Where should I locate a marijuana dispensary?

1) Introduction

Marijuana is newly legal in Canada, Toronto and Ontario is the largest city in Canada. Opening up a marijuana dispensary in Toronto sounds like a great business idea. Question is, where in Toronto should the new marijuana dispensary be located?

Research on this topic indicates that dispensaries are generally located in:

- Lower socio-economic neighborhoods
- In a neighborhood populated by strip clubs and alcohol dispensing venues
- Near a freeway offramp
- Near the city limits

2) Data

Toronto census data used to determine socio-economic status of neighborhoods (www.toronto.ca/city-government/data-research-maps/open-data), Toronto postal data (www.wikipedia.com) provides geo coordinates for neighborhood mapping, Foursquare (www.foursquare.com) provides information on local venues and general neighborhood characteristics, and GoogleMaps (maps.google.com) provides geo data on venues that are not adequately covered by Foursquare (e.g., strip clubs and marijuana venues). Foursquare data is lacking information on these venues, possibly due to self-report bias (many people don't want to announce their visits to strip clubs) and selection bias (a select population group are heavy Foursquare users, possibly not the same population group that frequents strip clubs). For this reason Foursquare data was augmented by GoogleMaps data.

3) Methodology

a) After reading in to a pandas dataframe, Toronto postal data cleaned and empty fields removed.

b) After reading in to a pandas dataframe, Toronto census dataframe combined with Toronto postal dataframe.

c) GoogleMaps data on existing Toronto marijuana venues and strip clubs read into dataframe

d) Features extracted from these datasets and plotted and examined for a discernable visual relationship between the location of marijuana venues and:

- Neighborhood socio-economic status
- Neighborhood size
- Neighborhood density
- Locations of strip clubs

Figure 1

Figure 1: Spending power of Toronto neighborhoods displayed as heat map, overlaid with locations of existing marijuana dispensaries (green dots)

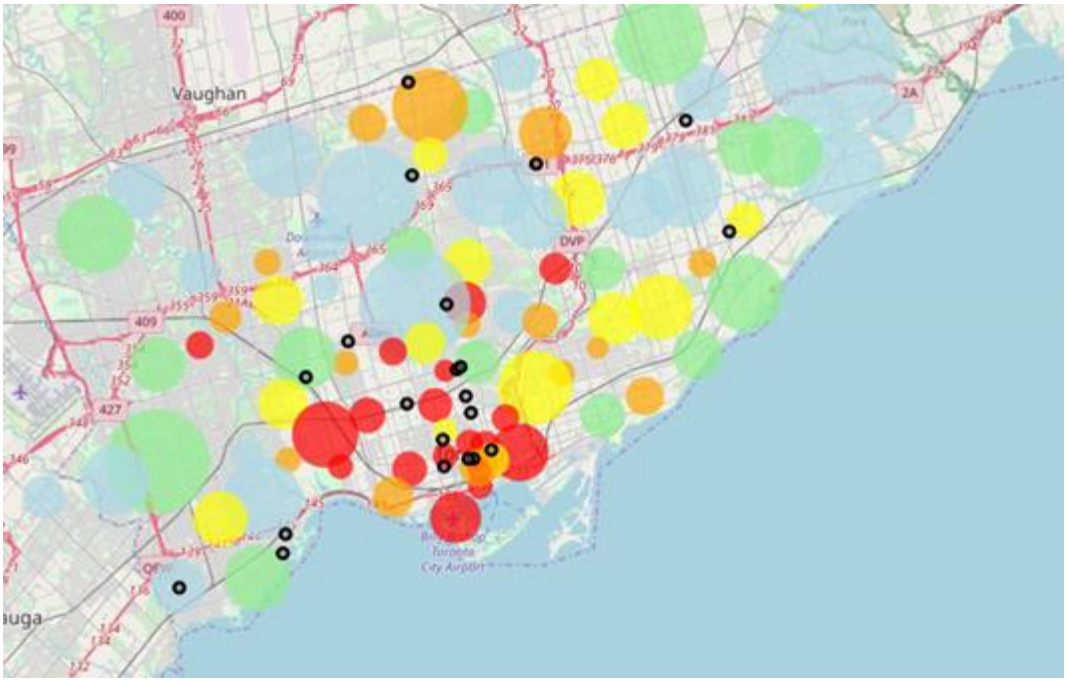

Figure 1

Figure 2: Neighborhood density color-coded (Red-Orange-Yellow-Green-Blue) in order of decreasing density, overlaid with locations of existing marijuana dispensaries (green dots).

Figure 2

Figure 3: Geographic relationship of marijuana dispensaries (green dots) and strip clubs (red dots).

e) Neighborhood characteristics, obtained from Foursquare data, used in Kmeans with a cluster size of five to determine optimal location for marijuana dispensary.

f) Elbow analysis performed on the clusters to determine optimal cluster size. As shown in the chart below, the optimal cluster size is 6.

g) Clustering re-run with optimized number of clusters, locations of marijuana venues and strip clubs overlaid on this map to arrive at a suitable location for the next marijuana dispensary.

Note: In order to minimize file size and file load time,

i) map of neighborhoods clustered into five groups overlaid with locations of existing venues available at https://drive.google.com/open?id=1Qqwdt3g04ph9qLapaCk8UHPEX6wIc4P1
ii) map of neighborhoods clustered into six groups overlaid with locations of existing venues available at https://drive.google.com/open?id=12QDfTXyZyZmLtNKAqrtwmKN7Qg5R2VtR

4) Results

Existing locations of marijuana venues and strip clubs overlaid on optimized clustering provided insight as to the best locale for a new marijuana dispensary. Neighborhood clustering into six groups (predicted as optimal by Elbow Analysis) provided a significant improvement in neighborhood identification.

5) Discussion

Geo-mapping neighborhood spending power, neighborhood population density, neighborhood similarity, and mapping locations of existing venues en toto provide guidance for optimally locating a

new marijuana dispensary. None of these features individually identified the optimal location; however, the combination of all these features (especially the location of existing venues) enable a potential investor to identify a new location.

Initially, Kmeans clustering was run using a cluster size of 5 to segregate the neighborhoods Using elbow analysis, it was determined that a cluster size of 6 was optimal. Rerunning the clustering algorithm with the optimal cluster size of 6 provided significant improvement in the ability of clusters to segment neighborhoods.

Geomapping was required to select the optimal location due to several of the identifying feature of marijuana dispensaries – proximity to freeway offramp and near city boundaries, as these features are easily discernable on a map.

Marijuana is an emerging market and there is a business need to optimize the location of new marijuana dispensaries. This project could be expanded by including other marijuana-friendly cities – Seattle, Long Beach, San Francisco, CA.

6) Conclusion

Geo-graphing neighborhoods and existing venues, along with geo-visual information such as freeway locations, shows that the area near Pearson Airport, with existing strip clubs, no existing marijuana dispensaries, freeways offramps and neighborhoods similar to those of existing marijuana dispensaries would be a good location for new dispensary.