Predict Grocery Demand for Mobile Grocery Store based on Cuisine Type for Franklin County Neighborhoods in OH

Jagadeesh Sreeram

Dec 27th 2019

Predicting Grocery Demand is valuable for Mobile Grocery Store

- Mobile Grocery stores target different neighborhoods and they need to load up their truck with right kind and quantity of inventory.
- Most of these inventory is perishable , so having loaded with inventory that can’t be sold is not good for the business. So, identifying the demand is key.
- Since this use case is mainly targeted towards restaurants , identifying right kind of restaurant and predicting demand based on these restaurants is key.
- Usecase is mainly targeted for neighborhoods of Franklin county

Data acquisition and cleaning

- Franklin neighborhood data is obtained from US Postal site. https://www.unitedstateszipcodes.org/zip-code-database/
- There were total 67 neighborhoods in Franklin county
- The neighborhood restaurant data is obtained from Foursquare using their Venue search API https://developer.foursquare.com/docs/api/endpoints
- Since the usecase demanded us to use below 5 grocery types, Indian,Chinese, Thai, Japanese and Korean
- There were some other restaurants categories suggest which needed to be cleanup.
- We got around 6700 total restaurants of above 5 types from Foursquare for all 67 neighborhoods in Franklin county.

Using ranking model for eah neighborhood based on restaurant category

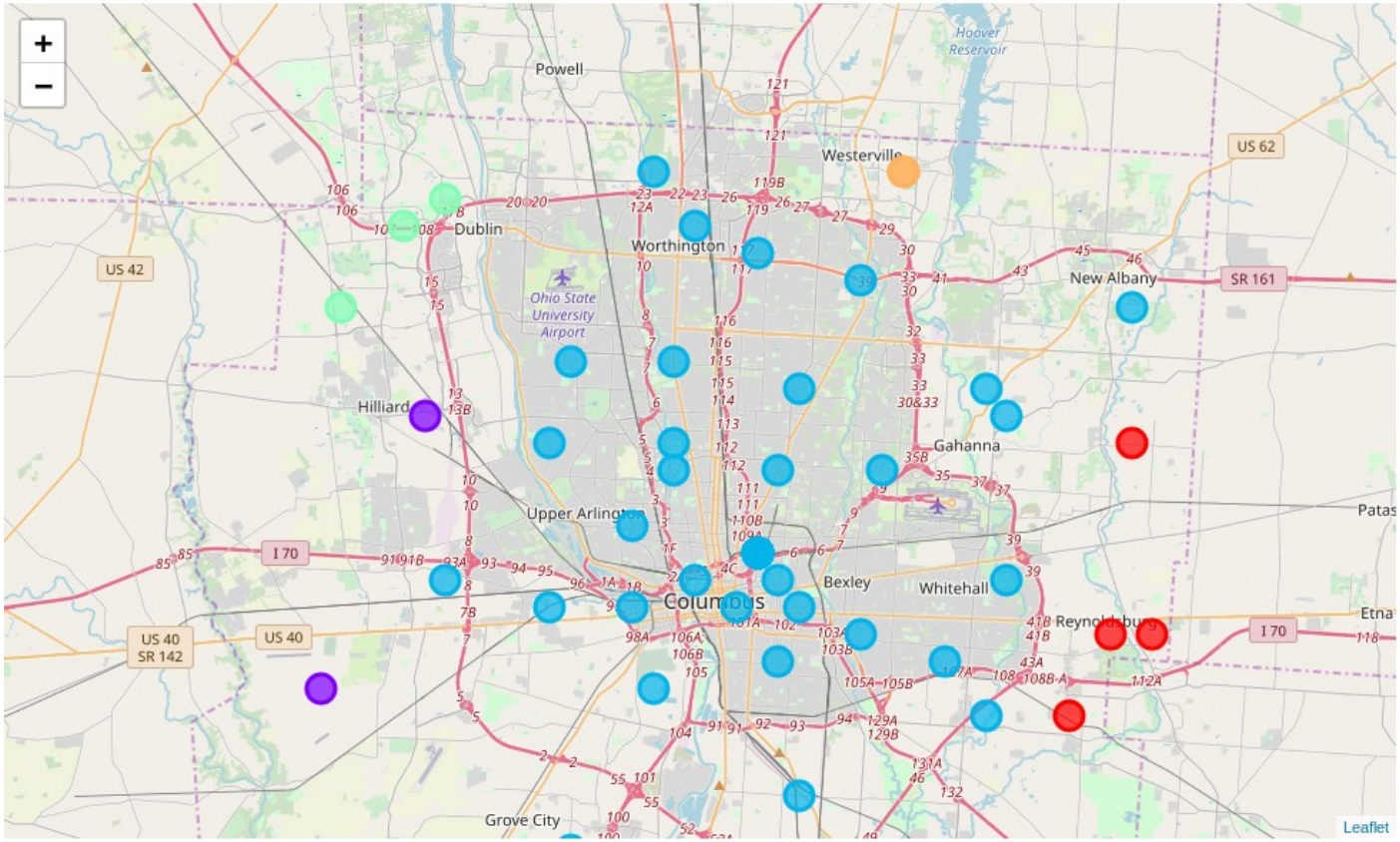
- Here we can see for each neighborhood we computed the weight-age for type of restaurants This gives good indication of which type of grocery among 5 are in more demand in which neighborhood.

	Neighborhood	Chinese Restaurant	Indian Restaurant	Japanese Restaurant	Korean Restaurant	Thai Restaurant
0	Amlin	0.470000	0.100000	0.160000	0.02	0.090000
1	Blacklick	0.460000	0.050000	0.200000	0.02	0.100000
2	Brice	0.480000	0.050000	0.190000	0.02	0.100000
3	Canal Winchester	0.500000	0.060000	0.170000	0.02	0.100000
4	Columbus	0.483478	0.058913	0.164565	0.02	0.098913

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Amlin	Chinese Restaurant	Japanese Restaurant	Indian Restaurant	Thai Restaurant	Korean Restaurant
1	Blacklick	Chinese Restaurant	Japanese Restaurant	Thai Restaurant	Indian Restaurant	Korean Restaurant
2	Brice	Chinese Restaurant	Japanese Restaurant	Thai Restaurant	Indian Restaurant	Korean Restaurant
3	Canal Winchester	Chinese Restaurant	Japanese Restaurant	Thai Restaurant	Indian Restaurant	Korean Restaurant
4	Columbus	Chinese Restaurant	Japanese Restaurant	Thai Restaurant	Indian Restaurant	Korean Restaurant
5	Dublin	Chinese Restaurant	Japanese Restaurant	Indian Restaurant	Thai Restaurant	Korean Restaurant

Clustering the Neighborhoods based on Grocery Demand

- Next we used K-means clustering algorithm to derive 5 different clusters for neighborhoods based on grocery demand.



Clustering the Neighborhoods based on Grocery Demand

- Also we were able to come up with demand percent for each type for a cluster. This will help the mobile store to just use 5 loading ratios of grocery and target those neighborhoods which has same demand.

	Latitude	Longitude	Chinese Grocery Percent	Japanese Grocery Percent	Indian Grocery Percent	Thai Grocery Percent	Korean Grocery Percent
Cluster Labels							
0	39.96	-82.81	48.25	17.50	5.75	10.00	2.00
1	39.98	-83.16	47.67	17.73	5.70	9.95	2.00
2	39.98	-82.97	49.83	16.33	6.83	9.00	2.00
3	40.09	-83.15	48.33	17.33	7.00	9.67	2.00
4	39.94	-82.95	48.33	17.33	6.00	9.67	2.00

Conclusion

- we were able to predict grocery demand based of restaurant type in neighborhood I.e 1st place is Chinese grocery with 50% , then in 2nd place Japanese with ~17% , 3rd Thai with ~10 % and 4th Indian with ~ 6 – 7% and finally 5th Korean which is pretty constant at 2%.
- Also we can see that demand at central part of the Franklin country is pretty much consistent and the differs around boarders. This also helps the Mobile grocery store to optimize their truck routes and direct their trucks to right neighborhoods nearby.

Future Improvements

- Currently the model only considers the number of restaurants in the vicinity. To make the model more accurate we should include other grocery demand factors like season , occupancy information of each restaurant, menu and its ingredients etc.