

## DATA 200 Spring 2020 Homework 10

Due May 8, 2020 at 11:59pm

**Note: Please only use python standard libraries covered in class lectures (i.e. json, csv etc). Please submit your solution for each problem as a standalone python file (.py) and NOT a Jupyter notebook. The TA should be able to run your program as is outside of Jupyter.**

1. Develop a python class **HTMLVerifier** to check an HTML document for proper opening and closing tags. In HTML, tags are defined in both opening and closing forms and must be balanced for browsers to properly render a web page. Here's a simple example of an HTML document:

```
<html>
  <head>
    <title>
      Example
    </title>
  </head>

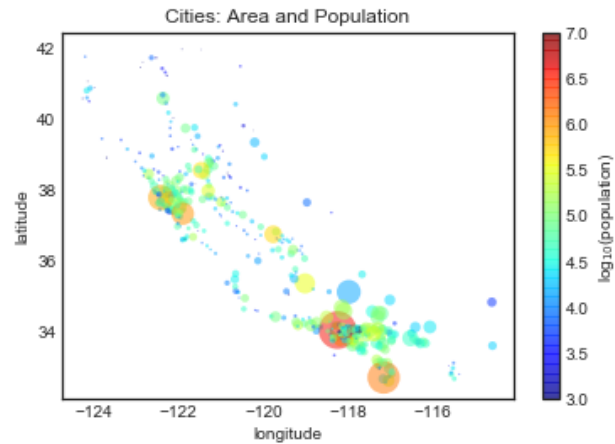
  <body>
    <h1>Hello, world</h1>
  </body>
</html>
```

The class should have 3 methods to perform the following tasks:

- Initialize an instance with a filename and read file content,
- Verify that the HTML tags are properly balanced in the document, and
- Print out any unbalanced tags in the document.

(15 points)

2. Using the data from ***cities.csv***, create a scatter plot of cities (as shown below) using python and matplotlib. The x-axis is the longitude of cities, the y-axis is the latitude of cities, the color represents the logarithm of city population and the size of the circles represents the total city area in square miles.



(10 points)