Hotel booking demand dataset

Content and needs :

Any hotel manager, will be very interested in knowing if a customer will cancel his hotel reservation, because that deeply correlates with the revenues our hotel can earn. In this project we will build a **model based on our dataset to predict the probability of the customer to cancel the booking or not.**

This data set contains booking information for a city hotel and a resort hotel, and includes information such as when the booking was made, length of stay, the number of adults, children, and/or babies, and the number of available parking spaces, etc

This dataset includes 119390 entries with 32 features , can be found at Kaggle in this link: https://www.kaggle.com/jessemostipak/hotel-booking-demand

The dataset is available as the .csv file. a sample of data is shown in the following table:

	hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number	arrival_date_day_of_month	stays_in_weekend_night
0	Resort Hotel	0	342	2015	July	27	1	
1	Resort Hotel	0	737	2015	July	27	1	
2	Resort Hotel	0	7	2015	July	27	1	
3	Resort Hotel	0	13	2015	July	27	1	
4	Resort Hotel	0	14	2015	July	27	1	
...	...	...	...	...	...	...	...	
119385	City Hotel	0	23	2017	August	35	30	
119386	City Hotel	0	102	2017	August	35	31	
119387	City Hotel	0	34	2017	August	35	31	
119388	City Hotel	0	109	2017	August	35	31	

Features :
hotel
is_canceled
lead_time
arrival_date_year
arrival_date_month
arrival_date_week_number
arrival_date_day_of_month
stays_in_weekend_nights
stays_in_week_nights
adults
children
babies

meal
country
market_segment
distribution_channel
is_repeated_guest
previous_cancellations
previous_bookings_not_canceled
reserved_room_type
assigned_room_type
booking_changes
deposit_type
agent
Company
days_in_waiting_list
customer_type
adr
required_car_parking_spaces
total_of_special_requests
reservation_status
reservation_status_date

Tools:

- ***numpy, pandas*** for data manipulation.
- ***matplotlib and seaborn,*** for plotting.
- ***sklearn*** for modling .
- The work will be done through ***Jupyter notebook*** .