GLOBAL CONSULTANTS

GUIDING CUSTOMERS TO SELECT BEST CITY

- Global Consultants predicts best place to customer when relocating to new city
- Using location data of Fursquare, company predicts best places near to the city chosen by customer
- Customer satisfaction is very important to company
- Limitation includes the limited amount of data received from Foursquare location data
- Company analyses and provides best option with the available data from Foursquare

BUSINESS PROBLEM

- Predict best places nearby city California
- Customer moves to new city to start a better life
- Customer will be highly satisfied by facilities that provide an option to earn a living and facilities or amenities to live
- Customer will get an idea about nearby venues and opportunities provided by those places.

DATA CLEANING AND FEATURE SELECTION

- The city's latitude and longitude values will be used to extract the venues
- we will filter the columns and create a dataframe from json file
- Remove unnecessary columns from the dataframe created .
- For our project the main feature would be the latitude and longitude values.
- Using latitude and longitude, the venues would be grouped into different clusters

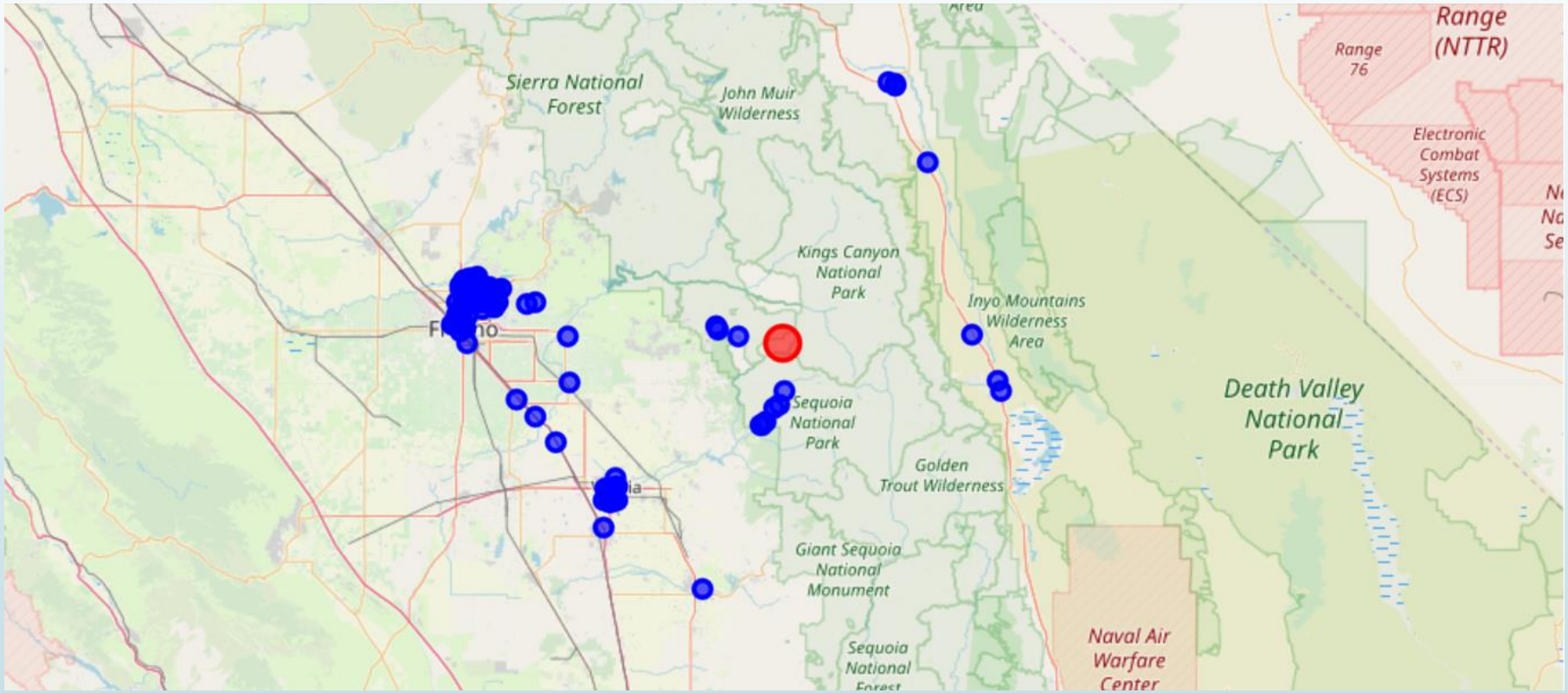
METHODOLOGY

- **Business Understanding**
  - Suggest the best location to a customer who would like to relocate to California City .
- **Analytic Approach**
  - Around 100 venues are identified using Foursquare location data.
  - K Means Clustering Algorithm is used to group the data in the dataframe to different clusters
  - Clusters are analysed to find suitable location for customer.
- **Exploratory Data Analysis**
  - We use geopy and get the latitude and longitude values of the city to explore
  - The data obtained from Foursquare is transformed to pandas dataframe.

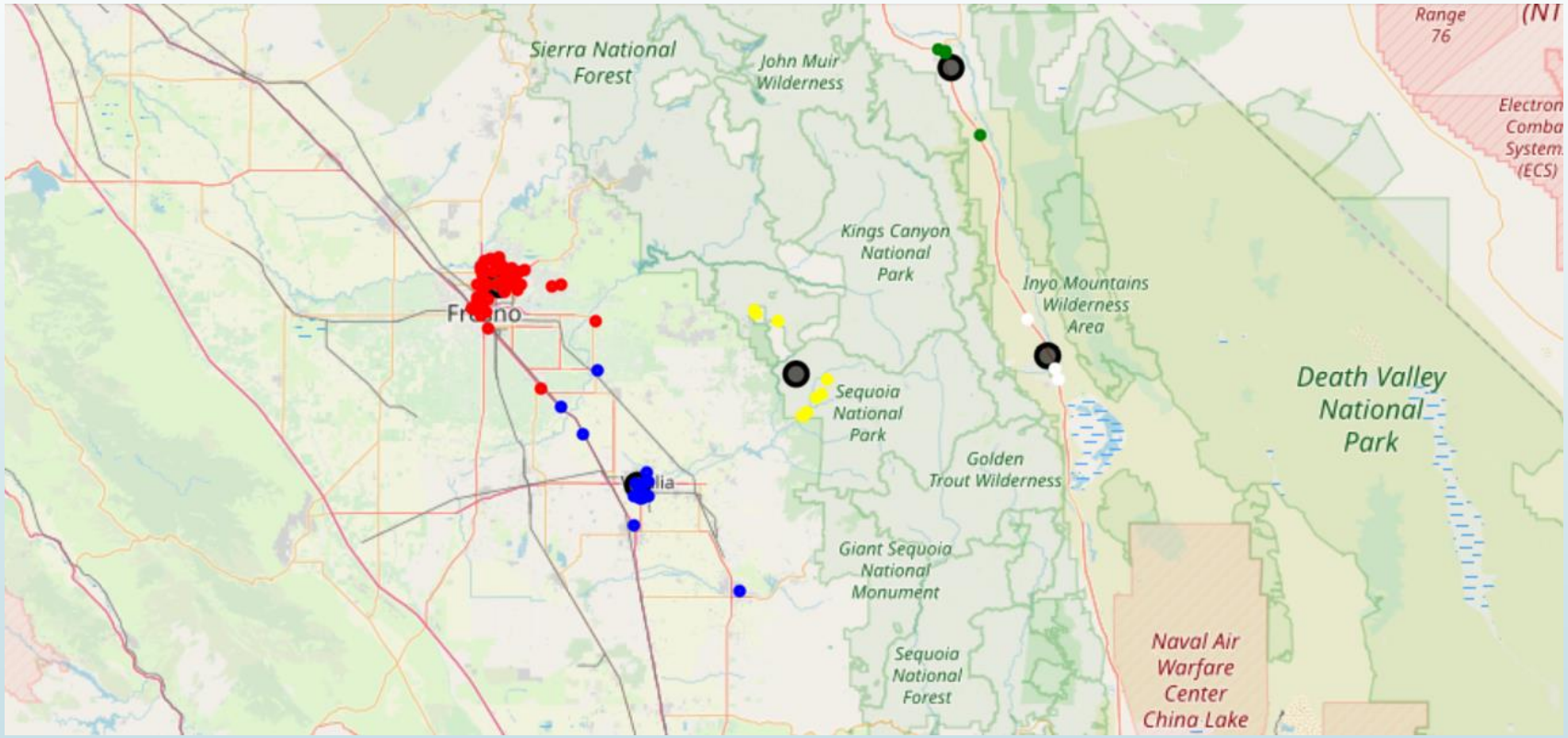
We will use folium libraries to create map of the city with venues

K Means clustering algorithm to group the nearby clusters

FOLIUM MAP OF CALIFORNIA



CLUSTERING CALIFORNIA CITY



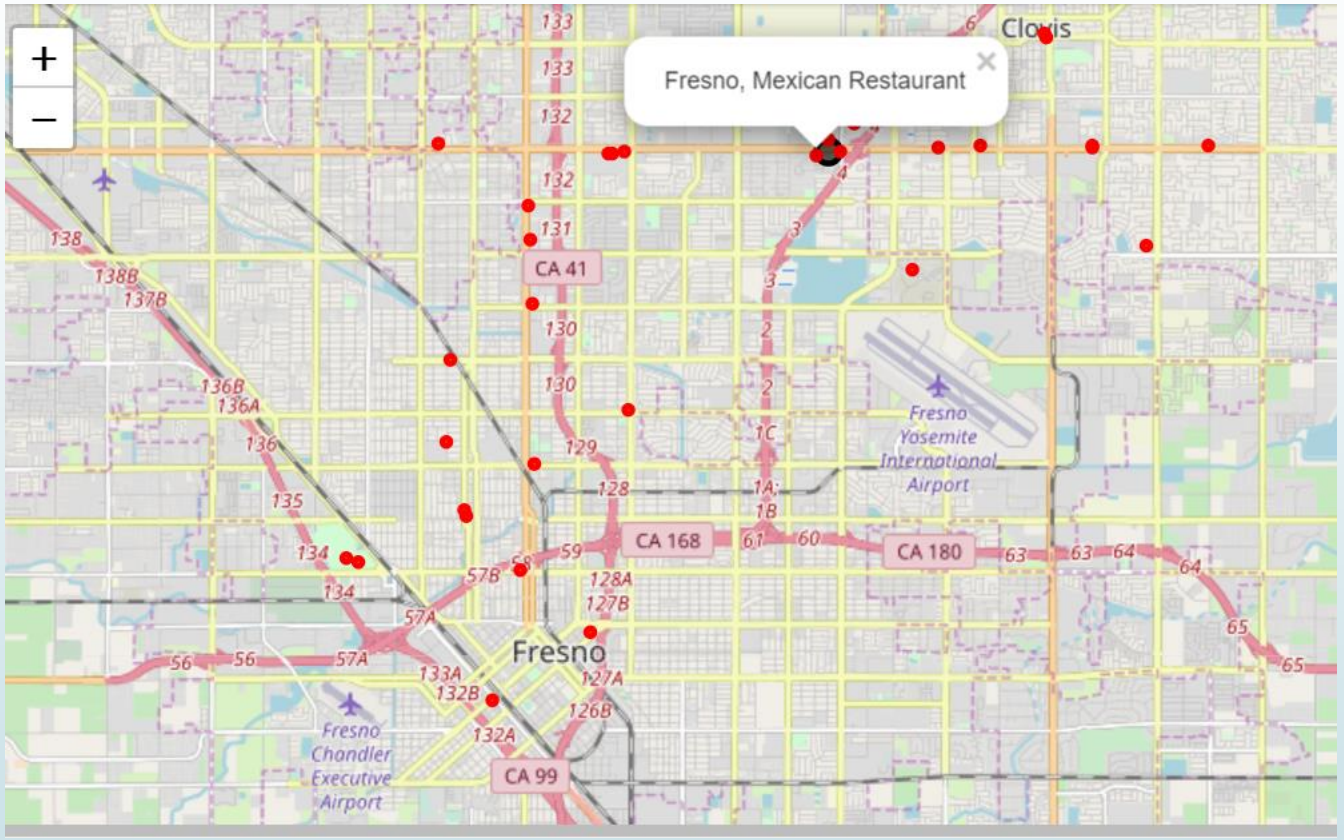
RESULTS

- Best and first option to live is cluster1
- From the map and dataframe it is clear that the maximum number of restaurants, movie theatres, grocery shops, coffee shops, shopping malls etc are available in cluster1.
- Second best option is cluster0
- Ignore cluster 2,3,4 as there is no much facility in venues in these clusters
- Cluster2 can be ignored as there are only three venues and they are bakery, BBQ and Coffee Shop.
- Cluster 3 should be ignored as there is only national park, mountain and scenic
- Cluster 4 can be ignored as there is only historic site and tourist information centre.

CLUSTER 1

```
[29]: dataframe_cluster1.categories.unique()
```

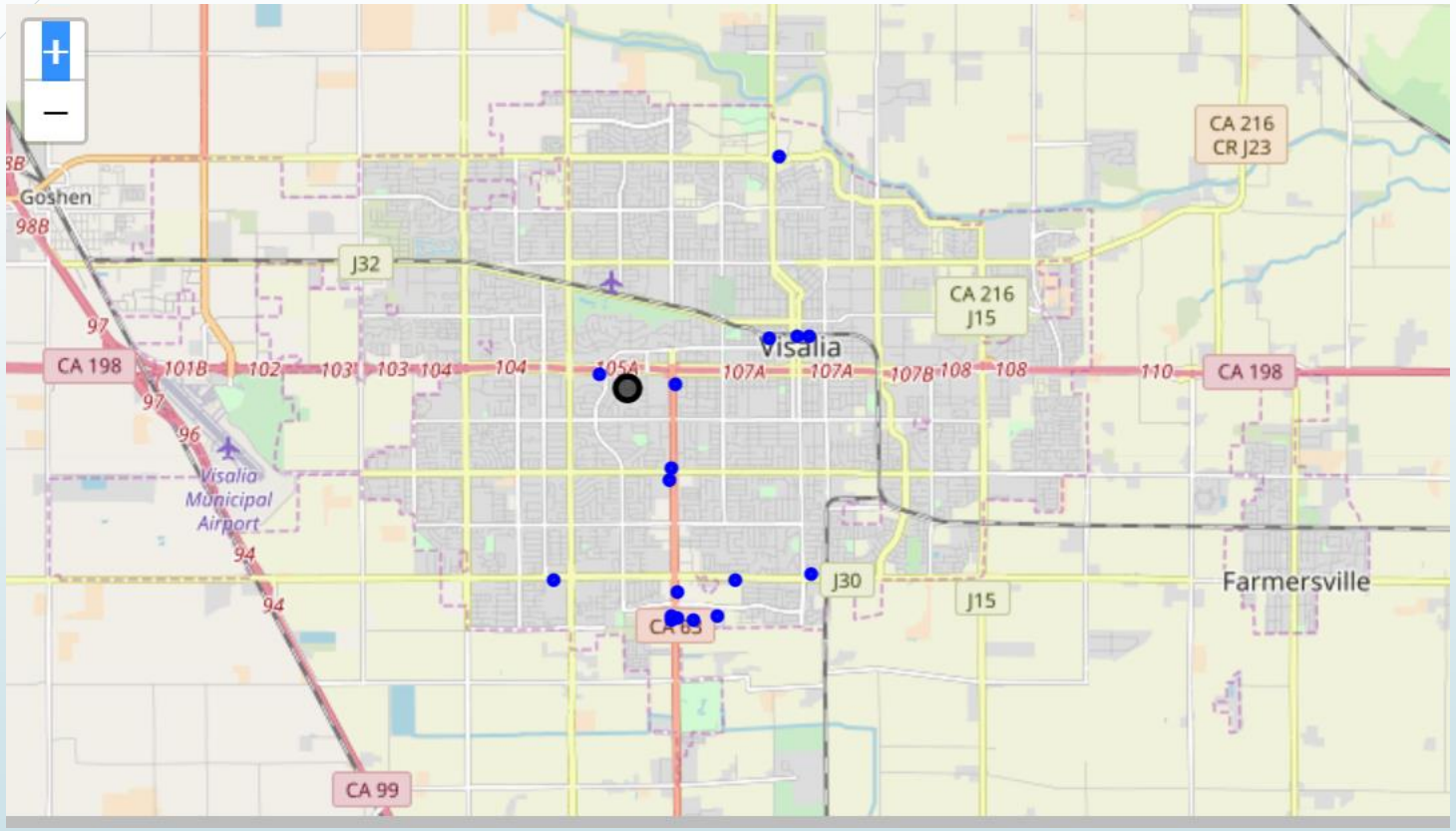
```
[29]: array(['New American Restaurant', 'Mexican Restaurant',
       'Frozen Yogurt Shop', 'American Restaurant', 'Sports Bar',
       'Grocery Store', 'Pizza Place', 'Sushi Restaurant',
       'Breakfast Spot', 'Burger Joint', 'Ice Cream Shop', 'Bakery',
       'Coffee Shop', 'Park', 'Water Park', 'Gym', 'Theater',
       'Steakhouse', 'Market', 'Brewery', 'Multiplex',
       'Mediterranean Restaurant', 'Garden Center', 'Zoo Exhibit', 'Bar',
       'Indian Restaurant', 'Fast Food Restaurant', 'Pet Store',
       'Gastropub', 'Shopping Mall', 'Arts & Crafts Store', 'Lounge',
       'Italian Restaurant', 'Zoo', 'Seafood Restaurant', 'Diner',
       'Sandwich Place', 'College Basketball Court'], dtype=object)
```


Fresno, Mexican Restaurant
Clovis
CA 41
CA 168
CA 180
CA 99
Fresno
Fresno Yosemite International Airport
Fresno Chandler Executive Airport

CLUSTER 0

```
dataframe_cluster0.categories.unique()
```

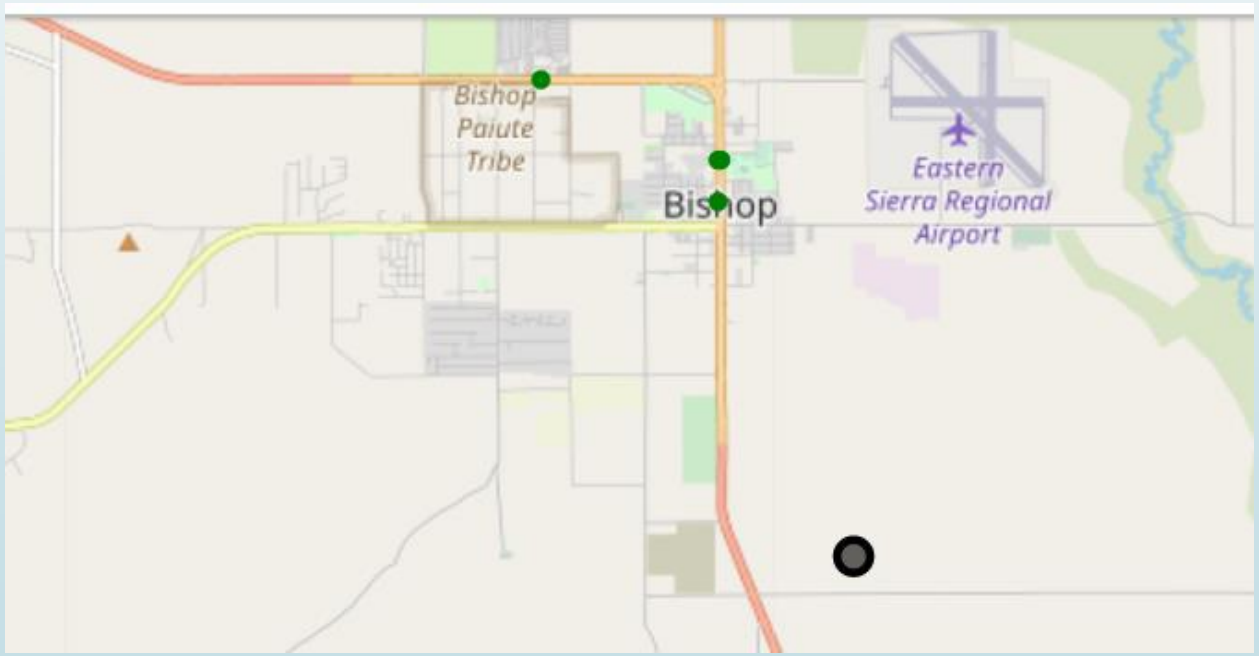
```
array(['Burger Joint', 'Gastropub', 'American Restaurant', 'Liquor Store',
       'Sushi Restaurant', 'Warehouse Store', 'Factory',
       'Japanese Restaurant', 'Donut Shop', 'Coffee Shop', 'Pizza Place',
       'Mexican Restaurant', 'Movie Theater', 'Pet Store',
       'Italian Restaurant'], dtype=object)
```


Goshen
98B
J32
CA 216
CR J23
CA 216
J15
Visalia
CA 198
101B
102
103
103
104
104
105A
107A
107A
107B
108
108
110
CA 198
97
97
96
Visalia
Municipal
Airport
94
94
J30
J15
Farmersville
CA 99

CLUSTER2

```
dataframe_cluster2.categories.unique()
```

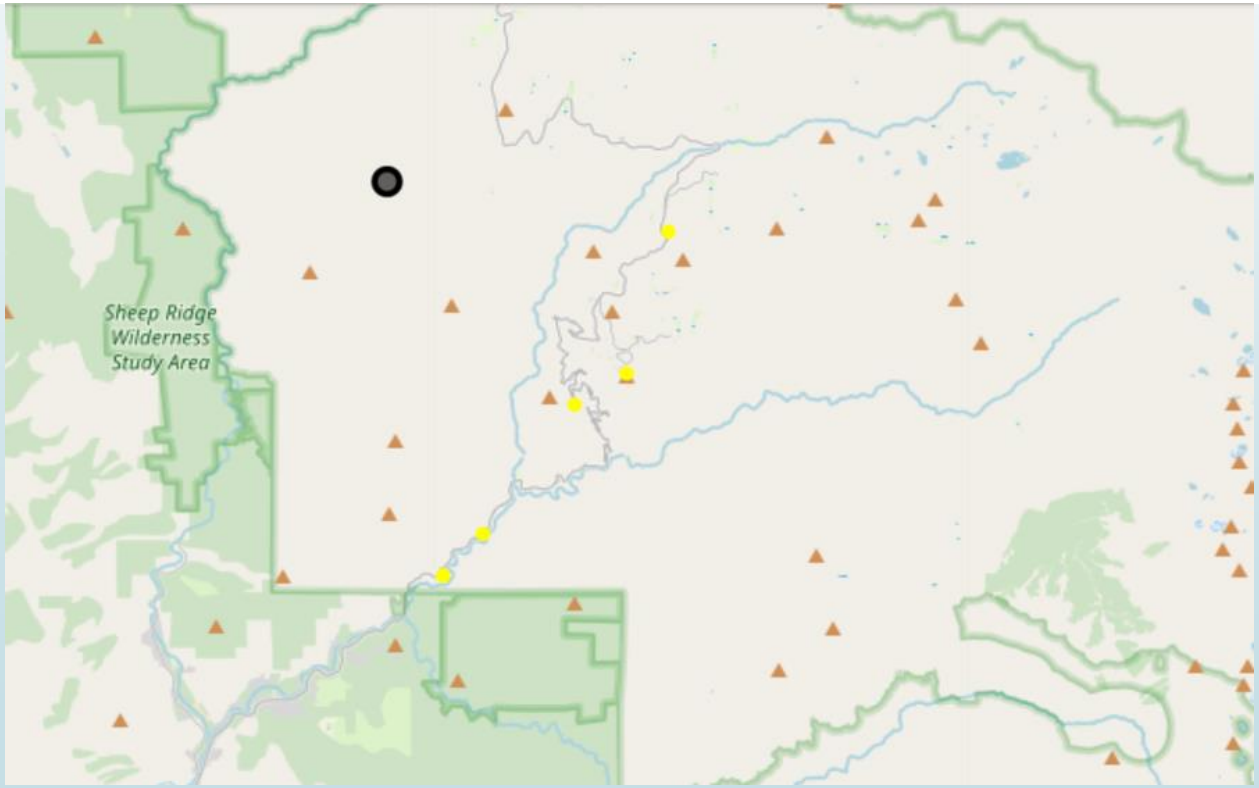
```
array(['BBQ Joint', 'Bakery', 'Deli / Bodega', 'Coffee Shop'],
      dtype=object)
```



CLUSTER 3

```
dataframe_cluster3.categories.unique()
```

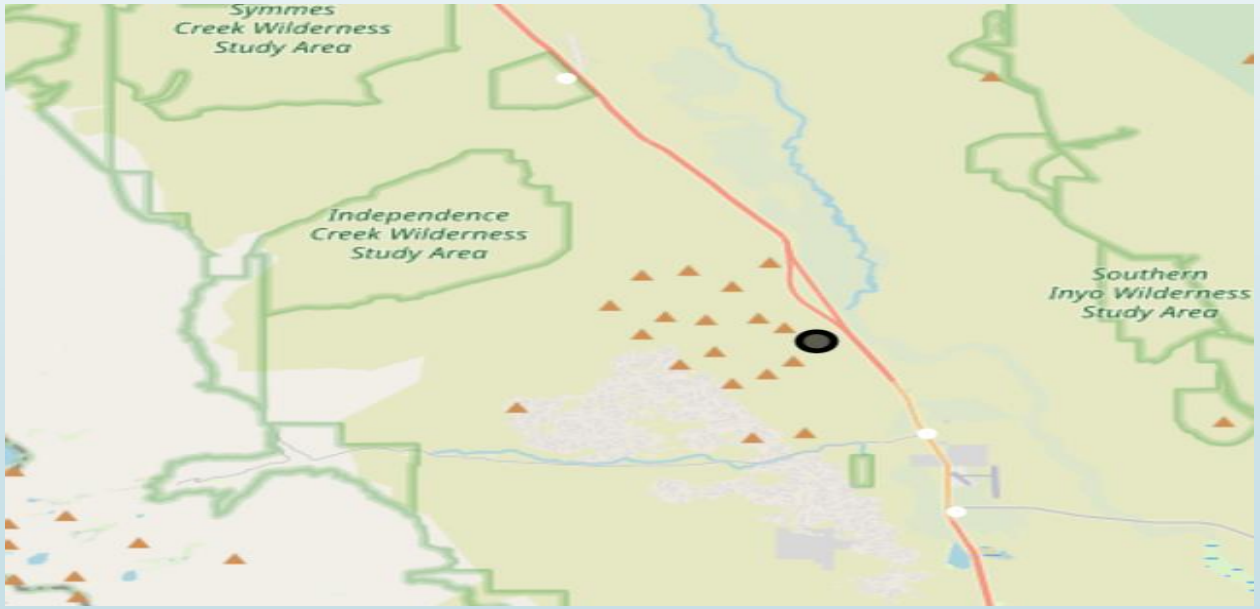
```
array(['Tree', 'National Park', 'Scenic Lookout', 'Mountain',
       'Tourist Information Center'], dtype=object)
```



CLUSTER 4

```
dataframe_cluster4.categories.unique()
```

```
array(['Historic Site', 'American Restaurant',
       'Tourist Information Center'], dtype=object)
```



DISCUSSION

- We have obtained two options to suggest to the customer who plans to move to California City.
- Cluster 1 is the best option where many restaurants, shopping mall, theatres, shops etc are available.
- Cluster0 is the second best option where few number of restaurants and shops, pizza place are available
- Cluster 2 ,3,4 can be ignored as there are not much facilities available.

CONCLUSION

- Cluster 1 is best place.
- Customer can choose Fresno(city) as best place to live and can earn money by doing part time or full time jobs.
- Restaurants, pubs, theatres, coffee shops reveals the cities in cluster 1 is best place for residence as all these facilities are established near residential areas .
- Cluster 0 can be second best option as the cities in cluster 0 has air transport(from map) and many shops and restaurants
- Cluster 2,3,4 can be tourist spots as scenic, mountain, historic places are available.
- This project is performed on limited data available from Foursquare.
- We would be able to provide more idea and suggestions if a good amount of data is available.