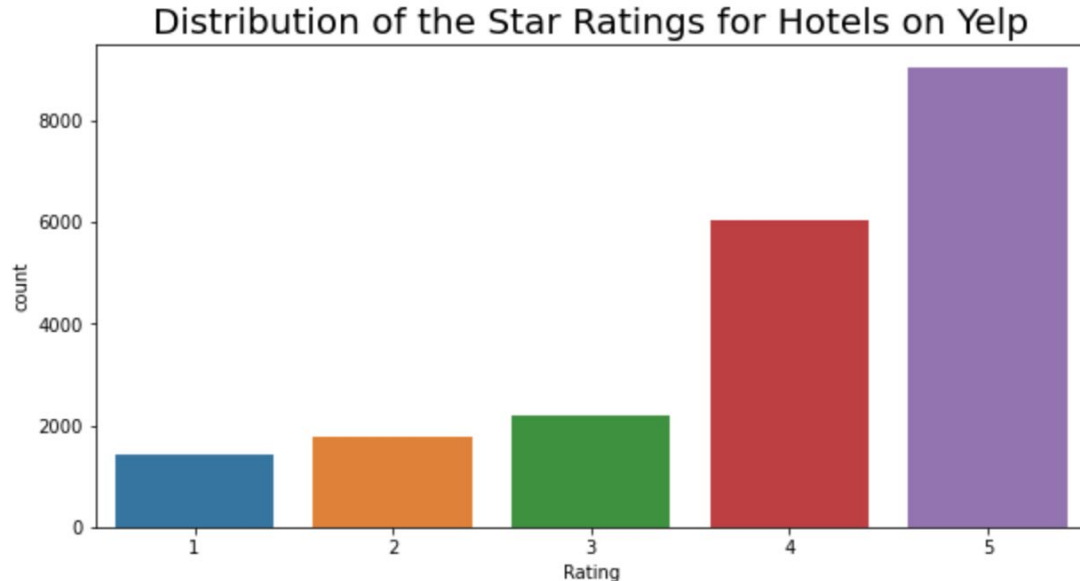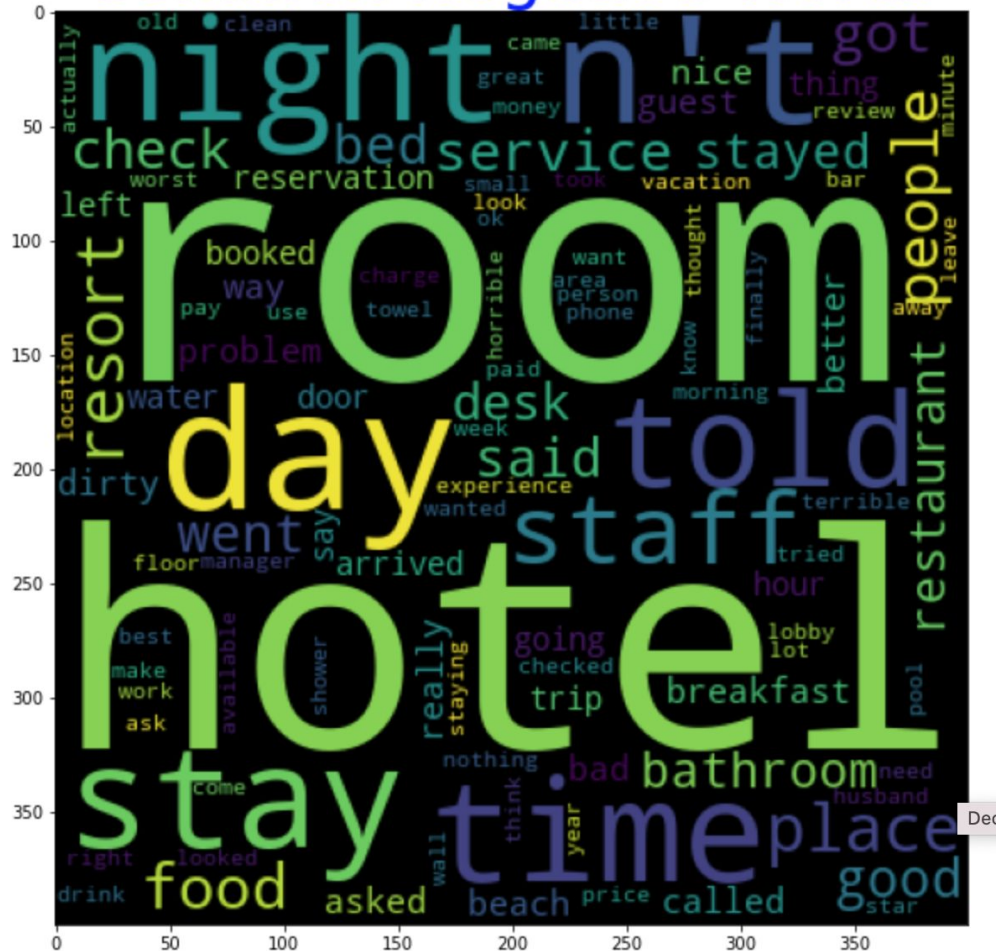# Hotel Satisfaction Reviews

Indicators of success

Who is Leaving Reviews?

Primarily satisfied customers.

Out of the 20,491 reviews we examined customers left 15,093 excellent reviews (4-5 stars). Which is 74% of our data.
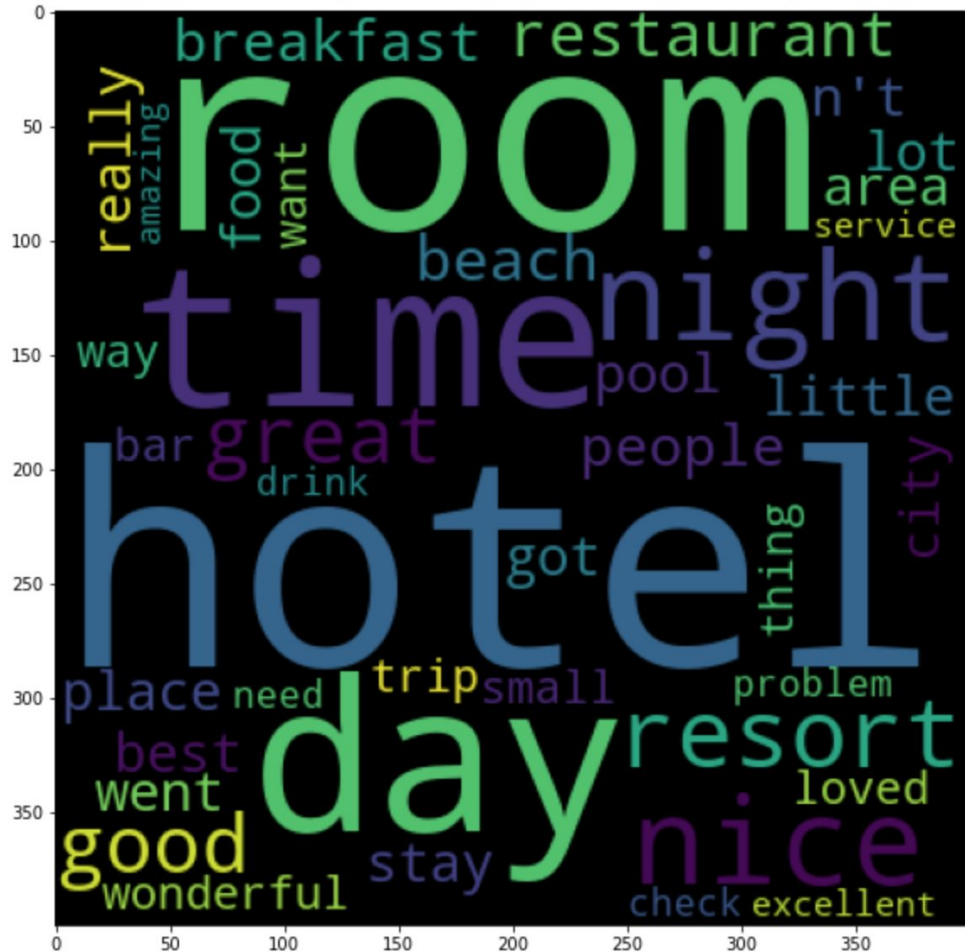

Distribution of the Star Ratings for Hotels on Yelp

Common Negative Words

Outside of common words that apply to all hotel reviews (room/hotel). We can see in our negative reviews they spoke frequently of service (835), place(649), and resort (633)
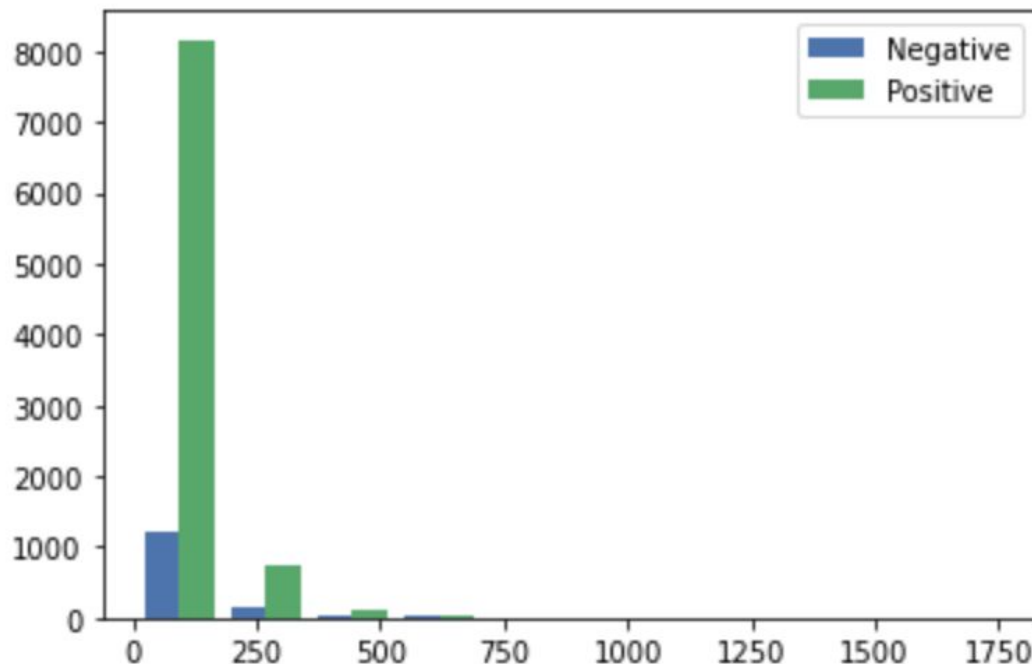
Common Positive Words

breakfast restaurant room n't lot area service beach night really amazing food want time way great bar people drink hotel little pool got thing city place need trip small problem best went good day resort loved wonderful stay nice check excellent

Outside of common words that apply to all hotel reviews (room/hotel). We can see in our positive reviews they spoke frequently of location (4,962), breakfast (4,370), service (4300), clean (3,935), and excellent (3792)

Distribution of Length of Words

When graphed together you can see they both skew towards more words. We had a higher quantity of positive reviews, but it appears the average amount of words is about the same. Positive reviews do have a slightly higher mean of 112 versus negative's 94.

Machine Learning

Based on a Naive Bayes Model with 98% accuracy we are able to examine which words held the most weight in determining the rating.

For our first model we wanted the most polarizing language so we just examined reviews with 5 star or 1 star reviews.

Machine Learning

We can see the common words for positive reviews were spotless, loved, delightful, brilliant, gem, pricey, eiffel, traditional, library, and wharf. This leads us to see that while cleanliness is one of the top predictors of if a customer will love their hotel the location is also very important.

We can see the common words for negative reviews were unhelpful, disgrace, reeked, unprofessional, rotten, stunk, disgusting, incompetent, urine, and vomiting. This leads us to see again cleanliness is one of the top predictors if a customer will love thier hotel. The second highest factor in a negative review is staff that is not friendly and accomodating.

2nd Model

We created a second model with all the reviews and our accuracy reduced to 58%.

The highest precision is for negative 1 star reviews. The highest recall is for 5 star reviews. While the highest f1 score is also for 5 star reviews.

Meaning 5 star reviews are not are very likely to be identified accurately when they are identified, but with the lower precision we know we are throwing in several reviews which do not belong there (which could be a lot of 4's since they would have similar language). While 1 star reviews have more extreme language so they're likely to be identified correctly, but with the low recall we know we are returning very few results.

In summary:

Customers are more likely to take time to write a glowing review.

Customers are most concerned with cleanliness.

They are secondarily concerned with location.