

IN CLASS PROBLEM 1:

Write a HTML parser named StatParser. Stat parser keeps a count of all tags encountered on a page and returns a **dictionary** with the tag name as a key and the count of how many times the tag was encountered in the document as a value.

Example output running testStatParser:

```
>>> testStatParser('http://www.cdm.depaul.edu')
{'html': 1, 'head': 1, 'meta': 24, 'link': 11, 'title': 1, 'script': 60, 'span':
 27, 'style': 4, 'body': 1, 'noscript': 2, 'iframe': 2, 'div': 94, 'ul': 28, 'li
': 170, 'a': 160, 'form': 1, 'input': 25, 'button': 3, 'img': 5, 'label': 2, 'i
': 15, 'nav': 2, 'h2': 11, 'h1': 2, 'p': 9, 'menu': 1, 'ie:menuitem': 1, 'footer
': 1, 'br': 3, 'strong': 3}
>>> |
```

IN CLASS PROBLEM 2:

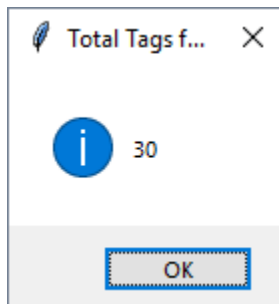
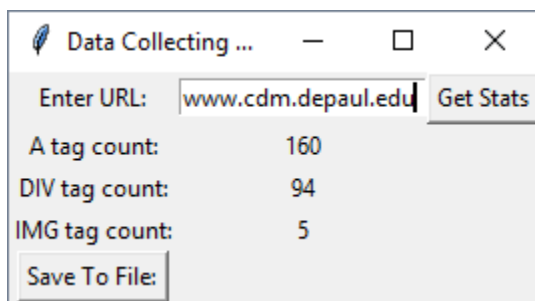
Take the StatParser from problem 1 and use it in the BrowserGUI class. BrowserGUI allows the user to browse a site and display the count of the most popular tags found on a page (a, div and img). There is also a button on the form that lets the user *append all of the tag stats* to a file named stats.txt.

When the user click the 'Get Stats' button a popup displays how many tags were found on the page and the labels showing how many a,div and img tags were encountered.

The text file stats with a line URL: and the url of the site visited. Then all the tags and their counts are listed on a separate line in the format tag:count.

You are not allowed to use testStatParser() function from Problem 1. You must use the StatParser class directly within the GUI.

See the following screenshots for an example of how the GUI behaves and how the file looks after the data is appended. **NOTE: NUMBERS MAY BE DIFFERENT. SITE CHANGES DAILY.**



Contents of the file after clicking "Save to file" twice:

URL:http://www.cdm.depaul.edu

html-1

head-1

meta-24

link-11

title-1

script-60

span-27

style-4

body-1

noscript-2

iframe-2

div-94

ul-28

li-170

a-160

form-1

input-25

button-3

img-5

label-2

i-15

nav-2

h2-11

h1-2

p-9

menu-1

ie:menuitem-1

footer-1

br-3

strong-3

URL:http://www.cdm.depaul.edu

html:1

head:1

meta:24

link:11

title:1

script:60

span:27

style:4

body:1

noscript:2

iframe:2

div:94

ul:28

li:170

a:160
form:1
input:25
button:3
img:5
label:2
i:15
nav:2
h2:11
h1:2
p:9
menu:1
ie:menuitem:1
footer:1
br:3
strong:3

