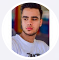
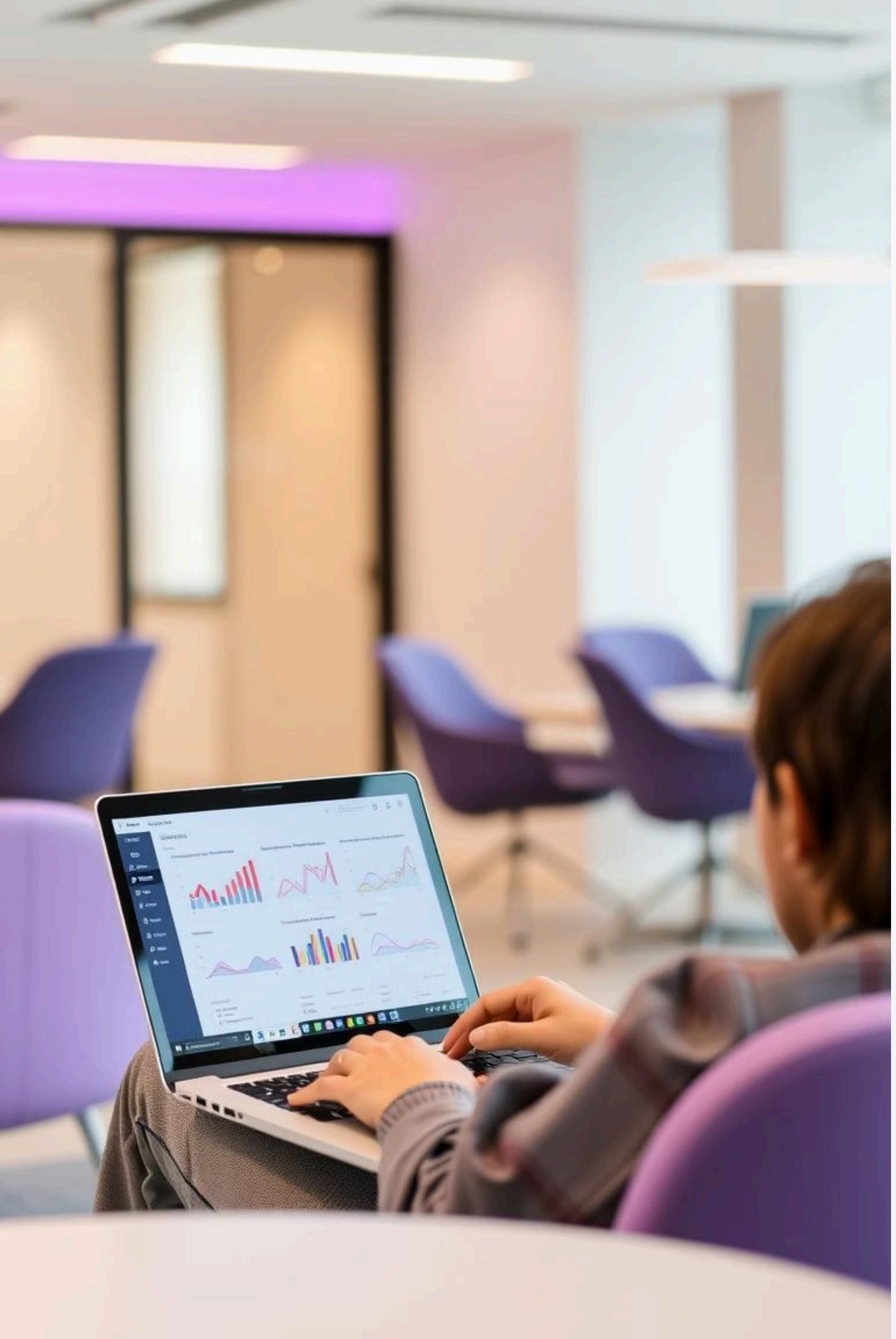
Hotel Analysis Project Overview

Team members and supervisor details

Platform: Google Colab, Streamlit, Hugging Face

Date: [Presentation Date]

 **by Mhammed Zain**



Project Motivation & Tools

Challenge

Hotels losing revenue from cancellations

Goal

Predict cancellations & enable proactive actions

Tools

Python, Power BI, Streamlit, Machine Learning

Project Objectives

Data Analysis

Understand booking patterns and cancellation trends

Key Insights

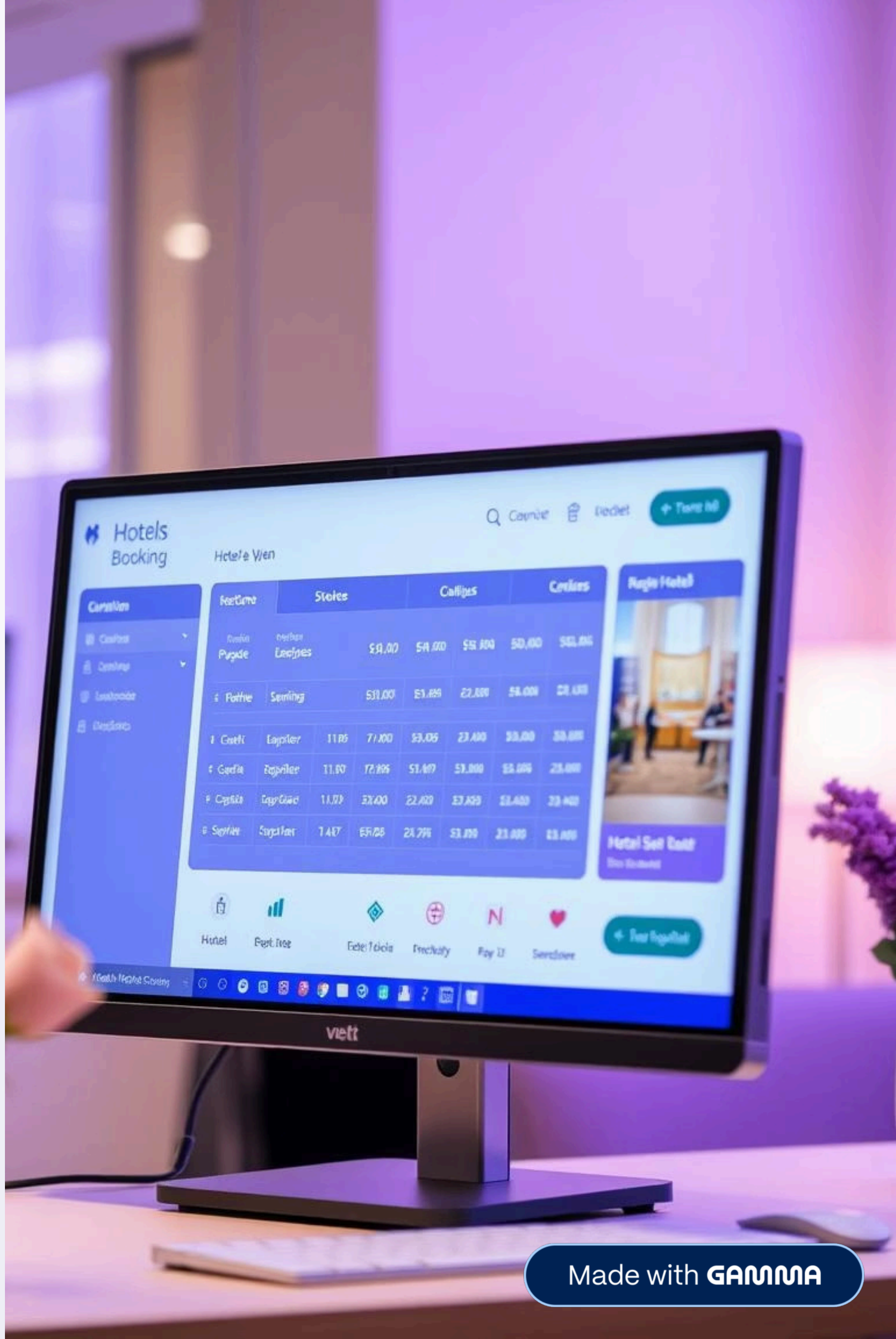
Identify factors impacting cancellations

Predictive Model

Build ML model to forecast cancellations

Deployment

Deliver application for real-time use



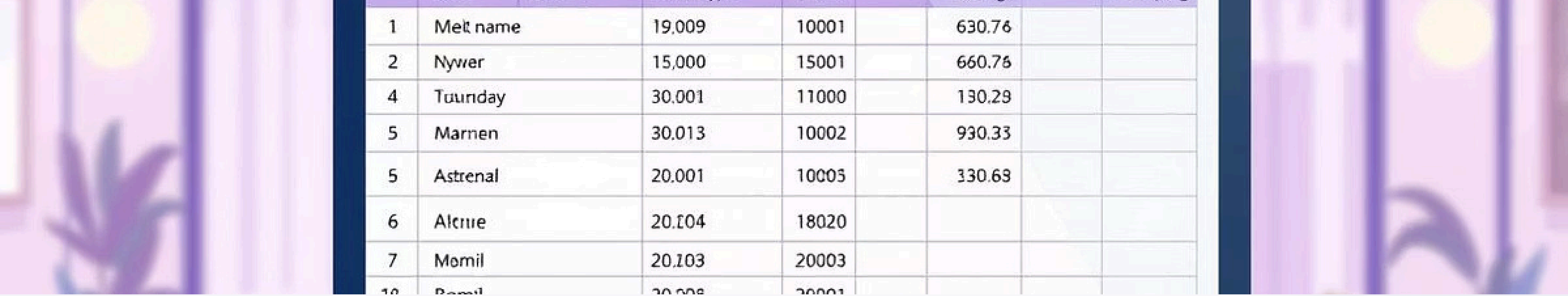
Tools & Environment

Platform & Deployment

- Google Colab for development
- Streamlit + Hugging Face Spaces for deployment

Technologies

- Python, Pandas, Seaborn, Matplotlib
- Scikit-Learn, XGBoost, LightGBM
- Power BI for visualization



1	Melt name	19,009	10001		630.76		
2	Nywer	15,000	15001		660.76		
4	Tuunday	30.001	11000		130.29		
5	Marnen	30.013	10002		930.33		
5	Astrenal	20.001	10005		330.63		
6	Alcrue	20.104	18020				
7	Momil	20.103	20003				

Data Overview

119,000+ booking records analyzed

- Booking dates and guests
- Room types and seasons
- Cancellation status

Insight #1: Guests Mostly Couples

1. **Booking Pattern**

   Predominantly 2 guests per booking

2. **Guest Type**

   Mostly couples, few with children

3. **Recommendations**

   Create honeymoon and couple packages

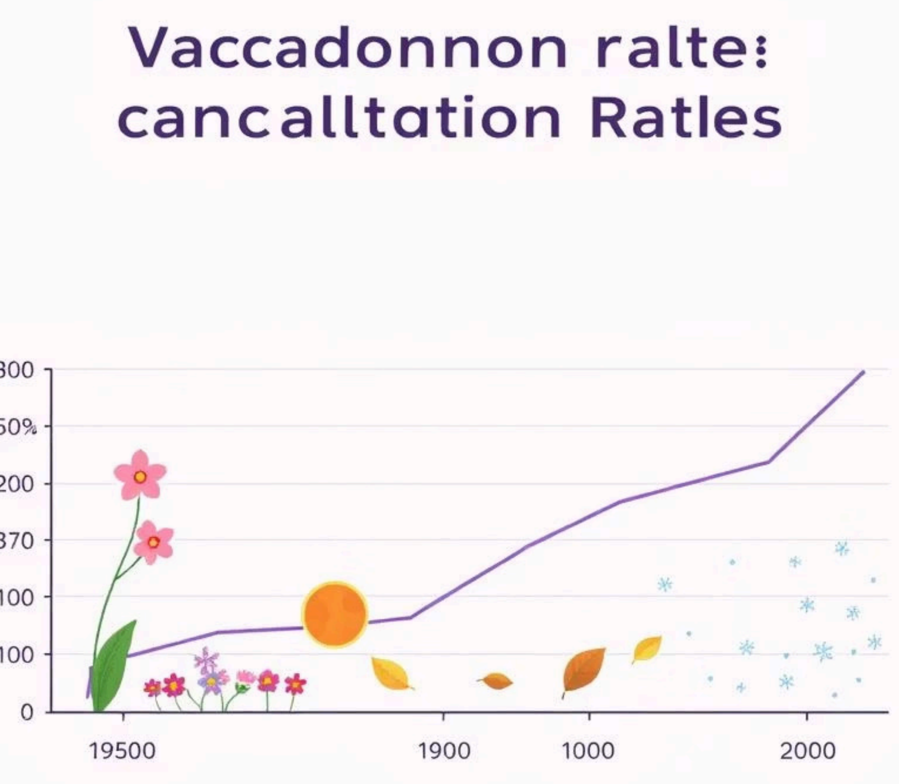
Insight #2: Peak Booking Months

Peak Season

August to October: End of summer, early fall

Recommendations

- Promote premium packages during peak months
- Upsell extra services to increase revenue

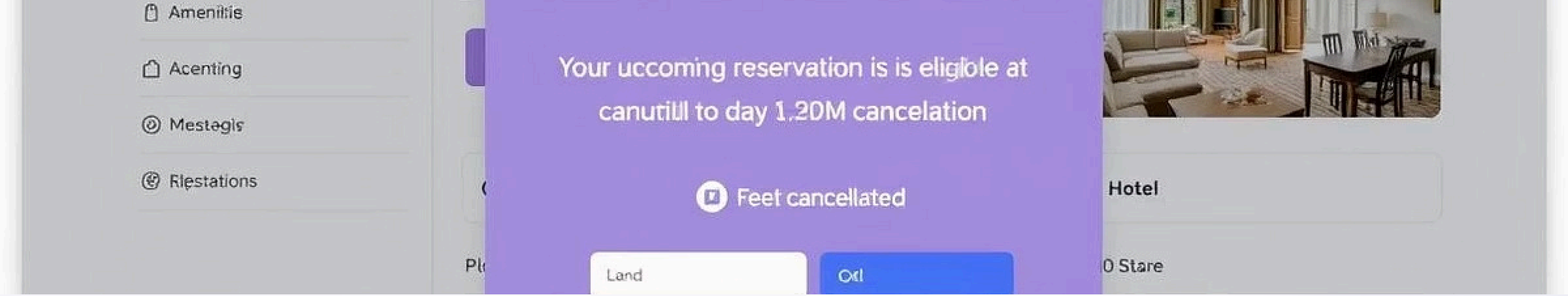


Insight #3: Seasonal Cancellation Trends

Summer	41.19% (Highest)
Winter	14.89% (Lowest)

Summer: limit free cancellations, offer discounts on non-refundable

Winter: allow flexibility, slightly increase prices



Insight #4: Booking Source & Cancellation

Most bookings online

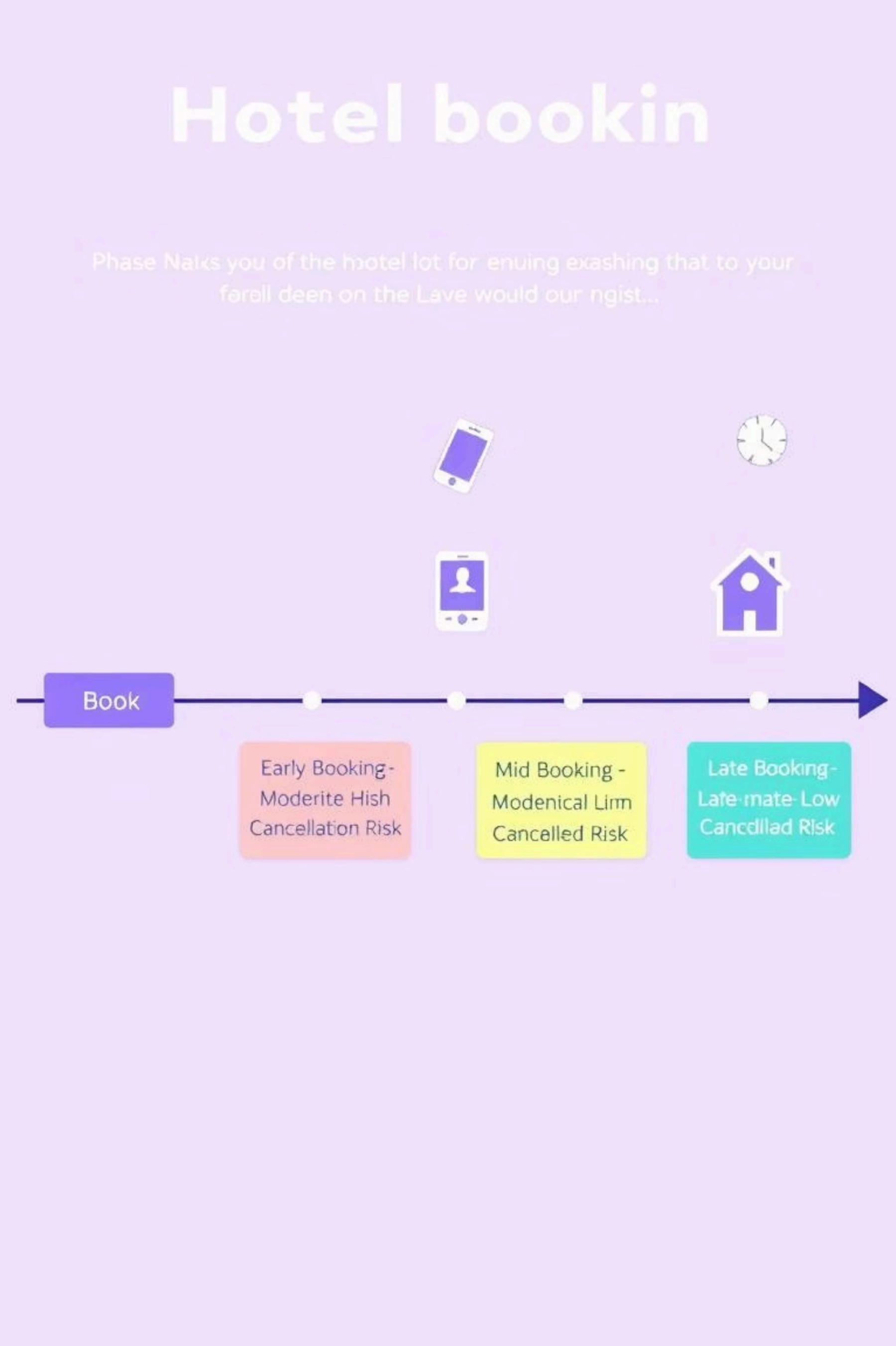
Weekend bookings cancel more

Policy Recommendation

Stricter rules for weekend bookings

Promotions

Loyalty discounts for direct bookings



Insight #5: Early Bookings Often Cancel

Early Bookings

Higher cancellation likelihood

Recommendations

Early non-refundable discounts

Engagement

Reminders to reduce cancellations

Insight #6: Repeat Guests Behavior

Guest Loyalty Patterns

Most repeat guests do not cancel their bookings.

They tend to visit during fall and winter seasons.

Recommendations:

- Develop targeted loyalty programs.
- Offer exclusive winter packages to returning guests.

Insight #7: Company/Business Sector

- **Low Cancellation Rate:** Business guests rarely cancel their bookings.
- **Corporate Partnerships:** Focus on building long-term collaborations with companies.
- **Tailored Packages:** Offer corporate packages and flexible invoicing options.

Insight #8: Parking Request Clients

Low Cancellation Rate

Guests requesting parking almost never cancel their bookings.

Value-Added Service

Offer guaranteed parking to enhance guest satisfaction.

Marketing Trust Signal

Highlight parking availability to boost booking confidence.

Insight #9: Room Preferences

- Majority prefer Room Type 1 and Meal Type 1 options
- Bookings favor medium-tier rooms balancing comfort and price
- Optimize inventory for the most popular room types
- Encourage upgrades to higher tiers with modest fees

Insight #10: Guest Type

- Majority of bookings come from new guests seeking first experience
- Returning guests are fewer but show higher booking reliability
- Enhance first-time guest experience to encourage loyalty
- Implement post-stay engagement to build lasting relationships

Insight #11: High Cancellation Guests

Guests who have canceled before are more likely to cancel again.

To reduce risk, implement stricter policies for these guests.

Consider requiring deposits to secure bookings reliably.

Machine Learning Pipeline Overview

1 **Data Cleaning & Preprocessing**

Remove duplicates, handle missing data, and format inputs for analysis.

2 **Feature Engineering**

Create and select key variables to improve model performance.

3 **Model Training**

Train algorithms on prepared data to recognize patterns.

4 **Evaluation**

Assess models using accuracy, precision, recall, and F1 scores.

5 **Model Comparison & Selection**

Compare algorithms; select best-performing model, such as LightGBM.

6 **Deployment**

Implement model in real-world system for ongoing predictions.

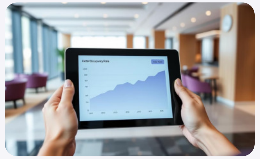
Models Compared

Model	Train Accuracy	Test Accuracy	Train F1	Test F1	Train ROC AUC	Test ROC AUC	Time (seconds)
ExtraTreesClassifier	0.9945	0.9723	0.9945	0.9791	0.9999	0.9713	33.77
RandomForestClassifier	0.9945	0.9719	0.9945	0.9788	0.9995	0.9697	48.10
DecisionTreeClassifier	0.9945	0.9654	0.9945	0.9739	0.9999	0.9638	2.53
BaggingClassifier	0.9875	0.9617	0.9875	0.9710	0.9991	0.9606	8.49
XGBClassifier	0.9047	0.8901	0.9059	0.9166	0.9719	0.8820	6.77
KNeighborsClassifier	0.9014	0.8830	0.8992	0.9089	0.9721	0.8863	92.00
GradientBoostingClassifier	0.8301	0.8240	0.8324	0.8631	0.9166	0.8193	30.46
AdaBoostClassifier	0.7788	0.7666	0.7783	0.8138	0.8628	0.7672	7.69
LogisticRegression	0.7736	0.7564	0.7684	0.8016	0.8552	0.7653	11.42
GaussianNB	0.5513	0.4103	0.2295	0.2295	0.7755	0.5488	0.58

Deployment of Cancellation Prediction Model

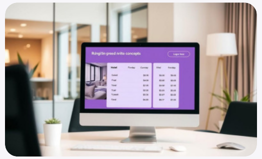
- **Interface:** Built with Streamlit for simplicity and accessibility.
- **Hosting:** Deployed on Hugging Face Spaces for reliable cloud access.
- **User Interaction:** Staff input booking data and receive instant predictions.
- **Use Case:** Enables hotel staff to make fast, informed decisions.

Business Impact of Cancellation Prediction



More Predictable Occupancy

Improved forecasting enables smarter planning and resource allocation.



Dynamic Pricing Offers

Tailored prices boost revenue by targeting customers effectively.

Reduced Refund Losses

Accurate predictions minimize costly unexpected cancellations and refunds.

Better Marketing Strategies

Data-driven insights optimize campaigns and increase guest engagement.

Project Summary & Key Outcomes

Insights Uncovered

Seasonality, guest type, and booking timing impact cancellations.

Powerful Prediction Model

Developed an accurate machine learning model for cancellations.

Effective Deployment

Created a user-friendly Streamlit app for easy staff use.

Actionable Recommendations

Provided strategies to reduce cancellations and boost loyalty.

Future Work

Payment Gateway Integration

Add secure payment processing to streamline bookings and cancellations.

Continuous Model Retraining

Use live hotel data to keep the prediction model accurate and up to date.

Expansion to More Hotels

Deploy the system across multiple hotels and locations for broader impact.

Thank You