Hotel Haven

Reducing Costly Cancellations with Early Risk Signals

Presented by: **Desiree’ Salvant**

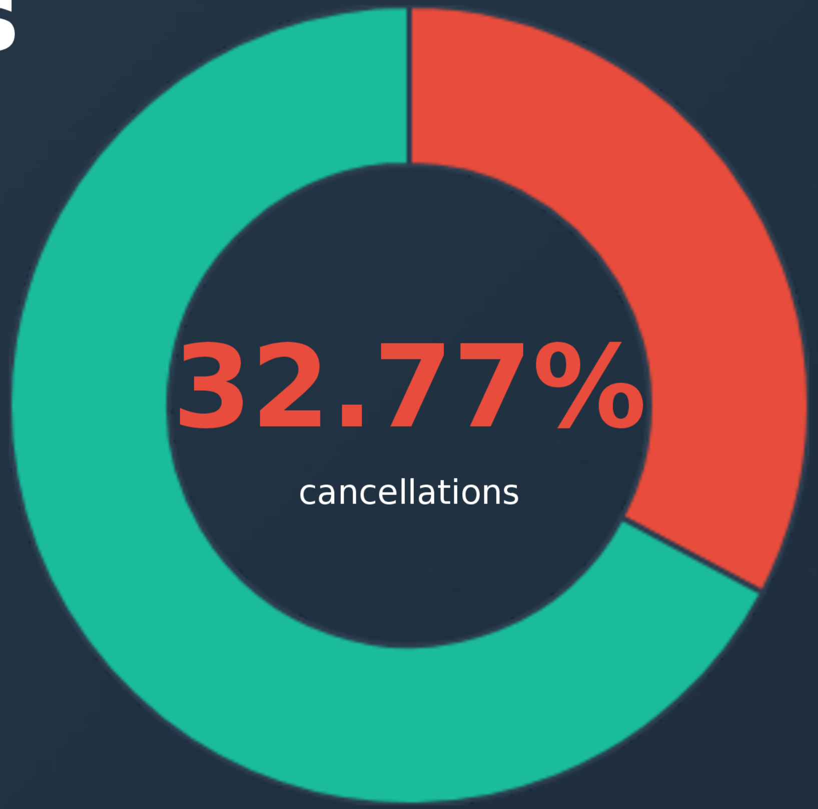
Unpredictable Cancellations Hurt Revenue and Operations

32.77% of all bookings cancel

That's approximately **one in three** reservations that don't materialize into stays.

Cancellations are frequent enough to materially impact revenue and guest experience.

- We can identify risk at booking time
- And implement targeted interventions before cancellations occur



With predictive analytics, we can identify the cancellations before they happen

What Cancellations Cost Us

36,285 bookings analyzed

11,889 canceled stays

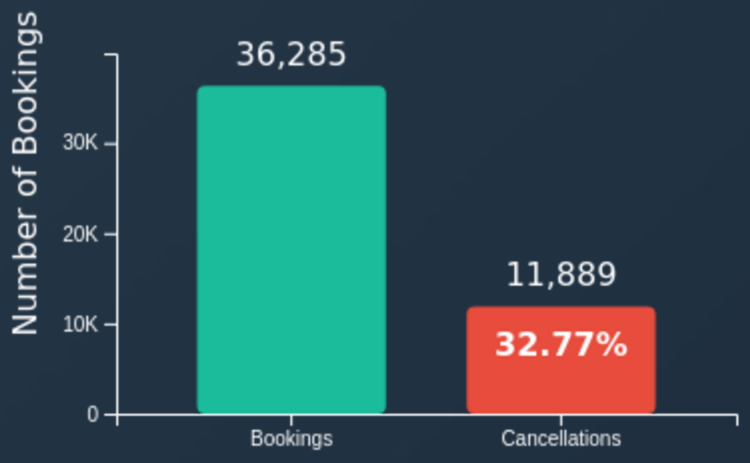
$1.23M at risk
Based on 11,889 cancellations × $103.42 average daily rate

Cancellations create:

Revenue
Unpredictable fluctuations

Operations
Staffing and resource allocation

Guest Experience
Empty rooms that could be filled



We can protect revenue with guest-friendly moves that don't require broad discounts

Flag At-Risk Bookings Early to Protect Revenue

High Risk
Score ≥ 0.70

What We Do:

- Proactive call/SMS
- +7d / −30d reassurance
- Price floors on hot dates
- Non-cash perks (room pick, late checkout)

Why It Works:

Stops avoidable cancels on valuable stays

Medium Risk
Score 0.36–0.69

What We Do:

- Auto email with change link
- Itinerary/reminder
- Push to loyalty sign-in

Why It Works:

Gentle nudge keeps intent without heavy discounts

Low Risk
Score < 0.36

What We Do:

- Standard confirmation only

Why It Works:

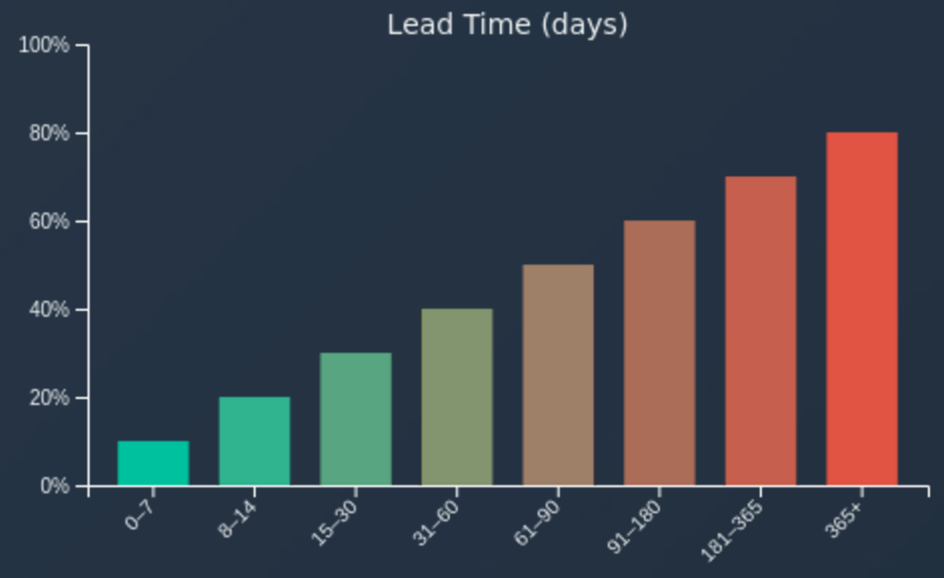
Saves effort and incentives for where they matter

🎙️*This grid is the playbook. The booking's risk band decides the action—no guesswork. High risk gets a proactive touch and price protection; medium gets automated reassurance; low risk gets the basics so we don't overspend.*

🎙️*Bands derived from calibrated risk on the holdout set; operating threshold ≈ 0.36*

Where Risk Concentrates

Cancellation Risk by Lead Time



Key Insight: Cancellation risk rises with longer lead time

Bookings with very long lead times (365+ days) exhibit a higher propensity for cancellation.

Cancellation Risk by market_segment_type

Key Insight: A few online/OTA channels carry the highest risk

OTAs like Expedia and Booking.com consistently demonstrate higher cancellation rates compared to others.

Focus interventions on **long-lead bookings** and **high-risk channels**

What We Built & Why

Target

Probability that a booking will be canceled **at the time of reservation**

(booking_status = cancel)

Key Features

- **lead_time**
  Days between booking and arrival
- **market_segment_type**
  OTA vs direct bookings
- **average_price (ADR)**
  Avg daily rate
- **weekend/weekday mix**
  Proportion of weekend/day nights
- **special_requests (Count 0- 5)**
  Guest requests presence/absence
- **car_parking_space**
  Parking space requested

Model Selection

After testing several algorithms, **LightGBM** was selected as the optimal model:

- LightGBM
- Histogram GBM
- Random Forest
- Logistic Regression

Why LightGBM?

- Superior accuracy and PR performance
- Fast processing of features
- Reliable after calibration
- Clear SHAP explanations

Test KPIs

ROC-AUC	PR-AUC	F1	Threshold
0.936	0.896	0.801	0.41

What Drives Cancellation Risk

HOTEL HAVEN

Analysis of the SHAP summary reveals the primary drivers influencing hotel booking cancellation risk: Drivers computed with SHAP on the LightGBM model

↑ higher risk ↔ neutral/mixed ↓ lower risk

These aren't just patterns—they're levers we can act on

Lead time ↑

Longer lead times significantly increase cancellation likelihood, as bookings made far in advance carry higher uncertainty.

Online/OTA ↑ risk

Bookings from OTAs like Expedia or Booking.com show higher cancellation risk, due to flexible booking terms that make cancellations simpler.

Weekend share × price ↔

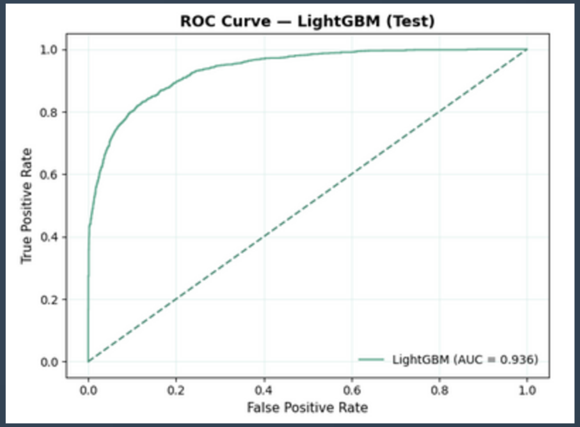
A higher proportion of weekend nights combined with increased ADR indicates pockets of elevated cancellation risk.

Commitment Signals ↓ risk

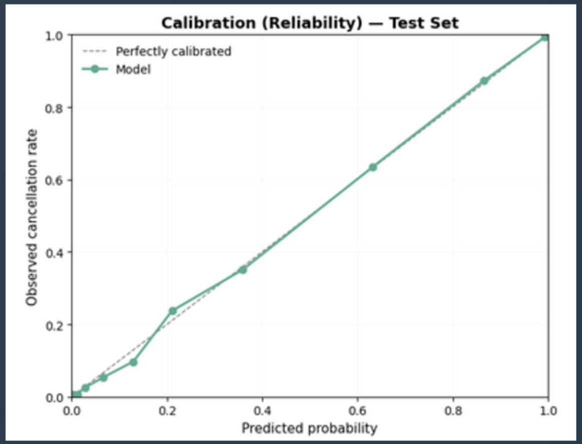
Special requests (and usually **parking requests**) correlate with lower cancel risk—guests are more committed.

How Good Are the Scores?

ROC Curve (Ranking Test)



Calibration Curve (Honesty Test)



ROC Curve: Overall ranking quality. **AUC 0.936**

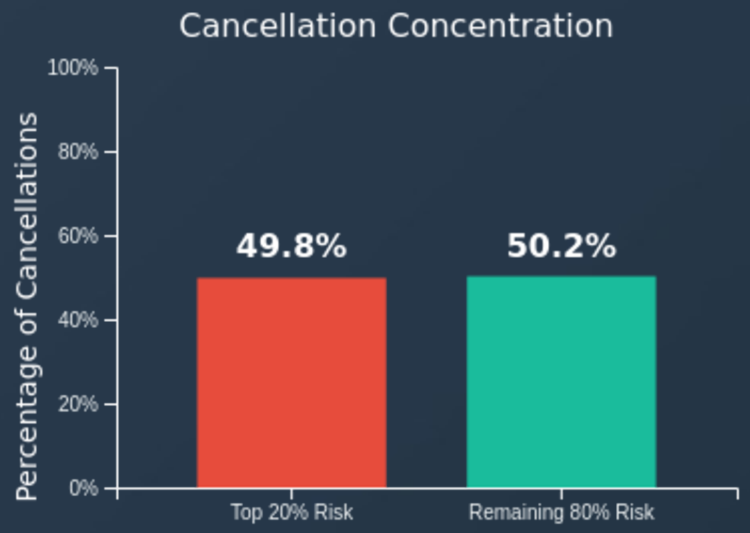
Good model: the model clearly distinguishes canceled vs not.

Calibration: Thresholds are trustworthy, **AP 0.896**

Well-calibrated model: high precision at workable recall.

A good model will help operations identify the right bookings to target with limited resources.

Precision Beats Blanket Discounts



Top 20% of riskiest bookings = **1,814** bookings

49.8% of all cancellations

come from the top 20% riskiest bookings

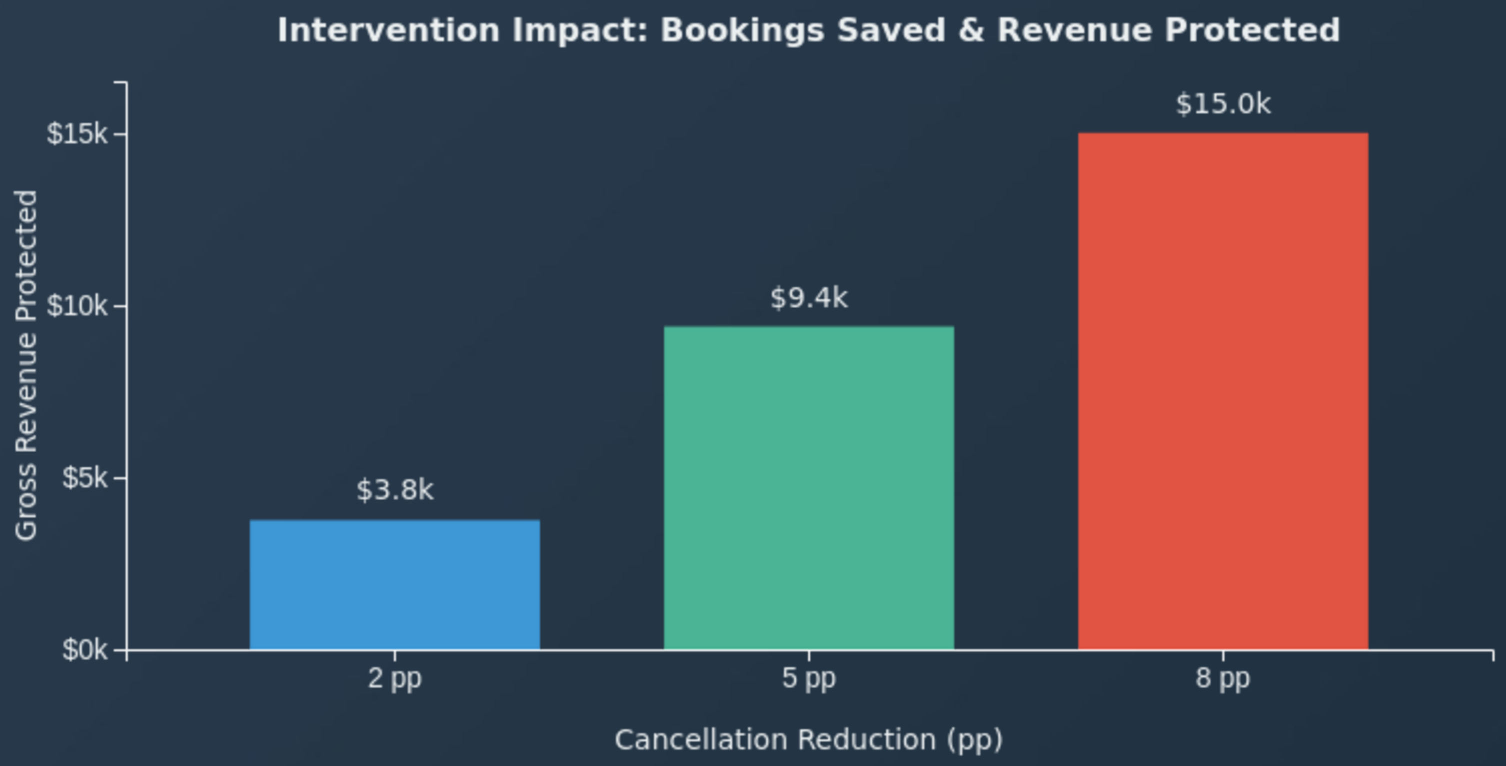
Targeted Efficiency
Focus resources where they'll have the biggest impact

Cost Optimization
Light-touch interventions capture big wins with small spend

Revenue Protection
Shield valuable stays from cancellations

Pilot Impact on the Top 20%

1,814 high-risk bookings (**~49.8%** of all cancels)



Net ROI Calculation

Net ROI = Gross protected − Program cost

What percentage point reduction equals your cost?

Even small reductions add up to significant revenue protection

pp = percentage points (e.g., 40% → 38% = 2 pp)

Assumptions: revenue protected ≈ ADR per booking (~$105, ~1 night); reduction applied to top-20% cohort only. “pp” = percentage points.

Pilot, Measure, Scale

Weeks 0-2

- Connect scoring to CRM
- Stand up dashboards: Risk Funnel, Save Rate, Net ROI.

Weeks 3-6

- Run playbooks on top 20% risk
- Track saved bookings & feedback

Weeks 7-12

- Scale to more dates/segments
- Report net ROI by risk band monthly

Our Ask

We request approval for a **two-property pilot program**, with further expansion contingent upon the verified net ROI demonstrated during this initial pilot phase.

This approach allows us to validate our solution in a controlled environment before scaling across all properties.

Next steps: Implement the pilot program and prepare for quarterly expansion based on ROI results.

HOTEL
HAVEN

THANK

YOU