

Ryanair Challenge. February 2024

The proposed challenge involves developing a predictive model to estimate the quantity of fresh food loaded onto aircraft once a day, prior to the commencement of operations for flights in November 2023.

This document describes the [evaluation criteria](#). All details about the challenge, including the dataset description and delivery requirements, are available in the [technical document](#). Please read both documents carefully.

1. Evaluation criteria

The evaluation criteria for this competition encompass both technical and business aspects, each contributing **50% to the final grades**.

- **Technical criteria (50% weight on the final grades)** are meant to assess the general ability of the teams to handle, interpret, and understand the data, build predictive models for the given problem and evaluate their performance according to an objective metric.
- **Business criteria (50% weight of the final grades)** are more general in scope, and they involve the ability to devise what value can be extracted from this data, formulate relevant business questions and try to find answers to them in the data. The quality of the presentation and the ability to communicate clear and powerful ideas on the final pitch will also be part of the business criteria. ***Please be aware that the announcement of finalists will not occur until shortly the closing ceremony, thus everyone should be prepared to present.***

Teams are required to submit:

1. **Technical materials:** (a) The code they developed to tackle the problem (ideally with comments, or jupyter notebooks alternating code and text explanations). (b) The output file. (c) The model performance metrics.
2. **Business materials:** An executive summary (3-5 pages) with their inspection strategy and value propositions.

The [technical criteria](#) will be assessed on the basis of the **first submission**, and the performance metrics obtained by the best model **submitted by [Google Forms](#)**. [Business criteria](#) will be assessed mostly from the executive summary.

2. Technical criteria (50% weight on the final grades)

The models submitted will be evaluated based on three chosen metrics: **MAPE, RMSE and MAE**. ***The superior predictive model will exhibit the smallest values across all the three metrics.***

In addition to this, the code should be clearly structured and results should be interlaced with explanations in jupyter notebooks. The notebooks should be clearly written, and explain the process followed starting from the raw dataset, cleaning and preprocessing, exploratory data analysis, model formulation, hyperparameter tuning (if needed), final metrics and discussion.

When developing your models, please pay special attention to:

- Overall understanding of the problem
- Exploratory data analysis
- Feature engineering
- Predictive models performance

Overall understanding

Ensure that you understand the meaning of each predictor variable in the different datasets: what it means, in which units is it expressed, how is this data registered, at which moment in the time or day and location, could it contain errors? could it contain outliers ? can we trust the data ? Using common sense, will a given predictor variable be useful to predict our target?

Exploratory Data Analysis

Getting acquainted with the datasets is a first necessary step before any modelling on the data takes place. Explore the data distribution, which variables are categorical and which are numerical, do we really understand the meaning of each variable? Are there any correlations among the variables? Are there predictor variables with missing values or outliers? Can we trust the values of the data? Try to formulate hypotheses and understand your datasets before further exploration is conducted. Create good visualizations that help develop your intuition and understand the patterns. If necessary, decide how to handle missing values by either data imputation or removing rows/columns from the dataset. Experiment with various levels of data aggregation to ensure the robustness of your predictions.

Feature engineering

Which features will you use in your predictive model? Is it legitimate to use all the provided data? Can you imagine how the model will be used in production?

Can you enrich your dataset with external information? At a minimum, you will need to combine multiple datasets and/or create new features to train models. Additionally, you should experiment with various aggregation and drill-down techniques to analyze the data effectively. Be creative: anything that you can build on the given data that might have a more direct connection to what you are trying to predict will improve your models performance.

Model performance

Performance metrics and the output file containing predictions should be submitted via **Google Forms**.

The proposed challenge involves developing a predictive model to estimate the quantity of fresh food loaded onto aircraft once a day (at line of flight level), prior to the commencement of operations for flights in November 2023. **The output file format is described in the Technical document.**

Remember: The models submitted will be evaluated based on three chosen metrics: **MAPE, RMSE and MAE. The superior predictive model will exhibit the smallest values across all the three metrics.**

Feel free to try different families of models, adjust their parameters, add regularization, go back to your preprocessing cycle and continue iterating, etc. **You can check your performance metrics (compared to other teams) anytime accessing this Google Sheets file using your IE email account: [Ranking of submissions](#).**

3. Business criteria (50% weight on the final grades)

Extracting value from data

For this task, you have to develop a predictive model to estimate the quantity of fresh food loaded onto aircraft once a day (at line of flight level), prior to the commencement of operations for flights in November 2023. Your task is to determine how to apply this model in a business case and how it can bring value to the company, aiming to maximize sales revenue while minimizing waste.

Knowing the expected sales for the 6 fresh product groups per line of flight level will allow the company to ensure provision of fresh food to customers according to their preferences, as well as avoid food waste.

1. Predict fresh food requirements at the line of flight level, where each aircraft follows a predefined schedule of flights.
2. Maximize sales by identifying customer patterns based on the composition of the line of flights, customer preferences, etc.
3. Minimize shortages of fresh food in alignment with client preferences.
4. Reduce food waste.

In this manner, the company can enhance customer satisfaction and increase revenue while also being environmentally friendly.

Your plan should not only fulfil these requirements but also propose innovative strategies to optimize how the company allocates food for its line of flights. We encourage you to analyse and suggest improvements to the current strategy, leveraging your data analysis and strategic thinking skills to enhance operational efficiency and customer satisfaction.

In addition to meeting the outlined criteria, your business plan should also demonstrate the measurable economic impact of the proposed improvements. This includes detailed estimations of potential revenue increases and environmental benefits. You are encouraged

to use additional sources and work with hypotheses to make these estimations. Your plan should provide a clear analysis of how your strategies could positively affect the company's bottom line while also contributing to environmental sustainability. This component is crucial in showcasing the real-world applicability and effectiveness of your proposal.

The evaluation will concentrate on the coherence and realism of your proposal, the precision of your impact calculations, and the clarity of your presentation. This challenge requires not only finding technical solutions but also considering their wider implications, including revenue impact, cost efficiency, and sustainability.

Special consideration will be given to proposals that, in addition to solving technical problems, contribute to enhancing airline operational efficiency and comprehensively address the economic, environmental, and customer satisfaction challenges linked to fresh food allocation on flights.

All participating teams will be assessed on a combination of these two factors **50% technical criteria and 50% business criteria**. A team with a model with poor performance might still qualify for the final if their business model and strategy is outstanding, so try to devote some time to both tasks. And remember to work in parallel and divide your team according to expertise and capacity.

Pitch clarity

The teams that are selected for the final phase will pitch their results in front of a jury. Their technical results will have already been assessed by the technical jury on the basis of their submitted notebooks and model performance. On the pitch you will need to transmit clear and powerful ideas that highlight your results and show your understanding of the problem, your ability to harness value from the data and your ideas to contribute to the problem under consideration. Focus on the large scale goals, while showing evidence that your technical skills are solid, but do not use your time to explain straightforward technical solutions, unless you think that they are really essential. ***Please be aware that the announcement of finalists will occur during the closing ceremony, thus everyone should be prepared to present.***