

# 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.



# Tech used

## Programming Language

Python – Used as the primary language for building data processing, analysis, and machine learning models.

## Data Processing & Analysis

- Pandas — For loading, cleaning, transforming, and analyzing hotel booking and revenue datasets.
- NumPy — For numerical computations and array operations.
- StandardScaler (Scikit-learn) — For feature scaling and normalization before applying ML models.

## Machine Learning & Forecasting

- Scikit-learn-for machine learning tasks like classification, regression, and clustering.
- KMeans — trained clustering model
- train\_test\_split, classification\_report, accuracy\_score — For training/testing and evaluating models.

Statsmodels (Holt-Winters)

- ExponentialSmoothing —

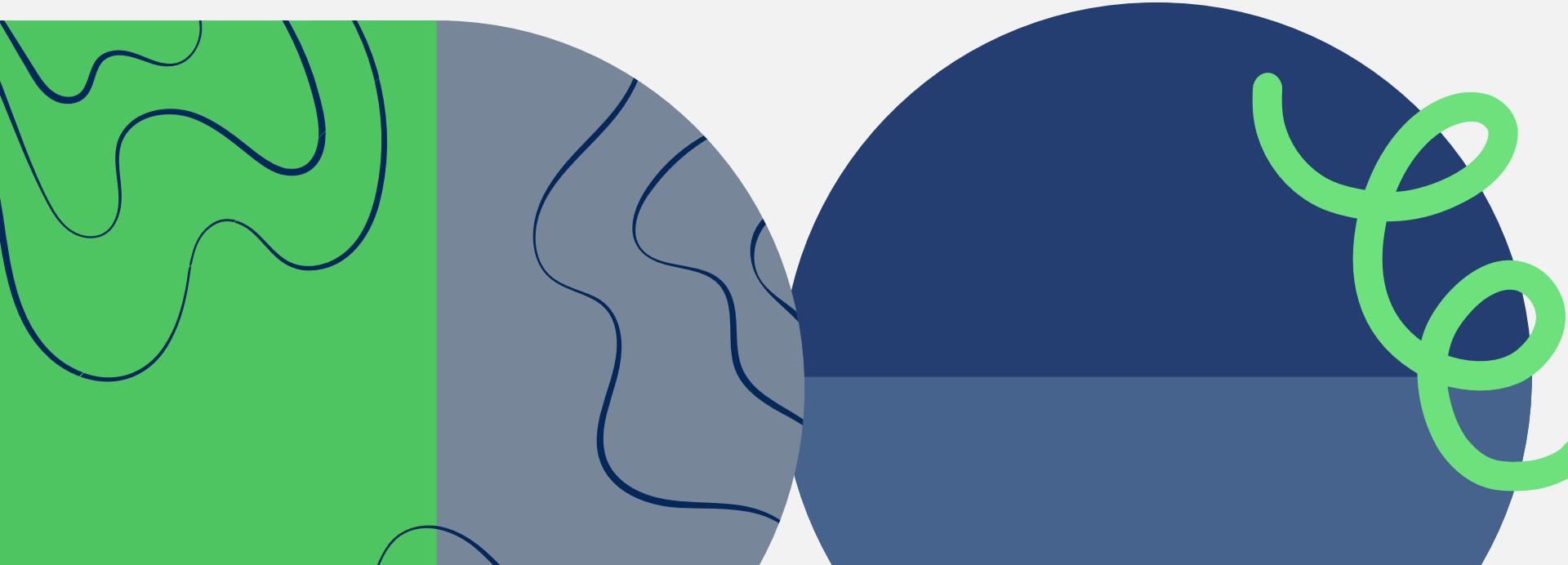
- Streamlit — Used to build interactive web applications directly in Python
- Matplotlib —Used for creating charts and graphs.

Streamlit Components

Power BI Embed —For embedding Power BI

## Power BI/Excel

For advanced visual analytics and business intelligence integration.





# 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

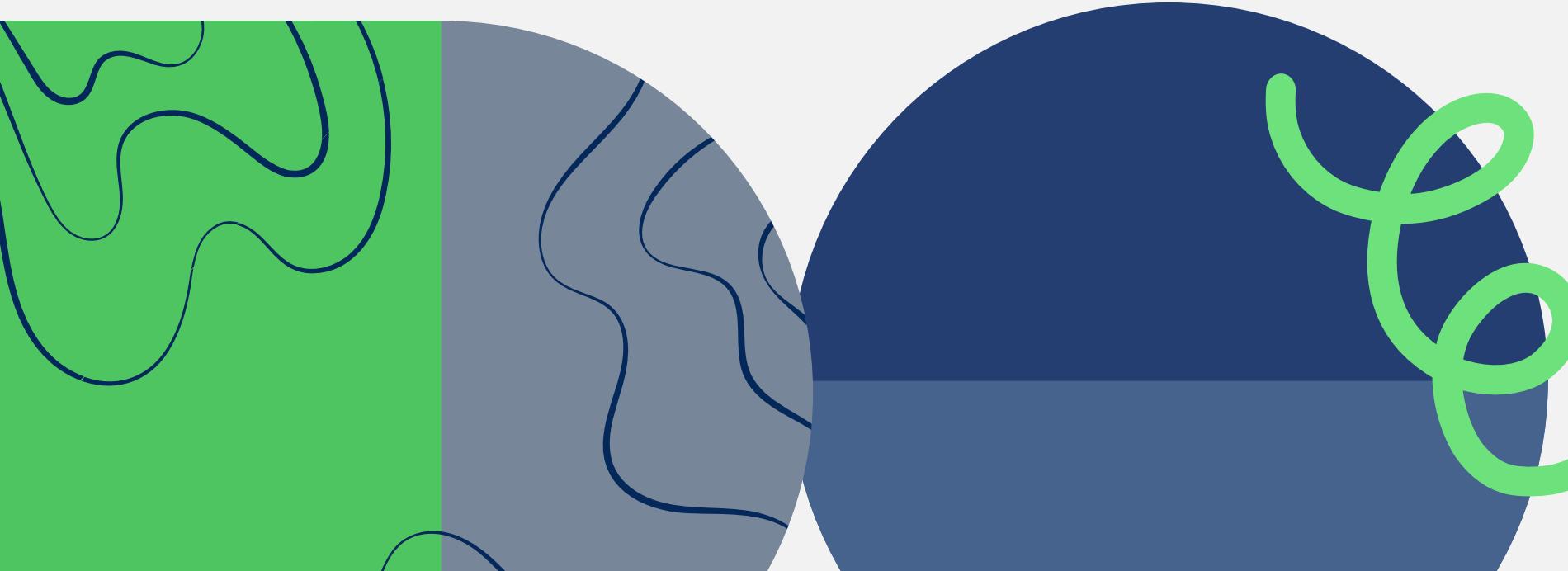
A detailed look at key phases in our project

## 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.



# Milestone 1

- 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

## 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)





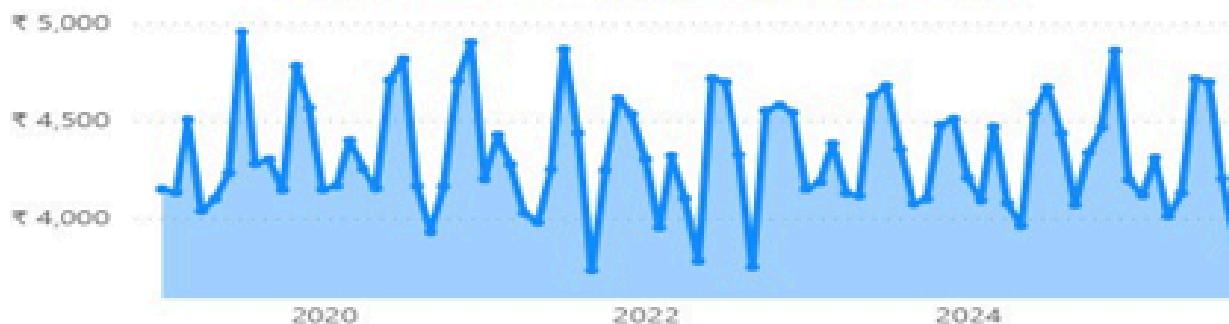
# Hotel Performance Overview

ADR  
Average Daily Rate  
**₹ 4.33K**

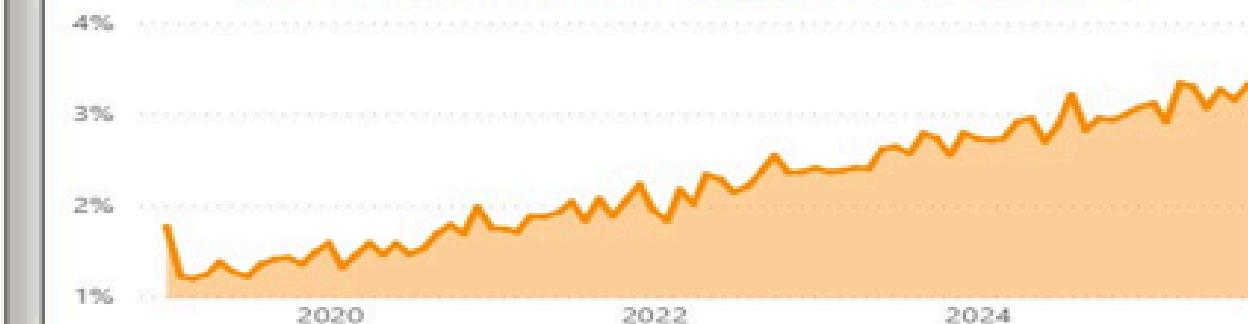
RevPAR  
Revenue per Available Room  
**₹ 97.56**

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



# 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



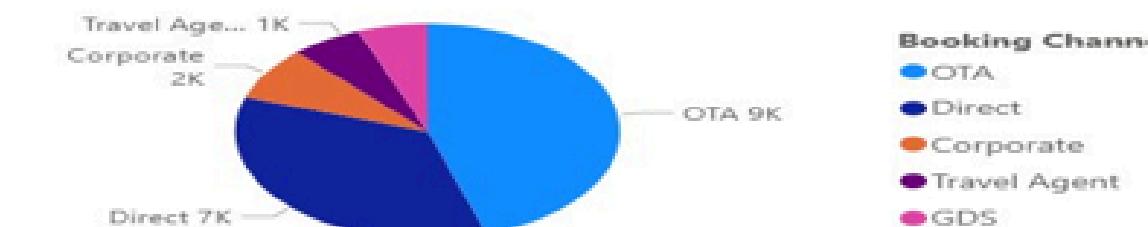
Guest Loyalty Segmentation



Guest Demographic by Top 5 Nationality



Guest Segmentation by Booking Source



## Filters

### Year

(Blank)	2021
2019	2022
2020	2023

### Booking Channel

- Corporate
- Direct
- GDS
- OTA
- Travel Agent

# milestone 2 and 3

## Revenue Health Metrics

- KPIs: Visualize ADR (₹4.33K), RevPAR (₹97.56), and Occupancy % (2.25%) trends over time.
- Functionality: Interactive filtering by Room Type and Booking Channel for instant segmentation of revenue drivers.

## Customer Insights

- Loyalty: 84% Repeat Guests.
- Segmentation: Identified key guest clusters by Purpose of Visit (e.g., Business, Family) and Loyalty Status (e.g., Platinum, Gold).
- Value: Understanding guest behavior is the foundation for the AI's Upsell Prediction Model.



# Forecasting & Cancellation Trends



Forecasted Occupancy %

**3.17%**

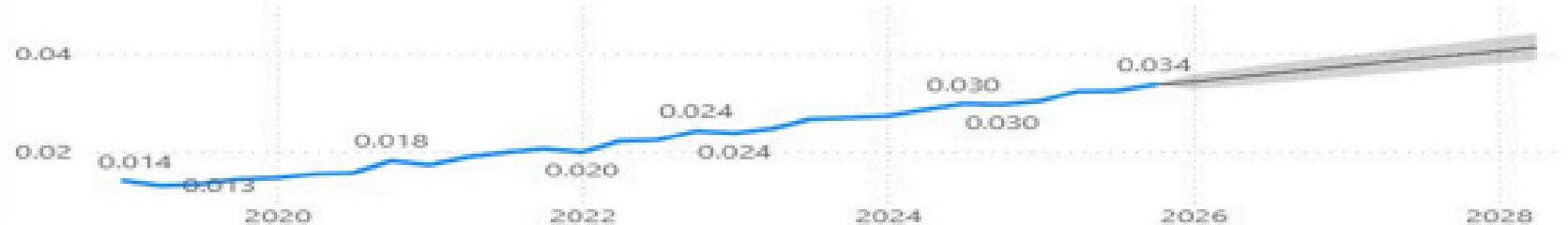
Forecasted Cancellations

**233**

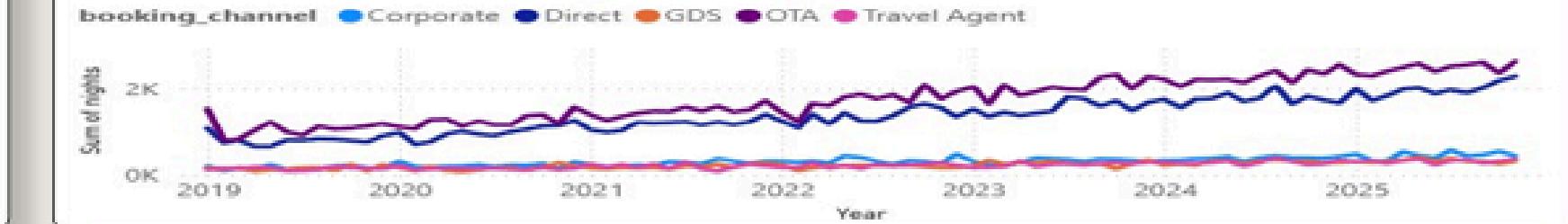
Average Lead Time (Days)

**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
Direct	36	383	315	337	360	350	348	384	389	353	365	372	305
GDS	2	46	61	51	43	69	60	61	67	64	76	55	50
OTA	40	471	434	391	461	433	432	470	452	483	470	476	385
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****Revenue Strategy Dashboard**

RevPar (Revenue Per Available Room)

**₹ 97.56**

ADR (Average Daily Rate)

**₹ 4.33K**

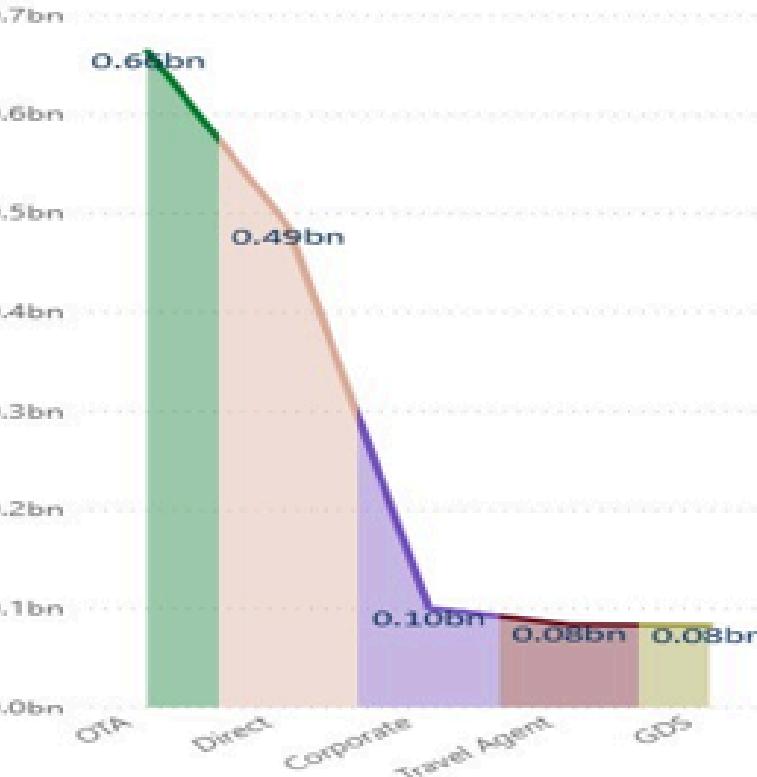
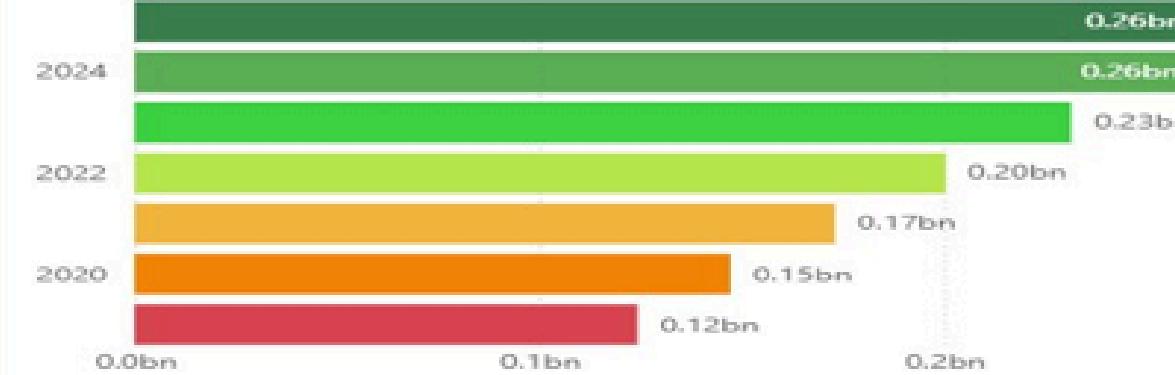
Cancellation %

**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/MM/YYYY

Hotel Branch

ServiceName  
 dining  
 spa  
 transport

# milestone 4 and 5

## Forecasting & cancellation trend

- Forecasting: Displays Forecasted Occupancy and Cancellation Trends (by channel/season).
- Loss Prevention: Cancellation Heatmap highlights high-risk periods/channels, enabling proactive retention efforts.

## Revenue strategy dashboard

- Pricing: Outputs Recommended Pricing Tiers adjusted by Season and Room Type from the Dynamic Pricing Model.
- Upsell: Quantifies Upselling Potential (ancillary services) and compares performance across Booking Channels.

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

Deploy :

## Revenue Forecasting – Smarter Revenue Trend Analysis

Avg. Revenue (Last 30 Days) ₹1,777,797

Projected Growth (Next 30d) 0.9%

Peak Forecast Revenue ₹1,900,922

### 30-Day Revenue Forecast

The chart displays a highly volatile line representing historical revenue from 2019 to 2026. A red line represents the projected growth for the next 30 days, showing a slight upward trend. The Y-axis is labeled 'Revenue (₹)' and ranges from 0 to 3M. The X-axis is labeled 'Date' and shows years from 2019 to 2026. A button at the bottom right says 'View Forecast Data'.

### Upsell Prediction – Enhanced Guest Segmentation

#### Guest Segmentation Visualization

##### Upsell Segmentation: ADR vs Revenue

This scatter plot illustrates the relationship between Average Daily Rate (ADR) on the X-axis and Total Revenue on the Y-axis. The data points are color-coded by guest segment, showing a positive correlation where higher ADR generally leads to higher total revenue. The X-axis ranges from 0 to 20k, and the Y-axis ranges from 0 to 250k.

# 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!**

