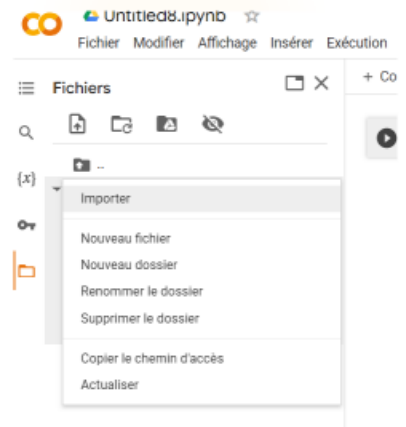


Développer des programmes avec Spark

Créer un dossier sous c: ou d: appelé ExercicesSpark qui contiendra les datasets utilisées

Exercice 1. Ecrire un programme qui compte la somme obtenu pour chaque nombre de stars (de 1 à 5) utilisés pour évaluer un film dans le fichier u.data.

Importer le fichier u.data sous le dossier de données sous google collab



Exercice 2. écrire un programme qui cherche le minimum de temperature pour une location, à partir de la base 1800.csv
Temperatures seront convertit en fahrenheit en utilisant une fonction.

Exercice 3. écrire un programme word-count.py qui calcule la fréquence de chaque mot dans le livre Book.txt.

Exercice 4: écrire un autre programme qui permet de modifier le programme précédent pour faire le tri du résultat suivant les fréquences de mots.

Exercice 5: écrire un programme qui permet de calculer la somme de montants dépensés par chaque consommateur en utilisant le fichier customer-orders.csv (1ere information: customerId, troisième information: montant dépensé)

Exercice 6: trier le résultat de l'exemple précédent suivant l'identification du consommateur

Exercice 7: trier le résultat de l'exemple précédent suivant le montant dépensé.

Exercice 8: écrire un programme qui permet de calculer le nombre moyen d'amis (friends) par age en l'appliquant sur le fichier friends.csv