

CS 1803

Individual Homework 8 – Web Table Reader

Due: Monday, April 4th, before 6 PM

Out of 100 points

Files to submit: 1. HW8.py

This is an INDIVIDUAL assignment!

Collaboration at a reasonable level will not result in substantially similar code. Students may only collaborate with fellow students currently taking CS 1803, the TA's and the lecturer. Collaboration means talking through problems, assisting with debugging, explaining a concept, etc. You should not exchange code or write code for others.

For Help:

- TA Helpdesk – Schedule posted on class website.
- Email TA's or use T-Square Forums

Notes:

- **Don't forget to include the required comments and collaboration statement (as outlined on the course syllabus).**
 - **Do not wait until the last minute** to do this assignment in case you run into problems.
-

Introduction:

The law firm of Dewey Cheatum & Howe post information about their foreclosure sales at the following URL daily:

<http://www.summet.com/dmsi/html/codesamples/dch.html>

You have been hired by a property investor to write a program that finds the "best deal" out of all of the upcoming foreclosure auctions. The "best deal" is defined to be the property which has the lowest ratio between the minimum bid amount and the original mortgage amount. For example, a house with a minimum bid amount of \$51,981.77 and an original mortgage amount of \$467,836.00 would have a ratio of 0.1109, and be a relatively good deal!

The property investor is hiring you to build a program he can run, but he also wants the source code! He specifies that the source code must be easy to read and well documented (with comments) so that he can hire another student after you graduate to maintain the program and potentially add extra features.

Table Web Scraper:

Write a function or method that will retrieve the table of data at the above URL. Once you have the data, calculate the ratio for each property. Note that many properties do not have a minimum bid amount available until near the end of the month. If a property does not have a minimum bid amount available, ignore it.

You may make use of the HTML table parser demonstrated in class and posted on the class course-notes (readTable.py), but you must indicate the source of that function with comments in your code. If you modify that function, be sure to use comments to explain how/why you modified the code so the TA's will know what extra work you have done.

A Snag:

You quickly write the program and send it via email to your client. The next day, he sends back an email claiming that the program is broken, asking you to fix it! After some investigation, you figure out that the law firm of Dewey Cheatum & Howe have outsourced their website to Cut Rate Web Sites, inc. Every morning, CRWS receives an email with a spreadsheet file from the law firm, and each morning a different intern converts the data to HTML and posts the new data on the website. Because of this, sometimes they exclude what they think is an "extra" column, or forget to format the currency correctly, or don't indicate properties where the bid amount is "Not Available". (On other days you get an almost qualified intern who fixes all of the above mentioned problems...) This means that you will have to make your program robust and able to handle malformed data, missing columns, etc... You will also have to test your program for several days to make sure it still works no matter which intern is posting the website today! (So get started early!)

Informational GUI

Write a simple GUI that has a button ("Get Info"). When the user clicks the button, the GUI will call your previously written code and display all of the information about the property with the lowest ratio visually to the user. You must at a minimum display the full address of the property, along with the minimum bid amount and original mortgage amount. You may use a simple Entry Widget or label. When your client (or the TA) runs your python file, your GUI should appear automatically.

Grading

You will earn points as follows for each element that works correctly according to the specifications.

Comments are used to explain the purpose of each function	10
Code is easily readable and understandable	10
Successfully downloads the webpage HTML	10
Correctly Parses the HTML table	10
Correctly retrieves needed data from the table	20
Functions correctly on both odd and even days of the month	10
The property with the lowest ratio is correctly identified	10
"get Info" button triggers the download/display	10
Information about the property is displayed correctly	10

+5 Bonus EC points are available for displaying ALL of the properties in a scrollable list such that they are ordered with the property having the lowest ratio first. Any properties that do not have a minimum bid amount available should be located at the end of the list.