**BI Assignment 3**

**Usaid Bhirya**

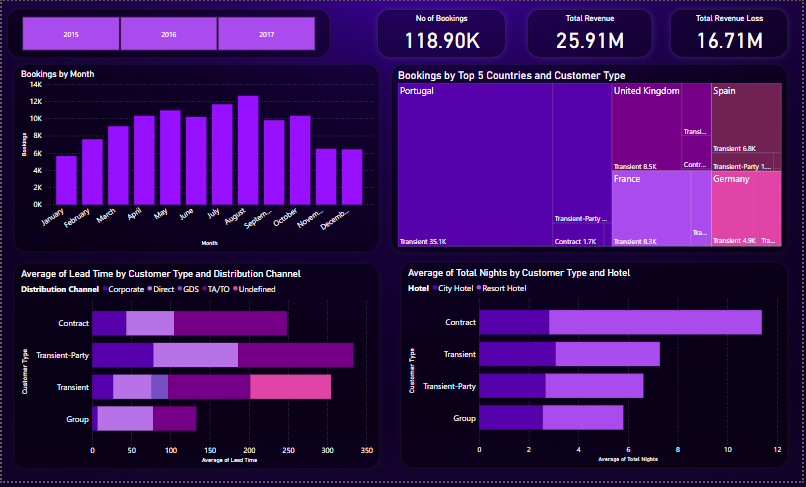
**Hasan Jawaid**

Our dataset was Hotel Booking <https://www.kaggle.com/datasets/mojtaba142/hotel-booking>.

The dataset contains booking information of two hotels - a resort hotel and a city hotel. It has 32 columns with details about the booking, such as whether the booking was canceled or not, number of days between the booking date and arrival date, month, week and year of arrival, number of nights stayed, number of adults, children and babies, type of meal booked, country of origin, market segment, distribution channel, whether the guest was a repeated guest or not, number of previous cancellations and previous bookings not canceled, type of room reserved and assigned, number of booking changes, deposit type, travel agency ID, waiting list duration, customer type, average daily rate, number of parking spaces required, number of special requests made, reservation status (whether the customer checked-in and checked-out or not) and the date on which the reservation status was set. Additionally, there are some columns for personal identification details, such as name, email, phone number and credit card number, which are not real.

For the data cleaning we used Jupyter Notebooks, we removed all the columns which had more than 50% missing values as well as columns that were not useful for our analysis. For the rows, we also removed the one’s with null values because those rows had a very small ratio. We also combined a few columns into one like the date one and created some new columns using the already given data to help with our analysis. We have attached the notebook in our submission for you to take a look at it.

The following are the screenshots of our Dashboard



Graphical user interface, application

Description automatically generated

A screenshot of a computer

Description automatically generated with low confidence