Checklist de base : Machine Learning

# Analyse de forme

Variable target : SARS-Cov-2 exam result

Lignes/colonnes : 5644/111

Types des variables : 74 quantitatives, 37 qualitatives

Analyse des valeurs manquantes :

* Deux groupes se distinguent :
  + ~ 76 % : Tests concernant d’autres types de virus
  + ~89 % : Analyse de paramètres sanguins

# Analyse de fond

Visualisation de la target :

* 10 % de positifs
* 90% de négatifs (très déséquilibré !)

Très important lors de l’exploration de données : TRACER les histogrammes des variables continues

Signification des variables :

Relation variables/target :