High school in San Diego

Sunil Yadavalli,
February 2020

Introduction

In today's world people keep switching between jobs for various reasons like career prospects, subject of interests, personal agenda etc. Thanks to the various channels of opportunities, we are able to find a good job that addresses our needs. However, this decision is governed by a lot of factors that takes priority and enable us to make a decision. One such factor, for people having school going kids, is to learn about the good schools in the neighborhood of the perspective job location. This is very important because we would want the kids get the best education and there is no compromise with their career.

Business Problem

Assume that someone got a job opportunity in the beautiful city of San Diego, California. The question is, where you would recommend the locality to stay so that there are enough high schools around with a minimal commute. Also, in the available options, which school would be the best option to opt for if there is a chance of admission? Can machine learning techniques like clustering, address this problem? Let's see.

Data

To address the above problem, we would need the below:

- List of the neighborhoods in San Diego
- The latitude and longitude of those neighborhoods
- Venue data, particularly the high schools. This data will be used to perform clustering

We will leverage the Foursquare API to collect the data. We then apply the machine learning techniques on this data and also utilize map visualization (Folium with heat map)

Below is the sample data

Identifier	SchoolName	ShortName	Address	PostalCode	Latitude	Longitude
4bbe5d61b083a5933ee5a1e9	San Diego SCPA	High School	2425 Dusk Dr	92139	32.679879	-117.048609
4b548e8ff964a520e5bf27e3	Helix Charter High School	High School	7323 University Ave	91942	32.754317	-117.037150
4bba201d7421a593db6bc340	Garfield High School	High School	1255 16th St	92101	32.718490	-117.149067
4dd293b61838a751965e6ca4	SDEMC	High School	1425 Russ Blvd Ste T112D	92101	32.719405	-117.151360
4bc8950d6501c9b6aa664029	San Diego High School	High School	1405 Park Blvd	92101	32.721871	-117.152180

In the further sections we will talk about the methodology of using this data and apply data analysis.