Capstone Project Report

Coursera- IBM

Submitted by:

Shaurya Gupta

BUSINESS PROBLEM

New York City is one of the widely recognized city in the world because of its quality of life, job opportunities, cultural diversity. People from all around the world shift to New York for the same. One of the issues faced by all is settling down in the right neighborhood at the right cost which has all the necessary facilities.

With authentic data on NYC properties along with data science techniques, one can make derive the useful information about current pricing in different localities of NYC while considering other factors of his choice.

This would help the potential client to make an informed decision about buying a suitable property. This could help potential clients looks to buy suitable property in NYC but are skeptical due to lack of knowledge and volatile market conditions.

The stakeholders of this project could be Government of NYC, Sellers, Buyers and Real estate agents.

**DATA**

Foursquare location data will be used via API .

**Type of data-**

JSON

**Description**

Location coordinates obtained by Foursquare API calls.

To determine the proximity of various amenities as per the client’s requirement, Foursquare location data is used. In this project, data of housing properties of New York City will be used.

**Source: -**

<https://foursquare.com/>

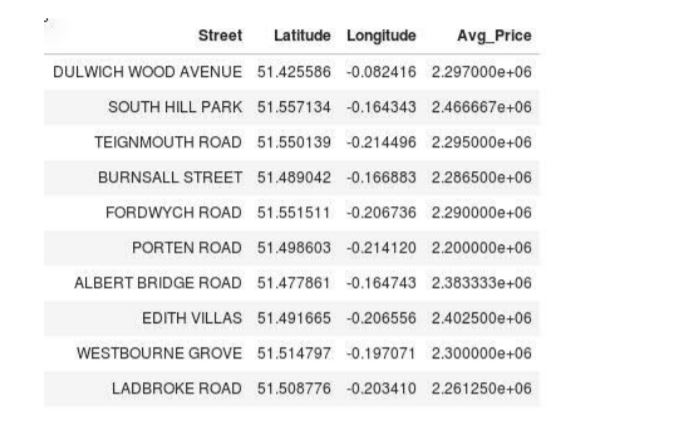
METHODOLOGY

The code developed as a part of this project does the following-

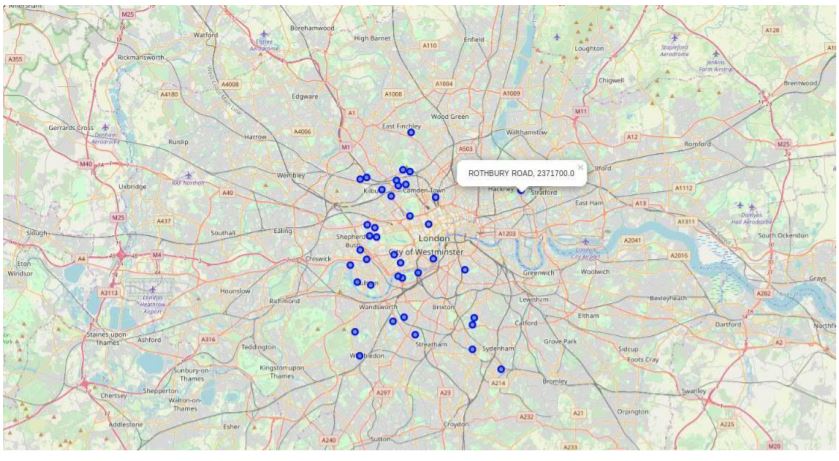
1. Parse the necessary data from the price paid dataset.
2. The data is cleansed and any data of sales agreements which predates 2016 is deleted from the dataset.
3. The data is further condensed by selecting it only for New York which is area of choice in this project.
4. Unique “Street names” in NYC where recent transactions for sale of property were done are filtered from the dataset.
5. Location coordinates (latitude, longitude) of these street names are fetched by making API calls to Foursquare. A separate one-time Python script was developed to fetch this data and store it in a CSV file.
6. The average price of property on each of these streets is determined by taking a mean on recent transactions of sale of property on respective streets.
7. Based upon the budget of the client, the current average prices are compared and all recommendations for the locations are made by plotting them on map of London. The locations popups are labelled with the respective street names and their average property price.
8. The recommended locations are further fed into Foursquare API calls to determine various venues in proximity to them. All reported venues are then tabulated and presented to the user.
9. Important facilities like Hospitals, Grocery stores, Elementary schools, High Schools are searched in vicinity of each location and then reported in a tabular form to the user.
10. The best recommended housing places are shown.

**RESULTS**

A list of such locations is presented to the user with location coordinates and most recent average prices.



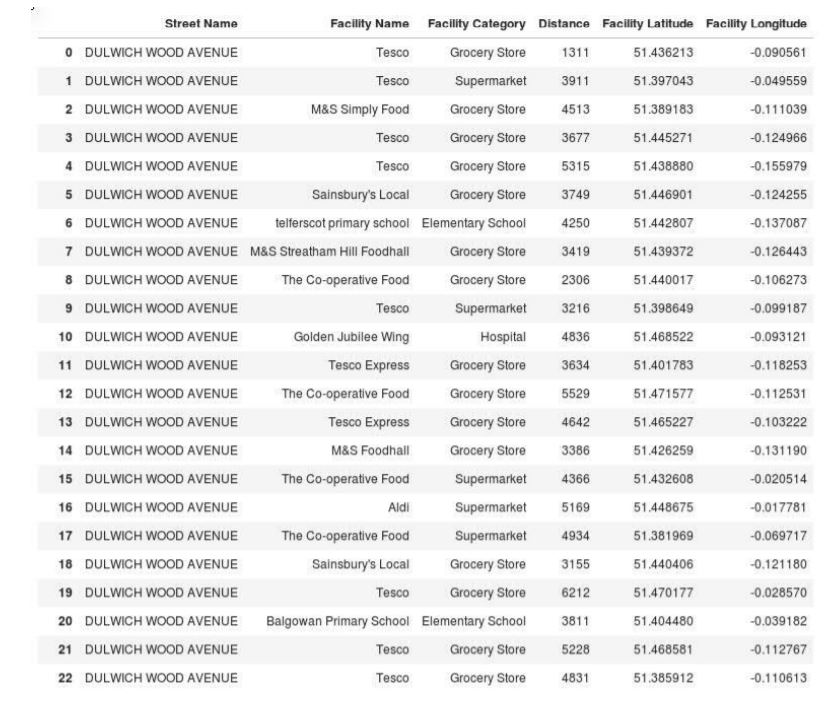
These recommended street names are plotted on the map of NYC with the average market prices.



Further, following venues are enlisted for the user to make an informed decision while choosing a location.



Important facilities are also presented to the user in a tabulated format to take care of his familial needs.



DISCUSSION

Based upon the findings in the results section, the user can take a conscious decision about choosing a street/ location based upon his/ her requirements.

The results section enlists 39 places where a prospective client can buy a property based upon his needs and choices. Such choices would be affected by the venues and facilities which are close to the property which match against his familial needs.

Few possible cases are: -

1.A prospective client with elders in the family would be inclined to choose a location where hospitals and grocery stores are located in close proximity.

2.A prospective client with kids in the family would choose a location where elementary and high schools are close-by. He would also like to choose a place with parks and other venues to accommodate his family are in the close vicinity.

3.A bachelor would be inclined to choose a property which has pubs, bars, entertainment places close to the property.

CONCLUSION

The decision of a buyer is influenced by the familial needs, personal biases. So, based upon the findings summarized in the results and discussion sections, following conclusions can be made: -

1.While making recommendations to a prospective client, it is imperative to know his/ her immediate needs and requirements besides the budget. This would help to catch his/ her attention.

2.Knowledge about the most recent market prices can be very helpful for the client and can help him take a decision.