Untitled1

June 26, 2020

[1]:
```
!pip install lxml
```

```
Collecting lxml
  Downloading https://files.pythonhosted.org/packages/55/6f/c87dffdd88a54d
d26a3a9fef1d14b6384a9933c455c54ce3ca7d64a84c88/lxml-4.5.1-cp36-cp36m-manylinux1_
x86_64.whl (5.5MB)
     |                  | 5.5MB 6.5MB/s eta
0:00:01             | 4.6MB 6.5MB/s eta 0:00:01
Installing collected packages: lxml
Successfully installed lxml-4.5.1
```

[2]:
```
import pandas as pd
import numpy as np
!pip install beautifulsoup4
```

```
Collecting beautifulsoup4
  Downloading https://files.pythonhosted.org/packages/66/25/ff030e24372656
16a1e9b25ccc864e0371a0bc3adb7c5a404fd661c6f4f6/beautifulsoup4-4.9.1-py3-none-
any.whl (115kB)
     |                  | 122kB 14.4MB/s eta 0:00:01
Collecting soupsieve>1.2 (from beautifulsoup4)
  Downloading https://files.pythonhosted.org/packages/6f/8f/457f4a5390eeae1cc3ae
ab89deb7724c965be841ffca6cfca9197482e470/soupsieve-2.0.1-py3-none-any.whl
Installing collected packages: soupsieve, beautifulsoup4
Successfully installed beautifulsoup4-4.9.1 soupsieve-2.0.1
```

[3]:
```
#Explore San Francisco for a potential city to open a bakery
```

[25]:
```
#download San Francisco cities from Wikipedia
url='https://en.wikipedia.org/wiki/Portal:San_Francisco_Bay_Area/Cities/Table'
df=pd.read_html(url)[1]
```

[26]:
```
df.head()
```

[26]:

	Name	Type	County	Population (2010)	Land area
	Name	Type	County	Population (2010)	Incorporated
0	Alameda	City	Alameda	73812	10.61
1	Albany	City	Alameda	18539	1.79

```
2  American Canyon  City          Napa              19454         4.84
3          Antioch  City  Contra Costa             102372        28.35
4         Atherton  Town     San Mateo               6914         5.02

                            Incorporated
  Unnamed: 5_level_1  Unnamed: 6_level_1
0               27.5      April 19, 1854
1                4.6  September 22, 1908
2               12.5     January 1, 1992
3               73.4    February 6, 1872
4               13.0  September 12, 1923
```

[27]:
```
df.shape
```

[27]:
```
(101, 7)
```

[28]:
```
#Rename columns
df.columns=['Name','Type','County','Population 2010', 'Incorporated','Land␣
↪area','Incorporated Unamed']
```

[29]:
```
df
```

[29]:
```
                Name  Type        County  Population 2010  Incorporated  \
0            Alameda  City       Alameda            73812         10.61
1             Albany  City       Alameda            18539          1.79
2    American Canyon  City          Napa            19454          4.84
3            Antioch  City  Contra Costa           102372         28.35
4           Atherton  Town     San Mateo             6914          5.02
..               …   …           …                …             …
96           Vallejo  City        Solano           115942         30.67
97      Walnut Creek  City  Contra Costa            64173         19.76
98           Windsor  Town        Sonoma            26801          7.27
99          Woodside  Town     San Mateo             5287         11.73
100       Yountville  Town          Napa             2933          1.53

     Land area Incorporated Unamed
0         27.5      April 19, 1854
1          4.6  September 22, 1908
2         12.5     January 1, 1992
3         73.4    February 6, 1872
4         13.0  September 12, 1923
..         …                 …
96        79.4      March 30, 1868
97        51.2    October 21, 1914
98        18.8        July 1, 1992
99        30.4   November 16, 1956
100        4.0    February 4, 1965
```

```
[101 rows x 7 columns]
```

[9]:
```
#clean dataframe
del df['Incorporated']
```

[10]:
```
del df['Land area']
```

[11]:
```
del df['Incorporated Unamed']
```

[12]:
```
df
```

[12]:

	Name	Type	County	Population 2010
0	Alameda	City	Alameda	73812
1	Albany	City	Alameda	18539
2	American Canyon	City	Napa	19454
3	Antioch	City	Contra Costa	102372
4	Atherton	Town	San Mateo	6914
..	…	…	…	…
96	Vallejo	City	Solano	115942
97	Walnut Creek	City	Contra Costa	64173
98	Windsor	Town	Sonoma	26801
99	Woodside	Town	San Mateo	5287
100	Yountville	Town	Napa	2933

[101 rows x 4 columns]

[13]:
```
#look for the most populated city
df.sort_values('Population 2010', ascending=False)
```

[13]:

	Name	Type	County	Population 2010
77	San Jose	City	Santa Clara	945942
76	San Francisco	City and county	San Francisco	805235
56	Oakland	City	Alameda	390724
31	Fremont	City	Alameda	214089
84	Santa Rosa	City	Sonoma	167815
..	…	…	…	…
49	Monte Sereno	City	Santa Clara	3341
100	Yountville	Town	Napa	2933
72	Ross	Town	Marin	2415
6	Belvedere	City	Marin	2068
16	Colma	Town	San Mateo	1792

[101 rows x 4 columns]

[30]:
```
import requests # library to handle requests
import pandas as pd # library for data analsysis
```

```
import numpy as np # library to handle data in a vectorized manner
import random # library for random number generation

!conda install -c conda-forge geopy --yes
from geopy.geocoders import Nominatim # module to convert an address into␣
 ↪latitude and longitude values

# libraries for displaying images
from IPython.display import Image
from IPython.core.display import HTML

# tranforming json file into a pandas dataframe library
from pandas.io.json import json_normalize

!conda install -c conda-forge folium=0.5.0 --yes
import folium # plotting library

print('Folium installed')
print('Libraries imported.')
```

```
Collecting package metadata (current_repodata.json): done
Solving environment: done

# All requested packages already installed.

Collecting package metadata (current_repodata.json): done
Solving environment: done

# All requested packages already installed.

Folium installed
Libraries imported.
```

[31]:
```
!pip install geocoder
!pip install geopandas
!pip install geopy
```

```
Requirement already satisfied: geocoder in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (1.38.1)
Requirement already satisfied: click in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geocoder)
(7.1.2)
Requirement already satisfied: requests in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geocoder)
(2.24.0)
Requirement already satisfied: six in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geocoder)
```

```
(1.15.0)
Requirement already satisfied: ratelim in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geocoder)
(0.1.6)
Requirement already satisfied: future in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geocoder)
(0.18.2)
Requirement already satisfied: chardet<4,>=3.0.2 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
requests->geocoder) (3.0.4)
Requirement already satisfied: certifi>=2017.4.17 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
requests->geocoder) (2020.6.20)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
requests->geocoder) (1.25.9)
Requirement already satisfied: idna<3,>=2.5 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
requests->geocoder) (2.9)
Requirement already satisfied: decorator in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
ratelim->geocoder) (4.4.2)
Requirement already satisfied: geopandas in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (0.8.0)
Requirement already satisfied: shapely in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geopandas)
(1.7.0)
Requirement already satisfied: fiona in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geopandas)
(1.8.13.post1)
Requirement already satisfied: pyproj>=2.2.0 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geopandas)
(2.6.1.post1)
Requirement already satisfied: pandas>=0.23.0 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geopandas)
(1.0.5)
Requirement already satisfied: six>=1.7 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
fiona->geopandas) (1.15.0)
Requirement already satisfied: cligj>=0.5 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
fiona->geopandas) (0.5.0)
Requirement already satisfied: click-plugins>=1.0 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
fiona->geopandas) (1.1.1)
Requirement already satisfied: munch in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
fiona->geopandas) (2.5.0)
```

```
Requirement already satisfied: click<8,>=4.0 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
fiona->geopandas) (7.1.2)
Requirement already satisfied: attrs>=17 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
fiona->geopandas) (19.3.0)
Requirement already satisfied: pytz>=2017.2 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
pandas>=0.23.0->geopandas) (2020.1)
Requirement already satisfied: python-dateutil>=2.6.1 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
pandas>=0.23.0->geopandas) (2.8.1)
Requirement already satisfied: numpy>=1.13.3 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from
pandas>=0.23.0->geopandas) (1.18.5)
Requirement already satisfied: geopy in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (1.22.0)
Requirement already satisfied: geographiclib<2,>=1.49 in
/home/jupyterlab/conda/envs/python/lib/python3.6/site-packages (from geopy)
(1.50)
```

[ ]:

[32]:
```
#Let's explore the most populated city in San Francisco
address = 'San Jose, CA'

geolocator = Nominatim(user_agent="ny_explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of San Jose are {}, {}.'.format(latitude,␣
 ↪longitude))
```

```
The geograpical coordinate of San Jose are 37.3361905, -121.8905833.
```

[33]:
```
#use geocoder to get location latitude and longitude points
from geopy.extra.rate_limiter import RateLimiter

# 1 - conveneint function to delay between geocoding calls
geocode = RateLimiter(geolocator.geocode, min_delay_seconds=1)
# 2- - create location column
df['location'] = df['County'].apply(geocode)
# 3 - create longitude, laatitude and altitude from location column (returns␣
 ↪tuple)
df['point'] = df['location'].apply(lambda loc: tuple(loc.point) if loc else␣
 ↪None)
# 4 - split point column into latitude, longitude and altitude columns
```

```
df[['latitude', 'longitude', 'altitude']] = pd.DataFrame(df['point'].tolist(),␣
↪index=df.index)
df
```

[33]:

```
                Name  Type        County  Population 2010  Incorporated  \
0            Alameda  City       Alameda            73812         10.61
1             Albany  City       Alameda            18539          1.79
2    American Canyon  City          Napa            19454          4.84
3            Antioch  City  Contra Costa           102372         28.35
4           Atherton  Town     San Mateo             6914          5.02
..               …   …           …                …            …
96           Vallejo  City        Solano           115942         30.67
97      Walnut Creek  City  Contra Costa            64173         19.76
98           Windsor  Town        Sonoma            26801          7.27
99          Woodside  Town     San Mateo             5287         11.73
100       Yountville  Town          Napa             2933          1.53

     Land area Incorporated Unamed  \
0         27.5      April 19, 1854
1          4.6  September 22, 1908
2         12.5     January 1, 1992
3         73.4    February 6, 1872
4         13.0  September 12, 1923
..         …                 …
96        79.4      March 30, 1868
97        51.2    October 21, 1914
98        18.8        July 1, 1992
99        30.4   November 16, 1956
100        4.0    February 4, 1965

                                                location  \
0    (Alameda County, California, United States of …
1    (Alameda County, California, United States of …
2    (Napa, Napa County, California, 94559, United …
3    (Contra Costa County, California, United State…
4    (San Mateo County, California, United States o…
..                                                 …
96   (Solano County, California, United States of A…
97   (Contra Costa County, California, United State…
98   (Sonoma County, California, United States of A…
99   (San Mateo County, California, United States o…
100  (Napa, Napa County, California, 94559, United …

                              point   latitude   longitude  altitude
0     (37.6090291, -121.899142, 0.0)  37.609029 -121.899142       0.0
1     (37.6090291, -121.899142, 0.0)  37.609029 -121.899142       0.0
2    (38.2971367, -122.2855293, 0.0)  38.297137 -122.285529       0.0
```

```
3    (37.9034806, -121.9175345, 0.0)  37.903481 -121.917535       0.0
4     (37.496904, -122.3330573, 0.0)  37.496904 -122.333057       0.0
..                                ...       ...         ...       ...
96   (38.2218938, -121.9163555, 0.0)  38.221894 -121.916355       0.0
97   (37.9034806, -121.9175345, 0.0)  37.903481 -121.917535       0.0
98   (38.5110803, -122.8473388, 0.0)  38.511080 -122.847339       0.0
99    (37.496904, -122.3330573, 0.0)  37.496904 -122.333057       0.0
100  (38.2971367, -122.2855293, 0.0)  38.297137 -122.285529       0.0

[101 rows x 12 columns]
```

[18]:
```
#Let's see where the counties are in relation to San Jose
SF_map = folium.Map(location=[latitude, longitude], zoom_start=8) # generate␣
 ↪map centred around the Conrad Hotel

# add a red circle marker to represent
folium.features.CircleMarker(
    [37.3361905, -121.8905833],
    radius=7,
    color='red',
    popup='San Jose',
    fill = True,
    fill_color = 'red',
    fill_opacity = 0.6
).add_to(SF_map)
# add popular spots to the map as blue circle markers
for lat, lng, label in zip(df['latitude'], df['longitude'], df['County']):
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='green',
        fill=True,
        fill_color='green',
        fill_opacity=0.7,
        parse_html=False).add_to(SF_map)

# display map
SF_map
```

[18]:
```
<folium.folium.Map at 0x7f1332cc68d0>
```

[19]:
```
#from CI Gateway Zip Code List look up cities in Santa Clara
```

[38]:
```
url='http://www.ciclt.net/sn/clt/capitolimpact/gw_ziplist.aspx?
 ↪ClientCode=capitolimpact&State=ca&StName=California&StFIPS=&FIPS=06085'
```

```
df1=pd.read_html(url)[2]
df1
```

[38]:

	Zip Code	City	County
0	94022	Los Altos	Santa Clara County
1	94022	Los Altos Hills	Santa Clara County
2	94023	Los Altos	Santa Clara County
3	94024	Los Altos	Santa Clara County
4	94024	Los Altos Hills	Santa Clara County
5	94035	Moffett Field	Santa Clara County
6	94035	Mountain View	Santa Clara County
7	94039	Mountain View	Santa Clara County
8	94040	Mountain View	Santa Clara County
9	94041	Mountain View	Santa Clara County
10	94042	Mountain View	Santa Clara County
11	94043	Mountain View	Santa Clara County
12	94085	Sunnyvale	Santa Clara County
13	94086	Sunnyvale	Santa Clara County
14	94087	Sunnyvale	Santa Clara County
15	94088	Onizuka Afb	Santa Clara County
16	94088	Sunnyvale	Santa Clara County
17	94089	Sunnyvale	Santa Clara County
18	94090	Sunnyvale	Santa Clara County
19	94301	Palo Alto	Santa Clara County
20	94302	Palo Alto	Santa Clara County
21	94303	East Palo Alto	Santa Clara County
22	94303	Palo Alto	Santa Clara County
23	94304	Palo Alto	Santa Clara County
24	94305	Stanford	Santa Clara County
25	94305	Palo Alto	Santa Clara County
26	94306	Palo Alto	Santa Clara County
27	94309	Stanford	Santa Clara County
28	94309	Palo Alto	Santa Clara County
29	94310	Palo Alto	Santa Clara County
30	95002	Alviso	Santa Clara County
31	95008	Campbell	Santa Clara County
32	95009	Campbell	Santa Clara County
33	95011	Campbell	Santa Clara County
34	95013	Coyote	Santa Clara County
35	95014	Monte Vista	Santa Clara County
36	95014	Permanente	Santa Clara County
37	95014	Cupertino	Santa Clara County
38	95015	Cupertino	Santa Clara County
39	95020	Gilroy	Santa Clara County
40	95021	Gilroy	Santa Clara County
41	95026	Holy City	Santa Clara County
42	95030	Los Gatos	Santa Clara County

```
43     95030     Monte Sereno  Santa Clara County
44     95031        Los Gatos  Santa Clara County
45     95032        Los Gatos  Santa Clara County
46     95035         Milpitas  Santa Clara County
47     95036         Milpitas  Santa Clara County
48     95037      Morgan Hill  Santa Clara County
49     95038      Morgan Hill  Santa Clara County
50     95042      New Almaden  Santa Clara County
51     95044  Redwood Estates  Santa Clara County
```

[39]:
```
#Clean up San Jose dataframe
df1.columns=['ZIP code','City','County']
```

[41]:
```
df1['State']='California'
```

[42]:
```
df1['Address']=df1['City']+','+df1['State']
```

[43]:
```
df1
```

[43]:
```
    ZIP code             City              County       State  \
0      94022        Los Altos  Santa Clara County  California
1      94022  Los Altos Hills  Santa Clara County  California
2      94023        Los Altos  Santa Clara County  California
3      94024        Los Altos  Santa Clara County  California
4      94024  Los Altos Hills  Santa Clara County  California
5      94035    Moffett Field  Santa Clara County  California
6      94035    Mountain View  Santa Clara County  California
7      94039    Mountain View  Santa Clara County  California
8      94040    Mountain View  Santa Clara County  California
9      94041    Mountain View  Santa Clara County  California
10     94042    Mountain View  Santa Clara County  California
11     94043    Mountain View  Santa Clara County  California
12     94085        Sunnyvale  Santa Clara County  California
13     94086        Sunnyvale  Santa Clara County  California
14     94087        Sunnyvale  Santa Clara County  California
15     94088      Onizuka Afb  Santa Clara County  California
16     94088        Sunnyvale  Santa Clara County  California
17     94089        Sunnyvale  Santa Clara County  California
18     94090        Sunnyvale  Santa Clara County  California
19     94301        Palo Alto  Santa Clara County  California
20     94302        Palo Alto  Santa Clara County  California
21     94303   East Palo Alto  Santa Clara County  California
22     94303        Palo Alto  Santa Clara County  California
23     94304        Palo Alto  Santa Clara County  California
24     94305         Stanford  Santa Clara County  California
25     94305        Palo Alto  Santa Clara County  California
26     94306        Palo Alto  Santa Clara County  California
```

27	94309	Stanford	Santa Clara County	California
28	94309	Palo Alto	Santa Clara County	California
29	94310	Palo Alto	Santa Clara County	California
30	95002	Alviso	Santa Clara County	California
31	95008	Campbell	Santa Clara County	California
32	95009	Campbell	Santa Clara County	California
33	95011	Campbell	Santa Clara County	California
34	95013	Coyote	Santa Clara County	California
35	95014	Monte Vista	Santa Clara County	California
36	95014	Permanente	Santa Clara County	California
37	95014	Cupertino	Santa Clara County	California
38	95015	Cupertino	Santa Clara County	California
39	95020	Gilroy	Santa Clara County	California
40	95021	Gilroy	Santa Clara County	California
41	95026	Holy City	Santa Clara County	California
42	95030	Los Gatos	Santa Clara County	California
43	95030	Monte Sereno	Santa Clara County	California
44	95031	Los Gatos	Santa Clara County	California
45	95032	Los Gatos	Santa Clara County	California
46	95035	Milpitas	Santa Clara County	California
47	95036	Milpitas	Santa Clara County	California
48	95037	Morgan Hill	Santa Clara County	California
49	95038	Morgan Hill	Santa Clara County	California
50	95042	New Almaden	Santa Clara County	California
51	95044	Redwood Estates	Santa Clara County	California

	Address
0	Los Altos,California
1	Los Altos Hills,California
2	Los Altos,California
3	Los Altos,California
4	Los Altos Hills,California
5	Moffett Field,California
6	Mountain View,California
7	Mountain View,California
8	Mountain View,California
9	Mountain View,California
10	Mountain View,California
11	Mountain View,California
12	Sunnyvale,California
13	Sunnyvale,California
14	Sunnyvale,California
15	Onizuka Afb,California
16	Sunnyvale,California
17	Sunnyvale,California
18	Sunnyvale,California
19	Palo Alto,California

```
20         Palo Alto,California
21    East Palo Alto,California
22         Palo Alto,California
23         Palo Alto,California
24          Stanford,California
25         Palo Alto,California
26         Palo Alto,California
27          Stanford,California
28         Palo Alto,California
29         Palo Alto,California
30            Alviso,California
31          Campbell,California
32          Campbell,California
33          Campbell,California
34            Coyote,California
35       Monte Vista,California
36        Permanente,California
37          Cupertino,California
38          Cupertino,California
39             Gilroy,California
40             Gilroy,California
41          Holy City,California
42          Los Gatos,California
43      Monte Sereno,California
44          Los Gatos,California
45          Los Gatos,California
46           Milpitas,California
47           Milpitas,California
48       Morgan Hill,California
49       Morgan Hill,California
50       New Almaden,California
51  Redwood Estates,California
```

[44]:
```
#use geocoder to get location latitude and longitude points for Santa Clara␣
 ↪neighborhoods
from geopy.extra.rate_limiter import RateLimiter

# 1 - conveneint function to delay between geocoding calls
geocode = RateLimiter(geolocator.geocode, min_delay_seconds=1)
# 2- - create location column
df1['location'] = df1['Address'].apply(geocode)
# 3 - create longitude, latitude and altitude from location column (returns␣
 ↪tuple)
df1['point'] = df1['location'].apply(lambda loc: tuple(loc.point) if loc else␣
 ↪None)
# 4 - split point column into latitude, longitude and altitude columns
```

```
df1[['latitude', 'longitude','altitude']] = pd.DataFrame(df1['point'].tolist(),␣
↪index=df1.index)
df1
```

[44]:

	ZIP code	City	County	State
0	94022	Los Altos	Santa Clara County	California
1	94022	Los Altos Hills	Santa Clara County	California
2	94023	Los Altos	Santa Clara County	California
3	94024	Los Altos	Santa Clara County	California
4	94024	Los Altos Hills	Santa Clara County	California
5	94035	Moffett Field	Santa Clara County	California
6	94035	Mountain View	Santa Clara County	California
7	94039	Mountain View	Santa Clara County	California
8	94040	Mountain View	Santa Clara County	California
9	94041	Mountain View	Santa Clara County	California
10	94042	Mountain View	Santa Clara County	California
11	94043	Mountain View	Santa Clara County	California
12	94085	Sunnyvale	Santa Clara County	California
13	94086	Sunnyvale	Santa Clara County	California
14	94087	Sunnyvale	Santa Clara County	California
15	94088	Onizuka Afb	Santa Clara County	California
16	94088	Sunnyvale	Santa Clara County	California
17	94089	Sunnyvale	Santa Clara County	California
18	94090	Sunnyvale	Santa Clara County	California
19	94301	Palo Alto	Santa Clara County	California
20	94302	Palo Alto	Santa Clara County	California
21	94303	East Palo Alto	Santa Clara County	California
22	94303	Palo Alto	Santa Clara County	California
23	94304	Palo Alto	Santa Clara County	California
24	94305	Stanford	Santa Clara County	California
25	94305	Palo Alto	Santa Clara County	California
26	94306	Palo Alto	Santa Clara County	California
27	94309	Stanford	Santa Clara County	California
28	94309	Palo Alto	Santa Clara County	California
29	94310	Palo Alto	Santa Clara County	California
30	95002	Alviso	Santa Clara County	California
31	95008	Campbell	Santa Clara County	California
32	95009	Campbell	Santa Clara County	California
33	95011	Campbell	Santa Clara County	California
34	95013	Coyote	Santa Clara County	California
35	95014	Monte Vista	Santa Clara County	California
36	95014	Permanente	Santa Clara County	California
37	95014	Cupertino	Santa Clara County	California
38	95015	Cupertino	Santa Clara County	California
39	95020	Gilroy	Santa Clara County	California
40	95021	Gilroy	Santa Clara County	California
41	95026	Holy City	Santa Clara County	California

```
42     95030        Los Gatos  Santa Clara County  California
43     95030     Monte Sereno  Santa Clara County  California
44     95031        Los Gatos  Santa Clara County  California
45     95032        Los Gatos  Santa Clara County  California
46     95035         Milpitas  Santa Clara County  California
47     95036         Milpitas  Santa Clara County  California
48     95037      Morgan Hill  Santa Clara County  California
49     95038      Morgan Hill  Santa Clara County  California
50     95042      New Almaden  Santa Clara County  California
51     95044  Redwood Estates  Santa Clara County  California

                      Address  \
0         Los Altos,California
1   Los Altos Hills,California
2         Los Altos,California
3         Los Altos,California
4   Los Altos Hills,California
5     Moffett Field,California
6     Mountain View,California
7     Mountain View,California
8     Mountain View,California
9     Mountain View,California
10    Mountain View,California
11    Mountain View,California
12        Sunnyvale,California
13        Sunnyvale,California
14        Sunnyvale,California
15      Onizuka Afb,California
16        Sunnyvale,California
17        Sunnyvale,California
18        Sunnyvale,California
19        Palo Alto,California
20        Palo Alto,California
21   East Palo Alto,California
22        Palo Alto,California
23        Palo Alto,California
24         Stanford,California
25        Palo Alto,California
26        Palo Alto,California
27         Stanford,California
28        Palo Alto,California
29        Palo Alto,California
30           Alviso,California
31         Campbell,California
32         Campbell,California
33         Campbell,California
34           Coyote,California
```

```
35      Monte Vista,California
36       Permanente,California
37        Cupertino,California
38        Cupertino,California
39           Gilroy,California
40           Gilroy,California
41        Holy City,California
42        Los Gatos,California
43     Monte Sereno,California
44        Los Gatos,California
45        Los Gatos,California
46         Milpitas,California
47         Milpitas,California
48      Morgan Hill,California
49      Morgan Hill,California
50      New Almaden,California
51  Redwood Estates,California
```

```
                                                location  \
0   (Los Altos, Santa Clara County, California, Un…
1   (Los Altos Hills, Santa Clara County, Californ…
2   (Los Altos, Santa Clara County, California, Un…
3   (Los Altos, Santa Clara County, California, Un…
4   (Los Altos Hills, Santa Clara County, Californ…
5   (Building 126, Cummins Avenue, Ames Research C…
6   (Mountain View, Santa Clara County, California…
7   (Mountain View, Santa Clara County, California…
8   (Mountain View, Santa Clara County, California…
9   (Mountain View, Santa Clara County, California…
10  (Mountain View, Santa Clara County, California…
11  (Mountain View, Santa Clara County, California…
12  (Sunnyvale, Santa Clara County, California, Un…
13  (Sunnyvale, Santa Clara County, California, Un…
14  (Sunnyvale, Santa Clara County, California, Un…
15                                                 None
16  (Sunnyvale, Santa Clara County, California, Un…
17  (Sunnyvale, Santa Clara County, California, Un…
18  (Sunnyvale, Santa Clara County, California, Un…
19  (Palo Alto, Santa Clara County, California, Un…
20  (Palo Alto, Santa Clara County, California, Un…
21  (East Palo Alto, San Mateo County, California,…
22  (Palo Alto, Santa Clara County, California, Un…
23  (Palo Alto, Santa Clara County, California, Un…
24  (Stanford, Santa Clara County, California, Uni…
25  (Palo Alto, Santa Clara County, California, Un…
26  (Palo Alto, Santa Clara County, California, Un…
27  (Stanford, Santa Clara County, California, Uni…
```

```
28  (Palo Alto, Santa Clara County, California, Un…
29  (Palo Alto, Santa Clara County, California, Un…
30  (Alviso, San Jose, Santa Clara County, Califor…
31  (Campbell, Santa Clara County, California, 950…
32  (Campbell, Santa Clara County, California, 950…
33  (Campbell, Santa Clara County, California, 950…
34  (Coyote, Santa Clara County, California, 95013…
35  (Monte Vista, Alta, Placer County, California,…
36  (Permanente, Stephen E. Abbors Trail, Loyola, …
37  (Cupertino, Santa Clara County, California, Un…
38  (Cupertino, Santa Clara County, California, Un…
39  (Gilroy, Santa Clara County, California, Unite…
40  (Gilroy, Santa Clara County, California, Unite…
41  (Holy City, Santa Clara County, California, 95…
42  (Los Gatos, Santa Clara County, California, Un…
43  (Monte Sereno, Santa Clara County, California,…
44  (Los Gatos, Santa Clara County, California, Un…
45  (Los Gatos, Santa Clara County, California, Un…
46  (Milpitas, Santa Clara County, California, Uni…
47  (Milpitas, Santa Clara County, California, Uni…
48  (Morgan Hill, Santa Clara County, California, …
49  (Morgan Hill, Santa Clara County, California, …
50  (New Almaden, Santa Clara County, California, …
51  (Redwood Estates, Santa Clara County, Californ…
```

	point	latitude	longitude	altitude
0	(37.3790629, -122.116578, 0.0)	37.379063	-122.116578	0.0
1	(37.3796627, -122.1374637, 0.0)	37.379663	-122.137464	0.0
2	(37.3790629, -122.116578, 0.0)	37.379063	-122.116578	0.0
3	(37.3790629, -122.116578, 0.0)	37.379063	-122.116578	0.0
4	(37.3796627, -122.1374637, 0.0)	37.379663	-122.137464	0.0
5	(37.41129325, -122.05416980310586, 0.0)	37.411293	-122.054170	0.0
6	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0
7	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0
8	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0
9	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0
10	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0
11	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0
12	(37.3688301, -122.0363496, 0.0)	37.368830	-122.036350	0.0
13	(37.3688301, -122.0363496, 0.0)	37.368830	-122.036350	0.0
14	(37.3688301, -122.0363496, 0.0)	37.368830	-122.036350	0.0
15	None	NaN	NaN	NaN
16	(37.3688301, -122.0363496, 0.0)	37.368830	-122.036350	0.0
17	(37.3688301, -122.0363496, 0.0)	37.368830	-122.036350	0.0
18	(37.3688301, -122.0363496, 0.0)	37.368830	-122.036350	0.0
19	(37.4443293, -122.1598465, 0.0)	37.444329	-122.159847	0.0
20	(37.4443293, -122.1598465, 0.0)	37.444329	-122.159847	0.0

```
21          (37.4688273, -122.1410751, 0.0)  37.468827 -122.141075       0.0
22          (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
23          (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
24           (37.427467, -122.1702445, 0.0)  37.427467 -122.170244       0.0
25          (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
26          (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
27           (37.427467, -122.1702445, 0.0)  37.427467 -122.170244       0.0
28          (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
29          (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
30           (37.426051, -121.9752373, 0.0)  37.426051 -121.975237       0.0
31          (37.2870626, -121.9448818, 0.0)  37.287063 -121.944882       0.0
32          (37.2870626, -121.9448818, 0.0)  37.287063 -121.944882       0.0
33          (37.2870626, -121.9448818, 0.0)  37.287063 -121.944882       0.0
34          (37.2164923, -121.7394174, 0.0)  37.216492 -121.739417       0.0
35          (39.1865648, -120.8327162, 0.0)  39.186565 -120.832716       0.0
36                (37.3189, -122.1111, 0.0)  37.318900 -122.111100       0.0
37          (37.3228934, -122.0322895, 0.0)  37.322893 -122.032290       0.0
38          (37.3228934, -122.0322895, 0.0)  37.322893 -122.032290       0.0
39          (37.0065078, -121.5631723, 0.0)  37.006508 -121.563172       0.0
40          (37.0065078, -121.5631723, 0.0)  37.006508 -121.563172       0.0
41          (37.1569939, -121.9788351, 0.0)  37.156994 -121.978835       0.0
42           (37.226611, -121.9746797, 0.0)  37.226611 -121.974680       0.0
43            (37.236333, -121.992458, 0.0)  37.236333 -121.992458       0.0
44           (37.226611, -121.9746797, 0.0)  37.226611 -121.974680       0.0
45           (37.226611, -121.9746797, 0.0)  37.226611 -121.974680       0.0
46          (37.4282724, -121.9066238, 0.0)  37.428272 -121.906624       0.0
47          (37.4282724, -121.9066238, 0.0)  37.428272 -121.906624       0.0
48           (37.130408, -121.6544974, 0.0)  37.130408 -121.654497       0.0
49           (37.130408, -121.6544974, 0.0)  37.130408 -121.654497       0.0
50          (37.1760567, -121.8207855, 0.0)  37.176057 -121.820785       0.0
51          (37.1563361, -121.9866223, 0.0)  37.156336 -121.986622       0.0
```

[45]:
```
#clean data drop Nan
df1=df1.drop([15])
df1
```

[45]:

	ZIP code	City	County	State
0	94022	Los Altos	Santa Clara County	California
1	94022	Los Altos Hills	Santa Clara County	California
2	94023	Los Altos	Santa Clara County	California
3	94024	Los Altos	Santa Clara County	California
4	94024	Los Altos Hills	Santa Clara County	California
5	94035	Moffett Field	Santa Clara County	California
6	94035	Mountain View	Santa Clara County	California
7	94039	Mountain View	Santa Clara County	California
8	94040	Mountain View	Santa Clara County	California
9	94041	Mountain View	Santa Clara County	California

```
10     94042    Mountain View  Santa Clara County  California
11     94043    Mountain View  Santa Clara County  California
12     94085        Sunnyvale  Santa Clara County  California
13     94086        Sunnyvale  Santa Clara County  California
14     94087        Sunnyvale  Santa Clara County  California
16     94088        Sunnyvale  Santa Clara County  California
17     94089        Sunnyvale  Santa Clara County  California
18     94090        Sunnyvale  Santa Clara County  California
19     94301        Palo Alto  Santa Clara County  California
20     94302        Palo Alto  Santa Clara County  California
21     94303   East Palo Alto  Santa Clara County  California
22     94303        Palo Alto  Santa Clara County  California
23     94304        Palo Alto  Santa Clara County  California
24     94305         Stanford  Santa Clara County  California
25     94305        Palo Alto  Santa Clara County  California
26     94306        Palo Alto  Santa Clara County  California
27     94309         Stanford  Santa Clara County  California
28     94309        Palo Alto  Santa Clara County  California
29     94310        Palo Alto  Santa Clara County  California
30     95002           Alviso  Santa Clara County  California
31     95008         Campbell  Santa Clara County  California
32     95009         Campbell  Santa Clara County  California
33     95011         Campbell  Santa Clara County  California
34     95013           Coyote  Santa Clara County  California
35     95014      Monte Vista  Santa Clara County  California
36     95014       Permanente  Santa Clara County  California
37     95014        Cupertino  Santa Clara County  California
38     95015        Cupertino  Santa Clara County  California
39     95020           Gilroy  Santa Clara County  California
40     95021           Gilroy  Santa Clara County  California
41     95026        Holy City  Santa Clara County  California
42     95030        Los Gatos  Santa Clara County  California
43     95030     Monte Sereno  Santa Clara County  California
44     95031        Los Gatos  Santa Clara County  California
45     95032        Los Gatos  Santa Clara County  California
46     95035         Milpitas  Santa Clara County  California
47     95036         Milpitas  Santa Clara County  California
48     95037      Morgan Hill  Santa Clara County  California
49     95038      Morgan Hill  Santa Clara County  California
50     95042      New Almaden  Santa Clara County  California
51     95044  Redwood Estates  Santa Clara County  California

                      Address  \
0         Los Altos,California
1   Los Altos Hills,California
2         Los Altos,California
3         Los Altos,California
```

```
4   Los Altos Hills,California
5     Moffett Field,California
6     Mountain View,California
7     Mountain View,California
8     Mountain View,California
9     Mountain View,California
10    Mountain View,California
11    Mountain View,California
12        Sunnyvale,California
13        Sunnyvale,California
14        Sunnyvale,California
16        Sunnyvale,California
17        Sunnyvale,California
18        Sunnyvale,California
19        Palo Alto,California
20        Palo Alto,California
21   East Palo Alto,California
22        Palo Alto,California
23        Palo Alto,California
24         Stanford,California
25        Palo Alto,California
26        Palo Alto,California
27         Stanford,California
28        Palo Alto,California
29        Palo Alto,California
30           Alviso,California
31         Campbell,California
32         Campbell,California
33         Campbell,California
34           Coyote,California
35      Monte Vista,California
36       Permanente,California
37         Cupertino,California
38         Cupertino,California
39            Gilroy,California
40            Gilroy,California
41         Holy City,California
42         Los Gatos,California
43     Monte Sereno,California
44         Los Gatos,California
45         Los Gatos,California
46          Milpitas,California
47          Milpitas,California
48      Morgan Hill,California
49      Morgan Hill,California
50      New Almaden,California
51  Redwood Estates,California
```

	location \
0	(Los Altos, Santa Clara County, California, Un…
1	(Los Altos Hills, Santa Clara County, Californ…
2	(Los Altos, Santa Clara County, California, Un…
3	(Los Altos, Santa Clara County, California, Un…
4	(Los Altos Hills, Santa Clara County, Californ…
5	(Building 126, Cummins Avenue, Ames Research C…
6	(Mountain View, Santa Clara County, California…
7	(Mountain View, Santa Clara County, California…
8	(Mountain View, Santa Clara County, California…
9	(Mountain View, Santa Clara County, California…
10	(Mountain View, Santa Clara County, California…
11	(Mountain View, Santa Clara County, California…
12	(Sunnyvale, Santa Clara County, California, Un…
13	(Sunnyvale, Santa Clara County, California, Un…
14	(Sunnyvale, Santa Clara County, California, Un…
16	(Sunnyvale, Santa Clara County, California, Un…
17	(Sunnyvale, Santa Clara County, California, Un…
18	(Sunnyvale, Santa Clara County, California, Un…
19	(Palo Alto, Santa Clara County, California, Un…
20	(Palo Alto, Santa Clara County, California, Un…
21	(East Palo Alto, San Mateo County, California,…
22	(Palo Alto, Santa Clara County, California, Un…
23	(Palo Alto, Santa Clara County, California, Un…
24	(Stanford, Santa Clara County, California, Uni…
25	(Palo Alto, Santa Clara County, California, Un…
26	(Palo Alto, Santa Clara County, California, Un…
27	(Stanford, Santa Clara County, California, Uni…
28	(Palo Alto, Santa Clara County, California, Un…
29	(Palo Alto, Santa Clara County, California, Un…
30	(Alviso, San Jose, Santa Clara County, Califor…
31	(Campbell, Santa Clara County, California, 950…
32	(Campbell, Santa Clara County, California, 950…
33	(Campbell, Santa Clara County, California, 950…
34	(Coyote, Santa Clara County, California, 95013…
35	(Monte Vista, Alta, Placer County, California,…
36	(Permanente, Stephen E. Abbors Trail, Loyola, …
37	(Cupertino, Santa Clara County, California, Un…
38	(Cupertino, Santa Clara County, California, Un…
39	(Gilroy, Santa Clara County, California, Unite…
40	(Gilroy, Santa Clara County, California, Unite…
41	(Holy City, Santa Clara County, California, 95…
42	(Los Gatos, Santa Clara County, California, Un…
43	(Monte Sereno, Santa Clara County, California,…
44	(Los Gatos, Santa Clara County, California, Un…
45	(Los Gatos, Santa Clara County, California, Un…

```
46  (Milpitas, Santa Clara County, California, Uni…
47  (Milpitas, Santa Clara County, California, Uni…
48  (Morgan Hill, Santa Clara County, California, …
49  (Morgan Hill, Santa Clara County, California, …
50  (New Almaden, Santa Clara County, California, …
51  (Redwood Estates, Santa Clara County, Californ…
```

	point	latitude	longitude	altitude
0	(37.3790629, -122.116578, 0.0)	37.379063	-122.116578	0.0
1	(37.3796627, -122.1374637, 0.0)	37.379663	-122.137464	0.0
2	(37.3790629, -122.116578, 0.0)	37.379063	-122.116578	0.0
3	(37.3790629, -122.116578, 0.0)	37.379063	-122.116578	0.0
4	(37.3796627, -122.1374637, 0.0)	37.379663	-122.137464	0.0
5	(37.41129325, -122.05416980310586, 0.0)	37.411293	-122.054170	0.0
6	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0
7	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0
8	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0
9	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0
10	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0
11	(37.3893889, -122.0832101, 0.0)	37.389389	-122.083210	0.0
12	(37.3688301, -122.0363496, 0.0)	37.368830	-122.036350	0.0
13	(37.3688301, -122.0363496, 0.0)	37.368830	-122.036350	0.0
14	(37.3688301, -122.0363496, 0.0)	37.368830	-122.036350	0.0
16	(37.3688301, -122.0363496, 0.0)	37.368830	-122.036350	0.0
17	(37.3688301, -122.0363496, 0.0)	37.368830	-122.036350	0.0
18	(37.3688301, -122.0363496, 0.0)	37.368830	-122.036350	0.0
19	(37.4443293, -122.1598465, 0.0)	37.444329	-122.159847	0.0
20	(37.4443293, -122.1598465, 0.0)	37.444329	-122.159847	0.0
21	(37.4688273, -122.1410751, 0.0)	37.468827	-122.141075	0.0
22	(37.4443293, -122.1598465, 0.0)	37.444329	-122.159847	0.0
23	(37.4443293, -122.1598465, 0.0)	37.444329	-122.159847	0.0
24	(37.427467, -122.1702445, 0.0)	37.427467	-122.170244	0.0
25	(37.4443293, -122.1598465, 0.0)	37.444329	-122.159847	0.0
26	(37.4443293, -122.1598465, 0.0)	37.444329	-122.159847	0.0
27	(37.427467, -122.1702445, 0.0)	37.427467	-122.170244	0.0
28	(37.4443293, -122.1598465, 0.0)	37.444329	-122.159847	0.0
29	(37.4443293, -122.1598465, 0.0)	37.444329	-122.159847	0.0
30	(37.426051, -121.9752373, 0.0)	37.426051	-121.975237	0.0
31	(37.2870626, -121.9448818, 0.0)	37.287063	-121.944882	0.0
32	(37.2870626, -121.9448818, 0.0)	37.287063	-121.944882	0.0
33	(37.2870626, -121.9448818, 0.0)	37.287063	-121.944882	0.0
34	(37.2164923, -121.7394174, 0.0)	37.216492	-121.739417	0.0
35	(39.1865648, -120.8327162, 0.0)	39.186565	-120.832716	0.0
36	(37.3189, -122.1111, 0.0)	37.318900	-122.111100	0.0
37	(37.3228934, -122.0322895, 0.0)	37.322893	-122.032290	0.0
38	(37.3228934, -122.0322895, 0.0)	37.322893	-122.032290	0.0
39	(37.0065078, -121.5631723, 0.0)	37.006508	-121.563172	0.0

```
40          (37.0065078, -121.5631723, 0.0)  37.006508 -121.563172       0.0
41          (37.1569939, -121.9788351, 0.0)  37.156994 -121.978835       0.0
42           (37.226611, -121.9746797, 0.0)  37.226611 -121.974680       0.0
43            (37.236333, -121.992458, 0.0)  37.236333 -121.992458       0.0
44           (37.226611, -121.9746797, 0.0)  37.226611 -121.974680       0.0
45           (37.226611, -121.9746797, 0.0)  37.226611 -121.974680       0.0
46          (37.4282724, -121.9066238, 0.0)  37.428272 -121.906624       0.0
47          (37.4282724, -121.9066238, 0.0)  37.428272 -121.906624       0.0
48           (37.130408, -121.6544974, 0.0)  37.130408 -121.654497       0.0
49           (37.130408, -121.6544974, 0.0)  37.130408 -121.654497       0.0
50          (37.1760567, -121.8207855, 0.0)  37.176057 -121.820785       0.0
51          (37.1563361, -121.9866223, 0.0)  37.156336 -121.986622       0.0
```

[46]:
```
#map cities in San Jose
SJ_map = folium.Map(location=[latitude, longitude], zoom_start=10) # generate␣
 ↪map centred around the Conrad Hotel

# add a red circle marker to represent
folium.features.CircleMarker(
    [37.3361905, -121.8905833],
    radius=7,
    color='red',
    popup='San Jose',
    fill = True,
    fill_color = 'red',
    fill_opacity = 0.6
).add_to(SJ_map)
#add cities
for lat, lng, label in zip(df1['latitude'], df1['longitude'], df1['City']):
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='green',
        fill=True,
        fill_color='green',
        fill_opacity=0.7,
        parse_html=False).add_to(SJ_map)


# display map of San Jose and it's cities
SJ_map
```

[46]:
```
<folium.folium.Map at 0x7f13310743c8>
```

[47]:
```
CLIENT_ID = 'GBATYYSJS45OFJJ2CJNIFQUNYXRMOUE2UWNCJDUOONES5HTS' # your␣
↪Foursquare ID
CLIENT_SECRET = 'N1FIOL3CFR2VTWTILBC1OZTB4T5S4GGMDR44BK3HNIU1OLML' # your␣
↪Foursquare Secret
VERSION = '20180604'
LIMIT = 30
print('done')
```

done

[48]:
```
#use Foursquare to get venues in San Jose cities
def getNearbyVenues(names, latitudes, longitudes, radius=500):

    venues_list=[]
    for name, lat, lng in zip(names, latitudes, longitudes):
        print(name)

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?
↪&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]['groups'][0]['items']

        # return only relevant information for each nearby venue
        venues_list.append([(
            name,
            lat,
            lng,
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],
            v['venue']['categories'][0]['name']) for v in results])

    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item␣
↪in venue_list])
    nearby_venues.columns = ['City',
                  'City Latitude',
                  'City Longitude',
                  'Venue',
```

```
                  'Venue Latitude',
                  'Venue Longitude',
                  'Venue Category']

    return(nearby_venues)
```

[49]:
```
#get venues by city
SC_venues = getNearbyVenues(names=df1['City'],
                                    latitudes=df1['latitude'],
                                    longitudes=df1['longitude'])
```

```
Los Altos
Los Altos Hills
Los Altos
Los Altos
Los Altos Hills
Moffett Field
Mountain View
Mountain View
Mountain View
Mountain View
Mountain View
Mountain View
Sunnyvale
Sunnyvale
Sunnyvale
Sunnyvale
Sunnyvale
Sunnyvale
Palo Alto
Palo Alto
East Palo Alto
Palo Alto
Palo Alto
Stanford
Palo Alto
Palo Alto
Stanford
Palo Alto
Palo Alto
Alviso
Campbell
Campbell
Campbell
Coyote
Monte Vista
Permanente
```

```
Cupertino
Cupertino
Gilroy
Gilroy
Holy City
Los Gatos
Monte Sereno
Los Gatos
Los Gatos
Milpitas
Milpitas
Morgan Hill
Morgan Hill
New Almaden
Redwood Estates
```

[50]:
```
print(SC_venues.shape)
SC_venues.head()
```

```
(1104, 7)
```

[50]:

	City	City Latitude	City Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Los Altos	37.379063	-122.116578	Satura Cakes	37.378835	-122.115860	Bakery
1	Los Altos	37.379063	-122.116578	Linden Tree	37.379216	-122.116963	Bookstore
2	Los Altos	37.379063	-122.116578	Tin Pot Creamery	37.378643	-122.118322	Ice Cream Shop
3	Los Altos	37.379063	-122.116578	State of Mind Public House & Pizzeria	37.380318	-122.115728	Pizza Place
4	Los Altos	37.379063	-122.116578	Asa	37.379490	-122.116821	Spanish Restaurant

[51]:
```
SC_venues.groupby('City').count()
```

[51]:

City	City Latitude	City Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Alviso	6	6	6	6	6	6
Campbell	90	90	90	90	90	90

Coyote	3	3	3	3
Cupertino	60	60	60	60
East Palo Alto	11	11	11	11
Gilroy	10	10	10	10
Holy City	4	4	4	4
Los Altos	90	90	90	90
Los Altos Hills	2	2	2	2
Los Gatos	33	33	33	33
Milpitas	60	60	60	60
Moffett Field	7	7	7	7
Monte Sereno	1	1	1	1
Monte Vista	4	4	4	4
Morgan Hill	60	60	60	60
Mountain View	180	180	180	180
New Almaden	4	4	4	4
Palo Alto	240	240	240	240
Redwood Estates	5	5	5	5
Stanford	54	54	54	54
Sunnyvale	180	180	180	180

City	Venue Longitude	Venue Category
Alviso	6	6
Campbell	90	90
Coyote	3	3
Cupertino	60	60
East Palo Alto	11	11
Gilroy	10	10
Holy City	4	4
Los Altos	90	90
Los Altos Hills	2	2
Los Gatos	33	33
Milpitas	60	60
Moffett Field	7	7
Monte Sereno	1	1
Monte Vista	4	4
Morgan Hill	60	60
Mountain View	180	180
New Almaden	4	4
Palo Alto	240	240
Redwood Estates	5	5
Stanford	54	54
Sunnyvale	180	180

[52]:
```
print('There are {} uniques categories.'.format(len(SC_venues['Venue Category'].
↪unique())))
```

There are 125 uniques categories.

[53]:
```
#analyze each city
# one hot encoding
SC_onehot = pd.get_dummies(SC_venues[['Venue Category']], prefix="",␣
 ↪prefix_sep="")

# add neighborhood column back to dataframe
SC_onehot['City'] =SC_venues['City']

# move neighborhood column to the first column
fixed_columns = [SC_onehot.columns[-1]] + list(SC_onehot.columns[:-1])
SC_onehot = SC_onehot[fixed_columns]

SC_onehot.head()
```

[53]:

	City	Airport	American Restaurant	Arcade	Art Gallery	\
0	Los Altos	0	0	0	0	
1	Los Altos	0	0	0	0	
2	Los Altos	0	0	0	0	
3	Los Altos	0	0	0	0	
4	Los Altos	0	0	0	0	

	Asian Restaurant	BBQ Joint	Bagel Shop	Bakery	Bank	…	\
0	0	0	0	1	0	…	
1	0	0	0	0	0	…	
2	0	0	0	0	0	…	
3	0	0	0	0	0	…	
4	0	0	0	0	0	…	

	Thrift / Vintage Store	Toy / Game Store	Track Stadium	Trail	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Train Station	Used Bookstore	Video Game Store	Vietnamese Restaurant	\
0	0	0	0	0	
1	0	0	0	0	
2	0	0	0	0	
3	0	0	0	0	
4	0	0	0	0	

	Wine Bar	Yoga Studio
0	0	0
1	0	0

```
2         0            0
3         0            0
4         0            0

[5 rows x 126 columns]
```

[54]:
```
#group each city and take the mean of the frequency of occurrence of each␣
↪category
SC_grouped = SC_onehot.groupby(df1['City']).mean().reset_index()
SC_grouped
```

[54]:

	City	Airport	American Restaurant	Arcade	Art Gallery	\
0	Alviso	0.0	0.0	0.000	0.0	
1	Campbell	0.0	0.0	0.000	0.0	
2	Coyote	0.0	0.0	0.000	0.0	
3	Cupertino	0.0	0.0	0.000	0.0	
4	East Palo Alto	0.0	0.0	0.000	0.0	
5	Gilroy	0.0	0.0	0.000	0.0	
6	Holy City	0.0	0.0	0.000	0.0	
7	Los Altos	0.0	0.0	0.000	0.0	
8	Los Altos Hills	0.0	0.0	0.000	0.0	
9	Los Gatos	0.0	0.0	0.000	0.0	
10	Milpitas	0.0	0.0	0.000	0.0	
11	Moffett Field	0.0	1.0	0.000	0.0	
12	Monte Sereno	0.0	0.0	0.000	0.0	
13	Monte Vista	0.0	0.0	0.000	0.0	
14	Morgan Hill	0.0	0.0	0.000	0.0	
15	Mountain View	0.0	0.0	0.000	0.0	
16	New Almaden	0.0	0.0	0.000	0.0	
17	Palo Alto	0.0	0.0	0.125	0.0	
18	Permanente	0.0	1.0	0.000	0.0	
19	Redwood Estates	0.0	0.0	0.000	0.0	
20	Stanford	0.0	0.5	0.000	0.0	
21	Sunnyvale	0.0	0.0	0.000	0.0	

	Asian Restaurant	BBQ Joint	Bagel Shop	Bakery	Bank	…	\
0	0.0	0.0	0.0	0.000000	0.0	…	
1	0.0	0.0	0.0	0.333333	0.0	…	
2	0.0	0.0	0.0	0.000000	0.0	…	
3	0.0	0.0	0.0	0.500000	0.0	…	
4	0.0	0.0	0.0	0.000000	0.0	…	
5	0.0	0.0	0.0	0.000000	0.0	…	
6	0.0	0.0	0.0	0.000000	0.0	…	
7	0.0	0.0	0.0	0.333333	0.0	…	
8	0.0	0.0	0.0	0.000000	0.0	…	
9	0.0	0.0	0.0	0.000000	0.0	…	
10	0.0	0.0	0.0	0.000000	0.0	…	

11	0.0	0.0	0.0	0.000000	0.0	…
12	0.0	0.0	0.0	0.000000	0.0	…
13	0.0	0.0	0.0	0.000000	0.0	…
14	0.0	0.0	0.0	0.000000	0.0	…
15	0.0	0.0	0.0	0.166667	0.0	…
16	0.0	0.0	0.0	0.000000	0.0	…
17	0.0	0.0	0.0	0.000000	0.0	…
18	0.0	0.0	0.0	0.000000	0.0	…
19	0.0	0.0	0.0	0.000000	0.0	…
20	0.0	0.0	0.0	0.000000	0.0	…
21	0.0	0.0	0.0	0.000000	0.0	…

	Thrift / Vintage Store	Toy / Game Store	Track Stadium	Trail	\
0	0.0	0.000000	0.0	0.0	
1	0.0	0.000000	0.0	0.0	
2	0.0	0.000000	0.0	0.0	
3	0.0	0.000000	0.0	0.0	
4	0.0	0.000000	0.0	0.0	
5	0.0	0.000000	0.0	0.0	
6	0.0	0.000000	0.0	0.0	
7	0.0	0.000000	0.0	0.0	
8	0.0	0.000000	0.0	0.0	
9	0.0	0.000000	0.0	0.0	
10	0.0	0.000000	0.0	0.0	
11	0.0	0.000000	0.0	0.0	
12	0.0	0.000000	0.0	0.0	
13	0.0	0.000000	0.0	0.0	
14	0.0	0.500000	0.0	0.0	
15	0.0	0.000000	0.0	0.0	
16	0.0	0.000000	0.0	0.0	
17	0.0	0.000000	0.0	0.0	
18	0.0	0.000000	0.0	0.0	
19	0.0	0.000000	0.0	0.0	
20	0.0	0.000000	0.0	0.0	
21	0.0	0.166667	0.0	0.0	

	Train Station	Used Bookstore	Video Game Store	Vietnamese Restaurant	\
0	0.0	0.0	0.0	0.0	
1	0.0	0.0	0.0	0.0	
2	0.0	0.0	0.0	0.0	
3	0.0	0.0	0.0	0.0	
4	0.0	0.0	0.0	0.0	
5	0.0	0.0	0.0	0.0	
6	0.0	0.0	0.0	0.0	
7	0.0	0.0	0.0	0.0	
8	0.0	0.0	0.0	0.0	
9	0.0	0.0	0.0	0.0	

```
10            0.0             0.0             0.0             0.0
11            0.0             0.0             0.0             0.0
12            0.0             0.0             0.0             0.0
13            0.0             0.0             0.0             0.0
14            0.0             0.0             0.0             0.0
15            0.0             0.0             0.0             0.0
16            0.0             0.0             0.0             0.0
17            0.0             0.0             0.0             0.0
18            0.0             0.0             0.0             0.0
19            0.0             0.0             0.0             0.0
20            0.0             0.0             0.0             0.0
21            0.0             0.0             0.0             0.0

    Wine Bar  Yoga Studio
0      0.000          0.0
1      0.000          0.0
2      0.000          0.0
3      0.000          0.0
4      0.000          0.0
5      0.000          0.0
6      0.000          0.0
7      0.000          0.0
8      0.000          0.0
9      0.000          0.0
10     0.000          0.0
11     0.000          0.0
12     0.000          0.0
13     0.000          0.0
14     0.000          0.0
15     0.000          0.0
16     0.000          0.0
17     0.125          0.0
18     0.000          0.0
19     0.000          0.0
20     0.000          0.0
21     0.000          0.0

[22 rows x 126 columns]
```

[55]:
```
#print city with top 5 venues
num_top_venues = 5

for city in SC_grouped['City']:
    print("----"+city+"----")
    temp = SC_grouped[SC_grouped['City'] == city].T.reset_index()
    temp.columns = ['venue','freq']
    temp = temp.iloc[1:]
```

```
    temp['freq'] = temp['freq'].astype(float)
    temp = temp.round({'freq': 2})
    print(temp.sort_values('freq', ascending=False).reset_index(drop=True).
↪head(num_top_venues))
    print('\n')
```

```
----Alviso----
                     venue  freq
0               Music Venue   1.0
1                   Airport   0.0
2  Mediterranean Restaurant   0.0
3               Post Office   0.0
4                      Pool   0.0


----Campbell----
            venue  freq
0  Ice Cream Shop  0.33
1         Bakery  0.33
2       Bookstore  0.33
3         Airport  0.00
4    Optical Shop  0.00


----Coyote----
                     venue  freq
0               Pizza Place   1.0
1                   Airport   0.0
2  Mediterranean Restaurant   0.0
3               Post Office   0.0
4                      Pool   0.0


----Cupertino----
                       venue  freq
0        Japanese Restaurant   0.5
1                     Bakery   0.5
2                    Airport   0.0
3  Middle Eastern Restaurant   0.0
4        Monument / Landmark   0.0


----East Palo Alto----
                     venue  freq
0         Sushi Restaurant   1.0
1                  Airport   0.0
2  Mediterranean Restaurant   0.0
```

```
3                      Pool   0.0
4                     Plaza   0.0


----Gilroy----
                 venue  freq
0       Deli / Bodega   0.5
1      Farmers Market   0.5
2             Airport   0.0
3  Mexican Restaurant   0.0
4                Pool   0.0


----Holy City----
                      venue  freq
0               Coffee Shop   1.0
1  Mediterranean Restaurant   0.0
2               Post Office   0.0
3                      Pool   0.0
4                     Plaza   0.0


----Los Altos----
            venue  freq
0  Ice Cream Shop  0.33
1     Pizza Place  0.33
2          Bakery  0.33
3         Airport  0.00
4    Optical Shop  0.00


----Los Altos Hills----
                      venue  freq
0        Spanish Restaurant   0.5
1                 Bookstore   0.5
2                   Airport   0.0
3  Mediterranean Restaurant   0.0
4                      Pool   0.0


----Los Gatos----
                      venue  freq
0                     Hotel  0.33
1             Breakfast Spot  0.33
2              Grocery Store  0.33
3              Optical Shop  0.00
4  Middle Eastern Restaurant  0.00
```

```
----Milpitas----
                 venue  freq
0           Restaurant   0.5
1         Cocktail Bar   0.5
2  Rental Car Location   0.0
3          Post Office   0.0
4                 Pool   0.0


----Moffett Field----
                 venue  freq
0  American Restaurant   1.0
1       Massage Studio   0.0
2                 Pool   0.0
3               Plaza   0.0
4          Pizza Place   0.0


----Monte Sereno----
            venue  freq
0    Gourmet Shop   1.0
1         Airport   0.0
2  Massage Studio   0.0
3            Pool   0.0
4           Plaza   0.0


----Monte Vista----
                      venue  freq
0        Spanish Restaurant   1.0
1                   Airport   0.0
2  Mediterranean Restaurant   0.0
3                      Pool   0.0
4                    Plaza   0.0


----Morgan Hill----
                venue  freq
0    Toy / Game Store   0.5
1  Mexican Restaurant   0.5
2             Airport   0.0
3      Massage Studio   0.0
4                Pool   0.0


----Mountain View----
                 venue  freq
```

```
0       Farmers Market  0.17
1          Coffee Shop  0.17
2        Deli / Bodega  0.17
3  Japanese Restaurant  0.17
4               Bakery  0.17


----New Almaden----
                      venue  freq
0                      Park   1.0
1                   Airport   0.0
2  Mediterranean Restaurant   0.0
3               Post Office   0.0
4                      Pool   0.0


----Palo Alto----
                venue  freq
0      History Museum  0.12
1                Park  0.12
2  Mexican Restaurant  0.12
3      Farmers Market  0.12
4   French Restaurant  0.12


----Permanente----
                  venue  freq
0  American Restaurant   1.0
1       Massage Studio   0.0
2                 Pool   0.0
3               Plaza    0.0
4          Pizza Place   0.0


----Redwood Estates----
                      venue  freq
0            Farmers Market   1.0
1                   Airport   0.0
2  Mediterranean Restaurant   0.0
3                      Pool   0.0
4                    Plaza    0.0


----Stanford----
                  venue  freq
0  American Restaurant   0.5
1     Greek Restaurant   0.5
2       Massage Studio   0.0
```

```
3                 Pool   0.0
4               Plaza    0.0


----Sunnyvale----
              venue  freq
0             Hotel  0.17
1     Grocery Store  0.17
2      Gourmet Shop  0.17
3  Toy / Game Store  0.17
4      Cocktail Bar  0.17
```

[ ]:
```
#The cities where bakery was on the top 5 venues were Campbell, Cupertino and␣
↪Los Altos, Mountain View
```

[ ]:
```
#Campell=.33, Cupertino=.5, Los Altos=.33, Mountain View=.17
```

[56]:
```
def return_most_common_venues(row, num_top_venues):
    row_categories = row.iloc[1:]
    row_categories_sorted = row_categories.sort_values(ascending=False)

    return row_categories_sorted.index.values[0:num_top_venues]
```

[60]:
```
#make dataframe for top 10 venues
num_top_venues = 10

indicators = ['st', 'nd', 'rd']

# create columns according to number of top venues
columns = ['City']
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

# create a new dataframe
SC_venues_sorted = pd.DataFrame(columns=columns)
SC_venues_sorted['City'] = SC_grouped['City']

for ind in np.arange(SC_grouped.shape[0]):
    SC_venues_sorted.iloc[ind, 1:] = return_most_common_venues(SC_grouped.
↪iloc[ind, :], num_top_venues)

SC_venues_sorted.head()
```

[60]:
```
             City 1st Most Common Venue 2nd Most Common Venue  \
0          Alviso           Music Venue           Yoga Studio
1        Campbell                Bakery        Ice Cream Shop
2          Coyote           Pizza Place           Yoga Studio
3       Cupertino                Bakery   Japanese Restaurant
4  East Palo Alto      Sushi Restaurant           Yoga Studio

  3rd Most Common Venue 4th Most Common Venue 5th Most Common Venue  \
0    Falafel Restaurant          Concert Hall     Convenience Store
1             Bookstore                 Diner     Electronics Store
2    Falafel Restaurant          Concert Hall     Convenience Store
3                Diner     Electronics Store   Dumpling Restaurant
4          Cocktail Bar           Comedy Club          Concert Hall

  6th Most Common Venue 7th Most Common Venue 8th Most Common Venue  \
0       Cosmetics Shop              Creperie         Deli / Bodega
1   Dumpling Restaurant            Donut Shop               Dog Run
2       Cosmetics Shop              Creperie         Deli / Bodega
3           Donut Shop               Dog Run              Dive Bar
4    Convenience Store        Cosmetics Shop              Creperie

  9th Most Common Venue 10th Most Common Venue
0          Dessert Shop                  Diner
1              Dive Bar            Yoga Studio
2          Dessert Shop                  Diner
3           Yoga Studio         Farmers Market
4         Deli / Bodega           Dessert Shop
```

[61]:
```
# import k-means from clustering stage to cluster cities
from sklearn.cluster import KMeans
# Matplotlib and associated plotting modules
import matplotlib.cm as cm
import matplotlib.colors as colors
```

[62]:
```
# set number of clusters
kclusters = 5

SC_grouped_clustering = SC_grouped.drop('City', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(SC_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]
```

[62]:
```
array([3, 2, 2, 2, 2, 2, 2, 2, 4, 2], dtype=int32)
```

[63]:
```
# add clustering labels
SC_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)


SC_merged = df1

# merge SC_grouped with dataframe to add latitude/longitude for each City
SC_merged = SC_merged.join(SC_venues_sorted.set_index('City'), on='City')

SC_merged.head()
```

[63]:
```
   ZIP code             City              County       State  \
0     94022        Los Altos  Santa Clara County  California
1     94022  Los Altos Hills  Santa Clara County  California
2     94023        Los Altos  Santa Clara County  California
3     94024        Los Altos  Santa Clara County  California
4     94024  Los Altos Hills  Santa Clara County  California

                      Address  \
0        Los Altos,California
1  Los Altos Hills,California
2        Los Altos,California
3        Los Altos,California
4  Los Altos Hills,California

                                           location  \
0  (Los Altos, Santa Clara County, California, Un…
1  (Los Altos Hills, Santa Clara County, Californ…
2  (Los Altos, Santa Clara County, California, Un…
3  (Los Altos, Santa Clara County, California, Un…
4  (Los Altos Hills, Santa Clara County, Californ…

                             point   latitude   longitude  altitude  …  \
0   (37.3790629, -122.116578, 0.0)  37.379063 -122.116578       0.0  …
1  (37.3796627, -122.1374637, 0.0)  37.379663 -122.137464       0.0  …
2   (37.3790629, -122.116578, 0.0)  37.379063 -122.116578       0.0  …
3   (37.3790629, -122.116578, 0.0)  37.379063 -122.116578       0.0  …
4  (37.3796627, -122.1374637, 0.0)  37.379663 -122.137464       0.0  …

   1st Most Common Venue 2nd Most Common Venue 3rd Most Common Venue  \
0                Bakery           Pizza Place        Ice Cream Shop
1    Spanish Restaurant             Bookstore           Yoga Studio
2                Bakery           Pizza Place        Ice Cream Shop
3                Bakery           Pizza Place        Ice Cream Shop
4    Spanish Restaurant             Bookstore           Yoga Studio

  4th Most Common Venue 5th Most Common Venue 6th Most Common Venue  \
```

```
0          Dessert Shop   Dumpling Restaurant            Donut Shop
1    Falafel Restaurant          Concert Hall     Convenience Store
2          Dessert Shop   Dumpling Restaurant            Donut Shop
3          Dessert Shop   Dumpling Restaurant            Donut Shop
4    Falafel Restaurant          Concert Hall     Convenience Store

  7th Most Common Venue 8th Most Common Venue 9th Most Common Venue  \
0               Dog Run              Dive Bar                 Diner
1        Cosmetics Shop              Creperie         Deli / Bodega
2               Dog Run              Dive Bar                 Diner
3               Dog Run              Dive Bar                 Diner
4        Cosmetics Shop              Creperie         Deli / Bodega

  10th Most Common Venue
0            Yoga Studio
1           Dessert Shop
2            Yoga Studio
3            Yoga Studio
4           Dessert Shop

[5 rows x 21 columns]
```

[64]:
```
SC_merged['Cluster Labels'].astype(int)
```

[64]:
```
0     2
1     4
2     2
3     2
4     4
5     1
6     2
7     2
8     2
9     2
10    2
11    2
12    2
13    2
14    2
16    2
17    2
18    2
19    2
20    2
21    2
22    2
23    2
```

```
24    1
25    2
26    2
27    1
28    2
29    2
30    3
31    2
32    2
33    2
34    2
35    4
36    1
37    2
38    2
39    2
40    2
41    2
42    2
43    0
44    2
45    2
46    2
47    2
48    2
49    2
50    2
51    2
Name: Cluster Labels, dtype: int64
```

[65]:
```
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11)

# set color scheme for the clusters
x = np.arange(kclusters)
ys = [i + x + (i*x)**2 for i in range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = [colors.rgb2hex(i) for i in colors_array]

# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(SC_merged['latitude'],␣
 ↪SC_merged['longitude'], SC_merged['City'], SC_merged['Cluster Labels']):
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
```

```
        color=rainbow[cluster-1],
        fill=True,
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```

[65]:
```
<folium.folium.Map at 0x7f12f5cf4c18>
```

[66]:
```
SC_merged.loc[SC_merged['Cluster Labels'] == 0,SC_merged.columns[[1] +␣
↪list(range(5, SC_merged.shape[1]))]]
```

[66]:
```
            City                                           location  \
43  Monte Sereno  (Monte Sereno, Santa Clara County, California,…

                           point   latitude   longitude  altitude  \
43  (37.236333, -121.992458, 0.0)  37.236333 -121.992458       0.0

    Cluster Labels 1st Most Common Venue 2nd Most Common Venue  \
43               0          Gourmet Shop           Yoga Studio

   3rd Most Common Venue 4th Most Common Venue 5th Most Common Venue  \
43    Falafel Restaurant           Comedy Club          Concert Hall

   6th Most Common Venue 7th Most Common Venue 8th Most Common Venue  \
43     Convenience Store        Cosmetics Shop              Creperie

   9th Most Common Venue 10th Most Common Venue
43         Deli / Bodega           Dessert Shop
```

[ ]:
```
#cluster 1 Gourmet Shops/Yoga
```

[67]:
```
SC_merged.loc[SC_merged['Cluster Labels'] == 1,SC_merged.columns[[1] +␣
↪list(range(5, SC_merged.shape[1]))]]
```

[67]:
```
            City                                           location  \
5   Moffett Field  (Building 126, Cummins Avenue, Ames Research C…
24       Stanford  (Stanford, Santa Clara County, California, Uni…
27       Stanford  (Stanford, Santa Clara County, California, Uni…
36     Permanente  (Permanente, Stephen E. Abbors Trail, Loyola, …

                                    point   latitude   longitude  altitude  \
5   (37.41129325, -122.05416980310586, 0.0)  37.411293 -122.054170       0.0
24           (37.427467, -122.1702445, 0.0)  37.427467 -122.170244       0.0
27           (37.427467, -122.1702445, 0.0)  37.427467 -122.170244       0.0
36                (37.3189, -122.1111, 0.0)  37.318900 -122.111100       0.0
```

```
    Cluster Labels 1st Most Common Venue 2nd Most Common Venue  \
5                1   American Restaurant           Yoga Studio
24               1   American Restaurant      Greek Restaurant
27               1   American Restaurant      Greek Restaurant
36               1   American Restaurant           Yoga Studio

   3rd Most Common Venue 4th Most Common Venue 5th Most Common Venue  \
5         Farmers Market          Concert Hall     Convenience Store
24           Yoga Studio    Falafel Restaurant          Concert Hall
27           Yoga Studio    Falafel Restaurant          Concert Hall
36        Farmers Market          Concert Hall     Convenience Store

   6th Most Common Venue 7th Most Common Venue 8th Most Common Venue  \
5         Cosmetics Shop              Creperie         Deli / Bodega
24     Convenience Store        Cosmetics Shop              Creperie
27     Convenience Store        Cosmetics Shop              Creperie
36        Cosmetics Shop              Creperie         Deli / Bodega

   9th Most Common Venue 10th Most Common Venue
5           Dessert Shop                 Diner
24         Deli / Bodega          Dessert Shop
27         Deli / Bodega          Dessert Shop
36          Dessert Shop                 Diner
```

[ ]:
```
#Cluster 2 American and Greek Restaurants/Yoga
```

[68]:
```
SC_merged.loc[SC_merged['Cluster Labels'] == 2,SC_merged.columns[[1] +␣
↪list(range(5, SC_merged.shape[1]))]]
```

[68]:
```
              City                                           location  \
0        Los Altos  (Los Altos, Santa Clara County, California, Un…
2        Los Altos  (Los Altos, Santa Clara County, California, Un…
3        Los Altos  (Los Altos, Santa Clara County, California, Un…
6    Mountain View  (Mountain View, Santa Clara County, California…
7    Mountain View  (Mountain View, Santa Clara County, California…
8    Mountain View  (Mountain View, Santa Clara County, California…
9    Mountain View  (Mountain View, Santa Clara County, California…
10   Mountain View  (Mountain View, Santa Clara County, California…
11   Mountain View  (Mountain View, Santa Clara County, California…
12       Sunnyvale  (Sunnyvale, Santa Clara County, California, Un…
13       Sunnyvale  (Sunnyvale, Santa Clara County, California, Un…
14       Sunnyvale  (Sunnyvale, Santa Clara County, California, Un…
16       Sunnyvale  (Sunnyvale, Santa Clara County, California, Un…
17       Sunnyvale  (Sunnyvale, Santa Clara County, California, Un…
18       Sunnyvale  (Sunnyvale, Santa Clara County, California, Un…
19       Palo Alto  (Palo Alto, Santa Clara County, California, Un…
20       Palo Alto  (Palo Alto, Santa Clara County, California, Un…
```

```
21   East Palo Alto  (East Palo Alto, San Mateo County, California,…
22        Palo Alto  (Palo Alto, Santa Clara County, California, Un…
23        Palo Alto  (Palo Alto, Santa Clara County, California, Un…
25        Palo Alto  (Palo Alto, Santa Clara County, California, Un…
26        Palo Alto  (Palo Alto, Santa Clara County, California, Un…
28        Palo Alto  (Palo Alto, Santa Clara County, California, Un…
29        Palo Alto  (Palo Alto, Santa Clara County, California, Un…
31         Campbell  (Campbell, Santa Clara County, California, 950…
32         Campbell  (Campbell, Santa Clara County, California, 950…
33         Campbell  (Campbell, Santa Clara County, California, 950…
34           Coyote  (Coyote, Santa Clara County, California, 95013…
37        Cupertino  (Cupertino, Santa Clara County, California, Un…
38        Cupertino  (Cupertino, Santa Clara County, California, Un…
39           Gilroy  (Gilroy, Santa Clara County, California, Unite…
40           Gilroy  (Gilroy, Santa Clara County, California, Unite…
41        Holy City  (Holy City, Santa Clara County, California, 95…
42        Los Gatos  (Los Gatos, Santa Clara County, California, Un…
44        Los Gatos  (Los Gatos, Santa Clara County, California, Un…
45        Los Gatos  (Los Gatos, Santa Clara County, California, Un…
46         Milpitas  (Milpitas, Santa Clara County, California, Uni…
47         Milpitas  (Milpitas, Santa Clara County, California, Uni…
48      Morgan Hill  (Morgan Hill, Santa Clara County, California, …
49      Morgan Hill  (Morgan Hill, Santa Clara County, California, …
50      New Almaden  (New Almaden, Santa Clara County, California, …
51  Redwood Estates  (Redwood Estates, Santa Clara County, Californ…

                              point   latitude   longitude  altitude  \
0    (37.3790629, -122.116578, 0.0)  37.379063 -122.116578       0.0
2    (37.3790629, -122.116578, 0.0)  37.379063 -122.116578       0.0
3    (37.3790629, -122.116578, 0.0)  37.379063 -122.116578       0.0
6   (37.3893889, -122.0832101, 0.0)  37.389389 -122.083210       0.0
7   (37.3893889, -122.0832101, 0.0)  37.389389 -122.083210       0.0
8   (37.3893889, -122.0832101, 0.0)  37.389389 -122.083210       0.0
9   (37.3893889, -122.0832101, 0.0)  37.389389 -122.083210       0.0
10  (37.3893889, -122.0832101, 0.0)  37.389389 -122.083210       0.0
11  (37.3893889, -122.0832101, 0.0)  37.389389 -122.083210       0.0
12  (37.3688301, -122.0363496, 0.0)  37.368830 -122.036350       0.0
13  (37.3688301, -122.0363496, 0.0)  37.368830 -122.036350       0.0
14  (37.3688301, -122.0363496, 0.0)  37.368830 -122.036350       0.0
16  (37.3688301, -122.0363496, 0.0)  37.368830 -122.036350       0.0
17  (37.3688301, -122.0363496, 0.0)  37.368830 -122.036350       0.0
18  (37.3688301, -122.0363496, 0.0)  37.368830 -122.036350       0.0
19  (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
20  (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
21  (37.4688273, -122.1410751, 0.0)  37.468827 -122.141075       0.0
22  (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
23  (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
```

```
25  (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
26  (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
28  (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
29  (37.4443293, -122.1598465, 0.0)  37.444329 -122.159847       0.0
31  (37.2870626, -121.9448818, 0.0)  37.287063 -121.944882       0.0
32  (37.2870626, -121.9448818, 0.0)  37.287063 -121.944882       0.0
33  (37.2870626, -121.9448818, 0.0)  37.287063 -121.944882       0.0
34  (37.2164923, -121.7394174, 0.0)  37.216492 -121.739417       0.0
37  (37.3228934, -122.0322895, 0.0)  37.322893 -122.032290       0.0
38  (37.3228934, -122.0322895, 0.0)  37.322893 -122.032290       0.0
39  (37.0065078, -121.5631723, 0.0)  37.006508 -121.563172       0.0
40  (37.0065078, -121.5631723, 0.0)  37.006508 -121.563172       0.0
41  (37.1569939, -121.9788351, 0.0)  37.156994 -121.978835       0.0
42   (37.226611, -121.9746797, 0.0)  37.226611 -121.974680       0.0
44   (37.226611, -121.9746797, 0.0)  37.226611 -121.974680       0.0
45   (37.226611, -121.9746797, 0.0)  37.226611 -121.974680       0.0
46  (37.4282724, -121.9066238, 0.0)  37.428272 -121.906624       0.0
47  (37.4282724, -121.9066238, 0.0)  37.428272 -121.906624       0.0
48   (37.130408, -121.6544974, 0.0)  37.130408 -121.654497       0.0
49   (37.130408, -121.6544974, 0.0)  37.130408 -121.654497       0.0
50  (37.1760567, -121.8207855, 0.0)  37.176057 -121.820785       0.0
51  (37.1563361, -121.9866223, 0.0)  37.156336 -121.986622       0.0

    Cluster Labels 1st Most Common Venue 2nd Most Common Venue  \
0                2                Bakery           Pizza Place
2                2                Bakery           Pizza Place
3                2                Bakery           Pizza Place
6                2           Coffee Shop   Japanese Restaurant
7                2           Coffee Shop   Japanese Restaurant
8                2           Coffee Shop   Japanese Restaurant
9                2           Coffee Shop   Japanese Restaurant
10               2           Coffee Shop   Japanese Restaurant
11               2           Coffee Shop   Japanese Restaurant
12               2          Cocktail Bar      Toy / Game Store
13               2          Cocktail Bar      Toy / Game Store
14               2          Cocktail Bar      Toy / Game Store
16               2          Cocktail Bar      Toy / Game Store
17               2          Cocktail Bar      Toy / Game Store
18               2          Cocktail Bar      Toy / Game Store
19               2        Farmers Market                 Park
20               2        Farmers Market                 Park
21               2      Sushi Restaurant           Yoga Studio
22               2        Farmers Market                 Park
23               2        Farmers Market                 Park
25               2        Farmers Market                 Park
26               2        Farmers Market                 Park
28               2        Farmers Market                 Park
```

```
29                2        Farmers Market                  Park
31                2                Bakery        Ice Cream Shop
32                2                Bakery        Ice Cream Shop
33                2                Bakery        Ice Cream Shop
34                2           Pizza Place           Yoga Studio
37                2                Bakery   Japanese Restaurant
38                2                Bakery   Japanese Restaurant
39                2        Farmers Market         Deli / Bodega
40                2        Farmers Market         Deli / Bodega
41                2           Coffee Shop    Falafel Restaurant
42                2         Grocery Store        Breakfast Spot
44                2         Grocery Store        Breakfast Spot
45                2         Grocery Store        Breakfast Spot
46                2          Cocktail Bar            Restaurant
47                2          Cocktail Bar            Restaurant
48                2      Toy / Game Store    Mexican Restaurant
49                2      Toy / Game Store    Mexican Restaurant
50                2                  Park           Yoga Studio
51                2         Farmers Market    Falafel Restaurant

   3rd Most Common Venue 4th Most Common Venue 5th Most Common Venue  \
0         Ice Cream Shop          Dessert Shop   Dumpling Restaurant
2         Ice Cream Shop          Dessert Shop   Dumpling Restaurant
3         Ice Cream Shop          Dessert Shop   Dumpling Restaurant
6         Breakfast Spot         Deli / Bodega                Bakery
7         Breakfast Spot         Deli / Bodega                Bakery
8         Breakfast Spot         Deli / Bodega                Bakery
9         Breakfast Spot         Deli / Bodega                Bakery
10        Breakfast Spot         Deli / Bodega                Bakery
11        Breakfast Spot         Deli / Bodega                Bakery
12          Gourmet Shop                 Hotel    Mexican Restaurant
13          Gourmet Shop                 Hotel    Mexican Restaurant
14          Gourmet Shop                 Hotel    Mexican Restaurant
16          Gourmet Shop                 Hotel    Mexican Restaurant
17          Gourmet Shop                 Hotel    Mexican Restaurant
18          Gourmet Shop                 Hotel    Mexican Restaurant
19                Arcade             Wine Bar        History Museum
20                Arcade             Wine Bar        History Museum
21          Cocktail Bar           Comedy Club          Concert Hall
22                Arcade             Wine Bar        History Museum
23                Arcade             Wine Bar        History Museum
25                Arcade             Wine Bar        History Museum
26                Arcade             Wine Bar        History Museum
28                Arcade             Wine Bar        History Museum
29                Arcade             Wine Bar        History Museum
31             Bookstore                Diner     Electronics Store
32             Bookstore                Diner     Electronics Store
```

33	Bookstore	Diner	Electronics Store	
34	Falafel Restaurant	Concert Hall	Convenience Store	
37	Diner	Electronics Store	Dumpling Restaurant	
38	Diner	Electronics Store	Dumpling Restaurant	
39	Falafel Restaurant	Comedy Club	Concert Hall	
40	Falafel Restaurant	Comedy Club	Concert Hall	
41	Comedy Club	Concert Hall	Convenience Store	
42	Hotel	Yoga Studio	Diner	
44	Hotel	Yoga Studio	Diner	
45	Hotel	Yoga Studio	Diner	
46	Yoga Studio	Comedy Club	Concert Hall	
47	Yoga Studio	Comedy Club	Concert Hall	
48	Yoga Studio	Electronics Store	Comedy Club	
49	Yoga Studio	Electronics Store	Comedy Club	
50	Falafel Restaurant	Concert Hall	Convenience Store	
51	Comedy Club	Concert Hall	Convenience Store	

	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	\
0	Donut Shop	Dog Run	Dive Bar	
2	Donut Shop	Dog Run	Dive Bar	
3	Donut Shop	Dog Run	Dive Bar	
6	Farmers Market	Electronics Store	Dumpling Restaurant	
7	Farmers Market	Electronics Store	Dumpling Restaurant	
8	Farmers Market	Electronics Store	Dumpling Restaurant	
9	Farmers Market	Electronics Store	Dumpling Restaurant	
10	Farmers Market	Electronics Store	Dumpling Restaurant	
11	Farmers Market	Electronics Store	Dumpling Restaurant	
12	Grocery Store	BBQ Joint	Farmers Market	
13	Grocery Store	BBQ Joint	Farmers Market	
14	Grocery Store	BBQ Joint	Farmers Market	
16	Grocery Store	BBQ Joint	Farmers Market	
17	Grocery Store	BBQ Joint	Farmers Market	
18	Grocery Store	BBQ Joint	Farmers Market	
19	Sushi Restaurant	Mexican Restaurant	French Restaurant	
20	Sushi Restaurant	Mexican Restaurant	French Restaurant	
21	Convenience Store	Cosmetics Shop	Creperie	
22	Sushi Restaurant	Mexican Restaurant	French Restaurant	
23	Sushi Restaurant	Mexican Restaurant	French Restaurant	
25	Sushi Restaurant	Mexican Restaurant	French Restaurant	
26	Sushi Restaurant	Mexican Restaurant	French Restaurant	
28	Sushi Restaurant	Mexican Restaurant	French Restaurant	
29	Sushi Restaurant	Mexican Restaurant	French Restaurant	
31	Dumpling Restaurant	Donut Shop	Dog Run	
32	Dumpling Restaurant	Donut Shop	Dog Run	
33	Dumpling Restaurant	Donut Shop	Dog Run	
34	Cosmetics Shop	Creperie	Deli / Bodega	
37	Donut Shop	Dog Run	Dive Bar	

38	Donut Shop	Dog Run	Dive Bar
39	Convenience Store	Cosmetics Shop	Creperie
40	Convenience Store	Cosmetics Shop	Creperie
41	Cosmetics Shop	Creperie	Deli / Bodega
42	Dumpling Restaurant	Donut Shop	Dog Run
44	Dumpling Restaurant	Donut Shop	Dog Run
45	Dumpling Restaurant	Donut Shop	Dog Run
46	Convenience Store	Cosmetics Shop	Creperie
47	Convenience Store	Cosmetics Shop	Creperie
48	Concert Hall	Convenience Store	Cosmetics Shop
49	Concert Hall	Convenience Store	Cosmetics Shop
50	Cosmetics Shop	Creperie	Deli / Bodega
51	Cosmetics Shop	Creperie	Deli / Bodega

	9th Most Common Venue	10th Most Common Venue
0	Diner	Yoga Studio
2	Diner	Yoga Studio
3	Diner	Yoga Studio
6	Donut Shop	Dog Run
7	Donut Shop	Dog Run
8	Donut Shop	Dog Run
9	Donut Shop	Dog Run
10	Donut Shop	Dog Run
11	Donut Shop	Dog Run
12	Cosmetics Shop	Creperie
13	Cosmetics Shop	Creperie
14	Cosmetics Shop	Creperie
16	Cosmetics Shop	Creperie
17	Cosmetics Shop	Creperie
18	Cosmetics Shop	Creperie
19	Diner	Donut Shop
20	Diner	Donut Shop
21	Deli / Bodega	Dessert Shop
22	Diner	Donut Shop
23	Diner	Donut Shop
25	Diner	Donut Shop
26	Diner	Donut Shop
28	Diner	Donut Shop
29	Diner	Donut Shop
31	Dive Bar	Yoga Studio
32	Dive Bar	Yoga Studio
33	Dive Bar	Yoga Studio
34	Dessert Shop	Diner
37	Yoga Studio	Farmers Market
38	Yoga Studio	Farmers Market
39	Dessert Shop	Diner
40	Dessert Shop	Diner

```
41          Dessert Shop                  Diner
42              Dive Bar         Deli / Bodega
44              Dive Bar         Deli / Bodega
45              Dive Bar         Deli / Bodega
46         Deli / Bodega          Dessert Shop
47         Deli / Bodega          Dessert Shop
48              Creperie         Deli / Bodega
49              Creperie         Deli / Bodega
50          Dessert Shop                  Diner
51          Dessert Shop                  Diner
```

[69]:
```
SC_merged.loc[SC_merged['Cluster Labels'] == 3,SC_merged.columns[[1] +␣
↪list(range(5, SC_merged.shape[1]))]]
```

[69]:
```
      City                                          location  \
30  Alviso  (Alviso, San Jose, Santa Clara County, Califor…

                             point   latitude   longitude  altitude  \
30  (37.426051, -121.9752373, 0.0)  37.426051 -121.975237       0.0

    Cluster Labels 1st Most Common Venue 2nd Most Common Venue  \
30               3           Music Venue           Yoga Studio

   3rd Most Common Venue 4th Most Common Venue 5th Most Common Venue  \
30    Falafel Restaurant          Concert Hall     Convenience Store

   6th Most Common Venue 7th Most Common Venue 8th Most Common Venue  \
30        Cosmetics Shop              Creperie         Deli / Bodega

   9th Most Common Venue 10th Most Common Venue
30          Dessert Shop                  Diner
```

[70]:
```
SC_merged.loc[SC_merged['Cluster Labels'] == 4,SC_merged.columns[[1] +␣
↪list(range(5, SC_merged.shape[1]))]]
```

[70]:
```
               City                                          location  \
1   Los Altos Hills  (Los Altos Hills, Santa Clara County, Californ…
4   Los Altos Hills  (Los Altos Hills, Santa Clara County, Californ…
35      Monte Vista  (Monte Vista, Alta, Placer County, California,…

                             point   latitude   longitude  altitude  \
1   (37.3796627, -122.1374637, 0.0)  37.379663 -122.137464       0.0
4   (37.3796627, -122.1374637, 0.0)  37.379663 -122.137464       0.0
35  (39.1865648, -120.8327162, 0.0)  39.186565 -120.832716       0.0

    Cluster Labels 1st Most Common Venue 2nd Most Common Venue  \
1                4    Spanish Restaurant             Bookstore
```

```
4                4    Spanish Restaurant            Bookstore
35               4    Spanish Restaurant          Yoga Studio

   3rd Most Common Venue 4th Most Common Venue 5th Most Common Venue  \
1             Yoga Studio    Falafel Restaurant          Concert Hall
4             Yoga Studio    Falafel Restaurant          Concert Hall
35          Cocktail Bar           Comedy Club          Concert Hall

   6th Most Common Venue 7th Most Common Venue 8th Most Common Venue  \
1      Convenience Store        Cosmetics Shop              Creperie
4      Convenience Store        Cosmetics Shop              Creperie
35     Convenience Store        Cosmetics Shop              Creperie

   9th Most Common Venue 10th Most Common Venue
1          Deli / Bodega           Dessert Shop
4          Deli / Bodega           Dessert Shop
35         Deli / Bodega           Dessert Shop
```

[ ]:
```
#end of notebook
```