

## **Capstone Project**

### **The Battle of Neighborhood**

Segmenting and Clustering Neighborhoods of San Francisco and

Los Angeles

By,

Charles C Gomes

## Contents

Introduction.....	3
Objective.....	4
Data.....	5
Conclusion.....	6

## **Introduction**

San Francisco and Los Angeles are two major cities in California.

Brief information about both cities:

- San Francisco: officially the City and County of San Francisco, is a city in, and the cultural, commercial, and financial center of, Northern California. San Francisco is the 13th-most populous city in the United States, and the fourth-most populous in California, with 883,305 residents as of 2018.
- Los Angeles: officially the City of Los Angeles and often known by its initials L.A., is the most populous city in California, the second most populous city in the United States, after New York City, and the third most populous city in North America. With an estimated population of nearly four million,[11] Los Angeles is the cultural, financial, and commercial center of Southern California.

## Objective

In this project, we will study in details the area classification using Foursquare data and machine learning segmentation and clustering. The aim of this project is to segment areas or neighborhood of San Francisco and Los Angeles based on the most common places captured from Foursquare.

Using segmentation and clustering, we hope we can determine:

1. The similarity or dissimilarity of both cities
2. Classification of area located inside the city whether it is residential, tourism places, or others

# Data

The data for neighborhoods is acquired from following sources

Los Angeles - <https://data.lacity.org/api/views/nwj3-ufba/rows.csv?accessType=DOWNLOAD>

San Francisco - <https://data.sfgov.org/api/views/xfcw-9evu/rows.csv?accessType=DOWNLOAD>

Additionally Foursquare data api was used for getting different kind of venues for segmentation and clustering.

## **Conclusion**

With the help of Foursquare API, we were able to capture the venue information and using venue information, we can figure out the similarities or dissimilarities of San Francisco and Los Angeles. We did classification of Neighbourhoods as Residential, tourism or Mixed. In conclusion, both cities San Francisco and Los Angeles have tourism as similarity as well as there are some residential areas. It is somewhat clear that in San Francisco, the residential and tourism neighbourhoods are mixed compare to Los Angeles.