

Automated IMDb Movie Data Scrapping and Analysis

Extract IMDb top 250 movie ratings and details using Python libraries.

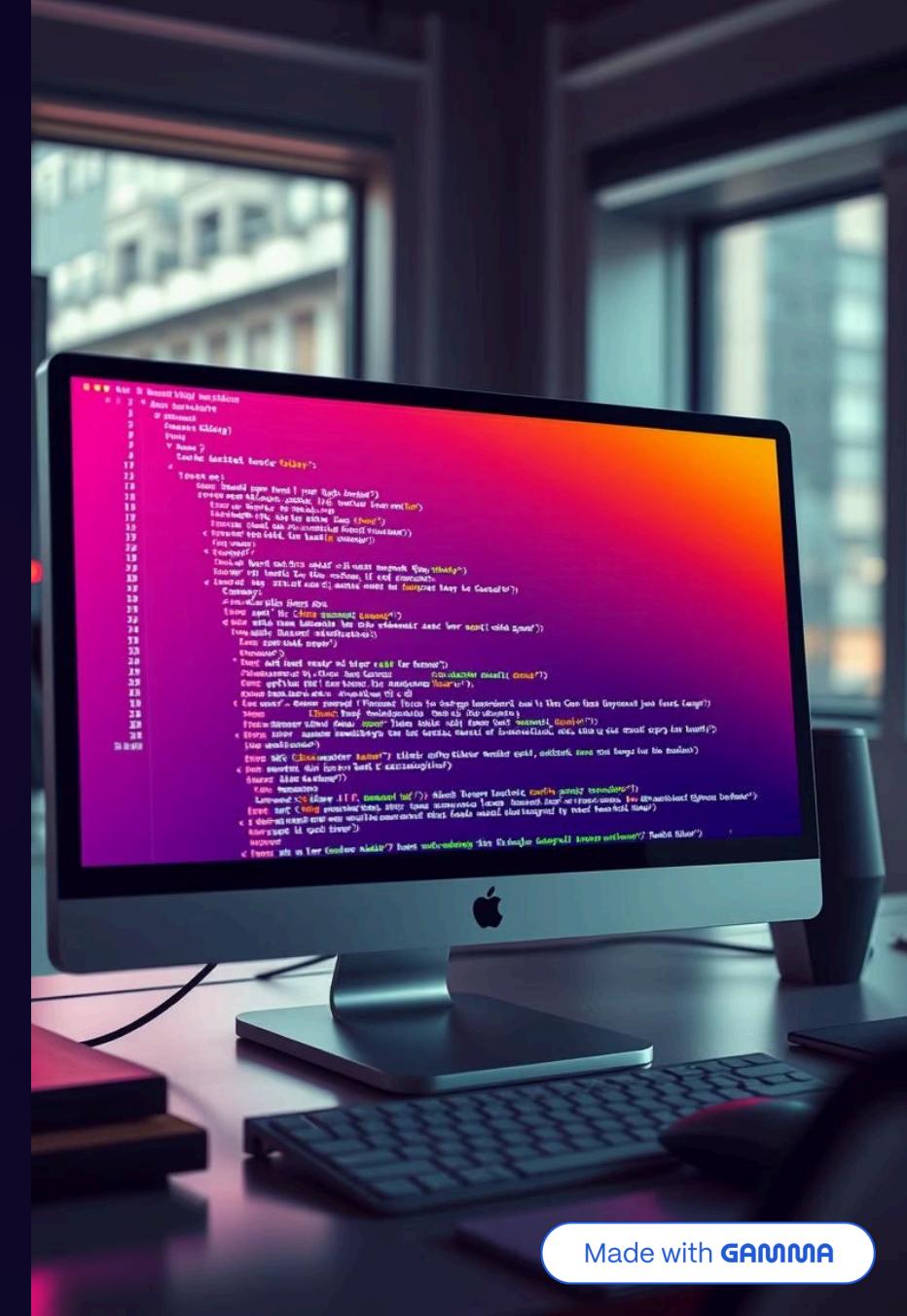
Presented by - Team Plot Twisters

Chirag kumar khuswah

Nishant Singh

Deepanshu

Harsh



Required Python Modules



requests

Make HTTP requests to URLs



html5lib

Parse HTML per web standards



bs4 (BeautifulSoup)

Extract data from HTML pages



pandas

Data manipulation and CSV export

```
77 commisien is fæst is beeldege tol:  
76     dest forestinent taylor();  
48     port li ts, "venet verdan;  
99     Meæssedes "Møglede",  
101     deti_rateing = (gille  
10     dass commisien looce (ð velcarue;  
17     cetsns (le tetar toll ve stras);  
78     rons biler prrlows apler (apcisest);  
46     (II;  
75     <ast (nette li roclacate; - wrbale siessource);  
49     <avtictest is resslovel);  
174     <ast ill seerehc: 15/74 (entl);  
733     )  
114     <ter -5 1db9 7;  
239     evet clerast, ();  
118     <ast (ast tatl netitnaleff)
```

```
<atr örofor -rawalle";  
<als apler lennerree (spilly);  
Cnlec:  
<itrefleectioetions: the_and_smtel li Cnalc);  
{  
    fall entle for lesction)  
    tadrata();  
    hrodrtandsice tater GND)  
    ent-oästerpiner (1);  
    cat:(is tr our langer factitas he fectapilly taels)  
    <Statister: Flser-Ind color
```

Accessing IMDb Webpage



Set URL

IMDb Top 250 movies page URL



Send Request

Use requests.get() to fetch HTML



Create Soup

Parse HTML with BeautifulSoup

Extracting Movie Data



Movies

Extract title columns from
HTML



Crew

Get star cast from title
attributes



Ratings

Extract rating values from HTML spans



Organizing Movie Details

Parse Movie Info

Separate place, title, year from text

Store Data

Save details in dictionaries inside a list

Frank Darabont (dir.), Tim Robbins, Morgan Freeman 8.9355175
Francis Ford Coppola (dir.), Marlon Brando, Al Pacino 9.0
Peter Jackson (dir.), Sean Connery, Ian McKellen 8.9355175
Christopher Nolan (dir.), Christian Bale, Heath Ledger 8.9355175
Henry Fonda, Lee J. Cobb 8.9355175
Steven Spielberg (dir.), Liam Neeson, Ralph Fiennes 8.9355175
Peter Jackson (dir.), Peter Jackson, Sam Rockwell 8.9355175
Sergio Leone (dir.), Clint Eastwood, Eli Wallach 8.9355175
Clint Eastwood, Brad Pitt 8.76284981
Tom Hanks, Robin Wright 8.75428571
Leonardo DiCaprio, Joseph Gordon-Levitt 8.75428571
Peter Jackson (dir.), Elijah Wood, Viggo Mortensen 8.75428571
Irvin Kershner (dir.), Keanu Reeves, Laurence Fishburne 8.75428571
Robert De Niro, Ray Liotta 8.64314049
Milos Forman (dir.), Jack Nicholson, Michael Caine 8.64314049
Toshiro Mifune, Takashi Shimura 8.591488615167
Morgan Freeman, Brad Pitt 8.591488615167
Roberto Benigni (dir.), Roberto Benigni, Nicoletta Braschi 8.591488615167
Alexandre Rodrigues, Leandro Hassum 8.591488615167
Jonathan Demme (dir.), Jodie Foster, Anthony Hopkins 8.591488615167
Frank Capra (dir.), James Stewart, Donna Reed 8.591488615167
Steven Spielberg (dir.), Tom Hanks, Matt Damon 8.591488615167
George Lucas (dir.), Mark Hamill, Harrison Ford 8.591488615167
Tom Hanks, Michael Clarke Duncan 8.591488615167
Daveigh Chase, Suzanne Pleshette 8.591488615167
Christopher Nolan (dir.), Matthew McConaughey, Anne Hathaway 8.591488615167
Kang-ho Song, Sun-kyun Lee 8.528458916667

Displaying Movie Data

Print Detail

Show place, title, year, stars, rating

Example Output

Formatted movie info in terminal

Data	Review	View
General		
Alignment	Number	Styles

D	E
year	star_cast
1994	Frank Darabont (dir.), Tim Robbins, Morgan Freeman
1972	Francis Ford Coppola (dir.), Marlon Brando, Al Pacino
2008	Christopher Nolan (dir.), Christian Bale, Heath Ledger
1974	Francis Ford Coppola (dir.), Al Pacino, Robert De Niro
1957	Sidney Lumet (dir.), Henry Fonda, Lee J. Cobb
1993	Steven Spielberg (dir.), Liam Neeson, Ralph Fiennes
2003	Peter Jackson (dir.), Elijah Wood, Viggo Mortensen
1994	Quentin Tarantino (dir.), John Travolta, Uma Thurman
2001	Peter Jackson (dir.), Elijah Wood, Ian McKellen
1966	Sergio Leone (dir.), Clint Eastwood, Eli Wallach
1994	Robert Zemeckis (dir.), Tom Hanks, Robin Wright
1999	David Fincher (dir.), Brad Pitt, Edward Norton
2010	Christopher Nolan (dir.), Leonardo DiCaprio, Joseph Gordon-Levitt
2002	Peter Jackson (dir.), Elijah Wood, Ian McKellen
1980	Irvin Kershner (dir.), Mark Hamill, Harrison Ford
1999	Lana Wachowski (dir.), Keanu Reeves, Laurence Fishburne
1990	Martin Scorsese (dir.), Robert De Niro, Ray Liotta
1975	Milos Forman (dir.), Jack Nicholson, Louise Fletcher
1995	David Fincher (dir.), Morgan Freeman, Brad Pitt
1954	Akira Kurosawa (dir.), Toshirô Mifune, Takashi Shimura
1946	Frank Capra (dir.), James Stewart, Donna Reed
1991	Jonathan Demme (dir.), Jodie Foster, Anthony Hopkins
1998	Steven Spielberg (dir.), Tom Hanks, Matt Damon
2002	Fernando Meirelles (dir.), Alexandre Rodrigues, Leandro Firmino

Saving Data to CSV

Create DataFrame

Convert list of dicts to pandas DataFrame

Export CSV

Save DataFrame to 'imdb_top_250_movies.csv'

Summary & Benefits

Automate Data Collection

Scrape IMDb ratings efficiently



Structured Data

Organize movie info for analysis

Easy Export

Save data for future use or sharing