In [1]:
```python
import requests
from bs4 import BeautifulSoup
```

In [2]:
```python
base_site = "https://editorial.rottentomatoes.com/guide/essential-movies-to-watch-now
```

In [3]:
```python
response = requests.get(base_site)
response.ok
```

Out[3]: True

In [4]:
```python
html = response.text
```

In [5]:
```python
soup = BeautifulSoup(html, "html.parser")
```

In [32]:
```python
soup.find("img")
```

Out[32]: `<img class="media-object" src="//graph.facebook.com/v2.2/100001988417196/picture"/>`

In [9]:
```python
divs = soup.find_all("div",{"class":"col-sm-18 col-full-xs countdown-item-content"})
```

In [10]:
```python
print(divs[1].text)
```

```
2001: A Space Odyssey (1968)   92%



#2



Adjusted Score: 108227%
Critics Consensus: One of the most influential of all sci-fi films -- and one of the
most controversial -- Stanley Kubrick's 2001 is a delicate, poetic meditation on the
ingenuity -- and folly -- of mankind.
Synopsis: An imposing black structure provides a connection between the past and the
future in this enigmatic adaptation of a short... [More]

Starring: Keir Dullea, Gary Lockwood, William Sylvester, Daniel Richter

Directed By: Stanley Kubrick
```

In [34]:
```python
headings = [div.find("h2") for div in divs]
headings[0].get_text()
```

Out[34]: '12 Angry Men (1957)   100%'

```
In [12]: movie_name = [movie_name.find("a").text for movie_name in headings]
         movie_name[:5]
```

```
Out[12]: ['12 Angry Men',
          '2001: A Space Odyssey',
          'The 400 Blows',
          'The Adventures of Priscilla, Queen of the Desert',
          'The Adventures of Robin Hood']
```

```
In [13]: len(movie_name)
```

```
Out[13]: 150
```

```
In [35]: year = \
         [year.find("span", {"class":"start-year"}).text.strip("()") \
          for year in headings]
```

```
In [36]: len(year)
```

```
Out[36]: 150
```

```
In [16]: score = [s.find("span",{"class":"tMeterScore"}).text for s in headings]
```

```
In [17]: directors = [div.find("div", class_ = "director") for div in divs]
```

```
In [18]: directors[4]
```

```
Out[18]: <div class="info director">
         <span class="descriptor">Directed By:</span> <a class="" href="//www.rottentomatoes.
         com/celebrity/michael_curtiz">Michael Curtiz</a>, <a class="" href="//www.rottentoma
         toes.com/celebrity/william_keighley">William Keighley</a></div>
```

```
In [19]: dir_link = directors[4].find_all("a")
         dir_link
```

```
Out[19]: [<a class="" href="//www.rottentomatoes.com/celebrity/michael_curtiz">Michael Curtiz
         </a>,
          <a class="" href="//www.rottentomatoes.com/celebrity/william_keighley">William Keig
         hley</a>]
```

```
In [20]: dir_name = [link.text for link in dir_link]
         dir_name
```

```
Out[20]: ['Michael Curtiz', 'William Keighley']
```

```
In [21]: " , ".join(dir_name)
```

```
Out[21]: 'Michael Curtiz , William Keighley'
```

In [22]:
```python
dirs = []

for d in directors:
    dir_link = d.find_all("a")
    dir_name = [link.text for link in dir_link]
    result = " , ".join(dir_name)
#     print(result)
    dirs.append(result)
```

In [23]:
```python
dirs[:6]
```

Out[23]:
```
['Sidney Lumet',
 'Stanley Kubrick',
 'François Truffaut',
 'Stephan Elliott',
 'Michael Curtiz , William Keighley',
 'Werner Herzog']
```

In [24]:
```python
dirs = \
[" , ".join([link.text for link in d.find_all("a")]) for d in directors]
```

In [37]:
```python
import pandas as pd

movie_info = pd.DataFrame()

movie_info["Movie_Title"] = movie_name
movie_info["Year"] = year
movie_info["Score"] = score
movie_info["Directors"] = dirs


movie_info
```

Out[37]:

	Movie_Title	Year	Score	Directors
0	12 Angry Men	1957	100%	Sidney Lumet
1	2001: A Space Odyssey	1968	92%	Stanley Kubrick
2	The 400 Blows	1959	99%	François Truffaut
3	The Adventures of Priscilla, Queen of the Desert	1994	94%	Stephan Elliott
4	The Adventures of Robin Hood	1938	100%	Michael Curtiz , William Keighley
...	...	...	...	...
145	Iron Man	2008	94%	Jon Favreau
146	It Happened One Night	1934	98%	Frank Capra
147	It's a Wonderful Life	1946	94%	Frank Capra
148	Jaws	1975	97%	Steven Spielberg
149	John Wick	2014	86%	Chad Stahelski

150 rows × 4 columns

In [ ]:

In [26]:
```python
pages = soup.find_all("a", class_="post-page-numbers")
pages = [i.get("href") for i in pages]
```

In [27]:
```python
pages_link = list(set(pages))
```

In [28]:
```python
pages_link.append(base_site)
```

In [29]:
```python
pages_link
```

Out[29]: ['https://editorial.rottentomatoes.com/guide/essential-movies-to-watch-now/2/',
          'https://editorial.rottentomatoes.com/guide/essential-movies-to-watch-now/']

In [ ]: