Coursera Capstone
IBM Applied Data Science

Open a New Restaurant in Downtown Portland

Oregon

By: Tuan Nguyen

July 2020

Business Problem

The objective of this capstone is to analyse and select the best locations in Downtown Portland to open a new restaurant.

Using data science methodology and machine learning techniques like clustering.This project aims to provide a solution the question of in Downtown Portland, where to open a new restaurant?.

Data

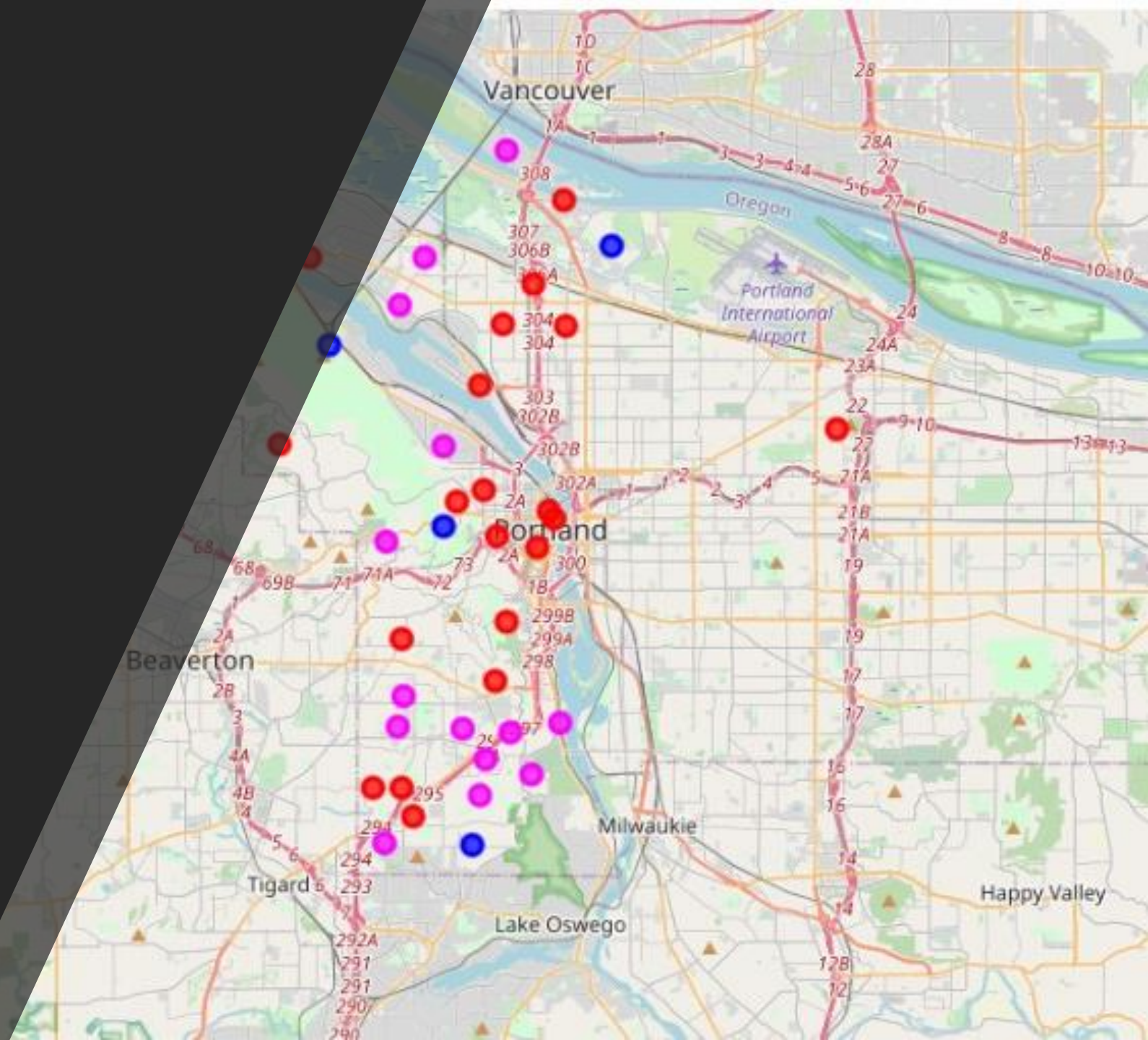
- List of neighborhoods in Portland. Downtown Portland specializes in South West and North West Portland.
- Latitude and longitude coordinates of those neighborhoods. This is required in order to plot the map and also to get the venue data.
- Venue data, particularly data related to restaurants. We will use this data to perform clustering on the neighborhoods.

Methodology

- The list of neighborhoods in Portland is in Wikipedia page (https://en.wikipedia.org/wiki/Neighborhoods_of_Portland,_Oregon)
- Foursquare API and geocoder to get all the information we need like latitude and longitude and other places like restaurants, schools, workplace and so on
- We will look at top 100 venues that are within a radius 1000 meters.

Results

- Cluster 0(Pink): Neighborhoods with moderate number of restaurants
- Cluster 1(Red): Neighborhoods with high number of restaurants
- Cluster 2(Blue): Neighborhoods with low number of restaurants



Discussion

- As observations noted from the note, most of the restaurants are concentrated in cluster 0 and 0. On the other hand cluster 0 has lower frequency of restaurants than cluster 1.
- While cluster 2 shows a very low number of restaurants. This also shows it is more competitive to open a restaurant in cluster 1.

Recommendations

- Therefore this project recommends shopkeepers to capitalize on these findings to open a new restaurant in neighborhoods in cluster 0, then 1, but not 2.

Conclusion

- In this project, we have gone through the process of identifying the business problem, specifying the data required, extracting and preparing the data, performing machine learning by clustering the data into 3 clusters based on their similarity, and providing recommendations to shopkeepers.
- To answer the business questions, the answer by this project is: The neighborhoods in cluster 0 is the most preferred location to open a new restaurant because of moderate frequency, unlikely cluster 1 is very competitive.