**1. Description of the problem and discussion of the background**

Moving from one part of the country to another has always been exciting for many people. This will provide opportunities to explore new places, make new friend and try new food. However, this excitement often entails some challenge and anxiety. One quite often thinks of which part of a new area to get an apartment, how the rent price differs across the neighborhoods, which neighborhood has more jobs that fit my abilities and skills. As a future data scientist, I have decided to analyze these questions and help the travelers to make the right decision.

Los Angeles (LA) is one of the biggest and most populous cities of the USA. Due to its Mediterranean climate, ethnic diversity and Hollywood entertainment industry, every year many people come and enjoy all the benefits the city offers [1]. For this capstone project, I have decided to explore central Los Angeles area which consists of 26 neighborhoods [2]. I will investigate the most common venues in each neighborhood. Then, the neighborhoods will be clustered based on similar venues. Also, I am going to incorporate the average rent price of apartments into analysis. Finally, all the analysis and observations will be depicted in a map which will provide visual information for the prospective travelers to central LA area.

**2. Description of the data and how it will be used to solve the problem**

The names of 26 neighborhoods will be scraped from a web page [2]. The latitude and longitude values of each neighborhood will be extracted from a web page [3] as well as [4]. As for the average rent prices for the neighborhoods, those will be obtained from the internet [5]. The geojson data that defines boundaries of the neighborhood is not quite readily available [6]. Therefore, the source data will be explored, cleaned and reshaped to meet the requirements of this project. The venues data will be extracted using Foursquare API. Once we have collected all the necessary data, we are ready to do the analysis and solve the existing problem. The data will be used to cluster the neighborhoods based on their similar venues. This will help the interested parties to choose a particular set of neighborhoods that fit their requirements best. Moreover, the choropleth map will be constructed based on the average rent price of apartments in each neighborhood which will help the travelers to make wise decision before settling down in a specific neighborhood. Finally, some recommendations will be given according to the observed analysis of the project.

**Bibliography**

1. <https://en.wikipedia.org/wiki/Los_Angeles>

2. <https://en.wikipedia.org/wiki/Central_Los_Angeles>

3. <http://wiki.stat.ucla.edu/socr/index.php/SOCR_Data_LA_Neighborhoods_Data>

4. <https://usc.data.socrata.com/dataset/Los-Angeles-Neighborhood-Map/r8qd-yxsr>

5. <https://www.rentcafe.com/average-rent-market-trends/us/ca/los-angeles/>

6. <https://s3-us-west-2.amazonaws.com/mappingla.com/downloads/neighborhoods/la_city.json?_ga=2.101578435.35324892.1592446928-1423563668.1590043322>