

Finding the Best Neighborhoods in Brooklyn

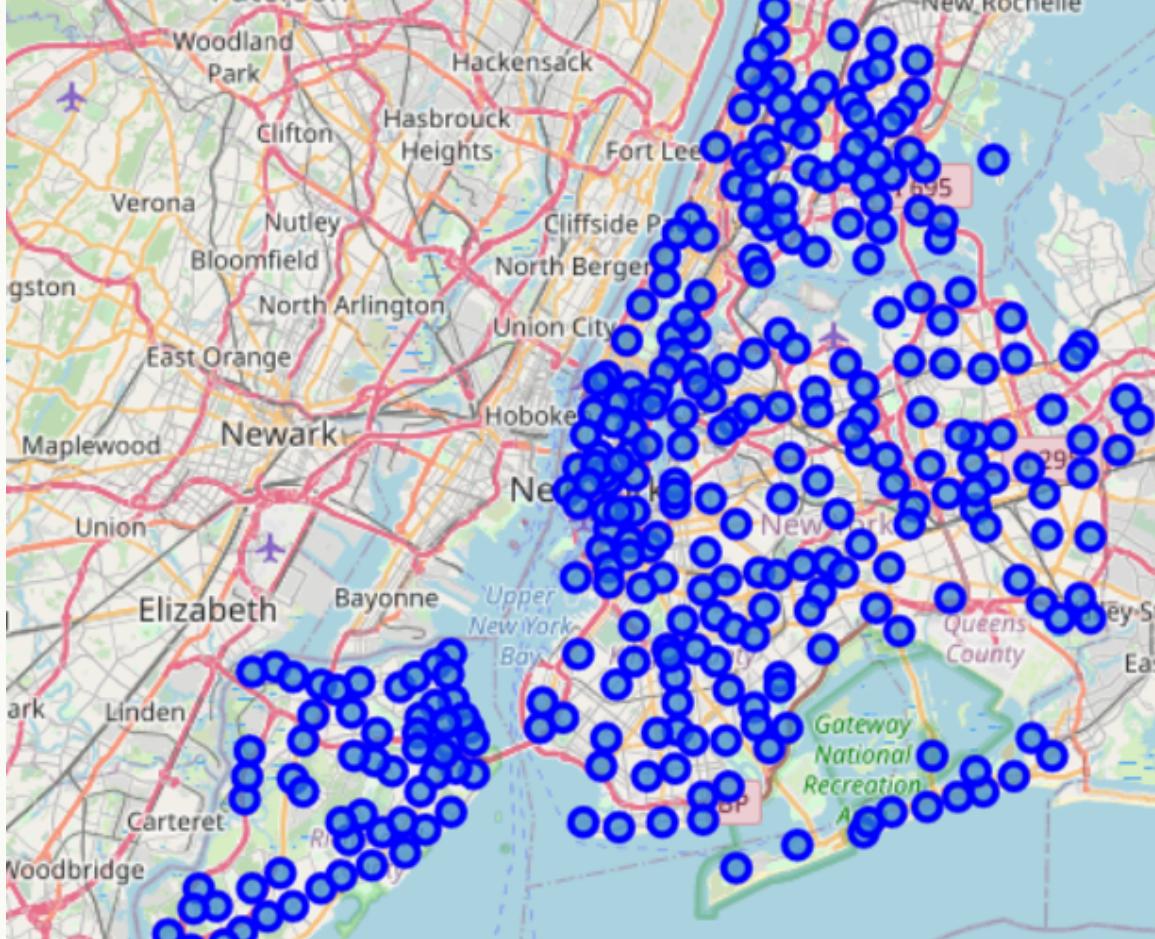
The more information potential migrants have the better

- Many different people value many different things
- For this project we will be looking at the top 10 venue types in Brooklyn these being things like Food, Transport, School, etc.
- This information will aid migrants in their journey to move to a new area
- This will not only be useful to regular residents but prospective entrepreneurs as well

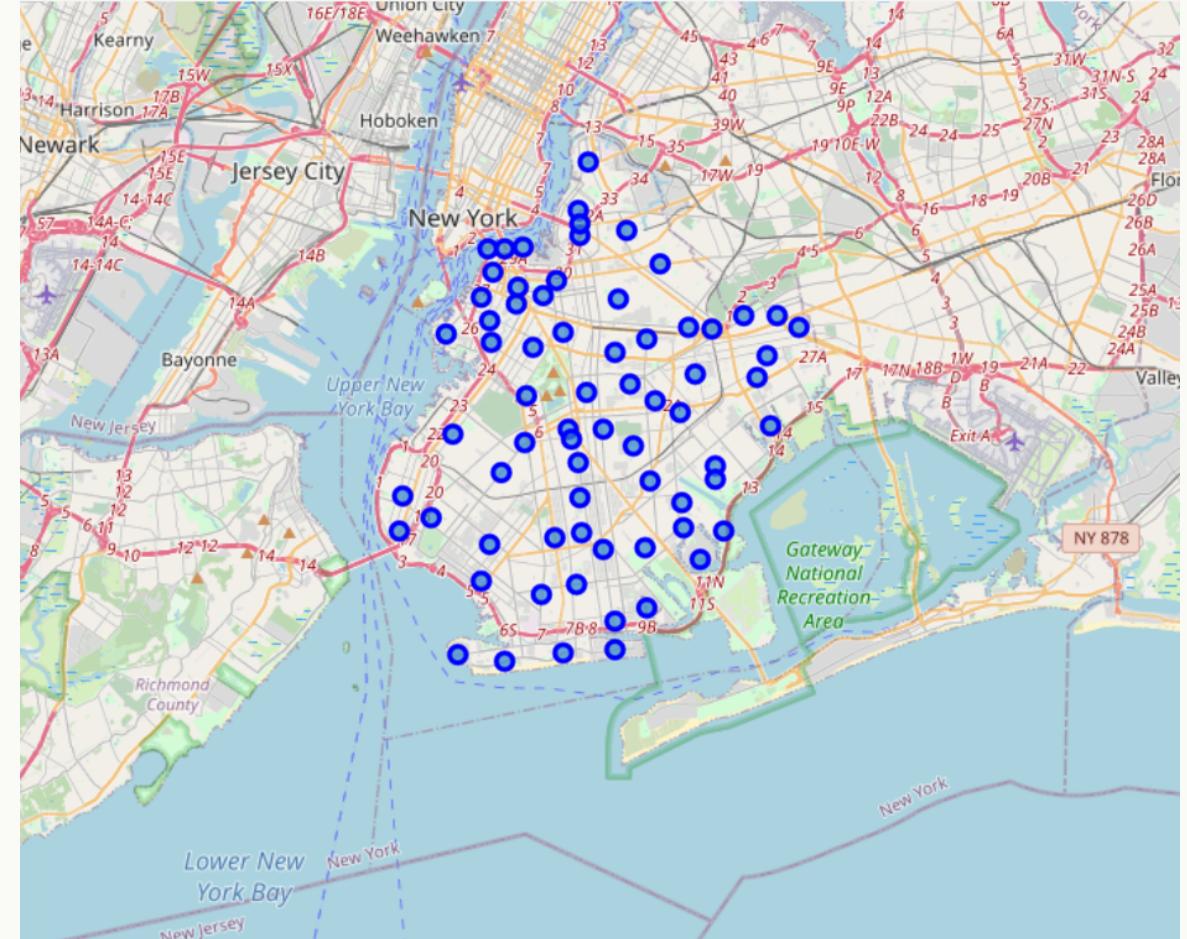
Data Selection and Cleaning

- Will be primarily using the New York data set provided in week three
- This will be section off and specified to the borough of Brooklyn instead of all of New York
- The other main source of data being Foursquare API this will be providing the top venue types and the totals within each neighborhood

Specify from New York to just Brooklyn



New York



Brooklyn

MADE WITH

beautiful.ai

Find the top venue types and the totals within the neighborhoods

Arts & Entertainment	College & University	Event	Food	Nightlife Spot	Outdoor Recreat
0	0	0	0	0	
0	0	0	0	0	
0	0	0	0	0	
0	0	0	0	0	
0	0	0	0	0	

appending the top types

```
|: for i, row in brooklyn_venues.iterrows():
    for c in categories_list:
        brooklyn_venues.loc[i, c[0]] = get_venues_coun
                                         broo
                                         radi
                                         print('{} ({}, {}) data gathering for is complete'
                                                brooklyn_venues.iloc[i].Neighborhood, brooklyn
                                                brooklyn_venues.iloc[i].Longitude))
```

finding the totals

Arts & Entertainment	College & University	Event	Food
11	26	7	208
8	51	8	114
21	38	9	180
55	129	24	121
15	93	6	105

totals added to the database

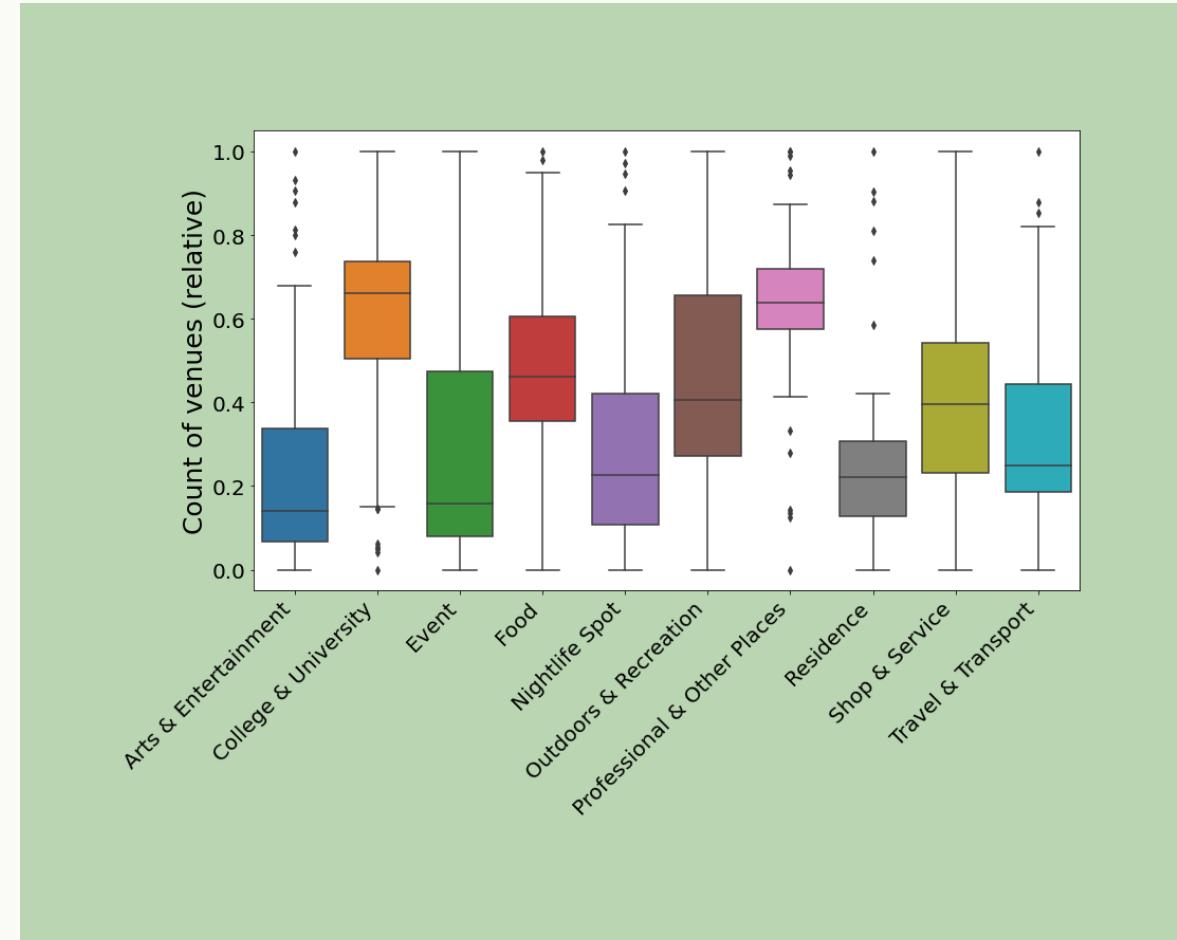
Normalize data and make box plot of Brooklyn

```
: from sklearn.preprocessing import MinMaxScaler  
X = brooklyn_venues.values[:,4:]  
scaled_dataset = MinMaxScaler().fit_transform(X)  
brooklyn_scaled = pd.DataFrame(scaled_dataset)  
brooklyn_scaled.columns = [c[0] for c in categor  
brooklyn_scaled.head()
```

08]:

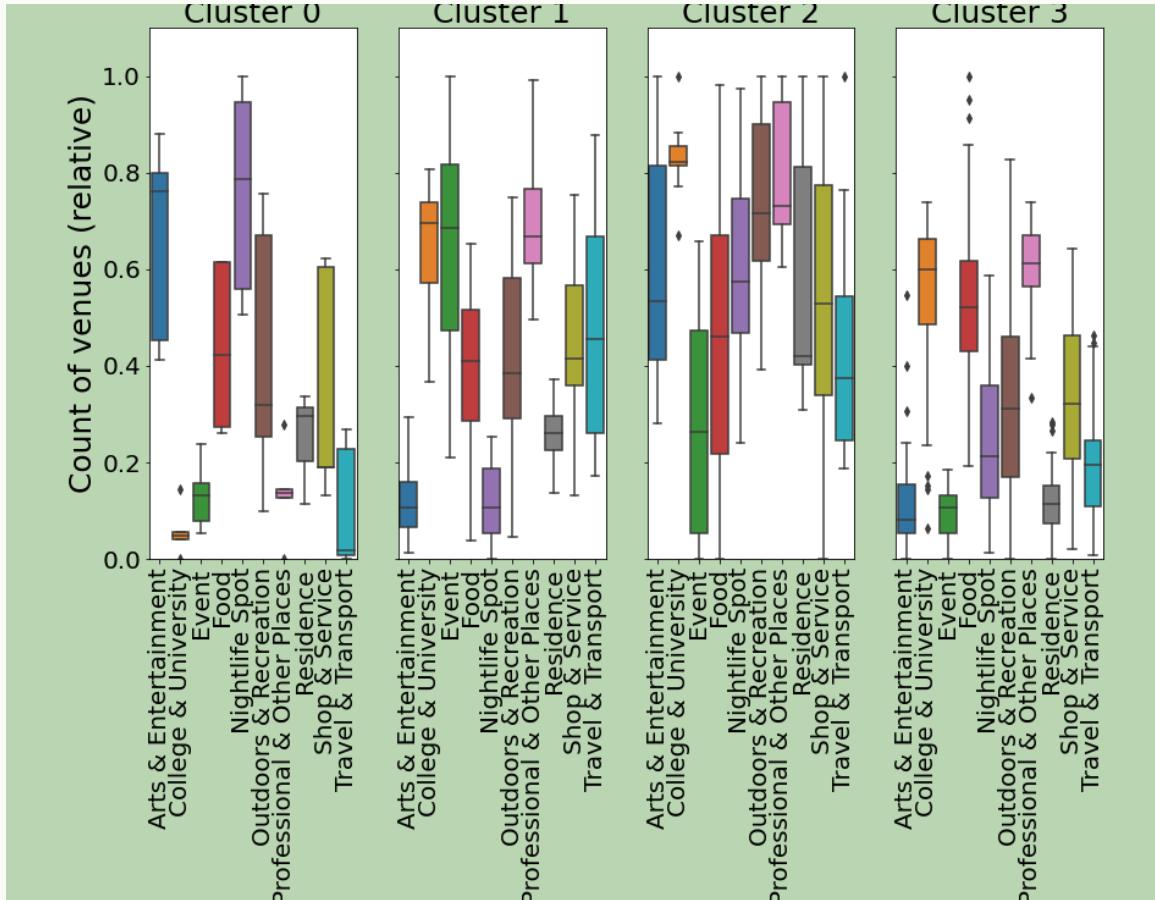
	Arts & Entertainment	College & University	Event
0	0.066667	0.062069	0.078947
1	0.026667	0.234483	0.105263
2	0.200000	0.144828	0.131579
3	0.653333	0.772414	0.526316
4	0.120000	0.524138	0.052632

normalize data

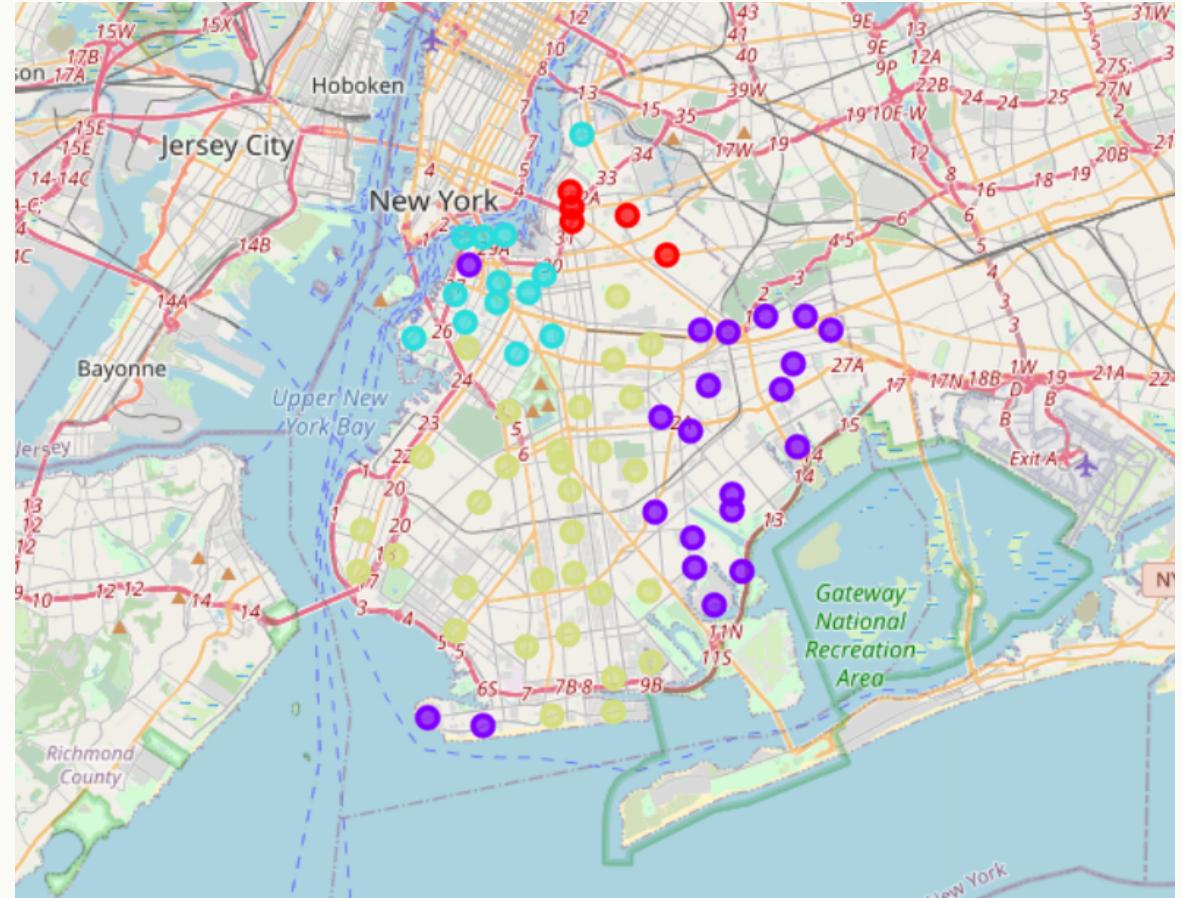


box plot of Brooklyn

Segment into clusters and map cluster data



box plot of each cluster



clusters mapped (red = 0, purple = 1, teal = 2, green/yellow = 3)

Conclusion based off results

- The main clusters of neighborhoods that have the most potential to move into are clusters 2 and 3 cluster 2 having the highest venue density, providing the most comfort, and cluster 3 having the lowest density, providing the best price and open business ventures
- This project can provide a large amount of utility to people looking to move to Brooklyn and could provide even more by being added on to.