Lab #3

Instructions

With lab #3, your group will write a python program that extends your work from lab #2 to extract data from all phones listed on the supplied webpage. This will be the culmination of your studies about Request, BeautifulSoup, HTML, and the reduction of redundant code.

Submission Requirements

- Python .py file with no markdown language. Only Python Code and comments.
- Group Assignment: Only 1 assignment needs to be submitted for each group.
- You should seek to reduce redundancy in your code.

Task

You are to scrape the data related to the phone found on the webpage listed below.

- Webpage: http://drd.ba.ttu.edu/isqs3358/labs/lab3/
- Your code ***MUST*** implement throttling
  - Please use a high and low variable that can be changed in the settings section.
- You will extract the following value into a CSV file in the order specified below:
  - Product_Id
  - Model
  - Product_Size
  - Storage
  - OS
  - # of Back Camera Features
  - # of Front Camera Features
  - Battery Capacity
- Your CSV should mimic the sample output data
- The order of the columns must be the same.
- Use the example for column name specifications.

Notes and Hints

- Your code from Lab #2 can be very useful as a starting point.
- Your team will need to first scrape the supplied webpage and enumerate the phones listed, to do a request for the child page.
  - Note, this request for the child page will be in the loop.
  - It will also help you variablize the root of your URLs
- HOWEVER, Your team CANNOT call each phone page with hard-coded urls, you must find the child page to call from the parent page listed above.
  - i.e. this simulates code that would work if we added 10, 20, 2000 more phones to the base page.
- Depending on your OS, you may see “blank” lines in your CSV, this is not an issue.

Example

- Included in lab3_example.zip found with this lab.