**Activity 2**

<div class="quote" itemscope="" itemtype="http://schema.org/CreativeWork">

<span class="text">“The world as we have created it is a process of our thinking. It cannot be changed without changing our thinking.”</span>

<span>by <small class="author">Albert Einstein</small>

<a href="/author/Albert-Einstein">(about)</a>

</span>

<div class="tags">Tags:

<a class="tag" href="/tag/change/page/1/">change</a>

<a class="tag" href="/tag/deep-thoughts/page/1/">deep-thoughts</a>

<a class="tag" href="/tag/thinking/page/1/">thinking</a>

<a class="tag" href="/tag/world/page/1/">world</a>

</div>

</div>

Locate the items that you are going to extract from the HTML strings and fill in the table below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable Name | Description | Single  or  Multiple items | Tag Name | Attribute | |
| Name | Value |
| quote\_text | Text of the quote | single | span | class | text |
| author | Name of the author |  |  |  |  |
| author\_url | URL of the author page |  |  |  |  |
| tags | Tags that assigned to the quote |  |  |  |  |

**Activity 3**

Observe the patterns of the URLs of different pages, and make a guess of the URL for page 1 and page 4.

|  |  |
| --- | --- |
| Page Number | Page URL |
| 1 |  |
| 2 | http://quotes.toscrape.com/page/2/ |
| 3 | http://quotes.toscrape.com/page/3/ |
| 4 |  |

What is the pattern of the URL?

1. http://quotes.toscrape.com/page/ + {page\_number}
2. http://quotes.toscrape.com/ + {page\_number}
3. No pattern