

Assignment

Scraping

Note:

- Only Python to be used – Any Library – Speed is the key (Selenium/Scrappy/API etc etc)
- Relevant Extra Columns or analysis will be appreciated
- Assignment details are not to be shared with anyone or on any platform.

- 1) The following link needs to be scraped

<https://www.noon.com/uae-en/sports-and-outdoors/exercise-and-fitness/yoga-16328/>

- 1) Output should be as per the example file attached. 200 products should be extracted at a minimum.
- 2) Can you incorporate proxies in your code
- 3) Date extracted should be in CSV
- 4) Also share the python code, no GitHub links or anything
- 5) Speed is key use appropriate Python libraries.
- 6) After extraction, analyse the data with your preferred python libraries and show your understanding of the data and your skill set. (Graphs & fruitful analysis will be appreciated)
 - a. Most expensive product
 - b. Cheapest Product
 - c. Number of Products from Each brand
 - d. Number of Products by Each seller