

Paul Tran
Applied Data Science Capstone
July 6, 2020

Place To Rent In Los Angeles.

I. Introduction

This is the report for final course of IBM Data Science Professional Certificate, 9 series course create by IBM. The final class have been equipped with the skills and the tools to use location data to explore a geographical location, over the course of weeks, students will have the opportunity to be as creative as they want and come up with an idea to leverage the Foursquare location data to explore or compare neighborhoods or cities of their choice or to come up with a problem that you can use the Foursquare location data to solve.

Los Angeles, the largest city of California, and one of the most populous city in the U.S. With its perfect Mediterranean climates, it is home of approximate 4 million people (1). The city also has many well known Universities such as UCLA, USC, and CalTech. Each year, hundreds thousands of people come to this city for good. Many of them are students, fresh graduated, working class, contractors, or even tourists. Therefore the demands for finding a place to settle down is essential.

The idea of comes from a process of a person try to figure out a place to rent after moving from an other city or even from a different country. It is common that people will looking for places that have reasonable prices and located closed by some kinds of venues such as restaurants, coffee show, and markets, ect. Therefore, what types of surrounding venues will affects the rent prices positively and negatively?

For this project, the main goal will be exploring the neighborhood of Los Angeles County in order to find out the correlation between the Counties's surrounding venues and the rents prices of it's neighborhood. The project will target:

- Students who attend at any Universities in Los Angeles county.
- Tenants who try to find a new place to rent.
- Business looking for a place to rent

- Tourist
- Renting agents or landlords who want to optimize their renting advertisement
- And last but not least, to this class's instructions and classmate who will grade this project.

II. Data Overview:

The data that being use in this project will available at "maps.latimes.com". Those datas provide the population density (measures the numbers of people per square mile) of each neighborhoods in Los Angeles. This dataset also be used to get the latitude and longitude of each neighborhood by using *geocoder.arcgis* library. Moreover, it also being used to visualize LA population's distribution, therefore audiences will have a broad idea of each locations.

In a addition, there were a dataset from "<https://usc.data.socrata.com/>" where it contains the median value of gross rent prices in an area, measured in dollars of from the year of 2010 to 2016. This median rent price measures the gross rent of that area, meaning that, it will include all the utility such as electricity, water, gas and sewage.

The geojson file of Los Angeles communities uses to indicate neighborhood boundaries on the map taken from "<https://boundaries.latimes.com/>". Surrounding venues for each city in LA county was searched using FourSquare API base on the latitude and longitude of each city gathered using *geocoder.arcgis* library from the "maps.latimes.com" dataset that being mentioned above. This data later will be used to compare, analyze and cluster.