

# Modeling Passenger Baggage Preferences in Air Travel

*Michel Robin, Bickerstaffe Faber and Magnin Antoine*

University of Lausanne, HEC

Machine Learning in Business Analytics

2024

Group E

04.05.2024

Abstract: | The following machine learning project focuses on...

## Introduction

- Overview and Motivation

In the rapidly evolving domain of air travel, passenger satisfaction and operational efficiency are paramount. A significant aspect of enhancing traveler experience lies in personalizing their journey, starting from the moment they book a flight to when they collect their baggage at their destination. One area ripe for innovation is the optimization of baggage handling and preferences, which directly impacts customer satisfaction and logistical operations.

The project aims to utilize machine learning techniques to predict and understand the diverse baggage preferences and behaviors of air travelers. This initiative seeks to explore various factors influencing passengers' baggage choices, such as trip duration, purpose, and individual preferences, to enhance service personalization and streamline airport operations.

The motivation behind this project stems from the challenges faced by airlines and airports in managing the efficiency of baggage processes while maintaining high standards of customer service. By accurately predicting baggage preferences, airlines can optimize resources, improve the accuracy of baggage handling, and reduce the instances of lost or mishandled luggage. Moreover, understanding these preferences enables tailored marketing strategies, potentially increasing revenue from ancillary services like paid baggage options.

- Related Work
- Research questions

## Data

- Sources
- Description
- Wrangling/cleaning
- Spotting mistakes and missing data (could be part of EDA too)
- Listing anomalies and outliers (could be part of EDA too)

```
# Example of a code block
```

## Exploratory data analysis

- Mapping out the underlying structure
- Identifying the most important variables
- Univariate visualizations
- Multivariate visualizations
- Summary tables

## **Supervised learning**

- Data splitting (if a training/test set split is enough for the global analysis, at least one CV or bootstrap must be used)
- Two or more models
- Two or more scores
- Tuning of one or more hyperparameters per model
- Interpretation of the model(s)

## **Unsupervised learning**

- Clustering and/or dimension reduction

## **Conclusion**

- Brief summary of the project
- Take home message
- Limitations
- Future work?