Michael Albers, Aubrey Kranz, Raegan Storin

Layer 1 (9/23) - Understanding the Problem

Accurately forecasting booking demand is a key task in the hospitality industry. Hotels face a constant challenge in balancing room availability, pricing strategies, and operational costs while maximizing revenue. Overbooking can damage a hotel's reputation, undermining customer satisfaction if their booking is canceled, while underbooking leads to a loss of potential revenue. The ultimate goal of any hotel is to maximize bookings without exceeding their capacity for guests.

The main problem addressed by this project is the unpredictability of booking capacity. If every person who booked a hotel room was guaranteed to follow through, it would be simple for hotels to predict how many people they can accommodate. However, a great deal of uncertainty stems from the fact that guests who book reservations might decide to cancel their reservation at any point before their arrival. Hotels must carefully balance demand and supply, often without the ability to forecast accurately, leading to overbooked or underbooked situations. Solving this problem is essential because hotels must adjust prices and room availability according to demand, and being able to predict these fluctuations allows for more profitable pricing models. This project aims to predict hotel booking capacity using data from 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.

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. This hypothesis is based on the assumption that there are patterns and relationships between the variables in the dataset that impact booking cancellations, such as lead time, previous cancellations, average daily rate, and seasonality. Our aim is to build a model that can predict hotel booking demand based on this historical data. To determine the success of the project, the model's ability to accurately predict effective booking capacity and the relative likelihood of cancellations will be tested against

historical data. We will additionally test the model in more specific scenarios, such as the difference in trends between resort versus city hotels, as well as peak or off-peak seasons for travel.

Regarding our schedule, all work is to be completed the night before the lab prior to the due date to take time during labs to look over our work, and see if there are any necessary adjustments prior to the due date. Aubrey will be leading the code work regarding the cleaning and modeling of the data. Michael will be leading more of the evaluation work regarding understanding the data and models before applying it to our idea on an optimal booking capacity. Finally, Reagan will be leading the work on documentation and presentation of our work.