Hotel Booking

DEPI (Data Analytics Power BI Final Project)

Group Code: ALX2_DAT2_G3
Team Members:

Rehab Ahmed Hassen
Reham Mohamed Ibrahim

Introduction:

The hotel industry is a vital part of the global economy, providing accommodation, services, and experiences to travelers for leisure and business purposes. As a customer-focused sector, hotels rely on data analysis to understand guest behavior, optimize operations, and enhance decision-making. By analyzing booking data, businesses can improve revenue strategies, identify trends, and deliver tailored services in this competitive and dynamic industry.

Data source:

- https://www.kaggle.com/datasets/mojtaba142/hotel-booking/data

Dataset Overview:

- The dataset contains hotel booking information for two hotels between the 1st of July 2015 and the 31st of August 2017 and includes booking details, customer information, and reservation status.

Main Objective of this Project:

- Provide a comprehensive overview of hotel booking performance, identify key factors influencing booking patterns and cancellation generation, and enable data-driven decision-making to optimize operations and reduce cancellations for both the City Hotel and Resort Hotel properties.

Week 1: Data cleaning and preprocessing:

Task: Data preprocessing: Data cleaning and preprocessing
Tool: Power BI
Deliverables:
Cleaned dataset ready for analysis.
Data preprocessing notebook.
Done By: Both team members.

1- Data Import into Power BI

2- Data inspection and cleaning

The dataset is a single CSV file.

- It contains hotel booking information (119390 observations) for a City hotel and a Resort hotel.
- Each observation represents a distinct hotel booking (by phone number).
- Including booking details (booking that effectively arrived and booking that were canceled), customer information, and reservation status. (119390 rows, 36 columns)

Column	Data Type	Description
Hotel	Text	The dataset contains the booking information of two hotels. One of the hotels is a resort hotel, and the other is a city hotel.
is_canceled	Whole number	A binary indicator Value indicating if the booking was canceled (1) or not (0)
lead_time	Whole number	Number of days that elapsed between the entering date of the booking into the PMS and the arrival date.
arrival_date_year	Whole number	Year of arrival date
arrival_date_month	Whole number	Month of arrival date with 12 categories: “January” to “December”
arrival_date_week_number	Whole number	Week number of the arrival date
arrival_date_day_of_month	Whole number	Day of the month of the arrival date
stays_in_weekend_nights	Whole number	Number of weekend nights (Saturday or Sunday) the guest stayed or booked to stay at the hotel
stays_in_week_nights	Whole number	Number of weeknights (Monday to Friday) the guest stayed or booked to stay at the hotel /Calculated by counting the number of weeknights
Adults	Whole number	Number of adults in the booking.
Children	Whole number	Number of children in the booking
Babies	Whole number	Number of babies in the booking
Meal	Text	The type of meal plan included in the booking (e.g., "BB" for bed & breakfast, "HB" for half board, "FB" for full board, and "SC" stands for self-catering).
Country	Text	Country of origin of the guest.
market_segment	Text	Market segment designation the booking belongs to • **Online TA**: Bookings made through online travel agencies like Booking.com or Expedia • **Direct**: Bookings made directly through the hotel's website or phone • **Corporate**: Bookings made through corporate contracts or business travel • **Offline TA/TO**: Bookings made through traditional travel agencies • **Complementary**: Complimentary stays (e.g., for hotel loyalty members) • **Aviation:** This market segment specifically refers to bookings made through airline partnerships or airport-related services. • **Groups**: market segment refers to bookings made for groups of people traveling together.
distribution_channel	Text	**Booking distribution channel.** • **"TA/TO":** (Travel Agency/Tour Operator) • **Direct**: Bookings made directly through the hotel's website, phone, or in-person • **Global Distribution Systems (GDS):** Used primarily by travel agents and airlines for booking hotel rooms alongside flights • **Corporate Contracts**: Direct agreements between hotels and companies for business travel
is_repeated_guest	Whole number	A binary indicator Value indicating if the booking name was from a repeated guest (**1**) or not (**0**)
previous_cancellations	Whole number	Number of previous bookings that were cancelled by the customer prior to the current booking
previous_bookings_not_canceled	Whole number	Number of previous bookings not cancelled by the customer prior to the current booking
reserved_room_type	Text	Code of room type reserved. Code is presented instead of designation for anonymity reasons

assigned_room_type	Text	Code for the type of room assigned to the booking. Sometimes the assigned room type differs from the reserved room type due to hotel operation reasons (e.g., overbooking) or by customer request. Code is presented instead of designation for anonymity reasons
booking_changes	Whole number	Number of changes/amendments made to the booking from the moment the booking was entered on the PMS until the moment of check-in or cancellation
deposit_type	Text	**No Deposit**: No payment is required at the time of booking **Non-refundable Deposit**: A portion of the total stay cost is paid at booking **A refundable deposit is** a type of payment required at the time of booking that can be returned to the guest under specific conditions. This deposit type balances security for the hotel with flexibility for the guest.
Agent	Whole number	ID of the travel agency that made the booking
Company	Whole number	ID of the company/entity that made the booking or responsible for paying the booking. ID is presented instead of designation for anonymity reasons
days_in_waiting_list	Whole number	Number of days the booking was in the waiting list before it was confirmed to the customer
customer_type	Text	**Group:** when the booking is associated to a group; **Transient:** when the booking is not part of a group or contract, and is not associated to other transient booking; **Transient party:** when the booking is transient, but is associated to at least another transient booking
Adr	Whole number	Average Daily Rate Per Night (Calculated by dividing the sum of all lodging transactions by the total number of staying nights)
required_car_parking_spaces	Whole number	Number of car parking spaces required by the customer
total_of_special_requests	Whole number	Total Number of special requests made by the customer (e.g., twin bed or high floor)
reservation_status	Text	**Check-Out:** customer has checked in but already departed; **No-Show:** customer did **not** check-in and didn't inform the hotel of the reason why. **Cancelled**
reservation_status_date	Date	Date at which the last status was set. This variable can be used in conjunction with the Reservation Status to understand when the booking was canceled or when did the customer checked-out of the hotel
Name	Text	Name of the Guest (Not Real)
Email	Text	Email (Not Real)
phone-number	Text	Phone number (not real)
credit_card	Text	last four digits of the guest's credit card number (not Real)

1. **Check data types**: Ensure correct data types are applied to columns.
2. Apply Trim, Clean, then remove duplicate rows for all selected columns to exclude any duplicates.
3. **Handle missing values** (as there are numerous possibilities that arise with such large number of recorded countries, agents and companies to be considered for predicting and replacing missing values, making it non practical).
    i. The country column: replace null values by Unknown; else, same values of country column
    ii. The agent and company columns: replace missing values by number **0**.
4. **Merge** arrival_date_month, arrival_date_day_of_month, and arrival_date_year in single column and renamed **arrival date**. Change data type to (**date)**.
5. **Correct reservation status date** in Power query by new conditional column:
   Correct reservation status date: = Table.AddColumn(#"Renamed Columns3", "Correct Reservation Status Date", each if [reservation_status] = "Check-Out" then [new reserv.st.date] else [reservation_status_date])
6. **Correction of reservation status for:**
    a) Check-Out guests with zero nights of stay to Day-Use
    b) No-Show guest with arrival date<> reservation status date
    In DAX, new column.

Challenges in this data:

- **Data Granularity:** The level of detail in the data will impact the types of analysis performed, for example:
  - The "Total Available Rooms" metric is critical for occupancy rate and RevPAR giving more significant analysis but is missing in this dataset.
  - The dataset lacks customer satisfaction scores directly, so inferences about satisfaction are limited.

- **Data Completeness:** Missing data can bias results, as null values in country, agent, and company columns. In this data:
  - In the country attribute: 488 missing values.
  - In the agent attribute, 16340 values are missing.
  - In the company attribute: 112593 missing values.
  - In the children attribute: 4 missing values.

- **Data inconsistency**:
  - Inconsistent reservation status date for number of Check-Out guests as compared to their nights of stay (arrival date = reservation status date).
  - Inconsistent reservation status for a single No-Show guest as compared to reservation status date earlier than arrival date.
  - Inconsistent reservation status for several Check-Out guests with zero nights of stay. (representing Day-Use guests).
  - Inconsistent some values of the is-repeated-guest column (recorded as 0) for several guests with previous bookings.

- **External Factors**: Consider external factors that may influence hotel performance (e.g., seasonality, economic conditions, local events).

- **Segmentation**: Analyze KPIs across different segments of customers, and booking channels to identify opportunities for optimization.

Week 2: Analysis Questions Phase

Tasks: Determine Data Analysis Questions:

- Determine all possible analysis questions that can be deducted from the given dataset and would be of interest to the organization's decision-makers.
- Determine possible KPIs and metrics of the hotel booking industry that can be calculated from the available data to highlight the performance and operating process of both hotels and the customer type and behavior.

Tools: Power BI.

Deliverables: A set of analysis questions that can be answered via the dataset.

Done By: Both team members.

Our Main KPIs:

Total number of bookings	Includes all bookings (all reservations made), regardless of outcome (cancellations, no-shows, check-ins), i.e., the total demand for rooms.
True Cancellation rate	Percentage of bookings canceled by guests before arrival. Use true cancellation rates to optimize policies (e.g., stricter cancellation windows).
No-Show rate	Percentage of bookings made by guest who neither arrive nor cancel their reservation. Mixing no-shows with true cancellations inflates the cancellation rate and obscures actionable insights.
True check-in rate	Percentage of bookings only that resulted in guests checking into the hotel as planned.
Day-use rate	Percentage of bookings assigned as Day-Use only.
ADR **(ADR for check-in & ADR for true cancelled)**	Average revenue earned for an occupied room on a given day. It indicated average price customers are willing to pay for a room. It focuses on occupied rooms that contributed to revenue. Calculating ADR excludes: 1. Day-Use because they do not involve overnight. 2. Canceled and No-Show Bookings because they do not generate revenue--
Average Lead time	Number of days that elapsed between the entering date of the booking into the PMS and the arrival date. It's calculated as average number of days after excluding outliers.
ALOS	reflects the actual average stay duration of guests who checked in, excluding canceled, no-show and day-use bookings.
Total guests	Those with current bookings that resulted in check-out and day use, and excluding cancelled and no-show.
Repeated guests	Retained customers are those who have made repeat bookings that resulted in a check-out or day use (completed stays) and excluded No-show and Cancelled for the given period of study.
Customer Retention Rate	Percentage of repeated guests from total guests.
Market Segment Performance	Analyzing booking metrics (check-in, cancellation and ADR) by different market segments to find out which market segments are most profitable and where to focus marketing efforts.
Distribution Channel Performance	Analyzing booking and revenue metrics (check-in, cancellation and ADR) by different distribution channels to find out which channels are most effective at driving bookings.

Other Unavailable KPIs in current data:

- **Occupancy Rate:** The percentage of available rooms that are occupied during a specific period (day, week, month, year). It'-s a measure how well the hotel is utilizing its available rooms.
- **Revenue Per Available Room (RevPAR):** A measure of a hotel's revenue-generating ability. It combines occupancy and ADR to provide a holistic view of revenue performance.
- **Customer Lifetime Value (CLTV)**: A prediction of the total revenue a customer is expected to generate during his relationship with the hotel to identify high-value customers and prioritize customer retention efforts.
- **Customer Satisfaction Score (CSAT) or Net Promoter Score (NPS):** A measure of customer satisfaction and loyalty, typically collected through post-stay surveys, that Indicates the overall guest experience and identifies areas for improvement.

Research Questions for our Hotel Booking Data:

Overall Performance and Trends

- What are the overall booking trends for the hotel over the available time period? (e.g., How do total bookings vary by month and year?)
- What is the average daily rate (ADR) and how does it fluctuate over time?
- What is the average length of stay (ALOS) and how does it correlate with other variables?

Booking Characteristics:

- How does the number of booking changes relate to the length of stay, lead time, or cancellation rate? (Do bookings with more changes tend to be canceled more often?)
- What is the average number of special requests per booking, and how does this relate to customer satisfaction or loyalty? (High number might indicate dissatisfaction).

Cancellation Analysis:

- What is the overall cancellation rate, and what factors contribute to cancellations?
- Do cancellation rates vary significantly by market segment, distribution channel, or customer type?
- Is there a relationship between lead time and cancellation rate? (Do bookings made further in advance have a higher cancellation rate?)
- Does the deposit type (No Deposit vs. Refundable vs. Non-Refundable) affect the cancellation rate?
- Are there differences in cancellation behavior between repeated and non-repeated guests?
- How do previous cancellations influence future bookings? (Is there a higher likelihood of cancellation for guests with previous cancellations?)

Customer Segmentation:

- Which market segments generate the most bookings and revenue?
- Are there differences in ADR, ALOS, or cancellation rates across different market segments?
- What are the characteristics of repeat guests, and how do they differ from new guests?
- Which countries generate the most bookings, and their booking patterns?

Distribution Channel Effectiveness:

- Which distribution channels (e.g., direct, online TA, corporate) are most effective at driving bookings and revenue?
- Are there differences in ADR or cancellation rates across different distribution channels?

Pricing and Revenue Management

- How does ADR vary by month or market segment?
- Are there opportunities to optimize pricing based on demand patterns or customer segmentation?

Does the number of days on the waiting list impact booking conversions or cancellation rates?

Operational Aspects

Is there a discrepancy between reserved and assigned room types? If so, what impact does this have on customer satisfaction?

DAX calculations:

New measures:	
Total # of booking	= COUNTROWS(hotel_booking) (COUNT& DISTINCTCOUNT here have same results as all records (all bookings) are distinct in our data.
True check-in	= COUNTX(FILTER(hotel_booking,hotel_booking[Reservation status 2]="Check-Out"), hotel_booking[Reservation status 2])
True check-in rate	= [True check-in]/[Total # of booking]
Total Day-Use	= COUNTX(FILTER(hotel_booking,hotel_booking[Reservation status 2]="Day-Use"), hotel_booking[Reservation status 2])
Day-Use rate	= [Total Day-Use]/[Total # of booking]
No-Show	= COUNTX(FILTER(hotel_booking,hotel_booking[Reservation status 2]="NO-Show"), hotel_booking[Reservation status 2])
True Canceled	= COUNTX(FILTER(hotel_booking,hotel_booking[Reservation status 2]="Canceled"),hotel_booking[Reservation status 2])
True Cancellation rate	= ([True Canceled]/[Total # of booking])
No-Show rate	= [No-Show]/[Total # of booking]
IQR for Average ADR:	QUARTILE1 = PERCENTILE.INC(hotel_booking[adr], 0.25) QUARTILE3 = PERCENTILE.INC(hotel_booking[adr], 0.75) IQR = [QUARTILE3] - [QUARTILE1] LOWER_LIMIT = [QUARTILE1] - 1.5 * [IQR] UPPER_LIMIT = [QUARTILE3] + 1.5 * [IQR]
Average ADR for check-in*	= CALCULATE(AVERAGE(hotel_booking[adr]),hotel_booking[Reservation status 2]="Check-Out",hotel_booking[adr]>=-15.77, hotel_booking[adr]<=211.07)
Average ADR for true canceled =	=CALCULATE(AVERAGE(hotel_booking[adr]),hotel_booking[Reservation status 2]="Canceled",hotel_booking[adr]>=-15.77, hotel_booking[adr]<=211.07)
IQR for lead time	QUARTILE1 Leadtime = PERCENTILE.INC(hotel_booking[lead_time], 0.25) QUARTILE3 Leadtime = PERCENTILE.INC(hotel_booking[lead_time], 0.75) IQR leadtime = [QUARTILE3 Leadtime] - [QUARTILE1 Leadtime] LOWER_LIMIT Leadtime = [QUARTILE1 Leadtime] - 1.5 * [IQR leadtime] UPPER_LIMIT Leadtime = [QUARTILE3 Leadtime] + 1.5 * [IQR leadtime]
average Lead time*	=CALCULATE(AVERAGE(hotel_booking[lead_time]),hotel_booking[lead_time]<=373, hotel_booking[lead_time]>=0)
Average lead time for true canceled	= CALCULATE([average leadtime],hotel_booking[Reservation status 2]="Canceled")
Average lead time for No-Show	=CALCULATE([average leadtime],hotel_booking [Reservation status 2]="No-Show")
Average special requests	= AVERAGE(hotel_booking[total_of_special_requests])

average Waitday	=CALCULATE(average(hotel_booking[days_in_waiting_list]),hotel_booking[days_in_waiting_list]<=0,hotel_booking[days_in_waiting_list]>=0)
Average waiting days for true canceled	= CALCULATE([average Waitday],hotel_booking[Reservation status 2]="Canceled")
Average waiting days for No-Show	= CALCULATE([average Waitday],hotel_booking[Reservation status 2]="No-Show")
ALOS	= SUM(hotel_booking[Actual Total Nights of stay])/[True check-in]
Total guests	=CALCULATE(DISTINCTCOUNT(hotel_booking[phone-number]),hotel_booking[is_canceled] = 0)
Repeated guests	=CALCULATE(DISTINCTCOUNT(hotel_booking[phonenumber]), FILTER(hotel_booking, hotel_booking[Guest type]="Repeated Guest" && hotel_booking[is_canceled]=0))
Customer Retention Rate	= DIVIDE([Repeated guests],[Total guests],0)
Top 10 check-in countries	=VAR checkin_countries =SUMMARIZE(hotel_booking,hotel_booking[Country],"checkin",[True check-in]) VAR TOP_checkin=TOPN(10,checkin_countries,[checkin],DESC) RETURN CALCULATE([True check-in],TOP_checkin)
Top 10 cancelling countries	= VAR cancelling_countries =SUMMARIZE(hotel_booking,hotel_booking[Country],"cancelling",[True Canceled]) VAR TOP_cancelling=TOPN(10,cancelling_countries,[cancelling],DESC) RETURN CALCULATE([True Canceled],TOP_cancelling)
Bookings of TOP 10 Day-Use countries	= SUMX(TOPN(10,SUMMARIZE('hotel_booking',hotel_booking[Country],"Day-Use",[Total Day-Use]),[Total Day-Use],DESC),[Total Day-Use])
Bookings of TOP 10 No-Show countries	=SUMX(TOPN(10,SUMMARIZE(hotel_booking,hotel_booking[Country],"No-Show",[No-Show]),[No-Show],DESC),[No-Show])
New column:	
Actual Total Nights of stay	= IF(hotel_booking[Reservation status 2]="Check-Out", hotel_booking[stays_in_week_nights]+hotel_booking[stays_in_weekend_nights],0) excluding nights recorded for cancelled or No-Show guests
Room changes	=IF(hotel_booking[reserved_room_type]<>hotel_booking[assigned_room_type],"Yes","No")
Reservation status 2	=IF(hotel_booking[reservation_status]="Check-Out"&& hotel_booking[stays_in_week_nights]=0 && hotel_booking[stays_in_weekend_nights]=0,"Day-Use", IF(AND(hotel_booking[reservation_status]="No-Show",hotel_booking[Arrival date]<>hotel_booking[Correct Reservation Status Date]), "Canceled",hotel_booking[reservation_status]))
Guest type	= IF(hotel_booking[previous_bookings_not_canceled]=0,"New Guest","Repeated Guest")
previous cancellation groups	= SWITCH(TRUE(), hotel_booking[previous_cancellations]=0,"0 cancellation", hotel_booking[previous_cancellations]<5,"< 5 cancellation", hotel_booking[previous_cancellations]>=5,"5+ cancellation")

lead time groups	`= IF(hotel_booking[lead_time]>=0 && hotel_booking[lead_time]<=373,`
	`SWITCH(TRUE(),`
	`hotel_booking[lead_time]<=30,"short",`
	`hotel_booking[lead_time]<=90,"medium",`
	`hotel_booking[lead_time]>90,"long",`
	`"373+ Days"),"Outlier")`

Booking changes groups	`= SWITCH(True(),`
	`hotel_booking[booking_changes]=0,"0",`
	`hotel_booking[booking_changes]<=5,"1-5",`
	`hotel_booking[booking_changes]>5,">5")`

New table:	
new table of hotel measures.	`By Enter data`

Calendar	`=CALENDARAUTO()`
	`extract new columns of Year, Month, Day and weekday.`
	Relate calendar table with hotel-booking table by two (1:*) relations as follows:
	`- Calendar date column to hotel arrival date column ------Active relationship.`
	`- Calendar date column to hotel corrected reservation status date column ---- Inactive relation-ship.`

HotelColour	`= DATATABLE("Hotel",STRING,`
	`"Colour",STRING,`
	`{{"Resort Hotel","#53670E"},`
	`{"CityHotel","#C56B62"}})`
	Relate HotelColour table with hotel-booking table by single (1:*) active relationship.
	`Add related new measure:`
	`Selectedcolor=CALCULATE(MAX(HotelColour[Colour]),FILTER(HotelColour,HotelColour[Hotel] =SELECTEDVALUE(hotel_booking[hotel])))`

*Lead times groups are filtered by outliers, such that they are not included in calculation because outliers can significantly skew categorization. Outliers above 373 days showed associations with zero or single previous cancellation, only 3 types of distribution channels: corporate, direct and TA/TO, adults <4, children and babies <3, and excluding direct customer type.

*Average was calculated for ADR after exclusion of outliers. No outliers below $-15.77 (lower limit) detected. Outliers above $211.07 (upper limit) showed association with zero or single previous cancellations, adults <5, children<4, babies <3. A single negative value revealed to belong to a case of direct booking type, multiple booking changes with upgrading of the reserved room.

Week 3: Dashboard Phase

Tasks: Build Dashboard:

Data Visualization**:** Create charts and graphs to visualize trends and patterns.

Build a Power BI dashboard that visualizes the analysis questions and calculated KPIs.

Tools: Power BI.

Deliverables: Interactive dashboard with slicers for hotel type, year and month

Done By: Both team members.

Data Visualization and dashboard:

- Line charts for trend analysis: check-in and cancellation by month and year and correlation with ADR.
- Bar charts for customer analysis.
- Pie charts for customer classification.

Overview tab:

- **Cards**: to highlight important KPIs and metrics: Total bookings, true check-in and cancellation rates, average lead times and ADR for check-in
- **Line charts**: to display the distribution of total bookings and breakdown by check-in and cancellation rates over months and years.
- **Bar charts**: provide insights into top check-in countries (and cancelling countries by tooltip) and the distribution of check-in and cancellation by week days.

Booking tab:

- **Cards**: to highlight important KPIs and metrics: True check-in and day use counts, true check-in rate, ALSO and ADR for check-in.
- **Line charts**: to display the distribution of check-in along months and years their correlation with corresponding ADR. Provided tooltip for weekday and weekend nights stay by year and month.
- **Pie chart**: represents the proportion of repeated and new guests to check-in count. Provided tooltip for booking customer analysis.
- **Bar charts**: provide insights into top check-in countries and compare the distribution of different types of check-in guests. Provided tooltip for booking customer analysis.

Cancellation tab:

- **Cards**: to highlight important KPIs and metrics: True cancelled and no-show counts, cancellation and No-show rates, ADR for cancelled and no-show, average lead times for cancelled and no-show, average waiting days for cancelled and no-show.
- **Line chart**: display the distribution of cancellation along months and years with their correlation to corresponding ADR.
- **Bar charts**: provide insights into top cancelling countries and compare the distribution of different types of cancelling guests. Provided tooltip for cancelling customer analysis 2.
- **Pie chart**: represents the proportion of different deposit types among cancelling guests. Provided tooltip for cancelling customer analysis 1.

Information tab:

Gives information about the data, main KPIs analyzed, key insights and points of improvement as recommendations.

Drill-through and tooltip tabs:

Booking customer:

- **Bar charts**: displaying booking guest analysis across: booking changes group, room changes, average lead times groups, distribution channels and market segments by corresponding ADR.
- **Pie chart**: represents the proportion of different deposit types among cancelling guests.

Cancelling customer 1:

- **Bar charts**: displaying cancelling guest analysis across distribution channels and market segments by corresponding ADR.
- **Pie chart**: represents the proportions of repeated and new guests to cancellation count.

Cancelling customer 2:

- **Bar charts**: displaying cancelling guests analysis across: room changes, previous cancellations, booking changes group, average lead times groups,
- **Pie chart**: represents the proportions of numbers of special requests among cancelling guests.

Weekday and weekend nights stay:

- **Line chart**: showing the distribution of weekday and weekend nights stay among the check-in guest by months and years.

Week 4: Final Presentation

Tasks: Final Presentation:

Prepare a report and presentation summarizing the project work, including data analysis, model development, and deployment, and identify areas for improvement.

Deliverables: Final report and presentation.

Done By: Both team members

Key insights:

In general:

- Two hotels with overall 119,390 bookings starting from 1 July 2015 till 31 August 2017, i.e. 2015 and 2017 are incomplete.
- The City Hotel had higher total bookings than Resort Hotel
- But the City has lower Check-in rate and higher cancellation rate than the Resort.
- Collectively, City had higher average lead times and average ADR than Resort.
- Both hotels showed 0 average waiting days.

Bookings (check-in):

- Portugal was the top check-in country for both hotels
- Monday was the top booking day in both hotels; while Saturday and Friday were least booking days.
- Both hotels had lower weekend nights stays than week day nights.
- Resort had longer average length of stay.
- For both 2016 was highest in check-in; while 2015 was the least.
- Both: In 2015: highest season October / off-peak season: November
- In 2016: highest season October & May/ off-peak season: January
- In 2017: highest season: May / off-peak season: June in Resort and January in City.
- Considering ADR for check-in: Both 2017 was highest in ADR, while 2016 was lowest in Resort; while in City ADR has a generally increasing trend over the 3 years starting from 2015.
- For City hotel: 2015, 2016: ADR and check-in had almost parallel curves, but 2017: ADR and check-in were parallel till the period from July to August, where they were inversely proportional (ADR increasing and check-in decreasing) and coincidentally noted that Portugal regressed as third top check-in country after Germany and United Kingdom in same period.
- So, ADR and check-in rate were almost proportionate in City hotel over the 3 years.
- This was not the case for Resort hotel. In 2015 and 2016, check-in and ADR were almost inversely proportional most of the time and proportionate at other times. In 2017, the general trend of ADR is increasing; while check-in revealed drop in June, where Portugal regressed to second top check-in country after United Kingdom in preceding month.
- Resort had higher Customer retention rate except for 2016 City was higher.
- Both hotels' guests were mainly:
    - New guests
    - Of transient type then transient party
    - TA/TO distribution channel
    - Online TA market segment then offline TA/TO
    - With **NO** booking changes and **NO** room changing groups
    - With short then long lead times group
    - Almost of **NO** deposit type

Cancellation:

- City had higher cancellation rate than Resort Hotel, associated with higher average ADR and average lead times than Resort.

- In Both PRT was highest cancelling country of guests.
- Both Sunday among highest cancellations, but Saturday and Friday were least cancelling days in Resort (as top check-out days); while Wednesday was least in City.
- For No-Show guests, both hotels had 0 waiting days and almost close average lead time, but City hotel had higher % of no-show guests and average ADR.
- Both hotels 2016 was highest cancelling rate; while 2015 was least.
- Both hotels: 2015 highest cancelling months: September & October; while least cancelling months: November & December.
  2016 October among highest cancelling months; while least cancelling month: January
  2017: different in top cancelling months; while January was among least cancelling months.
- Considering average ADR:
    - For City Hotel, ADR and cancellation rate had almost parallel curves over 3 years except for a short period (July to August in 2017) were inversely proportionate with no other corresponding factors detected.
    - For Resort Hotel: ADR and cancellation rate had almost parallel curves over 3 years (most evident in 2017).
- Both hotels cancelling guests are mainly of:
    - New guests
    - Transient type
    - **NO** deposit
    - TA/TO market segment
    - Online TA distribution channel
    - **No** booking changes or previous cancellations groups
    - **No** reserved room changes and zero waiting days.
    - **No** special request group
    - **long** lead times group and to lesser extent medium group

  (City showed high cancelling Non refund deposit and transient party customer type in 2015, both regressed to second place in following years.)

Summary: throughout the period of study

- Although City Hotel had higher total bookings than Resort Hotel, yet City Hotel manifested lower performance than Resort (lower Check-in rate, Higher cancellation & No-show rates, shorter ALOS). This performance showed significant association with higher average ADR and average lead times than Resort. Both hotels showed same distribution pattern of check-in & cancellation rates over the 3 years.
- PRT was the top booking and top cancelling country for both hotels.
- Week-end nights were least nights of stay for both hotels.
- City tends to have steadier demand throughout the year apart from end of year (winter). Resort tends to have lower demand rates during summer and winter, while higher during spring and autumn.
- Both hotels' check-in and cancellations were more related with No deposit types, new guests, Transient customer type, online TA market segment and TA/TO distribution channel collectively over the 3 years. In addition, the City Hotel showed higher cancellation rates with no-refund reservations in 2015, associated with high cancellations from transient party type. Both decreased back to 2nd place down to 2017.
- On the other hand, both hotels showed no significant relevance with other factors as previous cancellations, bookings changes, reserved room changes, length of waiting days nor number of special requests.
- Least booking customers were of contract and group types, as well as direct market segment, repeated guests, refundable and no-refund deposit, representing points of actions for possible opportunities for improvement.

Areas of improvement:

❖ **Demand-based Opportunities**	**Action**
1- **High Demand Periods**:	Dynamic pricing spikes during peak months Raise prices during peak seasons and lower them during off-peak periods to attract bookings.
2- Length of Stay (LOS) Resort guests stay longer than city hotel guests	**Minimum LOS**: Enforce minimum stay requirements during high-demand periods (e.g., festivals) to maximize occupancy. **Discounts for Longer Stays**: Encourage longer bookings during low-demand periods
3- Deposit/no refundable	**Optimize Pricing**: Offer discounts for non-refundable and refundable bookings to increase bookings.
4- Repeat Guests: Likely to pay more for personalized experiences.	Categorized pricing for loyal guests (e.g., VIP perks at higher rates).
5- **Weekend vs. Weekday**	Target higher rates for weekends.
6- Corporate/Direct Bookings:	- Offer premium pricing for corporate guests with flexible dates (less price-sensitive) and loyalty rewards for direct bookings.

❖ **Cancellation-based opportunities**	**Action**
A- **High-Risk Bookings**: 1- Guests with **long lead times** correlate with higher cancellations.	1- Offer discounts for shorter lead times (e.g., last-minute bookings) to reduce cancellations and fill unsold rooms 2-Charge non-refundable deposits/discounts for long-lead bookings. 3-Offer discounts for flexible cancellation policies to retain uncertain guests
2- Market Segment • Online Travel Agencies **(OTAs)**: Higher cancellations	Charge higher OTA rates but offer discounts for non-refundable bookings to offset cancellation risks
3- Deposit/non refundable	-Regulations for no-deposit booking -Optimize Pricing: Offer discounts for non-refundable bookings to increase bookings and reduce cancellations (esp. for city)
B- Reduce No-show	Require deposits or credit card guarantees for new guests

❖ **Customer Satisfaction**	Use reviews or survey scores (if available) or infer satisfaction from behavioral data (e.g., cancellations, repeat bookings).
❖ **General**	❖ Compensation Policy: Offer discounts, upgrades, or loyalty points for discrepancies ❖ Adopt Predictive Tools: Integrate tools like (Market Insight) to forecast demand and adjust prices accordingly.

	❖ A/B Test Pricing: Experiment with non-refundable vs. flexible rates for high-cancellation segments. ❖ Offer discounts for those on long waiting lists to reduce waiting list time and resolve overbooking: not applicable here because not an issue in our case. ❖ International bookings have higher cancellation rates: Dynamic pricing for international guests (e.g., non-refundable discounts).

N.B.:

- **PMS: Property Management System PMS:** A software that facilitates a hotel's reservation management and administrative tasks of front-office, such as booking reservations, guest check-in/checkout, room assignment, managing room rates, and billing. [1],[2]
- **Group bookings:** Bookings in which multiple rooms are reserved at once (booking 10 rooms or more per night for an agreed upon date or dates. Some others define more than 7 or more adults.
- **Transient Bookings:** Consist of individuals or groups that are occupying less than 10 rooms per night.
  They are walk-in guests, last-minute or bookers or simply people that require a very short term stay in hotel facility.
- **Third party:** Third-party hotel bookings are reservations made through intermediate platforms such as OTAs (online travel agencies), which act as a bridge between hotels and guests, allowing hotels to offer their services and guests to book conveniently.
- **"SC" stands for self-catering:** In hotel booking, this means that no meals are included in the booking, but guests have access to kitchen facilities to prepare their own meals
- **Direct Booking Channels:** By far the most valuable distribution channel for any hotel. The guest is dealing directly with hotel booking, through website, on email or over the phone with no payment commission paid.
- **Complimentary stay:** The guest received a free room night as: A loyalty program reward redemption, compensation for a service issue, a promotional offer or competition prize, or an employee benefit or friends/family rate.

Alternative resources:

- Perplexity
- ChatGPT
- Adobe Color: https://color.adobe.com
- Novypro: https://www.novypro.com/all_dashboards

References:

1. https://www.oracle.com/in/hospitality/what-is-hotel-pms/#:~:text=Traditionally%2C%20a%20hotel%20property%20management,managing%20room%20rates%2C%20and%20billing.
2. https://www.altexsoft.com/blog/hotel-property-management-systems-products-and-features/
3. https://engine.com/business-travel-guide/how-do-group-hotel-rates-work#:~:text=With%20almost%20every%20hotel%20the,agreed%20upon%20date%20or%20dates.
4. https://roompricegenie.com/transient-market-segment/#:~:text=Transient%20guests%20are%20one%20of,term%20stay%20in%20your%20facility.
5. https://insights.ehotelier.com/insights/2025/01/23/assessing-group-and-transient-bookings-which-is-better-for-your-hotel/#:~:text=Transient%20bookings%20offer%20flexibility%2C%20high,and%20long%2Dterm%20revenue%20visibility.
6. https://www.mews.com/en/blog/direct-bookings-vs-third-party reservations#:~:text=Third%20party%20hotel%20bookings%20are,and%20guests%20to%20book%20conveniently.