

# AI DRIVEN REVENUE ANALYSIS FOR HOTELS

# Team Members

Garv Jain

Hima Rose George

Nabamita Samui

Punna Gayathri

Rajshree Ghosh

Syed Anwar Rizvi

Under Guidance of  
Mentor Mukilan Selvaraj

# OVERVIEW

- INTRODUCTION
- PROJECT OVERVIEW
- MILESTONES
- DASHBOARD DEVELOPMENT
- DASHBOARD SNAPSHOTS
- INSIGHTS & FINDINGS
- CONCLUSION



# INTRODUCTION



- The objective is to give hotels real-time visibility into revenue drivers using modern BI and strong data models.
- The platform tracks Occupancy %, ADR, RevPAR, and guest segmentation, enabling branch-wise comparisons and forecasting of occupancy and revenue from historical data.
- By visualizing booking patterns and demographics, managers can find upselling opportunities, run targeted marketing, and adjust seasonal and room-type pricing for profitability.
- The solution scales across hotel chains and geographies, helping management make faster, smarter decisions with reliable insights.

# PROJECT OVERVIEW

## KEY FEATURES

- Tracks occupancy, guest demographics, and pricing performance
- KPI dashboards: Occupancy %, ADR, RevPAR, seasonal trends
- Guest segmentation, forecasting, and cancellation analysis
- Strategic insights for pricing, promotions, and upselling
- Traveler personas: business, family, solo, first-timers, loyal, high spenders
- Visuals on nationality, booking source, stay duration, loyalty & spend patterns

# MILESTONE 1

- Cleaned & transformed booking, customer, and room data
- Designed star schema with Fact & Dimension tables for reporting
- Ensured consistency for occupancy, revenue & guest dashboards

# MILESTONE 2

- Built dynamic Power BI dashboard for KPIs (Occupancy %, ADR, RevPAR)
- Added interactive filters: room type, location, booking source
- Enabled managers to drill into segment-specific insights

# MILESTONE 3

- Guest segmentation by purpose, demographics & loyalty (Business, Leisure, Groups, Repeat)
- Persona insights connect profiles with ADR, Occupancy & RevPAR
- Supports optimized pricing, promotions & loyalty strategies

# MILESTONE 4

- Built forecasting module for occupancy trends using historical booking data
- Analyzed cancellation behavior with lead time distributions & no-show patterns
- Designed visualizations: occupancy trend lines, cancellation heatmap, refund bar charts
- Enabled managers to anticipate demand and mitigate revenue loss from cancellations

# MILESTONE 5

- Developed revenue strategy dashboard for pricing & upselling opportunities
- Integrated upsell potential estimation (ancillary services, dining, packages)
- Implemented seasonal pricing tiers by room type & booking channel
- Added recommendation visuals for managers to optimize promotions & offers

# DATASET OVERVIEW

## Key Information:

- Source: Booking data from a city hotel and a resort hotel.
- Timeframe: July 2015 to August 2017.
- Record Count: Approximately 119,390 hotel bookings.
- Key Goal: Understand hotel booking trends, guest behavior, and factors influencing booking cancellations.

## Core Data Points:

DimAgentCompany	...
<input type="checkbox"/> $\sum$ agent	
<input type="checkbox"/> $\sum$ company	
<input type="checkbox"/> index	
<input type="checkbox"/> $\text{Occupancy\%}$	
<input type="checkbox"/> $\text{RevPAR}$	
<input type="checkbox"/> $\text{Total Customers}$	

DimCustomer
<input type="checkbox"/> country
<input type="checkbox"/> customer_type
<input type="checkbox"/> distribution_channel
<input type="checkbox"/> Guest Segment
<input type="checkbox"/> index
<input type="checkbox"/> is_repeated_guest
<input type="checkbox"/> market_segment
<input type="checkbox"/> previous_bookings_not_canceled
<input type="checkbox"/> previous_cancellations

DimMealRoom
<input type="checkbox"/> index
<input type="checkbox"/> meal
<input type="checkbox"/> Room Preference

FactBookings
<input type="checkbox"/> $\sum$ adr
<input type="checkbox"/> $\text{ADR Measure}$
<input type="checkbox"/> $\sum$ adults
<input type="checkbox"/> $\sum$ arrival_date_day_of_month
<input type="checkbox"/> arrival_date_month
<input type="checkbox"/> $\sum$ arrival_date_week_number
<input type="checkbox"/> $\sum$ arrival_date_year
<input type="checkbox"/> assigned_room_type
<input type="checkbox"/> $\sum$ babies
<input type="checkbox"/> $\sum$ booking_changes
<input type="checkbox"/> $\sum$ children
> <input type="checkbox"/> Date
<input type="checkbox"/> $\sum$ days_in_waiting_list
<input type="checkbox"/> deposit_type
<input type="checkbox"/> hotel
<input type="checkbox"/> index
<input type="checkbox"/> $\sum$ is_canceled



# POWER BI DATA PREPARATION & TRANSFORMATION

## Introduced Key Metrics:

- A custom column was added to calculate the total length of a guest's stay (total\_stay\_nights) by summing the stays\_in\_weekend\_nights and stays\_in\_week\_nights columns. This created a new metric for deeper analysis.

## Merged Columns:

Consolidated three separate date columns (arrival\_date\_year, arrival\_date\_month, arrival\_date\_day\_of\_month) into a single, unified arrival\_date column for efficient time-series analysis.

## Handled Missing Values:

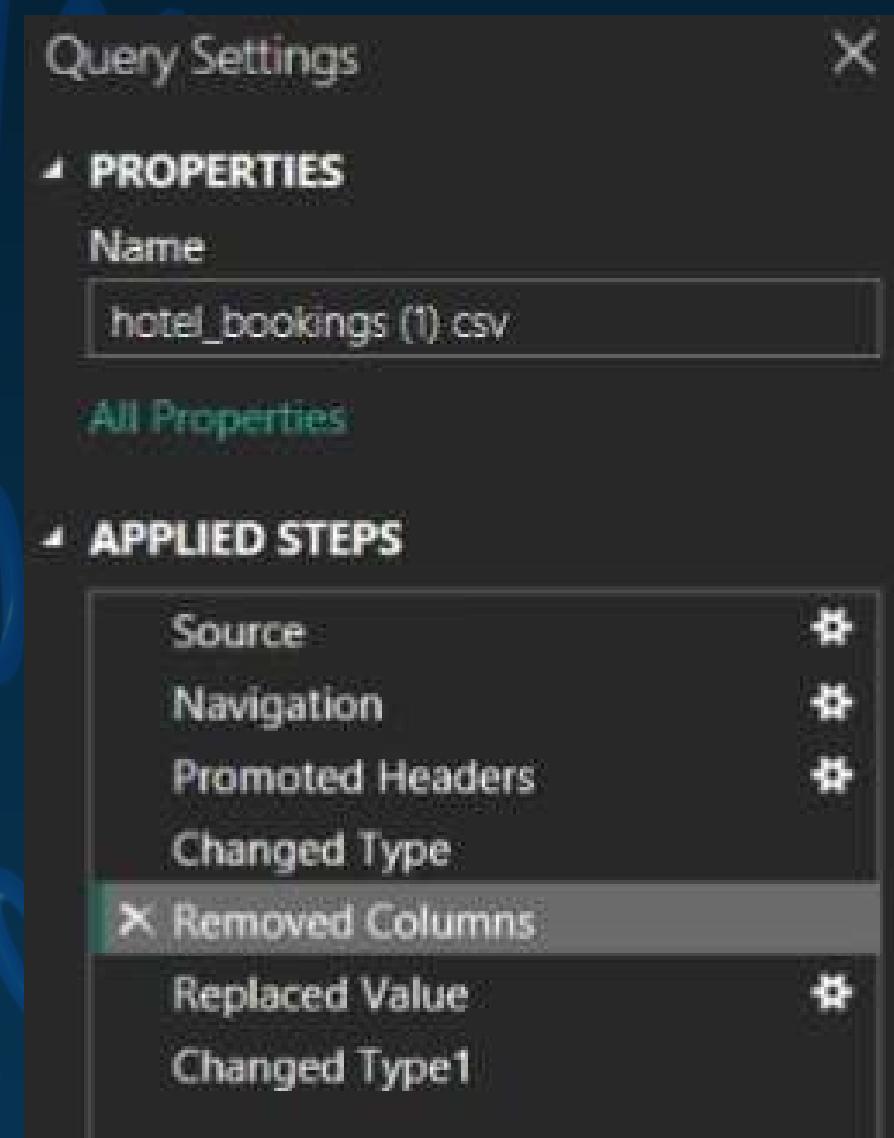
Addressed missing data in the country column and arrival\_date\_month and market\_segment and distribution\_channel by replacing all empty values to prevent errors and ensure all records are represented in the analysis.

## Corrected Data Types:

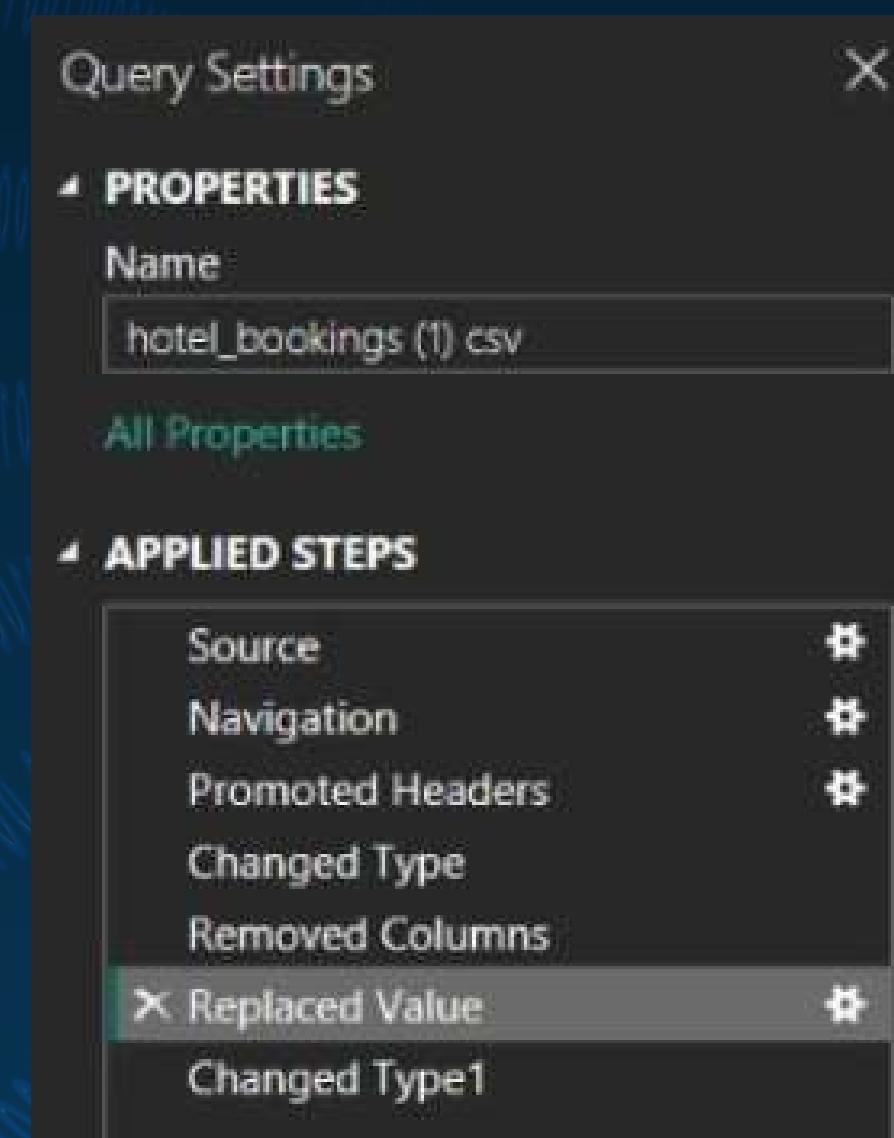
Identified and corrected incorrect data types for all numerical columns, including lead\_time, stays\_in\_weekend\_nights, adults, and adr. This ensured that the values could be used for accurate calculations and summaries.

# POWER BI SCREENSHOTS

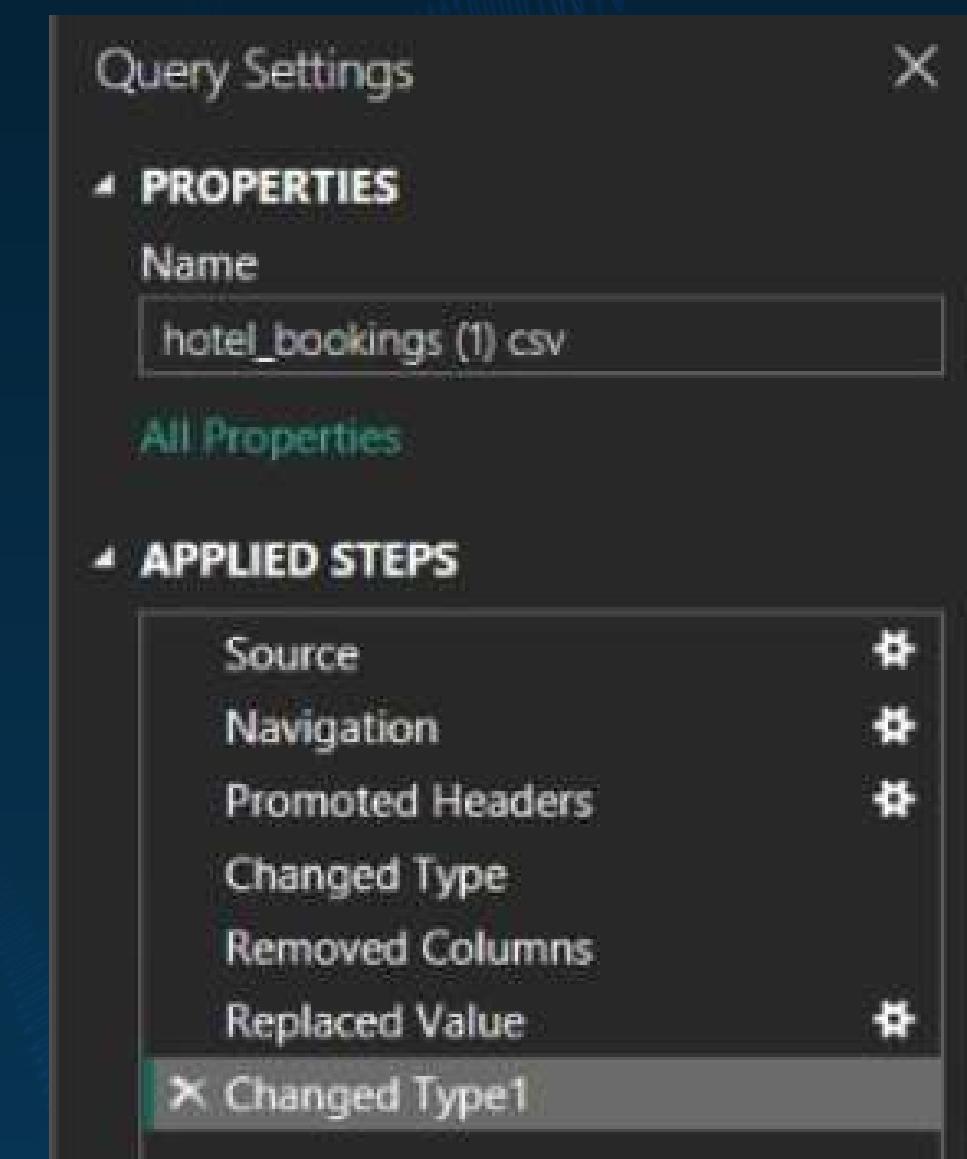
REMOVED COLUMNS



REPLACED NULLS



TRANSFORM COLUMN TYPES



# STAR SCHEMA & RELATIONSHIPS

FACTBOOKINGS → DIMCUSTOMER

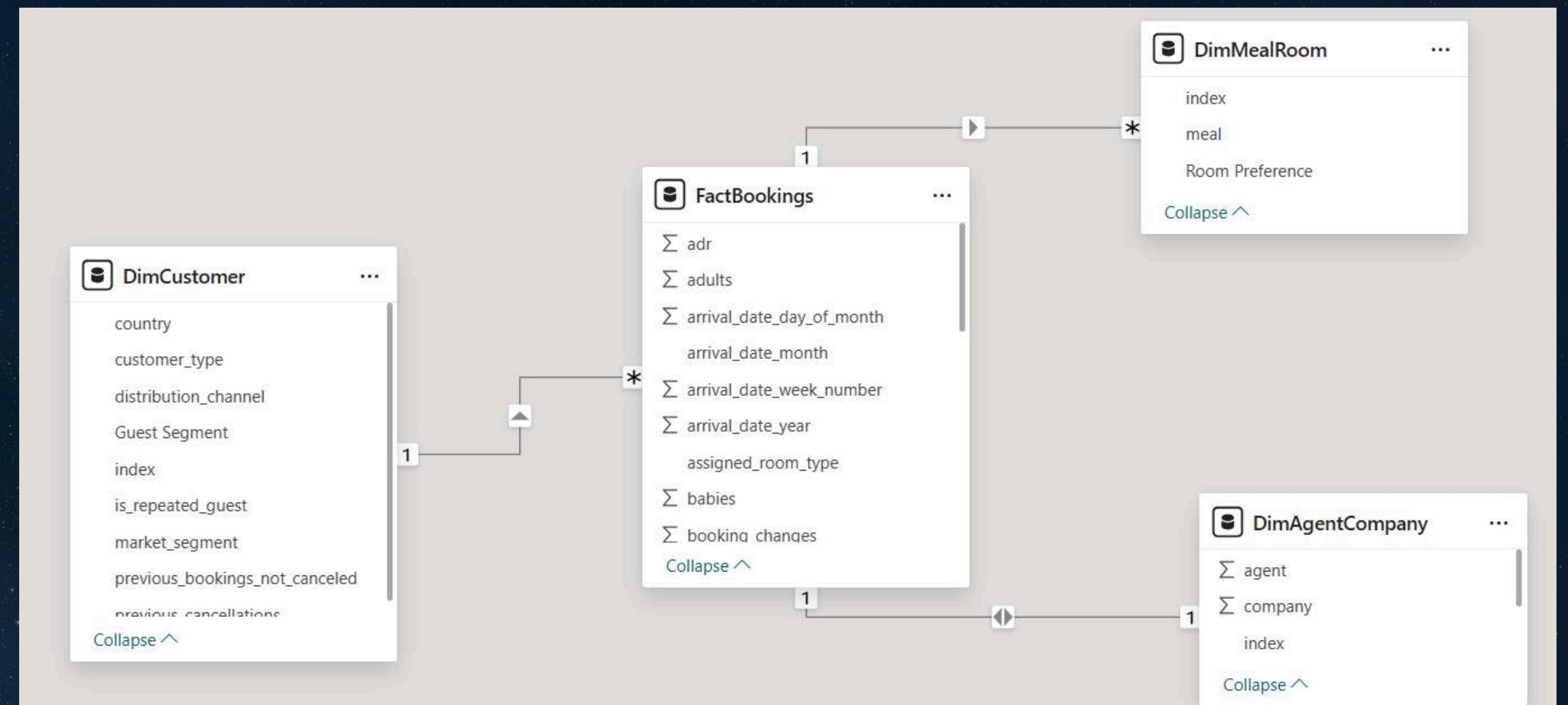
- RELATIONSHIP: 1-TO-MANY (ONE CUSTOMER CAN MAKE MANY BOOKINGS).
- CONNECTION VIA INDEX (CUSTOMERID).

FACTBOOKINGS → DIMMEALROOM

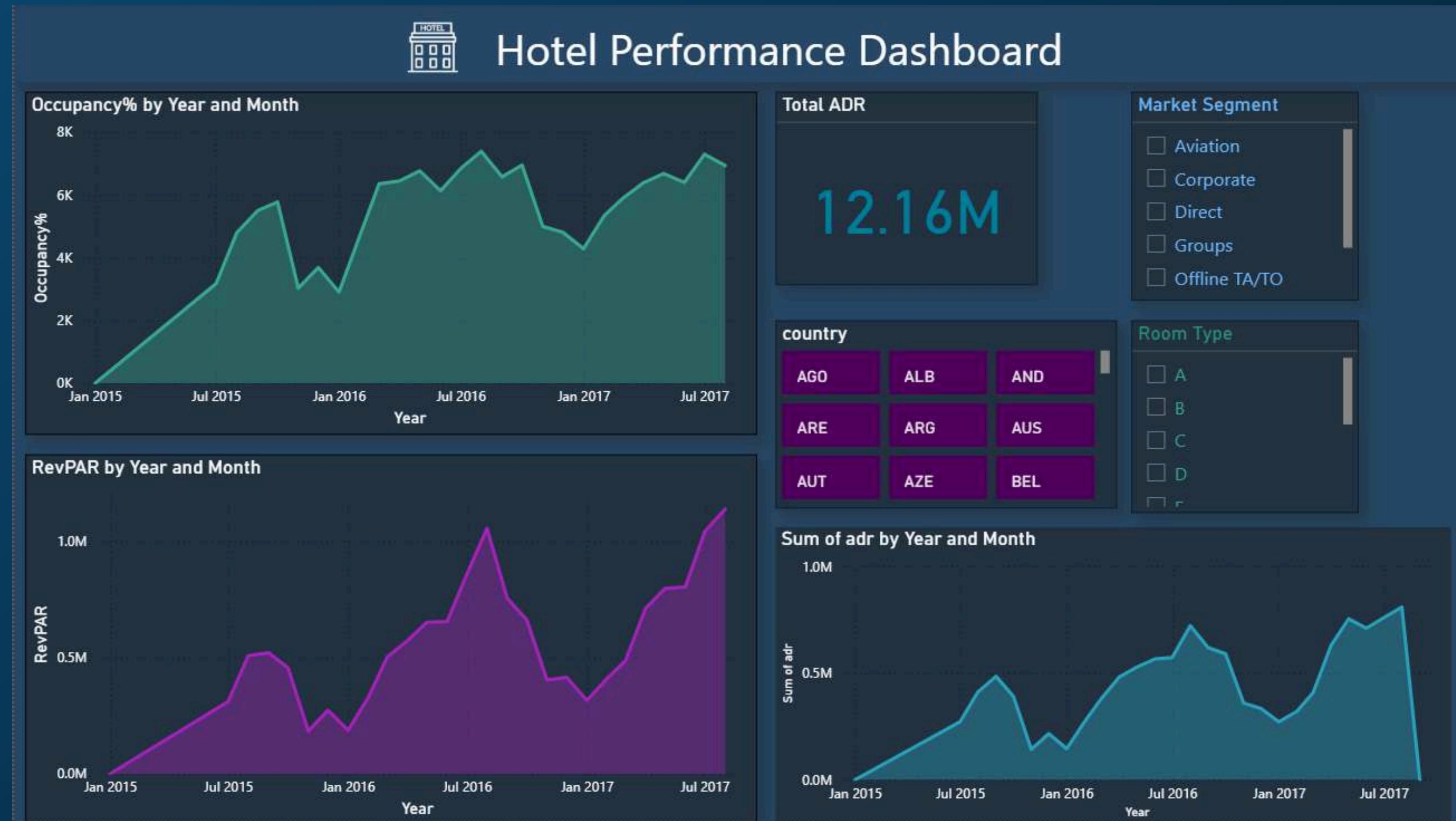
- RELATIONSHIP: 1-TO-MANY (ONE MEAL/ROOM TYPE CAN APPEAR IN MANY BOOKINGS).
- CONNECTION VIA MEAL / ASSIGNED\_ROOM\_TYPE.

FACTBOOKINGS → DIMAGENTCOMPANY

- RELATIONSHIP: 1-TO-MANY (ONE AGENT/COMPANY CAN HANDLE MANY BOOKINGS).
- CONNECTION VIA AGENT / COMPANY.



# INTERACTIVE DASHBOARDS



# OCCUPANCY & REVENUE PERFORMANCE ANALYSIS



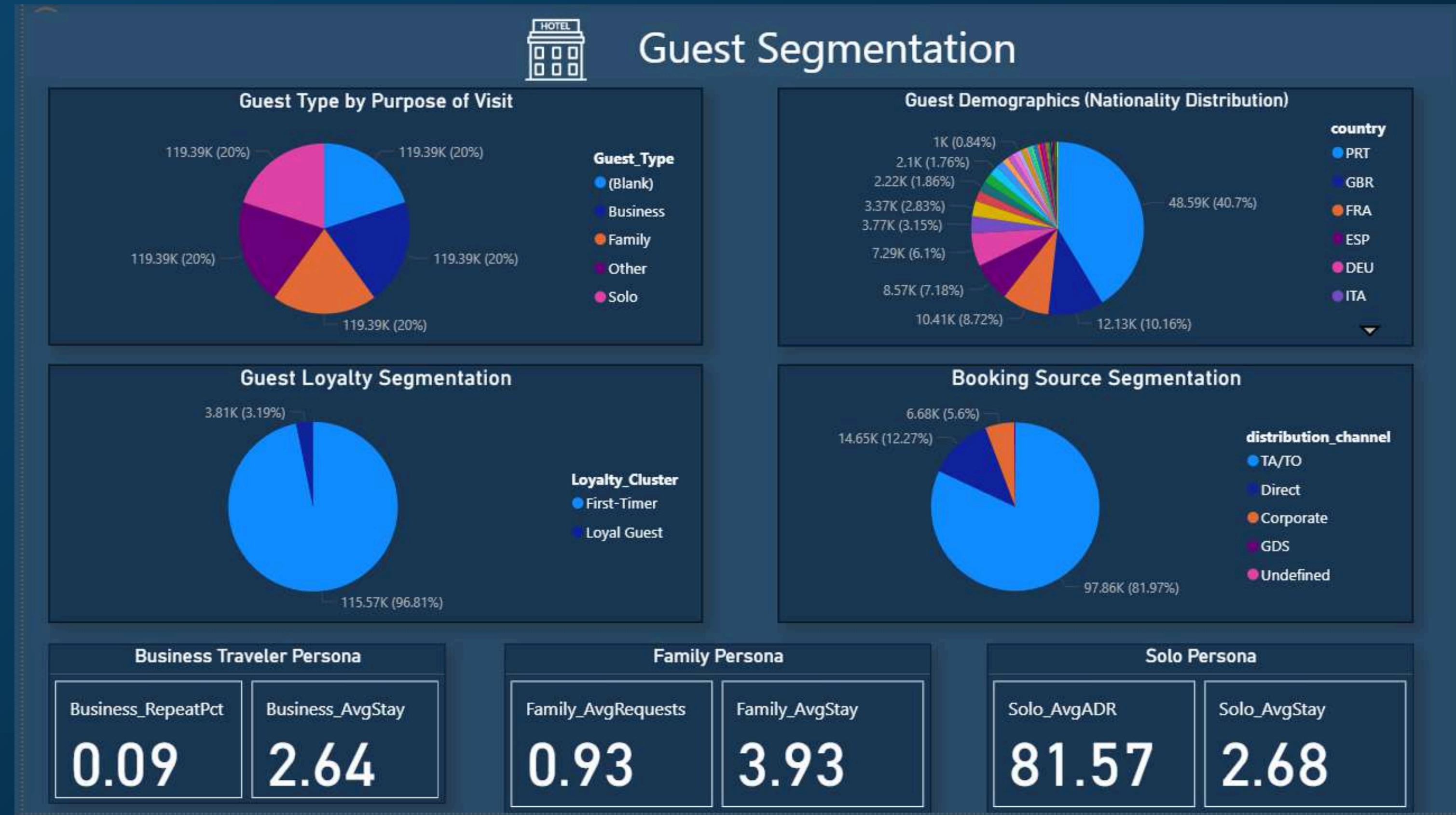
## Occupancy % Over Time

- Tracks monthly/annual demand with seasonal peaks.
- Guides staffing, inventory, and promotional planning.

## Revenue Performance

- ADR: Revenue per occupied room (pricing measure).
- RevPAR: Combines ADR + occupancy (overall revenue).
- Shows steady growth with dips in low-demand periods.

# INTERACTIVE DASHBOARD



# GUEST ANALYSIS MODE

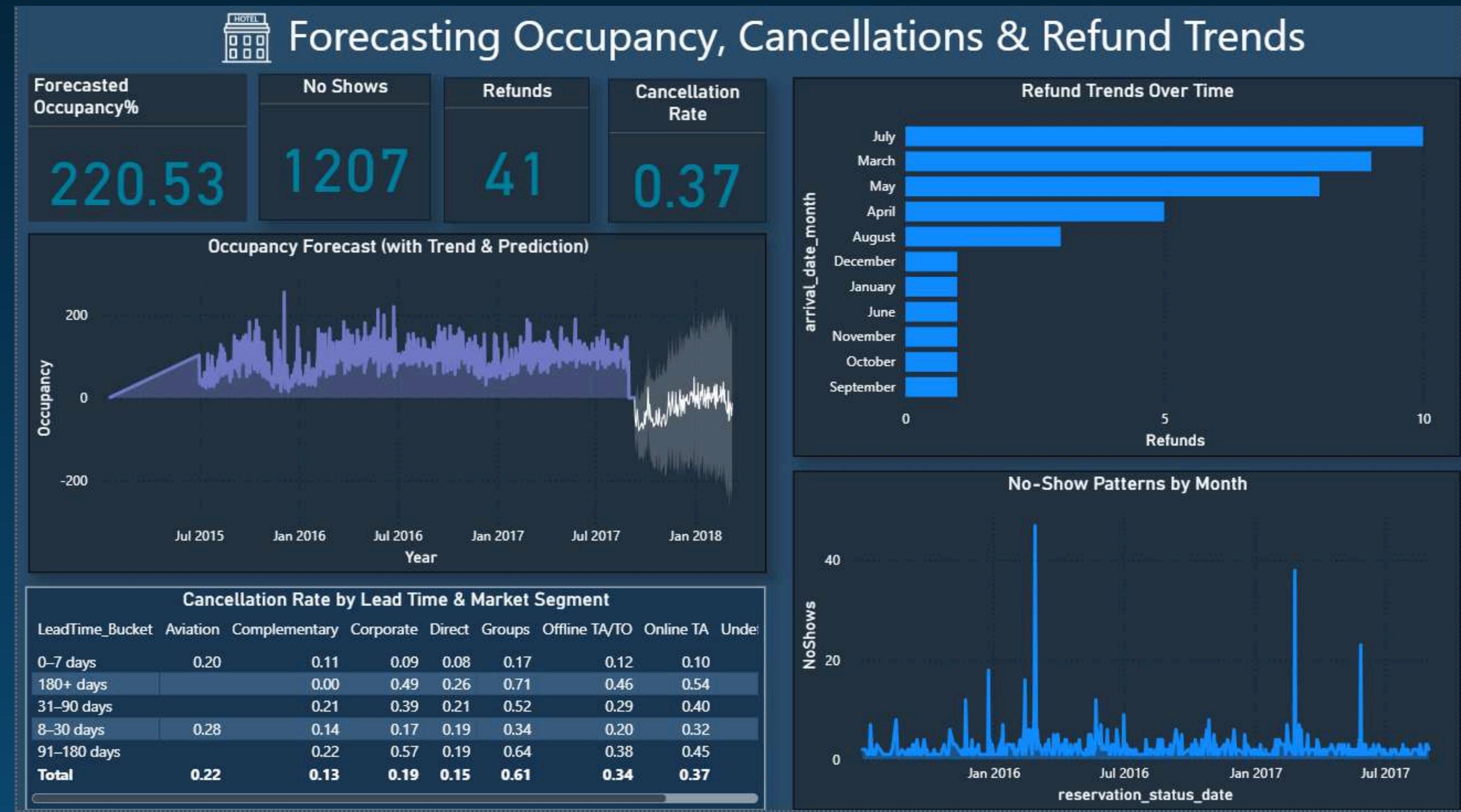
1. Implemented detailed guest segmentation to understand behavior and preferences.
2. Categorized guests by purpose of visit: Business, Family, Solo.
3. Added guest demographics analysis: nationality distribution.
4. Built booking source segmentation (Direct, OTA, Corporate, GDS).
5. Introduced loyalty module: First-timers, Loyal Guests, High Spenders.
6. Designed customer personas with KPIs: Avg Stay, Repeat %, ADR.

# INTERACTIVE FILTERING

One of the key features of this dashboard is its interactive filtering system.

- Filter by guest type (Business, Family, Solo, Other) to compare behaviors.
- Analyze nationality-based performance for regional insights.
- Explore loyalty patterns: First-Timers vs. Loyal Guests.
- Review booking channels (TA/TO, Direct, Corporate, GDS) to assess revenue sources.

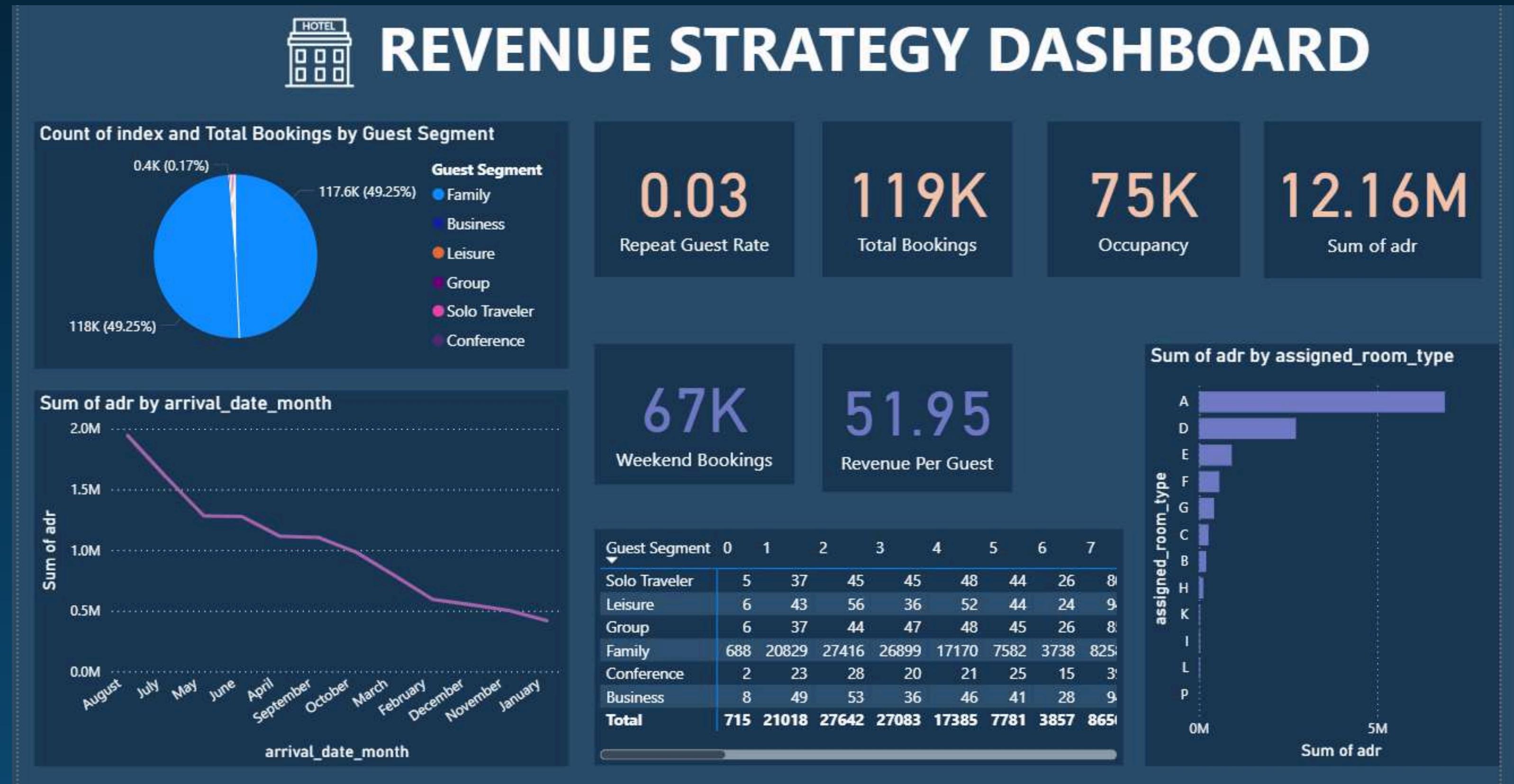
# INTERACTIVE DASHBOARDS



# FORECASTING OCCUPANCY, CANCELLATIONS & REFUND TRENDS

- Built an analytics dashboard providing real-time visibility into forecasted occupancy, cancellation rates, and refund trends.
- Forecasted occupancy shown through historical and prediction graphs, enabling managers to anticipate demand fluctuations.
- Cancellation rates analyzed by lead time and market segment, supporting precise policy adjustments and risk reduction.
- Refund trends tracked monthly, revealing periods of heightened risk and supporting proactive revenue protection.
- Detailed no-show patterns visualized, helping optimize booking conditions and resource planning.
- Delivers actionable insights for improving hotel revenue management, operational preparedness, and guest satisfaction

# INTERACTIVE DASHBOARD



# REVENUE STRATEGY

- Built a comprehensive dashboard offering a unified view of key revenue metrics—ADR (12.16M), total bookings (119K), occupancy (75K), and revenue per guest (51.95).
- Segmented guest booking patterns and total bookings by business, leisure, solo, family, group, and conference segments, spotlighting major revenue drivers.
- Analyzed potential upsell revenue (401.40) and identified 979 upsell candidates, enabling managers to create targeted packages and optimize ancillary income.
- Tracked ARR and repeat guest rate, enabling pricing strategies by season and guest type to maximize high-value bookings.
- Provided actionable insights to tailor promotions, strengthen guest retention, and drive overall profitability through data-driven decision-making and personalized marketing.

# AI-DRIVEN FORECASTING: OPTIMIZING FUTURE HOTEL PERFORMANCE

- **Goal:** Shift from reactive reporting to proactive decision-making using predictive AI.

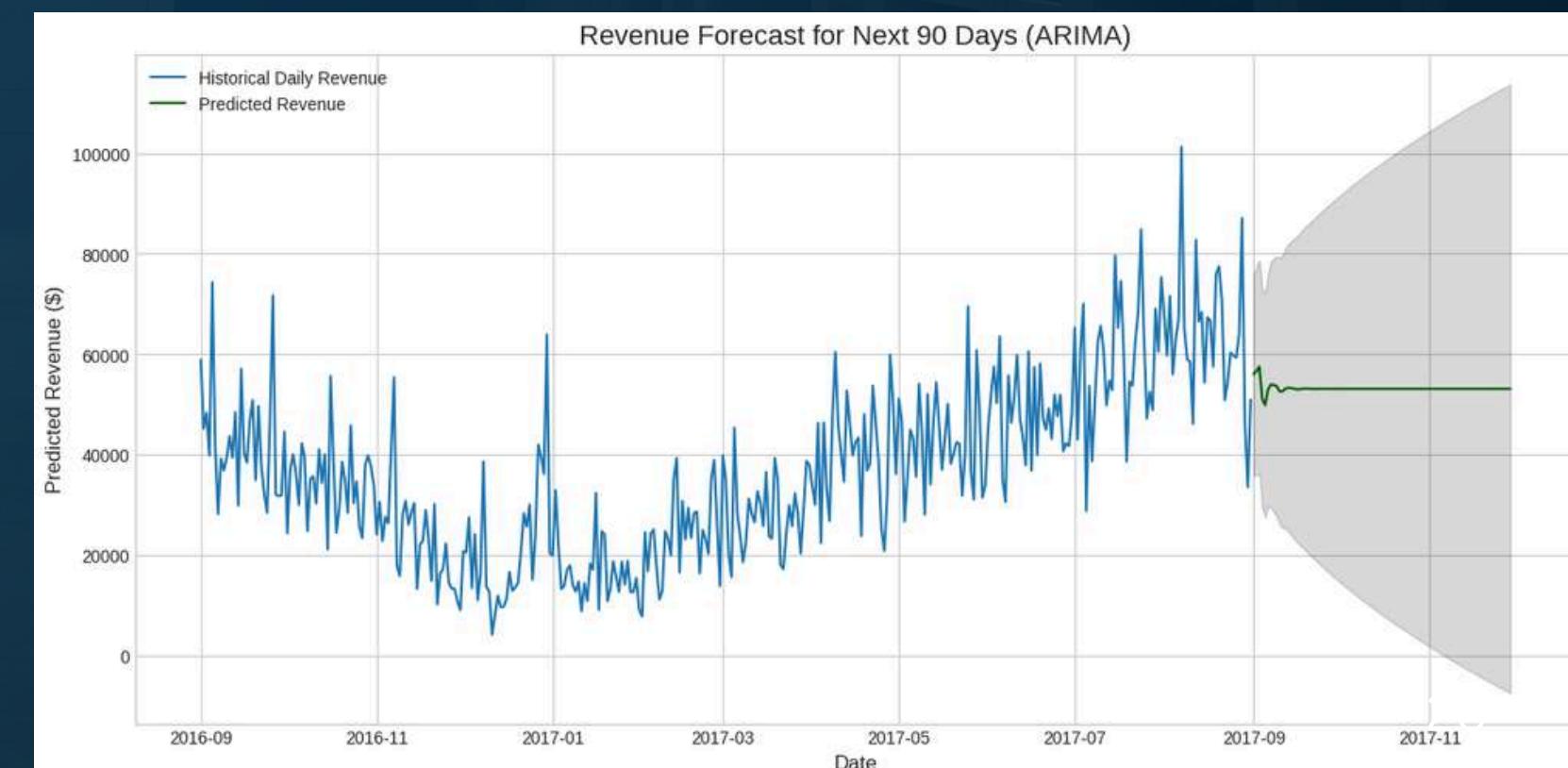
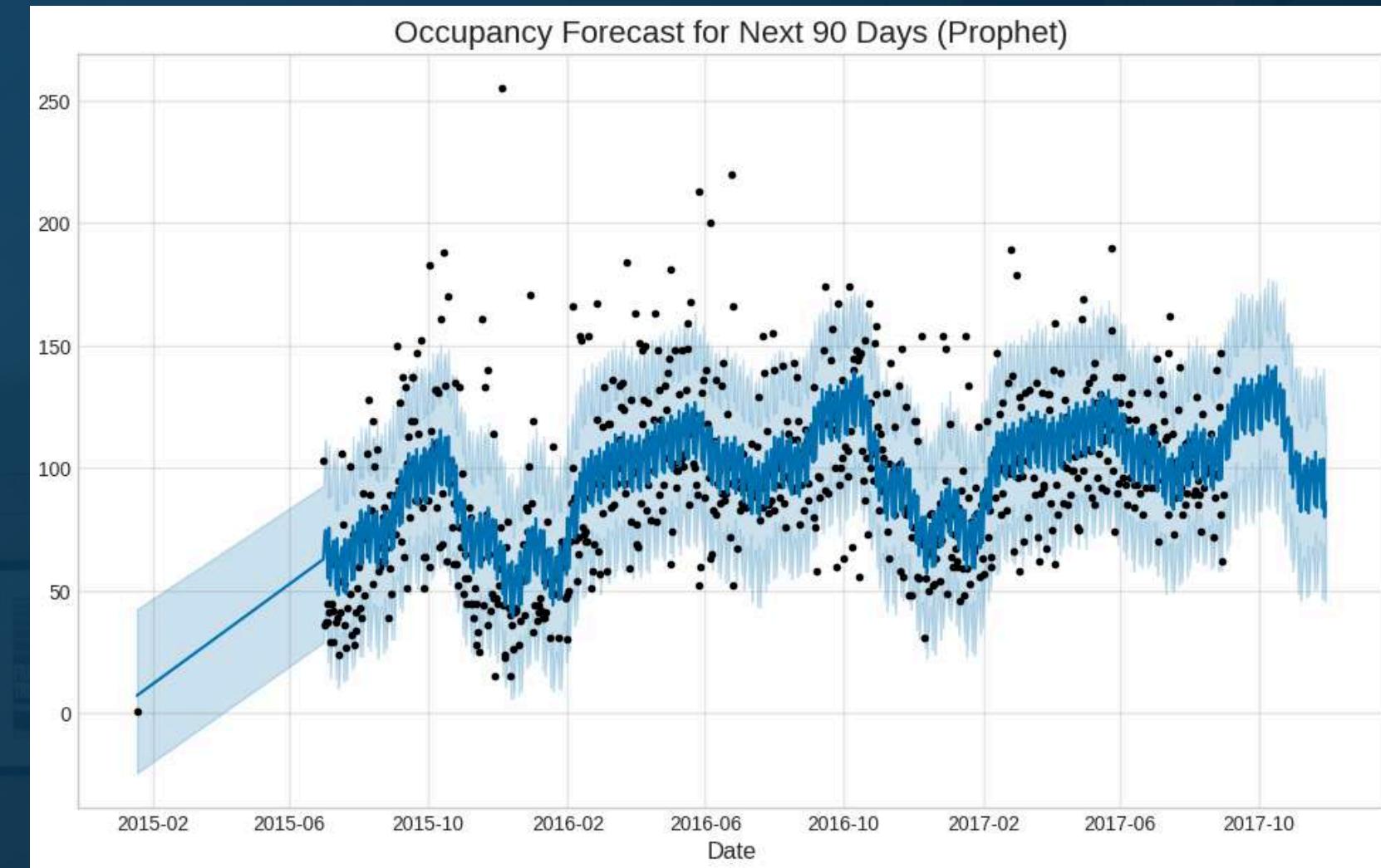
## Occupancy & Revenue Forecast (Prophet & ARIMA):

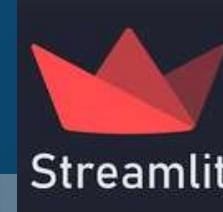
- **90-Day Projection:** Forecasts 10,254 total bookings and \$4,789,181.45 revenue.
- **Action:** Directly informs Yield Management (setting higher rates during peak times) and precise staff scheduling.

## Cancellation Risk Mitigation (Random Forest):

**High-Risk ID:** Identifies 4886 bookings (13.6% of the total) bookings with a  $\geq 75\%$  cancellation probability.

**Action:** Target these 4886 bookings with a proactive engagement campaign to mitigate revenue loss





# The Final Deliverable: Interactive Strategy Tool

- **Solution:** All predictive models are integrated into a user-friendly Streamlit Dashboard.
- **Key Features for Management:**
- **Adjustable Forecast:** Management can instantly change the number of forecast days (e.g., 30, 90, 365) to test different horizons.
- **Dynamic Risk List:** The dashboard provides the exact list of high-risk bookings (date, ADR, customer type) for immediate follow-up.
- **Model Performance:** Transparency via a live display of the Random Forest model's performance metrics.
- **Strategic Recommendation: Immediately implement a three-part strategy:**
- **Automate pricing** based on the **Revenue Forecast**.
- **Proactively staff** based on the **Occupancy Forecast**.
- **Target and retain** high-risk guests identified by the **Cancellation Model**.

# BUSINESS IMPACT

- **Revenue Optimization:** ARIMA's Revenue Forecast enables dynamic, smarter pricing against forecasted demand, significantly boosting hotel profitability (RevPAR/ADR).
- **Operational Efficiency:** Prophet's Occupancy Forecast ensures accurate staffing, while the Cancellation Model reduces costly overbooking and refunds.
- **Customer Retention:** Enhanced loyalty analysis and Cancellation Risk Identification allow for proactive outreach to high-value or high-risk guests, improving retention.
- **Targeted Marketing:** Use guest segmentation and risk profile to create tailored promotions (e.g., offer free services to high-risk bookings to secure them).
- **Distribution Strategy:** Channel analysis (OTA vs. Direct) guides negotiations and helps shift focus to increase direct bookings.
- **Data-Driven Decisions:** Managers make quick, evidence-based decisions using the unified dashboard, replacing time-consuming manual report generation.
- **Scalability & Consistency:** The predictive models and dashboard framework can be scaled across all hotel branches for chain-wide, consistent performance monitoring.

# CONCLUSION

- Cleaned and transformed raw booking, customer, and room data to ensure accuracy, consistency, and reliability.
- Designed a star schema model with Fact and Dimension tables, establishing correct relationships between customers, bookings, and hotel branches.
- Created a well-structured dataset that serves as a foundation for performance analysis and reporting.
- Developed interactive and dynamic dashboards to visualize key hotel metrics: Occupancy Percentage, Average Daily Rate (ADR), and Revenue per Available Room (RevPAR).
- Implemented advanced filters for room type, location, booking source, and time period, enabling flexible and in-depth analysis.
- Provided actionable, customizable insights to help monitor performance trends, optimize hotel operations, and support data-driven decision-making.

The background image shows a modern architectural structure with a dark blue-tinted glass facade. Above the glass, there are large, white, curved overhangs supported by a grid of steel beams. The overall aesthetic is clean and minimalist.

Thank You.