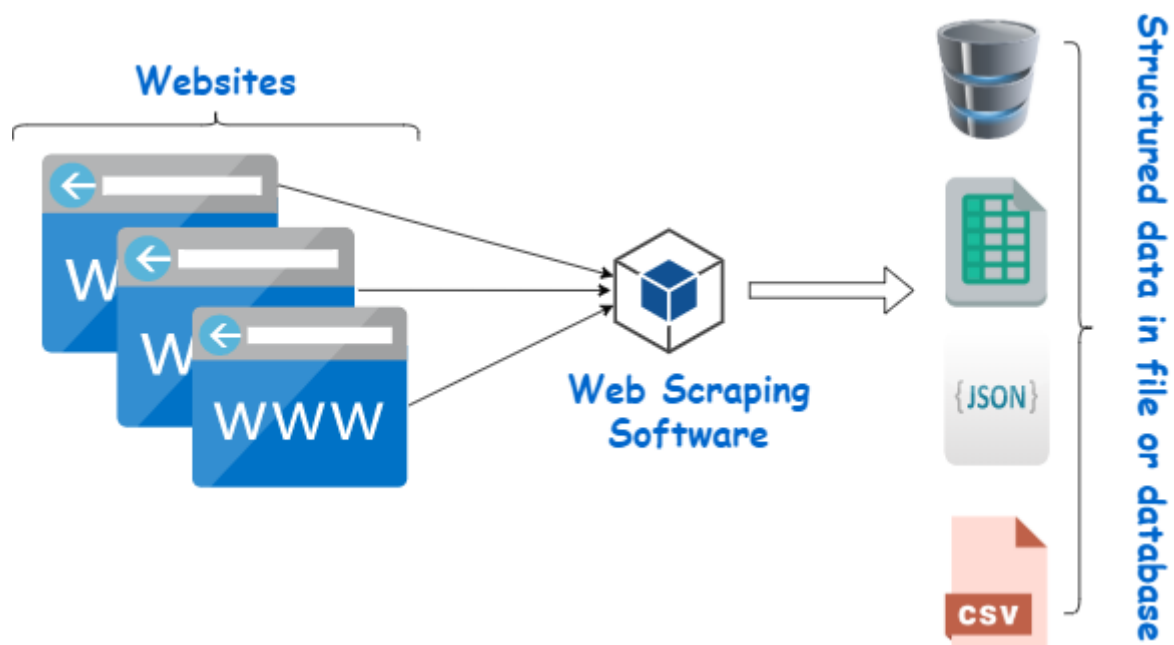


# Task 3

## Movie Data Web Scrapping and Analysis



Submitted by:  
Yousef Elshabrawy  
Philopateer Emad

# Used Libraries:

requests, scrapy, pandas, csv, Matplotlib

# Findings From Data Analysis:

```
Movie(s) with the highest IMDB rating:
      Movie Title  Movie Rating  IMDB
0  The Shawshank Redemption      9.3

Movie(s) with the highest moviemeter rating:
      Movie Title  Movie Rating mm
0  The Shawshank Redemption      4.42

Movie(s) with the lowest IMDB rating:
      Movie Title  Movie Rating  IMDB
26              Into the Wild      8.1
38      Lock, Stock and Two Smoking Barrels      8.1
46              Prisoners      8.1
52          Trainspotting      8.1
56      In the Name of the Father      8.1
62              Platoon      8.1
63          The Deer Hunter      8.1
68              Gran Torino      8.1
74  Pirates of the Caribbean: The Curse of the Bla...      8.1
75              Hotel Rwanda      8.1
76          Mary and Max      8.1
78              Warrior      8.1
84  Harry Potter and the Deathly Hallows: Part 2      8.1
85      Three Billboards Outside Ebbing, Missouri      8.1
89              La Haine      8.1
92          The Big Lebowski      8.1
97          The Third Man      8.1
98              Fargo      8.1
101          Raging Bull      8.1
103      La Battaglia di Algeri      8.1
108      The Best Years of Our Lives      8.1
114          The Grapes of Wrath      8.1
116      Million Dollar Baby      8.1
117              Ben-Hur      8.1
118              Persona      8.1
126          Tonari no Totoro      8.1
130              Room      8.1
131              The Help      8.1

Movie(s) with the lowest moviemeter rating:
      Movie Title  Movie Rating mm
129      Batman Begins      3.78
130              Room      3.78
131              The Help      3.78
132  The Great Dictator      3.78
133      Dial M for Murder      3.78
```

```
The most common genre is 'Drama' with 66 occurrences.
```

```
Number of movies released in each decade:
```

```
Decade
```

```
1940s      8
```

```
1950s     10
```

```
1960s      9
```

```
1970s     10
```

```
1980s     16
```

```
1990s     32
```

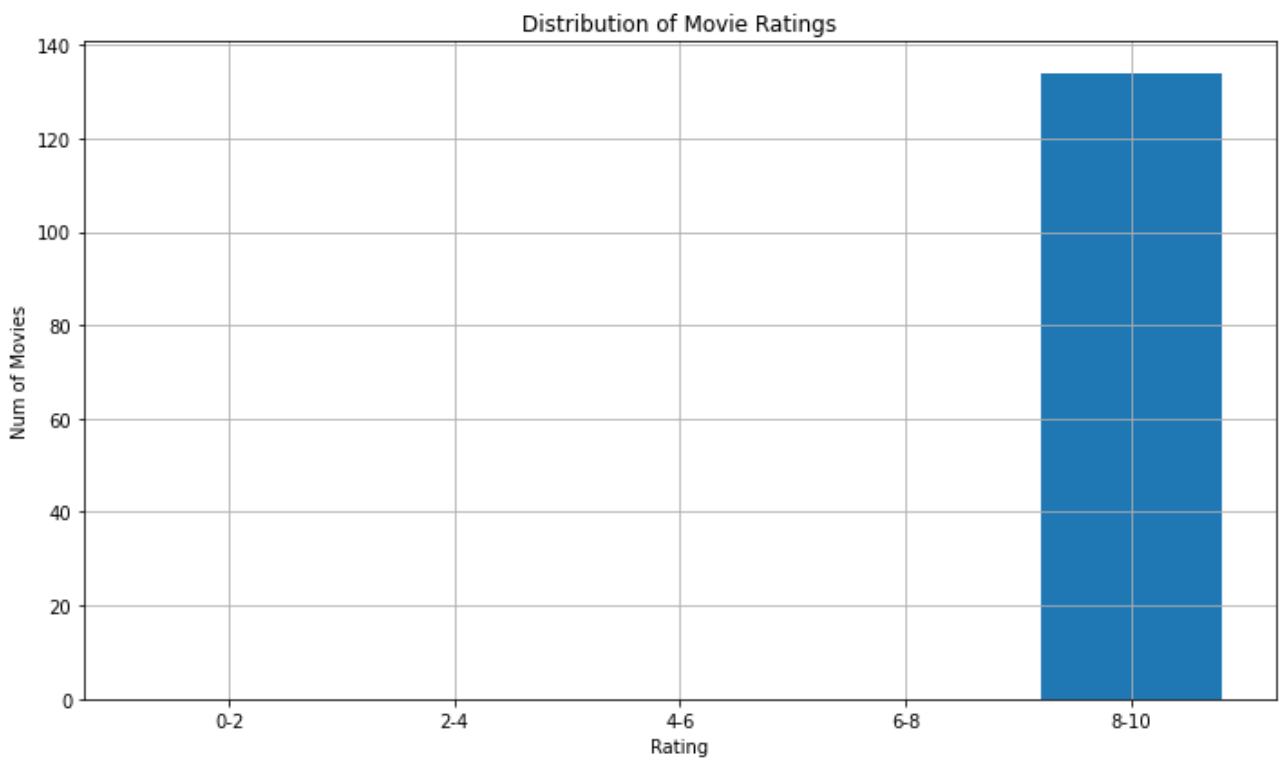
```
2000s     25
```

```
2010s     22
```

```
2020s      2
```

```
Name: count, dtype: int64
```

# Visualization for distribution of movie ratings



# Visualization for distribution of movie genre

