

Capstone Project Summary

Team Member's Name, Email and Contribution:

1. Tito Varghese
Email - tito.varghese1992@gmail.com
 - Removing the Outliers from the Dataframe
 - Factors Governing the Hotel Bookings
 - To predict whether or not a hotel was likely to receive a disproportionately high number of special requests?
 - Factors Governing the Cancellation
 - Find the best customer type with maximum Average Daily Rate in Different Year
2. Anmol Raj
Email – actuallyanmol@gmail.com
 - Removing the Duplicate rows from the Dataframe
 - Hotel with highest Revenue
 - Countries from which most guests are coming
 - Which Hotel has highest waiting time
 - Which distribution channel brings highest revenue generating deals
3. Lakshmi Keerthana
Email - keerthana826@gmail.com
 - Handling the missing values in the Dataframe
 - The most used Distribution Channel
 - The most preferred meal type by the customer
 - Market segment used by the guest in different hotels
 - Heat Correlation Map
4. Kamatam Harshith
Email - 1harshasai@gmail.com
 - Converting Columns to appropriate data type and adding important columns
 - Which hotel has a higher chance that its customer will return for another stay
 - Which hotel has higher booking cancellation rate
 - When is the best time in a year to book a hotel room
 - Type of customer booking the most

Please paste the GitHub Repo link.

GitHub Profile Link: - <https://github.com/Keerthana826>

GitHub Repository Link: - <https://github.com/Keerthana826/EDA-Capstone.git>

Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

Hotel Booking Analysis

In this Capstone Project on Exploratory Data Analysis we analyze, explore and get the insights into given hotel bookings csv dataset from the timeline 2015-2017. The data set that we have contains booking information for a city hotel and a resort hotel. We used python coding, pandas libraries and data visualization libraries.

Our approach to solving these problems were Loading the data from the drive into dataframe, cleaning the dataframe, Data Analysis using pandas library and Data Wrangling and Visualization

We have divided the project into four main parts. Hotel Wise Analysis, Distribution Channel Wise Analysis, Cancellation related Analysis and Time and stay related Analysis.

In Hotel wise analysis, we compared both the hotels focusing on revenue, number of bookings, waiting time, bookings cancellation and concluded which hotel has been better.

In Distribution channel wise analysis, we have seen what distribution channel and market segment are being used by the customers to book the hotel.

In Cancellation wise analysis, we see why the cancellation of bookings happen and concluded it has been mostly because of no car parking space.

In time and stay related analysis, we analyzed what type of customers are mostly visiting and at what time of the year do they visit and which countries most of the customers are from.

On the basis of EDA, we have identified why the customers mostly preferred City hotel over the resort hotel. The deals that are being provided by both the hotels and why booking cancellations happen. With the help of analysis we have concluded who the target customers are. Hotels can plan on improving the required parameters for more number of hotel bookings and to increase their demand.