COWORKING SPACE

A new collaborative space location in Los Angeles County, CA

The goal is to find a good place for a new COWORKING space in Los Angeles County, CA

The coworking industry is becoming a lucrative worldwide phenomenon, so now is the perfect time to break into the business.

There needs to be a niche for a new collaborative workspace business, and thankfully there are plenty. Perhaps the area has a high demand for coworking spaces beyond coffee shops and libraries.

Maybe the existing spaces in the area are too focused on one type of worker. On the other hand, maybe the other spaces are too broad.
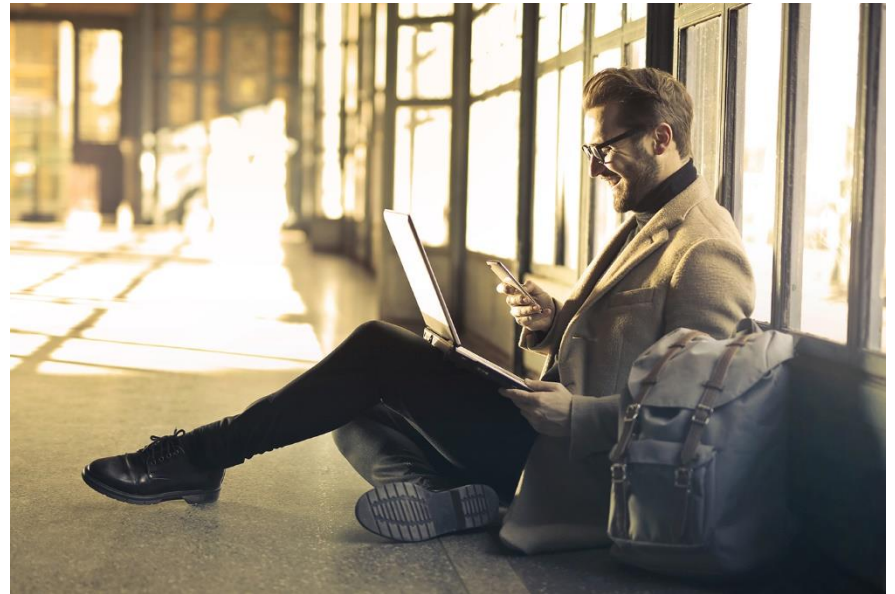
Exploring the data

	NEIGHBORHOOD	POPULATION PER SQMI	AVERAGE HOUSEHOLD SIZE	DIVERSITY INDEX	FOREIGN BORN POPULATION	MEDIAN INCOME	MEDIAN AGE
0	Koreatown	42611	2.7	0.602	68.0	30558.0	30.0
1	Westlake	38214	3.0	0.430	67.6	26757.0	27.0
2	East Hollywood	31095	3.0	0.578	66.5	29927.0	31.0
3	Pico-Union	25352	3.3	0.264	64.6	26424.0	27.0
4	Maywood	23638	4.1	0.069	55.2	41203.0	23.0

	POPULATION PER SQMI	AVERAGE HOUSEHOLD SIZE	DIVERSITY INDEX	FOREIGN BORN POPULATION	MEDIAN INCOME	MEDIAN AGE
count	265.000000	265.000000	265.000000	265.000000	265.000000	264.000000
mean	7571.471698	2.972075	0.492234	31.817358	68647.022642	33.178030
std	6925.987567	0.643715	0.160948	14.401015	31740.466505	5.942255
min	2.000000	1.500000	0.065000	4.400000	15003.000000	21.000000
25%	1601.000000	2.500000	0.380000	20.300000	48518.000000	28.000000
50%	6459.000000	2.900000	0.526000	31.800000	63039.000000	33.500000
75%	11266.000000	3.400000	0.628000	42.400000	81279.000000	37.000000
max	42611.000000	4.600000	0.755000	72.400000	207938.000000	48.000000

With info function we can see the types of values and NaNs

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 265 entries, 0 to 264
Data columns (total 7 columns):
NEIGHBORHOOD             265 non-null object
POPULATION PER SQMI     265 non-null int64
AVERAGE HOUSEHOLD SIZE  265 non-null float64
DIVERSITY INDEX         265 non-null float64
FOREIGN BORN POPULATION 265 non-null float64
MEDIAN INCOME           265 non-null float64
MEDIAN AGE              264 non-null float64
dtypes: float64(5), int64(1), object(1)
memory usage: 14.6+ KB
```

Here is the table with neighborhoods and key features

Methodology

In first step we have collected the required neighborhood data with features such as locations, population features, average income, diversity, age e.t.c. We will perform exploratory data analysis.

Second step in our analysis will be identifying currently operating Coworking Spaces (according to Foursquare categorization). We will use Folium maps to identify and analyze the Neighborhoods with Coworking Spaces information. We will try to find Correlations between neighborhood features.

In third step we will need data preparation Clustering for futere neighborhood segmentation. After, we will perform clustering using K-Means method with a certain number of clusters. Then, we will analyze the data based on cluster segmentation and compare it with the data of neighborhoods that have the highest number of currently operating Coworking Spaces.

In fourth step, by selecting important and meaningful information of features, we will narrow down the neighborhood list to only promising locations. We will be able evaluate our analysis by comparing promising location list the neighborhood list with currently operating Coworking Spaces.
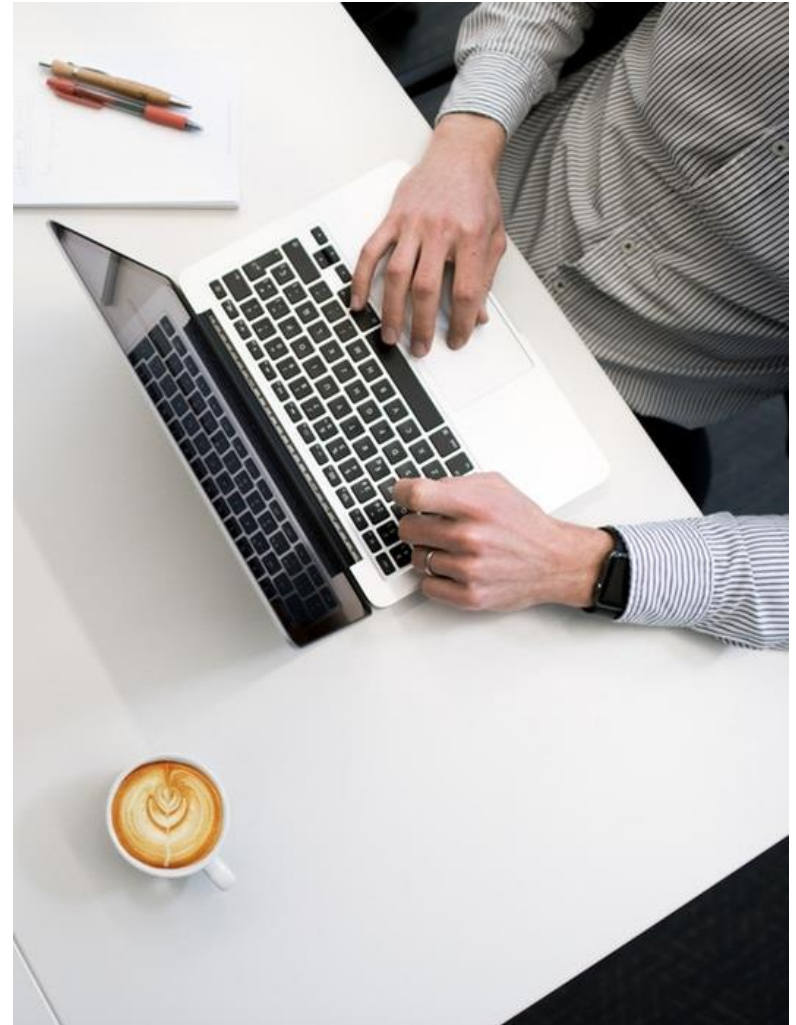
Current operating coworking places on a map

Plotting Coworking Space locations on a map to visualize for better understanding.
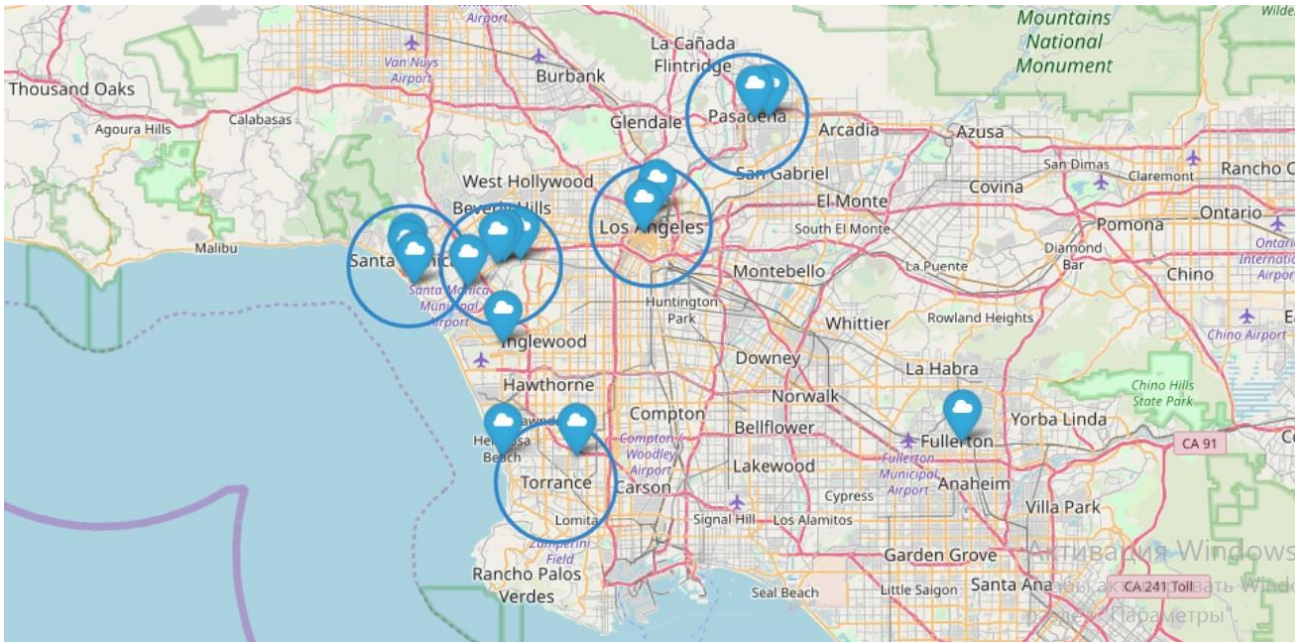
Analyzing Neighborhoods with Coworking Spaces information

Features Correlations

|	Unnamed: 0	POPULATION PER SQMI	AVERAGE HOUSEHOLD SIZE	DIVERSITY INDEX	FOREIGN BORN POPULATION	MEDIAN INCOME	MEDIAN AGE	CW_N
Unnamed: 0	1.000000	-0.941971	0.228393	0.532824	-0.188778	-0.043282	-0.074942	-0.039431
POPULATION PER SQMI	-0.941971	1.000000	-0.011151	-0.590618	0.060270	0.266449	-0.083081	0.130123
AVERAGE HOUSEHOLD SIZE	0.228393	-0.011151	1.000000	-0.048903	0.243529	0.097772	-0.395829	-0.233369
DIVERSITY INDEX	0.532824	-0.590618	-0.048903	1.000000	0.181470	-0.454514	0.430113	0.124096
FOREIGN BORN POPULATION	-0.188778	0.060270	0.243529	0.181470	1.000000	-0.874109	0.435546	-0.127935
MEDIAN INCOME	-0.043282	0.266449	0.097772	-0.454514	-0.874109	1.000000	-0.585762	0.067016
MEDIAN AGE	-0.074942	-0.083081	-0.395829	0.430113	0.435546	-0.585762	1.000000	0.412021
CW_N	-0.039431	0.130123	-0.233369	0.124096	-0.127935	0.067016	0.412021	1.000000

Analyzing Neighborhoods with Coworking Spaces information

Features distribution visualization
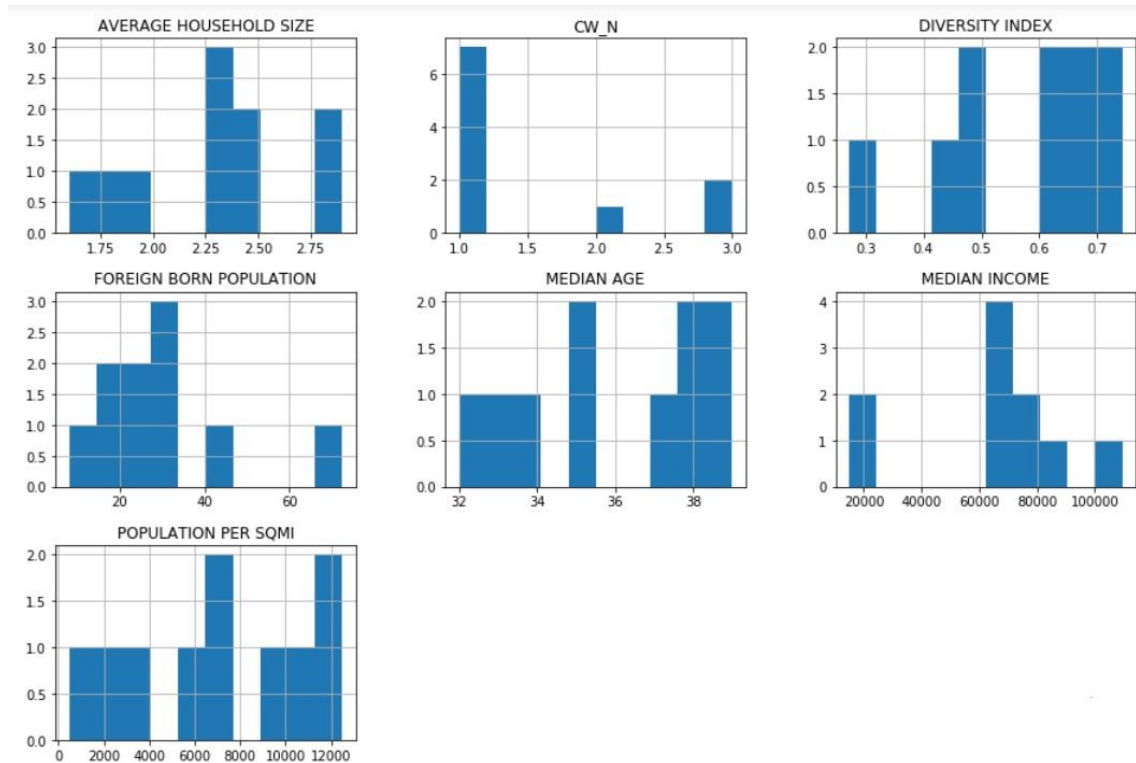
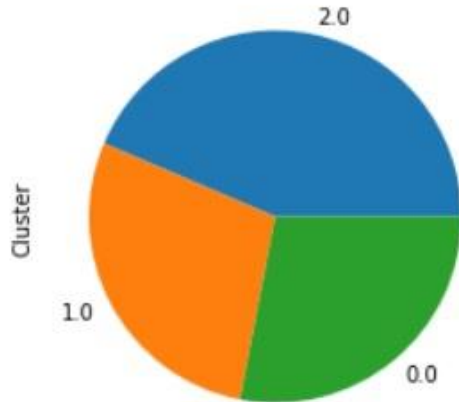# Visualizing neighborhoods based on median income

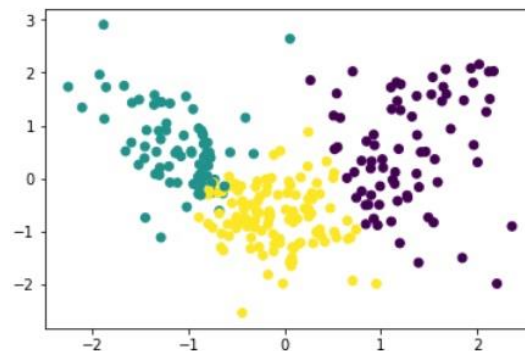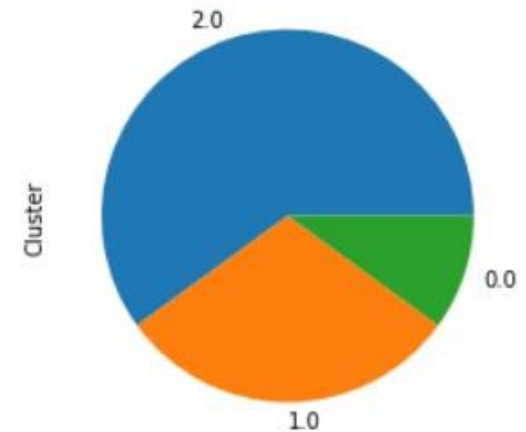Generating a choropleth map using the NEIGHBORHOOD and MEDIAN INCOME information.
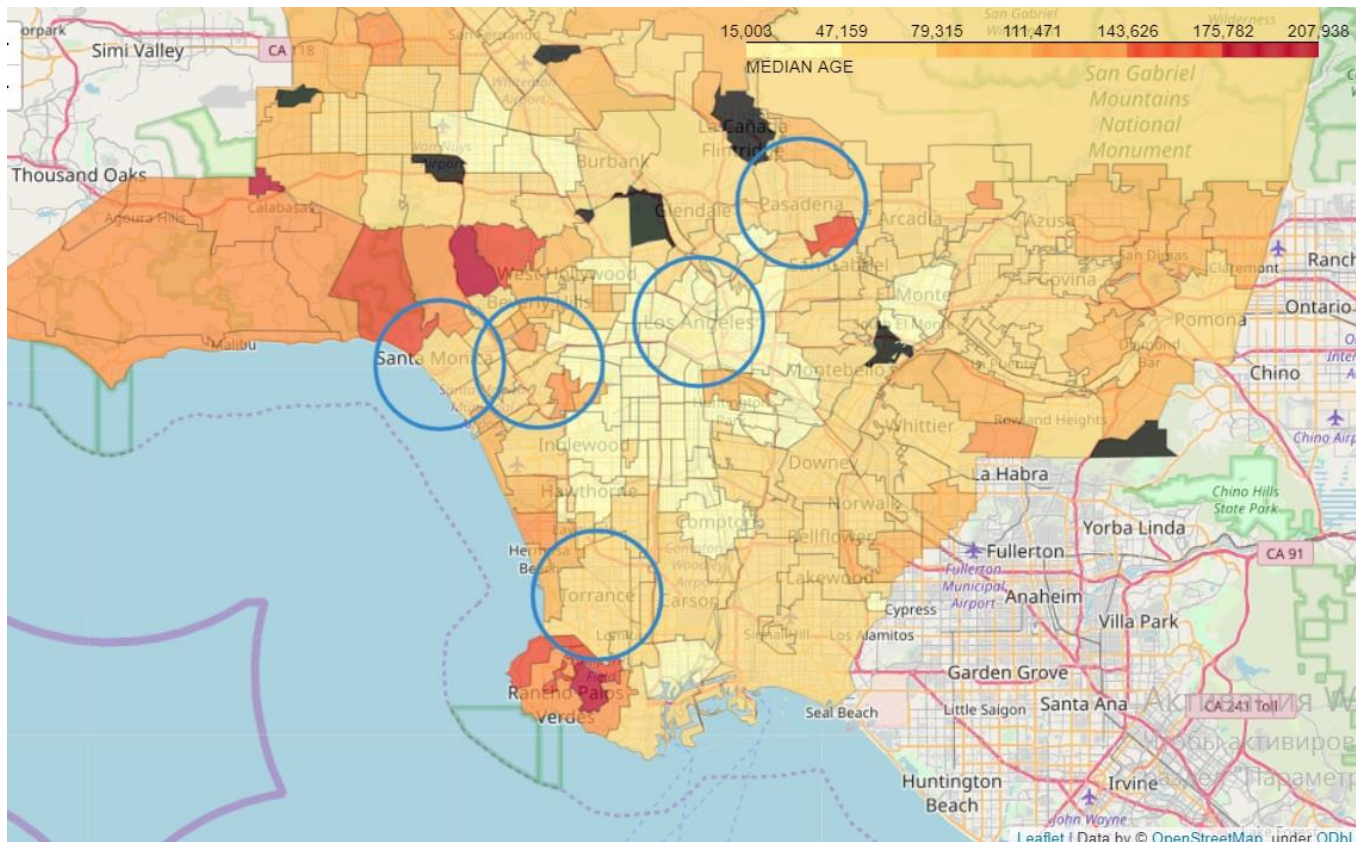
Clustering data

Let's perform clustering using K-Means method with number of clusters 3. Also, we will need to add a column with clusters labels to our dataset
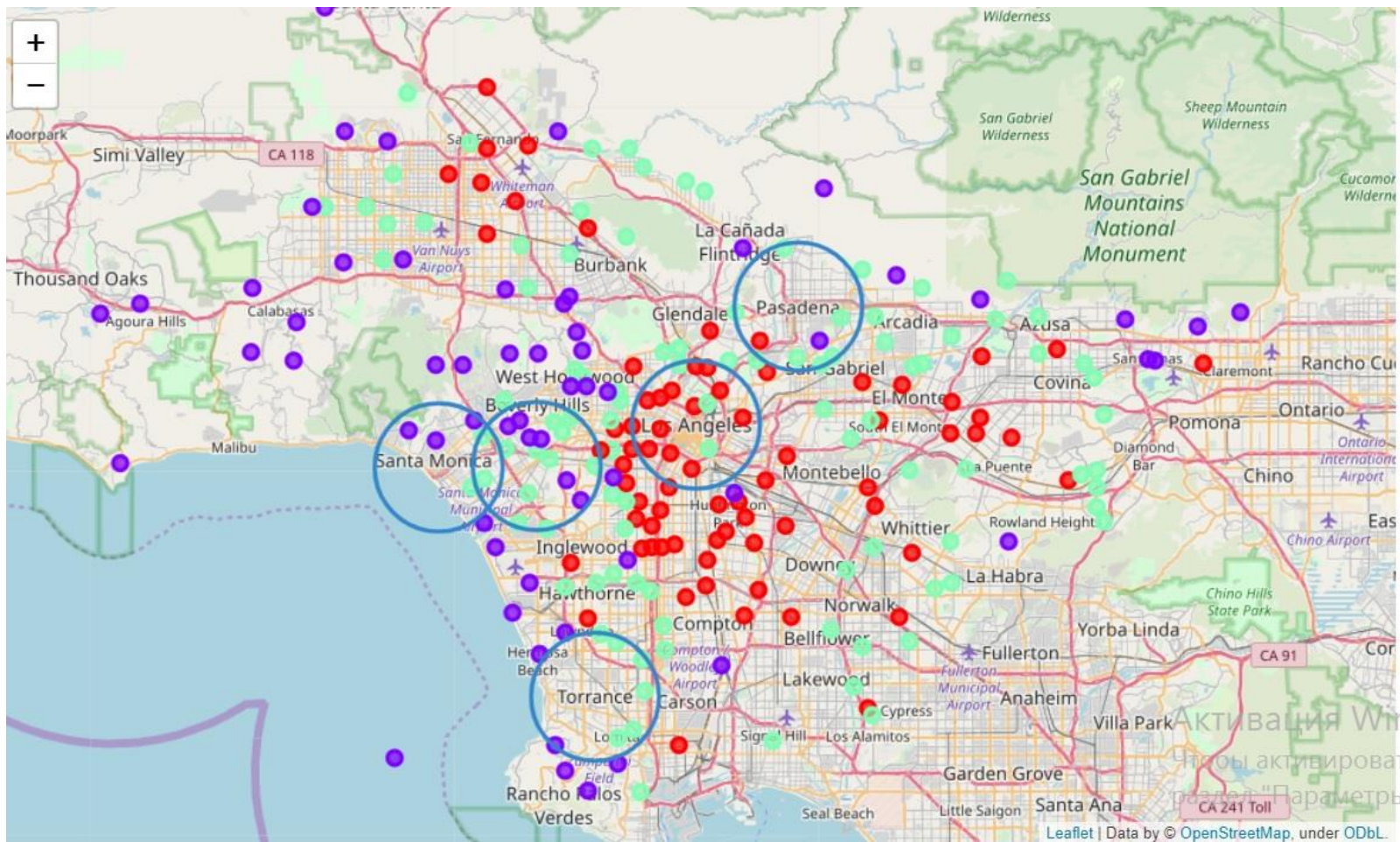
Visualizing clusters with pie plot.

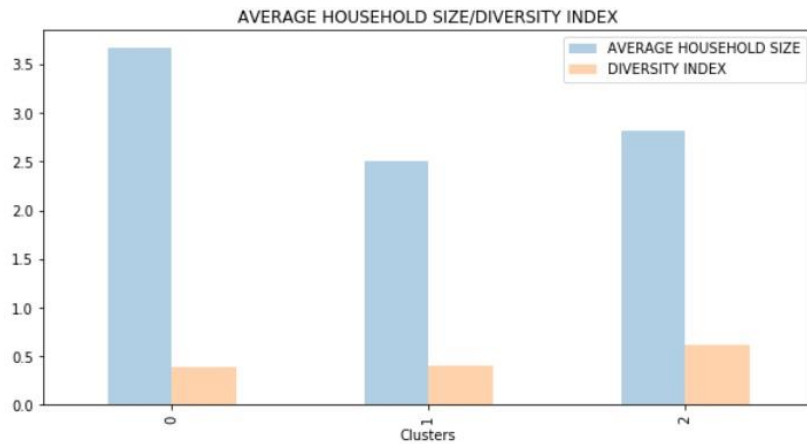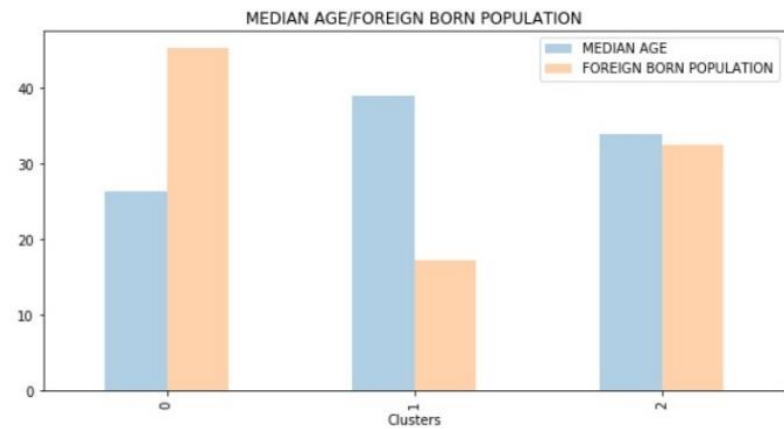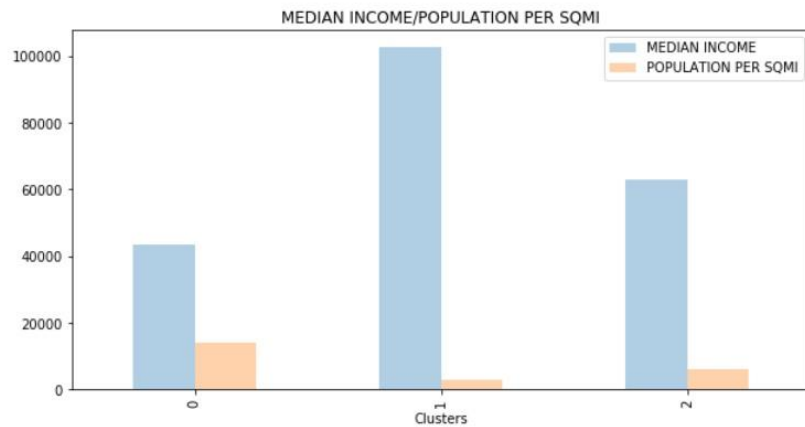Coworking Locations in Neighborhood Clusters
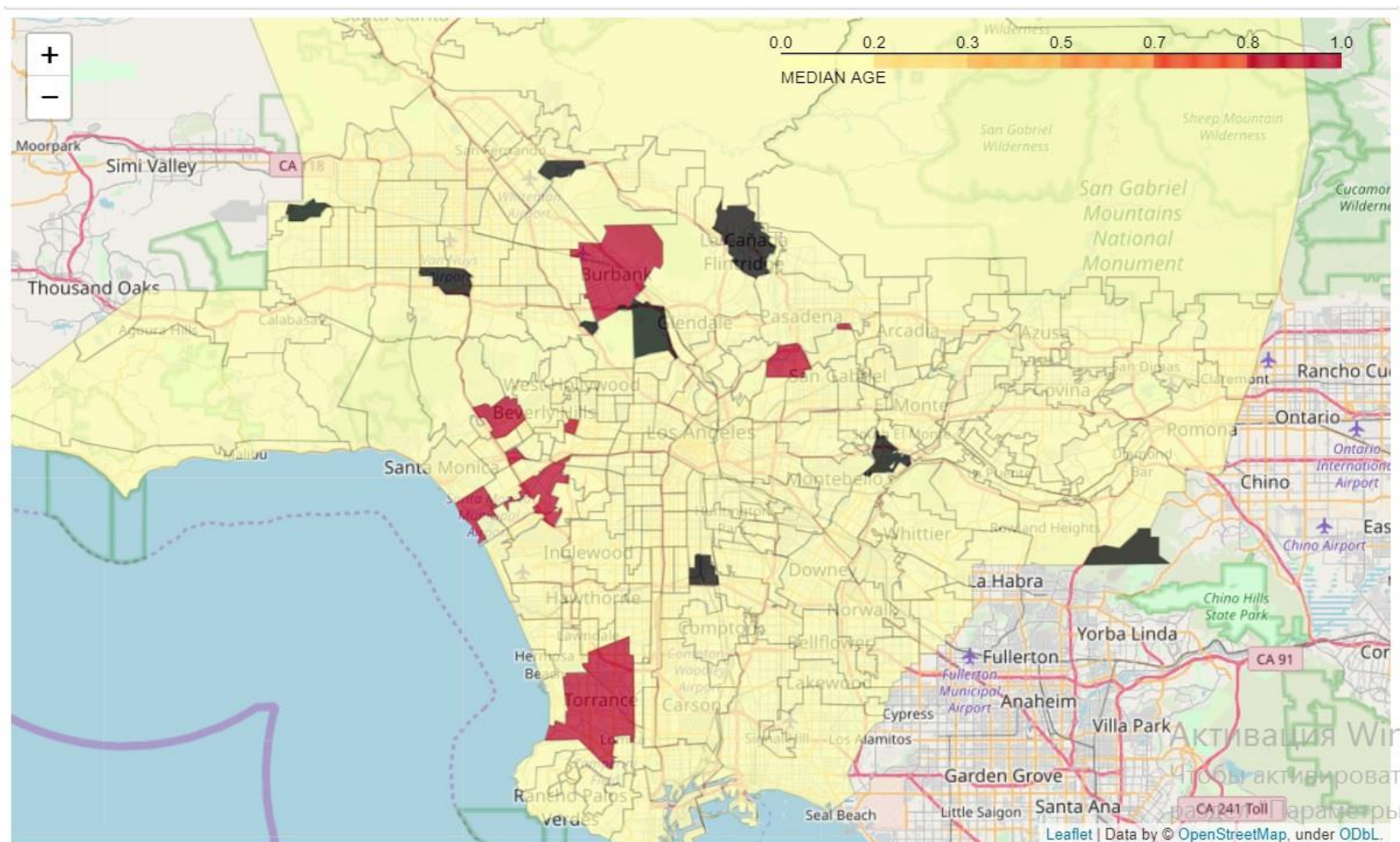

Visualize clustering decomposition

Plotting all clusters on the map

Features distribution by clusters

Plotting the Neighborhoods selected by Clusters and key features

Results

Here is the list of promising locations in Los Angeles County for a new Coworking Space. Some of them already have currently operating collaborative work space businesses, some of them not. The similarity of these neighborhoods features is high. The list has been narrowed down to top 10 promising locations.

	NEIGHBORHOOD	POPULATION PER SQMI	AVERAGE HOUSEHOLD SIZE	DIVERSITY INDEX	FOREIGN BORN POPULATION
0	Westwood	13036.0	2.0	0.543	31.3
1	Venice	11891.0	1.9	0.534	22.3
2	Carthay	9642.0	2.1	0.615	25.1
3	San Pasqual	8036.0	2.3	0.601	32.6
4	Culver City	7475.0	2.3	0.681	26.6
5	Rancho Park	7169.0	2.2	0.598	28.5
6	South Pasadena	7114.0	2.3	0.651	24.4
7	Torrance	6701.0	2.5	0.628	27.6
8	Burbank	5785.0	2.4	0.581	31.1

The main neighborhoods features are: High Median Income, Average household size, Foreign Born Population, Diversity, and Density. The Age Median value and Population Density will require extra attention.

Discussion

High Median Income neighborhood feature value should be higher than average or in our case more than 62000.

Average household size correlates very well Median Income and the suggested value should be less than 2.6.

Population Diversity Index value should be set to at least higher than average.

There Foreign Born Population feature value should be set as low.

The median Age feature can be set higher in some cases. However, it requires a specific analysis for university campuses with the high number of foreign born people, high density and low median age.
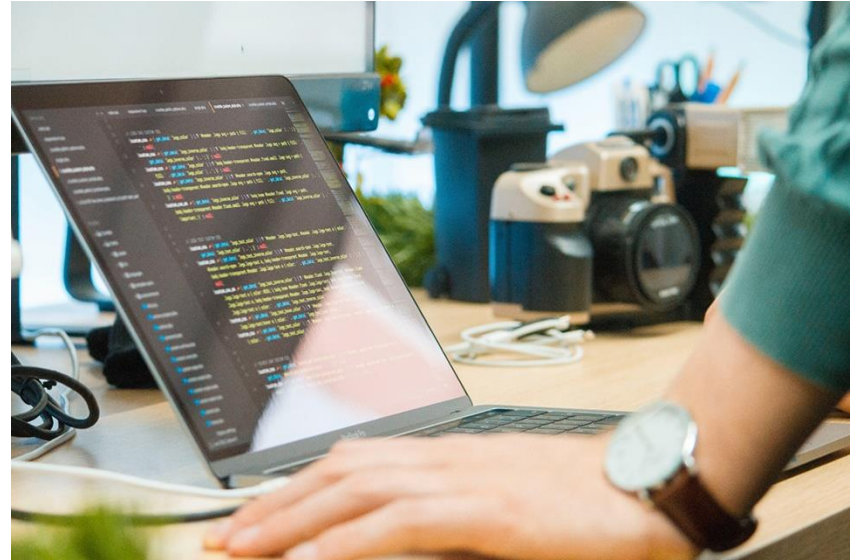
The Population per SqMi feature may cause false positive conclusions. The best suggestion would to narrow down the neighborhood list by Median Income first and eliminate the neighborhoods with low median income that may have a high population density. Then from the narrowed data we can drop wealthy suburbs that may have a low population density. It would be a good idea to compare this information with the Average household size.

Conclusion

The main purpose of this project was to identify the best neighborhoods to open a new Coworking Space. We have created a list of the most promising locations for this business by collecting the data from multiple sources, providing exploratory data analysis, data cleaning and preparation, performed clustering segmentation using K-Means method, analyzed important neighborhood features and provided suggested values for them. There is a room for a deeper analysis using other methods and adding new information data.

Final decision on optimal Coworking Space location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone. Taking into consideration additional factors.