CSE 587 – Data Intensive Computing

Problem 4 – Statistical analysis to support new data product

Author : Bhavika Jain
bhavikap@buffalo.edu

Objective: Do statistical analysis on tweets for apartmental rental and real estate collected from twitter. Armed with data based on analysis, provide recommendations to Real Direct to extend their business offerings.

Approach:

In order to do statistical analysis of tweets on apartment rental and real estate etc, I collected tweets over a period of one week from Twitter on topics like 'House buying', 'Apartment Rent', 'Apartment buying', 'House Rent' etc. I also collected tweets from timeliness of real estate agents like 'RealtorSSD', 'MIrealestate' etc in order to arm myself with more data for analysis.

Statistical analysis:

I parsed each of the tweets collected to search for terms like 'rent', 'buy' or 'sale'. For each of the 'rent', 'buy' and 'sale' tweets, I counted the number of tweets that are posted for each of the boroughs 'Manhattan', 'Brooklyn', 'Bronx', 'Queens' and 'Staten Island'.

Below plot shows the distribution of tweets for rent and buy across each of the Boroughs.
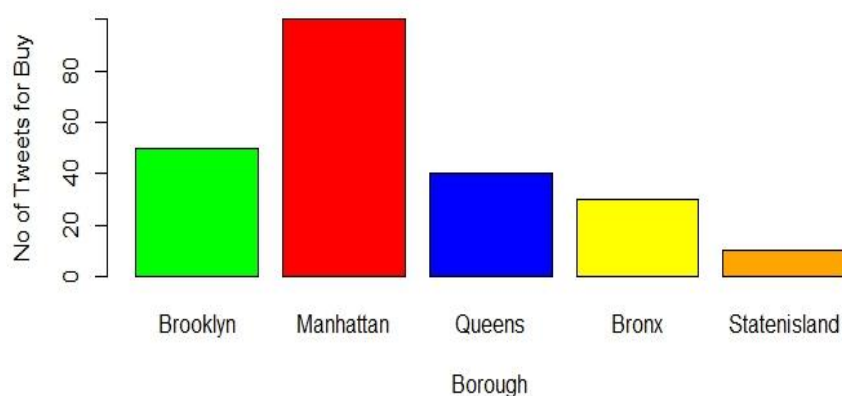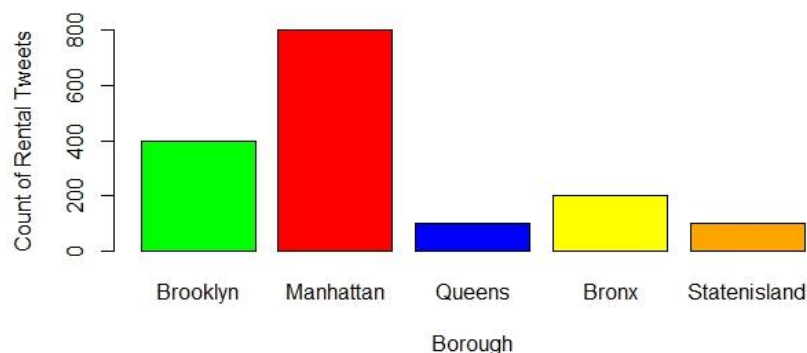

In the plots above, we can clearly see that the number of tweets for rental are more than for buying the property across all the Boroughs. Based on this statistics, I would recommend Real Direct to extend their business and create a new product offering apartment rental. Real Direct can also charge some percentage of rent as commission for any property rented through their platform from landlord/owners.

Further recommendations can be made based on the counts of tweets for 'rental property' posted on Twitter for NYC state and other states in USA.

As shown in the plots below, the count of tweets for both buying/rental is more in other states compared to NYC and also across all the Boroughs. Through this we can infer that other states are more popular for buying/apartment rental compared

Author : Bhavika Jain
bhavikap@buffalo.edu

to NYC. Therefore, I would recommend Real Direct to expand their business in other states of USA also and not just in NYC. This will eventually be beneficial for their business due to high chances of gaining profit.
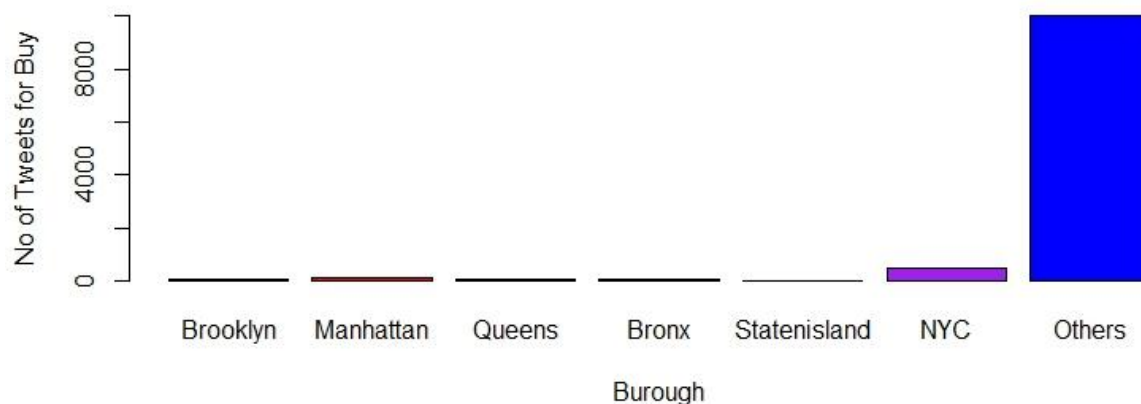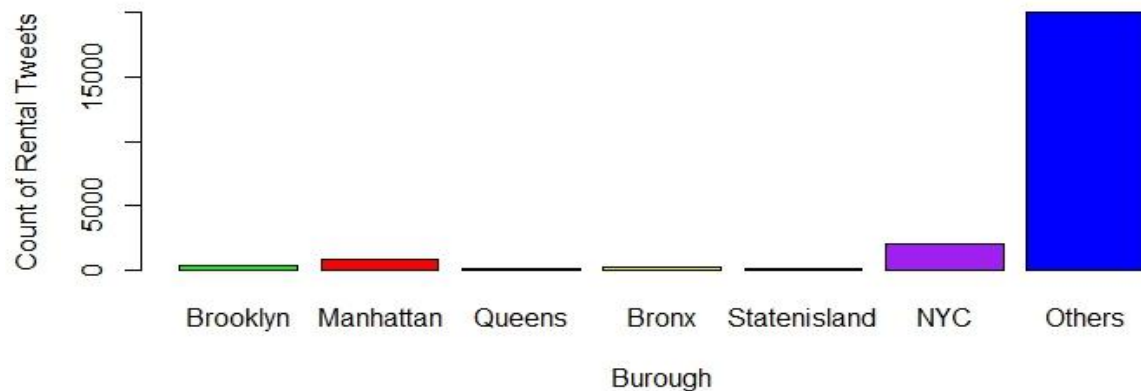
The property rentals, buying and selling mentions across locations along with New York City area has been shown in the map below. It is evident that a lot of people of are looking for properties to buy, to rent or to sell in popular cities like Los Angeles in California, Austin in Georgia and many other cities across United States of America and also in many developed and developing countries like India, Brazil, France, England.

The red dots in the map shown below are the places where the tweets originated from either by people living in that place or from local real estate agent. Either way, Real Direct can advertise their platform to those people and help in selling/renting/buying property.

Author : Bhavika Jain
bhavikap@buffalo.edu

Author : Bhavika Jain
bhavikap@buffalo.edu
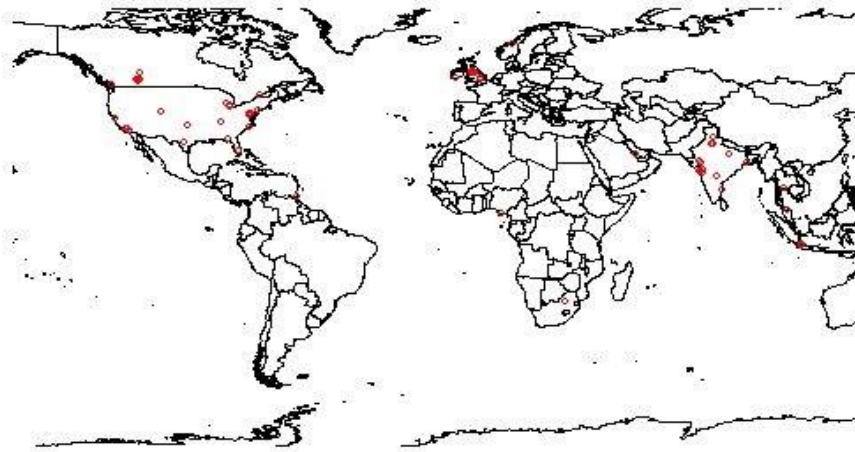
Conclusion:

Tweets can be further parsed to get the price and location of the property and these prices can be used to predict the revenue that Real Direct might be able to generate if they expand their business. We can also gather information from other sources like Facebook, Craigslist to get more data and make more meaningful conclusions and predictions.

Author: Bhavika Jain
bhavikap@buffalo.edu