> # Data Gathering

*<b>Sources of Data</b> A vast amount of historical data can be found in files such as:

- MS Word documents
- Emails
- Spreadsheets
- MS PowerPoints
- PDFs
- HTML
- and plaintext files

Public and Private Archives CSV, JSON, and XML files use plaintext, a common format, and are compatible with a wide range of applications

The Web can be mined for data using a web scraping application

The IoT uses sensors create data

Sensors in smartphones, cars, airplanes, street lamps, and home appliances capture raw data

Open Data and Private Data

1. Open Data

> The Open Knowledge Foundation describes Open Data as ”any content, information or data that people are free to use, reuse, and redistribute without any legal,

technological, or social restriction.”

2. Private Data

> Data related to an expectation of privacy and regulated by a particular country/government

Structured and Unstructured Data

1. Structured Data

> Data entered and maintained in fixed fields within a file or record Easily entered, classified, queried, and analyzed Relational databases or spreadsheets

2. Unstructured Data Lacks organization

> Raw data Photo contents, audio, video, web pages, blogs, books, journals, white papers, PowerPoint presentations, articles, email, wikis, word processing

documents, and text in general

Example of gathering image data using webcam

Note: Run this snippet using local jupyter notebook

In [7]:
```
!pip install opencv-python
```

```
Requirement already satisfied: opencv-python in c:\users\jens liam vista\anaconda3_\lib\site-packages (4.9.0.80)
Requirement already satisfied: numpy>=1.21.2 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from opencv-py
thon) (1.26.4)
```

In [28]:
```
import cv2

key = cv2.waitKey(1)
webcam = cv2.VideoCapture(0)

while True:
    try:
        check, frame = webcam.read()
        cv2.imshow("Capturing", frame)
        key = cv2.waitKey(1)

        if key == ord('s'):
            cv2.imwrite(filename='saved_img.jpg', img=frame)
            webcam.release()
            print("Processing image...")
            img = cv2.imread('saved_img.jpg', cv2.IMREAD_GRAYSCALE)
            img_new = cv2.imshow("Captured Image", img)
            cv2.waitKey(0)
            cv2.destroyAllWindows()
            print("Converting RGB image to grayscale...")
            # No need to convert to grayscale again, as it's already done above
            # gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
            print("Converted RGB image to grayscale...")
            print("Resizing image to 28x28 scale...")
            # Resize the already grayscale image
            img_resized = cv2.resize(img, (720, 720))
            print("Resized...")
            cv2.imwrite(filename='saved_img-final.jpg', img=img_resized)
            print("Image saved!")
            break

        elif key == ord('q'):
            print("Turning off camera.")
            webcam.release()
            print("Camera off.")
            print("Program ended.")
            cv2.destroyAllWindows()
            break

    except KeyboardInterrupt:
        print("Turning off camera.")
        webcam.release()
        print("Camera off.")
        print("Program ended.")
        cv2.destroyAllWindows()
        break
```

```
Processing image...
Converting RGB image to grayscale...
Converted RGB image to grayscale...
Resizing image to 28x28 scale...
Resized...
Image saved!
```

Example of gathering voice data using microphone

Note: Run the snippet of codes using local jupyter notebook

In [30]:
```
!pip3 install sounddevice
!pip3 install wavio
!pip3 install scipy
```

```
Requirement already satisfied: sounddevice in c:\users\jens liam vista\anaconda3_\lib\site-packages (0.4.6)
Requirement already satisfied: CFFI>=1.0 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from sounddevice)
(1.16.0)
Requirement already satisfied: pycparser in c:\users\jens liam vista\anaconda3_\lib\site-packages (from CFFI>=1.0->so
unddevice) (2.21)
Requirement already satisfied: wavio in c:\users\jens liam vista\anaconda3_\lib\site-packages (0.0.8)
Requirement already satisfied: numpy>=1.19.0 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from wavio)
(1.26.4)
Requirement already satisfied: scipy in c:\users\jens liam vista\anaconda3_\lib\site-packages (1.11.4)
Requirement already satisfied: numpy<1.28.0,>=1.21.6 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from s
cipy) (1.26.4)
```

In [36]:
```
# import required libraries
import sounddevice as sd
from scipy.io.wavfile import write
import wavio as wv

# Sampling frequency
freq = 44100

# Recording duration
duration = 5

# Start recorder with the given values
# of duration and sample frequency
recording = sd.rec(int(duration * freq),
 samplerate=freq, channels=2, device=1)

# Record audio for the given number of seconds
sd.wait()

# This will convert the NumPy array to an audio
# file with the given sampling frequency
write("recording0.wav", freq, recording)
```

```
# Convert the NumPy array to audio file
wv.write("recording1.wav", recording, freq, sampwidth=2)
```

Web Scraping

** Web scraping, web harvesting, or web data extractio** is data scraping used for extracting data from websites. The web scraping software may directly access the World Wide Web using the Hypertext Transfer Protocol or a web browser. While web scraping can be done manually by a software user, the term typically refers to automated processes implemented using a bot or web crawler. It is a form of copying in which specific data is gathered and copied from the web, typically into a central local database or spreadsheet, for later retrieval or analyis.

Image Scraping using BeautifulSoup and Request

In [38]:
```
!pip install bs4
!pip install requests
```

```
Collecting bs4
  Downloading bs4-0.0.2-py2.py3-none-any.whl.metadata (411 bytes)
Requirement already satisfied: beautifulsoup4 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from bs4) (4.
12.2)
Requirement already satisfied: soupsieve>1.2 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from beautiful
soup4->bs4) (2.5)
Downloading bs4-0.0.2-py2.py3-none-any.whl (1.2 kB)
Installing collected packages: bs4
Successfully installed bs4-0.0.2
Requirement already satisfied: requests in c:\users\jens liam vista\anaconda3_\lib\site-packages (2.31.0)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\jens liam vista\anaconda3_\lib\site-packages (fro
m requests) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from requests)
(3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from requ
ests) (2.0.7)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from requ
ests) (2024.2.2)
```

In [39]:
```
import requests
from bs4 import BeautifulSoup
```

```
def getdata(url):
 r = requests.get(url)
 return r.text

htmldata = getdata("https://www.google.com/")
soup = BeautifulSoup(htmldata, 'html.parser')
for item in soup.find_all('img'):
 print(item['src'])
```

```
/images/branding/googlelogo/1x/googlelogo_white_background_color_272x92dp.png
```

In [42]:
```
!pip install selenium
```

```
Requirement already satisfied: selenium in c:\users\jens liam vista\anaconda3_\lib\site-packages (4.18.1)
Requirement already satisfied: urllib3<3,>=1.26 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from urllib
3[socks]<3,>=1.26->selenium) (2.0.7)
Requirement already satisfied: trio~=0.17 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from selenium)
(0.25.0)
Requirement already satisfied: trio-websocket~=0.9 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from sel
enium) (0.11.1)
Requirement already satisfied: certifi>=2021.10.8 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from sele
nium) (2024.2.2)
Requirement already satisfied: typing_extensions>=4.9.0 in c:\users\jens liam vista\anaconda3_\lib\site-packages (fro
m selenium) (4.9.0)
Requirement already satisfied: attrs>=23.2.0 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio~=0.1
7->selenium) (23.2.0)
Requirement already satisfied: sortedcontainers in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio~=
0.17->selenium) (2.4.0)
Requirement already satisfied: idna in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio~=0.17->seleni
um) (3.4)
Requirement already satisfied: outcome in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio~=0.17->sel
enium) (1.3.0.post0)
Requirement already satisfied: sniffio>=1.3.0 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio~=0.
17->selenium) (1.3.0)
Requirement already satisfied: cffi>=1.14 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio~=0.17->
selenium) (1.16.0)
Requirement already satisfied: wsproto>=0.14 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio-webs
ocket~=0.9->selenium) (1.2.0)
Requirement already satisfied: pysocks!=1.5.7,<2.0,>=1.5.6 in c:\users\jens liam vista\anaconda3_\lib\site-packages
(from urllib3[socks]<3,>=1.26->selenium) (1.7.1)
Requirement already satisfied: pycparser in c:\users\jens liam vista\anaconda3_\lib\site-packages (from cffi>=1.14->t
rio~=0.17->selenium) (2.21)
Requirement already satisfied: h11<1,>=0.9.0 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from wsproto>=
0.14->trio-websocket~=0.9->selenium) (0.14.0)
```

In [48]:
```
!pip install selenium
import sys
sys.path.insert(0,'/usr/lib/chromium-browser/chromedriver')


from selenium import webdriver
from selenium.webdriver.common.by import By
import time
import requests
import shutil
import os
import getpass
import urllib.request
import io
import time
from PIL import Image
user = getpass.getuser()
chrome_options = webdriver.ChromeOptions()
chrome_options.add_argument('--headless')
chrome_options.add_argument('--no-sandbox')
chrome_options.add_argument('--disable-dev-shm-usage')
driver = webdriver.Chrome()
def scroll_to_end(driver):
    driver.execute_script("window.scrollTo(0, document.body.scrollHeight);")
    time.sleep(5)#sleep_between_interactions


def getImageUrls(name,totalImgs,driver):
    search_url = "https://www.google.com/search?q=cat&tbm=isch&ved=2ahUKEwjNn_Gn7YyFAxU3yDgGHQYQCesQ2-cCegQIABAA&oq=c
    driver.get(search_url)
    img_urls = set()
    img_count = 0
    results_start = 0

    while(img_count+results_start<totalImgs): #Extract actual images now
        scroll_to_end(driver)
        totalResults = driver.find_elements(By.CLASS_NAME,"Q4LuWd")
        print('total results:', len(totalResults))
        print(f"Found: {totalResults} search results. Extracting links from{results_start}:{totalResults}")
        for img in totalResults[results_start:totalImgs]:
            img.click()
            time.sleep(5)
            image = driver.find_element(By.CLASS_NAME,'iPVvYb')
            img_urls.add(image.get_attribute('src'))
            print(img_urls)
            img_count=len(img_urls)
            print(img_count)

    return img_urls

def downloadImages(folder_path,file_name,url):
    try:
        image_content = requests.get(url).content
    except Exception as e:
        print(f"ERROR - COULD NOT DOWNLOAD {url} - {e}")
    try:
        image_file = io.BytesIO(image_content)
        image = Image.open(image_file).convert('RGB')
        file_path = os.path.join(folder_path, file_name)
        with open(file_path, 'wb') as f:
            image.save(f, "JPEG", quality=85)
        print(f"SAVED - {url} - AT: {file_path}")
    except Exception as e:
        print(f"ERROR - COULD NOT SAVE {url} - {e}")

def saveInDestFolder(searchNames,destDir,totalImgs,driver):
    for name in list(searchNames):
        path=os.path.join(destDir,name)
        if not os.path.isdir(path):
            os.mkdir(path)
        print('Current Path',path)
        totalLinks=getImageUrls(name,totalImgs,driver)
        print('totalLinks',totalLinks)

    if totalLinks is None:
        print('images not found for :',name)

    else:
        for i, link in enumerate(totalLinks):
            file_name = f"{i:150}.jpg"
            downloadImages(path,file_name,link)

searchNames=['cat']
destDir=f'C:/Users/Jens Liam Vista/DATA SCIENCE'
totalImgs=5

saveInDestFolder(searchNames,destDir,totalImgs,driver)
```

```
Requirement already satisfied: selenium in c:\users\jens liam vista\anaconda3_\lib\site-packages (4.18.1)
Requirement already satisfied: urllib3<3,>=1.26 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from urllib
3[socks]<3,>=1.26->selenium) (2.0.7)
Requirement already satisfied: trio~=0.17 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from selenium)
(0.25.0)
Requirement already satisfied: trio-websocket~=0.9 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from sel
enium) (0.11.1)
Requirement already satisfied: certifi>=2021.10.8 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from sele
nium) (2024.2.2)
Requirement already satisfied: typing_extensions>=4.9.0 in c:\users\jens liam vista\anaconda3_\lib\site-packages (fro
m selenium) (4.9.0)
Requirement already satisfied: attrs>=23.2.0 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio~=0.1
7->selenium) (23.2.0)
Requirement already satisfied: sortedcontainers in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio~=
0.17->selenium) (2.4.0)
Requirement already satisfied: idna in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio~=0.17->seleni
um) (3.4)
Requirement already satisfied: outcome in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio~=0.17->sel
enium) (1.3.0.post0)
Requirement already satisfied: sniffio>=1.3.0 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio~=0.
17->selenium) (1.3.0)
Requirement already satisfied: cffi>=1.14 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio~=0.17->
selenium) (1.16.0)
Requirement already satisfied: wsproto>=0.14 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from trio-webs
ocket~=0.9->selenium) (1.2.0)
Requirement already satisfied: pysocks!=1.5.7,<2.0,>=1.5.6 in c:\users\jens liam vista\anaconda3_\lib\site-packages
(from urllib3[socks]<3,>=1.26->selenium) (1.7.1)
Requirement already satisfied: pycparser in c:\users\jens liam vista\anaconda3_\lib\site-packages (from cffi>=1.14->t
rio~=0.17->selenium) (2.21)
Requirement already satisfied: h11<1,>=0.9.0 in c:\users\jens liam vista\anaconda3_\lib\site-packages (from wsproto>=
0.14->trio-websocket~=0.9->selenium) (0.14.0)
Current Path C:/Users/Jens Liam Vista/DATA SCIENCE\cat
total results: 100
Found: [<selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055E
ABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.11")>, <selenium.webdriver.remote.webelement.WebEle
ment (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4
AD0A62090A6.e.13")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", el
ement="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.15")>, <selenium.webdriver.remote.webe
lement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4C
DD6233EC1C2F4AD0A62090A6.e.17")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5
bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.19")>, <selenium.webdrive
r.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9
AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.23")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6
f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.25")>, <sele
nium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F
82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.27")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.29")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055
EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.4")>, <selenium.webdriver.remote.webelement.WebEle
ment (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4
AD0A62090A6.e.32")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", el
ement="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.41")>, <selenium.webdriver.remote.webe
lement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4C
DD6233EC1C2F4AD0A62090A6.e.42")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5
bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.43")>, <selenium.webdrive
r.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9
AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.44")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6
f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.45")>, <sele
nium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F
82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.46")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.47")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055
EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.48")>, <selenium.webdriver.remote.webelement.WebEl
ement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F
4AD0A62090A6.e.49")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", e
lement="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.50")>, <selenium.webdriver.remote.web
element.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4
CDD6233EC1C2F4AD0A62090A6.e.52")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b
5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.54")>, <selenium.webdriv
er.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C
9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.56")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d
6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.58")>, <sel
enium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2
F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.60")>, <selenium.webdriver.remote.webelement.WebElement (sessio
n="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.62")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055
EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.64")>, <selenium.webdriver.remote.webelement.WebEl
ement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F
4AD0A62090A6.e.66")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", e
lement="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.68")>, <selenium.webdriver.remote.web
element.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4
CDD6233EC1C2F4AD0A62090A6.e.70")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b
5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.72")>, <selenium.webdriv
er.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C
9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.74")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d
6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.76")>, <sel
enium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2
F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.78")>, <selenium.webdriver.remote.webelement.WebElement (sessio
n="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.80")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055
EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.82")>, <selenium.webdriver.remote.webelement.WebEl
ement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F
4AD0A62090A6.e.84")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", e
lement="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.86")>, <selenium.webdriver.remote.web
element.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4
CDD6233EC1C2F4AD0A62090A6.e.88")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b
5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.90")>, <selenium.webdriv
er.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C
9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.40")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d
6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.93")>, <sel
enium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2
F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.95")>, <selenium.webdriver.remote.webelement.WebElement (sessio
n="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.97")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055
EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.99")>, <selenium.webdriver.remote.webelement.WebEl
ement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F
4AD0A62090A6.e.101")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786",
element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.103")>, <selenium.webdriver.remote.w
ebelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788
F4CDD6233EC1C2F4AD0A62090A6.e.105")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3
a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.107")>, <selenium.web
driver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4
E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.109")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd
5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.111")
>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6C
ED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.113")>, <selenium.webdriver.remote.webelement.WebElement
(session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A6
2090A6.e.115")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", elemen
t="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.117")>, <selenium.webdriver.remote.webelem
ent.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6
233EC1C2F4AD0A62090A6.e.119")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb
6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.121")>, <selenium.webdrive
r.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9
AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.123")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d
6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.125")>, <se
lenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B
2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.127")>, <selenium.webdriver.remote.webelement.WebElement (sess
ion="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A
6.e.129")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.
055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.131")>, <selenium.webdriver.remote.webelement.W
ebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC
1C2F4AD0A62090A6.e.133")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a178
6", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.135")>, <selenium.webdriver.remo
te.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9
A788F4CDD6233EC1C2F4AD0A62090A6.e.137")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30
a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.139")>, <seleniu
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.141")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.143")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.05
5EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.145")>, <selenium.webdriver.remote.webelement.Web
Element (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C
2F4AD0A62090A6.e.147")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a178
6", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.149")>, <selenium.webdriver.remo
te.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9
A788F4CDD6233EC1C2F4AD0A62090A6.e.151")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30
a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.153")>, <seleniu
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.155")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.157")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.05
5EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.159")>, <selenium.webdriver.remote.webelement.Web
Element (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C
2F4AD0A62090A6.e.161")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a178
6", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.163")>, <selenium.webdriver.remo
te.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9
A788F4CDD6233EC1C2F4AD0A62090A6.e.165")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30
a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.167")>, <seleniu
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.169")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.171")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.05
5EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.173")>, <selenium.webdriver.remote.webelement.Web
Element (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C
2F4AD0A62090A6.e.175")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a178
6", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.177")>, <selenium.webdriver.remo
te.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9
A788F4CDD6233EC1C2F4AD0A62090A6.e.179")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30
```

```
a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.181")>, <seleniu
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.183")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.185")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.05
5EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.187")>, <selenium.webdriver.remote.webelement.Web
Element (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C
2F4AD0A62090A6.e.189")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a178
6", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.191")>, <selenium.webdriver.remo
te.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9
A788F4CDD6233EC1C2F4AD0A62090A6.e.193")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30
a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.195")>, <seleniu
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.197")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.199")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.05
5EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.201")>, <selenium.webdriver.remote.webelement.Web
Element (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C
2F4AD0A62090A6.e.203")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a178
6", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.205")>, <selenium.webdriver.remo
te.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9
A788F4CDD6233EC1C2F4AD0A62090A6.e.207")>] search results. Extracting links from0:[<selenium.webdriver.remote.webeleme
nt.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD62
33EC1C2F4AD0A62090A6.e.11")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a
1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.13")>, <selenium.webdriver.re
mote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.
d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.15")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f
30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.17")>, <seleniu
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.19")>, <selenium.webdriver.remote.webelement.WebElement (session="d
5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.2
3")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EAB
D6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.25")>, <selenium.webdriver.remote.webelement.WebEleme
nt (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD
0A62090A6.e.27")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", elem
ent="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.29")>, <selenium.webdriver.remote.webele
ment.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD
6233EC1C2F4AD0A62090A6.e.4")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6
a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.32")>, <selenium.webdriver.r
emote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.
d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.41")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f
30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.42")>, <seleniu
```

```
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.43")>, <selenium.webdriver.remote.webelement.WebElement (session="d
5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.4
4")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EAB
D6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.45")>, <selenium.webdriver.remote.webelement.WebEleme
nt (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD
0A62090A6.e.46")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", elem
ent="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.47")>, <selenium.webdriver.remote.webele
ment.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD
6233EC1C2F4AD0A62090A6.e.48")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb
6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.49")>, <selenium.webdriver.
remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9A
C.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.50")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f
9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.52")>, <selen
ium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F8
2E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.54")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.56")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055
EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.58")>, <selenium.webdriver.remote.webelement.WebEl
ement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F
4AD0A62090A6.e.60")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", e
lement="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.62")>, <selenium.webdriver.remote.web
element.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4
CDD6233EC1C2F4AD0A62090A6.e.64")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b
5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.66")>, <selenium.webdriv
er.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C
9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.68")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d
6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.70")>, <sel
enium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2
F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.72")>, <selenium.webdriver.remote.webelement.WebElement (sessio
n="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.74")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055
EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.76")>, <selenium.webdriver.remote.webelement.WebEl
ement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F
4AD0A62090A6.e.78")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", e
lement="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.80")>, <selenium.webdriver.remote.web
element.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4
CDD6233EC1C2F4AD0A62090A6.e.82")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b
5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.84")>, <selenium.webdriv
er.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C
9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.86")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d
6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.88")>, <sel
```

```
enium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2
F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.90")>, <selenium.webdriver.remote.webelement.WebElement (sessio
n="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.40")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055
EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.93")>, <selenium.webdriver.remote.webelement.WebEl
ement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F
4AD0A62090A6.e.95")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", e
lement="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.97")>, <selenium.webdriver.remote.web
element.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4
CDD6233EC1C2F4AD0A62090A6.e.99")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b
5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.101")>, <selenium.webdri
ver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99
C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.103")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed
4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.105")>, <
selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C
1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.107")>, <selenium.webdriver.remote.webelement.WebElement (se
ssion="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A6209
0A6.e.109")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element
="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.111")>, <selenium.webdriver.remote.webeleme
nt.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD62
33EC1C2F4AD0A62090A6.e.113")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6
a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.115")>, <selenium.webdriver.
remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9A
C.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.117")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6
f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.119")>, <sel
enium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2
F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.121")>, <selenium.webdriver.remote.webelement.WebElement (sessi
on="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A
6.e.123")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.
055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.125")>, <selenium.webdriver.remote.webelement.W
ebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC
1C2F4AD0A62090A6.e.127")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a178
6", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.129")>, <selenium.webdriver.remo
te.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9
A788F4CDD6233EC1C2F4AD0A62090A6.e.131")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30
a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.133")>, <seleniu
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.135")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.137")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.05
5EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.139")>, <selenium.webdriver.remote.webelement.Web
Element (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C
2F4AD0A62090A6.e.141")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a178
6", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.143")>, <selenium.webdriver.remo
te.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9
A788F4CDD6233EC1C2F4AD0A62090A6.e.145")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30
a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.147")>, <seleniu
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.149")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.151")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.05
5EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.153")>, <selenium.webdriver.remote.webelement.Web
Element (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C
2F4AD0A62090A6.e.155")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a178
6", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.157")>, <selenium.webdriver.remo
te.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9
A788F4CDD6233EC1C2F4AD0A62090A6.e.159")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30
a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.161")>, <seleniu
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.163")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.165")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.05
5EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.167")>, <selenium.webdriver.remote.webelement.Web
Element (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C
2F4AD0A62090A6.e.169")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a178
6", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.171")>, <selenium.webdriver.remo
te.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9
A788F4CDD6233EC1C2F4AD0A62090A6.e.173")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30
a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.175")>, <seleniu
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.177")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.179")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.05
5EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.181")>, <selenium.webdriver.remote.webelement.Web
Element (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C
2F4AD0A62090A6.e.183")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a178
6", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.185")>, <selenium.webdriver.remo
te.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9
A788F4CDD6233EC1C2F4AD0A62090A6.e.187")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30
a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.189")>, <seleniu
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.191")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.193")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.05
```

```
5EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.195")>, <selenium.webdriver.remote.webelement.Web
Element (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C
2F4AD0A62090A6.e.197")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a178
6", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.199")>, <selenium.webdriver.remo
te.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9
A788F4CDD6233EC1C2F4AD0A62090A6.e.201")>, <selenium.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30
a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.203")>, <seleniu
m.webdriver.remote.webelement.WebElement (session="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E
50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.e.205")>, <selenium.webdriver.remote.webelement.WebElement (session
="d5abd5ed4d6f9f30a5d3a6b5bb6a1786", element="f.055EABD6CED7C1B2F82E50DD4E99C9AC.d.9A788F4CDD6233EC1C2F4AD0A62090A6.
e.207")>]
{'https://i.natgeofe.com/n/548467d8-c5f1-4551-9f58-6817a8d2c45e/NationalGeographic_2572187_square.jpg'}
1
{'https://cdn.britannica.com/70/234870-050-D4D024BB/Orange-colored-cat-yawns-displaying-teeth.jpg', 'https://i.natgeo
fe.com/n/548467d8-c5f1-4551-9f58-6817a8d2c45e/NationalGeographic_2572187_square.jpg'}
2
{'https://upload.wikimedia.org/wikipedia/commons/thumb/1/15/Cat_August_2010-4.jpg/1200px-Cat_August_2010-4.jpg', 'htt
ps://cdn.britannica.com/70/234870-050-D4D024BB/Orange-colored-cat-yawns-displaying-teeth.jpg', 'https://i.natgeofe.co
m/n/548467d8-c5f1-4551-9f58-6817a8d2c45e/NationalGeographic_2572187_square.jpg'}
3
{'https://upload.wikimedia.org/wikipedia/commons/thumb/1/15/Cat_August_2010-4.jpg/1200px-Cat_August_2010-4.jpg', 'htt
ps://cdn.britannica.com/34/235834-050-C5843610/two-different-breeds-of-cats-side-by-side-outdoors-in-the-garden.jpg',
'https://cdn.britannica.com/70/234870-050-D4D024BB/Orange-colored-cat-yawns-displaying-teeth.jpg', 'https://i.natgeof
e.com/n/548467d8-c5f1-4551-9f58-6817a8d2c45e/NationalGeographic_2572187_square.jpg'}
4
{'https://upload.wikimedia.org/wikipedia/commons/thumb/1/15/Cat_August_2010-4.jpg/1200px-Cat_August_2010-4.jpg', 'htt
ps://cdn.britannica.com/70/234870-050-D4D024BB/Orange-colored-cat-yawns-displaying-teeth.jpg', 'https://i.natgeofe.co
m/n/548467d8-c5f1-4551-9f58-6817a8d2c45e/NationalGeographic_2572187_square.jpg', 'https://cdn.britannica.com/34/23583
4-050-C5843610/two-different-breeds-of-cats-side-by-side-outdoors-in-the-garden.jpg', 'https://media.4-paws.org/5/b/
4/b/5b4b5a91dd9443fa1785ee7fca66850e06dcc7f9/VIER%20PFOTEN_2019-12-13_209-2890x2000-1920x1329.jpg'}
5
totalLinks {'https://upload.wikimedia.org/wikipedia/commons/thumb/1/15/Cat_August_2010-4.jpg/1200px-Cat_August_2010-
4.jpg', 'https://cdn.britannica.com/70/234870-050-D4D024BB/Orange-colored-cat-yawns-displaying-teeth.jpg', 'https://
i.natgeofe.com/n/548467d8-c5f1-4551-9f58-6817a8d2c45e/NationalGeographic_2572187_square.jpg', 'https://cdn.britannic
a.com/34/235834-050-C5843610/two-different-breeds-of-cats-side-by-side-outdoors-in-the-garden.jpg', 'https://media.4-
paws.org/5/b/4/b/5b4b5a91dd9443fa1785ee7fca66850e06dcc7f9/VIER%20PFOTEN_2019-12-13_209-2890x2000-1920x1329.jpg'}
SAVED - https://upload.wikimedia.org/wikipedia/commons/thumb/1/15/Cat_August_2010-4.jpg/1200px-Cat_August_2010-4.jpg
- AT: C:/Users/Jens Liam Vista/DATA SCIENCE\cat\
0.jpg
SAVED - https://cdn.britannica.com/70/234870-050-D4D024BB/Orange-colored-cat-yawns-displaying-teeth.jpg - AT: C:/User
s/Jens Liam Vista/DATA SCIENCE\cat\
1.jpg
```

```
SAVED - https://i.natgeofe.com/n/548467d8-c5f1-4551-9f58-6817a8d2c45e/NationalGeographic_2572187_square.jpg - AT: C:/
Users/Jens Liam Vista/DATA SCIENCE\cat\
2.jpg
SAVED - https://cdn.britannica.com/34/235834-050-C5843610/two-different-breeds-of-cats-side-by-side-outdoors-in-the-g
arden.jpg - AT: C:/Users/Jens Liam Vista/DATA SCIENCE\cat\
3.jpg
SAVED - https://media.4-paws.org/5/b/4/b/5b4b5a91dd9443fa1785ee7fca66850e06dcc7f9/VIER%20PFOTEN_2019-12-13_209-2890x2
000-1920x1329.jpg - AT: C:/Users/Jens Liam Vista/DATA SCIENCE\cat\
4.jpg
```

> # Web Scraping of Movies Information using BeautifulSoup

We want to analyze the distributions of IMDB and Metacritic movie ratings to see if we find anything interesting. To do this, we'll first scrape data for over 2000 movies.

In the image above, you can see that the URL has several parameters after the question mark:

release_date — Shows only the movies released in a specific year . sort — Sorts the movies on the page. sort=num_votes,desc translates to sort by number of votes in a descending orde

. page — Specifies the page numb

r. ref_ — Takes us to the the next or the previous page. The reference is the page we are currently on. adv_nxt and adv_prv are two possible values. They translate to advance to the next page, and advance to the previous page, respectivel

In [50]:
```python
from requests import get
url = 'https://www.imdb.com/search/title?release_date=2017&sort=num_votes,desc&page=1'
agent = {"User-Agent": 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/122.0
response = get(url, headers = agent)
print(response)
print(response.text[:500])
```

```
<Response [200]>
<!DOCTYPE html><html lang="en-US" xmlns:og="http://opengraphprotocol.org/schema/" xmlns:fb="http://www.facebook.com/2
008/fbml"><head><meta charSet="utf-8"/><meta name="viewport" content="width=device-width"/><script>if(typeof uet ===
'function'){ uet('bb', 'LoadTitle', {wb: 1}); }</script><script>window.addEventListener('load', (event) => {
        if (typeof window.csa !== 'undefined' && typeof window.csa === 'function') {
            var csaLatencyPlugin = window.csa('Content', {
```

Using BeautifulSoup to parse the HTML content

To parse our HTML document and extract the 50 div containers, we'll use a Python module called BeautifulSoup, the most common web scraping module for Python.

In the following code cell we will:

- Import the BeautifulSoup class creator from the package bs4.
- Parse response.text by creating a BeautifulSoup object, and assign this object to html_soup. The 'html.parser' argument indicates that we want to do the parsing using

Python's built-in HTML parser.

In [51]:
```python
from bs4 import BeautifulSoup
soup = BeautifulSoup(response.text, 'html.parser')
headers = {'Accept-Language': 'en-US,en;q=0.8'}
type(soup)
```

Out[51]:
```
bs4.BeautifulSoup
```

Before extracting the 50 div containers, we need to figure out what distinguishes them from other div elements on that page. Often, the distinctive mark resides in the class attribute. If you inspect the HTML lines of the containers of interest, you'll notice that the class attribute has two values: lister-item and mode-advanced. This combination is unique to these div containers. We can see that's true by doing a quick search (Ctrl + F). We have 50 such containers, so we expect to see only 50 matches:

In [52]:
```python
movie_containers = soup.find_all('li', class_ = 'ipc-metadata-list-summary-item')
print(type(movie_containers))
print(len(movie_containers))
```

```
<class 'bs4.element.ResultSet'>
50
```

find_all() returned a ResultSet object which is a list containing all the 50 divs we are interested in.

Now we'll select only the first container, and extract, by turn, each item of interest:

- The name of the movie
- The year of release.
- The IMDB rating.
- The Metascore.
- The number of votes

Extracting the data for a single movie

We can access the first container, which contains information about a single movie, by using list notation on movie_containers.

```
In [92]: first_movie = movie_containers[0]
```

```
In [93]: first_name = first_movie.h3.text
         first_name[3:]
```

```
Out[93]: 'Logan'
```

The year of the movie's release

```
In [55]: first_year = movie_containers[0].find('span', class_ = "sc-b0691f29-8 ilsLEX dli-title-metadata-item")
         first_year
```

```
Out[55]: <span class="sc-b0691f29-8 ilsLEX dli-title-metadata-item">2017</span>
```

```
In [56]: first_year = first_year.text
         first_year
```

```
Out[56]: '2017'
```

The IMDB rating

```
In [58]: first_imdb = movie_containers[0].find('span', class_ = "ipc-rating-star ipc-rating-star--base ipc-rating-star--imdb r
         first_imdb.text[:3]
```

```
Out[58]:  '8.1'
```

The Metascore

```
In [59]:  first_mscore = movie_containers[0].find('span', class_ = 'sc-b0901df4-0 bcQdDJ metacritic-score-box')
          first_mscore = first_mscore.text
          print(first_mscore)

         77
```

The number of votes

```
In [60]:  first_votes = movie_containers[0].find('span', class_ = 'ipc-rating-star--voteCount')
          first_votes.text[2:-1]

Out[60]:  '827K'
```

The script

```
In [95]:  # Lists to store the scraped data in
          names = []
          years = []
          imdb_ratings = []
          metascores = []
          votes = []
          # Extract data from individual movie container
          for container in movie_containers:
            names.append(container.find('h3', class_ = "ipc-title__text").text[3:])
            # print(container.find('h3', class_ = "ipc-title__text").text[3:])
            years.append(container.find('span', class_ = "sc-b0691f29-8 ilsLEX dli-title-metadata-item").text)
            # print(container.find('span', class_ = "sc-b0691f29-8 ilsLEX dli-title-metadata-item").text)
            imdb_ratings.append(container.find('span', class_ = "ipc-rating-star ipc-rating-star--base ipc-rating-star--imdb ra
            # print(container.find('span', class_ = "ipc-rating-star ipc-rating-star--base ipc-rating-star--imdb ratingGroup--i
            rate = container.find('span', class_='sc-b0901df4-0 bcQdDJ metacritic-score-box')
            if rate:
              metascores.append(rate.text)
            else:
              metascores.append(0)
            # print(container.find('span', class_ = 'ipc-rating-star--voteCount').text[2:-1])
            votes.append(container.find('span', class_ = 'ipc-rating-star--voteCount').text[2:-1])
```

```
print(len(names))
print(len(years))
print(len(imdb_ratings))
print(len(metascores))
print(len(votes)
)
```

```
50
50
50
50
50
```

In [96]:

```
import pandas as pd
test_df = pd.DataFrame({'movie': names,
'year': years,
'imdb': imdb_ratings,
'metascore': metascores,
'votes': votes
})
print(test_df.info())
test_df
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 50 entries, 0 to 49
Data columns (total 5 columns):
 #   Column     Non-Null Count  Dtype 
---  ------     --------------  ----- 
 0   movie      50 non-null     object
 1   year       50 non-null     object
 2   imdb       50 non-null     object
 3   metascore  50 non-null     object
 4   votes      50 non-null     object
dtypes: object(5)
memory usage: 2.1+ KB
None
```

Out[96]:

	movie	year	imdb	metascore	votes
0	Logan	2017	8.1	77	827K
1	Thor: Ragnarok	2017	7.9	74	813K
2	Guardians of the Galaxy Vol. 2	2017	7.6	67	756K
3	Dunkirk	2017	7.8	94	736K
4	Spider-Man: Homecoming	2017	7.4	73	716K
5	Wonder Woman	2017	7.3	76	698K
6	Get Out	2017	7.8	85	691K
7	Star Wars: Episode VIII - The Last Jedi	2017	6.9	84	670K
8	Blade Runner 2049	2017	8.0	81	658K
9	Baby Driver	2017	7.5	86	605K
10	It	2017	7.3	69	603K
11	Coco	2017	8.4	81	586K
12	Three Billboards Outside Ebbing, Missouri	2017	8.1	88	553K
13	Money Heist	2017–2021	8.2	0	529K
14	John Wick: Chapter 2	2017	7.4	75	509K
15	Justice League	2017	6.1	45	477K
16	The Shape of Water	2017	7.3	87	446K
17	Dark	2017–2020	8.7	0	440K
18	Jumanji: Welcome to the Jungle	2017	6.9	58	436K
19	Kingsman: The Golden Circle	2017	6.7	44	361K
20	Kong: Skull Island	2017	6.7	62	345K
21	Ozark	2017–2022	8.5	0	344K

	movie	year	imdb	metascore	votes
22	Pirates of the Caribbean: Salazar's Revenge	2017	6.5	39	344K
23	Beauty and the Beast	2017	7.1	65	333K
24	Mindhunter	2017–2019	8.6	0	333K
25	Lady Bird	2017	7.4	93	326K
26	13 Reasons Why	2017–2020	7.5	0	314K
27	Call Me by Your Name	2017	7.8	94	313K
28	The Greatest Showman	2017	7.5	48	310K
29	Alien: Covenant	2017	6.4	65	302K
30	Murder on the Orient Express	2017	6.5	52	295K
31	War for the Planet of the Apes	2017	7.4	82	280K
32	Wind River	2017	7.7	73	279K
33	The Punisher	2017–2019	8.4	0	263K
34	The Handmaid's Tale	2017–	8.4	0	257K
35	Fast & Furious 8	2017	6.6	56	253K
36	Life	2017	6.6	54	252K
37	Mother!	2017	6.6	76	249K
38	The Hitman's Bodyguard	2017	6.9	47	246K
39	I, Tonya	2017	7.5	77	242K
40	King Arthur: Legend of the Sword	2017	6.7	41	232K
41	Ghost in the Shell	2017	6.3	52	227K
42	Big Little Lies	2017–	8.4	0	223K
43	Darkest Hour	2017	7.4	75	220K

	movie	year	imdb	metascore	votes
44	The End of the F***ing World	2017–2019	8.0	0	218K
45	American Made	2017	7.1	65	207K
46	Atomic Blonde	2017	6.7	63	206K
47	The Mummy	2017	5.4	34	206K
48	Baywatch	2017	5.5	37	201K
49	Bright	2017	6.3	29	201K

In [97]:
```
test_df.to_csv('Movies.csv', index=False)
```

In [98]:
```
meow = pd.read_csv('Movies.csv')
meow.head()
```

Out[98]:

	movie	year	imdb	metascore	votes
0	Logan	2017	8.1	77	827K
1	Thor: Ragnarok	2017	7.9	74	813K
2	Guardians of the Galaxy Vol. 2	2017	7.6	67	756K
3	Dunkirk	2017	7.8	94	736K
4	Spider-Man: Homecoming	2017	7.4	73	716K

The script for multiple pages

In [64]:
```
from time import time
from time import sleep
from requests import get
from random import randint
from IPython.core.display import clear_output
from bs4 import BeautifulSoup

from IPython.core.display import clear_output
pages = ['1','2','3','4','5']
```

```
years_url = ['2017', '2018', '2019', '2020']

# Redeclaring the lists to store data in
names = []
years = []
imdb_ratings = []
metascores = []
votes = []

# Preparing the monitoring of the loop
start_time = time()
requests = 0

# For every year in the interval 2000-2017
for year_url in years_url:

    # For every page in the interval 1-4
    for page in pages:

        # Make a get request
        url = f'https://www.imdb.com/search/title/?release_date={year_url}-01-01,{year_url}-12-31&sort=num_votes,desc
        agent = {"User-Agent":"Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrom
        response = get(url,headers = agent)
        print(response.text[:500])
        #response = get('https://www.imdb.com/search/title?release_date=' + year_url +
        #'&sort=num_votes,desc&page=' + page, headers = headers)

        # Pause the loop
        sleep(5)

        # Monitor the requests
        requests += 1
        elapsed_time = time() - start_time
        print('Request:{}; Frequency: {} requests/s'.format(requests, requests/elapsed_time))
        clear_output(wait = True)

        # Throw a warning for non-200 status codes
        if response.status_code != 200:
            print('Request: {}; Status code: {}'.format(requests, response.status_code))

        # Break the loop if the number of requests is greater than expected
        if requests > 72:
            print('Number of requests was greater than expected.')
            break
        # Parse the content of the request with BeautifulSoup
        page_html = BeautifulSoup(response.text, 'html.parser')

        # Select all the 50 movie containers from a single page
        mv_containers = page_html.find_all('div', class_ = 'sc-ab6fa25a-3 bVYfLY dli-parent')

        # For every movie of these 50
        for container in mv_containers:
            # If the movie has a Metascore, then:
            if container.find('span', class_ = 'sc-b0901df4-0 bcQdDJ metacritic-score-box') is not None:
                # Scrape the name
                name = container.find('h3',class_='ipc-title__text').text[3:]
                names.append(name)

                # Scrape the year
                year = container.find('span', class_ = 'sc-b0691f29-8 ilsLEX dli-title-metadata-item').text
                years.append(year)

                # Scrape the IMDB rating
                imdb = container.find('span', class_ = 'ipc-rating-star ipc-rating-star--base ipc-rating-star--imdb r
                imdb_ratings.append(imdb)

                # Scrape the Metascore
                m_score = container.find('span', class_ = 'sc-b0901df4-0 bcQdDJ metacritic-score-box').text
                metascores.append(m_score)

                # Scrape the number of votes
                vote = container.find('span', class_ = 'ipc-rating-star--voteCount').text[2:-1]
                votes.append(vote)
```

```
<!DOCTYPE html><html lang="en-US" xmlns:og="http://opengraphprotocol.org/schema/" xmlns:fb="http://www.facebook.com/2
008/fbml"><head><meta charSet="utf-8"/><meta name="viewport" content="width=device-width"/><script>if(typeof uet ===
'function'){ uet('bb', 'LoadTitle', {wb: 1}); }</script><script>window.addEventListener('load', (event) => {
        if (typeof window.csa !== 'undefined' && typeof window.csa === 'function') {
            var csaLatencyPlugin = window.csa('Content', {

Request:20; Frequency: 0.130015380845493 22 requests/s
```

In [99]:
```
movie_ratings = pd.DataFrame({'movie': names,
'year': years,
'imdb': imdb_ratings,
'metascore': metascores,
'votes': votes
})
print(movie_ratings.info())
movie_ratings.head()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 50 entries, 0 to 49
Data columns (total 5 columns):
 #   Column     Non-Null Count  Dtype 
---  ------     --------------  ----- 
 0   movie      50 non-null     object
 1   year       50 non-null     object
 2   imdb       50 non-null     object
 3   metascore  50 non-null     object
 4   votes      50 non-null     object
dtypes: object(5)
memory usage: 2.1+ KB
None
```

Out[99]:

	movie	year	imdb	metascore	votes
0	Logan	2017	8.1	77	827K
1	Thor: Ragnarok	2017	7.9	74	813K
2	Guardians of the Galaxy Vol. 2	2017	7.6	67	756K
3	Dunkirk	2017	7.8	94	736K
4	Spider-Man: Homecoming	2017	7.4	73	716K

In [66]:
```
movie_ratings.tail(10)
```

Out[66]:

	movie	year	imdb	metascore	votes
775	The Hunt	2020	6.5	50	128K
776	Greyhound	2020	7.0	64	114K
777	Hamilton	2020	8.3	88	112K
778	Eurovision Song Contest: The Story of Fire Saga	2020	6.5	50	102K
779	I'm Thinking of Ending Things	2020	6.6	78	99K
780	Project Power	2020	6.0	51	97K
781	Spenser Confidential	2020	6.2	49	97K
782	Underwater	2020	5.9	48	97K
783	Minari	2020	7.4	89	96K
784	News of the World	2020	6.8	73	95K

In [68]:
```
movie_ratings.to_csv('movie_ratings.csv')
```

> Data Preparation

In [76]:
```
movie_ratings = pd.read_csv('movie_ratings.csv')
movie_ratings.index += 1
```

In [77]:
```
movie_ratings['year'].unique()
```

Out[77]:
```
array([2017, 2018, 2019, 2020], dtype=int64)
```

In [78]:
```
movie_ratings.dtypes
```

```
Out[78]: Unnamed: 0      int64
         movie          object
         year            int64
         imdb          float64
         metascore       int64
         votes          object
         dtype: object
```

```
In [79]: movie_ratings['year'] = movie_ratings['year'].astype(int)
```

```
In [80]: movie_ratings['year'].unique()
```

```
Out[80]: array([2017, 2018, 2019, 2020])
```

```
In [83]: movie_ratings.dtypes
```

```
Out[83]: Unnamed: 0      int64
         movie          object
         year            int32
         imdb          float64
         metascore       int64
         votes          object
         dtype: object
```

```
In [84]: movie_ratings['year'] = movie_ratings['year'].astype(int)
```

```
In [85]: movie_ratings['year'].unique()
```

```
Out[85]: array([2017, 2018, 2019, 2020])
```

```
In [86]: movie_ratings.dtypes
```

```
Out[86]: Unnamed: 0      int64
         movie          object
         year            int32
         imdb          float64
         metascore       int64
         votes          object
         dtype: object
```

```
In [87]: movie_ratings.head(10)
```

Out[87]:

	Unnamed: 0	movie	year	imdb	metascore	votes
1	0	Logan	2017	8.1	77	827K
2	1	Thor: Ragnarok	2017	7.9	74	813K
3	2	Guardians of the Galaxy Vol. 2	2017	7.6	67	756K
4	3	Dunkirk	2017	7.8	94	736K
5	4	Spider-Man: Homecoming	2017	7.4	73	716K
6	5	Wonder Woman	2017	7.3	76	698K
7	6	Get Out	2017	7.8	85	691K
8	7	Star Wars: Episode VIII - The Last Jedi	2017	6.9	84	670K
9	8	Blade Runner 2049	2017	8.0	81	658K
10	9	Baby Driver	2017	7.5	86	605K

```
In [88]: movie_ratings.tail(10)
```

Out[88]:

	Unnamed: 0	movie	year	imdb	metascore	votes
776	775	The Hunt	2020	6.5	50	128K
777	776	Greyhound	2020	7.0	64	114K
778	777	Hamilton	2020	8.3	88	112K
779	778	Eurovision Song Contest: The Story of Fire Saga	2020	6.5	50	102K
780	779	I'm Thinking of Ending Things	2020	6.6	78	99K
781	780	Project Power	2020	6.0	51	97K
782	781	Spenser Confidential	2020	6.2	49	97K
783	782	Underwater	2020	5.9	48	97K
784	783	Minari	2020	7.4	89	96K
785	784	News of the World	2020	6.8	73	95K

In [89]:
```
movie_ratings
```

Out[89]:

	Unnamed: 0	movie	year	imdb	metascore	votes
1	0	Logan	2017	8.1	77	827K
2	1	Thor: Ragnarok	2017	7.9	74	813K
3	2	Guardians of the Galaxy Vol. 2	2017	7.6	67	756K
4	3	Dunkirk	2017	7.8	94	736K
5	4	Spider-Man: Homecoming	2017	7.4	73	716K
...	...	...	...	...	...	...
781	780	Project Power	2020	6.0	51	97K
782	781	Spenser Confidential	2020	6.2	49	97K
783	782	Underwater	2020	5.9	48	97K
784	783	Minari	2020	7.4	89	96K
785	784	News of the World	2020	6.8	73	95K

785 rows × 6 columns

In [ ]: