

Intoduction/Business Problem:

For this project, we will be assuming the role of an IT company asking ourselves the following question; how can we best implement advertisement for ourselves in the greater Los Angeles area. We will be attempting to reach the greatest number of potential clients, and to nuance our advertisement 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 the aforementioned data from areas in the county of Los Angeles. By requesting from the FourSquare API and using these geographic coordinates we obtained information for venues nearby each Zipcode location. We then obtained the means of each venue category for each Zipcode. Using these means we grouped the zipcodes into 9 groups (or kclusters) and displayed them using a map. We also include a table 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

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