

HOTEL ANALYTICS REPORT: LAS VEGAS STRIP

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

I. Introduction and Data Overview:

1. Project Goal:

The goal of this project is to analyze reviews of hotels in Las Vegas by applying tools such as Pandas, NumPy, and Matplotlib to process, explore, and visualize the data. Through this, it aims to discover the behavior of customers, the performance of the hotel, and also what key factors affect the review score. The results of the analysis are expected to provide useful information and insights that can support hotel managers and tourism industry stakeholders in making strategic decisions to improve service quality and optimize business operations.

2. Data Overview and Missing Value handling:

The original LasVegasTripAdvisorReviews dataset consisted of 539 observations and 22 variables, reflecting travelers' hotel reviews. The main variables in the dataset include Score, Hotel stars, Traveler type, Member years, Helpful votes, Amenities (Pool, Gym, Casino, Free Internet), etc.

During the data inspection and cleaning, the two columns "Unnamed:15" and "Status" were removed due to the high percentage of missing data (539 and 526) and no analytical significance. Besides, rows that lack essential fields, such as Hotel Name, Score, Traveler Type, User's Continent, have been removed because these variables are essential for core business questions.

Other missing values (Amenities, Number of Membership Years, Length of Stay) are deliberately retained. Removing them will likely remove a large number of valid reviews and may skew the results in later analyses. Since these fields don't really affect every business question, retaining them helps preserve the integrity of the dataset and maximize the amount of usable information. After this processing, the dataset is left with 491 observations and 20 variables.

3. Data cleaning summary:

The data cleaning process is carried out in a step-by-step manner to ensure that the dataset is ready for further analysis. First, columns that are irrelevant or contain too many missing values are removed to avoid confounding the information. Next rows that lack important variables are processed selectively. Some banana distortions but of an arithmetic nature, such as Hotel stars, are converted to numerical form for filtering and calculation. Finally, the data is re-checked to make sure there are no formatting errors or abnormal values left. Thanks to the elimination process, the dataset becomes more consistent, clear, and relevant for further analysis steps.

II. Key Findings:

1. Hotel performance insights:

a. Top 5 hotels with the Highest Average Score:

Based on all review data, the results received show that the leading hotels have an average result of over 4.4, reflecting a very high level of satisfaction from travelers. Wynn Las Vegas came out on top with 4.625, followed by The Venetian Las Vegas Hotel (4.5833) and Marriott's Grand Chateau (4.5652). The other two positions belong to Encore at Wynn Las Vegas (4.5417) and Wyndham Grand Desert (4.4348). These hotels

are all in the group of high-end accommodation facilities, often providing full facilities and quality services, contributing to a positive experience for customers. This result shows that major brands and luxury hotels in Las Vegas tend to maintain a stable quality of service and receive positive reviews from visitors.

b. Top 10 based on Reviews from EU users:

Besides, when looking at reviews from European travelers separately, the results show that the highest-rated hotels are still focused on high-end hotels, with outstanding service and convenient locations. Specifically, Wynn Las Vegas continues to lead with an average score of 4.625, followed by The Venetian Las Vegas Hotel (4.5833), etc. In the last two positions in the top 10 are Bellagio Las Vegas and The Cosmopolitan Las Vegas, with scores ranging from 4.18 to 4.38. This result shows that the top hotels in Las Vegas not only attract domestic tourists but also make a good impression on international visitors, especially in the European market. This is an important basis for managers to come up with strategies to promote and improve services that are more in line with the expectations of international customers.

c. Bottom 5 hotels among Full - Amenity hotels:

Despite being fully equipped with high-end amenities, some hotels still receive lower average ratings than the general premises. The lowest five locations included Treasure Island – TI Hotel & Casino, Tropicana Las Vegas – A Double Tree by Hilton Hotel, Tuscany Las Vegas Suites & Casino, and several other hotels with an average score below 4.2. These results show that despite providing adequate amenities, it is not enough to ensure a high level of satisfaction. Other factors such as service quality, actual experience, or customer expectations may play other important roles. This shows that comfort is only a necessary condition, but customer service is a sufficient condition for a high score.

d. Review volume and Country contribution for Top hotels:

With the results received, the hotels with the highest reviews were all 24, reflecting the high popularity and wide interest from users. For each of the top hotels, the three countries that contributed the most reviews were the USA, UK, and Canada with some hotels having an additional presence from Australia, India, Saudi Arabia or Thailand. This result shows that hotels in Las Vegas not only attract large numbers from Europe but also from other countries, especially the North American region.

e. Review distribution and Engagement by Continent:

An analysis of the number of reviews by continent showed that North America was the largest contributor with 287 reviews, accounting for a large proportion of the entire data set. The next regions are Europe (115), Oceania (39), and Asia (36), while Africa and South America recorded only seven reviews each.

In terms of satisfaction and engagement, South America has the highest average score (4.43), followed by North America (4.16) and Oceania (4.15). In terms of the number of helpful votes, North America leads with an average of 33.65, followed by South America (33.43) and Asia (32.69). This result shows that visitors from North America not only make up a large proportion of the number but also tend to leave detailed and helpful reviews. At the same time, high levels of satisfaction from regions such as South America and Oceania also hint at the potential for further development in markets beyond North America and Europe.

f. Review Behavior for hotels without Free Internet:

In the group of hotels that do not offer free internet, the three countries with the highest number of reviews are the USA (6), the UK (3) and Canada (2). This shows that travelers from major markets still choose to stay at some hotels that lack basic amenities, possibly due to location, price or brand factors.

In terms of engagement, Monte Carlo Resort & Casino is the standout hotel in this category with an average Helpful votes score of 39.91, which is significantly higher than the overall average. This result shows that, despite the lack of Internet elements, some hotels are still capable of creating worthwhile experiences that are considered useful and worth sharing by users.

g. Top 5 hotels by Room Capacity and Service Quality:

Based on the results received, the five hotels with a large number of rooms and meeting quality criteria such as 4 stars or higher, free internet, gyms and swimming pools are all among the high-end hotels in Las Vegas. The Venetian Las Vegas Hotel leads the way with over 4,000 rooms and an average review score of 4.58, followed by Bellagio Las Vegas (3,933 rooms, 4.21), Caesars Palace (3,348 rooms, 4.13), The Palazzo Resort Hotel Casino (3,025 rooms, 4.38), and The Cosmopolitan Las Vegas (2,959 rooms, 4.18). All of the above hotels are 5-star and have a high number of reviews. This result shows that large-scale hotels still maintain their service quality and attractiveness, while meeting the utility needs of tourists well.

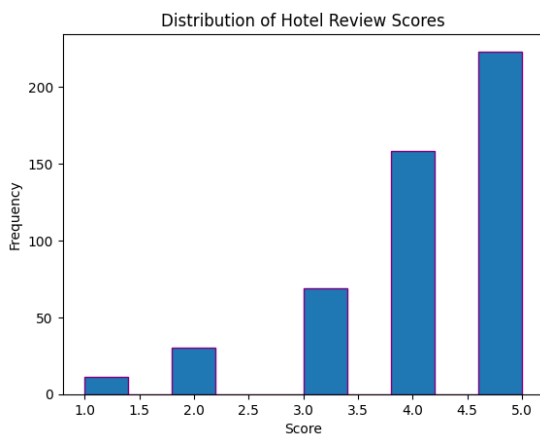
h. Temporal pattern in Review Activity:

According to the results of the analysis, the three months – January, February and June – had the highest number of reviews, with about 42 reviews per month. For weekdays, travelers tend to contribute their reviews on Wednesdays, Tuesdays, and Sundays with 82, 75, and 79. In addition, the most popular length of stay is from Dec - Feb, indicating that the trend of travel to Las Vegas for the majority of travelers is in the winter and early part of the year.

In terms of user behavior, the most common group of tourists is couples, showing the city's characteristic as an ideal destination for couples. These trends show that reviews focus on peak travel times, while revealing demographics and user experience-sharing behaviors.

2. Visualization:

a. Distribution of hotel Review Scores:

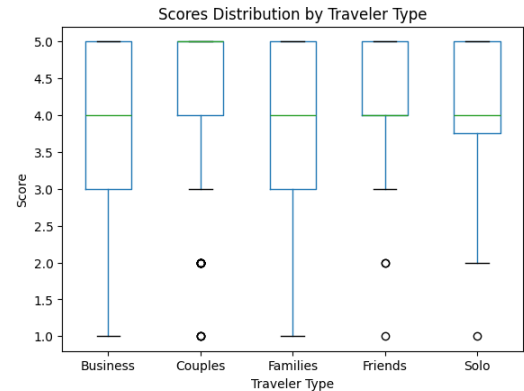


The review score frequency chart shows that the majority of travelers rate the hotel as high. Scores of 4.0 and 5.0 account for the largest proportion, about 160 and 220 reviews. Meanwhile, scores of 1.0 and 2.0 are low frequency, suggesting that the level of dissatisfaction is rare. This distribution reflects the trend of positive user reviews and shows that hotels in Las Vegas generally respond well to traveler expectations.

b. Scores distribution by Traveler Type:

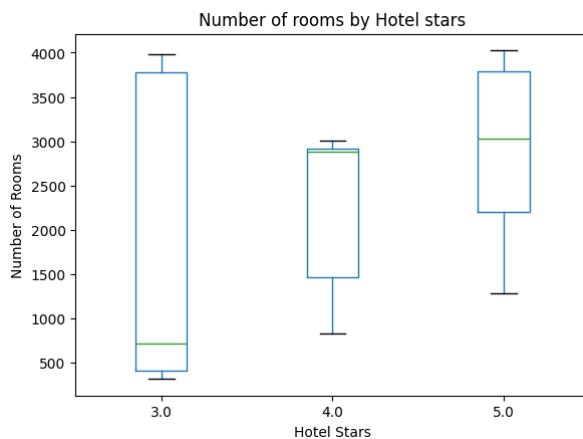
The boxplot chart shows that the score distribution is generally high, with the median ranging from 4.0 to 5.0, but there are marked differences between customer groups. The Couples group showed the highest level of satisfaction when at least half of the customers gave it an absolute 5-star rating, while the Solo

group had the highest consistency in the experience. In contrast, the Business and Families groups have the widest range of dispersion scores, ranging from 1.0 to 5.0, showing that the quality of service for these two subjects is not really stable. In addition, the appearance of outliers at levels 1.0 and 2.0 in the remaining groups warns of single negative experiences that need to be focused on improving to optimize brand reputation. It can be seen that, although the Histogram shows a general positive trend, Boxplot has clearly dissected the lack of stability in the experience of Business and Families customers.



c. Number of Room by hotel Stars:

This chart shows a marked difference in the scale of the number of rooms between star classes, with 5-star



hotels tending to possess the largest and most stable number of rooms with a median of about 3000 rooms. Notably, the group of 3-star hotels. For this group, there is an extremely large differentiation, although the average is below the low level (less than 1000 rooms), but the box range extends up to 4000 rooms, showing the existence of both small residences and large-scale budget hotel complexes. Meanwhile, in 4-star hotels, there is a deviation to the upside, with the median convenience close to the limit of about 3000 rooms, showing that the majority of hotels in this segment

often maintain a large scale to optimize service capacity.

III. Conclusion:

Through analysis, it can be seen that Las Vegas is still a destination with very high service quality, especially attracting customers who are couples from the North American market. However, to optimize revenue and brand reputation, managers need to focus more on improving the experience for Business and Family, two groups with highly fluctuating satisfaction levels. The standardization of meeting room services, home amenities and specialized support elements will help mitigate this instability.

Besides, maintaining the advantage of scale is also a point worth considering. 5-star hotels are doing a great job of combining large scale and quality, but large-scale 3-star hotels should also be considered to avoid lagging behind smaller competitors in the same segment. Finally, managers should know how to grasp and take advantage of the international effect. With a large amount of interest from Europe and other regions, promoting international promotion is a potential strategy and direction.