

What do users actually want?



The dataset is the reviews that Airbnb users left when they finished a reservation. The focus is to analyze the top 200 active listings and the reviews about them. I also subset the ratings that are lower than 85 (100 scale) and see the difference among them. Active listings means they have average 8+ reviews per months, and ratings are over 93.

Most people usually give higher ratings when doing online reviews, partially because they are afraid to backfire on themselves. However, most of them also express their opinions in the reviews. My research focuses on mining user-generated content (UGC) and hopefully gain insights from those reviews.

| Word | Count (approx.) |
|-------------|-----------------|
| place | 270 |
| great | 265 |
| clean | 260 |
| room | 255 |
| airport | 245 |
| close | 235 |
| rice | 225 |
| would | 215 |
| comfortable | 205 |
| host | 195 |
| recommend | 185 |
| night | 175 |
| everything | 165 |
| good | 155 |
| house | 145 |
| flight | 135 |
| definitely | 125 |
| friendly | 115 |
| perfect | 105 |
| home | 95 |
| helpful | 85 |
| easy | 75 |
| hosts | 65 |
| near | 55 |
| bed | 45 |
| gave | 35 |
| location | 25 |
| welcoming | 15 |
| needed | 10 |

[illegible][illegible]

| Word | Frequency |
|--------------------------------|-----------|
| would, 'definitely', 'stay' | 0.0006 |
| would, 'highly', 'recommend' | 0.0005 |
| definitely, 'come', 'back' | 0.0004 |
| highly, 'recommend', 'staying' | 0.0003 |
| another, 'great', 'stay' | 0.0002 |
| would, 'come', 'back' | 0.0002 |
| first, 'time', 'using' | 0.0002 |
| would, 'definitely', 'book' | 0.0002 |
| d, 'definitely', 'stay' | 0.0002 |
| visiting, 'new', 'york' | 0.0002 |
| would, 'absolutely', 'stay' | 0.0002 |
| definitely, 'would', 'stay' | 0.0002 |
| ca, "n't", 'beat' | 0.0002 |
| feel, 'like', 'home' | 0.0002 |
| great, 'little', 'place' | 0.0002 |
| ca, "n't", 'wait' | 0.0002 |
| bus, 'stop', 'right' | 0.0002 |
| could, "n't", 'ask' | 0.0002 |
| feels, 'like', 'home' | 0.0002 |
| last, 'minute', 'booking' | 0.0002 |
| short, 'walk', 'away' | 0.0002 |
| good, 'night', 'sleep' | 0.0002 |
| would, 'definitely', 'return' | 0.0002 |
| made, 'sure', 'everything' | 0.0002 |
| would, 'totally', 'recommend' | 0.0002 |
| best, 'host', 'ever' | 0.0002 |

- Have polarized comments, such as, "would definitely" and "certainly wouldn't"