Using machine learning to find location for opening a new restaurant in Michigan.

Awet Tsegai

IBM Data Science Capstone Project
Awet Tsegai

INTRODUCTION AND BACKGROUND OF THE PROBLEM

- Having a successful restaurant depends on multiple factors.
- One of the key factors for a successful restaurant is choosing the location.
- In this project we will use three variables of a neighborhood to predict optimum location.

Neighborhood characteristics

In this project we will consider three variables in predicting the optimum location for opening a restaurant.

- The presence of other businesses
- The population size.
- Income of the neighborhood

Target Audience

Entrepreneurs who want to open Restaurant in Michigan

DATA

 Neighborhood, population size of the neighborhood and income per capital will be downloaded from Wikipedia

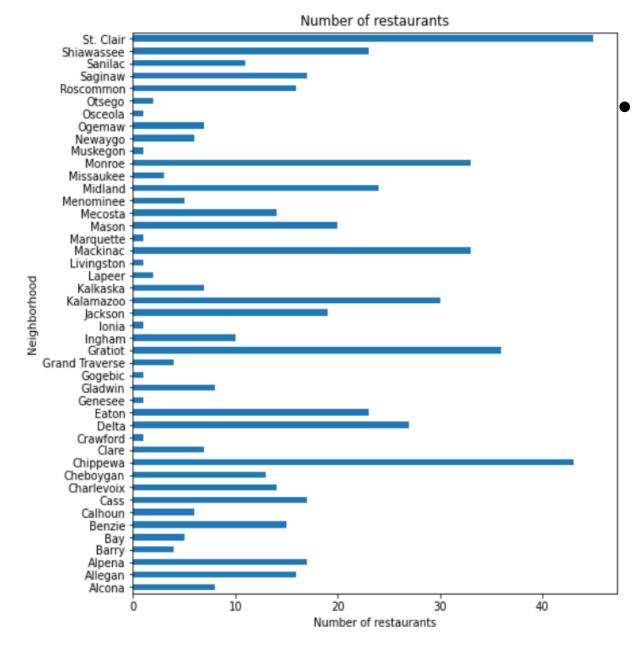
(83, 3)

	Neighborhood	Income	Population
0	Oakland	56138	1202362
1	Leelanau	32194	21708
2	Livingston	31609	180967
3	Washtenaw	31316	344791
4	Charlevoix	28403	25949

DATA

 Coordinates (Latitude and longitude) of the neighborhoods will be collected using
 Geocoder package

	Neighborhood	Income	Population	Latitude	Longitude
1	Leelanau	32194	21708	44.938491	-85.811823
2	Livingston	31609	180967	41.968310	-86.542530
3	Washtenaw	31316	344791	42.253229	-83.838777
4	Charlevoix	28403	25949	45.316530	-85.260720
5	Midland	28363	83629	43.617080	-84.245060

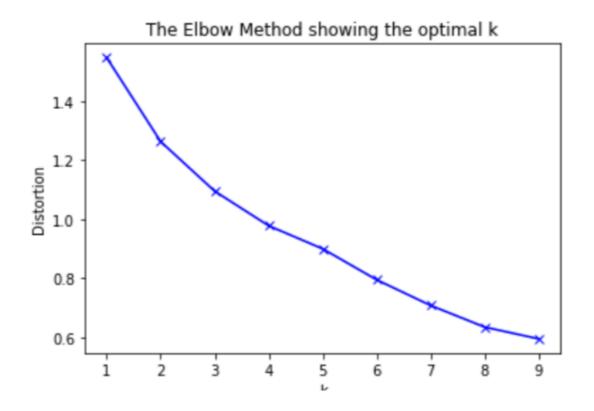


METHOD

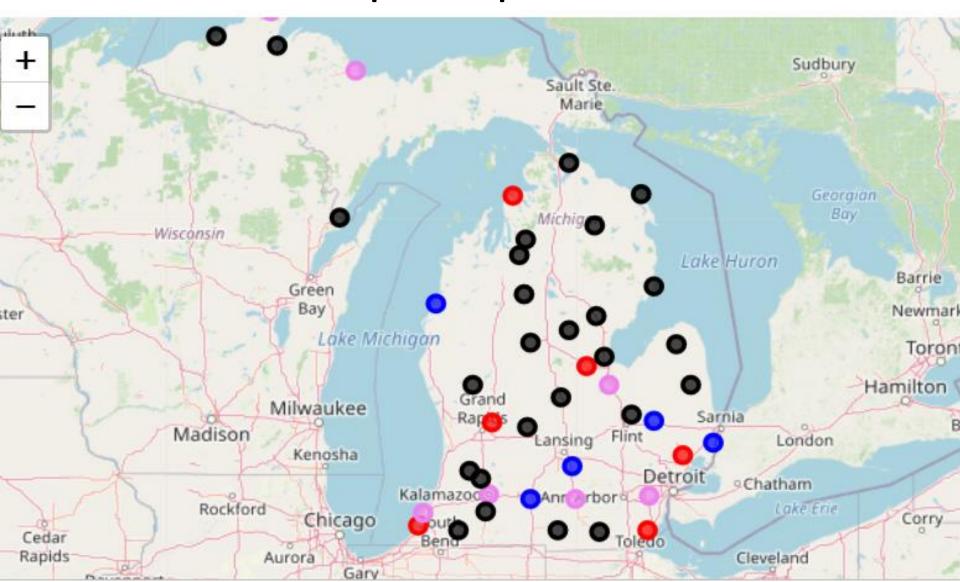
Number Of
Restaurants in
each
neighborhood
will be Counted

METHOD

- Machine learning: K-mean to cluster neighborhood in to four clusters
 - K value set 4 using Elbow Method

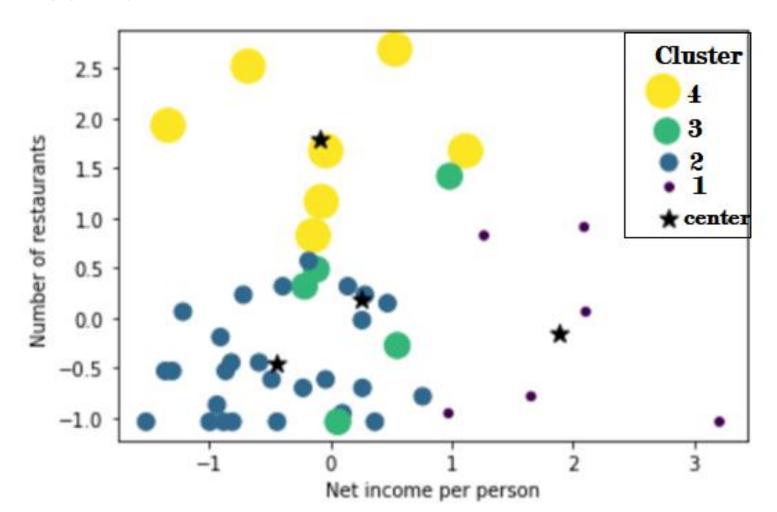


Michigan map with clusters superimposed



RESULTS

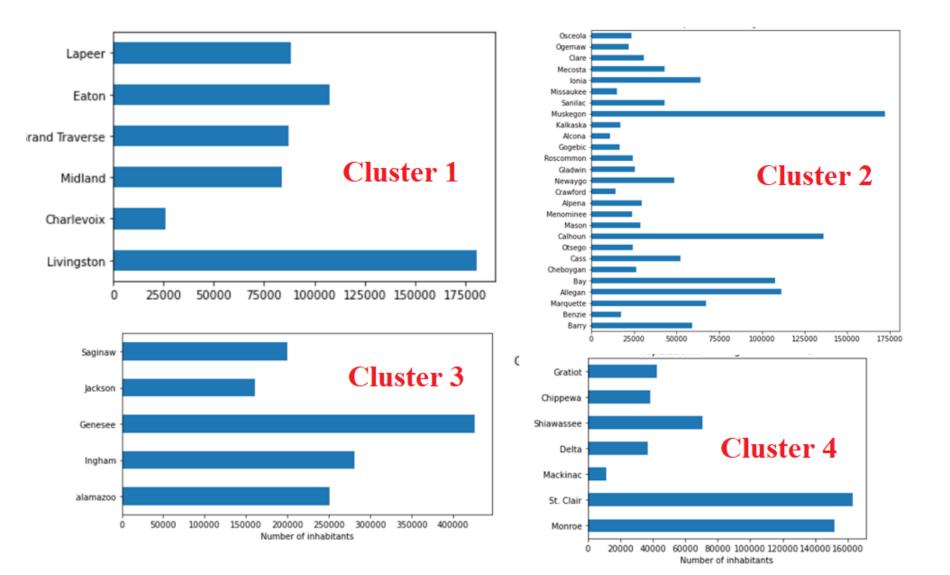
Four clusters number of restaurants and income



RESULTS

- From the result it could be noted Cluster 1 has high income population but fewer number of restaurants. While Cluster 4 has the highest number of restaurants and average income.
- Majority of the neighborhood is in Cluster 2 and they have low income and fewer restaurants. Neighborhoods in Cluster 3 have relatively higher number of income and more restaurants as compared to Cluster 2.

Clusters and number of population



Discussion

- Neighborhoods in Cluster 1 have the optimum location to open a restaurant.
 - Highest income.
- Though, the number of restaurants are few
 - it has a population as much as the other two clusters.
- From cluster 3, **Livingstone** highest population
 - Livingstone could be considered the as the first option.

DISCUSSION

- Second option Cluster 3
 - Highest population with mid income.
- Neighborhoods in Cluster 4 as the third option
 - high number of restaurants though with low income
- Cluster 2 are the least to be considered
 - Low number of restaurants and low income population.
 - This cluster also comprises the majority of the neighborhood.

Conclusion

- From the results it can be concluded that majority of the neighborhoods in Michigan might not be an optimum location to open a restaurant.
- Therefore, a carful consideration is needed when selecting a neighborhood to open a new restaurant.
- From this analysis we found Livingstone to be the most optimum neighborhood to open a restaurant.