22nd Feb Assignment

April 19, 2023

1 Assignment 21

Go to this given URL and solve the following questions URL: https://www.youtube.com/@PW-Foundation/videos

Q1. Write a python program to extract the video URL of the first five videos.

Ans.

[48]:
```
import requests
from bs4 import BeautifulSoup
```

[49]:
```
# Define the URL of the YouTube channel
url = 'https://www.youtube.com/@PW-Foundation/videos'
```

[50]:
```
# Send a GET request to the URL and get the HTML content
response = requests.get(url)
content = response.content
```

[51]:
```
# Parse the HTML content using BeautifulSoup
soup = BeautifulSoup(content, 'html.parser')
```

[57]:
```
video_links = []
for link in soup.find_all('a', href=True):
    if len(video_links) < 5:
        if(link['href']=='/'):
           video_links.append("https://www.youtube.com"+link['href'])
        else:
            video_links.append(link['href'])

# Print the video links
for video_link in video_links:
    print(video_link)
```

```
https://www.youtube.com/
https://www.youtube.com/
https://www.youtube.com/about/
https://www.youtube.com/about/press/
https://www.youtube.com/about/copyright/
```

Q2. Write a python program to extract the URL of the video thumbnails of the first five videos.

[58]:
```
import requests
from bs4 import BeautifulSoup

# specify the URL of the YouTube channel
url = "https://www.youtube.com/@PW-Foundation/videos"

# send a GET request to the URL and get the HTML content
response = requests.get(url)
html_content = response.content

# parse the HTML content using BeautifulSoup
soup = BeautifulSoup(html_content, "html.parser")

# extract the thumbnail URLs of the first five videos
thumbnail_urls = []
for video in soup.find_all("a", class_="yt-simple-endpoint style-scope␣
 ↪ytd-grid-video-renderer"):
    thumbnail_url = video.find("img")["src"]
    thumbnail_urls.append(thumbnail_url)
    if len(thumbnail_urls) >= 5:
        break

# print the extracted thumbnail URLs
print(thumbnail_urls)
```

```
[]
```

Q3. Write a python program to extract the title of the first five videos.

[59]:
```
import requests
from bs4 import BeautifulSoup

# specify the URL of the YouTube channel
url = "https://www.youtube.com/@PW-Foundation/videos"

# send a GET request to the URL and get the HTML content
response = requests.get(url)
html_content = response.content

# parse the HTML content using BeautifulSoup
soup = BeautifulSoup(html_content, "html.parser")

# extract the titles of the first five videos
video_titles = []
for video in soup.find_all("a", class_="yt-simple-endpoint style-scope␣
 ↪ytd-grid-video-renderer"):
```

```
    title = video.find("yt-formatted-string", class_="style-scope␣
 ↪ytd-grid-video-renderer")
    video_title = title.text.strip()
    video_titles.append(video_title)
    if len(video_titles) >= 5:
        break

# print the extracted video titles
print(video_titles)
```

```
[]
```

Q4. Write a python program to extract the number of views of the first five videos.

[61]:
```
import requests
from bs4 import BeautifulSoup

# specify the URL of the YouTube channel
url = "https://www.youtube.com/@PW-Foundation/videos"

# send a GET request to the URL and get the HTML content
response = requests.get(url)
html_content = response.content

# parse the HTML content using BeautifulSoup
soup = BeautifulSoup(html_content, "html.parser")

# extract the number of views of the first five videos
video_views = []
for video in soup.find_all("a", class_="yt-simple-endpoint style-scope␣
 ↪ytd-grid-video-renderer"):
    views = video.find("span", class_="style-scope ytd-grid-video-renderer")
    if views:
        video_view = views.text.strip()
    else:
        video_view = "N/A"
    video_views.append(video_view)
    if len(video_views) >= 5:
        break

# print the extracted video views
print(video_views)
```

```
[]
```

Q5. Write a python program to extract the time of posting of video for the first five videos.

[62]:
```
import requests
from bs4 import BeautifulSoup

# specify the URL of the YouTube channel
url = "https://www.youtube.com/@PW-Foundation/videos"

# send a GET request to the URL and get the HTML content
response = requests.get(url)
html_content = response.content

# parse the HTML content using BeautifulSoup
soup = BeautifulSoup(html_content, "html.parser")

# extract the time of posting of the first five videos
video_posting_time = []
for video in soup.find_all("a", class_="yt-simple-endpoint style-scope␣
 ↪ytd-grid-video-renderer"):
    posting_time = video.find("span", class_="style-scope␣
 ↪ytd-grid-video-renderer")
    if posting_time:
        video_posting = posting_time.text.strip()
    else:
        video_posting = "N/A"
    video_posting_time.append(video_posting)
    if len(video_posting_time) >= 5:
        break

# print the extracted video posting time
print(video_posting_time)
```

```
[]
```

[ ]: