**Exploring Indian food venues in the city of Toronto**

**1. Introduction**

**1.1 Background**

Whenever a person enters a new city, they would be interested in some of the best places the city has to offer. A person may want to know how good a given restaurant is good in their vicinity. This extra information can help them decide which food venues are the best in the new city. A restaurant with the highest rating can surely help to draw more customers as diners can make informed decisions about places to dine or visit.

The city of Toronto is densely populated by 6,417,516 spread across a total area of 630.2 sq Km. There are quite a lot of venues located in this densely populated city which lies in the Ontario province in Canada. This project explores food venues in Toronto based on user ratings. Furthermore, a map of the venues with areas of interest indicated by color give more findings and information about these venues in general.

**1.2 Interested audience**

The target audience for this project could be varied. Firstly, any person who is visiting the city of Toronto, Canada can use these maps to quickly select the places which they may like to visit based on their budget and other user ratings. Secondly, a company can use this information to create a regularly updated website that allows individuals to gain a better liking or familiarity to Indian venues. This functionality can even be expanded to other places other than restaurants.

**2. Data**

**2.1 Data Sources**

To get location and other information about various venues in Toronto, the data about the geographical coordinates was obtained using wikidata’s list of postal codes in Canada. The data with respect to recommendations was fetched using the Foursquare’s explore API, which gives venues recommendations of the top 100 venues in the region within a radius of 500m.

The

**2.2 Data Cleaning**

The given data was cleaned to remove all the empty rows containing missing information with respect to a neighborhood, borough or postal code.

**3. Methodology and Exploratory Data Analysis**

Firstly, we use the Foursquare API to fetch the venue information based on ratings for a given locality for all venues up to a distance of 0.5 kilometers from the center of Toronto. After cleaning the data, the dataset is combined with the latitude, longitude values and the venue names. We then remove any outliers after careful inspection for any neighborhood. The final data includes the columns Neighborhood, Indian Restaurant- which indicates proximity to the city center, Cluster Label, Neighborhood Latitude, Neighborhood Longitude, Venue, Venue Latitude, Venue Longitude and Venue Category.

Using this data, we are better able to analyze the places based on ratings and color coding so that at a simple glance we are able to differentiate between preferences a person may like to have to give more information about the venue. This can also be extended to provide personalized recommendations about locations based on clustering, which is an unsupervised machine learning approach to gather useful inferences.

**3.1 Categories**

The analysis begins by looking at a set of all venues in the city of Toronto that range from entertainment to food venues. The entire city of Toronto has 10 boroughs and 103 neighborhoods. The maximum number of neighborhoods are 19 and are found in the North York borough. The least number of neighborhoods are 5 and these are found in the boroughs East Toronto, York and East York respectively.

**3.2 Clustering**

Finally, all the venues are clustered using K-means clustering based on their location and category of Indian restaurants to identify similar venues and the relationship amongst them.

**4. Results and Discussion**

After collecting data from the Foursquare API and postal codes, we got a list of 230 different

Venues after removing the outliers. The latitude and longitude values as well as their names to combine them and remove all the outliers.

We identified that from the total set of venues, majority of them were recreational parks. The results were further filtered to match venues by Indian restaurants.

A company can use this information to build an online website/mobile application, to provide

users with up to date information about various venues in the city based on the search criteria

(name, rating and price).

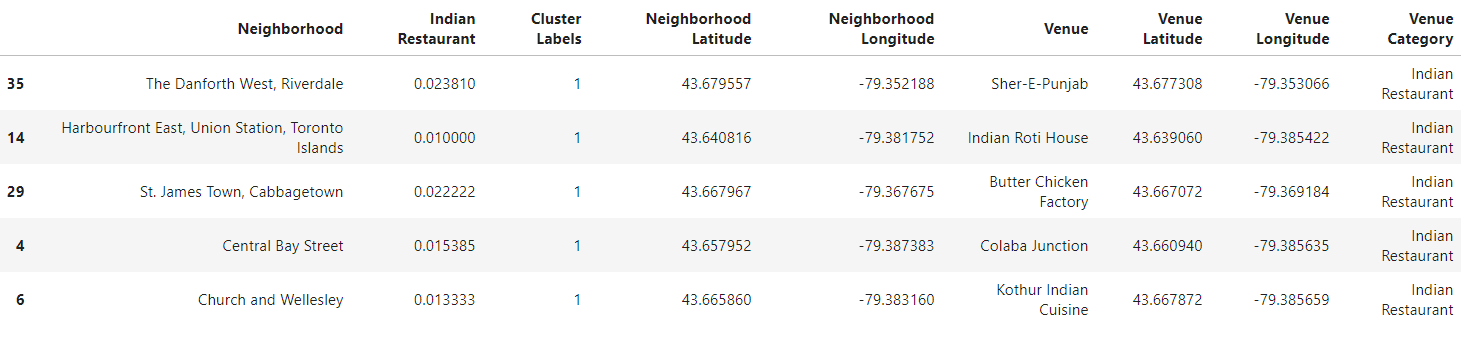


Fig1: Cluster 1 with the most number of Indian restaurants among the three clusters

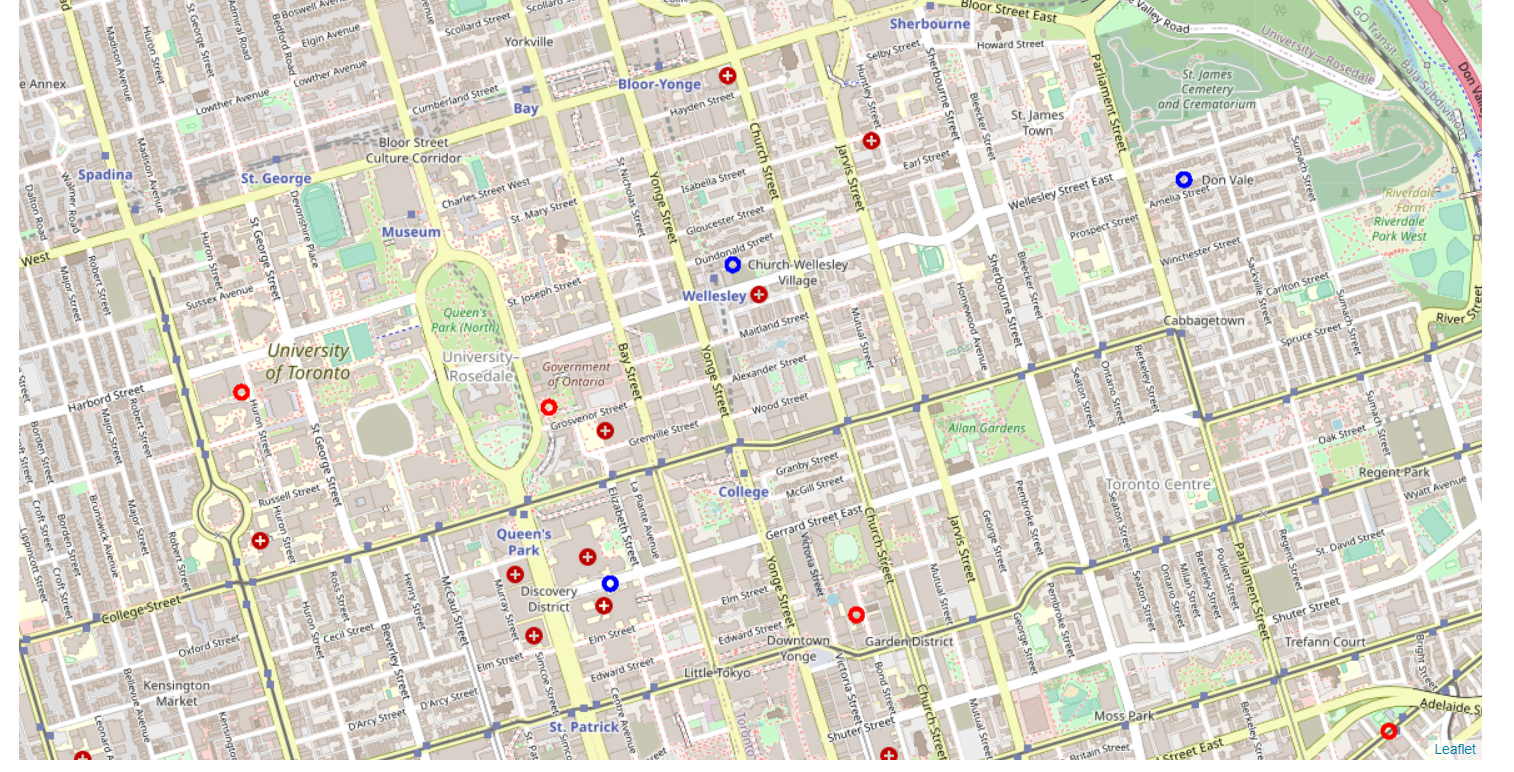


Fig2: A closer view of Indian restaurants among the three clusters in the city of Toronto

**5. Conclusion**

The purpose of this project was to explore the places that a person visiting Toronto could

explore the food outlets. The venues have been identified using Foursquare API and have been

plotted on the map. A lot of the Indian Restaurants can be found in Cluster1 which is around Davisville, Central Bay Street, St.James Town, Danforth West and Church & Wellesley area. A lesser number of restaurants are found in Cluster 2 in the North Midtown of Yorkville.