Battle of Neighborhoods

- An entrepreneur wishes to set up a profitable health care facility in New York State
- Is health care facility a profitable venture? Are other business corporations interested in such a venture?
- There are several kinds of health care facilities. Which type of health care facility most popular?
- As it turns out Diagnostic and Treatment Centers are most popular, the question of location arises. Which neighborhood is most profitable for setting up a Diagnostic and Treatment Center.
- Data from New York State Health department and Foursquare API can be and have been used to answer these questions.
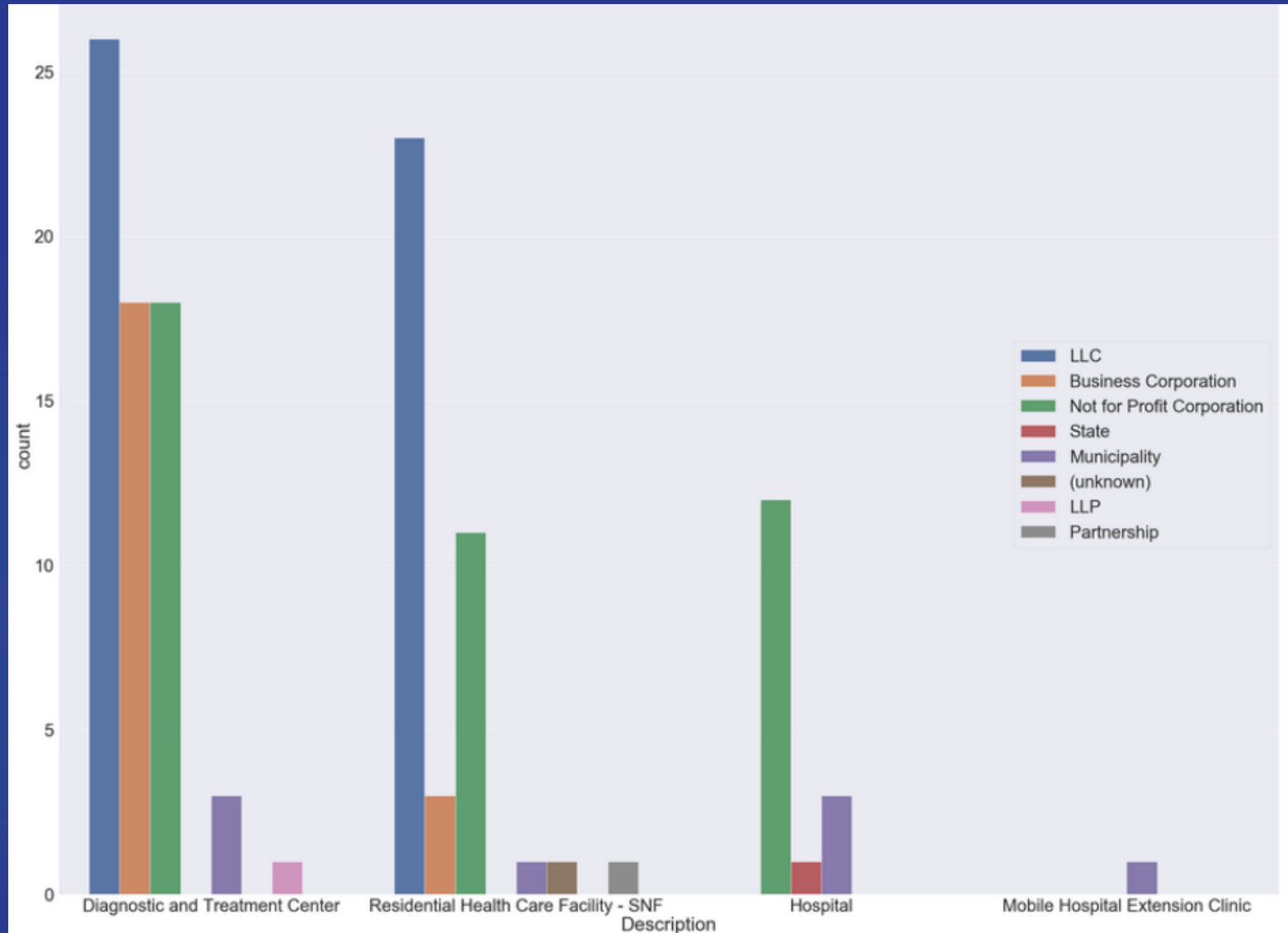
Methodology

Data
Cleaning

Exploratory
Analysis

Machine
Learning

Inferential
Analysis

Data acquisition and cleaning

- Diagnostic and Treatment Centers, and other Health Facility data scraped from https://health.data.ny.gov/Health/Health-Facility-General-Information/vn5v-hh5r/data
- In total, 1,420 rows and 36 features in the raw dataset.
- Duplicate, highly similar or highly correlated features were dropped.
- Dropped the rows in which co-ordinates (Latitude or Longitude) are NaN
- Cleaned data contains 10 features

Ownership Type	(unknown)	Business Corporation	County	Individual	LLC	LLP	Municipality	Not for Profit Corporation	Partnership	Public Benefit Corporation	State
Hospital_Year											
1901	0	10	24	1	69	1	14	245	0	4	7
1960	0	0	0	0	0	0	0	0	1	0	0
1970	0	0	0	1	1	0	0	1	0	0	0
....	..	..	..	..	..	..	..	..	..	..	..

Which is the most popular type of health facility from business point of view?
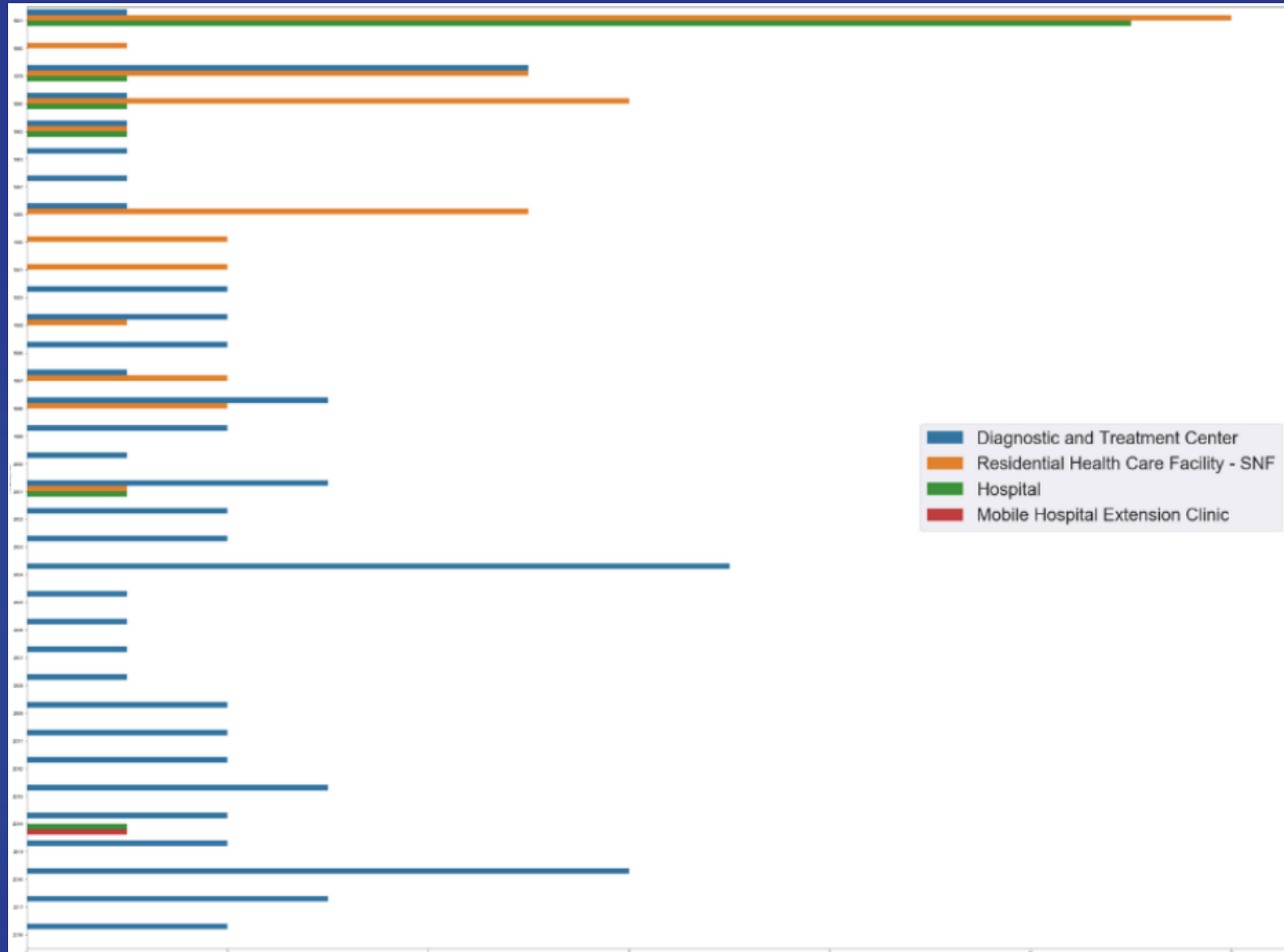
Which type of business has focused on which type of health care facility?

Methodology and Exploratory Analysis

- Foursquare API is used to explore the neighborhoods around the diagnostic centers and segment them.

- Top 10 most frequently occurring venues in each neighborhood were fetched.

- K-Means Machine Learning Algorithm has been used for clustering the health care facilities in to 5 different clusters.

- Based on the information obtained on the surrounding venues in each cluster, a plan will be proposed for the development of a new diagnostic and treatment center.

- A Cluster Map of Brooklyn with the hospitals and their clusters superimposed on it is created.

Machine Learning

Mapping and clustering
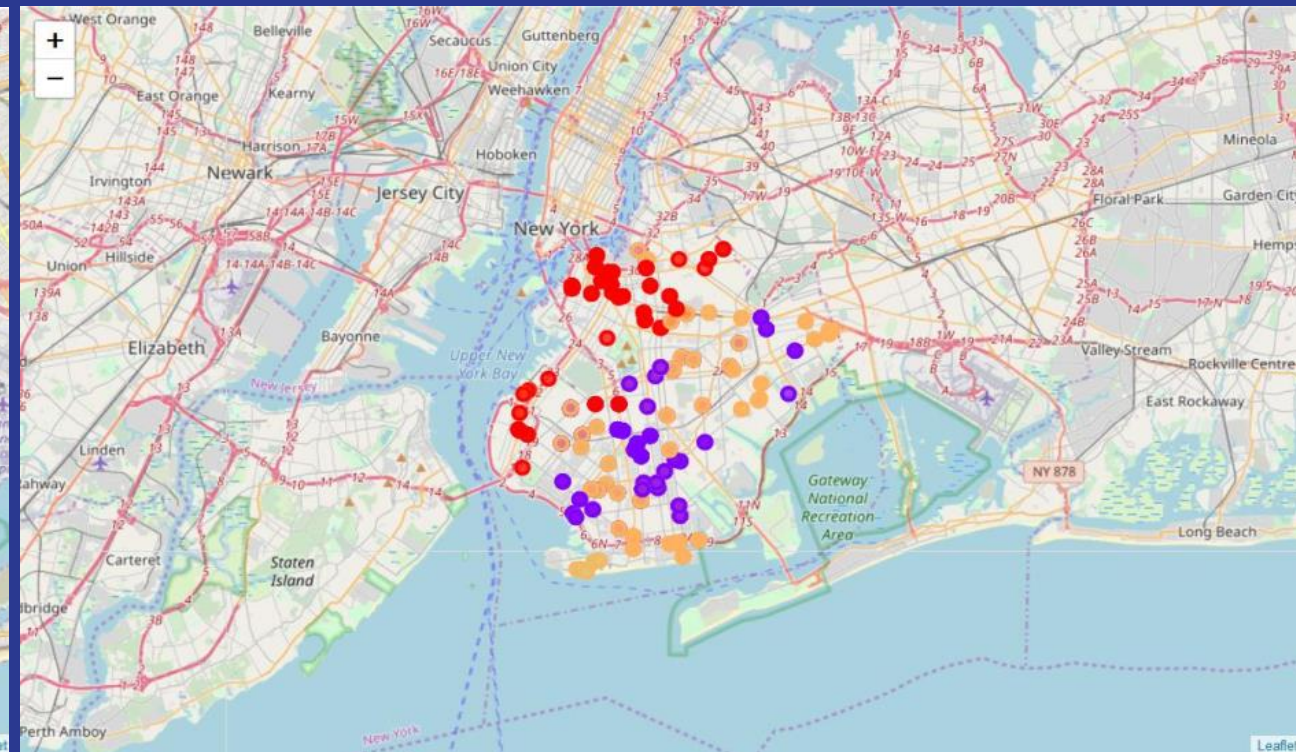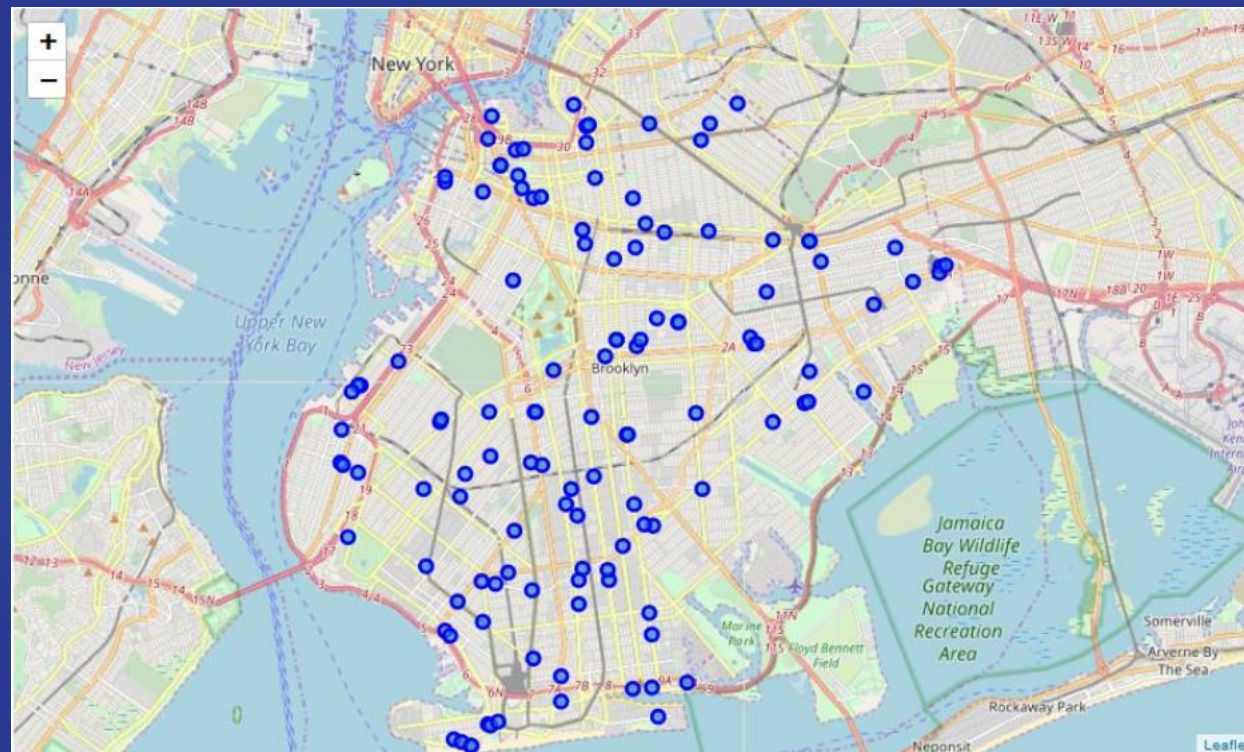
Results and conclusion

- Brooklyn neighborhood has the highest number of health care facilities in it.

- The surrounding venues in the neighborhoods in cluster 5 mainly consist of cafe, restaurants, banks, pharmacy, bus stations, and some residential buildings too.

- There could be significant competition posed by other diagnostic centers in this cluster but this cluster has the most promising surroundings to establish a new diagnostic and treatment center.

Thank You!

GAURI.SARAN100@GMAIL.COM