

A Project Work Report on

"HOTEL BOOKING DEMAND & CANCELLATION ANALYSIS"

*Uncovering booking trends and cancellation
behavior to drive data-informed hotel strategies*

Prepared by:

DHRUV

TABLE OF CONTENTS

	TITLE	PAGE NO.
1	EXECUTIVE SUMMARY	1
2	PROJECT OVERVIEW	2
2.1	Purpose of the Analysis	2
2.2	Scope of the Dataset	2
2.3	Analytical Approach	2
2.4	Tools Used	3
2.5	Outcome of the Setup Phase	3
3	DATA PREPARATION	4
3.1	Purpose of Data Preparation	4
3.2	Data Cleaning Process	4
3.3	Feature Engineering	5
3.4	Data Validation Summary	5
4	EXPLORATORY ANALYSIS & KPIs	6
4.1	KPI Framework	6
4.2	KPI Analysis — Group A: Booking Volume & Trends	6
4.3	KPI Analysis — Group B: Profit & Revenue Metrics	6
4.4	KPI Analysis — Group C: Guest Behaviour	7
4.4	KPI Analysis — Group D: Cancellation Intelligence	7
5	DASHBOARD VISUALIZATION & STRATEGIC INSIGHTS	8
5.1	Overview of Dashboard Development	8
5.2	Dashboard 1 — Booking & Revenue	8
5.2.1	KPI Summary	8
5.2.2	Monthly Booking Trend	9
5.2.3	Revenue & ADR Trend	9
5.2.4	Seasonal Booking Distribution	10
5.2.5	Revenue Comparison by Hotel Type	10

5.2.6	Market Segment Contribution	11
5.2.7	Bookings & Revenue by Country	11
5.3	Dashboard 2 — Guest Behaviour & Cancellation Insights	12
5.3.1	Cancellation KPIs	12
5.3.2	Guest Composition Breakdown	13
5.3.3	Booking Distribution by Channel	13
5.3.4	Lead Time Comparison by Hotel Type	13
5.3.5	Cancellation Rate by Market Segment	14
5.3.6	Lead Time vs Cancellation Rate	14
5.3.7	Monthly Cancellation Trends	15
6	CONCLUSION & RECOMMENDATIONS	16
6.1	Conclusion	16
6.2	Business Recommendations	16

1. EXECUTIVE SUMMARY

This report presents a comprehensive analysis of hotel booking patterns, revenue performance, guest behaviour, and cancellation risks using SQL-based data processing and Power BI visualisation. The dataset covers ~119,000 bookings from City and Resort Hotels, offering a broad view of customer trends across multiple seasons and booking channels.

City Hotels consistently generate higher revenue due to greater booking volume, while Resort Hotels maintain steady ADR driven by leisure-oriented travel. Strong seasonality is observed, with Summer and Spring showing peak demand. Guest planning behaviour shows long lead times (~104 days), indicating early travel planning and higher exposure to cancellations.

Cancellations remain a critical concern, with ~37% of bookings cancelled and an estimated revenue loss of over \$16M. Online Travel Agents and Group segments contribute heavily to this risk, while Direct and Corporate channels demonstrate more stable and committed booking behaviour.

The findings highlight clear opportunities to optimize pricing, strengthen channel strategy, improve demand forecasting, and enhance customer retention. This analysis provides the foundation for implementing targeted policy changes, improving booking reliability, and maximizing revenue across different travel seasons and segments.

2. PROJECT OVERVIEW

2.1 Purpose of the Analysis

This project examines hotel booking data to uncover patterns in booking behaviour, revenue performance, guest demographics, and cancellations.

The goal is to understand booking trends, revenue patterns, guest behaviour, and cancellation risks to help hotels make better data-driven decisions around pricing, demand forecasting, and customer retention.

2.2 Scope of the Dataset

The analysis uses a consolidated dataset from two property types - City Hotel and Resort Hotel, covering a wide range of booking attributes.

Dataset Snapshot

- Total Records: ~119,000 bookings
- Period Covered: July 2015 – August 2017
- Hotel Types: City and Resort
- Key Variables:
 - Booking dates & lead time
 - Guest composition
 - Stay duration
 - Market segments & channels
 - ADR (Average Daily Rate)
 - Cancellation indicators
 - Geographic information

This dataset provides comprehensive visibility into operational trends and customer behaviour.

2.3 Analytical Approach

Structured Data Setup

- SQL database created: *hotel_analysis*
- Main dataset loaded into: *hotel_bookings*
- Lookup tables integrated for:
 - Market segments
 - Distribution channels

- Country-to-continent mapping

Data Validation Activities

- Record count verification
- Sample record checks
- Consistency checks for dates, categories, and numerical fields

These steps ensured a clean and reliable foundation before performing KPI analysis and dashboard creation.

2.4 Tools Used

Excel	Used for initial data scanning, field profiling, and early consistency checks.
MySQL	Used for data cleaning, transformation, mapping, and KPI derivation. This includes handling missing values, creating analytical features, and validating data structures.
Power BI	Used to build interactive dashboards, visualize trends, and present insights in a business-ready format.

2.5 Outcome of the Setup Phase

With the data structured, validated, and enriched through SQL, the analysis was able to advance into KPI development, exploratory insight generation, and dashboard visualization with accuracy and clarity.

3. DATA PREPARATION

3.1 Purpose of Data Preparation

This phase focused on converting raw booking records into a structured, analysis-ready dataset. The objective was to ensure accuracy, consistency, and reliability by resolving data quality issues, handling null values, and standardizing fields before KPI development.

3.2 Data Cleaning Process

Guest & Stay Details

- Corrected missing or invalid guest counts (adults, children, babies).
- Removed records with zero or negative stay durations.

Booking Source Classification

- Standardized booking sources by deriving:
booking_source, source_id, agent_flag, company_flag
- Enabled segmentation of bookings into *Agent, Company, or Direct* categories.

Country & Continent Mapping

- Replaced blank or null country fields with '*Unknown*'.
- Mapped each country to its corresponding *continent* using a lookup table.

Financial & Duration Checks

- Corrected negative ADR values.
- Validated *lead_time* ranges to eliminate unrealistic entries.

Repeat Guest Validation

- Cross-checked *is_repeated_guest* against booking history.
- Added logical flags to improve repeat guest accuracy.

Sanity & Consistency Tests

- Reviewed distinct values in categorical fields such as
hotel, meal, deposit_type, market_segment.
- Ensured no nulls existed in critical analytical columns.

3.3 Feature Engineering

Objective:

To enrich the dataset with analytical fields that improve trend detection, segmentation, and KPI computation.

Derived Date Fields

- Combined arrival day, month, and year into a single *arrival_date* column.
- Converted *reservation_status_date* into a valid SQL date (*reservation_status_dt*).
- Verified that all dates fall within the dataset range (July 2015 – August 2017).

Stay & Guest Metrics

- Created *stay_length* = *stays_in_weekend_nights* + *stays_in_week_nights*
- Created *guest_count* = *adults* + *children* + *babies*
- Useful for occupancy, group size, and revenue calculations.

Cancellation Indicator

- Derived *cancellation_flag* from *is_canceled* to simplify cancellation-related KPIs in SQL and Power BI.

Booking Date Calculation

- Added *booking_date* = *arrival_date* – *lead_time*
- Enables analysis of booking patterns, lead-time behaviour, and demand forecasting.

Season Classification

- Grouped months into *Winter*, *Spring*, *Summer*, *Autumn*.
- Simplifies identification of season-driven demand patterns.

Clean Repeat Guest Indicator

- Improved repeat guest identification by refining *repeat_status_clean* through historical booking checks.
- Helps measure loyalty and returning customer behaviour.

3.4 Data Validation Summary

All derived and cleaned fields were tested for:

- i) Valid ranges ii) Data type consistency iii) Null-free completeness

The dataset was confirmed to be fully prepared for KPI development, exploratory analysis, and dashboard visualization.

4. EXPLORATORY ANALYSIS & KPIs

4.1 Purpose of the Analysis

This phase focuses on deriving *quantitative metrics* that summarize booking demand, revenue performance, guest behaviour, and cancellation patterns.

SQL-based aggregations were used to convert raw booking records into structured KPIs that support trend analysis and decision-making.

4.2 KPI Framework

To ensure clarity and consistency, KPIs were grouped into four key analytical categories:

1. **Booking Volume & Trends:** Measures overall demand, seasonality, and year-over-year movement.
2. **Profit & Revenue Metrics:** Evaluates pricing effectiveness and revenue contribution across hotel types.
3. **Guest Behavior Indicators:** Highlights booking patterns, stay characteristics, and planning behavior.
4. **Cancellation Intelligence:** Identifies drivers, risk factors, and operational impact of cancellations.

4.3 KPI Analysis — Group A: Booking Volume & Trends

Focus:

To measure how booking activity changes over time and across seasons.

Key Findings:

- Total Bookings: ~119,000 bookings analysed
- Yearly Trends: Moderate growth from 2015 → 2016; slight dip in early 2017
- Monthly Trends: Highest volume in *July–August* (peak travel period)
- Seasonal Trends: *Summer and Spring* drove strongest demand

Business Insight:

Peak summer activity confirms seasonality enabling targeted pricing and inventory planning.

4.4 KPI Analysis — Group B: Profit & Revenue Metrics

Focus:

Assess profitability and understand ADR (Average Daily Rate) movement.

Key Findings:

- ADR: Higher for City Hotels with a steady month-over-month increase
- Total Revenue: City Hotels generated larger revenue volumes due to higher bookings
- Hotel-Type Comparison:
 - Resort Hotels show stable ADR
 - Lower volume reduces total revenue contribution

Business Insight:

City Hotels lead revenue through volume, while Resort Hotels benefit from steady pricing. Strategies can be tailored by hotel type.

4.5 KPI Analysis — Group C: Guest Behaviour

Focus:

Analyse stay duration, group size, and planning characteristics.

Key Findings:

- Average Stay Length: 3–4 nights; slightly longer at resorts
- Guest Composition: Mostly adults; limited family-oriented stays
- Lead Time: Average lead time *~104 days*; resort guests book further in advance

Business Insight:

Resort bookings show planned, leisure-driven customer behaviour, suggesting opportunities for advance-booking promotions.

4.6 KPI Analysis — Group D: Cancellation Intelligence

Focus:

Identify patterns behind cancellation behaviour and its operational impact.

Key Findings:

- Cancellation Rate: *~37%* of bookings cancelled
- Yearly Trend: 2016 recorded the highest cancellation volume
- Lead-Time Effect: Longer lead times correlated with higher cancellation probability
- Market Segments: OTAs (Online Travel Agents) showed the most cancellations
- Geography: Portugal contributed a high cancellation volume

Business Insight:

High cancellations from OTA channels and long-lead bookings indicate the need for stricter cancellation policies and improved guest engagement for high-risk segments.

5. DASHBOARD VISUALIZATION & STRATEGIC INSIGHTS

5.1 Overview of Dashboard Development

An interactive Power BI dashboard was created to translate SQL-derived KPIs into visual insights.

The dashboard provides a consolidated view of *booking demand, revenue performance, guest behaviour, and cancellation risks*, enabling quick decision-making for hotel management.

5.2 Dashboard 1 — Booking & Revenue



Dashboard 1 provides a consolidated view of overall hotel performance by combining key booking, revenue, seasonal, and market-segment insights.

It highlights where demand is strongest, how revenue fluctuates across months and hotel types, and which customer segments contribute the most.

The dashboard serves as the primary tool for understanding *booking trends, revenue drivers, seasonal patterns, and market mix performance*.

5.2.1 KPI Summary

Total Bookings	Total Revenue	Realized Revenue	Average ADR	Average Stay Length	Average Guest Count
119K	\$42.72M	\$26.0M	101.83	3.43	1.97

This section highlights the core metrics reflecting hotel performance:

- Total Bookings: ~119K bookings, representing overall demand volume.
- Total Revenue: Estimated using $ADR \times \text{stay length}$ to capture total business value.
- Realized Revenue: Revenue from non-cancelled bookings, indicating actual earnings.
- Average Daily Rate (ADR): ~101.8, measuring pricing effectiveness per occupied room.
- Average Stay Length: ~3.43 nights, indicating typical occupancy duration.
- Average Guest Count: ~1.97 guests per booking, showing the general party size.

These KPIs form the foundation for understanding revenue contribution, occupancy behaviour, and pricing dynamics.

5.2.2 Monthly Booking Trend

Insights

- City Hotels maintain significantly higher monthly bookings than Resort Hotels.
- Strong seasonality is observed, with peaks during *Oct–Nov* and *Apr–Jun*.
- Resort Hotel demand remains stable, while City Hotels experience sharper fluctuations driven by business travel cycles.



Recommendations

- Optimize pricing during peak months to maximize revenue.
- Offer off-season packages for Resort Hotels to stabilize low-demand periods.
- Strengthen forecasting for City Hotels due to higher demand volatility.

5.2.3 Revenue & ADR Trend

Insights

- Revenue peaks in *Aug–Sep*, driven by strong booking activity and higher ADR.
- ADR movements closely mirror revenue trends, confirming the impact of pricing on revenue.



- Jan–Feb consistently underperform, marking clear off-season months.

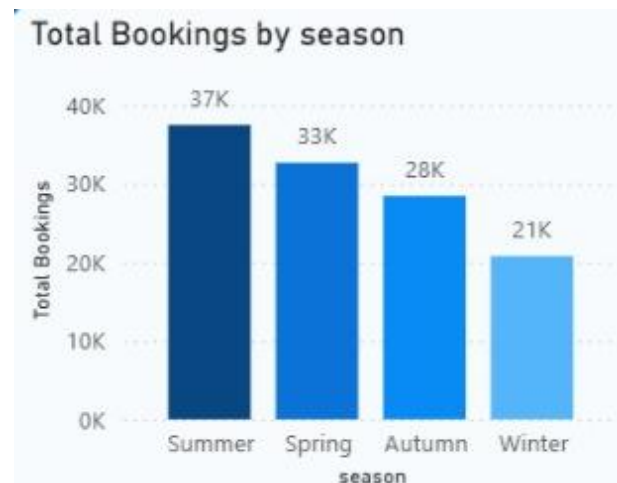
Recommendations

- Introduce premium pricing bundles during high-ADR periods (Aug–Sep).
- Use promotional pricing in Jan–Feb to improve occupancy.
- Align marketing spend with months showing high ADR response.

5.2.4 Seasonal Booking Distribution

Insights

- Summer records the highest booking volume, confirming vacation-driven demand.
- Spring performs moderately well, showing early-year travel interest.
- Winter is the weakest season for both hotel types.



Recommendations

- Leverage Summer for premium packages and upsell opportunities.
- Use Winter discounts and targeted promotions to lift occupancy.
- Adjust staffing and inventory levels based on seasonal demand cycles.

5.2.5 Revenue Comparison by Hotel Type

Insights

- City Hotels generate *\$14.4M+* in realized revenue, outperforming Resort Hotels (*\$11.6M+*).
- Resort Hotels maintain steadier ADR but have fewer bookings.
- The revenue gap reflects strong business-driven demand for City Hotels.



Recommendations

- Expand corporate and business-traveller offerings for City Hotels.
- Use seasonal upsells and bundled experiences for Resort Hotels.
- Allocate marketing budgets based on hotel-type revenue efficiency.

5.2.6 Market Segment Contribution

Insights

- OTAs dominate booking contribution and act as the strongest demand channel.
- Offline Travel Agencies and Group bookings offer meaningful secondary volume.
- Direct and Corporate channels remain underutilized.

Recommendations

- Strengthen OTA partnerships to maintain visibility and volume.
- Promote direct-booking benefits to improve margin and loyalty.
- Expand corporate and group-travel initiatives for diversified revenue.



5.2.7 Bookings & Revenue by Country

Insights

- Europe contributes the majority of bookings and revenue.
- North and South America show moderate inflow.
- Asia and Africa are emerging segments with smaller volumes.

Recommendations

- Increase targeted campaigns in high-performing European regions.
- Build awareness programs to grow mid-tier American markets.
- Explore partnerships in Asia and Africa to capture new demand pockets.



5.3 Dashboard 2 — Guest Behaviour & Cancellation Insights



Dashboard 2 focuses on customer behaviour and booking reliability.

It presents insights on guest demographics, lead-time patterns, cancellation rates, and channel-level booking stability.

The dashboard helps identify *cancellation risks*, *guest planning behaviour*, *retention challenges*, and *the operational impact of cancellations*, enabling more accurate forecasting and policy decisions.

5.3.1 Cancellation KPIs



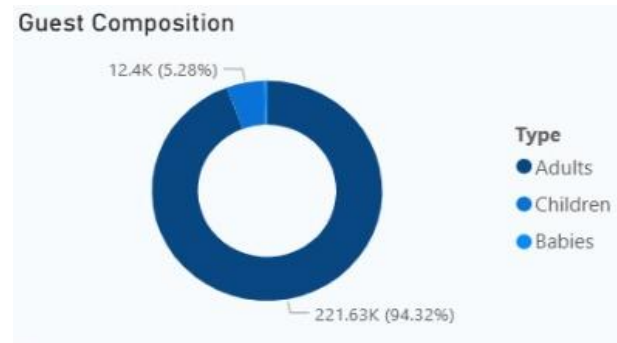
This section highlights the core metrics that capture guest behaviour patterns and the overall impact of cancellations on hotel performance.

- Total Cancellations: 44K cancellations impacting occupancy stability.
- Cancellation Rate: ~37%, indicating significant revenue leakage.
- Revenue Lost: ~\$16.73M due to cancellations.
- Average Lead Time: 104 days, increasing cancellation risk.
- Repeat Guests: 8.4%, showing low guest loyalty.
- Special Requests: 41%, indicating high personalization expectations.

5.3.2 Guest Composition Breakdown

Insights

- Adults form the majority (94%).
- Children represent ~5%; babies <1%.
- Demand is primarily driven by adult or couple travellers.



Recommendations

- Prioritize adult-centric premium services.
- Improve family-friendly offerings to grow the children segment.
- Provide optional low-cost infant amenities where needed.

5.3.3 Booking Distribution by Channel

Insights

- Travel Agents /Tour Operatoring channels dominate (97K+ bookings).
- Direct bookings form a strong secondary channel (14K+).
- Corporate and GDS channels remain underutilized.



Recommendations

- Strengthen OTA relationships to maintain volume.
- Increase direct-booking incentives (website exclusives, loyalty points).
- Build corporate partnerships for steady, low-risk demand.

5.3.4 Lead Time Comparison by Hotel Type

Insights

- City Hotels: 110-day lead time — long, planned bookings.
- Resort Hotels: 93-day lead time — more spontaneous leisure travel.
- Longer lead time increases cancellation exposure.



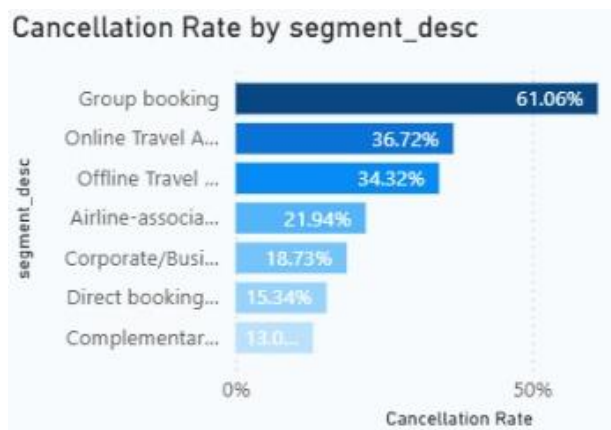
Recommendations

- Apply stricter cancellation policies for long lead-time bookings.
- Promote early-bird offers for Resort Hotels to stabilize demand.
- Use lead-time pricing (premium for last-minute, discounts for early planners).

5.3.5 Cancellation Rate by Market Segment

Insights

- Group bookings show the highest cancellation rate (~61%).
- OTA and Offline TA bookings have elevated risk.
- Direct & Corporate channels show low cancellation rates.



Recommendations

- Implement stricter cancellation terms for high-risk segments.
- Grow direct bookings for more stable demand.
- Expand corporate contracts with commitment-based pricing.

5.3.6 Lead Time vs Cancellation Rate

Insights

- City Hotels: early bookings (~110 days) → higher cancellations (~42%).
- Resort Hotels: shorter planning window (~92 days) → lower cancellations (~28%).
- Booking volume similar, but cancellation behaviour different.



Recommendations

- Reinforce deposit-based rules for long lead-time bookings.
- Encourage stable early bookings through incentives.
- Adjust pricing models to reflect lead-time sensitivity.

5.3.7 Monthly Cancellation Trends

Insights

- Peaks in Oct–Nov and Apr–Jun indicate predictable high-risk months.
- Lowest cancellations in Dec–Jan, showing strong commitment.
- Patterns align with booking volume cycles.



Recommendations

- Tighten cancellation windows during peak cancellation months.
- Offer off-peak flexibility to boost confidence in booking.
- Use trends to optimize staffing, pricing, and occupancy strategy.

6. CONCLUSION & RECOMMENDATIONS

6.1 Conclusion

The analysis shows strong seasonal demand patterns, with City Hotels leading in revenue due to higher booking volume and Resort Hotels maintaining steady ADR. Cancellations remain the major challenge, with ~37% cancellations and over \$16M in lost revenue—driven largely by OTA and long lead-time bookings.

Overall, these insights provide a clear foundation for improving pricing, channel strategy, forecasting, and customer experience.

6.2 Business Recommendations

1. Strengthen Cancellation Risk Management

- Introduce stricter cancellation windows for OTA and Group bookings.
- Implement lead-time-based policies (advance deposits, staggered cancellation fees).
- Monitor high-risk countries and segments with targeted communication.

2. Improve Direct Booking Ecosystem

- Offer loyalty rewards, exclusive website pricing, and value-added benefits.
- Build targeted campaigns to shift customers from high-cost OTA channels.
- Simplify booking workflows and deliver personalized offers to improve conversions.

3. Optimize Channel Strategy

- Deepen collaborations with high-performing OTAs to maintain peak-season visibility.
- Expand corporate partnerships through negotiated contracts and bundled services.
- Develop specialized packages to grow Group and event-related bookings.

4. Leverage Seasonality for Revenue Planning

- Use peak seasons (Summer, Spring) for premium pricing and upsell opportunities.
- Deploy Winter/off-season promotional campaigns to stabilize occupancy.
- Adjust staffing and inventory planning around predictable demand cycles.

5. Scale High-Performing Segments; Revive Underperforming Ones

- Promote premium offerings for City Hotel business travellers.
- Introduce family-oriented packages to expand the Children segment.
- Explore emerging regions (Asia, Africa) with targeted marketing to build awareness.