Récapitulatif

1. Scraping des 5 premières pages de citations :

Mise en place une boucle pour parcourir les 5 premières pages du site de citations.
Filtration des citations par raports aux top premiers tags "love", "inspirational", "life", et "humor"

2. Connexion au site et récupération du token :

On accède à la page de connexion pour récupérer le token renvoyé par la requête Post
Envoie des informations de connexion avec le token pour authentifier la session.
Après la connexion, extraction du cookie de session contenant le token de connexion.

3. Ajout des citations avec le tag "books" :

Grâce à la connexion, on peut scrap les 2 premières pages de citations avec le tag "books"
Ensuite stockage de ces citations dans les citations

Filtrage des doublons :

Utilisation de panda (creation d'un dataframe) à partir de la liste des citations pour supprimer les doublons basés sur le texte des citations.

4. Écriture des résultats dans un fichier csv :

Écriture des citations filtrées et sans doublons dans un fichier csv au nom de results.csv
Ajout grâce au script du token de connexion à la fin du fichier csv

Ajout de l'étape pour écrire le token de connexion dans le fichier Csv.

Modifs :
Changements apportés en cours de route:

1. Initialisation et Scraping des citations:

J'ai modifié le script pour garder uniquement le texte des citations et les tags, en excluant l'auteur

2. Authentification et récupération du token:

J'ai ajouté une étape pour récupérer le token avant d'envoyer les informations de connexion
Après la connexion, j'ai extrait le cookie de session pour obtenir le token de connexion

3. Utilisation du cookie de session:

J'ai utilisé ce cookie de session pour authentifier les requêtes supplémentaires nécessaires pour récupérer les citations avec le tag "books".

4. Écriture dans le fichier CSV:

J'ai ajouté une étape pour écrire le token de connexion dans le fichier csv à la fin du fichier