


Hotel Booking Cancelation



Michael Albers, Aubrey Kranz, Raegan Storin



Index

01 | Proposal Overview

02 | Research Question and Hypothesis

03 | Data and Methods Plan

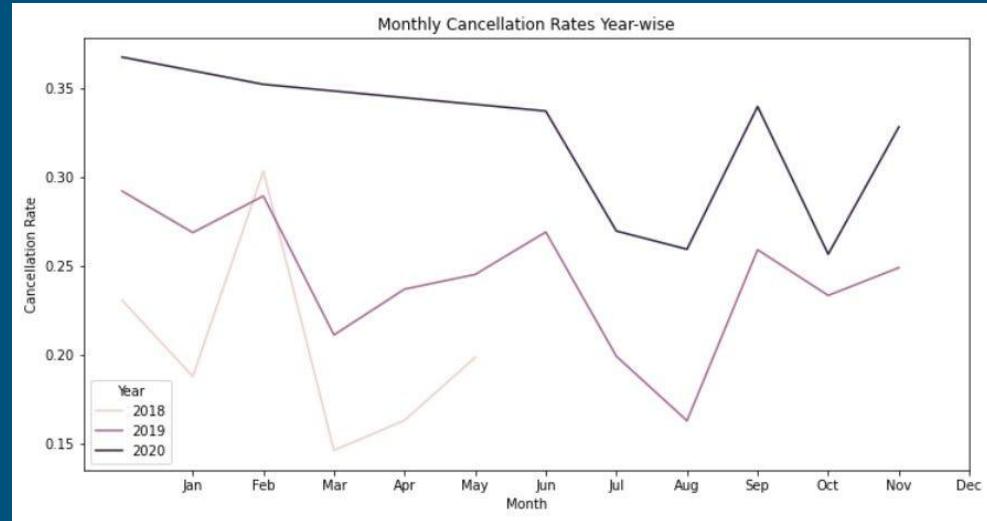
04 | Schedule

Proposal Overview

What is the problem?

Hotel overbooking can decrease customer satisfaction and negatively impact reputation while underbooking reduces potential revenue.

The goal would be to minimise hotel underbooking without overbooking and exceeding capacity.



Research Question and Hypothesis/ Assessment

How can we improve the method of booking or cancellation policies to accurately accomodate for booking cancellations?

Or

Our hypothesis is that we will be able to accurately predict a hotel's optimal booking capacity by taking into account known factors of historical data.

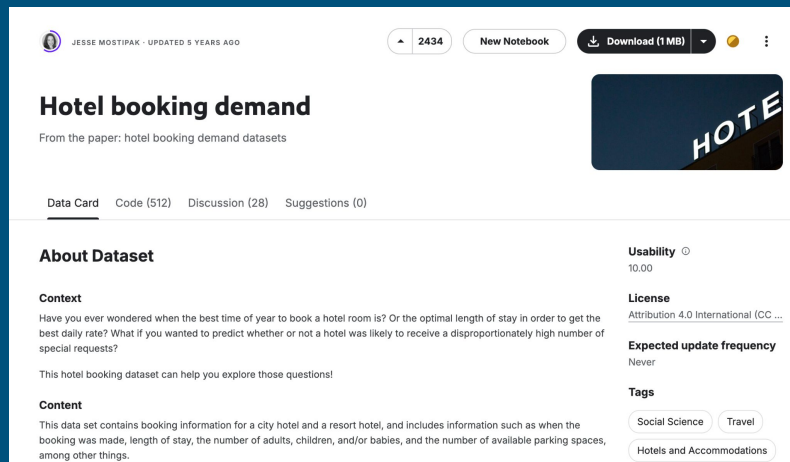
Data and Methods Plan

Our aim is to build a model that can predict hotel booking demand based on a Kaggle hotel booking demand dataset.

By analyzing patterns found in this data, we project seek to provide insights that can help hotels optimize their available bookings.

Variables included

is_canceled	meal	agent
lead_time	Country	company
arrival_date_year	market_segment	days_in_waiting_list
arrival_date_month	distribution_channel	customer_type
arrival_date_week_number	is_repeated_guest	adr
arrival_date_day_of_month	previous_cancellations	required_car_parking_spaces
stays_in_weekend_nights	previous_bookings_not_canceled	total_of_special_requests
stays_in_week_nights	Reserved_room_type	reservation_status
adults	assigned_room_type	reservation_status_date
children	booking_changes	
babies	deposit_type	



The screenshot shows the Kaggle dataset page for 'Hotel booking demand'. At the top, it indicates the dataset was updated 5 years ago and has 2434 views. The title 'Hotel booking demand' is prominently displayed, followed by the source 'From the paper: hotel booking demand datasets'. Below the title, there are tabs for 'Data Card' (selected), 'Code (512)', 'Discussion (28)', and 'Suggestions (0)'. The 'About Dataset' section provides context, asking if the user has wondered about the best time to book a hotel room or the optimal length of stay. It also mentions that the dataset can help explore questions about booking patterns. The 'Content' section states that the data set contains booking information for a city hotel and a resort hotel, including details like booking date, length of stay, number of adults, children, and babies, and the number of available parking spaces. On the right side, there are sections for 'Usability' (10.00), 'License' (Attribution 4.0 International), 'Expected update frequency' (Never), and 'Tags' (Social Science, Travel, Hotels and Accommodations).

Schedule

All work is to be expected to be completed the night before the lab prior to the due date, according to our internal deadlines.

- Part 1: 9/17 (Presentation 9/23)
- Part 2: 10/11 (Presentation 10/18)
- Part 3: 11/6 (Draft Paper)
- Part 4: 12/3 (Presentation 12/10)

Part 1

Today < > September 2024						
SUN Sep 1	MON 2	TUE 3	WED 4	THU 5	FRI 6	SAT 7
8	9	10	11	12	13	14
15	16	17 Finish Presentation	18	19	20 Discuss in class	21
22	23 Presentation	24	25	26	27	28

Sources

<https://medium.datadriveninvestor.com/a-python-data-analysis-project-to-understand-hotel-cancellations-fb3f0fee6eea>

<https://www.kaggle.com/datasets/jessemostipak/hotel-booking-demand/code>