Problem 4

Problem Description:

This problem involves analysis of Tweets and the Rolling Sales data set to provide a profitable solution to Real Direct to expand their business. For this problem, data was fetched from Twitter using the keywords "apartment", "condo" and "house". More than 20000 tweets were collected over a three-day period. This data was cleaned to remove special characters from the text and the text was then converted to lower case for better analysis. The rolling sales data was loaded into R and cleaned. In both the data sets we look for keywords related to apartment rentals and sales and analyze the popularity.
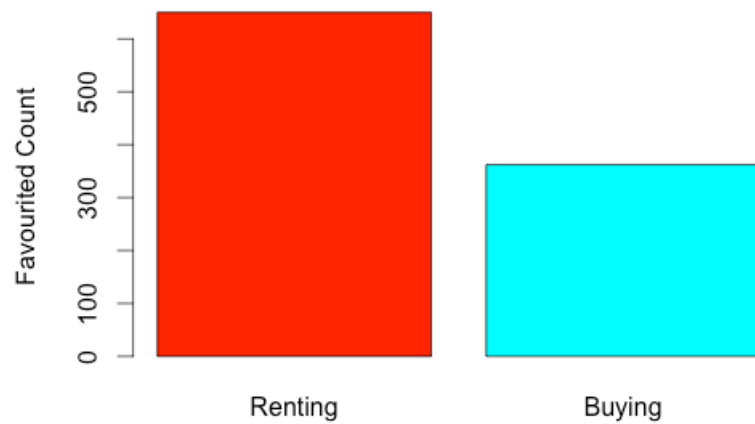
Analysis:

Twitter Data:

The bar plot below shows the popularity of keywords related to renting and buying houses. We can clearly see that more people are talking about renting compared to buying.
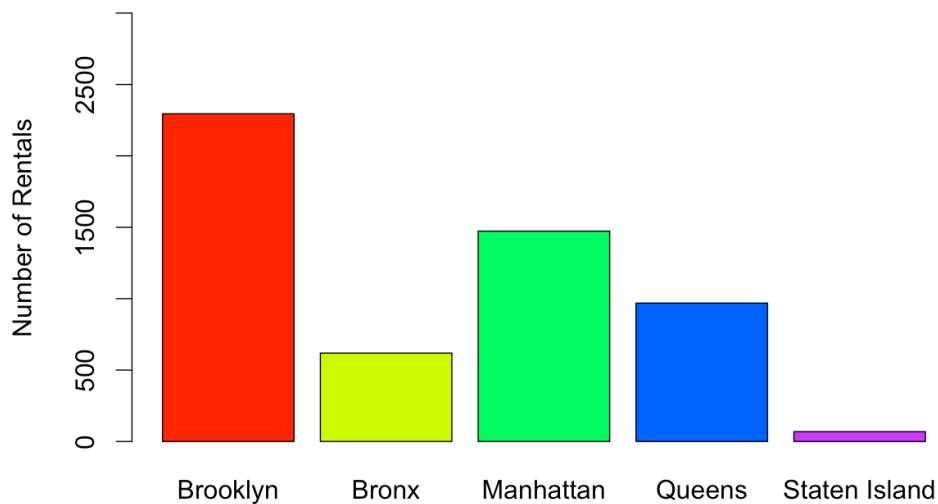

Similarly, it has also been observed that more people have favourited tweets related to renting compared to buying.

Rolling Sales Data:

The rolling sales data was classified by the keyword "RENTALS" to find the rental properties in the data set. In the below plot it can be seen that there are more rental properties in Brooklyn.

Conclusion:

It is obvious from the above data that more people are talking about apartment rents and renting than buying or selling properties. The Rolling Sales data shows that there are more properties available in the Brooklyn area for renting. Hence Real Direct should start offering apartment renting as a new product in this area to boost their profits.