



**Identifying the neighborhood of
Brooklyn to understand new business
openings.**

Introduction

- To open a new business a variety of insights are required such as :

- 1) Where can a business be opened?
- 2) What kind of business can be opened in a particular area?
- 3) What is the success rate of the business?

To answer the questions we need to identify the business already present in an area and study them.



Business Problem

A business can grow and develop when there are not much of a competition and when the product produced is of great quality. In any particular area there can be a lot of people with same business and products.

Data

Loading the data from json file

```
!]: with open('newyork_data.json') as json_data:  
    newyork_data = json.load(json_data)
```

```
!]: newyork_data
```

```
!]: {'type': 'FeatureCollection',  
     'totalFeatures': 306,  
     'features': [{'type': 'Feature',  
                    'id': 'nyu_2451_34572.1',  
                    'geometry': {'type': 'Point',  
                                 'coordinates': [-73.84720052054902, 40.89470517661]},  
                    'geometry_name': 'geom',  
                    'properties': {'name': 'Wakefield',  
                                   'stacked': 1,  
                                   'annoline1': 'Wakefield',  
                                   'annoline2': None,  
                                   'annoline3': None,  
                                   'annoangle': 0.0,  
                                   'borough': 'Bronx',  
                                   'bbox': [-73.84720052054902,  
                                           40.89470517661,  
                                           -73.84720052054902,  
                                           40.89470517661]}},  
                    {'type': 'Feature',
```

Visualization

- Here is a map of Brooklyn which will be clustered.



Clustering

- Brooklyn after clustering



Identifying common venues

	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Bath Beach	Chinese Restaurant	Pharmacy	Bubble Tea Shop	Pizza Place	Donut Shop	Fast Food Restaurant	Gas Station	Italian Restaurant	Restaurant	Peruvian Restaurant
1	Bay Ridge	Spa	Italian Restaurant	Pizza Place	Greek Restaurant	Bar	American Restaurant	Bagel Shop	Pharmacy	Middle Eastern Restaurant	Ice Cream Shop
2	Bedford Stuyvesant	Coffee Shop	Pizza Place	Café	Bar	Del / Bodega	Gourmet Shop	Park	Bagel Shop	BBQ Joint	Cocktail Bar
3	Bensonhurst	Chinese Restaurant	Italian Restaurant	Pizza Place	Sushi Restaurant	Ice Cream Shop	Donut Shop	Park	Cha Chaan Teng	Liquor Store	Road
4	Bergen Beach	Harbor / Marina	Park	Playground	Athletics & Sports	Baseball Field	Farmers Market	Ethiopian Restaurant	Event Space	Factory	Falafel Restaurant



Methodology

- Brooklyn data is given to K mean for clustering with a cluster of 5.
- It creates 5 clusters and identifies the cluster with maximum number of businesses.
- We can visualize and see what can be opened in any cluster.

Result and Conclusion

- Cluster 1,4 and 5 are ideal to open any new well established businesses such as Dairy, Bakery, Flower shop, Clothing store etc, whereas any new idea or innovation can be tested in cluster 2 and 3 as the amount of well established businesses are already present and the success rate of new business is high (if any new idea) as the population is high in that region.