# The Battle of Neighborhoods (Week 2)

## Recommend a Neighborhood to Open an Indian Restaurant in San Jose, CA

## Table of contents



## Business Context and The Problem

San Jose is a large city surrounded by rolling hills and the largest city in Northern California. It is a multicultural and very diverse due to a large population of migrants from various parts of world. This has resulted in the establishment of various authentic ethnic restaurants in the city, ranging from Indian, Korean, Mexican, and Chinese, to Ethiopian.

In this project we will try to find an optimal neighborhood to open a restaurant in San Jose. Specifically, this report will be targeted to stakeholders interested in opening an **Indian restaurant** in **San Jose**, CA. Since there are lots of restaurants in San Jose, we will try to find **locations with less Indian cuisines**. We are also particularly interested in **neighborhoods known for restaurants but not with many Indian cuisines in vicinity**. We would also check the selected neighborhoods for any related Indian activities like Indian grocery shops, temples etc which probably help the Indian cuisine to do more business.

## Data

The following data sources will be used in this project to do the relevant analysis:

1. Neighborhoods in San Jose
   Wikipedia: https://en.wikipedia.org/wiki/Category:Neighborhoods_in_San_Jose,_California
2. Geocoder – Geocoder library and Nominatim API to determine the latitude and longitude for each of the neighborhoods
3. Location Data - Foursquare API will be leveraged to get the list of current restaurants and other most popular venues for each of those neighborhoods

### Neighborhood Data

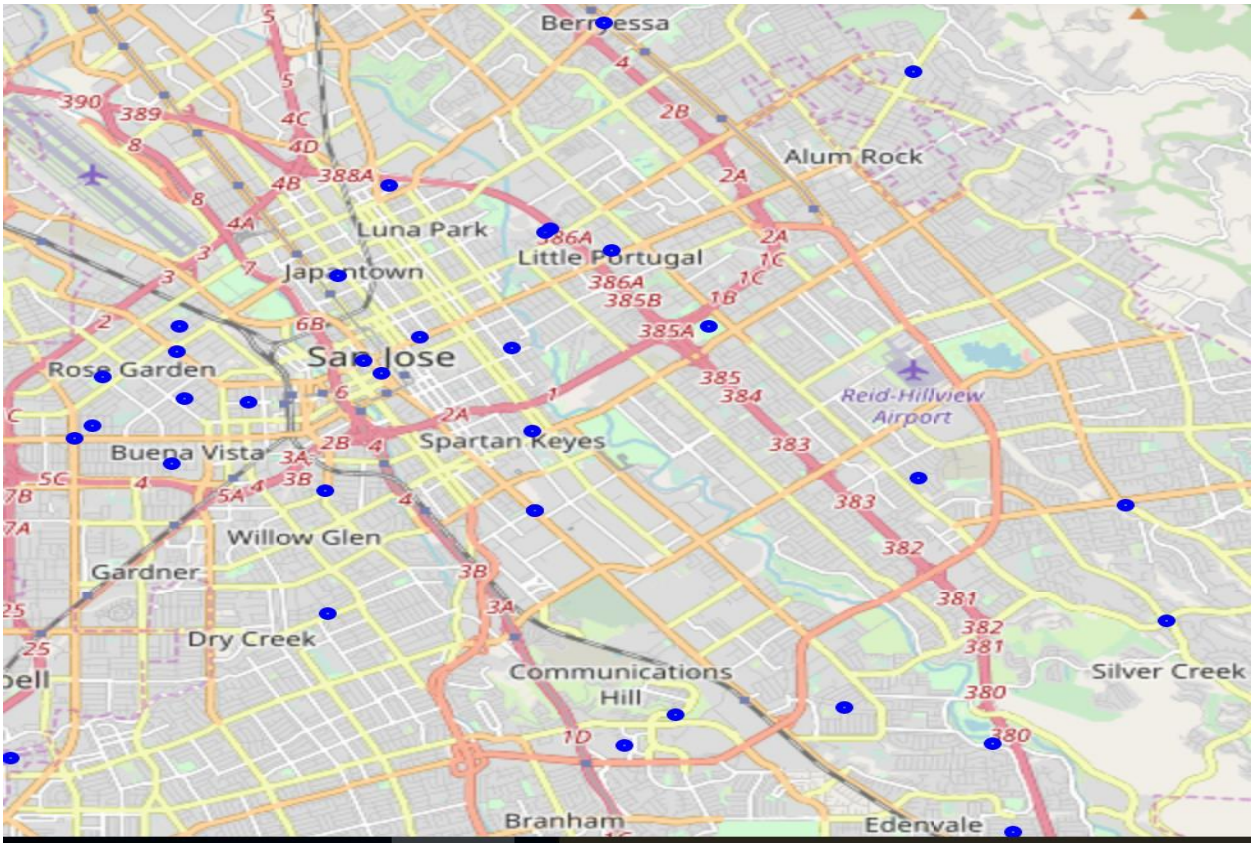
Fortunately, all the neighborhoods are listed in the Wikipedia in the link given in data secton and we will directly take this list into python and create a dataframe for these neighborhoods.

# The Battle of Neighborhoods (Week 2)

| Neighborhoods in San Jose, CA | | |
|---|---|---|
| The Alameda | East San Jose | San Pedro Square |
| Almaden Valley | Edenvale | Santa Teresa |
| Alum Rock | Evergreen | Santana Row |
| Alviso | Japantown | Seven Trees |
| Berryessa | King and Story | College Park |
| Blossom Valley | Little Portugal | Shasta Hanchett Park |
| Buena Vista | Luna Park | Silver Creek Valley |
| Burbank | Meadowfair | SoFA District |
| Cambrian Park | Midtown San Jose | South San Jose |
| Chinatown | Naglee Park | Spartan Keyes |
| Communications Hill | North San Jose | St. Leo |
| Coyote Valley | North San Jose Innovation District | West San Carlos |
| Downtown San Jose | Palm Haven | West San Jose |
| East Foothills | Rose Garden | Willow Glen |

## Geocoder Data

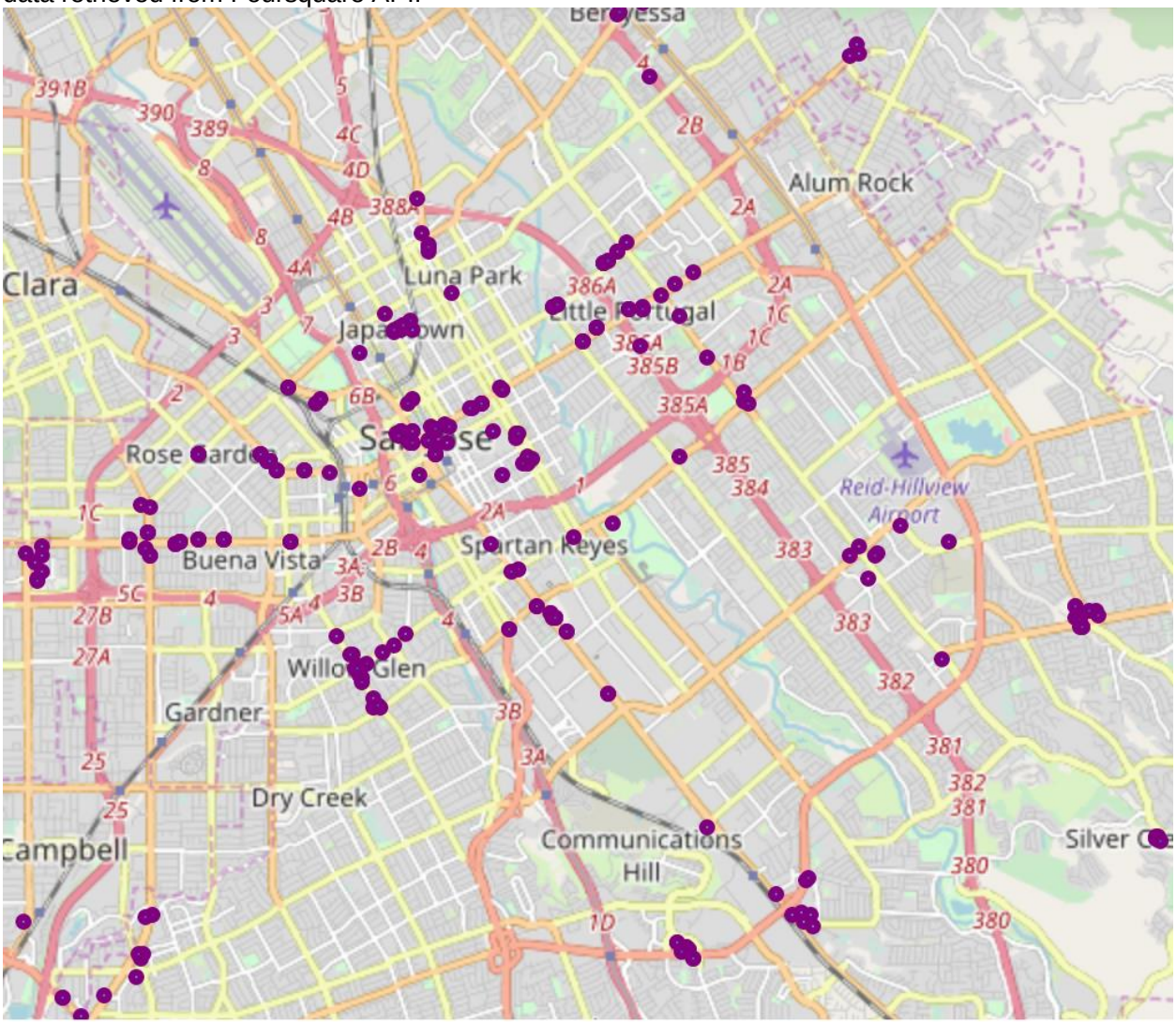
Geocoder library and Nominatim api will be leveraged to get the Geocoding data for each of the neighborhoods that will be analyzed in this project. Below is the map of neighborhood data after retrieving the latitude and longitudes.



# The Battle of Neighborhoods (Week 2)

## Location Data

A list of all restaurants and top venues along with their categories for each of these neighborhoods will be retrieved through the Foursquare API https://api.foursquare.com. For e.g, below is the restaurant data retrieved from Foursquare API:




# Methodology

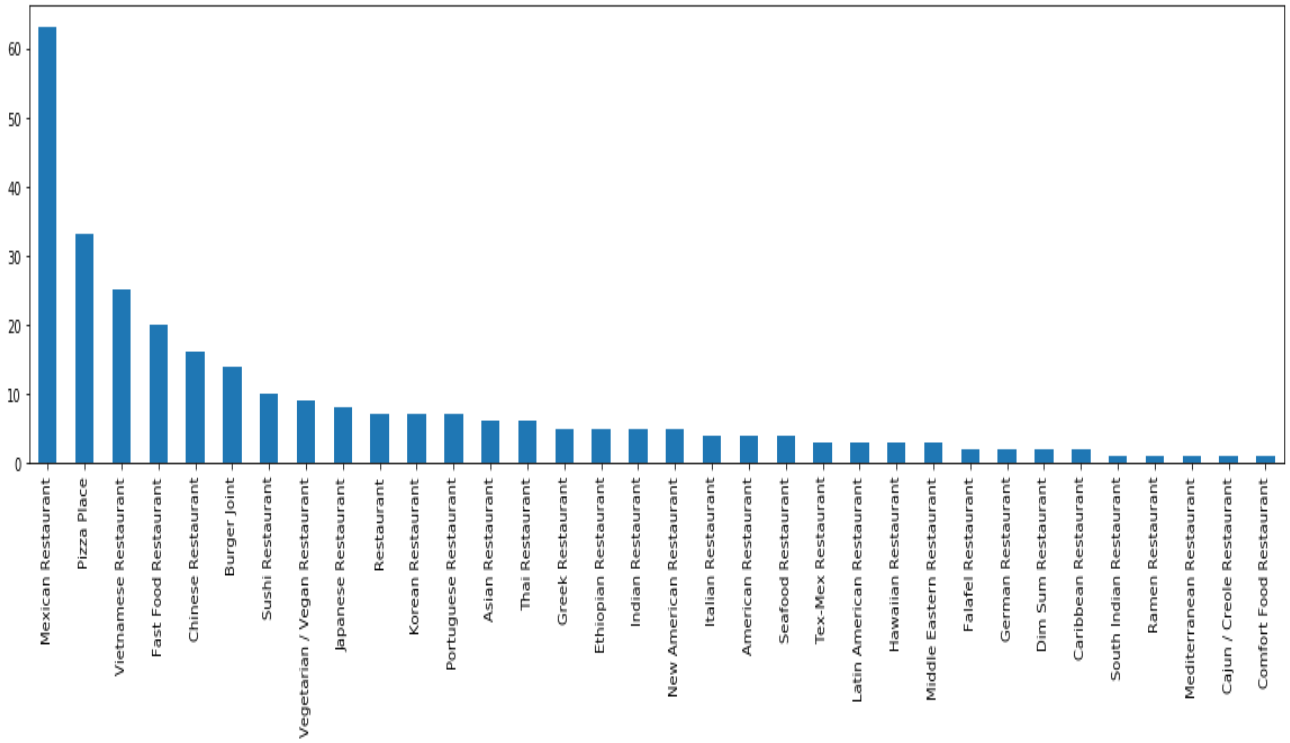
Once the data containing the list of restaurants and top venues and their categories retrieved, we will do the following activities:

1. Perform Data Wrangling to cleanup and reformat the data
2. Retrieve Retaurant and related categories from Venue category
3. See the distribution of Restaurants by Cuisine type
4. Identify neighborhoods with Indian restaurants
5. Find neighborhoods known for Restaurants
6. Pick a neighborhood that has the least ratio of Indian restaurants to overall restaurants

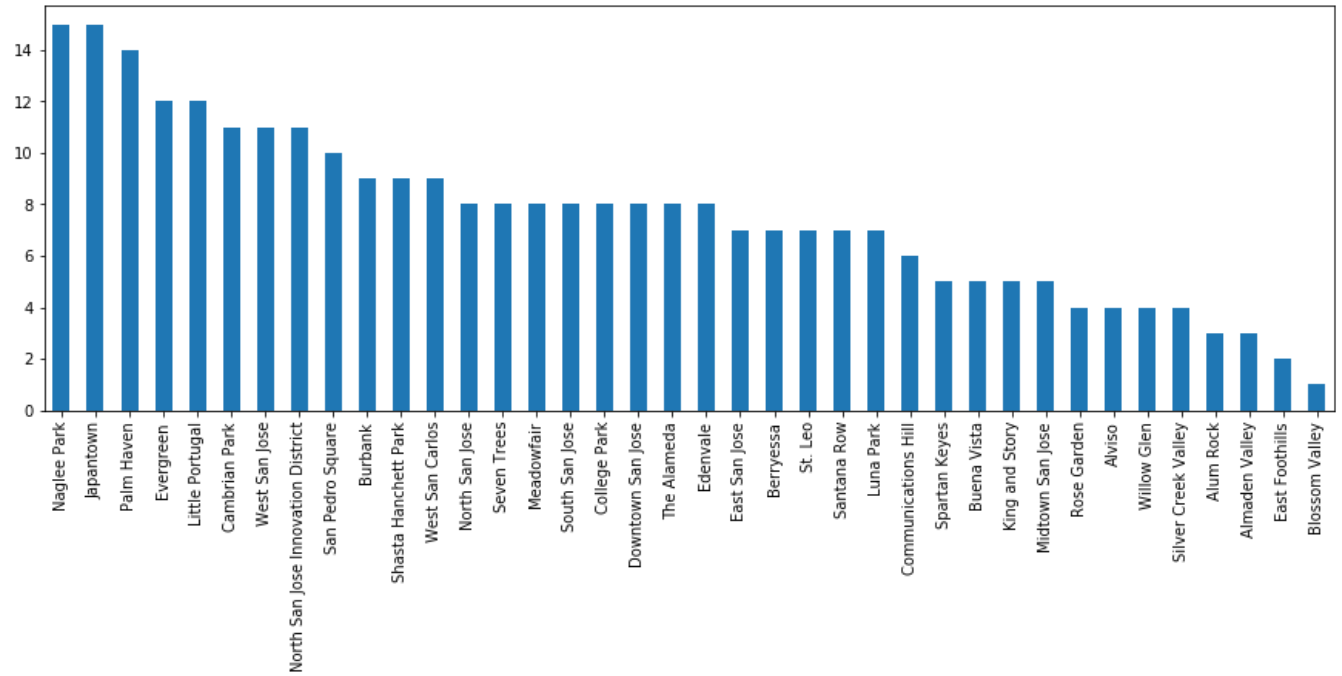
Recommend a best neighborhood to the Stakeholder for opening their new Indian restaurant

# The Battle of Neighborhoods (Week 2)

## Distribution of Restaurants by Cuisine type:




### Find which neighborhoods have more # of Restaurants:



# The Battle of Neighborhoods (Week 2)

## Identify neighborhoods with Indian restaurants:

| | Neighborhood | Neighborhood Latitude | Neighborhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category |
|---|---|---|---|---|---|---|---|
| **48** | Cambrian Park | 37.275156 | -121.940299 | Royal Taj | 37.269457 | -121.944910 | Indian Restaurant |
| **88** | Evergreen | 37.313290 | -121.790650 | Peacock Indian Cafe | 37.313507 | -121.793007 | Indian Restaurant |
| **105** | Japantown | 37.348010 | -121.896430 | Sagar Sweets | 37.345903 | -121.900682 | Indian Restaurant |
| **157** | Naglee Park | 37.337120 | -121.873120 | Punjab Cafe | 37.339317 | -121.883750 | Indian Restaurant |
| **186** | Palm Haven | 37.315552 | -121.898154 | New Indian Cuisine | 37.310389 | -121.897145 | Indian Restaurant |

## Results and Discussion

Our analysis shows that there are great number of restaurants in San Jose. There are pockets of low restaurant density as well. Highest concentration of restaurants was detected on 3 neighborhoods: 1) Naglee Park 2) Japantown 3) Palm Haven, so we focused our attention to these areas initially. Surprisingly, there aren't many Indian restaurants in these neighborhoods.

Those location candidates can be explored further. It is entirely possible that there is a very good reason for small number of Indian restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

## Conclusion

Purpose of this project was to identify San Jose neighborhoods with low ratio of Indian restaurants in order to aid stakeholders in narrowing down the search for optimal neighborhood for an Indian restaurant. By calculating #restaurants distribution from Foursquare data we have identified 3 neighborhoods that justify further analysis and can be used as starting points for final exploration by stakeholders.

Final decission 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 like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood including parking etc.