Proposal for new Product Service line in *realdirect.com* based on Analysis of multi sourced data

Data Sources

In order to base the proposal for a new product, the data was taken from 3 different sources; twitter, rolling sales data of NY boroughs, and real estate market research data from Zillow.com.

Each of these sources serves a particular purpose for the proposal formation. Using analysis of tweets, the customer need and sentiment for rental spaces and related products can be gauged [*Alexander Pak, Patrick Paroubek*]. This can give a sense of the kind of product customers may desire. However, in order to the process to be random, the data collected should not be coerced in any manner and should be purely generated by the customers by their own accord. Also the user base should also be random, as restricting the analysis to only a subset of twitter users (either by content or preferences) would lead the data to be skewed highly in favor of a particular preference.

By analyzing the rolling sales data, boroughs that have high sale prices or low unit sales can be determined. These boroughs, due to their high prices would be better candidates for the rental space product line as an average customer is likely to pursue rental spaces than buy the units within these boroughs.

Zilllow.com is a popular real estate website. They have publicly released their market research data. This data analysis different markets throughout the United States an overview of national and local real estate and rental markets. This public data considers various metrics such as Sale prices, Rental Listing prices, price-to-rent ratio etc. analysis of some of these metrics can give an insight into the marketing and sale potential of a given market for the targeted region.

1] Twitter Analysis

Twitter, due its large volume, can be useful in estimating customer needs. In order to get appropriate tweets, twitter library was used to extract tweets pertaining rental and real estate related searched in New York. The Tweets were collected over a week and then their text corpus analyzed to find user word usage pattern. For this a word cloud was used as it gives a clear visual analysis of the most repeated terms/words during this time period. In order to make the data clean, stop words were removed and then words were reduced to their roots to form a more general analysis of the data. The word cloud formed is shown.
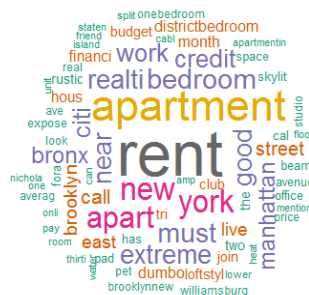
The most common terms related to real estate are found to be *rent, apartment, studio* while terms like *near, space, skylit, budget* indicate some of the user preferences. In locations, New York is the most common followed by *Brooklyn, manhattan, staten island.* Although this is not a quantitative measure of user behavior, this does give a correct measure of the most common needs for a user and justifies the reasoning for following a new product aimed at rental real estates.

2] Zillow Research data

The Zillow data was analyzed to determine the year on year projected growth rate in rental markets across the United States. The projected growth rate for New York is higher than the national average and higher than many other metropolitan areas like Washington DC, San Diego, Chicago etc. A high growth rate would in turn imply a higher valuation and returns from rental property. This makes the introduction of a rental real estate product viable as property owners will look into ways to get returns from rentals
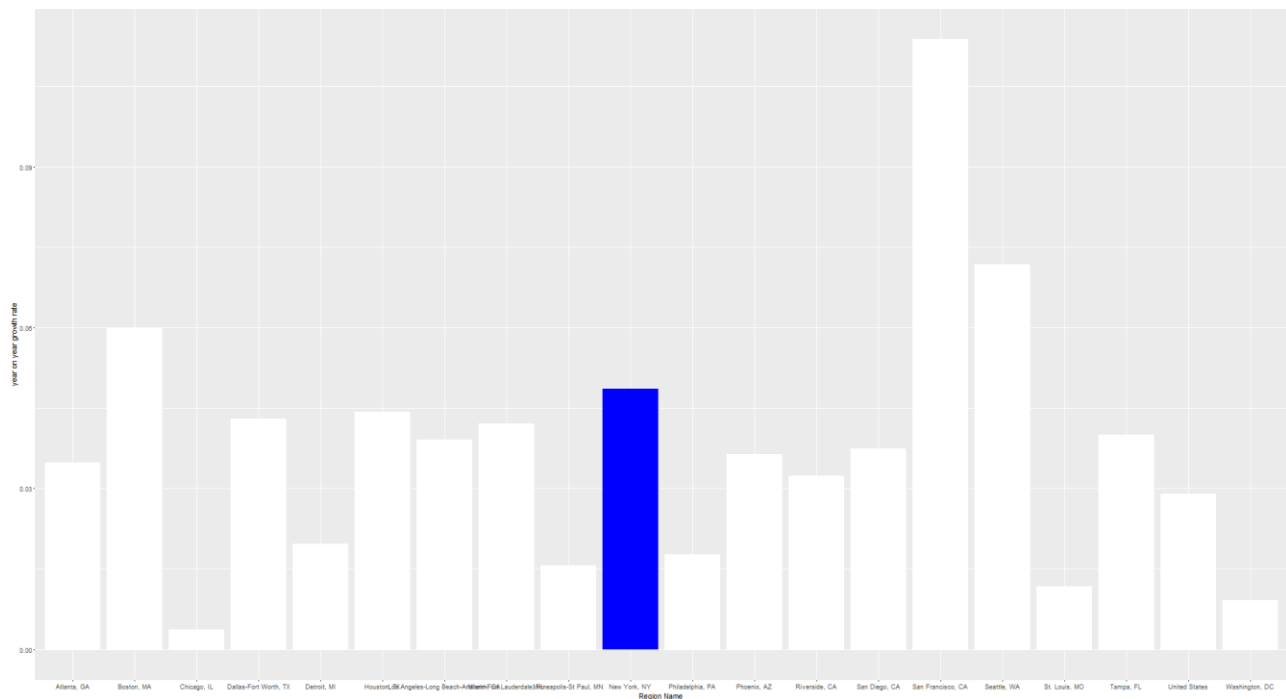

The price to rent ratio* is a metric associated at the individual home level, where the estimated home value is divided by 12 times its estimated rental value at the current market evaluation. This gives a median of all home-level price to rent ratios of a given region. It was observed that among all the major metropolitans for which this metric was estimated, New York has the third highest price to rent ratio. This implies that for a possible tenant looking for a property, it is cheaper to rent the space then buy it at the selling price. This also implies that for against a given property rate , the rent is cheapest in New York.
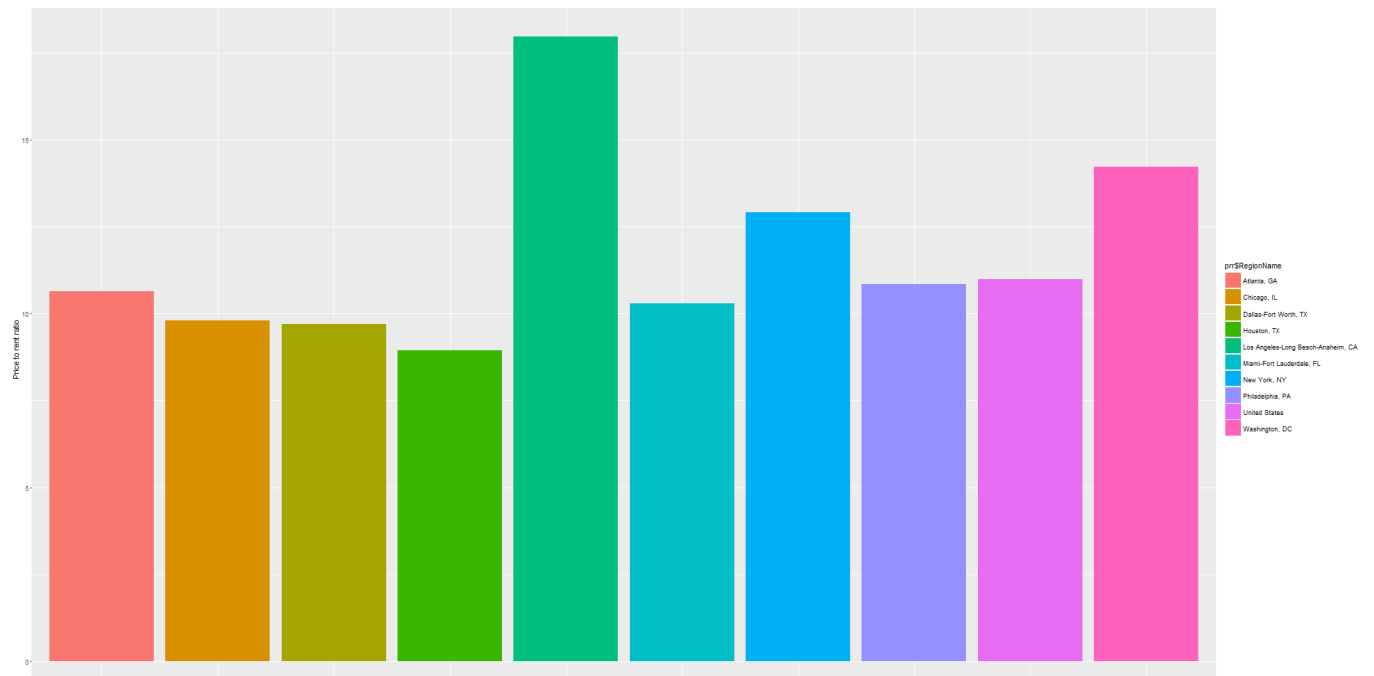
*estimated within the Zillow Data

The market health index is an indicator of the current rental market health of a region relative to other markets nationwide. An analysis of the largest metropolitans showed that New York scored highest in the index, which along with the higher growth rate prediction gives enough evidence to support the claim that rental market could be both a growing and a safe marketing strategy for realdirect.com
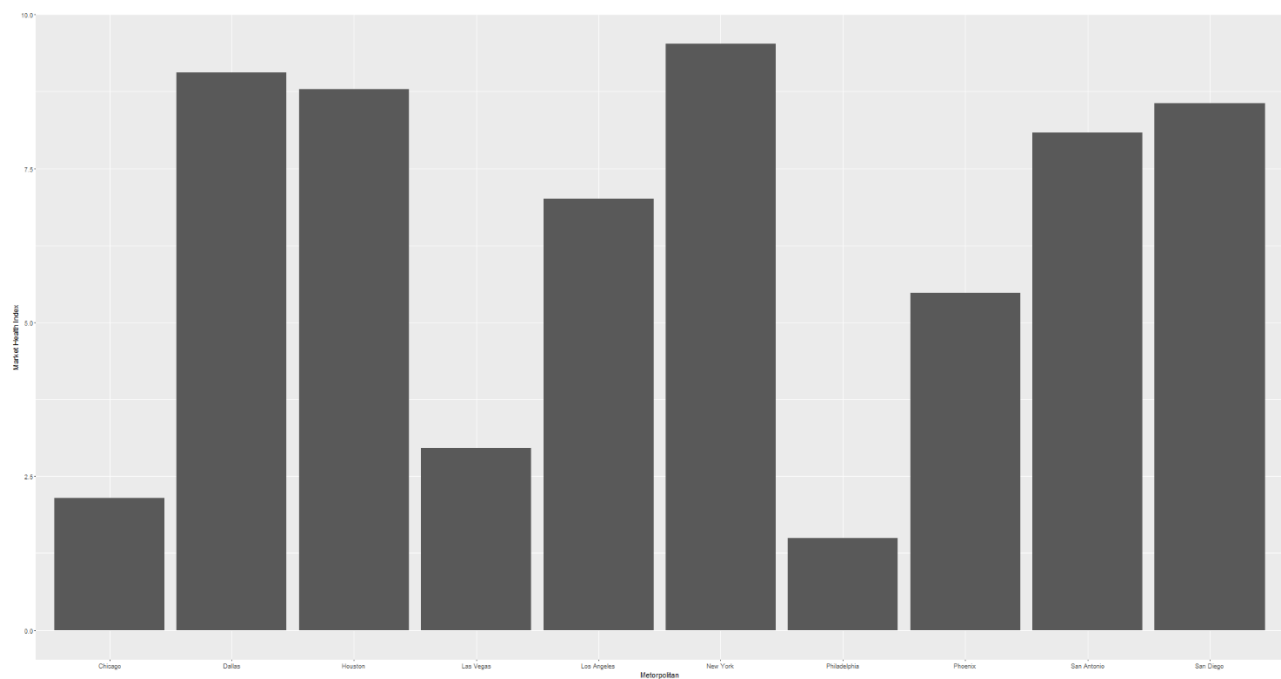
3] Analysis of Borough rolling sales data

Analysis of the borough rolling sales data shows the mean sale prices within each borough. This in turn can helps in evaluating the market prices for the boroughs in New York. With such high prices, it will be pertinent to introduce rental market to get access to market of people that want to live in New York but do not have access the capital to directly buy real estate.