

Author: Jered Kalombo

Task: YouTube Skills Video Scraping

Script: PhaseTwo/Scraping/youtube_skills_scraper.ipynb

Data: PhaseTwo/Data/skills_resources.csv

Methodology:

- Used Selenium with Chrome WebDriver to scrape YouTube search results.
- Targeted skills needed from the new_england_indeed_job_skills list with the highest frequency: Excel, Data analysis, Microsoft Office, CPR, Python, crm, and budgeting.
- For each skill, performed a search for "skill + tutorial" to focus on instructional videos.
- Collected the following video attributes:
 - Video Title
 - Video URL
 - Skill
 - Source (YouTube)
 - Stored the results in a Pandas DataFrame and exported to CSV (skills_resources.csv).

Notes:

- This scraper relies on a real browser session (Selenium) to render dynamic JavaScript content, which requests or BeautifulSoup alone cannot reliably access.
- Pauses (time.sleep) were added between searches to avoid rate-limiting and excessive requests to YouTube.
- Only got results for 4 of the skills listed above.