

Project Progress Report 4

Objective:

Gather correct information about the product and compare its price across different websites/retailers so that the customer can surf through good deals easily and effectively.

What I did last week:

I worked on displaying the scraped data on my web page. For that, I need to setup the backend environment and integrate it with frontend. So worked on both of them, like setting up environment and integrating it with frontend. Integrated filter component and started working on scraping of another website.

What I did this week:

The main components I worked on this week is:

- Tried to resolve the span tag issue
- Scraping the data from website using puppeteer
- Worked on sorting component

Last week I worked on integrating the backend of my project with frontend and displaying the scraped data from amazon website on my web page. This time I worked on the span tag issue on my webpage while showing the results. Also worked on the scraping data from another website. The data is scraped using puppeteer web scraper. (Still working on the display part of my web page). Also, while integrating the code, I am facing some issue with the end-to-end service after integrating the new code I developed. So, still working on that. So, on GitHub I'm submitting the frontend part of Amazon website and the Pipes (which I used as a solution for the span tag issue)

What Next:

To work the display part of my website for the data which is scraped from another website and also to start working on the other modules like sorting, to increase the functionality.

References:

[1]<https://stackoverflow.com/questions/72457467/how-to-remove-span-tag-and-class-name-after-scraping-whereas-i-want-to-scrape-o>

[2]<https://www.appsloveworld.com/coding/python3x/136/removing-span-tags-in-python>

[3]<https://www.section.io/engineering-education/build-a-web-scraping-using-cheerio/>