

Extraction des données du site *Books to Scrape*.

Cher Sam,

Je vous écris pour expliquer le code que j'ai écrit pour notre projet.

Le code extrait des informations sur les livres d'un site Web appelé "books.toscrape.com" et les transforme en un format structuré pouvant être enregistré dans un fichier CSV. Le code est écrit en langage de programmation Python et utilise trois fonctions principales : extraire(extract), transformer(transform) et charger(load).

La fonction "**extract**" extrait le contenu d'une page Web à l'aide de la bibliothèque Python "requests" et analyse le contenu HTML à l'aide de la bibliothèque "**BeautifulSoup**".

La fonction "**transform**" transforme les données extraites en un **Pandas DataFrame**(Plugin) en sélectionnant des éléments spécifiques du contenu HTML et en les enregistrant dans un dictionnaire.

La fonction "**load**" enregistre les données transformées dans un fichier CSV.

Le code extrait d'abord les informations d'un seul livre et les enregistre dans un fichier CSV. Ensuite, il extrait les informations de tous les livres de la catégorie Mystère et les enregistre dans un autre fichier CSV. Enfin, il extrait les informations de tous les livres de toutes les catégories et les enregistre dans un troisième fichier CSV.

J'espère que cette explication vous aidera à mieux comprendre le code.

Cordialement, Huck Cameron