Milestone 0: Project Proposal

Team Members: Phuong Thao Nguyen, Edmund Hui, Promod Rajaguru

Project title: When Should I Book My Tickets? Analysis and Prediction of Optimal Flight Booking Times

Abstract

Our study investigates flight durations, geographic influences, and temporal trends across the US to pinpoint optimal moments for purchasing airline tickets. Analyzing factors like seasonality, demand, flight duration, and location, our research aims to unveil key features dictating ticket pricing. This insight caters to diverse groups, empowering individuals like students, business professionals, and families to make informed travel plans. We employ a predictive approach, essential for the dynamic nature of airfare. Utilizing models such as ARIMA for time series analysis and Machine Learning algorithms like Random Forests and GBMs, we forecast changes in flight prices. This forecasting facilitates timely and cost-effective travel decisions, helping travelers capitalize on the best available deals for leisure, education, or business, especially during peak travel periods.

Data Description

Our data is a 31GB CSV file which contains purchasable flight tickets found on Expedia between the dates of 2022-04-16 and 2022-10-05. Flight paths represented are exclusive to the following airports ATL, DFW, DEN, ORD, LAX, CLT, MIA, JFK, EWR, SFO, DTW, BOS, PHL, LGA, IAD, OAK. There are 27 structured columns in the data providing detailed information about any given flight including:

- Basic Info: Flight dates, Origin and Destination, Flight Price
- Ticket Info: Refundability, Number of Stops, Cabin Class
- Plane Info: Airline Name, Airline Code, Plane Type, Empty Seats

Relevant Links

Dataset Link: https://www.kaggle.com/datasets/dilwong/flightprices

Collection Process: https://github.com/dilwong/FlightPrices