

Social Food Ordering App

Business Problem

In workplaces, employees have lunch hours where they can have a break and get something to eat before going back to work. Employees can do either of the following:

- Use the workplace's in-house catering service.
- Bring the food from home and eat inside.
- Go out and eat at dine-in restaurants.
- Go out get their food and come back to the workplace and eat.

We want to create an app that helps the employees place an order with a nearby restaurant, pay online, and go pick up the food when it is ready while avoiding staying in the line.

For the app to be successful we should incentivize both consumers and merchants to participate. Restaurant owners need to be convinced that they will get more customers and hence more profit. Customers need to be convinced that they have a good amount of choices nearby to use the app.

For the start, we need to talk to restaurant owners to participate. This is an expensive operation because we need to give them hardware, spend hours selling the product, and use manpower to help them with operations. Some of these restaurants may churn if there are not enough customers and for our company, it doesn't make sense to invest in restaurants that don't get enough customers to cover initial investment costs.

To have a better chance of success, we want to target the area with a dense amount of restaurants, to begin with. This has multiple benefits:

- Group of restaurants together shows high demand in the area. If we convince these restaurants to participate, we probably get a good amount of users on our platform.
- Sending salespeople to talk to the restaurant owners is costly. If restaurants are placed close together we can make better use of our sales team and send them to cover more restaurants in each sales pitch tour.
- Operations cost as well as sales cost is high. We have to send equipment and install it for our merchants. The closer these merchants are, the more we can provide our services each day.

Our goal is to find target neighborhoods in each city that are good candidates to start expanding our business. Neighborhoods with a dense number of restaurants are a good start for us.

Data

Foursquare data for restaurants

To find areas in a city that has a dense amount of restaurants, we are going to use Foursquare data to find the restaurants and use their latitude and longitude to find their locations.

These locations will be the input for our clustering algorithm to find dense clusters that would be good targets for our initial business investment in each city.

We'll use **venues** api from FourSquare and search for nearby places given a latitude and longitude. These data will be aggregated for each neighbourhood to create a good repository for the city.

URL that is going to be used for getting the data:

GET [https://api.foursquare.com/v2/venues/search?ll=\[LatLong\]](https://api.foursquare.com/v2/venues/search?ll=[LatLong])

Response is going to be:

```
{
  "meta": {
    "code": 200,
    "requestId": "5ac51d7e6a607143d811cecb"
  },
  "response": {
    "venues": [
      {
        "id": "5642aef9498e51025cf4a7a5",
        "name": "Mr. Purple",
        "location": {
          "address": "180 Orchard St",
          "crossStreet": "btwn Houston & Stanton St",
          "lat": 40.72173744277209,
          "lng": -73.98800687282996,
          "labeledLatLngs": [
            {
              "label": "display",
              "lat": 40.72173744277209,
              "lng": -73.98800687282996
            }
          ],
          "distance": 8,
          "postalCode": "10002",
          "cc": "US",
          "city": "New York",
          "state": "NY",
          "country": "United States",
          "formattedAddress": [
            "180 Orchard St (btwn Houston & Stanton St)",
            "New York, NY 10002",
            "United States"
          ]
        },
        "categories": [

```

```

    {
      "id": "4bf58dd8d48988d1d5941735",
      "name": "Hotel Bar",
      "pluralName": "Hotel Bars",
      "shortName": "Hotel Bar",
      "icon": {
        "prefix": "https://ss3.4sqi.net/img/categories_v2/travel/hotel_bar_",
        "suffix": ".png"
      },
      "primary": true
    },
    "venuePage": {
      "id": "150747252"
    }
  }
]
}

```

For the purpose of this project, I'm only interested in the latitude and longitude of the restaurants.

Toronto Neighbourhoods Coordinates

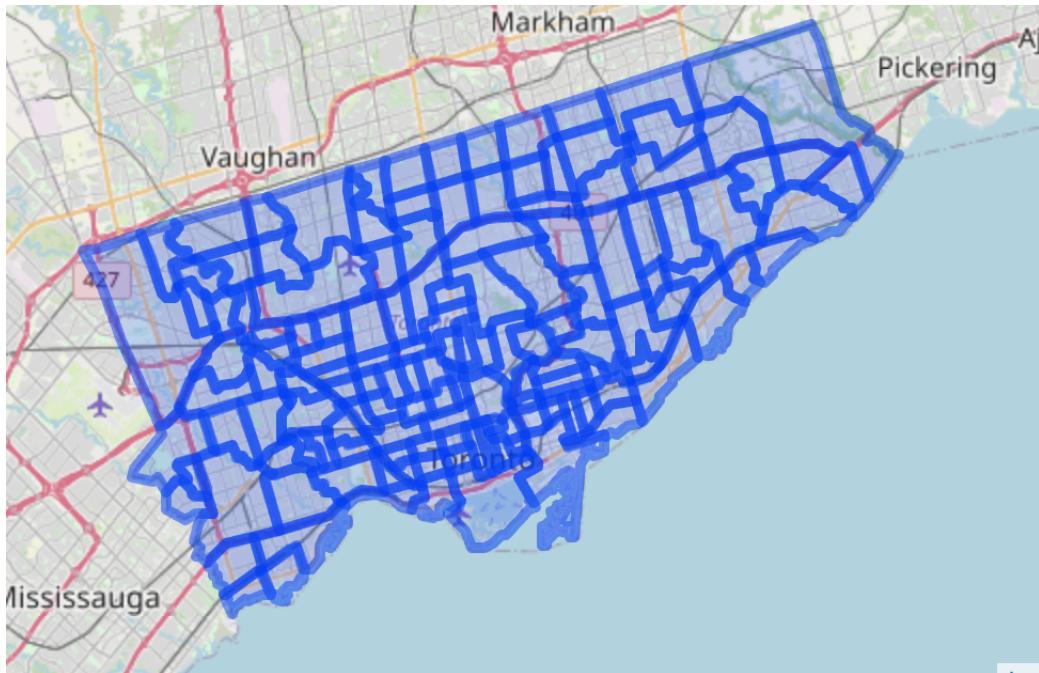
Before using Foursquare API, we need to split Toronto into neighbourhoods and look for restaurants in each neighbourhood. This is because Foursquare returns a maximum of 50 answers for the search query. With this split we can get max 50 restaurants per neighbourhood and at the end aggregate all together.

The data for Toronto neighbourhoods is available at <https://open.toronto.ca/dataset/neighbourhoods/>. The challenge is that the data is in GeoJson format which is in a polygon shape. We need to use some libraries to find centroid for these polygons and use them in our search API.

Methodology

Toronto Neighborhood Locations

Data gathering is the first step in this project. We are using geojson data from the open data portal of Toronto. The geojson format shows the shape of the neighbourhoods in terms of a polygon.



This shape is not going to be useful in our analysis. We can look at the central point of this polygon and use that as a point of reference.

Restaurants in each Neighbourhood

Next step is to get all of the restaurants around the central location of each neighbourhood. After adding all of the data together we will begin our analysis.

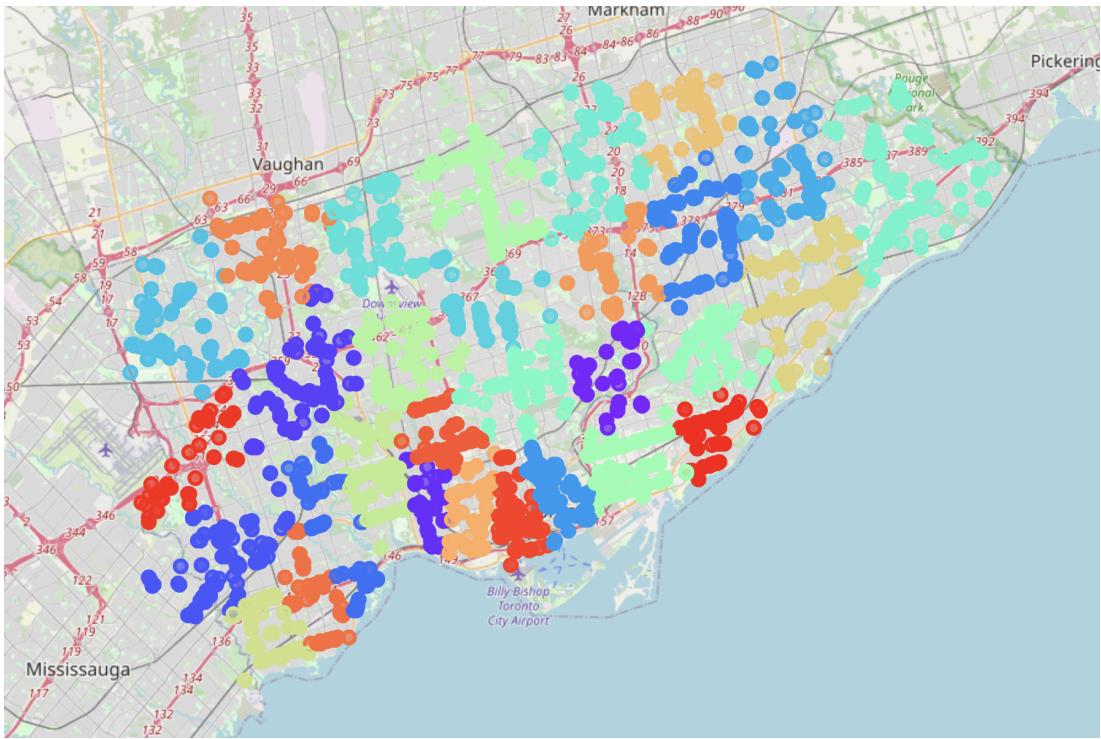
Clustering

For our business, we are going to allocate 30 salespeople to handle Toronto. We need to allocate one salesperson to sets of nearby restaurants. This makes it efficient for them to go to all of the restaurants with the minimum amount of commute.

We are using unsupervised machine learning algorithm K-Means to cluster the restaurants together. We are going to use the restaurant locations and try to create clusters so that adjacent restaurants fall into the same cluster.

Results

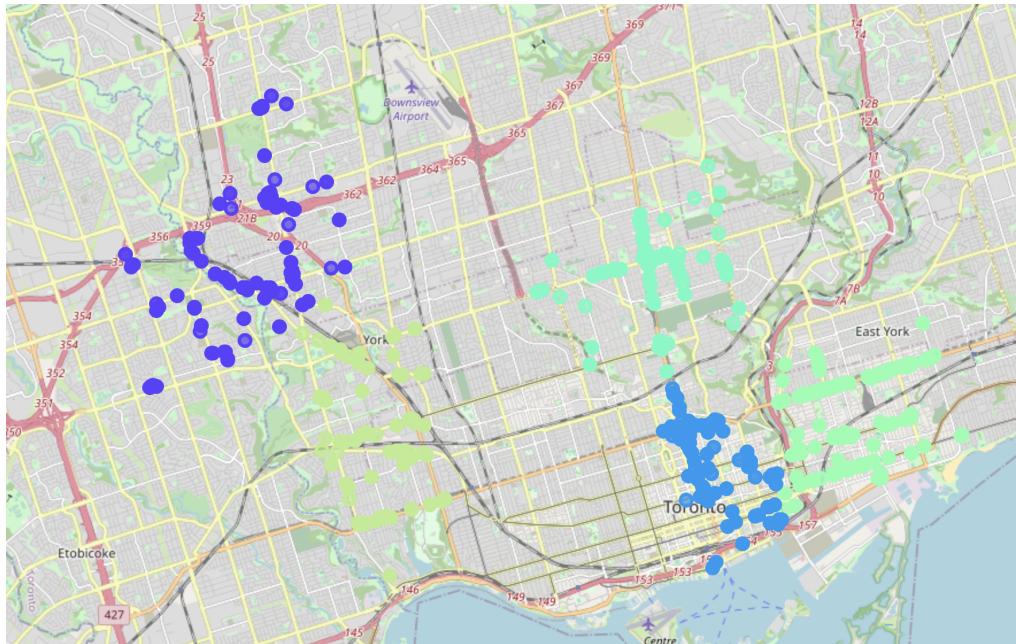
After running the algorithm, we splitted the restaurants into 30 clusters



Discussion

Based on the results, we can see that some areas have a dense amount of restaurants compared to others. Question is which of these clusters should we target first? Clusters with a denser number of restaurants provide more business opportunities. On the other hand, they require more experience from the salesperson because they have to take care of more restaurants in those locations.

Not all of the 30 salespeople are available at the moment of the application launch. We are going to hire them over time. Our first budget is for 5 salespeople. It is better if we allocate them to clusters with more restaurants. Based on the analysis the following clusters have more restaurants:



Conclusion

With our current data, our team is equipped with the knowledge of which restaurants they should start looking for to start the business. Revenue from these 5 clusters can help recruiting more people to attend other clusters.