**Pardes India Inc.**

**Optimal Location for opening an Indian Restaurant**

in

**Vicinity of Schaumburg, IL**

**Rajesh Dhar**

19 March 2019

**Table of Contents**

## **[Introduction](#_Introduction)**

## [**Data**](#_Data)

## [Methodology](#_Methodology_1)

## [Data Analysis](#_Data_Analysis)

## [Data Collection - Areas to explore](#_Areas_to_Explore)

## [Analyze the data](#_Analyze_the_Data)

## [Identify the candidate locations](#_Identify_the_candidate)

## [Results and Discussion](#_Results_and_Discussion)

## [Conclusion](#_Conclusion)

## **Introduction**

"Pardes India Inc." is a large Indian restaurant chain operating out of Mid-West USA. They already have few restaurants in Chicago metropolitan area and are looking to open Indian restaurants in many Chicago Suburbs. One of the key area they want to explore is Northwest suburb of Chicago and they have identified few suburbs in vicinity of Schaumburg IL. This area has sizeable Indian population and they believe opening an Indian restaurant there is a great business opportunity.

Schaumburg IL, is a suburb in the Northwest of Chicago. It is a multi-cultural suburb, famous for Woodfield Mall and has offices of many multi-national corporations. "Pardes India Inc." wants to identify a right location which is not crowded with Indian restaurants and has reasonable Indian population and the household income of the area is on higher side. Hence we will find locations that are good candidates for opening a new Indian restaurant.

## **Data**

For this project, we used different sources of data. Some of this data was obtained by scraping various websites and some of it was fetched by using Foursquare API.

First, we identified the suburbs in the vicinity of Schaumburg. Next we got all the unique zip codes in these suburbs. Then, we got some demographic details about these zip codes and finally we got the restaurant details in the identified zip codes. While collecting this information, we also got location coordinates of suburbs, zip codes as well as of the restaurants.

We used following sources to get the needed data.

* Scrape Zipmat.net to get the zip codes for areas in vicinity of Schaumburg, IL.
* Scrape zipdatamaps.com to get the Asian Population percentage, Total Population and Median Household income of each Zip Code.
* Geopy "geocode.Geolake" to get the location coordinates for suburb and each zip code.
* Geopy “geocode.Nominatim” to get the addresses from the location coordinates.
* Use Foursquare API's to get the list of all restaurants and identify whether a restaurant is India or not.

## **Methodology**

In this project, we identified the zip codes in vicinity of Schaumburg that have low density of Indian restaurants. We also got the demographic details of each zip code to ensure we have enough Indian/Asian population in the zip code. Given that Indian cuisine is typically little costlier than other cuisines, we will also check the median Household Income in each zip code. and and the community.

In the first step, we will get the location details of all the suburbs we wanted to explore. And, with each suburb we got the details of all zip codes in the suburb. In case, zip code was part of more than one suburb, we assigned it to the suburb which was nearest to it. We also scraped the population of the zip code, percentage of Asian population as well as the Median Household income.

In the second step, we did analysis to find the density of all restaurants as well as density of Indian restaurants in each zip code. We also used maps (including Heat maps) to identify few promising areas.

In the third step, we identified zip codes that have less density of Indian restaurants (less than 5%), where household Income is greater than 70K and Asian Population is at least 5% of the total population.

Finally, we clustered these locations to identified the center of addresses that can be used to do further analysis to pinpoint exact location for Indian Restaurant in discussion with the stakeholders.

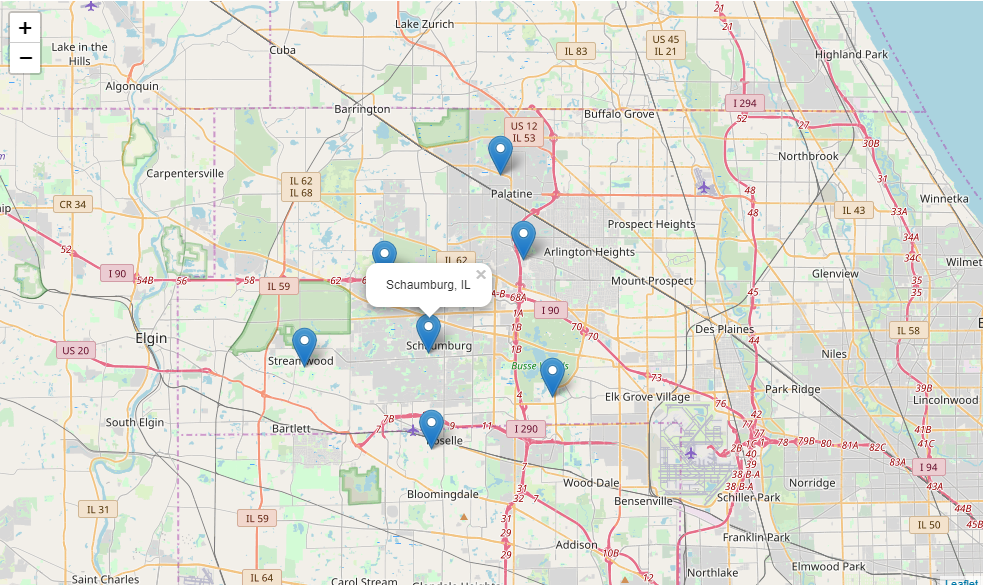
## **Data Analysis**

### **Data Collection - Areas to Explore**

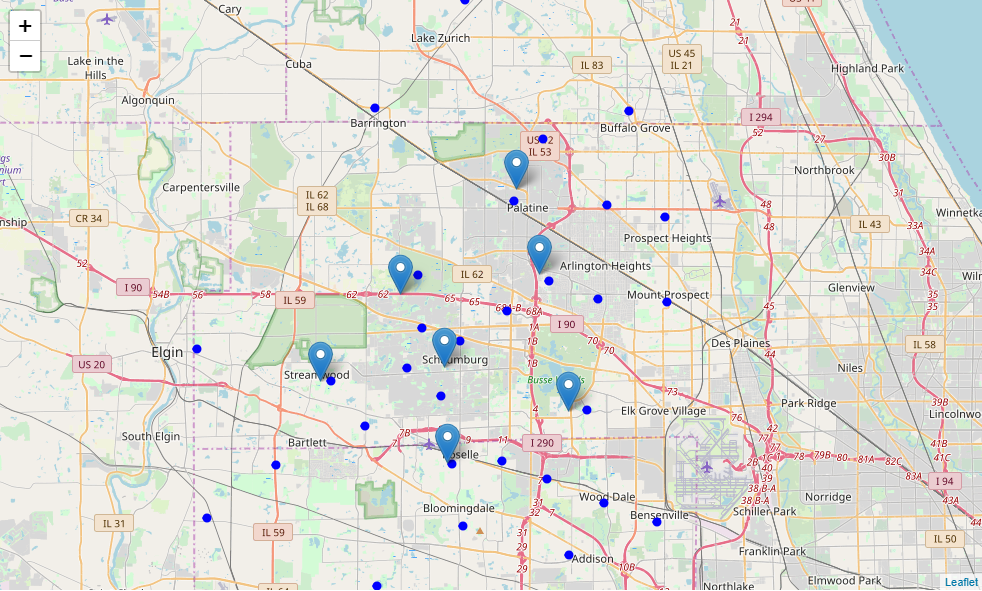
First we found which suburbs are in the vicinity of Schaumburg, IL. Quick Wikipedia search showed following 7 suburbs in the vicinity.

**Schaumburg, Hoffman Estates, Streamwood, Roselle, Elk Grove Village, Rolling Meadows and Palatine**

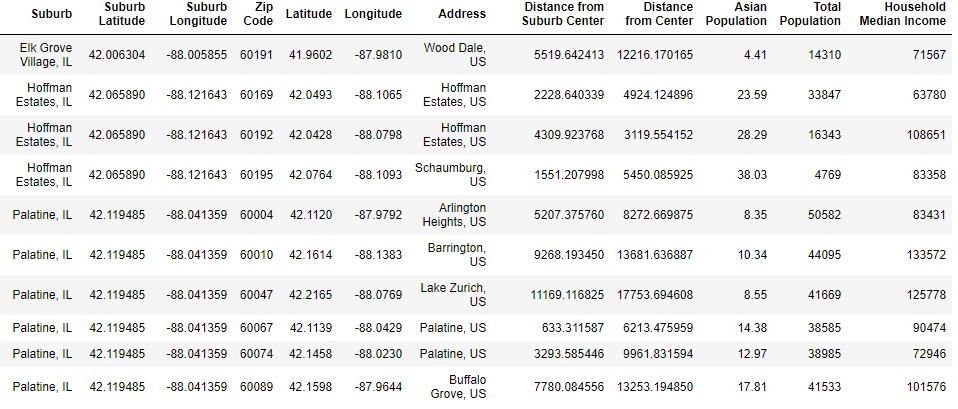
For these suburbs, we got the latitude and longitude details and plotted them on a map.



Next, we got the unique zip codes in these suburbs. We got 30 unique zip codes and average zip codes per suburb was 4.29. We also plotted these zip codes on map to see how they align with suburbs.

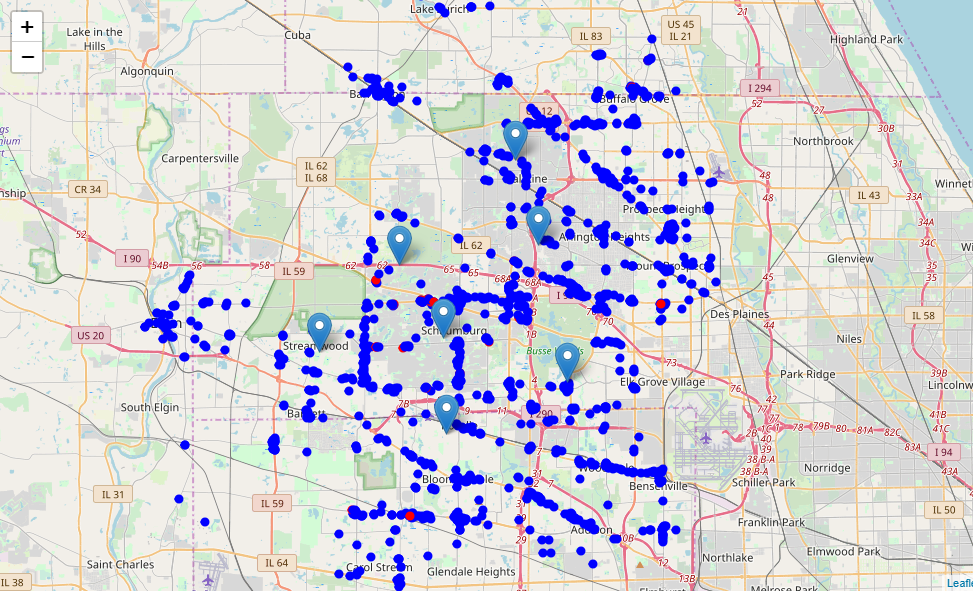


We also got some demographic information (Total Population, Asian Population and Median Household income for these zip codes.

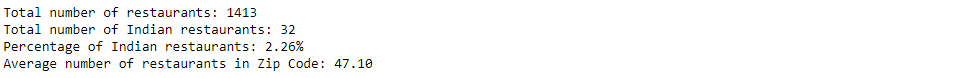


Finally, we got list of restaurants in those zip codes within 3KM radius of and also identified the Indian restaurants among them.

We plotted on map “all restaurants” and “Indian restaurants” to see where they are located. Indian restaurants were marked in “red” while all others were marked in “blue”.



We found there are large number of restaurants in these zip codes and Indian restaurants constitute only 2.26%.

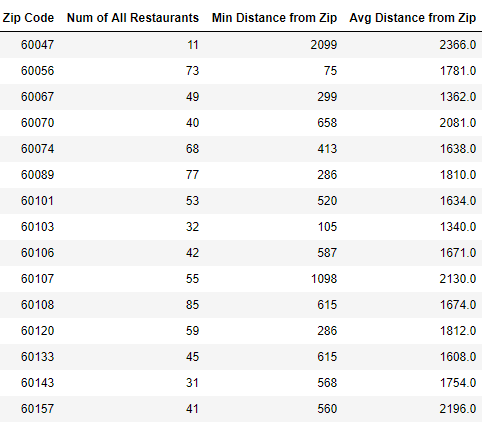


### **Analyze the Data**

Following data analysis was performed to

* Find how many restaurants are in each zip code and their minimum and average distance from center of zip code.
* Find how many Indian restaurants are in each zip code and their minimum and average distance from center of zip code.
* Calculate the percentage of Indian restaurants in each Zip Code
* Plot the heat map of all restaurants and Indian restaurants only.

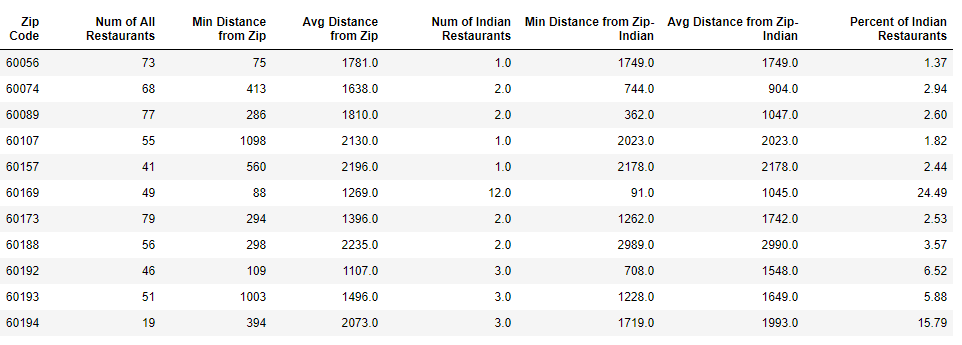
We found that most of the zip codes had more than 50 restaurants and the average distance of all restaurants was 571m from the center of zip code.



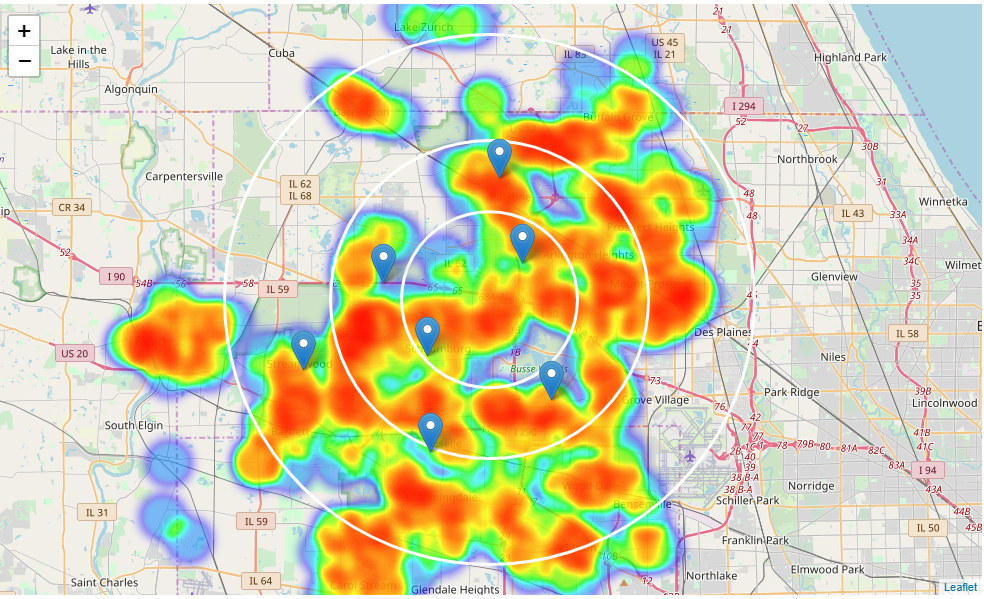
We also found that Indian restaurants were located only in 11 of 30 zip codes and zip code 60169 has 12 Indian restaurants (40%) of all Indian restaurants in all zip codes. We also found average distance of Indian restaurants was 1368m from the center of zip codes.



We calculated the percentage of Indian restaurants in zip codes where they exist. Zip code 60169 has 24% Indian restaurants. Overall there were 4 zip codes where percentage of Indian restaurants is more than 5%.

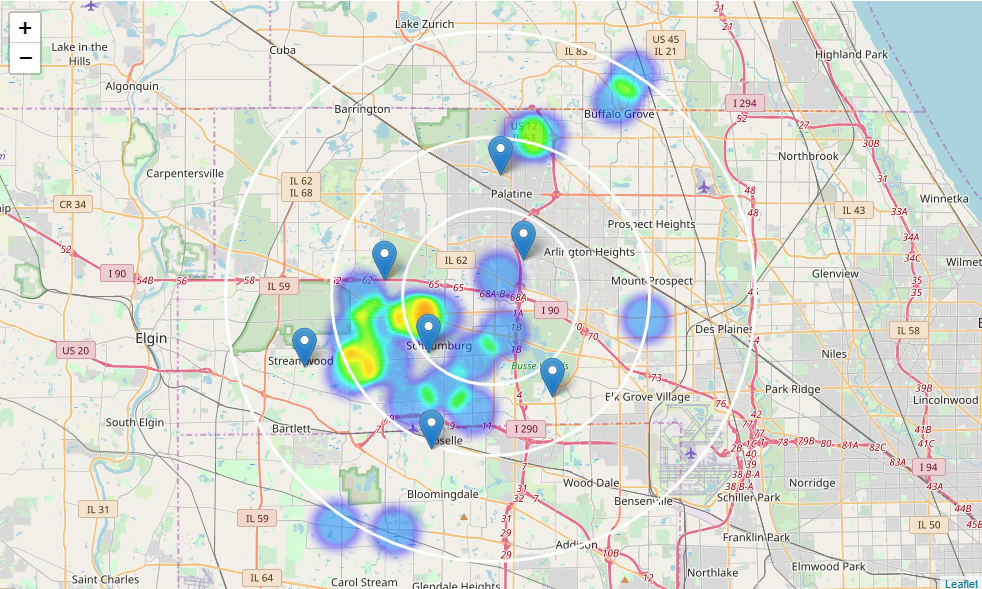


We plotted heat maps of all restaurants to see the density of restaurants.



We noticed there aren’t many restaurants in north east of Schaumburg.

We also plotted heat map of Indian restaurants and noticed that most of the restaurants are located in the triangle of Schaumburg, Hoffman Estates and Streamwood.

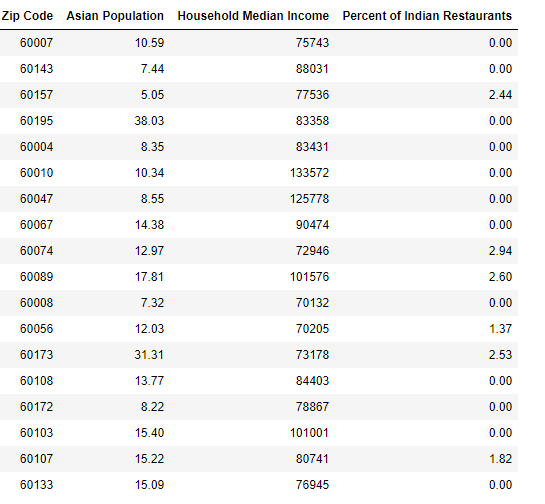


### **Identify the candidate locations**

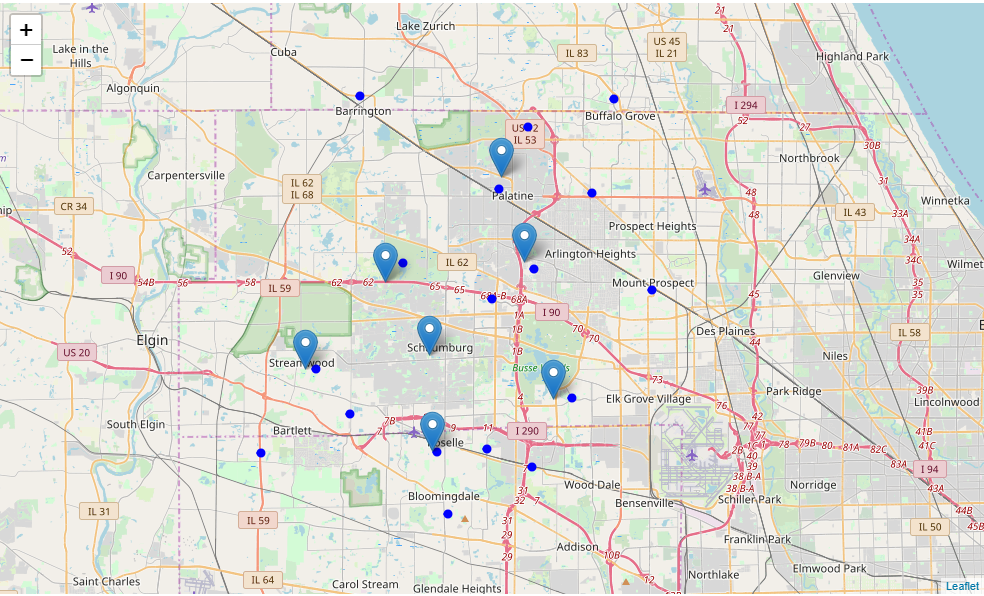
Next we narrowed zip codes on following criteria.

* Percentage of Indian restaurants is less than 3%
* Asian population is more than 5% of total population
* Household Median Income is above 70K

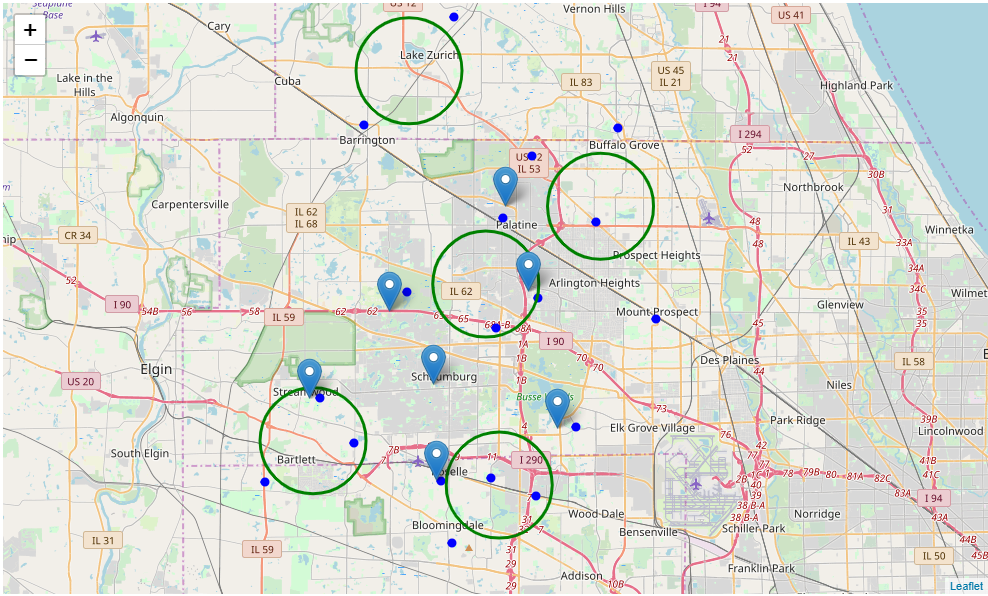
We got 18 zip codes that match the above criteria.



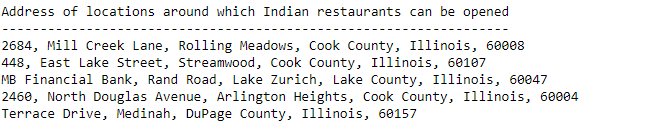
We plotted these zip codes on the map to get sense of how far they from each suburb.



We applied clustering on the location coordinates of these 18 zip codes to get 5 zones that can be candidates for opening a new location and plotted them on a map.



Finally, we got address of center of these 5 zones. This will be the starting point for stakeholders to make final decision after considering other criteria’s as well.



## **Results and Discussion**

Our Analysis shows there are 30 unique zip codes in suburbs in the vicinity of Schaumburg. These zip codes have close to 1400 restaurants. Of these there are only 32 Indian restaurants, which constitute 2.26% of all restaurants. Average number of restaurants in each zip code is around 47. Average distance of restaurants from center of zip code is 571m.

Of the 30 zip codes only 11 zip codes have Indian restaurants and average distance of these restaurants from center of zip code is 1368m.

Highest number of restaurants were detected in all directions from Schaumburg center except in North West. Indian restaurants are mostly concentrated in a triangle of Schaumburg, Hoffman Estates and Streamwood. We also looked the demographics of the zip codes to get the total population, percentage of Asian Population and Median Household Income.

We then narrowed down the number of zip codes to where percentage of Indian restaurants is less than 3%, Asian population is at least 5% and Median Household income is greater than 70K. This identified 18 target zip codes. We than clustered these zip codes to identify centers of points of interest. We also get the address of these centers by reverse geocoding.

Result of all this analysis is 5 zones which can be used as a starting point for finding the best location for opening an Indian restaurant. However, this is only an initial analysis and should be further analyzed by looking at other data e.g., crime rate, age grouping, education, commercial activity, real estate, fun areas etc. of the zip codes. Additional analysis might lead to different conclusions.

## 

## **Conclusion**

Purpose of this project was to identify areas in vicinity of Schaumburg with low number of Indian restaurants to help "Pardes India Inc." in narrowing down the search for optimal location for a new Indian restaurant. By calculating restaurant density distribution from Foursquare data we have first identified zip codes that warrant further analysis. Clustering of these zip codes was performed to create major zones of interest and addresses of those zone centers were created to be used as starting points for final exploration.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors listed above.