In [1]:
```python
from selenium import webdriver
import time
import pandas as pd
from selenium.common.exceptions import NoSuchElementException
from selenium.webdriver.common.by import By
```

1. Scrape the details of most viewed videos on YouTube from Wikipedia. Url =
https://en.wikipedia.org/wiki/List_of_most-viewed_YouTube_videos (https://en.wikipedia.org/wiki/List_of_most-viewed_YouTube_videos) You need to find following details: A) Rank B) Name C) Artist D) Upload date E) Views

In [51]:
```python
driver=webdriver.Chrome()
```

In [52]:
```python
driver.get("https://en.wikipedia.org/wiki/List_of_most-viewed_YouTube_videos")
```

In [53]:
```python
rank=[]
name=[]
artist=[]
upload_date=[]
views=[]
```

In [54]:
```python
name_tags=driver.find_elements(By.XPATH,"//*[@id='mw-content-text']/div[1]/table[1]/tbody/tr/td[1]
for i in name_tags:
    if i.text is None:
        name.append("--")
    else:
        name.append(i.text)
print(len(name),name)
```

```
30 ['"Baby Shark Dance"[7]', '"Despacito"[10]', '"Johny Johny Yes Papa"[18]', '"Bath Song"[19]',
'"Wheels on the Bus"[20]', '"See You Again"[21]', '"Shape of You"[26]', '"Phonics Song with Two W
ords"[29]', '"Uptown Funk"[30]', '"Gangnam Style"[31]', '"Learning Colors – Colorful Eggs on a Fa
rm"[36]', '"Dame Tu Cosita"[37]', '"Axel F"[38]', '"Masha and the Bear – Recipe for Disaster"[3
9]', '"Baa Baa Black Sheep"[40]', '"Lakdi Ki Kathi"[41]', '"Sugar"[42]', '"Counting Stars"[43]',
'"Roar"[44]', '"Waka Waka (This Time for Africa)"[45]', '"Shree Hanuman Chalisa"[46]', '"Humpty t
he train on a fruits ride"[47]', '"Sorry"[48]', '"Thinking Out Loud"[49]', '"Perfect"[50]', '"Dar
k Horse"[51]', '"Let Her Go"[52]', '"Faded"[53]', '"Girls Like You"[54]', '"Lean On"[55]']
```

In [55]:
```python
artist_tags=driver.find_elements(By.XPATH,"//*[@id='mw-content-text']/div[1]/table[1]/tbody/tr/td[
for i in artist_tags:
    if i.text is None:
        artist.append("--")
    else:
        artist.append(i.text)
print(len(artist),artist)
```

```
30 ["Pinkfong Baby Shark - Kids' Songs & Stories", 'Luis Fonsi', "LooLoo Kids - Nursery Rhymes an
d Children's Songs", 'Cocomelon - Nursery Rhymes', 'Cocomelon - Nursery Rhymes', 'Wiz Khalifa',
'Ed Sheeran', 'ChuChu TV Nursery Rhymes & Kids Songs', 'Mark Ronson', 'Psy', 'Miroshka TV', 'Ultr
a Records', 'Crazy Frog', 'Get Movies', 'Cocomelon - Nursery Rhymes', 'Jingle Toons', 'Maroon 5',
'OneRepublic', 'Katy Perry', 'Shakira', 'T-Series Bhakti Sagar', 'Kiddiestv Hindi - Nursery Rhyme
s & Kids Songs', 'Justin Bieber', 'Ed Sheeran', 'Ed Sheeran', 'Katy Perry', 'Passenger', 'Alan Wa
lker', 'Maroon 5', 'Major Lazer Official']
```

In [56]:
```python
date=driver.find_elements(By.XPATH,"//*[@id='mw-content-text']/div[1]/table[1]/tbody/tr/td[4]")
for i in date:
    if i.text is None:
        upload_date.append("--")
    else:
        upload_date.append(i.text)
print(len(upload_date),upload_date)
```

30 ['June 17, 2016', 'January 12, 2017', 'October 8, 2016', 'May 2, 2018', 'May 24, 2018', 'April
6, 2015', 'January 30, 2017', 'March 6, 2014', 'November 19, 2014', 'July 15, 2012', 'February 2
7, 2018', 'April 5, 2018', 'June 16, 2009', 'January 31, 2012', 'June 25, 2018', 'June 14, 2018',
'January 14, 2015', 'May 31, 2013', 'September 5, 2013', 'June 4, 2010', 'May 10, 2011', 'January
26, 2018', 'October 22, 2015', 'October 7, 2014', 'November 9, 2017', 'February 20, 2014', 'July
25, 2012', 'December 3, 2015', 'May 31, 2018', 'March 22, 2015']

In [57]:
```python
view=driver.find_elements(By.XPATH,"//*[@id='mw-content-text']/div[1]/table[1]/tbody/tr/td[3]")
for i in view:
    if i.text is None:
        views.append("--")
    else:
        views.append(i.text)
print(len(views),views)
```

30 ['14.82', '8.49', '6.94', '6.79', '6.34', '6.33', '6.30', '5.90', '5.28', '5.22', '5.15', '4.7
2', '4.66', '4.59', '4.14', '4.11', '4.08', '4.05', '4.03', '3.99', '3.94', '3.87', '3.83', '3.7
9', '3.77', '3.76', '3.69', '3.66', '3.64', '3.64']

In [58]:
```python
df=pd.DataFrame([])
df["Name"]=name
df["Artist_Name"]=artist
df["Upload_Date"]=upload_date
df["Views"]=views
```

In [59]: `df`

Out[59]:

	Name	Artist_Name	Upload_Date	Views
0	"Baby Shark Dance"[7]	Pinkfong Baby Shark - Kids' Songs & Stories	June 17, 2016	14.82
1	"Despacito"[10]	Luis Fonsi	January 12, 2017	8.49
2	"Johny Johny Yes Papa"[18]	LooLoo Kids - Nursery Rhymes and Children's Songs	October 8, 2016	6.94
3	"Bath Song"[19]	Cocomelon - Nursery Rhymes	May 2, 2018	6.79
4	"Wheels on the Bus"[20]	Cocomelon - Nursery Rhymes	May 24, 2018	6.34
5	"See You Again"[21]	Wiz Khalifa	April 6, 2015	6.33
6	"Shape of You"[26]	Ed Sheeran	January 30, 2017	6.30
7	"Phonics Song with Two Words"[29]	ChuChu TV Nursery Rhymes & Kids Songs	March 6, 2014	5.90
8	"Uptown Funk"[30]	Mark Ronson	November 19, 2014	5.28
9	"Gangnam Style"[31]	Psy	July 15, 2012	5.22
10	"Learning Colors – Colorful Eggs on a Farm" [36]	Miroshka TV	February 27, 2018	5.15
11	"Dame Tu Cosita"[37]	Ultra Records	April 5, 2018	4.72
12	"Axel F"[38]	Crazy Frog	June 16, 2009	4.66
13	"Masha and the Bear – Recipe for Disaster"[39]	Get Movies	January 31, 2012	4.59
14	"Baa Baa Black Sheep"[40]	Cocomelon - Nursery Rhymes	June 25, 2018	4.14
15	"Lakdi Ki Kathi"[41]	Jingle Toons	June 14, 2018	4.11
16	"Sugar"[42]	Maroon 5	January 14, 2015	4.08
17	"Counting Stars"[43]	OneRepublic	May 31, 2013	4.05
18	"Roar"[44]	Katy Perry	September 5, 2013	4.03
19	"Waka Waka (This Time for Africa)"[45]	Shakira	June 4, 2010	3.99
20	"Shree Hanuman Chalisa"[46]	T-Series Bhakti Sagar	May 10, 2011	3.94
21	"Humpty the train on a fruits ride"[47]	Kiddiestv Hindi - Nursery Rhymes & Kids Songs	January 26, 2018	3.87
22	"Sorry"[48]	Justin Bieber	October 22, 2015	3.83
23	"Thinking Out Loud"[49]	Ed Sheeran	October 7, 2014	3.79
24	"Perfect"[50]	Ed Sheeran	November 9, 2017	3.77
25	"Dark Horse"[51]	Katy Perry	February 20, 2014	3.76
26	"Let Her Go"[52]	Passenger	July 25, 2012	3.69
27	"Faded"[53]	Alan Walker	December 3, 2015	3.66
28	"Girls Like You"[54]	Maroon 5	May 31, 2018	3.64
29	"Lean On"[55]	Major Lazer Official	March 22, 2015	3.64

In [ ]:

2. Scrape the details team India's international fixtures from bcci.tv. Url = https://www.bcci.tv/ (https://www.bcci.tv/). You need to find following details: A) Series B) Place C) Date D) Time Note: - From bcci.tv home page you have reach to the international fixture page through code.

In [73]: `driver=webdriver.Chrome()`

In [74]: `driver.get("https://www.bcci.tv/")`

In [75]:
```python
series=[]
place=[]
date=[]
time=[]
```

In [78]:
```python
series_tags=driver.find_elements(By.XPATH,"/html/body/div[5]/div[2]/div/div/div[2]/div/div[2]/div/
for i in series_tags:
    if i.text is None:
        series.append("--")
    else:
        series.append(i.text)
print(len(series),series)
```

1 ['INDIA TOUR OF SRI LANKA T20 SERIES 2024']

In [79]:
```python
place_tags=driver.find_elements(By.XPATH,"/html/body/section/div/div/div/div/div/div[2]/div[3]/div
for i in place_tags:
    if i.text is None:
        place.append("--")
    else:
        place.append(i.text)
print(len(place),place)
```

1 ['Pallekele International Cricket Stadium, Pallekele']

In [80]:
```python
date_tags=driver.find_elements(By.XPATH,"/html/body/section/div/div/div/div/div/div[2]/div[3]/div[
for i in date_tags:
    if i.text is None:
        date.append("--")
    else:
        date.append(i.text)
print(len(date),date)
```

1 ['30 Jul']

In [81]:
```python
time_tags=driver.find_elements(By.XPATH,"/html/body/section/div/div/div/div/div/div[2]/div[3]/div[
for i in time_tags:
    if i.text is None:
        time.append("--")
    else:
        time.append(i.text)
print(len(time),time)
```

1 ['19:00 IST']

In [ ]:

3. Scrape the details of State-wise GDP of India from statisticstime.com. Url = http://statisticstimes.com/ (http://statisticstimes.com/) You have to find following details: A) Rank B) State C) GSDP(18-19)- at current prices D) GSDP(19-20)- at current prices E) Share(18-19) F) GDP($ billion) Note: - From statisticstimes home page you have to reach to economy page through code.

In [24]:
```python
driver=webdriver.Chrome()
```

In [25]:
```python
driver.get("http://statisticstimes.com/")
```

In [26]:
```python
rank=[]
state=[]
GDP=[]
GDP_current=[]
GDP_previous=[]
share=[]
```

In [27]:
```python
economy=driver.find_element(By.XPATH,"/html/body/div[2]/div[1]/div[2]/div[2]")
economy.click()
```

In [28]:
```python
india=driver.find_element(By.XPATH,"/html/body/div[2]/div[1]/div[2]/div[2]/div/a[3]")
india.click()
```

In [29]:
```python
gdp=driver.find_element(By.XPATH,"/html/body/div[2]/div[2]/div[2]/ul/li[1]/a")
gdp.click()
```

In [30]:
```python
r=driver.find_elements(By.XPATH,"//td[@class='data1']")
for i in r:
    if i.text is None:
        rank.append("--")
    else:
        rank.append(i.text)
print(len(rank),rank)
```

```
68 ['1', '2', '3', '4', '5', '6', '7', '8', '9', '10', '11', '12', '13', '14', '15', '16', '17',
'18', '19', '20', '21', '22', '23', '24', '25', '26', '27', '28', '29', '30', '31', '32', '33',
'', '1', '2', '3', '4', '5', '6', '7', '8', '9', '10', '11', '12', '13', '14', '15', '16', '17',
'18', '19', '20', '21', '22', '23', '24', '25', '26', '27', '28', '29', '30', '31', '32', '33',
'']
```

In [31]:
```python
st=driver.find_elements(By.XPATH,"//td[@class='name']")
for i in st:
    if i.text is None:
        state.append("--")
    else:
        state.append(i.text)
print(len(state),state)
```

```
68 ['Maharashtra', 'Tamil Nadu', 'Karnataka', 'Uttar Pradesh', 'Gujarat', 'West Bengal', 'Rajasth
an', 'Andhra Pradesh', 'Telangana', 'Madhya Pradesh', 'Kerala', 'Delhi', 'Haryana', 'Odisha', 'Bi
har', 'Punjab', 'Assam', 'Chhattisgarh', 'Jharkhand', 'Uttarakhand', 'Jammu & Kashmir', 'Himachal
Pradesh', 'Goa', 'Tripura', 'Chandigarh', 'Puducherry', 'Meghalaya', 'Sikkim', 'Manipur', 'Arunac
hal Pradesh', 'Nagaland', 'Mizoram', 'Andaman & Nicobar Islands', 'India', 'Maharashtra', 'Tamil
Nadu', 'Karnataka', 'Uttar Pradesh', 'Gujarat', 'West Bengal', 'Rajasthan', 'Andhra Pradesh', 'Te
langana', 'Madhya Pradesh', 'Kerala', 'Delhi', 'Haryana', 'Bihar', 'Odisha', 'Punjab', 'Assam',
'Chhattisgarh', 'Jharkhand', 'Uttarakhand', 'Jammu & Kashmir', 'Himachal Pradesh', 'Goa', 'Tripur
a', 'Chandigarh', 'Puducherry', 'Meghalaya', 'Manipur', 'Sikkim', 'Arunachal Pradesh', 'Nagalan
d', 'Mizoram', 'Andaman & Nicobar Islands', 'India']
```

In [32]:
```python
gdp=driver.find_elements(By.XPATH,"//*[@id='table_id']/tbody/tr/td[5]")
for i in gdp:
    if i.text is None:
        GDP.append("--")
    else:
        GDP.append(i.text)
print(len(GDP),GDP)
```

```
33 ['3,108,022', '2,071,286', '1,978,094', '1,975,595', '1,928,683', '1,329,238', '1,193,489',
'1,148,471', '1,124,204', '1,092,964', '934,542', '881,336', '868,905', '662,886', '650,302', '61
7,192', '411,454', '410,525', '358,863', '267,143', '193,352', '172,162', '84,266', '62,550', '4
6,096', '43,810', '38,785', '37,557', '36,594', '34,775', '31,038', '27,824', '10,371']
```

In [33]:
```python
shr=driver.find_elements(By.XPATH,"//*[@id='table_id']/tbody/tr/td[6]")
for i in shr:
    if i.text is None:
        share.append("--")
    else:
        share.append(i.text)
print(len(share),share)
```

33 ['13.17%', '8.78%', '8.38%', '8.37%', '8.17%', '5.63%', '5.06%', '4.87%', '4.76%', '4.63%',
'3.96%', '3.73%', '3.68%', '2.81%', '2.76%', '2.62%', '1.74%', '1.74%', '1.52%', '1.13%', '0.8
2%', '0.73%', '0.36%', '0.27%', '0.20%', '0.19%', '0.16%', '0.16%', '0.16%', '0.15%', '0.13%',
'0.12%', '0.04%']

In [34]:
```python
tags=driver.find_elements(By.XPATH,"//*[@id='table_id']/tbody/tr/td[4]")
for i in tags:
    if i.text is None:
        GDP_current.append("--")
    else:
        GDP_current.append(i.text)
print(len(GDP_current),GDP_current)
```

33 ['-', '2,364,514', '2,269,995', '2,258,040', '2,230,609', '1,531,758', '1,365,849', '1,303,52
4', '1,308,034', '1,246,471', '1,046,188', '1,014,688', '984,055', '753,177', '751,396', '676,16
4', '493,167', '464,399', '393,722', '303,781', '224,226', '191,728', '93,672', '72,636', '54,28
5', '49,643', '42,697', '42,756', '-', '39,630', '35,643', '-', '-']

In [35]:
```python
tags_2=driver.find_elements(By.XPATH,"//*[@id='table_id']/tbody/tr/td[8]")
for i in tags_2:
    if i.text is None:
        GDP_previous.append("--")
    else:
        GDP_previous.append(i.text)
print(len(GDP_previous),GDP_previous)
```

33 ['-', '1,572,601', '1,423,229', '1,423,358', '-', '904,088', '842,621', '820,894', '767,936',
'660,363', '-', '672,247', '634,027', '503,871', '-', '493,283', '318,559', '321,945', '-', '213,
378', '139,171', '142,800', '-', '-', '-', '-', '26,524', '-', '-', '-', '-', '-', '-']

In [36]:
```python
df=pd.DataFrame([])
df['Rank']=rank[:33]
df['State']=state[:33]
df['Share']=share[:33]
df['GDP']=GDP[:33]
df['GDP_current']=GDP_current[:33]
df['GDP_previous']=GDP_previous[:33]
```

In [37]: `df`

Out[37]:

	Rank	State	Share	GDP	GDP_current	GDP_previous
0	1	Maharashtra	13.17%	3,108,022	-	-
1	2	Tamil Nadu	8.78%	2,071,286	2,364,514	1,572,601
2	3	Karnataka	8.38%	1,978,094	2,269,995	1,423,229
3	4	Uttar Pradesh	8.37%	1,975,595	2,258,040	1,423,358
4	5	Gujarat	8.17%	1,928,683	2,230,609	-
5	6	West Bengal	5.63%	1,329,238	1,531,758	904,088
6	7	Rajasthan	5.06%	1,193,489	1,365,849	842,621
7	8	Andhra Pradesh	4.87%	1,148,471	1,303,524	820,894
8	9	Telangana	4.76%	1,124,204	1,308,034	767,936
9	10	Madhya Pradesh	4.63%	1,092,964	1,246,471	660,363
10	11	Kerala	3.96%	934,542	1,046,188	-
11	12	Delhi	3.73%	881,336	1,014,688	672,247
12	13	Haryana	3.68%	868,905	984,055	634,027
13	14	Odisha	2.81%	662,886	753,177	503,871
14	15	Bihar	2.76%	650,302	751,396	-
15	16	Punjab	2.62%	617,192	676,164	493,283
16	17	Assam	1.74%	411,454	493,167	318,559
17	18	Chhattisgarh	1.74%	410,525	464,399	321,945
18	19	Jharkhand	1.52%	358,863	393,722	-
19	20	Uttarakhand	1.13%	267,143	303,781	213,378
20	21	Jammu & Kashmir	0.82%	193,352	224,226	139,171
21	22	Himachal Pradesh	0.73%	172,162	191,728	142,800
22	23	Goa	0.36%	84,266	93,672	-
23	24	Tripura	0.27%	62,550	72,636	-
24	25	Chandigarh	0.20%	46,096	54,285	-
25	26	Puducherry	0.19%	43,810	49,643	-
26	27	Meghalaya	0.16%	38,785	42,697	26,524
27	28	Sikkim	0.16%	37,557	42,756	-
28	29	Manipur	0.16%	36,594	-	-
29	30	Arunachal Pradesh	0.15%	34,775	39,630	-
30	31	Nagaland	0.13%	31,038	35,643	-
31	32	Mizoram	0.12%	27,824	-	-
32	33	Andaman & Nicobar Islands	0.04%	10,371	-	-

In [ ]:

4. Scrape the details of trending repositories on Github.com. Url = https://github.com/ (https://github.com/) You have to
find the following details: A) Repository title B) Repository description C) Contributors count D) Language used
Note: - From the home page you have to click on the trending option from Explore menu through code.

In [109]: `driver=webdriver.Chrome()`

In [110]: `driver.get("https://github.com/")`

In [111]:
```python
repository_title=[]
repository_description=[]
contributors_count=[]
muted_link=[]
language_used=[]
```

In [112]:
```python
explore=driver.find_element(By.XPATH,"/html/body/div[1]/div[1]/header/div/div[2]/div/nav/ul/li[4]"
explore.click()
```

In [113]:
```python
trending=driver.find_element(By.XPATH,"/html/body/div[1]/div[1]/header/div/div[2]/div/nav/ul/li[4]
trending.click()
```

In [114]:
```python
RN=driver.find_elements(By.CLASS_NAME,"text-normal")
for i in RN:
    if i.text is None:
        repository_title.append("--")
    else:
        repository_title.append(i.text)
print(len(repository_title),repository_title)
```

19 ['', 'sindresorhus /', 'opendatalab /', 'ollama /', 'freeCodeCamp /', 'freeCodeCamp /', 'jwash
am /', 'drawdb-io /', 'lllyasviel /', 'miss-mumu /', 'lllyasviel /', 'localsend /', 'alpkeskin
/', 'mingrammer /', 'practical-tutorials /', 'gchq /', 'nomic-ai /', 'SagerNet /', 'recloudstream
/']

In [115]:
```python
des=driver.find_elements(By.XPATH,"//p[@class='col-9 color-fg-muted my-1 pr-4']")
for i in des:
    if i.text is None:
        repository_description.append("--")
    else:
        repository_description.append(i.text)
print(len(repository_description),repository_description)
```

17 ['😎 Awesome lists about all kinds of interesting topics', 'A one-stop, open-source, high-qua
lity data extraction tool, supports PDF/webpage/e-book extraction.一站式开源高质量数据提取工具，支持
PDF/网页/多格式电子书提取。', 'Get up and running with Llama 3.1, Mistral, Gemma 2, and other large
language models.', "freeCodeCamp.org's open-source codebase and curriculum. Learn to code for fre
e.", 'API Documentation Browser', 'A complete computer science study plan to become a software en
gineer.', 'Free, simple, and intuitive online database design tool and SQL generator.', '公务员从
入门到上岸，最佳程序员公考实践教程', 'Focus on prompting and generating', 'An open-source cross-plat
form alternative to AirDrop', 'An automated e-mail OSINT tool', '🎨 Diagram as Code for prototyp
ing cloud system architectures', 'Curated list of project-based tutorials', 'The Cyber Swiss Army
Knife - a web app for encryption, encoding, compression and data analysis', 'GPT4All: Chat with L
ocal LLMs on Any Device', 'The universal proxy platform', 'Android app for streaming and download
ing media.']

In [116]:
```python
ml=driver.find_elements(By.XPATH,"//a[@class='Link Link--muted d-inline-block mr-3']")
for i in ml:
    if i.text is None:
        muted_link.append("NAN")
    else:
        muted_link.append(i.text)

print(len(muted_link),muted_link)
```

36 ['315,268', '27,369', '3,439', '257', '81,895', '6,259', '396,183', '36,441', '34,526', '2,30
2', '300,458', '75,499', '15,292', '1,090', '5,511', '551', '4,039', '371', '38,840', '5,230', '3
9,093', '2,024', '4,429', '541', '36,135', '2,353', '187,738', '24,516', '27,719', '3,124', '68,0
20', '7,473', '16,853', '2,052', '5,960', '496']

In [117]:
```python
for i in range(1,len(muted_link),2):
    contributors_count.append(muted_link[i])

print(len(contributors_count),contributors_count)
```

```
18 ['27,369', '257', '6,259', '36,441', '2,302', '75,499', '1,090', '551', '371', '5,230', '2,02
4', '541', '2,353', '24,516', '3,124', '7,473', '2,052', '496']
```

In [118]:
```python
language=driver.find_elements(By.XPATH,"//span[@itemprop='programmingLanguage']")
for i in language:
    if i.text is None:
        language_used.append("NAN")
    else:
        language_used.append(i.text)
print(len(language_used),language_used)
```

```
14 ['Python', 'Go', 'TypeScript', 'Ruby', 'JavaScript', 'Python', 'Python', 'Dart', 'Go', 'Pytho
n', 'JavaScript', 'C++', 'Go', 'Kotlin']
```

In [120]:
```python
df=pd.DataFrame([])
df["Repository_NAME"]=repository_title[:10]
df["Description"]=repository_description[:10]
df["Language"]=language_used[:10]
df["Contributed"]=contributors_count[:10]
```

In [121]:
```python
df
```

Out[121]:

	Repository_NAME	Description	Language	Contributed
0		😎 Awesome lists about all kinds of interesting...	Python	27,369
1	sindresorhus /	A one-stop, open-source, high-quality data ext...	Go	257
2	opendatalab /	Get up and running with Llama 3.1, Mistral, Ge...	TypeScript	6,259
3	ollama /	freeCodeCamp.org's open-source codebase and cu...	Ruby	36,441
4	freeCodeCamp /	API Documentation Browser	JavaScript	2,302
5	freeCodeCamp /	A complete computer science study plan to beco...	Python	75,499
6	jwasham /	Free, simple, and intuitive online database de...	Python	1,090
7	drawdb-io /	公务员从入门到上岸，最佳程序员公考实践教程	Dart	551
8	lllyasviel /	Focus on prompting and generating	Go	371
9	miss-mumu /	An open-source cross-platform alternative to A...	Python	5,230

In [ ]:

5. Scrape the details of top 100 songs on billiboard.com. Url = https:/www.billboard.com/ (http://www.billboard.com/)
You have to find the following details: A) Song name B) Artist name C) Last week rank D) Peak rank E) Weeks on
board Note: - From the home page you have to click on the charts option then hot 100-page link through code.

In [148]:
```python
driver=webdriver.Chrome()
```

In [149]:
```python
driver.get("http://www.billboard.com/")
```

In [152]:
```python
song_name=[]
artist_name=[]
last_week_rank=[]
peak_rank=[]
week_on_board=[]
```

In [153]:
```python
song=driver.find_elements(By.XPATH,"//h3[@id='title-of-a-story']")
for i in song:
    if i.text is None:
        song_name.append("--")
    else:
        song_name.append(i.text)
print(len(song_name),song_name)
```

45 ['Shaboozey Tosses Back Third Week at No. 1 on Hot 100 With 'A Bar Song (Tipsy)'', 'Olympics Clarify That the Opening Ceremony's Drag Performance Was Not Mocking 'The Last Supper'', 'Why More Songs Are Succeeding at More Radio Formats In the Streaming Era', 'Jimin Becomes Second BTS Member to Hit No. 1 as a Soloist on Billboard Global Charts, With 'Who'', 'Billboard's Music Industry Events Calendar', 'Billboard Hot 100 Top 10 Countdown for Aug. 3, 2024 | Billboard News', 'Selena Gomez Shuts Down Plastic Surgery Speculation: 'Leave Me Alone'', 'K. Michelle Signs to BBR Music Group/BMG Nashville', 'Trending', 'Fan Army Face-Off 2024: Vote Now In Round 2', 'Rory Feek Marries His Daughter's Teacher, Surprises Her With Special 'I Do' Song at Wedding', 'Tenacious D's Kyle Gass Deletes Trump Apology Instagram Post', 'Megan Fox Gives Birth, Bunnie XO Struggles With Infertility in Emotional Jelly Roll and MGK 'Lonely Road' Video', 'Angela Aguilar and Christian Nodal Are Married', 'Charli XCX Takes Fans Through a Day In Her Life | Billboard Cover', 'Charli XCX Takes Fans Through a Day In Her Life | Billboard Cover', 'Billboard Hot 100 Top 10 Countdown for July 27, 2024 | Billboard News', 'Which Rappers Support Trump's Bid for President? | Billboard Unfiltered', 'ENHYPEN Plays 'Who Gets the Most Hype' | Billboard News', 'Exclusive Look: ATEEZ's TOWARDS THE LIGHT : WILL TO POWER Tour | All Access | Billboard News', 'What if Streaming Royalties Favored Listening to Songs in their Entirety? ', 'Shaboozey Tosses Back Third Week at No. 1 on Hot 100 With 'A Bar Song (Tipsy)'', 'Legal Fights, Transparency & Neutrality: DiMA's CEO On Improvements Streamers Suggest for the MLC', 'Justin Bieber Can't Stop Smiling As He Cradles Hailey Bieber's Baby Bump', 'Snoop Dogg Does Affirmations With Suni Lee, Cheers on Simone Biles at 2024 Olympics', 'Looking Ahead: A Release Calendar of Upcoming Albums in 2024', 'Snoop Dogg Steps Out in Custom Gold Skechers to Cheer on Team USA at Olympics: 'All Gold Everything'', 'Swae Lee Faces Backlash After Telling Fans Not to Vote for Kamala Harris', 'Big Sean Signs Management Deal With Brandon Silverstein's S10 Entertainment', 'Queen, Taylor Swift, Michael Bublé and the Albums With the Most Weeks at No. 1 on Top Catalog Albums', 'Watch Flavor Flav Meet & Hug Dr. Jill Biden While Cheering Women's Water Polo at the 2024 Olympics', 'THE DAILY', "Just Stop Oil Activists Who Threw Tomato Soup at Van Gogh's 'Sunflowers' Get Prison Time", "Robert Downey Jr. Back as Doctor Doom for Two 'Avengers' Movies", 'Kaley Cuoco's Daughter Tildy Is Glowing With Happiness in a Doggy Dress', "Chad Dillon, Co-Owner Of Lil Baby's Seafood Restaurant, Shot And Killed In Atlanta Shooting", 'How to Watch Men's 3×3 Basketball at the 2024 Summer Olympics in Paris Without Cable', 'GloRilla Celebrates Birthday With "Fiery, Feisty, Ferocious" Photo Shoot', 'Nike's Women's Olympic Uniform for USA Track and Field Criticized for Being Too Revealing: The Controversy Explained', '', '', 'Follow Us', '', 'The Daily', 'Have a Tip?']

In [161]:
```python
details=[]
```

In [162]:
```python
detail_tags=driver.find_elements(By.XPATH,"/html/body/div[3]/main/div[2]/div[3]")
for i in detail_tags:
    title=i.text
    details.append(title)
```

In [163]:
```python
details
```

Out[163]: ['Video\nVIEW ALL\nPOPULAR ON BILLBOARD\nCharli XCX Takes Fans Through a Day In Her Life | Billboard Cover\nCharli XCX Takes Fans Through a Day In Her Life | Billboard Cover\nBillboard Hot 100 Top 10 Countdown for July 27, 2024 | Billboard News\nWhich Rappers Support Trump's Bid for President? | Billboard Unfiltered\nENHYPEN Plays 'Who Gets the Most Hype' | Billboard News\nExclusive Look: ATEEZ's TOWARDS THE LIGHT : WILL TO POWER Tour | All Access | Billboard News']

In [ ]:

6. Scrape the details of Highest selling novels. A) Book name B) Author name C) Volumes sold D) Publisher E) Genre
Url - https://www.theguardian.com/news/datablog/2012/aug/09/best-selling-books-all-time-fifty-shades-grey-compare (https://www.theguardian.com/news/datablog/2012/aug/09/best-selling-books-all-time-fifty-shades-grey-compare)

In [6]:
```python
driver=webdriver.Chrome()
```

In [7]:
```python
driver.get("https://www.theguardian.com/news/datablog/2012/aug/09/best-selling-books-all-time-fift
```

In [8]:
```python
details=[]
```

In [9]:
```python
detail_tags=driver.find_elements(By.XPATH,"/html/body/div/div[2]/div[2]/div/div[2]/div/table")
for i in detail_tags:
    title=i.text
    details.append(title)
```

In [11]: details

Out[11]: ['Top 100 best selling books of all time\nUK sales only. Click heading to sort table. Download th
is data\nRank\nTitle\nAuthor\nVolume Sales\nPublisher\nGenre\nSOURCE: NIELSEN BOOK SCAN\n1 Da Vin
ci Code,The Brown, Dan 5,094,805 Transworld Crime, Thriller & Adventure\n2 Harry Potter and the D
eathly Hallows Rowling, J.K. 4,475,152 Bloomsbury Children\'s Fiction\n3 Harry Potter and the Phi
losopher\'s Stone Rowling, J.K. 4,200,654 Bloomsbury Children\'s Fiction\n4 Harry Potter and the
Order of the Phoenix Rowling, J.K. 4,179,479 Bloomsbury Children\'s Fiction\n5 Fifty Shades of Gr
ey James, E. L. 3,758,936 Random House Romance & Sagas\n6 Harry Potter and the Goblet of Fire Row
ling, J.K. 3,583,215 Bloomsbury Children\'s Fiction\n7 Harry Potter and the Chamber of Secrets Ro
wling, J.K. 3,484,047 Bloomsbury Children\'s Fiction\n8 Harry Potter and the Prisoner of Azkaban
Rowling, J.K. 3,377,906 Bloomsbury Children\'s Fiction\n9 Angels and Demons Brown, Dan 3,193,946
Transworld Crime, Thriller & Adventure\n10 Harry Potter and the Half-blood Prince:Children\'s Edi
tion Rowling, J.K. 2,950,264 Bloomsbury Children\'s Fiction\n11 Fifty Shades Darker James, E. L.
2,479,784 Random House Romance & Sagas\n12 Twilight Meyer, Stephenie 2,315,405 Little, Brown Book
Young Adult Fiction\n13 Girl with the Dragon Tattoo,The:Millennium Trilogy Larsson, Stieg 2,233,5
70 Quercus Crime, Thriller & Adventure\n14 Fifty Shades Freed James, E. L. 2,193,928 Random House
Romance & Sagas\n15 Lost Symbol,The Brown, Dan 2,183,031 Transworld Crime, Thriller & Adventure\n
16 New Moon Meyer, Stephenie 2,152,737 Little, Brown Book Young Adult Fiction\n17 Deception Point
Brown, Dan 2,062,145 Transworld Crime, Thriller & Adventure\n18 Eclipse Meyer, Stephenie 2,052,87
6 Little, Brown Book Young Adult Fiction\n19 Lovely Bones,The Sebold, Alice 2,005,598 Pan Macmill
an General & Literary Fiction\n20 Curious Incident of the Dog in the Night-time,The Haddon, Mark
1,979,552 Random House General & Literary Fiction\n21 Digital Fortress Brown, Dan 1,928,900 Trans
world Crime, Thriller & Adventure\n22 Short History of Nearly Everything,A Bryson, Bill 1,852,919
Transworld Popular Science\n23 Girl Who Played with Fire,The:Millennium Trilogy Larsson, Stieg 1,
814,784 Quercus Crime, Thriller & Adventure\n24 Breaking Dawn Meyer, Stephenie 1,787,118 Little,
Brown Book Young Adult Fiction\n25 Very Hungry Caterpillar,The:The Very Hungry Caterpillar Carle,
Eric 1,783,535 Penguin Picture Books\n26 Gruffalo,The Donaldson, Julia 1,781,269 Pan Macmillan Pi
cture Books\n27 Jamie\'s 30-Minute Meals Oliver, Jamie 1,743,266 Penguin Food & Drink: General\n2
8 Kite Runner,The Hosseini, Khaled 1,629,119 Bloomsbury General & Literary Fiction\n29 One Day Ni
cholls, David 1,616,068 Hodder & Stoughton General & Literary Fiction\n30 Thousand Splendid Suns,
A Hosseini, Khaled 1,583,992 Bloomsbury General & Literary Fiction\n31 Girl Who Kicked the Hornet
s\' Nest,The:Millennium Trilogy Larsson, Stieg 1,555,135 Quercus Crime, Thriller & Adventure\n32
Time Traveler\'s Wife,The Niffenegger, Audrey 1,546,886 Random House General & Literary Fiction\n
33 Atonement McEwan, Ian 1,539,428 Random House General & Literary Fiction\n34 Bridget Jones\'s D
iary:A Novel Fielding, Helen 1,508,205 Pan Macmillan General & Literary Fiction\n35 World Accordi
ng to Clarkson,The Clarkson, Jeremy 1,489,403 Penguin Humour: Collections & General\n36 Captain C
orelli\'s Mandolin Bernieres, Louis de 1,352,318 Random House General & Literary Fiction\n37 Soun
d of Laughter,The Kay, Peter 1,310,207 Random House Autobiography: General\n38 Life of Pi Martel,
Yann 1,310,176 Canongate General & Literary Fiction\n39 Billy Connolly Stephenson, Pamela 1,231,9
57 HarperCollins Biography: The Arts\n40 Child Called It,A Pelzer, Dave 1,217,712 Orion Autobiogr
aphy: General\n41 Gruffalo\'s Child,The Donaldson, Julia 1,208,711 Pan Macmillan Picture Books\n4
2 Angela\'s Ashes:A Memoir of a Childhood McCourt, Frank 1,204,058 HarperCollins Autobiography: G
eneral\n43 Birdsong Faulks, Sebastian 1,184,967 Random House General & Literary Fiction\n44 North
ern Lights:His Dark Materials S. Pullman, Philip 1,181,503 Scholastic Ltd. Young Adult Fiction\n4
5 Labyrinth Mosse, Kate 1,181,093 Orion General & Literary Fiction\n46 Harry Potter and the Half-
blood Prince Rowling, J.K. 1,153,181 Bloomsbury Science Fiction & Fantasy\n47 Help,The Stockett,
Kathryn 1,132,336 Penguin General & Literary Fiction\n48 Man and Boy Parsons, Tony 1,130,802 Harp
erCollins General & Literary Fiction\n49 Memoirs of a Geisha Golden, Arthur 1,126,337 Random Hous
e General & Literary Fiction\n50 No.1 Ladies\' Detective Agency,The:No.1 Ladies\' Detective Agenc
y S. McCall Smith, Alexander 1,115,549 Little, Brown Book Crime, Thriller & Adventure\n51 Island,
The Hislop, Victoria 1,108,328 Headline General & Literary Fiction\n52 PS, I Love You Ahern, Cece
lia 1,107,379 HarperCollins General & Literary Fiction\n53 You are What You Eat:The Plan That Wil
l Change Your Life McKeith, Gillian 1,104,403 Penguin Fitness & Diet\n54 Shadow of the Wind,The Z
afon, Carlos Ruiz 1,092,349 Orion General & Literary Fiction\n55 Tales of Beedle the Bard,The Row
ling, J.K. 1,090,847 Bloomsbury Children\'s Fiction\n56 Broker,The Grisham, John 1,087,262 Random
House Crime, Thriller & Adventure\n57 Dr. Atkins\' New Diet Revolution:The No-hunger, Luxurious W
eight Loss P Atkins, Robert C. 1,054,196 Random House Fitness & Diet\n58 Subtle Knife,The:His Dar
k Materials S. Pullman, Philip 1,037,160 Scholastic Ltd. Young Adult Fiction\n59 Eats, Shoots and
Leaves:The Zero Tolerance Approach to Punctuation Truss, Lynne 1,023,688 Profile Books Group Usag
e & Writing Guides\n60 Delia\'s How to Cook:(Bk.1) Smith, Delia 1,015,956 Random House Food & Dri
nk: General\n61 Chocolat Harris, Joanne 1,009,873 Transworld General & Literary Fiction\n62 Boy i
n the Striped Pyjamas,The Boyne, John 1,004,414 Random House Childrens Books G Young Adult Fictio
n\n63 My Sister\'s Keeper Picoult, Jodi 1,003,780 Hodder & Stoughton General & Literary Fiction\n
64 Amber Spyglass,The:His Dark Materials S. Pullman, Philip 1,002,314 Scholastic Ltd. Young Adult
Fiction\n65 To Kill a Mockingbird Lee, Harper 998,213 Random House General & Literary Fiction\n66
Men are from Mars, Women are from Venus:A Practical Guide for Improvin Gray, John 992,846 HarperC
ollins Popular Culture & Media: General Interest\n67 Dear Fatty French, Dawn 986,753 Random House
Autobiography: The Arts\n68 Short History of Tractors in Ukrainian,A Lewycka, Marina 986,115 Peng
uin General & Literary Fiction\n69 Hannibal Harris, Thomas 970,509 Random House Crime, Thriller &
Adventure\n70 Lord of the Rings,The Tolkien, J. R. R. 967,466 HarperCollins Science Fiction & Fan
tasy\n71 Stupid White Men:...and Other Sorry Excuses for the State of the Natio Moore, Michael 96
3,353 Penguin Current Affairs & Issues\n72 Interpretation of Murder,The Rubenfeld, Jed 962,515 He
adline Crime, Thriller & Adventure\n73 Sharon Osbourne Extreme:My Autobiography Osbourne, Sharon
959,496 Little, Brown Book Autobiography: The Arts\n74 Alchemist,The:A Fable About Following Your
Dream Coelho, Paulo 956,114 HarperCollins General & Literary Fiction\n75 At My Mother\'s Knee

```
...:and Other Low Joints O\'Grady, Paul 945,640 Transworld Autobiography: The Arts\n76 Notes from
a Small Island Bryson, Bill 931,312 Transworld Travel Writing\n77 Return of the Naked Chef,The Ol
iver, Jamie 925,425 Penguin Food & Drink: General\n78 Bridget Jones: The Edge of Reason Fielding,
Helen 924,695 Pan Macmillan General & Literary Fiction\n79 Jamie\'s Italy Oliver, Jamie 906,968 P
enguin National & Regional Cuisine\n80 I Can Make You Thin McKenna, Paul 905,086 Transworld Fitne
ss & Diet\n81 Down Under Bryson, Bill 890,847 Transworld Travel Writing\n82 Summons,The Grisham,
John 869,671 Random House Crime, Thriller & Adventure\n83 Small Island Levy, Andrea 869,659 Headl
ine General & Literary Fiction\n84 Nigella Express Lawson, Nigella 862,602 Random House Food & Dr
ink: General\n85 Brick Lane Ali, Monica 856,540 Transworld General & Literary Fiction\n86 Memory
Keeper\'s Daughter,The Edwards, Kim 845,858 Penguin General & Literary Fiction\n87 Room on the Br
oom Donaldson, Julia 842,535 Pan Macmillan Picture Books\n88 About a Boy Hornby, Nick 828,215 Pen
guin General & Literary Fiction\n89 My Booky Wook Brand, Russell 820,563 Hodder & Stoughton Autob
iography: The Arts\n90 God Delusion,The Dawkins, Richard 816,907 Transworld Popular Science\n91
"Beano" Annual,The 0 816,585 D.C. Thomson Children\'s Annuals\n92 White Teeth Smith, Zadie 815,58
6 Penguin General & Literary Fiction\n93 House at Riverton,The Morton, Kate 814,370 Pan Macmillan
General & Literary Fiction\n94 Book Thief,The Zusak, Markus 809,641 Transworld General & Literary
Fiction\n95 Nights of Rain and Stars Binchy, Maeve 808,900 Orion General & Literary Fiction\n96 G
host,The Harris, Robert 807,311 Random House General & Literary Fiction\n97 Happy Days with the N
aked Chef Oliver, Jamie 794,201 Penguin Food & Drink: General\n98 Hunger Games,The:Hunger Games T
rilogy Collins, Suzanne 792,187 Scholastic Ltd. Young Adult Fiction\n99 Lost Boy,The:A Foster Chi
ld\'s Search for the Love of a Family Pelzer, Dave 791,507 Orion Biography: General\n100 Jamie\'s
Ministry of Food:Anyone Can Learn to Cook in 24 Hours Oliver, Jamie 791,095 Penguin Food & Drink:
General']
```

7. Scrape the details most watched tv series of all time from imdb.com. Url = https://www.imdb.com/list/ls095964455/ (https://www.imdb.com/list/ls095964455/) You have to find the following details: A) Name B) Year span C) Genre D) Run time E) Ratings F) Votes

```
In [164]: driver=webdriver.Chrome()
```

```
In [165]: driver.get("https://www.imdb.com/list/ls095964455/")
```

```
In [ ]:
```

8. Details of Datasets from UCI machine learning repositories. Url = https://archive.ics.uci.edu/ (https://archive.ics.uci.edu/) You have to find the following details: A) Dataset name B) Data type C) Task D) Attribute type E) No of instances F) No of attribute G) Year Note: - from the home page you have to go to the Show All Dataset page through code.

```
In [4]: driver=webdriver.Chrome()
```

```
In [27]: driver.get("https://archive.ics.uci.edu/")
```

```
In [29]: search_bar=driver.find_element(By.XPATH,"/html/body/div/div[1]/div[1]/header/div/div[1]/div/form/d
         search_bar.send_keys("Show All Dataset")
```

```
In [31]: search_button=driver.find_element(By.XPATH,"/html/body/div/div[1]/div[1]/header/div/div[1]/div/for
         search_button.click()
```

```
In [32]: detail=[]
```

```
In [33]: detail_tags=driver.find_elements(By.XPATH,"/html/body/div/div[1]/div[1]/main/div/div[1]")
         for i in detail_tags:
             title=i.text
             detail.append(title)
```

In [34]: `detail`

Out[34]: ["Basketball dataset\nDonated on 7/1/2019\nIt's data collected from different volunteers that are done in a basketball practice: dribbling, pass, shoot, picking the ball, and holding the ball.\nDataset Characteristics\nTime-Series\nSubject Area\nGames\nAssociated Tasks\nClassification\nFeature Type\nInteger\n# Instances\n10000\n# Features\n7\nDataset Information\nAdditional Information\nThere are different trials. For which, pass, shoot and pick up the ball have 5. and hold and dribble there are 2. \nFirst of all, we gathered the 4 users who were willing to be our test samples. \nThen, one by one we made them do the following 5 activities: Pass, hold the ball, shoot pick up the ball, and dribble. \nEach activity had a different way of gathering its corresponding data. \n\nFor holding the ball, we made the volunteer stand in one place in a holding position. \nOnce ready, we run the app. After 5 seconds we stop it and save the data with the userâ€™s first initial, \nthe activity and the number of the trial. For this label we did a total of 3 trials for each person.\n\nNext we started collecting the data of passing. The volunteer starts with the ball in a holding position. \nNext we run the app, for which after 3 seconds we tell the volunteer to pass the ball to one of us, once \nfinish we stop the app. For this label we did a total of 5 trails for each person. \n\nThen, we collected the data of dribbling. The volunteers start with the ball in holding position. Then, 3 seconds\nafter we run the app we tell him to dribble and after 5 he started the dribbling we said to stop. \nOnce he stops we go and stop the app. For this label we did a total of 3 trials for each person.\n\nContinuing with the activity of shooting, we let the volunteer get ready in a holding position with the ball, and \nthen we run the app. For which, he shoots immediately after we start the application. Once finish, we stop the app. \nFor this label we did a total of 5 trials for each person. \n\nFinally, we gather the data of picking the ball. For this data, we just start the app, \nand after 3 seconds the user picks up the ball, and at the 6 seconds we stop the app gathering this range of data.\nFor this label we did a total of 5 trails for each person. \n\nFinally we did a different way of collecting the data. For which one user in a set of time did each activity spontanously. Every time he did an activity we \ncollected the time for which he did it in a chronometer, and in a video all for which started closely at the same time. \nSHOW LESS\nHas Missing Values?\nNo\nVariable Information\nwe use acceloremeter measures x y z in ms2 and gyroscope measures r phi theta. for X the data is in g\nReviews\nThere are no reviews for this dataset yet.\nLOGIN TO WRITE A REVIEW"]

In [ ]: