

Finding the Safest area in India for whom want to relocate in india

Saris kiattithapanayong

April 11, 2020

1 Introduction

This project aims to select the safest area in india based on the total crimes, explore the neighborhoods of that borough to find the 10 most common venues in each neighborhood and finally cluster the neighborhoods using k-mean clustering.

This report will be targeted to people who are looking to relocate to India. Inorder to finalise a neighborhood to hunt for an apartment, safety is considered as a top concern when moving to a new place. If you don't feel safe in your own home, you're not going to be able to enjoy living there. The crime statistics will provide an insight into this issue.

We will focus on the safest Area and explore its neighborhoods and the most common venues in each neighborhood so that the best neighborhood suited to an individual's needs can be selected.

1.1 Interest

Expats who are considering to relocate to India will be interested to identify the safest borough in London and explore its neighborhoods and common venues around each neighborhood.

2 Data Acquisition and Cleaning

2.1 Data Sources

The data acquired for this project is a combination of data from three sources. The first data source of the project uses a https://www.kaggle.com/rajanand/crime-in-india#10_Property_stolen_and_recovered.csv that shows the crime per area in India.

The second source of data is scraped from https://www.mapsofworld.com/lat_long/india-lat-long.html page that contains the area with each area's coordinates This page contains additional information about the area

2.2 Data Selection and Cleaning

Source 1 cleaning: sum up each type of crime in each area can see in following picture

	Area_Name	Year	Group_Name	Sub_Group_Name	Cases_Property_Recovered	Cases_Property_Stolen	Value_of_Property_Recovered	Value_of_Property_Stolen
0	Andaman & Nicobar Islands	2001	Burglary - Property	3. Burglary	27	64	755858	1321961
1	Andhra Pradesh	2001	Burglary - Property	3. Burglary	3321	7134	51483437	147019348
2	Arunachal Pradesh	2001	Burglary - Property	3. Burglary	66	248	825115	4931904

	Area	Year	Group	Sub-group-name	Cased_recovered	Cased_stolen	Value_recovered	Value_stolen
0	Andaman & Nicobar Islands	2001	Burglary - Property	3. Burglary	27	64	755858	1321961
1	Andhra Pradesh	2001	Burglary - Property	3. Burglary	3321	7134	51483437	147019348
2	Arunachal Pradesh	2001	Burglary - Property	3. Burglary	66	248	825115	4931904
3	Assam	2001	Burglary - Property	3. Burglary	539	2423	3722850	21466955
4	Bihar	2001	Burglary - Property	3. Burglary	367	3231	2327135	17023937

Cased_stolen							
Sub-group-name	1. Dacoity	2. Robbery	3. Burglary	4. Theft	5. Criminal Breach of Trust	6. Other Property	7. Total Property Stolen & Recovered
Area							
Andaman & Nicobar Islands	9	71	722	965	64	0	1831
Andhra Pradesh	1620	6185	79444	226763	3581	3818	321411
Arunachal Pradesh	227	714	2382	4362	380	251	8316
Assam	3599	5723	27600	67772	6204	11882	122780
Bihar	10002	21016	32898	119387	12069	10548	205920

	Area	Cased_stolen1. Dacoity	Cased_stolen2. Robbery	Cased_stolen3. Burglary	Cased_stolen4. Theft	Cased_stolen5. Criminal Breach of Trust	Cased_stolen6. Other Property	Cased_stolen7. Total Property Stolen & Recovered	Total
0	Andaman & Nicobar Islands	9	71	722	965	64	0	1831	3662
1	Andhra Pradesh	1620	6185	79444	226763	3581	3818	321411	642822
2	Arunachal Pradesh	227	714	2382	4362	380	251	8316	16632
3	Assam	3599	5723	27600	67772	6204	11882	122780	245560
4	Bihar	10002	21016	32898	119387	12069	10548	205920	411840

	Area	Dacoity	Robbery	Burglary	Theft	Criminal_breach_of_trust	other	Total
0	Andaman & Nicobar Islands	9	71	722	965	64	0	3662
1	Andhra Pradesh	1620	6185	79444	226763	3581	3818	642822
2	Arunachal Pradesh	227	714	2382	4362	380	251	16632
3	Assam	3599	5723	27600	67772	6204	11882	245560
4	Bihar	10002	21016	32898	119387	12069	10548	411840
5	Chandigarh	22	266	2427	15423	352	1370	39720
6	Chhattisgarh	1336	4231	37774	52217	1276	3022	199712
7	Dadra & Nagar Haveli	29	26	344	633	143	146	2642
8	Daman & Diu	29	14	423	461	54	47	2056
9	Delhi	320	5291	20421	186615	4047	28653	490694
10	Goa	45	235	2871	5928	336	479	19788
11	Gujarat	3225	11381	49852	179942	11748	10882	534060
12	Haryana	1002	4498	36315	96321	6158	7374	303336
13	Himachal Pradesh	42	204	7469	7773	821	691	34000
14	Jammu & Kashmir	159	1119	14085	20501	787	802	74906
15	Jharkhand	5157	7328	14003	56772	2747	2427	176868
16	Karnataka	2154	12946	61364	161535	3873	5612	494968
17	Kerala	1079	6091	38143	50503	2652	12358	221652
18	Lakshadweep	3	0	31	101	1	48	342
19	Madhya Pradesh	1430	18420	116445	218058	4306	8103	733524
20	Maharashtra	6517	26061	144520	453712	15896	41701	1376814
21	Manipur	39	73	650	3212	272	1546	11584
22	Meghalaya	630	840	1548	4605	193	546	16724
23	Mizoram	36	91	4017	8347	206	749	26892
24	Nagaland	103	762	871	2752	104	815	10814

]:

	Area	Dacoly	Robbery	Burglary	Thef	Criminal_breach_of_trust	other	Total
0	Andaman & Nicobar Islands	9	71	722	965		64	0 1831
1	Andhra Pradesh	1620	6185	79444	226763		3581	3818 321411
2	Arunachal Pradesh	227	714	2382	4362		380	251 8316
3	Assam	3599	5723	27600	67772		6204	11882 122780
4	Bihar	10002	21016	32898	119387		12069	10548 205920
5	Chandigarh	22	266	2427	15423		352	1370 19860
6	Chhattisgarh	1336	4231	37774	52217		1276	3022 99856
7	Dadra & Nagar Haveli	29	26	344	633		143	146 1321
8	Daman & Diu	29	14	423	461		54	47 1028
9	Delhi	320	5291	20421	186615		4047	28653 245347

SOURCE 2 CLEANING: BRING EACH AREA COORDINATES

	location	latitude	longitude	Loc
0	Abhayapuri	26°24'N	90°38'E	
1	Abohar	30°10'N	74°10'E	
2	Abu	24°41'N	72°50'E	
3	Achalpur	21°22'N	77°32'E	
4	Acher	23°10'N	72°32'E	
...
1631	Yadgir	16°45'N	77°5'E	
1632	Yamuna	25°30'N	81°53'E	
1633	Yamunagar	30°7'N	77°17'E	
1634	Yavatmal	20°20'N	78°15'E	
1635	Zaskar Mts.	33°15'N	77°30'E	

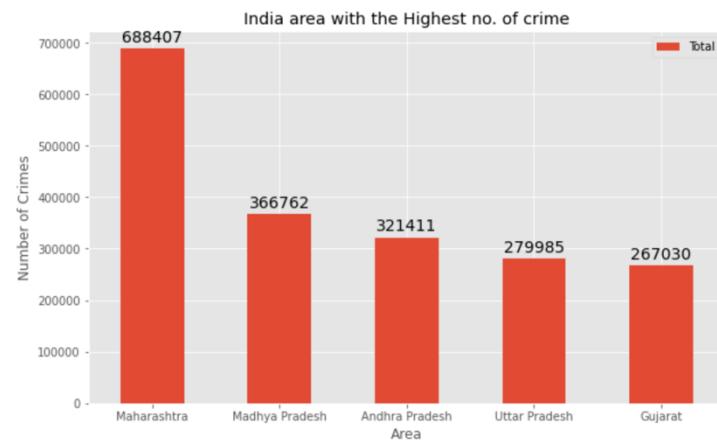
Then merge source 1 and 2 together

	Area	Dacolty	Robbery	Burglary	Thef	Criminal_breach_of_trust	other	Total	latitude	longitude
0	Andhra Pradesh	1620	6185	79444	226763		3581	3818	321411	18°00'N 79°00'E
1	Arunachal Pradesh	227	714	2382	4362		380	251	8316	28°00'N 95°00'E
2	Assam	3599	5723	27600	67772		6204	11882	122780	26°00'N 93°00'E
3	Bihar	10002	21016	32898	119387		12069	10548	205920	25°05'N 85°40'E
4	Bihar	10002	21016	32898	119387		12069	10548	205920	25°00'N 86°00'E
5	Chandigarh	22	266	2427	15423		352	1370	19860	30°43'N 76°47'E
6	Chhattisgarh	1336	4231	37774	52217		1276	3022	99856	22°00'N 82°00'E
7	Dadra & Nagar Haveli	29	26	344	633		143	146	1321	20°05'N 73°00'E
8	Delhi	320	5291	20421	186615		4047	28653	245347	28°38'N 77°17'E
9	Goa	45	235	2871	5928		336	479	9894	15°33'N 73°59'E

3. Methodology

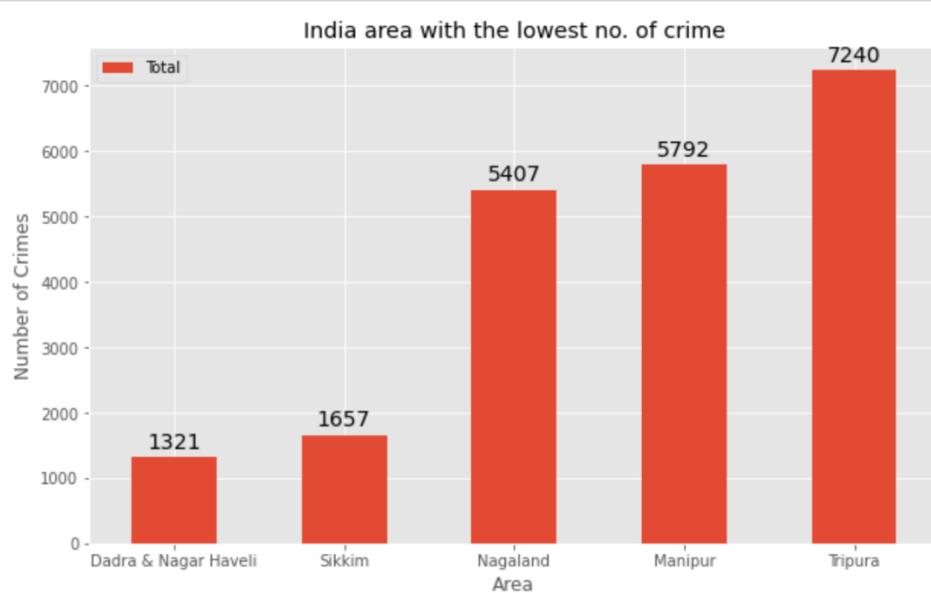
- Area with the highest crime rate from year 2001 – 2010

Comparing five boroughs with the highest crime rate during the year 2001 to 2010 it is evident that Maharashtra has the highest crimes recorded



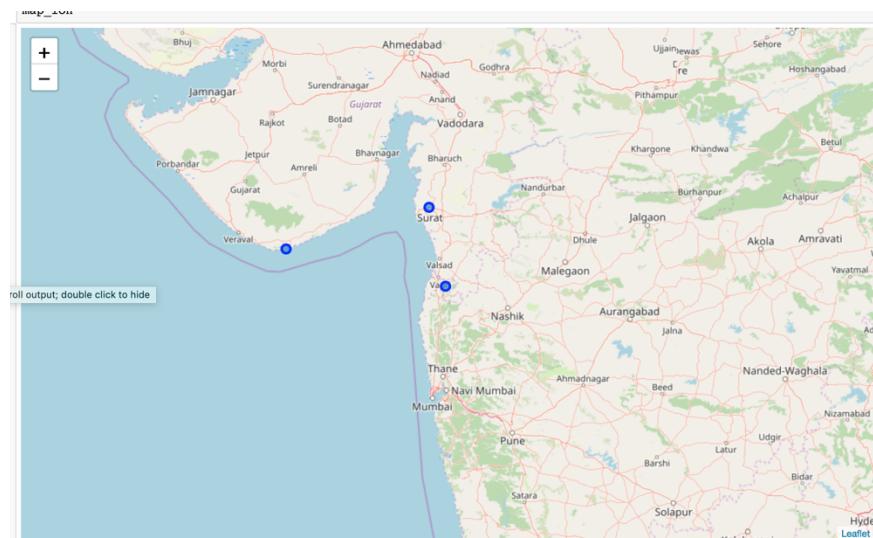
-Areas with the lowest crime rates

Comparing five boroughs with the lowest crime rate during the year 2001 to 2010, Dadra and naveli has the lowest recorded crimes



Neighborhoods in Dara and naveli

There are 3 neighborhoods in the royal borough of Dara Naveli, they are visualised on a map using folium on python



4. Results

After running the K-means clustering we can access each cluster created to see which neighborhoods were assigned to each of the two clusters.

Neighborhood	Area	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	10th Most Common Venue
0 Daman District	Dadra & Nagar Haveli	21.204498	72.799187	1	Fast Food Restaurant	Art Gallery	Hotel	Food Truck	Fish & Chips Shop	Café	Bus Station	Bakery	Asian Restaurant	ATM
1 Diu District	Dadra & Nagar Haveli	20.715953	70.978500	0	Hotel	Fish & Chips Shop	Bus Station	ATM	Food Truck	Fast Food Restaurant	Café	Bakery	Asian Restaurant	Gr
2 Dadra and Nagar Haveli District	Dadra & Nagar Haveli	20.273450	73.006472	0	Food Truck	Café	Bakery	Asian Restaurant	Hotel	Fish & Chips Shop	Fast Food Restaurant	Bus Station	Art Gallery	Gr

Analysis

Analyse each of the clusters to identify the characteristics of each cluster and the neighborhoods in them.

Examine the first cluster

```
|: kut_merged[kut_merged['Cluster Labels'] == 0]
```

```
|:
```

Neighborhood	Area	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	10th Most Common Venue
1 Diu District	Dadra & Nagar Haveli	20.715953	70.978500	0	Hotel	Fish & Chips Shop	Bus Station	ATM	Food Truck	Fast Food Restaurant	Café	Bakery	Asian Restaurant	Gal
2 Dadra and Nagar Haveli District	Dadra & Nagar Haveli	20.273450	73.006472	0	Food Truck	Café	Bakery	Asian Restaurant	Hotel	Fish & Chips Shop	Fast Food Restaurant	Bus Station	Art Gallery	A

The first cluster has one neighborhood which consists of Venues such as Hotel, and Food Truck.

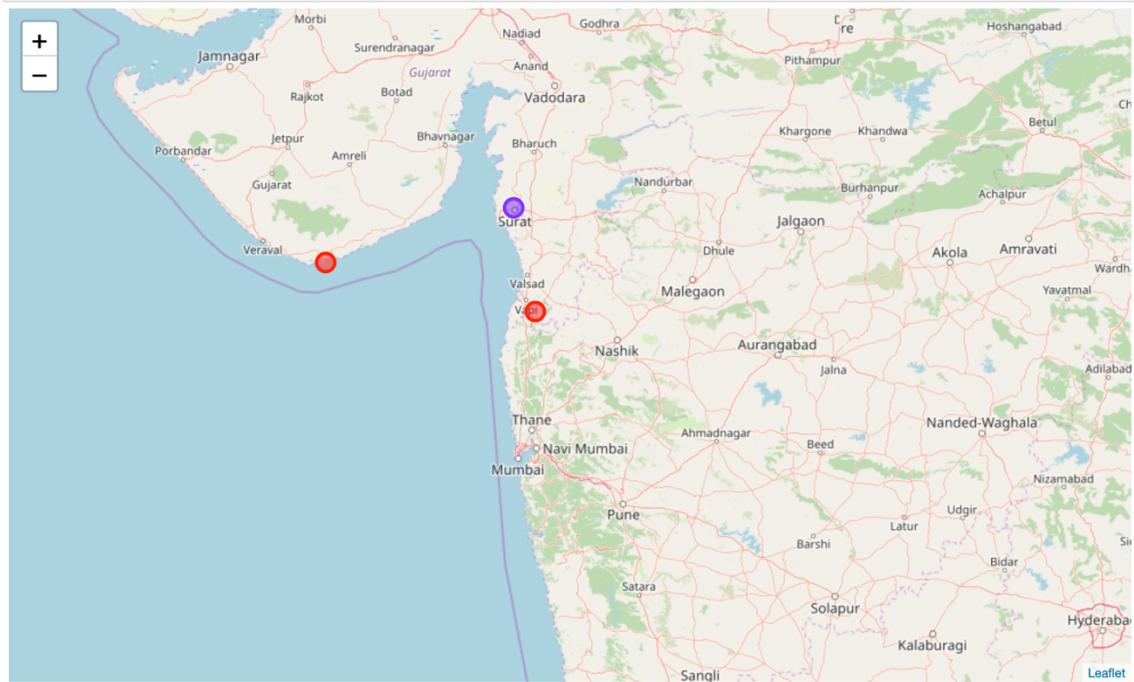
```
|: kut_merged[kut_merged['Cluster Labels'] == 1]
```

```
|:
```

Neighborhood	Area	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	10th Most Common Venue
0 Daman District	Dadra & Nagar Haveli	21.204498	72.799187	1	Fast Food Restaurant	Art Gallery	Hotel	Food Truck	Fish & Chips Shop	Café	Bus Station	Bakery	Asian Restaurant	ATM

The second cluster has one neighborhood which consists of Venues such as Fast Food Restaurant, Hotel, and Art Gallery.

Visualize the neighbourhood



Each cluster is color coded for the ease of presentation, we can see that majority of the neighborhood falls in the red cluster which is the first cluster.

5. Discussion

The aim of this project is to help people who want to relocate to the safest Area in India, expats can chose the neighborhoods to which they want to relocate based on the most common venues in it. For example if a person is looking for a neighborhood with art gallery you should choose cluster 1 but if a person is looking for a neighborhood with a Hotel or a lot of restaurant you should go to district in cluster 0

6.Conclusion

This project helps a person get a better understanding of the neighborhoods with respect to the most common venues in that neighborhood. It is always helpful to make use of technology to stay one step ahead i.e. finding out more about places before moving into a neighborhood. We have just taken safety as a primary concern to shortlist the safest borough of London. The future of this project includes taking other factors such as cost of living in the areas into consideration to shortlist the borough, such as filtering areas based on a predefined budget.