Objet : Programme d’extraction des prix

Bonjour Sam !

J’ai terminé le programme d’extraction des prix. Il couvre l’intégralité des étapes ETL via les fonctions suivantes :

* Extraction des données demandées :

Le programme scanne l’intégralité des ouvrages répertoriés sur le site Books To Scrape, cible les données demandées et les extrait en 3 étapes. Premièrement, l’extraction de l’ensemble des catégories est réalisé. En second temps, la liste complète des ouvrages contenus dans chaque catégorie est extraite. Enfin, les données demandées pour chaque ouvrage sont extraites.

* Transformation des données :

Le programme transforme les données brutes afin d’extraire l’URL de chaque ouvrage, la quantité disponible, le nombre de reviews ainsi que pour nommer les fichiers images.

* Chargement des données :

Les données extraites sont chargées dans un fichier .csv et dans un dossier /Images créé dans le répertoire du programme.

Pour réaliser l’ensemble de ces étapes, il suffit d’exécuter le fichier main.py.