Searching for optimal locations for cloud kitchens in Mumbai

building a set of locations that maximize serviceable population
while ensuring minimum service time

By Ameya Kulkarni, as part of Applied Data Science Capstone Project

Image source: https://get.apicbase.com/reduce-costs-cloud-kitchen/

Why a cloud kitchen model?

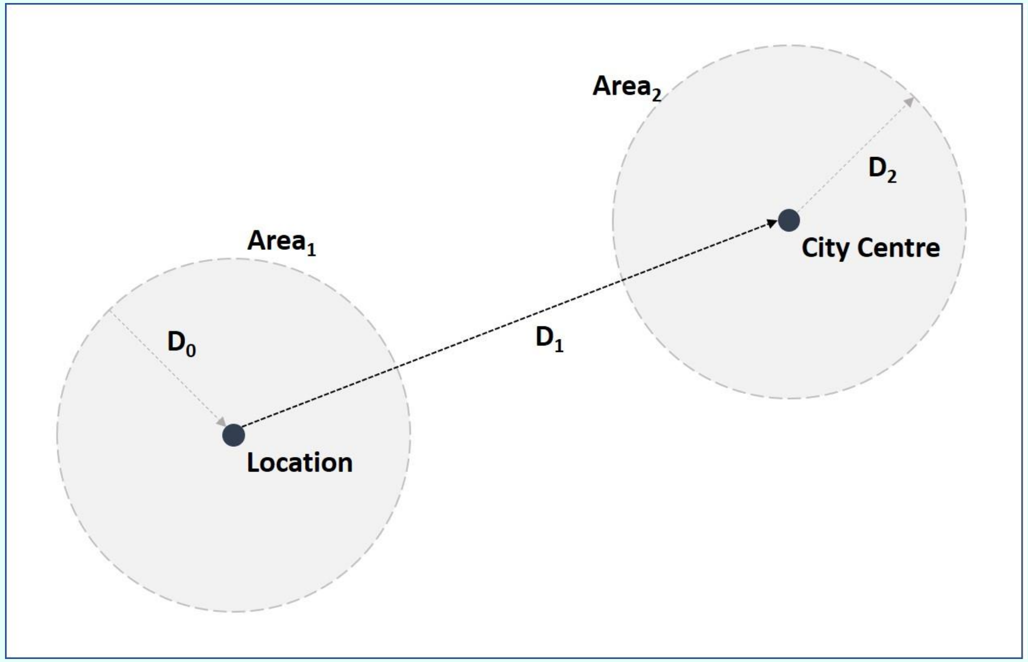
- The cloud kitchen model allows business to provide customers with multiple cuisines under different brands operating from the same physical location
- The model optimizes delivery logistics for the business, thereby reducing operational expenses
- With consumer behavior changing all around the globe due to the covid-19 pandemic, it may be the right time to switch from a traditional in-dining restaurant model to a delivery-only restaurant model

Location is the game

- Through this project we hope to efficiently identify locations suitable for setting up cloud kitchen
- Customizing the menu or cuisines offered would be an ongoing exercise, where the business must simply experiment with different chefs and cuisines and strategize based on the customer’s responses
- A report by Google and Boston Consulting Group revealed that most customers look for variety in cuisines while using food delivery apps
- This leaves only the other aspect of the cloud kitchen as the variable to optimize: location

Rating locations based on service time

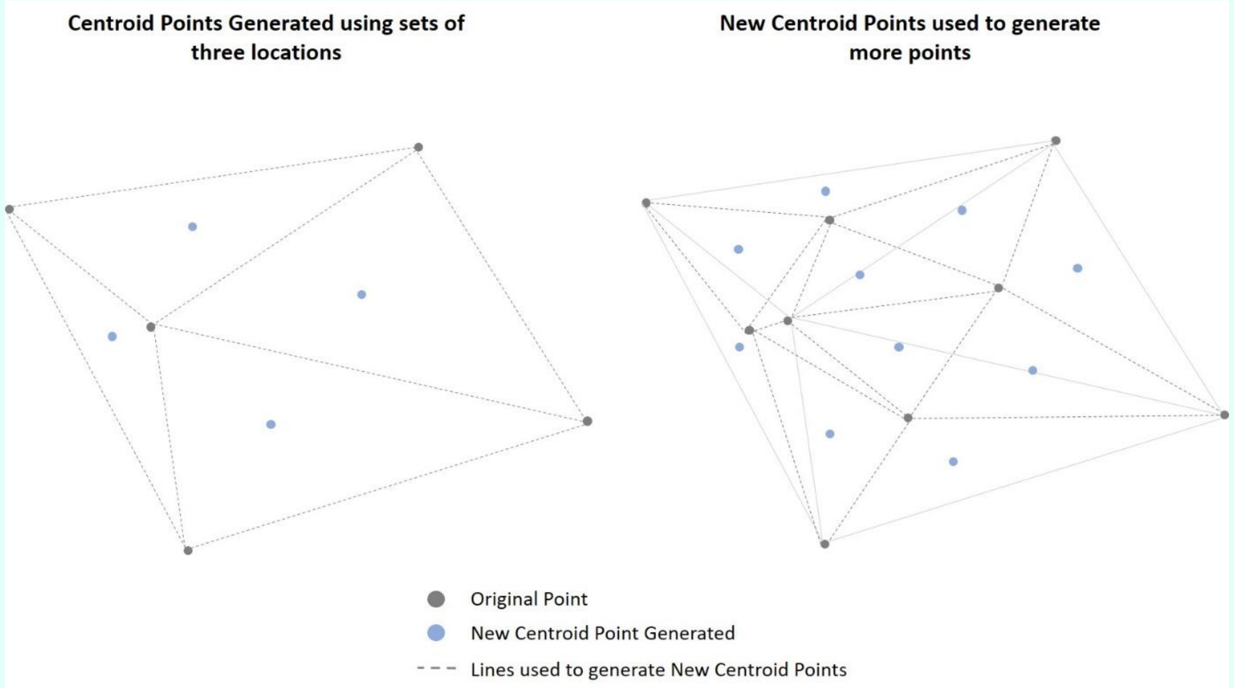
- By assuming a parabolic distribution of population around the city center coordinates, we can compute the percentage of the city center population that the venue location can service within 45 min.
- This combined with the minimum possible service time for a given set of city centers provides us with a rating



$$rating = \frac{\text{(minimum service time for a given set of city centres)}}{\text{(service time of a location for given set of city centres)}}$$

Finding the minimum service time

- In order to search the area bounded by the city center locations, we compute triangle centroids multiple times
- An triangle that's too large or too small is ignored
- Accuracy depends of depth of search, and as points grow exponential, so does the search time



Note: Not all triangle generates are used to generate new centroid points, each triangle has to meet an upper limit of side length and a lower limit on area

Best locations for cloud kitchens!

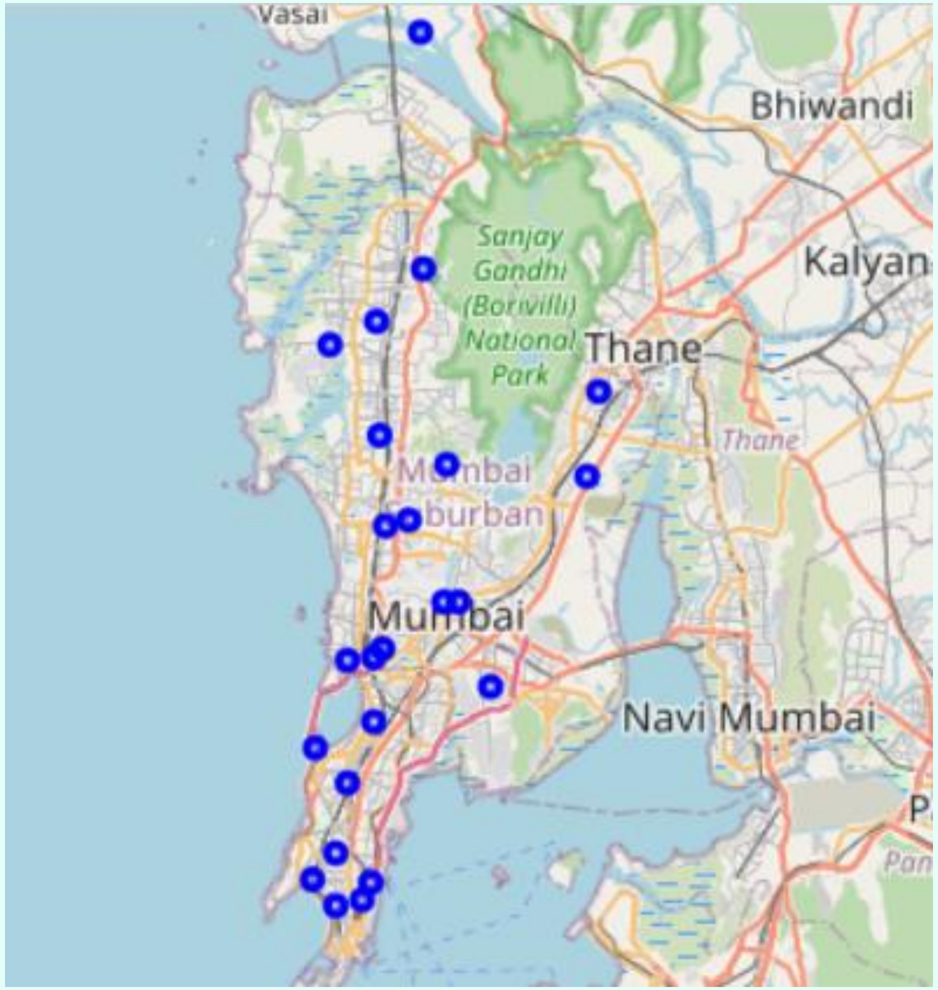
- Venue location data is used to train a model for more rapidly scoring multiple location points throughout the city area
- Top rated locations are easy to search for, rating close to 1
- By looking at top rated locations we can create a set that services the entire population with minimum number venues within a 45 min time limit

Data and model attributes

- For this analysis, we required two datasets,
    - Restaurant location details
        - Latitude and longitude
        - This was obtained using the Foursquare location data
    - City center details:
        - Latitude and longitude
        - Population
        - This was obtained scrapping multiple website
- Three inputs for the multivariate polynomial regression model,
    - Number of city centers within serviceable distance
    - Average distance between location and the city centers
    - Average potential serviceable population
- Target variable is the rating of a location, as defined earlier

City centers in Mumbai (Wards)

- Total number of city centers = 24
- Total population of all city centers = 1,19,76,439
  - Average city center population = 4,99,018
- Total area = 483 sq. km
  - Average city center area = 20 sq. km
- Average inter-city center distance = 16 km
- Average delivery time from all venues = 1.02 hr
- Average delivery time from venues within 15 km = 0.54 hr
- Average delivery time by category (venues within 15 km)
  - Dine-in: ~ 0.54 hr | Café: ~0.55 hr | Eatery = ~0.55 hr



Venue locations in Mumbai (Foursquare)

- Total number of restaurants = 1,129
- Average number city centers within serviceable distance (15 km) = 9
- Average distance to city centre within serviceable distance (15 km) = 7.2 km

Venue Type	Count	Average count of City Centres within 15 km	Average distance to City Centres within 15 km
Café	103	9	7.3
Dine in	844	8	7.2
Eatery	182	9	7.1

Linear Regression model with 3 inputs performed best

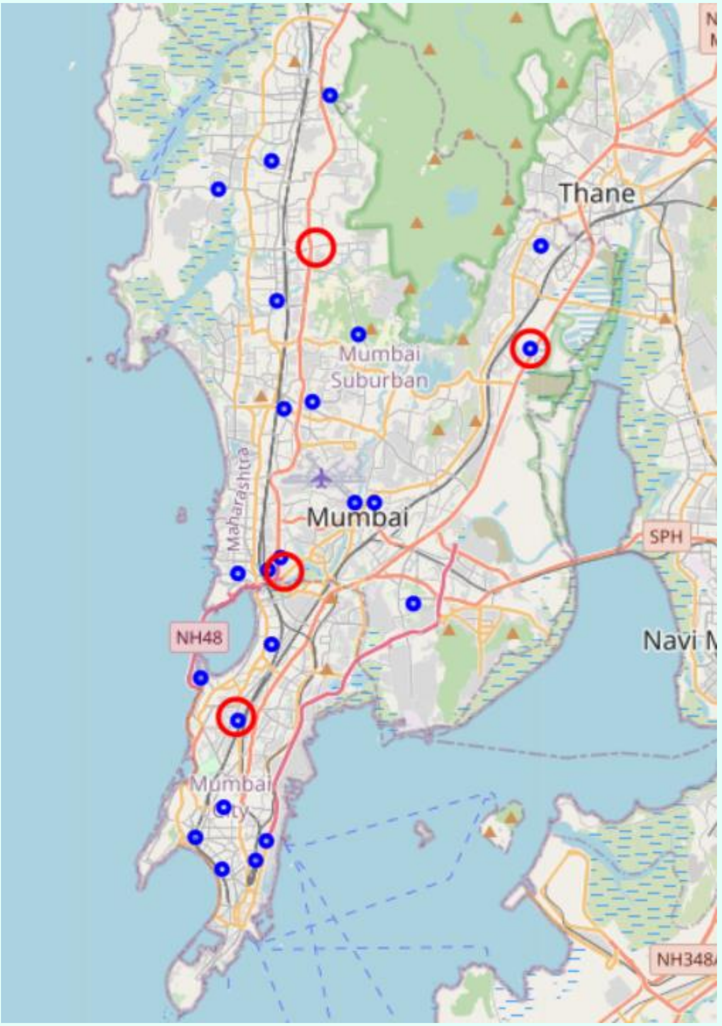
- We tested Linear Regression and Ridge Regression model, reduced model attributes
- The Linear Regression model with 3 inputs performed better than the other combinations.
- Testing the model of 500 samples gave a mean squared error of __
- Models tested
    A. Linear Regression, Polynomial of order 6, 3 inputs
    B. Linear Regression, Polynomial of order 6, 2 inputs
    C. Ridge Regression (alpha = 0.1), Polynomial of order 6, 3 inputs
    D. Ridge Regression (alpha = 0.1), Polynomial of order 6, 2 inputs

Model Evaluation	(A)	(B)	(C)	(D)
R^2	0.995311	0.960776	0.994489	0.960772
R^2	0.992297	0.958375	0.991564	0.958402
MSE	0.000456	0.003815	0.000536	0.003815
MSE	0.000636	0.003437	0.000696	0.003434

Parameter Evaluation (R^2)	(A)	(B)	(C)	(D)
Number of city centres within serviceable distance	0.913689	0.913689	0.913687	0.913687
Average distance between location and the city centres	0.317266	0.317266	0.317266	0.317266
Average potential serviceable population	0.194955		0.194954	

4 cloud kitchens to service maximum population

- City Centre Wise Details (considering best locations within 15 km)
    - Average service time = 0.67 hr
    - Average distance to best locations = 8 km
    - Proportion of population unserved = 20.82%
- Best Location Wise Details (considering city centres within 15 km)
    - Average service time = 0.70 hr
    - Average distance to city centre = 9
    - Average inter-best location distance = 18 km
- Currently the location selection has ensured a minimum of service time of 0.67 hr versus our initial benchmark of 0.75 hr



Cuisine recommendations were fairly standard for all locations

For each location identified, we computed the top 5 venue categories that are frequently available at the nearby (within 15 km) city centres, bottom 5 venue categories that are infrequently available, and list of venue categories that are unable at any of the city centres.

Location	Top 5	Bottom 5	Absent
1	Indian Restaurant, Café, Fast Food Restaurant, Chinese Restaurant, Restaurant	German Restaurant, Modern European Restaurant, Fish & Chips Shop, Brazilian Restaurant, Multicuisine Indian Restaurant	Dhaba, English Restaurant, Cafeteria, Molecular Gastronomy Restaurant, Parsi Restaurant, Steakhouse, Chaat Place, Portuguese Restaurant
2	Indian Restaurant, Café, Fast Food Restaurant, Chinese Restaurant, Restaurant	Dhaba, German Restaurant, Multicuisine Indian Restaurant, Gluten-free Restaurant, English Restaurant	Cafeteria, Molecular Gastronomy Restaurant, Parsi Restaurant, Steakhouse, Chaat Place, Portuguese Restaurant
3	Indian Restaurant, Café, Fast Food Restaurant, Chinese Restaurant, Restaurant	English Restaurant, Chaat Place, Burrito Place, Dumpling Restaurant, Parsi Restaurant	Dhaba
4	Indian Restaurant, Café, Chinese Restaurant, Fast Food Restaurant, Bakery	North Indian Restaurant, Creperie, Afghan Restaurant, Comfort Food Restaurant, Mughlai Restaurant	Dhaba, English Restaurant, Burrito Place, Dumpling Restaurant

Discussion (1/2)

- Model allows us to select a location at random and quickly gauge its optimality for a setting up a cloud kitchen model
- The key aspect in creating set of *best locations* was approaching it through the lens of population serviceability at a given city centre: the goal was to ensure that we don’t deliberately overserve any city centre
- For the 4 best locations generated, the proportion of unserved population is 20.82%, primarily due to an outlier city center that is 31 km away from the rest of the set
- Distance-travel time computations and population density functions we have assumed leave much to be desired

Discussion (2/2)

- Cuisine recommendation engine considers the presence of a type of venue category/cuisine near the city centre as being analogous to the population showing preference for it
- Today’s food delivery start-ups are flush with this information, and thus can use predictive analytics to identify trends to make more accuracy recommendations on present and upcoming consumer cuisine preferences.
- As the cloud kitchen model is able to host multiple brands with different cuisines, it is up to the owner to try and explore which cuisine combinations work best for a given location

Conclusion

- As cloud kitchen hosts multiple brands with different cuisines, it can have a delivery executive make multiple deliveries in a single journey
- This makes the demand for delivery executives less sensitive to fluctuations in order volume
- Cloud kitchen model makes owning the delivery operation more competitive versus relying on food delivery start-ups
- The project structure and model can easily be tweaked to optimise location selection for other types of problems like selecting locations for setting up hospitals or schools