Opening a bakery in a convenient neighborhood in

New York City

Approaches to open or buy a bakery.

- In a business area: to provide other businesses like restaurants, cafes, sandwich places, etc.
- In a residential area: to provide services directly to people.

Question to determine the best place for a bakery.

- What indicators determine a borough is a good borough?

- What indicators determine a neighborhood is a good neighborhood ?

Indicator considered in this case.

- Boroughs with more neighborhoods.
- Venues with few bakeries around.
- Population
- Density
- Gross Domestic Product (GDP)

Source of data.

- Borough, Neighborhoods of New York City with their latitudes and longitudes https://cocl.us/new_york_dataset.

- The link with resource for developers with codes of type of businesses in Foursquare API venues. https://developer.foursquare.com/docs/resources/categories.

- Wikipedia table with demographic information of New York City's boroughs (population, density, and Gross Domestic Product). https://en.wikipedia.org/wiki/Demographics_of_New_York_City.

- Excel worksheet with information extracted from www.point2homes.com. Population and Average Household Income in venues.

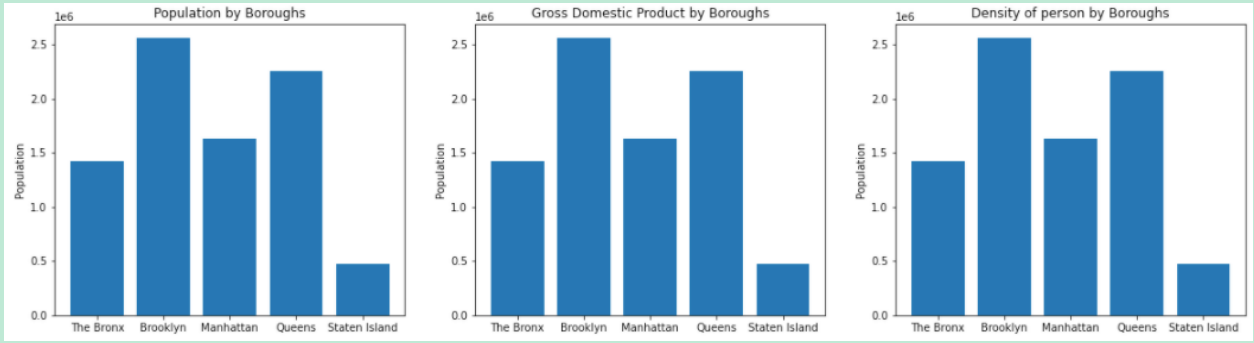
Exploring a borough.

- Brooklyn was the borough selected because it is the borough with more neighborhoods, more population, more density, and the best GDP

Grouping by borough and counting neighborhoods in every borough

```
[10]: # We group by Borough and count the Neighborhoods in every one of them.
      neighborhoods_in_boroughs = df_ny.groupby(['Borough'])['Neighborhood'].count()
      neighborhoods_in_boroughs
```

```
[10]: Borough
      Bronx            52
      Brooklyn         70
      Manhattan        40
      Queens           81
      Staten Island    63
      Name: Neighborhood, dtype: int64
```



Two options to be considered to make a choice about a best place for a bakery.

- Coney Island venue is an option to be considered because is a place with few bakeries around in a broad area
- Clustering Method shows different options based on population and average of income in venues. A balance between this two indicator represent good options

```
[45]: venues_option = brooklyn_neigbor[brooklyn_neigbor['Population'].between(75000, 125000) \
                    & brooklyn_neigbor['Average Household Income'].between(70000, 110000) ]
      venues_option
```

[45]:

	Neighborhood	Population	Average Household Income	Clusters
1	Bay Ridge	85791.0	90550.57	3
6	Brighton Beach	78775.0	81900.00	3
10	Bushwick	102607.0	70401.52	3
11	Canarsie	100844.0	86568.00	3
17	Crown Heights	118623.0	79791.13	3
27	Flatbush	99558.0	75780.00	3
32	Georgetown	95666.0	108156.00	3
33	Gerritsen Beach	83119.0	87360.00	3
35	Gravesend	112900.0	77822.67	3
41	Manhattan Beach	78775.0	81900.00	3
44	Midwood	113280.0	81581.01	3
45	Mill Basin	93534.0	108156.00	3
63	Sunset Park	79113.0	72787.62	3

Any of this venues could be a good place for a bakery

Conclusion

On my viewpoint, a depth of a study, always depends on data available for analysis.
Even though data used in this case help to make good decisions, more data could be necessary to make a deeper analysis.
Variables like income of bakeries around the area selected, rate of crime in the area, etc. could be considered, but all of these depend on the stakeholder's requirements.