

Évaluation

Code du cours : A51-BB

Date : 30 juin 2025 à 18h00

Durée : 1h 15 (de 18h00 à 19h15)

Notation : sur 30

Pondération : 30%

Vous devez remettre via Léa un seul document (en format pdf) contenant aussi bien toutes les lignes de code que toutes les réponses aux questions. Il est très important d'indiquer clairement le numéro de la question à laquelle vous répondez.

Note : Avant la remise, vérifiez la qualité visuelle et l'entièreté du document. Respectez l'heure limite de remise, les documents remis en retard, ne seront pas acceptés.

Partie I : Questions de compréhension (10 points = 5×2)

Répondez à chacune de ces questions de façon concise (en 2 phrases).

- Q1. Énumérez les quatre principales techniques d'apprentissage machine.
- Q2. Face à un problème d'apprentissage machine, comment pouvons-nous anticiper s'il s'agira d'utiliser une technique d'apprentissage supervisée ou une technique d'apprentissage non-supervisée ?
- Q3. Quelles sont les étapes typiques dans un projet d'apprentissage machine ?
- Q4. Qu'est qu'une modalité ?
- Q5. Le coefficient d'asymétrie d'une distribution est nul. Que peut-on dire de la distribution ?

Partie II : Description des données (20 points = 10× 2 points)

Dans cet exercice, nous utiliserons l'ensemble de données `RecolteVins.xlsx` disponible sur Léa pour effectuer une analyse descriptive. Nous analyserons des variables quantitatives et qualitatives pour obtenir des informations utiles.

1. Charger les données
2. Déterminez le nombre d'individus et de variables et identifiez le type des variables
3. Calculez la moyenne et l'écart-type de la variable Intensity
4. Déterminez les modalités de la variable Region
5. Calculez l'effectif de chaque modalité de Region
6. Visualiser le diagramme circulaire de la variable Region
7. Déterminez les caractéristique de tendance centrale de la variable Spice
8. Représentez l'histogramme de la variable Spice
9. Calculez le coefficient d'assymétrie de la variable Spice
10. Interprétez les résultats