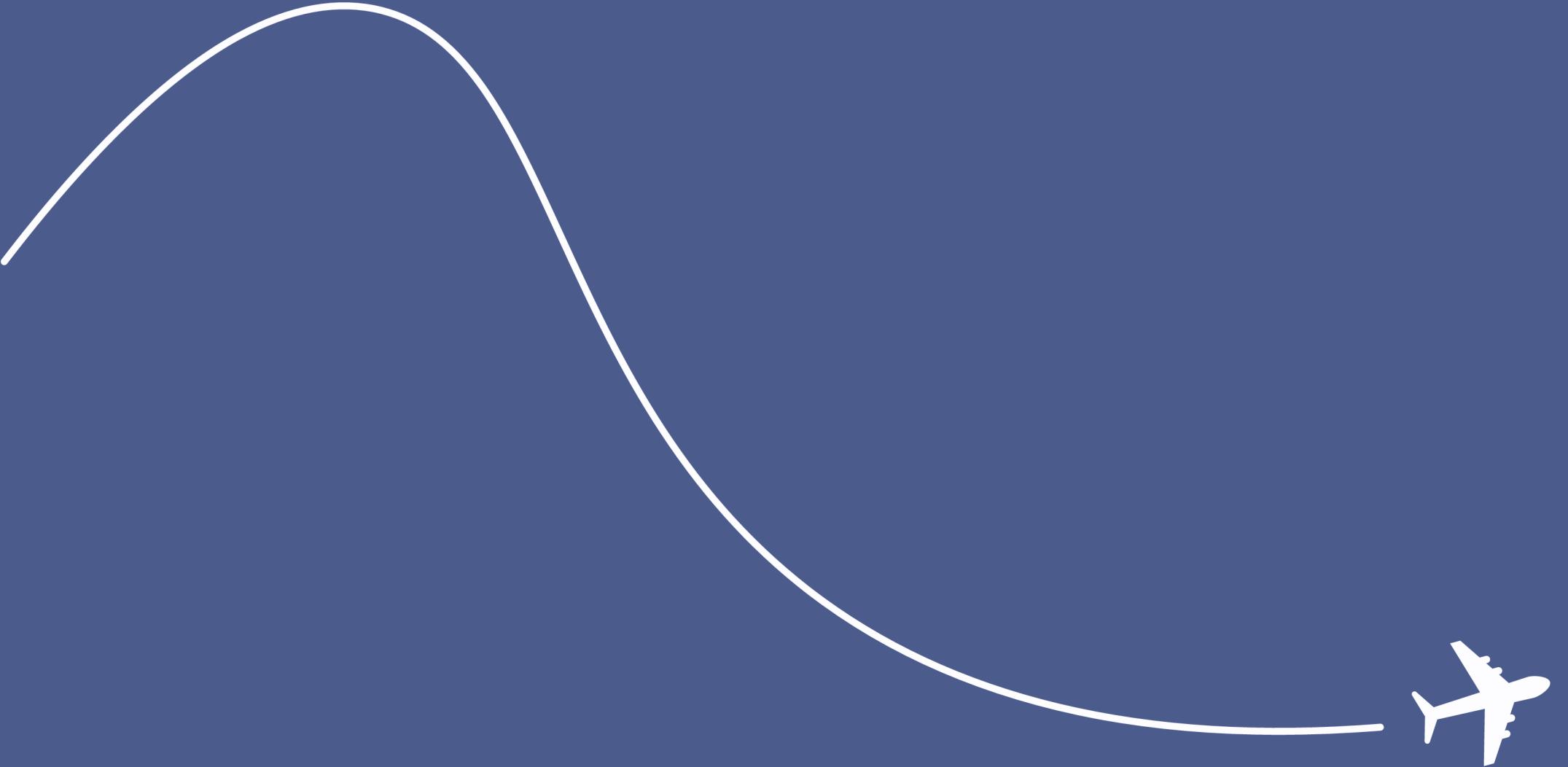


Ticket Price Predictor

Amirreza, Bru, Jonathan, Victor





Jimbo Airlines

**Sky-High Ambitions,
Down-to-Earth Fares**

Project Overview

- Jimbo Airlines is a new **low-cost carrier** entering the Indian aviation market, set to begin operations in the 1st quarter of next year.
- Its primary goal is to gain market share and build a strong customer base by **leveraging predictive analytics** for an aggressive ticket pricing strategy.
- Desired Outcomes:** rapid market penetration, build brand recognition and establish a competitive advantage

Data Selection and Preparation

Scope: Flight bookings between India's top 6 cities .

Period: February 11 to March 31, 2022.

Key Features:

- Take off and landing times (binned).
- Departure and arrival cities.
- Number of stopovers.
- Days until departure (booking to trip date).
- Flight duration and price.
- Airline names (Indigo, SpiceJet, AirAsia, Air India, among others)

Preparation Steps:

- Standardized dates, times, data types, and column names.
- Removed irrelevant columns.

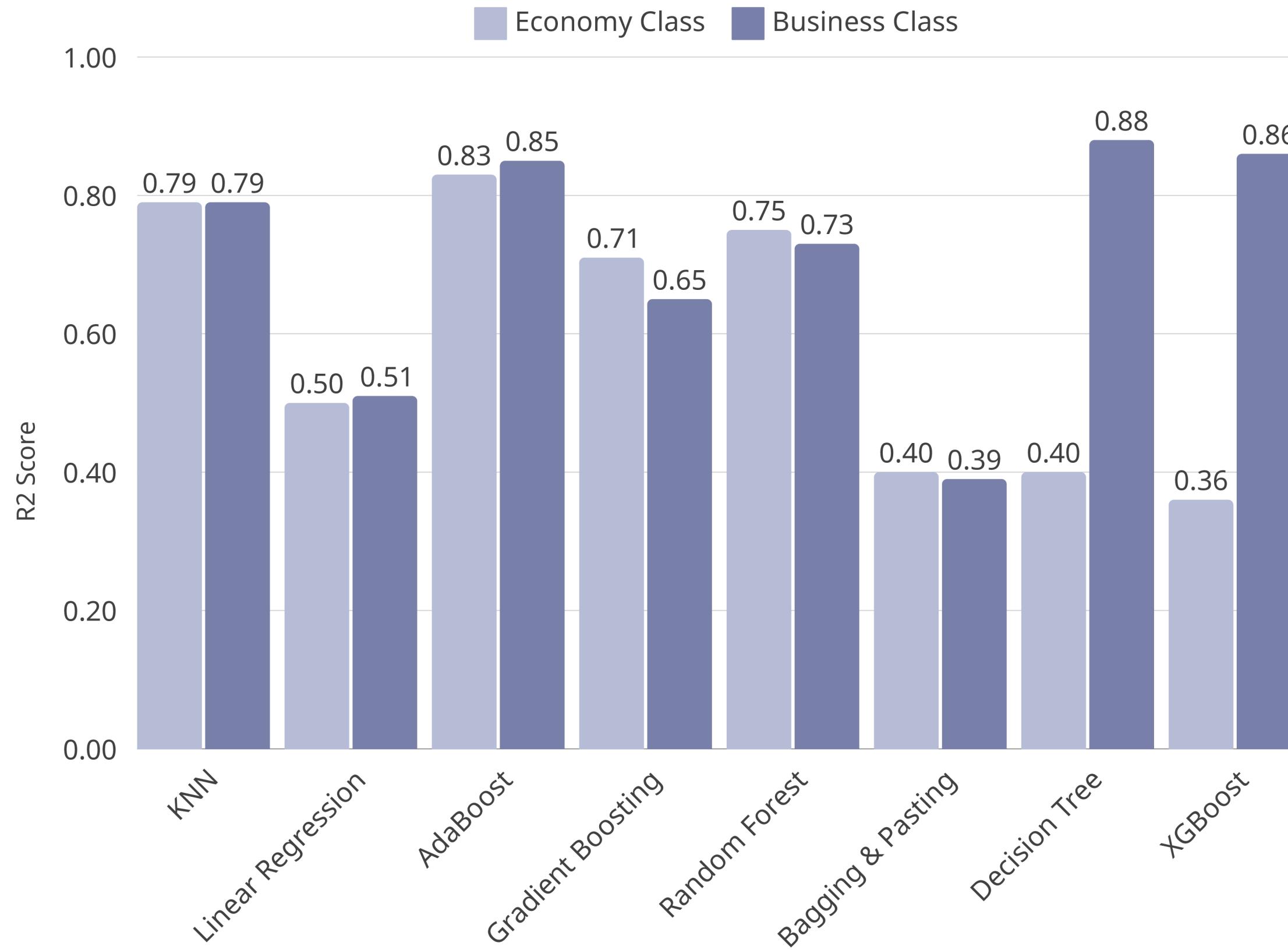
Feature Engineering and Selection

Tableau Dashboard

Categorical Columns: Converted classification columns into dummy variables.

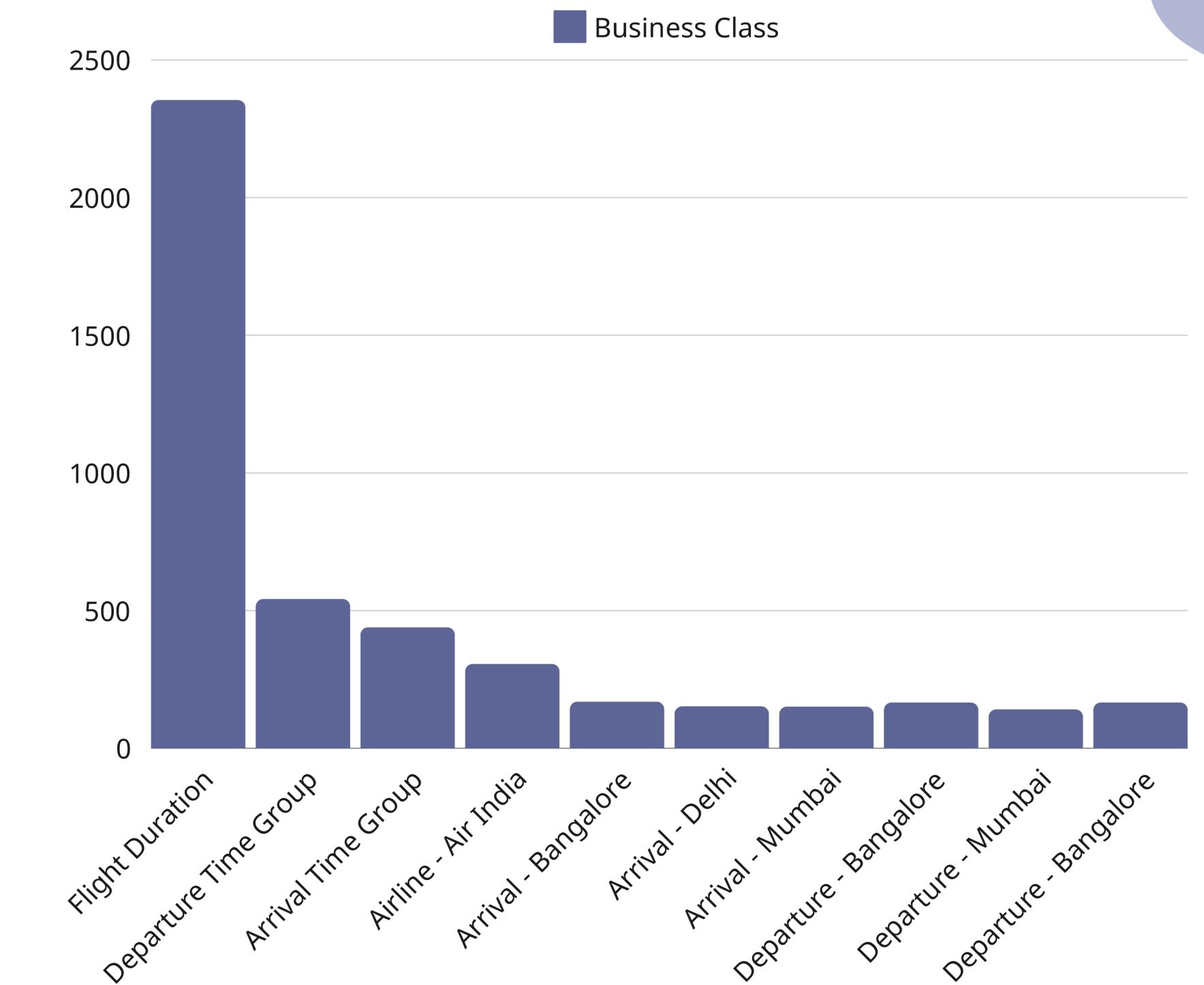
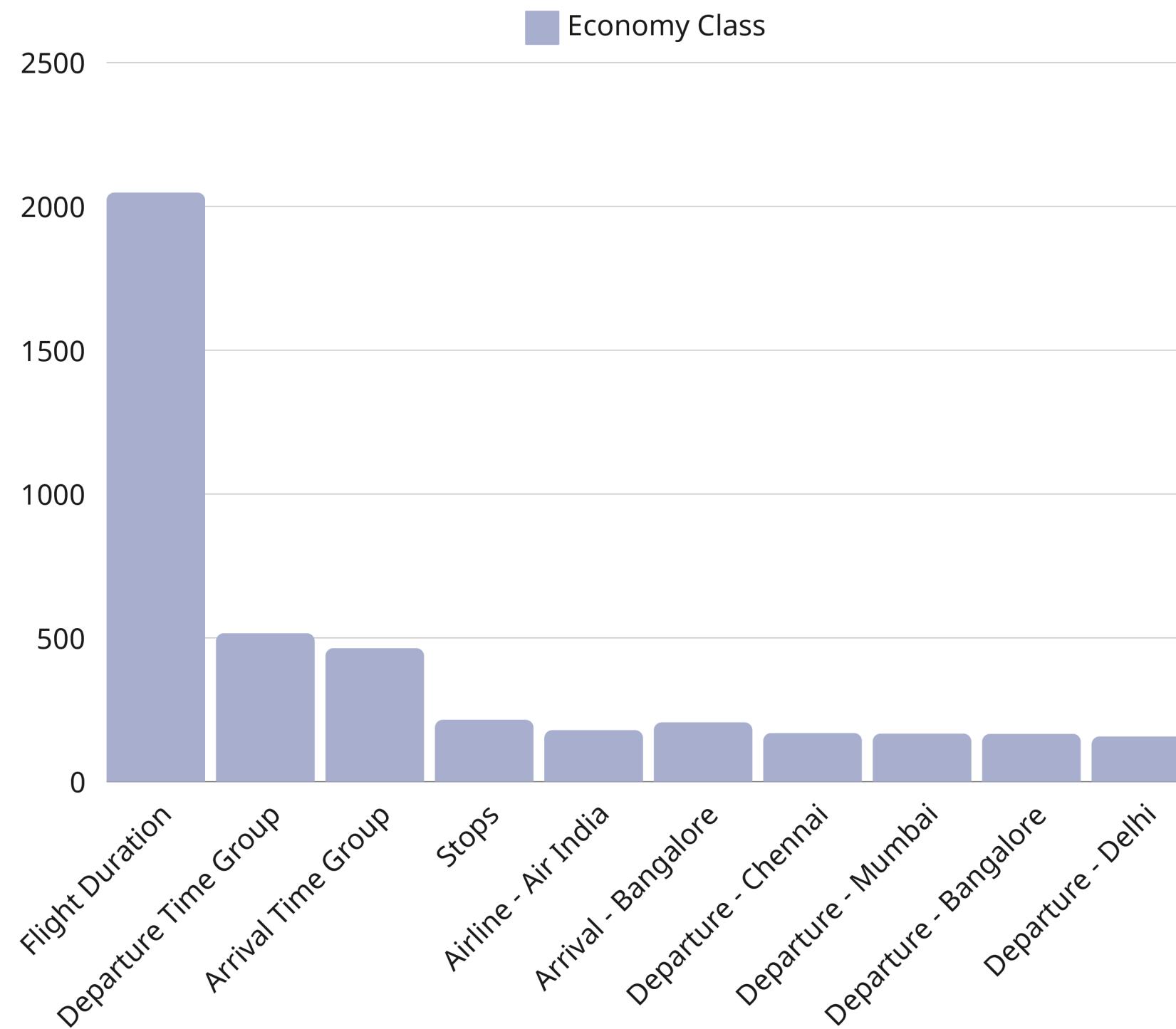
Numerical Columns: Applied MinMaxScaler for normalization.

Model Building and Evaluation



Hyperparameter Tuning and Model Optimization

Feature Importance



Key Findings and Insights

Extreme Gradient Boosting Regressor

First Model on Business Dataset

Business R²: 0.8561

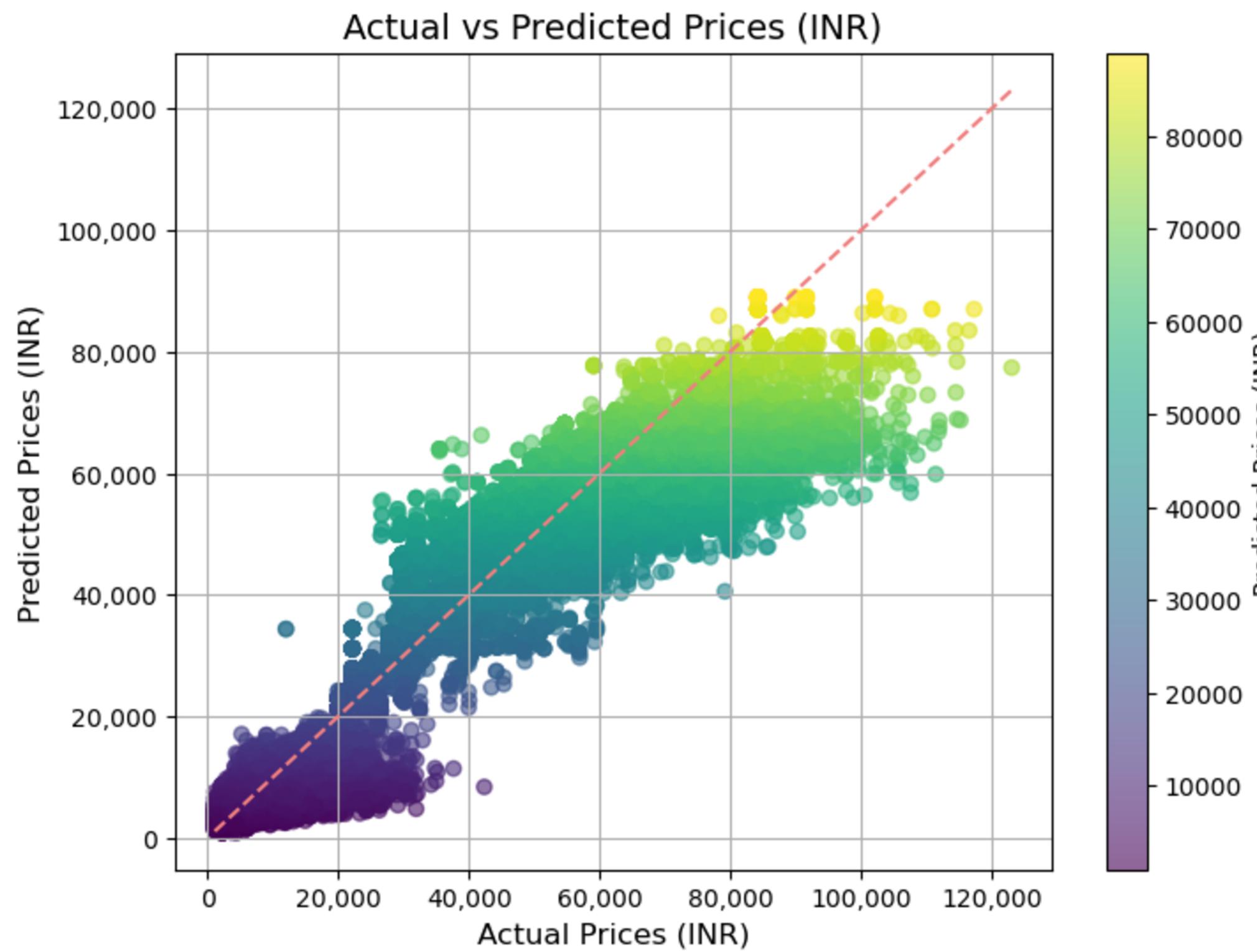
Business RMSE: 4932.6074

Second Model on Economy Dataset

Economy R²: 0.3603

Economy RMSE: 2983.9987

100 INR = 1,12 EUR ≈ 1 EUR



Key Findings and Insights

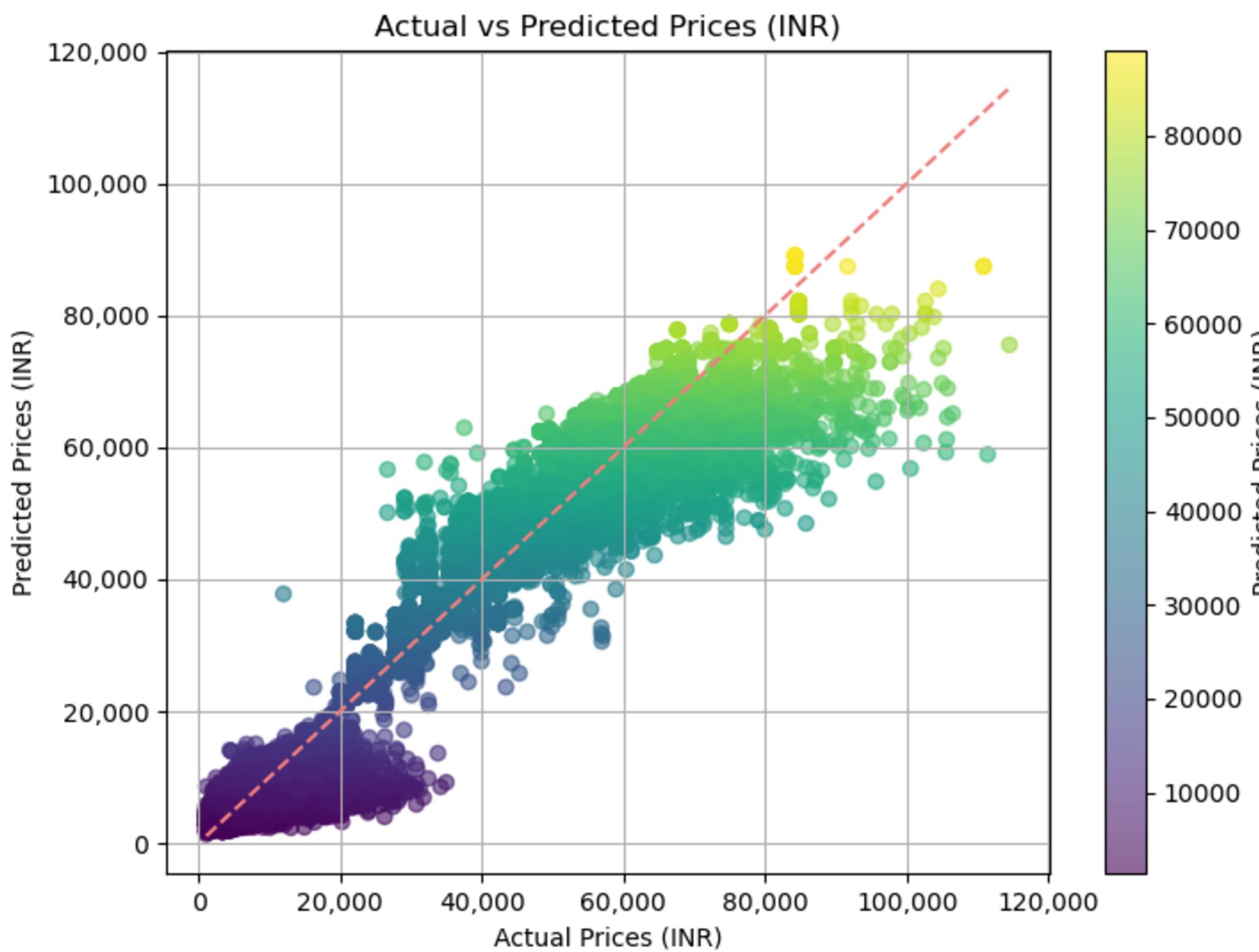
Extreme Gradient Boosting Regressor

On a single balanced dataset

R²: 0.9653

RMSE: 4634.4477

100 INR = 1,12 EUR \approx 1 EUR



Future Improvements

Data Enhancements for Price Analysis

Seasonal Demand: Incorporate a year's worth of flight price data.

Airline Features: Standardize service-related data (luggage policies, meals).

Economic Factors: Include fuel prices and inflation rates.

Optimization Techniques: Leverage advanced tuning (Bayesian, Optuna).

Real-World Application

Model Application & Automation

Prediction Focus: Forecast flight prices in India for February and March.

Automation: Develop an API for real-time price retrieval.

Model Updates: Retrain periodically with fresh data for sustained accuracy.

Challenges:

Black Box Problem: Difficulty explaining ML/AI results, good or bad.

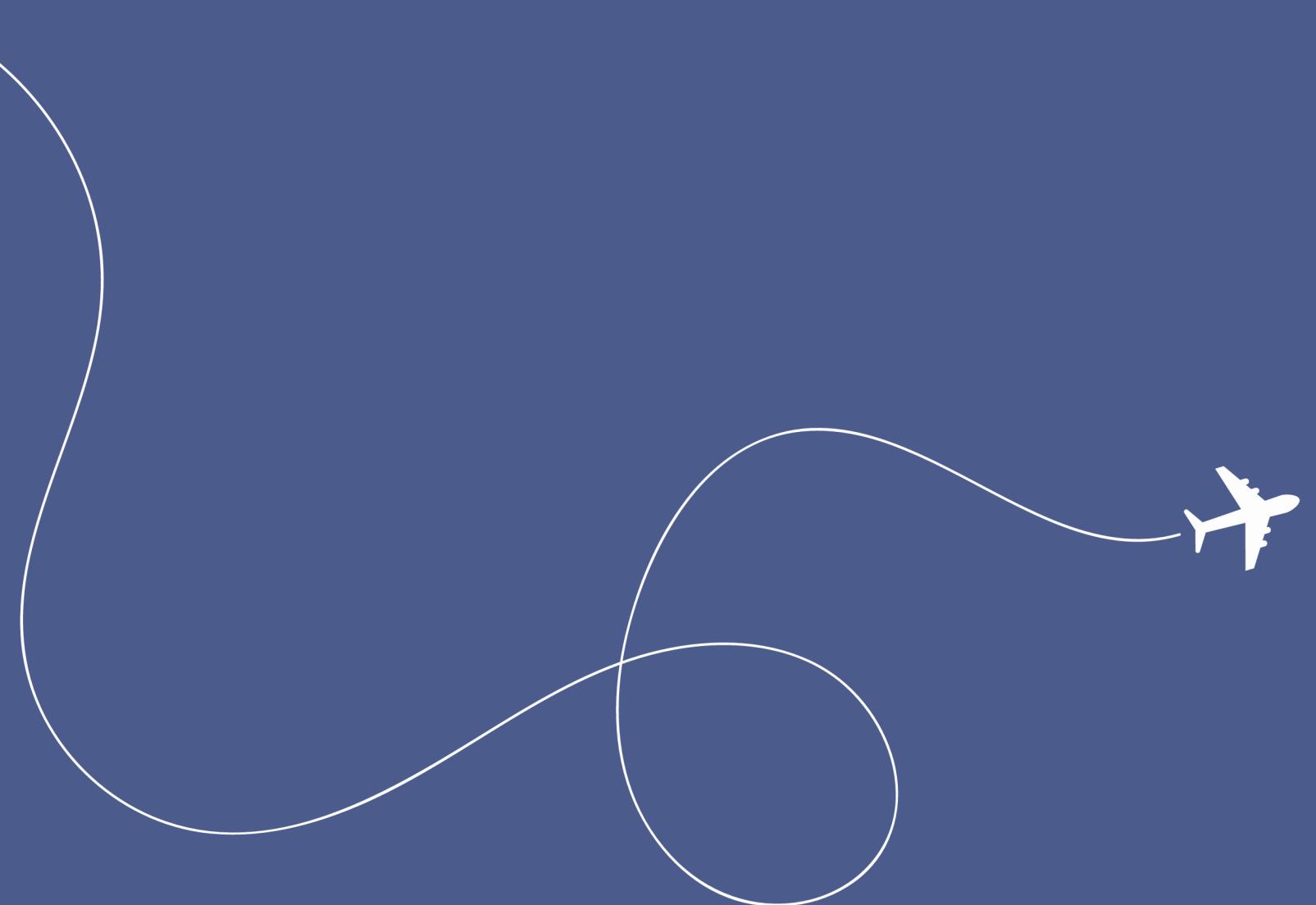
Time Management: Balancing tasks and deadlines.

Code Rigidity: Debugging complexities.

Learnings:

Hands-on learning: Each member builds their own models.

Communication tools: Slack, Trello, and daily meetings.



Thank you!