**Plan du cours (7h) pour UPPA:**

***Matin :***

1. Présentation de HUPI ([lien](https://docs.google.com/presentation/d/174tYuT-vgmIGe0oYeBzj5zQN5cG4vty2/edit#slide=id.p52)) : 1h
2. Prise en main de Python ([lien](https://colab.research.google.com/drive/13bj8nRPiOhRb7OtOpzQ2rFQuS-R8rd1p)) : 1h
3. Présentation de l’objectif de la formation ([lien](https://docs.google.com/document/d/19cktY2LE_mph0rS-FpGRmUn21f_rcTo8081IBwgqywc/edit?tab=t.0)) : 30 min
   1. Présentation des sujets d’évaluation

*pausa (30 min)*

1. Présentation des méthodes d’analyse de données ([lien](https://docs.google.com/presentation/d/12ZMXYqdU9PccPGeIqXBnGigndFtM7dVOT0FnLnJDwTY/edit#slide=id.g31f5a4735e9_0_724)) : 1h 30
   1. Analyse univariée
   2. Analyse bivariée
   3. Tests de corrélation
2. Présentation des méthodes de traitement des valeurs manquantes ([lien](https://docs.google.com/presentation/d/1KJdMpnrBnaPCk8vNrkLegYauPDraRmb33f32qTvDW9Q/edit#slide=id.g31f5a4735e9_0_178)) : 30 min

***Après-midi :***

1. TP Analyse des données : 2h
   1. Lecture des données
   2. Statistiques
   3. Analyse univariée
   4. Analyse bivariée
   5. Tests de corrélation
   6. Visualisation

***Aide aux étudiants :***

* Dataset “diabetes.csv” :
  + Détection et traitement des données manquantes (les zéros) :
    - Les variables sont les suivantes (Glucose', 'BloodPressure', 'SkinThickness', 'Insulin', 'BMI').
    - Exception : Les variables pregnancies ainsi que OutCome peuvent contenir des 0, celle ci représente le nombre de grossesses ou bien si les personnes sont diabétique, ce qui est donc pertinent.
* Dataset “house.csv” : préparation des données
  + Convertir les données au bon format
  + df['price'] = df['price'].str.replace(",",".")
  + df['price'] = df['price'].astype(float)
  + df['date'] = df['date'].str[-4:]
  + df['date'] = df['date'].astype(int)
  + Traitement des valeurs aberrantes
* Dataset “wine.csv” : jeux de données déséquilibrée
  + Solution : créer des classes de qualité

Support pour des techniques de modélisation : <https://docs.google.com/presentation/d/1ivSoq1UOwHaP_so5ATYtc9abTuedoNwv/edit#slide=id.g2161c2ca195_0_149>