

# IBM Applied Data Science

## Coursera Capstone Final Project

*Project - Exploring venues in Chandigarh, India for identifying suitable venue based on location, price range and rating*

*By : Ashish Patil*

*June 2020*

# Business Problem Statement

- Location of a venue plays an important role in the success of business to the owner
- For a customer, the venue location matters a lot based on his/her status, paying capacity and choice
- Objective : To explore the venues in city of Chandigarh, India based on their category, price range, rating and locality and make this information available to users
- Business Question:
  - Any person who is visiting Chandigarh, India can use the project to quickly filter and select the venues that would suit their budget.
  - A tourism company can utilize this project information to create a website or a mobile application, which would be updated on a regular basis, to provide value added service to their clients. Third,
  - A businessman or a company who wants to start their own restaurant/cafe/hotel can utilize this information to identify the suitable location for their new startup.

# Data

- Data Required
  - List of venues in Chandigarh, India
  - Latitude and Longitudes of venues, so that they can be identified and plotted on maps
  - Venue average price range and ratings provided by users who had visited these venues earlier
- Sources of Data
  - Foursquare API (<https://api.foursquare.com/v2>) – For name, category and lat-long
  - Zomato API (<https://developers.zomato.com/api>) – For name, address, rating, price range and lat-long

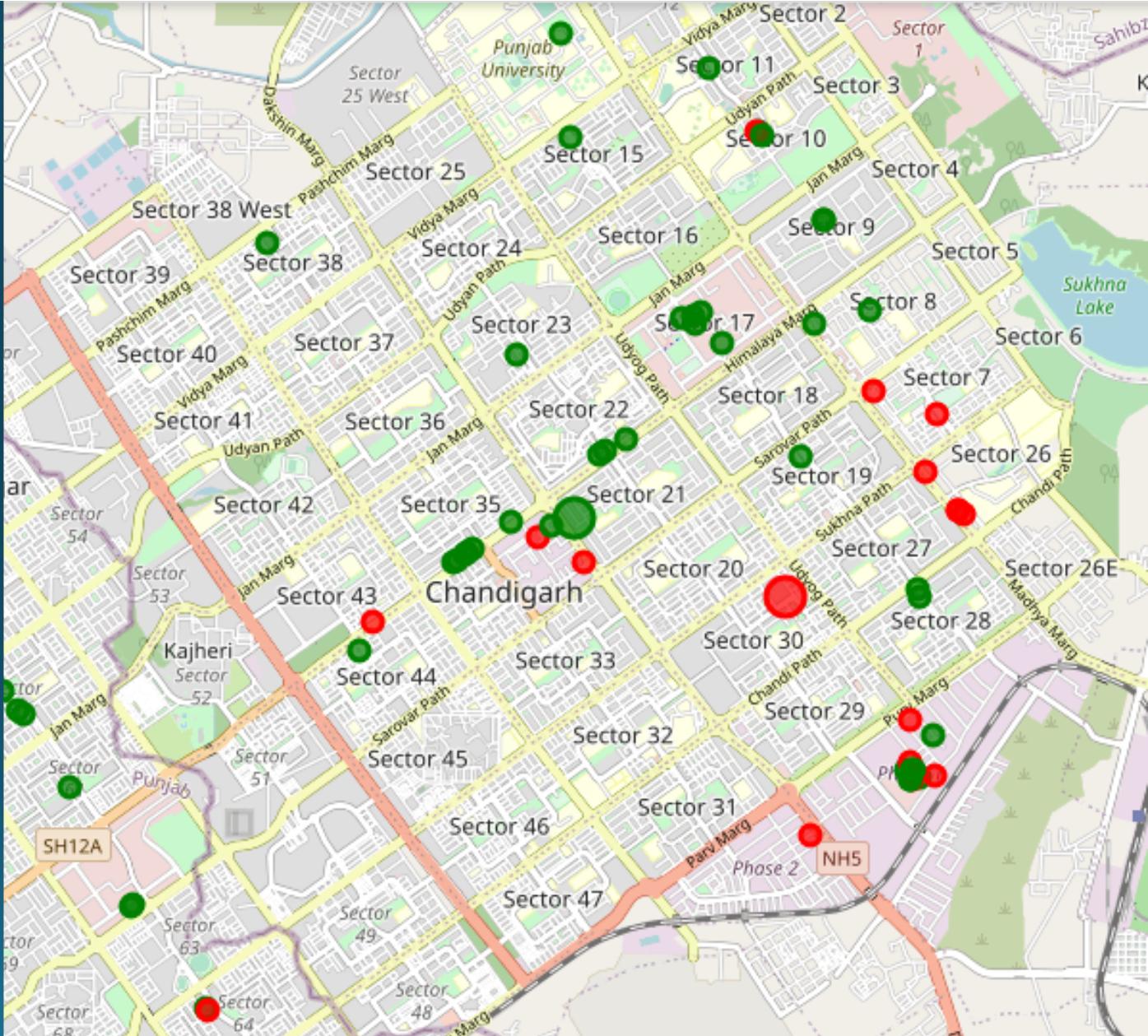
# Methodology

- Get latitude and longitude coordinates of Chandigarh City, India using Geopy.Geocoders package
- Use the Chandigarh City coordinates with Foursquare API to fetch venue data (Name, Category, Latitude & Longitude)within the radius of 5 Km
- Use Zomato API to fetch venue data (Name, Latitude & Longitude, Rating, Price Range)
- Combine the venue data fetched from two sources based on similarity in Name, Latitude and Longitude
- Clean the data by dropping the outliers identified in the data combining process
- After the combined dataset with clean data is formed, use it for further analysis
- Plot the venues on map based on rating, price range to derive meaningful information
- Perform clustering on the clean data by using k-means clustering
- Visualize the clusters in a map using Folium

# Results

Categorized the data in 2 Clusters:

- Cluster-1 (Green):
    - It is spread across the whole city
    - Includes majority of venues with different category
    - These venues have mean price range of 1.68 and rating spread of around 3.82.
  - Cluster -2 (Red):
    - This is very sparsely spread
    - Has very limited venues of specific category
    - These venues have mean price range of 3.12 and rating spread around 4.04



# Discussion

- During feting process from Foutsquare and Zomato API, we received data for 130 venues
- After cleaning of data we left we 63 venues for further analysis
- Majority of the venues in final dataset were Cafes and Indian Restaurants
- Though the rating range is 1-5, but majority of venues have rating close to 4. It means majority of the venues serves the good quality of food and service
- When these venues were plotted on map, we found that Sector 17, 35 and Elante Mall have higher density of venues.
- We discovered that venues located near Sector 17 and 35 are relatively priced lower than venues in Sector 7 and 26.
- A mix of low priced and high-priced venues exists in Elante Mall.

# Recommendations & Conclusion

- Customer Perspective

- We can conclude that Sector 35, 17 and Elante Mall have venues with good quality of food and service and yet they have wide price range. Hence there are venues available for different category (Middle to High class) of consumers.
- Venues in Sector 7 and 26 can only be afforded by upper-middle and high-class customers.

- Business Perspective

- In Sector 35, 17 and Elante Mall venues are densely available, hence there will be high competition if anyone plans to setup new venue.
- However, in spite of high competition the probability of success is very high due to heavy foot fall of customers in these sectors. Hence the risk is relatively low.
- We can also explore other sectors in the proximity of sector 35, 17 and Elante Mall for setup of new venue. As these sectors have potential to get developed in future

# Thank You