

Bin Yang (by2303)
Xiangyi Liu (xl3048)
Yihan Feng (yf2555)
Zhongqi Yue (zy2442)

Analysis and Prediction of Hotel Booking and Cancellation

The dataset provides information about the duration of stay, the time of booking, and the visit's daily rate. We are interested in the most popular time during the year for hotel booking and the association between the daily rate and the length of stay. For customers, what factors contribute to a better daily rate. For hotel managers, what kind of order is more likely to be canceled.

We intend to create a website introducing the dataset, including interactive plots displaying the variables in the dataset, and creating shiny apps for users. In addition, we will perform statistical analysis using several supervised learning techniques and display the results.

We plan to perform statistical analyses using several supervised learning techniques. We are interested to see whether there is an association between certain variables and hotel cancellation. We first plan to create the correlation plots of the variables and detect any significant correlation. After selecting interested variables, we will fit several supervised learning models and select the best performing model.

In the exploratory data analysis, we will plot the variables and see whether we could see the correlation between the distribution of cancellation and any seasonality trend.

We expect coding challenges include, data cleaning, supervised learning models (variable selection), and HTML coding.

Data was obtained directly from the hotels' PMS databases' servers by executing a TSQL query on SQL Server Studio Manager. (Antonio, et.al)

(<https://www.kaggle.com/jessemostipak/hotel-booking-demand>)

November 7	Form a team and submit a proposal	Written proposal document
November 10-13	Project review meeting	Brainstorm questions that could be answered by the dataset
Nov 13 - Dec 5	Report	<ol style="list-style-type: none">1. Overview of the dataset2. EDA and visualization3. Prediction model on order cancellation

Bin Yang (by2303)
Xiangyi Liu (xl3048)
Yihan Feng (yf2555)
Zhongqi Yue (zy2442)

Nov 28 - Dec 5	Webpage and screencast	1. Webpage build
		2. Made explanatory video

Reference

Antonio, N., Almeida, A., & Nunes, L. (2018, November 29). Hotel booking demand datasets. Retrieved November 06, 2020, from <https://www.sciencedirect.com/science/article/pii/S2352340918315191>