



# Applied Data Science Capstone

THE BATTLE OF NEIGHBOURHOODS (WEEK 2)

# Introduction

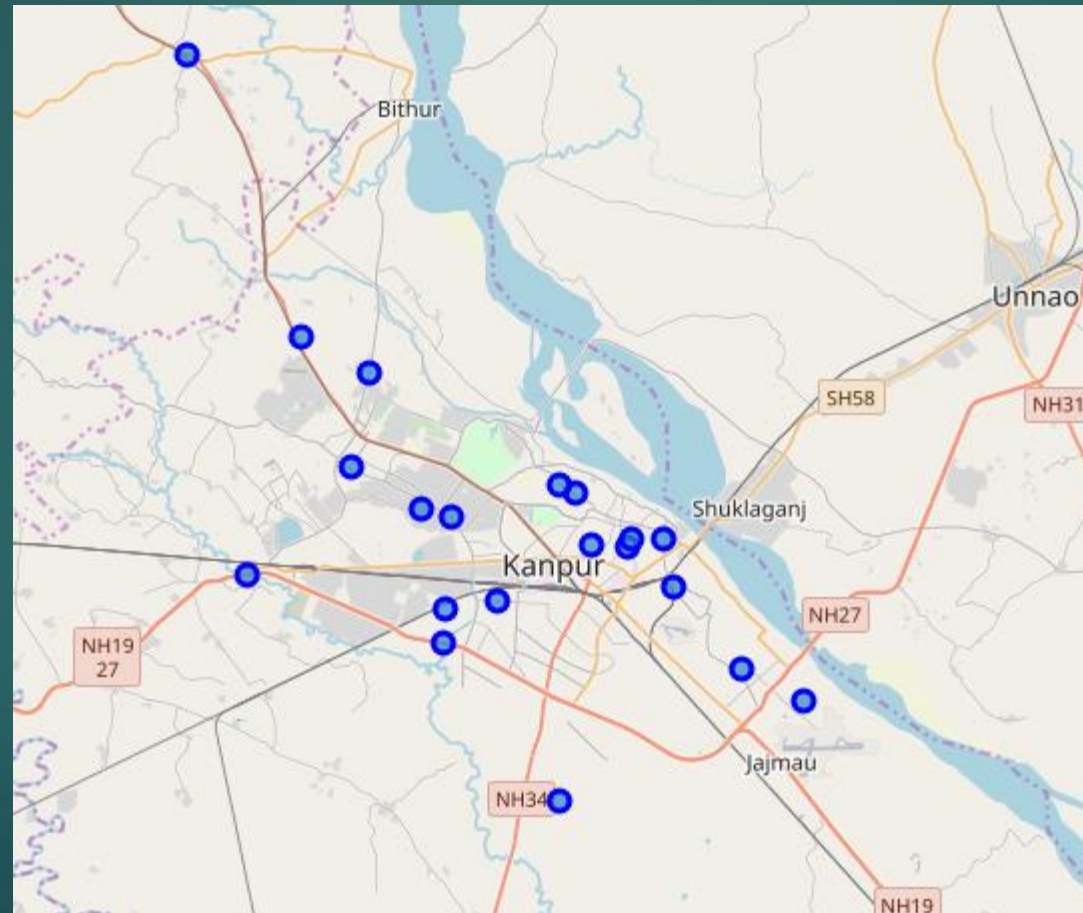
- ▶ Mall industry has emerged as one of the growing sector in India.
- ▶ Visiting shopping malls has become a great way to relax and enjoy for the visitors during weekends and holidays. They can do grocery shopping, dine at restaurants, shop at the various fashion outlets, watch movies and perform many more activities. Shopping malls are like a one-stop destination for all types of shoppers.
- ▶ For retailers, the central location and the large crowd at the shopping malls provides a great distribution channel to market their products and services.
- ▶ Property developers are also taking advantage of this trend to build more shopping malls to cater to the demand.
- ▶ Of course, as with any business decision, opening a new shopping mall requires serious consideration and is a lot more complicated than it seems. Particularly, the location of the shopping mall is one of the most important decisions that will determine whether the mall will be a success or a failure.
- ▶ **This project is particularly useful to property developers and investors looking to open or invest in new shopping malls in the capital city of India i.e. Kanpur. It will help the developer to decide which areas which have potential to become a relaxing center for the people of the city and financially wise decision for the investors.**

# Data

- ▶ Data of 21 neighborhoods in Kanpur is scraped from [Wikipedia](#) with the help of Python requests and BeautifulSoup packages.
- ▶ Foursquare API is used to get the venue data for the neighborhoods. Now since the response contains many fields which are not required, hence proper cleaning of the response is required to remove irrelevant information.

# Methodology

- Possible locations for the shopping mall in the neighbourhoods in Kanpur :



- ▶ Using FourSquare API, it was found that 41 unique venue categories in Kanpur, with 'Shopping Mall' being one of them.

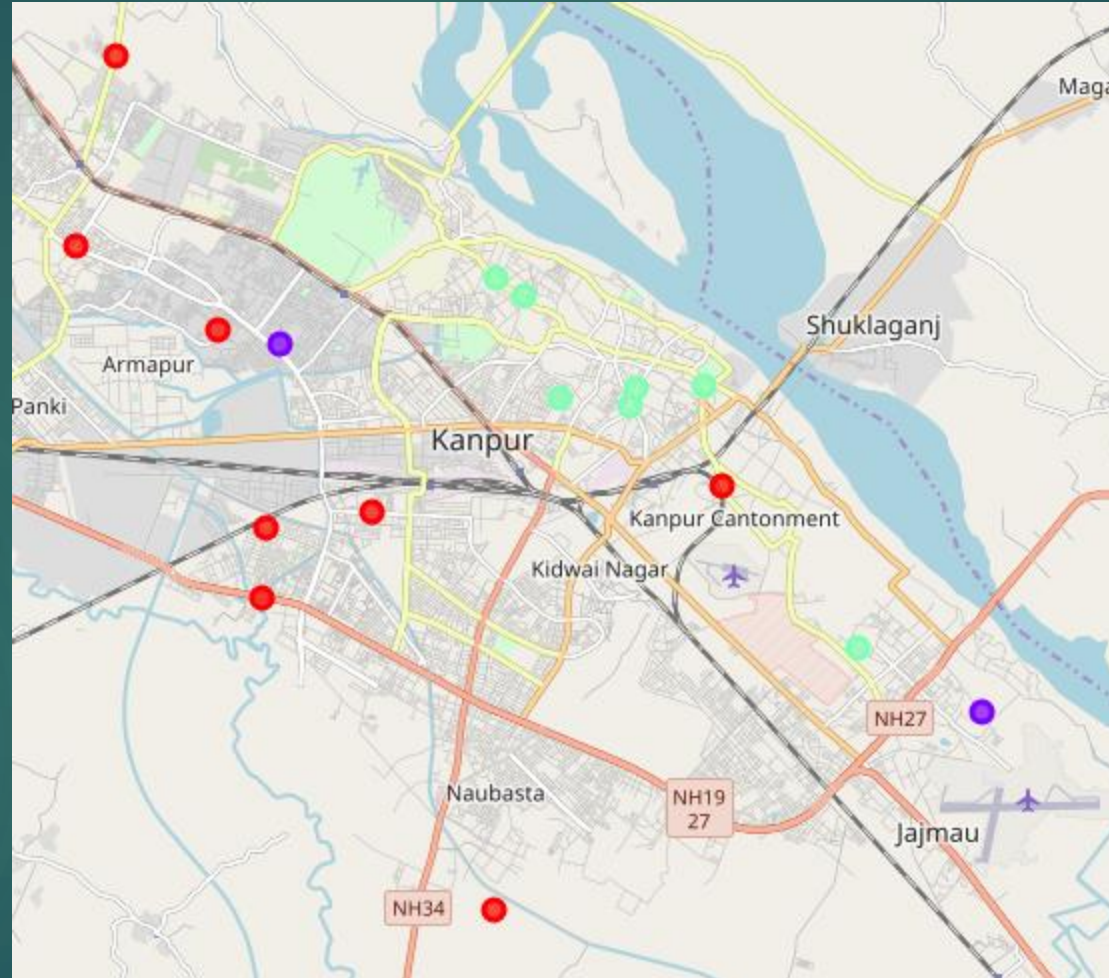
```
array(['Breakfast Spot', 'Business Service', 'ATM', 'Diner', 'Bar', 'Spa',  
      'Bowling Alley', 'Shopping Mall', 'Bookstore', 'Hotel',  
      'Indian Restaurant', 'Golf Course', 'Convenience Store',  
      'Food Court', 'Cricket Ground', 'Dessert Shop', 'Café',  
      'Fast Food Restaurant', 'Clothing Store', 'Pizza Place',  
      'Multiplex', 'Coffee Shop', 'Market', 'Bus Station',  
      'Jewelry Store', 'Train Station', 'Airport', 'Snack Place',  
      'Bakery', 'Pool', 'Garden Center', 'Cosmetics Shop', 'Men's Store',  
      'Chinese Restaurant', 'Camera Store', 'Furniture / Home Store',  
      'Accessories Store', 'Park', 'Tea Room', 'Gym', 'Ice Cream Shop'],  
      dtype=object)
```

- ▶ Data was prepared using OneHotEncoding to explore each neighbourhood across all the venue categories
- ▶ K-Means was applied to get different clusters based on frequency of shopping mall in the neighbourhood :

	Neighborhood	Shopping Mall	Cluster Labels	Latitude	Longitude
0	Barra, Kanpur	0.0000	0	26.436512	80.284110
1	Birhana Road, Kanpur	0.1250	2	26.428658	80.386283
2	Bithoor	0.0000	0	26.519671	80.258794
3	Chaman Ganj, Kanpur	0.0625	2	26.467179	80.334996
4	Govind Nagar	0.0000	0	26.449520	80.302740



► Output after clustering :



# Results

## ► Cluster -1 :

	Neighborhood	Shopping Mall	Cluster Labels	Latitude	Longitude
0	Barra, Kanpur	0.0	0	26.436512	80.284110
14	Rawatpur	0.0	0	26.477770	80.276340
13	Ratan Lal Nagar	0.0	0	26.447290	80.284820
12	Padri Lalpur	0.0	0	26.388380	80.323780
7	Kalyanpur, Uttar Pradesh	0.0	0	26.490680	80.252150
8	Kanpur Cantonment	0.0	0	26.453790	80.362830
4	Govind Nagar	0.0	0	26.449520	80.302740
2	Bithoor	0.0	0	26.519671	80.258794

► Cluster - 2 :

	Neighborhood	Shopping Mall	Cluster Labels	Latitude	Longitude
6	Kakadeo	0.25	1	26.47543	80.28700
5	Jajmau	0.25	1	26.41890	80.40744

► Cluster – 3 :

	Neighborhood	Shopping Mall	Cluster Labels	Latitude	Longitude
15	The Mall, Kanpur	0.050000	2	26.468932	80.359776
9	Latouche Road, Kanpur	0.047619	2	26.466097	80.347061
10	McRobertganj	0.052632	2	26.482900	80.329010
11	Meston Road, Kanpur	0.050000	2	26.468861	80.348330
3	Chaman Ganj, Kanpur	0.062500	2	26.467179	80.334996
1	Birhana Road, Kanpur	0.125000	2	26.428658	80.386283
16	VIP Road, Kanpur	0.083333	2	26.485630	80.323988



# Discussion

- ▶ As per the clusters observed after K-Means, cluster 1 is the best candidate to open a shopping mall.
- ▶ Since, the data is only for 21 neighbours, it will be a good test to see how the model performs when trained with data of more neighbours. This can be seen has future enhancement in the current project.

# Conclusion

- ▶ Most of the shopping malls in Kanpur City are concentrated in cluster 2 with moderate in cluster 2 and none in cluster 1, hence for the new builders, it will be a good opportunity to open shopping mall in cluster 1 since there will be no competition which lead to benefit for both the vendors and for the people living in the near by areas.