Determining Cluster of Tehsils in Delhi for Setting up Fashion Retail Stores

1   INTRODUCTION

According to McKinsey & Company, India's apparel market alone is estimated to be of worth $59.3 Billion by 2022, making it the sixth largest in the world, comparable to the United Kingdom's ($65 billion) and Germany's ($63.1 billion) [1]. The growth rate of fashion industry in India is driven by its increasingly growing population and increasing spending behaviour of people. Factors like literacy rate of people, population and gender also highly affect the success of a fashion brand/retail stores in a particular location [2]. This project attempts to find clusters of Tehsils in Delhi where an existing/new Fashion Brands can set up their retail stores.

2   DATA

Details about Population, Literacy Rate and Sex Ratio of people living in each Tehsil of each District of Delhi is collected from the website of Census India [https://www.censusindia.co.in/states/delhi]. FourSquare API is used to get location coordinates of each Tehsil, venues in the neighbourhood of each Tehsil, user ratings of those venues etc. This data will be used to define clusters of Tehsils in which a fashion retail store can be set up.
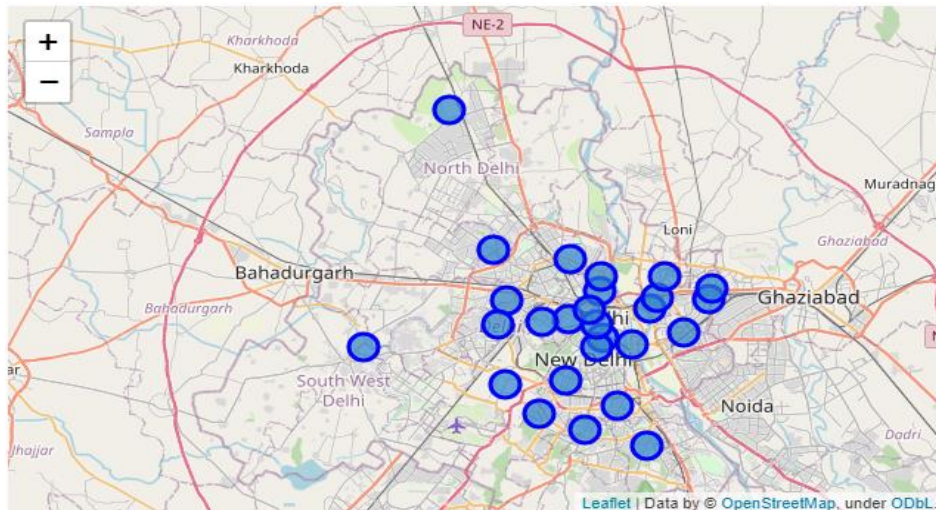
3   METHODOLOGY

3.1   Fetching Data

First we will import data about Tehsils of Delhi and fetch latitude and Longitude of each Tehsil through 'geocoder' class of 'Geopy' package of python.
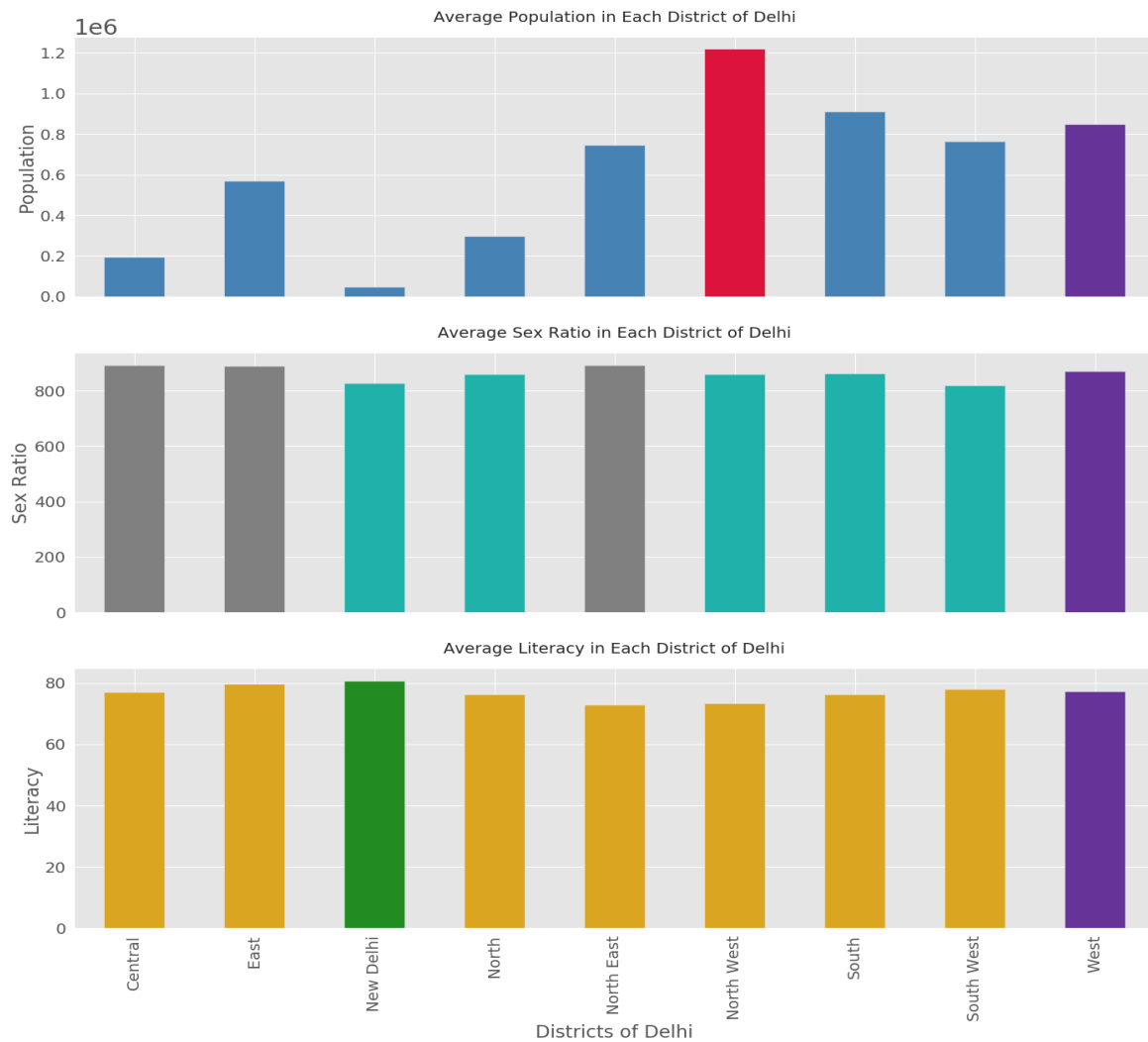
3.2   Visualizing Delhi on Map

Now, we will visualize all Tehsils on Map of Delhi as shown below:

Fig 1: Map of Delhi showing Locations of Tehsils

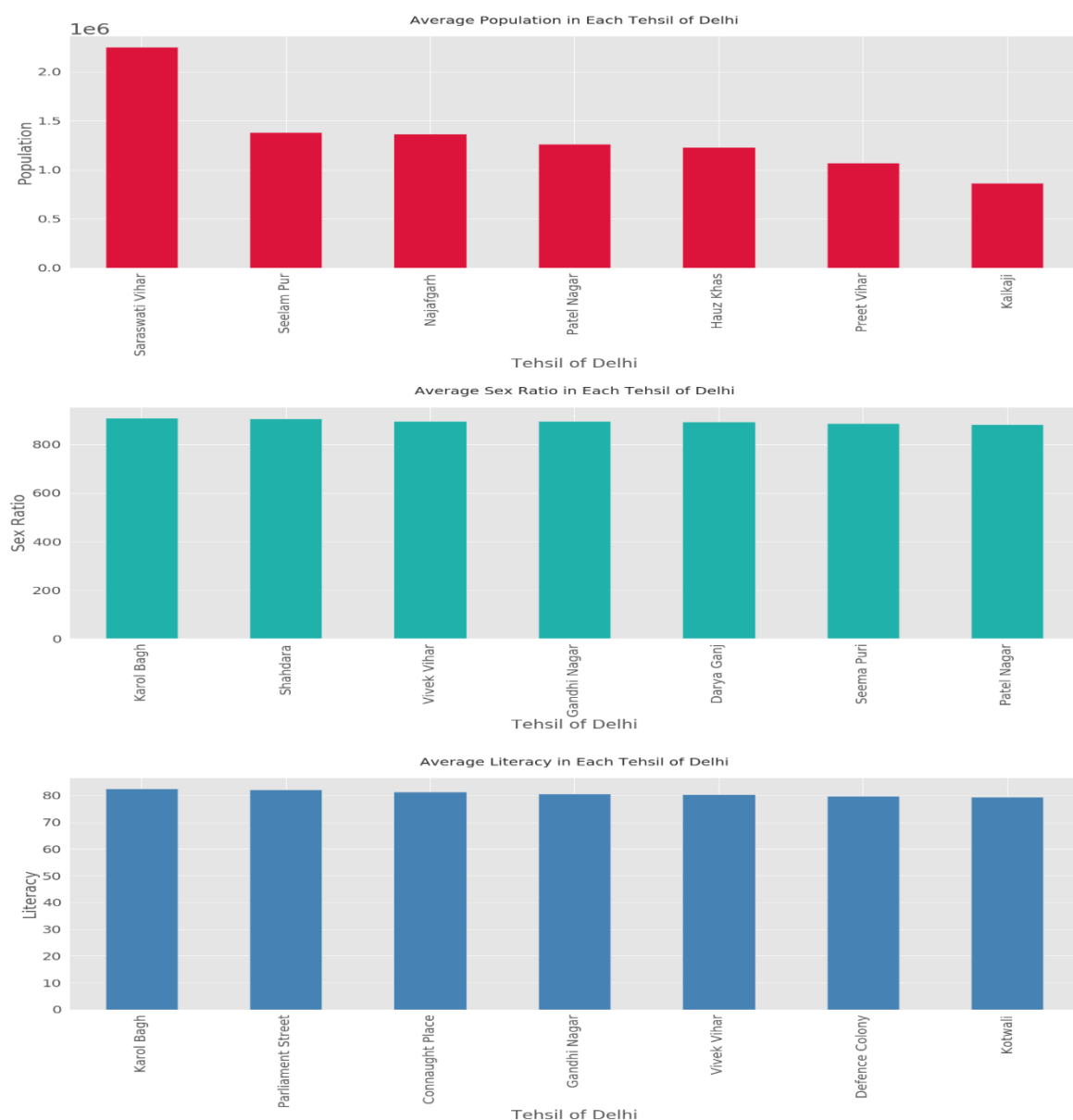## 3.3   Exploring Population, Literacy Rate and Sex Ratio by Districts of Delhi

Fig 2: Average Population, Sex Ratio and Literacy of each District of Delhi

From Fig 2, we observe that:

a) North West Delhi has highest population followed by South Delhi and then, West Delhi

b) Central Delhi, East Delhi and North East Delhi have highest Sex Ratio

c) New Delhi has highest Literacy Rate followed by East Delhi

d) Tehsils in West Delhi can be considered to be good locations for fashion store as they have good population size, good sex ratio and descent literacy rate

e) For those fashion stores that are aiming for niche market, Tehsils in East Delhi can be better locations to target as it has low population size, high literacy rate and moderate sex ratio.

3.4 Exploring Population, Literacy Rate and Sex Ratio by Tehsils of Delhi

Fig 3: Tehsils of Delhi having highest Average Population, Sex Ratio and Literacy

From Fig 3, we can observe that:

a. Karol Bagh has highest Literacy rate and Sex ratio. But its population size is not even in top seven
b. Tehsils which have high population do not have high Sex Ratio and Literacy rate

3.5   Exploring Venues in the Tehsils

Now, we will fetch 100 venues from FourSquare API for each Tehsil of Delhi within the radius of 1 km from their location as shown in Fig 4.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Chanakya Puri	28.59506	77.18573	Amour Bistro	28.601569	77.185923	Café
1	Chanakya Puri	28.59506	77.18573	Sanadige	28.601969	77.187020	Karnataka Restauran
2	Chanakya Puri	28.59506	77.18573	Nehru Park | नेहरू पार्क (Nehru Park)	28.591798	77.192860	Park
3	Chanakya Puri	28.59506	77.18573	Moti Mahal Delux	28.601677	77.187106	Indian Restauran
4	Chanakya Puri	28.59506	77.18573	Lázeez Affaire	28.602237	77.186044	Indian Restauran
5	Chanakya Puri	28.59506	77.18573	Sidewok	28.602500	77.186422	Asian Restauran
6	Chanakya Puri	28.59506	77.18573	FEZ Restaurant	28.602451	77.186323	Moroccan Restauran
7	Chanakya Puri	28.59506	77.18573	Sagar Ratna	28.595551	77.195764	Indian Restauran
8	Chanakya Puri	28.59506	77.18573	nU.Delhi QBA	28.601498	77.186332	Pub
9	Chanakya Puri	28.59506	77.18573	Fujiya	28.601526	77.186129	Chinese Restauran

Fig 4: Venues in Chanakyapuri, one of the Tehsils of Delhi

	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
Chanakya Puri	15	15	15	15	15	15
Civil Lines	12	12	12	12	12	12
Connaught Place	91	91	91	91	91	91
Darya Ganj	11	11	11	11	11	11
Defence Colony	54	54	54	54	54	54
Delhi Cantonment	4	4	4	4	4	4
Gandhi Nagar	5	5	5	5	5	5
Hauz Khas	94	94	94	94	94	94
Kalkaji	11	11	11	11	11	11
Karol Bagh	37	37	37	37	37	37
Kotwali	5	5	5	5	5	5
Model Town	14	14	14	14	14	14
Najafgarh	1	1	1	1	1	1
Narela	4	4	4	4	4	4
Pahar Ganj	50	50	50	50	50	50
Parliament Street	72	72	72	72	72	72

Fig 5: Number of Venues in some of the Tehsils of Delhi

## 3.6   Analysing Each Tehsil of Delhi

After analysis of venues, we will find out top 10 venues in each Tehsil of Delhi that people visit frequently as shown in Fig 6.

	Neighborhood	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	Chanakya Puri	Indian Restaurant	Restaurant	Asian Restaurant	Café	Chinese Restaurant	Hotel	Ka Re
1	Civil Lines	Indian Restaurant	Snack Place	Italian Restaurant	Grocery Store	Donut Shop	Pizza Place	Ch Re
2	Connaught Place	Indian Restaurant	Hotel	Café	Bar	Chinese Restaurant	Coffee Shop	Lo
3	Darya Ganj	Café	Indian Restaurant	Art Gallery	Fast Food Restaurant	Stadium	Art Museum	Pla
4	Defence Colony	Indian Restaurant	Italian Restaurant	Fast Food Restaurant	Café	Market	Afghan Restaurant	Fre Re
5	Delhi Cantonment	Convenience Store	Coffee Shop	Historic Site	Shopping Mall	Vegetarian / Vegan Restaurant	French Restaurant	Dir
6	Gandhi Nagar	Dessert Shop	Burger Joint	Indian Restaurant	Train Station	Hotel Pool	French Restaurant	Dh
7	Hauz Khas	Coffee Shop	Indian Restaurant	Café	Chinese Restaurant	Bar	Lounge	Ma
8	Kalkaji	Fast Food Restaurant	Indian Restaurant	Sandwich Place	Hotel	Market	Park	Piz Pla
9	Karol Bagh	Hotel	Indian Restaurant	Snack Place	Coffee Shop	Fast Food Restaurant	Food & Drink Shop	Ba

Fig 6: Top 10 Venues in each Tehsil of Delhi that People visit frequently

3.7 Clustering Tehsils

Now, we will group tehsils into cluster and identify potential ones among them to target for. For this, we will run k-means algorithm to form clusters among tehsil (here, we are forming 8 clusters [see Fig 8]).
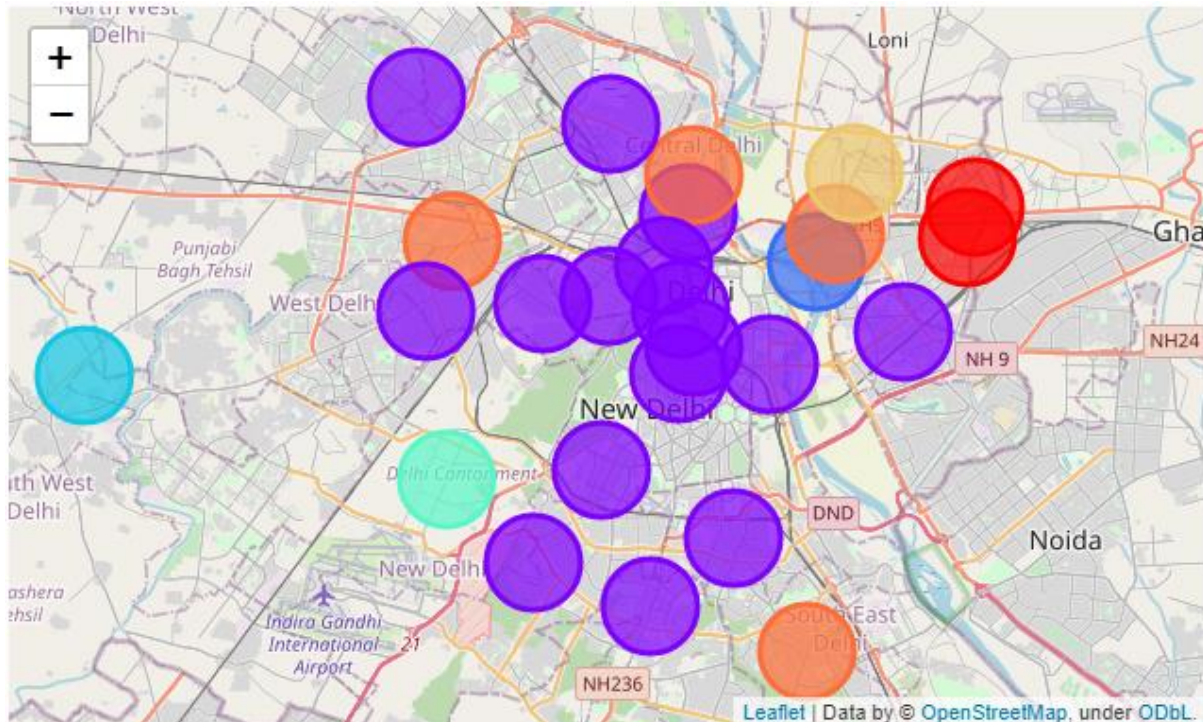
Fig 7: Map of Delhi showing Clusters of Tehsils

3.8 Examining Clusters

3.8.1 Cluster 1: East Delhi and North East Delhi

In cluster 1, tehsils that have Pizza Place, Indian Restaurant and, Shop and Service are clubbed together (See Fig 8).

Borough	Postal Code	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Com Venu
East	2081	28.6687	77.319	0	Indian Restaurant	Pizza Place	Diner	Track	Pool
North East	2079	28.6783	77.3221	0	Indian Restaurant	Light Rail Station	Pizza Place	Diner	Shop Serv

Fig 8: Cluster 1

3.8.2 Cluster 2: New Delhi, Central Delhi, North Delhi, East Delhi, South-West Delhi, West Delhi, South Delhi and North-West Delhi

In cluster 2, Tehsils with Indian/Chinese/Italian Restaurant, cafe/coffee shop, hotel and pubs are clubbed together. This is the biggest cluster found through k-means algorithm(See Fig 9).

	Borough	Postal Code	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Cor Ven
0	New Delhi	2071	28.5951	77.1857	1	Indian Restaurant	Restaurant	Asian Restaurant	Caf
1	New Delhi	2084	28.6339	77.2197	1	Indian Restaurant	Hotel	Café	Bar
2	New Delhi	2083	28.6255	77.2142	1	Hotel	Indian Restaurant	Chinese Restaurant	Bar
3	Central	2088	28.6283	77.2473	1	Café	Indian Restaurant	Art Gallery	Fas Res
4	Central	2086	28.6505	77.1887	1	Hotel	Indian Restaurant	Snack Place	Cof Sho
5	Central	2087	28.646	77.2149	1	Hotel	Indian Restaurant	Fast Food Restaurant	Pizz Pla
6	North	2074	28.6767	77.2177	1	Indian Restaurant	Snack Place	Italian Restaurant	Gro Sto
8	North	2075	28.6598	77.2085	1	Indian Restaurant	Jewelry Store	Miscellaneous Shop	Bre Spc
10	East	2082	28.639	77.296	1	Pizza Place	Indian Restaurant	South Indian Restaurant	Gro Sto
17	South West	2094	28.5649	77.1613	1	Pizza Place	Chinese Restaurant	Café	Cof Sho
18	West	2090	28.6478	77.1645	1	Pizza Place	Bakery	Park	Ligh Sta

Fig 9: Cluster 2

3.8.3 Cluster 3 and Cluster 4

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
2	Dessert Shop	Burger Joint	Indian Restaurant	Train Station	Hotel Pool	French Restaurant	Dhaba	Diner

Fig 10: Cluster 3

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
3	Resort	Vegetarian / Vegan Restaurant	Furniture / Home Store	Dhaba	Diner	Donut Shop	Event Space	Falafel Restaurant

Fig 11: Cluster 4

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
5	ATM	Furniture / Home Store	Multiplex	Dhaba	Diner	Donut Shop	Event Space	Falafel Restaurant

Fig 12: Cluster 6

The Cluster 3, 4 and 6 (see Fig 10, 11 & 12) have similar venues. So, they can be clubbed together. These clusters have Tehsils belonging to East Delhi, South – West Delhi and North – West Delhi.

3.8.4 Cluster 8

	Borough	Postal Code	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue
7	North	2076	28.689	77.2194	7	Metro Station	Snack Place	Pizza Place	Fast F Resta
12	North East	2077	28.6701	77.2712	7	Train Station	Shipping Store	Platform	Light Statio
19	West	2089	28.6677	77.1313	7	Fast Food Restaurant	Indian Restaurant	Pizza Place	Garde Cente
23	South	2097	28.5366	77.2609	7	Fast Food Restaurant	Indian Restaurant	Sandwich Place	Hotel

Fig 13: Cluster 6

In Cluster 6, Tehsils with Pizza Place and Fast Food Restaurant are clubbed together.
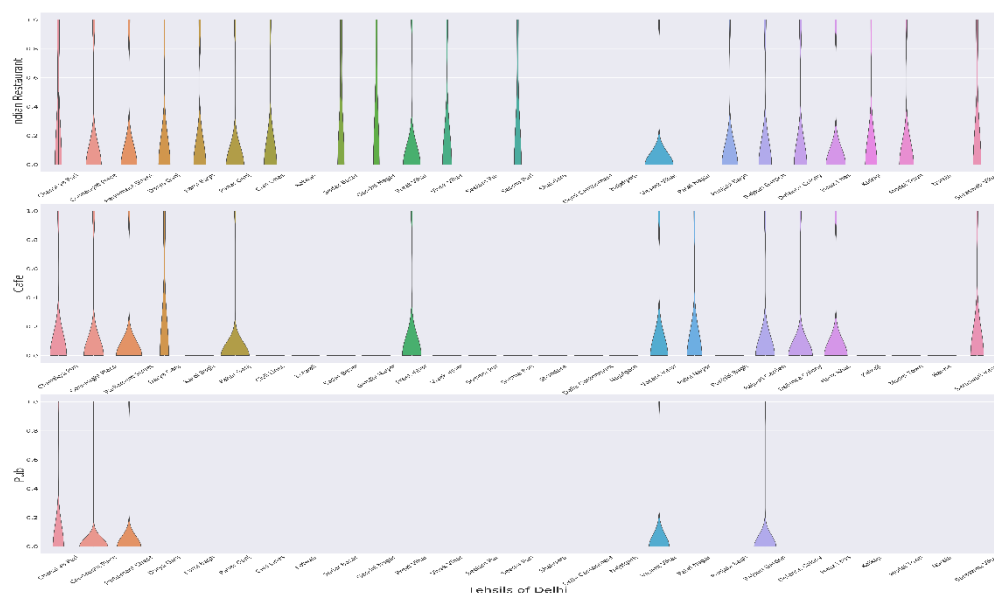
3.8.5   Cluster 5 and Cluster 7

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 Comr Venue
4	Convenience Store	Coffee Shop	Historic Site	Shopping Mall	Vegetarian / Vegan Restaurant	French Restaurant	Diner	Donu Shop

Fig 14: Cluster 5

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
6	ATM	Clothing Store	Gift Shop	Chocolate Shop	Motorcycle Shop	Park	History Museum	Historic Site

Fig 15: Cluster 7

The Cluster 5 and 7 (see Fig 14 & 15) have similar venues. So, they can be clubbed together. These clusters have Tehsils belonging to South – West Delhi and North – East Delhi.

3.9   Finding and Visualizing Prospective Tehsils Based on a Criteria

Suppose a client wants to set up fashion retail stores at Tehsils in Delhi which has Indian Restaurant, Cafe and Pubs. I've used a violin plot from the seaborn library - it is a great way to visualise frequency distribution datasets, they display a density estimation of the underlying distribution.

Based on above violin plots (Fig 16), Chanakya Puri, Connaught Place, Parliament Street, Vasant Vihar and Rajouri Garden each of which lie in cluster 1 (Fig 17).
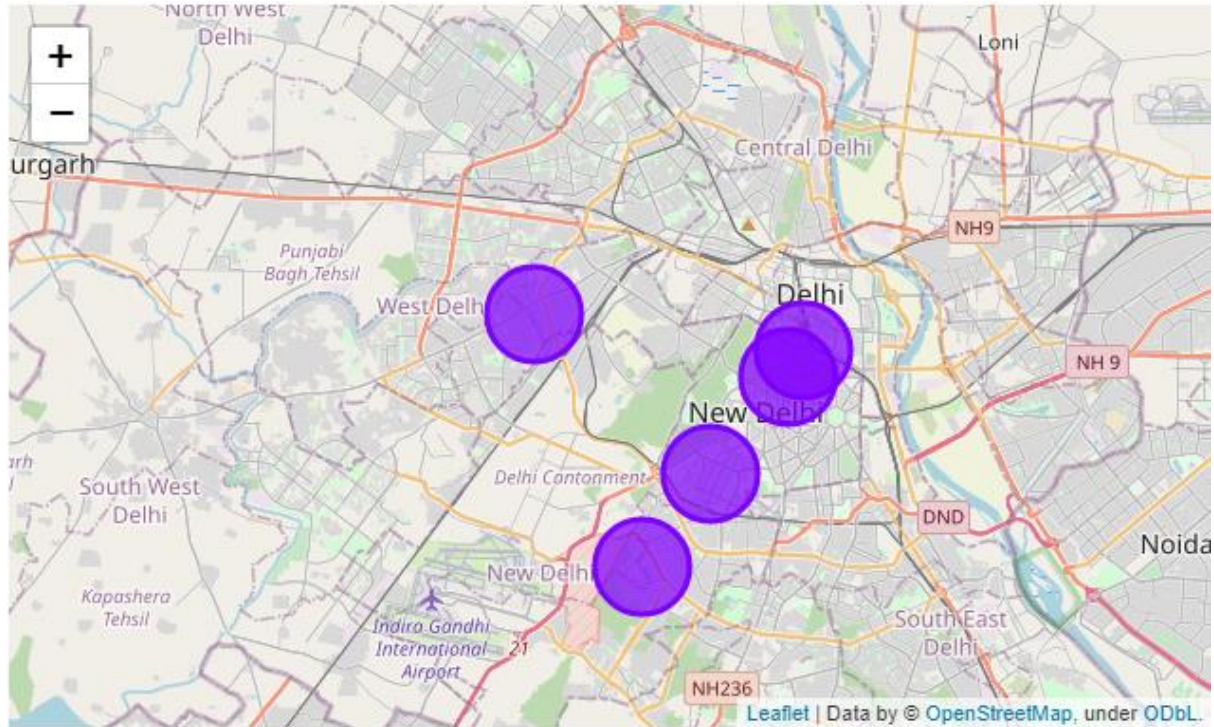
Fig 17: Map of Delhi showing Location of Chanakya Puri, Connaught Place, Parliament Street, Vasant Vihar and Rajouri Garden in Cluster 1

As West Delhi has good population size, good literacy rate and descent gender ratio, we can narrow down our choice for prospective locations for a fashion retail store to Rajouri Garden as it is in West Delhi District.

4   CONCLUSION

We conclude that based on certain parameters or criteria, we can select one or two clusters of Tehsils of Delhi through the above methodology for the establishment of a retail store.

There may be other ways this analysis could have been performed based on different methodolgy, different algorithm and perhaps different data sources. But I chose this method as the dataset related to tehsils in Delhi is readily available through Census India, and also, K-Means algorithm is easily able to be applied on this dataset.

5   LIMITATIONS

1. Area of each Tehsil can also be used for analysis.

2. Data related to centennials and millennials can be employed for decision-making and, thus can be used for further analysis.

3. Fashion Brands famous or trending in a particular tehsil can also be helpful to decide better locations for establishment of fashion retail stores.

6 REFERENCES

[1] Arvind Singhal, "India's Fashion and Retail Sector: An Overview", Brand Growth Inspiration, 12th April 2018

[2] Amed, Imran et. al., "How India's Ascent could change the Fashion Industry", McKinsey & Company – Our Insight, February 2019