

SAMAA OSAMA SALLAM

**WEEK 1 OF CELLULA INTERNSHIP
EDA PRESENTATION – HOTEL
RESERVATION DATA**



EXPLORATORY DATA ANALYSIS – HOTEL RESERVATION DATASET OVERVIEW

- 36,285 rows × 17 columns covering hotel reservation records.
- Includes details on guests, room types, pricing, booking status, and special requests.
- Key fields: number of guests, lead time, average price, and cancellation status.
- Date range based on date_of_reservation, allowing time-based trend analysis.
- Enables insights into guest behavior, booking patterns, and cancellation trends.

KEY FEATURES IN THE DATASET

- Booking ID, number of guests, lead time
- Meal type, room type, market segment
- Special requests, price, booking status

	Booking_ID	number of adults	number of children	number of weekend nights	number of week nights	type of meal	car parking space	room type	lead time	market segment	type	repeated	P-C	P-not-C	average price	special requests	date of reservation	booking status
0	INN00001	1	1	2	5	Meal Plan 1	0	Room_Type_1	224	Offline	0	0	0	88.00	0	10/2/2015	Not_Canceled	
1	INN00002	1	0	1	3	Not Selected	0	Room_Type_1	5	Online	0	0	0	106.68	1	11/6/2018	Not_Canceled	
2	INN00003	2	1	1	3	Meal Plan 1	0	Room_Type_1	1	Online	0	0	0	50.00	0	2/28/2018	Canceled	
3	INN00004	1	0	0	2	Meal Plan 1	0	Room_Type_1	211	Online	0	0	0	100.00	1	5/20/2017	Canceled	
4	INN00005	1	0	1	2	Not Selected	0	Room_Type_1	48	Online	0	0	0	77.00	0	4/11/2018	Canceled	
...	
36280	INN36282	2	0	0	2	Meal Plan 2	0	Room_Type_1	346	Online	0	0	0	115.00	1	9/13/2018	Canceled	
36281	INN36283	2	0	1	3	Meal Plan 1	0	Room_Type_1	34	Online	0	0	0	107.55	1	10/15/2017	Not_Canceled	
36282	INN36284	2	0	1	3	Meal Plan 1	0	Room_Type_4	83	Online	0	0	0	105.61	1	12/26/2018	Not_Canceled	
36283	INN36285	3	0	0	4	Meal Plan 1	0	Room_Type_1	121	Offline	0	0	0	96.90	1	7/6/2018	Not_Canceled	
36284	INN36286	2	0	0	5	Meal Plan 1	0	Room_Type_4	44	Online	0	0	0	133.44	3	10/18/2018	Not_Canceled	

36285 rows × 17 columns

DATA CLEANING STEPS

01 : STANDARDIZED COLUMN NAMES

- Converted all column names to lowercase and replaced spaces with underscores for consistency:

02 : RENAMED AMBIGUOUS COLUMNS FOR CLARITY

- 'p-c' → 'Preservation_Canceled'
- 'p-not-c' → 'Preservation_not_Canceled'
- 'repeated' → 'Visited_Before'

03 : CHECKED FOR MISSING VALUES:

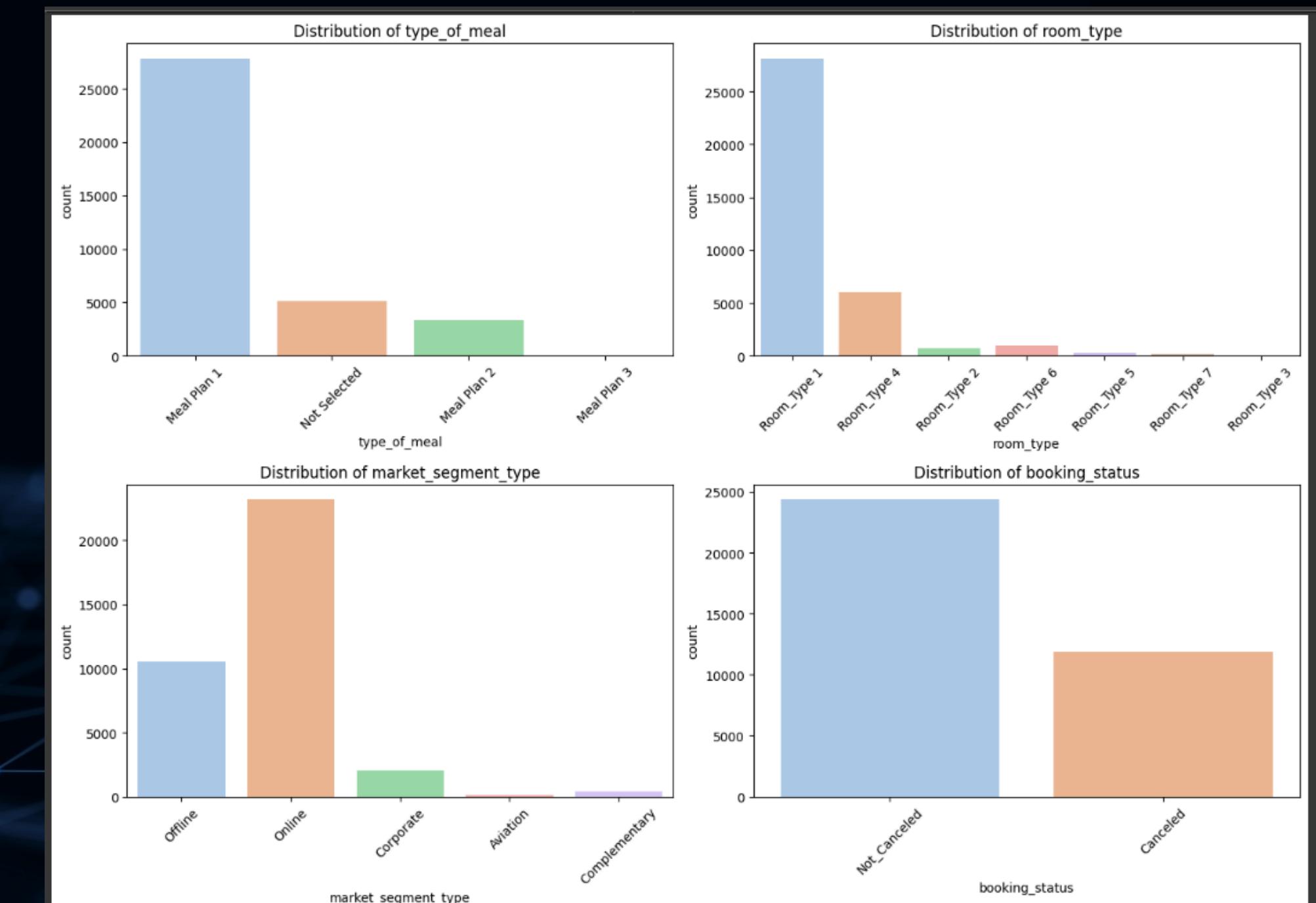
- Nothing Found

04 : CHECKED FOR DUPLICATE RECORDS:

- Nothing Found

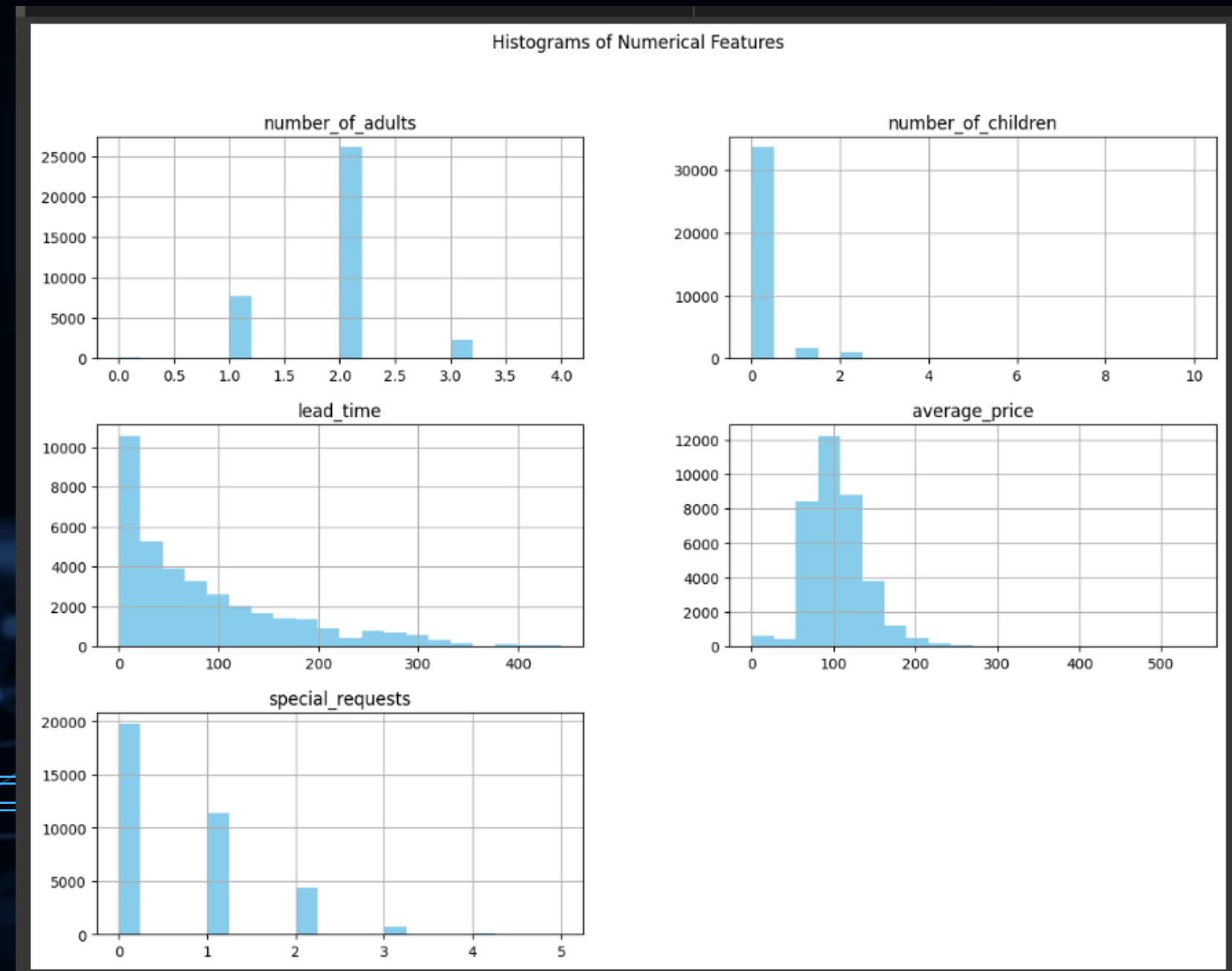
UNIVARIATE ANALYSIS - CATEGORICAL DATA

- Meal_Type 1 is the dominant choice among guests, accounting for more than half of all bookings.
- Room_Type 1 is the most frequently booked room
- Online bookings are the most common market segment
- The majority of bookings are not canceled, with a cancellation rate of ~ 22%.



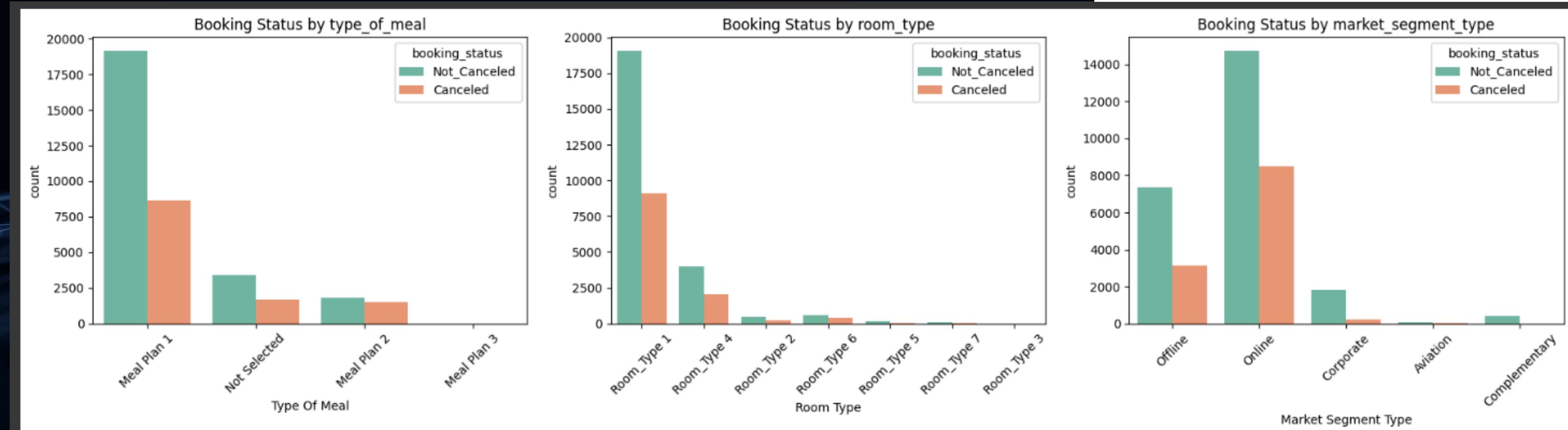
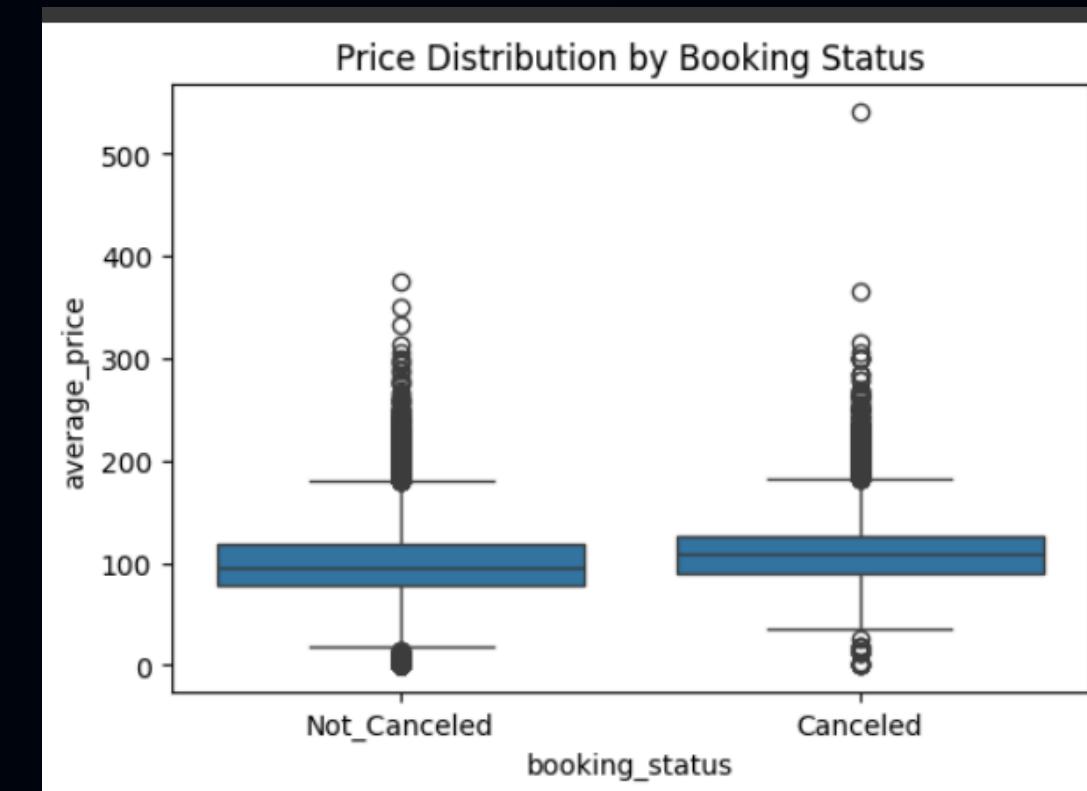
UNIVARIATE ANALYSIS - NUMERICAL DATA

- Most bookings are made for 2 adults, with very few involving 3 or more.
- Children are not included in most bookings
- Lead time (days before check-in) is right-skewed — most bookings are made close to check-in, but a few are made well in advance (100+ days).
- Average price per reservation is mostly between 50 and 150 units
- Special requests are rare — over 60% of bookings include 0 requests,

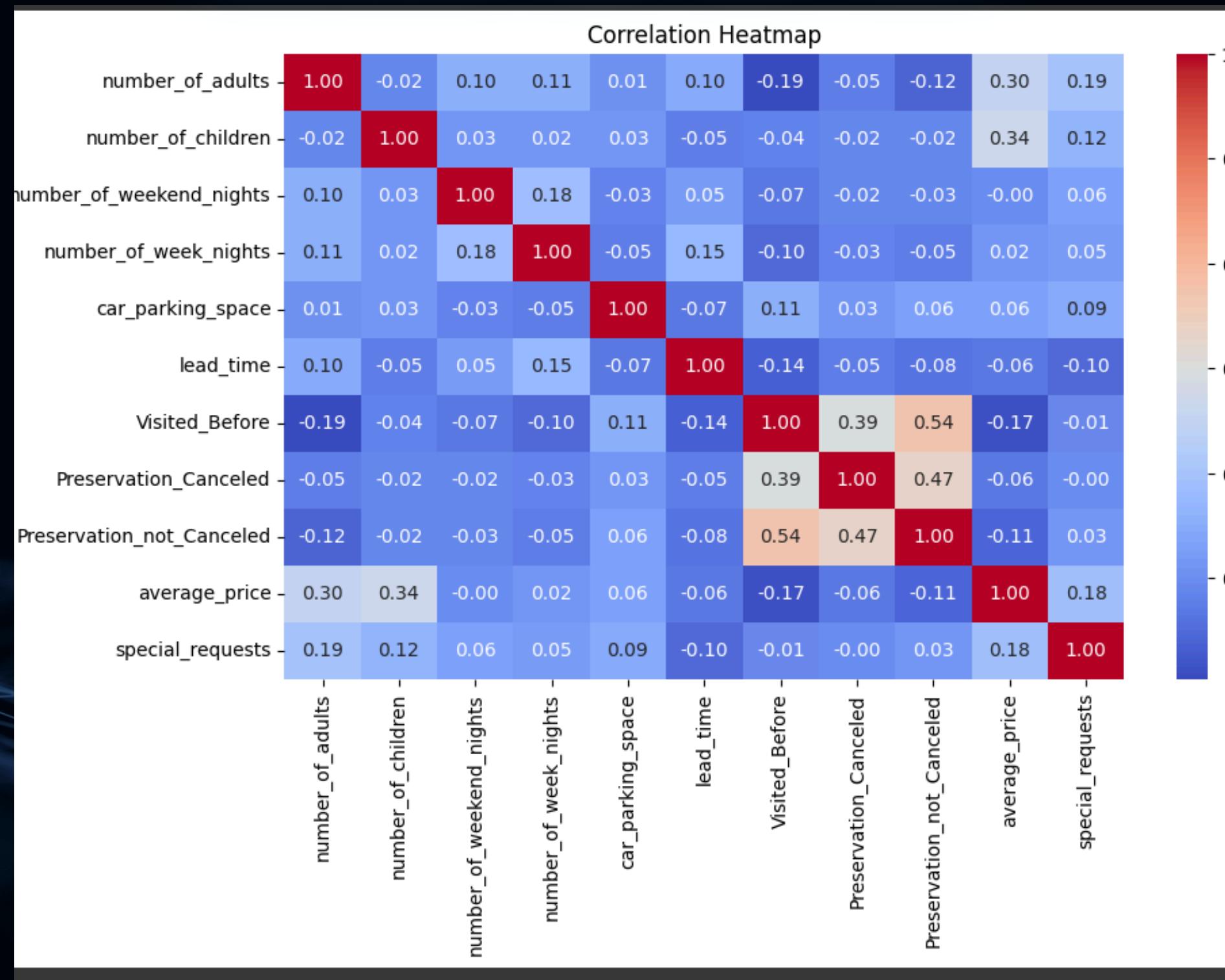


BIVARIATE ANALYSIS

- Booking status by Type of Meal
- Booking status by Room Type
- Booking status by Market Segment
- Price Distribution by Booking Status

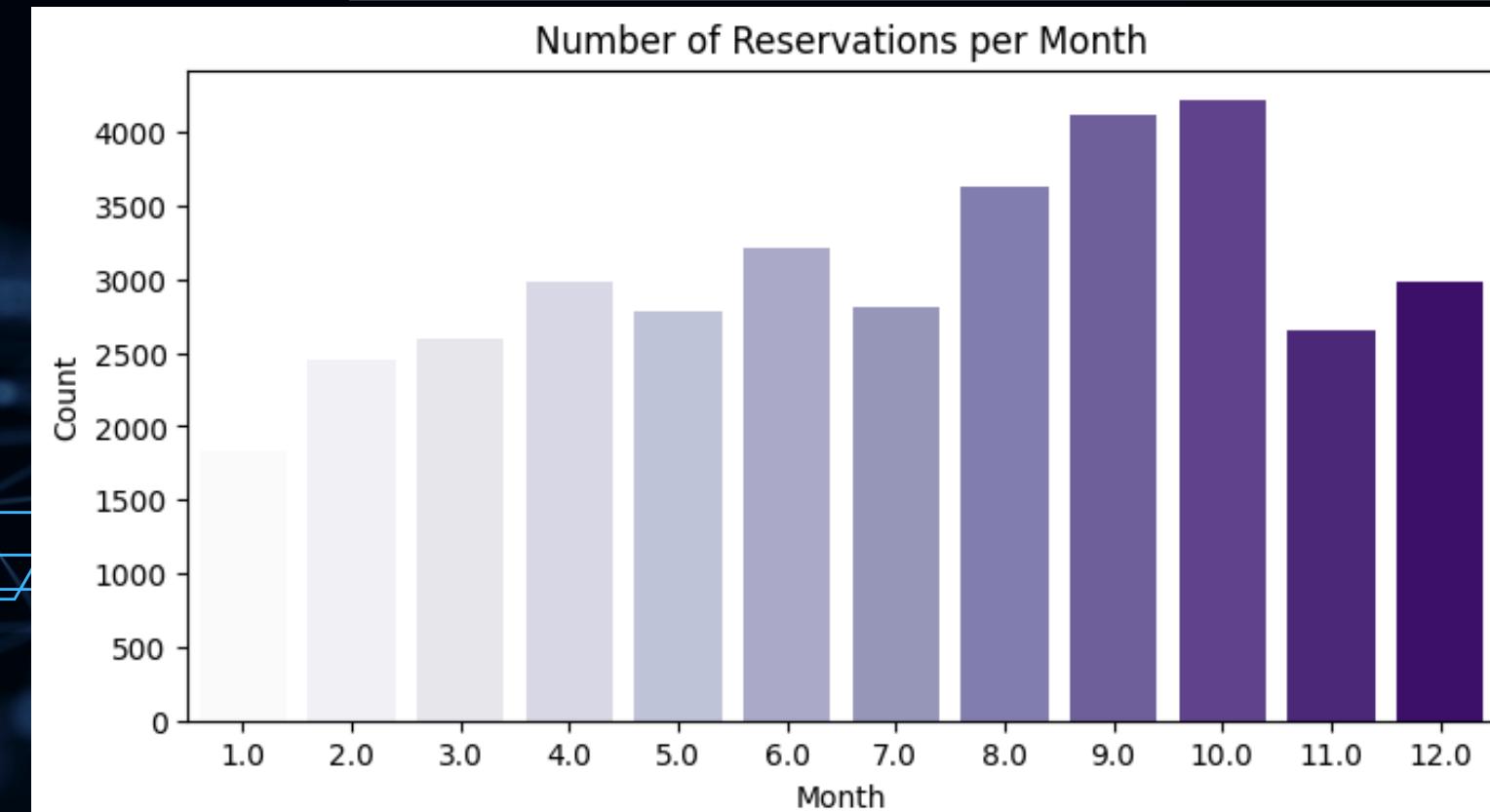
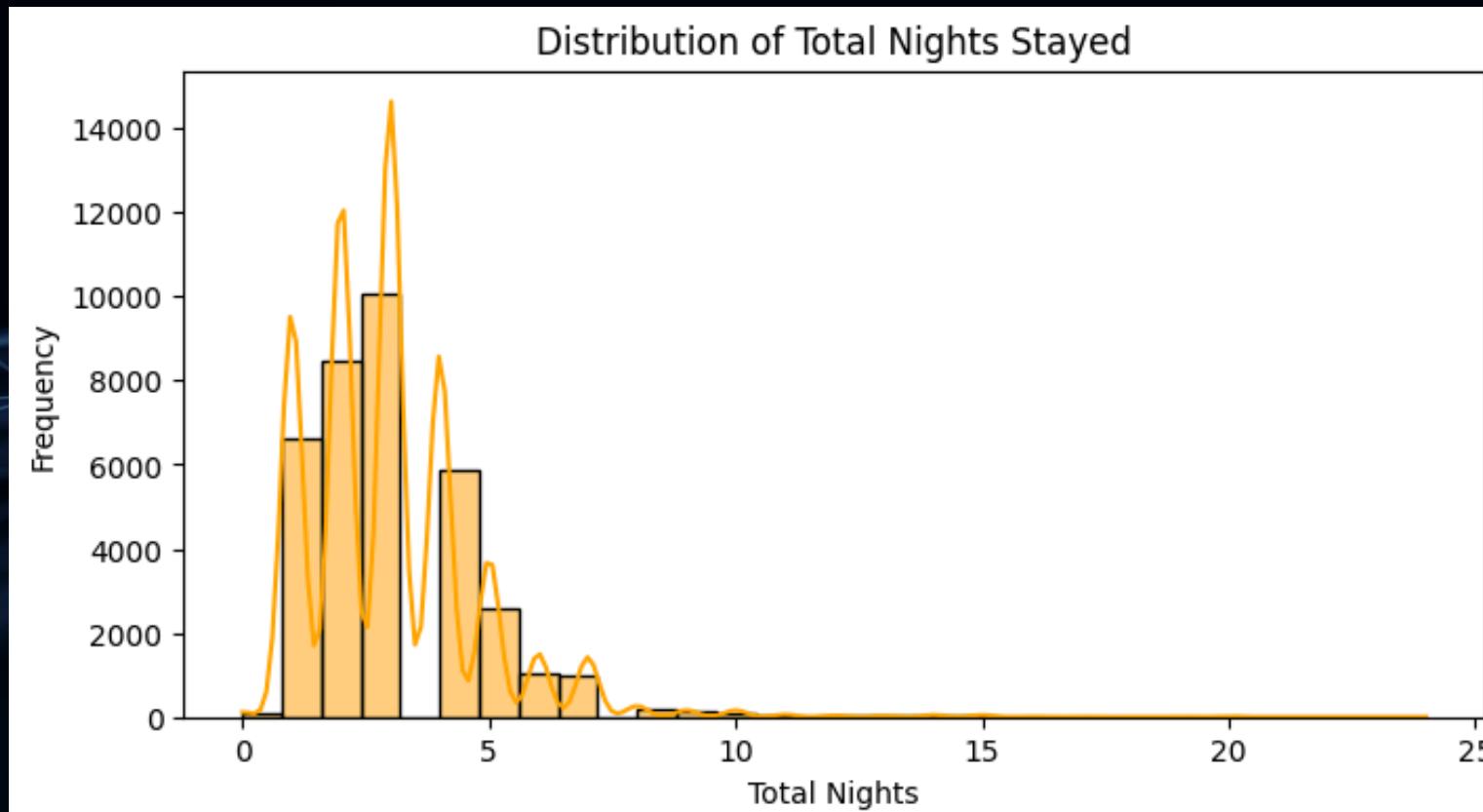
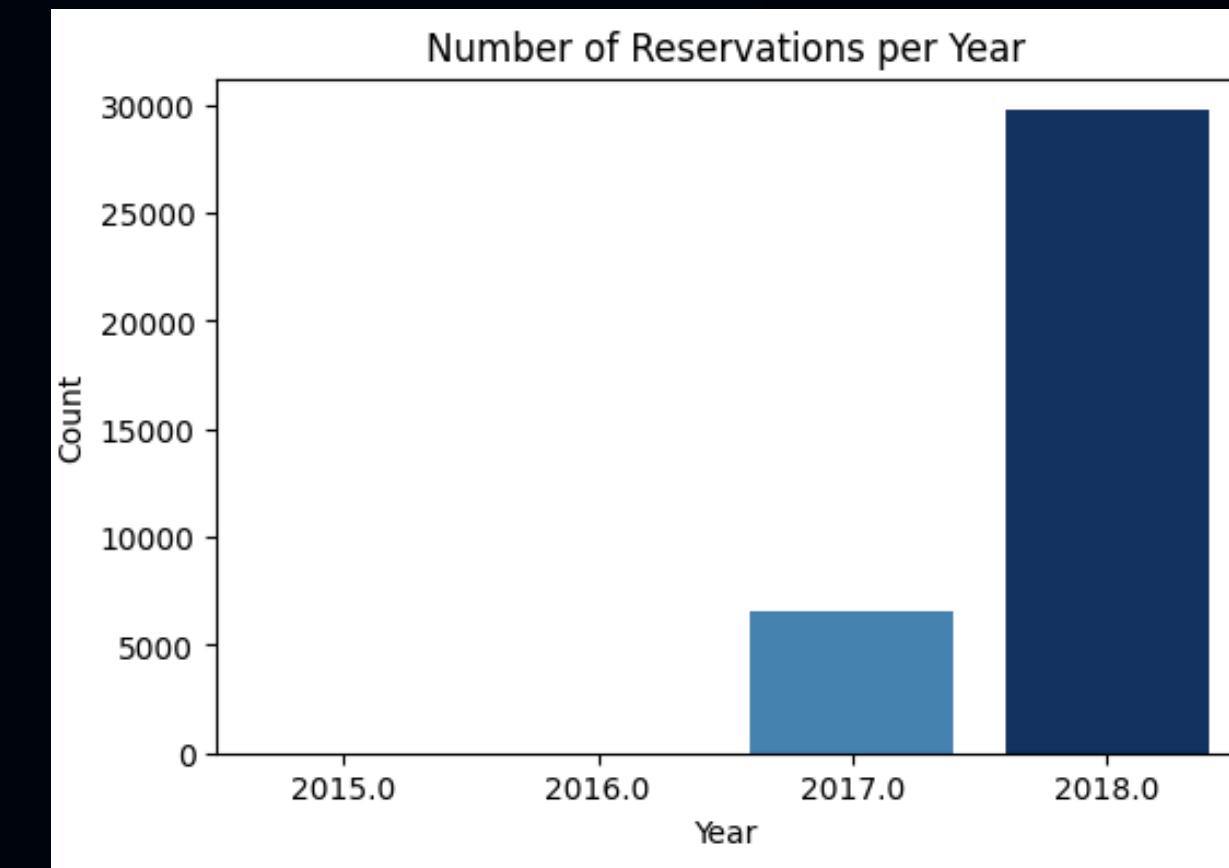


CORRELATION HEATMAP



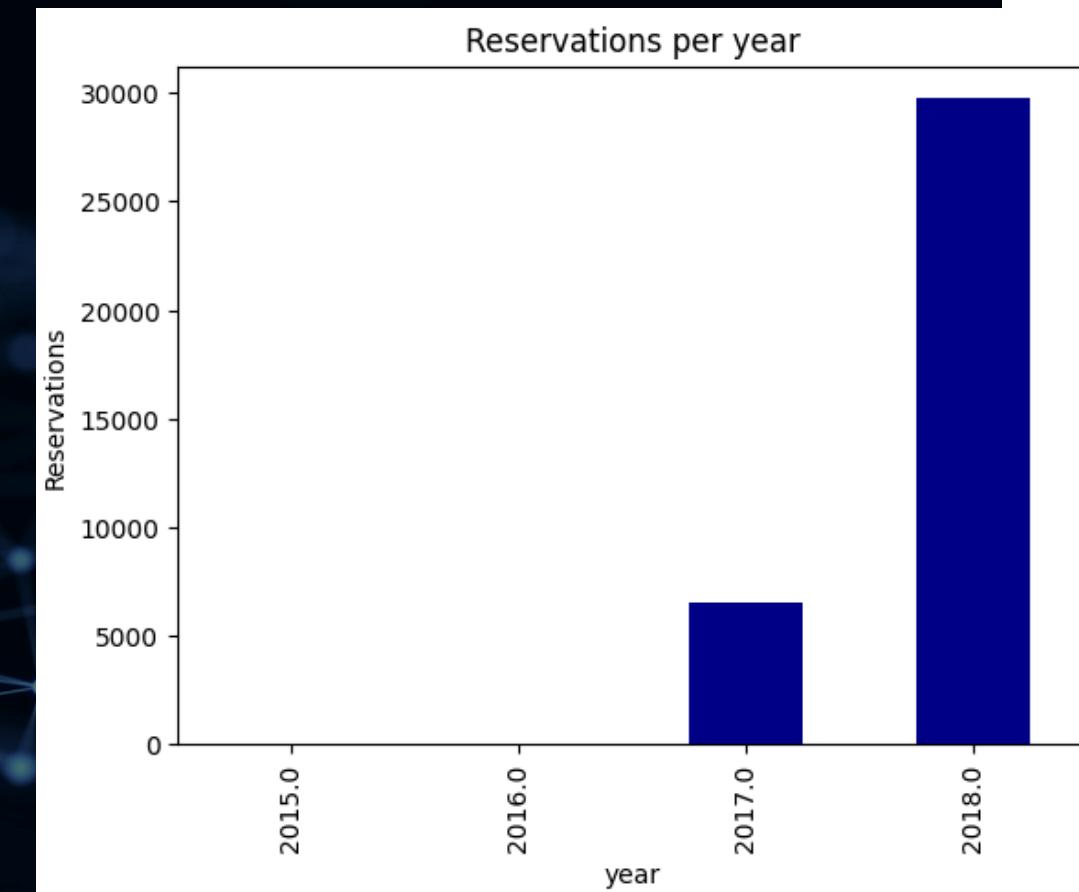
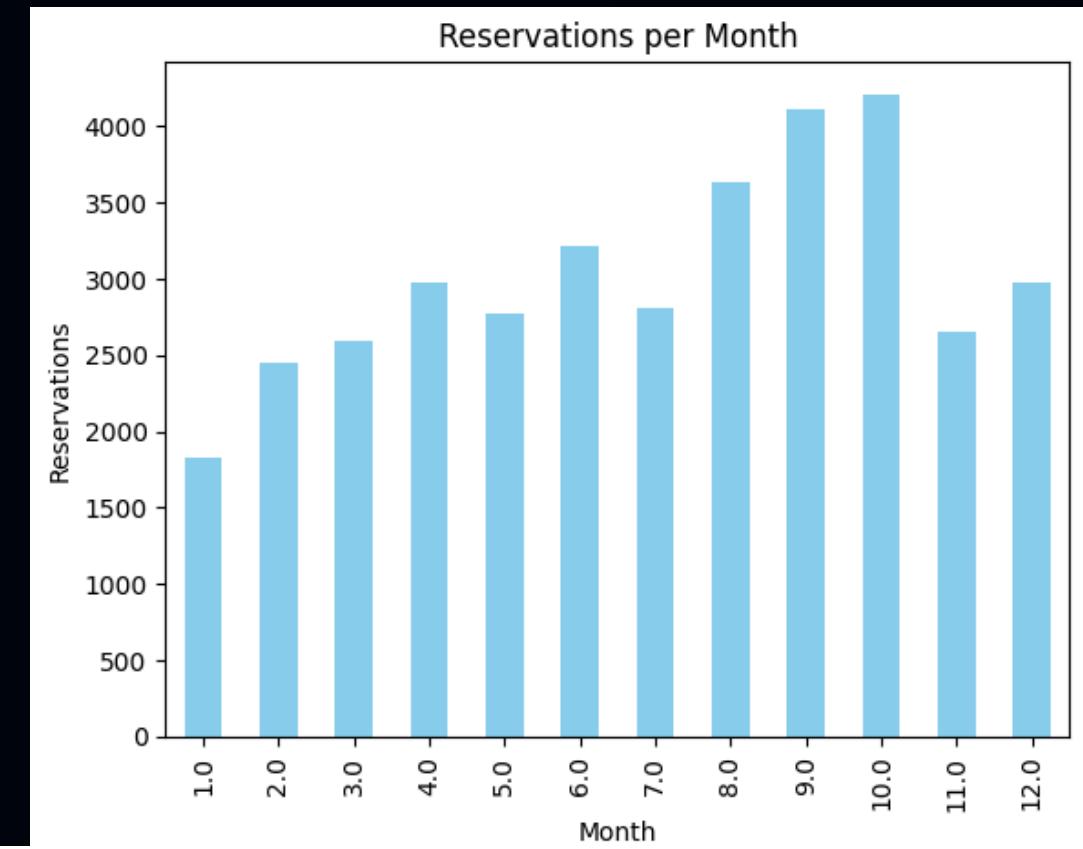
FEATURE ENGINEERING

- Converted reservation date into datetime format Insights
- `reservation_year`
- `reservation_month`
- `total_nights = number_of_weekend_nights + number_of_week_nights`



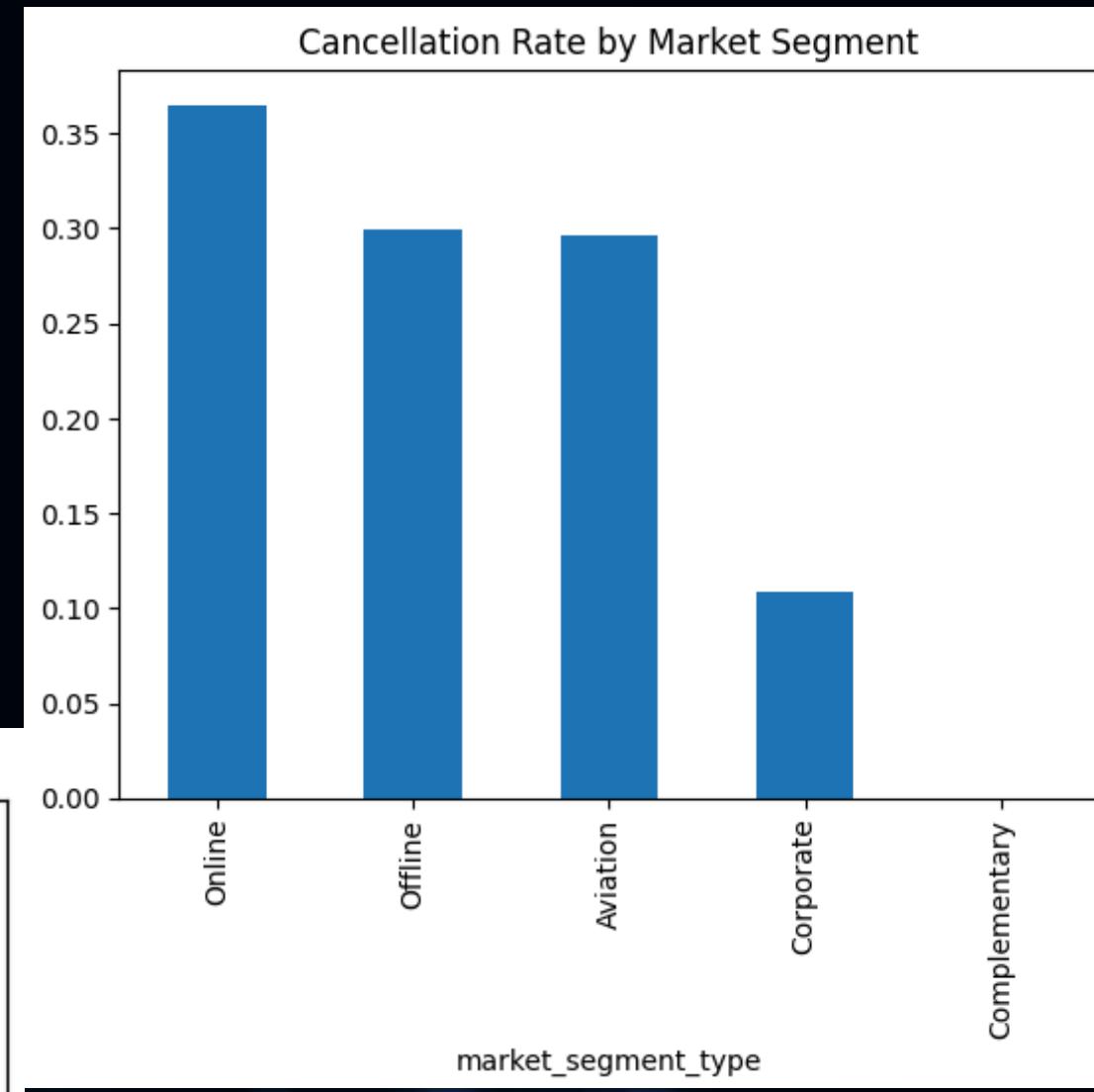
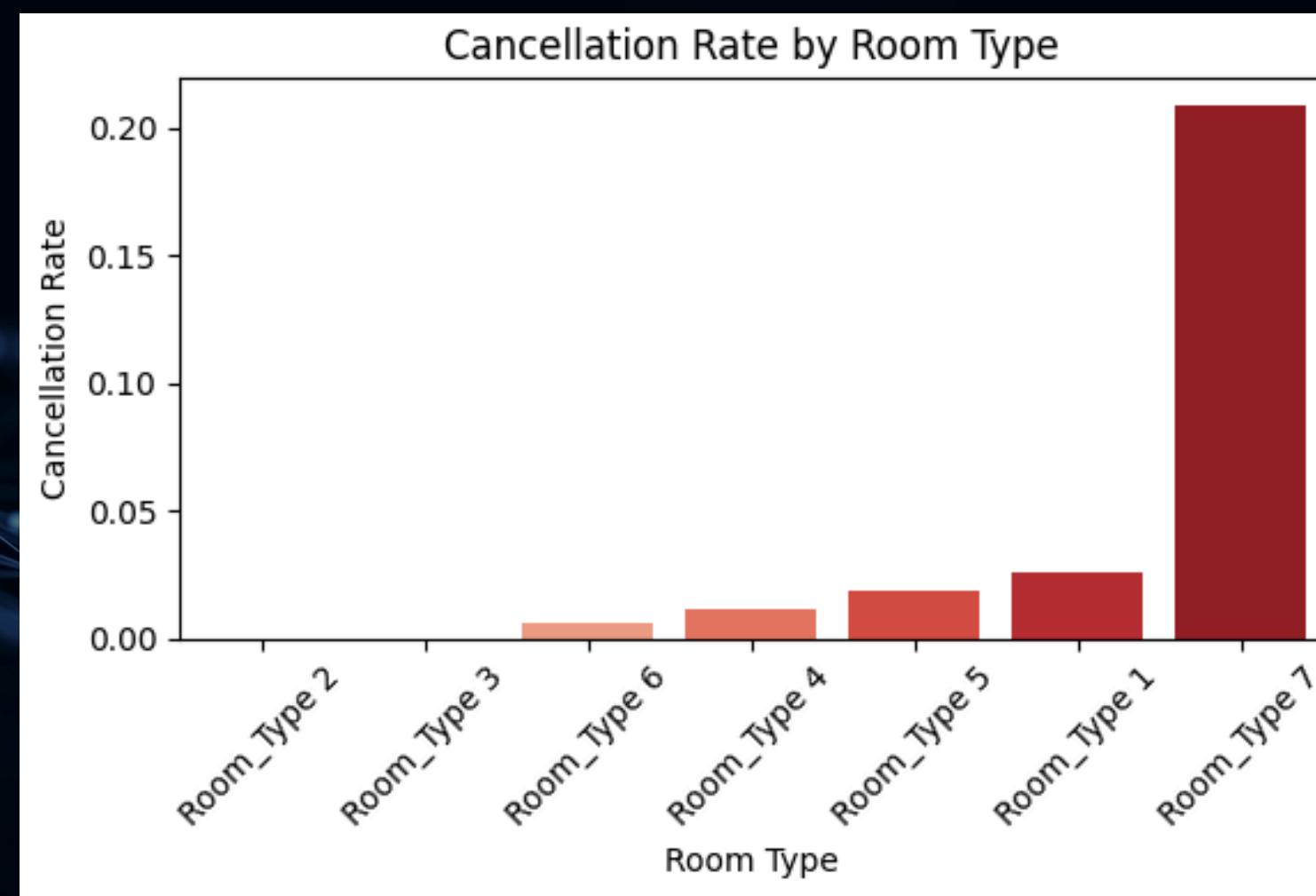
GENERAL INSIGHTS

- Most reservations are made during the summer months (september - october), peaking in October.
- 2018 had the highest number of reservations with 29,730 bookings, accounting for ~82% of the dataset.



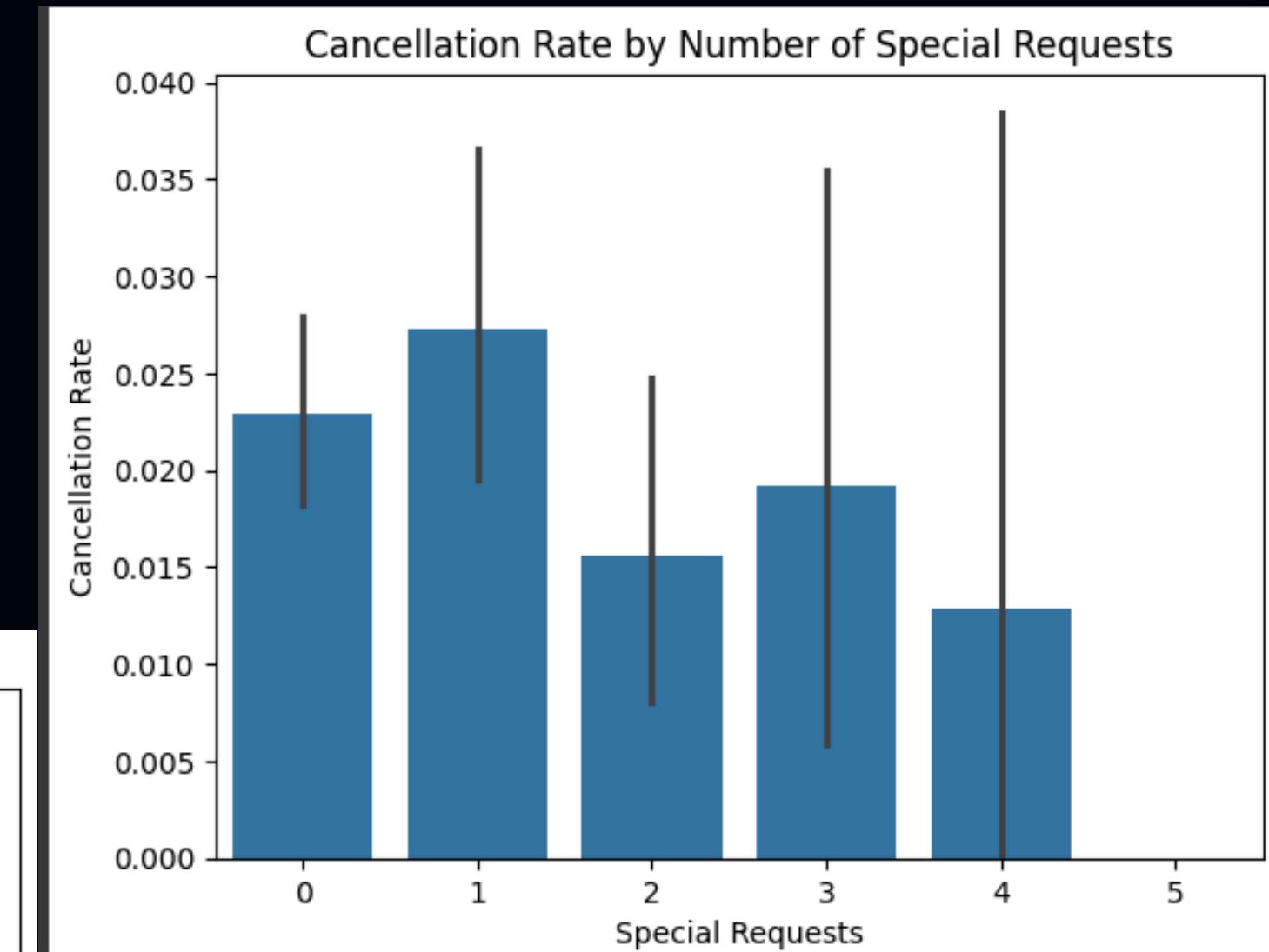
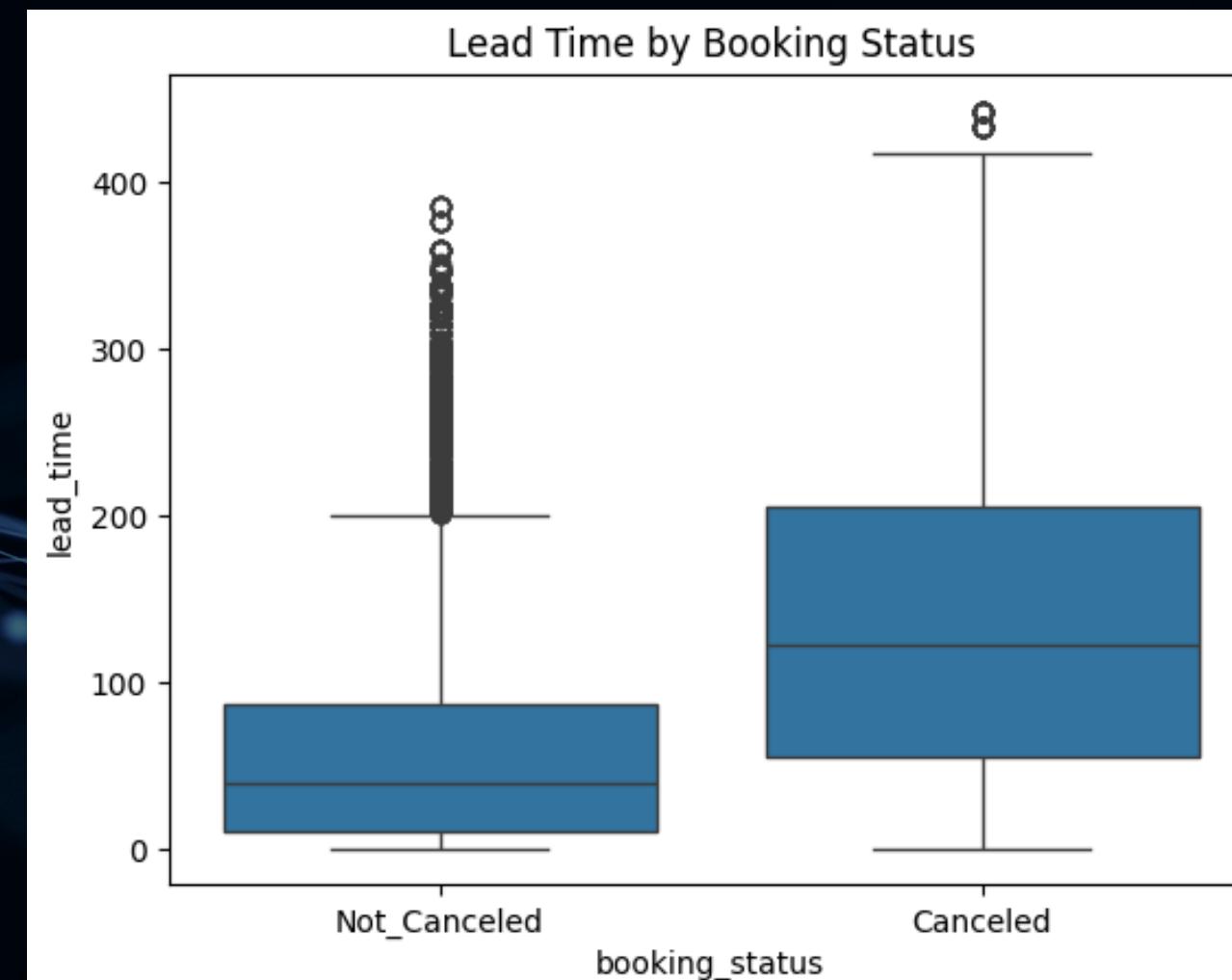
CANCELLATION INSIGHTS

- Bookings from the "Online" market segment have the highest cancellation rate (~33%)
- Room_Type 7 has the highest cancellation rate at ~21%



CANCELLATION INSIGHTS

- Guests who make special requests are significantly less likely to cancel.
- Higher lead times are correlated with a higher cancellation probability.



THANK YOU

