

Presented by: NTI G2

september ,
09/2025



HOTEL RESERVATIONS

Can you predict if customer is going to cancel the reservation ?



Agenda:

- A THE PROBLEM
- B SOLUTION
- C PROJECT GOAL
- D DATA DEFINITION
- E METHODOLOGIES
- F WORKFLOW
- G CONCLUSION
- H RESOURCES



THE PROBLEM



- Hotels face a big challenge with cancellations and no-shows. This leads to wasted rooms, poor resource planning, and revenue loss.
- Our dataset shows booking details (adults, children, room type, lead time, special requests, etc.) that can help us understand why customers cancel



SOLUTION

- We can use Machine Learning models (Decision Trees, Random Forest, SVM, etc.) to predict whether a reservation will be cancelled or not.
- In our project, we trained different models. Random Forest gave the best results with accuracy around 90%. We also checked feature importance to see which factors matter most (like lead time, price, and special requests)

PROJECT GOAL

Deliver an intelligent solution that helps hotels:

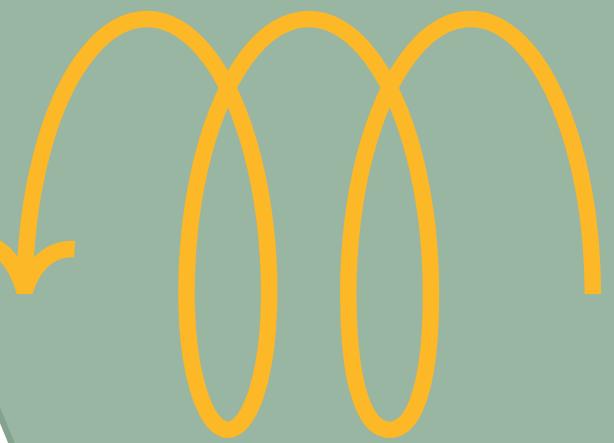
The goal is to build a predictive model that helps hotels forecast cancellations before they happen

This will allow hotels to make better decisions in pricing, room allocation, and staff planning
EX:

- If the model predicts high cancellation, the hotel can overbook to reduce losses.



ABOUT DATA



The dataset has 36,275 records and 19 features. It contains both numerical and categorical variables , We made The dataset cleaned (no missing values). Most bookings are for 2 adults, few with children, and usually for short stays.



The data is divided into four sections →

Guests Information	no_of_adults	repeated_guest	no_of_children
Booking & Stay	no_of_week_nights	Lead Time	arrival_date
Services & Room	market_segment_type	room_type_reserved	type_of_meal_plan
Outcome	booking_status		

METHODOLOGIES

WITH
PYTHON
LIBRARIES

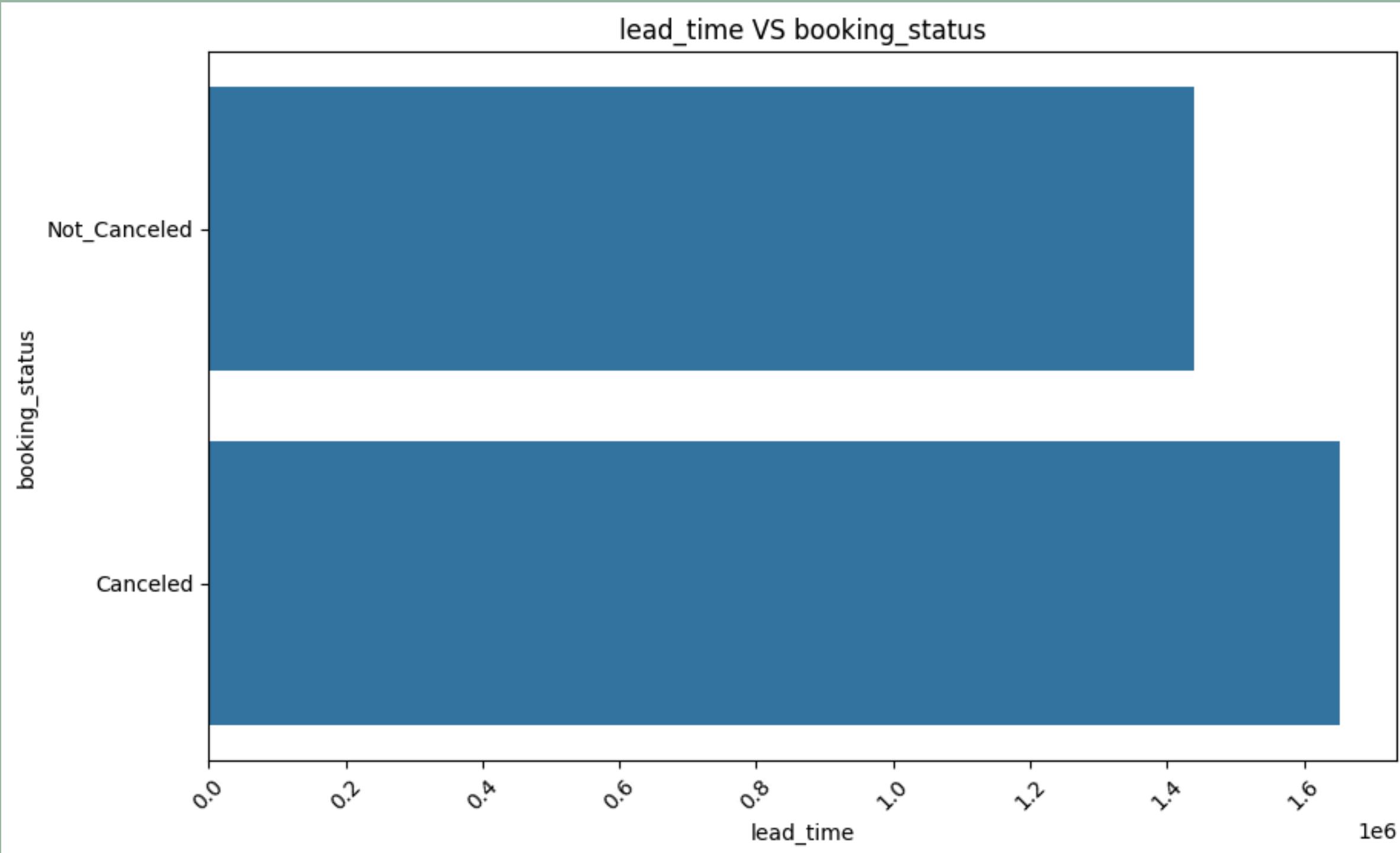
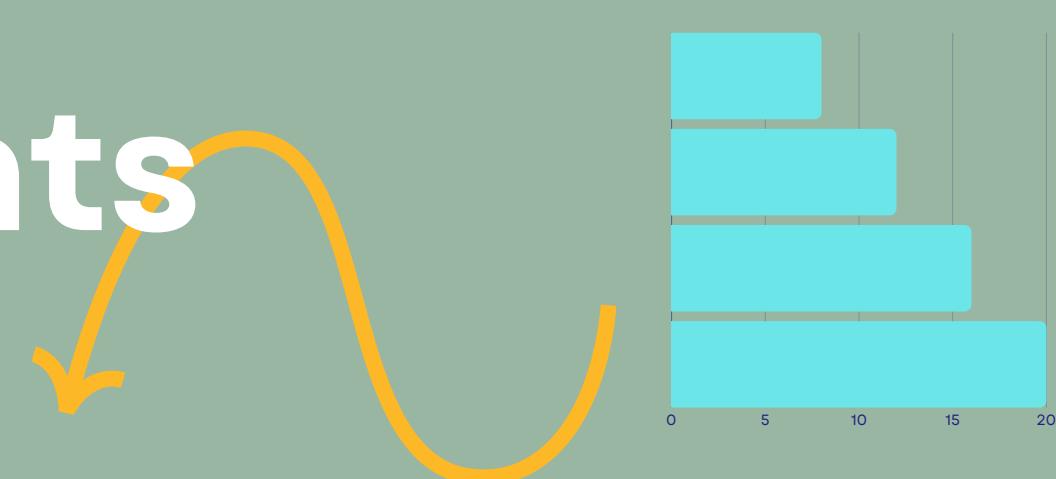
THE MOST
IMPORTANT
WORKS

- Reading and understanding data-> Understanding data aspects
- Data Analysis -> Explore sales by time, product, and region.
- Data Cleaning -> Handle missing values & anomalies



- Business Insights-> Support smarter inventory and sales decisions
- Modeling → Build classification (sales levels) & anomaly detection models.
- Evaluation → Measure accuracy & reliability of predictions.

SOME ANALYSIS Insights

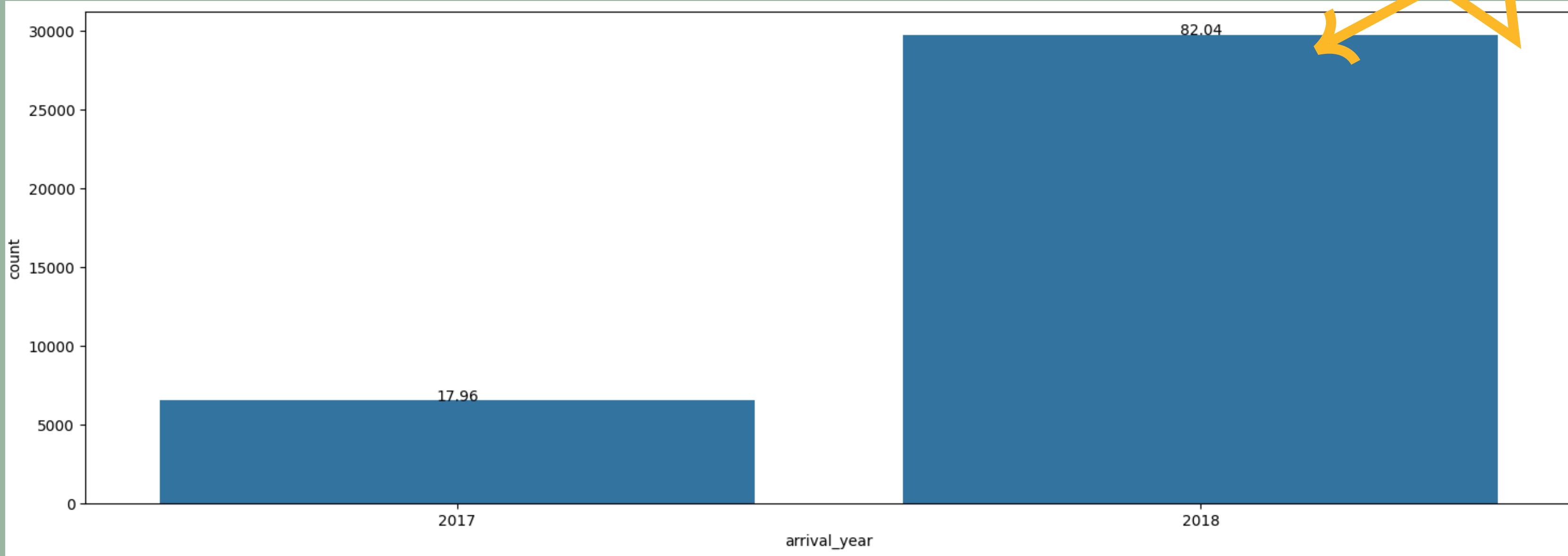


1. lead time Vs booking status by(Bar Chart)

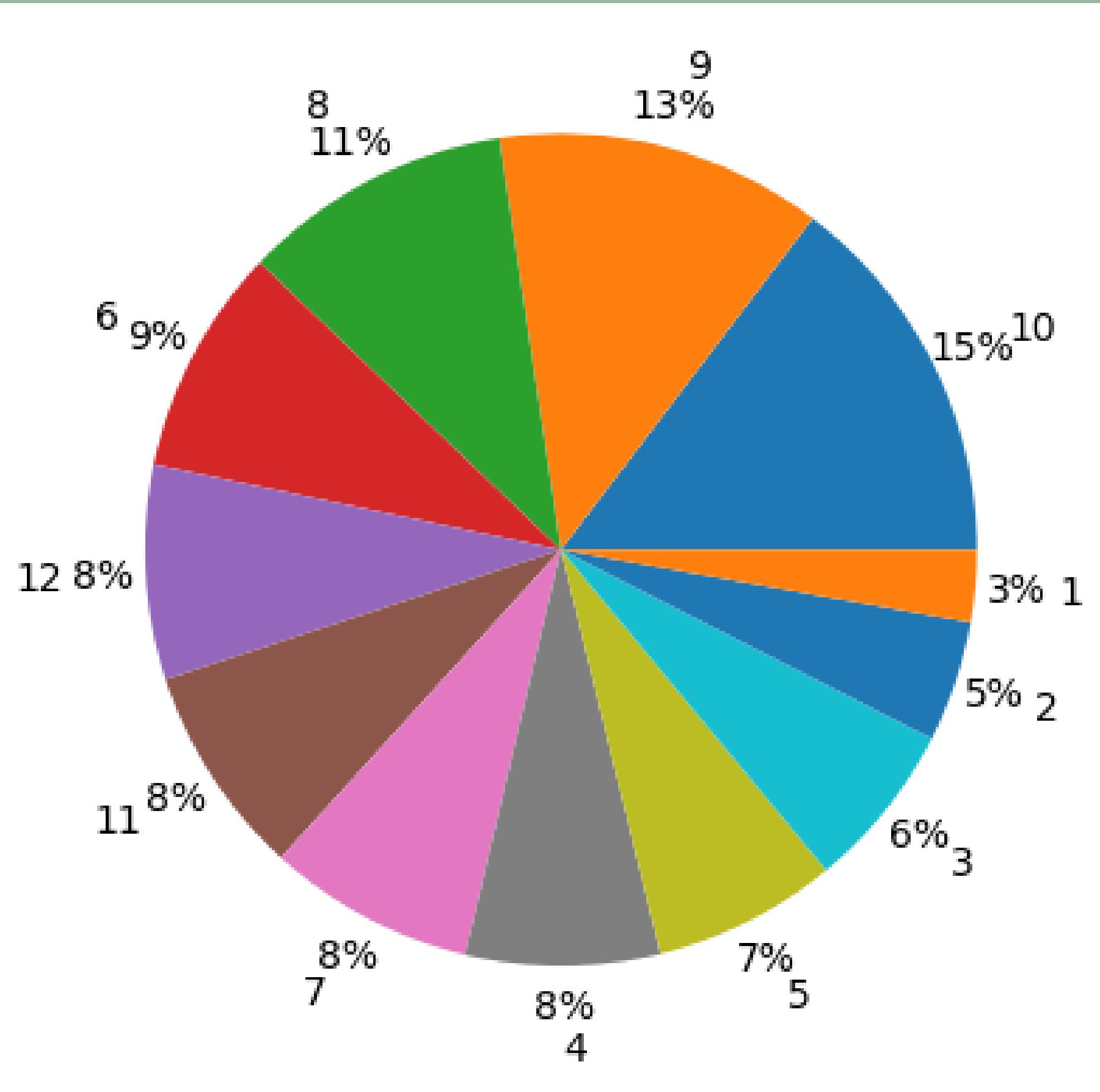
This chart shows the relation between lead time and booking status.

(If someone books 6 months before arrival, chances of cancellation are higher).

2. The most busiest year (bar Chart)



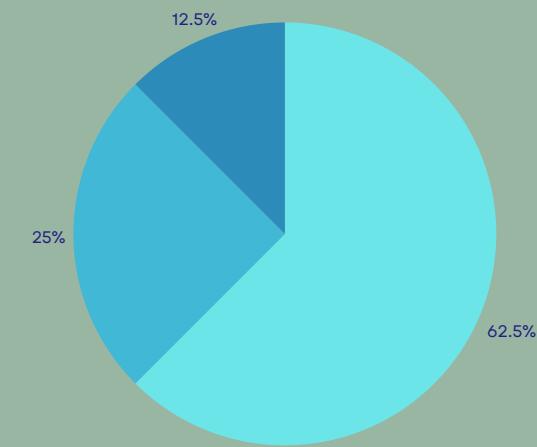
 This chart highlights the most year the hotel has guests.



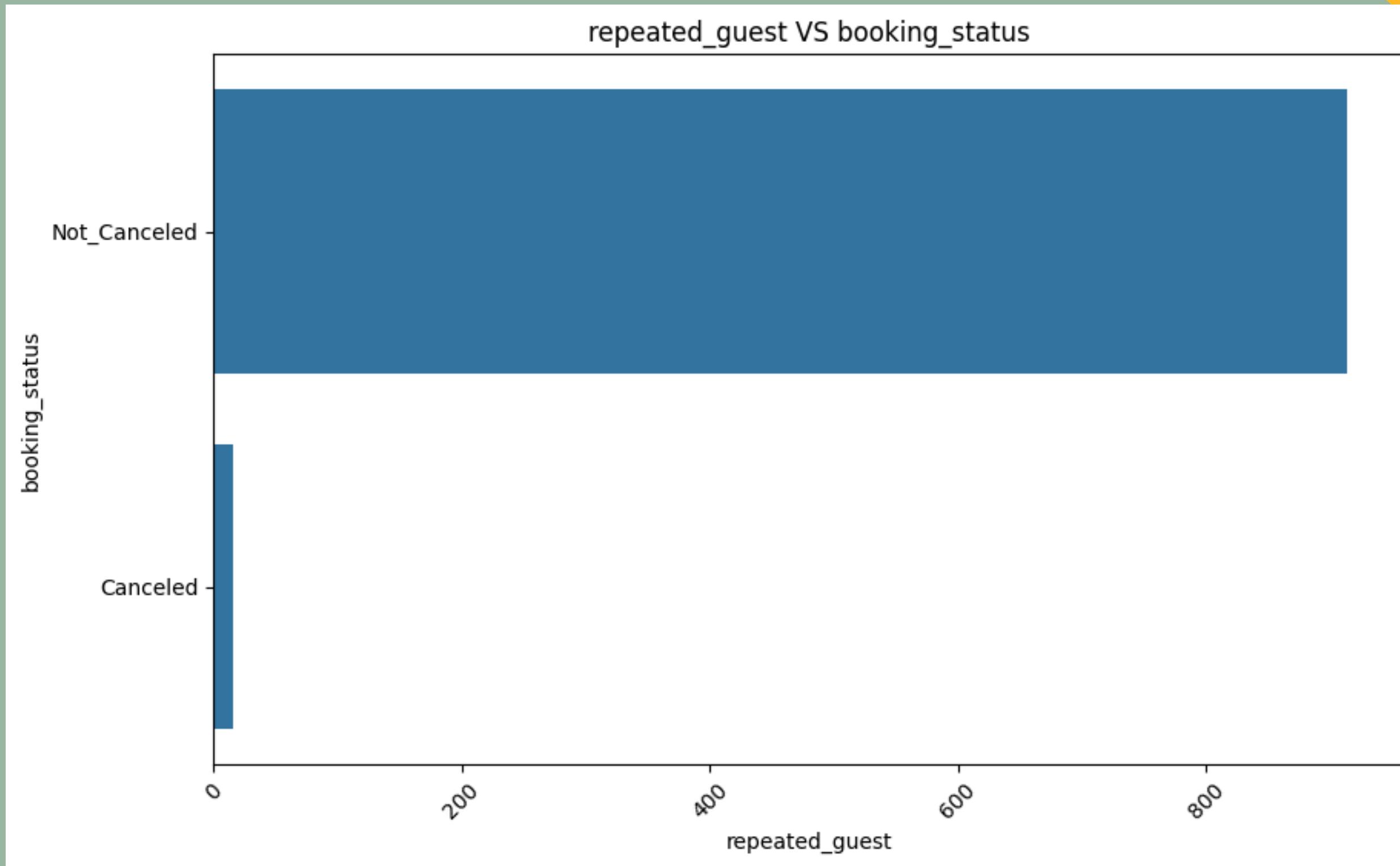
3. The busiest month(Pie Chart)

► This chart shows the most month the hotel has guests.

Maybe in October there was a holiday or an official occasion

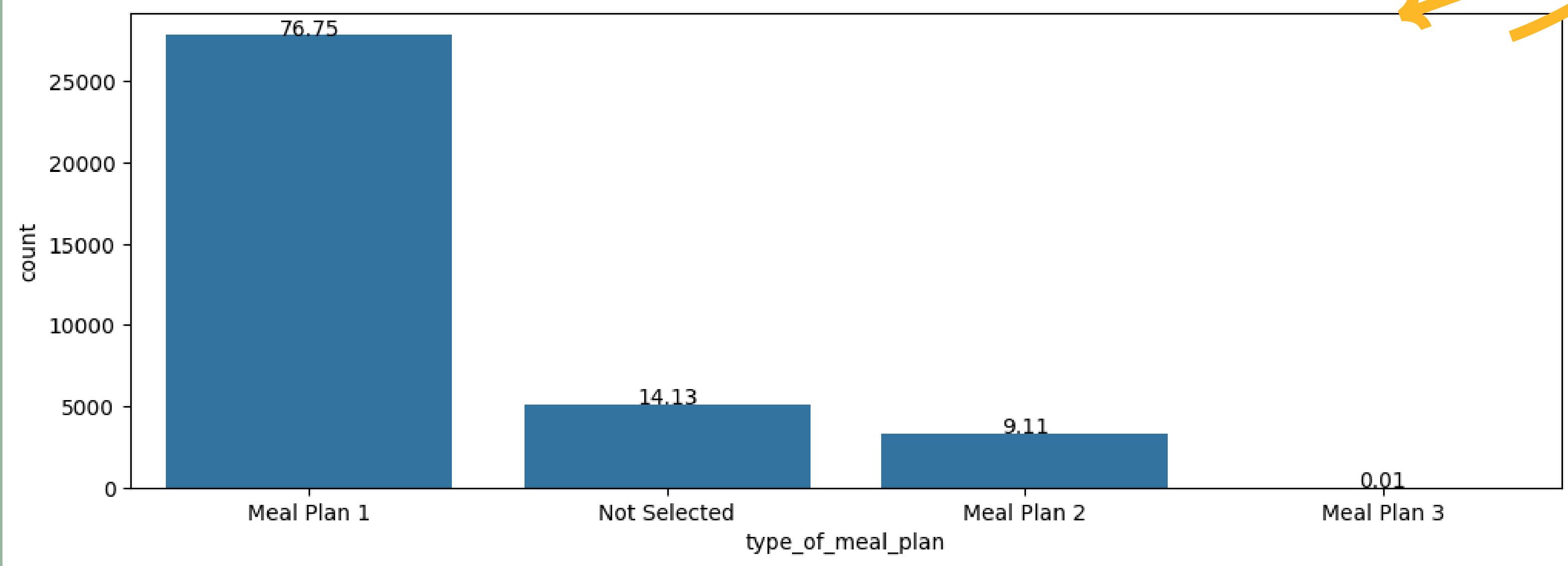


4. Repeated guest VS booking status (bar chart)



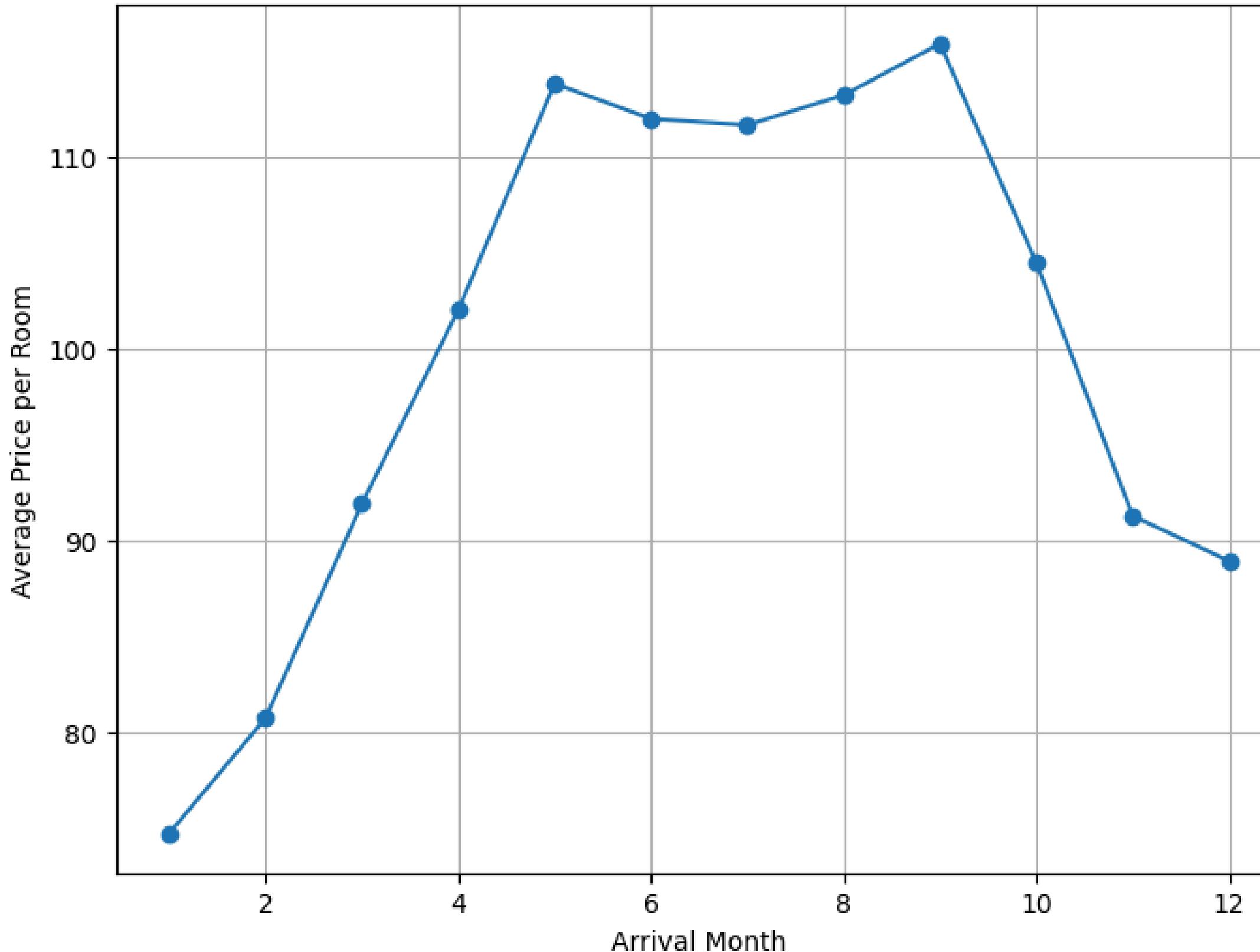
⚡ This chart shows that most of the previous guests come back again. This indicates that they liked the services of this hotel.

5. Count of meals types(Bar Chart)



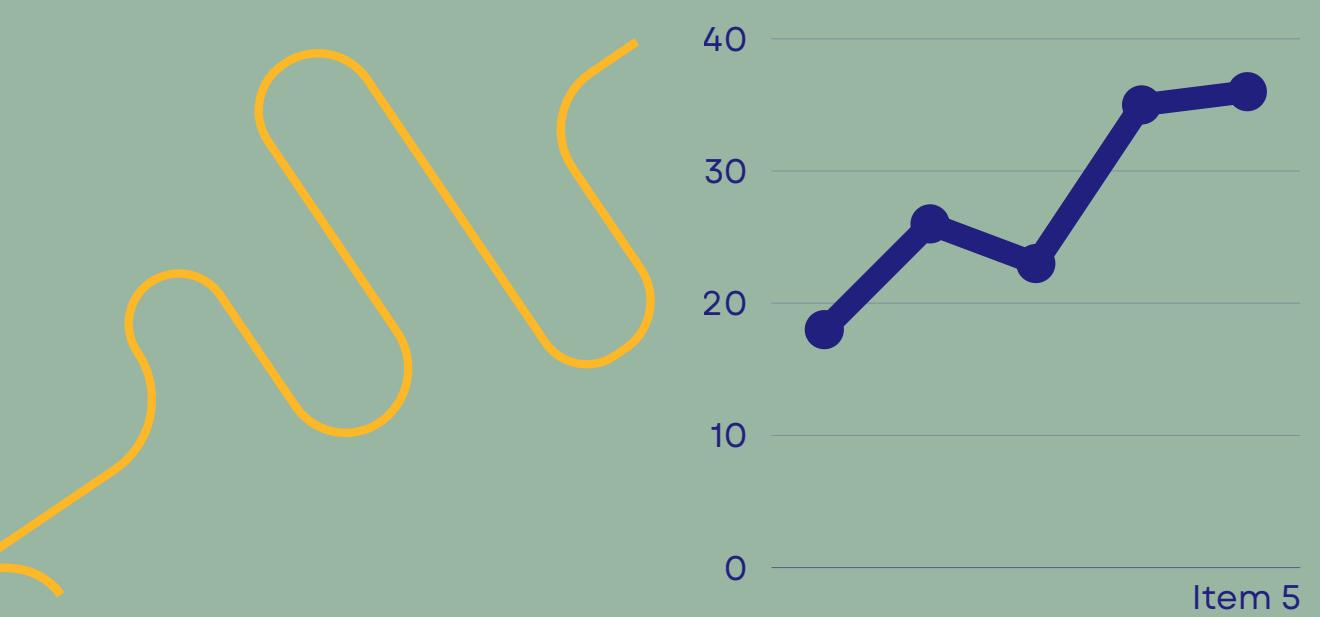
This chart It indicates that most of the reservations were for meal type No. 1. This means that most of the visitors liked it and they should continue to make it.

Average Price per Room by Arrival Month



7. Arrival month VS Avg price per room(line Plot)

The line shows that in September the rooms have the highest price, Of course, this time is an official holiday, so they must take advantage of it





Market segment type VS Booking status(count plot)

WORKFLOW

- ENG Amina → EDA&Analysis,
Presentation
- ENG Mai → Preprocessing, Data
cleaning
- ENG Ahmed → Building ML models,
Streamlit



CONCLUSION

Our project is projected to:

Predicting cancellations helps hotels improve their efficiency and revenue management.

- Random Forest gave the best performance in our tests and gave us 90% accuracy.
- The most important features were: lead time, special requests, and average price per room.
- Even using only a few key features, the model still performed well.



RESOURCES

[HTTPS://DOCS.PYTHO
N.ORG/3/](https://docs.pytho
n.org/3/)

[HTTPS://MATPLOTLI
B.ORG/](https://matplotlib.
org/)

[https://www.kaggle.com/datasets/ahsa
n81/hotel-reservations-classification-
dataset](https://www.kaggle.com/datasets/ahsa
n81/hotel-reservations-classification-
dataset)

[HTTPS://SCIKITLEARN.ORG
/STABLE/](https://scikitlearn.org
/stable/)

<https://pandas.pydata.org/>

[HTTPS://NUMPY.ORG/](https://numpy.org/)

[HTTPS://STREAMLIT.IO](https://streamlit.io/)



THANK YOU

