2 .Write a python program to display IMDB's Top rated 100 movies' data (i.e. name, rating, year of release)and make data frame.

In [3]:
```python
!pip install bs4
```

```
Requirement already satisfied: bs4 in ./opt/anaconda3/lib/python3.9/site-packa
ges (0.0.1)
Requirement already satisfied: beautifulsoup4 in ./opt/anaconda3/lib/python3.9
/site-packages (from bs4) (4.10.0)
Requirement already satisfied: soupsieve>1.2 in ./opt/anaconda3/lib/python3.9/
site-packages (from beautifulsoup4->bs4) (2.2.1)
```

In [4]:
```python
import unidecode
```

In [5]:
```python
!pip install requests
```

```
Requirement already satisfied: requests in ./opt/anaconda3/lib/python3.9/site-
packages (2.26.0)
Requirement already satisfied: idna<4,>=2.5 in ./opt/anaconda3/lib/python3.9/s
ite-packages (from requests) (3.2)
Requirement already satisfied: charset-normalizer~=2.0.0 in ./opt/anaconda3/li
b/python3.9/site-packages (from requests) (2.0.4)
Requirement already satisfied: certifi>=2017.4.17 in ./opt/anaconda3/lib/pytho
n3.9/site-packages (from requests) (2021.10.8)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in ./opt/anaconda3/lib/py
thon3.9/site-packages (from requests) (1.26.7)
```

In [6]:
```python
from bs4 import BeautifulSoup
```

In [7]:
```python
import requests
```

In [8]:
```python
url = "https://www.imdb.com/search/title/?groups=top_100&sort=user_rating,des
```

In [9]:
```python
r = requests.get(url)
r
```

Out[9]: <Response [200]>

In [10]:
```python
soup = BeautifulSoup(r.content)
```

```
In [1]:    #soup
```

```
In [51]:   titles = []
           year = []
```

```
In [54]:   for i in soup.find_all("h3",class_="lister-item-header"):
               y = i.text.strip().replace("\n","").split("(")
               titles.append(y[0])
               year.append(y[1])
```

```
In [55]:   titles
```

```
Out[55]:   ['1.The Shawshank Redemption',
            '2.The Godfather',
            '3.The Dark Knight',
            '4.The Lord of the Rings: The Return of the King',
            "5.Schindler's List",
            '6.The Godfather Part II',
            '7.12 Angry Men',
            '8.Pulp Fiction',
            '9.Inception',
            '10.The Lord of the Rings: The Two Towers',
            '11.Fight Club',
            '12.The Lord of the Rings: The Fellowship of the Ring',
            '13.Forrest Gump',
            '14.Il buono, il brutto, il cattivo',
            '15.The Matrix',
            '16.Goodfellas',
            '17.The Empire Strikes Back',
            "18.One Flew Over the Cuckoo's Nest",
            '19.Interstellar',
            '20.Cidade de Deus',
            '21.Sen to Chihiro no kamikakushi',
            '22.Saving Private Ryan',
            '23.The Green Mile',
            '24.La vita è bella',
            '25.Se7en',
            '26.Terminator 2: Judgment Day',
            '27.The Silence of the Lambs',
            '28.Star Wars',
            '29.Seppuku',
            '30.Shichinin no samurai',
            "31.It's a Wonderful Life",
            '32.Gisaengchung',
            '33.Whiplash',
            '34.The Intouchables',
            '35.The Prestige',
            '36.The Departed',
            '37.The Pianist',
            '38.Gladiator',
```

'39.American History X',
'40.The Usual Suspects',
'41.Léon',
'42.The Lion King',
'43.Nuovo Cinema Paradiso',
'44.Hotaru no haka',
'45.Back to the Future',
'46.Apocalypse Now',
'47.Alien',
'48.Once Upon a Time in the West',
'49.Psycho',
'50.Rear Window',
'1.The Shawshank Redemption',
'2.The Godfather',
'3.The Dark Knight',
'4.The Lord of the Rings: The Return of the King',
"5.Schindler's List",
'6.The Godfather Part II',
'7.12 Angry Men',
'8.Pulp Fiction',
'9.Inception',
'10.The Lord of the Rings: The Two Towers',
'11.Fight Club',
'12.The Lord of the Rings: The Fellowship of the Ring',
'13.Forrest Gump',
'14.Il buono, il brutto, il cattivo',
'15.The Matrix',
'16.Goodfellas',
'17.The Empire Strikes Back',
"18.One Flew Over the Cuckoo's Nest",
'19.Interstellar',
'20.Cidade de Deus',
'21.Sen to Chihiro no kamikakushi',
'22.Saving Private Ryan',
'23.The Green Mile',
'24.La vita è bella',
'25.Se7en',
'26.Terminator 2: Judgment Day',
'27.The Silence of the Lambs',
'28.Star Wars',
'29.Seppuku',
'30.Shichinin no samurai',
"31.It's a Wonderful Life",
'32.Gisaengchung',
'33.Whiplash',
'34.The Intouchables',
'35.The Prestige',
'36.The Departed',
'37.The Pianist',
'38.Gladiator',
'39.American History X',
'40.The Usual Suspects',
'41.Léon',
'42.The Lion King',
'43.Nuovo Cinema Paradiso',
'44.Hotaru no haka',
'45.Back to the Future',

```
       '46.Apocalypse Now',
       '47.Alien',
       '48.Once Upon a Time in the West',
       '49.Psycho',
       '50.Rear Window']
```

In [86]:
```python
title_of_movies = []
```

In [87]:
```python
for i in range(0,50):
    y = titles[i].split(".")
    title_of_movies.append(y[1])
```

In [88]:
```python
title_of_movies
```

```
Out[88]:  ['The Shawshank Redemption',
           'The Godfather',
           'The Dark Knight',
           'The Lord of the Rings: The Return of the King',
           "Schindler's List",
           'The Godfather Part II',
           '12 Angry Men',
           'Pulp Fiction',
           'Inception',
           'The Lord of the Rings: The Two Towers',
           'Fight Club',
           'The Lord of the Rings: The Fellowship of the Ring',
           'Forrest Gump',
           'Il buono, il brutto, il cattivo',
           'The Matrix',
           'Goodfellas',
           'The Empire Strikes Back',
           "One Flew Over the Cuckoo's Nest",
           'Interstellar',
           'Cidade de Deus',
           'Sen to Chihiro no kamikakushi',
           'Saving Private Ryan',
           'The Green Mile',
           'La vita è bella',
           'Se7en',
           'Terminator 2: Judgment Day',
           'The Silence of the Lambs',
           'Star Wars',
           'Seppuku',
           'Shichinin no samurai',
           "It's a Wonderful Life",
           'Gisaengchung',
           'Whiplash',
           'The Intouchables',
           'The Prestige',
           'The Departed',
           'The Pianist',
           'Gladiator',
           'American History X',
           'The Usual Suspects',
           'Léon',
           'The Lion King',
           'Nuovo Cinema Paradiso',
           'Hotaru no haka',
           'Back to the Future',
           'Apocalypse Now',
           'Alien',
           'Once Upon a Time in the West',
           'Psycho',
           'Rear Window']

In [91]:  year_movie = []
```

```python
for i in range(0,50,1):
    z = year[i].replace(")","")
    year_movie.append(z)
```

```python
year_movie
```

```
Out[93]:  ['1994',
          '1972',
          '2008',
          '2003',
          '1993',
          '1974',
          '1957',
          '1994',
          '2010',
          '2002',
          '1999',
          '2001',
          '1994',
          '1966',
          '1999',
          '1990',
          '1980',
          '1975',
          '2014',
          '2002',
          '2001',
          '1998',
          '1999',
          '1997',
          '1995',
          '1991',
          '1991',
          '1977',
          '1962',
          '1954',
          '1946',
          '2019',
          '2014',
          '2011',
          '2006',
          '2006',
          '2002',
          '2000',
          '1998',
          '1995',
          '1994',
          '1994',
          '1988',
          '1988',
          '1985',
          '1979',
          '1979',
          '1968',
          '1960',
          '1954']
```

```
In [94]:  rating = []
```

```python
for i in soup.find_all("div", class_="inline-block ratings-imdb-rating"):
    rating.append(i.text.replace("\n",""))
```

```python
rating
```

```
Out[96]: ['9.3',
          '9.2',
          '9.0',
          '9.0',
          '9.0',
          '9.0',
          '9.0',
          '8.9',
          '8.8',
          '8.8',
          '8.8',
          '8.8',
          '8.8',
          '8.8',
          '8.7',
          '8.7',
          '8.7',
          '8.7',
          '8.6',
          '8.6',
          '8.6',
          '8.6',
          '8.6',
          '8.6',
          '8.6',
          '8.6',
          '8.6',
          '8.6',
          '8.6',
          '8.6',
          '8.6',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5',
          '8.5']
```

In [99]:
```python
print(len(title_of_movies),len(year_movie),len(rating))
```

50 50 50

```
import pandas as pd
df = pd.DataFrame({"Movie":title_of_movies,"Year":year_movie,"Rating":rating}
df.index += 1
df
```

	Movie	Year	Rating
1	The Shawshank Redemption	1994	9.3
2	The Godfather	1972	9.2
3	The Dark Knight	2008	9.0
4	The Lord of the Rings: The Return of the King	2003	9.0
5	Schindler's List	1993	9.0
6	The Godfather Part II	1974	9.0
7	12 Angry Men	1957	9.0
8	Pulp Fiction	1994	8.9
9	Inception	2010	8.8
10	The Lord of the Rings: The Two Towers	2002	8.8
11	Fight Club	1999	8.8
12	The Lord of the Rings: The Fellowship of the Ring	2001	8.8
13	Forrest Gump	1994	8.8
14	Il buono, il brutto, il cattivo	1966	8.8
15	The Matrix	1999	8.7
16	Goodfellas	1990	8.7
17	The Empire Strikes Back	1980	8.7
18	One Flew Over the Cuckoo's Nest	1975	8.7
19	Interstellar	2014	8.6
20	Cidade de Deus	2002	8.6
21	Sen to Chihiro no kamikakushi	2001	8.6
22	Saving Private Ryan	1998	8.6
23	The Green Mile	1999	8.6
24	La vita è bella	1997	8.6
25	Se7en	1995	8.6
26	Terminator 2: Judgment Day	1991	8.6
27	The Silence of the Lambs	1991	8.6
28	Star Wars	1977	8.6
29	Seppuku	1962	8.6

30	Shichinin no samurai	1954	8.6
31	It's a Wonderful Life	1946	8.6
32	Gisaengchung	2019	8.5
33	Whiplash	2014	8.5
34	The Intouchables	2011	8.5
35	The Prestige	2006	8.5
36	The Departed	2006	8.5
37	The Pianist	2002	8.5
38	Gladiator	2000	8.5
39	American History X	1998	8.5
40	The Usual Suspects	1995	8.5
41	Léon	1994	8.5
42	The Lion King	1994	8.5
43	Nuovo Cinema Paradiso	1988	8.5
44	Hotaru no haka	1988	8.5
45	Back to the Future	1985	8.5
46	Apocalypse Now	1979	8.5
47	Alien	1979	8.5
48	Once Upon a Time in the West	1968	8.5
49	Psycho	1960	8.5
50	Rear Window	1954	8.5

In [ ]: