WEB SCRAPING

Web scraping is a software technique of extracting information from websites in an automated fashion. There maybe a situation where you find data from the web, and there is no direct way to download it, web scraping will be saviour and using python makes it easier to extract the data into a useful form that can be processed and analyzed.

Here, we will scrap data from wekipedia, E-commerce website, Amazon.com.

```
Python libraries used for scraping data are:

a)BeautifulSoup : prettify

b)Requests

c)selenium

d)scrapy
```

In [1]:

```
# import required libraries
from bs4 import BeautifulSoup
import requests
```

In [2]:

```
# url from web scraping
url = 'http://www.simplilearn.com'
```

In [3]:

```
# access the url with request object
result = requests.get(url)

# scrape the web page content
webpage= result.content

# create a soup object for parsing web page using html parser
sl_soup = BeautifulSoup(webpage,'html.parser')
```

In [4]:

```
# close the result object
result.close()
```

In [5]:

```
# view the content of the soup object
sl_soup.contents
```

Out[5]:

```
['html',
 <html dir="ltr" lang="en-US" xmlns="https://www.w3.org/1999/xhtml">
 <head>
 <meta content="text/html; charset=utf-8" http-equiv="Content-Type"/>
 <meta content="IE=edge" http-equiv="X-UA-Compatible"/>
 <title>Simplilearn | Online Courses - Bootcamp &amp;amp; Certification Platform</title>
 <link href="https://ssl.google-analytics.com/" rel="preconnect"/>
 <link href="https://stats.g.doubleclick.net/" rel="dns-prefetch">
 <link href="https://www.google.com/" rel="dns-prefetch"/>
 <link href="https://cdn.cookielaw.org/" rel="preconnect"/>
 <link href="https://js-agent.newrelic.com" rel="preconnect"/>
 <link href="https://accounts.google.com/" rel="dns-prefetch"/>
 <link href="https://googletagmanager.com/" rel="preconnect"/>
 <link href="https://cdn.gumlet.com/" rel="preconnect"/>
 <script type="text/javascript">
   ;window.NREUM||(NREUM={});NREUM.init={privacy:{cookies_enabled:true}};window.NREUM||(NREUM={}),__nr_require=f
unction(t,e,n){function r(n){if(!e[n]){var i=e[n]={exports:{}};t[n][0].call(i.exports,function(e){var i=t[n][1]
[e];return r(i||e)},i,i.exports)}return e[n].exports}if("function"==typeof __nr_require)return __nr_require;for
```

In [6]:

```
# use prettify method to view well formatted output (html tags)
print(sl_soup.prettify())
```

```
<!DOCTYPE html>
<html dir="ltr" lang="en-US" xmlns="https://www.w3.org/1999/xhtml">
 <head>
  <meta content="text/html; charset=utf-8" http-equiv="Content-Type"/>
  <meta content="IE=edge" http-equiv="X-UA-Compatible"/>
  <title>
   Simplilearn | Online Courses - Bootcamp &amp;amp; Certification Platform
  </title>
  <link href="https://ssl.google-analytics.com/" rel="preconnect"/>
  <link href="https://stats.g.doubleclick.net/" rel="dns-prefetch">
   <link href="https://www.google.com/" rel="dns-prefetch"/>
   <link href="https://cdn.cookielaw.org/" rel="preconnect"/>
   <link href="https://js-agent.newrelic.com" rel="preconnect"/>
   <link href="https://accounts.google.com/" rel="dns-prefetch"/>
   <link href="https://googletagmanager.com/" rel="preconnect"/>
   <link href="https://cdn.gumlet.com/" rel="preconnect"/>
   <script type="text/javascript">
    ;window.NREUM||(NREUM={});NREUM.init={privacy:{cookies_enabled:true}};window.NREUM||(NREUM={}),__nr_require=
function(t,e,n){function r(n){if(!e[n]){var i=e[n]={exports:{}};t[n][0].call(i.exports,function(e){var i=t[n][1]
```

In [7]:

```
# view the head of the soup object
print(sl_soup.head())
```

```
[<meta content="text/html; charset=utf-8" http-equiv="Content-Type"/>, <meta content="IE=edge" http-equiv="X-UA-
Compatible"/>, <title>Simplilearn | Online Courses - Bootcamp &amp;amp; Certification Platform</title>, <link href
="https://ssl.google-analytics.com/" rel="preconnect"/>, <link href="https://stats.g.doubleclick.net/" rel="dns-
prefetch">
<link href="https://www.google.com/" rel="dns-prefetch"/>
<link href="https://cdn.cookielaw.org/" rel="preconnect"/>
<link href="https://js-agent.newrelic.com" rel="preconnect"/>
<link href="https://accounts.google.com/" rel="dns-prefetch"/>
<link href="https://googletagmanager.com/" rel="preconnect"/>
<link href="https://cdn.gumlet.com/" rel="preconnect"/>
<script type="text/javascript">
  ;window.NREUM||(NREUM={});NREUM.init={privacy:{cookies_enabled:true}};window.NREUM||(NREUM={}),__nr_require=fu
nction(t,e,n){function r(n){if(!e[n]){var i=e[n]={exports:{}};t[n][0].call(i.exports,function(e){var i=t[n][1]
[e];return r(i||e)},i,i.exports)}return e[n].exports}if("function"==typeof __nr_require)return __nr_require;for
(var i=0;i<n.length;i++)r(n[i]);return r}({1:[function(t,e,n){function r(t){try{c.console&&console.log(t)}catch
(e){}}var i,o=t("ee"),a=t(24),c={};try{i=localStorage.getItem("__nr_flags").split(","),console&&"function"==type
of console.log&&(c.console=!0,i.indexOf("dev")!==-1&&(c.dev=!0),i.indexOf("nr_dev")!==-1&&(c.nrDev=!0))}catch(s)
{}c.nrDev&&o.on("internal-error",function(t){r(t.stack)}),c.dev&&o.on("fn-err",function(t,e,n){r(n.stack)}),c.de
v&&(r("NR AGENT IN DEVELOPMENT MODE"),r("flags: "+a(c,function(t,e){return t}).join(", "))},{}],2:[function(t,
```

In [8]:

```
# view the title of the web page
print(sl_soup.title)
```

```
<title>Simplilearn | Online Courses - Bootcamp &amp;amp; Certification Platform</title>
```

In [9]:

```
# Find all the links present the web page
for href in sl_soup.findAll('a',href=True):
    print (href['href'])
```

```
https://www.simplilearn.com/job-guarantee-programs (https://www.simplilearn.com/job-guarantee-programs)
https://www.simplilearn.com/resources (https://www.simplilearn.com/resources)
https://www.simplilearn.com/corporate-training (https://www.simplilearn.com/corporate-training)
https://www.simplilearn.com/become-our-trainer (https://www.simplilearn.com/become-our-trainer)
https://www.simplilearn.com/simplirecruit-hire-employees-from-us (https://www.simplilearn.com/simplirecruit-hire
-employees-from-us)
https://www.simplilearn.com/reviews (https://www.simplilearn.com/reviews)
https://www.simplilearn.com (https://www.simplilearn.com)

#post-graduate-programs
https://www.simplilearn.com/ice9/assets/Simplilearn_Terms_and_Conditions_Final.pdf (https://www.simplilearn.com/
ice9/assets/Simplilearn_Terms_and_Conditions_Final.pdf)
ai-for-decision-making-program
product-management-certification-training-course
vapt-vulnerability-assessment-penetration-testing-certification
general-management-certification-training-course
ms-in-artificial-intelligence
professional-brand-strategy-certification
professional-certificate-digital-marketing
```

In [ ]:

```
```

In [10]:

```
ourdata = requests.get('https://en.wikipedia.org/wiki/Artificial_intelligence')
```

In [11]:

```
ourdata.text
```

Out[11]:

```
'<!DOCTYPE html>\n<html class="client-nojs" lang="en" dir="ltr">\n<head>\n<meta charset="UTF-8"/>\n<title>Artifi
cial intelligence - Wikipedia</title>\n<script>document.documentElement.className="client-js";RLCONF={"wgBreakFr
ames":false,"wgSeparatorTransformTable":["",""],"wgDigitTransformTable":["",""],"wgDefaultDateFormat":"dmy","wgM
onthNames":["","January","February","March","April","May","June","July","August","September","October","Novembe
r","December"],"wgRequestId":"4ef5ac8a-f327-4788-a92f-a7a4af14a0d9","wgCSPNonce":false,"wgCanonicalNamespac
e":"","wgCanonicalSpecialPageName":false,"wgNamespaceNumber":0,"wgPageName":"Artificial_intelligence","wgTitl
e":"Artificial intelligence","wgCurRevisionId":1131103814,"wgRevisionId":1131103814,"wgArticleId":1164,"wgIsArti
cle":true,"wgIsRedirect":false,"wgAction":"view","wgUserName":null,"wgUserGroups":["*"],"wgCategories":["All art
icles with dead external links","Articles with dead external links from October 2022","Articles with permanently
dead external links","CS1 errors: missing periodical","Articles with short description","Short description is di
fferent from Wikidata",\n"Wikipedia indefinitely semi-protected pages","Use dmy dates from August 2022","All art
icles lacking reliable references","Articles lacking reliable references from December 2022","All articles that
may have off-topic sections","Wikipedia articles that may have off-topic sections from November 2022","CS1 Finni
sh-language sources (fi)","CS1: Julian–Gregorian uncertainty","CS1: long volume value","Webarchive template wayb
ack links","Pages using Sister project links with hidden wikidata","Articles with Internet Encyclopedia of Philo
sophy links","Articles with BNE identifiers","Articles with BNF identifiers","Articles with GND identifiers","Ar
ticles with J9U identifiers","Articles with LCCN identifiers","Articles with LNB identifiers","Articles with NDL
identifiers","Articles with NKC identifiers","Artificial intelligence","Cybernetics","Formal sciences","Data sci
```

In [12]:

```
bsp = BeautifulSoup(ourdata.text,'xml')
```

In [13]:

```
bsp
```

Out[13]:

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE html>
<html class="client-nojs" dir="ltr" lang="en">
<head>
<meta charset="UTF-8"/>
<title>Artificial intelligence - Wikipedia</title>
<script>document.documentElement.className="client-js";RLCONF={"wgBreakFrames":false,"wgSeparatorTransformTabl
e":["",""],"wgDigitTransformTable":["",""],"wgDefaultDateFormat":"dmy","wgMonthNames":["","January","Februar
y","March","April","May","June","July","August","September","October","November","December"],"wgRequestId":"4ef5
ac8a-f327-4788-a92f-a7a4af14a0d9","wgCSPNonce":false,"wgCanonicalNamespace":"","wgCanonicalSpecialPageName":fals
e,"wgNamespaceNumber":0,"wgPageName":"Artificial_intelligence","wgTitle":"Artificial intelligence","wgCurRevisio
nId":1131103814,"wgRevisionId":1131103814,"wgArticleId":1164,"wgIsArticle":true,"wgIsRedirect":false,"wgActio
n":"view","wgUserName":null,"wgUserGroups":["*"],"wgCategories":["All articles with dead external links","Articl
es with dead external links from October 2022","Articles with permanently dead external links","CS1 errors: miss
ing periodical","Articles with short description","Short description is different from Wikidata",
"Wikipedia indefinitely semi-protected pages","Use dmy dates from August 2022","All articles lacking reliable re
ferences","Articles lacking reliable references from December 2022","All articles that may have off-topic sectio
ns","Wikipedia articles that may have off-topic sections from November 2022","CS1 Finnish-language sources (f
```

In [14]:

```
titles= bsp.select('title')
```

In [15]:

```
titles
```

Out[15]:

```
[<title>Artificial intelligence - Wikipedia</title>]
```

In [16]:

```
titles[0].getText()
```

Out[16]:

```
'Artificial intelligence - Wikipedia'
```

In [17]:

```
sub_titles= bsp.select('h2')
```

In [18]:

```
sub_titles
```

Out[18]:

```
[<h2 id="mw-toc-heading">Contents</h2>,
 <h2><span class="mw-headline" id="History">History</span></h2>,
 <h2><span class="mw-headline" id="Goals">Goals</span></h2>,
 <h2><span class="mw-headline" id="Tools">Tools</span></h2>,
 <h2><span class="mw-headline" id="Applications">Applications</span></h2>,
 <h2><span class="mw-headline" id="Intellectual_Property">Intellectual Property</span></h2>,
 <h2><span class="mw-headline" id="Philosophy">Philosophy</span></h2>,
 <h2><span class="mw-headline" id="Future">Future</span></h2>,
 <h2><span class="mw-headline" id="In_fiction">In fiction</span></h2>,
 <h2><span class="mw-headline" id="Scientific_diplomacy">Scientific diplomacy</span></h2>,
 <h2><span class="mw-headline" id="See_also">See also</span></h2>,
 <h2><span class="mw-headline" id="Explanatory_notes">Explanatory notes</span></h2>,
 <h2><span class="mw-headline" id="References">References</span></h2>,
 <h2><span class="mw-headline" id="Further_reading">Further reading</span></h2>,
 <h2><span class="mw-headline" id="External_links">External links</span></h2>,
 <h2>Navigation menu</h2>]
```

In [19]:

```
sub_titles[0].getText()
```

Out[19]:

```
'Contents'
```

In [20]:

```
for i in sub_titles:
    print(i.text)
```

```
Contents
History
Goals
Tools
Applications
Intellectual Property
Philosophy
Future
In fiction
Scientific diplomacy
See also
Explanatory notes
References
Further reading
External links
Navigation menu
```

In [21]:

```
bsp.select('link')
```

Out[21]:

```
[<link href="/w/load.php?lang=en&amp;modules=ext.cite.styles%7Cext.uls.interlanguage%7Cext.visualEditor.desktopA
rticleTarget.noscript%7Cext.wikimediaBadges%7Cjquery.makeCollapsible.styles%7Cskins.vector.styles.legacy%7Cwikib
ase.client.init&amp;only=styles&amp;skin=vector" rel="stylesheet"/>,
 <link href="/w/load.php?lang=en&amp;modules=site.styles&amp;only=styles&amp;skin=vector" rel="stylesheet"/>,
 <link href="//upload.wikimedia.org" rel="preconnect"/>,
 <link href="//en.m.wikipedia.org/wiki/Artificial_intelligence" media="only screen and (max-width: 720px)" rel
="alternate"/>,
 <link href="/static/apple-touch/wikipedia.png" rel="apple-touch-icon"/>,
 <link href="/static/favicon/wikipedia.ico" rel="icon"/>,
 <link href="/w/opensearch_desc.php" rel="search" title="Wikipedia (en)" type="application/opensearchdescription
+xml"/>,
 <link href="//en.wikipedia.org/w/api.php?action=rsd" rel="EditURI" type="application/rsd+xml"/>,
 <link href="https://creativecommons.org/licenses/by-sa/3.0/" rel="license"/>,
 <link href="https://en.wikipedia.org/wiki/Artificial_intelligence" rel="canonical"/>,
 <link href="//meta.wikimedia.org" rel="dns-prefetch"/>,
 <link href="//login.wikimedia.org" rel="dns-prefetch"/>,
 <link href="mw-data:TemplateStyles:r1129693374" rel="mw-deduplicated-inline-style"/>,
 <link href="mw-data:TemplateStyles:r1033289096" rel="mw-deduplicated-inline-style"/>,
```

In [22]:

```
bsp.select('img')
```

Out[22]:

```
[<img alt="Page semi-protected" data-file-height="512" data-file-width="512" decoding="async" height="20" src
="//upload.wikimedia.org/wikipedia/en/thumb/1/1b/Semi-protection-shackle.svg/20px-Semi-protection-shackle.svg.pn
g" srcset="//upload.wikimedia.org/wikipedia/en/thumb/1/1b/Semi-protection-shackle.svg/30px-Semi-protection-shack
le.svg.png 1.5x, //upload.wikimedia.org/wikipedia/en/thumb/1/1b/Semi-protection-shackle.svg/40px-Semi-protection
-shackle.svg.png 2x" width="20"/>,
 <img alt="Anatomy-1751201 1280.png" data-file-height="1088" data-file-width="1280" decoding="async" height="85"
src="//upload.wikimedia.org/wikipedia/commons/thumb/7/7a/Anatomy-1751201_1280.png/100px-Anatomy-1751201_1280.pn
g" srcset="//upload.wikimedia.org/wikipedia/commons/thumb/7/7a/Anatomy-1751201_1280.png/150px-Anatomy-1751201_12
80.png 1.5x, //upload.wikimedia.org/wikipedia/commons/thumb/7/7a/Anatomy-1751201_1280.png/200px-Anatomy-1751201_
1280.png 2x" width="100"/>,
 <img alt="" class="thumbimage" data-file-height="1020" data-file-width="1020" decoding="async" height="220" src
="//upload.wikimedia.org/wikipedia/commons/thumb/c/c8/Didrachm_Phaistos_obverse_CdM.jpg/220px-Didrachm_Phaistos_
obverse_CdM.jpg" srcset="//upload.wikimedia.org/wikipedia/commons/thumb/c/c8/Didrachm_Phaistos_obverse_CdM.jpg/3
30px-Didrachm_Phaistos_obverse_CdM.jpg 1.5x, //upload.wikimedia.org/wikipedia/commons/thumb/c/c8/Didrachm_Phaist
os_obverse_CdM.jpg/440px-Didrachm_Phaistos_obverse_CdM.jpg 2x" width="220"/>,
 <img alt="" class="thumbimage" data-file-height="443" data-file-width="469" decoding="async" height="208" src
="//upload.wikimedia.org/wikipedia/commons/thumb/9/9e/GFO_taxonomy_tree.png/220px-GFO_taxonomy_tree.png" srcset
="//upload.wikimedia.org/wikipedia/commons/thumb/9/9e/GFO_taxonomy_tree.png/330px-GFO_taxonomy_tree.png 1.5x, //
```

In [23]:

```
! pip install wikipedia
```

```
Requirement already satisfied: wikipedia in c:\users\lenovo\anaconda3\lib\site-packages (1.4.0)
Requirement already satisfied: requests<3.0.0,>=2.0.0 in c:\users\lenovo\anaconda3\lib\site-packages (from wikipedi
a) (2.28.1)
Requirement already satisfied: beautifulsoup4 in c:\users\lenovo\anaconda3\lib\site-packages (from wikipedia) (4.1
1.1)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\lenovo\anaconda3\lib\site-packages (from requests<
3.0.0,>=2.0.0->wikipedia) (1.26.4)
Requirement already satisfied: idna<4,>=2.5 in c:\users\lenovo\anaconda3\lib\site-packages (from requests<3.0.0,>=
2.0.0->wikipedia) (2.10)
Requirement already satisfied: charset-normalizer<3,>=2 in c:\users\lenovo\anaconda3\lib\site-packages (from reques
ts<3.0.0,>=2.0.0->wikipedia) (2.1.1)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\lenovo\anaconda3\lib\site-packages (from requests<3.
0.0,>=2.0.0->wikipedia) (2020.12.5)
Requirement already satisfied: soupsieve>1.2 in c:\users\lenovo\anaconda3\lib\site-packages (from beautifulsoup4->w
ikipedia) (2.2.1)
```

In [24]:

```
import wikipedia
from pprint import pprint
```

In [25]:

```
wikipedia.search('python')
```

Out[25]:

```
['Python',
 'Python (programming language)',
 'Monty Python',
 'Setuptools',
 'Python (codename)',
 'Ball python',
 'Reticulated python',
 'Burmese python',
 'Python (missile)',
 'Colt Python']
```

In [26]:

```
wikipedia.search('Artificial Intelligence')
```

Out[26]:

```
['Artificial intelligence',
 'Artificial general intelligence',
 'A.I. Artificial Intelligence',
 'Artificial intelligence art',
 'Applications of artificial intelligence',
 'History of artificial intelligence',
 'Swarm intelligence',
 'Ethics of artificial intelligence',
 'Philosophy of artificial intelligence',
 'Artificial intelligence in healthcare']
```

In [27]:

```
ai = wikipedia.page(wikipedia.search('Artificial Intelligence')[0])
```

In [28]:

```
ai
```

Out[28]:

```
<WikipediaPage 'Artificial intelligence'>
```

In [29]:

```
dir(ai)
```

Out[29]:

```
['_WikipediaPage__continued_query',
 '_WikipediaPage__load',
 '_WikipediaPage__title_query_param',
 '__class__',
 '__delattr__',
 '__dict__',
 '__dir__',
 '__doc__',
 '__eq__',
 '__format__',
 '__ge__',
 '__getattribute__',
 '__gt__',
 '__hash__',
 '__init__',
 '__init_subclass__',
 '__le__',
 '__lt__',
 '__module__',
 '__ne__',
 '__new__',
 '__reduce__',
 '__reduce_ex__',
 '__repr__',
 '__setattr__',
 '__sizeof__',
 '__str__',
 '__subclasshook__',
 '__weakref__',
 'categories',
 'content',
 'coordinates',
 'html',
 'images',
 'links',
 'original_title',
 'pageid',
 'parent_id',
 'references',
 'revision_id',
 'section',
 'sections',
 'summary',
 'title',
 'url']
```

In [30]:

```
ai.title
```

Out[30]:

```
'Artificial intelligence'
```

In [31]:

```
ai.summary
```

Out[31]:

```
'Artificial intelligence (AI) is intelligence—perceiving, synthesizing, and inferring information—demonstrated by m
achines, as opposed to intelligence displayed by animals and humans. Example tasks in which this is done include sp
eech recognition, computer vision, translation between (natural) languages, as well as other mappings of inputs. Th
e Oxford English Dictionary of Oxford University Press defines artificial intelligence as:\n\nthe theory and develo
pment of computer systems able to perform tasks that normally require human intelligence, such as visual perceptio
n, speech recognition, decision-making, and translation between languages.\n\nAI applications include advanced web
search engines (e.g., Google), recommendation systems (used by YouTube, Amazon and Netflix), understanding human sp
eech (such as Siri and Alexa), self-driving cars (e.g., Waymo), automated decision-making and competing at the high
est level in strategic game systems (such as chess and Go).\nAs machines become increasingly capable, tasks conside
red to require "intelligence" are often removed from the definition of AI, a phenomenon known as the AI effect. For
instance, optical character recognition is frequently excluded from things considered to be AI, having become a rou
tine technology.Artificial intelligence was founded as an academic discipline in 1956, and in the years since has e
xperienced several waves of optimism, followed by disappointment and the loss of funding (known as an "AI winter"),
followed by new approaches, success and renewed funding. AI research has tried and discarded many different approac
hes since its founding, including simulating the brain, modeling human problem solving, formal logic, large databas
es of knowledge and imitating animal behavior. In the first decades of the 21st century, highly mathematical-statis
tical machine learning has dominated the field, and this technique has proved highly successful, helping to solve m
any challenging problems throughout industry and academia.The various sub-fields of AI research are centered around
particular goals and the use of particular tools. The traditional goals of AI research include reasoning, knowledge
representation, planning, learning, natural language processing, perception, and the ability to move and manipulate
objects. General intelligence (the ability to solve an arbitrary problem) is among the field\'s long-term goals. To
solve these problems, AI researchers have adapted and integrated a wide range of problem-solving techniques – inclu
ding search and mathematical optimization, formal logic, artificial neural networks, and methods based on statistic
s, probability and economics. AI also draws upon computer science, psychology, linguistics, philosophy, and many ot
her fields.\nThe field was founded on the assumption that human intelligence "can be so precisely described that a
machine can be made to simulate it".\nThis raised philosophical arguments about the mind and the ethical consequenc
es of creating artificial beings endowed with human-like intelligence; these issues have previously been explored b
y myth, fiction and philosophy since antiquity. Computer scientists and philosophers have since suggested that AI m
ay become an existential risk to humanity if its rational capacities are not steered towards beneficial goals.'
```

In [32]:

```
ai.url
```

Out[32]:

```
'https://en.wikipedia.org/wiki/Artificial_intelligence'
```

In [33]:

```
ai.images
```

Out[33]:

```
['https://upload.wikimedia.org/wikipedia/commons/f/f3/AI_Patent_families_for_functional_application_categories_and_
sub_categories.png',
 'https://upload.wikimedia.org/wikipedia/commons/7/7a/Anatomy-1751201_1280.png',
 'https://upload.wikimedia.org/wikipedia/commons/e/e4/Artificial_neural_network.svg',
 'https://upload.wikimedia.org/wikipedia/commons/8/87/Capek_play.jpg',
 'https://upload.wikimedia.org/wikipedia/commons/5/51/Computer_Retro.svg',
 'https://upload.wikimedia.org/wikipedia/commons/2/26/Deep_Learning.jpg',
 'https://upload.wikimedia.org/wikipedia/commons/c/c8/Didrachm_Phaistos_obverse_CdM.jpg',
 'https://upload.wikimedia.org/wikipedia/commons/6/69/EM_Clustering_of_Old_Faithful_data.gif',
 'https://upload.wikimedia.org/wikipedia/commons/9/9e/GFO_taxonomy_tree.png',
 'https://upload.wikimedia.org/wikipedia/commons/4/41/Global_thinking.svg',
 'https://upload.wikimedia.org/wikipedia/commons/1/13/Joseph_Ayerle_portrait_of_Ornella_Muti_%28detail%29%2C_calcul
ated_by_Artificial_Intelligence_%28AI%29_technology.jpg',
 'https://upload.wikimedia.org/wikipedia/commons/b/b8/Kismet_robot_at_MIT_Museum.jpg',
 'https://upload.wikimedia.org/wikipedia/commons/8/8b/Nuvola_apps_kalzium.svg',
 'https://upload.wikimedia.org/wikipedia/commons/6/6e/ParseTree.svg',
 'https://upload.wikimedia.org/wikipedia/commons/e/ec/ParticleSwarmArrowsAnimation.gif',
 'https://upload.wikimedia.org/wikipedia/commons/3/37/People_icon.svg',
 'https://upload.wikimedia.org/wikipedia/commons/c/c5/Shadow_Hand_Bulb_large.jpg',
 'https://upload.wikimedia.org/wikipedia/commons/b/b0/Smart_traffic_light.webp',
 'https://upload.wikimedia.org/wikipedia/commons/f/fc/Smart_traffic_light_intersection.webp',
 'https://upload.wikimedia.org/wikipedia/commons/c/cd/Socrates.png',
 'https://upload.wikimedia.org/wikipedia/commons/f/fa/Wikibooks-logo.svg',
 'https://upload.wikimedia.org/wikipedia/commons/f/ff/Wikidata-logo.svg',
 'https://upload.wikimedia.org/wikipedia/commons/f/fa/Wikiquote-logo.svg',
 'https://upload.wikimedia.org/wikipedia/commons/0/0b/Wikiversity_logo_2017.svg',
 'https://upload.wikimedia.org/wikipedia/commons/2/20/%C3%84%C3%A4retuvastuse_n%C3%A4ide.png',
 'https://upload.wikimedia.org/wikipedia/en/4/4a/Commons-logo.svg',
 'https://upload.wikimedia.org/wikipedia/en/8/8a/OOjs_UI_icon_edit-ltr-progressive.svg',
 'https://upload.wikimedia.org/wikipedia/en/1/1b/Semi-protection-shackle.svg',
 'https://upload.wikimedia.org/wikipedia/en/9/96/Symbol_category_class.svg',
 'https://upload.wikimedia.org/wikipedia/en/d/db/Symbol_list_class.svg',
 'https://upload.wikimedia.org/wikipedia/en/e/e2/Symbol_portal_class.svg',
 'https://upload.wikimedia.org/wikipedia/en/0/06/Wiktionary-logo-v2.svg']
```

In [34]:

```
data = requests.get("https://www.amazon.in/?&ext_vrnc=hi&tag=googhydrabk1-21&ref=pd_sl_96612yg6jw_e&adgrpid=60571832564&hvpone=&h
```

In [35]:

```
data
```

Out[35]:

```
<Response [200]>
```

In [36]:

```
amazon = BeautifulSoup(data.text,'html.parser')
```

In [37]:

```
amazon
```

Out[37]:

```
<!DOCTYPE html>
<html class="a-no-js" data-19ax5a9jf="dingo" lang="en-in"><!-- sp:feature:head-start -->
<head><script>var aPageStart = (new Date()).getTime();</script><meta charset="utf-8"/>
<!-- sp:end-feature:head-start -->
<!-- sp:feature:csm:head-open-part1 -->
<!-- sp:end-feature:csm:head-open-part1 -->
<!-- sp:feature:cs-optimization -->
<meta content="on" http-equiv="x-dns-prefetch-control"/>
<link href="https://images-eu.ssl-images-amazon.com" rel="dns-prefetch"/>
<link href="https://m.media-amazon.com" rel="dns-prefetch"/>
<link href="https://completion.amazon.com" rel="dns-prefetch"/>
<!-- sp:end-feature:cs-optimization -->
<!-- sp:feature:csm:head-open-part2 -->
<!-- sp:end-feature:csm:head-open-part2 -->
<!-- sp:feature:aui-assets -->
<link href="https://m.media-amazon.com/images/I/11EIQ5IGqaL._RC|01ZTHTZObnL.css,41C-I1lXVwL.css,31ufSReDtSL.css,
013z33uKh2L.css,017DsKjNQJL.css,0131vqwP5UL.css,41EWOOlBJ9L.css,11TIuySqr6L.css,01ElnPiDxWL.css,11Qjwq-j69L.css,
01Dm5eKVxwL.css,01IdKcBuAdL.css,01v-XAlI+2L.css,21P6CS3L9LL.css,01oDR3IULNL.css,41cg-OrH1jL.css,01XPHJk60-L.css,
```

In [38]:

```
title = amazon.select('title')
```

In [39]:

```
title
```

Out[39]:

```
[<title>Online Shopping site in India: Shop Online for Mobiles, Books, Watches, Shoes and More - Amazon.in</title>]
```

In [ ]: