Commuter Friendly Neighborhoods of San Francisco, CA, USA

Golden Gate Bridge, San Francisco

IBM Applied Data Science Capstone Project
The Battle of Neighborhoods

Usha Manoharan

July 2020

Business Problem

San Francisco (fondly called "The City" by the locals), is a beautiful and culturally diverse city. It has many parks with lush greenery, coastline trails with views of the Pacific Ocean, world class museums and ample options for dining. People from the Chinese, Japanese, Latino, LGBTQ, Hippie, Italian, Southeast Asian cultures have all made San Francisco their home for many decades resulting in culturally vibrant neighborhoods. Above all, many industries and particularly high-tech companies like Airbnb, Dropbox, Facebook, Salesforce, Uber…to name a few, have their presence in San Francisco. Top notch universities like Stanford and Berkeley are in the proximity which help in bringing talented individuals closer to the city. All these factors attract people from all walks of life to the city to live and make a living.

To provide a smooth and stress free transition to a person moving into San Francisco, I would like to address the problem of identifying the best "commuter friendly" neighborhoods where people can live and use economical options such as buses and trains to commute to the neighborhoods for work, pleasure, dining and other activities of interest within the city.

Introduction

The goal of this capstone project is to identify the neighborhoods of San Francisco that are best suited for people to commute using public transportation.

My approach is to search the web and obtain a list of zip codes of San Francisco, then find their geographical co-ordinates to use as input to the Foursquare API to retrieve transit related location data. To keep it simple, I will be looking for venues that relate to buses or trains. I will use the K-means algorithm to cluster the zip codes into segments with similar venue categories. I will also use the Folium library to visually represent the zip codes and the commuter friendly clusters in a map of San Francisco.

The outcome of this analysis should lead to the identification of zip codes in San Francisco that provide similar options to commute by bus or train. I think this analysis will benefit a wide variety of people who is seeking economical modes of transportation in the city of San Francisco. Fresh graduates taking up new jobs, tourists visiting the city, immigrants trying to move in and establish themselves, city/transportation officials interested in knowing about the distribution and frequency of bus/train stations in the city are all some of the examples that I can think of.

Data

Let's take a closer look at the data and its source(s), used in this project.

Zip codes of San Francisco

I downloaded an excel file from the web at https://www.melissa.com/v2/lookups/cityzip/city/?name=san+francisco, which contains the list of zip codes of San Francisco.

	ZIP
45	94143
46	94137
47	94080
48	94083

Figure 1: Sample list of zip codes of San Francisco

Geographical co-ordinates

Next, I used the "pgeocode" library from https://pypi.org/project/pgeocode/ to obtain the latitude and longitude for each of the listed zip codes.

	zipcode	latitude	longitude
45	94143	37.7631	-122.4586
46	94137	37.7749	-122.4194
47	94080	37.6574	-122.4235
48	94083	37.6547	-122.4077

Figure 2: zip codes with latitude and longitude information

Location Data from Foursquare

I used the strings "bus" and "train" in the foursquare search query API, to obtain the venues related to buses and trains for each of the zip codes.

	Zipcode	Zipcode Latitude	Zipcode Longitude	Venue Name	Venue Category	Venue Address	Venue Latitude	Venue Longitude
0	94163	37.7749	-122.4194	MUNI Bus Stop - Van Ness & Market	Bus Station	Van Ness Ave.	37.775657	-122.419270
1	94163	37.7749	-122.4194	Muni Bus 5474	Bus Line	NaN	37.774795	-122.420017
2	94163	37.7749	-122.4194	MUNI bus 8143	Bus Line	NaN	37.774355	-122.420631
3	94163	37.7749	-122.4194	MUNI bus #8198	Bus Line	NaN	37.776567	-122.419050
4	94163	37.7749	-122.4194	MUNI Bus Stop - Mission & 11th	Bus Line	Mission St	37.774149	-122.417307
...	...	...	...	...	...	...	...	...

Figure 3: Venues obtained from Foursquare API for search query "bus"