Categorical Analysis of Restaurants in LA Downtown and their Online Delivery System

Roopashree Bhat

Background and Objective

Due to increasing population, traffic and drastic increase in demand for online food delivery both commercial as well as residential areas, the focus of this project is:

- To analyze all restaurant venues and extract their major delivery providers if any, to get information on major competitors and other potential information which may be beneficial for future deliver providers

- To analyze and categorize different cuisines which might provide valuable input for online delivery providers and to new startups planning to open new restaurants

Data Acquisition

- Geo spatial data of Los Angeles county was provided by datasets of LA county in this link http://boundaries.latimes.com/set/la-county-neighborhoods-v5/ in form a JSON file

- JSON file was extracted and converted into Data frame Information such as region, Neighborhood and their respective coordinates was extracted

- Dataframe is further filtered to get data corresponding to central-LA region and within this only 4 main Downtown Neighborhoods namely Chinatown, Downtown, Pico-Union, Westlake are chosen

- Foursquare location API was used to process the above neighborhoods to get various restaurant venues. The obtained JSON response was processed to get delivery provider information and cuisine category
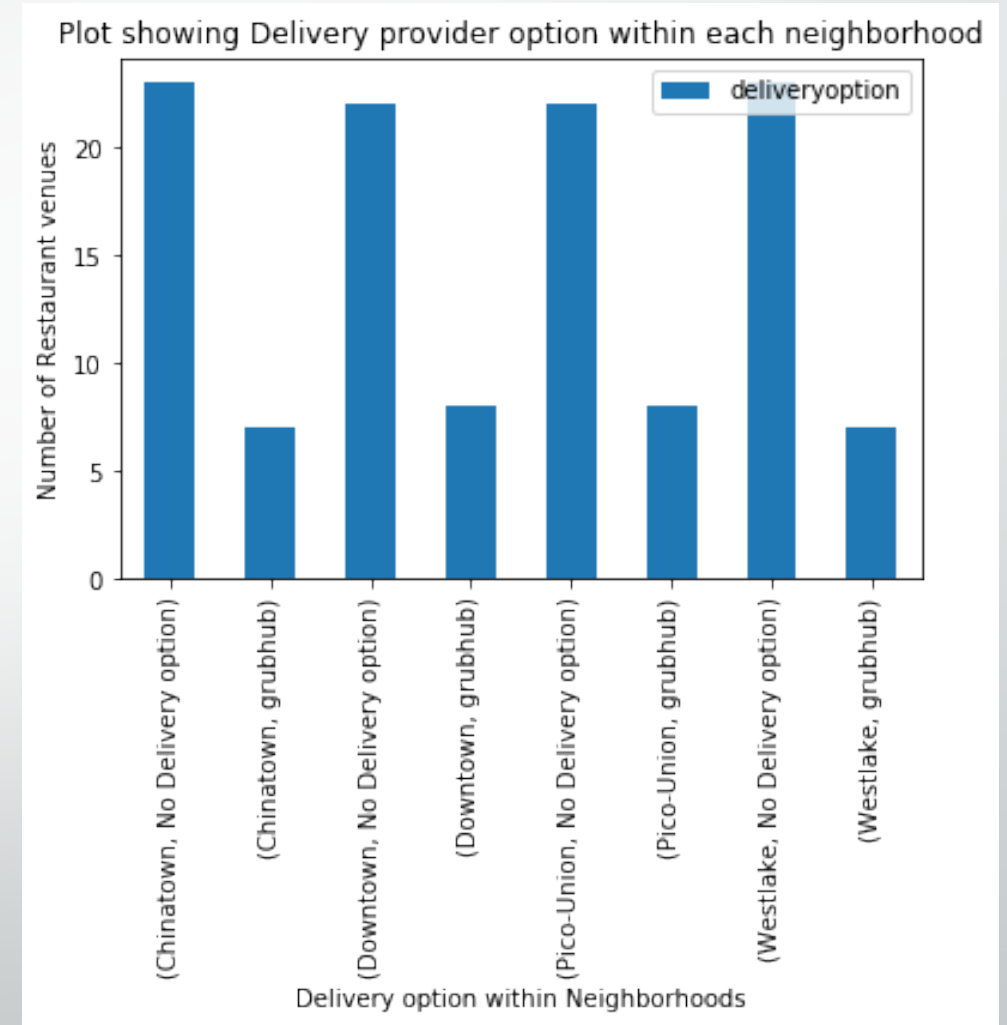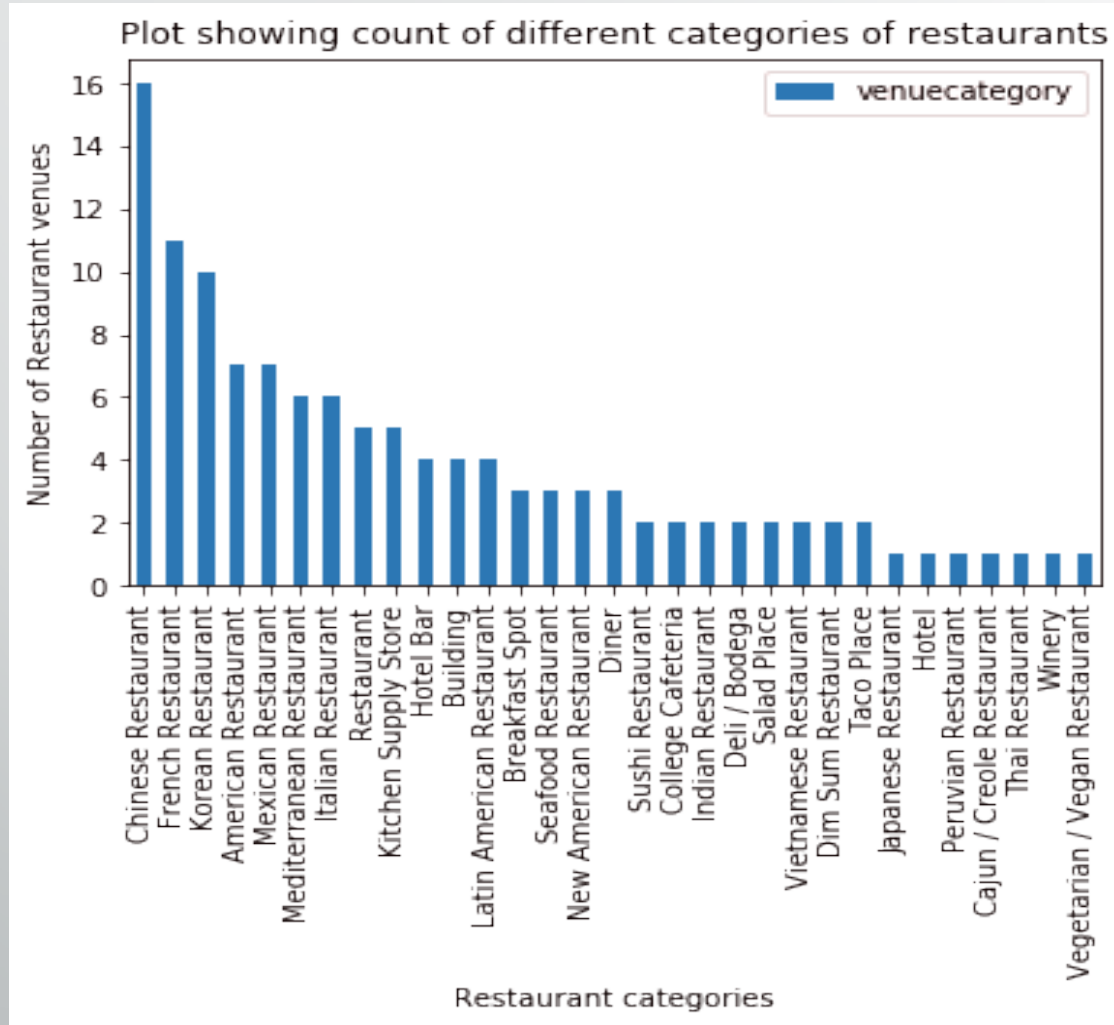
Dataframe Structure (Sample)

	Neighborhood	Latitude	Longitude	Venueid	venuename	venueLatitude	venueLongitude	venuecategory	deliveryoption
0	Arlington Heights	34.0525	-118.32	52c77412498ee9acc1e90889	Liwan Restaurant & Hookah Lounge	34.05035	-118.437779	Mediterranean Restaurant	grubhub
1	Arlington Heights	34.0525	-118.32	4c00d9439cf52d7f636714e7	Hae Jang Chon Korean BBQ Restaurant	34.06389	-118.306075	Korean Restaurant	No Delivery option
2	Arlington Heights	34.0525	-118.32	4aff640ef964a520173822e3	Young Kyung Restaurant	34.05246	-118.30332	Chinese Restaurant	No Delivery option
3	Arlington Heights	34.0525	-118.32	4c7435bd2db5236a2eadb979	All Family Restaurant	34.05367	-118.323318	Korean Restaurant	No Delivery option

Preprocessing and Data Analysis

- All column data is checked for Null values and irreverent data is dropped

- Categorical data is converted into numeric using one-hot encoding

- Restaurant Categories are grouped and analyzed with respect to total count of venues

- Delivery Provider option within each neighborhood is also grouped and analyzed

Data Analysis Bar Plot

Clustering of Restaurants

- K-Means clustering was employed to cluster into 10 clusters
- Merged Dataframe including Cluster labels

	Neighborhood	Latitude	Longitude	Venueid	venuename	venueLatitude	venueLongitude	venuecategory	deliveryoption	labels
0	Chinatown	34.067573	-118.224722	49ebc74af964a5202b671fe3	Yang Chow Restaurant	34.062926	-118.238059	Chinese Restaurant	No Delivery option	4
1	Chinatown	34.067573	-118.224722	4415b1c5f964a520fd301fe3	Full House Seafood Restaurant	34.066155	-118.237916	Chinese Restaurant	No Delivery option	4
2	Chinatown	34.067573	-118.224722	52c77412498ee9acc1e90889	Liwan Restaurant & Hookah Lounge	34.050354	-118.437779	Mediterranean Restaurant	grubhub	1
3	Chinatown	34.067573	-118.224722	49c2818cf964a520f9551fe3	Boda Restaurant	34.075601	-118.217365	Vietnamese Restaurant	No Delivery option	0
4	Chinatown	34.067573	-118.224722	4aa8416bf964a5206c5020e3	Traxx Restaurant	34.056232	-118.236183	New American Restaurant	No Delivery option	0

Results and Discussions

Properties of clusters and its future potential is summarized below

cluster	Cuisines to be supported by Application	Delivery provider	success probability
0	Miscellaneous	None	Moderate
1	Miscellaneous	grubhub	Low
2	American	None	High
3	Mexican	None	High
4	New American	None	High
5	Mediterranean	grubhub	Low
6	Korean	None	High
7	Chinese	None	High
8	French	grubhub	Low
9	French	None	Moderate

Conclusion

- To start a new delivery provider service to current restaurants, Chinese, Korean, American and Mexican restaurants will prove to be highly beneficial due to their large numbers and due to lack of any current major delivery provider competitors

- French Restaurants will be moderately beneficial as some French restaurants are already using grubhub service provider and the rest may follow them too

- Cluster 0 corresponding to miscellaneous cuisines may be profitable as they don't have current delivery provider but might be challenging as multiple cuisines need to be supported

- Cluster 1 restaurants will be least profitable as they have existing provider and also complex due to multiple categories

- From perspective of providing new cuisines / specialties in the location Chinese / Korean / French / Mexican / American may not be beneficial as they are already present in large numbers. Indian, Japanese, Peruvian, Vietnamese are in very few numbers. Starting these cuisines may be beneficial