Data Visualisation with ggplot2

Project Deliverable

- Your deliverable will be an **HTML file** generated from an R notebook with your solution.

Instructions

Background Information

You are a Data Scientist working for Hilton International Hotels and need to help the management decide on the strategies that will lead to revenue growth.

You have a dataset containing information on the various chain of hotels, including when customers made the booking, length of stay, the number of adults, children, or babies, and the number of available parking spaces, among other things.

Given this dataset, perform data exploration, data cleaning, and analysis to come with appropriate recommendations. As a start, while performing data analysis, you ask the following questions from the given dataset. Make sure that you have a visualization for each question.

- How many bookings were canceled?
- What was the booking ratio between resort hotels and city hotels?
- What was the percentage of booking for each year?
- Which was the busiest month for hotels?
- Most guests came from which country?
- How long do most people stay in hotels?
- Which was the most booked accommodation type (Single, Couple, Family)?

You will be required to use the following notebook template:

Dataset Overview

- This data set contains booking information for a city hotel and a resort hotel. It includes information such as when customers made the booking, length of stay, the number of adults, children, or babies, and the number of available parking spaces, among other things.
- We have removed personally identifying information from the dataset.
- Dataset Download Link: https://bit.ly/2WvQbhJ
- NB: You need to download this file.

Acknowledgments

The data is originally from the article; *Hotel Booking Demand Datasets* by Nuno Antonio, Ana Almeida, and Luis Nunes for Data in Brief, Volume 22, February 2019.