

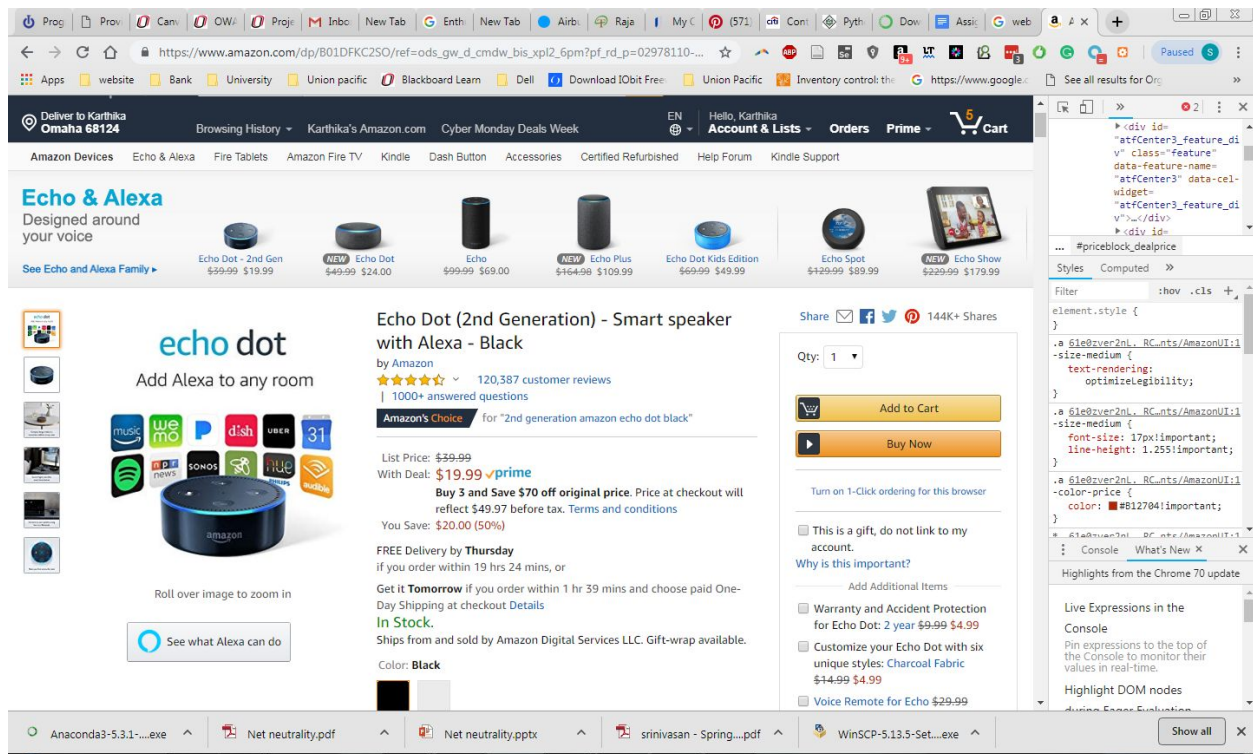
Assignment 2 (Due by 12/13/2018 11:59 PM)

Maximum Points-50

Web Scraping Using Python

This assignment is intended to get you familiar with web scraping. Web Scraping is used to get the HTML page content, for example if I ask you to compare the best price for Google pixel phone during Black Friday, imagine how will you check for best price. Ideal solution is you visit the top retailer websites such as Amazon, Newegg, and many other websites. Rather than doing it manually we can use scraping script to automate this process. The idea is that using Python library we can get the HTML for the page we are interested. Now if we know the correct tags, i.e. the price of Google Pixel we can just fetch it.

For this assignment, the scrapping code using Python will be provided as a boilerplate after the class on 11/28 (lab 10). You are expected to modify the code such that you choose a particular website and you want to extract a particular attribute of the item. For example if I choose (https://www.amazon.com/dp/B01DFKC2SO/ref=ods_gw_d_cmdw_bis_xpl2_6pm?pf_rd_p=02978110-1fff-4425-9cb1-85dee8951250&pf_rd_r=ED18ASZBDPNTPJ6Y8PZ) and my script should return just the price of the product, i.e. 19.99.



You can be creative by choosing any other attributes, or choosing multiple attributes, such as price, "you save" etc. If you are able to scrape website and display any particular attribute, you will receive full credit for this assignment.

Assignment Submission Instructions (Very Important)

1. Please create a new folder inside your CIST1300-001 Gitlab repository, next put your script file “which should end with .py” inside the repository.
2. Once you push your files to gitlab, Next thing is you need to submit a word document or PDF which should explain which website you are scraping and what value are you expecting as output. Finally the word document or PDF should have a screenshot of your script running and giving the expected value.
3. Complete reflection for Assignment 2 on slack, please remember to use the appropriate slack channel for reflection.

I won't be able to grade, if you miss step 1 or step 2, please complete both steps.

Grading Rubric-

Code - 25 pts

Word Document (Explaining all the details & Screenshot)- 20 pts

Reflection - 5 pts