**Report DevOps Assignment**

**Name- Ayush Bhandari**

**Program- B.Tech CSE (DevOps) || [Sem-8]**

**Contact- 8439167656**

**Collage- UPES, Dehradun, Uttarakhand**

**GitHub Link for source code-** [**Link**](https://github.com/Ayushgittt/DevOps_Assignment)

This project demonstrates a multi-stage containerized application that performs dynamic web scraping using Node.js + Puppeteer and serves the scraped data through a Flask API.

**Task 1: Packages & Technologies Required**

**Packages Used**

**Node.js (Scraper)**

* puppeteer: Automates Chromium browser for scraping

**Python (Server)**

* Flask: Lightweight web framework to serve JSON data

**System Dependencies (for Puppeteer & Chromium)**

* chromium, libx11-xcb1, fonts-liberation, xdg-utils, etc.

**Technologies**

* **Node.js v18-slim**: Used in the first stage for scraping
* **Python 3.10-slim**: Used in the final stage for serving
* **Docker Multi-stage Build**: Optimizes image size
* **Chromium**: Headless browser for Puppeteer
* **Flask**: Python server to expose scraped content

**Task 2: How to Run Locally (Without Docker)**  
**Step 1: Install Node.js Dependencies**

npm install

**Step 2: Scrape the Website**

npm start https://webscraper.io/test-sites/e-commerce/allinone/computers/laptops

**Step 3: Create Python Virtual Environment & Install Flask**

python3 -m venv Project

source Project/bin/activate

pip install -r requirements.txt

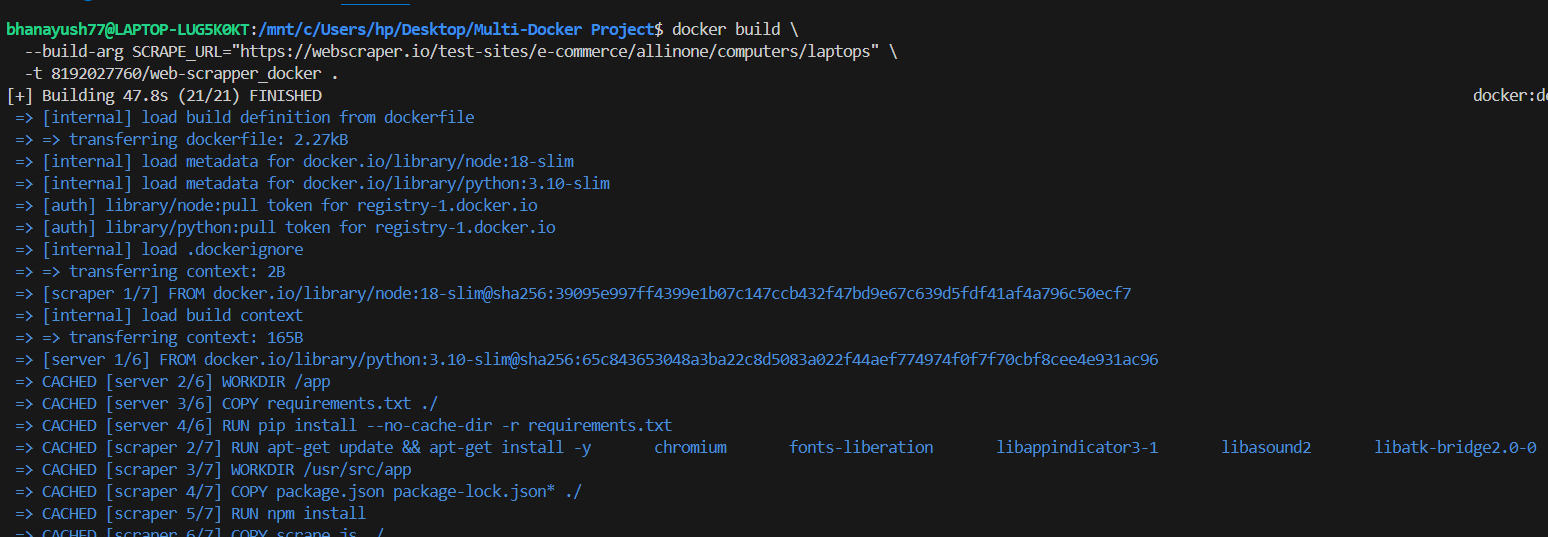
**Step 4: Run the Flask Server**

Python3 server.py

Visit <http://localhost:5000> to view the scraped data.

**Task 3: How to Build the Docker Image (** [**Link**](https://github.com/Ayushgittt/DevOps_Assignment) **)**

Source code for Dockerfile is in my GitHub repository.

****

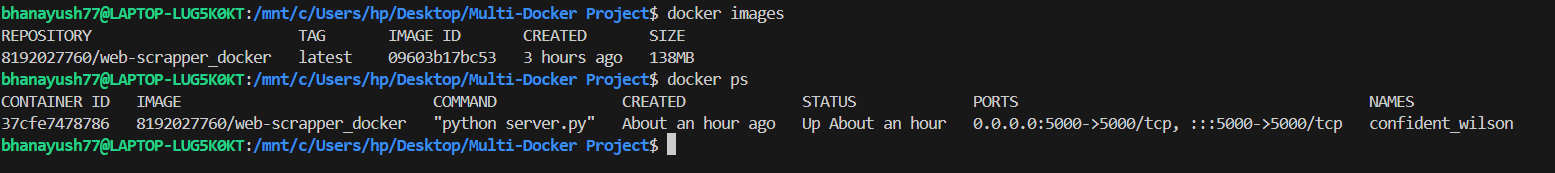
**Task 4: How to Run the Container**

****

Open to browser and go to:

[**http://localhost:5000**](http://localhost:5000) ****

This show the size of image which is small:-

****