BY NIKITA

IBM CAPSTONE PROJECT: EXPLORING RESTAURANTS IN AN INDUSTRIAL DISTRICT

CONTENTS

- Introduction
- Business problem
- Data
- Methodology
- Results
- Discussion and conclusion

INTRODUCTION

- Etobicoke is an industrial district in Canada. Due to the fact that it is industrialised, the population density is lower than the surrounding districts. However, restaurants and shopping malls are located in this district.

BUSINESS PROBLEM

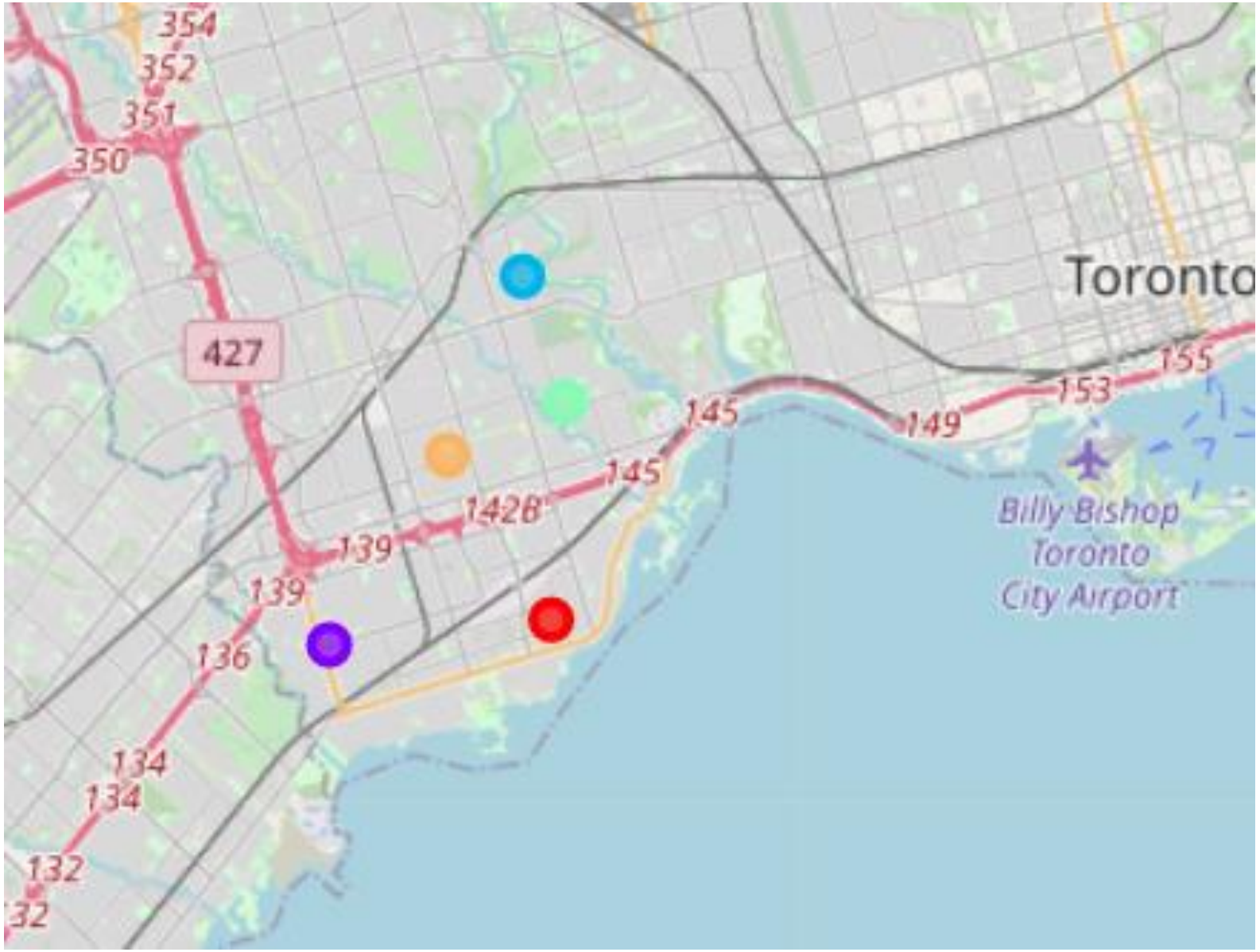
- Given that Etobicoke is such an industrialised area, it is likely that the district is not saturated with restaurants.
- If restaurants are clustered, then as more restaurants open, the market share of the existing restaurants in the surrounding area decreases, which has the potential to impact profitability of these existing venues.
- If the restaurants are not clustered, there is less risk of losing market share.
- Thus, the target audience of this study is existing restaurant owners looking to protect their existing business or to further expand their franchise.

DATA

- Sources:
  - Neighbourhoods: these will be obtained from Wikipedia using webscraping tools in Python.
  - Geocoding: the geographic coordinates of each neighbourhood will be obtained using Google Geocoding API
  - Restaurants: the geographic coordinates of each restaurant will be obtained using Google Geocoding API and Foursquare API

METHODOLOGY

- To extract the list of neighbourhoods of Etobicoke, web scraping was used. This was done through Python. The data was cleaned and formatted as a dataframe.
- The geographical coordinates of each neighbourhood was determined using Foursquare API and the Geocoder package. The data was formatted as a dataframe and visualised.
- Furthermore, clustering was performed to analyse the presence of venues in a particular neighbourhood. The method used was k-means. K-means clustering identifies k number of centroids, and then allocates every data point to the nearest cluster.
- The neighborhoods were clustered into 5 clusters. The results would indicate which neighborhoods have had a concentration of certain types of restaurants.



RESULTS

Based on the results, the most commonly occurring types of restaurants in these clusters include:

- Bakeries
- Wing joints
- Sandwich shops
- American restaurants

However, one cluster (green) was more populated by shopping venues rather than restaurants

CONCLUSION

- Thus, based on the results observed, bakeries, wing joints, sandwich shops and American restaurants should aim to protect their existing market share within their neighbourhoods. A potential area for expansion, where almost no other restaurants exist, is in the area defined as cluster 3, and the most market share can be gained here if expanding or starting a new restaurant.