

```
#!/usr/bin/env python3
```

```
# Limitations:
```

```
# The program is limited by the number of command line arguments supported by the  
operations system.
```

```
# The program assumes a 'utf-8' encoding by the website.
```

```
#
```

```
# This program requires Python 3
```

```
#
```

```
# How to execute this program:
```

```
#
```

```
# Save this file as parsehtml.py
```

```
#
```

```
# Run the program by providing a list of web sites on the command line, for example:
```

```
# parsehtml.py https://www.perry-brandiezs.com https://wikipedia.com
```

```
http://www.columbia.edu/~fdc/sample.html
```

```
#
```

```
# Sample output:
```

```
# >parsehtml.py https://www.perry-brandiezs.com https://wikipedia.com
```

```
http://www.columbia.edu/~fdc/sample.html
```

```
# https://www.perry-brandiezs.com 6
```

```
# https://wikipedia.com 318
```

```
# http://www.columbia.edu/~fdc/sample.html 47
```

```
#
```

```
# Github
```

```
# https://github.com/pbrandiezs/parsehtml
```

```
#
```

```
# Author: Perry Brandiezs 6/9/2021
```

```
import sys
```

```
import urllib.request
```

```
# Remove the calling program from the string
```

```
arg_str = ' '.join(sys.argv[1:])
```

```
# Create a list of the websites
```

```
site_list = list(arg_str.split(" "))
```

```
# Open each site, and count href occurrences
```

```
for site in site_list:
```

```
    with urllib.request.urlopen(site) as response:
```

```
        html = response.read()
```

```
        # print (html)
```

```
        encoding = 'utf-8'
```

```
        html_string = html.decode(encoding)
```

```
        print (site + " " + str(html_string.count('href')))
```