

Bonjour Sam,

Le version bêta du système permettant l'extraction des prix de Books to Scrape est aboutie. Comme convenu, vous trouverez en pièce jointe les éléments sollicités dans notre dernier échange (lien vers le repository et zip des fichiers obtenus.).

Concernant la pipeline E.T.L, voici quelques points qui j'espère éclaireront votre argumentaire auprès du responsable.

### **La pipeline E.T.L :**

C'est un ensemble de processus qui vise à l'extraction, la transformation et au chargement de données.

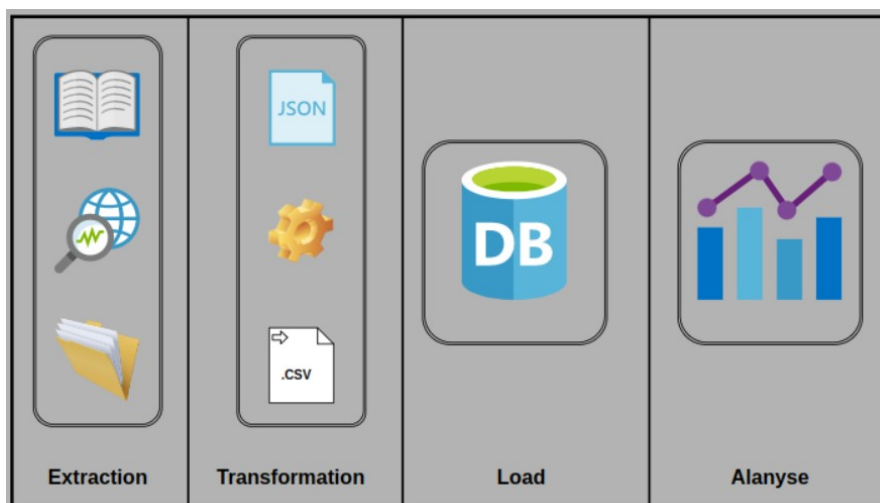
Concrètement, dans le cas nous concernant, il s'est agi, d'extraire l'intégralité des données du catalogue du revendeur Books to Scrape sur leur site en ligne. Une fois ces données brutes obtenues, nous les avons transformés dans un format nous permettant d'exploiter facilement les éléments obtenus. À la suite de quoi il ne nous restait plus qu'à télécharger ces données , ici au format « .csv ».

### **Le but de la pipeline E.T.L :**

Un procédé automatisé, exécuté quotidiennement, couplé à un système d'analyse des résultats obtenus, nous permettrait :

- La constitution d'une base de données au plus près de nos besoins.
- La surveillance avec une grande efficience l'activité de nos concurrents.
- Le dégagement de tendances de marchés, aussi bien du côté des consommateurs que des professionnels.

Nous pourrions ainsi, grâce à ce type de procédés, augmenter notre réactivité et nos capacités d'anticipation.



En espérant que ces éléments vous conviennent.  
Cordialement.