

```
In [1]: from bs4 import BeautifulSoup
import requests
import re
import pandas as pd
from imdb import Cinemagoer
```

```
In [2]: url = 'https://www.imdb.com/list/ls000096828/'
response = requests.get(url)
soup = BeautifulSoup(response.text, 'lxml')
```

```
In [3]: s = soup.find('script', type='application/ld+json')
```

```
In [4]: temp = pd.DataFrame(data = json.loads(s.text)["about"]["itemListElement"])
```

```
In [5]: rank = []
name = []
year = []
rating = []
```

```
In [6]: ia = Cinemagoer()
```

```
In [7]: i = 1
for movie in temp["url"]:
    num = str(re.findall('\d+', movie)[0])
    mv = ia.get_movie(num)
    rank.append(i)
    name.append(mv["localized title"])
    year.append(mv["year"])
    rating.append(mv["rating"])
    i += 1
```

```
In [8]: df = pd.DataFrame(data = {"rank": rank, "name": name, "year": year, "rating": rating})
```

```
In [9]: df.to_csv("parsed_ranking.csv", index = False)
```