

# IBM Applied Data Science Capstone

## ***Opening a New Coffee Shop in Chennai***

By: Ashok Venkatachalam

September 2020

# Business Problem

- Location of the coffee Shop is one of the most important decisions that will determine whether the Shop will be a success or a failure
- Objective: To analyze and select the best locations in the city of Chennai, India to open a new coffee Shop
- This project is timely as the city is currently short of coffee shop and the lifestyle of the people is changing and people are preferring to go to coffee shop for meeting and get together with friends.
- Business question  
In the suburbs of Chennai, if investor is looking to open a new coffee shop, where would you recommend that they open it?

# Data

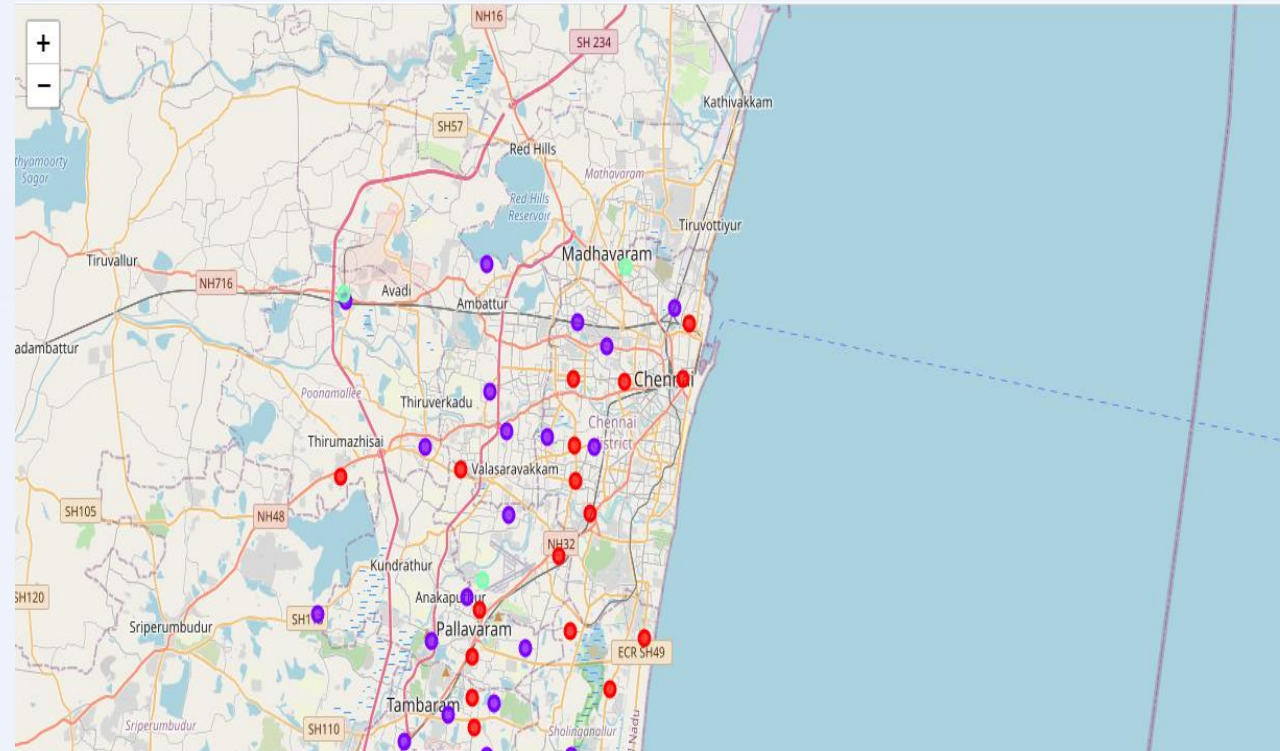
- Data required
  - List of neighbourhoods in Chennai
  - Latitude and longitude coordinates of the neighbourhoods
  - Venue data, particularly data related to Food Joints
- Sources of data
  - Wikipedia page for neighbourhoods  
([https://en.wikipedia.org/wiki/Category:Suburbs\\_of\\_Chennai](https://en.wikipedia.org/wiki/Category:Suburbs_of_Chennai))
  - Geocoder package for latitude and longitude coordinates
  - Foursquare API for venue data

# Methodology

- Web scraping Wikipedia page for neighbourhoods list
- Get latitude and longitude coordinates using Geocoder
- Use Foursquare API to get venue data
- Group data by neighborhood and taking the mean of the frequency of occurrence of each venue category
- Filter venue category by Food Joints
- Perform clustering on the data by using k-means clustering
- Visualize the clusters in a map using Folium

# Results

- Categorized the neighbourhoods into 3 clusters :
  - Cluster 0: Neighbourhoods with moderate number of Food Joints
  - Cluster 1: Neighbourhoods with low number to no existence of shopping malls
  - Cluster 2: Neighbourhoods with high concentration of Food Joints





# Discussion

- Most of the Food joints are concentrated in the central area of the city
- Highest number in cluster 2 and moderate number in cluster 0
- Cluster 1 has very low number to no Food joints in the neighbourhoods
- Oversupply of Food joints mostly happened in the central area of the city, with the suburb area still have very few shopping malls

# Recommendations

- Open new Food joints in neighbourhoods in cluster 0 moderate competition and good footfall.
- Can also open in neighbourhoods in cluster 1 with no competition if have unique selling propositions to stand out from the competition
- Avoid neighbourhoods in cluster 2, already high concentration of food joints and intense competition

# Conclusion

- Answer to business question: The neighbourhoods in cluster 0 are the most preferred locations to open a new coffee shop
- Findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a new coffee shop



# Thank you!

