# Hotel Review Analysis Report

Business Objective:  
The purpose of this report is to conduct a comprehensive text analysis of hotel customer reviews to identify top customer concerns, highlight key strengths, and recommend areas for improvement. By leveraging sentiment analysis and topic modelling, the insights gained will help prioritize service enhancements to improve customer satisfaction and overall experience.

1. Data Preprocessing and Cleaning

The raw hotel review data underwent several preprocessing steps to ensure consistency and quality of the analysis. The following steps were performed:

* Removal of special characters, numbers, and punctuation.
* Conversion of text to lowercase.
* Removal of stopwords.
* Lemmatization to convert words into their base forms.

## 2. Tokenization

The text data was tokenized into individual words to facilitate further analysis. This process helped in breaking down the reviews into smaller components, making it easier to identify patterns and trends.

## 3. Exploratory Data Analysis (EDA)

Initial exploratory data analysis revealed that the majority of the reviews carried positive sentiment. The most common words in the dataset were fairly general, such as "good," "stay," and "service," which highlighted the need for more refined topic modelling.

## 4. Sentiment Analysis

Sentiment analysis was conducted to classify reviews into positive, negative, and neutral categories. Additionally, a new column was added to the dataset to calculate review length and investigate whether review length played a role in sentiment distribution. The analysis indicated that longer reviews tended to express more detailed experiences, often associated with either very positive or very negative sentiments.

## 5. Topic Modelling with LDA

Latent Dirichlet Allocation (LDA) was used to uncover hidden topics within the reviews. The initial model with 5 topics produced overly general themes. To enhance the granularity of insights, the number of topics was increased to 10. This adjustment allowed for better identification of customer concerns and key strengths.

## 6. Visual Analysis

Visualization of topic distributions using PyLDAvis provided an interactive way to explore the relationship between topics and keywords. This helped to better understand the dominant themes within the reviews and how they were associated with sentiment categories.

## 7. Segregated Sentiment-Based Topic Modelling

To further improve the analysis, the dataset was segmented by sentiment categories before performing topic modelling. This approach yielded more targeted insights into the concerns expressed in negative reviews and the strengths highlighted in positive reviews.

## 8. Custom Stopwords Removal

To avoid skewed results, custom stopwords such as "hotel," "room," and other frequently occurring but non-informative words were removed. This step enhanced the clarity of the extracted topics.

## 9. Identifying the key strengths in Topics

### 1. Key Strengths Identified from Positive Reviews

Topic #1: Friendly and Helpful Staff  
Top Words: staff, great, clean, friendly, helpful, historic  
Insight: Customers consistently praised the **friendly and helpful attitude of the staff**, especially in historic and high-traffic areas like the French Quarter.

Topic #3: Scenic Views and Pool Facilities  
Top Words: great, view, nice, pool, location, perfect  
Insight: Guests enjoyed **beautiful views and clean pool facilities**, indicating these amenities positively impact customer satisfaction.

Topic #7: Breakfast Quality and Free Parking  
Top Words: good, great, breakfast, parking, restaurant, free  
Insight: Complimentary **breakfast and free parking** were appreciated by customers, especially in budget-friendly hotels.

Topic #9: Prime Location and Restaurant Options  
Top Words: great, location, restaurant, beach, enjoyed

Insight: Hotels located near **beaches, restaurants, and city attractions** received positive feedback, especially in destinations like San Francisco and Waikiki.

### 2. Top Concerns Identified from Negative Reviews

Topic #4: Dirty Rooms and Smell Issues  
Top Words: dirty, smell, bed, floor, water, old  
Insight: The most common complaint was **dirty rooms** and **unpleasant smells**, particularly in older hotels.

Topic #6: Shower and Luggage Issues  
Top Words: shower, luggage, fix, microwave, dirty  
Insight: Guests frequently reported **broken showers and dirty amenities** like microwaves, which negatively impacted their stay.

Topic #8: Poor Front Desk Service  
Top Words: desk, bad, bathroom, booked, bed  
Insight: **Poor front desk service** and issues with **room bookings** were recurring concerns across various hotels.

Topic #10: Gym Facilities and Cleanliness  
Top Words: gym, cleanliness, food, decor, size  
Insight: The **cleanliness of gym facilities** and **food quality** were identified as areas needing improvement.

## 10. Recommendations

Based on the analysis, the following recommendations are proposed to enhance customer satisfaction:

* **Improve Cleanliness Standards:** Implement stricter cleaning protocols for rooms, bathrooms, and gym facilities.
* **Upgrade Maintenance Services:** Address frequent issues with showers, microwaves, and other room amenities promptly.
* **Enhance Front Desk Training:** Provide customer service training to improve front desk efficiency and communication.
* **Expand Breakfast Options:** Upgrade breakfast selections to offer more variety and better quality.
* **Leverage Strengths in Marketing:** Highlight friendly staff, prime locations, and free amenities in promotional materials.

## 11. Conclusion

This report identifies both positive and negative aspects of customer experiences in hotels. By focusing on cleanliness, customer service, and maintenance, hotels can significantly improve guest satisfaction. Leveraging existing strengths in marketing campaigns will further enhance the hotel's reputation and customer loyalty.