

Projet Web Scraping ~ QUOTES

Site web à scraper : <http://quotes.toscrape.com>

Guide Markdown: <https://www.markdownguide.org/basic-syntax>

Contrainte : Utilisez le notebook Jupyter ou Google Colab pour ce projet rapport y compris. Avec le markdown vous pouvez écrire du texte et faire la mise en page. Le tableau de correspondance aussi peut être fait en markdown TOUT PROJET QUI N'EST PAS UN NOTEBOOK SERA REFUSER !

Tâches :

1. Définir le web scraping et les cas d'usage
2. Quels sont les étapes du web scraping ?
3. Faire une cartographie du site à analyser comme dans l'exemple suivant



4.

- a. Identifier les données à récupérer
- b. Identifier les balises html
- c. Faire une table de correspondance avec les éléments précédents

Données	Html
Titre	Chemin + balise correspondante

5. Ecrire un script python pour récupérer l'ensemble des données de la page
 - a. D'abord une page de description détaillée de votre solution
 - b. Le script python qui implémente votre solution
6. Que peut-on faire avec les données récupérées en termes d'analyse ?
Proposez des cas d'études

Dans votre rapport :

1. La cartographie du site
2. Les étapes et résultats de vos recherches (question 4)
3. Votre ou vos solutions pour récupérer les données
4. Les études de cas possible