**HOTEL BOOKING ANALYSIS**

Mrityunjay Singh Chandel, Shubham Jaiswal

(DATA SCIENCE TRAINEES)

**ALMABETTER, BANGALORE**

**Abstract :**

Have you ever wondered when the best time of year to book a hotel room is? Or the optimal length of stay in order to get the best daily rate? What if you wanted to predict whether or not a hotel was likely to receive a disproportionately high number of special requests? This hotel booking dataset can help you explore those questions!

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, among other things. All personally identifying information has been removed from the data.

Explore and analyze the data to discover important factors that govern the bookings.

**Following are the few questions that we aim to answer through our analysis:**

1. Which type of hotel is mostly preferred by guests?
2. Which agent made the most bookings?
3. What is the percentage of cancellation?
4. What is the percentage of repeated guests?
5. What are the arrival months in city hotel?
6. What is the number of weekend nights that the customer has stayed or booked in city hotel?
7. Which type of meal is preferred in City hotels?
8. In which year more people arrived in City hotels?
9. In which year more number of people arrived in Resort Hotels?
10. What are the arrival months in Resort Hotel?
11. Which type of breakfast is preferred in Resort Hotel?
12. What is the number of weekend nights that the customer has stayed or booked the resort hotel?
13. Which channel is most often used to book hotels early?
14. Which distribution channel generates more revenue for hotels?
15. Which channel has a longer average turnaround time?

## A quick glance at data shows that this dataset has around 119390 observations in it with 32 columns and it is a mix between categorical and numeric values.

## We notice that there are 32 columns in the data frame and that some columns like 'children', 'company', 'country' and 'agent' have null values.

## We followed various steps which are mentioned below in our analysis:

## Taking a look at our data. :

## Cleaning the data.

## Removing duplicate rows if any.

## Handling missing values.

## Converting columns into suitable data types.

## Adding Important Columns.

**Exploratory Data Analysis -** In this section, we will detail our analysis to the questions of interest mentioned above and gain preliminary insights through exploratory data analysis and visualization.

We follow the steps given below for the data analysis:

* Plotting the heat map.
* Analysis of which hotel is preferred by the guests.
* Analysis of which agent made the most bookings.
* Analysis of percentage of cancellation.
* Analysis of percentage of repeated guests.
* Analysis of City Hotels.
* Analysis of Resort Hotels.
* Analyzing the Distribution Channels.