Explore Santa Clara : A Traveller's Guide to Best Businesses in Town

Introduction:

A traveller coming to explore the town of Santa Clara would like to know which businesses are better with the purpose to either have a pleasant trip, buy a house and also know how much congested the city is. The purpose of this project is to help people in gathering an insight based on business ratings and traffic congestion in the city. It will help people making smart and efficient decision on deciding which places to focus on in the city.

Lots of people are migrating to Santa Clara as multiple tech companies like Google, Apple, LinkedIn are in the vicinity. They try to search for houses or come to explore the city or on business conventions. They try to get a knowhow of the city. This project is for those people who wish to know about the businesses in the city and also get information on the congestion in the city based on the traffic incidents in the city.

This project aim to create an analysis of features for people visith Santa Clara to search the best businesses as a comparative analysis between zipcodes. The features include traffic congestion data to give a better understanding of travel times around the city,

It will help people to get awareness of the area before moving to a new city, state, country or place for their work or to start a new fresh life.

Problem Which Tried to Solve:

The major purpose of this project, is to suggest a better streets and zip codes in a new city for the person who are shiffting there. Social presence in society in terms of like minded people. Connectivity to the airport, bus stand, city center, markets and other daily needs things nearby.

1. Sorted list of businesses in terms of business ratings in descending order.
2. Sorted list of pincodes in terms of number of traffic incidents.

The Location:

Santa Clara is a popular destination for people who decide to visit the Bay Area in terms of business and tourism. The Big tech companies are in the vicinity of the city. As it attracts a lot of jobs, it is one it is one the most diverse and multicultural areas in the Bay Area, being home to various religious

groups and places of worship. Although resulting immigration has become a hot topic over the past few years with more governments seeking more restrictions on immigrants and refugees, the general trend of immigration into Canada has been one of on the rise.

Foursquare API:

This project would use Four-square API as its prime data gathering source as it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about a business.

Work Flow:

Using credentials of Foursquare API features of zipcode, streetname, business name and category of business with business rating of the neighborhoods would be mined. Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 100 and the radius parameter would be set to 500.

Clustering Approach:

To compare the different pincodes, we decided to explore the areas, segment them, and group them into clusters to find similar neighborhoods in the city of Santa Clara. To be able to do that, we need to cluster data which is a form of unsupervised machine learning: k-means clustering algorithm

Libraries Which are Used to Develop the Project:

Pandas: For creating and manipulating dataframes.

Folium: Python visualization library would be used to visualize the neighborhoods cluster distribution of using interactive leaflet map.

Scikit Learn: For importing k-means clustering.

JSON: Library to handle JSON files.

XML: To separate data from presentation and XML stores data in plain text format.

Geocoder: To retrieve Location Data.

Bing : To retrieve Traffic Incident Data.

Yelp : To get business rating data.

Matplotlib: Python Plotting Module.