



- **Nom :**
 - **Prénom :**
- Groupe :**
Durée : 1 h 15

Copiez le dataset `Iris.csv` ensuite répondez aux questions ci-dessous en créant un Jupyter Notebook où vous effectuerez les analyses demandées. Écrivez également vos réponses au dos de cette feuille pour la remise finale.

*Veillez nommer votre fichier Jupyter Notebook en respectant la convention suivante : **nom_prenom.ipynb***

Q1. Aperçu des données

- Importez le dataset Iris dans Python.
- Affichez les 10 premières et dernières lignes du dataset.
- Afficher la dimension et les types de données des colonnes ?

Q2. Statistiques descriptives

- Calculez les statistiques descriptives pour les colonnes numériques : moyenne, médiane, écart-type, minimum, maximum.
- Créez une table montrant la répartition des échantillons par espèce de fleur.

Q3. Vérification des données manquantes et nettoyage

- Vérifiez si le dataset contient des valeurs manquantes.
- Proposez une stratégie pour gérer les données manquantes.

Q4. Visualisation exploratoire

1. Créez un histogramme pour visualiser la répartition de la longueur des sépales. La distribution est-elle symétrique ?
2. Réalisez un scatter plot (nuage de points) pour observer la relation entre la longueur et la largeur des pétales. Que pouvez-vous conclure?
3. Existe-t-il une corrélation notable entre certaines variables numériques ? Calculez la matrice de corrélation entre les variables numériques et interprétez les résultats.
4. Réalisez un diagramme en barres pour représenter le nombre de fleurs par espèce.
5. Que peut-on conclure sur la distribution des espèces dans ce dataset ?
6. Convertissez les colonnes catégorielles qui contiennent des informations répétitives en format approprié.