Capstone Project

The Battle of Neighborhoods

Introduction

Investors are always looking for new ventures to increase their worth and one of the most popular small business ideas is a restaurant. As we all know, one of the most important factors in a business is its location and many factors can affect this choice. A poor decision can result in a huge loss of finances. I believe entrepreneurs in the restaurant industry will be interested in this study.

For this project, I will use Population and Crime data, and the FourSquare tool to acquire the locations of restaurants and a setlist of general venues that promote traffic for Chicago and Toronto. My intention is to find any connection between proximity to regular pedestrian traffic and the relative safety of different areas in the city.

Data

For this investigation, I have acquired major crime records from the official websites of the police departments in Toronto and Chicago. this was made simple by the fact that these sites provide their databases freely in a few formats.
Population data was acquired through Census websites for Toronto and Chicago

Methodology

The crime index is highly generalized and often does not indicate the level of crime in an area as perceived by the inhabitants. Because they interact more closely within a city, the level of crime in different areas will also affect businesses. For this, I use folium and choropleth maps to highlight areas of low to high crime, as would be perceived by the population.

After showing the areas of low to high crime on the map, I will then plot the business areas, and use k-means to find the center of the main clusters.

I will also use DBSCAN to perform more intelligent clustering of businesses and restaurants. The idea is to establish a visual pattern, with help from machine learning, to decide on areas that may be a potential good choice for the opening of a small business.

Results

The maps clearly show a correlation between general business locations and restaurants. There is also more than a vague correlation between areas of low crime and high business activity, especially in the case of Chicago.

Toronto

Image 1 shows the business cluster centers of Toronto, found using K-means. Red circles show the centers for restaurants, and blue-green ones show the centers for general businesses. The shaded areas on the map represent the level of crime for each zone, with bluer shaded meaning low crime, red shades mean more.

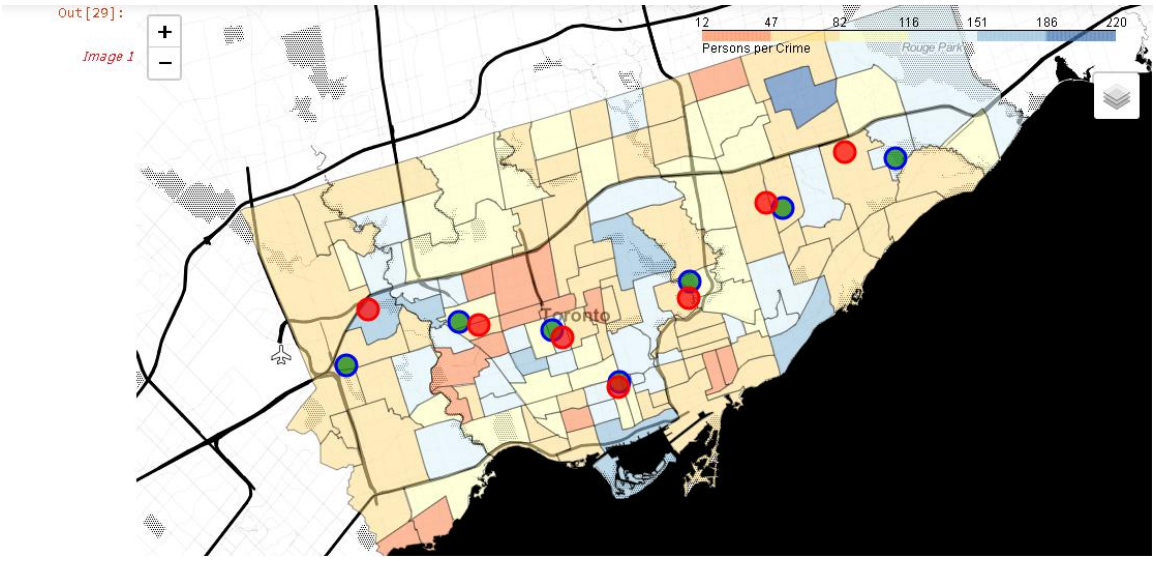


Image 2 shows the main restaurant clusters in Toronto, acquired using DBSCAN. This gives a good idea of which areas are more densely packed with restaurants.

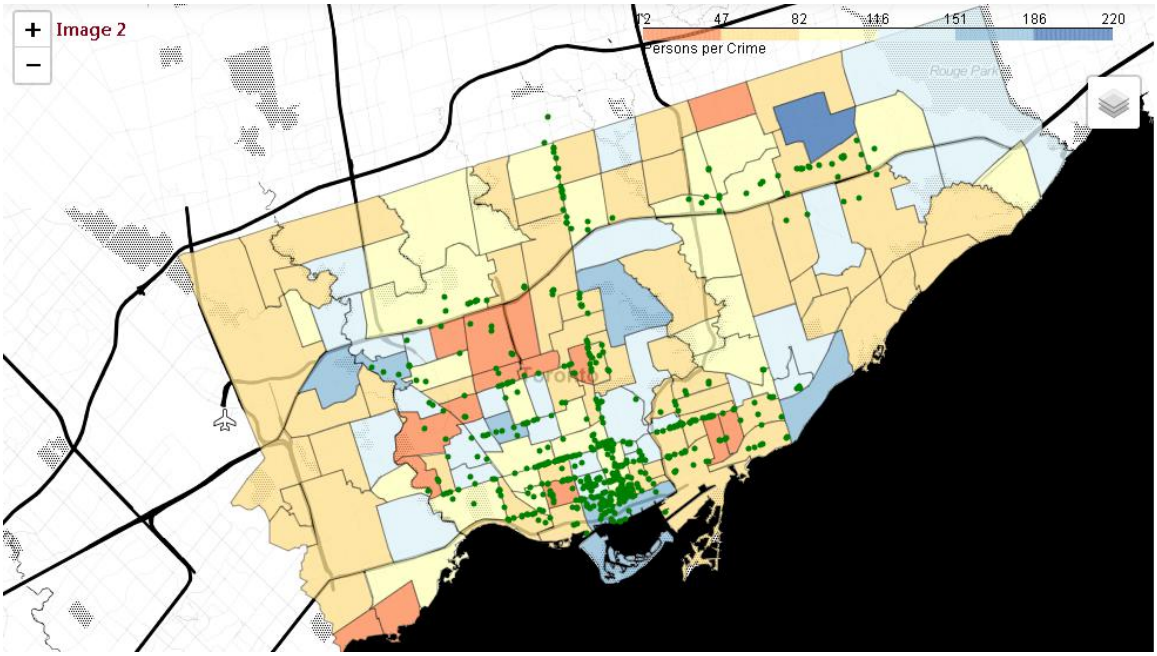


Image 3 contains the major clusters for both restaurants and general businesses. It can be seen that both clusters coincide, and also its more clear that businesses generally follow areas of low crime.

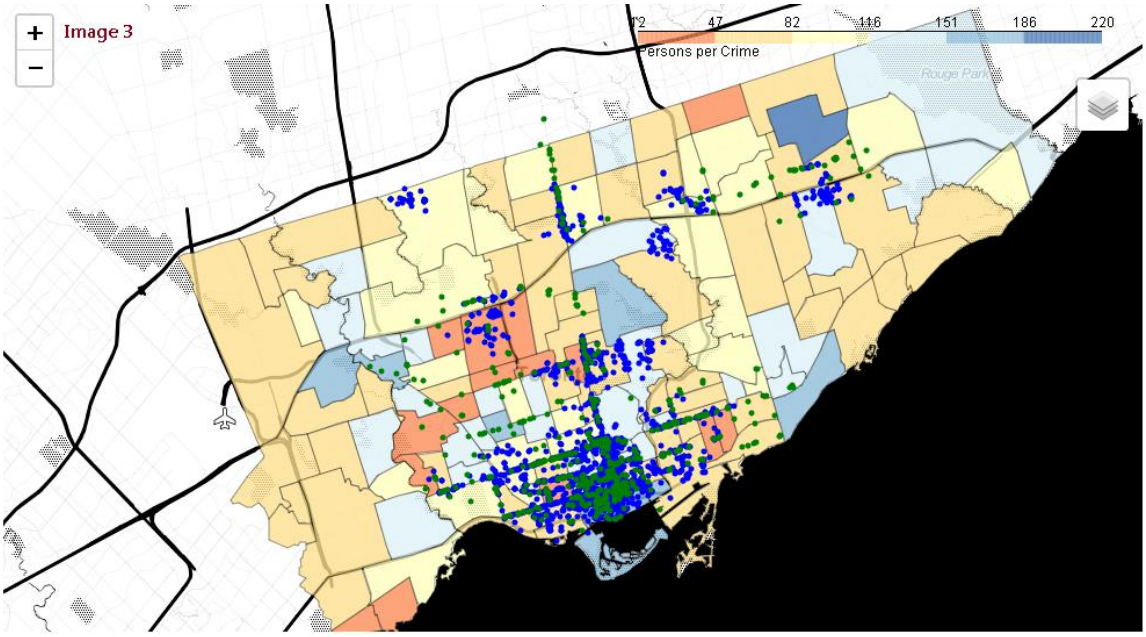
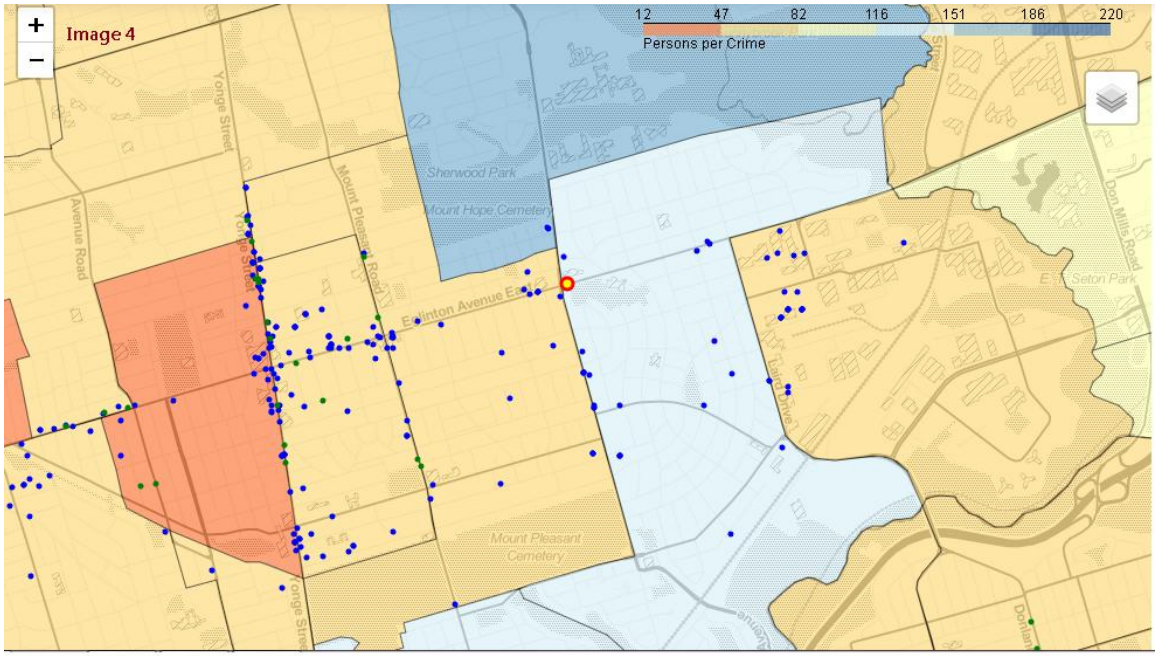
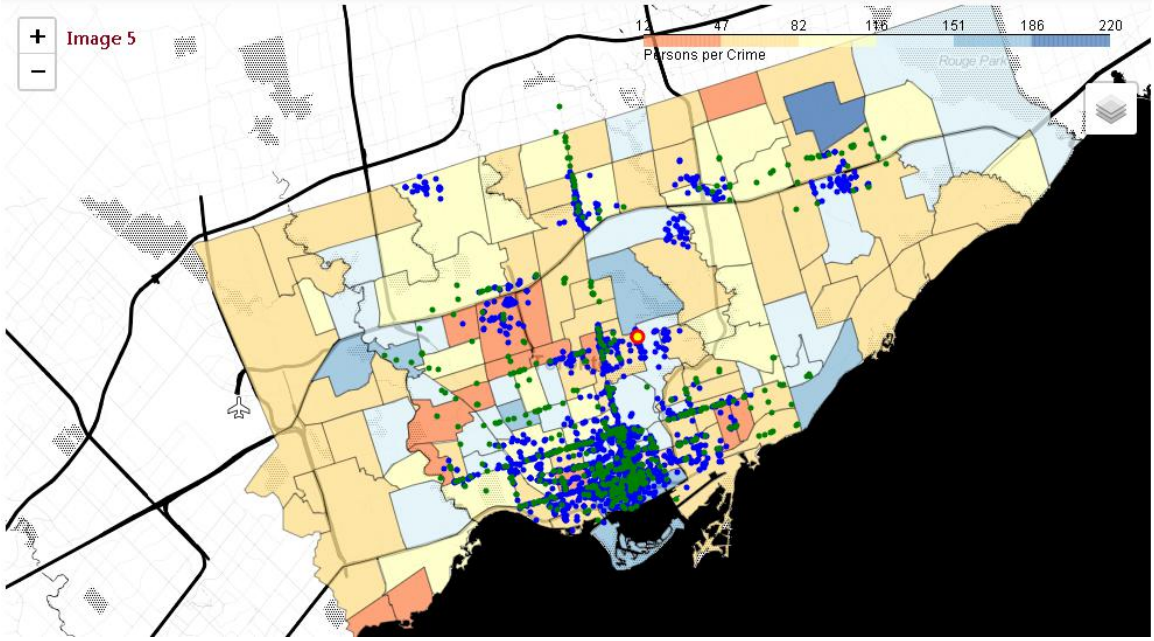


Image 4 and Image 5 show the spot I have selected for the starting a business, based on visual inspection of the map and the group clusters. It is located in an area containing no restaurant cluster, so competition will not be fierce. It is also close to many active businesses, meaning it can benefit from the traffic generated as a result.





Chicago

Figure 1 shows the business cluster centers of Chicago, found using K-means. Red circles show the centers for restaurants, and blue-green ones show the centers for general businesses. The shaded areas on the map represent the level of crime for each zone, where bluer shades mean less crime, red shades mean more.

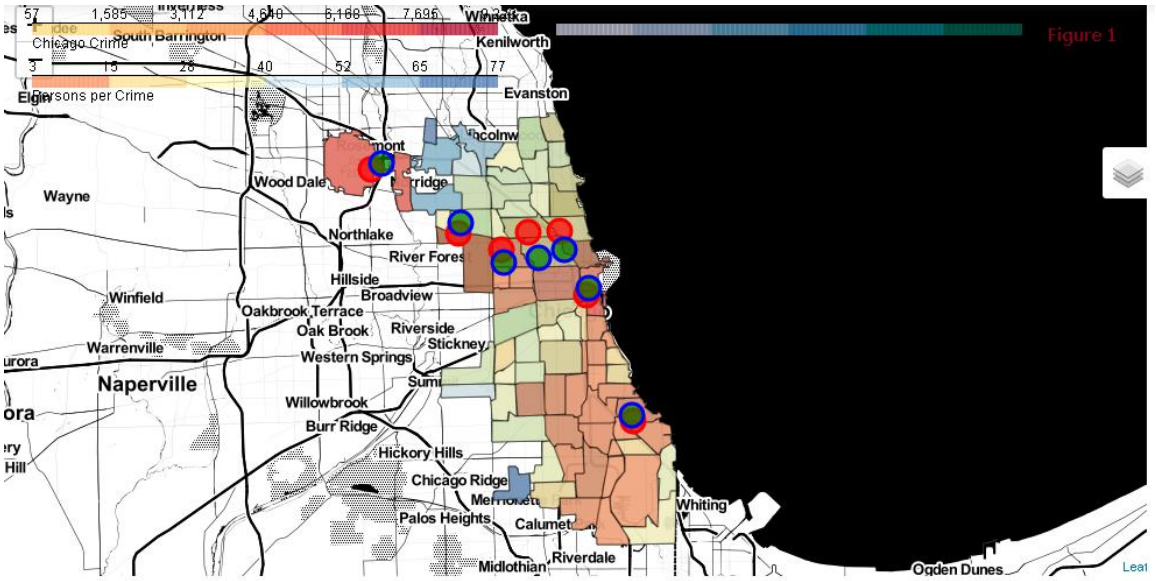


Figure 2 shows the main restaurant clusters in Chicago, acquired using DBSCAN. This gives a good idea of which areas are more densely packed with restaurants.

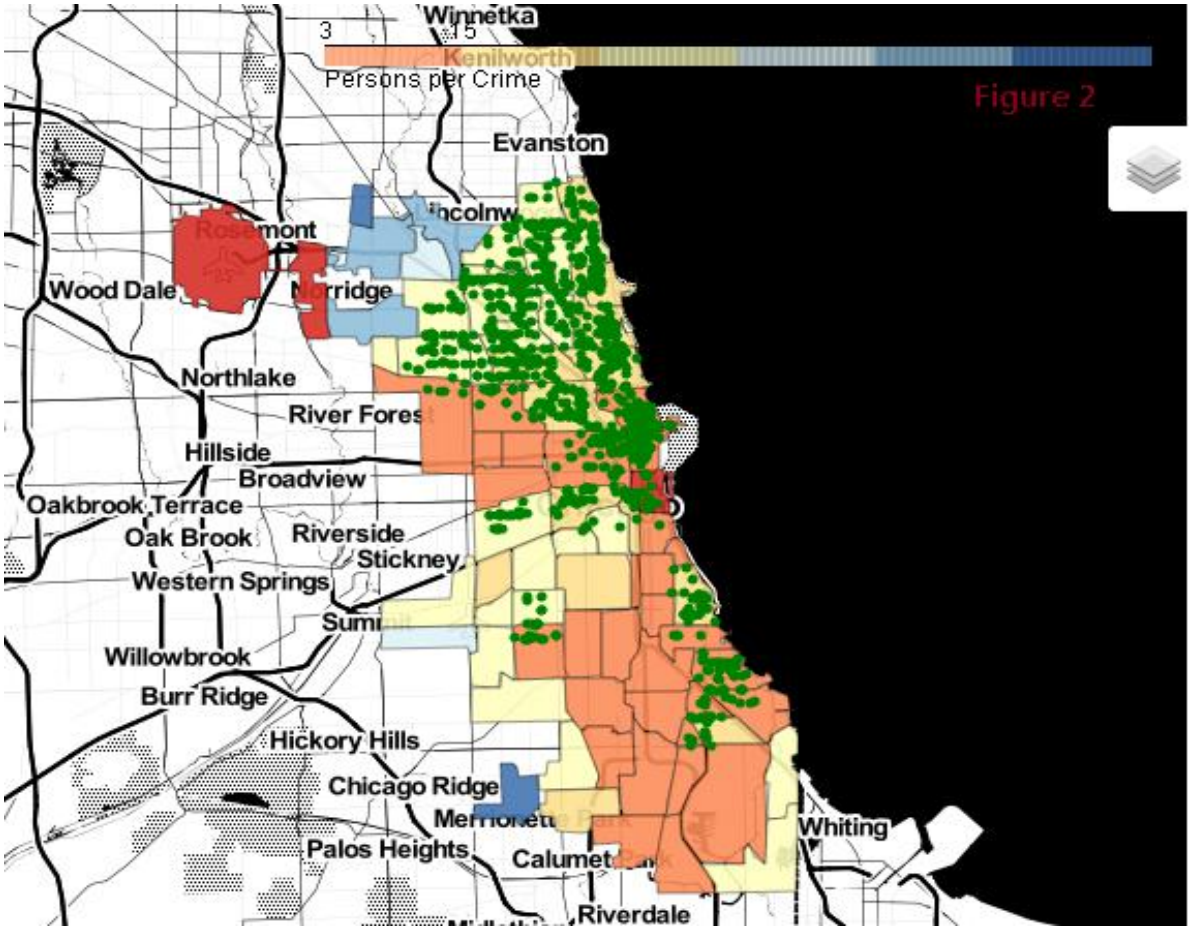


Figure 3 contains the major clusters for general businesses. Take note of how strongly they align with the lower crime areas of Chicago.

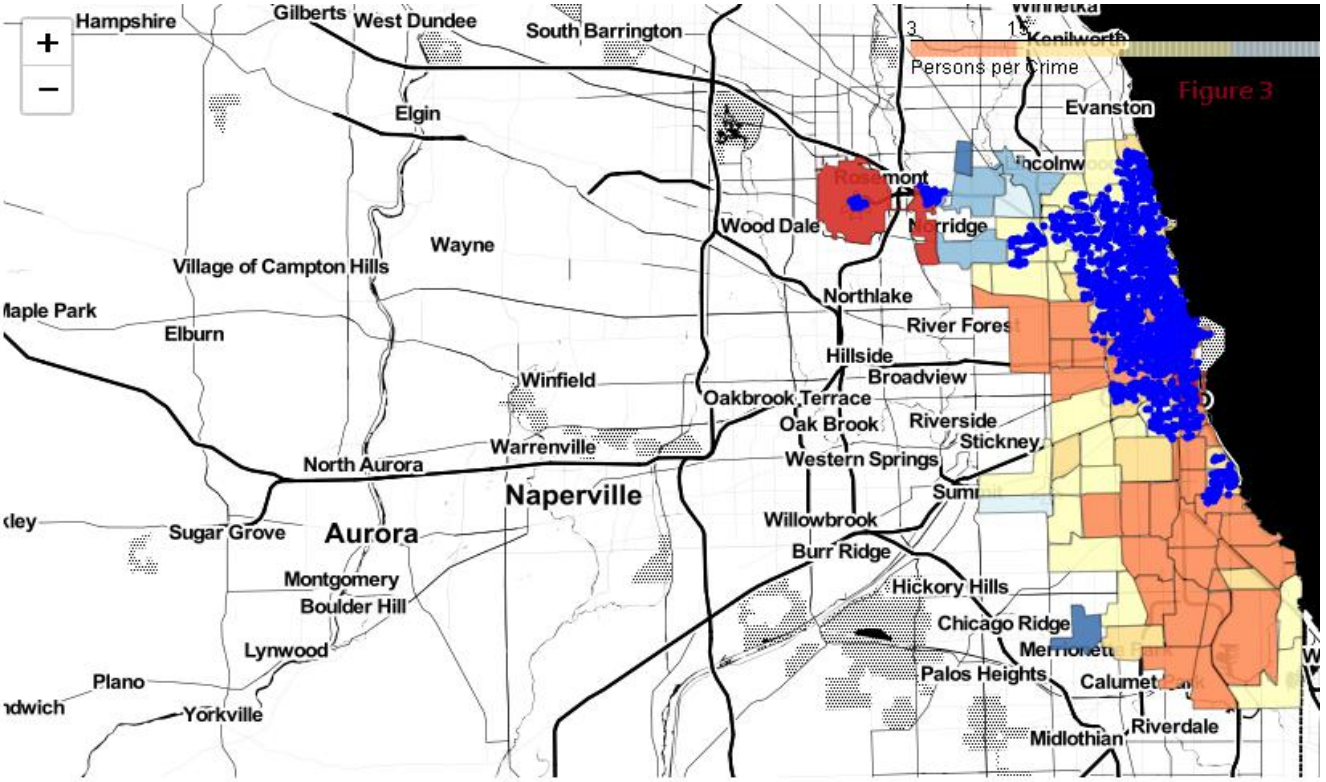


Figure 4 shows the spot I have selected for the starting a business in Chicago, based on visual inspection of the map and the group clusters. It is located in an area containing no restaurant cluster, so competition will not be fierce. It is also close to many active businesses and the main business center of Chicago, meaning it can benefit from the traffic generated as a result. It is also in an area of relatively low crime.

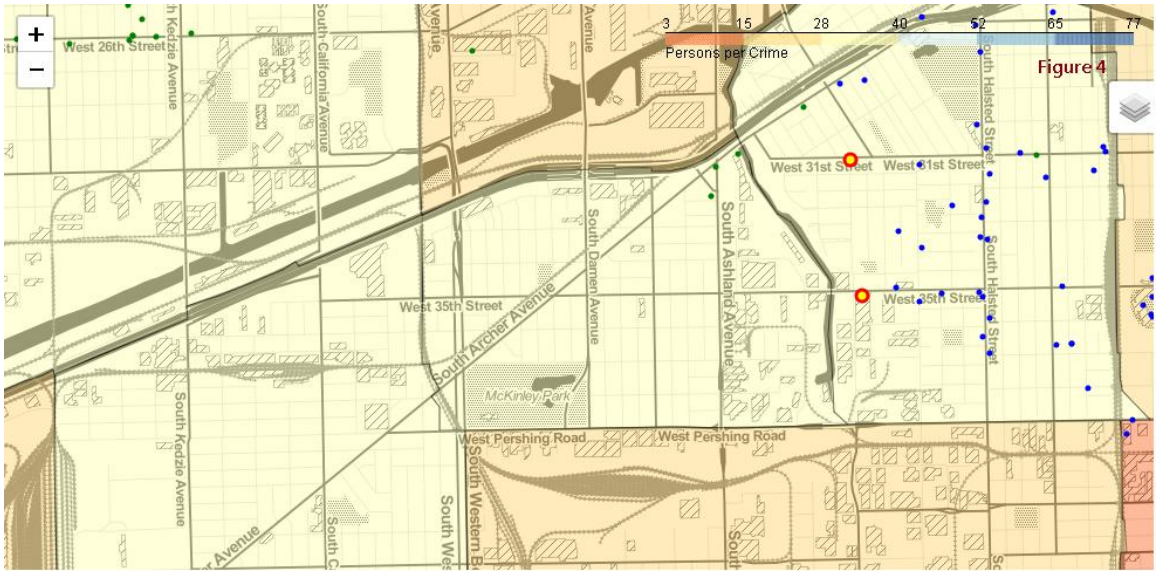
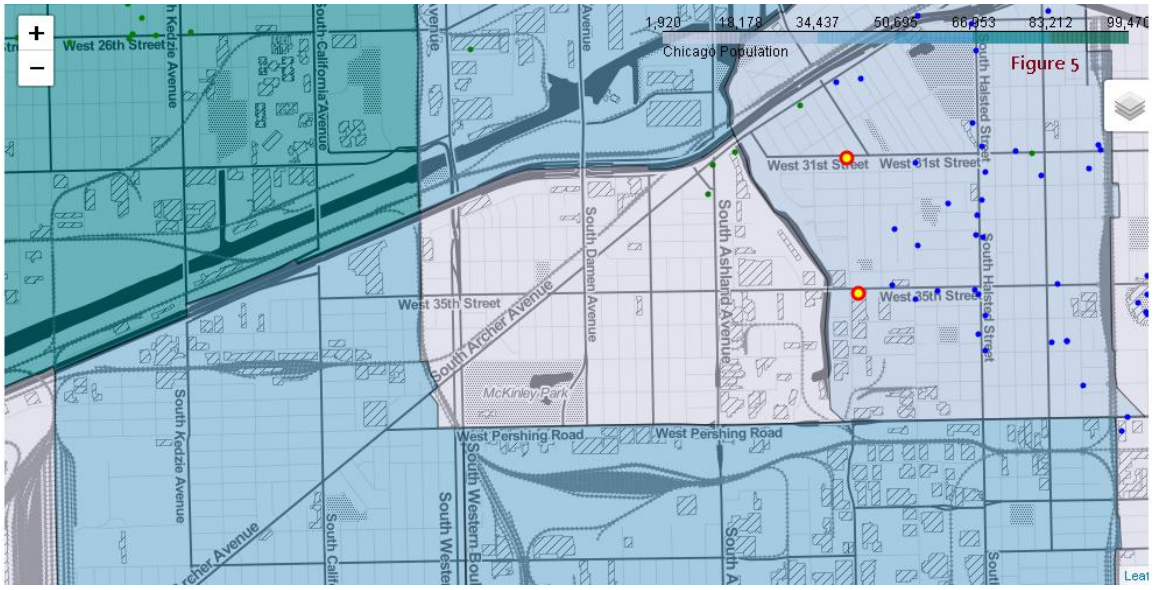
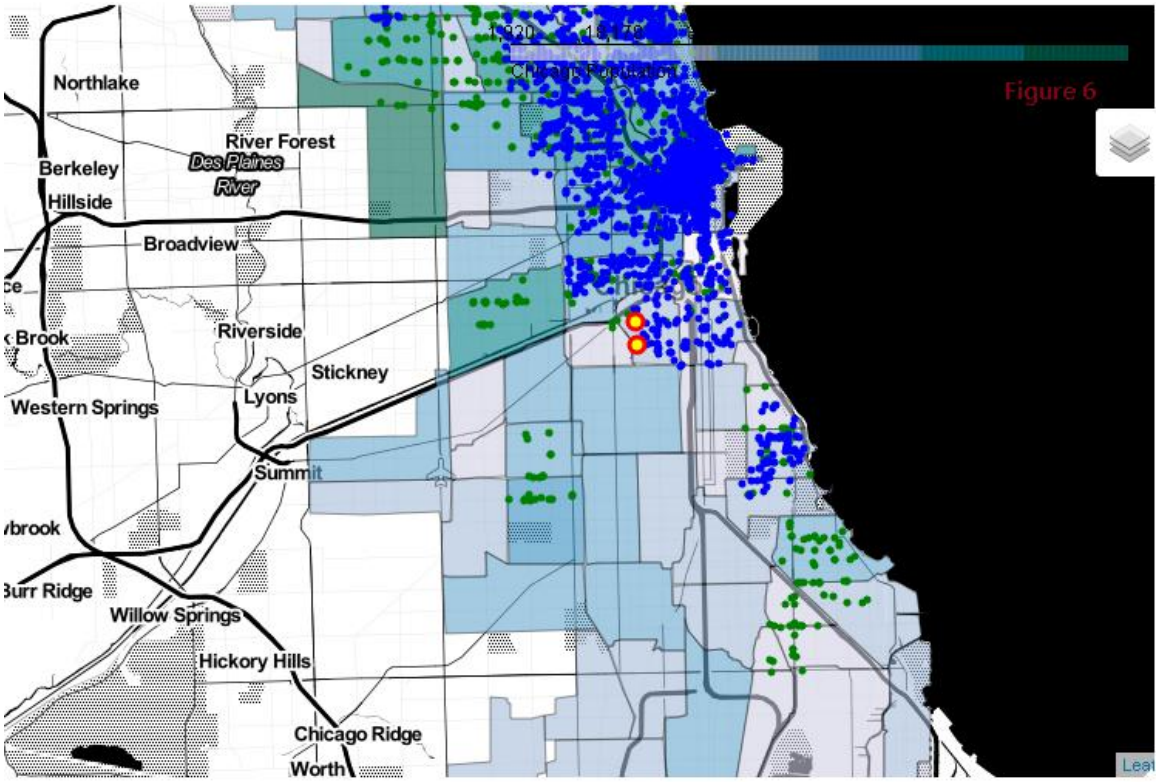


Figure 5 and Figure 6 shows the population of the surrounding areas. This also works in favor of opening a restaurant or fast food business at this location.





Discussion

I think it safe to infer that areas of low crime and high business activity are associated with the success of restaurants. Using both forms of machine learning revealed strong relationships between restaurant locations and the two features.

Low crime rates played a more significant role than I expected. I suppose it is safe to assume that if persons are able to move about more safely and freely, they are more likely to patronize businesses in that location.

Traffic generated from persons simply having to be in an area due to work, leisure activities and shopping is also quite significant, as seen from the maps. It was quite clear that clusters of high general business activity are usually accompanied by a cluster of restaurants.

Using this observation, I selected areas close to clusters of general business activity, that lacked an accompanying restaurant cluster, while remaining In a zone of low-crime rate.

Recommendations

While I believe the findings of my project to be sound, I recommend further investigation into the rent of the areas suggested. This data was not considered but it may have the power to change a 4th or 5th choice into a 1st or 2nd.

Conclusion

Based on my findings:
1. In the city of Toronto, anywhere along the corners of Eglinton Avenue East and Bayview Avenue is my suggestions for opening a fast food outlet or Restaurant.

2. In the city of Chicago, west 31st and 35th Streets, between South Ashland Avenue and South Halsted Street, are my suggestions for opening a restaurant or fast food outlet.