

Introduction:

For this project, we will be assuming the role of a company asking ourselves the following question; how can we best target clients in the greater Los Angeles area. We will be attempting to reach the greatest number of potential clients, and to nuance our advertisements to appeal to the greatest number of businesses in certain areas.

Data:

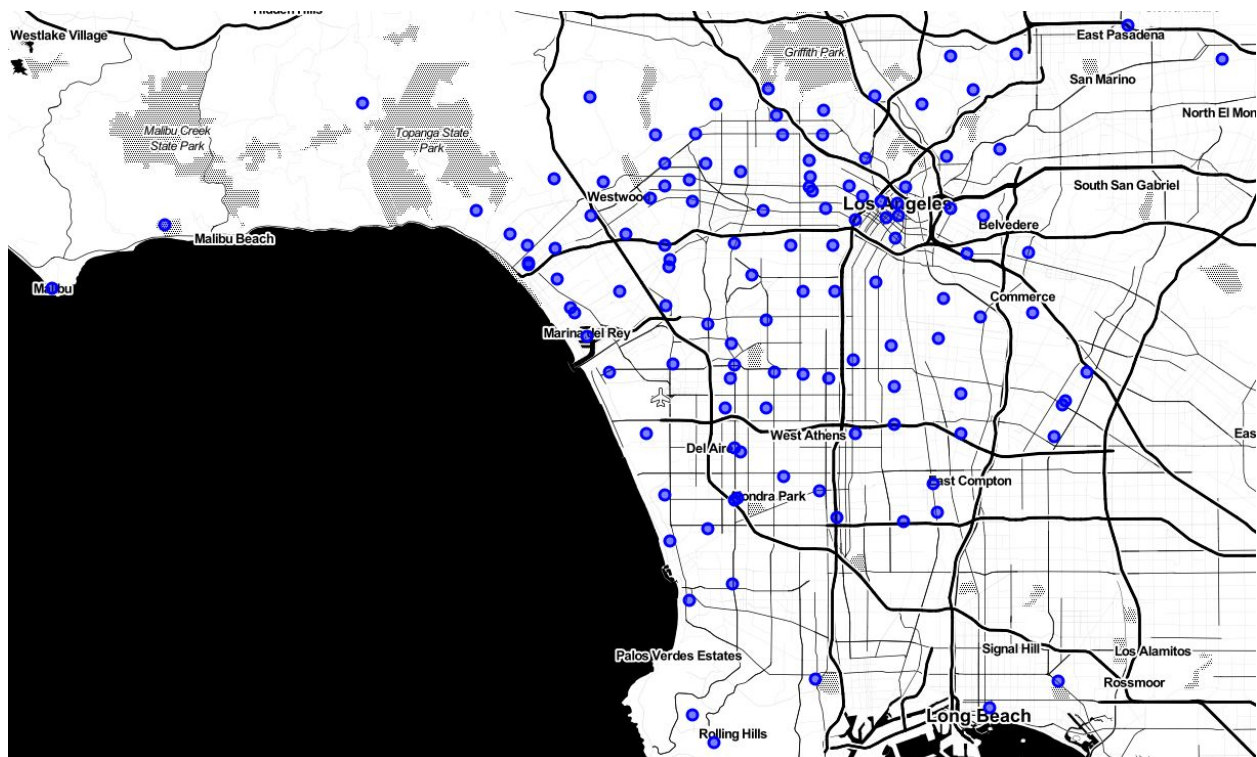
Geospatial and ZIPCode data were gathered free of charge from the Geonames geographical database at

<https://www.geonames.org/postal-codes/US/CA/california.html>.

From these we pulled only data from areas in the county of Los Angeles. By requesting from the FourSquare API and using these geographic coordinates we obtained information for business venues nearby each ZIPCode location.

Methodology:

To begin, we plotted the coordinates on a folium map of Los Angeles. We can see that we have a good spread of points across the county.



In order to avoid overlapping, we decide to limit our search of businesses to 500 meters around each of these locations. We use the Foursquare API to create a dataframe of which includes, the business name, the type of business, and the ZIPCode in which it resides. We create another dataframe from this which shows the average amount of each business type per ZIPCode. Finally, we use the common types of businesses around each ZIPCode to put them into 8 groups (or kclusters). In other words, each group contains ZIPCodes with similar types of business.

Results:

We display the different clusters with a dataframe comprised of:

- Code: ZIPCode of the area
- Place: City where the ZIPCode is found
- Latitude: North-South Coordinate
- Longitude: East-West Coordinate
- Cluster Labels: Group assignment based on nearby venues
- #1-5 Venue: Top 5 most common venues in the area

We also include a color coded map displaying each location and the area its associated businesses occupy.

Discussion/Conclusion:

Most of the ZIPCodes were placed into one group, cluster 5. This cluster consists mostly of various kinds of restaurants. Other notable groups include clusters 1 and 7, which consisted of mostly parks and trails/beaches, respectively. These two clusters are probably areas that can be avoided.

This data allows us to not only see the types of business around each ZIPCode, but the larger areas in which those businesses are concentrated. Therefore we can determine the most beneficial places to advertise depending on our companies goals and in what way we should proceed in doing so.