CS 288 2019S Section 004
Homework 6 - Milestone 1

Due: At the beginning of class on Thursday April 18, 2019.

Write a UNIX Shell script that repeats the following steps every minute for an hour or so. Make sure you do this on a weekday between 9:30am and 4pm EST while the NASDAQ Stock Exchange (NASDAQ) is open:

1. Set a shell variable to the current date & time using the format: `yyyy-mm-dd-hh-mm-ss.html`
2. Use `curl` or `wget` to download the 100 Nasdaq Most Active Stocks Web page and rename it to the string you've obtained in the previous step.
3. Use TagSoup, or any similar tool, to generate a `.xhtml` that corresponds to the downloaded `.html` file. TagSoup is written in Java. As such, you have to have a Java JRE installed on your system in order to run it. The command to run TagSoup is:
   ```
           java -jar tagsoup-1.2.1.jar --files yyyy-mm-dd-hh-mm-ss.html
   ```
   Use `curl` or `wget` to download the `tagsoup-1.2.1.jar` file if it isn't present in the current directory.
4. Run a Python script that uses the `xml.dom.minidom` module to traverse the `.xhtml` document and extract the relevant values.
5. Store the extracted values in a `.csv` file as comma-separated values based on the Sample Input/Output given below.

Your deliverables are the shell and Python scripts. Before submitting your solutions via Moodle, zip the scripts and name the archive, if your name is Harry Houdini, for example, HW6a_HarryHoudini.zip. You may use gzip if that's more convenient.

Sample Input/Output:

The following row, for example:

1	Advanced Micro Devices (AMD)	73,492,234	$19.54	0.43	2.25

should produce the following output where the first line is reserved as a header containing a list of field names:

```
exchange,symbol,company,volume,price,change
Nasdaq,AMD,Advanced Micro Devices,73492234,19.54,0.43
```

Note that three of the original fields required trivial transformations, and the first field (*rank*) and the last field (*% change*) were discarded because they can be computed easily.