

# **IBM DATA SCIENCE PROFESSIONAL CERTIFICATE CAPSTONE PRESENTATION**

**TOPIC: LOCATING THE BEST PLACE IN DELHI NCR  
REGION (INDIA) TO OPEN A CHINESE FOOD JOINT.**

**By ASHISH KUMAR GAUTAM**



# INTRODUCTION:

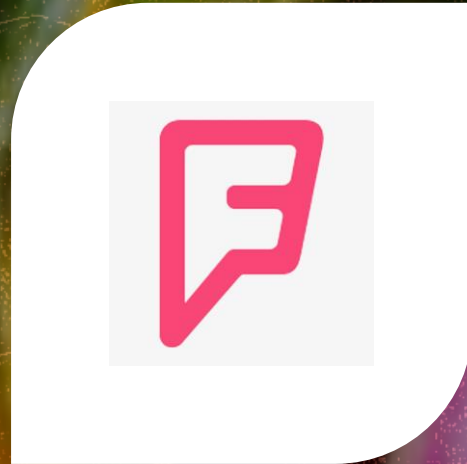
- The inspiration behind this problem is pretty simple. I'm in love with Chinese cuisine. So I kind of thought of a way to incorporate this final report in and around that. The idea is to find a few Chinese Food Spots in and around town and classify them with parameters that are available.

## BUSINESS PROBLEM:

- Identifying critical parameters to open a new Chinese Food Joint in Delhi-NCR area.



# DATA SOURCES



FOUR SQUARE API



ZOMATO API



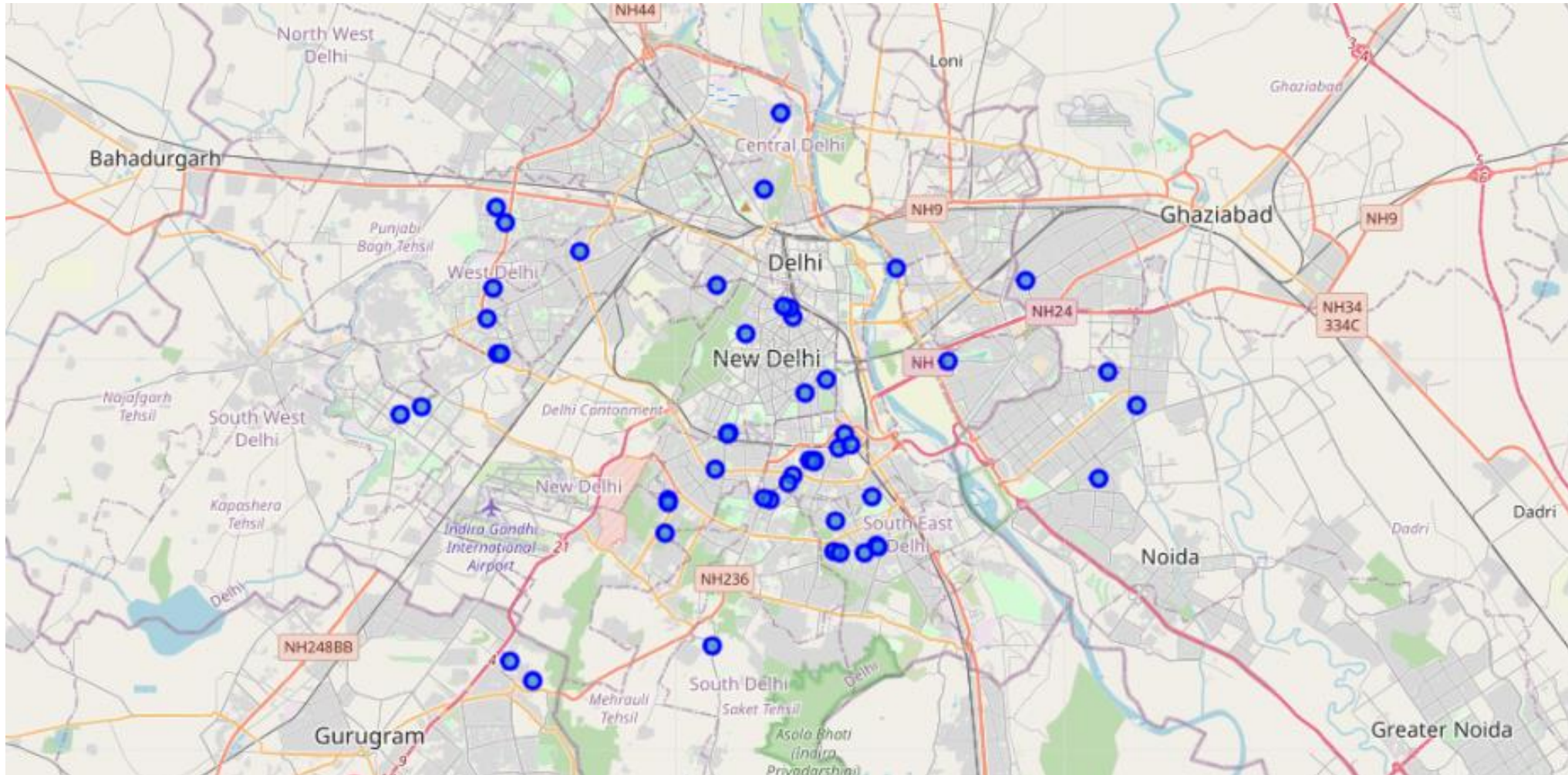
# METHODOLOGY

- Initially, I queried 'Chinese' and obtained results from the foursquare api in and around Delhi-NCR. The radius was kept 5000 (large to get enough results for a conclusion and not too large that we deviate too far from city centre)

## RESULTANT DATAFRAME

	name	location.lat	location.lng	Category
0	The Chinese	28.630593	77.220618	Chinese Restaurant
1	Fujiya Japanese n Chinese restaurant	28.623762	77.199564	Asian Restaurant
2	Chinese Bite	28.584280	77.191714	Chinese Restaurant
3	Chinese Garden, Yashwant Place	28.584106	77.191610	Restaurant
4	Chinese corner	28.642798	77.186733	Chinese Restaurant
5	Soy So Chinese	28.600367	77.226541	Indian Restaurant
6	Chinese Hut	28.583872	77.243970	College Administrative Building
7	Pioneer - The Chinese Restaurant	28.634521	77.216965	Chinese Restaurant

# PLOTTING DATA ON MAP TO SEE THE SPREAD



# RETRIEVAL FROM ZOMATO API

- **Next Data from Zomato API is collected using the same latitude longitudes of the obtained results from the foursquare API**

## RESULTANT DATAFRAME

	name1	location.lat1	location.lng1	AvgCostfor2	PriceRange	AverageUserRating	Address
0	Kinbuck 2	28.6321597190	77.2166951001	1500	3	4.1	P-10/90, 1st & 2nd Floor, Outer Circle, Connau...
1	Fujiya	28.6015914075	77.1860203519	1500	3	4.0	12/48, Malcha Marg Market, Chanakyapuri, New D...
2	Chinese Bite	28.5842797140	77.1916550025	750	2	3.1	12 & 13, Yashwant Place Commercial Complex, Ch...
3	China Garden	28.6340639657	77.2172831744	1900	3	4.2	G 4/5/6, Marina Arcade, Connaught Place, New D...
4	Chinese Corner	28.6405233153	77.1846685186	250	1	3.0	Near Safal Pure Veg, Old Rajinder Nagar, New D...
5	Oh So Stoned!	28.5673629837	77.3211758584	400	1	4.0	4th Floor, DLF Mall of India, Sector 18, Noida
6	The Chinese Hut	28.6533996311	77.2729292884	300	1	3.5	M-134, Jagat Ram Park, Laxmi Nagar, New Delhi
7	The Chinese & Thai Restaurant	28.5614604055	77.1928516030	350	1	3.9	85-B, Humayunpur, Safdarjung, New Delhi
8	Kinbuck 2	28.6321597190	77.2166951001	1500	3	4.1	P-10/90, 1st & 2nd Floor, Outer Circle, Connau...
9	Zen	28.6343029136	77.2193112597	1600	3	3.5	B-25, Connaught Place, New Delhi
10	The Chinese Hut	28.6533996311	77.2729292884	300	1	3.5	M-134, Jagat Ram Park, Laxmi Nagar, New Delhi
11	Chinese Corner	28.6405233153	77.1846685186	250	1	3.0	Near Safal Pure Veg, Old Rajinder Nagar, New D...
12	The Chinese & Thai Restaurant	28.5614604055	77.1928516030	350	1	3.9	85-B, Humayunpur, Safdarjung, New Delhi

# MERGING FOURSQUARE AND ZOMATO DATA

## RESULTANT DATAFRAME

	AvgCostfor2	PriceRange	AverageUserRating	name	location.lat	location.lng	Category
0	1500	3	4.1	The Chinese	28.630593	77.220618	Chinese Restaurant
1	1500	3	4.0	Fujiya Japanese n Chinese restaurant	28.623762	77.199564	Asian Restaurant
2	750	2	3.1	Chinese Bite	28.584280	77.191714	Chinese Restaurant
3	1900	3	4.2	Chinese Garden, Yashwant Place	28.584106	77.191610	Restaurant
4	250	1	3.0	Chinese corner	28.642798	77.186733	Chinese Restaurant
5	400	1	4.0	Soy So Chinese	28.600367	77.226541	Indian Restaurant
6	300	1	3.5	Chinese Hut	28.583872	77.243970	College Administrative Building
7	350	1	3.9	Pioneer - The Chinese Restaurant	28.634521	77.216965	Chinese Restaurant
8	1600	3	3.5	Zen Chinese, Connaught Place	28.633658	77.219495	Chinese Restaurant
9	300	1	3.2	Chinese Inn	28.649473	77.267306	Chinese Restaurant
10	300	1	3.5	Chinese van C-Block EOK	28.559443	77.256739	Chinese Restaurant
11	2000	4	4.1	chinese chilli sizzler	28.549471	77.240234	Chinese Restaurant
12	500	2	3.2	chinese express	28.656441	77.124744	Chinese Restaurant
13	550	2	4.2	Chinese Kitchen	28.536727	77.253087	Chinese Restaurant



# DATA WRANGLING

- Data Wrangling is done to use most of the data, eg One Hot Encoding of Category, Names are dropped as they are of no use etc.

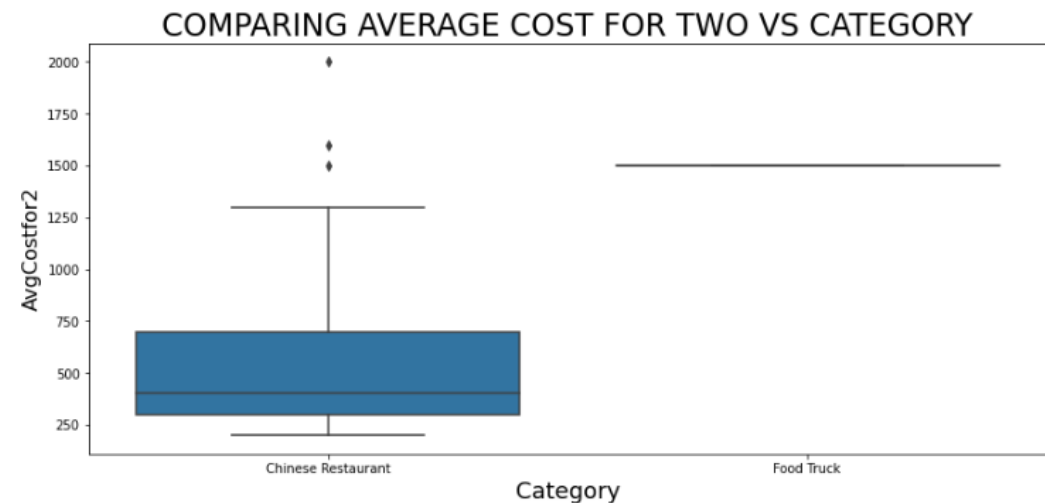
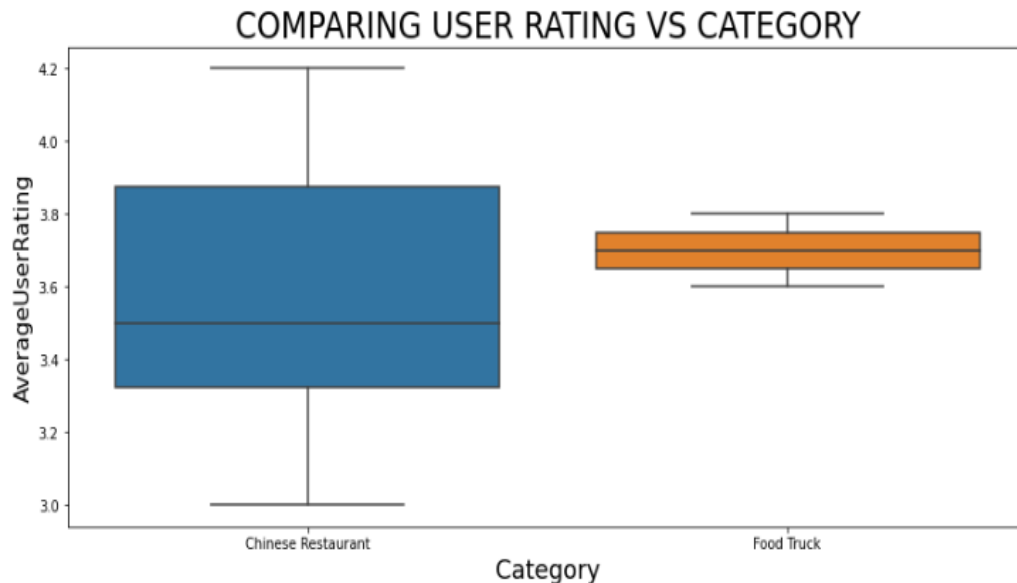
## RESULTANT DATAFRAME

	Asian Restaurant	Bakery	Bar	Chinese Restaurant	College Administrative Building	College Cafeteria	Food Truck	Indian Chinese Restaurant	Indian Restaurant	Restaurant	AvgCostfor2	PriceRange	AverageUserRating	location.lat	location.lng
0	0	0	0	1	0	0	0	0	0	0	1500	3.0	4.1	28.630593	77.220618
1	1	0	0	0	0	0	0	0	0	0	1500	3.0	4.0	28.623762	77.199564
2	0	0	0	1	0	0	0	0	0	0	750	2.0	3.1	28.584280	77.191714
3	0	0	0	0	0	0	0	0	0	1	1900	3.0	4.2	28.584106	77.191610
4	0	0	0	1	0	0	0	0	0	0	250	1.0	3.0	28.642798	77.186733
5	0	0	0	0	0	0	0	0	1	0	400	1.0	4.0	28.600367	77.226541
6	0	0	0	0	1	0	0	0	0	0	300	1.0	3.5	28.583872	77.243970
7	0	0	0	1	0	0	0	0	0	0	350	1.0	3.9	28.634521	77.216965
8	0	0	0	1	0	0	0	0	0	0	1600	3.0	3.5	28.633658	77.219495
9	0	0	0	1	0	0	0	0	0	0	300	1.0	3.2	28.649473	77.267306
10	0	0	0	1	0	0	0	0	0	0	300	1.0	3.5	28.559443	77.256739
11	0	0	0	1	0	0	0	0	0	0	2000	4.0	4.1	28.549471	77.240234
12	0	0	0	1	0	0	0	0	0	0	500	2.0	3.2	28.656441	77.124744
13	0	0	0	1	0	0	0	0	0	0	550	2.0	4.2	28.536727	77.253087
14	0	0	0	0	0	0	1	0	0	0	1500	3.0	3.6	28.573691	77.228528
15	0	0	0	1	0	0	0	0	0	0	1300	3.0	3.9	28.567631	77.220950



# DATA ANALYSIS

- Data was grouped in categories and two most frequent categories were compared using boxplot.
- We can easily notice that food trucks have higher rating and costlier than dine-in Chinese restaurants. Also spread is also lower in food trucks.



# DATA NORMALIZATION AND MODELING

- Data was normalized using standard scaler and then clustering was done using K-Means into the group of 4

## RESULTANT DATAFRAME

	AvgCostfor2	PriceRange	AverageUserRating	name	location.lat	location.lng	Category	label
0	1500	3.0	4.1	The Chinese	28.630593	77.220618	Chinese Restaurant	1
1	1500	3.0	4.0	Fujiya Japanese n Chinese restaurant	28.623762	77.199564	Asian Restaurant	1
2	750	2.0	3.1	Chinese Bite	28.584280	77.191714	Chinese Restaurant	0
3	1900	3.0	4.2	Chinese Garden, Yashwant Place	28.584106	77.191610	Restaurant	1
4	250	1.0	3.0	Chinese corner	28.642798	77.186733	Chinese Restaurant	0
5	1600	3.0	3.5	Zen Chinese, Connaught Place	28.633658	77.219495	Chinese Restaurant	1
6	400	1.0	4.0	Soy So Chinese	28.600367	77.226541	Indian Restaurant	2
7	300	1.0	3.5	Chinese Hut	28.583872	77.243970	College Administrative Building	2
8	350	1.0	3.9	Pioneer - The Chinese Restaurant	28.634521	77.216965	Chinese Restaurant	0
9	300	1.0	3.2	Chinese Inn	28.649473	77.267306	Chinese Restaurant	0
10	300	1.0	3.5	Chinese van C-Block EOK	28.559443	77.256739	Chinese Restaurant	0

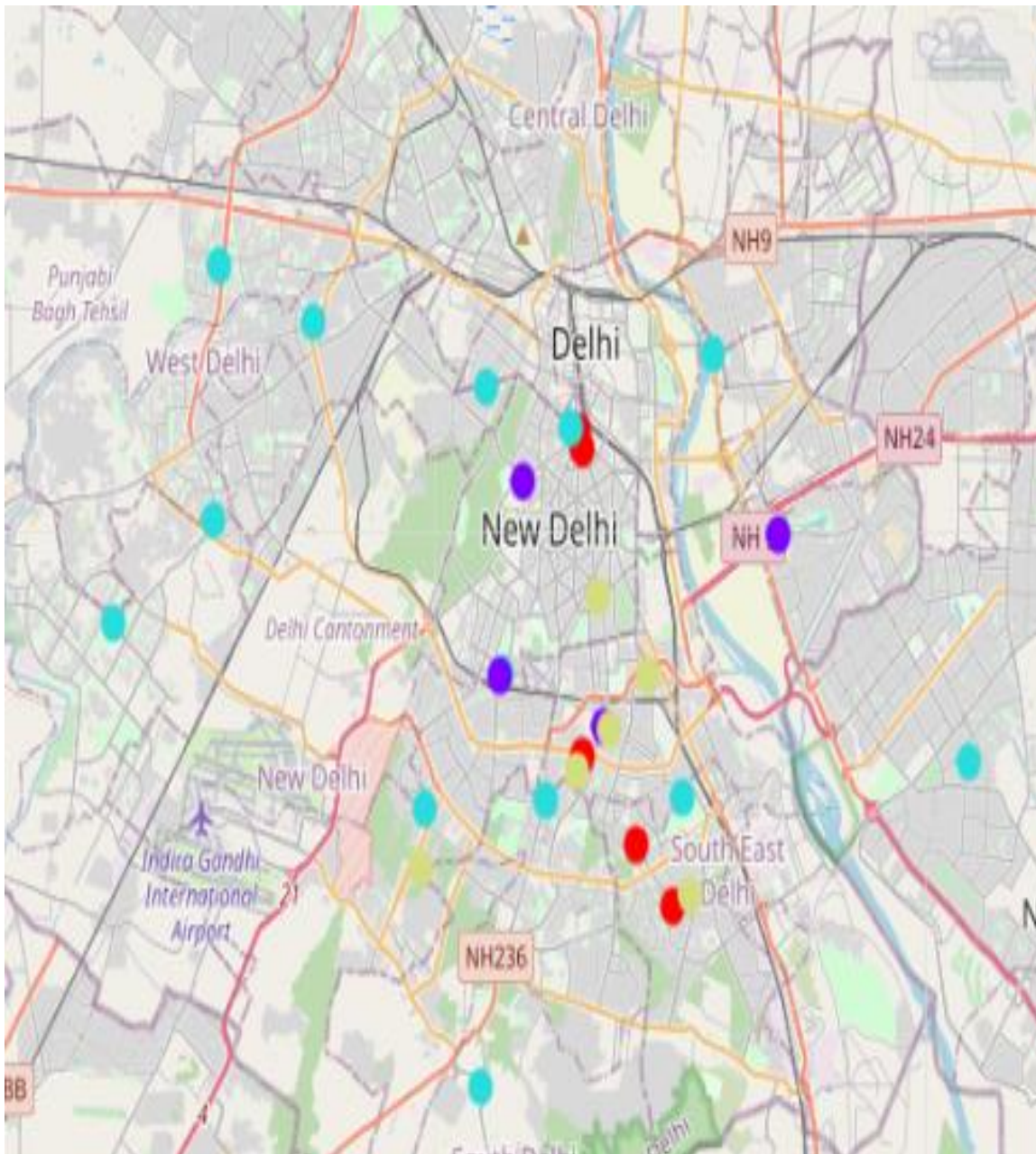


# DATA NORMALIZATION AND MODELING

- Data was normalized using standard scaler and then clustering was done using K-Means into the group of 4

## RESULTANT DATAFRAME

	AvgCostfor2	PriceRange	AverageUserRating	name	location.lat	location.lng	Category	label
0	1500	3.0	4.1	The Chinese	28.630593	77.220618	Chinese Restaurant	1
1	1500	3.0	4.0	Fujiya Japanese n Chinese restaurant	28.623762	77.199564	Asian Restaurant	1
2	750	2.0	3.1	Chinese Bite	28.584280	77.191714	Chinese Restaurant	0
3	1900	3.0	4.2	Chinese Garden, Yashwant Place	28.584106	77.191610	Restaurant	1
4	250	1.0	3.0	Chinese corner	28.642798	77.186733	Chinese Restaurant	0
5	1600	3.0	3.5	Zen Chinese, Connaught Place	28.633658	77.219495	Chinese Restaurant	1
6	400	1.0	4.0	Soy So Chinese	28.600367	77.226541	Indian Restaurant	2
7	300	1.0	3.5	Chinese Hut	28.583872	77.243970	College Administrative Building	2
8	350	1.0	3.9	Pioneer - The Chinese Restaurant	28.634521	77.216965	Chinese Restaurant	0
9	300	1.0	3.2	Chinese Inn	28.649473	77.267306	Chinese Restaurant	0
10	300	1.0	3.5	Chinese van C-Block EOK	28.559443	77.256739	Chinese Restaurant	0



# **LABELS OBTAINED**

**0**

**RED-HIGH USER RATING AND MEDIUM PRICE RANGE**

**1**

**PURPLE - HIGH PRICE RANGE AND HIGH USER RATING**

**2**

**CYAN - LOW COST + LOW RATING**

**3**

**YELLOW - LOW COST + MEDIUM RATING**



# DISCUSSIONS

- Label 0 seems like the best option if you have a medium budget. (Fried chicken spots)
- We can see clusters of Purple- Label 1 [Food truck, BBQ joint etc.] around south and east Delhi– being posh places, where people like to spend more. So if you have budget you should go for these locations.
- Label 2 has the lowest rating. So no point being in this category even though it is most crowded category.
- Label 3 is for Low Price Medium rating can be a good choice if you are running low on budget.



# CONCLUSION

Several parameters of the problem were analyzed and discussed successfully, and the question where and what type of Chinese Restaurant should you open was answered.

