

Predicting the Best Locations for a New Restaurant in the City of Detroit

Matthew Ruder

May 1, 2019

1. Introduction

1.1 Background

The city of Detroit has for decades been an example of urban decay. Detroit is known as the Motor City because of the presence of many auto manufacturers, dating back Henry Ford and the advent of the assembly line. Behind this economic engine, Detroit rose to be one of the richest cities in the United States in the early 20th century. However, racial inequalities began to boil over in the civil rights era, spurring white flight to the suburbs which resulted in the decades long decline of the city. The population decline of the city has continued for decades, leading to high rates of crime, inequality, and poverty. All of this culminated in 2013 when Detroit became the largest city to ever declare bankruptcy.

Since the rock-bottom of bankruptcy, the city has begun to rebound. Current civic leadership and private investment have resulted in a renewed optimism in the city. In 2017, Detroit opened a new downtown sports arena for hockey and basketball. In 2018, a ground was broken for what will be the tallest building in Detroit when it is completed. Ford Motor Company bought the derelict Michigan Central Station, long the symbol of Detroit's decline, and promised to restore the building as a beacon of the city's turnaround. While these large scale projects are promising and give hope to Detroit's return to prominence, on a smaller level, locals and others from out of town have been able to help aid the city's recovery through investment of local businesses, including restaurants. Due to the low price of real estate and cost of living, many aspiring chefs have left places like New York City to open up their own restaurants in Detroit. This has resulted in an influx of young talent and Detroit has begun to be recognized as a great food city. However, restaurants are still seen as a risky business venture no matter the city. Therefore, being able to use available data to aid in where to open a new restaurant could mitigate some of this risk.

1.2 Problem

While someone may want to open or invest a new restaurant, factors such as crime or the presence and types of other restaurants for a given area may affect success of the new establishment. Therefore, purpose of this project is to use these data to predict the best locations to open a new restaurant in Detroit.

1.3 Interest

Potential investors and restaurateurs would be interested in learning where to open a new establishment. Using publicly available data will aid in reducing the risk in opening a new restaurant while potentially maximizing potential returns as Detroit continues to recover.

2. Description of Data

2.1 City of Detroit Data

There is a multitude of data available from the city of Detroit via the city's open data portal (<https://data.detroitmi.gov>). The name and location of all the neighborhoods in Detroit can be found. Datasets regarding crime from December 2016 through present available, which will be useful in determining the overall safety of an area. Sales of property through the 1970's also exist as a dataset with almost a million transactions recorded. This could be useful in determining areas that would be affordable to buy a property, though we will limit analysis of this through December 2016 to match the crime dataset.

2.2 Foursquare Data

Foursquare has a robust and large dataset that can be used to find the number and types of restaurants that already exist in an area. The analysis will seem to maximize areas that have some restaurant presence but not an oversaturated area.

3. Methodology

This analysis will consist of four parts. First, we will use the Detroit neighborhood dataset to locate each neighborhood in Detroit. Next, we will begin analysis by clustering each neighborhood together to find neighborhoods that are similar together. Then we will use the city's public safety data to find incidence of crime for each cluster. Finally, using Foursquare data, we will locate the number of restaurants near each cluster center. Using this information, we will make recommendations regarding where to open a new restaurant in Detroit.

To start, we need to find where each neighborhood in Detroit is. To do this we will utilize the Detroit Neighborhoods dataset from the city of Detroit. Since this dataset only has polygonal boundaries for visualization, we will only use the neighborhood names and acres fields. We will use the NHood_Name as this is likely what the neighborhood is known as currently. We know from loading this dataset that there are 208 unique neighborhoods.

To obtain approximate latitude/longitude coordinates, we will use the Arcgis geocoder in Python to get the latitude and longitude of each neighborhood. We will construct

each query as “[neighborhood name] Neighborhood, Detroit, MI”. From each query, we obtain the latitude and longitude of each neighborhood. We now have a dataframe of the acreage with the latitude and longitude. We know from Table 1 that the size of each neighborhood has a wide range, with the smallest being just 26 acres and the largest with 3026 acres. With all this information, we are able to have our basic dataset for analysis and map each location (Figure 1). Visual inspection of the map indicates that some neighborhoods might not have been found correctly and have the same latitude/longitude pair. We can drop the duplicates from the dataset, which leave us with 138 unique neighborhoods.

Table 1: Descriptive statistics for each Detroit Neighborhood

	Acres	Latitude	Longitude
count	208.000000	208.000000	208.000000
mean	440.723619	42.332674	-83.303577
min	26.492368	35.488835	-111.930492
max	3026.947549	42.985260	-70.933360

Figure 2: Initial Sum of Squares curve for optimal K size

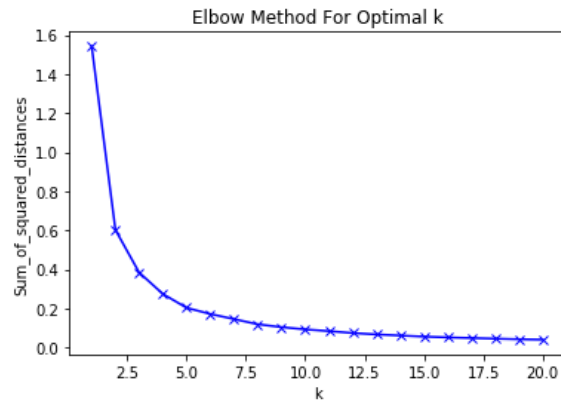


Figure 3: Sum of Squares curve for optimal K size

K-means clustering was utilized with an optimal cluster size of 15. The number of neighborhoods in each cluster ranged from 5 to 17. This makes sense given what we know about the acreage of each neighborhood. Each neighborhood cluster was plotted on the map of Detroit separately (Figure 4) and with just the cluster center (Figure 5).

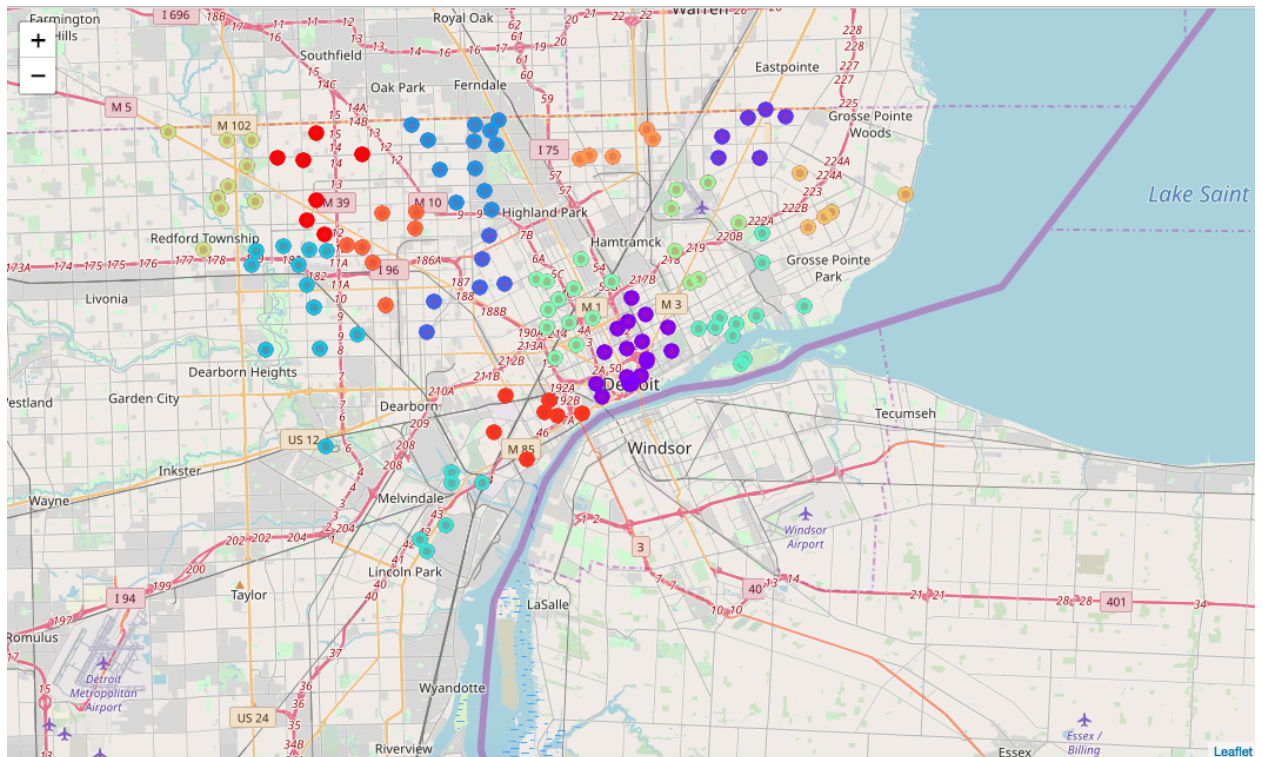


Figure 4: All neighborhoods clustered

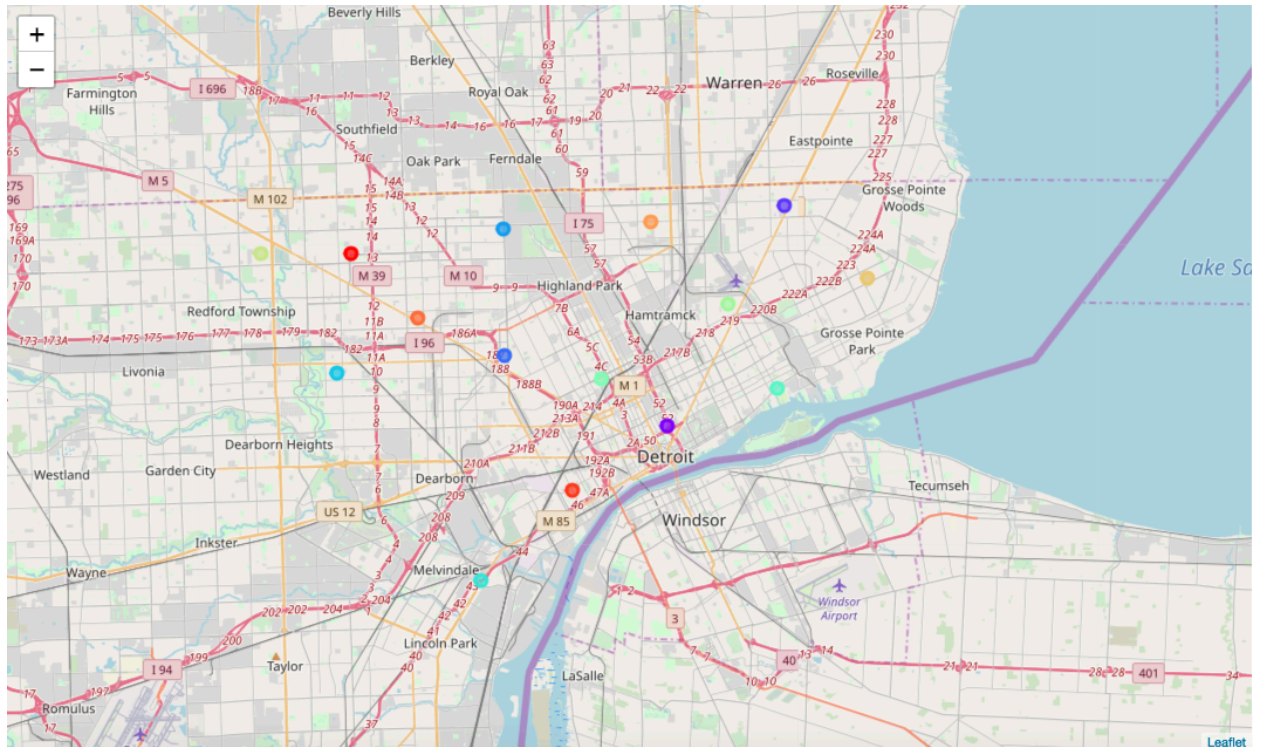


Figure 5: Cluster Centers

4.2 Public Safety Analysis

The Detroit Police Department incident log is publicly available from the City. This dataset is all crime in the city since December 16, 2017. Due to the volume of incidents (almost 200000 in the dataset), we looked to find if there was a representative month to analyze (Figure 6). Since crime appears to be relatively constant across all months, we selected the past month as a snapshot of what is happening. This limits our number of incidents to just 2280.

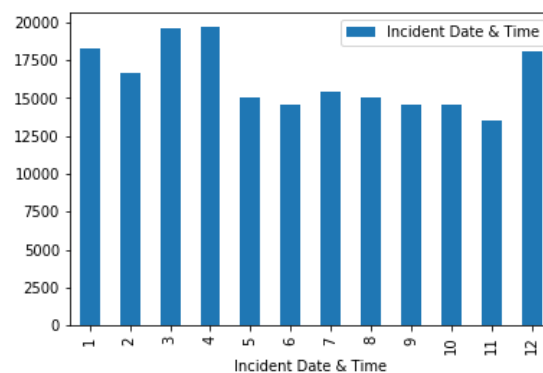


Figure 6: Number of crime incidents by month since December 16, 2017

The distance of each crime was compared to the center of each cluster. We selected all incidents within a half mile radius of the center to assess the safety of each cluster

(Table 2). The top 5 neighborhood finalists were selected because they were the least amount of incidents. These neighborhoods were replotted and visually analyzed (Figure 7). Cluster 6 and Cluster 6 were eliminated since they are near the Detroit city limits and our client would like to be as close to downtown as possible. The remaining neighborhoods are recodified as Southwest, Southeast, and Northeast for simplicity.

Table 2: Number of Crime Incidents Within 0.5 Miles in Past Month

Cluster	Lat	Lon	Incidents
6	42.277712	-83.154569	7
9	42.397591	-83.009826	16
7	42.361056	-82.981332	23
14	42.316804	-83.101450	26
10	42.419292	-83.284067	28
8	42.365287	-83.084392	35
11	42.408907	-82.928162	36
3	42.375067	-83.141000	42
2	42.440001	-82.977042	43
13	42.391567	-83.191584	45
12	42.433206	-83.055197	48
0	42.419454	-83.231023	50
5	42.367863	-83.239070	55

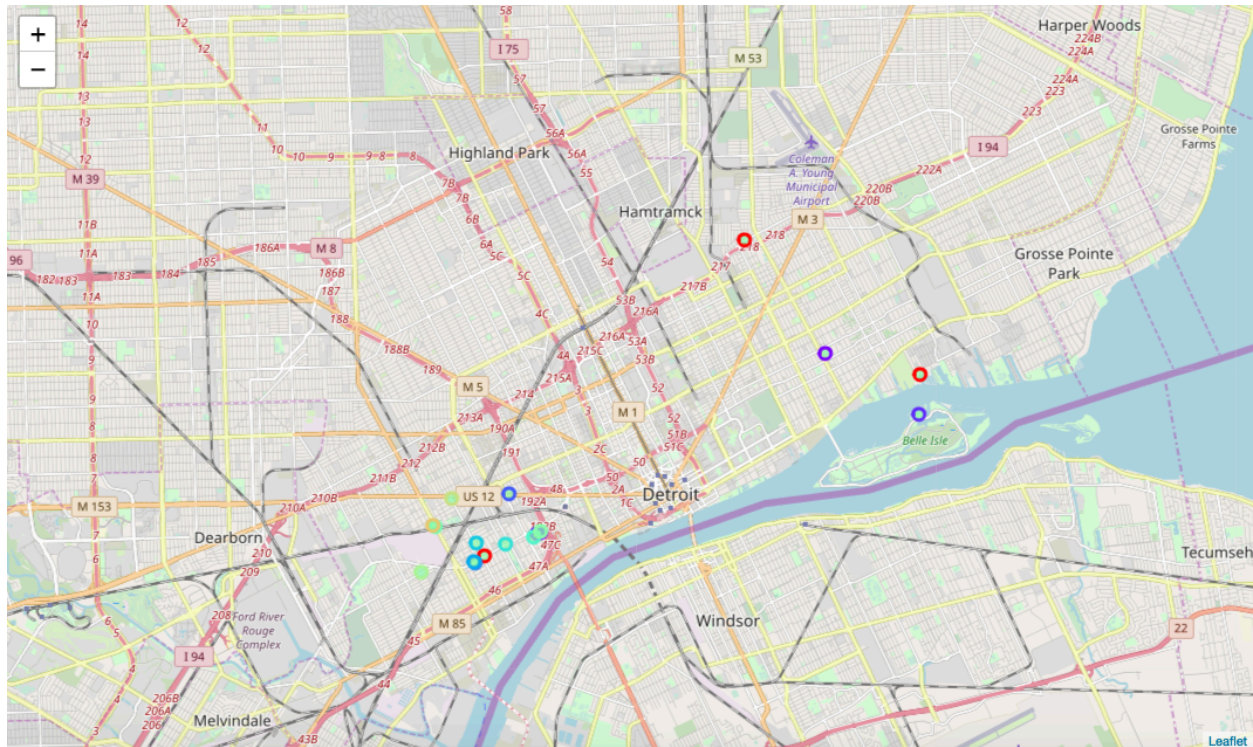


Figure 8: Mapped existing restaurants near each cluster

5. Results and Discussion

Our analysis shows that there are 3 high potential neighborhood clusters in the City of Detroit to develop a new restaurant. We started with 200 unique neighborhoods and clustered them into 15 unique areas. The clusters were spread all over the city and led us to look at other factors to narrow the search. The city of Detroit's crime dataset is a convenient way to quickly assess the safety of a given cluster. Analyzing the crime for each cluster allowed us to reduce the number of candidate neighborhoods from 15 to 5. Using geography, we were able to remove another 2 clusters from analysis, leaving us with 3 candidate areas: Southwest, Northeast, and Southeast. Each cluster has its pros and cons.

The Northeast cluster is not close to the downtown core. It also only has one restaurant near its core. While it has the lowest incidence of crime, this may be indicative of lack of any development when paired with the restaurant density. It would not be recommended as an option based on the data.

The Southwest cluster is in close proximity to the downtown area. It also has a network of existing restaurants in the area, which could indicate that it is a thriving area that could support more restaurants. Additionally, the types of restaurants it has are mostly Mexican, meaning that other developments could be successful and stand out in this

area. It also is close to new developments in the area like the Ford rehabilitation of the Michigan Central Station and will likely be a thriving area very soon. Crime is also relatively low. The number of restaurants in the area could also be seen as a negative, as there would be more competition, especially if the desired restaurant type is Mexican.

The Southeast cluster geographically is situated on the other side of downtown, opposite of the Southwest cluster. There are currently not many restaurants in this area, so it could be ripe for development. Perhaps the best characteristic is close to Belle Isle, which thrives during when it is not winter, as it is Michigan's most popular state park with approximately 4 million people visiting annually. It also is the site of the Detroit Grand Prix every spring. Crime-wise is about equal as the Southwest cluster. It also includes neighborhoods such as Indian village that are healthy neighborhoods and proximal to the affluent suburb of Grosse Pointe.

6. Conclusions

The purpose of this investigation was to find locations that might be ripe for development of new restaurants to aid potential investors and/or restauranteurs in selecting a site to open their new establishment. By clustering neighborhoods together, we were able to do find similar neighborhoods and allow us to narrow our search. We then used indexed each neighborhood cluster by incidence of crime to narrow our search to 5 candidate neighborhoods. Because of the desire of our investors to be close to the downtown core, this eliminated two additional clusters, leaving us with 3 candidate neighborhood clusters. The final determination of where to open the new restaurant will be left up to the stakeholders based on other factors such as microneighborhood factors, availability and affordability of real estate, other developments in the area, promity to downtown, and areas of interest that will drive traffic to the restaurant. That said, there seem to be few downsides to the Southeast cluster and would be our recommendation for a new restaurant. However future developments near the Southwest cluster make it intriguing in the future.