Airline Demand Forecasting Report

By

Siqi Zhang

Introduction

Demand forecasting is important to the success of the airline business. Forecasting models aim to maximize revenue by matching demand to available seats. In this project, our group developed four advanced models utilizing additive and multiplicative models to forecast the demand on a given departure date depending on the booking date.

Methodology

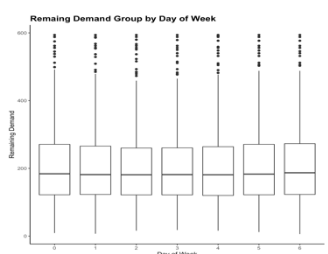
We divided the training dataset into different groups based on different days prior and/or day of week and trained our model on the training dataset, then validated the accuracy of our model on the validation dataset.

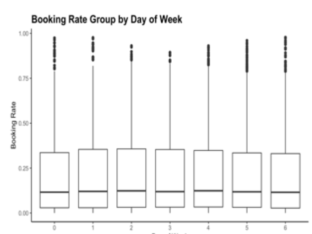
First, we added new variables to the training dataset. We created **days prior** which shows the number of days left before departure based on departure date and booking date variable. Then we got the **final demand** which was the cum bookings when days prior = 0. we added **day of week**, was added which corresponded to the booking weekday. We added **remaining demand** which equals to final demand – cumulative booking (the number of cumulative bookings for the given departure date). Finally, we added the **booking rate** which was the cumulative bookings / final demand.

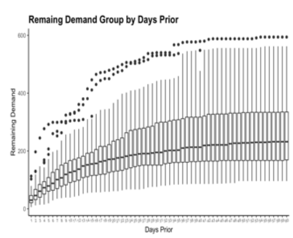
Our Advanced model utilizes the Additive & Multiplicative models:

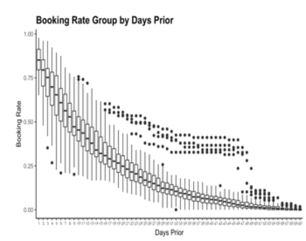
- Additive forecast = Historical remaining demand + cumulative booking
- Multiplicative forecast = Historical booking rate + cumulative booking

Then we started to pick method for forecast model. From the four graphs below, we can find that no matter in which way to group data (group by days prior or day of week), there always have some outliers which far from the main track of data, and the distribution of data points are asymmetric (the closer to departure date the less of the remaining demand), so using median to calculate the grouped data will be a better choice which are not affected by extreme values. (See below)









Therefore, we developed four models using the training dataset to forecast final demand:

- additive model with median remaining demand grouped by days prior
- additive model with median remaining demand grouped by day of week and days prior
- multiplicative model with median booking rate grouped by days prior
- multiplicative model with median booking rate grouped by day of week and days prior.

After that, we used **MAD** to calculate the error for each model and pick one model with minimize error which will be used on the validation data. Finally, we merged our model on validation data, then we got the forecast final demand data. And we calculated **MASE** to validate the accuracy of the model, which equals to around 60%.