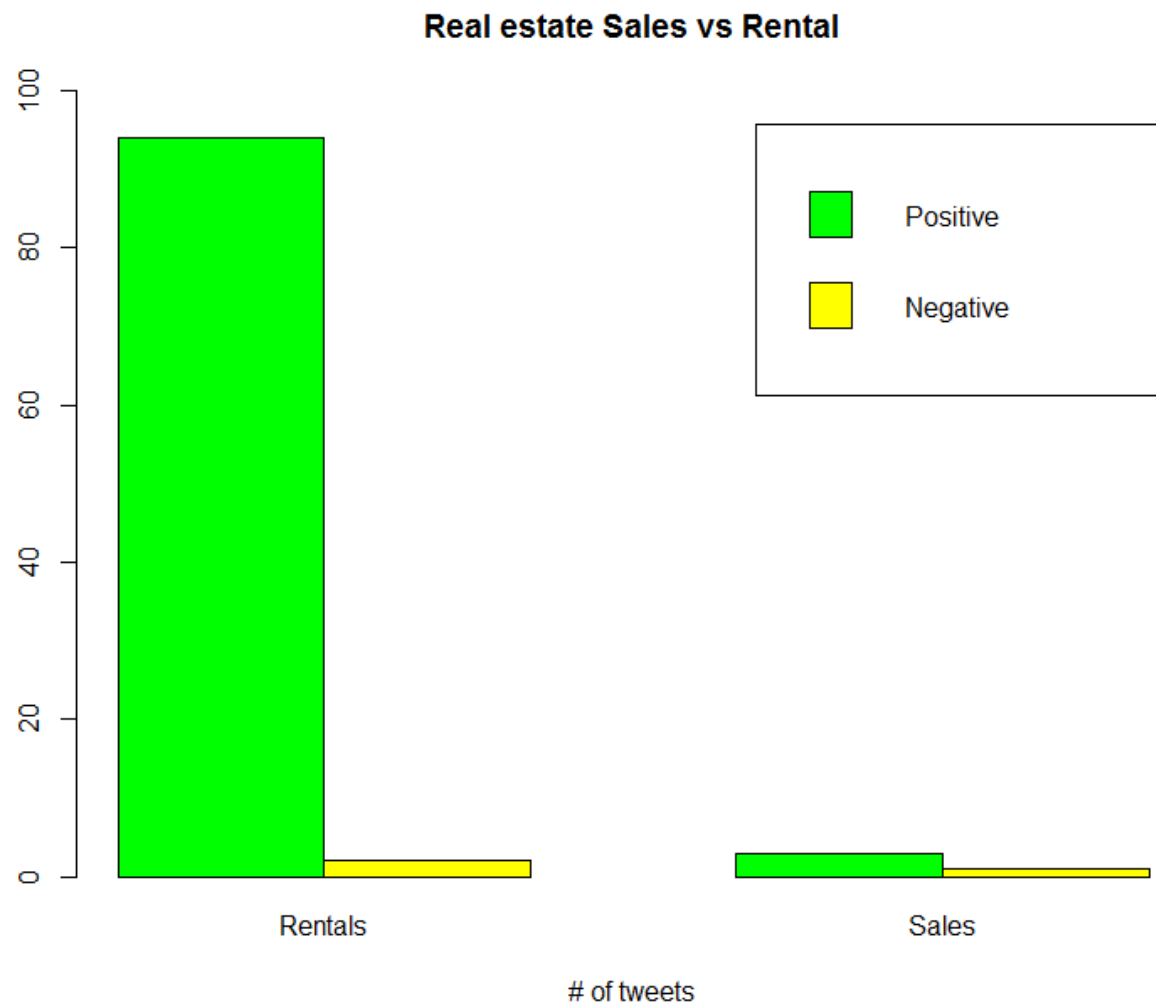


Problem 4: Data Product

Main Task: Combine different data sets and perform statistical analysis for business recommendation.

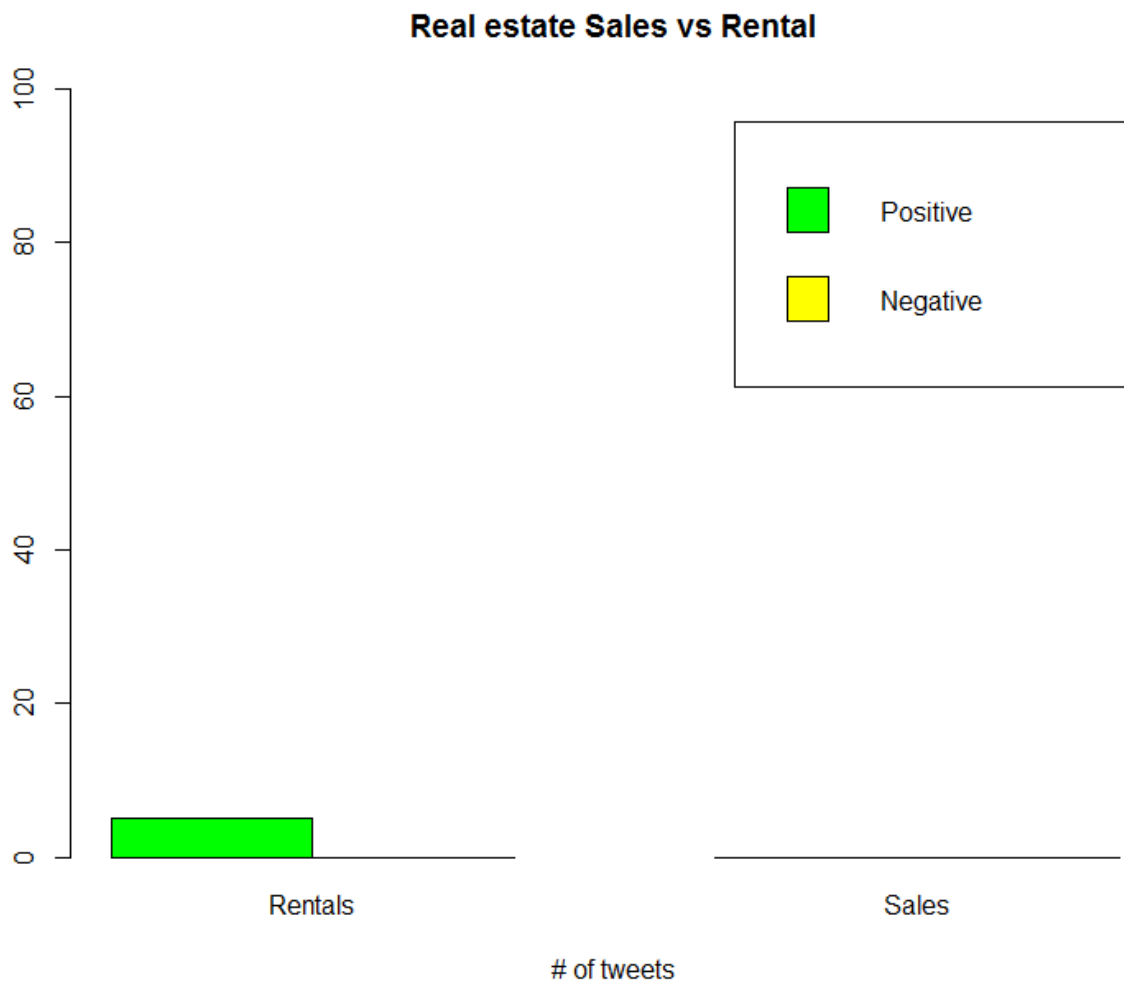
To obtain a proper analysis of what exactly is the opinion of people when given a choice amongst Property rental or Property sale, the following approach was followed. Firstly tweets were extracted. Then for each tweet, the respective number of positive and negative words were found. This was done by comparing the tweet text from the two files namely positive.txt and negative.txt which are saved in local disk. After finding the number of positive and negative words, a cumulative score of each tweet was calculated by subtracting number of Positive words with the Number of Negative words. Then plotting the score of each tweet.

Fetches tweets with the track “Sale apartment in New York” and “Rent apartment in New York”. Below is the plot obtained:



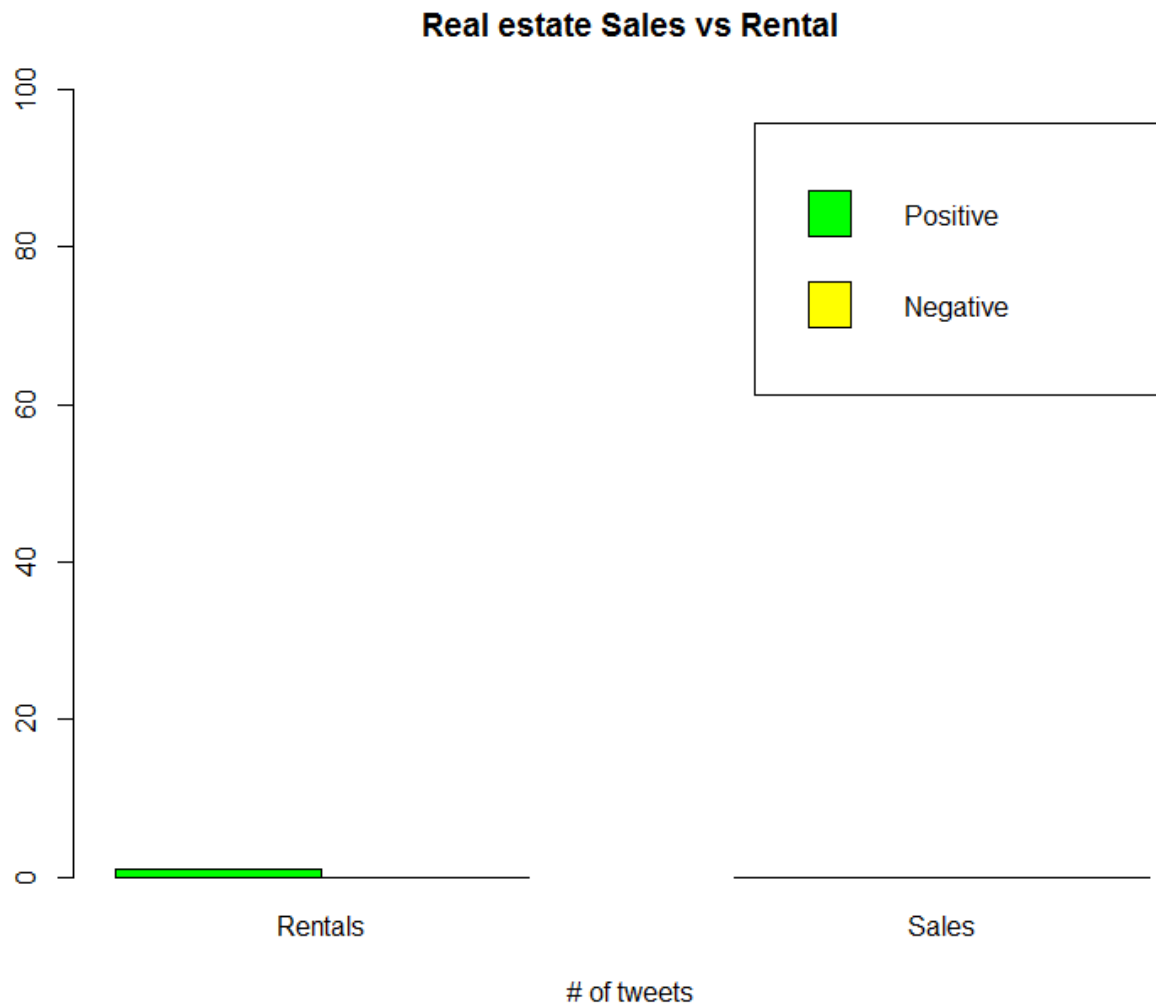
{The number of positive rental tweets are more than positive Sales tweets. Also the total score of Rental is a lot more than that of Sale}

Fetches tweets with the track “Selling property in New York” and “Rent property in New York”. Below is the plot obtained:



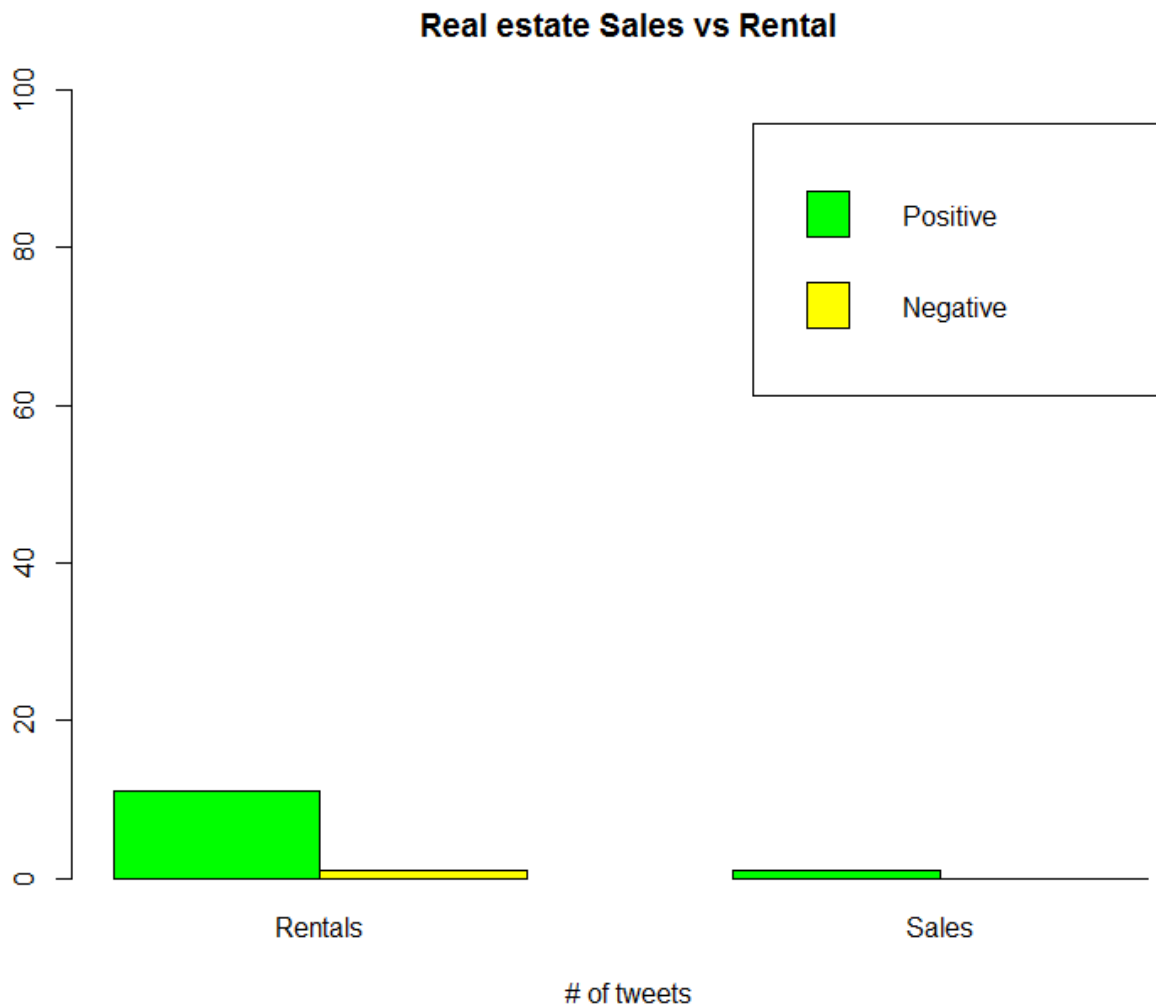
{The number of positive rental tweets are more than positive Sales tweets.}

Fetches tweets with the track “Sale an apartment in Manhattan” and “Rent an apartment in Manhattan”. Below is the plot obtained:



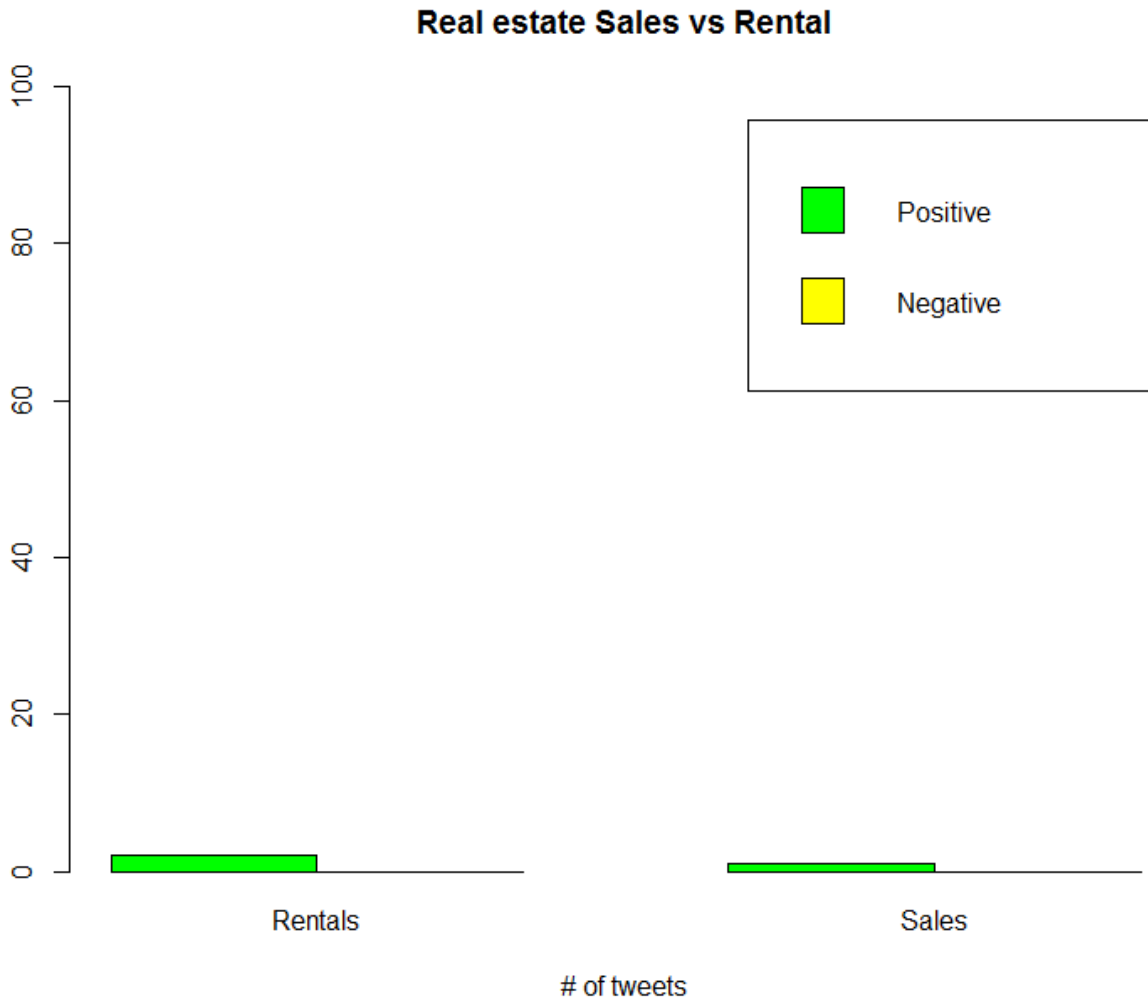
{The number of positive rental tweets are more than positive Sales tweets.}

Fetches tweets with the track “Sale an apartment in new York city” and “Rent an apartment in new York city”. Below is the plot obtained:



{The number of positive rental tweets are more than positive Sales tweets. Also the total score of Rental is a lot more than that of Sale}

Fetches tweets with the track “Sale property in new York city” and “Rent property in new York city”. Below is the plot obtained:



{The number of positive rental tweets are more than positive Sales tweets.}

Reference:

Made positive_words_list.txt and negative_words_list.txt by taking words from the below links.

<http://www.enchantedlearning.com/wordlist/positivewords.shtml>

<http://www.enchantedlearning.com/wordlist/negativewords.shtml>