

CSc 120: Fall 2019

Assignment 01: Part 4

Start: Fri Aug 30, 2019

Due: 8:00 PM, Tue Sept 3 2019

Important: Follow the directions given *exactly*: your code will be graded by a software script, so any deviation from the program specification can result in a significant loss of credit. If you are unsure about something, ask for clarification in office hours or Piazza (however, see the class's academic integrity policy about sharing your code or using code written by others).

Please pay attention to the [programming style guidelines](#) for this class. For this assignment you will be notified of style violations but not penalized for them; style violations will be penalized in subsequent assignments.

Problem 1.

This problem involves analysis of data for the US Women's National Soccer Team. For the purposes of this assignment we will use the file [data.csv](#) (based on the data available [here](#)). Your program will use this file to compute the player(s) who won the "Athlete of the Year" award the greatest number of times between start and stop years specified by the user.

Write a program, in a file **soccer.py**, that behaves as follows:

1. It reads in the data file specified above. You can assume that this file is in the same directory as your program.
2. It queries the user for a start year and a stop year using (in this order):

```
input("start: ")
and
```

```
input("stop: ")
```

3. It computes a list of the name(s) of the players who won "Athlete of the Year" the greatest number of times between start_year and stop_year (inclusive). If there is a tie, all of the players who tied are included in the list.
4. It sorts this list of names and prints them out using the statement:

```
print(name)
```

No output is generated if the list of Athlete of the Year players is empty (e.g., if the start year provided is later than the stop year, or if the start and stop years fall outside the data range available).

File format

The data file is in "csv", or "comma-separated values", format, which means that the different data fields on each line are separated by commas. Any line that begins with the character '#' is a comment and should be

ignored. The first line of the file is a comment that describes the different data fields. The “Athlete of the Year” field is field 5 (starting from field 0).

Programming hints

For the purposes of this assignment, the simplest way to process the data file is to use **split()** on the non-comment lines, then access the various fields by indexing into the resulting list as needed. (Python also has a [CSV module](#) that is useful for reading more complex csv files. You can use that for this assignment if you want, but you’ll have to figure out how to use it.)

I found it convenient to use a dictionary to map player names to athlete-of-the-year counts.

You can sort a list `L` in a couple of different ways:

- `L.sort()` sorts `L` in place. It does not return a value.
- `sorted(L)` leaves `L` unchanged and returns a sorted version of `L`.

More documentation on sorting is available [here](#).

Examples

Input file (contains start, stop values)	Output
input1.txt	Abby Wambach Mia Hamm
input2.txt	Abby Wambach Alex Morgan Carli Lloyd Lauren Holiday Tobin Heath
input3.txt	Carli Lloyd
input4.txt	Alex Morgan Julie Ertz
input5.txt	(no output)