighting in other neighborhoods. This is not representative of the entire zip code area, so I would search for materials that provide a more detailed foot traffic breakdown. To gain unique customer perspectives, I would find a tool that conducts sentiment analysis on the location-related reviews of the existing Trader Joe's and competitors' stores. This information might help me make a more nuanced decision due to a better understanding of pain points in the existing store locations. This way, I could narrow my recommendation down to a specific block within the recommended neighborhood.

References

Fink, B. (2021). *Here's why there's not a Trader Joe's where you live.* Allrecipes. Retrieved from

https://www.allrecipes.com/article/where-trader-joes-opens-stores

LOJIC Open Data Portal. (2020). *ZIP code population-weighted centroids.* Retrieved from

https://data.lojic.org/datasets/HUD::zip-code-population-weighted-centroids/about

Name Census. (n.d.). *List of 214 ZIP codes in New York, New York.* Retrieved from

https://namecensus.com/zip-codes/new-york/city/new-york

NYC Open Data. (2022). *Bi-Annual Pedestrian Counts.* Retrieved from

https://www.nyc.gov/html/dot/downloads/pdf/bi-annual-ped-count-readme.pdf

OpenCage. (n.d.). OpenCage Geocoding API Python Tutorial. Retrieved from

https://opencagedata.com/tutorials/geocode-in-python

Terzigni, C. (2023). *25 things you didn't know about Trader Joe's products.* The Daily Meal.

Retrieved from https://www.thedailymeal.com/1315862/trader-joes-products-facts

Trader Joe's. (2022). *Sustainability.* Retrieved from

https://www.traderjoes.com/home/sustainability

Trader Joe's. (n.d.). *Trader Joe's locations in New York.* Retrieved from

https://locations.traderjoes.com/ny

U.S. Census Bureau. (2021). *American Community Survey: 5-Year Estimates: Data Profiles 5-*

Year. Retrieved from https://api.census.gov/data/2021/acs/acs5/profile

Wikimedia Foundation. (2023). *Inverse distance weighting*. Wikipedia. Retrieved from

https://en.wikipedia.org/wiki/Inverse_distance_weighting

Zippia (2023). *Trader Joe's demographics and statistics [2023]*. Retrieved from

https://www.zippia.com/trader-joe-s-jobs/demographics/#:~:text=34-,Trader%20JOE%27s%https://www.zippia.com/trader-joe-s-jobs/demographics/#:~:text=34-,Trader%20JOE%27s%20Age%20Breakdown,-The%20average%20age20Age%20Breakdown,-The%20average%20age