Hotel Reservations Analysis

MKT 6337.003 - Predictive Analytics for Data Science - S25

**Group 6**

Choksi, Ria A

Jana, Sunayana

Parida, Priyadarshan

Qazi, Khurem

Rajappan, Soniya

##### Research Question

1. *What novel question(s) do you aim to answer in your project?*

* What factors contribute most significantly to hotel reservation cancellations?
* Can customer and booking characteristics predict the likelihood of a cancellation?
* What insights can we draw to make better pricing and promotion decisions?

1. *Who will be your potential audience?*

* Hotel operations managers, marketing managers, finance officers, and executives in the hospitality industry.

1. *Why would they be interested in your question(s)?*

* Marketing managers can use insights to implement targeted strategies to reduce cancellations, such as offering incentives or improving customer satisfaction.
* Operations managers can optimize hotel occupancy with the available prediction.
* Finance Team: Improved forecasting of occupancy rates and optimized pricing strategies.

##### Data

1. *What is the data source?* [ScienceDirect - Hotel Bookings](https://www.sciencedirect.com/science/article/pii/S2352340918315191) | [Data](https://ars.els-cdn.com/content/image/1-s2.0-S2352340918315191-mmc2.zip)
2. *What is the data period?* Data period: July 2015 - August 2017
3. *What is the level of observation?* Level of observation: Individual hotel booking
4. *Do you have repeated observations for a given party?* Yes, there are repeated observations for customers who are repeated guests or have multiple bookings during the data period.

##### Model

1. *What is the outcome of interest (Y variable)?* Outcome: is\_canceled (binary variable indicating whether a booking was canceled).
2. *What covariates or predictors (X variables) do you plan to include in your model?*

* *Customer demographics:* adults, children, babies | *Booking details:* lead\_time, stays\_in\_weekend\_nights, stays\_in\_week\_nights, reserved\_room\_type, assigned\_room\_type, booking\_changes, deposit\_type, days\_in\_waiting\_list, total\_of\_special\_requests | *Market and distribution:* market\_segment, distribution\_channel, country | *Historical behavior:* is\_repeated\_guest, previous\_cancellations, previous\_bookings\_not\_canceled | *Pricing:* adr (Average Daily Rate)

1. *What statistical model(s) do you plan on using?* Logistic Regression, Decision Tree and Random Forest, Gradient Boosting (e.g., XGBoost)

##### Results/Conclusions

1. *What are the final results and conclusions?*

* Identify key factors driving cancellations and provide a predictive model with high accuracy.
* Insights will help hotels implement strategies to reduce cancellations, such as adjusting lead times, flexible deposit policies, or targeting specific customer segments with promotions.

1. *What conclusions can you expect, or reach based on your analyses?*

* Customers with longer lead times, higher ADR, or no deposit are more likely to cancel.
* Repeated guests and those with special requests are less likely to cancel.
* Predictive models will enable hotels to manage cancellations and improve operational efficiency proactively.