

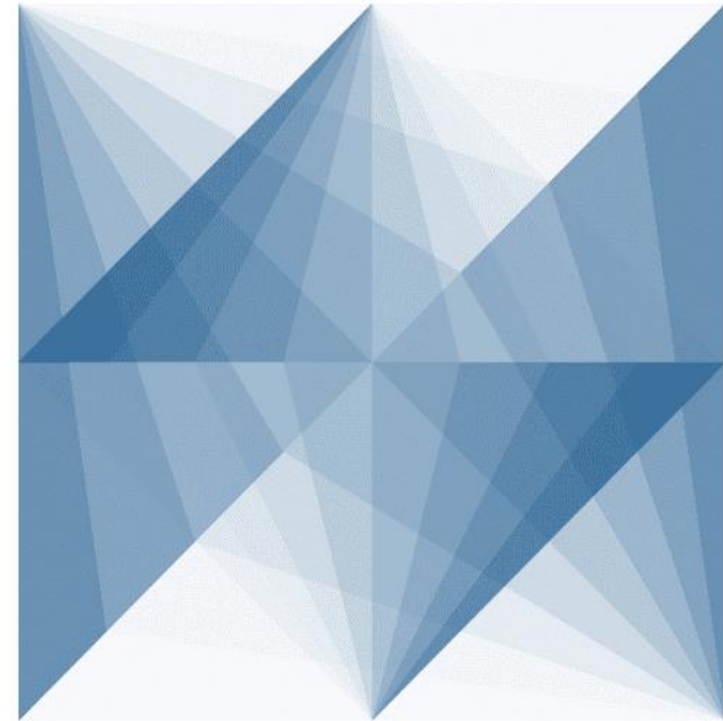
Hotel Reviews Project

By

Elsayed Aly & Nour Shosharah

Supervisor

Doaa Mahmoud Abdel-aty



Data science challenge

Reveal data secrets

Index:-

- BIO
- INTRODUCTION
- DATA CONTENT
- EDA
- BUSINESS QUESTIONS
- Business Solutions
- Sentiment analysis
- MODELING
- RECOMMENDATION SYSTEM
- DEPLOYMENT



Bio:-

- **Elsayed Ahmed Elsayed**
- **Marketing specialist**
- **NGO Organization**
- **Live in Kuwait**
- **LinkedIn: [Elsayed Aly](#)**



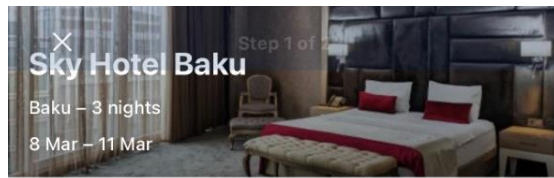
- **Nour Shosharah**
- **Bachelor's Degree in Computer engineering**
- **Natural Language Processing Engineer**
- **Live in Syria**
- **LinkedIn: [NOUR CHOCHARAH](#)**



Have you ever booked hotel online?

Talk about your experience if you did

How to write a review on Booking.com?



Rate this property:

How was your stay?

1 2 3 4 5 6 7 8 9 10

Bad Exceptional

How were the facilities?

1 2 3 4

Was it comfortable?

1 2 3 4

How was the location?

1 2 3 4

Was it clean?

1 2 3 4

Was it good value?

1 2 3 4

How was the staff?

1 2 3 4

Step 1 of 2

How was the location?

1 2 3 4

Was it clean?

1 2 3 4

Was it good value?

1 2 3 4

How was the staff?

1 2 3 4

Tell us about your trip

Purpose of your trip?

Leisure

Traveler type?

Solo traveler

Continue

Finish

Submitting this review confirms acceptance of our [Terms and Conditions](#). Also check out our [Review Guidelines](#).



Could you tell us a little more?

What did you like about the property?

What didn't you like about the property?

Write a short sentence to sum up your stay.

Share photos from your stay

Room Bathroom Lobby Outside

Finish

Submitting this review confirms acceptance of our [Terms and Conditions](#).

Introduction:-

We have over 500,000 reviews for 1,492 luxury hotels in Europe, and the data we have is a mix of hotels and reviewers' data.

For every review or every record in the data, we have in the hotel data (address, name, average score, and number of people who have done it).

The reviewer data (nationality, negative review, positive review, review date, after that tag column, contains some information related to the references like the number of nights, type of trip, etc.).

Objective:-

Work in two parts

- Hotel: Find the best hotel in Europe for tourists.
- Reviewers: They analyze his reviews and discover the most reasons that spoil his experience or make him happy.

And in the end, to recommend what is best for the guests and the hotel.

Data Content:-

```

RangeIndex: 515738 entries, 0 to 515737
Data columns (total 17 columns):
 #   Column                                     Non-Null Count  Dtype
---  -
 0   Hotel_Address                             515738 non-null object
 1   Additional_Number_of_Scoring              515738 non-null int64
 2   Review_Date                              515738 non-null object
 3   Average_Score                             515738 non-null float64
 4   Hotel_Name                               515738 non-null object
 5   Reviewer_Nationality                     515738 non-null object
 6   Negative_Review                           515738 non-null object
 7   Review_Total_Negative_Word_Counts        515738 non-null int64
 8   Total_Number_of_Reviews                  515738 non-null int64
 9   Positive_Review                           515738 non-null object
10   Review_Total_Positive_Word_Counts        515738 non-null int64
11   Total_Number_of_Reviews_Reviewer_Has_Given 515738 non-null int64
12   Reviewer_Score                           515738 non-null float64
13   Tags                                      515738 non-null object
14   days_since_review                         515738 non-null object
15   lat                                        512470 non-null float64
16   lng                                        512470 non-null float64

```

After Preprocessing

```

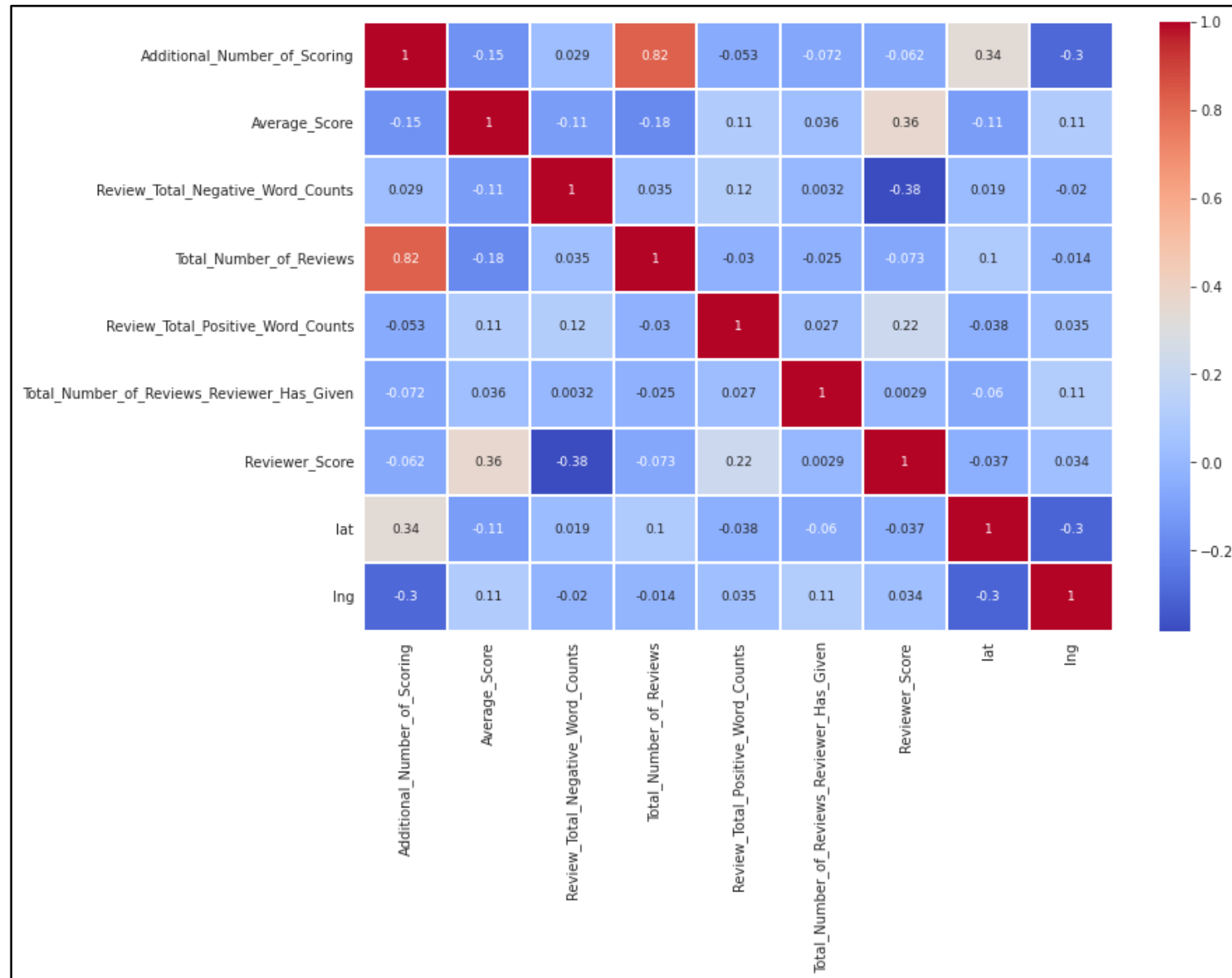
17   regions                                515212 non-null object
18   rooms_type                             515212 non-null object
19   guests                                 515212 non-null object
20   trip_type                              515212 non-null object
21   stayed                                 515212 non-null object
22   submit                                 515212 non-null object
23   With a pet                             515212 non-null object
24   extra_room_type                         515212 non-null object
25   country                                515212 non-null object
26   tourist                                515212 non-null int64
27   year                                    515212 non-null int64
28   month                                   515212 non-null int64
29   day                                    515212 non-null int64

```

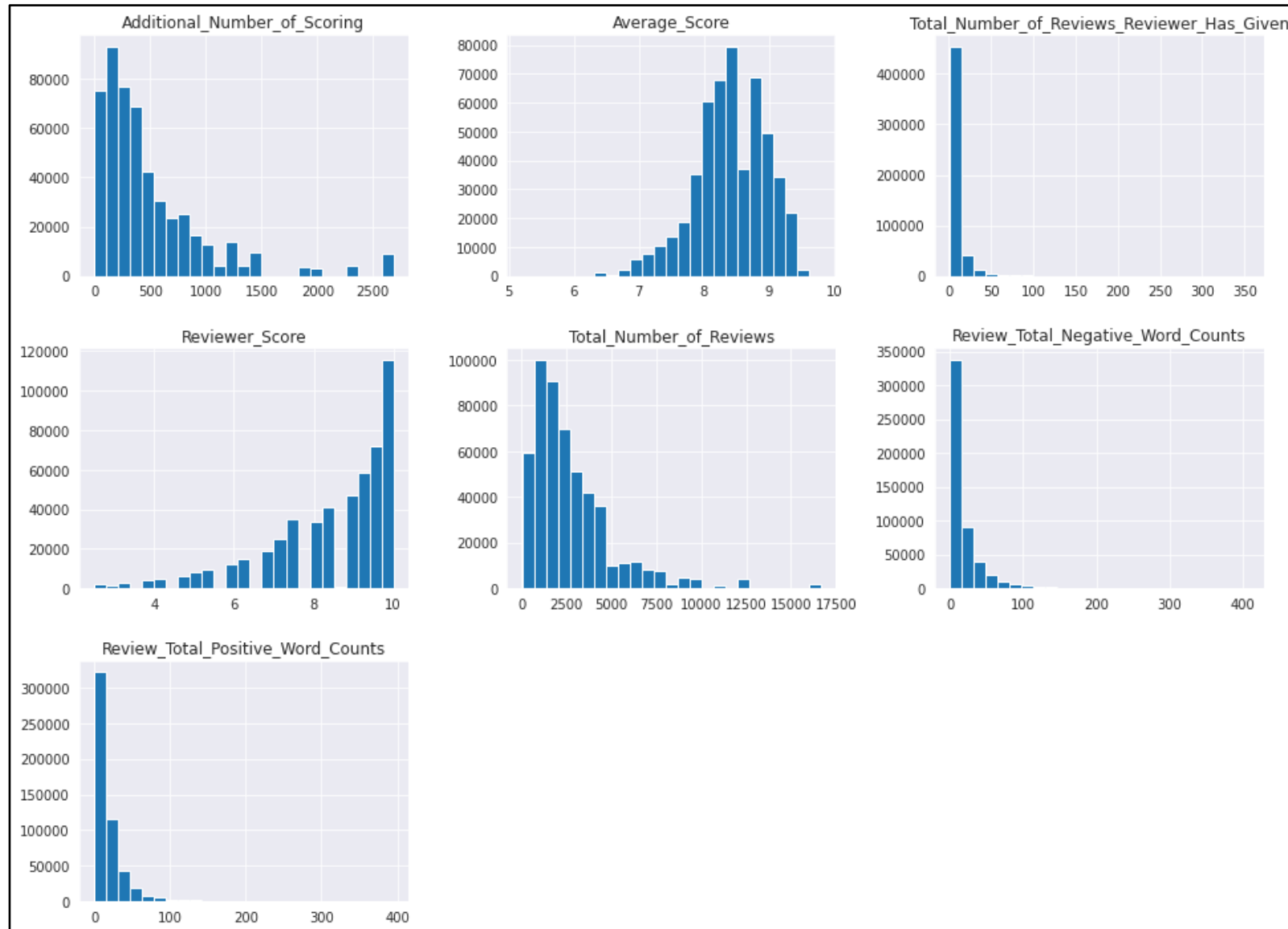

Exploratory data analysis



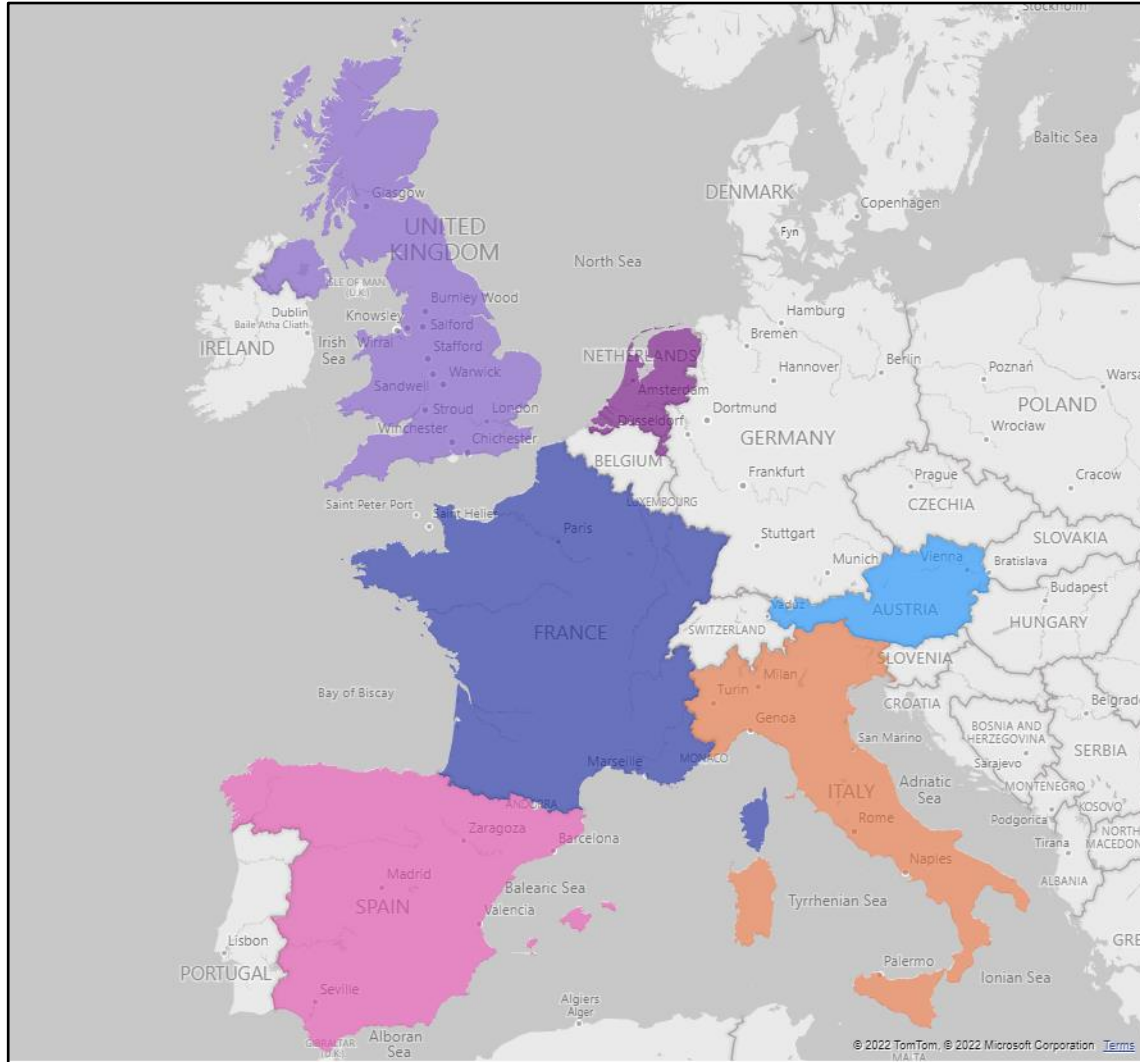
Data Exploration:-



Data Exploration:-



Data Exploration:-

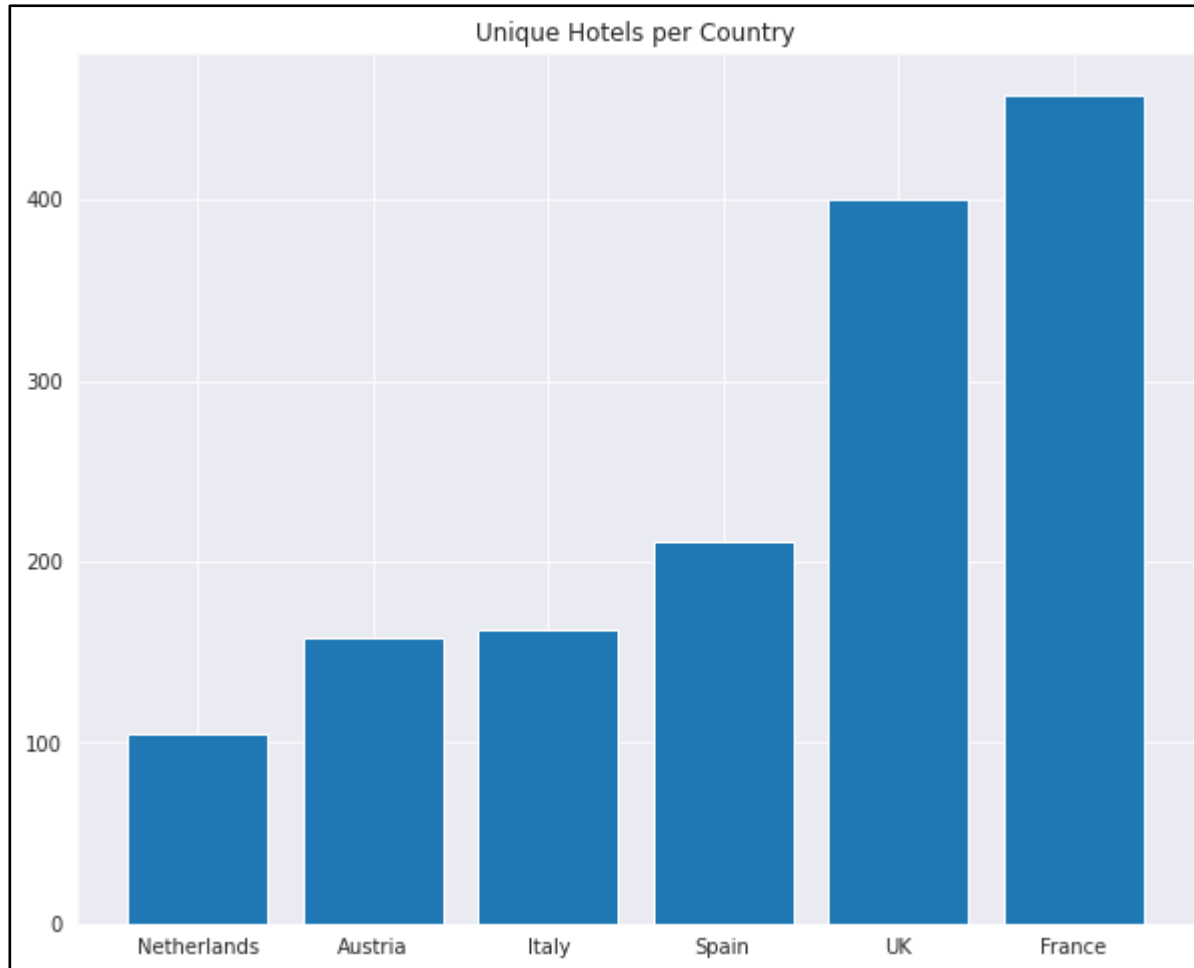


Countries with the number of reviews



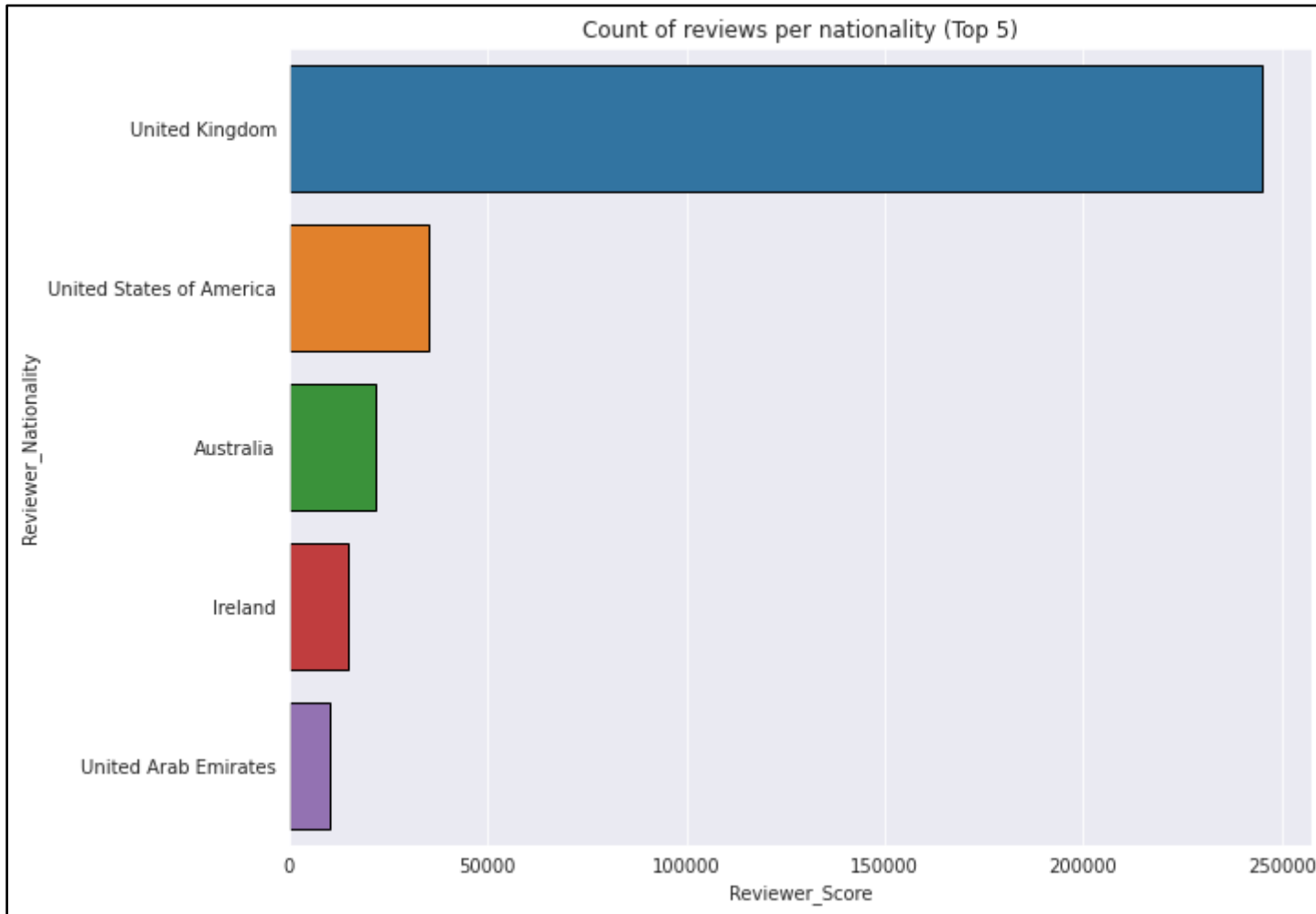
UK	262298
Spain	60149
France	59413
Netherlands	57211
Austria	38937
Italy	37204

Data Exploration:-



Number of unique hotels by country
It is noted that France has the most hotels, but the United Kingdom has the most reviews

Data Exploration:-



The most 5 nationalities

0	UK	245110
1	USA	35349
2	Australia	21648
3	Ireland	14814
4	UAE	10229

Business Questions

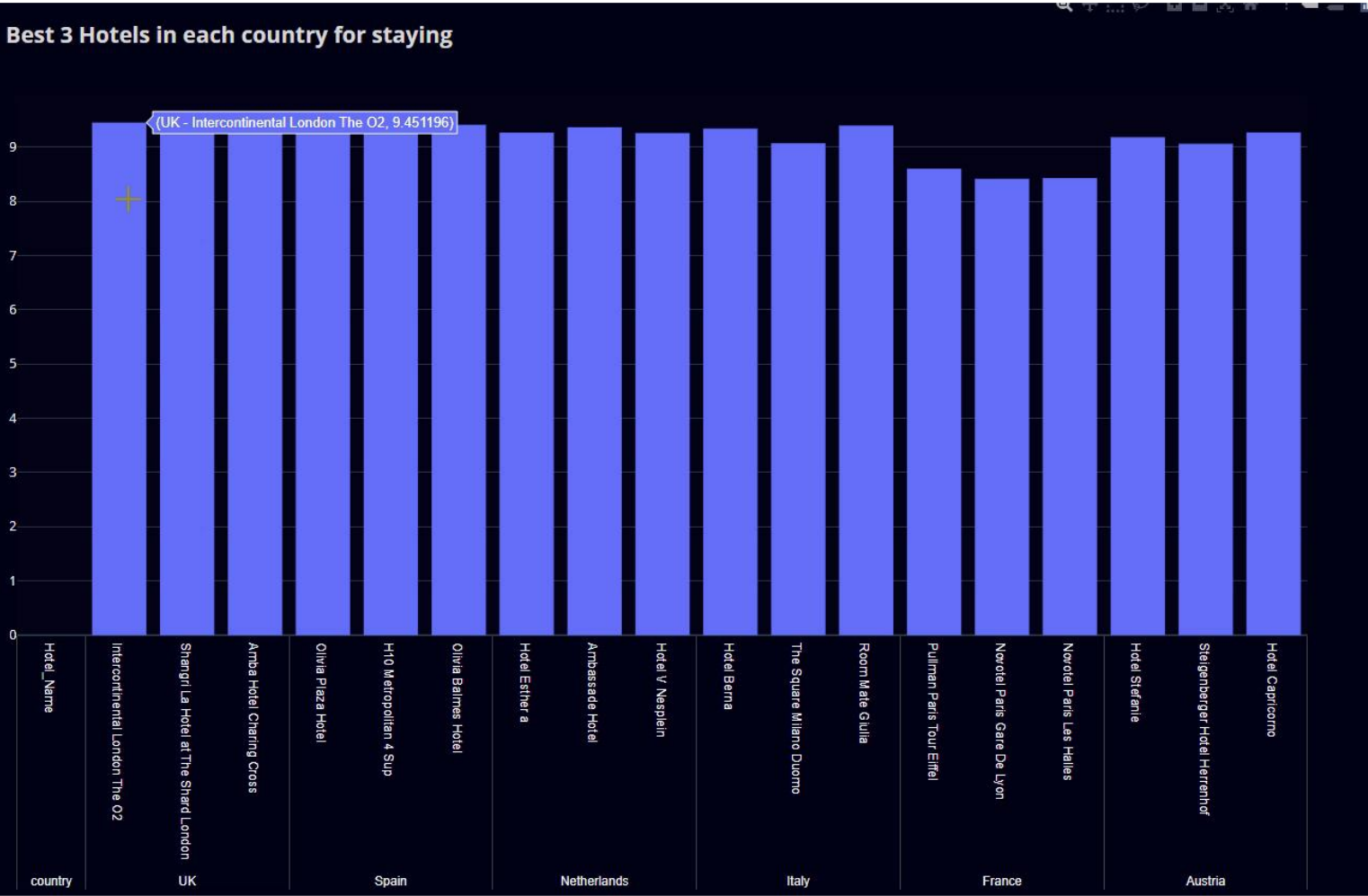


Business Questions:-

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 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.

Intercontinental **London** the o2



Olivia plaza hotel **Barcelona**



Hotel esther **Amsterdam**



Hotel Berna **Milano**



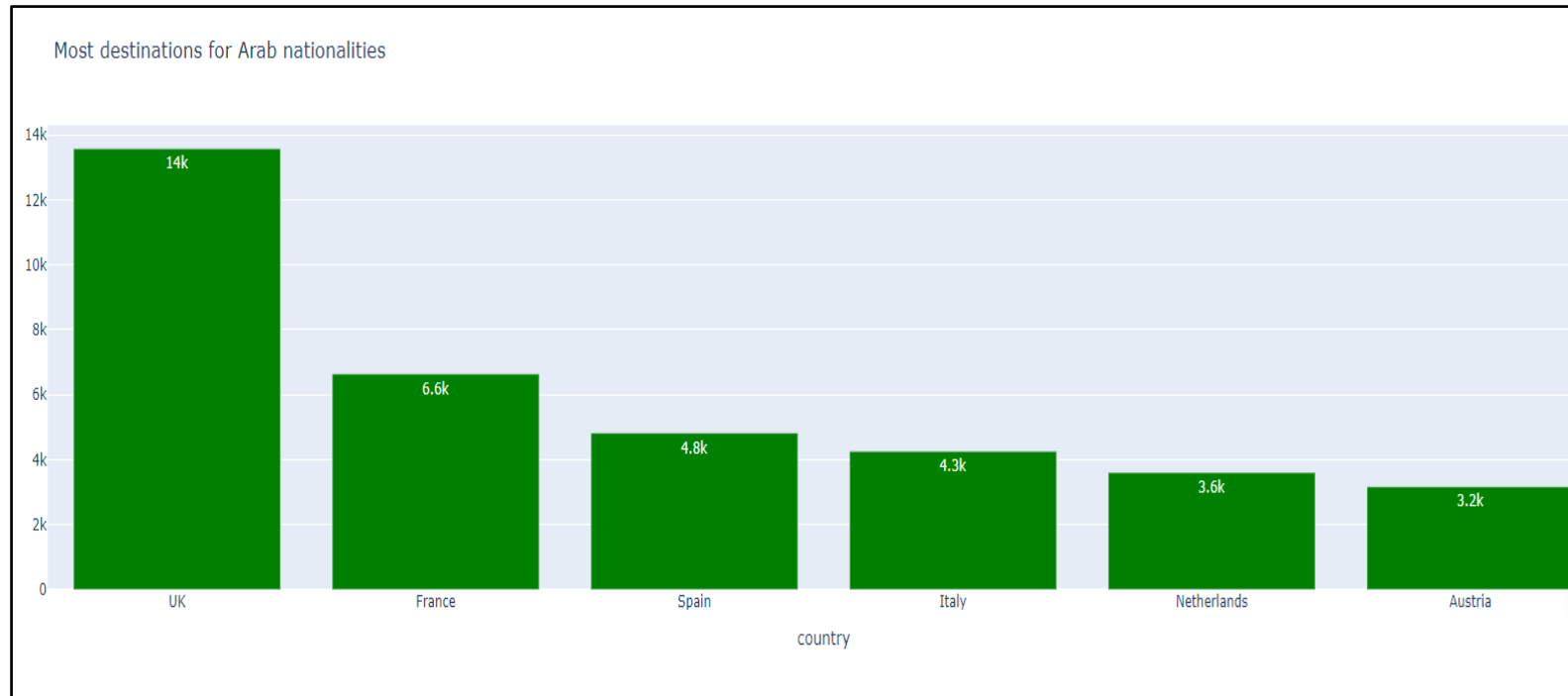
Pullman **Paris** Tour Eiffel



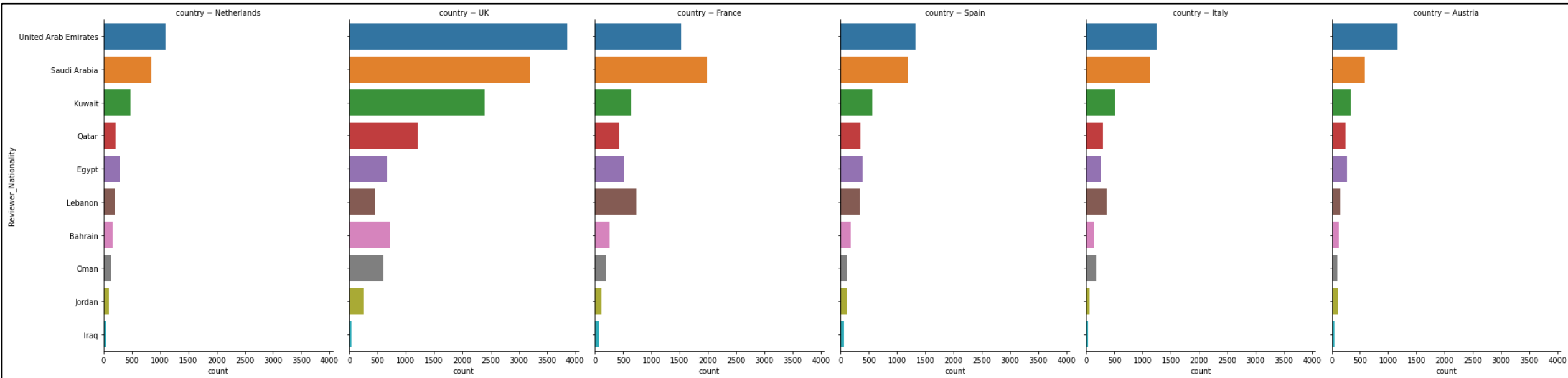
Hotel Stefanie **vienna**



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

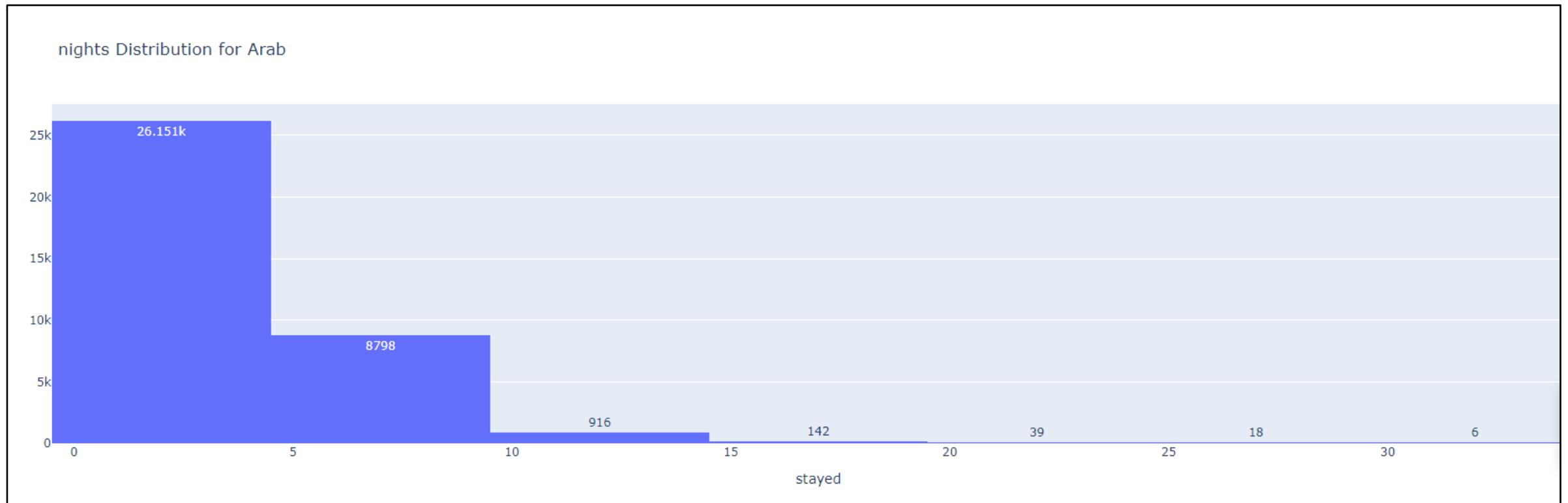


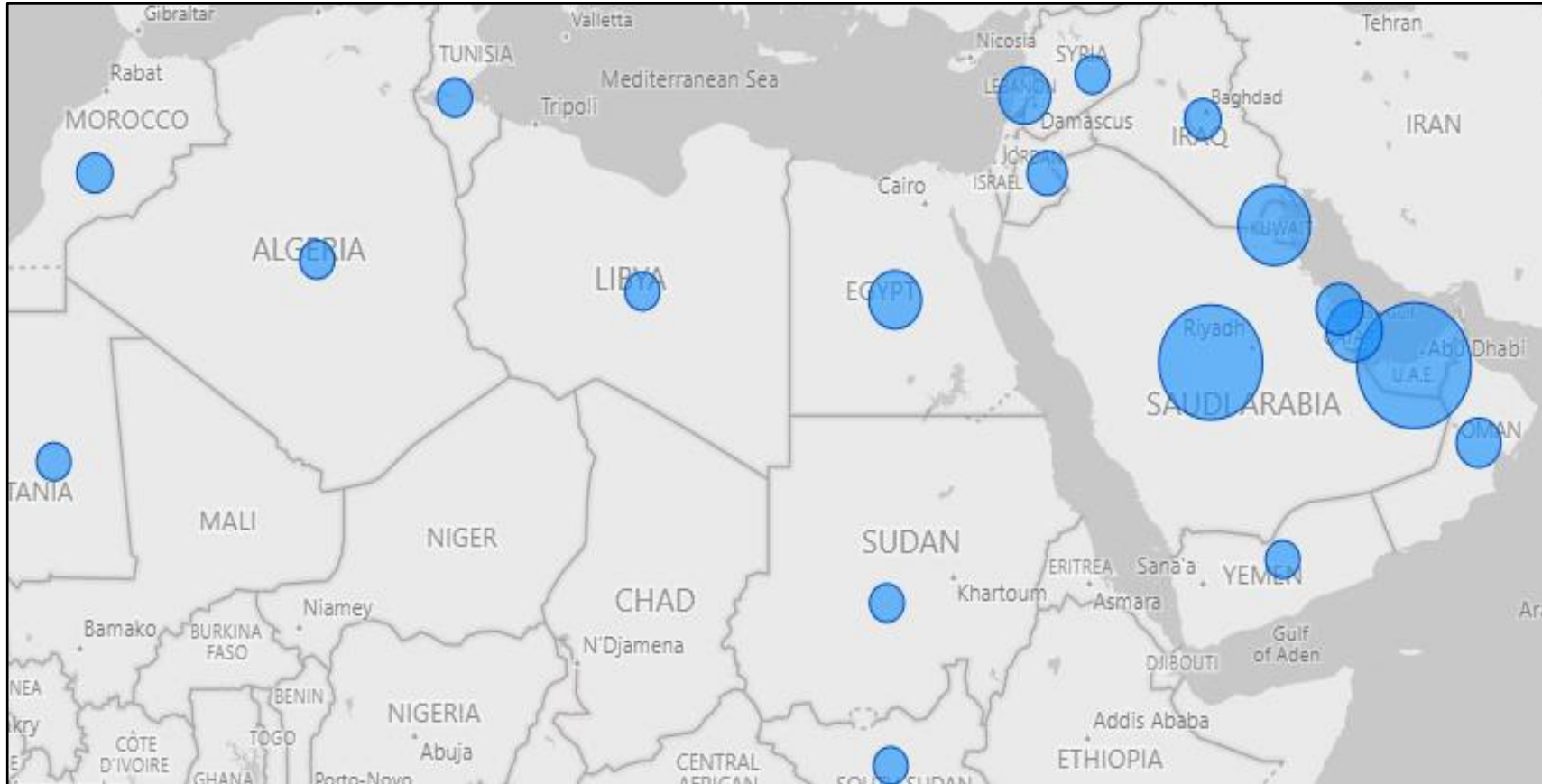
The most destinations
the United Kingdom,
then France and Spain.



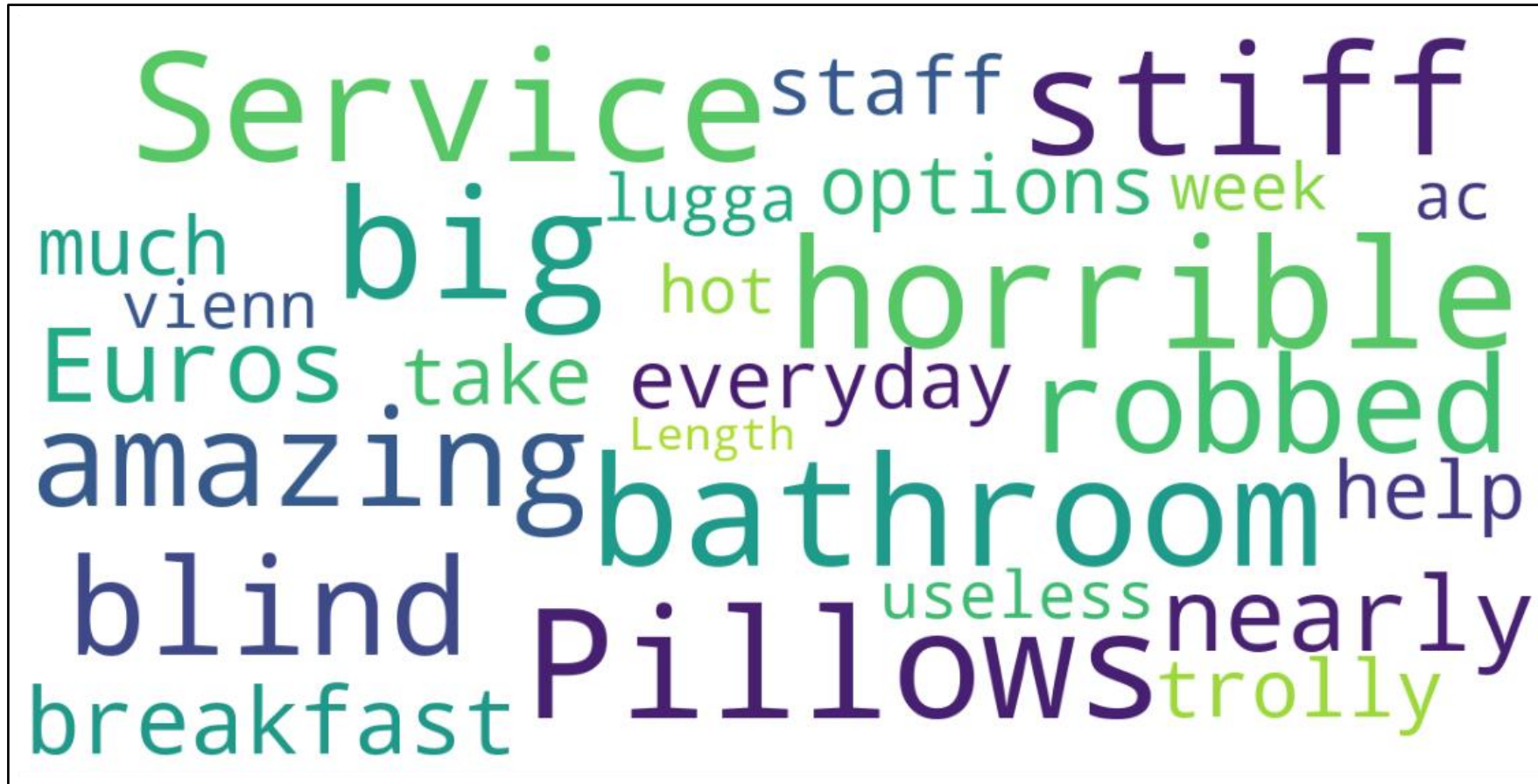
The most Arab nationalities are the Emiratis, then the Saudis in all countries
But the Saudis are the most in France

The average nights for Arabs stay from one to five days





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,

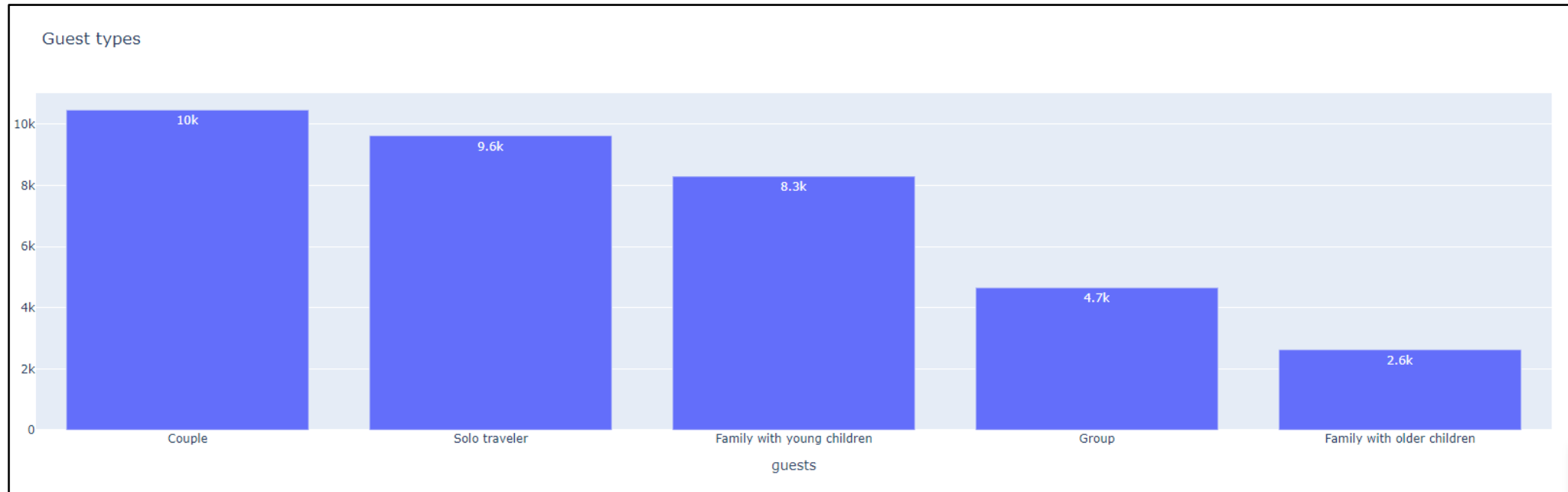


most comment negative comments of Arab guests :
pillows, blinds, bathrooms, and robbed



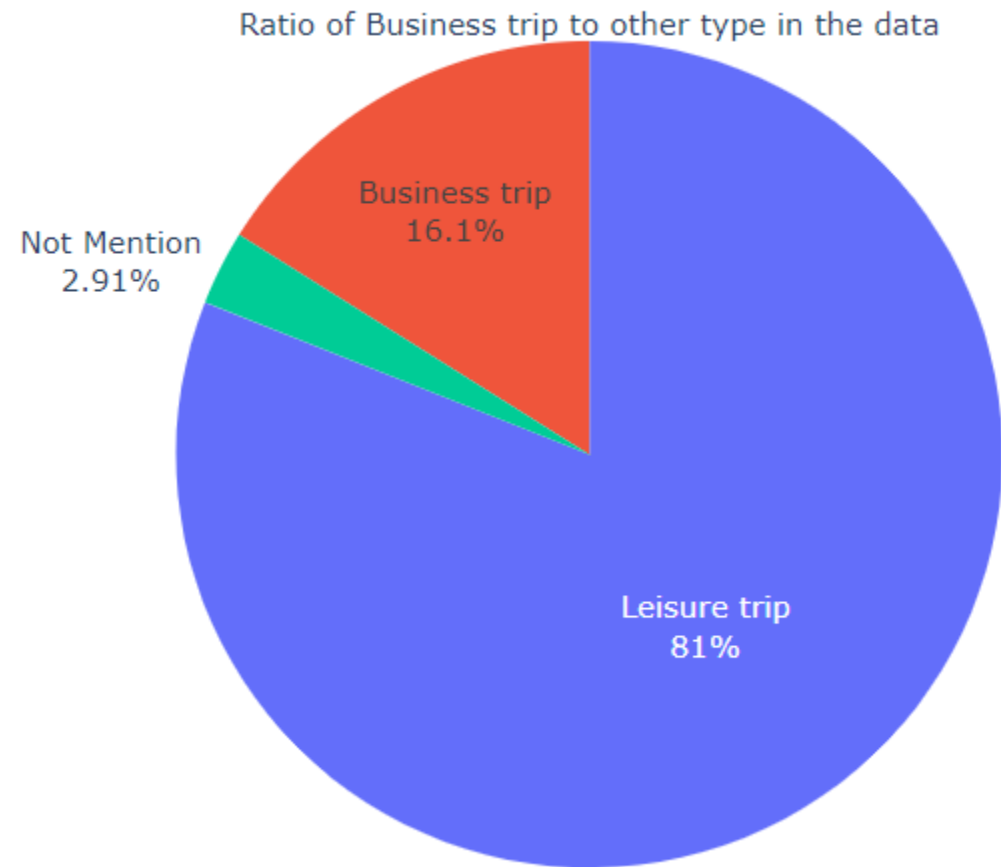
Most positive comments from Arab guests: location, cleanliness, hotel staff, parks.

Types of guests for Arab nationalities

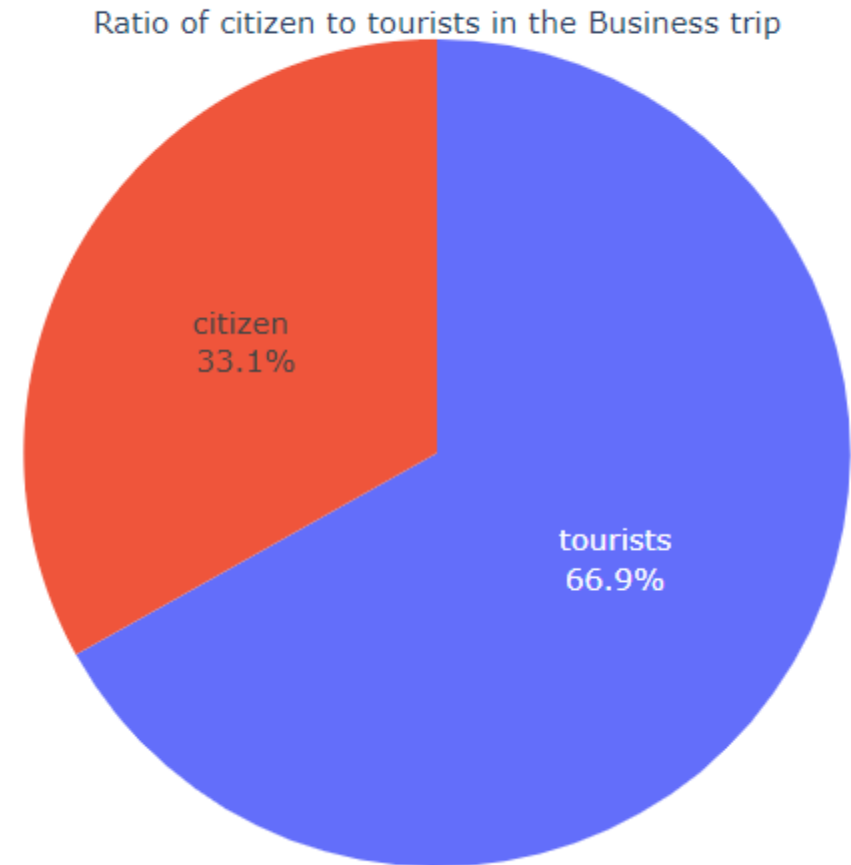


**What about *business trips*, the average days of stay
and the most reserved rooms, and do they have
special requests during booking, and what are the
most negative and positive reviews?**

Ratio of business trips to other types of trips



Ratio of tourists to citizens on business trips





negative comments talked about noise and workers, and some comments related to the car, such as the parking lot or the garage

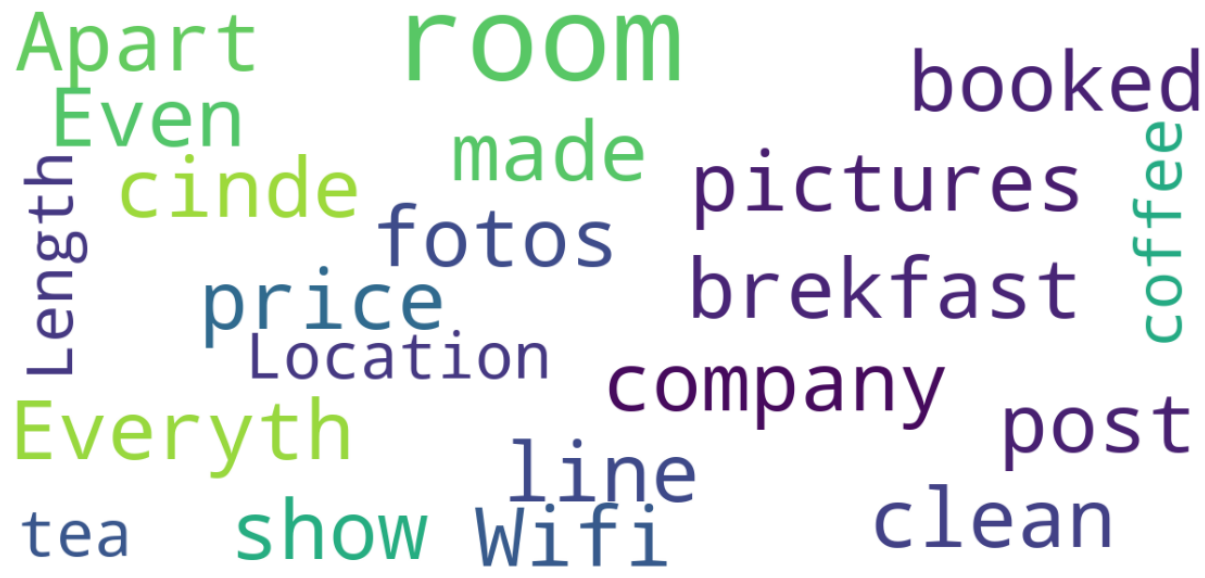
peaceful room lovelyLength
Breakfast areas
Public style modern
clean concept bright original nice
meetinglocation renovated
building hotel enough care

Positive comments were about admiring the style of the hotel, the design of the building, the room decor, and the breakfast

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?**

Couple – negative

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.



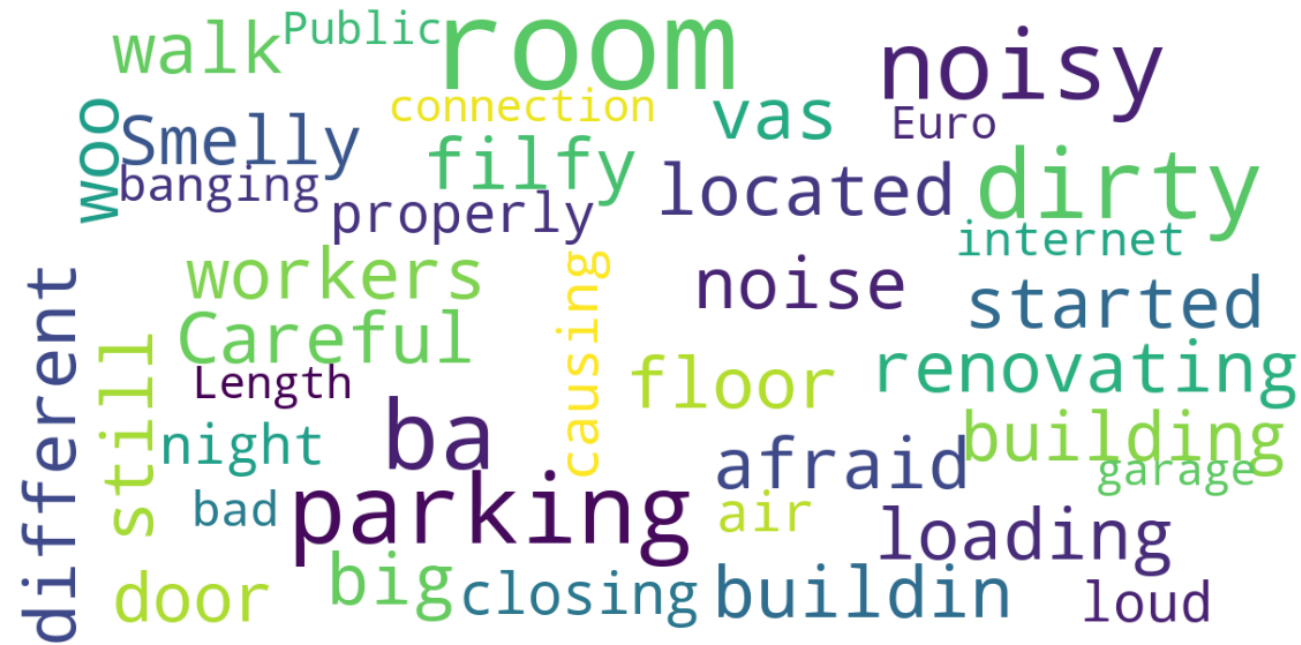
Couple – positive

the location of the hotel, the cheap price, and the location outside the hotel, and romantic.



Solo – negative

the negative comments were more than positive. The negatives were noise, bad smell, uncleanliness, and parking.

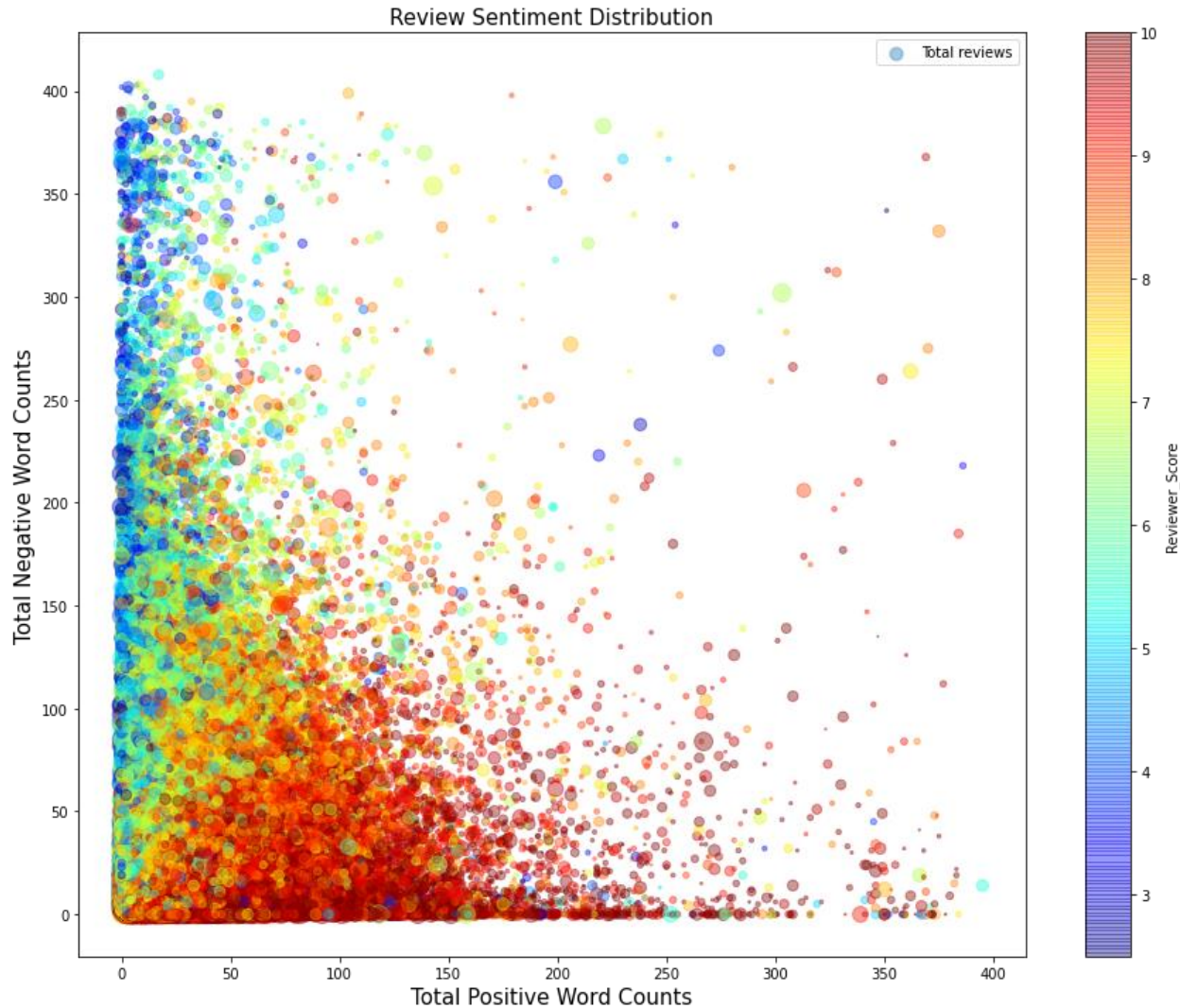


Solo – positive

the location ,
breakfast and the
room.



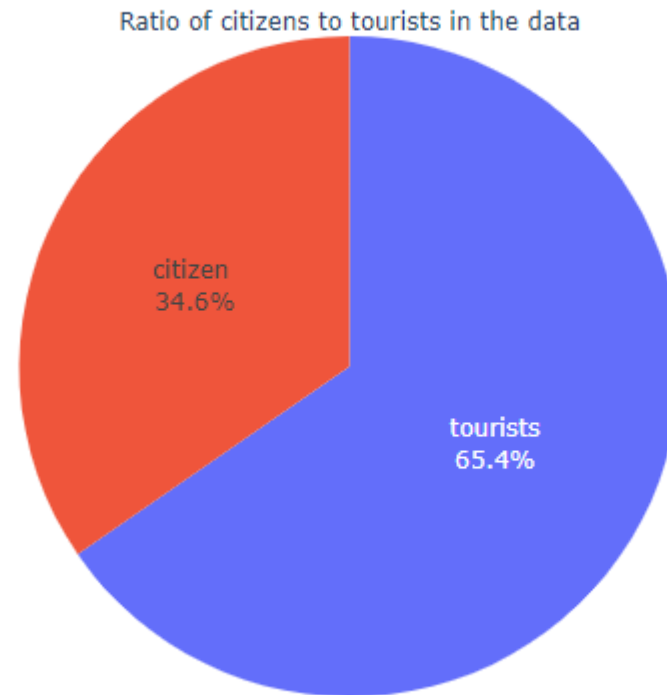
Are the **number of guest review **words** indicated on the **score**, and what are the most frequently used words in positive and negative reviews, and the most sentences or most effective?**



The number of words is not directly related to the evaluation, but we noticed in the negative words that the number of words as many is different from the positive

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

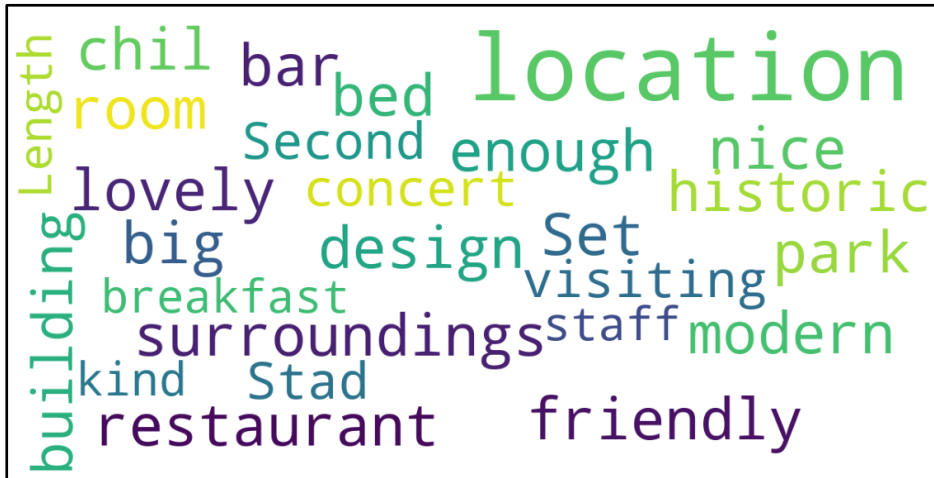
Ratio of tourists to citizens in all data
Citizens are those who stay in a hotel of the
same nationality



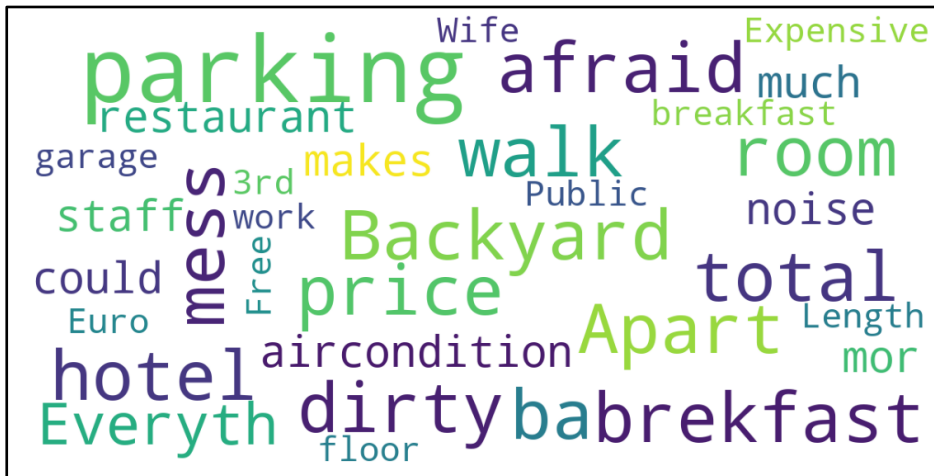
Is there a difference between the behavior of those who wrote a review for the **first time and others?**

First exp

Negative

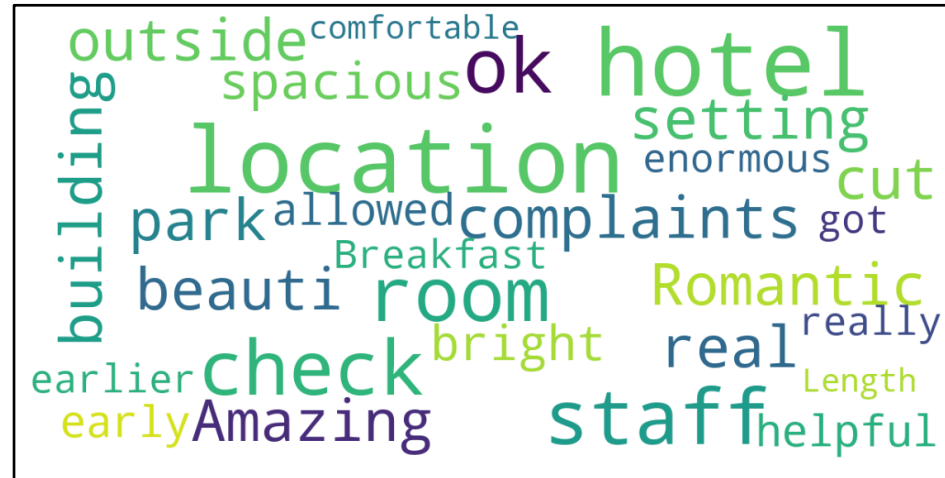


Positive



experience

Negative



Positive

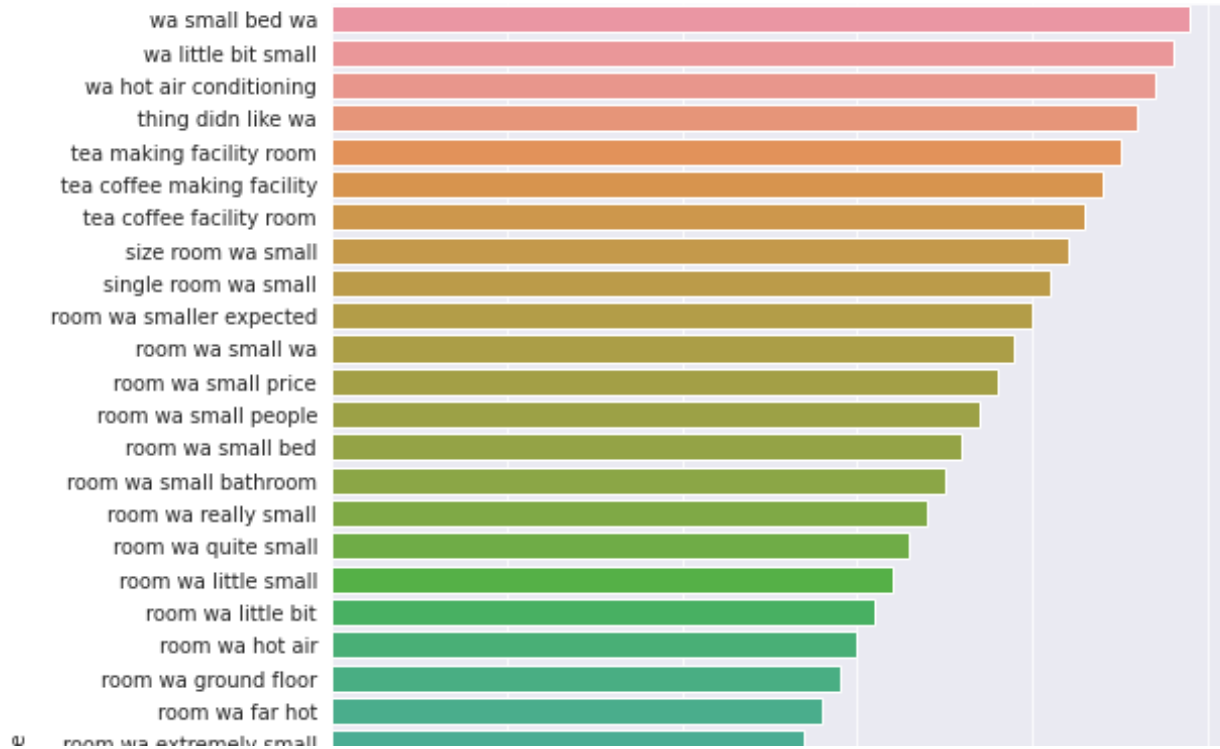


Sentiment analysis

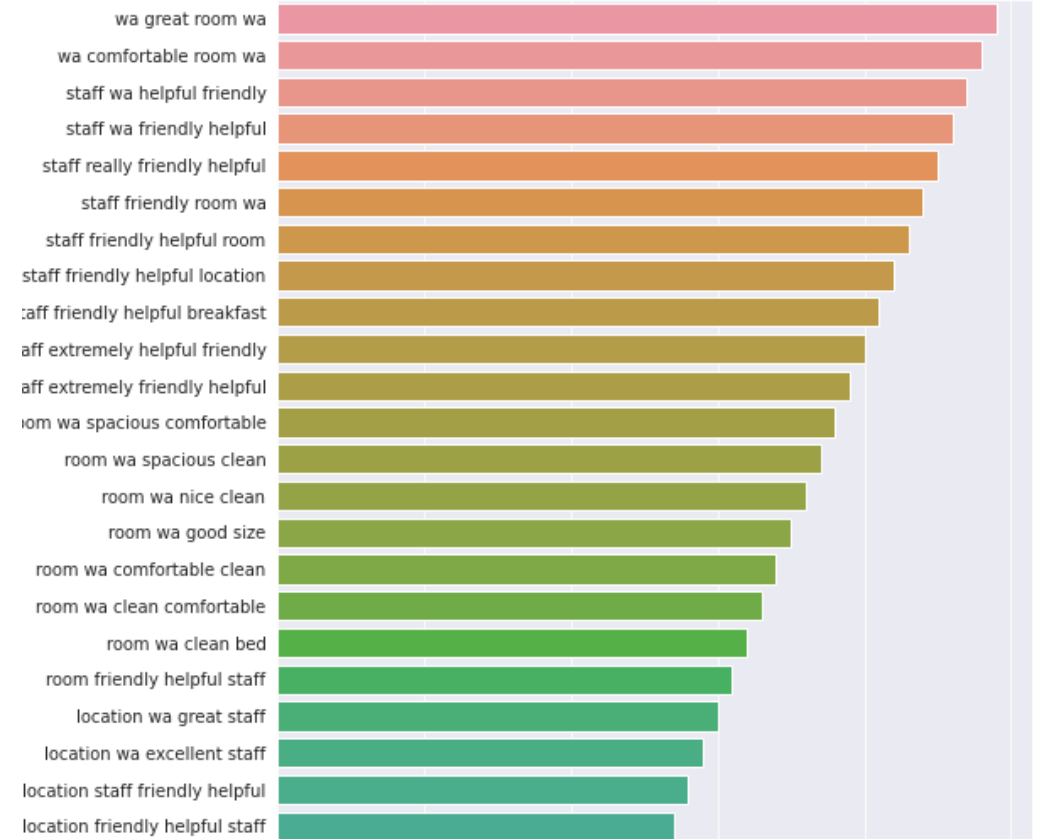




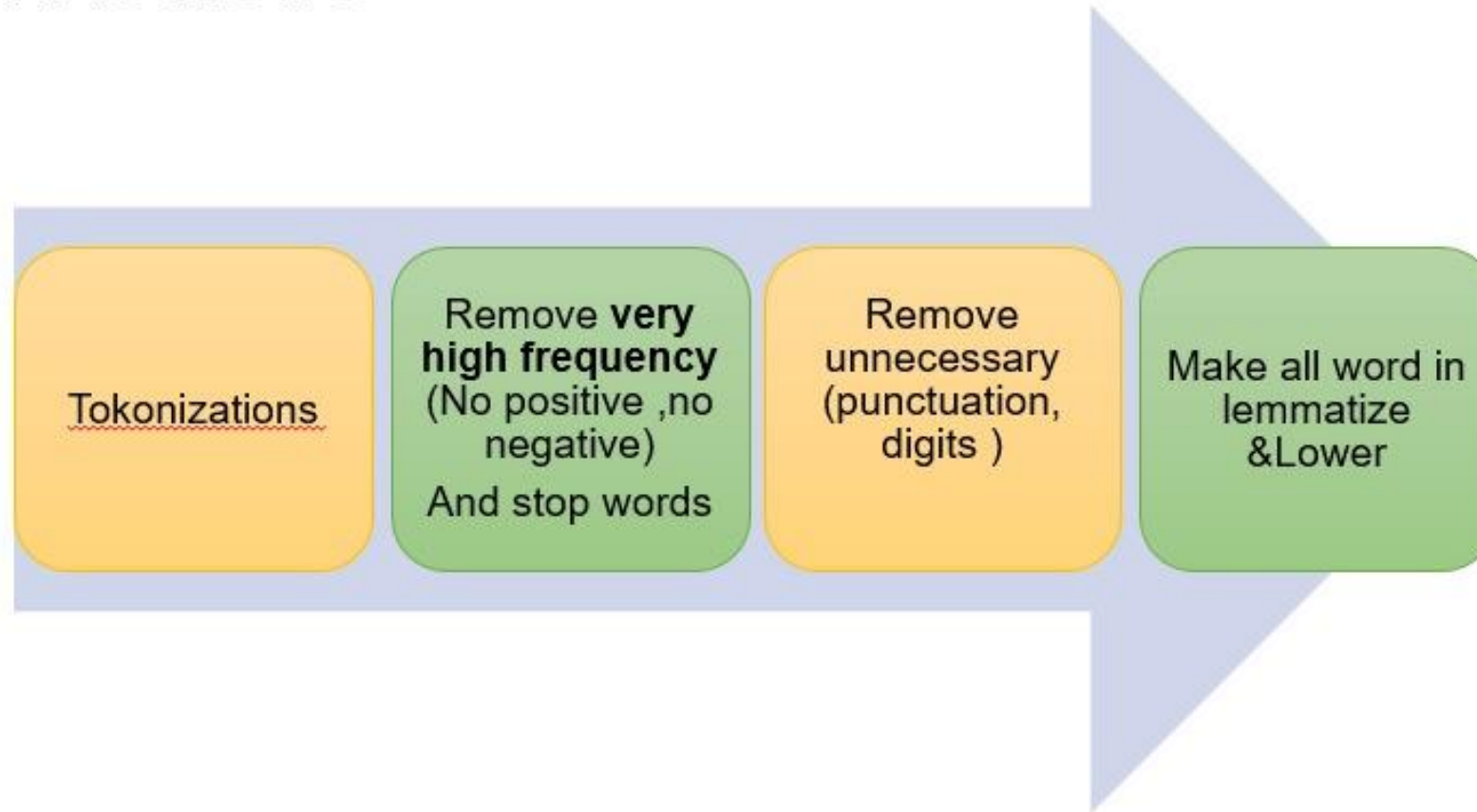
50 sentences of negative reviews



50 sentences of positive reviews

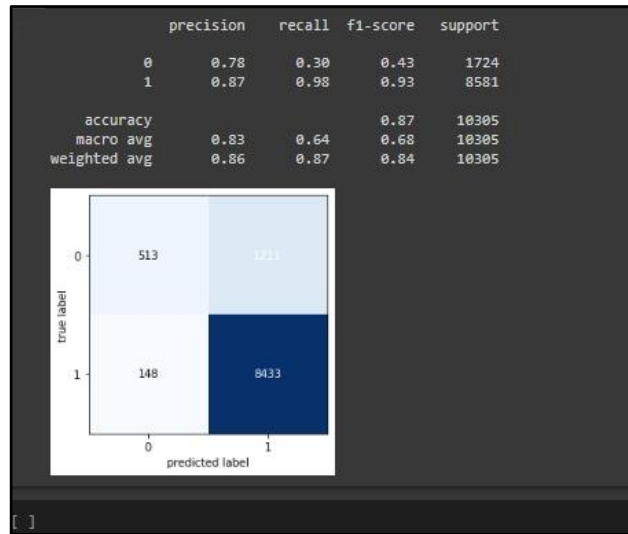


Clean Reviews

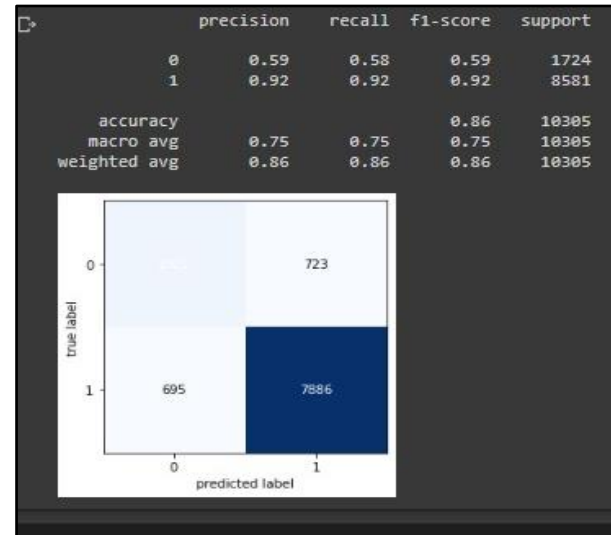


Modeling

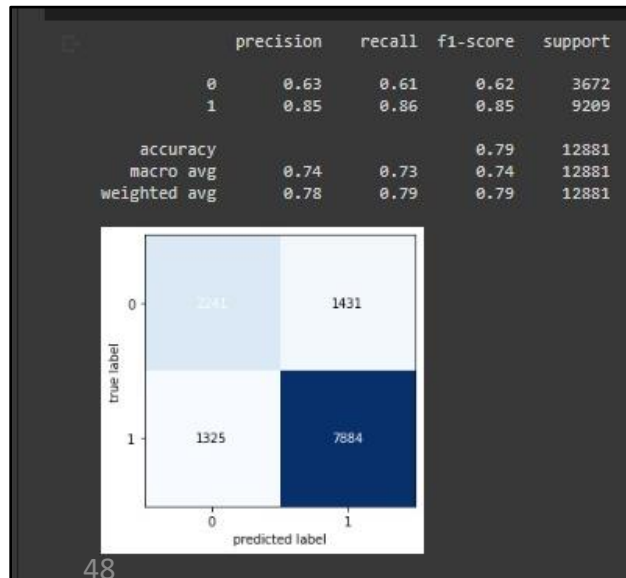
Random forest without SMOT



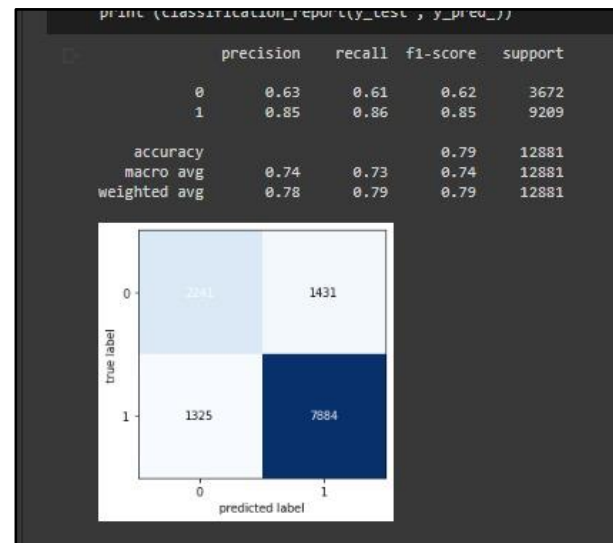
Random forest with SMOT



Random forest with ADASYN

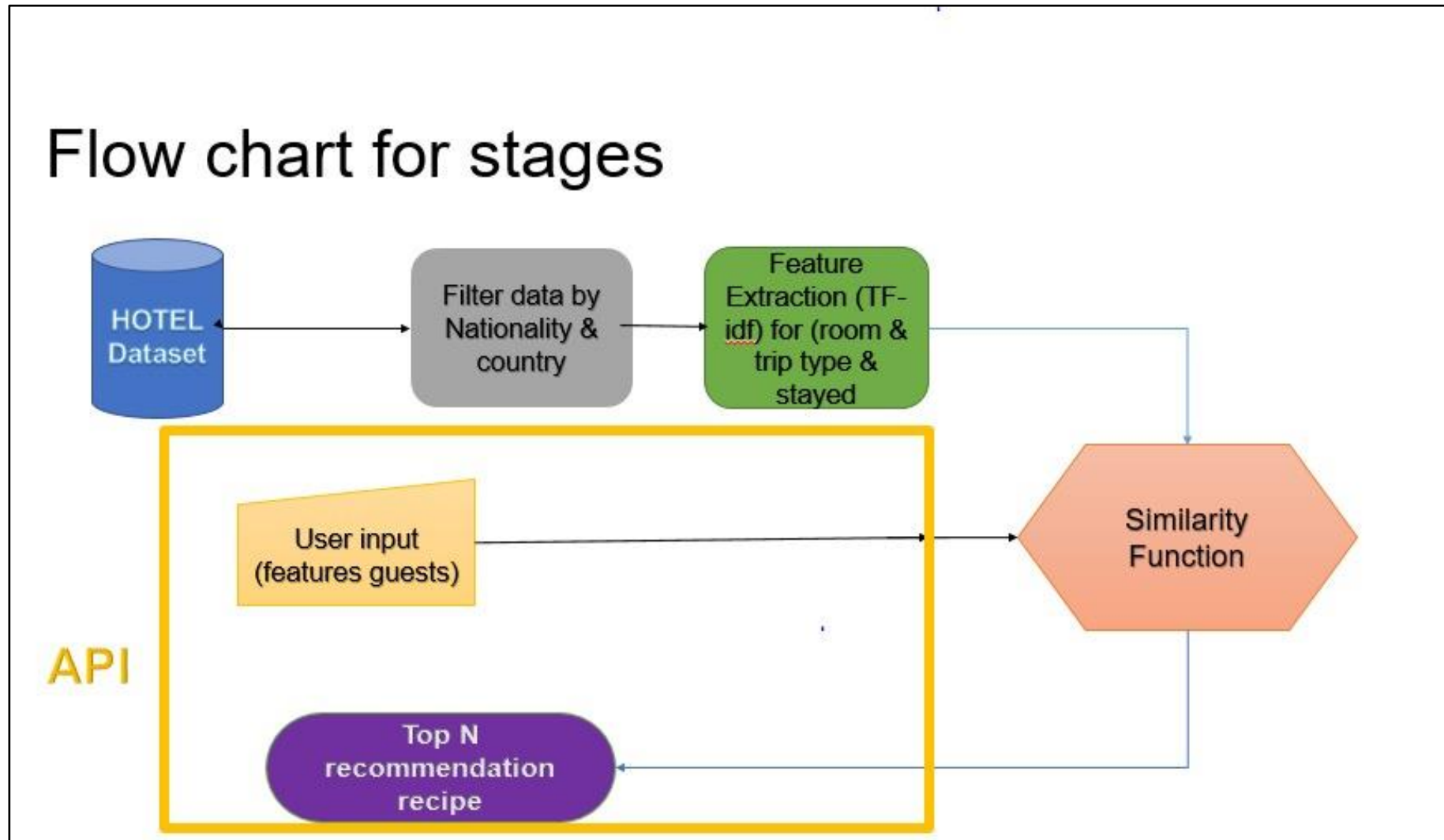


Feature selection with smote



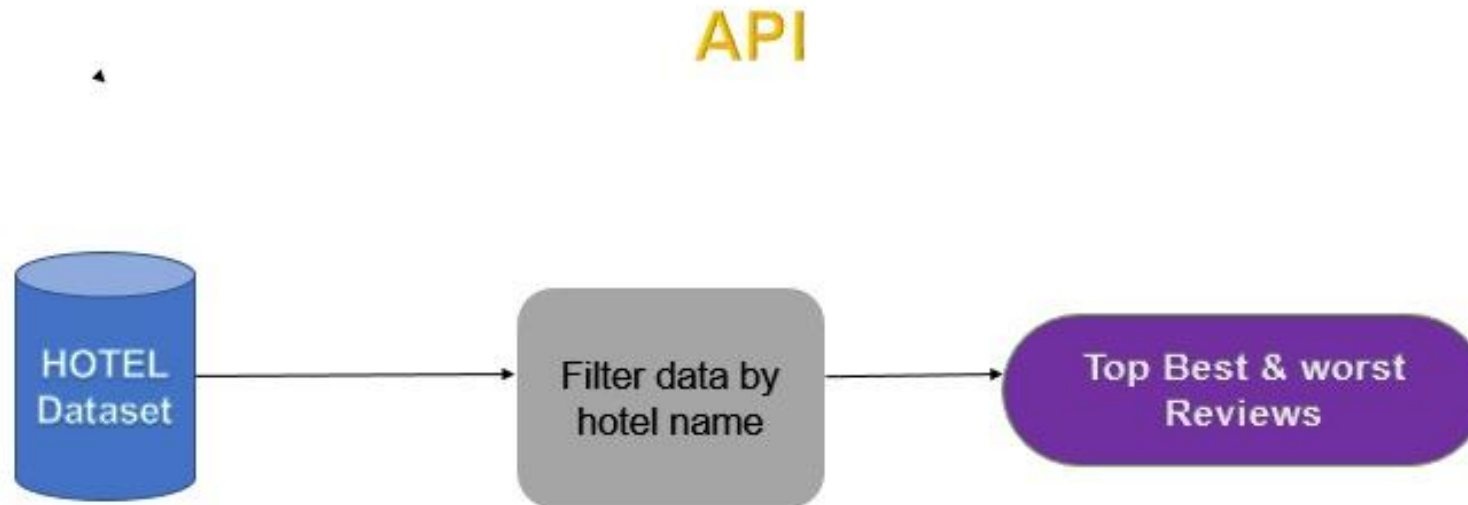
Recommendation System

Recommendation System:-



Recommendation System:-

Flow chart for guests Talks



Deployment



Deployment:- The Best for you

Select Page

guest_featuers

Informations about the client :

from

Russia

to

Netherlands

trip_type

☒ Leisure trip

☐ Business trip

guests

Couple


rooms_type

Duplex Double Room

stayed

0

31

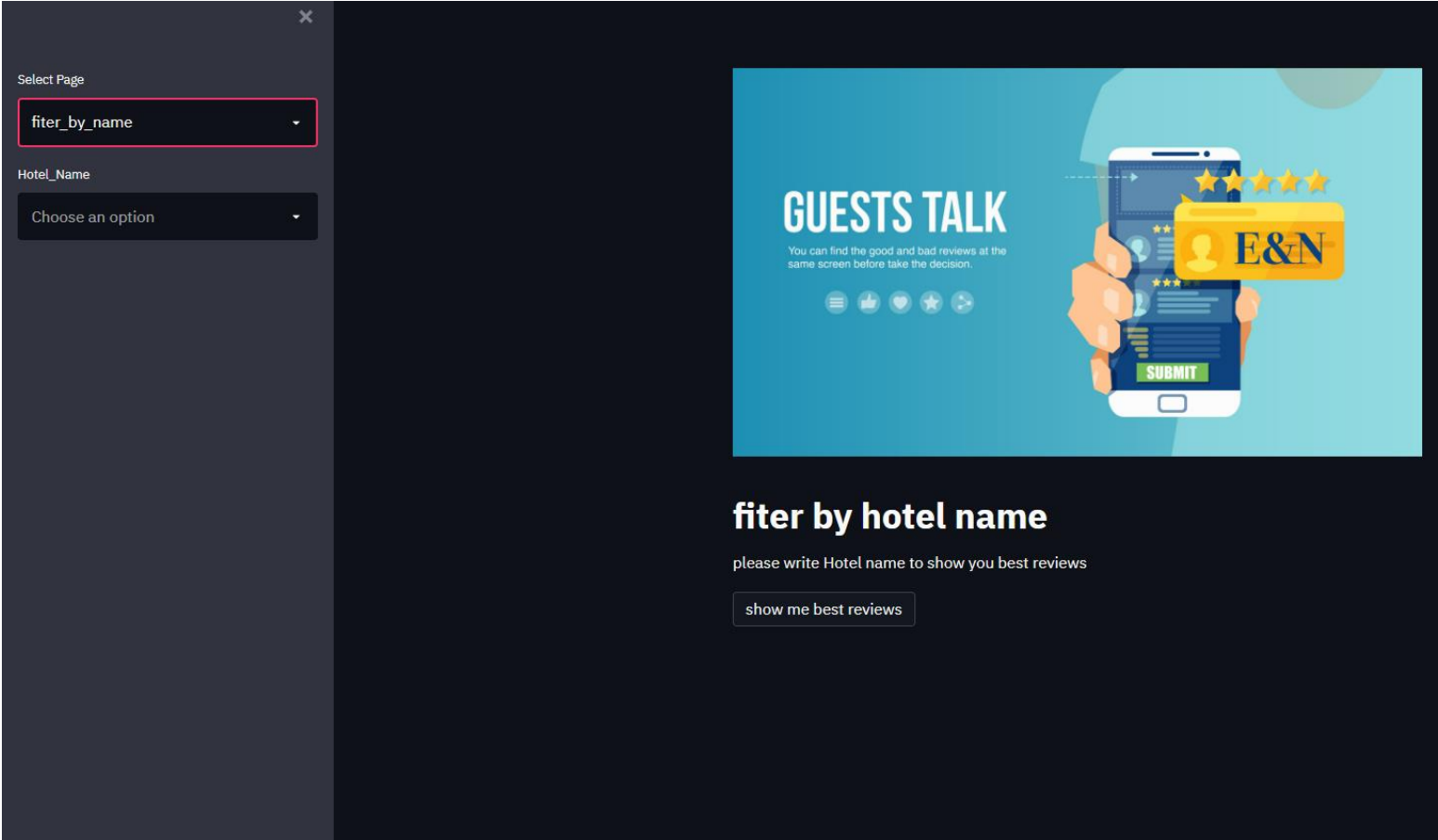


Recommended by guest_featuers

Sir/Mme , YOU need to fill all neccesary informations in order to get a reply to your best stay!

show me best Recommendations

Deployment:- Guests Talk

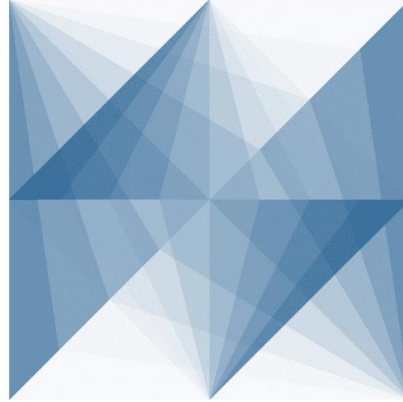


[LINK](#)

Conclusion:-

- Business analysis in this project can change all global tourism economy, providing tourism managers with deeper insights, and infiltrating into all aspects of tourist travels
- Answer to all complaints, know all compliment they have.
- Make clients satisfied.
- This data can often come as a good complementary source in data science projects in order to extract more learning features and increase the predictive power of the models.
- The experiences of the guests vary according to the country, the time of stay, the hotel staff, and from one person to another at the same time and place, but there are some basic things such as: Cleanliness - Photos must be real rooms with description of room size and bed size- noise - help from hotel staff, etc.
- The data detracts from some of the basic features that may be available, such as (room price- age of reviewer- gender - date of stay, etc.) With this data, the model can be greatly improved and find more ideas to serve more guests
- The work doesn't end, the ideas don't stop, it is possible to work and work a lot to analyze deeper and discover more features, but time.





Data science challenge

Reveal data secrets

Special Thanks to **Doaa Mahmoud Abdel-aty**
for support and give us chance

Elsayed Aly & Nour Shosharah

e.ahmed.gm@gmail.com

Nourshosharah@gmail.com