Coursera Capstone Project

By Navya S Nambiar

INTRODUCTION

Problem Description

- My young cousin will move to New York City to start studying at Manhattan college engineering ,we face a lot of problem in find suitable flat for rent special he doesn't eligible to stay at university housing. so that I try to find best flat for rent around undergrounds that can walking to the station.
- My cousin need the following conditions at new flat .
  - Apartment with min 1 bedroom and Max bedrooms
  - Rent not to exceed $1500/month
  - Max 10 minutes to walling till underground metro station
  - Fully live place a lot of restaurants , gum, cafes, and more at night live to be save for living
- Our comparison will be with Stratford city at London last beautiful place was visit this year "MOXY LONDON STRATFORD"

Audience

- First Audience will be any person who plans to move to New York City or any other major cities.
- And the second one is Data Science training that needs a powerful case study to train his / her skill

DATA

The heart of our project

Data Sources and Processing

- Manhattan neighborhoods were obtained from Wikipedia and organized by Neighborhoods with geodata via Nominatim for mapping with Folium.
  - https://en.wikipedia.org/wiki/List_of_New_York_City_Subway_stations_in_Manhattan
- List of Subway stations was obtained via Wikipedia, NY Transit web site and Google map
  - https://www.google.com/maps/search/manhattan+subway+metro+stations/@40.7837297,-74.1033043,11z/data=!3m1!4b1
- List of apartments for rent was consolidated from web-scraping real estate sites for MH. The geolocation(lat,long) data was found with algorithm coding and using Nominatim.
  - http://www.rentmanhattan.com/index.cfm?page=search&state=results
- Folium map was the basis of mapping with various features to consolidate all data in ONE map where one can visualize all details needed to make a selection of apartment

METHODOLOGY

The heart of our project

Methodology

- Our Methodology is primarily based on described our data on MAP, in order to facilitate locations for rent
- All information will consolidate on one MAP, where my cousin can see all rental statistics in one place
- Map exhibit two famous places around apartment and subway station records

Data science Applied tools

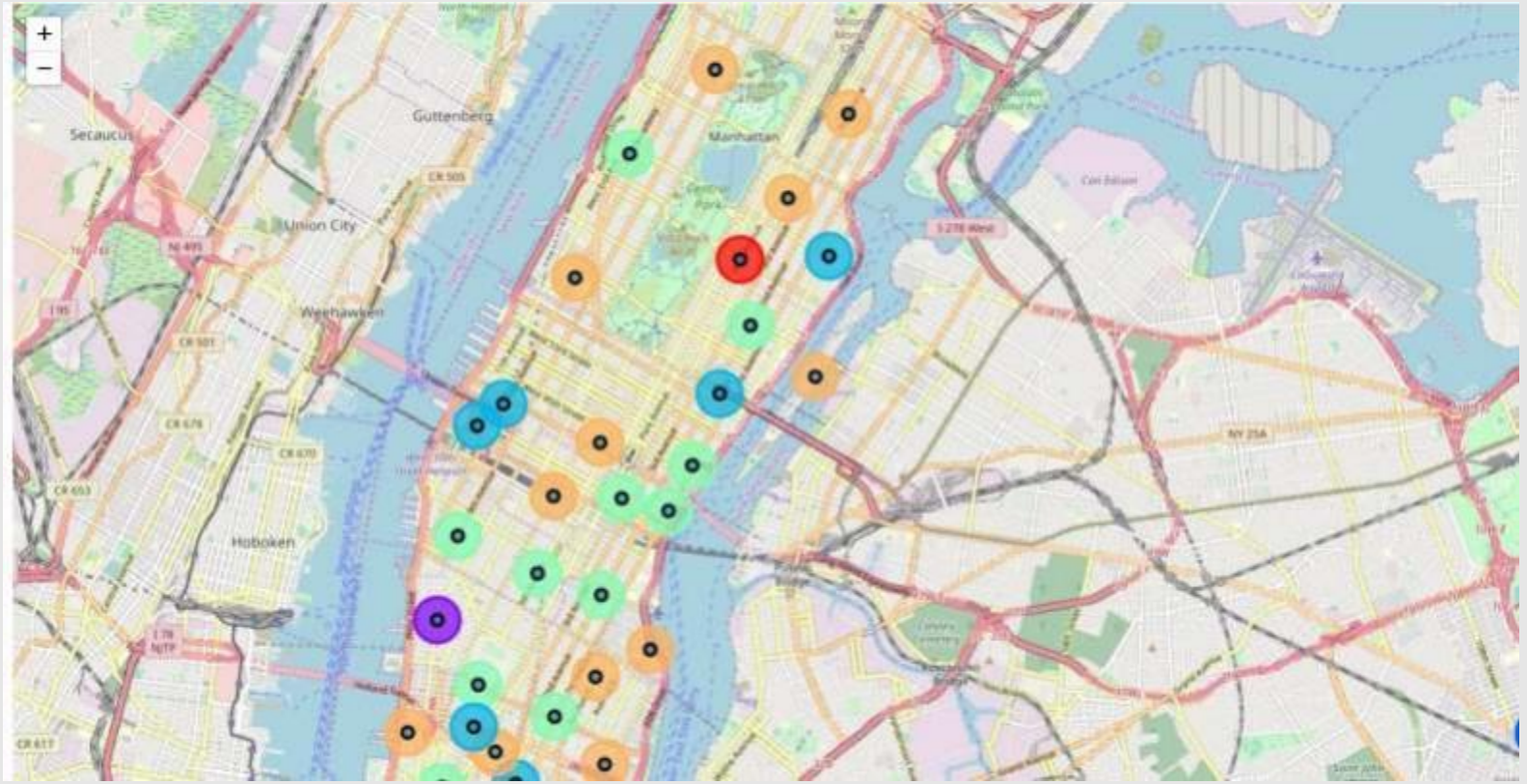
- Web-scraping of sites is used to consolidate data-frame information which was saved as csv files for convenience and to simply the report.
- Geodata was obtained by coding a program to use Nominatim to get latitude and longitude of subway stations and also for each apartments for rent listed.
- Geopy_distance and Nominatim were used to establish relative distances.
- Seaborn graphic was used for general statistics on rental data.
- Matplotlib used CM module that provides a large set of colormaps, functions for registering new colormaps and for getting a colormap by name, and a mixin class for adding color mapping functionality.
- Sklearn used KMeans for applying cluster on generated Data
- Maps with popups labels allow quick identification of location, price and feature, thus making the selection very easy

Execution

EXECUTION

Run our codebase

Manhattan University Map

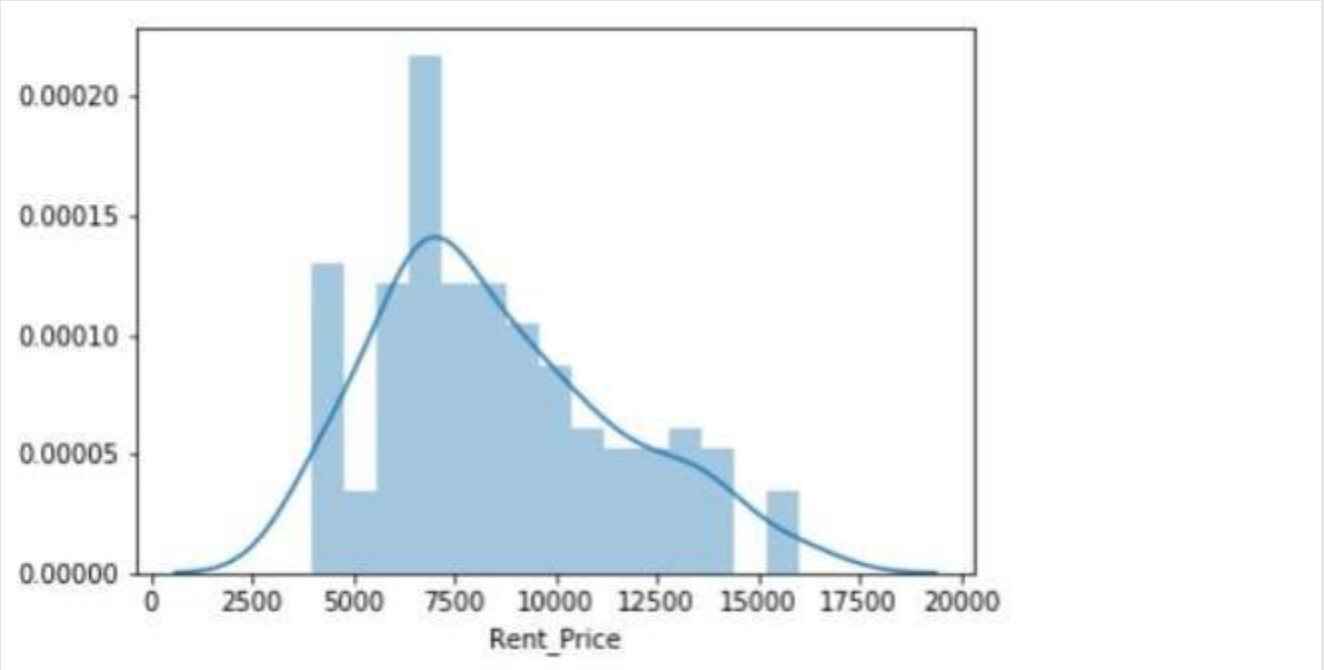


Offred flats for rent

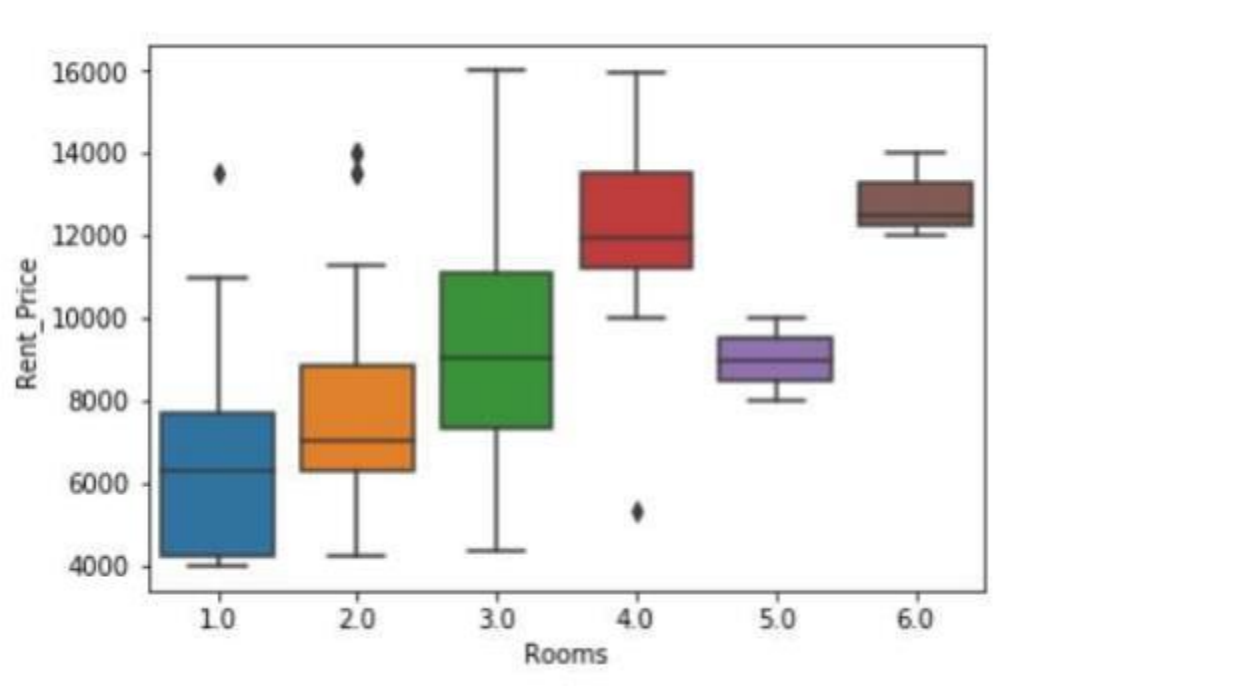
Out[55]:

	Address	Area	Price_per_ft2	Rooms	Area-ft2	Rent_Price	Lat	Long
0	West 105th Street	Upper West Side	2.94	5.0	3400	10000	40.799771	-73.966213
1	East 97th Street	Upper East Side	3.57	3.0	2100	7500	40.788585	-73.955277
2	West 105th Street	Upper West Side	1.89	4.0	2800	5300	40.799771	-73.966213
3	CARMINE ST.	West Village	3.03	2.0	1650	5000	40.730523	-74.001873
4	171 W 23RD ST.	Chelsea	3.45	2.0	1450	5000	40.744118	-73.995299

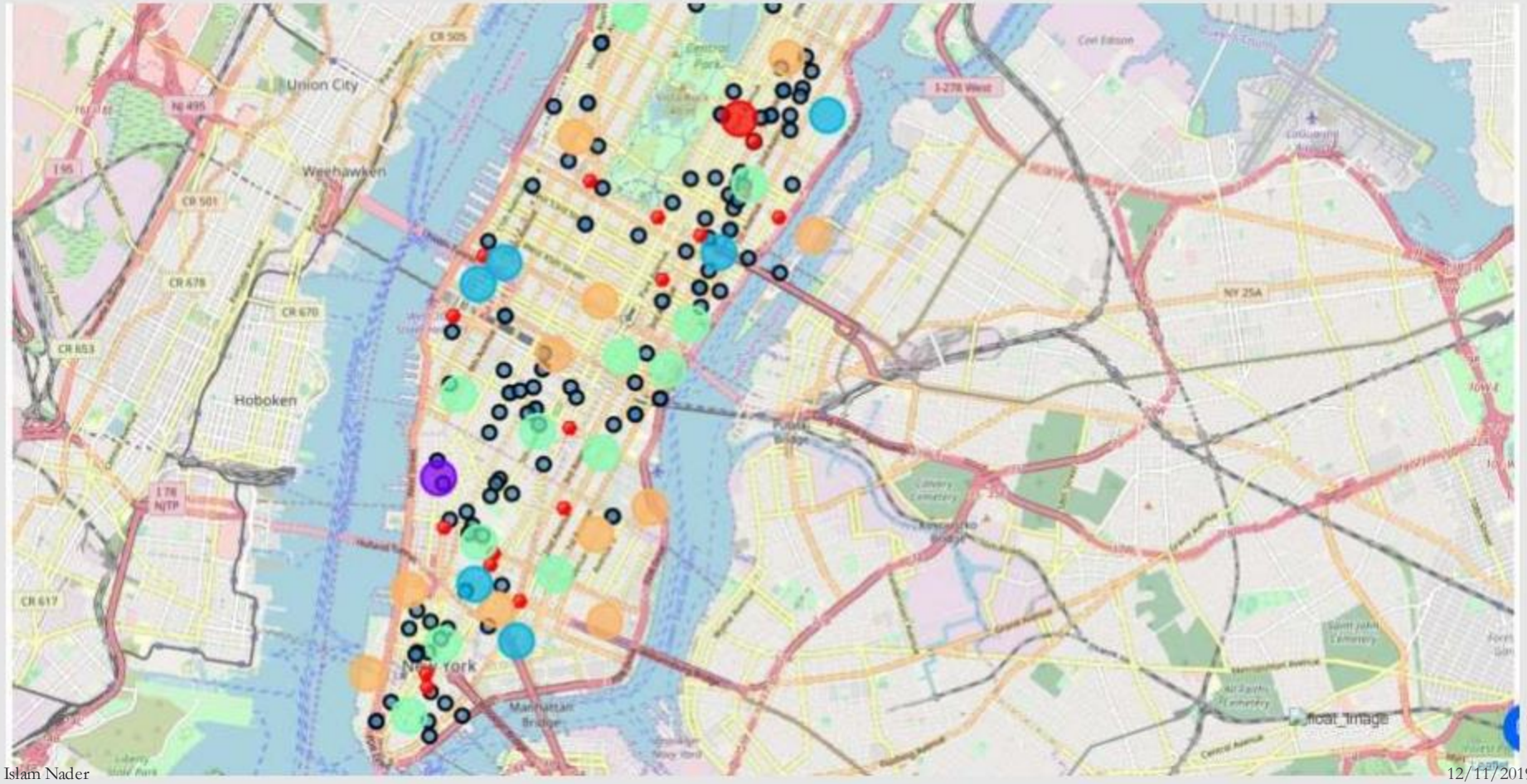
Offred flats for rent per rent price



Offred flats for rent price vs No. of rooms



Final Flat for rent and subway stations



Map key

- Subway
- Flat

CONCLUSION

Conclusion

- In general, I am thankful IBM for wonderfully Certification Course that improve all my thinning and data analysis thinning
- Folium is a very powerful technique to consolidate information and make the analysis and decision with integration with Foursquare API . I would recommend for use in similar situations .
- One must keep abreast of new tools for DS that continue to appear for app
- For my all class-met it's recommenced to continue study **Advanced Data Science with IBM Specialization**

.