**Description of the problem and discussion of the background:**

I have received a job offer in San Francisco, CA. This city is a central hub for lots of start-up companies. A few of my friends from college are already living there. And they have assured me about good public transportation of this city (Muni Train & BART train – can be used to commute to go anywhere in the city).

However, for me personally, I like to stay away from crowded city life – would rather like to live a neighborhood that is peaceful and quiet after work hours. I am also a photographer and like to go on long hike during the weekends.

For this exercise – I would explore the boroughs and neighborhoods in San Francisco city and pick one neighborhood that is close to beach or has hiking trails or photographic scenarios – along with a bus stop/train stop for commuting to work.

**Description of Data – How it will be used to solve problems**

I am using two sets of information here:

1. First set of data is been download from San Francisco, CA – city government website in csv format. This data set has all the Boroughs (as Plan District) and Neighborhoods of the city listed along with other information. After downloading this data set, I have imported it into my work using Panda and converted into a data frame, kept the columns I need to get an idea of the names of the boroughs and their associated neighborhoods in/around San Francisco, CA.

Here’s a link to the source of the data set: <https://catalog.data.gov/dataset?organization=city-of-san-francisco&tags=housing>

1. I am using the names of the neighborhoods/boroughs from this data set to call the Foursquare API’s, to get location info and associated venues around these neighborhoods. My goal is to find a neighborhood that has public transportation (bus/Muni train), hiking trails, close to the beach and has tons of photographic views.

(I am only submitting the codes that meets my criteria of choosing a location to live. Anyone can use these codes, change the name of boroughs/neighborhoods and find a location that they are interested in – for example: your priorities could be coffee shops and tons of restaurants etc.)

Link to GitHub: <https://github.com/shonkhochil/Coursera_Capstone/tree/master/Week%204%20Assignment%20Completed>

**Report:**

**Introduction:**

I have received a job offer in San Francisco, CA. This city is a central hub for lots of start-up companies. A few of my friends from college are already living there. And they have assured me about good public transportation of this city (Muni Train & BART train – can be used to commute to go anywhere in the city).

However, for me personally, I like to stay away from crowded city life – would rather like to live a neighborhood that is peaceful and quiet after work hours. I am also a photographer and like to go on long hike during the weekends.

For this exercise – I would explore the boroughs and neighborhoods in San Francisco city and pick one neighborhood that is close to beach or has hiking trails or photographic scenarios – along with a bus stop/train stop for commuting to work.

Data Source:

I am using two sets of information here:

1. First set of data is been download from San Francisco, CA – city government website in csv format. This data set has all the Boroughs (as Plan District) and Neighborhoods of the city listed along with other information. After downloading this data set, I have imported it into my work using Panda and converted into a data frame, kept the columns I need to get an idea of the names of the boroughs and their associated neighborhoods in/around San Francisco, CA.

Here’s a link to the source of the data set: <https://catalog.data.gov/dataset?organization=city-of-san-francisco&tags=housing>

1. I am using the names of the neighborhoods/boroughs from this data set to call the Foursquare API’s, to get location info and associated venues around these neighborhoods. My goal is to find a neighborhood that has public transportation (bus/Muni train), hiking trails, close to the beach and has tons of photographic views.

(I am only submitting the codes that meets my criteria of choosing a location to live. Anyone can use these codes, change the name of boroughs/neighborhoods and find a location that they are interested in – for example: your priorities could be coffee shops and tons of restaurants etc.)

Methodology:

List of Step taken to analyze this data:

1. Downloading data from San Francisco City Government website in csv format (<https://catalog.data.gov/dataset?organization=city-of-san-francisco&tags=housing>)
2. Importing data into panda data frame, wrangling and exploration to identify boroughs and neighborhoods.
3. Explore one neighborhood at a time with a goal to identify most common venues in that neighborhoods.
4. Use Foursquare API’s to import venues in JSON format.
5. Wrangle JSON format and create a panda data frame
6. Use K-Means clustering to identify similar neighborhoods.
7. Finally pick the neighborhood which has access to public transportation, an ocean view, hiking trails, away from the city and tons of photographic views/opportunities.

Result:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Borough** | **Neighborhood** | **1st Most Common Venue** | **2nd Most Common Venue** | **3rd Most Common Venue** | **4th Most Common Venue** | **5th Most Common Venue** |
| Richmond | Inner Richmond | Grocery Store | Café | Chinese Restaurant | Korean Restaurant | Sushi Restaurant |
| Richmond | Outer Richmond | Café | Chinese Restaurant | Bakery | Japanese Restaurant | Sporting Goods Shop |
| Richmond | Pacific Heights | Furniture / Home Store | Park | Gym | Scenic Lookout | Trail |
| Richmond | Presidio Heights | Cosmetics Shop | American Restaurant | Golf Course | Coffee Shop | Bakery |
| **Richmond** | **Seacliff** | **Scenic Lookout** | **Bus Stop** | **Trail** | **Golf Course** | **Beach** |
| Richmond | Western Addition | Café | Cosmetics Shop | American Restaurant | Coffee Shop | Gym |

Conclusion:

I am selecting Richmond Borough and Seacliff neighborhood as my future home because of Scenic Lookout, Trail, Gold Course, Beach and access to public transportation.

1. A link to your Notebook on your GitHub repository pushed showing your code. (15 marks)

<https://github.com/shonkhochil/Coursera_Capstone/blob/master/Week%204%20Assignment%20Completed/Week4%20Capstone%20Find%20a%20Home%20in%20San%20Francisco%20California%20RH.ipynb>

1. Your choice of a presentation or blogpost. (10 marks)