

Hotel Revenue Analysis



Meet Our Dedicated Team Members

Priyanshu Priya

Ravichandra Reddy

D Arun Kumar

Ritesh Musale

Saranya L.S.

Subaharsini.R

**Under guidance of
Mentor Mukilan Selvaraj**





Overview of Hotel Revenue Analysis

- Introduction
- Project Overview
- Milestone
- Data Cleaning
- Data Modeling
- Occupancy & Revenue Metrics
- Forecasting and Cancellation Trends
- Revenue Strategy Dashboard
- AI For Upselling ,Forecasting and cancellation
- Conclusion

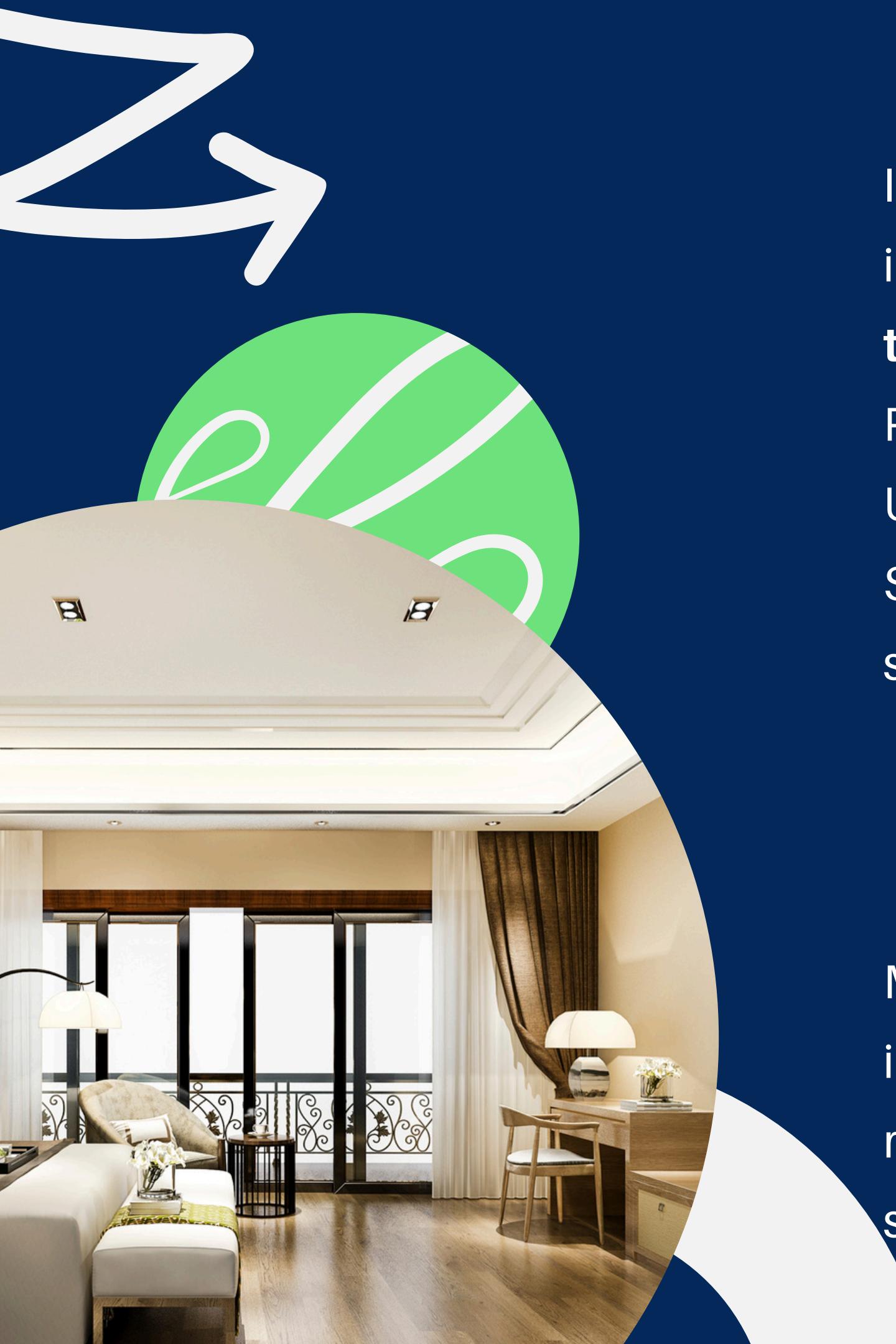
Introduction

In today's data-driven hospitality industry, success depends on how intelligently hotels interpret their data. This project **transforms traditional analysis into AI-powered hotel intelligence**, combining Python and Power BI to deliver a 360° view of performance.

Using **Python libraries** like Streamlit, Pandas, NumPy, and Statsmodels, the system analyzes and visualizes key revenue drivers such as Occupancy, ADR, and RevPAR, while AI insights uncover:

- Upsell opportunities to increase guest spending,
- Cancellation patterns to prevent revenue loss, and
- Revenue forecasts to anticipate market trends.

Meanwhile, the interactive Power BI Dashboard complements these insights with rich visuals and dynamic reports — providing management with a unified, intelligent platform for decision-making, strategy planning, and sustainable growth.





Project Overview

This project focuses on analyzing hotel data through structured modeling, visualization, and AI-driven insights. It covers end-to-end processes from data ingestion and star schema modeling to occupancy and revenue analysis, guest segmentation, and forecasting. The final outcome is an interactive Revenue Strategy Dashboard that highlights upsell opportunities, seasonal pricing strategies, and key KPIs to support data-driven decision-making for hotel managers.



Project Milestones

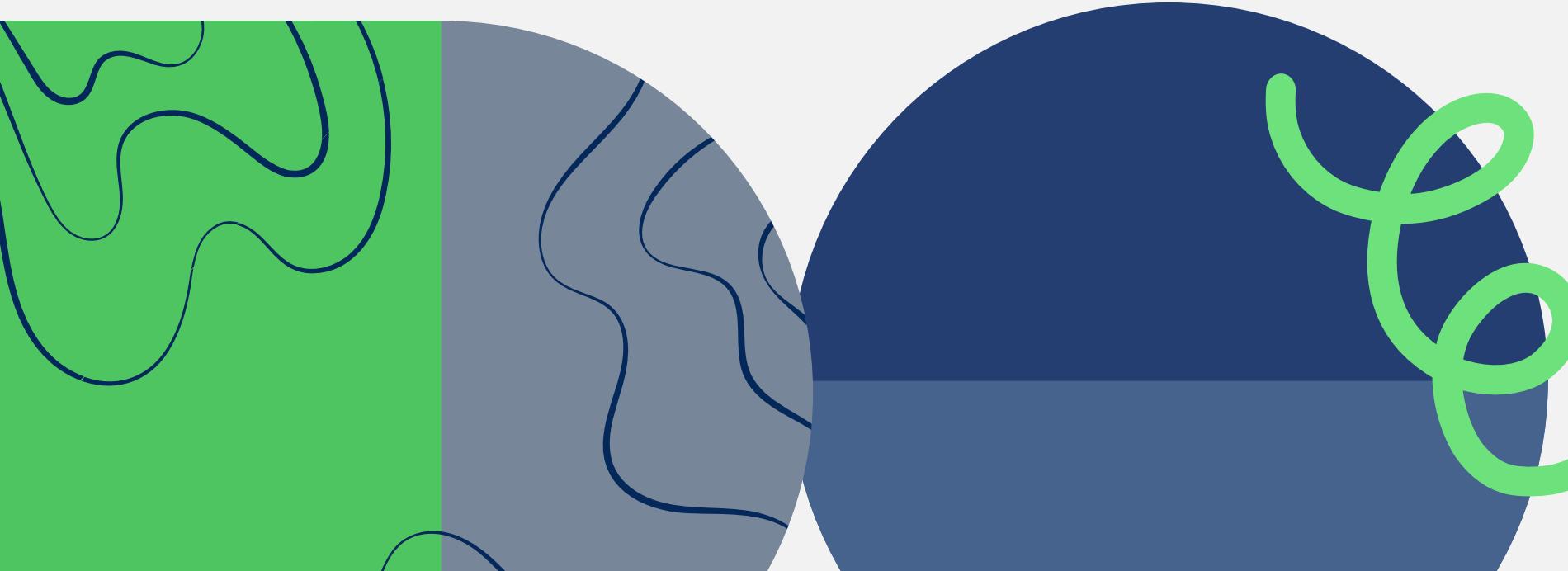
A detailed look at key phases in our project

Milestone: 1

Clean and transformed booking and room data ingested into a star schema. Relationship among customer, booking, and hotel branches are tested and validated.

Milestone:2

Occupancy % and revenue metrics (RevPAR, ADR) are visualized over time. Interactive filtering by room type, location, and booking source is fully functional



Milestone:3

Guest segmentation is operational based on purpose of visit, demographics, and loyalty. Persona insights are available in dashboard view

Milestone:4

Forecasting of future occupancy and cancellations using historical data is complete. Dashboard includes booking lead time analysis and refund/cancellation heatmaps.

Milestone:5

Revenue strategy dashboard includes recommendations for pricing tiers, upsell offers, and seasonal promotions.

Data Cleaning

Remove Columns
= Table.RemoveColumns(#"Removed Duplicates", "preferred_hotel_id")

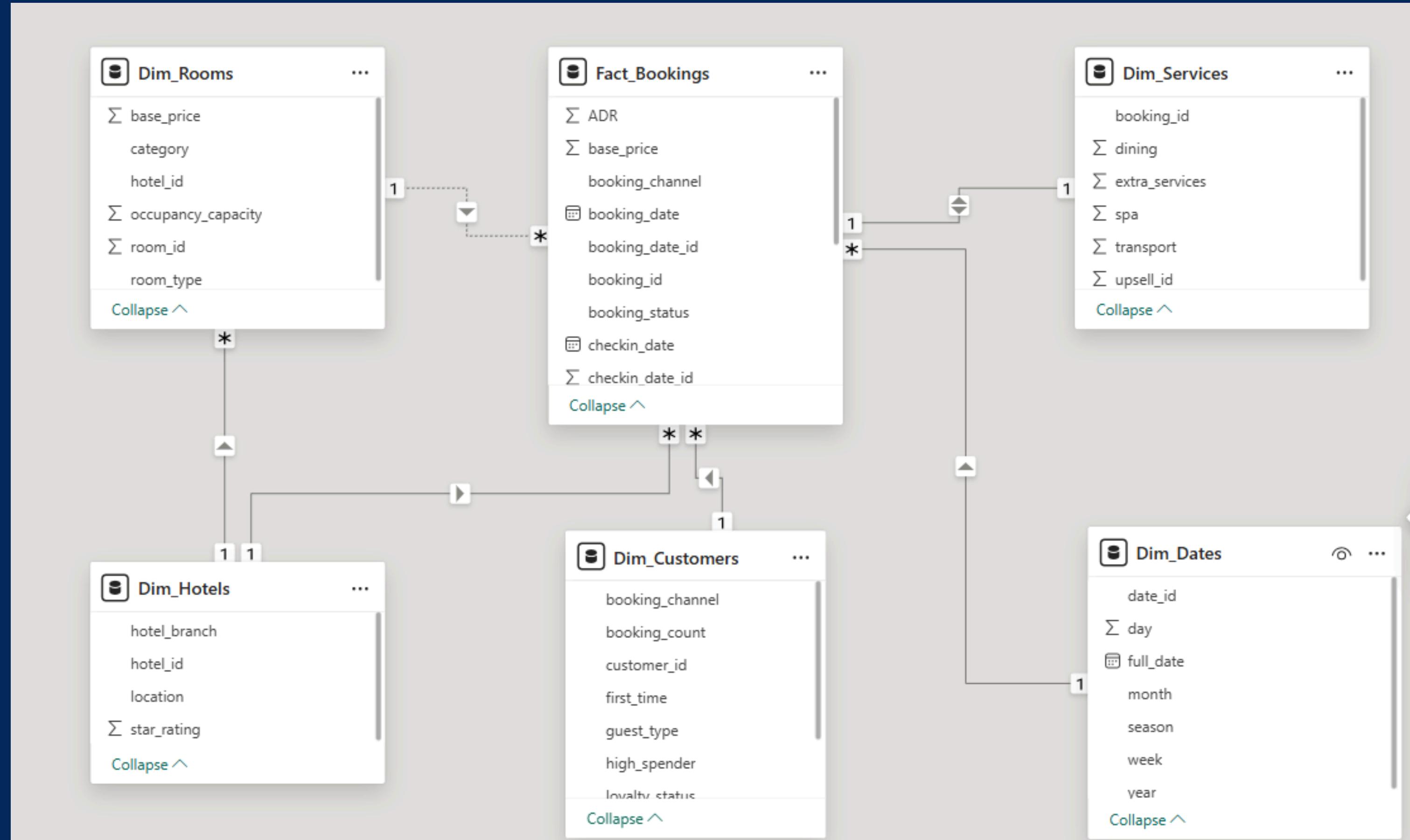
	¹ ₂ ₃ customer_id	A ^B _C guest_type	A ^B _C nationality	A ^B _C loyalty_status	A ^B _C booking_channel	¹ ₂ ₃ booking_count	¹ ₂ ₃ total_revenue
	Valid	100%	Valid	100%	Valid	100%	Valid
	Error	0%	Error	0%	Error	0%	Error
	Empty	0%	Empty	0%	Empty	0%	Empty
1	1 Family	Australia	None	OTA		1	
2	2 Business	Italy	Silver	GDS		5	
3	3 Business	India	Platinum	Direct		0	
4	4 Business	India	Silver	Direct		8	
5	5 Family	Saudi Arabia	None	OTA		3	
6	6 Business	India	Gold	Direct		6	
7	7 Solo	United Kingdom	Gold	Direct		5	
8	8 Family	Singapore	None	OTA		1	
9	9 Family	India	None	GDS		1	
10	10 Solo	India	Gold	Direct		2	
11	11 Business	China	Silver	OTA		8	
12	12 Solo	Canada	None	OTA		6	
13	13 Family	India	Silver	Travel Agent		3	
14	14 Solo	UAE	Silver	OTA		0	
15	15 Solo	United States	None	Direct		5	
16	16 Business	India	None	Direct		3	
17	17 Business	India	None	OTA		0	
18	18 Family	India	None	OTA		3	
19	19 Family	Australia	None	OTA		2	
20	20 Business	UAE	None	OTA		4	
21	21 Business	Netherlands	None	Direct		4	
22	22 Business	India	None	Direct		1	
23	23 Family	Spain	None	Direct		1	
24	24 Solo	United Kingdom	Silver	OTA		5	
25	25 Family	India	Silver	Travel Agent		2	

Dataset Introduction

- Hotel Revenue Analysis Dataset with 5 CSV files
- Fact Tables: fact_bookings
- Dimension Tables: dim_date , dim_hotels , dim_rooms
- Structured as a Star Schema for Power BI cleaning, transformation & dashboards
- Imported fact & dimension tables into Power BI
- Checked for missing values, errors & duplicates
- Standardized data types (Date, Text, Number)
- Removed redundant columns
- Ensured key consistency (hotel_id, room_id, date_id)



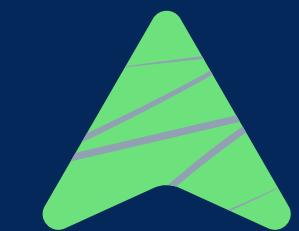
Data Modeling



Step to build star schema

- Built Star Schema with fact & dimension tables
- Linked tables via primary–foreign key relationships
- Fact: bookings, revenue, occupancy, cancellations
- Dimension: date, hotel, room details
- Model supports KPIs (ADR, RevPAR, utilization, trends)

Why Star Schema Model is Important



Simplifies Reporting

Organizes data into facts and dimensions, making it easier to understand for business users.



Supports Business KPIs

Enables calculation of metrics like ADR, RevPAR, occupancy, and cancellations across multiple dimensions



Flexibility

Allows slicing and dicing data by **time, hotel, location, or room category** for deeper insights.



Introduction to milestone 2



Hotel revenue performance is measured by three metrics: Occupancy % (rooms sold vs. available), ADR (average revenue per room sold), and RevPAR (revenue per available room).

Compute Occupancy %, ADR, and RevPAR.
Visualize them in Power BI on a single chart.
Add interactive filters (room type, location, booking source).

Dashboard



Hotel Performance Overview

ADR

Average Daily Rate

₹ 4.33K

RevPAR

Revenue per Available Room

₹ 97.56

Occupancy %

2.25%

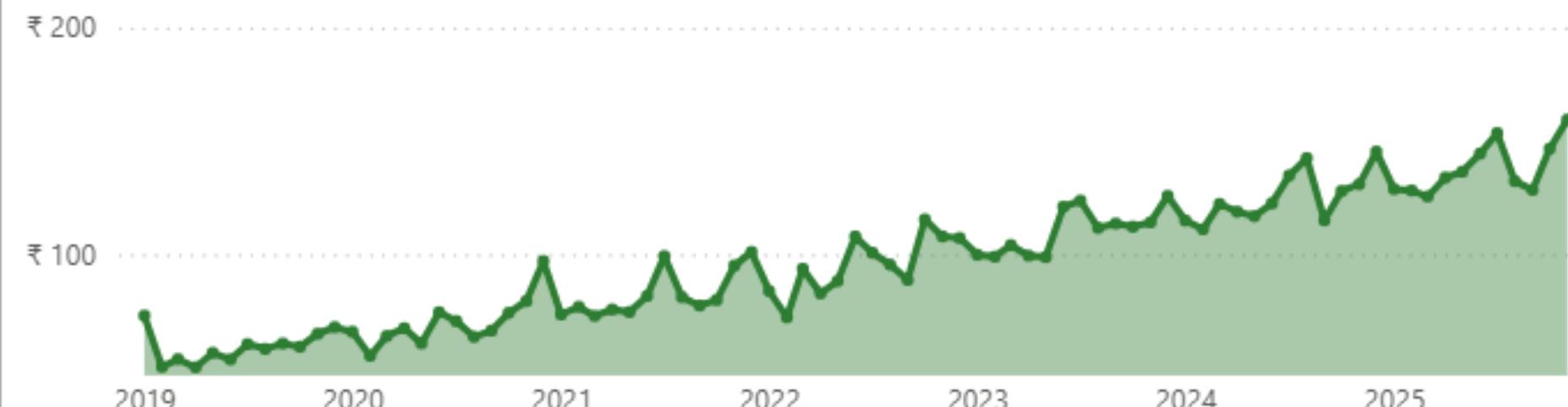
Daily Performance Trends By ADR



Daily Performance Trends By Occupancy %



Daily Performance Trends By RevPAR



Filters

Location

All

Booking Channel

- Corporate
- Direct
- GDS
- OTA
- Travel Agent

Room Type

- Deluxe
- Executive
- Family
- Standard
- Suite

Steps to achieve

Step 1: Prepare Data Model

Step 2: Create Measures

In Power BI (DAX)

Step 3: Build Visualization

Insert a Line chart.

X-axis = Days Till Completion (continuous).

Y-axis = ADR, RevPAR, Occupancy %.

Put Occupancy % on secondary axis.

Step 4: Add Interactivity

Insert slicers:

[room type][location][booking channel]

Step 5: Format Dashboard

Title: "Occupancy %, ADR, and RevPAR over Days Till Completion".



Introduction to milestone 3

The Hotel Guest Analysis dashboard segments guests by purpose of visit, demographics, and loyalty. It shows key metrics like total guests, repeat stays, average stay, and ADR, along with guest distribution by type, nationality, booking channel, and loyalty status to build quick persona insights.

Dashboard



Hotel Guest Analysis

Total Guests

19K

Repeat Guests %

84%

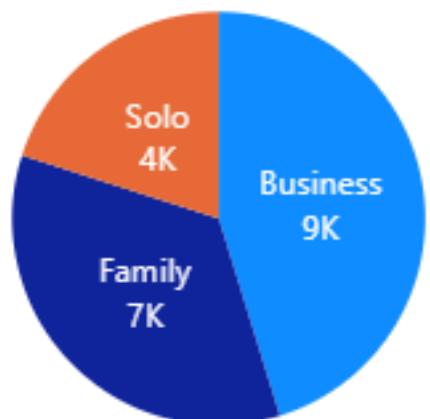
Average Stay Duration

5.45

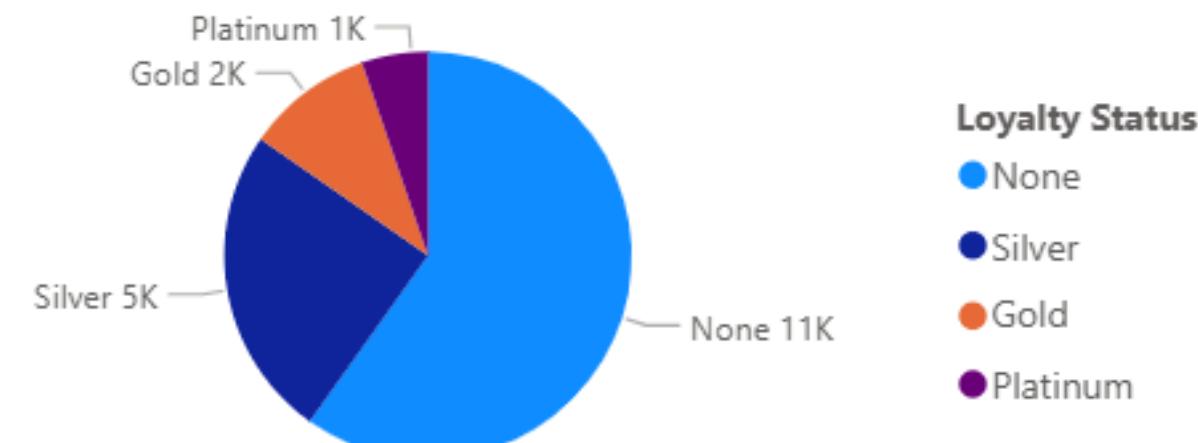
ADR by Segment

4.34K

Guest Type by Purpose of Visit



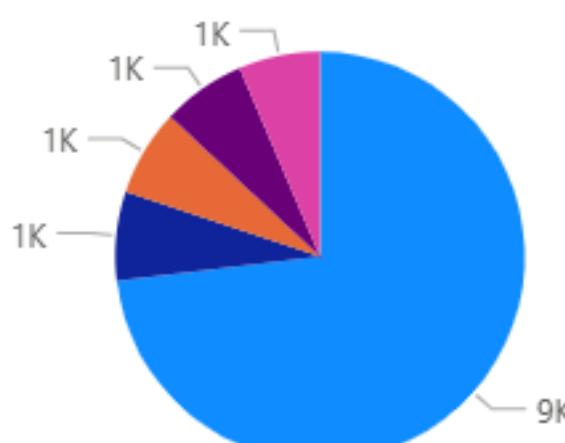
Guest Loyalty Segmentation



Filters

Year	
(Blank)	2021
2019	2022
2020	2023

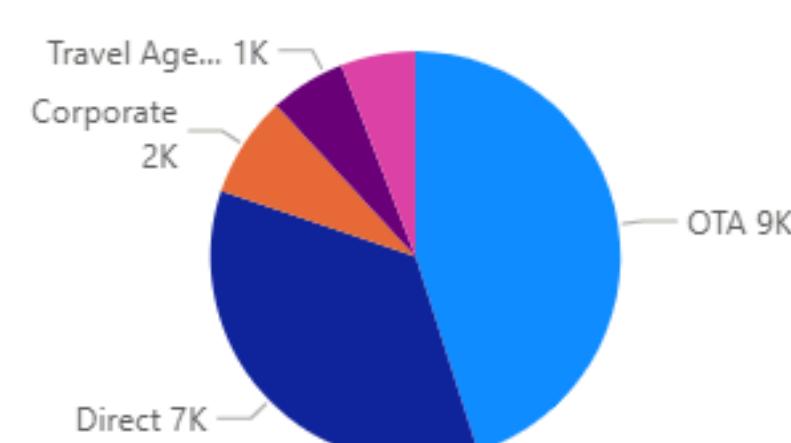
Guest Demographic by Top 5 Nationality



Nationality

- India
- United Kingdom
- Spain
- UAE
- United States

Guest Segmentation by Booking Source



Booking Channel

- Corporate
- Direct
- GDS
- OTA
- Travel Agent

Booking Channel	Value
Corporate	1
Direct	1
GDS	1
OTA	1
Travel Agent	1

Steps to achieve

1. Insert Pie Chart

2. Assign Fields

(e.g. Guest Type) → Legend.

(e.g., Guest Count) → Values.

3. Format the Pie Chart

4 . Repeat for Other Pie Charts

Copy the first pie chart and just swap the Legend field: Loyalty Status , Nationality , Booking Channel

5 . Apply Filters

Add Slicers for Year, Booking Channel, etc.



Introduction to milestone 4

The Forecasting & Cancellation Trends Dashboard is designed to analyze hotel booking performance and predict future trends. It provides a clear view of occupancy, cancellations, and booking patterns across different channels. By combining historical data with forecasting, the dashboard helps identify risk areas, optimize room utilization, and reduce revenue loss from cancellations.



Dashboard



Forecasting & Cancellation Trends



Forecasted Occupancy %

3.17%

Forecasted Cancellations

233

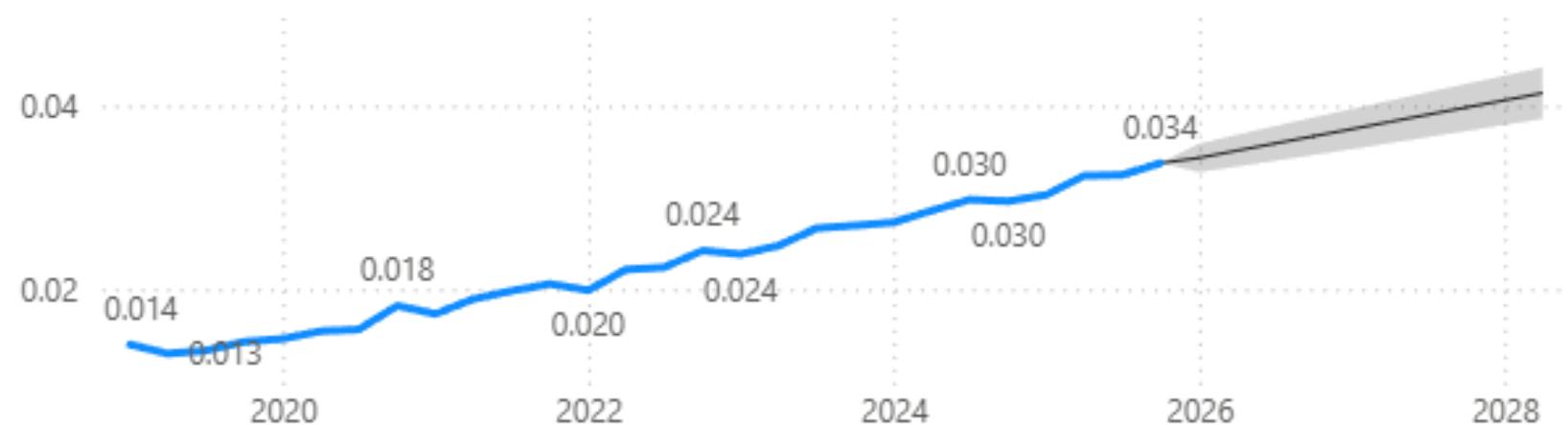
Average Lead Time (Days)

18

Cancellation Count

12K

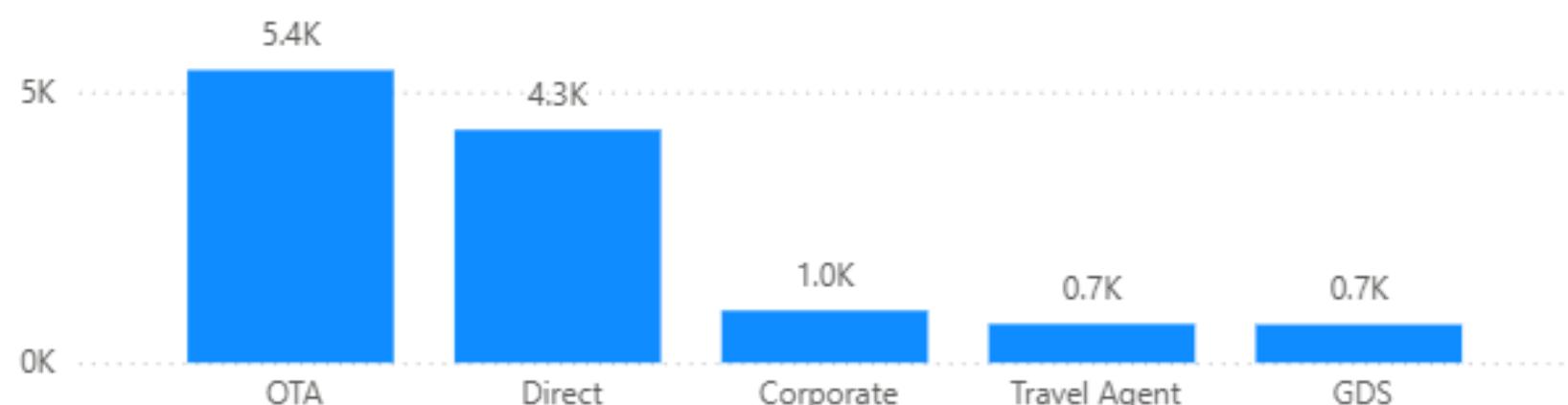
Occupancy % with Forecast



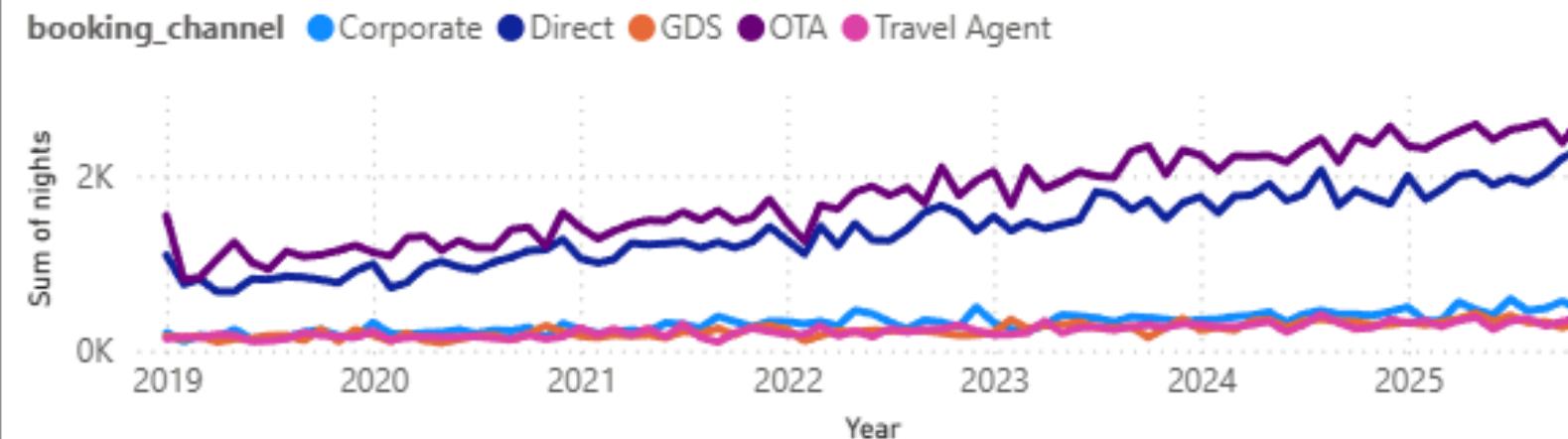
Monthly Cancellation Trends by Channel

booking_channel	1	2	3	4	5	6	7	8	9	10	11	12	Total	
Corporate	6	76	61	78	70	91	83	95	77	80	93	74	78	962
Direct	36	383	315	337	360	350	348	384	389	353	365	372	305	4297
GDS	2	46	61	51	43	69	60	61	67	64	76	55	50	705
OTA	40	471	434	391	461	433	432	470	452	483	470	476	385	5398
Travel Agent	9	62	55	58	56	62	48	57	70	60	61	64	53	715

Cancellation Count by Booking Channel



Booking Trend by Channel Over Time



Steps to achieve

- Data Model Setup – Linked Fact_Bookings with Dates, Rooms, Hotels, and Customers for structured reporting.
- Created Key Measures – Occupancy %, Forecasted Cancellations, Avg. Lead Time, Cancellation Count, and Total Nights using DAX.
- Built KPIs (Cards) – Displayed forecasted occupancy, cancellations, average lead time, and cancellation count.
- Visualized Trends – Added line charts with forecasting, heatmap for cancellations by channel, bar chart by channel, and booking trends over time.



Introduction to milestone 5

The Revenue Strategy Dashboard provides a comprehensive view of hotel performance by combining key revenue metrics with upselling and pricing insights. It highlights opportunities to maximize revenue through better channel management, guest segmentation, and seasonal pricing.

Dashboard



Revenue Strategy Dashboard

RevPar (Revenue Per Available Room)

₹ 97.56

ADR (Average Daily Rate)

₹ 4.33K

Cancellation %

64%

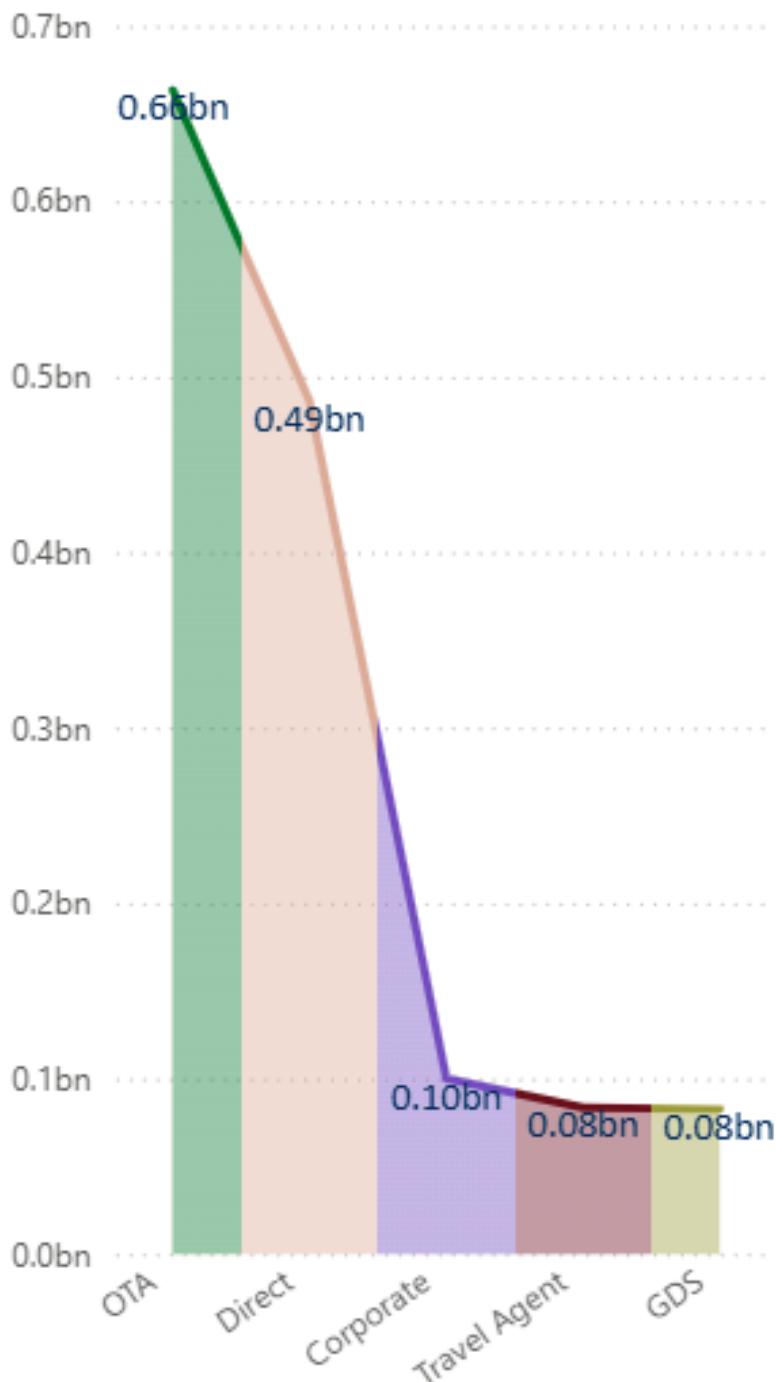
Occupancy %

2.25%

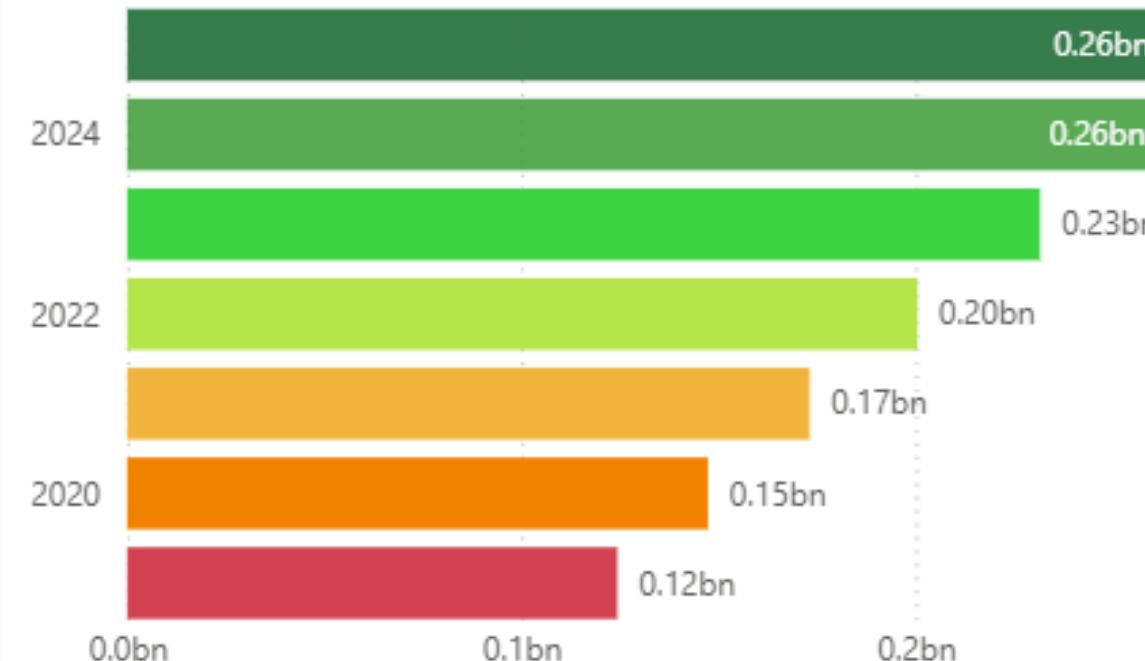
Upsell %

1.00

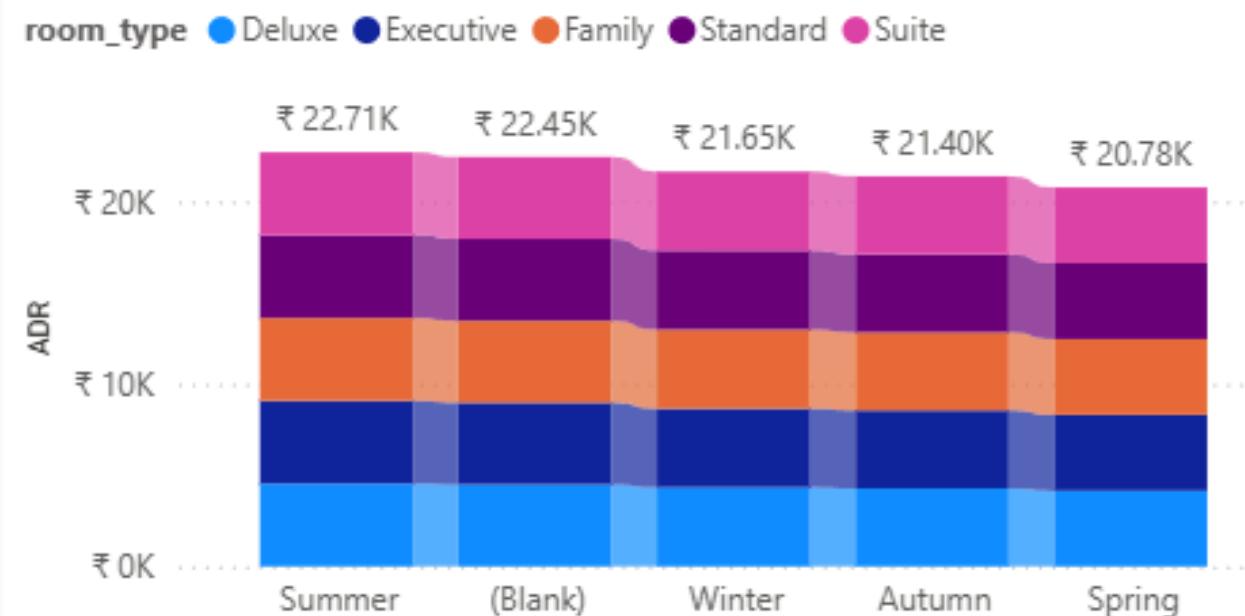
Upsell Revenue by Booking Channel



Upselling Potential(Spa, Dining, Transport)



Recommended Pricing Tiers by Season



Filters

Guest Type

- Business
- Family
- Solo

DD/WW/YYYY

- All

Hotel Branch

- All

ServiceName

- dining
- spa
- transport

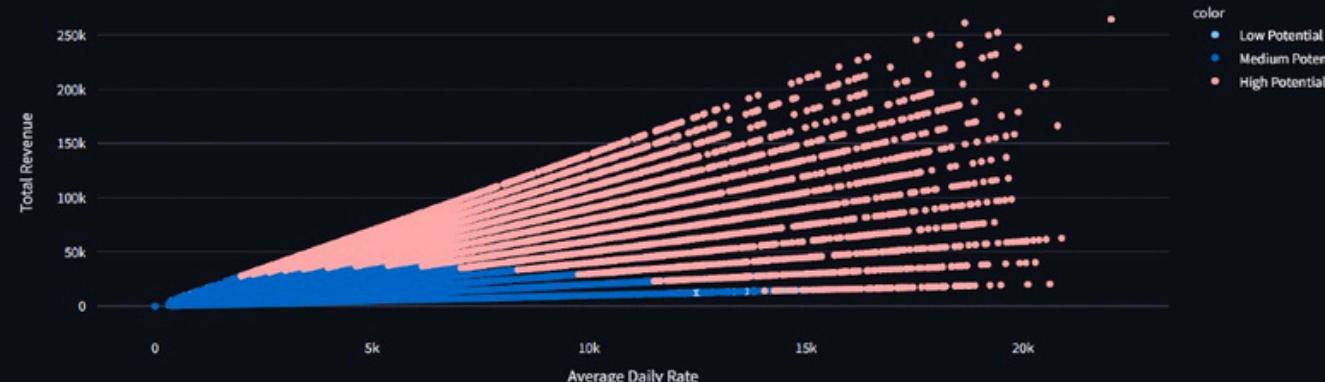
Steps to achieve

- Track Key Revenue KPIs – Monitor RevPAR, ADR, Occupancy, Cancellation %, and Upsell %.
- Analyze Booking Channels – Compare upsell revenue performance across OTA, Direct, Corporate, Travel Agent, and GDS.
- Identify Upselling Potential – Evaluate additional revenue opportunities from Spa, Dining, and Transport services.
- Optimize Seasonal Pricing – Use recommended pricing tiers by season and room type to adjust rates dynamically.
- Filter for Insights – Slice data by Guest Type, Hotel Branch, and Service Category to tailor strategies.

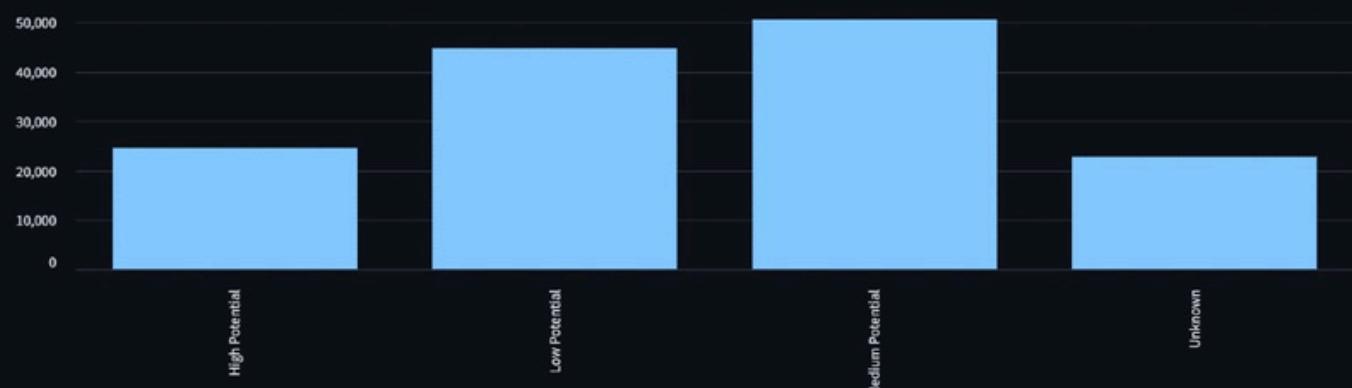
Upsell Prediction – Enhanced Guest Segmentation

Guest Segmentation Visualization

Upsell Segmentation: ADR vs Revenue



Upsell Potential Distribution



Top 10 Guests for Upselling

	customer_id	ADR	revenue	nights	upsell_label
49781	18543	22033.08	264397	12	High Potential
132614	18543	22033.08	264397	12	High Potential
2551	1201	18661.43	261260	14	High Potential
85384	1201	18661.43	261260	14	High Potential
13488	10444	19416	252408	13	High Potential
96321	10444	19416	252408	13	High Potential
14348	19891	17868.29	250156	14	High Potential
97181	19891	17868.29	250156	14	High Potential
115314	13061	19204	249652	13	High Potential
32481	13061	19204	249652	13	High Potential

Insights:

- High Potential Guests: Most profitable customers – ideal for premium offers or upgrades.
- Medium Potential Guests: May respond well to small-value add-ons (e.g., breakfast deals).
- Low Potential Guests: Budget-conscious or short-stay visitors – minimal upsell opportunity.

Upload Booking Data

Upload Excel/CSV

Drag and drop file here

Limit 200MB per file • XLSX, CSV

Browse files

Merged.xlsx

12.2MB

Navigate

- Introduction
- Upsell Prediction
- Cancellation Prediction
- Revenue Forecasting
- Power BI Dashboard

Hotel AI Dashboard

Smart Insights for Hotel Revenue Optimization

Welcome to the Hotel AI Dashboard — your one-stop solution for data-driven decision-making in hotel management.

This interactive dashboard uses Machine Learning to analyze booking, guest, and revenue data to help you make smarter, faster decisions.

Key Features

- Upsell Prediction → Identify guests most likely to accept premium offers
- Cancellation Prediction → Detect high-risk bookings before they cancel
- Revenue Forecasting → Predict next 30 days' revenue with confidence intervals
- Power BI Integration → Explore interactive visual insights

Project Objectives

- Maximize Occupancy Rate
- Optimize Average Daily Rate (ADR)
- Improve Revenue Forecasting Accuracy
- Enhance Operational Decision-Making

Data Requirements

Upload your Bookings dataset (CSV or Excel) to start exploring insights.

The dataset should ideally include:

- checkin_date, checkout_date
- lead_time, ADR, nights
- booking_status, revenue, and guest details

Quick Data Preview

	booking_id	customer_id	room_id	hotel_id	checkin_date	checkout_date	checkin_date_id	booking_date	booking_date_id	lead_time	nights
0	1	4174	4662	61	2024-01-21 00:00:00	2024-01-23 00:00:00	1847	2023-12-29 00:00:00	1824	23	2
1	2	4346	215	3	2024-09-05 00:00:00	2024-09-08 00:00:00	2075	2024-08-20 00:00:00	2059	16	3
2	3	11024	1612	22	2020-02-02 00:00:00	2020-02-04 00:00:00	398	2020-01-16 00:00:00	381	17	2
3	4	8408	5390	71	2024-03-13 00:00:00	2024-03-20 00:00:00	1899	2024-02-19 00:00:00	1876	23	7
4	5	7313	5552	73	2022-06-18 00:00:00	2022-06-28 00:00:00	1265	2022-05-25 00:00:00	1241	24	10



Cancellation Prediction – Smart Booking Risk Model

Model Accuracy: 0.94

Classification Report:

precision recall f1-score support

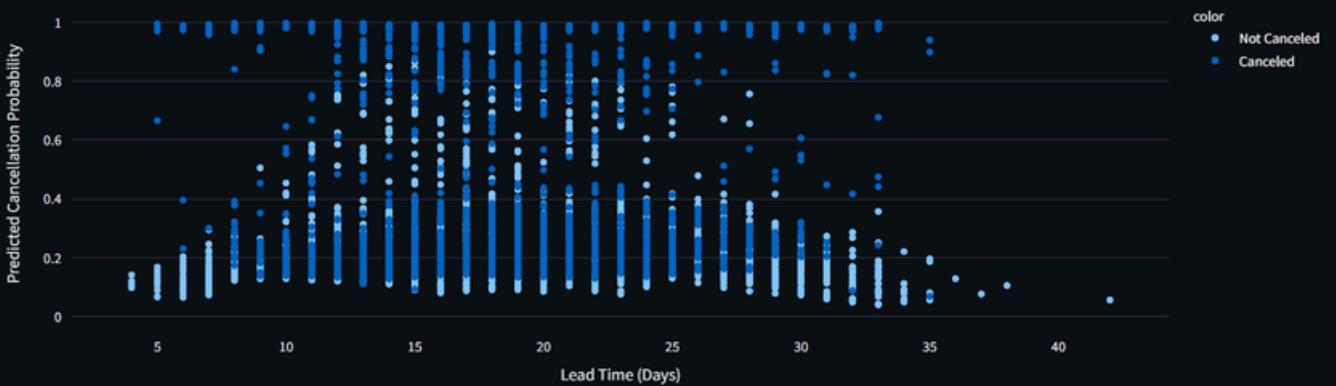
0	0.92	1.00	0.96	17886
1	0.99	0.76	0.86	6114

accuracy	0.94	24000		
macro avg	0.96	0.88	0.91	24000
weighted avg	0.94	0.94	0.93	24000

Feature Importance



Cancellation Probability vs Lead Time



Predict Cancellation for New Booking

Lead Time (days)

ADR (Average Daily Rate)

Number of Nights

Predicted Cancellation Probability: 99.0%

High Risk: Recommend prepayment or reminder.

Hotel Dynamic Pricing Model

Objective:

Automate hotel room pricing using historical booking data to predict demand and analyze optimal prices for maximizing revenue.



What We Used

- Python – Data processing & model pipeline

Libraries:

- pandas, numpy – data cleaning & feature creation
- scikit-learn – train-test split
- LightGBM – demand prediction model
- Power BI – visualize final recommended prices



Key Steps

- Data Preparation: Loaded and normalized booking data from CSV; parsed dates, created features like lead time & stay length.
- Feature Engineering: Converted booking info to numeric demand.
- Modeling: Trained a LightGBM regression model on rate, lead time, stay length, hotel & room IDs.
- Price Optimization: Simulated ±10% price changes → selected the price with highest expected revenue.
- Integration: Exported results as pricing_recommendations.csv for use in Power BI dashboards.

Conclusion

- Forecasting and cancellation predictions enable proactive demand management and minimize revenue loss.
- Guest segmentation identifies loyal customers, first-time visitors, and high spenders for targeted marketing.
- The Revenue Strategy Dashboard uncovers upselling opportunities and recommends optimal seasonal pricing.
- Combines predictive analytics and artificial intelligence to deliver actionable insights, empowering managers to make data-driven decisions and ensure long-term growth.
- Integrated AI-driven data modeling and analysis provide a complete view of hotel performance.

**Thank You
for Your
Attention!**

