

Summary:

1. Amazon's Website generate random UUID as id of product's div tag
2. Sometimes the product name & price're under 6 elements or under 8 elements
3. These elements're either div or span
4. HTML Elements doesn't supports number to be first char in id, i.e., #0a is invalid but #a0 is valid
5. Since some Google's UUID starts with number, we've to add escape character e.g.
6. Instead of #0a we've to use #\30 a & since '\' is escape char, we've to use #\\30 a
7. Using Beautiful Soup to process the HTML which is stored in variable, since fetching data from web-page slows down the process
8. Playwright is used to interact with Web-Page (Clicking Buttons & Entering Text)
9. Beautiful Soup is used to Parse HTML

Output:

```
File Edit Selection View Go Run ... PROBLEMS DEBUG CONSOLE OUTPUT TERMINAL PORTS Data Mining & Warehousing Enter Electronic Device(e.g., Laptop, Mobile, Smartwatch, etc.): Laptop Starting Scrapping ... Main WebPage Visited Products Page Visited Page 1 Total Products found on the Webpage: 22 Products Successfully Retrieved: 22 Page 2 Total Products found on the Webpage: 22 Products Successfully Retrieved: 22 Page 3 Total Products found on the Webpage: 22 Products Successfully Retrieved: 22 Page 4 Total Products found on the Webpage: 22 Products Successfully Retrieved: 19 Page 5 Total Products found on the Webpage: 22 Products Successfully Retrieved: 20 Page 6 Total Products found on the Webpage: 22 Products Successfully Retrieved: 21 Page 7 Total Products found on the Webpage: 22 Products Successfully Retrieved: 21 Page 8 Total Products found on the Webpage: 22 Products Successfully Retrieved: 21 Ln 26, Col 1 Spaces: 4 UTF-8 CRLF () Python_3.12.2 Background Prettier Launchpad < OUTLINE > TIMELINE 0 △ 0 0 0 File Edit Selection View Go Run ... PROBLEMS DEBUG CONSOLE OUTPUT TERMINAL PORTS Data Mining & Warehousing powershell - Practical-1 + Main WebPage Visited Total Products found on the Webpage: 17 Products Successfully Retrieved: 17 Page 14 Total Products found on the Webpage: 17 Products Successfully Retrieved: 14 Page 15 Total Products found on the Webpage: 17 Products Successfully Retrieved: 15 Page 16 Total Products found on the Webpage: 17 Products Successfully Retrieved: 16 Page 17 Total Products found on the Webpage: 16 Products Successfully Retrieved: 14 Page 18 Total Products found on the Webpage: 19 Products Successfully Retrieved: 17 Page 19 Total Products found on the Webpage: 16 Products Successfully Retrieved: 16 Page 20 Total Products found on the Webpage: 2 Products Successfully Retrieved: 2 No Additional Pages Total Complete Product Count: 357 Time Taken: 296.63 Ln 26, Col 1 Spaces: 4 UTF-8 CRLF () Python_3.12.2 Background Prettier Launchpad < OUTLINE > TIMELINE 0 △ 0 0 0
```

Data Screenshots:

