To: Startrek Hotels

From: Anamika Singh, Anurag Lahon, Rubina Shaik

Subject: Factors affecting booking cancellations, special requests, and number of bookings.

Date: 05/10/2020

This summary reports on the factors that influence booking cancellations, special requests, and number of bookings for the resort hotel and the city hotel.

EXECUTIVE SUMMARY

Major Findings

- The analysis of the hotel data shows that the repeated guests are only 3% of the total number of guests. (see Exhibit A).

- The bookings are least (around 5%) from November to January (see Exhibit B).

- Based on random forest model for predicting cancellations, deposit type, Average daily rate and lead time affects the booking cancellation most (see Exhibit C).

- Based on random forest model for predicting special requests, lead time, market segment and reservation status date are some of the factors which impacts special requests (see Exhibit D).

Recommendations for Action

- The repeated guests are only 3%, which points to a need for change in marketing and hospitality strategy.

- More offers on booking should be provided from the month of November- January as the bookings are least in these months.

- Booking cancellation model will help to Identify the likelihood of bookings being cancelled and makes it possible for hotel managers to take measures to avoid these potential cancellations, such as offering services, discounts, or other perks.

- The booking cancellation prediction model enables hotel managers to mitigate revenue loss derived from booking cancellations and to mitigate the risks associated with overbooking.

- Special request model will contribute to reduce uncertainty in the inventory allocation and pricing decision process by predicting the likelihood of getting a request from customers.

- We recommend using random forest model to predict booking cancellations as it gives 93% accuracy, whereas logistic regression, lasso and ridge gave approximately 80% accuracy.

Analytical Overview

- Exploratory data analysis was performed on all the variables to analyze the data and to find the correlation between them.

- Multiple logistic regression was used to build the model to predict booking cancellations and special requests.

- Other machine learning algorithms were used like lasso, ridge, and random forest to get the better predictions.

- Time series model was used to predict the number of booking using Holtwinters.

- Model was tuned to get better accuracy for predictions.

- To validate the model, data was split into training and testing dataset and we also tried k-fold cross validation.

Documentation Page

Exhibit A. Plot between repeated guests and number of bookings.

Is Repeated Guest

Exhibit B. Bookings by month

Bookings by Month

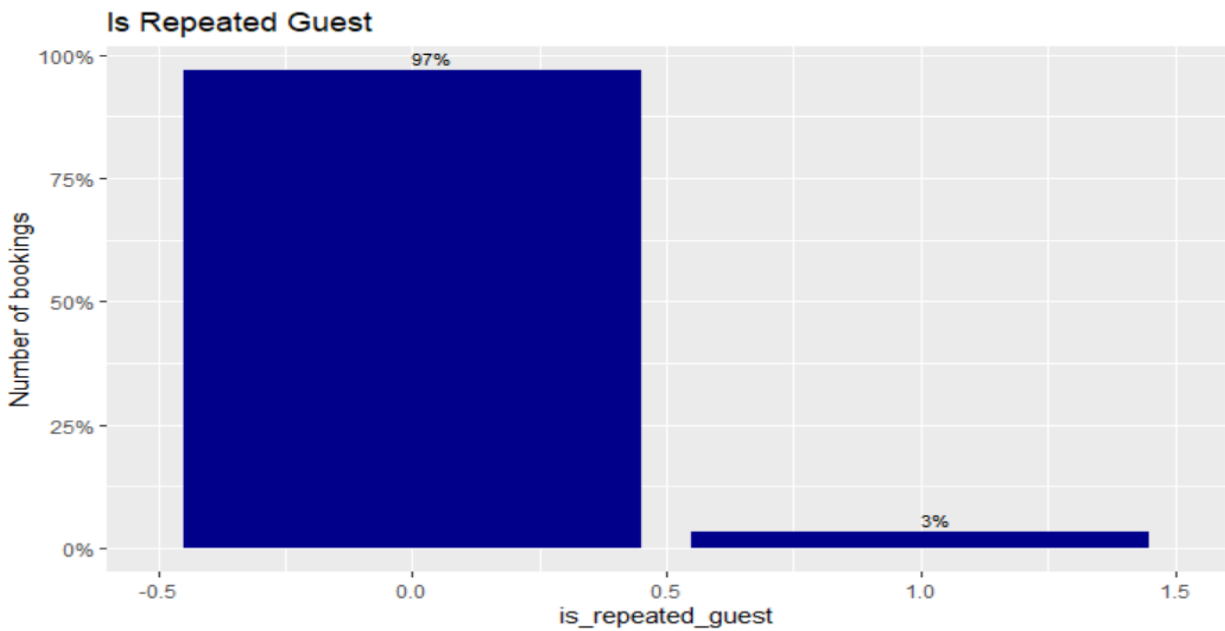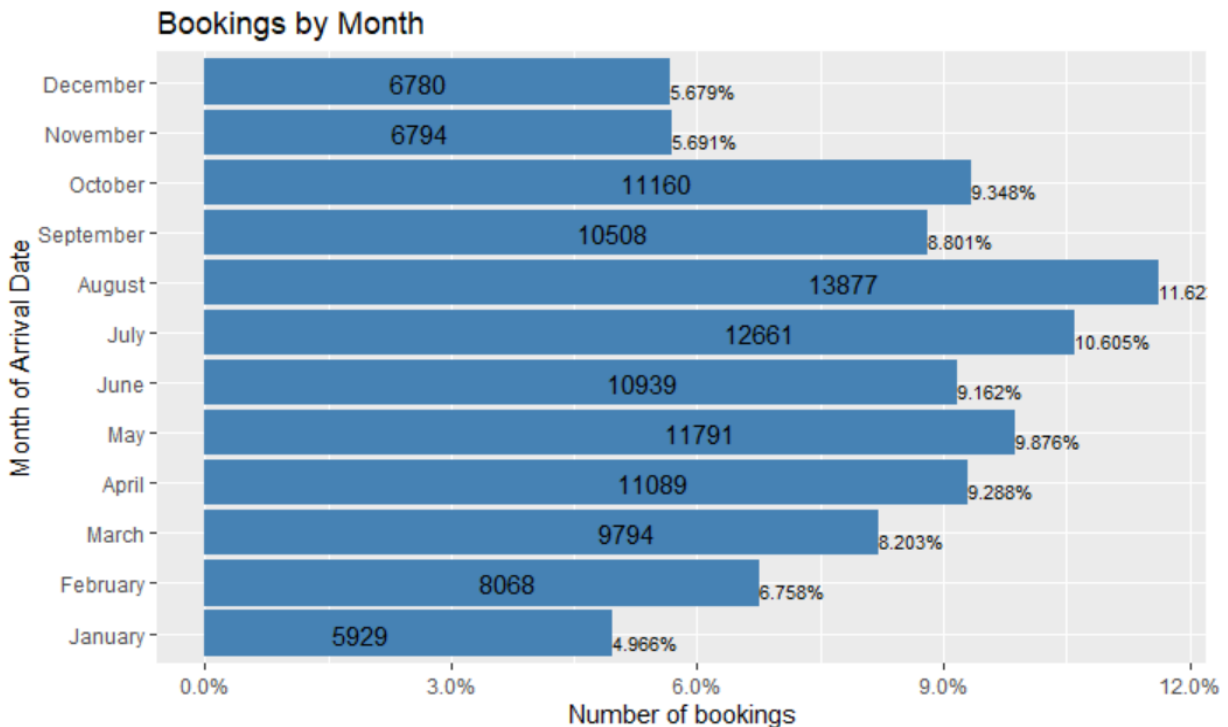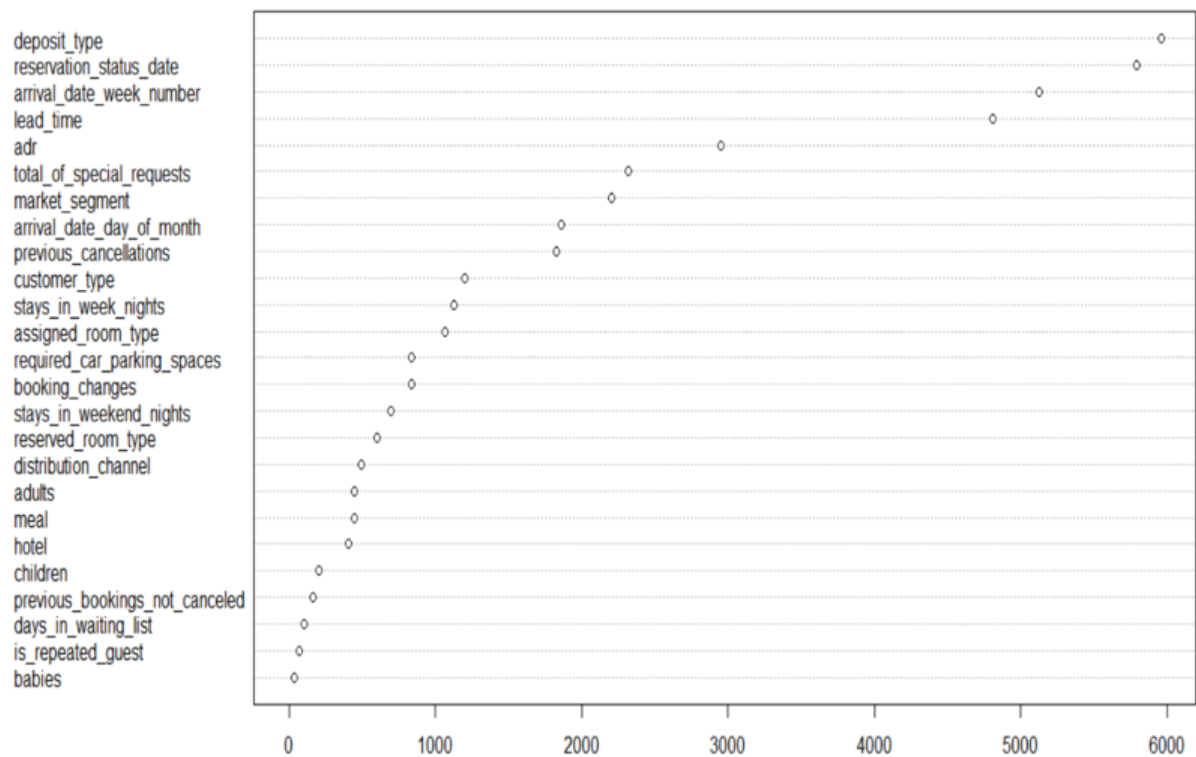**Exhibit C.** Random forest model for booking cancellations

Exhibit D. Random forest model for Special requests