Name: Chathankulam Gokul Satheesan

Class: SE Comps A

Project: Web Scraping of MOOC Websites and displaying on a responsive website.

Description: I created a responsive website named Coursway for comparing available courses on 3 MOOC websites – Coursera,Vskills and FutureLearn based on their pricing. Front end was developed using html and css. The scraped data is saved into the database application i.e. DB Browser, using sqlite commands. This data is then fetched and passed into Coursway. The backend part of Coursway was executed using Flask, jinja templates and sqlite3. Web Scraping was done using dedicated python packages like BeautifulSoup4, requests, lxml and regular expressions. Requests package helped in sending requests to desired websites to scrape data from their webpage. BeautifulSoup4 helped in parsing the extracted data in html form into usable form making scraping efficient and easy. Regex was used to find specific data for matching the prices from the database to the website. A delay time was introduced as a precaution for getting blocked from MOOC websites.

Packages used: BeautifulSoup4, Requests, sqllite3, regex.

Thank You.