

Data Analysis of Global Tourism

Project 1 - Group 6:
Anh Pham, Hetal Prajapati,
Kathy Nguyen, Kelvin Osei Assibey,
Miguel Soriano

November 19, 2024



Introduction

Objective:

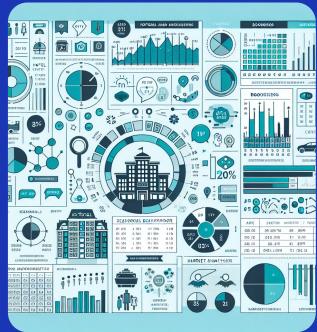
To analyze travel and hotel data to uncover insights addressing key research questions about tourist behavior and market trends



Why Tourism?

- Understand consumer behavior in hotel bookings through tourism data analysis.
- Utilize trends to predict seasonal patterns for strategic marketing and business planning.
- Inform pricing strategies and enhance decision-making for travel and hospitality stakeholders.

Data Exploration



**Primary
Hotel
Dataset:
Kaggle.com**



**Reference
Tourism
Dataset:
Kaggle.com**



**Reference
Google
Places API:
[cloud.google.c
om](https://cloud.google.com)**



**Reference
Geoapify
API:
Geoapify.com**

Hotel Dataset Cleaning



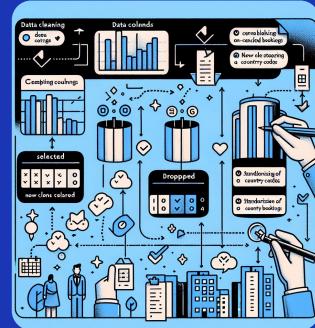
1. Focused column selection



2. Created new columns



3. Refined dataset, excluded canceled bookings & nulls



4. Renamed country codes to full names

Research Questions

Anh

Popular Destinations

What are the most popular destinations for travelers?

Miguel

Trip Duration Variability

How does the duration of trips vary by destination?

Hetal

Hotel Ratings and Bookings

What is the relationship between hotel ratings and the frequency of bookings?

Kelvin

Seasonal Tourism Trends

Are there seasonal trends in tourism?

Kathy

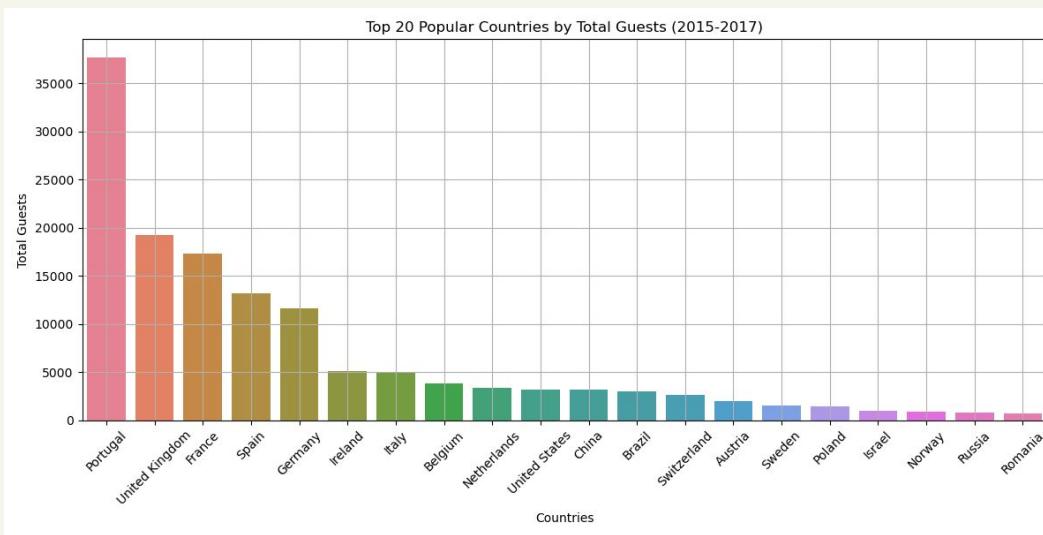
Hotel Price Fluctuations

How do hotel prices fluctuate over time in different regions?

Popular Destinations

3 key metrics:

- Total Guests
- Average Stay
- Total Guest Nights



Total Guests

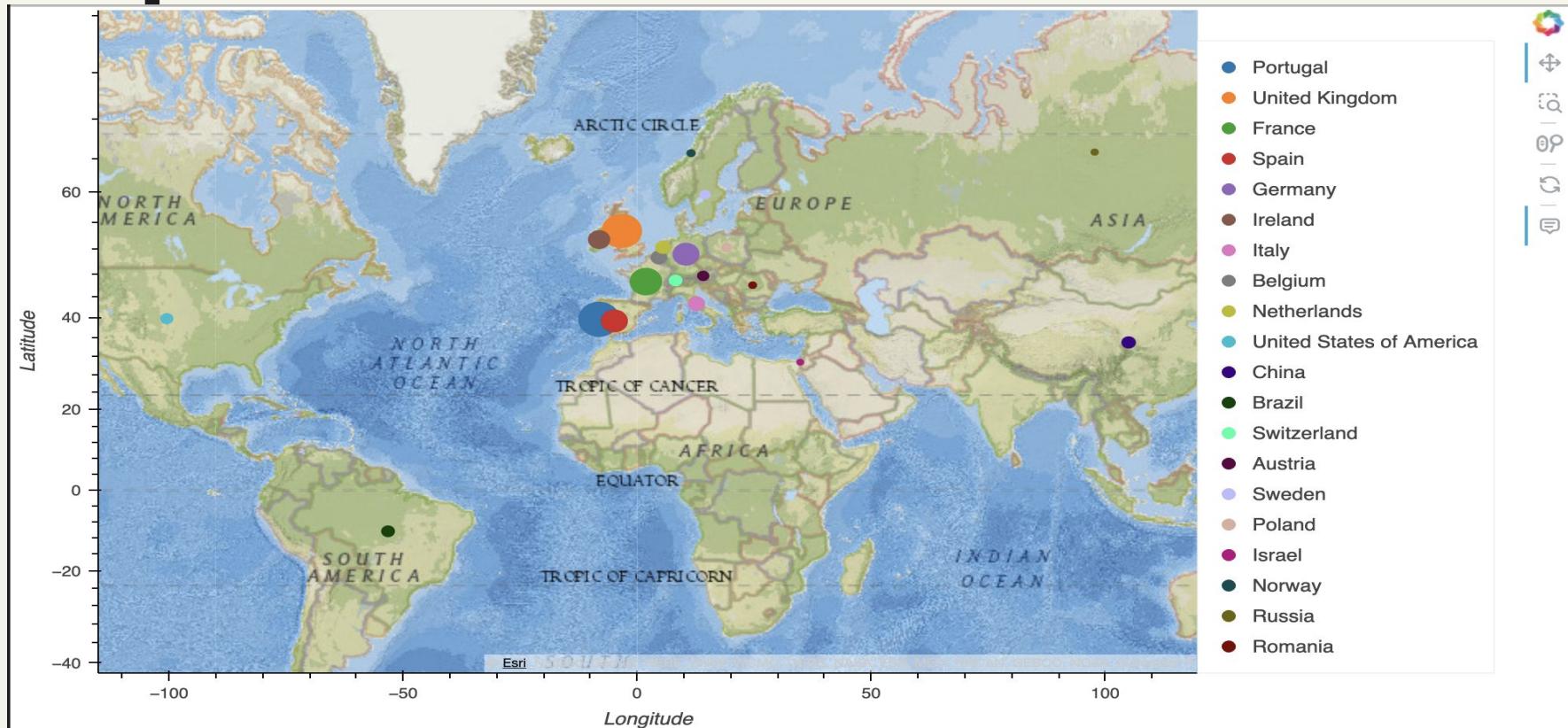
Portugal - tops the list: cultural and historical appeal, affordable travel options, and increasing international awareness

United Kingdom follows closely - major international hub, 50% of Portugal guest arrivals

France, Spain, and Germany make up the next few spots - rich cultural offerings, major cities, and significant infrastructure

These 20 countries account for 94% of the total guests across all 164 countries in the dataset.

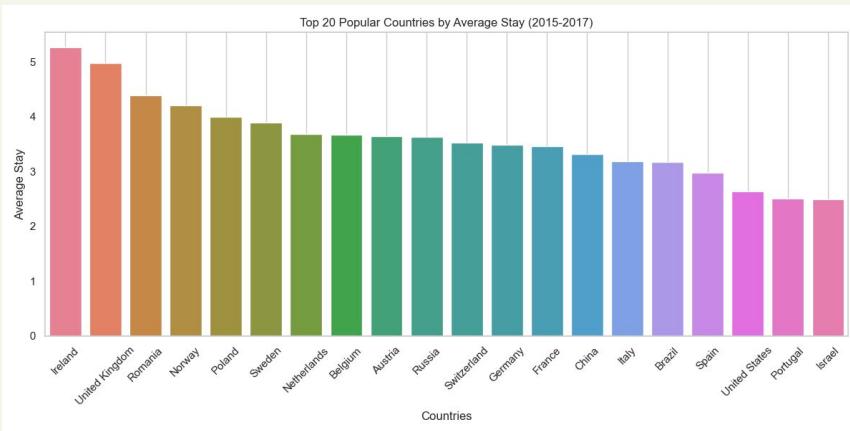
Popular Destinations



Popular Destinations

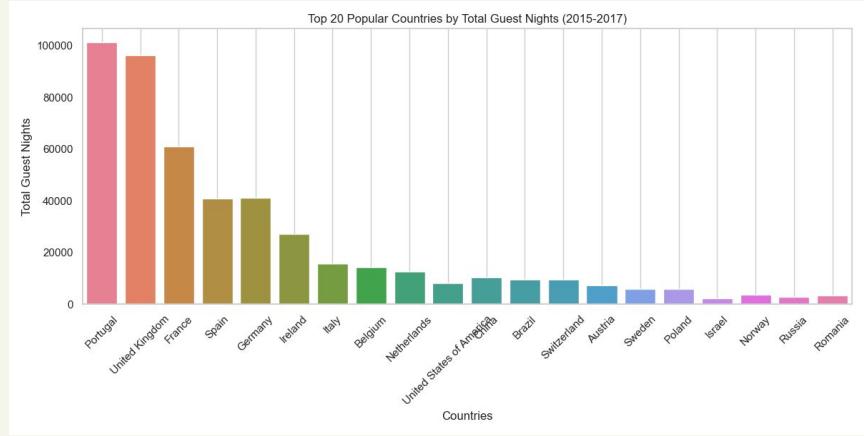
Average Stay

A longer **average stay** might indicate that tourists are investing more time in exploring the destination, perhaps for a more **immersive or niche experience**.

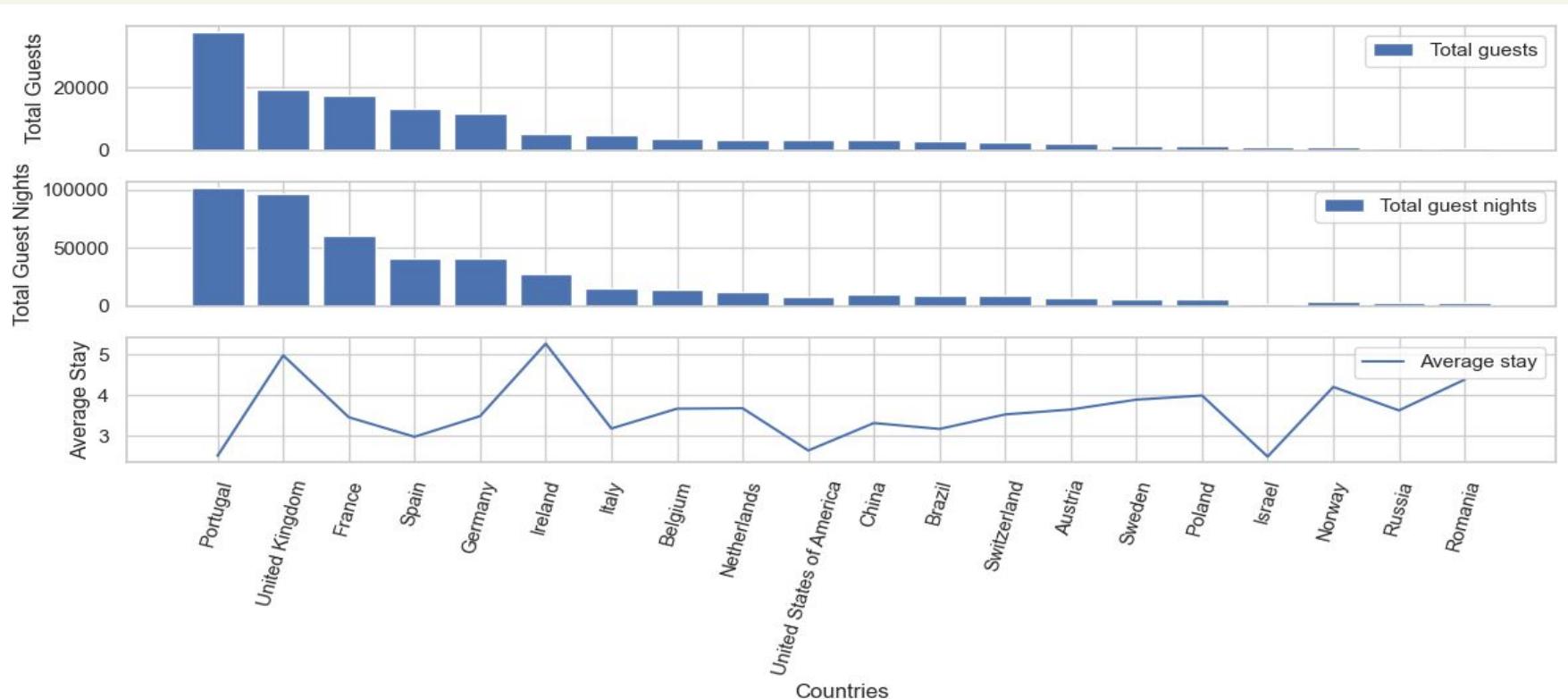


Total Guest Nights

A high total number of guest nights indicates that a destination is **attracting not only a lot of tourists but also** encouraging them to **stay for longer periods**.



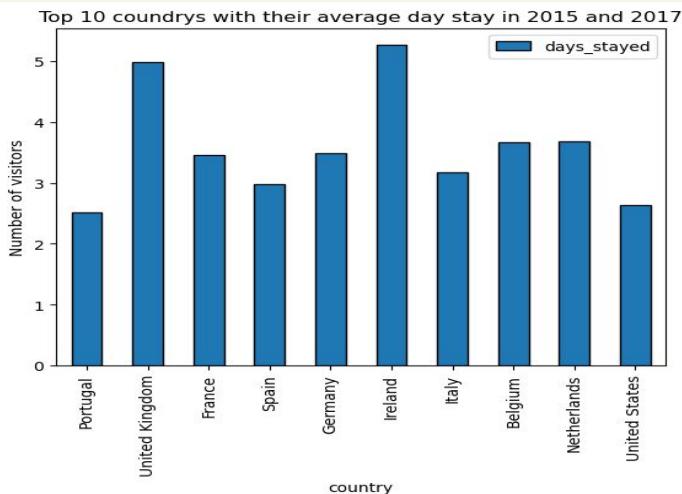
Popular Destinations



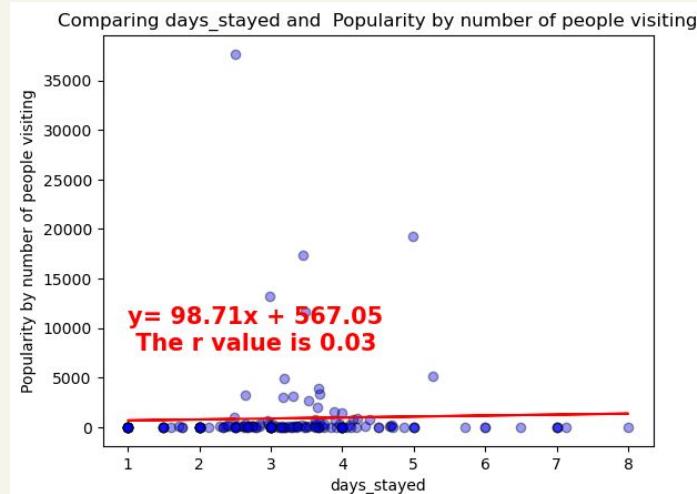
Trip Duration

- Popularity of destination does not correlate to how much time visitors will stay in a country.
- Both the top 10 countries and overall countries have no significant linear correlation between popularity of a country and the number of average days stayed in that country.(graph 2)
- Although portugal is overall the most visited country according to our data, it ranked the lowest on average for day stayed at 2.5 days, for top 10 most visited countries.
- The top ranked country for average day stayed is Ireland at 5.3 days, among the top 10 countries.(graph 1)

Graph 1



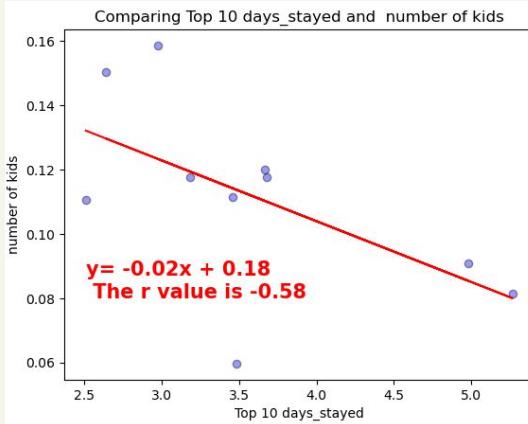
Graph 2



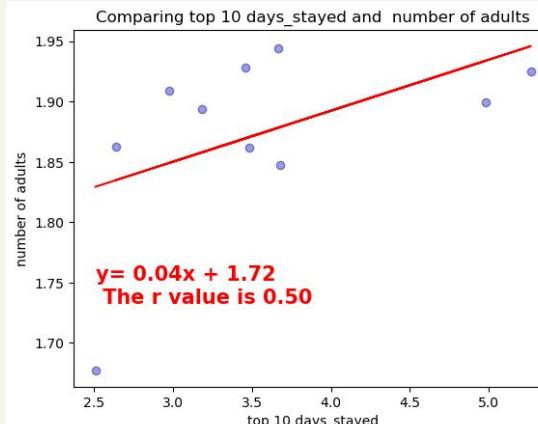
Trip Duration Variability

- Taking into account all the countries, there was no specific variable that had much of an affect to the duration of a trip.
- When looking more specifically at the top 10 visited countries in our data, we can see a variable with a significant measurable linear correlation to the duration of a trip. This variable is whether kids join a trip or not. When kids are involved in a trip, there is a negative linear correlation of -.58.(Graph 1)
- The more adults in a traveling group, the longer the trip duration. This variables had a positive linear correlation of .50.(Graph 2)
- The average rate for hotels in the top 10 countries has a measurable negative linear correlation(-0.51) with the duration of trip.(Graph 3) The higher rate the shorter the stay.

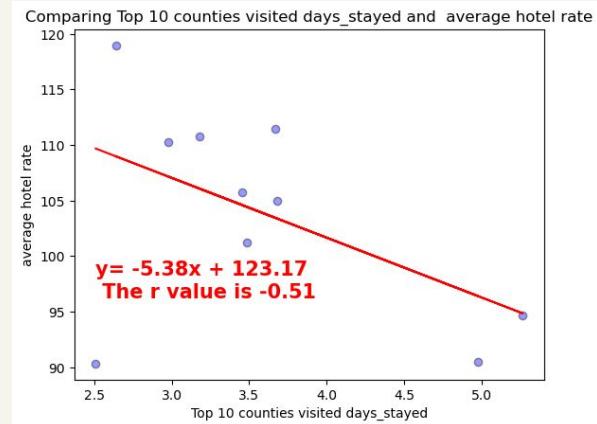
Graph 1



Graph 2



Graph 3



Hotel Ratings and Bookings

Relationship between Average rating and Booking Frequency.

1. Correlation Coefficient: 1.0

Indicates a perfect positive correlation.

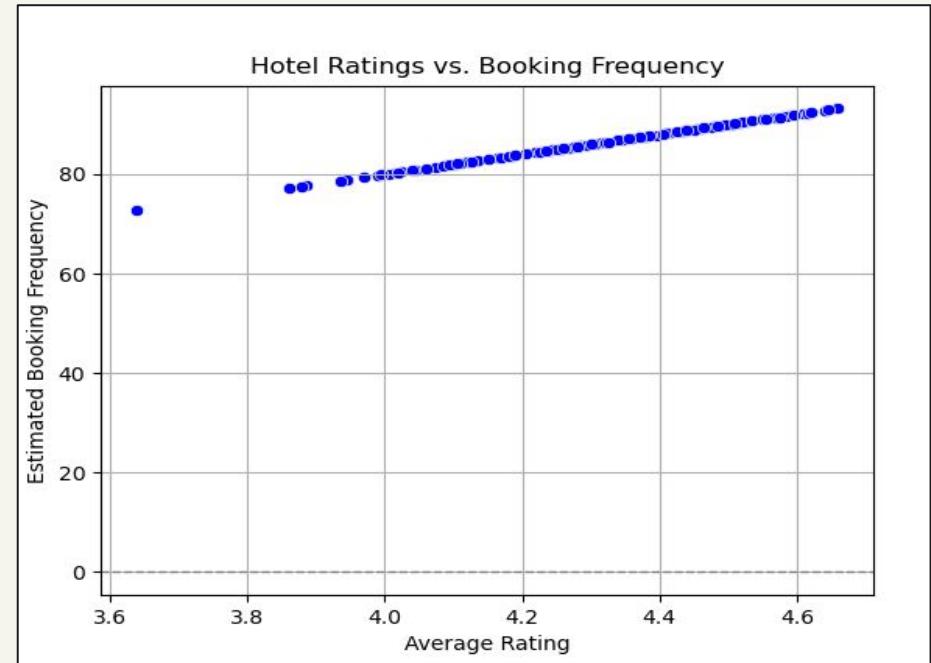
2. Implications:

Higher hotel ratings lead to more frequent bookings.

Important for hotels to maintain high ratings to increase booking frequency.

3. Visual Representation:

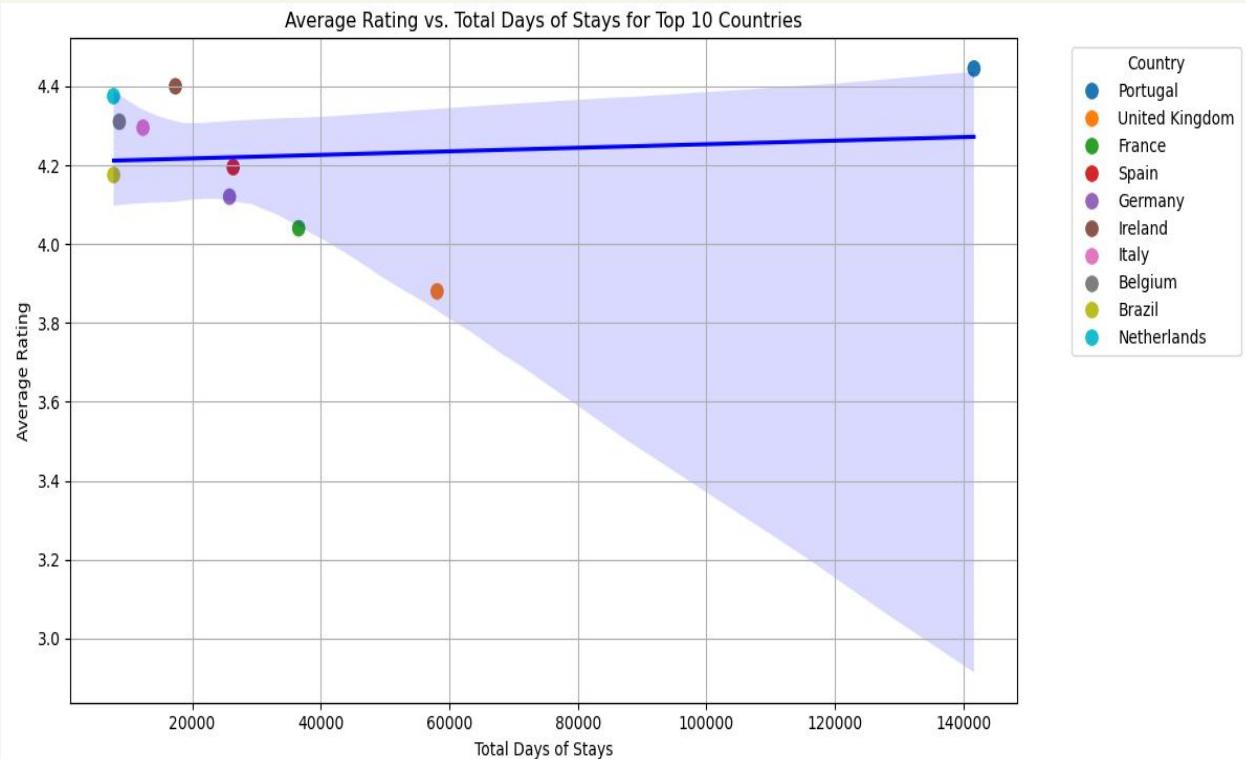
Scatter plot shows a perfectly linear trend.



Hotel Ratings and Bookings

The number of higher total days of stays correlate with a higher average rating for top 10 countries:

- There is a weak positive relationship between average hotel rating and total days of stays.
- The regression line indicates a slight upward trend, but data points are widely dispersed.
- Portugal stands out with the highest average rating among the top 10 countries.



Seasonal Tourism Trends

- **Seasonal patterns**

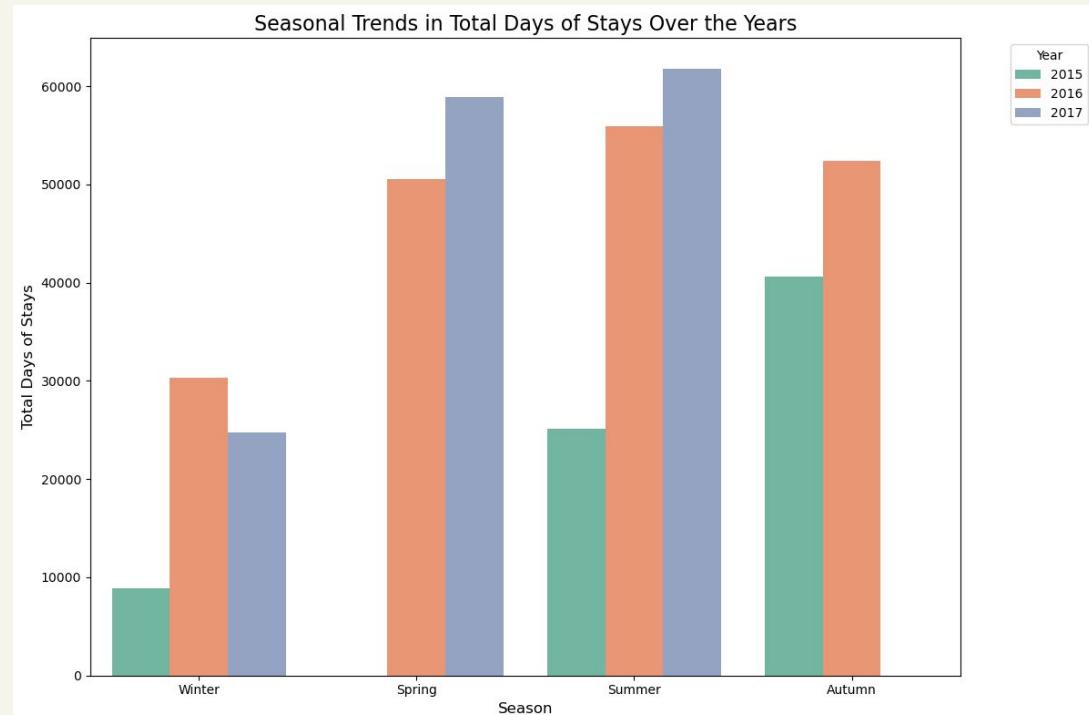
All seasons recorded an increase in total number of days from 2015 to 2017.

- **Differences across seasons**

All the seasons have either been appreciating or stable regarding the total number of stays.

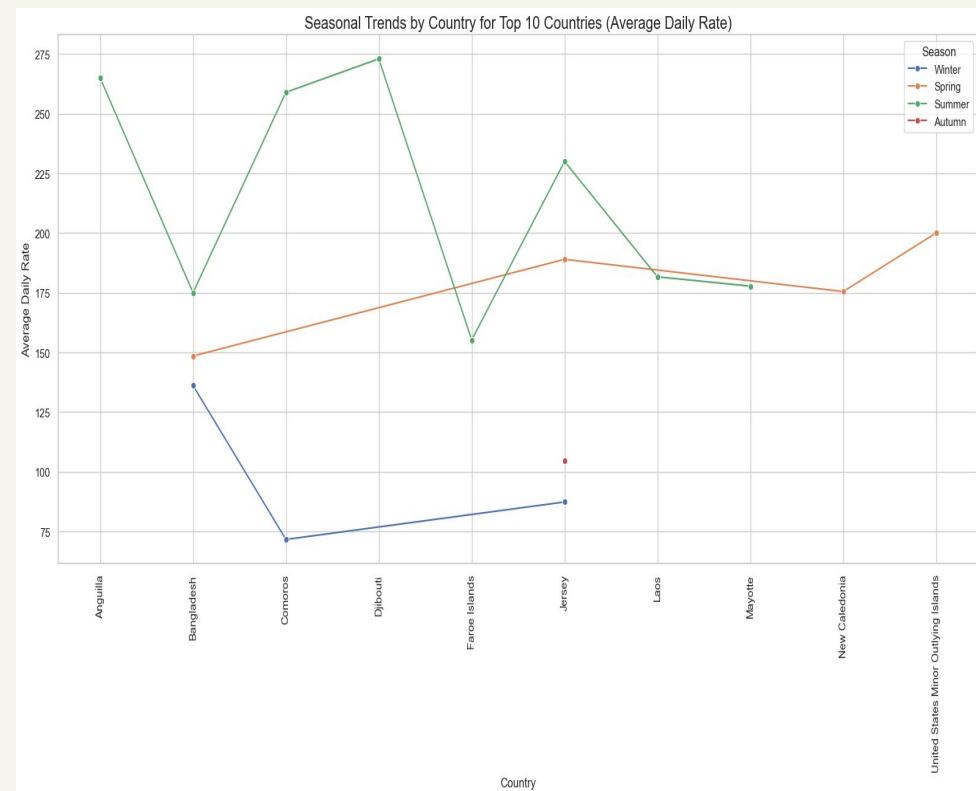
- **Changes over time**

Summer seasons recorded massive increase in total days of stay over the 3-year period ranging from 25140 to 6180

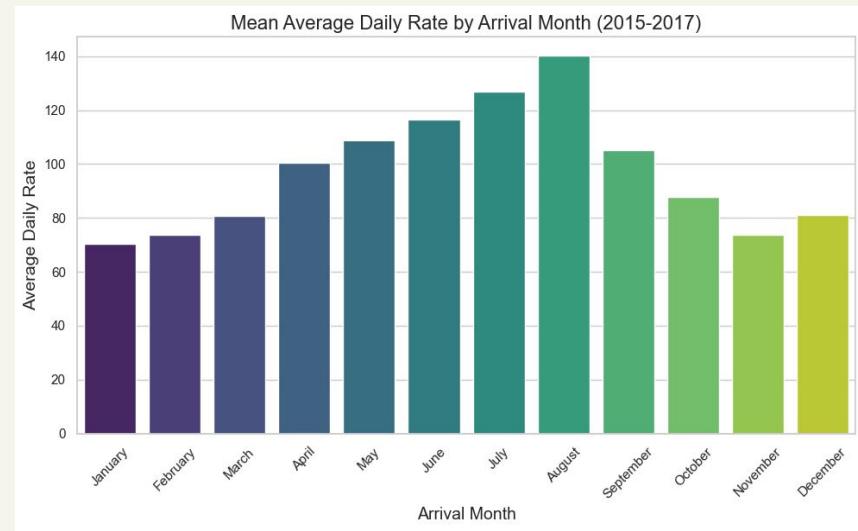
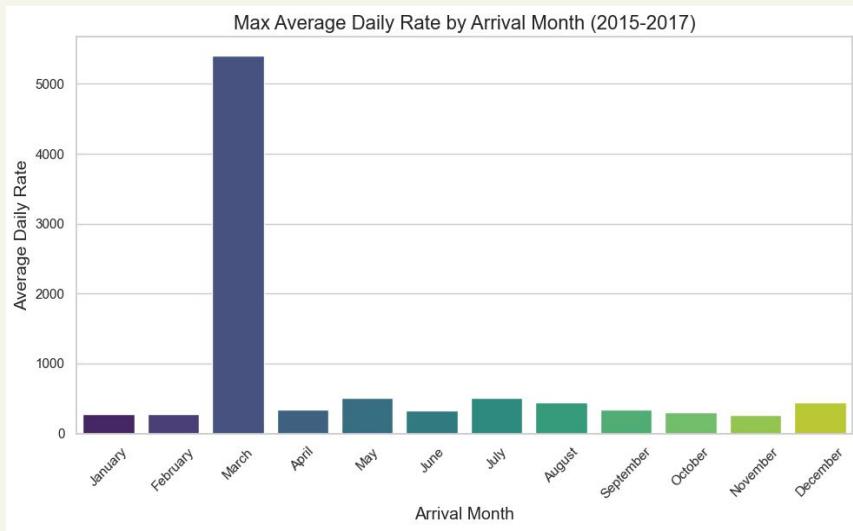


Seasonal Tourism Trends

- **High ADR countries:** These 10 countries typically show a higher ADR across all seasons, which is a direct reflection of both higher hotel prices and potentially more premium accommodations.
- **Seasonal Variations in ADR :** summer sees higher ADR in countries that are popular during warmer months.
- **Consistency Across Seasons:** Some countries might show consistent ADR across seasons. Jersey had a consistent ADR within the 3-year period recording 87.31 in winter, 189.00 in spring, 230 in summer and 104.67 autumn.



Hotel Price Fluctuations



Using Max ADRs:

- March reaching a max of 5400 (likely due to outliers or peak travel events). So the Mean ADR was utilized to show a true trend between the months.

Using Mean ADRs:

- July & August consistently high ADRs, reflecting peak travel seasons in many regions.

Monthly ADR Trends

Hotel Price Fluctuations

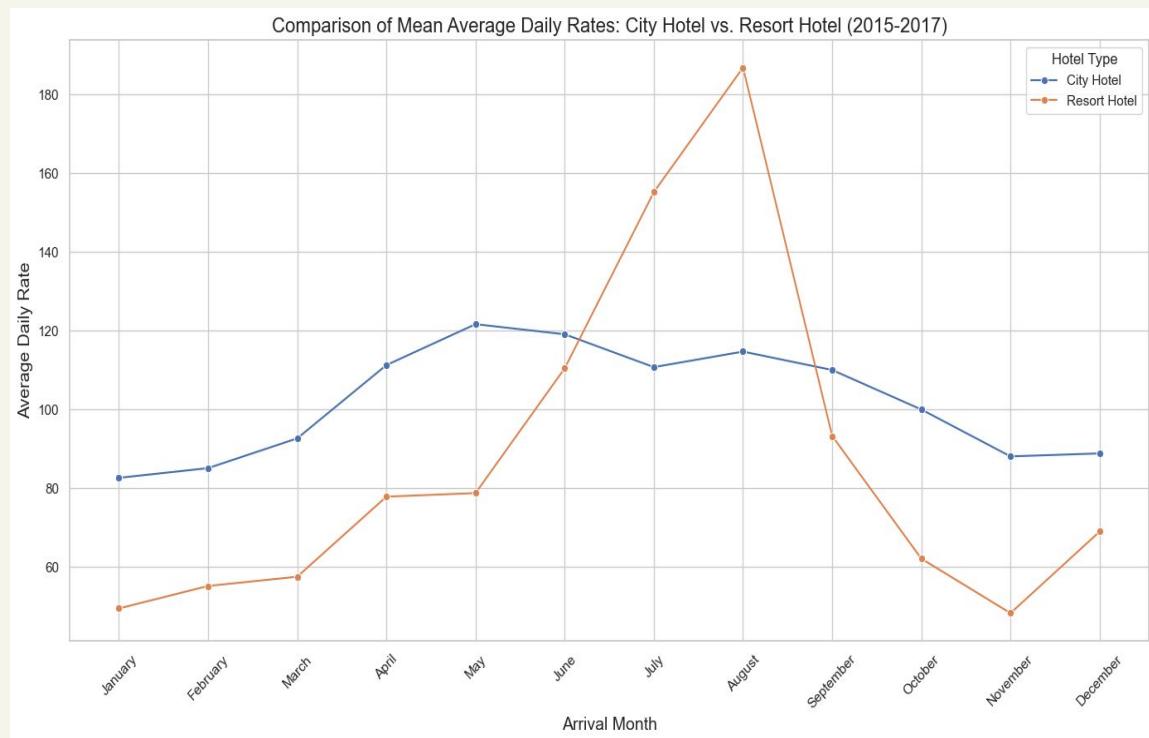
ADR: City Hotels Versus Resort

City Hotels:

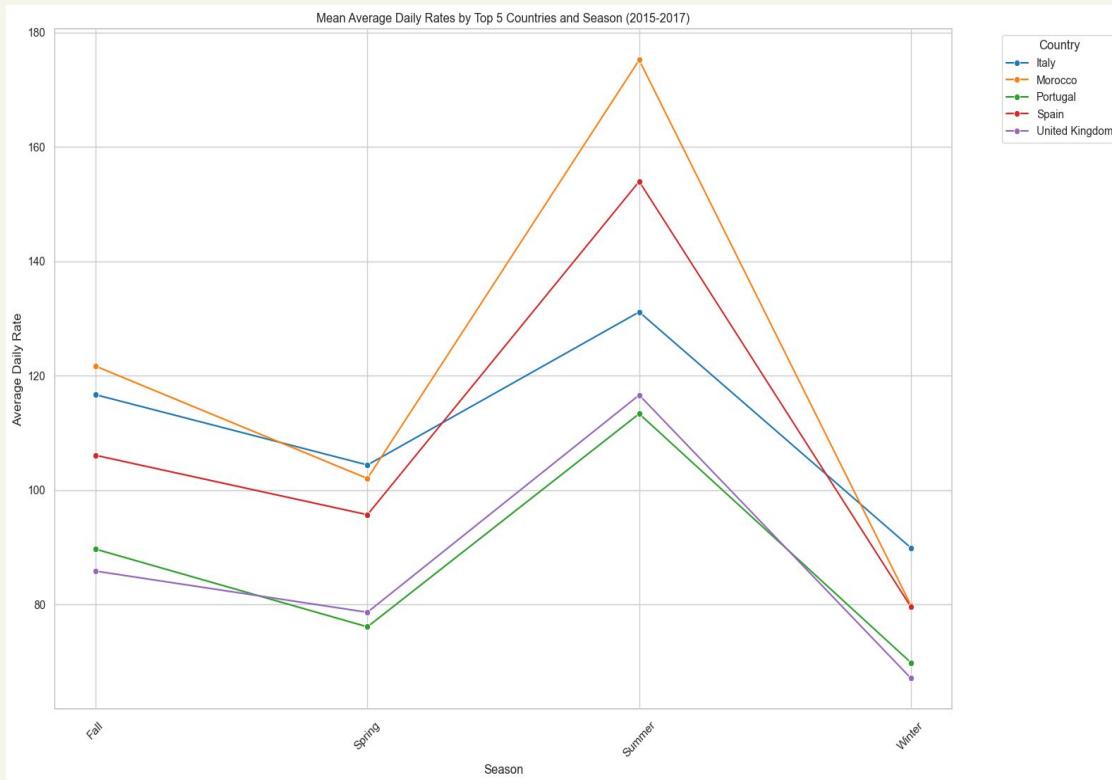
- Relatively stable ADR throughout the year.
- Peaks in April to July.
- Higher ADR in April and May, showing increased business or leisure travel.

Resort Hotels:

- More pronounced changes in ADR.
- Significant peaks in July and August, showing higher demand for leisure travel destinations.



Hotel Price Fluctuations



**Highest Mean ADR:
Top 5 Countries**

Summer Trends:

- Morocco: highest ADR 175.21
- Spain: 2nd highest at 153.98
- Italy: Competitive ADR 131.14
- UK: ADR of 116.62

Spring Trends:

- Italy leads with 104.39
- Morocco follows 102.02
- Spain and UK moderate rates (95.68 and 78.62)

Fall Trends:

- Morocco leads at 121.66
- Italy and Spain follows (116.67 and 106.08)

Winter Trends:

- Significant decline in ADRs across all countries.
- UK and Portugal lowest ADR.
- Italy highest ADR for winter.

Conclusion

Recommendations

1. Create immersive, personalized guest experiences.
2. Implement dynamic pricing strategies and targeted marketing to capitalize on seasonal demands.
3. Enhance service quality and amenities to improve hotel ratings.

- ★ Portugal leads in guest arrivals, but Ireland tops average stays.
- ★ Tourist destination popularity relies on average stay and guests nights.
- ★ Trip duration is affected by the type of group travel and hotel rates.
- ★ Strong hotel ratings directly boost bookings.
- ★ Summer sees a rise in total stays, with Morocco and Spain leading ADR.
- ★ City hotels maintain stable ADR while resort hotels show peak demand in the summer.

Thank you!

Questions & Answers?

