

SI 507 Final Project Proposal:
Munjot Singh
Munjotks

Data Source:

Crawling [and scraping] multiple pages in a site you haven't used before

- I plan to crawl and scrape multiple Amazon pages/products based on the product the user chooses to search within the user interface. The user would try to search for a product, get the list of results by number of reviews, cost or stars. The user could then choose to look up the description of product of their choosing or exit the program. I will use plotly to plot and display cost vs. stars and number of reviews vs. price for the specific results.

Challenge Score: 8

- Based on the number of pages and information I will be crawling and scraping, I access the challenge score to be an 8. I will also be incorporating a database and using graphs in my demo.

Description:

- I am able to crawl the pages really well and I will be able to create a cache dict for the specific urls the user has already searched for.

Presentation:

- I will be creating a user interface and showcase the data for the user using plotly (cost vs .stars vs. reviews vs. price)