Geolocating Ideal venues for Indian Restaurants in San Diego

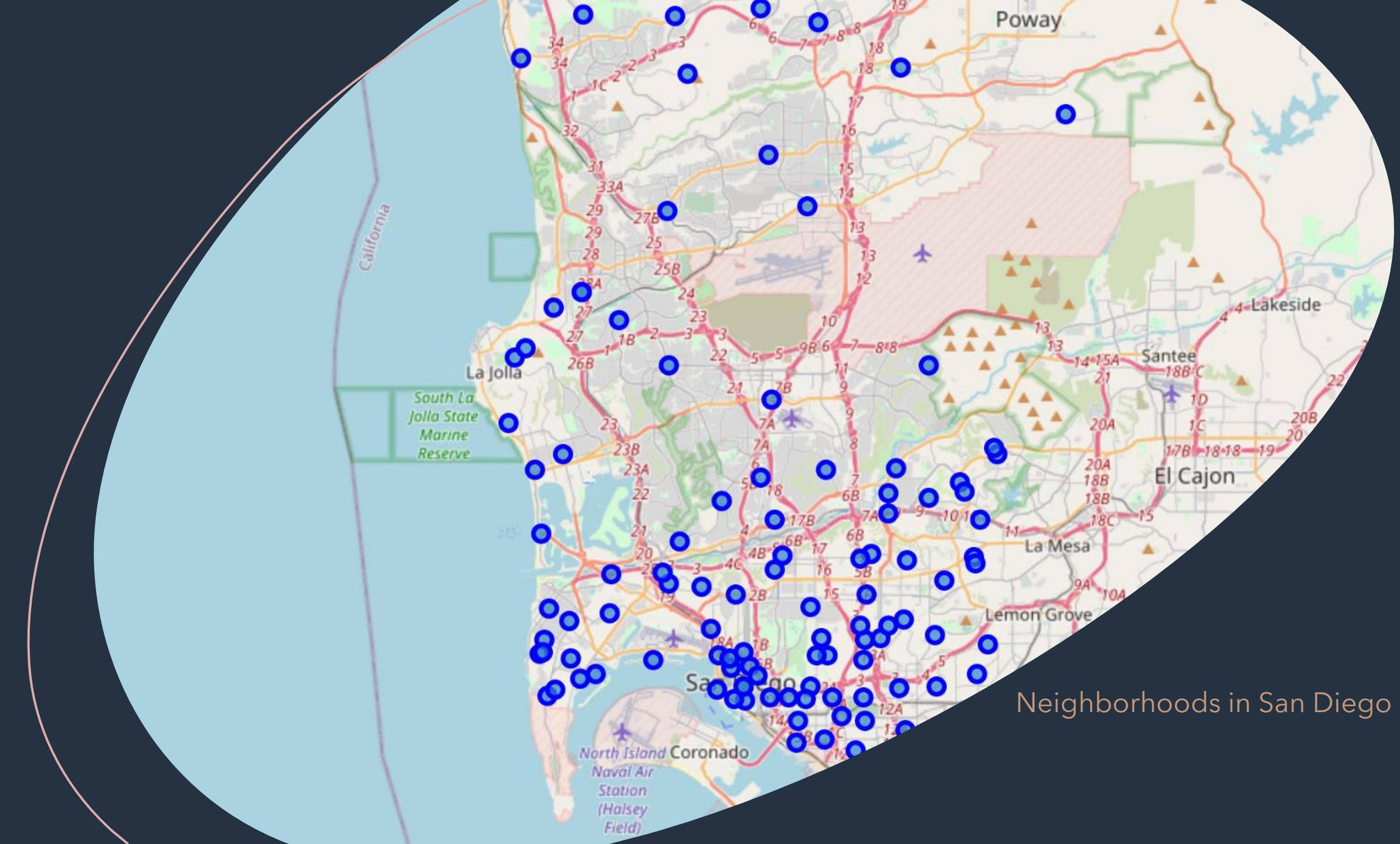
Divyaprakash Dhurandhar

Business Problem

- Growing demand for Indian food and restaurants across the United States.
- We assume that a client in the city of San Diego wants to open an Indian restaurant.
- To maximize revenue and business success, insights into existing Indian restaurants and peer competition are required.
- The project's main objective is to find ideal spots in the city where Indian restaurants can be set up.

Data

- The neighborhood data is scraped from a Wikipedia webpage https://en.wikipedia.org/wiki/Category:Neighborhoods_in_San_Diego
- The latitude and longitude of the neighborhoods are retrieved using the geocoder API.
- The venue data is found out by passing in the required parameters to the FourSquare API, and creating another DataFrame to contain all the venue details along with the respective neighborhoods.

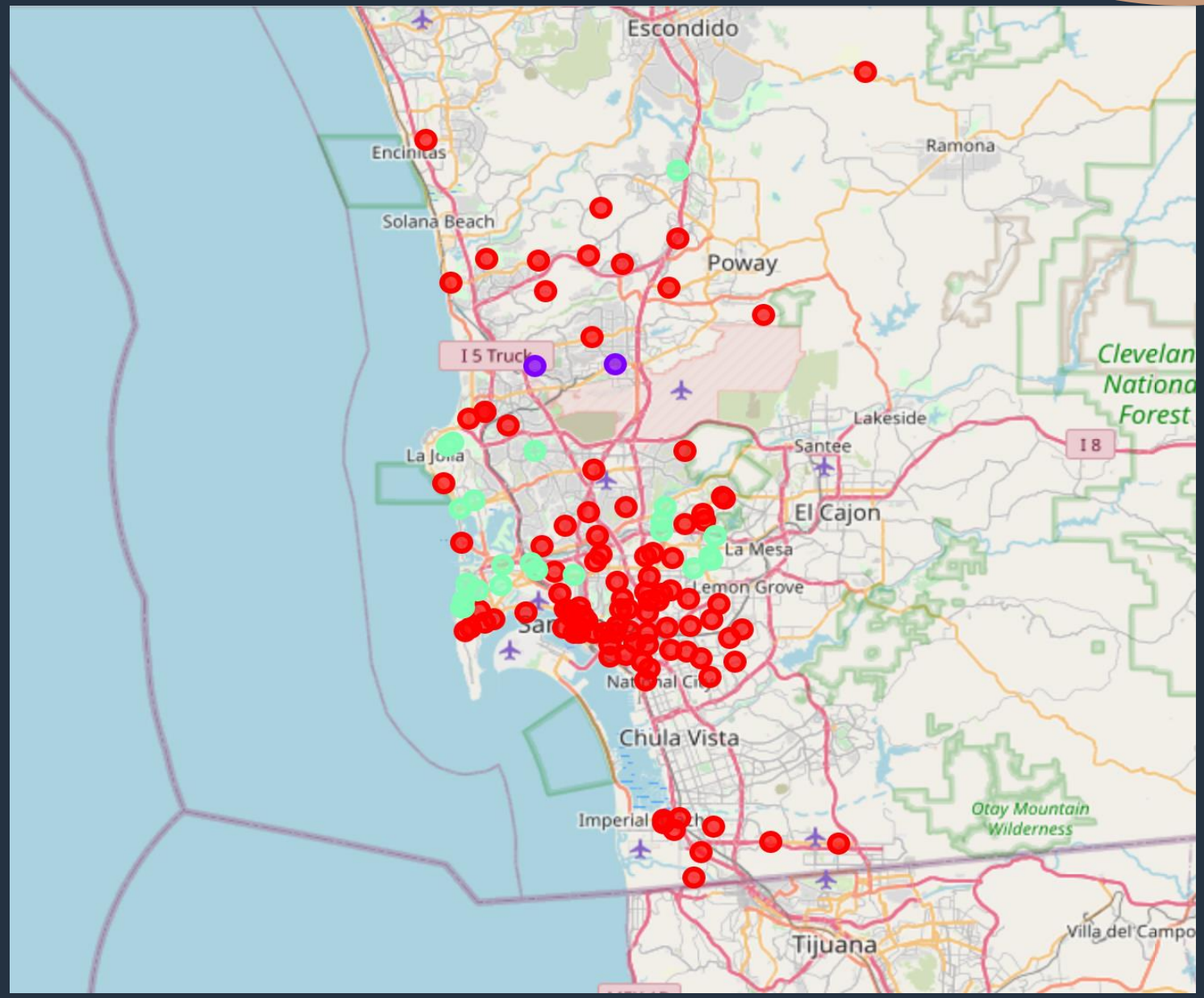


Neighborhoods in San Diego

Results

- The results from the k-means clustering show that we can categorize the neighborhoods into 3 clusters based on the frequency of occurrence for Indian Restaurants:
- Cluster 0: Neighbourhoods with a low number of Indian Restaurants
- Cluster 1: Neighbourhoods with a high number of Indian Restaurants
- Cluster 2: Neighbourhoods with a moderate concentration of Indian Restaurants

Clusters representing the density of Indian Restaurants in neighborhoods



Conclusion

- In this project, we have gone through the process of identifying the business problem.
- Conducting EDA, performing machine learning by clustering the data into 3 clusters based on their similarities.
- Providing recommendations to the relevant stakeholders i.e., people in the business of restaurants and investors regarding the best locations to open a new Indian restaurant.