

Capstone project

By : ANSHUL GUPTA

Contents:

1. Introduction Section:

1.1 Discussion of the "background situation" leading to the problem at hand:

1.2 Problem to be resolved

1.3 Audience for this project.

2. Data Section:

2.1 Data of Current Situation (current residence place) 2.2 Data required to resolve the problem

2.3 Data sources and data manipulation

3. Methodology section:

3.1 Process steps and strategy to resolve the problem

3.2 Data Science Methods, machine learning, mapping tools and exploratory data analysis.

4. Results Section: Discussion of the results and how they help to take a decision.

5. Discussion Section-Elaboration and discussion on any observations and/or recommendations for improvement.

6. Conclusion Section-Decision taken and Report Conclusion.

1. INTRODUCTION

1.1SCENARIO AND BACKGROUND

I have chosen CHANDIGARH, INDIA (my hometown) as the location to be used as current location for this project.

Being a growing town with many breweries, malls, gyms, clubs, restaurants, shops, cafes opening with a rapid pace, my hometown surely serves a great place to live

1.2PROBLEM TO BE RESOLVED

Many youngsters are migrating to foreign countries in search of jobs, better education and thus want to find an apartment at least similar to one they used to live.

Thus the problem is to find similar apartment in MANHATTAN, NEW YORK, USA with similar amenities which were around in Chandigarh

-as a reference I have included the map of venues in Chandigarh.

1.3AUDIENCE FOR THIS PROJECT

I think that this project will be useful to every person considering moving to foreign countries similar to USA, basically anyone migrating to developed nation.

Also I think this will help travel agencies partner with appropriate entities offering accommodation, homestays, lodging, etc.

Also this will surely develop my data science skills

The use of foursquare data and mapping techniques along with data analysis and visualization will help solve the core problem

2. DATA SECTION

- **DATA OF CURRENT PLACE (CHANDIGARH HERE)**

Data from FOURSQUARE will be used to identify the main and appropriate venues near Chandigarh.

This data will be superimposed on the map of Chandigarh, INDIA and thus will serve as reference to find desired location in MANHATTAN, NY.

- **DATA NEEDED TO SOLVE THE PROBLEM**

To make the best prediction of the required apartment in Manhattan, data we need:

- Information of the neighbourhoods in Manhattan with their geo data (coordinates-latitude and longitude)
- Information of the apartments listed for renting out in Manhattan
- List of basic facilities provided at each neighbourhood (top 5)
- List of metro/bus stations near and in each neighbourhood.

3. METHODOLOGY

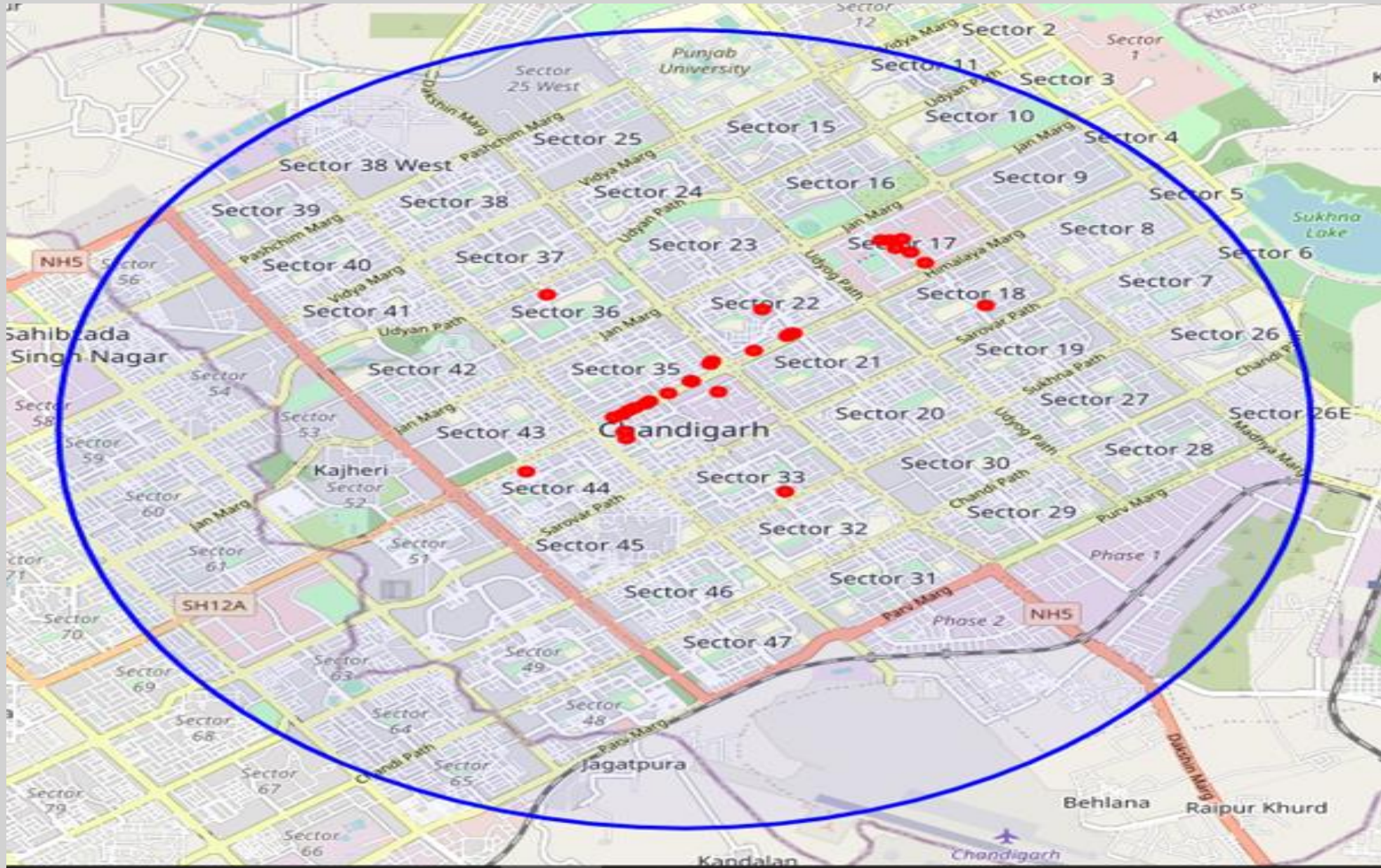
3.1 STEPS AND STRATEGY USED TO SOLVE THE PROBLEM

- FourSquare API to get the coordinates of the venues around the current residence (chandigarh) .
- compared those venues with clustered neighbourhood.
- Following rules were kept in mind while filtering out apartments in manhattan from which we will finally choose apartments:
 1. Must be close to 1 or more metro stations.
 2. Must be in the neighbourhood containing similar venues
 3. Rent price must be lower or equal to our budget .
 4. No of rooms

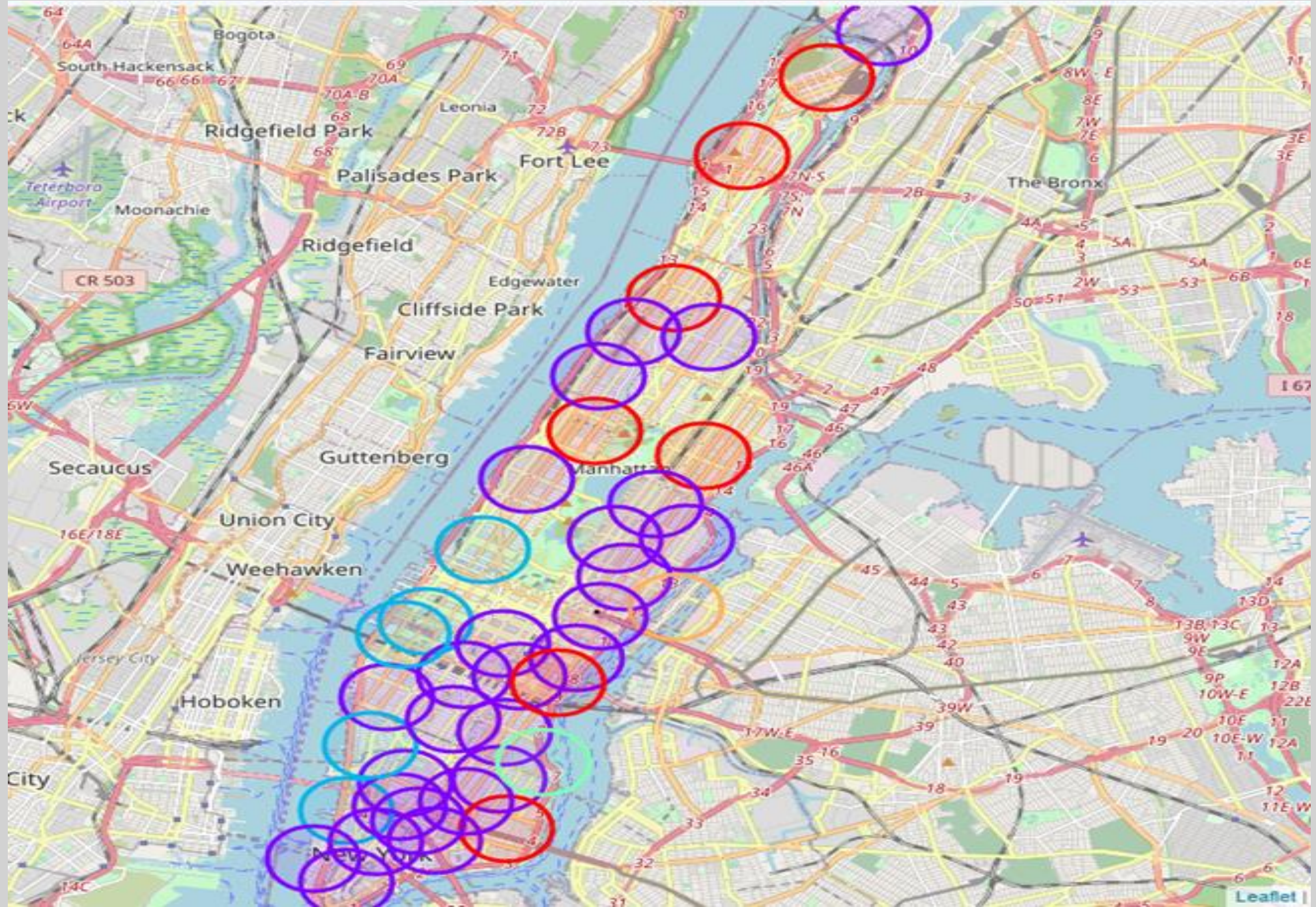
Maps gave us easy way to identify apartments by checking the requirements.

4. RESULTS

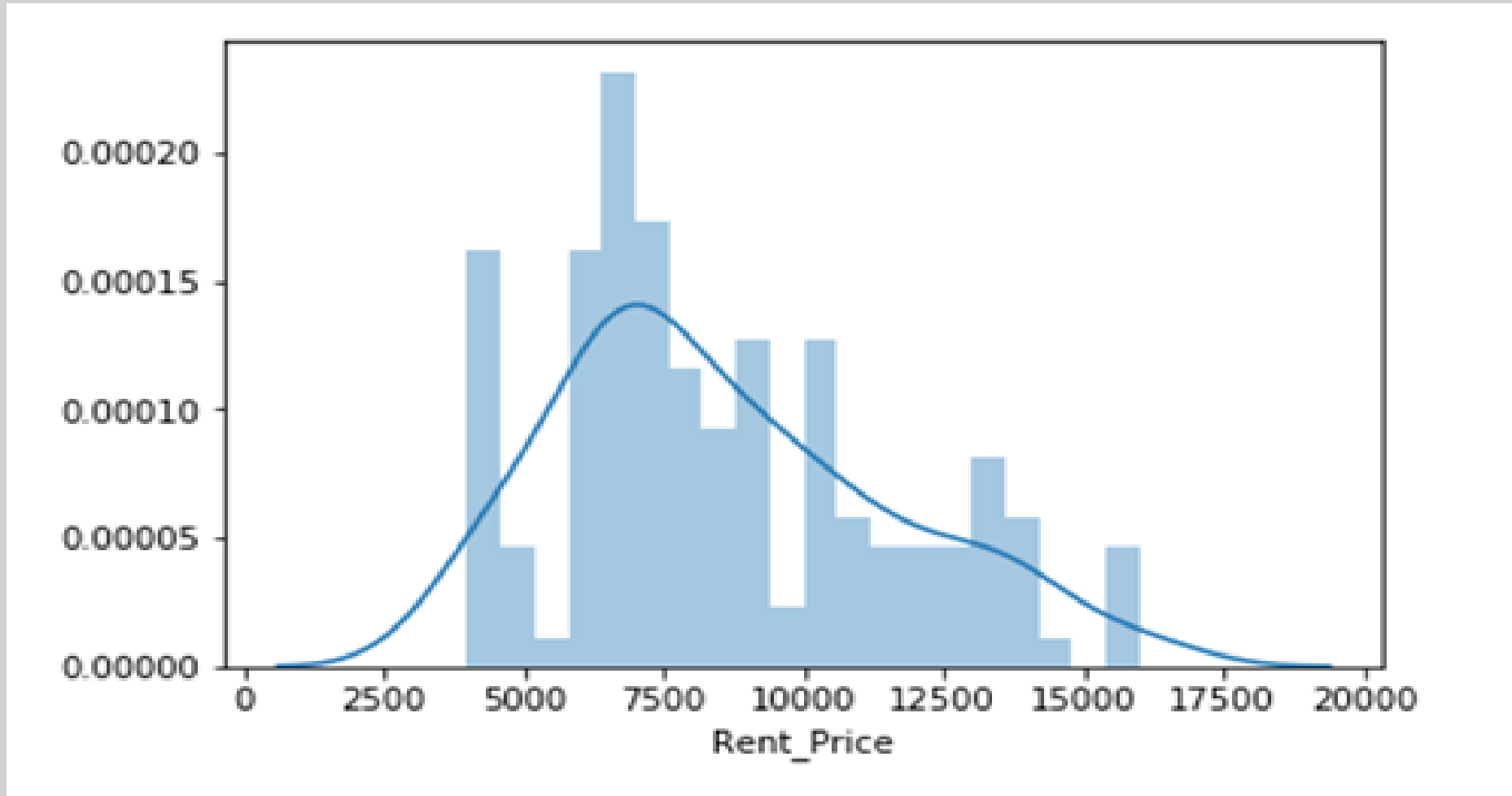
4.1 MAP OF CHANDIGARH WITH VENUES



4.2 MAP OF MANHATTAN, NEW YORK WITH CLUSTERED NEIGHBORHOODS ALONG VENUES

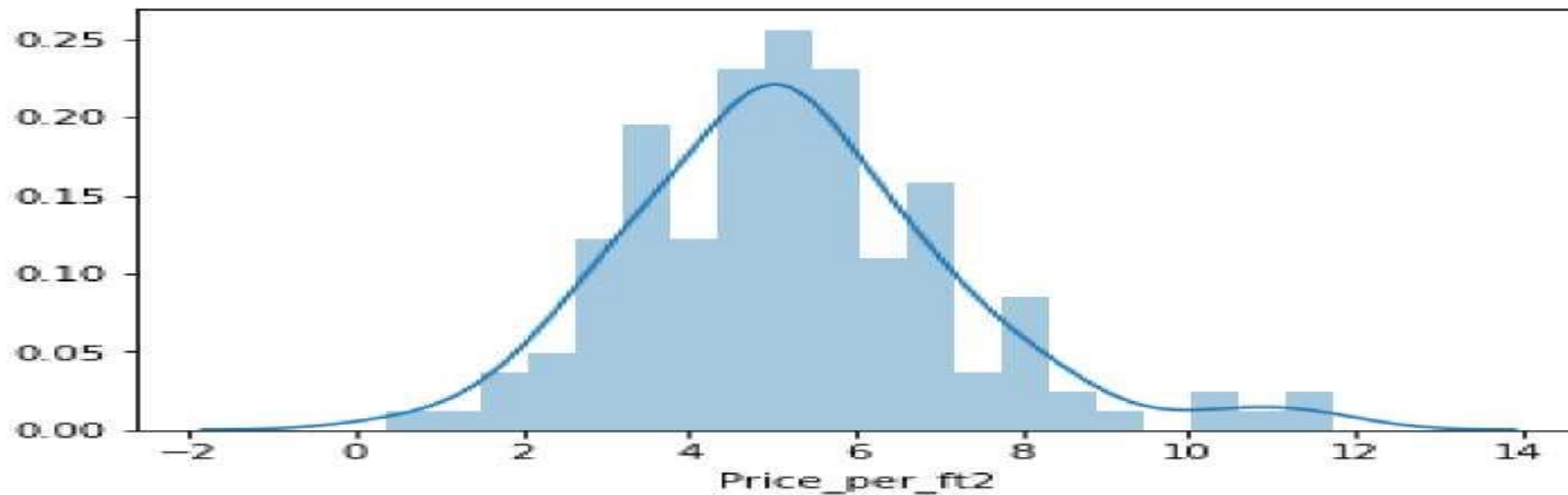


4.3 ANALYSING GEODATA OF APARTMENTS ON RENT IN MANHATTAN

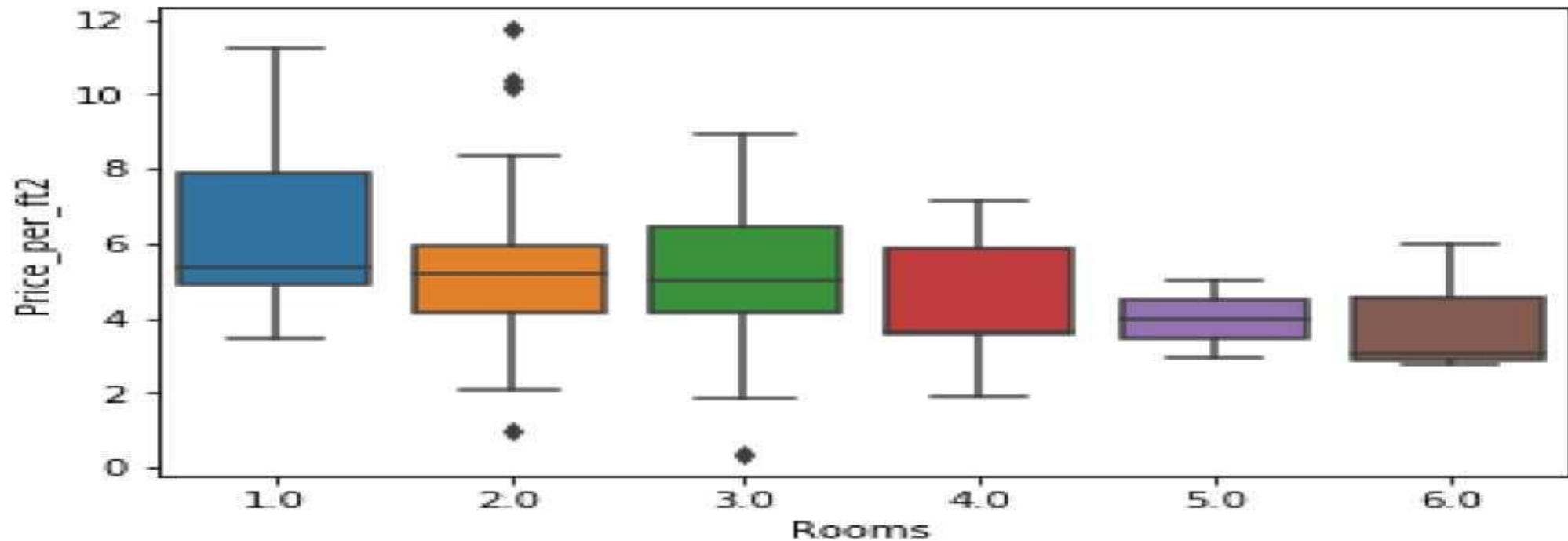


Observation from above plot- maximum rental apartments have rent close from 6000 to 7000 US dollar which is greater than my budget but being developed nation with better facilities they'll definitely charge more.

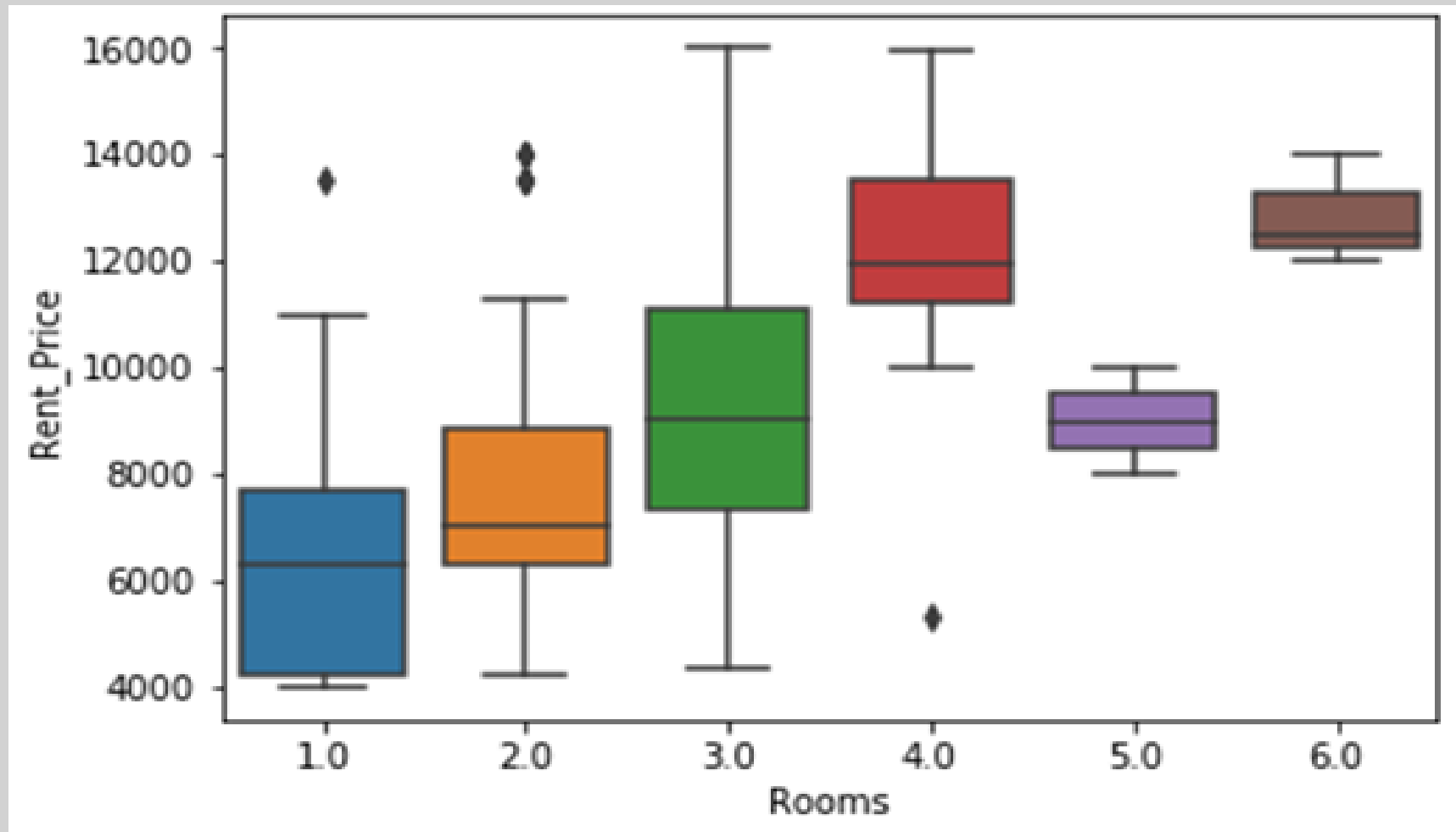
2. Plotting univariate distribution of Price per square foot



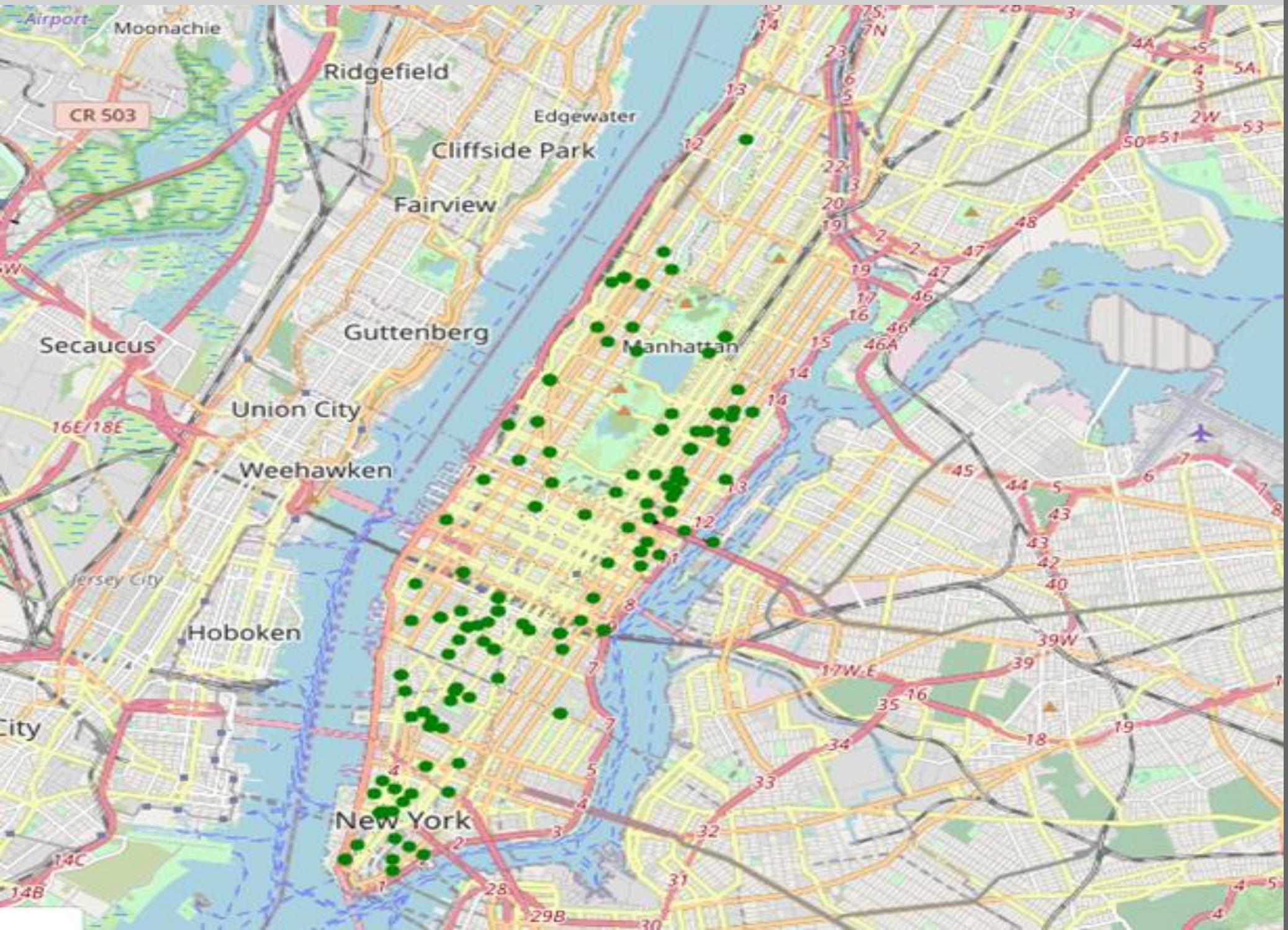
3. Boxplot of rooms vs price per square foot



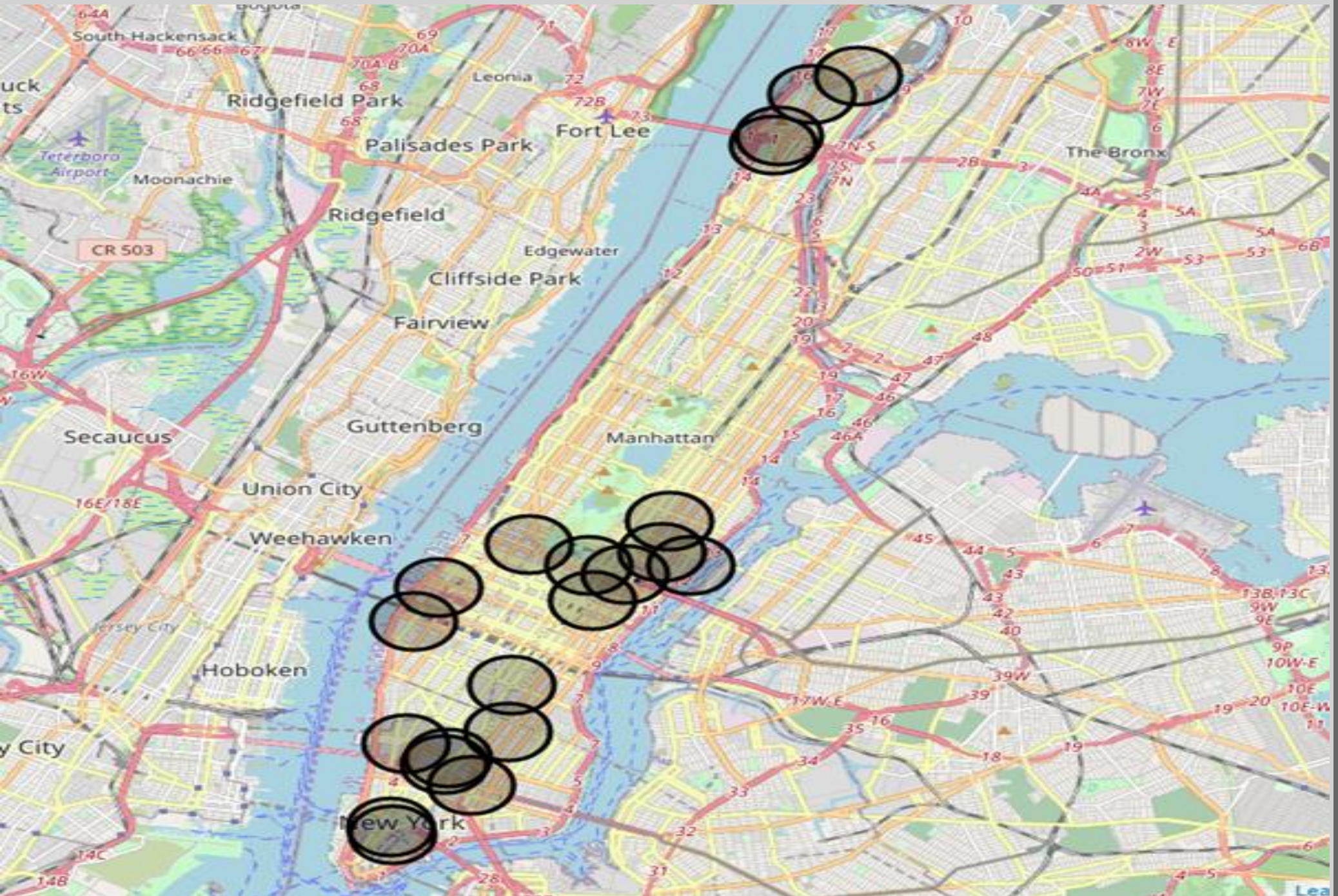
4. Boxplot of rooms vs rental price



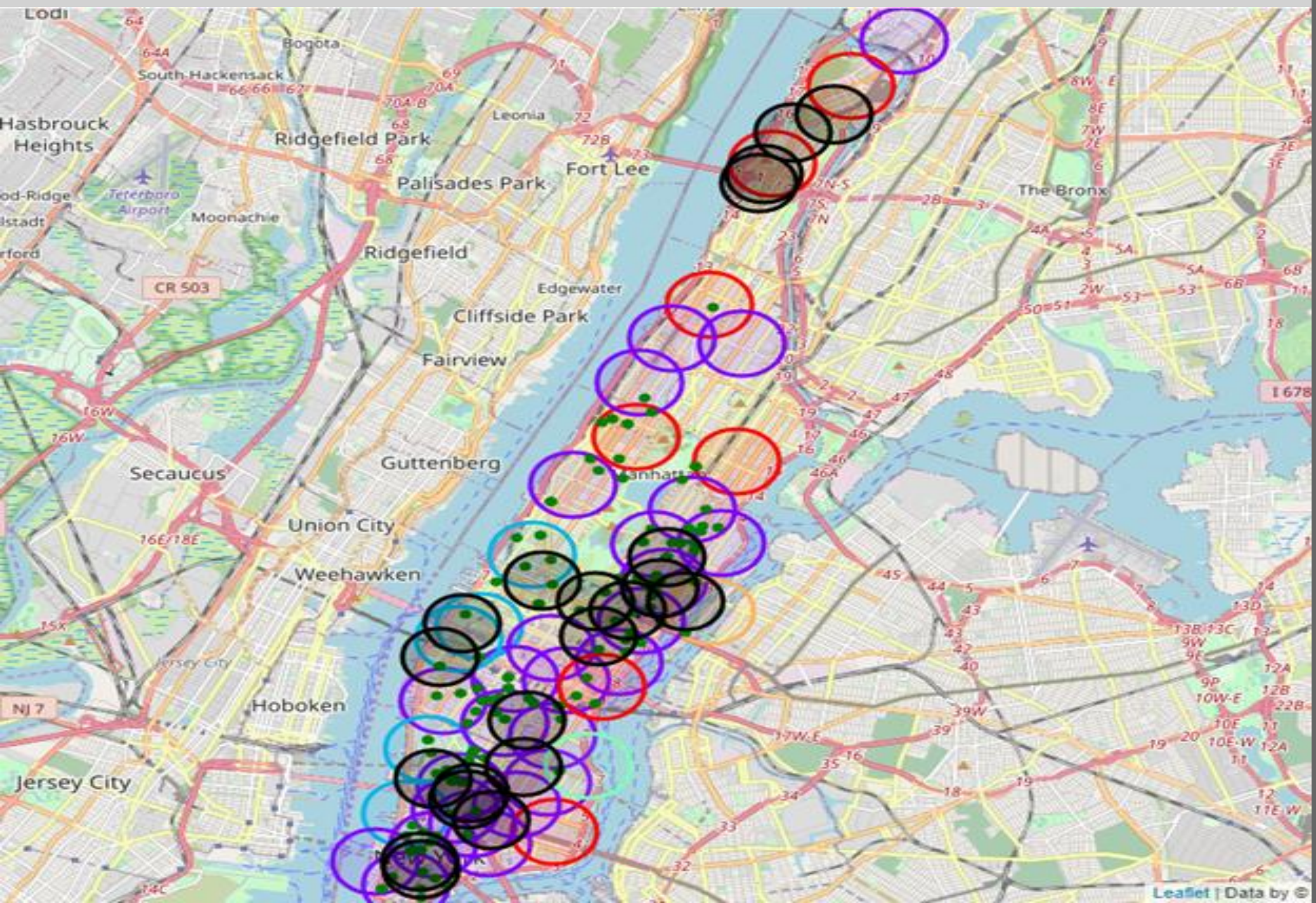
4.4 APARTMENTS ON RENT IN MANHATTAN, NY



4.6 MAP OF MANHATTAN WITH METRO STATIONS MARKED



FINAL MAP



THE FINAL MAP IS COMBINED MAP WITH ALL THE NECESSARY MARKERS :

- BLACK MARKERS TELL THE INFORMATION OF METRO STATIONS
- GREEN MARKERS TELL THE INFORMATION OF APARTMENTS ON RENT
- RED MARKERS TELL THE INFORMATION OF CLUSTER=0 NEIGHBORHOODS
- PURPLE MARKERS TELL THE INFORMATION OF CLUSTER=1 NEIGHBORHOODS
- SKY BLUE MARKERS TELL THE INFORMATION OF CLUSTER=2 NEIGHBORHOODS

5. DISCUSSION

- The model can help someone moving to manhattan to choose an apartment but to generalize it more , we need to take into account more diverse data , from different countries and different kind of areas.**
- Different person might require more facilities and bigger Apartment with more number of rooms and thus the model can help them also.**
- One can observe that the apartments on rent are not available in north of manhattan which is isolated a little.**
- Most apartments on rent plus metro stations as well as cluster venues can be found in southern part**

6. CONCLUSION

In this study , I found out how to find an apartment in New York, With atleast similar amenities as in chandigarh but better in ways such as metro stations nearby !



The above labelled apartment 211 East 70th street (4250\$ rent price) with nearest 72street Subway Station in LENOX HILL neighbourhood(cluster=1) is the one I have chosen personally by considering various personal as well as general factors and I can confidently say that the chosen apartment is the best deal at best price in my opinion.

THANK YOU