

# Coursera - Capstone Project

IBM Applied Data Science Capstone

## Food Business In Pune, Maharashtra

By: Shivali Sur

# Introduction

- Pune city is home to a number of entrepreneurs.
- General population in Pune come from a diverse background and comprise of different locations, cultures, age groups and professions.
- The project is aimed at analysing the data available in the public domain to answer important questions while starting a business.

# Problem Statement

- A budding entrepreneur is willing to setup an Ice Cream Parlor in the City.
- We would like to understand the data available in the public domain to understand the trends across the city and take an informed decision about where to setup the Ice Cream Shop.

# Data

- Neighbourhood data is scraped from Wikipedia
  - [https://en.wikipedia.org/wiki/Category:Neighbourhoods\\_in\\_Pune](https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Pune)
- Location data is retrieved from Geocode APIs
  - <https://geocode.arcgis.com>
- Existing businesses across multiple neighbourhoods is extracted from ForeSquare
  - <https://api.foursquare.com/>

# Methodology

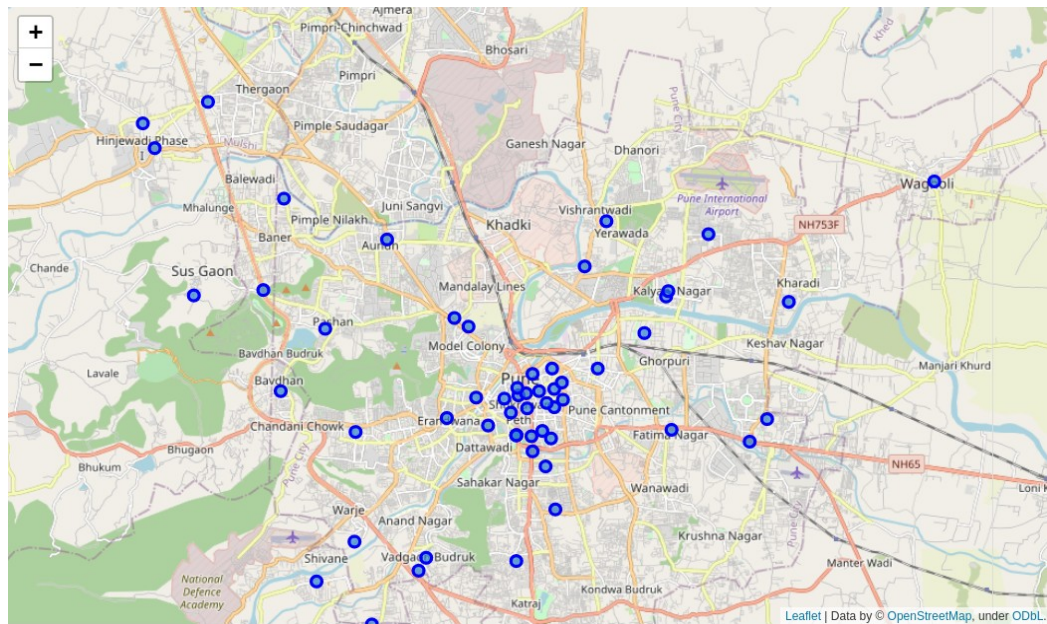
- Scrape neighbourhood data from Wikipedia

Neighborhoods	
0	Appa Balwant Chowk
1	Aundh, Pune
2	Balewadi
3	Baner
4	Bavdhan

# Methodology

- Extract location data and create a map using Geocode

	Neighborhoods	Latitude	Longitude
0	Appa Balwant Chowk	18.516479	73.853868
1	Aundh, Pune	18.563450	73.812270
2	Balewadi	18.576020	73.779830
3	Baner	18.548200	73.773160
4	Bavdhan	18.517542	73.778529
5	Bhavani Peth, Pune	18.512976	73.865289
6	Blue Ridge Town Pune	18.581196	73.985460
7	Budhwar Peth, Pune	18.516890	73.856170
8	Chakan, Pune	18.512448	73.856574
9	Dapodi	18.494410	74.398570
10	Deccan Gymkhana	18.515850	73.840610
11	Dhankawadi	18.466280	73.853260
12	Dhayari	18.447020	73.807570
13	Erandwane	18.509650	73.831240



# Methodology

- Get existing business details across neighbourhoods through FourSquare APIs

	Neighborhoods	Neighborhoods Latitude	Neighborhoods Longitude	Venue Name	Venue Category	Venue Latitude	Venue Longitude
0	Appa Balwant Chowk	18.516479	73.853868	Sujata Mastani	Ice Cream Shop	18.511793	73.852145
1	Appa Balwant Chowk	18.516479	73.853868	Bhagat Tarachand	Indian Restaurant	18.514332	73.851317
2	Appa Balwant Chowk	18.516479	73.853868	Hotel Madhuban	Tea Room	18.519248	73.848688
3	Appa Balwant Chowk	18.516479	73.853868	Raja Dinkar Kelkar museum	History Museum	18.510744	73.854389
4	Appa Balwant Chowk	18.516479	73.853868	Lal Mahal	Historic Site	18.518720	73.856556

# Methodology

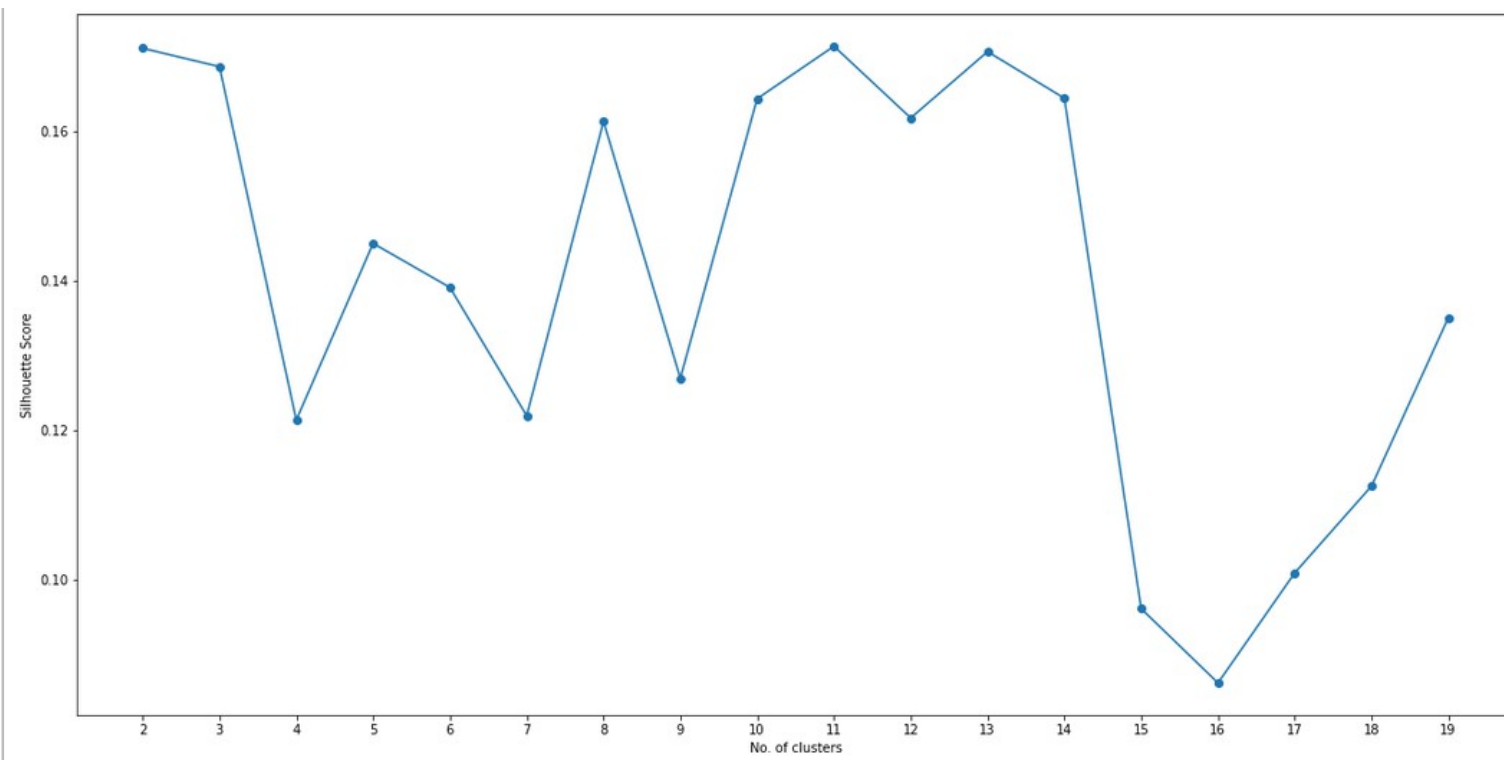
- Analyze neighbourhoods and respective businesses, sort top 10 categories.

	Neighborhoods	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Appa Balwant Chowk	Indian Restaurant	Vegetarian / Vegan Restaurant	Dessert Shop	Fast Food Restaurant	Snack Place	Tea Room	Café	Ice Cream Shop	Seafood Restaurant	Restaurant
1	Aundh, Pune	Indian Restaurant	Dessert Shop	Restaurant	Shopping Mall	Coffee Shop	Bakery	Fast Food Restaurant	Café	Snack Place	Gym
2	Balewadi	Indian Restaurant	Café	Asian Restaurant	Fast Food Restaurant	Ice Cream Shop	Brewery	Chinese Restaurant	Shopping Mall	Seafood Restaurant	Sandwich Place
3	Baner	Indian Restaurant	Department Store	Vegetarian / Vegan Restaurant	Ice Cream Shop	Beer Garden	Convenience Store	Coffee Shop	Mountain	Food	Café
4	Bavdhan	Indian Restaurant	Café	Smoke Shop	Fast Food Restaurant	Asian Restaurant	Restaurant	Bakery	Pool	Shop & Service	Sandwich Place



# Methodology

- Analyze Clusters based on their Silhouette Score



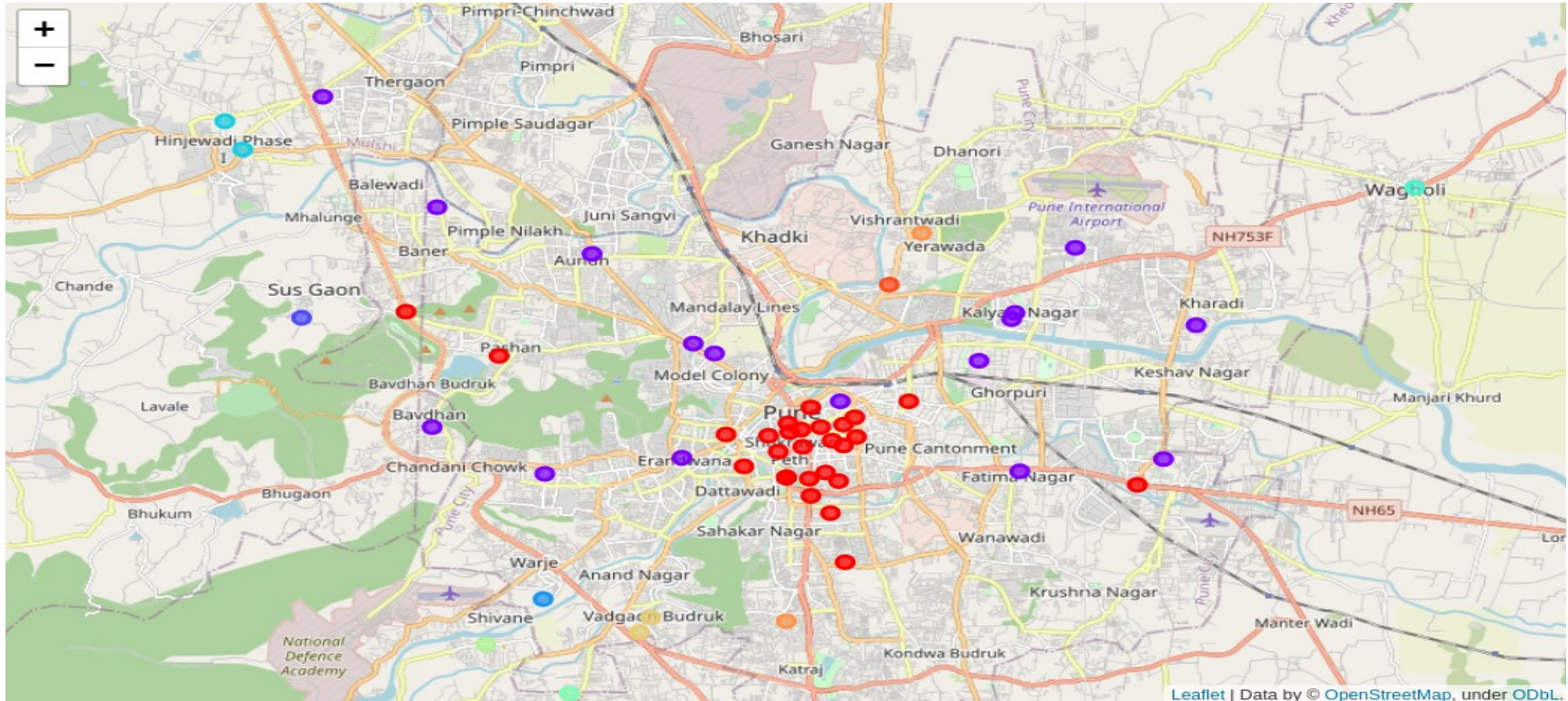
# Methodology

- Figure out top 3 common businesses across the neighbourhoods. Also, examine the clusters using KMeans

	Neighborhoods	Latitude	Longitude	Cluster Labels	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	8th Most Common Venue	9th Most Common Venue
0	Appa Balwant Chowk	18.516479	73.853868	0	Indian Restaurant	Vegetarian / Vegan Restaurant	Dessert Shop	Fast Food Restaurant	Snack Place	Tea Room	Café	Ice Cream Shop	Seafood Restaurant
1	Aundh, Pune	18.563450	73.812270	1	Indian Restaurant	Dessert Shop	Restaurant	Shopping Mall	Coffee Shop	Bakery	Fast Food Restaurant	Café	Snack Place
2	Balewadi	18.576020	73.779830	1	Indian Restaurant	Café	Asian Restaurant	Fast Food Restaurant	Ice Cream Shop	Brewery	Chinese Restaurant	Shopping Mall	Seafood Restaurant
3	Baner	18.548200	73.773160	0	Indian Restaurant	Department Store	Vegetarian / Vegan Restaurant	Ice Cream Shop	Beer Garden	Convenience Store	Coffee Shop	Mountain	Food
4	Bavdhan	18.517542	73.778529	1	Indian Restaurant	Café	Smoke Shop	Fast Food Restaurant	Asian Restaurant	Restaurant	Bakery	Pool	Shop & Service

# Methodology

- Cluster Mapping



# Observations

- Most common food joints are of the following categories:
  - Snack Place
  - Ice Cream Shop
  - Fast Food Restaurant
- The city being in India, Indian Restaurants are in abundance, hence the entry is ignored
- There exists clusters which don't have an Ice Cream Shop.

# Limitations and Future Work

- To more accurately determine the neighbourhood to start a business, existing data can be compared against more datasets such as
  - Demographic data of different neighbourhoods
  - Expense trends across the city
  - Cost of business in different neighbourhoods

# Result

- Based on analysis, we can determine the most common and popular food categories and joints.
- Also, this information helps us determine neighbourhoods where no such establishments are present.
- These neighbourhoods can be targeted to open up an Ice Cream Shop as competition is low, and one can tap into the customer base which otherwise have to travel to nearby neighbourhoods to get their deserts.