### **Task Description**

#### **Objective 🎯**

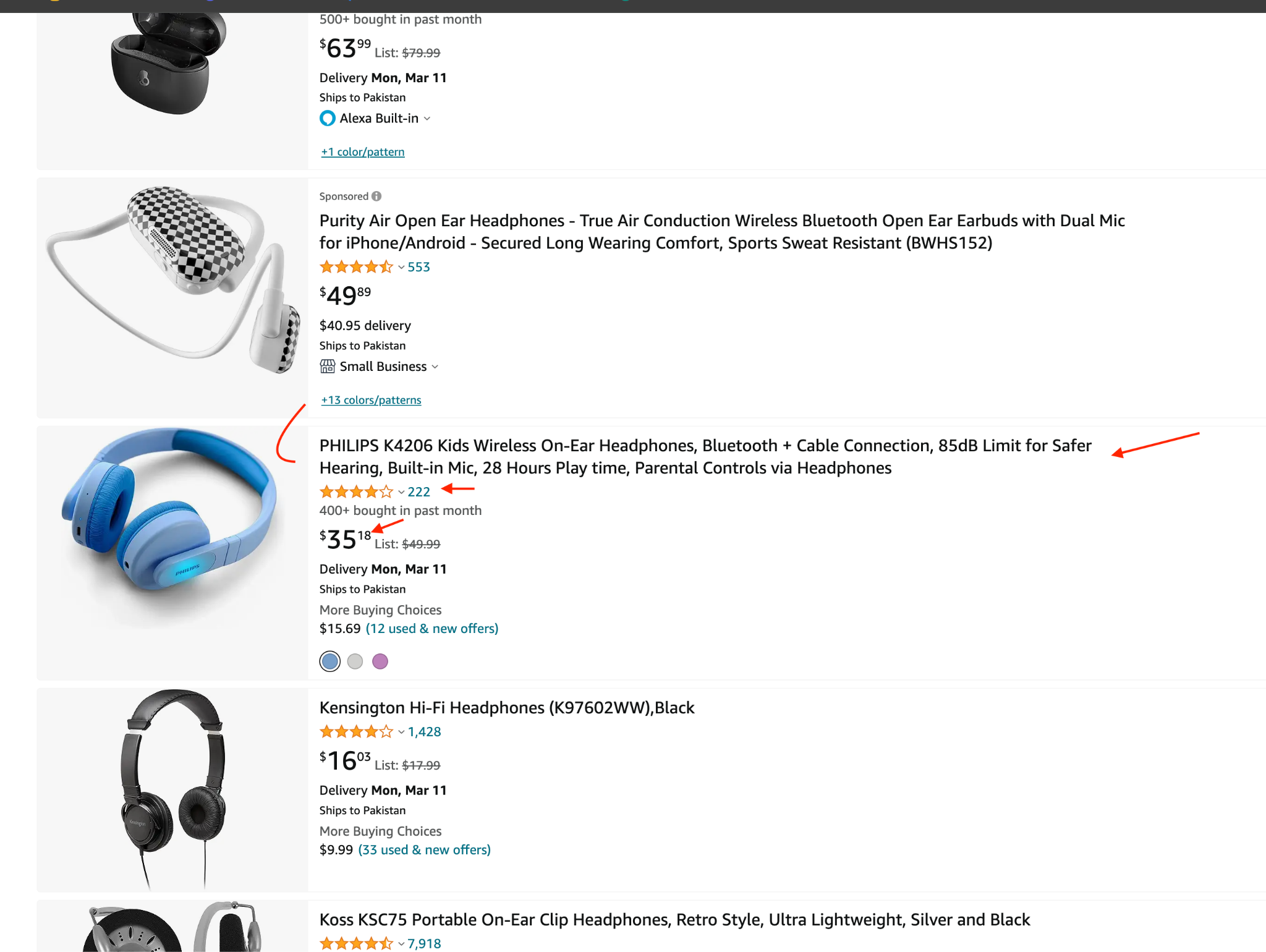
<https://chatgpt.com/c/67699e84-6850-8002-b3c2-6b05cf42d804>

https://github.com/praneethravuri/Amazon-Product-Information-Scraper/blob/main/amazon\_scraper.py

Develop a Python script to scrape product details from Amazon for a set of search queries. The script should be designed with modularity, efficiency, and readability in mind.

#### **Requirements 📃**

1. **Read Search Queries: 📕**
   * Input file: user\_queries.json
   * Structure: ["headphones", "smartphones", ...] (10 queries in total)
2. **Scraping: 🎉**
   * For each query, scrape product details from the first 20 pages of Amazon search results.
   * URL format:
     + Example: <https://www.amazon.com/s?k=headphones>



* + Extracted details for each product:
    - Title
    - Total reviews
    - Price
    - Image URL

**Scrape other fields - It will be a plus**

* + Consider additional fields like creation and update timestamps for each scraped record, and a field indicating the scrape date.

1. **Output:**
   * Save the scraped data into JSON files named after the search query (e.g., headphones.json).
   * Each output file should contain an array of product details.
2. **Code Organization:**
   * Apply DRY and Separation of Concerns principles.
   * Suggested structure:
     + A module for reading input queries.
     + A scraping module that defines a class or function for fetching and parsing web pages.
     + A data model representing the product.
     + A module for saving the scraped data to JSON files.
3. **Error Handling:**
   * Implement error handling for network requests, parsing errors, and file operations.
4. Design a UI
   * **Design a Frontend React App that shows the scraped data in the tabular form.**
     + Deploy on Vercel.
     + Share the link in the email
     + Be creative as much as you can.

#### **Additional Considerations**

* Keep it simple.
* Don’t use AI or Chat GPT 👏 Use Stackoverflow 👍

​​**Deliverables**

Candidates should submit:

1. The complete source code, adhering to the specified requirements and code organization principles.
2. A brief documentation explaining their design decisions, how to run the script, and any dependencies.
3. Sample output files for each of the search queries provided in the user\_queries.json file.

**Other Questions**

What do you know about Microsoft Azure and AWS? Have you ever worked with cloud technologies? If so please describe.

What is FIFO in aws queue? What is the purpose of Queue?

Have you ever setup the CI/CD pipeline? Do you know how to deploy the Lambda functions/Create them?

**Email your submission to:**

Set Email Subject to  
  
Python Developer Lead | Your Full Name

**In the email - Don’t forget to answer the following questions ☘️**

Why do you want to join Codeaza Technologies? ❓

Why should we hire you? What makes you unique from other candidates? 🧑‍🍼

Tell 3 bad habits and 3 good habits about yourself? 🦡

**muhammadasim@codeaza.org**

**Note: You will be evaluated based on your technical skills, with attention to detail, how you understand and approach the problem.**

**Technical skills are only 20% of what you will be doing, the rest are soft skills.**

**Good Luck - We are expecting you to be part of an exciting fast growing startup. 🚀**