Locations for New Indian Restaurants in San Jose

1.0 Introduction

1.1 Background

The market for Indian food is significant and growing in San Jose, California. Drawn to IT and other vocations pertaining to opportunities in technology, around ten percent of the people living in San Jose are Indian. Furthermore, Indian food is very tasty and can be appreciated by people of any ethnicity. Therefore demand for Indian food will continue to grow and there exists, in addition to the profit motive, an unparalleled opportunity to promote authentic Indian cuisine in the bay area.

1.2 Problem

Having acknowledged the opportunity that exists in the bay area, we can see that many possible locations exist where an entrepreneur can start an Indian restaurant. There exist many good locations in San Francisco, Oakland, and San Jose. For the purposes of this Data Science initiative however, we will restrict ourselves to potential venues in San Jose. This is a good restriction because the cost of opening a business is lower in San Jose as opposed to San Francisco. Furthermore, unlike Oakland, San Jose has a higher percentage of Indian people living there. Ultimately, by seeing the efficacy of out methods in San Jose, we can use similar techniques to decide upon the location of future restaurants in other Cities.

1.3 Interest

We will optimize our locations according to two criteria. First, we want to make sure there are few restaurants nearby so our new business will have geographic prominence. Once we do this, we will want to pay special attention to areas with no Indian restaurants within the proximity of our selected location. Finally, once these two conditions are met, we want to find locations as close as possible to downtown San Jose, an area with a higher concentration of tourists. More people are also likely to spontaneously discover our restaurant in downtown San Jose when exploring or just out for a night of fun.

2.0 Data

2.1 Data sources

We construct a lattice centered around Downtown San Jose. The points on the lattice will represent the potential neighbourhoods. We will obtain the data from the following sources to obtain the relevant information.

- Automatically create the center of potential venues and use the **Google Maps Geocoder API** to find the addresses of those areas
- Use the **Foursquare API** to find the number of restaurants and their type and location in every neighbourhoood
- Use **Google Maps Geocoder API** to find the approximate address of Downtown San Jose (e.g. around Market and 1st probably)

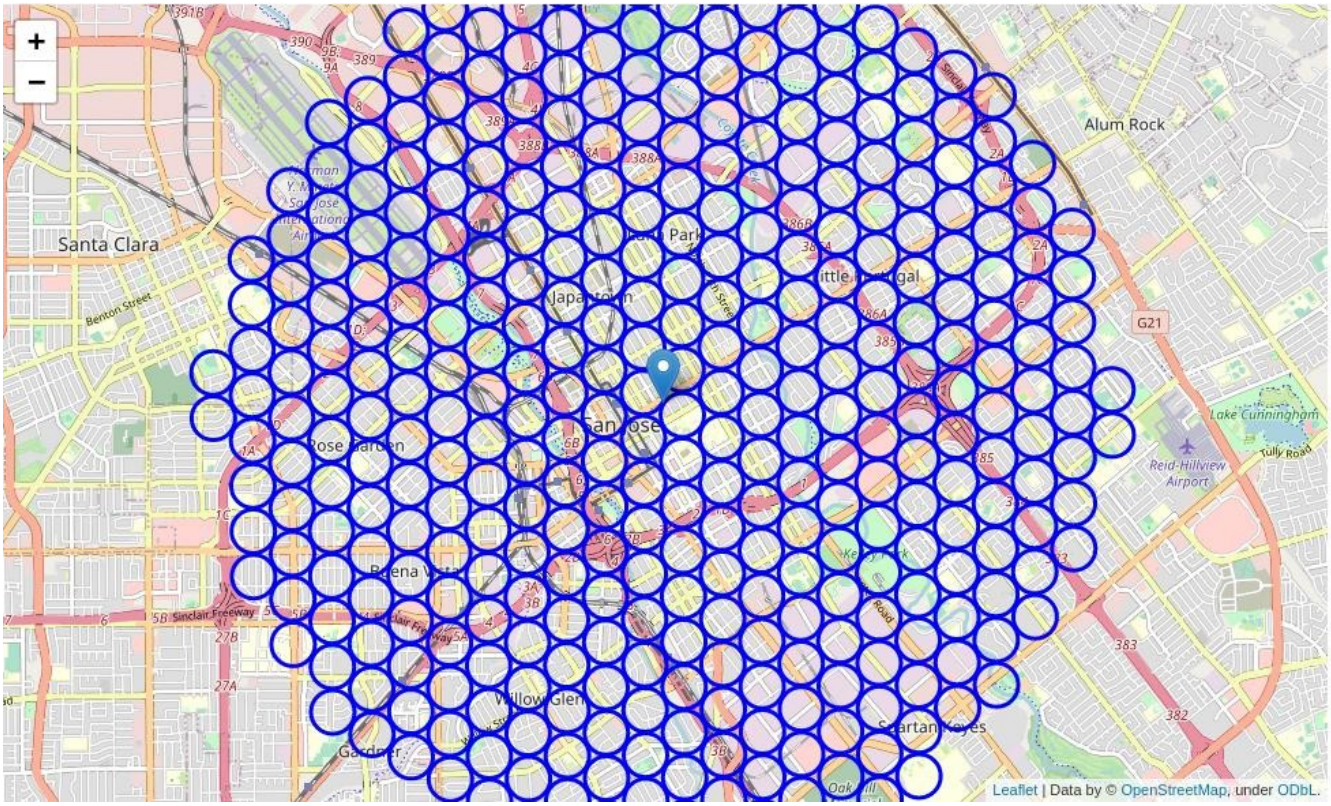
2.2 Feature selection

Considering the problem definition, lets identify the relevant features that will influence our estimate of a prospective venue's desirability :

- How many existing restaurants are in the neighbourhood (any type)
- Quantity of Indian restaurants in the neighbourhood
- Distance of Indian restaurants from our prospective venue
- Distance between venue address and downtown San Jose

3.0 Methodology

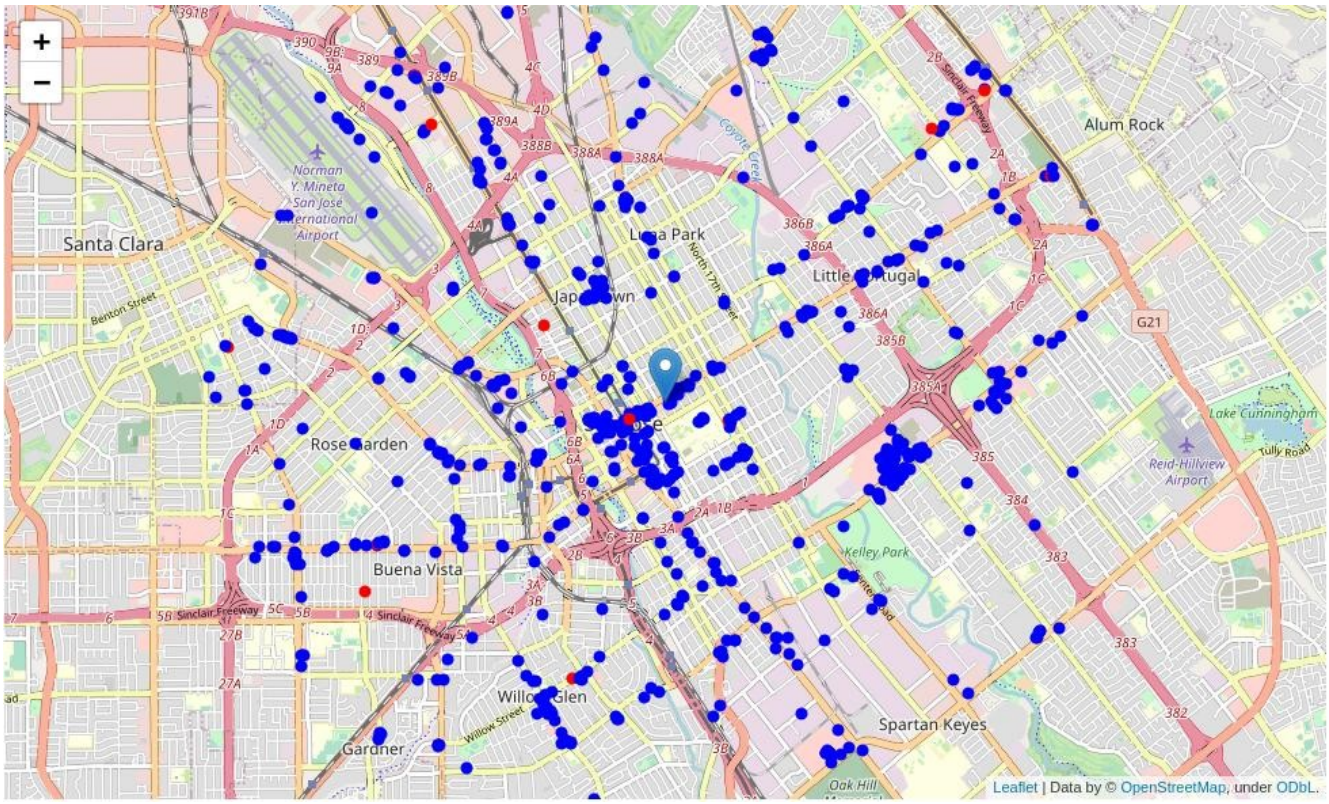
Here, we will focus on finding areas of San Jose that have low density of restaurants, particularly Indian restaurants. We will limit our analysis to 3.75 miles around City Hall.



Initially, we collect the required data : *location and type of every restaurant* in the proximity of city hall (a roughly 4 mile radius sweep centered at downtown San Jose). We also used the relevant Foursquare codes to find Indian restaurants.

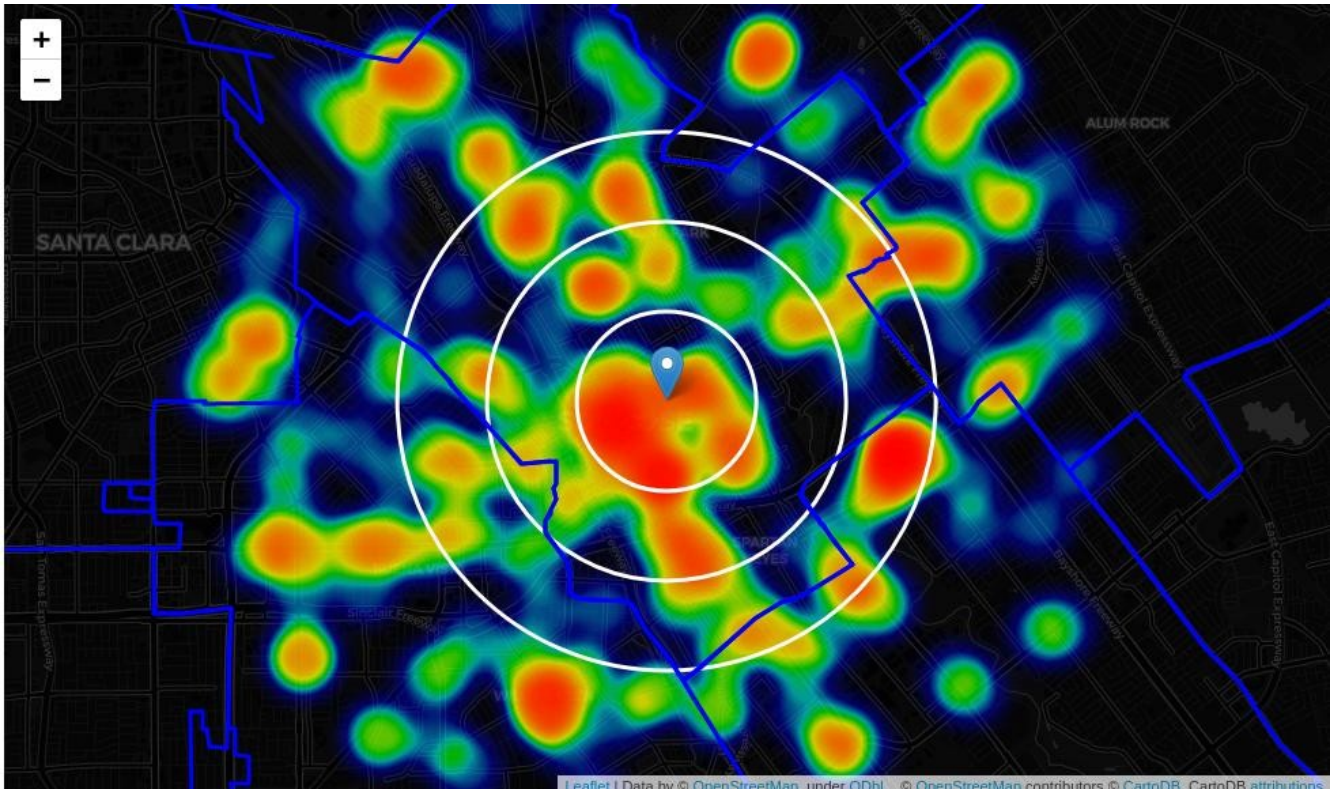
Now, we are ready to create visualisations of **density** pertaining restaurants. Using Folium and other tools, we will take advantage of the visual appeal of **heatmaps** and the intuitive power and simplicity of **k-means clustering** to identify ideal neighbourhoods that would serve as a good starting point for a potential entrepreneur/restauranteur looking to get into the business of catering Indian cuisine.

We will make sure that the areas we are looking at have no more than **two** restaurants within a radius of **250 meters** and **zero** Indian restaurants within a radius of **400 meters**.



4.0 Results

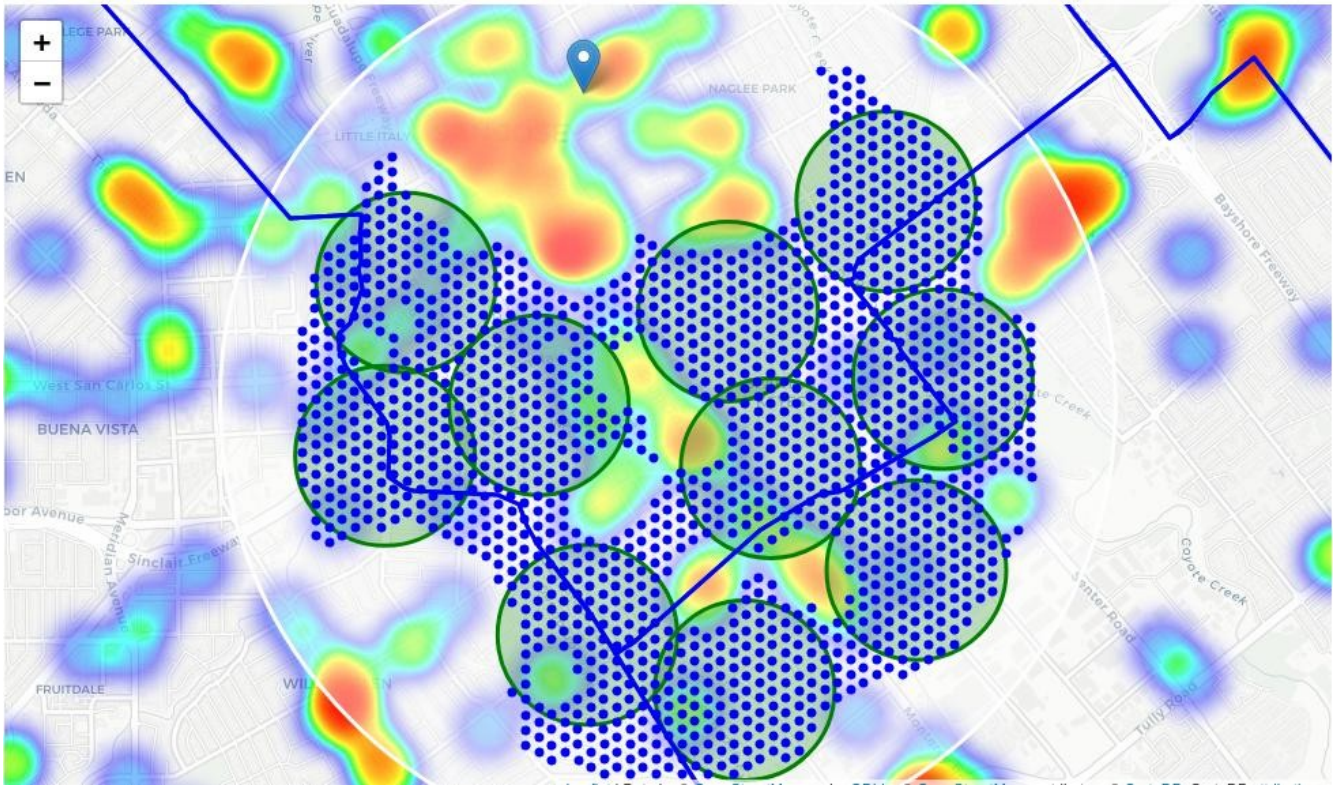
The analysis shows that despite the large number of restaurants in San Jose, there are pockets of low density somewhat close to city Hall. Most Indian restaurants were north of San Jose, hence we focused most of our attention on the borough of Spartan Keyes, to the South of city hall. Due to the community of artists, students, and tourists, along with recent development of buildings from commercial to residential, this serves as an ideal neighbourhood to start an Indian restaurants. Another interesting borough was Washington-Guadalupe, however we focused most of our attentions on Spartan Keyes.



Subsequently, we created a closely spaced grid of location candidates and filtered for those with few restaurants nearby and no Indian restaurants nearby. Then, after clustering these zones, we used reverse geocoding to find approximate addresses as starting points for more detailed local analysis based on other factors.

5.0 Discussion

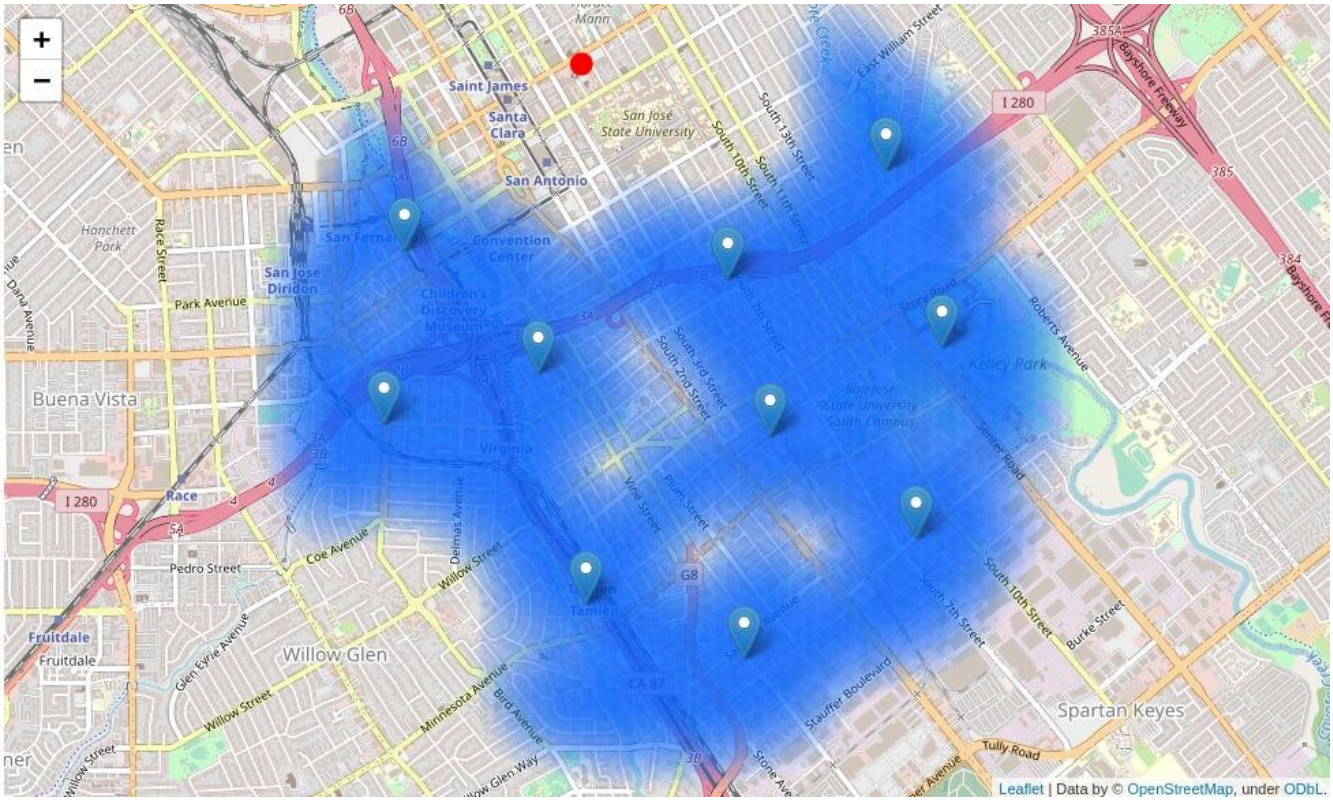
This results in 10 zones containing the greatest potential new restaurant locations based on the parameters we filtered for. Naturally, these are not all optimal locations. There could easily exist other reasons to invalidate these locations. Furthermore, additional locations that are not in this area could also be excellent candidates for restaurants (close to the highway but far from city hall). The techniques used in this project only serve to illustrate one possible way for identifying desirable locations for a new venue.



6.0 Conclusion

6.1 Preliminary findings

The purpose of this project was to identify San Jose areas close to City Hall with a low number of restaurants (particularly Indian restaurants) to aid entrepreneurs, investors, speculators, and restaurateurs in narrowing down the search for optimal locations for a new Indian restaurants. By calculating restaurant density distribution from Foursquare data, we have identified general neighbourhoods that justify further analysis (Washington-Guadalupe and Spartan Keyes), then generate extensive collection of locations satisfying basic requirements regarding nearby venues. Clustering these locations helps to identify critical zones of interest (most potential locations) and addresses of those centroids are reverse geocoded to serve as starting points for final explorations by interested parties.



6.2 Future considerations

Final decisions on the optimal location will be made by shareholders based on specific characteristics of neighbourhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to public transportation, for example), noise levels, prices, and social dynamics pertaining to each neighbourhood.