

The next store location for Decathlon in Delhi

1. Introduction

Decathlon is a French sporting goods retailer. With over 1,500 stores in 57 countries, it is the largest sporting goods retailer in the world. Its holding company was formerly known as Oxyane. In 2017, the company posted a total revenue of US \$ 12.8 billion.

Decathlon has been rapidly expanding in many countries. At present, Decathlon has opened around 50 stores across India, with about 4 stores in Delhi itself. As per a news article published in 2014, Decathlon has an ambition of opening 100 stores in India. Decathlon has also invested significantly on online sale of its products.

India has a burgeoning middle class and it is infact growing at a rapid pace. India's middle and upper class in Tier 1 cities is estimated to be around 57% and is estimated to increase to about 85% in 2030. Aspirational brands like Decathlon sees a great opportunity in this market.

INDIAN HOUSEHOLD INCOME DEMOGRAPHICS (Tier 1 city)			
	2019		2030*
Category	Household income/month	% of population	% of population
Low Poverty Level	<90\$	43	15
Poor	< 250\$		
Lower middle class	<1200\$	54	78
Upper middle class	< 3000\$		
Affluent class	>3000\$	3	7
Very Rich	>12,000		

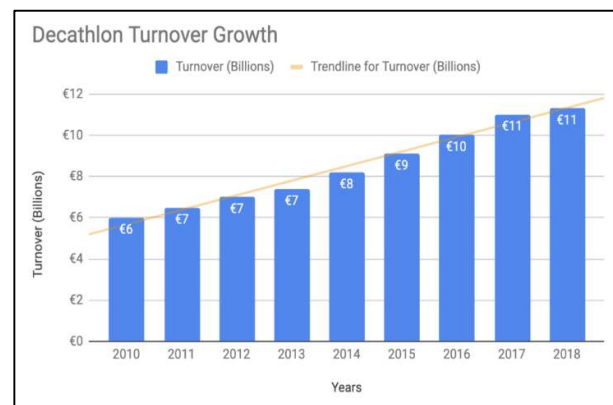
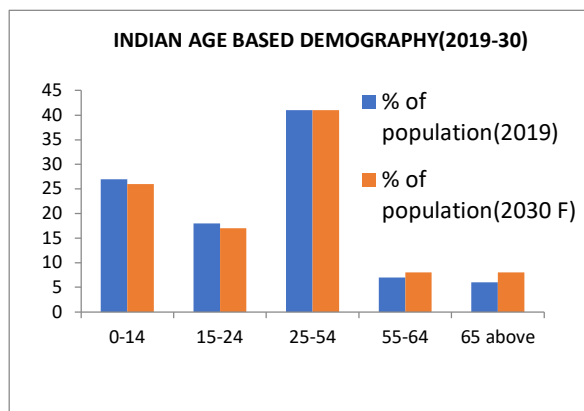


Fig-1 – Income brackets within India and growing sales of Decathlon

2. Business Requirement

Decathlon has four stores in Delhi presently. The plan is to open a fifth store but a decision needs to be taken where the location should be. The location of the new store needs to be a already popular destination where -people visit in large number. It can be selected based on availability of public destinations like shopping malls, restaurants etc. So here are the broad criteria for selection of the place-

1. The location should be a popular one
2. It should be far from the existing Decathlon store
3. There should be plenty of infrastructure around

3. Data Section

The approach is mainly to locate popular shopping malls in the city of Delhi. This will be done by using data from Wikipedia.

Link- https://en.wikipedia.org/wiki/List_of_shopping_malls_in_India#Delhi

First five Columns from the datasets are shown below-

Name	Location	Year	Size (gross leasable area)
Ambience Mall, Vasant Kunj	Vasant Kunj	2008	1,200,000 sq ft (110,000 m2)
Vegas Mall	Sector-14, Dwarka	2019	650,000 sq ft (60,000 m2)
Parsvnath Mall, Azadpur	Azadpur metro station	2005	600,000 sq ft (56,000 m2)
Select Citywalk	Sector-6, Pushp Vihar, Saket	2007	600,000 sq ft (56,000 m2)
Pacific Mall, Tagore Garden	Tagore Garden, Subhash Nagar	2011	600,000 sq ft (56,000 m2)

Table-1 – Shopping mall locations around Delhi

The second source of data used in the assignment is location of Decathlon stores in Delhi. This is being retrieved from another website. This data will be manually converted into a table. Also, some additional stores mentioned in Delhi Suburbs like in Gurgaon and Noida has been removed, and the study has been kept specific to the city of Delhi.

Link-<https://www.tiendeo.in/stores/delhi/decathlon>

The following table shows the data obtained from the website

Store	Address
Decathlon Khelgaon	Khel Gaon New Delhi
Decathlon CBD Shahdara	Shahdara
Decathlon Tagore Garden	Tagore Garden, Najafgarh
Decathlon Rohini	Rohini, New Delhi

Table2- Decathlon Store locations in Delhi

The third data source to be used is www.foursquare.com to explore the neighbourhood of the area selected for setting up a new store. This will be done using a REST API.

4. Methodology Section

4.1. Approach

The business problem defined above was solved based on following ideas-

1. The shopping malls in Delhi give a good idea about the popular localities in the city, generally surrounded by cafes, markets, hotels and often frequently visited by people with good disposable incomes.
2. Since Decathlon has few stores in Delhi, the best criteria to identify a locality is that it should be far from the existing stores to avoid customer redundancy so that an untapped geographical market segment can be reached.

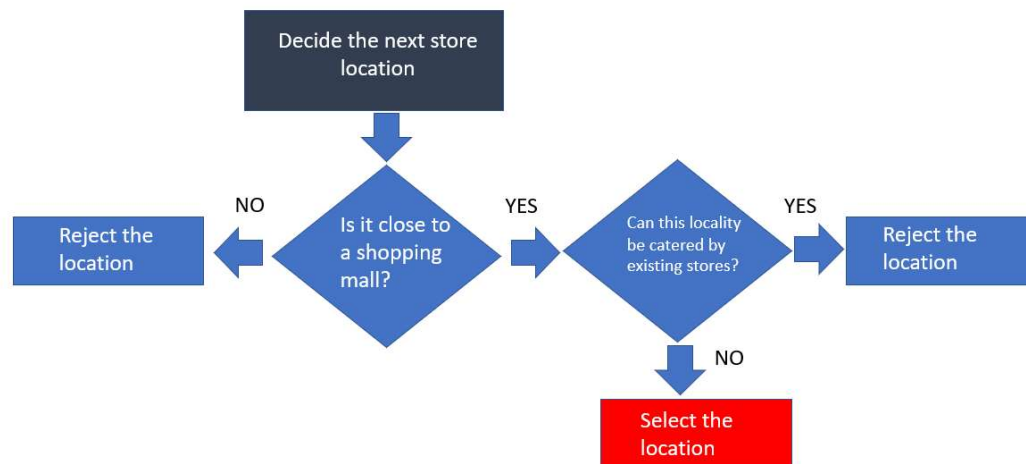


Fig 2- Flowchart showing the principle for deciding the next store location

4.2. Workflow

In this section, the details of the capstone project will be discussed. In a nutshell, following workflow was used to find the best location for the next store of Decathlon.

1. Data of Shopping malls across all neighbourhoods of Delhi loaded
2. Using GEOCODER retrieved latitude- longitude for all shopping malls in Delhi
3. Removed all extra columns from the Data frame that are not required
4. Data of all Decathlon store locations in Delhi loaded
5. Using GEOCODER retrieved latitude-longitude of the Decathlon stores
6. Plotted Decathlon stores and Shopping malls on a Delhi map
7. Using K Means clustering method divided all the shopping mall locations into FOUR major shopping and lifestyle zones
8. Based on distance (or distance index) from centroid of each cluster to each existing store, a comparative analysis of the relative distance of each cluster was ranked
9. The farthest popular location recommended for setting up the next store
10. Using Foursquare API, the neighbourhood of the selected locality explored for cafes, hotels etc

4.3. Exploratory Data Analysis & Data Wrangling

The data collected was not voluminous and simple, so it was quite simple to explore the data visually. However, some plots were used to visualize the data.

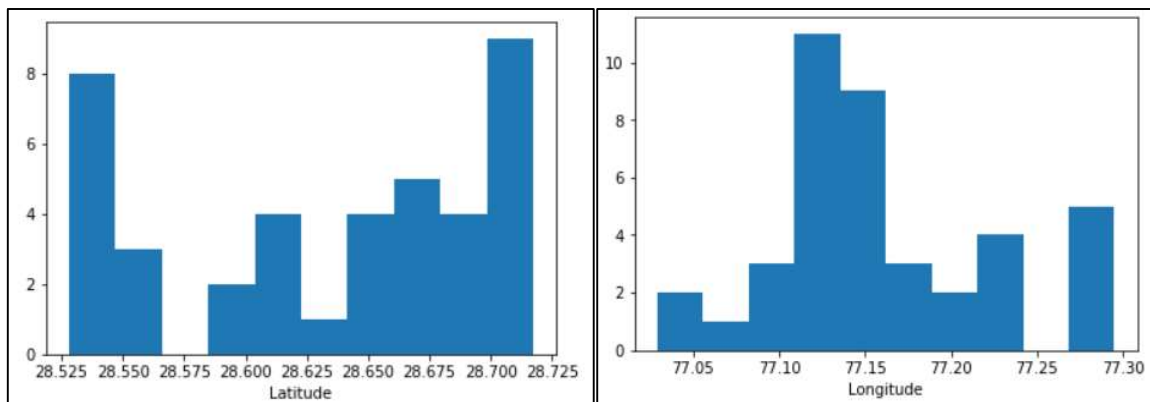
1. Data Description-

	Latitude	Longitude	Zone
count	40.000000	40.000000	40.000000
mean	28.632681	77.159411	0.900000
std	0.069833	0.069068	1.057331
min	28.527903	77.029238	0.000000
25%	28.551420	77.117074	0.000000
50%	28.647980	77.152640	1.000000
75%	28.696052	77.198090	1.250000
max	28.717453	77.294960	3.000000

2. Checking for missing values- There were 4 missing values in the Year column but it was not filled as the column was not useful.

<code>df_mall.isnull().sum()#Searching Nan values</code>	
Name	0
Location	0
Year	4
Size (gross leasable area)	0
Latitude	0
Longitude	0
Zone	0
dtype: int64	

3. Histogram of Latitude and Longitude- The histogram and latitude and longitude which shows there are clusters existing laterally but not longitudinally



4. Dropping columns- Some unnecessary columns were removed from Table 1 like Year and Floor area.

4.4. Data Visualization

Using folium, a map was created to observe the locations on a Delhi Map.

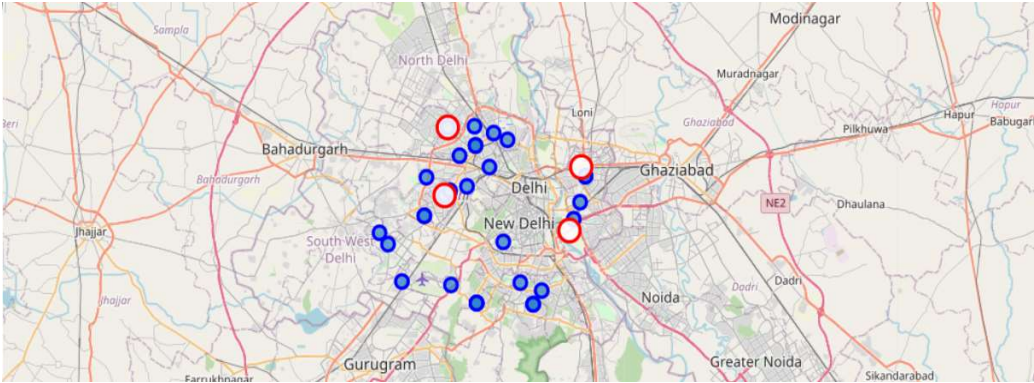


Fig 3- Delhi map with Shopping mall locations (blue) and Decathlon store locations(white)

After clustering the Shopping mall locations were clustered into 4 zones. The representation is as below-

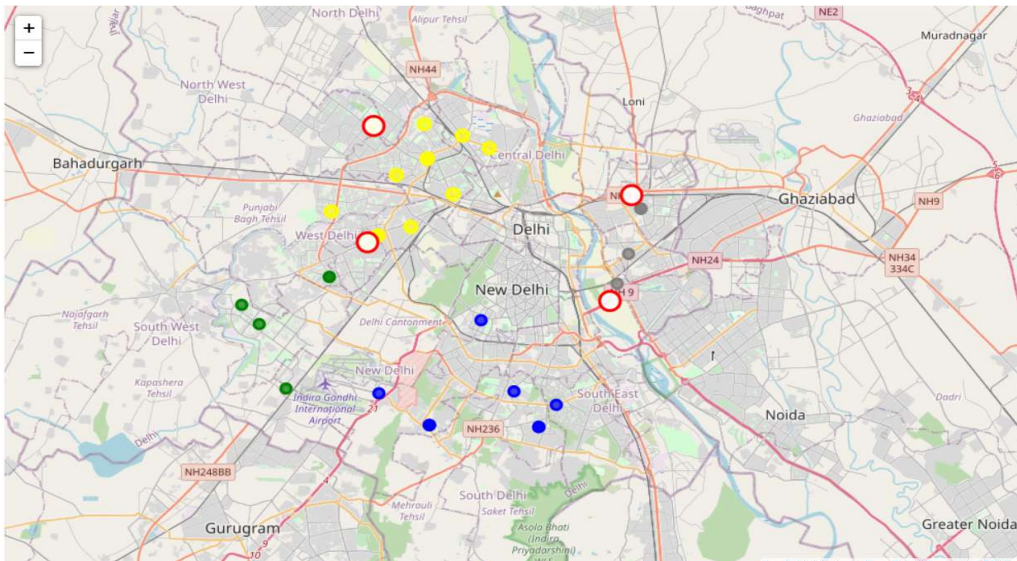


Fig 3- Delhi map with Shopping mall locations represented with different colours for clusters

4.5. Data Clustering

Using K Means unsupervised clustering technique, all shopping malls were divided into 4 clusters. Further the centroid (latitude/longitude) for each of the clusters were retrieved.

	Name	Location	Latitude	Longitude	Zone
0	Ambience Mall, Vasant Kunj	Vasant Kunj	28.529249	77.154134	1
1	Vegas Mall	Sector-14, Dwarka	28.604197	77.029238	2
2	Parsvnath Mall, Azadpur	Azadpur metro station	28.710353	77.175901	0
3	Select Citywalk	Sector-6, Pushp Vihar, Saket	28.527903	77.226713	1
4	Pacific Mall, Tagore Garden	Tagore Garden, Subhash Nagar	28.643764	77.112845	0
5	Moments Mall	Kirti Nagar	28.653281	77.141773	0

Centroid Coordinates

```
[ [28.69111491 77.13509586]
  [28.53966315 77.1868041 ]
  [28.59842046 77.06072112]
  [28.64952837 77.29023249]]
```

4.6. Summarizing Data for Interpretation

Since distance is an important parameter to determine farthest popular locations, a table was created by calculating Euclidian distance using coordinates. Since latitude and longitude cannot be directly used for directly finding distance qualitatively, it will be referred as 'Distance Index' to understand the nearness-farness in a qualitative sense.

	Store	Address	Latitude	Longitude	Zone0	Zone1	Zone2	Zone3
0	Decathlon Khelgaon	Khel Gaon New Delhi	28.606650	77.274260	0.162791	0.110163	0.213698	0.045756
1	Decathlon CBD Shahdara	Shahdara	28.673333	77.289025	0.154953	0.168276	0.240280	0.023836
2	Decathlon Tagor Garden	Tagore Garden,Najafgarh	28.643764	77.112845	0.052318	0.127698	0.069087	0.177481
3	Decathlon Rohini	Rohini,New Delhi	28.716209	77.117074	0.030895	0.189818	0.130575	0.185553

Using the following table total distance for each column for Zone0,1,2 and 3 were calculated.

5. Results

Based on the workflow discussed, the distance index for the four zones were determined.



Fig-4- The distance index from the 4 zones

The plot shows that West Delhi is the best place to set up a new store in Delhi as it is both popular destination and farthest from the existing stores. South Delhi can be the second most favourable location.

West Delhi neighbourhood was further explored for other public amenities, to ensure any biases in the findings as the only factor to identify popular cluster zones were done using shopping mall data. Using Foursquare Rest API, the other amenities were looked into using Venue search.

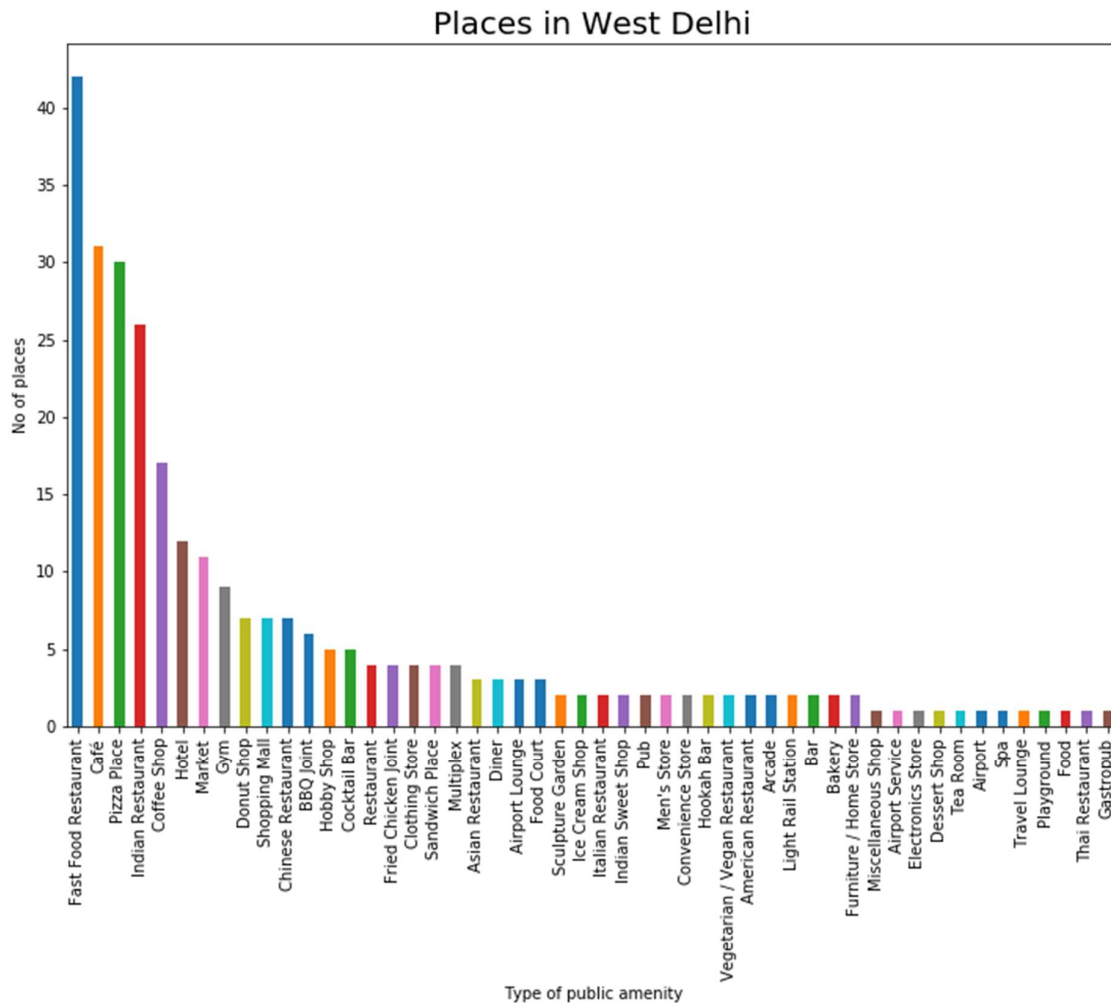


Fig-5- The other amenities in West Delhi

The plot represents that West Delhi is full of Cafes, Restaurants, hotels, gyms and other public places and seems to be a very popular destination.

6. Conclusion

In this capstone project, with publicly available data from Four Square, Wikipedia and other sources, we have been able to devise a methodology to determine location of a store for future expansion using a simple logic. The same technique can be improvised and further sharpened by using more comprehensive datasets by including parameters like Average Age, Median income, rental of commercial space etc. of the neighbourhood.