

# Capstone Project

Comparative view  
of 2 large  
metropolitan  
cities of India

(Chennai & Mumbai)





# Intent

- Chennai and Mumbai - both are large coastal metropolitan cities in India with lot of history. Both the cities have their unique lifestyle, types of industries, economical drivers and lifestyle of people.
- The comparison will show how these two large metropolitan cities are different and what are the similarities.
- We will use available data from the web to identify major clusters in the city and use that information to create a point of view about the kind of economy may flourish or be better suited for these two cities.



# Environment & Setup



For area & neighborhood data



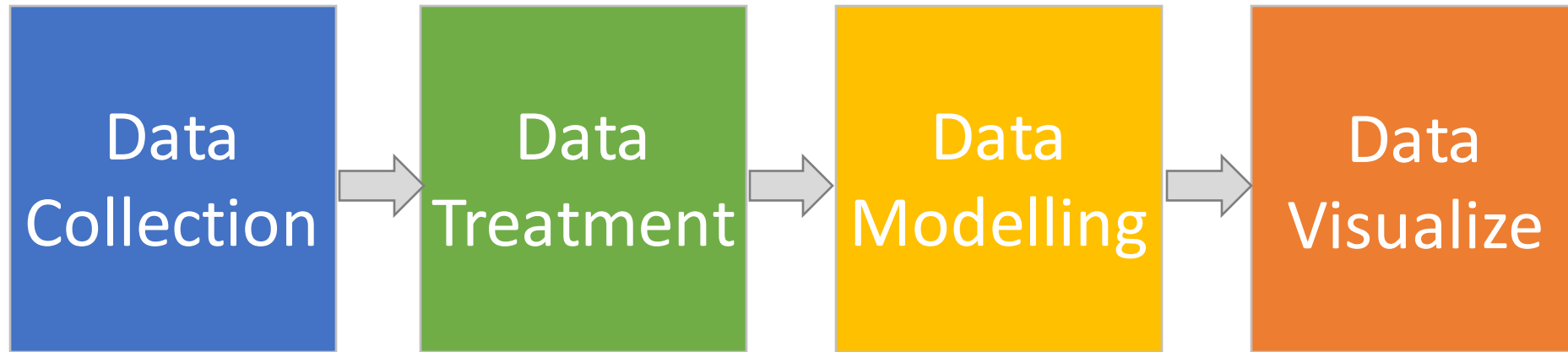
Programming Language



For venue & poi data

[This Photo](#) by Unknown author is licensed under [CC BY-SA](#).

# Stages of Study



# Data Collection



Area data was scraped from Wikipedia pages for Chennai and Mumbai

Wiki page

- Utharathi Vela
- North Usman Road
- South Usman Road
- Bazulath Road
- Chennai bypass road (puzhal-perungalathur)
- Habbuthali Road
- Arcot Road (SH-113)
- Mount-Poonamallee Road (SH-55)
- Nungambakkam High Road
- Outer Ring Road
- Peters Road
- Sander Patel Road
- Smith Road
- Whites Road
- Pallavaram - Kundrathur - Poonamallee Road
- Kundrathur to porur road
- Pallavaram - Thirupakkam Radial Road
- Velachery Main Road
- Vanagaram-Ambattur-Puzhal Road
- Tambaram Mudichur road (SH-48)
- East Coast Road (SH-48)
- Rajiv Gandhi Salai (Old Mahabalipuram Road) (SH-48A)
- Chennai-Trivelloor High Road (NH-205)
- Erukankerry High Road (NH-5)
- Grand Southern Trunk Road (GST Road) (NH-45)
- GST Road
- Avadi- Poonamallee- St. Thomas mount High road (SH-55)
- Old Mahabalipuram Road (OMR)
- 200 feet Radial Road
- Poonamallee - Thiruninravur (SH-50)
- Poonamallee - Pattabiram High road (SH-206)
- Wallajah Road, Chennai
- Grand Northern Trunk Road ("GNT Road") (NH-5)

Area	Location	Latitude	Longitude
Adambakkam	South and East Chennai	12.988	80.2047
Adyar	South and East Chennai	13.0012	80.2555
Alandur	South and East Chennai	12.9975	80.2006
Alapakkam	West Chennai	13.049	80.1673
Alwarthirunagar	West Chennai	13.0426	80.184
Ambattur	West Chennai	13.1143	80.1548
Aminjikarai	West Chennai	13.0908	80.2245
Anna Nagar	West Chennai	13.085	80.2101
Annamur	West Chennai	13.1184	80.1246
Arumbakkam	West Chennai	13.0724	80.2102
Ashok Nagar	West Chennai	13.0373	80.2123
Avadi	West Chennai	13.1067	80.097
Ayappakkam	West Chennai	13.0625	80.1372
Basin Bridge	North Chennai	13.1014	80.2704
Basant Nagar	South and East Chennai	13.0003	80.2987
Broadway	North Chennai	19.2274	72.9739
Central	North Chennai	13.0825	80.2755
Chetpet	South and East Chennai	13.0714	80.2417
Choolai	North Chennai	13.0919	80.2542
WMCA Colony	West Chennai	13.0665	80.2147
Defense Colony	West Chennai	13.0262	80.2022
Egmore	South and East Chennai	13.0732	80.2809
Recon.	North Chennai	13.15148	80.3555



Chennai Area Data

	Area	Location	Latitude	Longitude
0	Adambakkam	South and East Chennai	12.9880	80.2047
1	Adyar	South and East Chennai	13.0012	80.2565
2	Alandur	South and East Chennai	12.9975	80.2006
3	Alapakkam	West Chennai	13.0490	80.1673
4	Alwarthirunagar	West Chennai	13.0426	80.1840

Mumbai

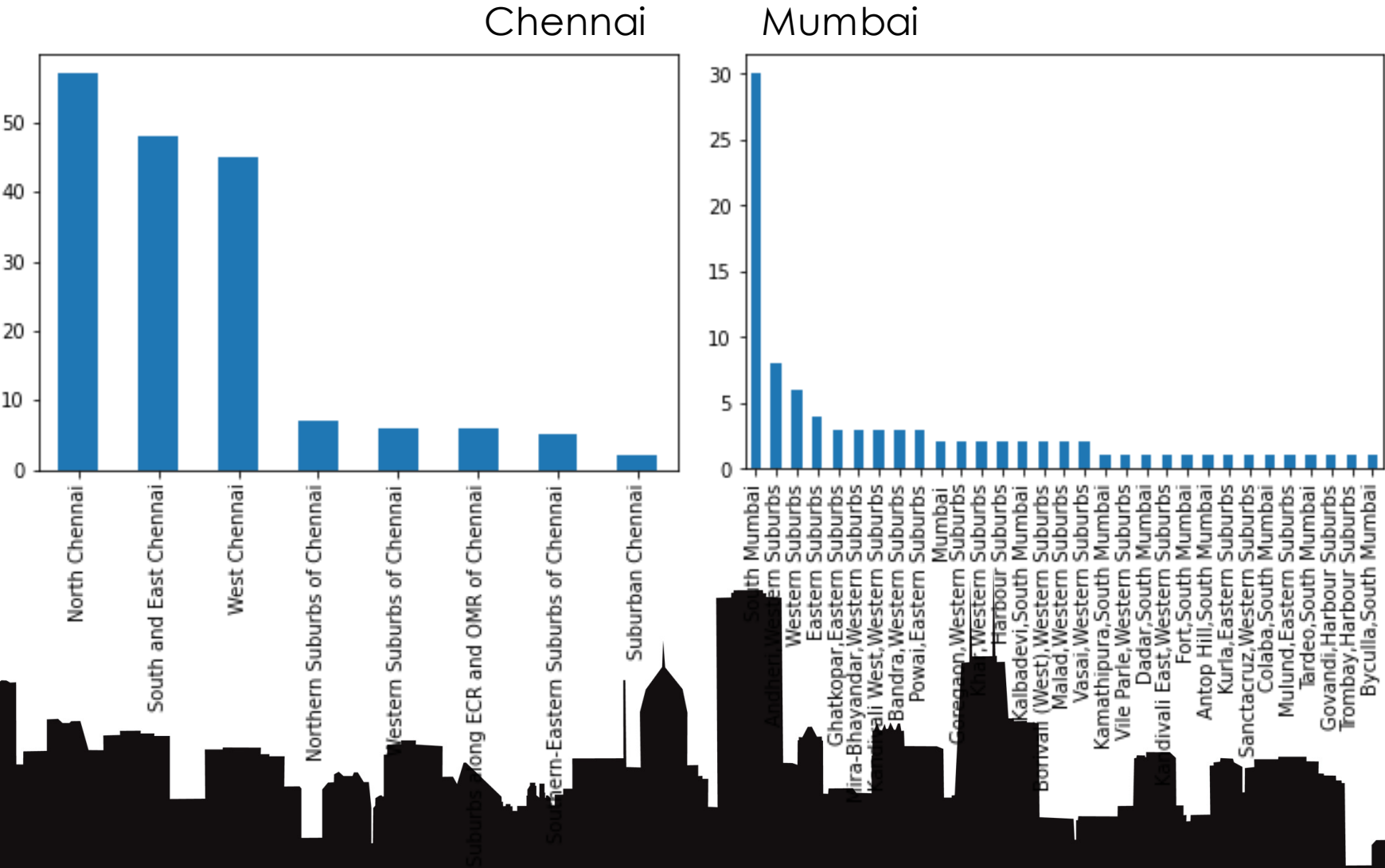
	Area	Location	Latitude	Longitude
0	Amboli	Andheri, Western Suburbs	19.129300	72.843400
1	Chakala, Andheri	Western Suburbs	19.111388	72.860833
2	D.N. Nagar	Andheri, Western Suburbs	19.124085	72.831373
3	Four Bungalows	Andheri, Western Suburbs	19.124714	72.827210
4	Lokhandwala	Andheri, Western Suburbs	19.130815	72.829270



# Data Collection



Area data was scaped from Wikipedia pages for Chennai and Mumbai





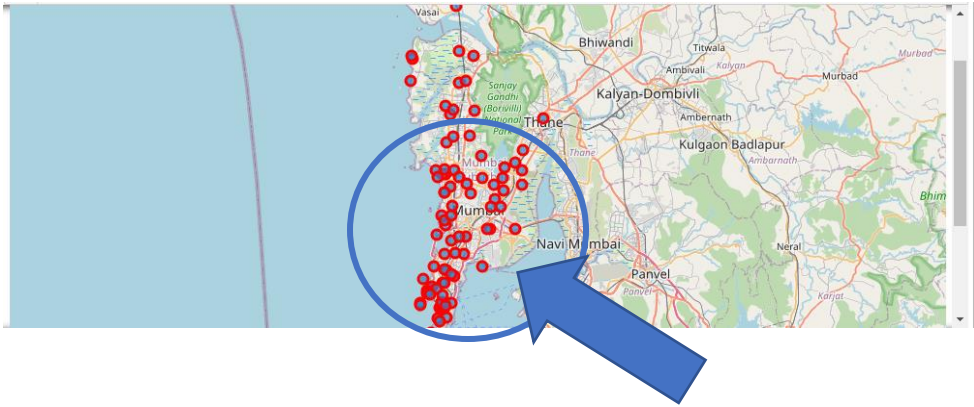
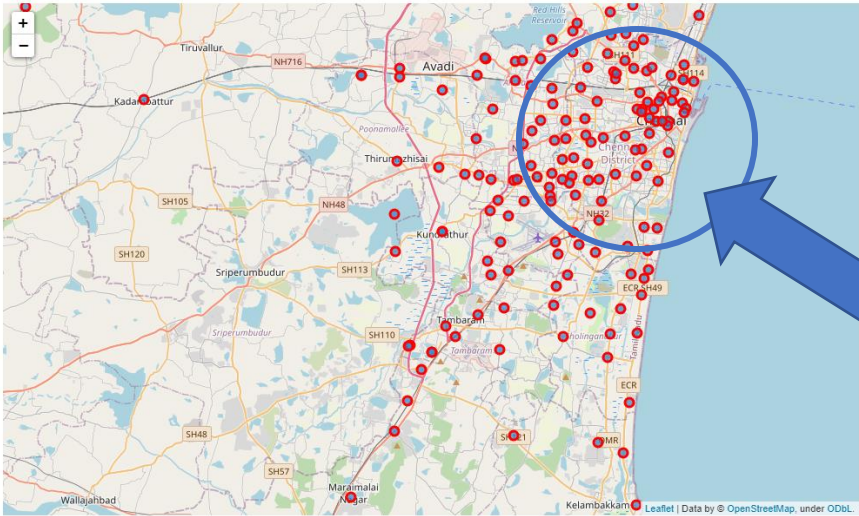
# Data Collection



Area data was scraped from Wikipedia pages for Chennai and Mumbai

Chennai

Mumbai



**Geo data  
mappable &  
accurate**



# Data Collection

Geo coordinates data was scraped from FourSquare Developer API for Chennai and Mumbai

## FSQ/developer

Explore Our Endpoints

Regular API Endpoints

Method	Endpoint Group	Endpoint	Usage	Required?
GET	venues	<a href="#">search</a>	Search for Venues	—
GET	venues	<a href="#">explore</a>	Get Venue Recommendations	—
POST	venues	<a href="#">select</a>	Report Venue Selection	—
GET	venues	<a href="#">trending</a>	Get Trending Venues	—
GET	venues	<a href="#">suggestcompletion</a>	Suggest Completions	—
GET	venues	<a href="#">likes</a>	Get Users Who Liked a Venue	—
GET	venues	<a href="#">categories</a>	Get Venue Categories	—
GET	venues	<a href="#">similar</a>	Get Similar Venues	—
GET	venues	<a href="#">nextvenues</a>	Get Next Venues	—
GET	venues	<a href="#">listed</a>	Get Lists a Venue is On	—

```
1 venues_in_chn = getNearbyVenues(df_chn['Area'], df_chn['Latitude'], df_chn['Longitude'])
```

Venue data by area from FSQ

```
1 venues_in_mum = getNearbyVenues(df_mum['Area'], df_mum['Latitude'], df_mum['Longitude'])
```

Venue data by area from FSQ



**FOURSQUARE**

<https://foursquare.com/>





# Data Treatment

Transforming data so that it can be used for modelling in the next stage

## Grouping Venues by categories

Venue Category	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue
ATM	Thirumullaivoyal north	13.2989	80.3203	HDFC Bank ATM
Afghan Restaurant	Pallavaram	13.0969	80.2865	Yaa Mohaideen Briyani
African Restaurant	Gopalapuram	13.0489	80.2586	Nando's
American Restaurant	Neelankarai	13.0850	80.2547	Tryst Cafe (Baker Street)
Antique Shop	Koyambedu	13.0694	80.1948	koyambedu fruit market
...	...	...	...	...

## One Hot Encoding

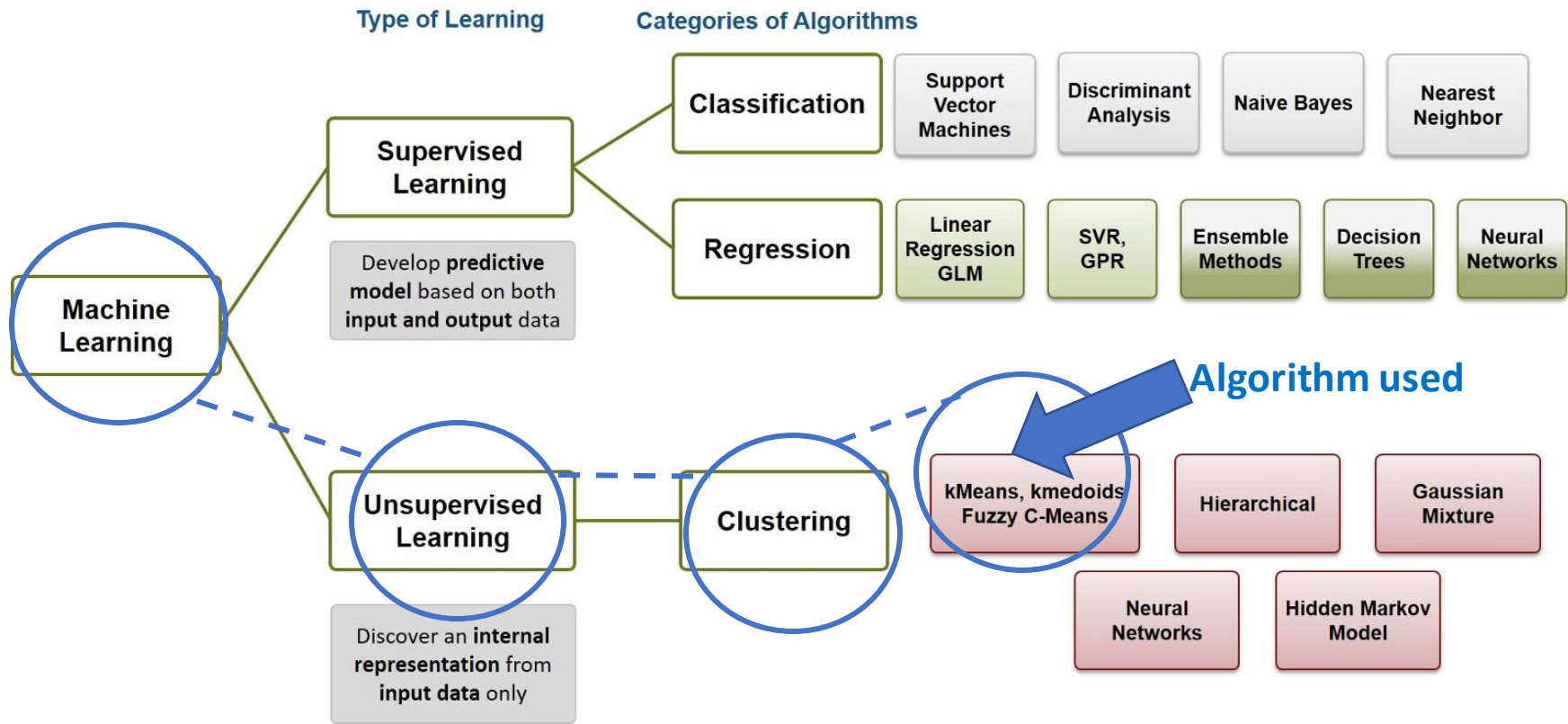
	ATM	Afghan Restaurant	African Restaurant	American Restaurant	Antique Shop	Art Gallery	Art Museum	Asian Restaurant	Astrol...
0	0	0	0	0	0	0	0	0	
1	0	0	0	0	0	0	0	0	
2	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0	
4	0	0	0	0	0	0	0	0	

## Top venues

	Neighbourhood	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
0	Adambakkam	Department Store	Dessert Shop	Gym	Indian Restaurant	Ice Cream Shop	Park	North Indian Restaurant
1	Adyar	Indian Restaurant	Café	Pizza Place	Electronics Store	Food Court	Bakery	Middle Eastern Restaurant
2	Alandur	Train Station	Indian Restaurant	Breakfast Spot	Kebab Restaurant	Metro Station	Park	Music Store
3	Alankulam	Fast Food	Snack	Indian	Park	Nightclub	North Indian	Office



# Data Modelling



Source : <https://medium.com/@taniyaghosh29/machine-learning-algorithms-what-are-the-differences-9b71df4f248f>

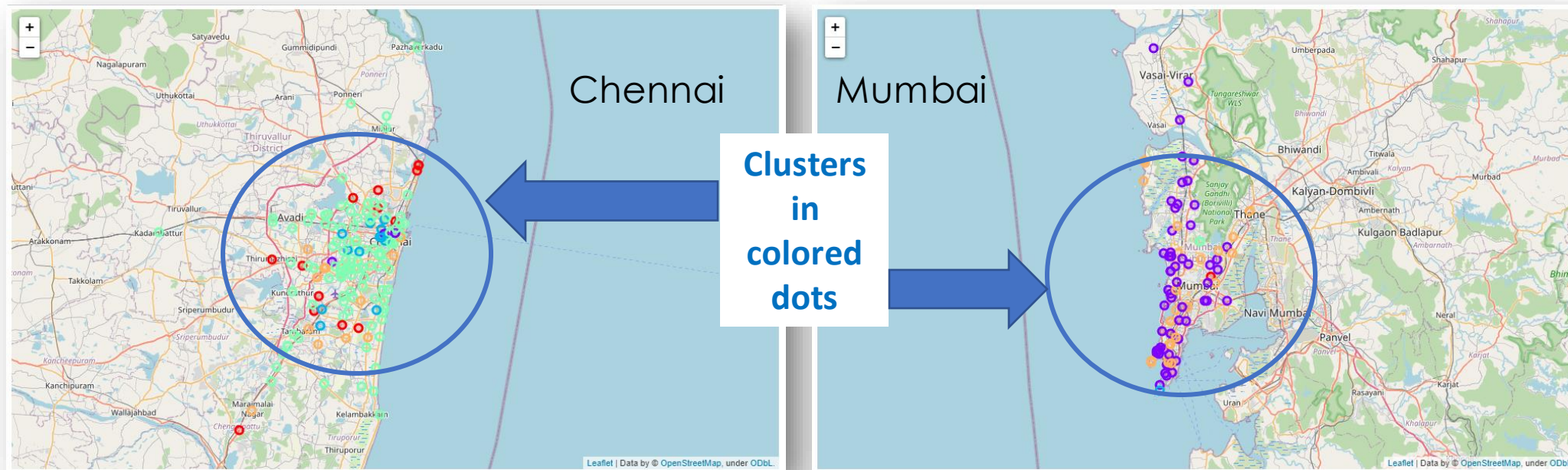


# Data Visualize



folium

Post Modelling visualization -- Venues are clustered and mapped





# Observations

	Similarities	Dissimilarities
1	Venues are of similar types/categories - as you would expect the needs of large size cities	Chennai preference seemed to be developing around fast food whereas Mumbai has been a mix of restaurants, markets and hotels
2	City seemed to have started around the main beach/coast and then developed progressively inward	Multiple and more references of vegetarian & vegan food styles in Chennai compared to Mumbai
3	Points of interest and venues are clustered towards the main/old city and are scattered scantily inwards	More outdoor lifestyle options in Mumbai compared to Chennai
4	Train stations scattered throughout the city in both cities	Chennai seems to have developed in concentric circular circles whereas Mumbai has developed around a linear development of settlements



# Conclusion

- ❖ Chennai and Mumbai are 2 of the 4 large metropolitan cities in India, both have rich history and have been developing.
- ❖ The cities started small and concentrated inwards closer to the coast/beach and have developed expanding inwards.
- ❖ These cities continue to grow and expand, on account of increasing economic activity as shown based on the venue listing from FourSquare.
- ❖ There are subtle lifestyle differences as observed by presence/absence of certain features. Travel options seemed to be common with inland train station/network within the cities.
- ❖ With this trend of growth, both the cities can be expected to become more significant in area and more points of interest/venues coming up around the outer periphery of the cities which continues to expand. Both the cities offer ideal ground for newer businesses and economic growth.

