HTML-to-Markdown

# How to use it?

* Extract the archive in a location
* Choose the html file of your choice and place it in the same directory (just to avoid issues with path and filenames)
* Run the following command:

#python htmlparser.py **<name\_of\_the\_html\_file.html>**

* The Markdown version of the HTML file is displayed

# What are the Features supported?

* Heading tags :

<h1>, <h2> ….tags

* Ordered list tags:

<ul>

<li>first one</li>

<li>second one</li>

</ul>

* Unordered list tags

<ul>

<li>first one</li>

<li>second one</li>

</ul>

* Nested unordered list tags

<ul>

<li>first one</li>

<ul>

<li>nested level 1</li>

<li>nested level 1</li>

</ul>

<li>second one</li>

</ul>

* Nested Ordered list tags **don’t** **work** as expected yet – working on it. Probably will need to stack the current bullet number at each nesting level and popping when required

<ol>

<li>first one</li>

<ol>

<li>nested level 1</li>

<li>nested level 1</li>

</ol>

<li>second one</li>

</ol>

* Link tags:

<a href = “hhtp://www.google.com”>a link here </a>

* Code tags – They are displayed but are **not well formatted** well given that each line of code is not in a specific tag causing me issues in identifying the breakpoint.

If I can make language specific code, it might be implementable given that each line of code has a specific start pattern.

(Say) for Shell script it can be ‘$’ or ‘#’