Web Scraping with Beautiful soups and requests

In [60]:
```python
# Step 1: Import libraries
import requests
from bs4 import BeautifulSoup
import csv
```

In [62]:
```python
# URL of the movies data page
url = "https://scrapethissite.com/pages/simple/"

# Send a GET request to fetch the HTML content
response = requests.get(url)

# Check if the request was successful
if response.status_code == 200:
    print("Page fetched successfully!")
else:
    print(f"Failed to fetch the page. Status code: {response.status_code}")

# Parse the page content using BeautifulSoup
soup = BeautifulSoup(response.text, 'html.parser')
```

Page fetched successfully!

In [64]:
```python
# Find all movie entries on the page
movies = soup.find_all('div', class_='movie')

# Initialize a list to store movie data
movies_data = []

# Loop through each movie entry and extract details
for movie in movies:
    # Extract the movie title
    title = movie.find('h2', class_='movie-title').text.strip()

    # Extract the release year
    year = movie.find('span', class_='movie-year').text.strip()

    # Extract the director (after the label)
```

```python
        director = movie.find('p', string=lambda text: "Director:" in text).text.replace('Director:', '').strip()

        # Extract the budget (if available)
        budget = movie.find('p', string=lambda text: "Budget:" in text)
        budget = budget.text.replace('Budget:', '').strip() if budget else "N/A"

        # Extract the gross revenue (if available)
        gross = movie.find('p', string=lambda text: "Gross:" in text)
        gross = gross.text.replace('Gross:', '').strip() if gross else "N/A"

        # Append the data to the list
        movies_data.append({
            'Title': title,
            'Year': year,
            'Director': director,
            'Budget': budget,
            'Gross Revenue': gross
        })

# Display the scraped data
for movie in movies_data:
    print(movie)
```

```python
In [66]: # Write data to a CSV file
         with open('movies_data.csv', 'w', newline='', encoding='utf-8') as file:
             writer = csv.DictWriter(file, fieldnames=['Title', 'Year', 'Director', 'Budget', 'Gross Revenue'])
             writer.writeheader()
             writer.writerows(movies_data)

         print("Data saved to movies_data.csv")
```

```
Data saved to movies_data.csv
```