Introduction au machine learning

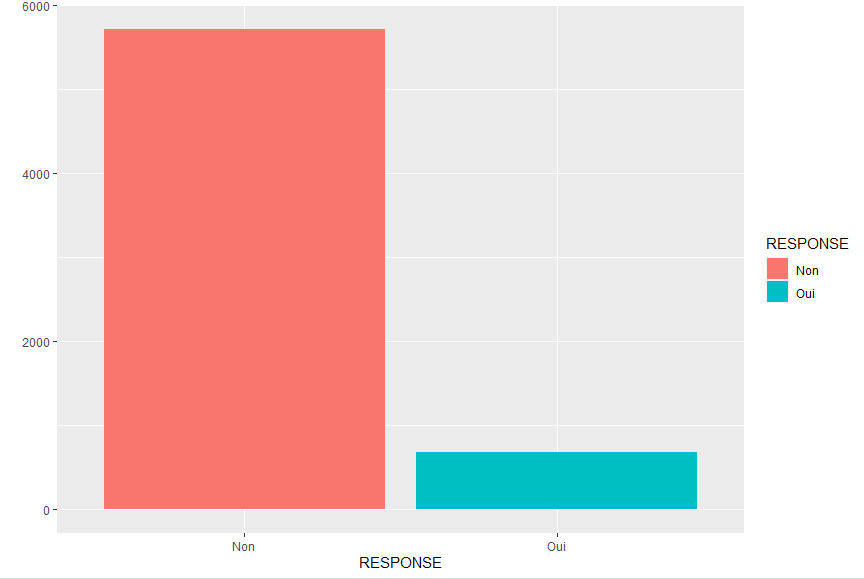
Maxime Prévôt & François Martin

# Exploration et visualisation des données

6400 instances

Non Oui

5721 679



# Pré-traitement des données éventuels

Classification supervisée

Test\_tree 1 Non Oui

2119 281

Test\_tree 2 Non Oui

2400 0

Test\_tree 3 Non Oui

2400 0

Taux\_succes1 0.81875

Taux\_succes2 0.8933333

Taux\_succes3 0.8933333

Matrice de confusion tree1

test\_tree1

Non Oui

Non 1914 230

Oui 205 51

Rappel

[1] 0.1992188

Specificité

[1] 0.8927239

Précision

[1] 0.1814947

Tx de vrai negatif

[1] 0.9032563

Matrice de confusion tree1

Rappel

[1] 0.1992188

Specificité

[1] 0.8927239

Précision

[1] 0.1814947

Tx de vrai negatif

[1] 0.9032563

Matrice de confusion tree1

Rappel

[1] 0.1992188

Specificité

[1] 0.8927239

Précision

[1] 0.1814947

Tx de vrai negatif

[1] 0.9032563

Matrice de confusion tree2

test\_tree2

Non Oui

Non 2144 0

Oui 256 0

Matrice de confusion tree3

test\_tree3

Non Oui

Non 2144 0

Oui 256 0

Courbe roc et indice AUC

test\_rpart("information", 5, TRUE, "orange")

dt\_class

Non Oui

Non 2144 0

Oui 256 0

AUC = 0.5

test\_rf(300, 3, FALSE, "red")

rf\_class

Non Oui

Non 2144 0

Oui 256 0

AUC = 0.634613492595616

test\_knn(20, 1, TRUE, "green")

Non Oui

Non 2139 5

Oui 252 4

AUC = 0.586350170534048

test\_svm("linear", FALSE, "red")

svm\_class

Non Oui

Non 2144 0

Oui 256 0

AUC = 0.569228442747202

test\_nb(0, FALSE, FALSE, "red")

nb\_class

Non Oui

Non 2076 68

Oui 230 26

AUC = 0.649616298390859

NaiveBayes = GOAT