

Data Extraction from HTML Page

Marks Distribution:

Marks will be awarded on the following parameters:

- Code
 - Errors logged in the excel sheet
 - video of the code
 - Viva
 - Extra testcases (Other web pages)
-

Extract the following data from the given web page.

- Image URLs and the number of images
 - Here you need to print all the URLs of the image
 - All the images links will be there in “img” tag under “src” attribute
 - Display the count of images present on the webpage
- Background colours
 - Display all the background colours available on the webpage
 - Background colours are indicated with an attribute “background-color”
 - Display the colours in sorted order (remove duplicate colours)
 - Display the count of different colours present on the webpage
- Text present in between the li tags
 - Display the text that is present in between the list tags
 - List tags are represented using “li” tag
 - Display the text
 - Display the count of texts present on in between the list tags

Note :

- Observe the required tags carefully.

Input format :

- The webpage is provided in the folder
- The first line in each testcase contains the following words
 - image (image urls and count)
 - background (background colours and count)
 - list (text between the list tags)

Output format :

- Text followed by the count in the last line

