



# PROPOSAL OF EUROPE HOTEL REVIEWS

from Booking.com

## Abstract

If you want to visit Spain, Austria, Italy, France, the Netherlands or the United Kingdom, whether in a leisure or business stay, and if you are alone or with friends, we will help you discover the best through more than 500 K reviews of hotels in Europe through the Booking.com.

**Elsayed & Nour**

e.ahmed.gm@gmail.com  
Nourshosharah@gmail.com

## INDEX

1 .....	<u>INTRODUCTION:</u>
2 .....	<u>DATA CONTENT: -</u>
3 .....	<u>BUSINESS QUESTIONS: -</u>
4 .....	<u>BUSINESS SOLUTIONS: -</u>
7 .....	<u>PLAN FOR MODELING AND DEPLOYMENT: -</u>

## Introduction

At the first, we thank Dr. Doaa for this challenge, because it always makes us work to learn and benefit, and we thank the group and all participants in the challenge

As for the project, the data was taken from Kaggle via this link

<https://www.kaggle.com/datasets/jiashenliu/515k-hotel-reviews-data-in-europe>

it's a large group of guest reviews in Western Europe hotel, more than 500,000 reviews for 1493 luxury hotels, as for the guests we have more than 200 nationalities, and about 50% of the reviews or data are for hotels in the United Kingdom, and for each review in the data we have a positive review and negative review.

We will explore the data, clean it up, work on the analysis, and then choose the appropriate prediction model.

## Data Content:

The csv file contains 17 fields and 515,000 records. The description of each field is as below:

1. Hotel\_Address: Address of hotel.
2. Review\_Date: Date when reviewer posted the corresponding review.
3. Average\_Score: Average Score of the hotel, calculated based on the latest comment in the last year.
4. Hotel\_Name: Name of Hotel
5. Reviewer\_Nationality: Nationality of Reviewer
6. Negative\_Review: Negative Review the reviewer gave to the hotel. If the reviewer does not give the negative review, then it should be: 'No Negative'
7. ReviewTotalNegativeWordCounts: Total number of words in the negative review.
8. Positive\_Review: Positive Review the reviewer gave to the hotel. If the reviewer does not give the negative review, then it should be: 'No Positive'
9. ReviewTotalPositiveWordCounts: Total number of words in the positive review.
10. Reviewer\_Score: Score the reviewer has given to the hotel, based on his/her experience
11. TotalNumberofReviewsReviewerHasGiven: Number of Reviews the reviewers has given in the past.
12. TotalNumberof\_Reviews: Total number of valid reviews the hotel has.
13. Tags: Tags reviewer gave the hotel.
14. dayssincereview: Duration between the review date and scrape date.
15. AdditionalNumberof\_Scoring: There are also some guests who just made a - scoring on the service rather than a review. This number indicates how many valid scores without review in there.
16. lat: Latitude of the hotel
17. lng: longitude of the hotel

### **We generated new features based on our data such as:-**

- 18- regions: Divide the nationalities of the reviewers into continents and regions.
- 19- rooms\_type: Room types such as a double room, a single bed, or a special room
- 20- guests: kind of guests, single, couple, or with their family, and so on
- 21- trip\_type: leisure trip or a business trip
- 22- stayed: number of nights of stay
- 23- submit: review submit the phone or something else.
- 24- With a pet: Does he have pets or not
- 25- extra\_room\_type: Special requests with the room.
- 26- country: hotel country
- 27- tourist: ourist or is he from the same country

## Business Questions:

- 1- what is the best Hotel in each country for staying?
- 2- What are the best destination and best hotels for Arab guests, and what were their reviews?
- 3- What about business trips, the average days of stay and the most reserved rooms, and do they have special requests during booking, and what are the most negative and positive reviews?
- 4- What about the type of guests, whether single or couple, what are the most important points that make them happy with the experience, and also what are the most important requests of those who have pets?
- 5- Are the number of guest review words indicated on the score, and what are the most frequently used words in positive and negative reviews, and the most sentences or most effective?
- 6- What is the ratio of tourists to non-tourists in the data?
- 7- Is there a difference between the behavior of those who wrote a review for the first time and others?

## Business Solutions:

### what is the best Hotel in each country for staying?

The best 3 hotels in each country were selected based on the guests' reviews and according to the number of guests for the same hotel in the data. We did not rely on the average score because it is better to rely on the guests in the data who reviewed.

### What are the best destination and best hotels for Arab guests, and what were their reviews?

- About the Arab guests, we have 20 different nationalities, their number reaches more than 36 thousand assessments, most of them are from the Arab Gulf region, and in order, the Emirates, then Saudi Arabia, then Kuwait and then Qatar,
- As for the countries, the most destinations were the United Kingdom, then France and Spain.
- Average number of nights stay 3 nights.
- The rating is from 75. to 95. Emiratis are the most numerous in all countries except France, the Saudis are more.
- The most comment negative comments of Arab guests are: pillows, blinds, bathrooms, and robbed
- The most positive comments from Arab guests are: location, cleanliness, hotel staff, parks.

What about business trips, the average days of stay and the most reserved rooms, and do they have special requests during booking, and what are the most negative and positive reviews?

- We have more than 80,000 guests for business
- Duration of stay from one night to 3 nights, the most frequent one is one night.
- There is no preference for room types, but most are Double Room.
- Extra requests do not prefer to have a window in the room and prefer a balcony with a view.
- Most of the negative comments talked about noise and workers, and some comments related to the car, such as the parking lot or the garage.
- Positive comments were about admiring the style of the hotel, the design of the building, the room decor, and the breakfast.
- About 30% of business guest data from the same country as the hotel and most of them are from the UK.

What about the type of guests, whether single or couple, what are the most important points that make them happy with the experience, and what are the most important requests of those who have pets?

- With pets they were happy because of the breakfast, the building, room service and the gardens and spa
- With pets they were unhappy because of the carpets, the parking, over price, cleanliness
- Most of the guests are couples or with family, and couples are 50% of the data.
- The couples were not satisfied with the breakfast, tea, and coffee services, and many of them talked about the Wi-Fi service and the rooms, and some mentioned that the pictures of the hotel were not the truth.
- But the couples were satisfied with the location of the hotel, the cheap price, and the location outside the hotel, and romantic.
- As for solo guests, the negative comments were more than positive. The negatives were noise, bad smell, uncleanliness, and parking.
- As for the positive, the location was breakfast and the room.
- There are no special requests for with pets.

Are the number of guest review words indicated on the score, and what are the most frequently used words in positive and negative reviews, and the most sentences or most effective?

There's no strong relationship between the number of words and the reviewer score of references because many reviews only written no negatives or no positives, Therefore, we couldn't find a relationship and we couldn't delete the small reviews because they're important and express the guest's satisfaction or dissatisfaction.

What is the ratio of tourists to non-tourists in the data?

Through the tourist column, the data filtered, whereby a person who has a nationality and wrote a review of a hotel in the same country is not a tourist, and if otherwise, he is a tourist, in the data We have 35% are non-tourists and 65% are tourists.

Is there a difference between the behavior of those who wrote a review for the first time and others?

- We made a data frame for those who were just a reviewer for the first time because they have slightly different feelings than those who have experience in assessment, and we will monitor their behavior through reviews.
- First-time reviewers had Various comments as they were dissatisfied with the parking and garage and had a problem with cleanliness and mess, and some reviewers were complaining about the air conditioning.
- As for the positive reviews, first-reviewers liked the hotel's surroundings, historical appearance, parks, concerts, and hospitality services such as drinks and breakfast.

## Plan for Modeling and Deployment:

### Modelling

- Regression model :based on reviews and user properties predict score(10-0)
- Classification model: Based on same features of regression model Predict satisfaction of reviewer (Satisfied - neutral - non Satisfied)

### Recommendation system:

- Recommend hotel to a user based on user inputs such as type of trip – Leisure or Business, type of travel – Solo or family, type of room, number of days of stay.
- Recommend hotel based on Nationality provided by user .

### Deployment:

- Create web app to for recommendation system



# Thank you

**Elsayed Aly & Nour Shosharah**

**Email**

**e.ahmed.gm@gmail.com**

**Nourshosharah@gmail.com**