

# **Capstone Project - The Battle of Neighborhoods: Relocating to New York City in 2021**

**Applied Data Science Capstone Project by IBM / Coursera**  
**Author: Natalia Volkova**

**Date: April 8th, 2021**

# 1. Introduction

1. Start of the global pandemic in March 2020.
2. People relocating.
3. Mass exodus reported from big cities, at the same time lots of people from smaller cities migrating to big cities hoping to get a job.
4. New York City is the most populous city in the United States and Manhattan is the most populous of New York City's 5 boroughs.
5. Looking at Manhattan neighborhoods to relocate.

## 2. Business Problem

To provide support to homebuyers which neighborhood in Manhattan would be a good place to start looking for a new home?

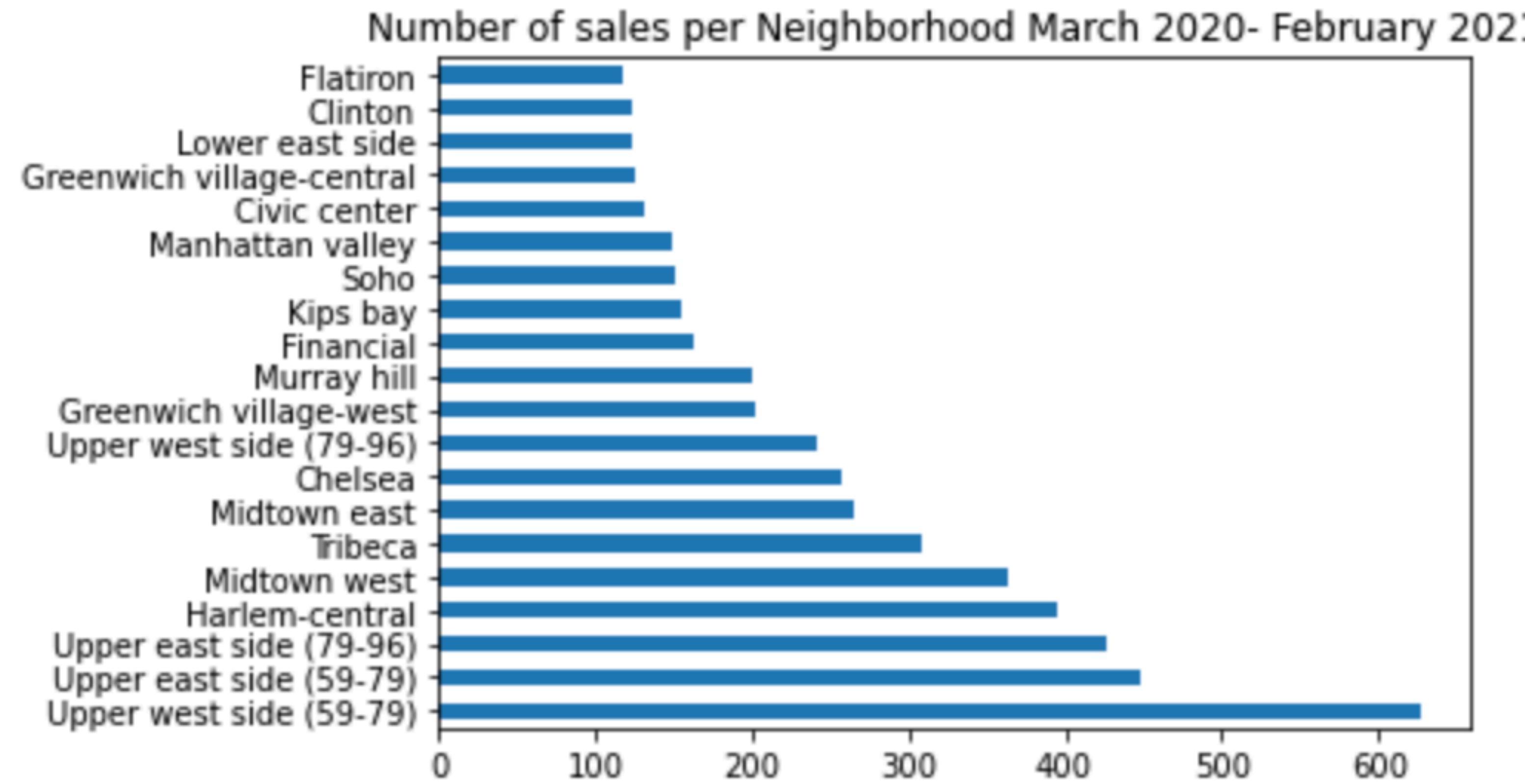
## 2. Solution

To recommend the neighborhood, I got the data of the recent Manhattan Sales from March 2020 to February 2021 and then explore neighborhoods for amenities and essential facilities.

### 3. Data and Methodology

- Data on Manhattan sales from March 2020 to February 2021 from <https://www1.nyc.gov/site/finance/taxes/property-annualized-sales-update.page> , New York City data latitudes and longitudes from [https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset) and data on amenities and essential facilities from FourSquare API interface.
- Methodology:
  - Collect necessary data
  - Explore and understand data
  - Modeling
  - Conclusion

## 4. Bar Plot Number of Sales in Manhattan per Neighborhood



Majority of sales in Manhattan happened in Upper West Side and Upper East Side.

## 4. K - Means Clustering

Since majority of all residential sales last year happened in two neighborhoods, I am going to look at Upper West Side and Upper East Side of Manhattan to get an idea about what those two areas have to offer.

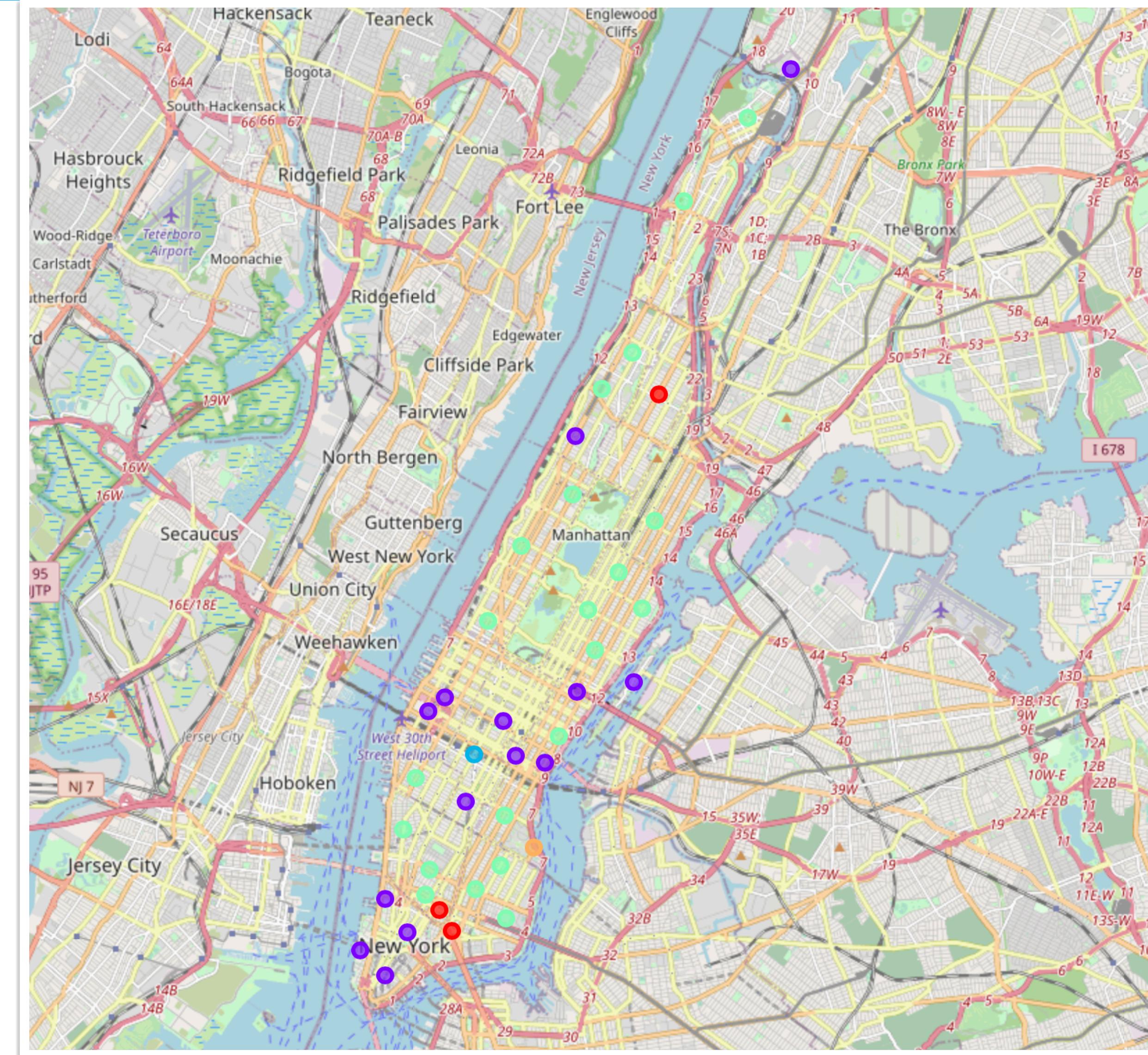
Cluster Neighborhoods.

- Manhattan neighborhoods are classified into one of five clusters based on similarity to each other in terms of categories of venues

Run K - Means to cluster the neighborhoods into 5 clusters.

Visualize the results on map.

# 4. K - Means Clustering Map



## 5. Outcome

Majority of sales in New York Manhattan Borough from March 2020 to February 2021 happened in Upper West Side and Upper East Side, they are in green colored cluster number 3.

Both areas have great choices of restaurants, bakeries, coffee shops, museums, yoga studios and gyms.

Great advantage of those two areas is close proximity to Central Park, walking distance 10-15 minutes depending on exact location in addition to many playgrounds and river walks on the outer edges of both neighborhoods.

It looks like it can be great place for families or single people.

Further recommendation in choosing one neighborhood over the other would be considering the budget size and examining prices of desired housing size.