# IBM Data Science - NTL

## Capstone Project- The Battle of Neighborhoods

## Optimal Location for Asian Restaurant in London

# Introduction

London is the capital and largest city of England and the United Kingdom. Opening restaurant in the capital city like London can be challenging as you need to make huge investment but before making such investments you want  to be certain about the place to enjoy maximum patrons. London has a large population of people from different foreign countries from Asia, Australia, America, Middle east. The 2011 census recorded that 2,998,264 people or 36.7% of London's population are foreign-born making London the city with the second largest immigrant population, behind New York City. Ethnicity is one of the many factors that play a role in food choices so factors such as the kind of demographics who live there (Racial make-up, ethnic groups) can give investors a good start off. In this project, we aim to find ideal location for opening Asian restaurant in London through analysis of demographics of London to choose best borough and explore neighborhoods of that borough.

# Data

1. Demonstrate the Ethnic make-up of London(2011 Census). This can also be obtained from below link:

   https://en.wikipedia.org/wiki/Demography_of_London

2. Find list of all the boroughs of London. This data can be obtained from link below https://en.wikipedia.org/wiki/London_boroughs

3. Find demography of London which will give more details about Racial make-up of London boroughs (2011 Census). This data can be obtained from below link : https://en.wikipedia.org/wiki/Demography_of_London

4. Find all Neighbourhoods of Newham which can be obtained through web scraping from link below https://en.wikipedia.org/wiki/London_Borough_of_Newham#Districts

5. Geographical co-ordinates of Boroughs of London and Districts of Newham was obtained with the help of Geopy Library (Geocoding Web Services).

6. I will use Foursquare location data (Foursquare API) to explore neighborhoods of Newham and get number of restaurants within defined radius of each neighborhood.

# Exploratory Data Analysis

```
In [3]: london_ethnic_fig=pd.DataFrame(tables[0])
        london_ethnic_fig.head()
```

Out[3]:

| | Ethnic Group | 1991[6] | | 2001[7] | | 2011[8] | | Change 2001–2011 |
|---|---|---|---|---|---|---|---|---|
| | Ethnic Group | Number | % | Number | % | Number | % | % |
| 0 | White: British[Note 1] | NaN | NaN | 4287661.0 | 59.79% | 3669284 | 44.89% | 14.43% |
| 1 | White: Irish | 256470.0 | 3.83% | 220468.0 | 3.07% | 175974 | 2.15% | 20.19% |
| 2 | White: Gypsy or Irish Traveller[Note 2] | NaN | NaN | NaN | NaN | 8196 | 0.10% | NaN |
| 3 | White: Other[Note 1] | NaN | NaN | 594854.0 | 8.29% | 1033981 | 12.65% | 73.82% |
| 4 | White: Total | 5333580.0 | 79.80% | 5103203.0 | 71.15% | 4887435 | 59.79% | 4.23% |

```
In [4]: print("There are",len(london_ethnic_fig.columns), "columns in the dataframe")
        print(london_ethnic_fig.columns)

There are 8 columns in the dataframe
MultiIndex([(     'Ethnic Group', 'Ethnic Group'),
            (          '1991[6]',       'Number'),
            (          '1991[6]',            '%'),
            (          '2001[7]',       'Number'),
            (          '2001[7]',            '%'),
            (          '2011[8]',       'Number'),
            (          '2011[8]',            '%'),
            ('Change 2001-2011',            '%')],
           )
```

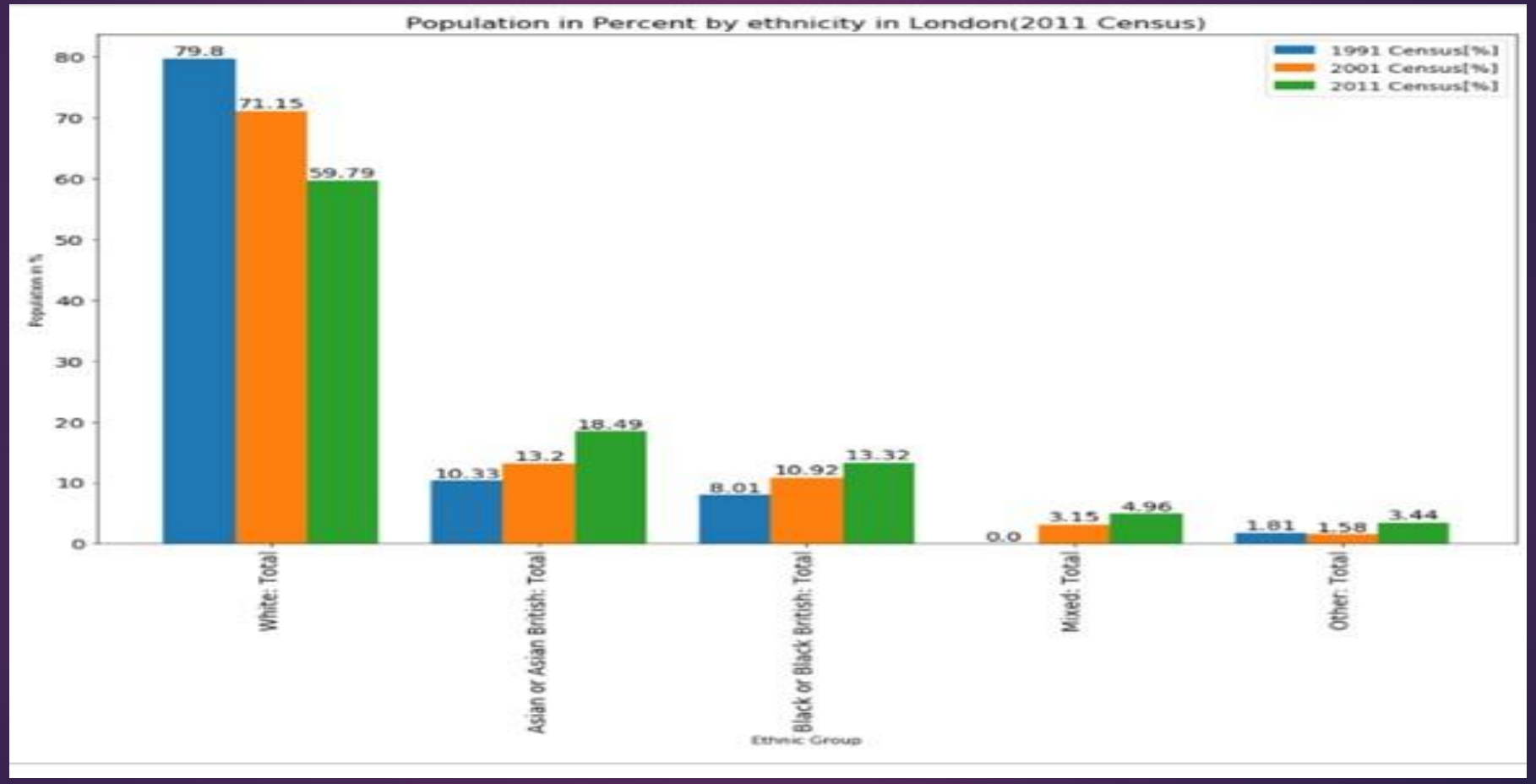
# Percentage (%) of ethnic group

```
In [10]: london_ethnic_fig1 = london_ethnic_fig[london_ethnic_fig['Ethnic Group'].s
         london_ethnic_fig2 = london_ethnic_fig1[['Ethnic Group','1991 Census[%]',
         london_ethnic_fig2.reset_index(drop=True, inplace=True)
         london_ethnic_fig2.set_index('Ethnic Group', inplace=True)
         london_ethnic_fig2 = london_ethnic_fig2.drop(['Total'])
         london_ethnic_fig2
```

Out[10]:

| Ethnic Group | 1991 Census[%] | 2001 Census[%] | 2011 Census[%] |
|---|---|---|---|
| White: Total | 79.80 | 71.15 | 59.79 |
| Asian or Asian British: Total | 10.33 | 13.20 | 18.49 |
| Black or Black British: Total | 8.01 | 10.92 | 13.32 |
| Mixed: Total | 0.00 | 3.15 | 4.96 |
| Other: Total | 1.81 | 1.58 | 3.44 |

# Population in percentages by ethnicity in London



# Boroughs of London

```
In [15]: London_df.head()
```

Out[15]:

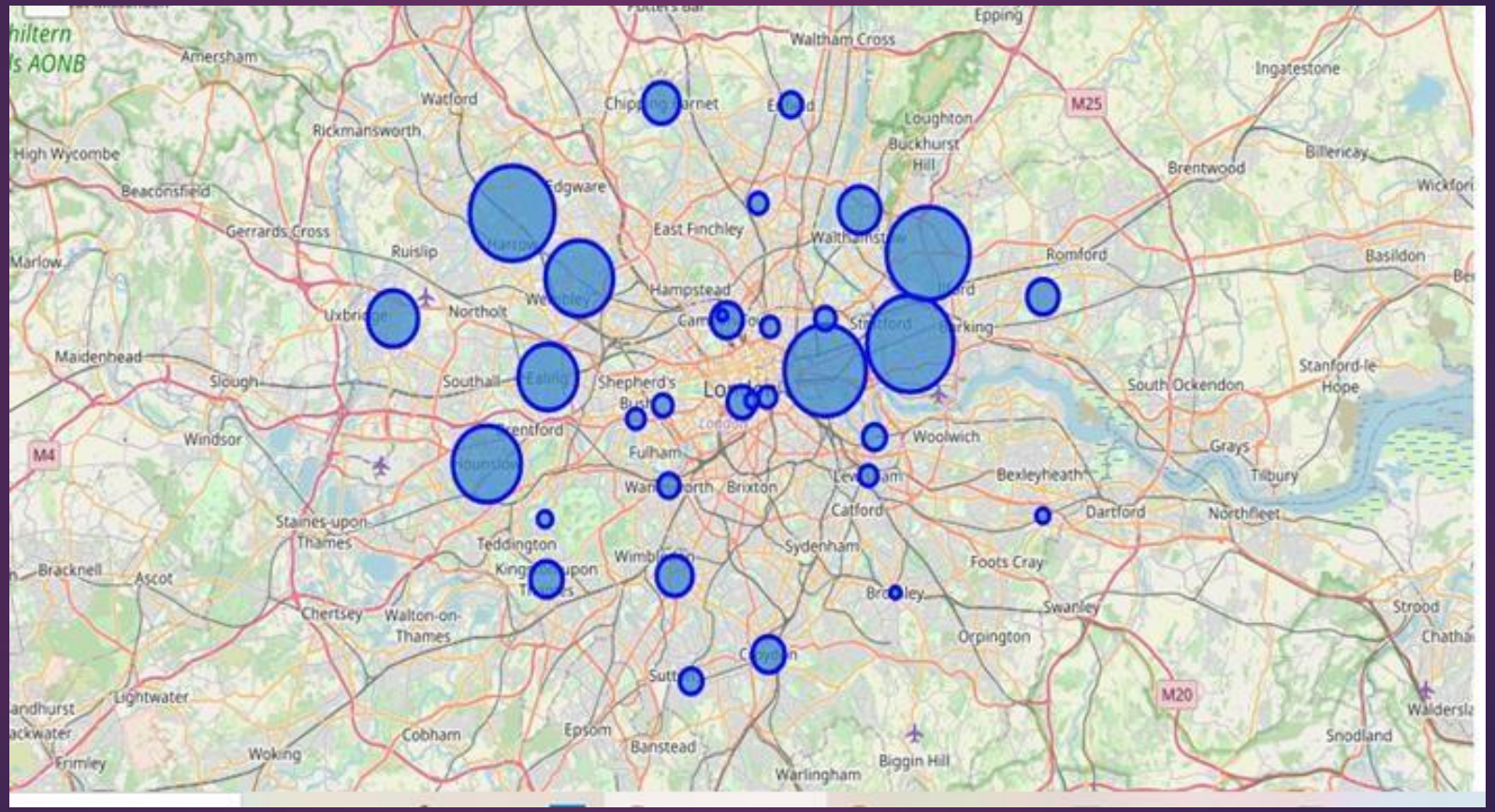
| | London_Borough |
|---|---|
| 0 | Camden |
| 1 | Greenwich |
| 2 | Hackney |
| 3 | Hammersmith |
| 4 | Islington |

Out[25]:

| | London_Borough | White | Mixed | Asian | Black | Other |
|---|---|---|---|---|---|---|
| 0 | Barnet | 64.1 | 4.8 | 18.5 | 7.7 | 4.8 |
| 1 | Barking and Dagenham | 58.3 | 4.2 | 15.9 | 20.0 | 1.6 |
| 2 | Bexley | 81.9 | 2.3 | 6.6 | 8.5 | 0.8 |
| 3 | Brent | 36.3 | 5.1 | 34.1 | 18.8 | 5.8 |
| 4 | Bromley | 84.3 | 3.5 | 5.2 | 6.0 | 0.9 |
| 5 | Camden | 66.3 | 5.6 | 16.1 | 8.2 | 3.8 |
| 6 | Croydon | 55.1 | 6.6 | 16.4 | 20.2 | 1.8 |

# Racial make-up of London borough (2011 Census)

# Asian ethnic proportion in London



# Superimposed neighbourhoods on the map of Newham

# Venues nearby neighborhoods of Newham

| | District | District Latitude | District Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category |
|---|---|---|---|---|---|---|---|
| 0 | Beckton | 51.51608 | 0.059426 | East london Gymnastics Club | 51.514107 | 0.060155 | Gym / Fitness Center |
| 1 | Beckton | 51.51608 | 0.059426 | Lidl | 51.515982 | 0.054794 | Supermarket |
| 2 | Beckton | 51.51608 | 0.059426 | Home Bargains | 51.516790 | 0.062967 | Discount Store |
| 3 | Beckton | 51.51608 | 0.059426 | Lituanica | 51.516442 | 0.062927 | Grocery Store |
| 4 | Beckton | 51.51608 | 0.059426 | Pets at Home | 51.520473 | 0.070494 | Pet Store |

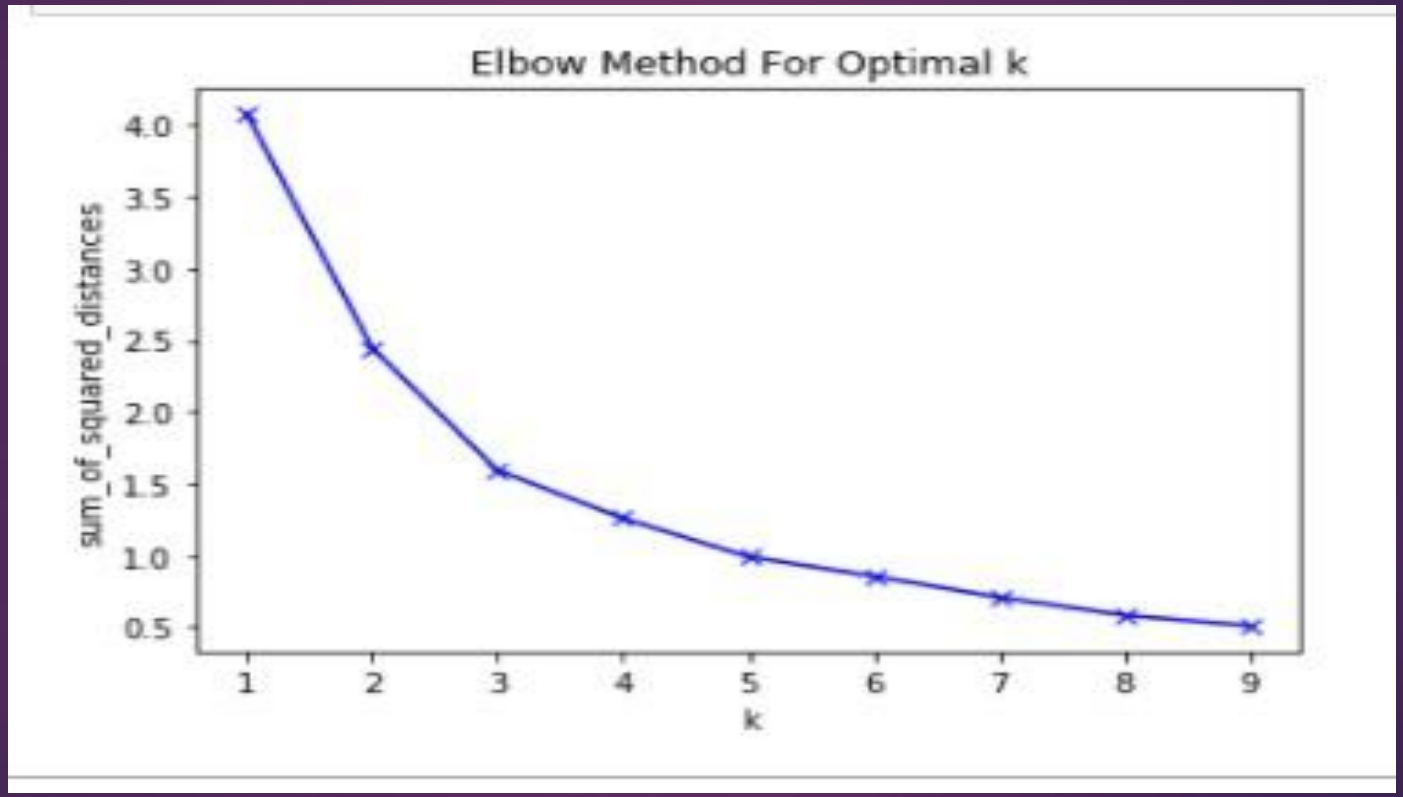
# Dataframe after One- hot encoding

| | District | American Restaurant | Asian Restaurant | Bed & Breakfast | Bistro | Breakfast Spot | Bulgarian Restaurant | Burger Joint | Café | Chinese Restaurant | Chocolate Shop | Coffee Shop | Comfort Food Restaurant | Dessert Shop | Diner | Dc Restau |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Beckton | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | |
| 1 | Beckton | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | |
| 2 | Beckton | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | |
| 3 | Canning Town | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | |
| 4 | Canning Town | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |

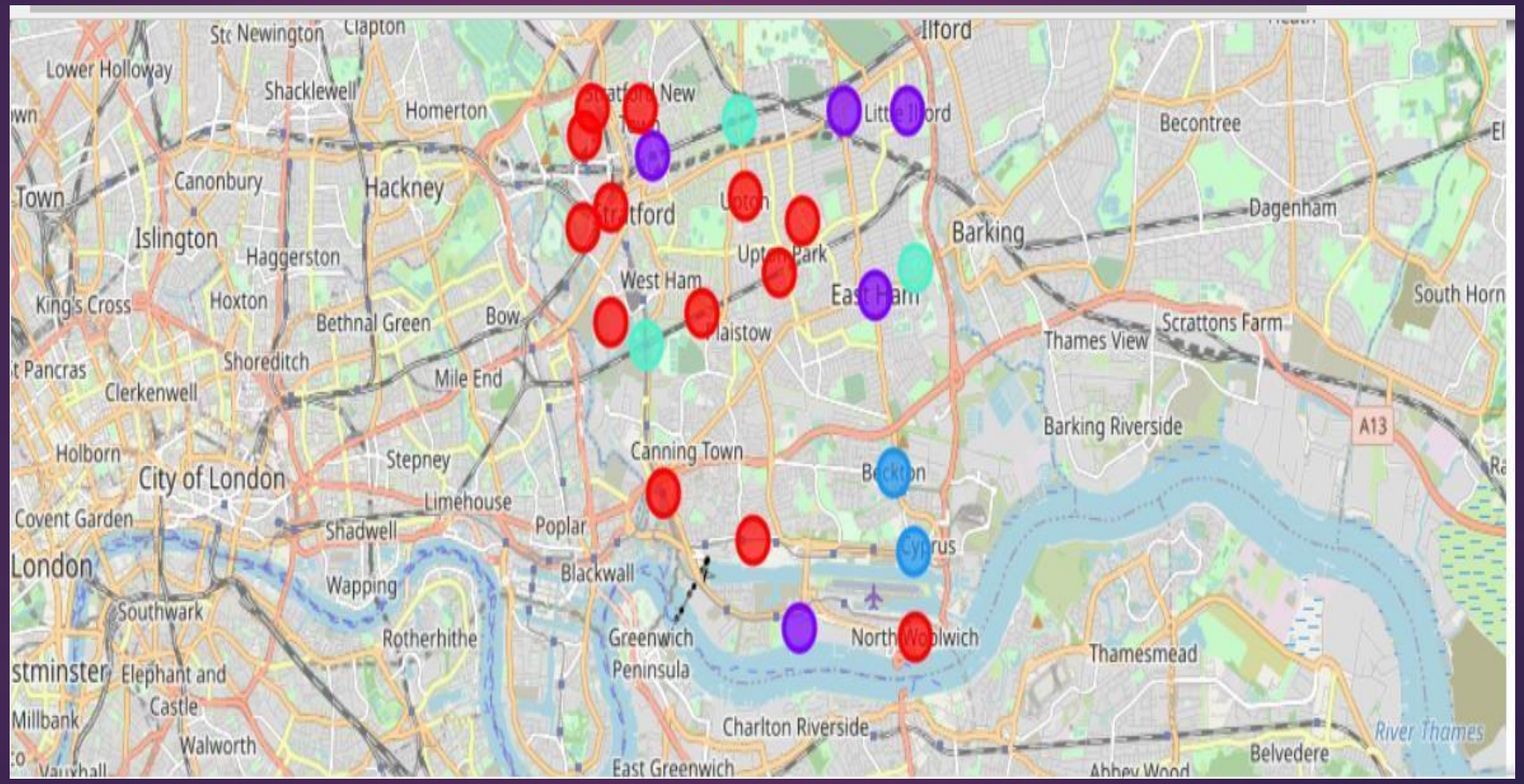
# top 10 Most common venues in each neighborhood

| | District | 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 | Beckton | Coffee Shop | Café | Comfort Food Restaurant | Food & Drink Shop | Fish & Chips Shop | Fast Food Restaurant | English Restaurant | Eastern European Restaurant | Donut Shop | Doner Restaurant |
| 1 | Canning Town | Coffee Shop | Café | Fast Food Restaurant | Sandwich Place | Italian Restaurant | Diner | Food & Drink Shop | Turkish Restaurant | Breakfast Spot | Burger Joint |
| 2 | Custom House | Coffee Shop | Café | Tapas Restaurant | Chinese Restaurant | Restaurant | Bistro | Italian Restaurant | Lebanese Restaurant | Middle Eastern Restaurant | American Restaurant |
| 3 | Cyprus | Coffee Shop | Comfort Food Restaurant | Food & Drink Shop | Fish & Chips Shop | Fast Food Restaurant | English Restaurant | Eastern European Restaurant | Donut Shop | Doner Restaurant | Diner |
| 4 | East Ham | Indian Restaurant | Fast Food Restaurant | Coffee Shop | Sandwich Place | Pizza Place | Vegetarian / Vegan Restaurant | English Restaurant | Eastern European Restaurant | Donut Shop | Doner Restaurant |

# Elbow Method



# Neighborhoods of Newham clustering



# Examining Cluster 1

| | District | Borough | 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 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | East Ham | Newham | 51.532963 | 0.055320 | 1 | Indian Restaurant | Fast Food Restaurant | Coffee Shop | Sandwich Place | Pizza Place | Vegetarian / Vegan Restaurant | English Restaurant | Eastern European Restaurant | Don Sh |
| 7 | Little Ilford | Newham | 51.550298 | 0.062522 | 1 | Indian Restaurant | Fast Food Restaurant | Ice Cream Shop | Restaurant | Vegetarian / Vegan Restaurant | Comfort Food Restaurant | English Restaurant | Eastern European Restaurant | Don Sh |
| 8 | Maryland | Newham | 51.546053 | 0.005922 | 1 | Pizza Place | Café | Coffee Shop | Indian Restaurant | Burger Joint | Mediterranean Restaurant | English Restaurant | Eastern European Restaurant | Don Sh |
| 13 | Silvertown | Newham | 51.501363 | 0.038518 | 1 | Coffee Shop | Sandwich Place | Restaurant | Vegetarian / Vegan Restaurant | Mexican Restaurant | Asian Restaurant | Bistro | Café | Chine Restaura |
| 21 | Manor Park | Newham | 51.550330 | 0.048580 | 1 | Indian Restaurant | Restaurant | Vegetarian / Vegan Restaurant | Comfort Food Restaurant | Fish & Chips Shop | Fast Food Restaurant | English Restaurant | Eastern European Restaurant | Don Sh |

In cluster 1, Indian restaurants, Fast food restaurants and Vegetarian / Vegan Restaurant are most common.

# Examining Cluster 2

```
cluster_2 = Newham_merged.loc[Newham_merged['Cluster Labels'] == 2, Newham_merged.columns[0:15]]
cluster_2
```

| | District | Borough | 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 Comi Venu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Beckton | Newham | 51.516080 | 0.059426 | 2 | Coffee Shop | Café | Comfort Food Restaurant | Food & Drink Shop | Fish & Chips Shop | Fast Food Restaurant | English Restaurant | Eastern European Restaurant | Donut Shop | I Resta |
| 3 | Cyprus | Newham | 51.508478 | 0.063969 | 2 | Coffee Shop | Comfort Food Restaurant | Food & Drink Shop | Fish & Chips Shop | Fast Food Restaurant | English Restaurant | Eastern European Restaurant | Donut Shop | Doner Restaurant | |

In cluster2, Coffee shop and Vegetarian / Vegan Restaurant are most common venues.

# Discussion

Due to the diversity of the Newham in each neighborhood, there is an assortment of most common venues and there are numerous ethnic restaurants as well. Our analysis is focused on finding optimal neighborhood for opening Asian restaurant so to understand the clusters let us find out which neighborhood has the most common venues related to Asian ethnicity. From cluster 0, Custom House, Plashet, Upton Park, Upton, Silvertown are the neighborhoods with the highest number of Asian restaurants. In cluster 1, Indian Restaurant is most common across all the neighborhoods and these are not crowded with other Asian cuisines. Cluster 2 is not famous for Asian cuisine hence opening an Asian restaurant in these neighborhoods will not be profitable. In Cluster 3 Forest Gate and Wallend has Asian restaurant in top 2 most common venue.

# Conclusion

One application of Clustering Algorithm, k-Means or others, to a multi-dimensional dataset, a very inquisitive result can be curated which helps to understand and visualize the data. The neighborhoods of Newham borough are very briefly segmented into four clusters based on the most common venue hence when looking for a restaurant location, one must consider who else is doing business in the neighborhood. If there are already many restaurants with the same concept of ethnic cooking, then it will not be a profitable deal to choose that location such neighborhoods are mostly appearing in cluster 0. While neighborhoods in cluster 1 are most common for Asian ethnic venue but at the same time, these are less crowded with Asian restaurants. To enjoy maximum patrons in the restaurant, the neighborhoods from cluster 1 are assumed the best choice to open Asian restaurant. The results of this project can be improved and made more inquisitive by considering neighborhoods of other boroughs which have high proportion of Asian population. The scope of this project can be expanded further to choose best borough for opening Asian or other ethnic concept restaurants and suggest a new vendor a profitable location in a diverse city like London.