Hotel Analytics Report

29 January 2026

Anna Holle

5-Star or Flop? What Drives Hotel Reviews on the Las Vegas Strip

Data on hotel reviews can give valuable insights into the successes and failures of hotels. Data can also give an insight into the mind of the customer. What does a customer value the most, and what do they value the least? Can hotels do anything to influence the rating received in a customer review, or are some types of customers predestined to leave bad reviews no matter what?

The Data

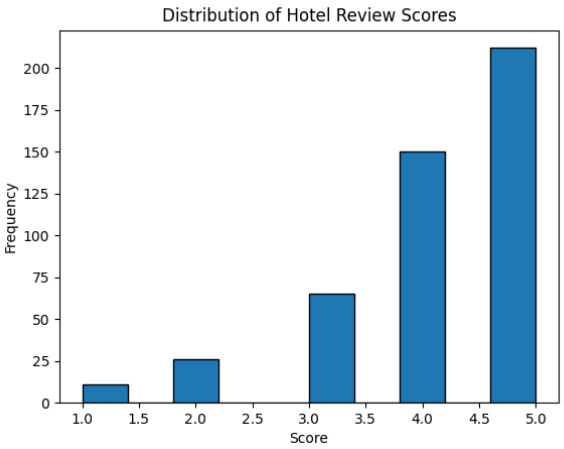
I was given a dataset of Las Vegas TripAdvisor Reviews. This data set contained information on reviews collected on the travel platform TripAdvisor. It contains information about the users that leave reviews, such as their home country, the number of reviews they've left on the platform, and how many years they have been a member of TripAdvisor. It also contains information about the hotel being reviewed and some general information about the trip. For example, the dataset tracks the amenities available in each hotel, its star rating, the number of rooms available, and the score given by the reviewer. In addition, information about the traveler type, time period of the stay, and the time the review was posted are also recorded in the dataset.

First, I cleaned the data. The dataset originally had a column for status that was mostly missing. This column was probably intended to record whether or not the reviewer has a loyalty status at the hotel, but due to the 526 missing values in this column, I removed it from the analysis. Additionally, I removed an unnamed column that was also mostly empty.

The second aspect of data cleaning focused on ensuring the User Country information was accurate. I noticed that some locations were misclassified as countries (Hawaii), and there were some minor entry errors that had to be corrected (such as an extra space or inputting the nationality instead of the country name). I also had to check the data types and change some variables from categorical variables to floats in order to make an analysis possible at all.

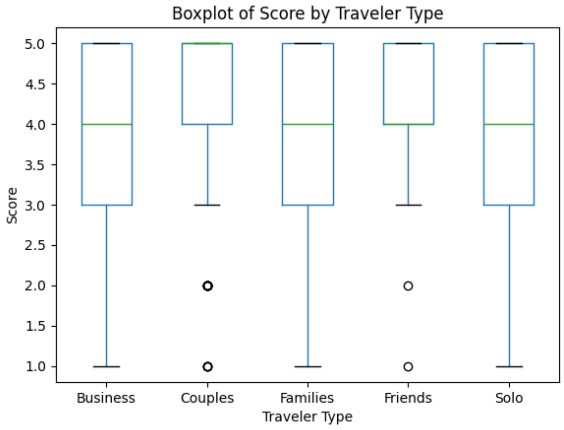
After cleaning, the dataset had 20 columns and 464 entries. It contains information on 21 unique hotels on the Las Vegas Strip. The users of the platform are from 43 unique countries, and some are avid users of TripAdvisor with many reviews posted, while others have only posted the hotel review.

Key Findings

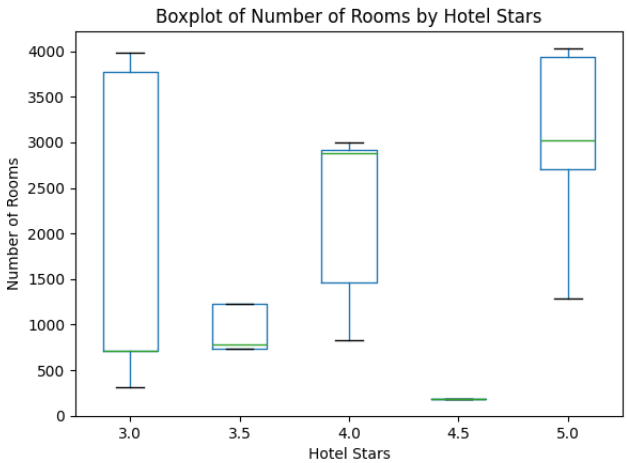


Some exploratory analysis shows that overall, the reviews in this dataset are quite positive. The hotels are scored on a scale from 1 to 5, with 5 being the best score. The mean score was 4.134 out of 5. The dataset contains relatively few low scores. This surprised me because I assumed that a person who had a bad experience at a hotel would be more likely to review it than someone who had a good experience.

One of my key findings is that most traveler types score hotels similarly. As shown by the boxplot to the right, the median score for all traveler types, except for couples, was 4. The median score for couples was 5. Overall, couples and friends had a tighter and higher range of scores, perhaps because of the nature of the trip, with some extreme values located at 1 and 2 stars.



Another interesting finding is the relationship between the number of rooms in a hotel and the Hotel stars. Hotels are typically rated by stars to indicate the level of amenities or experience that a guest can expect. Interestingly, a trend can be seen in that hotels with a lower star rating generally have fewer rooms, with an exception for 4.5 star hotels, which have the median least amount of rooms in the data set. 5 star hotels have the highest median number of rooms.



I also created a subset of “high-quality” hotels. These hotels all had 4 or 5 stars, an average score greater than or equal to 4, free internet, a gym, and a pool. Out of the 21 hotels in the dataset, 12 of the hotels meet this criteria. This may help explain the overwhelmingly positive reviews in the dataset.

Model Performance Summary

For this analysis, I did not have to create any models. This was an exploratory analysis that focused on looking at data and visualizing it rather than making a model.

Conclusion

In conclusion, the analysis suggests that hotel reviews on the Las Vegas Strip are primarily driven by overall hotel quality rather than by traveler type or trip characteristics. Reviews in the dataset were overwhelmingly positive, with an average score above 4 out of, indicating that guests generally report satisfactory experiences with their hotel stays in Las Vegas. This positivity may be partly explained by the fact that many of the hotels in the dataset already meet high standards in terms of star ratings and amenities. Since these hotels are located on the Strip rather than downtown or on the outskirts, they occupy prime real estate and aim to deliver the best experience for their guests.

While different traveler types were examined, there was little variation in review scores across groups, suggesting that couples, solo travelers, families, and friends tend to evaluate hotels in similar ways. Couples did rate hotels slightly higher on average, but the overall consistency across traveler types indicates that the perceived quality of the hotel experience outweighs individual travel context.

Overall, the findings suggest that amenities and baseline hotel standards are key drivers of positive reviews, rather than who the traveler is or when they travel. On the Las Vegas Strip, where competition is high and expectations are elevated, hotels that consistently deliver high-quality experiences are more likely to receive strong reviews, making a 5-star stay far more common than a flop.