Aufgabe 1: Bookstore-Scraper

Baut als Gruppe einen Webscraper für <https://books.toscrape.com/>

Drei Wege sind denkbar:

A)

* Baut einen Scraper, der alle Seiten mit Bücher-Einträgen durchgeht und sich die folgenden Informationen holt: Buchtitel, Bewertung (Sterne) und Preis.
* Speichert die Informationen in eine geeignete Form, etwa in eine CSV (oder eine Datenbank).
* Ladet die CSV in einen Dataframe (oder betreibt eine SQL-Analyse in DBeaver oder Pgadmin) und beantwortet Fragen wie: Was kostet ein durchschnittliches Buch auf der Webseite? Wie sind die Bücher im Schnitt bewertet? Wie viele Bewertungen mit fünf Sternen gibt es? Wie viele Bewertungen mit einem Stern?

B)

* Baut einen Scraper, der Bücher nach Kategorien holt (etwa Classics, Sequential Art etc). Holt Euch die Informationen wie in A) zu den Kategorien.
* Speichert die Informationen in eine geeignete Form so, dass die Kategorienunterschiede erhalten bleiben.
* Betreibt Datenanalyse mit den Daten und beschäftigt euch mit Fragen wie in A, allerdings mit dem Zusatz, dass ihr nun Kategorien miteinander vergleichen könnt. Damit eröffnen sich so Fragen wie: Wurden Kinderbücher im Schnitt besser bewertet als Bücher zu Religion? Kosten Science-Fiction-Bücher im Schnitt weniger als Klassiker? Usw.

C)

* Baut einen Scraper, der alle Seiten mit Bücher-Einträgen durchgeht und den Link zu jedem einzelnen Buch aufruft, um sich von dort (!) alle nötigen Infos **inklusive** der Beschreibungstexte zu den Inhalten der Bücher zu holen.
* Speichert die gesamten Informationen in txt-Dateien, wobei die txt-Dateien am besten nach den Titeln der Bücher benannt werden sollten. Also pro Buch eine txt-Datei.

Erarbeitet idealerweise zwei von diesen drei Wegen und macht eine gemeinsame Präsentation fertig. Begebt Euch aber auch gern zusätzlich auf Abwege und habt Spaß dabei! 😊