INF 510

Homework 5 (60 points)

Due Friday, April 19th at 11:59pm (via blackboard)

*AKA "Project" Milestone #1*

Basically, your task for this is to code up interfaces for the data sources you specified in HW3 (and plan to use for the future HW/project milestones as specified in class).

You are to turn in a Python script (NOT a notebook!) that can access all three of your data sources. You are to specify which source via the command line. If there are configurations beyond the specific data source (i.e. how many days, months, items, years, etc. to scrape) you should specify command-line parameters for those as well (In the example below, we're only grabbing 3 items)

Your script should, when invoked, output the following: A brief debugging output describing what the data source IS, where you're getting it from (URL, API endpoint, etc.) and what, exactly, you're obtaining. Also include a brief output of a few items from the data source. You do not need to output ALL data items (nor should you!)

Name your script: LASTNAME_FIRSTNAME_hw5.py (you will LOSE points if you don't do this!)

Example output:

```
$ python AKOMOLEDE_TOBI_hw5.py source1 –players=3

*** Outputing data source 1 ***

This data source is some ESPN recruiting information from ESPN.com.  We're
grabbing the player name, position, college and ranking.

URL:
http://www.espn.com/collegesports/football/recruiting/databaseresults/_/sportid/24/class/2012/sort/stars/order/true/starsfilter/GT/ratingfilter/GT/statuscommit/Commitments/statusuncommit/Uncommited

Sample output:

Mario Edwards, DE, 5 Stars, Florida State

Johnathan Gray, RB, 5 Stars, Texas

Dorial Green-Beckham, WR, 5 Stars, Missouri

$
```

For now, you don't necessarily need to put the data you scrape into a nice object model, etc. (that's the next assignment!). You can do that now if you'd like, but it is acceptable to just output the data. The point is you need to have your infrastructure to obtain the data up and running. We'll work on modeling the data next. **DUE FRIDAY, April 19th at 11:59pm** (via blackboard)

The rubric for HW5 is as follows:

- Can successfully access all three data sources: +40
    - Debugging output describes (for all sources):
        - What the data source is: +5
        - Where you're getting it from: +5
        - What exactly you're obtaining: +5
    - Source is specified via the command line: +2.5
    - Output includes a few items from the data source: +2.5
    - Code is poorly documented: -5
- Can access only two data sources: +25
    - Debugging output describes (for all sources):
        - What the data source is: +5
        - Where you're getting it from: +5
        - What exactly you're obtaining: +5
    - Source is specified via the command line: +2.5
    - Output includes a few items from the data source: +2.5
    - Code is poorly documented: -5
- Can access only one data source: +15
    - Debugging output describes (for all sources):
        - What the data source is: +5
        - Where you're getting it from: +5
        - What exactly you're obtaining: +5
    - Source is specified via the command line: +2.5
    - Output includes a few items from the data source: +2.5
- Doesn't run (i.e. tracebacks): +0
- Script is not named LASTNAME_FIRSTNAME_hw5.py: -20% of final score