Italian & Asian Gourmet (IAG) Restaurant Chain in Toronto[1]

Andrija Mihoci

Executive Summary

As a part of our business expansion we investigate the possibility to open a chain of IAG restaurants in Toronto. Our data science team is responsible to accomplish the following goals: (i) finding and analyzing clusters of Italian and Asian restaurants in Toronto, (ii) reporting on identified central interest points for our business expansion and restaurant type and (iii) providing a list of streets at focal points that will be used in the negotiations with local entrepreneurs. This report summarizes the main findings of the data science group.

Keywords: business expansion, data science, Toronto business data, cluster analysis

Highlights: Company aims expanding to a new city | Data science team supports management with necessary expertise | Data on restaurants in Toronto paramount for the campaign success | Priority: extracting sufficient and relevant information

1 Introduction

An Italian & Asian Gourmet (IAG) inter-culinary expansion strategy combines advantages of both cuisines and businesses at the same time, ultimately increasing the value for all our stakeholders. For example, customers that are used to one kind of cuisine can enjoy more culinary variety. Our company can furthermore substantially reduce operating costs while buying established local restaurants, essentially 'outsourcing' all resources: we call it 'free operation flow'. Our restaurant managers and personnel may be recruited from the local restaurants, keeping the employment and reputation intact.

A popular strategy is to place a chain of three IAG restaurants with carefully selected restaurant type in a given city. The first two are placed in the mid-point of the identified clusters of established Italian as well as Asian restaurants. The third IAG restaurant is placed in the midpoint of the two locations and serves in first line as an important connecting node. This restaurant facilitates operations of the other two restaurants (personnel and logistics) and may even provide expansion and branding prospects.

The IAG inter-culinary expansion strategy therefore demands precise identification of local interest points and consumption habits at a given area. At these spots, we can easily start our negotiations with the local community, especially when buying local restaurants and while arranging them in the IAG style. The largest benefits are achieved in terms of cost savings. The existing restaurants are slightly re-arranged to accommodate both culinary styles while keeping the loyal customers.

Our strategy is initially very expensive and calls for an outstanding persuasion technique. A good proportion of the needed capital can be buffered over time, for example by hiring existing

[1]Disclaimer: This report simulates a business idea and serves as a term paper submitted at the IBM Coursera Data Science module.

personnel at the acquired local restaurants. While furthermore negotiating with owners and personnel of several restaurants in the vicinity of the focal points, we may actually make great business deals.

The main objectives of the data science team reported here are to provide valuable insights related to: (i) Italian and Asian restaurant cluster identification, (ii) central focal point and restaurant type detection and (iii) creating a list of potential streets for our visit and the IAG-arrangement. Based on the empirical study we provide evidence about the identified clusters of established restaurants and provide precise locations of focal points in Toronto. Suggestions for serving particular restaurant types are identified. A list of relevant streets close to the three spots is provided, which facilitates our IAG inter-culinary expansion strategy.

The structure of this paper is as follows: after the data and methodology are described in Sections 2 and 3, the key empirical results are presented in Section 4. The discussion and conclusion Sections 5 and 6 finally summarize the main data science findings related to our IAG business strategy.

2 Data

The main purpose of the location targeting is to find precise geographical data about existing restaurants in the city of Toronto. Here latitude and longitude information as well as the the name of the restaurants play an important role. After the external data is collected, the activities in the data preparation stage ensure reliable input to machine learning algorithms. Properly storing output data finally helps in the interpretation of results and the pursuit of our business activities.

The data of our project are correspondingly divided into three groups: (a) location data of neighborhoods in Toronto that are accessible on Wikipedia.org and cognitiveclass.io, (b) data about local businesses in Toronto, source: Foursquare.com and (c) datasets created by our data science team that are suitable to address the requirements of the IAG inter-culinary expansion strategy.

There are two necessary requirements about all datasets that should be met within this project. All data should firstly pass the set of our high-quality standards. Here the focus is on reproducibility of all data entries in a very short time frame with a very precise description of all steps. The second requirement is that the algorithms and apps should be fast enough in retrieving and processing new/current data. This point is especially important few hours before the start of our IAG-expansion in Toronto.

At present there are in total 103 postal codes in Toronto with associated neighborhoods. For convenience we will refer to them as neighborhoods in the remainder of this paper. To give you a first impression about the location dataset, consider an excerpt of five Toronto neighborhoods, see Table 1. All locations of the Toronto location dataset are for convenience displayed in Figure 1. The distance between the two most distant locations equals approximately 40km (NE to SW) and 20km (NW to SE), see the interactive map at gmap-pedometer.com for more information.

Since we are interested in creating an AIG restaurant chain, we focus exclusively on currently registered restaurants. In 98 neighborhoods there are 55 restaurant types present. For each

Postal Code	Latitude	Longitude
M3A	43.75	-79.33
M4A	43.73	-79.32
M5A	43.65	-79.36
M6A	43.72	-79.46
M7A	43.66	-79.39

Table 1: Selected neighborhoods in Toronto.

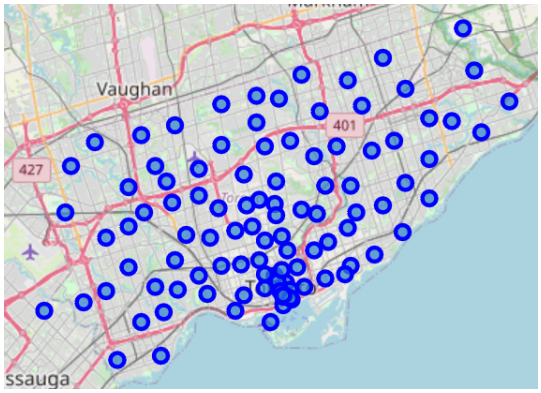

Figure 1: Neighborhoods in Toronto.

restaurant type we measure its attractiveness by its relative proportion in the total restaurant population of a given neighborhood. By reporting the average attractiveness level (average proportion) across all 98 neighborhoods we finally identify the top 10 restaurant types and show the results in Table 2.

Restaurant Type	Attractiveness
Italian Restaurant	0.11
Japanese Restaurant	0.06
Sushi Restaurant	0.06
Indian Restaurant	0.06
Middle Eastern Restaurant	0.05
French Restaurant	0.05
American Restaurant	0.04
Thai Restaurant	0.04
Mediterranean Restaurant	0.04
Chinese Restaurant	0.04

Table 2: Average proportion across 98 neighborhoods.

The leading type is Italian, which is followed by several Asian restaurant types. After this results have been presented, the management followed the suggestion of the data science team and started to draft the ideal AIG restaurant type. A consensus has been made to focus on a core Italian restaurant with elements of Japanese and/or Sushi cousine. This selection will make our three restaurants in Toronto relatively flexible to accommodate prevailing customer preferences.

3 Methodology

Given the aims of our company and the high-dimensional nature of the datasets, our data science team proposes to select cluster analysis as a suitable dimension reduction technique in the modeling phase. A k nearest neighbors algorithm appears to accommodate many desirable benefits for our attractiveness and location investigation and will be the main tool in the empirical analysis.

In modeling the data, our scientists will start with the full data matrix that contains all restaurant type attractiveness variables for all neighborhoods. In addition to the attractiveness variables we include the geographical information through the latitude and longitude information. In the graphical illustrations, the neighborhoods names will be used and thus needs to be made available by the algorithm.

The algorithm essentially searches for best k (high-dimensional) data points in the data cloud. At every iteration step, the distances of each point to the closest imposed data point are evaluated. This means that at every step the central point moves until it remains relatively stable. The initialization and the algorithm steps are often repeated in order to ensure robust results relative to initial value selection.

Selecting a suitable k will be done by augmenting the resulting clusters. This is common practice due to the unsupervised nature of the clustering procedure. Giving names to the clusters will be proposed by our research team as this requires practical experience. Due to the subjective aspect of these two decisions, all results will be jointly discussed between all participating stakeholder groups.

After this stage we can safely use the selected cluster center data (latitude and longitute of the points) as our three location principal focal points. Our team will then identify the closest neighboring Italian, Japanese and Sushi restaurants at each focus point. This list will be utilized in the negotiation process for the acquisition and the arrangement of local Toronto restaurants in the IAG-style.

4 Results

After conducting the cluster analysis with several starting values for k, the data science team decided to work with $k = 2$ clusters. While employing analysis with more clusters, the identified 'Italian' cluster remains unchanged which supports this suggestion. We depict both clusters in Figure 2. Observe that the 'Italian' cluster covers slightly more than a quarter of all Toronto neighborhoods (26 out of 98) with a length of 15km and a width of 10km.

The average attractiveness of the first cluster is by far highest for the Italian restaurant type, see Table 3. The second cluster is identified as an area that has slightly more Japanese and Sushi restaurants as compared to the Italian ones.

With reference to the centroid of the 'Italian' cluster we our app reports its geographical center location: 'The 1st focal point "Oakwood Avenue" has coordinates $\{43.69, -79.44\}$'. Looking at the second cluster, we report on the centroid as follows: 'The 2nd focal point "Bayview Avenue" has coordinates $\{43.71, -79.38\}$'. And the third focal point, the midpoint between the two centroids, displays: 'The 3rd focal point "Avenue Road" has coordinates $\{43.70, -79.41\}$'.

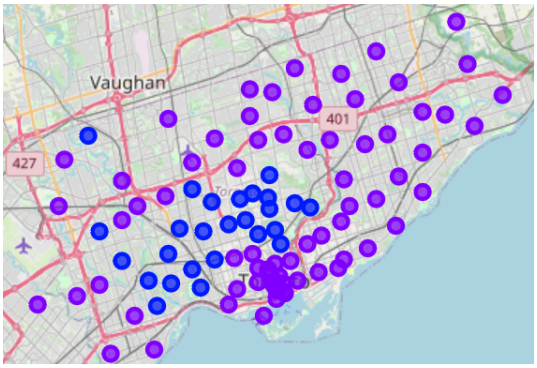

Figure 2: Restaurant clusters in Toronto.

Restaurant Type	Cluster 1	Cluster 2
Italian Restaurant	0.30	0.05
Japanese Restaurant	0.05	0.07
Sushi Restaurant	0.06	0.06

Table 3: Average attractiveness of restaurant types.

An illustration of all three target places is provided in Figure 3. The identified streets are Oakwood Avenue, Bayview Avenue and the Avenue Road. The latest street is located between the first two potential restaurant locations. All identified streets are accessible from the Eglinton avenue. Due to their central geographical location, all IAG restaurants may be easily connected from other neighborhoods.

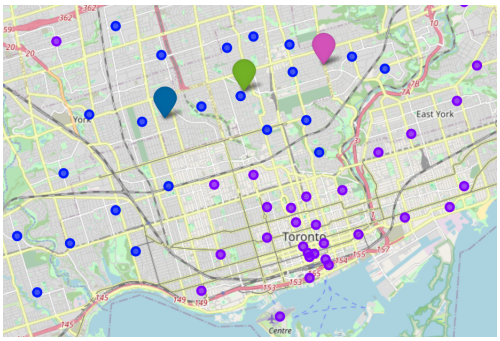

Figure 3: Identified focal points for IAG expansion in Toronto.

5 Discussion

The possibility of opening an IAG restaurant chain in Toronto looks quite promising. The following are our main empirical findings:

(i) Restaurant types - in Toronto there are many restaurant types present. This supports the idea of the IAG management of offering more variety to the customers. Looking at the attractiveness of special types, the results indicate a strong presence of Italian, Japanese and Sushi restaurants. A flexible restaurant setup, with core Italian and potentially changing second cuisine (Japanese and/or Sushi) makes the AIG a desirable choice.

(i) Clusters - there is a very strong and compact Italian restaurant cluster present. To offer a multi-culinary experience to people in the associated neighborhoods may add more value while keeping the most loyal customers. The other part of the city will most likely prosper from the interchanging of the Japanese and the Sushi cuisine.

(ii) Locations - by geographically connecting all three focal points (Oakwood Avenue, Bayview Avenue and Avenue Road) via the Eglinton avenue, we can optimize our distribution costs and faster react to the customer demands in real time. The time needed to travel from Avenue Road to other places is approximately 15 minutes: a special request of a customer in one place can be served by the personnel from the other place in short time and vice versa.

6 Conclusion

This paper provides the main findings for an Italian & Asian Gourmet (IAG) restaurant expansion in Toronto. The added value of this project can be summarized as follows: (i) two clusters of Italian and Asian (Japanese and Sushi) restaurants have been identified in Toronto, (ii) there are three identified central interest points for the business expansion and (iii) a list of inter-connected streets at focal points is provided. Businesses at these streets can be visited in the negotiations with local entrepreneurs. Given the reproducibility of results, future operations can easily be supported by statistical and machine learning techniques.

Due to the high-dimensional nature of the geographical and economic data, cluster analysis provides here valuable insights into the spatial data structure. Two clusters of restaurants have been particularly identified, and their central points are associated to inter-connected streets. Presented time and distance measurement data additionally support a potential smooth business expansion in Toronto.