IBM CAPSTONE PROJECT

The Battle of Neighborhoods:

Cluster Analysis of Manhattan Real Estate Market

GODKOWICZ PAWEŁ

- Business problem section

How we can provide suport to potential buyers, investors, to purchese a suitable real estate in Manhattan in this uncertain economic and financial scenerio.

- Solution

Clustering algorithms Manhattan neighborhood in order to recommend venues and the current average price of real estate where buyers can make a real estate investments.

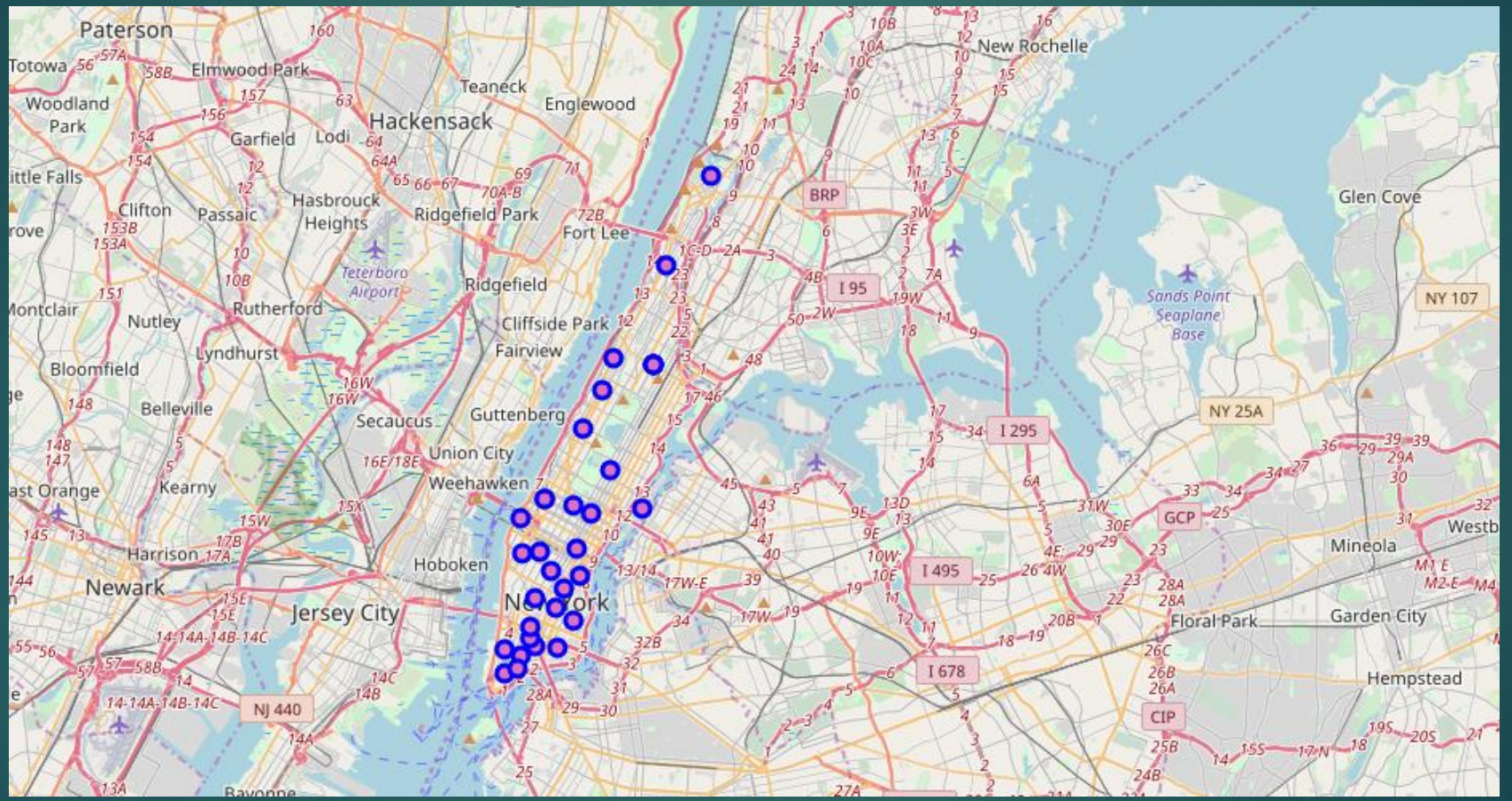
Data:

Merging data on Manhattan properties and the relative price paid data from NYC Department of Finance: https://www1.nyc.gov/site/finance/taxes/property-annualized-sales-update.page and data on amenities and essential facilities surrounding such properties from FourSquare API interface.

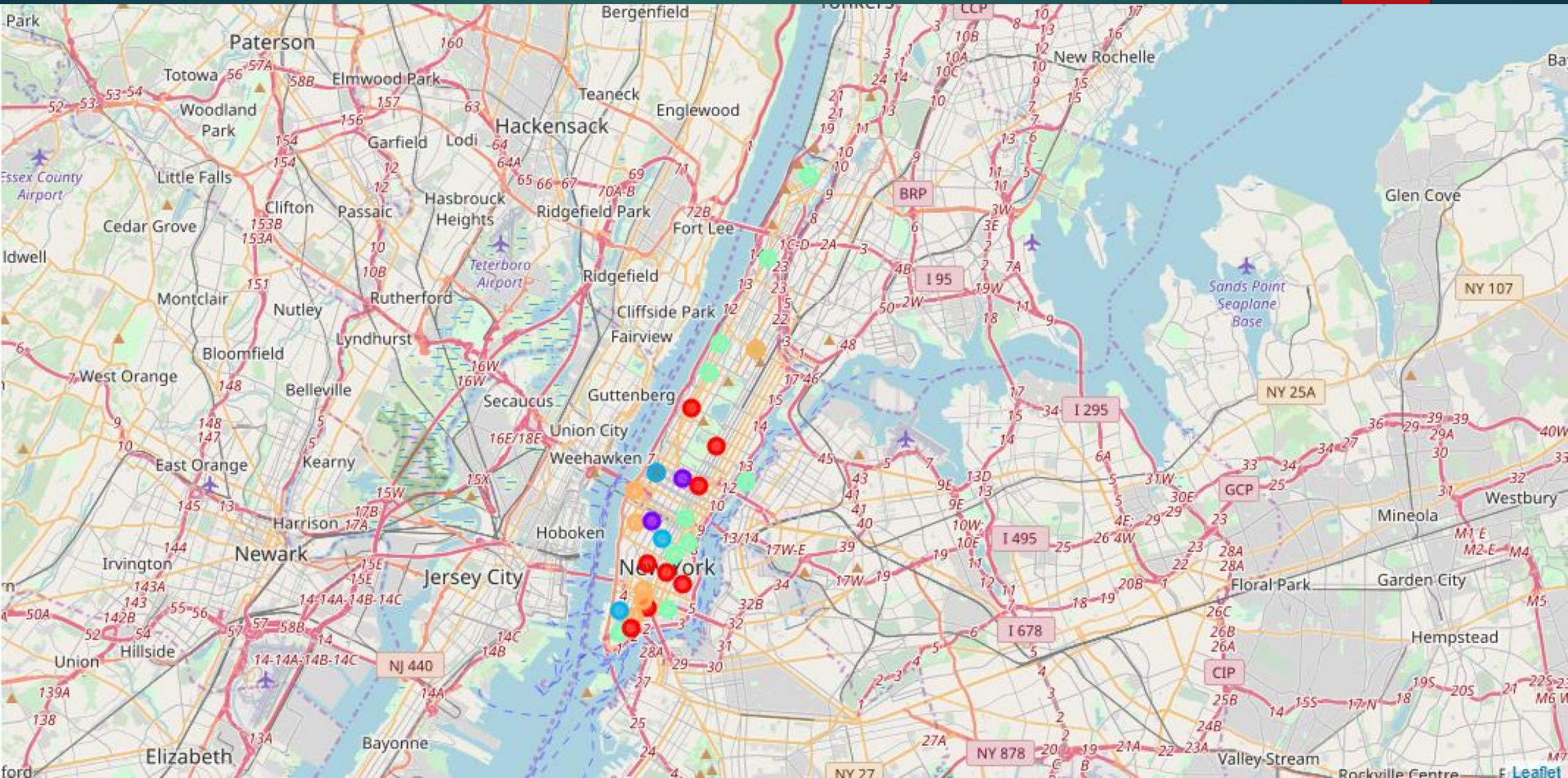
Methotology:

1. Collect Inspection Data;
2. Explore and Understand Data;
3. Data preparation and preprocessing;
4. Modeling

Map of Manhattan neighborhoods



K-Means clustering



Conclusion:

By analyzing the results according to our five clusters, we can see that all clusters could praise an optimal range of facilities and amenities. The first pattern we are referring to, i.e. Cluster 1 which have the bigest average price may target investor, who value the neighborhood of hotels, Italian restaurants, gym, concert hall, maybe for potential Italian investors. Cluster 2 also have high average price, but with a larger neighborhood of the restaurant. The CLusters 0, 3, 4 may target potential buyers who are more interest to live in area full of restaurants, shops.