**PROBLEM DESCRIPTION**

The objective of our project is to predict hotel booking cancellations. Booking cancellations have a significant impact on the demand-management decisions in the hospitality industry. To reduce this negative impact on the hotel's revenue, we intend to predict booking cancellations.

1)How many bookings were cancelled?Does this depend on the time of the year?

2)Which is the busiest month for hotels? Could we predict when the hotel was likely to receive a disproportionately high number of special requests?

3)When is the best time of the year to book a hotel room/What is the optimal length of the stay in order to get the best daily rate?

4)Is the number of booking changes/amendments made to a booking deterministic of a cancellation?

5)How does price vary per night over the year?

6)Does the booking cancellation depend on the travel agency that made the booking?

7)Could we cluster/segment bookings based on their features?

8)Do repeated guests show a lower chance of booking cancellation?

The dataset chosen describes hotel demand/bookings data.It comprises two hotels (a resort hotel, a city hotel).It has 32 columns wherein each observation indicates a hotel booking. It includes features such as arrival date,lead time(number of days between booking and check in), number of adults,children,babies,type of meal booked,country of origin,previous cancellations,previous bookings not cancelled by the user,type of room reserved,booking changes made after the booking, deposit type, travel agent that made the booking,ADR(Average Daily Rate), number of special requests,number of days the booking was in the waiting list, customer type and so on.