­­­­

CS 495: Introduction to Web Science

Fall 2013

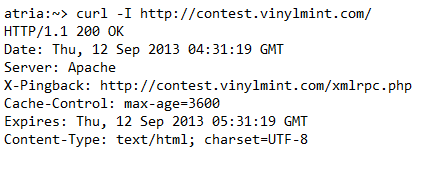
Assignment 1

Onapha Rattanachottiteepakorn

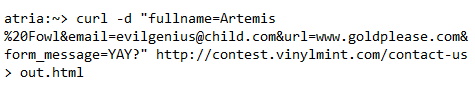
Submission Date: September 12, 2013

**Question 1:** Demonstrate that you know how to use "curl" well enough to correctly POST data to a form. Show that the HTML response that is returned is "correct" (e.g., save it to a file and then view that file in a browser and take a screen shot).

To illustrate the curl function for this assignment, I used the website called Vinylmint, accessible via <http://contest.vinylmint.com>. The server responded with status code of 200. The information on the website displayed as followed:



In order to show that curl’s ability to post data to a form, I choose the form on the website’s contact us page with and entered the following command line:



The result of this line was recorded in the out.html file.

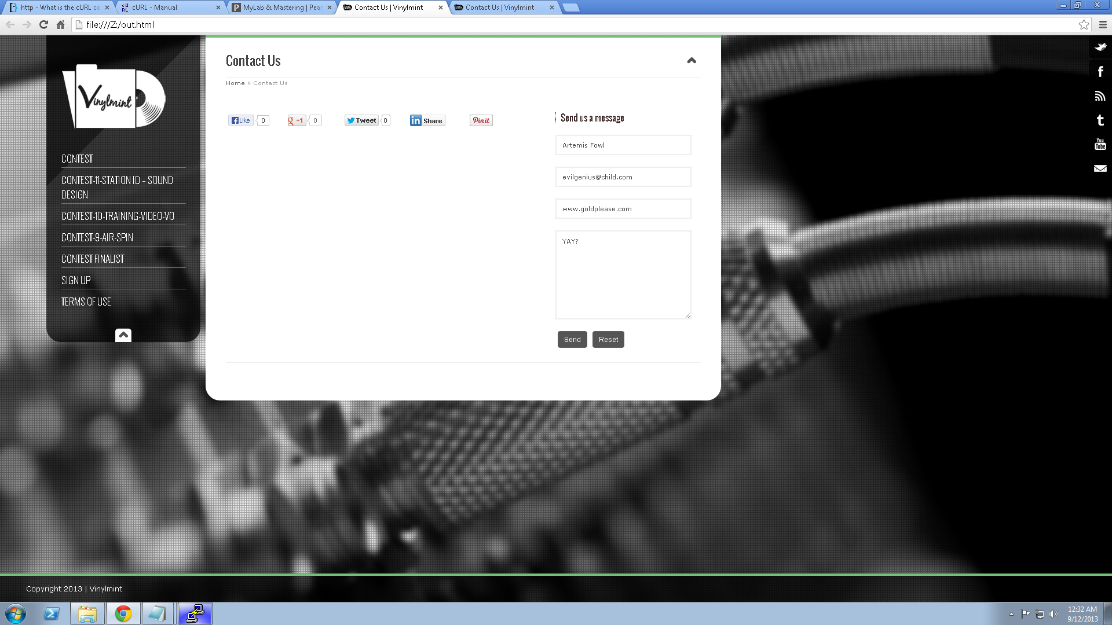


Figure 1: Full screen caption of out.html with data from the command line

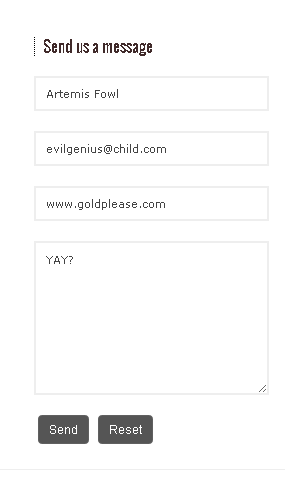


Figure 2: The filled in form from Figure 1.

**Question 2**: Write a Python program that:

1. Takes one argument, like "Old Dominion" or "Virginia Tech"
2. Takes another argument specified in seconds (e.g., "60" for one minute).
3. takes a URI as a third argument, such as:

* http://scores.espn.go.com/ncf/scoreboard?confId=80&seasonYear=2013&seasonType=2&weekNumber=2
* http://scores.espn.go.com/ncf/scoreboard?confId=80&seasonYear=2013&seasonType=2&weekNumber=1
* http://scores.espn.go.com/ncf/scoreboard?confId=80&seasonYear=2012&seasonType=2&weekNumber=1

1. Downloads the URI, finds the game corresponding to the teamargument, prints out the current score (e.g., "Old Dominion 38, East Carolina 52), sleeps for the specified seconds, and then repeats (until control-C is hit).

You can use any source for college football box scores that you'd like.

The program takes 4 arguments and check for the length of the arguments. If the total parameters is 4, the program proceeds, while if the condition was not met, the program exits with usage message. Variable called soup is assigned via BeautifulSoup function. The program loops through soup to find the div class “mod-content”, which are node for each of the recorded game. If the name of the entered school is found, find classes for “team visitor” and “team home” and then proceeds to print out the name of the visiting team, its score, home team, and home team’s score.

The attached python script is attached and called getscore.py.

**Question 3**: Consider the "bow-tie" graph in the Broder et al. paper (fig 9): http://www9.org/w9cdrom/160/160.html

Now consider the following graph:

A --> B

B --> C

C --> D

C --> A

C --> G

E --> F

G --> C

G --> H

I --> H

I --> J

I --> K

J --> D

L --> D

M --> A

M --> N

N --> D

For the above graph, give the values for:

IN: M, A

SCC: C, G

OUT: D, K, H

Tendrils: J, B

Tubes: L, I, N

Disconnected: E, F

E

L

N

M

J

F

K

I

H

G

D

C

B

A

Figure 3: The graph showing the relationship between each node.

Works Cited:

<http://superuser.com/questions/149329/what-is-the-curl-command-line-syntax-to-do-a-post-request>

<http://curl.haxx.se/docs/manual.html>

<http://www9.org/w9cdrom/160/160.html>