

Wrocław, 29th of May 2023

Dear colleagues,

We have prepared for you a small project solution to which should be delivered to us in one week from the time when you get an e-mail from us. Please be aware that solutions sent after this deadline will not be accepted (5th of May 2023).

Project description

The goal of the project is to prepare a regression model predicting so called Take-Off Weight of an aircraft. The initial training set contains a list of flights of Ryanair from 2016-10-01 till 2016-10-15 (training.csv). You are given a validation dataset (validation.csv) where the Take-Off Weight must be predicted. The evaluation will be conducted by comparing the results provided by you with values known to Ryanair. RMSE is to be used as a base measure of a model quality.

You are expected to deliver the results as a CSV file. Additionally, a description containing exploratory data analysis of the dataset and the method must be attached.

Description of columns in the dataset

DepartureDate – date of departure
DepartureYear – year of departure
DepartureMonth – month of departure
DepartureDay – day of departure
FlightNumber – flight number
DepartureAirport – departure airport code
ArrivalAirport – arrival airport code
Route – route
ActualFlightTime – flight time
ActualTotalFuel – burnt fuel
ActualTOW – take-off weight (kg)
FlownPassengers – flown pax
BagsCount – flown bags
FlightBagsWeight – flown bags weight

Regards,
Data & Analytics Team