<u>Description of the HTML structure of a WEB page, case of 'Sharp Objects '</u>

- Title of the book :

The HTML <h1> or <h2> element containing the title of the book. In the "Sharp Objects " example , it is located inside a <div> tag with the col-sm-6 product_main class .

-Book Price:

structure : <p class=" price_color ">£47.82</p>

The price of the book is contained in a <p> tag with the price_color class ("£47.82" in our example).

-Available stock:

The HTML <p> element with the instock class availability containing information on the availability of the book ("In stock (15 available)" in our example).

-Number of Stars:

The stars assigned to the book are represented by a <p> tag with the star-rating class, where the specific class ("Four in our example) indicates the number of stars. CSS styles are used to determine the number of stars assigned.

-Book Description:

HTML structure : <meta name="description" content="description of the book." />

The description of the book is included in a < meta > tag with the name ="description" attribute. This tag is found in the < head > section of the document.

-UPC:

The UPC is present in a <th> tag followed by its value in a <td> tag.
("e00eb4fd7b871a48" in our example)

-Product Type:

HTML structure : < th >Product Type</ th > <td> value_of_product_type </td>

The product type is indicated in a <th> tag followed by its value in a <td> tag
(Books in our example).

- Tax :

HTML structure : < th >Tax</ th > <td> value_of_tax </td>

Tax information is included in a <th> tag followed by its value in a <td> tag
(£0.00 in our example).

- Number Number of Reviews :

HTML structure: < th >Number of Reviews </ th > <td>value </td>

The number of reviews is included in a <th> tag followed by its value in a <td>
tag (0 in our example).

-Price (excl. tax):

HTML structure : < th >Price (excl. tax)</ th > <td> value_of_price_excl_tax
</td>

The price excluding tax is in a <th> tag followed by its value in a <td> tag
(£47.82 in our example).

-Price (incl. tax):

HTML structure : < th >Price (incl. tax)</ th > <td> value_of_price_incl_tax
</td>

The tax-inclusive price is included in a <th> tag followed by its value in a <td>
tag (£47.82 in our example).

Study limitation: In our work, we encountered limitations in scraping the 'author name' which was mentioned on the images of the books and not as information in the html structure of a web page of a book.