# Business Case with Machine Learning
## Unsupervised Learning: Clustering

**MSc Business Analytics**

Prof. Marc Torrens

TAs: Gal·la Garcia-Castany

# The grade is based on 3 factors

**1.** **Technical Solution & Analysis – 40%**

- The student has appropriately followed the instructions.

- The student has applied the methodology and basic ML practices explained in class.

**2.** **Business Interpretation & Recommendations – 40%**

- The student is able to interpret ML models and their business implications.

- The student is able to propose business recommendations based on the findings.

**3.** **Design of Deliverable – 20%**

- The student communicates his/her results clearly, briefly and with a good design.
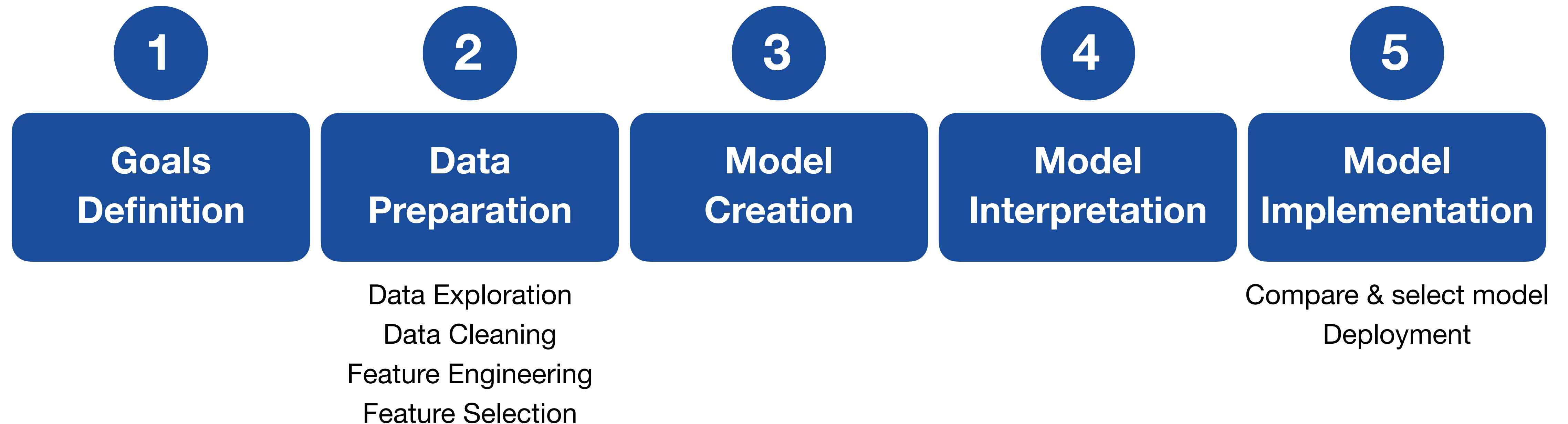
# Deliverables

**1.** **PDF**

- The objective of the slide deck is to present a business presentation explaining what has been done, why it was done, and its economic impact. Answering the questions of the case.

**2.** **Jupyter Notebook**

- Full code used with clear documentation to facilitate reproducibility and further analysis, addressing all the questions in the assignment.

# About the case

- Follow the ML lifecycle steps to the Online Retail case.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| **Goals Definition** | **Data Preparation** | **Model Creation** | **Model Interpretation** | **Model Implementation** |

**Data Preparation:**
Data Exploration
Data Cleaning
Feature Engineering
Feature Selection

**Model Implementation:**
Compare & select model
Deployment

# About the case

- An e-commerce company aims to optimize its marketing and customer retention strategies. Currently, all customers receive the same promotions and communications, but the results are not optimal. To improve campaign effectiveness, the company has decided to segment its customers into groups with similar characteristics and purchasing behaviors.

- To achieve this segmentation, use the "Online Retail" dataset, which contains transaction data from an online store serving customers from different countries. This dataset includes variables such as invoice number, product code and description, quantity purchased, transaction date, unit price, customer ID, and country of residence.

# About the case

- The main objective of this case is to apply clustering techniques to segment customers based on their purchasing behavior. Through this analysis, we aim to:
  - Identify different customer segments based on purchase frequency, spending patterns, and transaction volume.
  - Analyze the characteristics of each segment to understand consumption patterns and specific needs.
  - Propose personalized marketing strategies for each segment, such as targeted promotions, product recommendations, or loyalty programs.
  - Optimize data-driven business decision-making to improve profitability and customer satisfaction.

- Respond to the proposed questions while keeping these objectives in mind.

# Data Preparation

- 1. Load the dataset "Online Retail.xlsx" and conduct an initial exploration of the data.
- 2. Handle missing values appropriately.
- 3. Analyze correlations between numerical features and select appropriate independent variables for the model.
- 4. Perform feature engineering (feature selection, feature transformation, feature creation, and evaluation).
- 5. Apply necessary data transformations, such as normalization or standardization.

# Model Creation

- 1. Choosing a Clustering Algorithm: What clustering algorithm (e.g., K-Means, DBSCAN, Hierarchical Clustering) would you choose for this case, and why? Compare at least two algorithms in terms of suitability.

- 2. Determining the Optimal Number of Clusters: How would you determine the optimal number of customer segments? Explain and apply at least one method (e.g., the elbow method, silhouette score).

# Model Interpretation

- 1. Once the clustering model is trained, how would you analyze and interpret the resulting customer segments?
- 2. What business insights can be drawn from the characteristics of each cluster?

# Model Implementation

- 1. How can the company leverage the clustering results to improve its marketing strategies? Provide specific examples of personalized marketing actions for different customer segments.

# Business Case with Machine Learning
## Unsupervised Learning: Clustering Algorithms

**Business Analytics**

Prof. Marc Torrens

TAs: Gal·la Garcia-Castany