INTRODUCTION:

In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening an restaurant and school in Chennai, India. Chennai is a city which home for some of the most delicious foods in southern india. A wide variety of cuisines are consumed in Chennai.Being a metro city home to people from multiple cultures it is important to have multiple cuisines in Chennai. Apart from the local cuisines, the most sought after cuisines are Italian and Chinese. In this project we try to identify the distribution of Italian restaurants in different localities

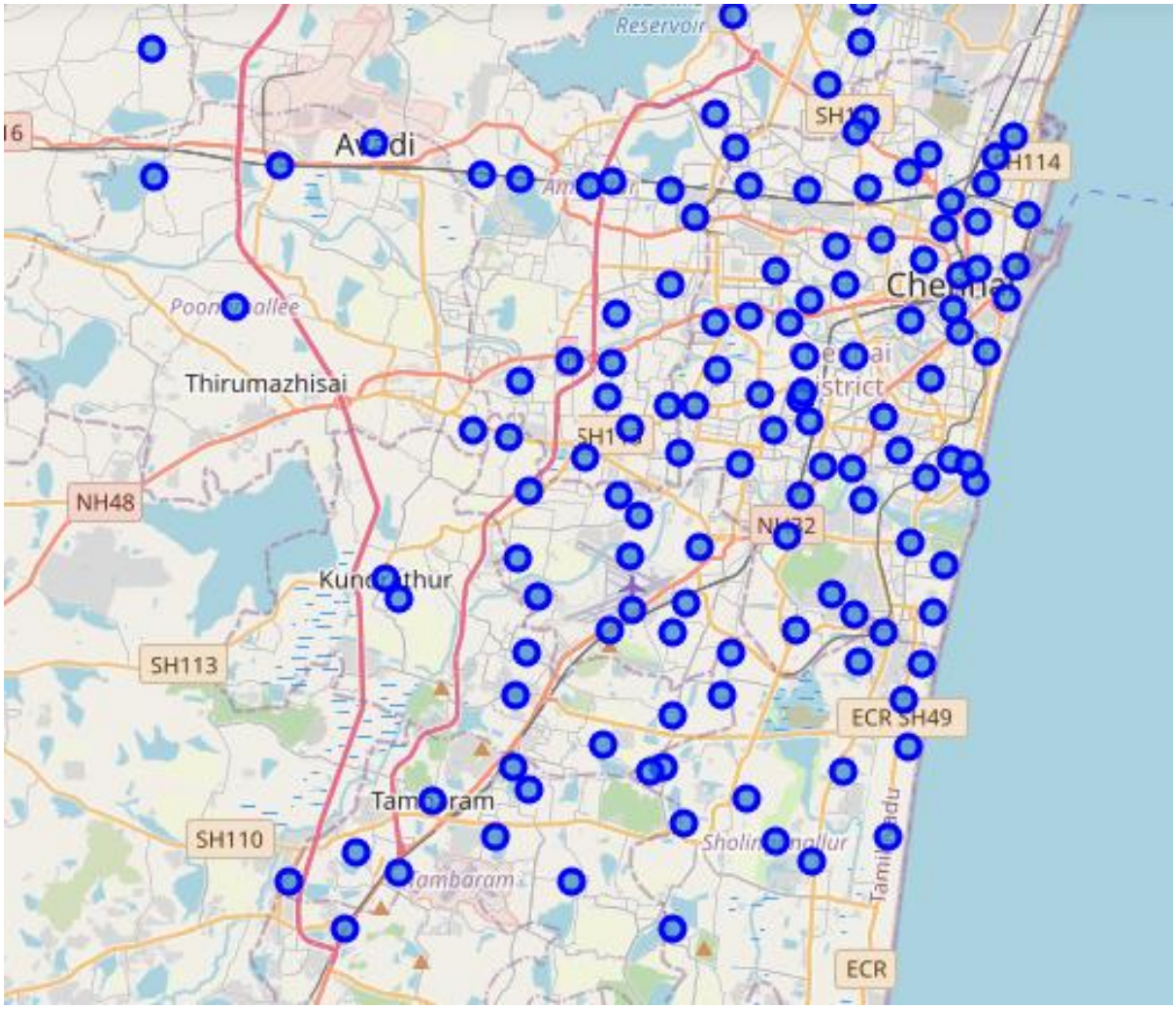
DATA ACQUISITION AND CLEANING

Based on definition of our problem, factors that will influence our decision are: number of existing restaurants in the neighborhood (any type of restaurant) number of and distance to Italian restaurants in the neighbourhood. Then we will classify these restaurants within the neighbourhood into clusters to visually understand proximity of restaurants and localities. Localities which have thinner clusters and those which have a wealthier group will be identified as he target localities. The weath distribution of localitis is not available in a tabular data format. Hence we will use local knowledge of the maker of the program to analyse this dat. To get the detailed information in each locality by using Foursquare API Besides we use other packages, likes:

- Pandas
- Numpy
- Json
- Geopy
- Matplotlib
- Shapely
- Pyproj
- Sklearn
- Folium Requests

METHODOLOGY

At first we import all the available localities into a python dataframe. The map coordinates for those localities are imported from Geocoder library. The overall map of Chennai localities is shown below



With the above data and the localities map coordinates, we import the data of all available restaurants in each locality. This will help us narrow down the Italian restaurants in a locality.

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Adyar\n	65	65	65	65	65	65
Alandur\n	51	51	51	51	51	51
Alapakkam\n	43	43	43	43	43	43
Alwarpet\n	100	100	100	100	100	100
Alwarthirunagar\n	65	65	65	65	65	65
Ambattur\n	13	13	13	13	13	13
Aminjikarai\n	100	100	100	100	100	100
Anna Nagar\n	100	100	100	100	100	100
Annanur\n	7	7	7	7	7	7
Arumbakkam\n	100	100	100	100	100	100

The above image is a part of the total dataframe. The total no of localities exceeds 150 places so it is not possible to show all the localities. The program was expected to find a limit of 100 restaurants in a locality, so the localities showing 100 restaurants have more than the mentioned number.

We also study the most common restaurant with each locality. This will give us a top level idea of which cuisine is most common within the whole city and also within localities.

	Neighborhood	1st Most Common Cuisine	2nd Most Common Cuisine	3rd Most Common Cuisine	4th Most Common Cuisine	5th Most Common Cuisine
0	Adyar\n	Indian Restaurant	Café	Chinese Restaurant	Fast Food Restaurant	Bakery
1	Alandur\n	Indian Restaurant	Café	Fast Food Restaurant	Pizza Place	South Indian Restaurant
2	Alapakkam\n	Indian Restaurant	Asian Restaurant	Pizza Place	Vegetarian / Vegan Restaurant	Fast Food Restaurant
3	Alwarpet\n	Indian Restaurant	Café	Chinese Restaurant	Restaurant	Italian Restaurant
4	Alwarthirunagar\n	Indian Restaurant	Fast Food Restaurant	Pizza Place	Vegetarian / Vegan Restaurant	Bakery
5	Ambattur\n	Indian Restaurant	Food Court	Pizza Place	Vegetarian / Vegan Restaurant	Asian Restaurant
6	Aminjikarai\n	Indian Restaurant	Fast Food Restaurant	Café	Vegetarian / Vegan Restaurant	Chinese Restaurant
7	Anna Nagar\n	Indian Restaurant	Fast Food Restaurant	Vegetarian / Vegan Restaurant	Café	Chinese Restaurant

We have plotted only the top 5 cuisines. It is very evident that the most common cuisine is Indian. This includes both south indian and north indian. The remaining common cuisines are Chinese, Italian and pizza joints. Although fast food & café appear predominantly in the list, they cannot be classified as cuisines. So we will ignore these 2 in our analysis.

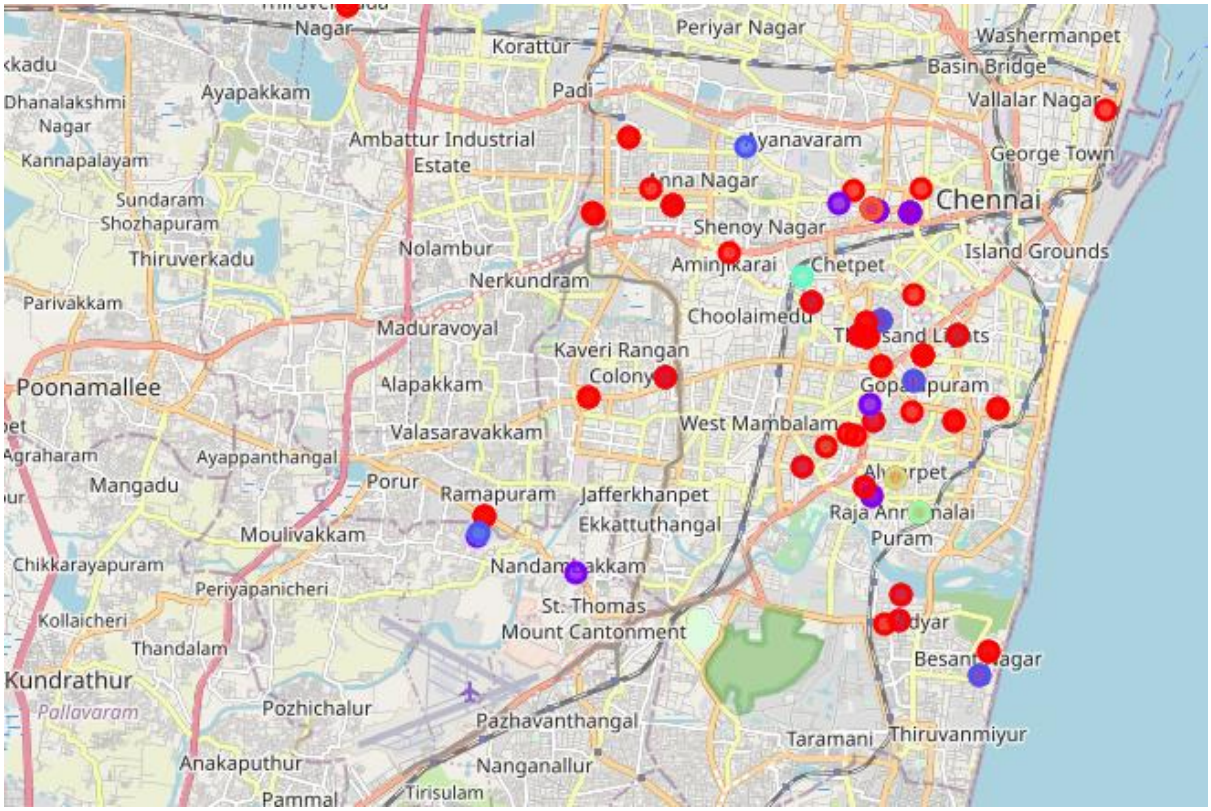
Based on the dataframe, we segregate the Italian restaurants in every locality. These restaurants are clustered as unsupervised data. This data will be helpful to plot in a map and study which locality has minimum or no Italian restaurant in wealthier localities.

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Adyar\n	9	9	9	9	9	9
Alandur\n	4	4	4	4	4	4
Alapakkam\n	1	1	1	1	1	1
Alwarpet\n	23	23	23	23	23	23
Alwarthirunagar\n	4	4	4	4	4	4
Ambattur\n	1	1	1	1	1	1
Aminjikarai\n	3	3	3	3	3	3
Anna Nagar\n	9	9	9	9	9	9
Annanur\n	1	1	1	1	1	1
Arumbakkam\n	3	3	3	3	3	3

Based on the above dataframe, it is understandable that some localities have a lot of Italian restaurants like Alwarpet, Anna nagar, Egmore, Nungambakkam & Adyar. These areas are generally considered the commercial hubs or wealthy residential suburbs.

Next we classify the restaurants into clusters to identify the proximity of Italian restaurants with each other. Although for this project we have used a KMeans clustering algorithm, the end result is not satisfactory as this has not been able to identify clusters properly.

However based on the existing locality data, we can still make conclusions from this data. We plot the restaurant locations on a map of Chennai to arrive at our conslusions



CONCLUSIONS

Based on the data we have analysed we arrive at the following conclusions

- Even thou Italian and Chinese cuisines are favoured by the people the number of restaurants serving authentic Italian is very low.
- Only a few localities have a good number of Italian restaurants which can serve good food.
- Commercial localities and wealthier residential suburbs have the most number of Italian restaurants. If we start a restaurant in such localities, the foot falls will be lower.
- The wealthier localities with no Italian restaurants are OMR & ECR. These will host a lot of wealthy people considering they are the localities with most IT companies. So we can expect higher foot falls.
- As a backup option, mambalam & kodambakkam can also be considered since these are commercial localities so foot falls maybe higher.