To: Misuk Lee

From: Mansi Agarwal

**Objectives**

We aim to analyze Airline Bookings data to understand the demand curve and forecast the demand for future using various forecasting methods. Further, we calculate MASE (Mean Absolute Scaled Error) for each of these models and recommend the most suitable method for Airline Bookings demand forecasting.

**Data**

We use Training data set (that comprise of Departure Date, Booking Date and Cumulative bookings) in order to understand the booking curve for the months of May, June and July while our validation data is for the month July (that comprise of Departure Date, Booking Date, Cumulative bookings, Final Demand and Naïve Forecast). In both Training and Validation data we calculate “Days Prior: The number of days between booking and departure dates” and “Departure Day: The day of the week for departure date”. Since we have Validation data only for 28 Days Prior, we have removed the data for rest of the days to improve performance.

**Forecasting Methods**

In this assignment, we have used Additive, Multiplicative and Multiplicative using Daily growth models to predict the demand on the validation data for a particular booking date. From the validation data, we have removed the data for Days Prior = 0. Brief description of the models are as follows:

1. Additive Model – In this model, we calculate the remaining Demand from Training data for a particular Days Prior and Departure Day by subtracting Cumulative bookings and bookings on Days Prior = 0. We then use the Validation data to find **4 Nearest Neighbors** (by subtracting Cumulative Bookings in Training and Validation data) for a particular Days Prior and Departure day in Training data. The average of the remaining demand from these four Nearest Neighbors is then added to the Cumulative bookings in the Validation data to arrive at the Final Forecast.
2. Multiplicative Model – In this model, we calculate the booking rate from Training data for a particular Days Prior and Departure Day by dividing Cumulative bookings by bookings on Days Prior = 0. We then use the Validation data to find **4 Nearest Neighbors** (by subtracting Cumulative Bookings in Training and Validation data) for a particular Days Prior and Departure day in Training data. The average of the booking rate from these four Nearest Neighbors is then used to divide the Cumulative bookings in the Validation data to arrive at the Forecast.
3. Multiplicative Model using Daily Growth - In this model, we calculate the booking rate from Training data for a particular Days Prior by dividing Cumulative bookings for adjacent days (for ex. Days Prior=0 / Days Prior=1). We then use the Validation data to find **4 Nearest Neighbors** (by subtracting Cumulative Bookings in Training and Validation data) for a particular Days Prior and Departure day in Training data. The average of the booking rate from these four Nearest Neighbors is then used to divide the Cumulative bookings in the Validation data to arrive at the Forecast.

MASE is then calculated against the Naïve Forecast to compare these models and arrive at the recommendation.

**Functions**

1. **airlineForecast** – Takes Forecast Method, Training and Validation data as input and return MASE and Forecast on the validation data.
2. **calculateDPDD** – This function calculates Days Prior and Departure day for given Departure and Booking dates.
3. **calculateDemand** – This function calculates remaining demand/booking rate for the respective Forecast method chosen as an input.
4. **calculateForecast** – This function finds 4 Nearest Neighbors for the Validation data and then depending upon the Forecast method selected by the user calculates the Forecast.
5. **calculateMase** – calculates MASE using Final Demand, Forecast and Naïve Forecast.

**Recommendations**

MASE for our Additive Model is 73.34%, Multiplicative is 85.66% for Multiplicative using Daily Growth is 498.49%. Therefore, we recommend using Additive Model for Airlines Booking demand Forecasting.