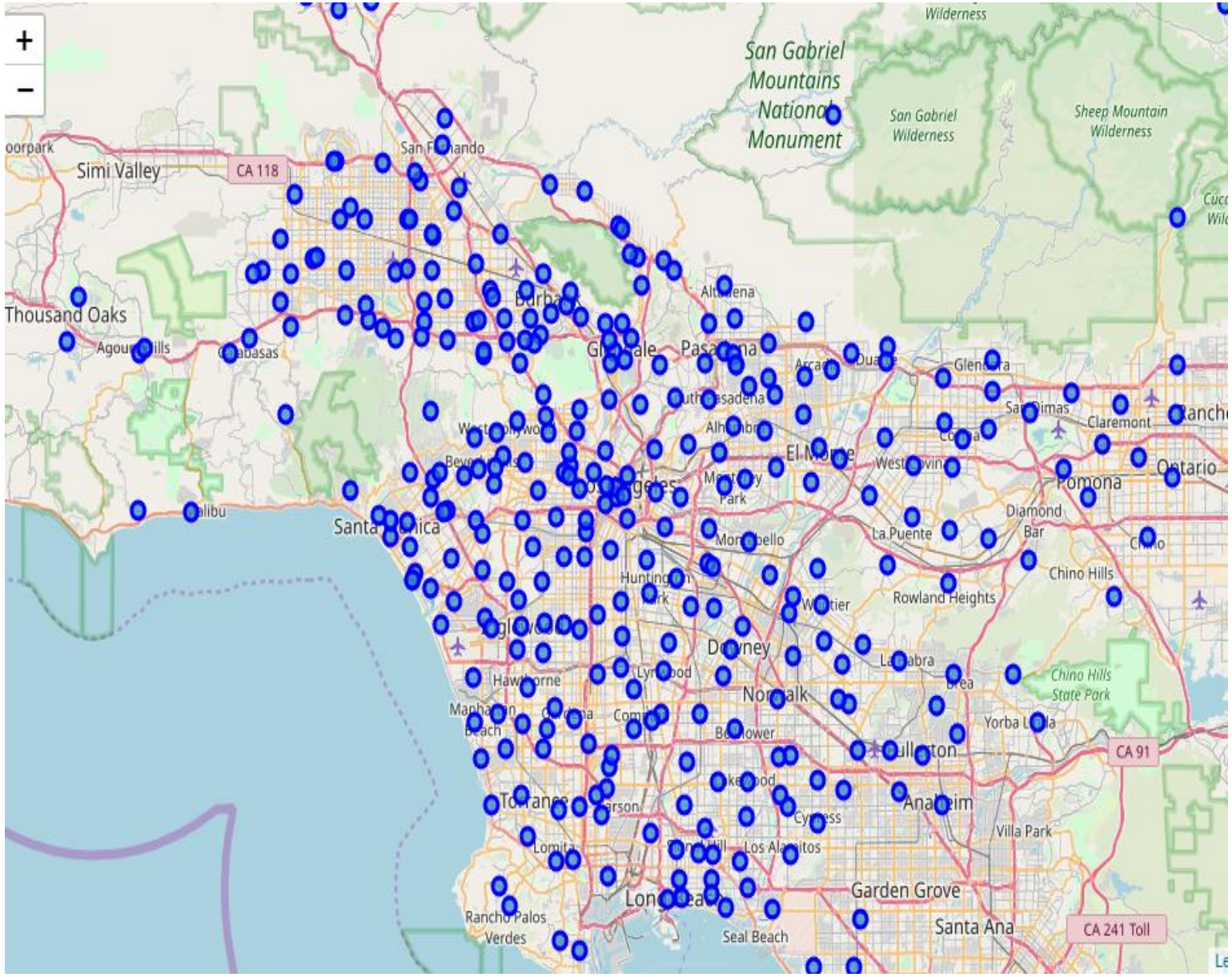
The battle of neighborhoods

Explore restaurant in LA County



Introduction

- A multi cuisine restaurant chain owner is looking for opening a restaurant in Los Angeles county
- **Our goal is to find appropriate location for the Restaurant**

Business problem

- The food market in LAis highly competitive
- Need a thorough market analysis
- Looking for a place with a highest demand and low competition

Competitors

- Fast food (hot dogs, bagels, ice cream, burgers, pizza, etc.)
- Mexican restaurants
- American Restaurants
- Asian restaurants (Thai, Chinese, Korean, Indian etc.)
- Coffee shops
- Middle Eastern restaurants

Factors to study

- All cities in Los Angeles county
- Segmentation of neighborhoods (by cities and zip code)

The objective of this project is to deliver a recommendation of which neighborhood of LA will be the best choice to build the restaurant

Data description

Based on definition of our problem, factors that will influence our decission are:

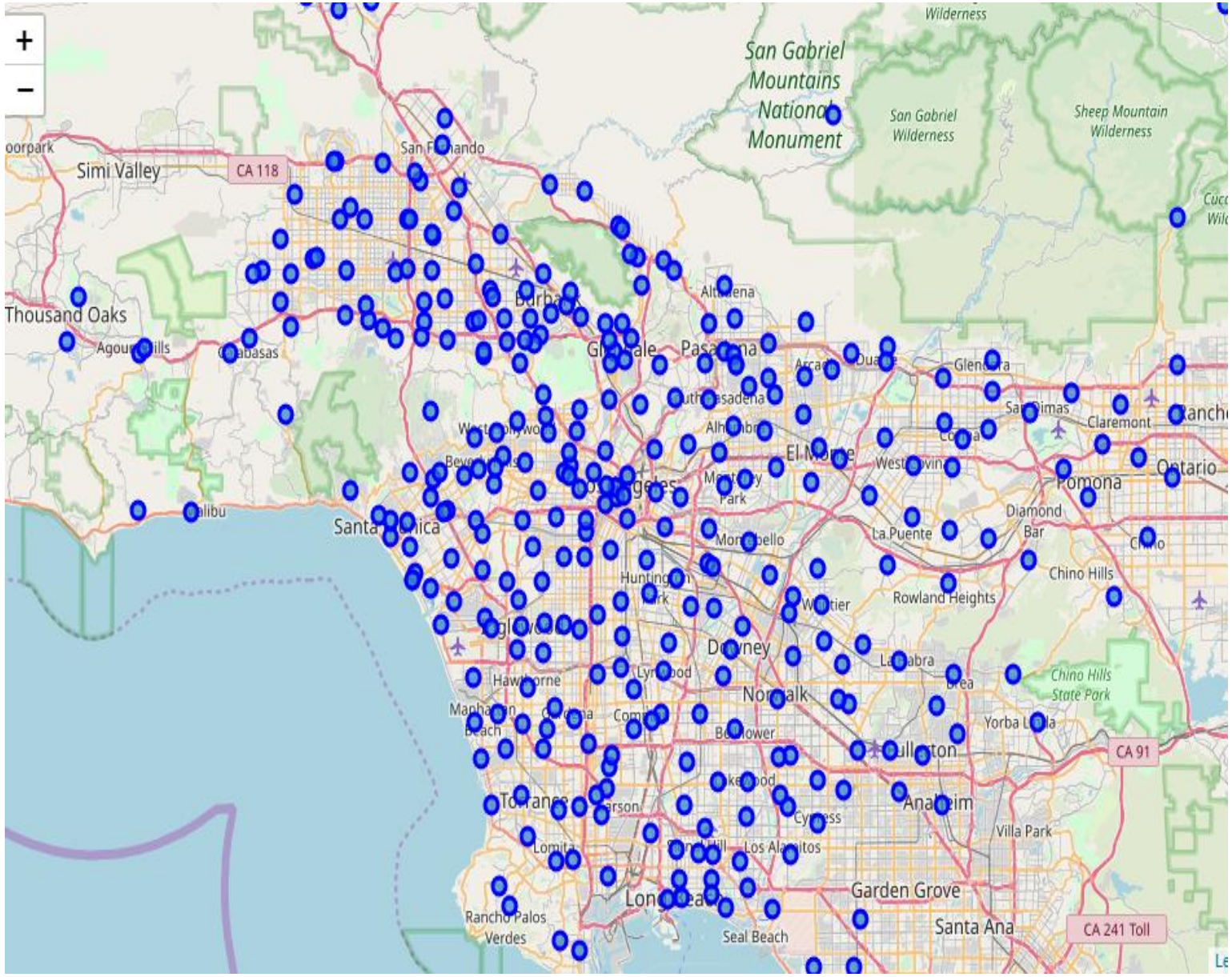
- number of existing restaurants in the neighborhood (any type of restaurant)
- number of and distance to specific restaurant type in the neighborhood, if any
- We decided to use regularly spaced grid of locations, centered around a zip code, to define our neighborhoods.
- Following data sources will be needed to extract/generate the required information:
  - number of restaurants and their type and location in every neighborhood will be obtained using **Foursquare API**
  - the region vs city mapping will be done using web scraping from a web page published by LA Times (http://maps.latimes.com/neighborhoods/neighborhood/list/)
  - coordinate of LA will be obtained using some open data sources made available by the city of Los Angeles (ZIP_Codes_and_Postal_Cities.csv)

Methodology

- In this project we will direct our efforts on detecting what type of restaurants are available in each areas of LA County that have low restaurant density. We will limit our analysis to area 500 mtr around each zip code in LA County. In first step we have collected the required data: location and type (category) of every restaurant in LA County. We have to filter the data from foursqare to select only Restaurants / Food joints.
- Second step in our analysis will be calculation and exploration of 'restaurant density' across different areas of LA
- In third and final step we will focus on most promising areas and within those create clusters of locations that meet some basic requirements established in discussion with stakeholders. We will present map of all such locations but also create clusters (using k-means clustering) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

Analytic approach

- 1. We load and explore the data from ZIP_Codes_and_Postal_Cities.csv file from LA County website and did a web scrapping from http://maps.latimes.com/neighborhoods/neighborhood/list/.
- 2. Transform the data into a pandas dataframe.
- 3. This dataframe contains the geographical coordinates of Los Angeles county neighborhoods.
- 4. This data will be used to get Venues data from Foursquare with its API.
- 5. We used Geopy and Folium libraries to create a map of Los Angeles county with neighborhoods



Methodology

Los Angeles visualization

Zip_Code	Afghan Restaurant	African Restaurant	American Restaurant	Argentinian Restaurant	Asian Restaurant	Australian Restaurant	Bakery	Bar	Beer Bar	Bistro	Brazilian Restaurant
90001	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
90003	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
90004	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
90005	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
90006	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
90007	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
90008	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.100000	0.000000	0.000000	0.000000	0.000000
90009	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
90010	0.000000	0.000000	0.000000	0.000000	0.055556	0.000000	0.055556	0.000000	0.027778	0.000000	0.055556
90011	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
90012	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.064516	0.000000	0.000000	0.000000	0.000000
90013	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.269231	0.000000	0.000000	0.000000
90014	0.000000	0.000000	0.037037	0.000000	0.000000	0.000000	0.000000	0.148148	0.000000	0.000000	0.000000
90015	0.000000	0.000000	0.058824	0.000000	0.000000	0.000000	0.029412	0.088235	0.029412	0.000000	0.000000
90016	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
90017	0.000000	0.000000	0.041667	0.000000	0.000000	0.000000	0.000000	0.041667	0.000000	0.000000	0.000000
90018	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.500000	0.000000	0.000000	0.000000	0.000000
90019	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
90020	0.000000	0.000000	0.000000	0.000000	0.034483	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
90021	0.000000	0.000000	0.000000	0.000000	0.076923	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
90022	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000

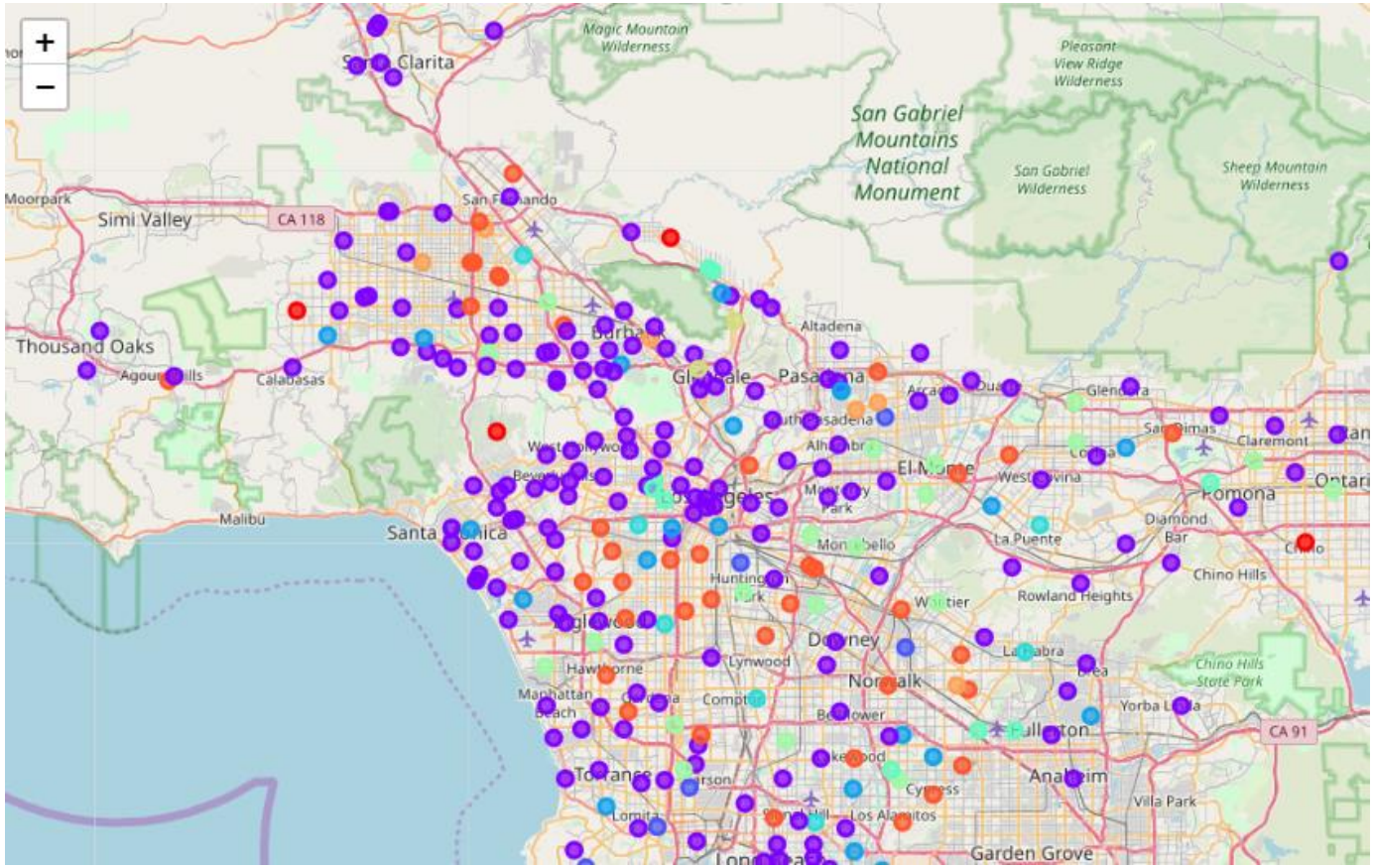
```
----90006----
                 venue  freq
0     Korean Restaurant  0.67
1           Pizza Place  0.08
2  Fast Food Restaurant  0.08
3           Coffee Shop  0.08
4    Chinese Restaurant  0.08


----90007----
                    venue  freq
0             Coffee Shop  0.27
1              Food Truck  0.18
2                   Café  0.09
3     Caribbean Restaurant  0.09
4  Mediterranean Restaurant  0.09
```

Aggregated the restaurant type frequency per neighborhood by Zip Code

Zip_Code	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
90001	Pizza Place	Donut Shop	Fast Food Restaurant	Mexican Restaurant	Donburi Restaurant	Dongbei Restaurant	Dumpling Restaurant	Eastern European Restaurant	English Restaurant	Ethiopian Restaurant
90003	Pizza Place	Fast Food Restaurant	Taco Place	Southern / Soul Food Restaurant	Fish Market	Doner Restaurant	Dongbei Restaurant	Donut Shop	Dumpling Restaurant	Eastern European Restaurant
90004	Cocktail Bar	Mexican Restaurant	Pizza Place	Sushi Restaurant	Fish Market	Dongbei Restaurant	Donut Shop	Dumpling Restaurant	Eastern European Restaurant	English Restaurant
90005	Korean Restaurant	Restaurant	Karaoke Bar	Café	Sushi Restaurant	Mexican Restaurant	Japanese Restaurant	Steakhouse	Noodle House	Halal Restaurant
90006	Korean Restaurant	Pizza Place	Fast Food Restaurant	Coffee Shop	Chinese Restaurant	Food	Donut Shop	Dumpling Restaurant	Eastern European Restaurant	English Restaurant
90007	Coffee Shop	Food Truck	Korean Restaurant	Caribbean Restaurant	Italian Restaurant	Mexican Restaurant	Mediterranean Restaurant	Café	Eastern European Restaurant	English Restaurant
90008	Fast Food Restaurant	Indian Restaurant	Chinese Restaurant	New American Restaurant	Pizza Place	Bakery	Southern / Soul Food Restaurant	Seafood Restaurant	Mexican Restaurant	Ethiopian Restaurant

Find the most popular restaurant in each neighborhood



Exploring the neighborhoods in LA County

- Los Angeles geographical coordinates data has to be utilized as input for the Foursquare API, that has been leveraged to provision venues information for each neighborhood. We used the Foursquare API data to explore neighborhoods in Los Angeles County.

Now create 10 different cluster

Results

- Our analysis shows that although there is a great number of restaurants in Los Angeles County , there are pockets of low restaurant density towwards the north and eastern parts of the county. Highest concentration of restaurants was detected in Los Angeles (downtown) and southern part.
- We were able to distinctly identify the clusters where certain kinds of cuisines are popular predominantly (for example for Mexican, Korean, Bakery). This will help us determine some location to open a certain kind of cuisine. Of course, it does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas within LA but not crowded with existing restaurants Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.