Lupe Covarrubias

SpringBoard Data Science Bootcamp

Capstone Two: Project Proposal

In this project we will aim to increase knowledge of cheapest flights by identifying the key factors influencing flight fares. Airfare can be highly expensive and price can depend on many factors. We would like to know when is the best time to purchase a flight in order to find the cheapest fare. We need to find the most important factors that contribute to the daily price of plane tickets and how those factors raise or lower the price. This includes destination, departure and arrival time, number of stops, airline, etc. How can we determine the perfect combination of factors to find the lowest price for a plane ticket?

We will look at the historical features of the dataset and see how they have changed over time and how those features have influenced ticket price over time to get an idea of their general relationship with flight price. Once we gain insight to those features we will investigate the key factors that directly influence ticket price through graphs. We will eventually use a linear regression model to predict flight prices based on those key features. There are many factors that are involved in price fluctuation; Price fluctuates daily so this is just an overall estimate of how cheap or expensive flights would be in general and not a day-to-day estimate of the cheapest flights. Our stakeholders in this operation are Director of operations - John Smith and the Database Manager - Jane Doe. Our data source for this project will be from Kaggle and is listed here:

https://www.kaggle.com/datasets/yashdharme36/airfare-ml-predicting-flight-fares .