from bs4 import BeautifulSoup

import requests

import pandas as pd

# 2. find url and store it in a variable

url = "https://www.imdb.com/chart/top"

# 3. download html with a get request

headers = {'Accept-Language': 'en-US,en;q=0.8'}

response = requests.get(url, headers = headers)

response.status\_code # 200 status code means OK!

# 4.1. parse html (create the 'soup')

soup = BeautifulSoup(response.content, "html.parser")

# 4.2. check that the html code looks like it should

soup

print(soup.div.name) tag’s name  
print(soup.div.attrs) all  the attributes of the tag

print(soup.div.string) get the string inside of the tag

for child in soup.ul.children:  
    print(child)

webpage = webpage\_response.content

soup = BeautifulSoup(webpage, "html.parser")

turtle\_links = soup.find\_all("a")

links = []

#go through all of the a tags and get the links associated with them:

for a in turtle\_links:

  links.append(prefix+a["href"])

#Define turtle\_data:

turtle\_data = {}

#follow each link:

for link in links:

  webpage = requests.get(link)

  turtle = BeautifulSoup(webpage.content, "html.parser")

  #Add your code here:

  turtle\_name = turtle.select(".name")[0]

  turtle\_data[turtle\_name] = []

print(turtle\_data)

<https://airportcodes.aero/name/B>