

Investing Restaurant in Los Angeles

Yuhai Li

Business Problem

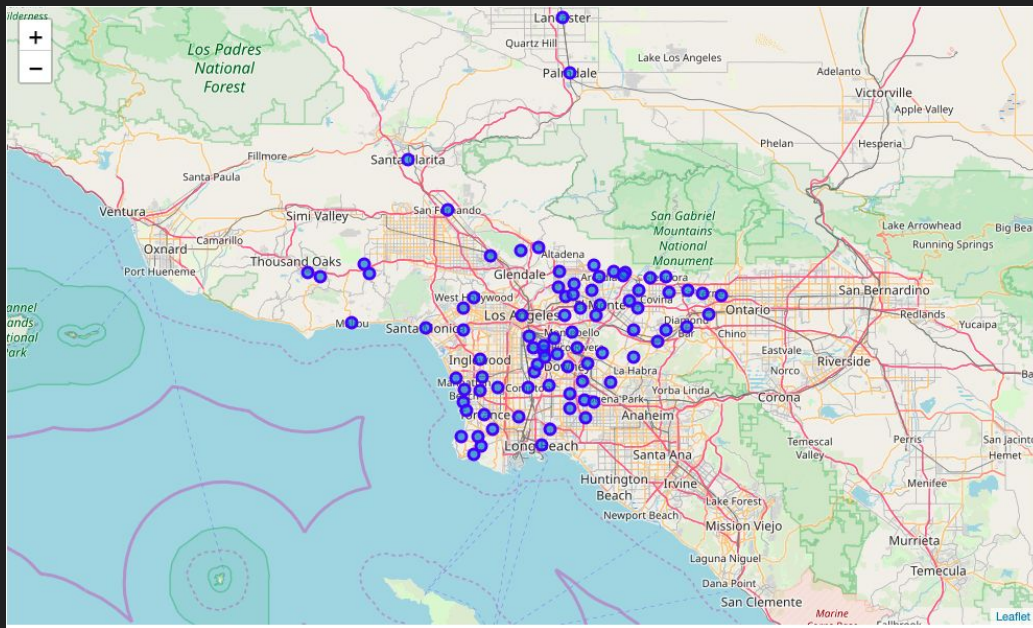
An investor would like to open a new restaurant in Los Angeles. Unfortunately, the investor doesn't know Los Angeles very well and actually has no idea about where to locate the restaurant. Whether it is a good idea to open it in downtown or some specific districts.

Data

- A list of cities in Los Angeles County from Wikipedia:
https://en.wikipedia.org/wiki/List_of_cities_in_Los_Angeles_County,_California
- Foursquare location data to take info about venues from each district. Then I'll be able to make segmentation and profiling of districts and look at places best for opening new restaurant

Methodology

1. I have names of cities in Los Angeles and put them on a Warsaw map:



Methodology

2. Having data obtained from Foursquare API (list of nearby venues with their categories), I've used mainly Pandas to prepare data - wrangle them, join, make one-hot encoding:

	Neighborhood	ATM	Accessories Store	African Restaurant	Airport	American Restaurant	Argentinian Restaurant	Art Gallery	Arts & Crafts Store	Arts & Entertainment	Asian Restaurant	Athletics & Sports	Australian Restaurant	Dea
0	Agoura Hills	0.000000	0.0	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.000000	0.0	
1	Alhambra	0.000000	0.0	0.0	0.0	0.000000	0.0	0.033333	0.0	0.0	0.000000	0.000000	0.0	
2	Arcadia	0.000000	0.0	0.0	0.0	0.058824	0.0	0.000000	0.0	0.0	0.000000	0.058824	0.0	
3	Artesia	0.000000	0.0	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.033333	0.000000	0.0	
4	Avalon	0.033333	0.0	0.0	0.0	0.033333	0.0	0.000000	0.0	0.0	0.000000	0.000000	0.0	
5	Azusa	0.000000	0.0	0.0	0.0	0.033333	0.0	0.000000	0.0	0.0	0.000000	0.000000	0.0	
6	Baldwin Park	0.000000	0.0	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.066667	0.000000	0.0	
7	Bell	0.038462	0.0	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.000000	0.0	
8	Bell Gardens	0.000000	0.0	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.000000	0.0	
9	Bellflower	0.000000	0.0	0.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.000000	0.000000	0.0	

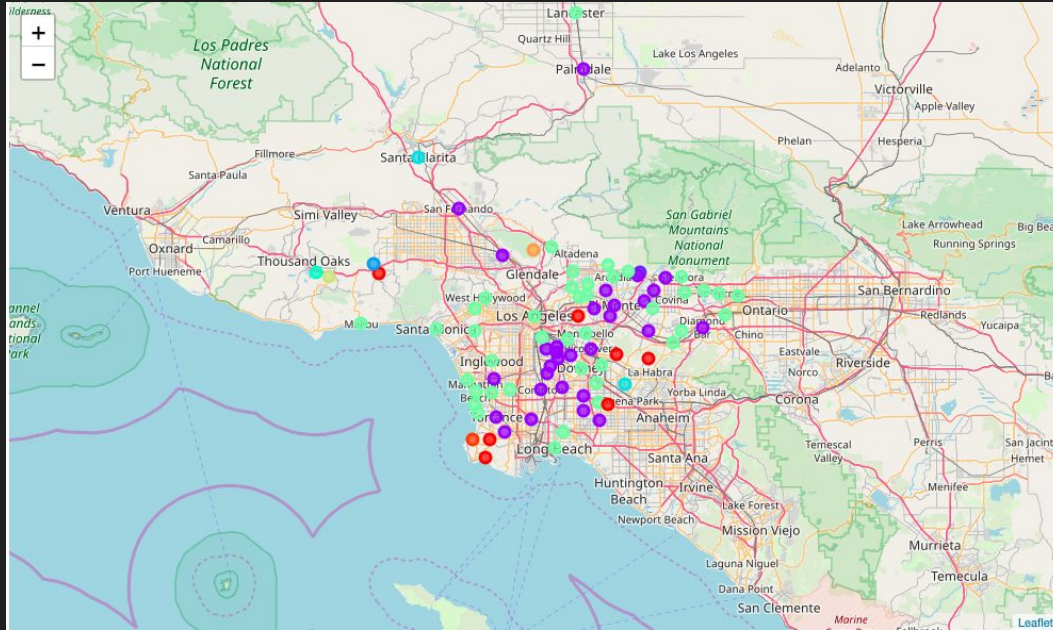
Methodology

3. Finding most common venues in each city:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Agoura Hills	Locksmith	Thai Restaurant	Home Service	Yoga Studio	Financial or Legal Service	Farmers Market	Fast Food Restaurant	Field	Filipino Restaurant	Fish & Chips Shop
1	Alhambra	Bakery	Café	Seafood Restaurant	Burger Joint	Thai Restaurant	Gastropub	Sushi Restaurant	Bookstore	Lounge	Market
2	Arcadia	Food Truck	Optical Shop	Racetrack	Athletics & Sports	Restaurant	Fountain	American Restaurant	Track	Park	Night Market
3	Artesia	Indian Restaurant	Chinese Restaurant	Sandwich Place	Bubble Tea Shop	Thai Restaurant	Hotel	Salon / Barbershop	Filipino Restaurant	Bank	Dessert Shop
4	Avalon	Hotel	Seafood Restaurant	Mexican Restaurant	Pizza Place	ATM	Café	Boat or Ferry	Brewery	Burger Joint	Candy Store
5	Azusa	Mexican Restaurant	Pizza Place	Liquor Store	Big Box Store	Coffee Shop	Bank	Burger Joint	Bakery	Mobile Phone Shop	Mediterranean Restaurant
6	Baldwin Park	Mexican Restaurant	Discount Store	Burger Joint	Bank	Pharmacy	Pizza Place	Asian Restaurant	Fast Food Restaurant	Buffet	Men's Store
7	Bell	Mexican Restaurant	Pizza Place	ATM	Diner	Burger Joint	Chinese Restaurant	Shoe Store	Sandwich Place	Coffee Shop	Convenience Store
8	Bell Gardens	Mexican Restaurant	Burger Joint	Park	Sporting Goods Shop	Fried Chicken Joint	New American Restaurant	Fast Food Restaurant	Liquor Store	Coffee Shop	Grocery Store
9	Bellflower	Trail	Sandwich Place	Concert Hall	Yoga Studio	Filipino Restaurant	Farm	Farmers Market	Fast Food Restaurant	Field	Financial or Legal Service

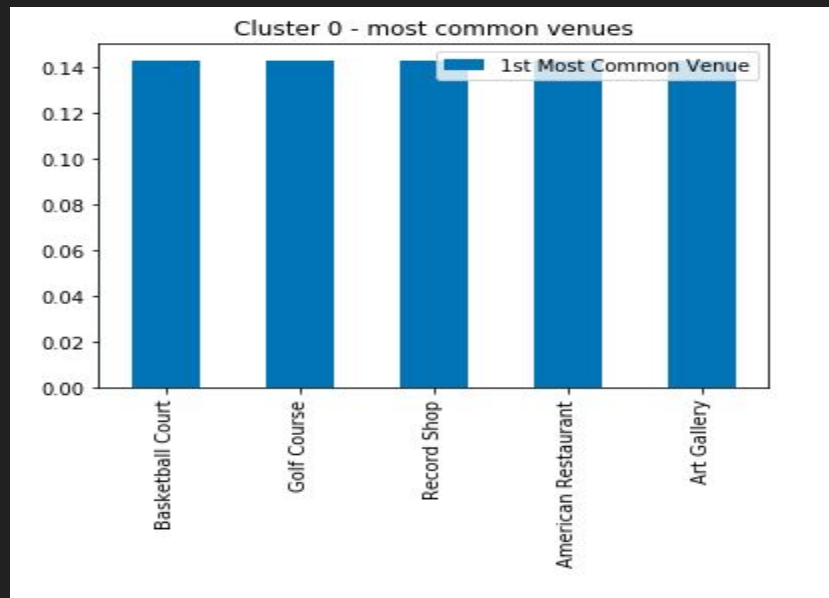
Methodology

4. After that preprocessing stage I've used K-means algorithm from scikit to find similar clusters similar to each other inside clusters, and as much as possible different among clusters:



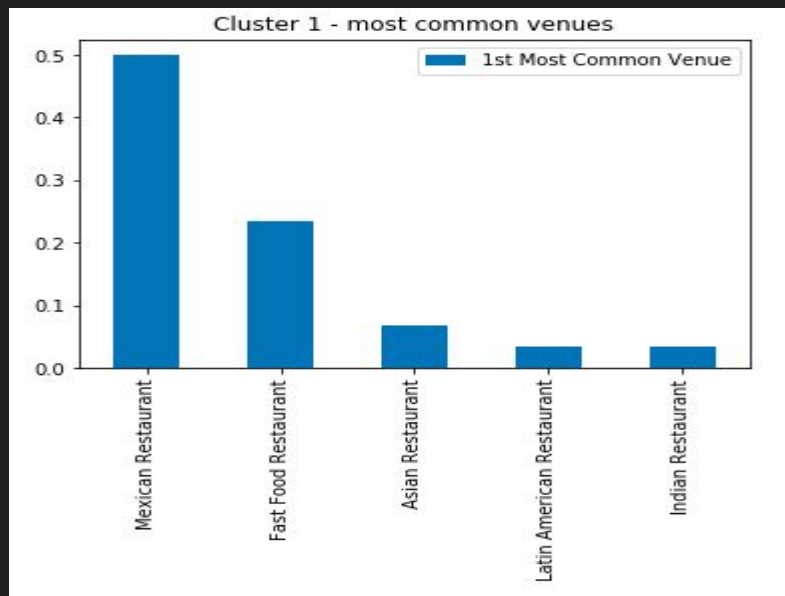
Results

Cluster 0 - There are a lot of people. It is good for American food.



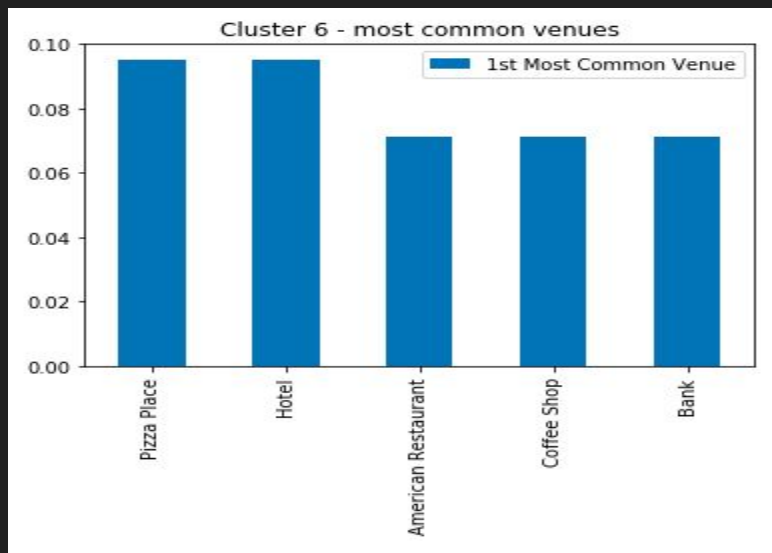
Results

Cluster 1 - Good place for restaurant. Top 3 are Mexican food, fast food and Asian food.



Results

Cluster 6 - Hotel zone. Good for fast food, like pizza and cafe, or American food.



Observations

Perhaps the most attracting place to open a restaurant is cluster 1.

However, it may also be a good idea to invest on restaurant in cluster 6 which is hotel zone because there are a large number of people move through every day.

Conclusion

The analysis shows that clustering districts look reasonable (I know Warsaw), so we can conduct similar analysis for other cities all over the world without knowing them and in automatic manner find some similar clusters, profile them and take decisions as a investor.