Hotel Booking Analysis – Summary Report

Name: Adhyatm Mishra

Dataset: hotel\_bookings.csv

Tools Used: Pandas, NumPy, Matplotlib, Seaborn, Statsmodels, Scipy

# Data Preparation:

- Dropped irrelevant columns: agent, company  
- Filled null values:  
 • children → 0  
 • country → 'PRT' (most frequent)  
- Created derived fields: total\_guests, total\_nights, total\_revenue  
- Standardized and parsed date columns (arrival\_date, reservation\_status\_date)  
- Outliers kept intentionally to preserve data integrity for business analysis

# 🔍 19 Business Questions – Insights

## 1. What influences ADR the most?

Strongest correlation: total\_revenue and total\_nights  
Guests with longer stays and more special requests tend to have higher ADR

## 2. Do guests who book earlier tend to request more changes?

Very weak correlation between lead\_time and booking\_changes  
 No practical relationship observed.

## 3. Are there pricing or booking differences across countries?

Significant ADR variation by country. Smaller or niche countries may have higher ADR.

## 4. Is there a pattern in room upgrades or reassignment?

~12.5% bookings involve room reassignment. Room Type A shows most mismatches (likely default type)

## 5. Are reserved room types consistently matched with assigned room types?

87.5% match rate. Room reassignment is relatively rare but notable.

## 6. What are the most common guest demographics?

Majority book for 2 guests (adults)  
Top nationalities: Portugal, UK, France, Spain, Germany

## 7. Do guest types influence booking behavior?

Transient guests: High ADR, high cancellations  
Contract guests: Lower ADR, longer stays, more stable

## 8. How does lead time vary by customer type and country?

Contract guests book earlier. International guests plan ahead more than local (PRT) guests.

## 9. Are longer lead times associated with fewer changes or cancellations?

Slight link to higher cancellations  
No significant effect on booking modifications

## 10. How does duration of stay vary?

Longer stays: Contract/TA customers  
Shortest: Complimentary or corporate guests

## 11. How often do guests get upgrades?

~12.5% are upgraded (reserved ≠ assigned). Possibly due to overbooking or upselling.

## 12. Are special request guests likely to modify/cancel?

Weak positive correlation  
More special requests → Slightly more changes and longer stays

## 13. Which segments or channels yield consistent revenue?

Online TA: Highest ADR but high cancellations  
Direct bookings: High ADR & reliability

## 14. Factors most correlated with higher ADR?

Top correlations:  
 - total\_revenue (+)  
 - total\_nights (+)  
 - special\_requests (+)

## 15. Which customer types/segments contribute most revenue?

Contract and Transient customers = highest revenue per booking  
Online TA and Direct channels top revenue drivers

## 16. Do early bookings or specific countries yield higher ADR?

Mid-range lead times (1–6 months) yield higher ADR  
ESP, FRA, DEU guests pay more than Portuguese

## 17. Are high ADR guests more demanding?

Positive correlation between ADR and special requests

## 18. Do guests from different countries behave differently?

Local guests (PRT) = shorter stays/lead time  
Intl. guests = longer stays/lead time

## 19. Do guests making changes also request services or cancel?

Statistically significant: Guests making changes are more likely to cancel or request services