# EXPLORATORY DATA ANALYSIS - Analyse exploratoire des données

Objectif : Comprendre au maximum les données dont on dispose pour définir une stratégie de modélisation

## Analyse de Forme

* Variable target :
* Lignes et colonnes :
* Types de variables :
* Analyse des valeurs manquantes :

## Analyse de Fond

* Visualisation de la target (histogramme/boxplot)
* Compréhensions des différentes variables
* Visualisation des relations features-target
* Indentification des outliers

# PREPROCESSING

Objectif : Transformer le dataset pour le mettre dans un format propice au Machine Learning

* Elimination des colonnes inutiles
* Creation du Train set / Test set
* Elimination des NaN : dropna(),imputation, colonnes « vides »
* Encodage
* Suppression des outliers néfastes au modèle
* Feature Selection
* Feature Engineering
* Feature Scaling

# MODELLING

Objectif : Développer un modèle de machine learning qui réponde à l’objectif final

* Définir une fonction d’évaluation
* Entrainement de différents modèles
* Optimisation avec GridSearchCV
* (Optionnel) Analyse des erreurs et retour au preProcessing / EDA
* Learnig Curve et prise de décision