



AERO DATA

DATA SCIENCE COMPETITION

Decode The Data Define The Future

RULEBOOK

What's the Challenge?

The challenge is to predict Flight Ticket Prices using a real-world dataset of Indian domestic flights (March–June 2019). Participants must build Machine Learning Models that can accurately estimate ticket prices based on features like airline, journey date, source, destination, duration, number of stops, and more.

Judging Criteria

1. Model Performance (40%)
 - Accuracy of predictions based on evaluation metrics (e.g., RMSE, MAE, R²).
 - Generalization ability (train vs. test performance).
2. Feature Engineering & Preprocessing (30%)
 - Handling of missing values, outliers, and categorical/numerical variables.
 - Creativity in generating meaningful features.
3. Approach & Methodology (30%)
 - Clarity in problem understanding.
 - Justification for model selection and techniques used.
 - Proper experimentation and comparisons.

Rewards

The Top 3 Winners will receive a Certificate of Excellence as official recognition of their achievement, along with a special shoutout on all social media handles, where they'll be featured.

Beyond the prizes, every participant benefits from valuable opportunities, including hands-on experience with a real-world ML dataset, a portfolio-worthy project to showcase, and the chance to learn, compete, and grow as a data scientist.

Submission Guidelines

- Submission via Form
- Last Date of Final Submission: 18 September

Problem Statement & Dataset Information

Airline ticket prices are highly dynamic and influenced by multiple factors such as time of journey, airline, source, destination, duration, and number of stops. For travelers, this unpredictability makes planning difficult, while for data scientists, it presents an exciting real-world problem. In this challenge, participants are tasked with building machine learning models that can accurately predict flight ticket prices using the given dataset of domestic flights (March–June 2019). The objective is to apply data preprocessing, feature engineering, and model development to minimize prediction error and achieve the best performance metrics.

Competition Info

Start Date: 13 September

End Date-18 September

Eligibility: Open for all

Hosted by: Analytical Arena -Data science club

Download Dataset- <https://tinyurl.com/Aero-Data>

Submission Form- <https://forms.office.com/r/xyxYz9Brhw>