

Assignment-V

(Based on Data Scraping)

Solve both questions using BeautifulSoup and Selenium

Q1: For each of the best seller book mentioned on the first two pages of the following link

<https://www.amazon.in/gp/bestsellers/books/>

Scrape the following information:

- i. Publisher
- ii. Language
- iii. Paperback
- iv. ISBN-10
- v. ISBN-13
- vi. Weight
- vii. Dimensions
- viii. First Three Customer Reviews

Q2: Consider the following Link:

https://www.imdb.com/search/title/?release_date=2020&sort=num_votes,desc&page=1

The page contains information about movies released in the year 2020 sorted according to number of votes. Scrape the following information for the first 1000 movies.

- (i) Movie name
- (ii) IMDB Rating
- (iii) Description
- (iv) Votes
- (v) Director(s)
- (vi) Stars
- (vii) Metascore
- (viii) Movie Genre
- (v) Total Duration