

Lab5-1_HTMLElementsAndWebRequests
SYS-320
Ben W
SP24

Deliverables:

1. Basic use of Invoke-WebRequest getting the number of links

```
1 # Powershell Script
2 # SYS-320 - Ben W
3
4 $scraped_page =Invoke-WebRequest -TimeoutSec 10 http://localhost/tobescraped.html
5
6 $scraped_page.Links.Count
```

```
<
PS C:\xampp\apache\logs> C:\Users\champuser\Documents\sys320git\PowershellLabs\Lab05-1-Cla
1
```

2. Get the links as html elements

```
1 # Powershell Script
2 # SYS-320 - Ben W
3
4 $scraped_page =Invoke-WebRequest -TimeoutSec 10 http://localhost/tobescraped.html
5
6 $scraped_page.Links
```

```
<
PS C:\xampp\apache\logs> C:\Users\champuser\Documents\sys320git\PowershellLabs\Lab05-1-C

innerHTML : HTML Tags.
innerText : HTML Tags.
outerHTML : <A href="https://www.w3schools.com/tags/">HTML Tags. </A>
outerText : HTML Tags.
tagName   : A
href      : https://www.w3schools.com/tags/
```

3. Get the links and display only the URL and its text.

```
1 # Powershell Script
2 # SYS-320 - Ben W
3
4 $scraped_page =Invoke-WebRequest -TimeoutSec 10 http://localhost/tobescraped.html
5
6 $scraped_page.Links | select-Object outerText, href
```

```
<
PS C:\xampp\apache\logs> C:\Users\champuser\Documents\sys320git\PowershellLabs\Lab05-1-Cla

outerText  href
-----
HTML Tags. https://www.w3schools.com/tags/
```

4. Get outer text of every element with the tag h2

```
1 # PowerShell Script
2 # SYS-320 - Ben W
3
4 $scraped_page = Invoke-WebRequest -TimeoutSec 10 http://localhost/tobescraped.html
5
6 $h2s = $scraped_page.ParsedHtml.body.getElementsByTagName("h2") | Select outerText
7
8 $h2s
```

```
PS C:\xampp\apache\logs> C:\Users\champuser\Documents\sys320git\PowerShellLabs\Lab05-1-ClassActivity\Deliverable5.ps1
outerText
-----
Hi, am second div with h2, span, and br tags
To learn more about tage: ...
```

5. Print innerText of every div element that has the class as "div-1"

```
1 # PowerShell Script
2 # SYS-320 - Ben W
3
4 $scraped_page = Invoke-WebRequest -TimeoutSec 10 http://localhost/tobescraped.html
5
6 $divs = $scraped_page.ParsedHtml.body.getElementsByTagName("div") | ? {$_.getAttributeNode("class").value -like "div-1"} | Select innerText
7
8 $divs
```

```
PS C:\xampp\apache\logs> C:\Users\champuser\Documents\sys320git\PowerShellLabs\Lab05-1-ClassActivity\Deliverable5.ps1
innerText
-----
Hi, am first Div
```