

Coursera Capstone

IBM Applied Data Science

***Setting up a new Bar in the city of Los Angeles,
United States of America***

By: Aniruddh Mallya
8 August 2020



Business Problem

- The location of setting up public infrastructure/venues is a major factor that decides as to whether a business would be successful or not.

➤ Problem Statement:

The main objective of this Capstone is to answer the question as to **which locations in Los Angeles (LA) would be most preferable to set up a new bar business.**

- On comparison of the spread of bars across various locations with the corresponding population density in LA, we observe patterns that make this project relevant and timely.
- By making use of various data science methodologies and Machine Learning tools, this paper aims to provide a conclusive proof to the underlying problem statement.

Data Required:

- List of neighborhoods/districts in Los Angeles.
- Latitude and Longitudes coordinates
- Venue data corresponding to Bars.

Sources of Data:

- Wikipedia page of neighborhoods

https://en.wikipedia.org/wiki/List_of_districts_and_neighborhoods_of_Los_Angeles

- Geopy and Geocoder packages for latitude and longitude coordinates
- Foursquare API to obtain venues data

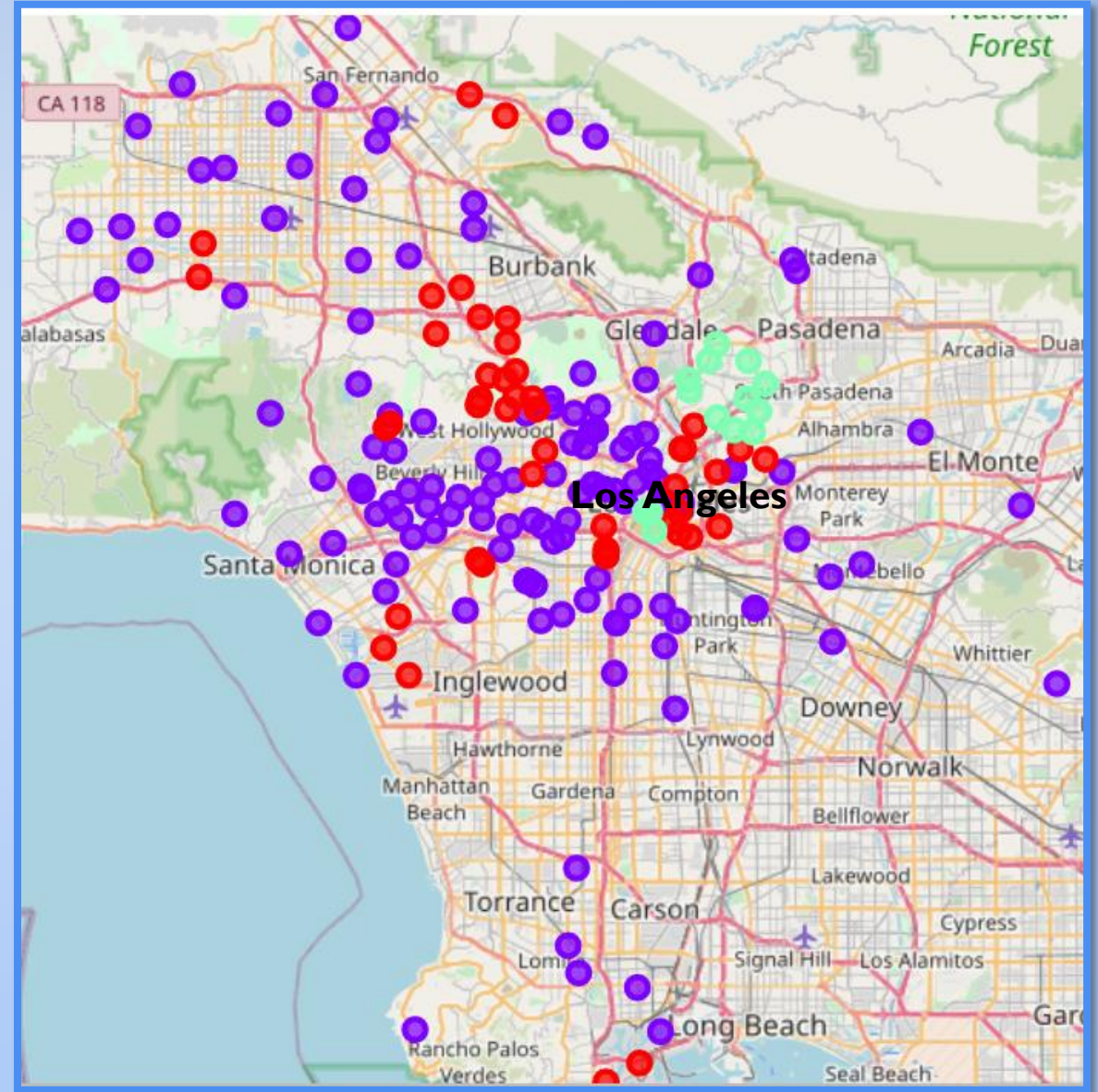
Methodology:

- ➔ Web Scraping (Wikipedia)
- ➔ Importing of Python Libraries
- ➔ Data extraction from Foursquare API
- ➔ K-Map Clustering Algorithm
- ➔ Elbow Curve visualization
- ➔ Folium map for plotting, labeling and cluster visualization

Results:

The neighborhoods were segmented to 3 clusters:

- Cluster 0: Moderate density – Red color
- Cluster 1: Low density – Purple color
- Cluster 2: High density – Pale Green color



Discussions:

- Most of the bars are concentrated in the central area of the city
- Suburbs have fewer number of bars, so that people residing here have to travel farther for their drinks.
- Sometimes this can lead to the situation of unnecessary crowding at specific locations, oversupply of items such as drinks and lead to an imbalance in the supply chain.

Recommendations:

- ❖ As far as possible, try and avoid neighborhoods in 'Cluster 2', since there's already a high concentration of the number of bars implying very high competition.
- ❖ Since there is far less competition in 'Cluster 1' neighborhoods, new businesses have a great opportunity to quickly establish themselves and capitalize.
- ❖ For the existing businesses that are well established, diversifying can contribute to improved productivity. However, in order to avoid taking too much of a risk, 'Cluster 0' with low-moderate density would be seemingly a good option.

Conclusion:

- In my opinion, cluster 1 would be best preferred location to set up a 'bar' business, which answers the main business question.
- Thus, the findings of this paper will help the relevant stakeholders to capitalize on their opportunities on high potential locations while avoiding overcrowded areas with high competition to open a new bar.

Thank
you!