CS506 Final Project: PlaceMe & Housing in Great Boston

Kunlin Cai, Yating Yang, Yihao Shi

Our Goal

we found four topics that connect with Boston housing. In order to relate the information that can help PlaceMe Living, we provide detailed data, analysis and recommendation associated with each topic.

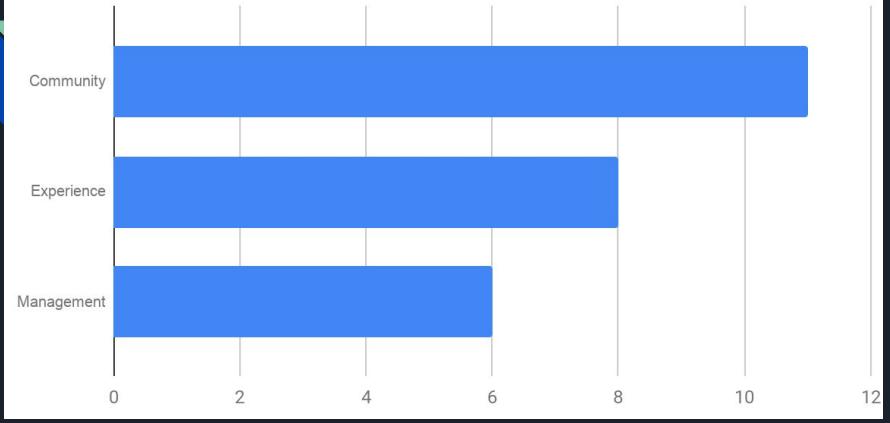
Part 1: Housing price V.S. Housing Market

1. Data from : bostonpads.com
2. Understand Data:
    - Real Time Availability Rate
    - Real Time Vacancy Rate
3. Correlation with Rent Price:
    - 1b1b price and Real Time Availability Rate: **0.1065869552**
    - 1b1b price and Real Time Vacancy Rate: **-0.4392974535**
4. Recommendation

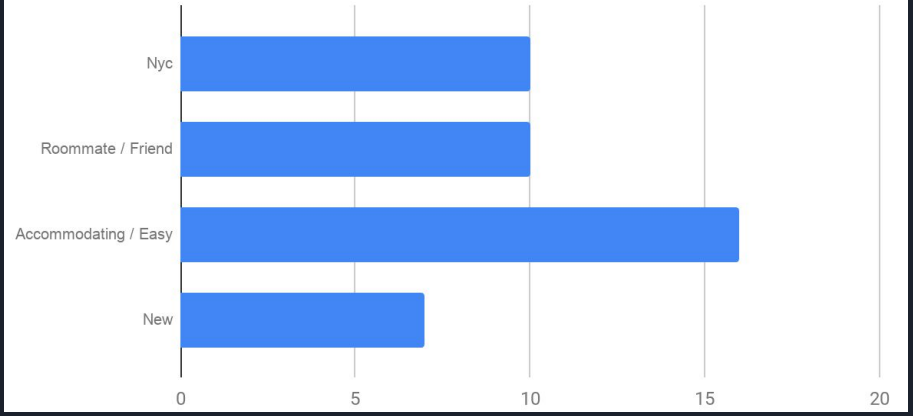
Part 2: Keywords Frequency Analysis to Reviews

1. Collect Reviews from Facebook, Google and Yelp
2. Transfer reviews to keywords and count frequencies
3. Analysis the keywords
    a. Positive reviews
        i. PlaceMe
        ii. Bedly
        iii. Bungalow
    b. Negative reviews
4. Recommendations

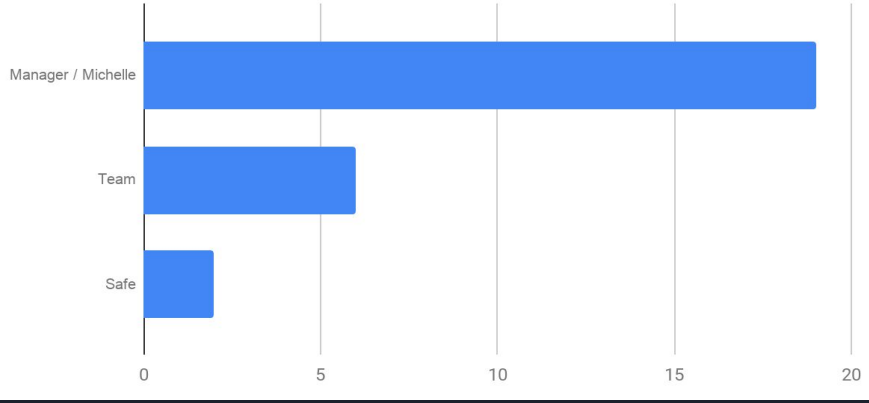
Commons positive review



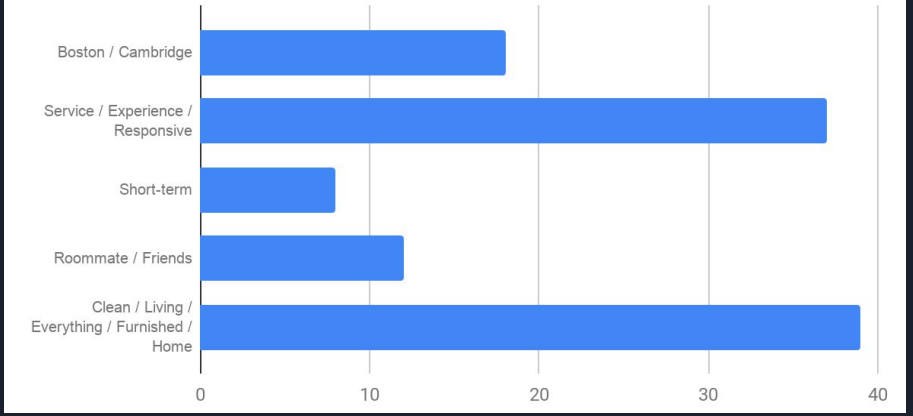
Bedly Review

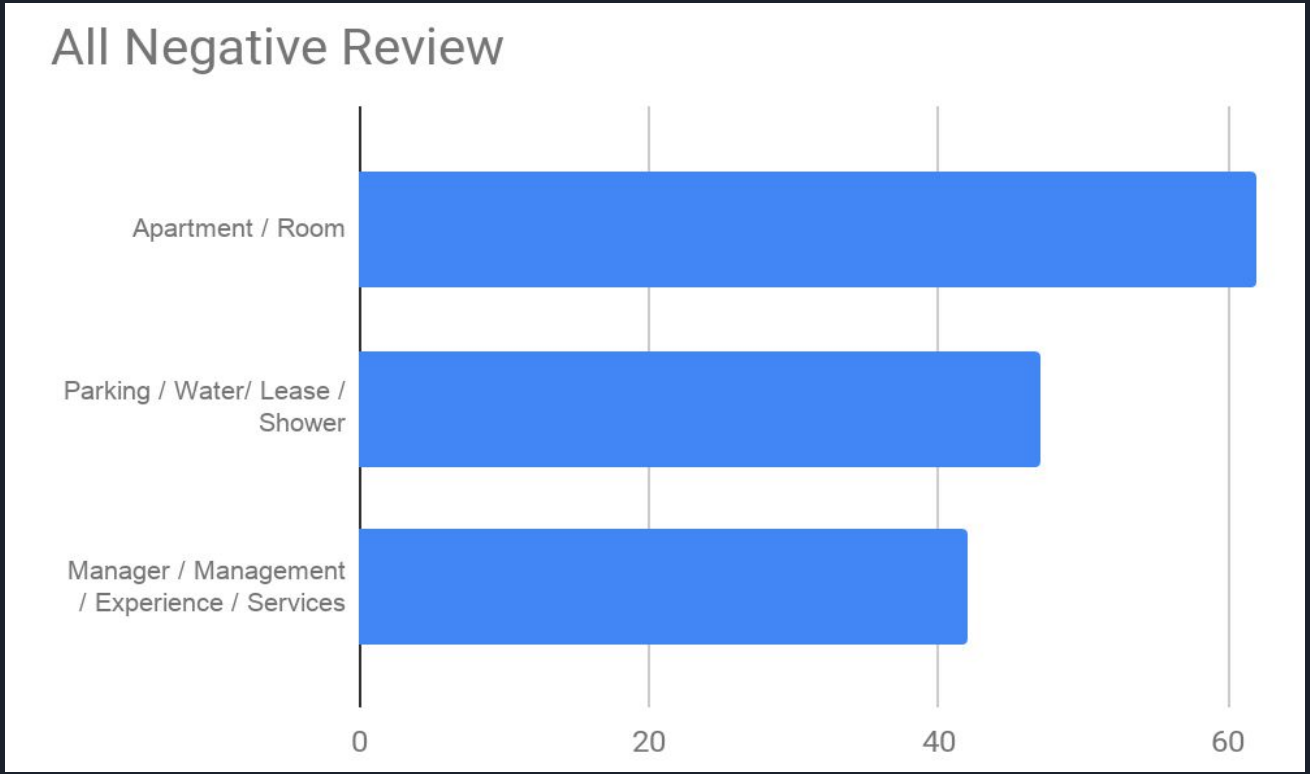


Bungalow Review



PlaceMe Review




All Negative Review
Apartment / Room
Parking / Water/ Lease / Shower
Manager / Management / Experience / Services
0
20
40
60

Part 3: Potential Customer and their expected prices.

1. Choosing data feature and Collect Data from ipums
2. Build models to predict the willingness to pay for rents of a house.
3. Best result by **Logistic Regression + SVM**
4. Recommendations

Results:

Methods:	**Accuracy:**
Logistic Regression	0.5960555149744339
Random Forest	0.5558802045288532
MLP((20,15))	0.5924032140248356
SVM (C = 1.5)	0.5989773557341125
LinearSVM(C = 3)	0.5953250547845143
KNN(K = 47)	0.5938641344046749
SVM + Logistic Regression	0.6194945848375452
SVM + Random Forest	0.5949458483754513
SVM + SVM	0.6180505415162455
SVM + LinearSVM	0.6151624548736462
LogisticRegression + LogisticRegression	0.6209386281588448
LogisticRegression + SVM	**0.6274368231046932**
LogisticRegression + Random Forest	0.5891696750902528
LogisticRegression + LinearSVM	0.6166064981949458

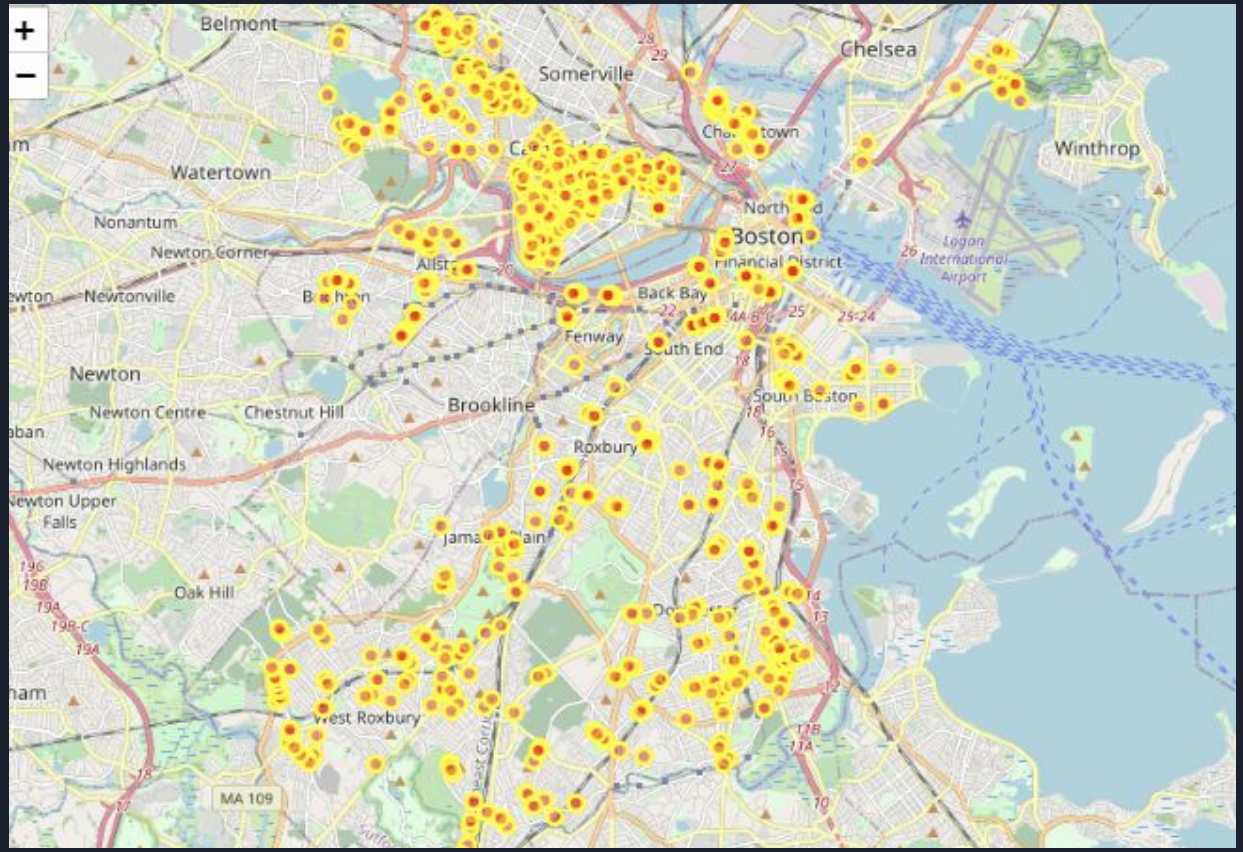
Part 4: Visualization of Short-Term-Rent Houses in the Boston Area

1. Data from boston.gov/cambridge.gov
2. Translate address to ['Latitude' 'Longitude'] (geocoder)

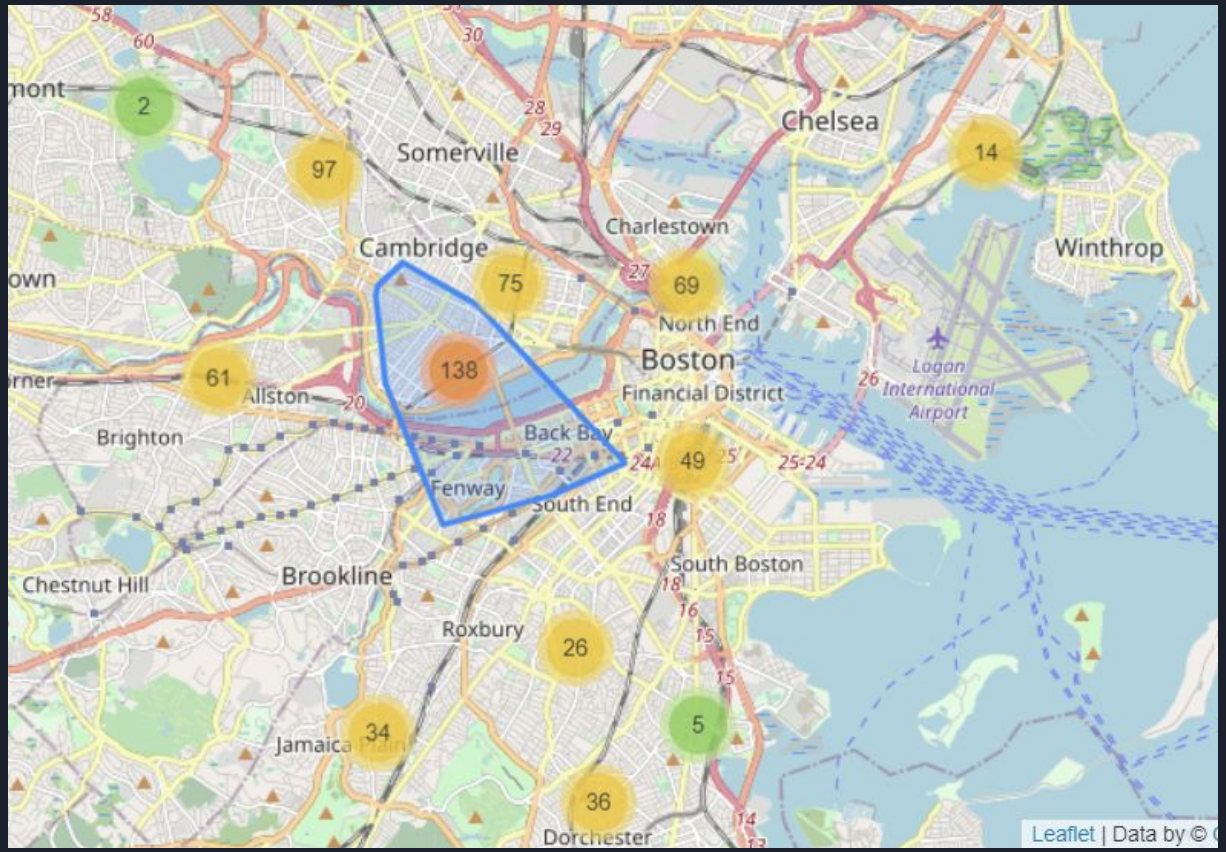
   -- Due to time limitation, we just use the first 500 data entries of the Boston dataset.

3. Draw MarkerCluster & CircleMarker on the map (Folium)

CircleMarker



MarkerCluster



Thanks!