Optimizing Film Selection: A Cutting-Edge Movie Recommendation System

Key Objectives

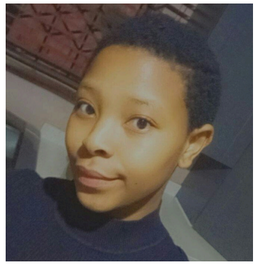
- Recommendation Accuracy
- Genre-Based Recommendations
- Profitability Analysis
- Language-Based Insights
- Actor and Genre Influence

Expected outcome

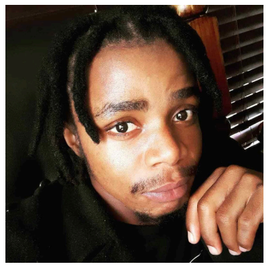
- A robust recommendation engine that enhances user satisfaction and engagement.
- Insights into correlations between movie attributes (e.g., genre, language, actors) and financial success.
- A platform that provides accurate and relevant movie suggestions, contributing to revenue generation.

Meet the team

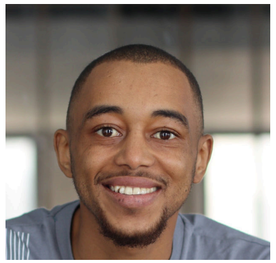
Neo Motaung



Sanelisiwe Ndlovu



Sonwabile Ximbi



Yamkela Jojo

Zenani Manqele

User Persona

Bio

Sarah is a film enthusiast who enjoys spending her free time watching movies. She often struggles to find new films that match her taste and preferences.

Pain

- Wasting time on films she doesn't enjoy
- Feeling frustrated with generic movie recommendations
- Missing out on hidden gems due to lack of personalized suggestions

Sortify Goal

To provide a movie recommendation system that can accurately predict her movie preferences and provide her with tailored suggestions for an enjoyable viewing experience.

Sortify Benefits

- Discovering new favorite movies effortlessly
- Saving time searching for films to watch
- Enhanced movie-watching experience with personalized recommendations

The Data

Dataset Overview:

- Consists of several million 5-star ratings from users of MovieLens.

Special Version:

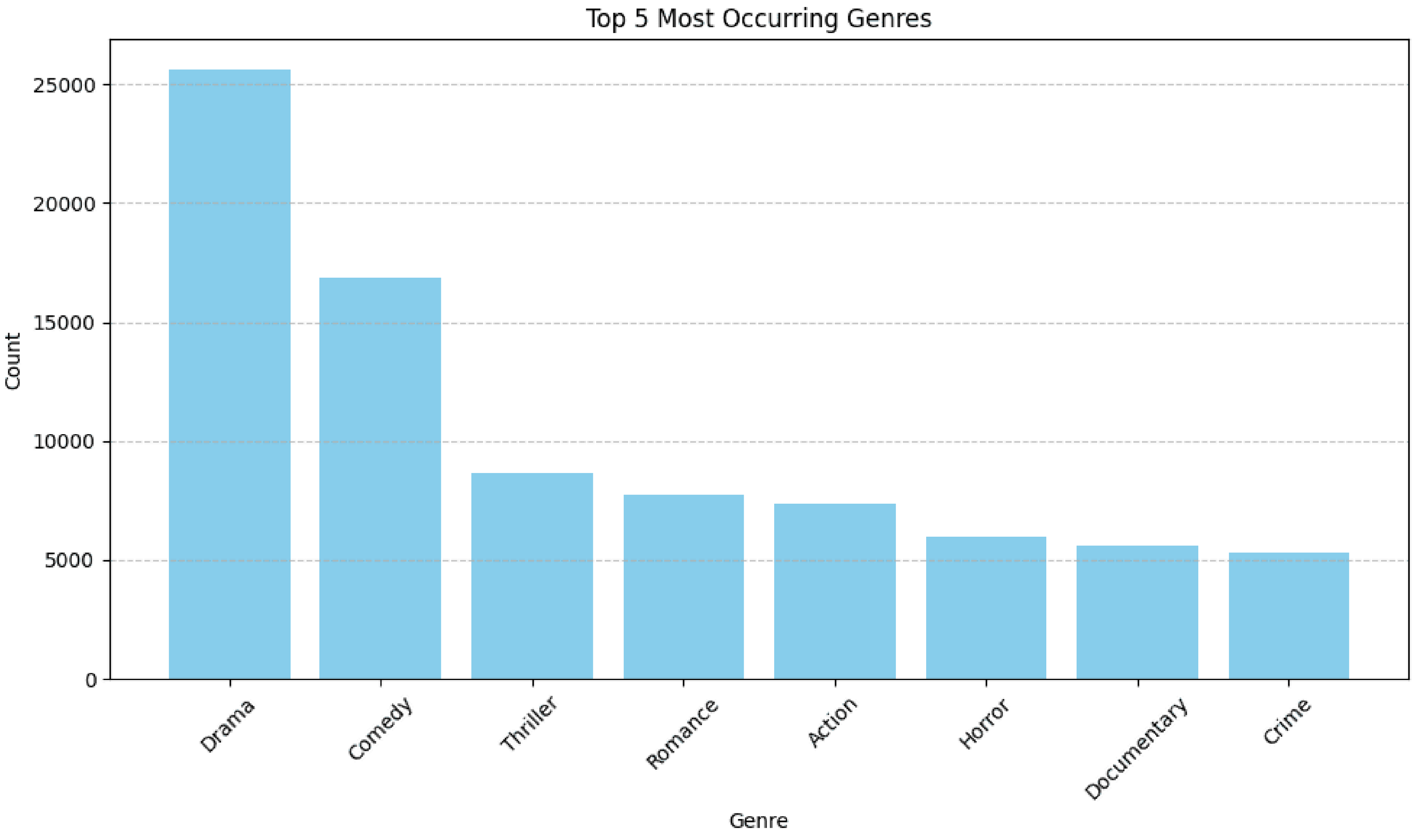
- Enriched with additional data.

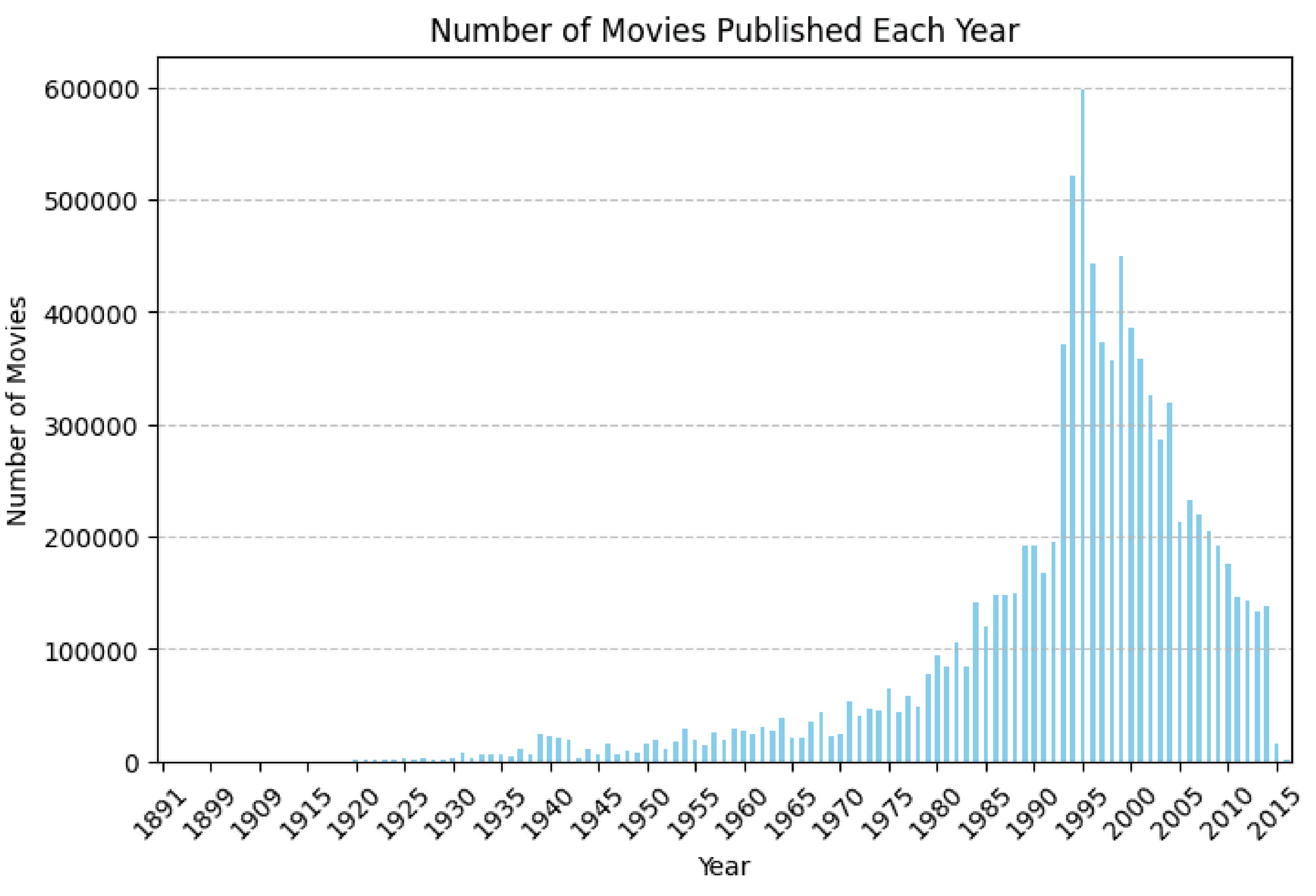
Source:

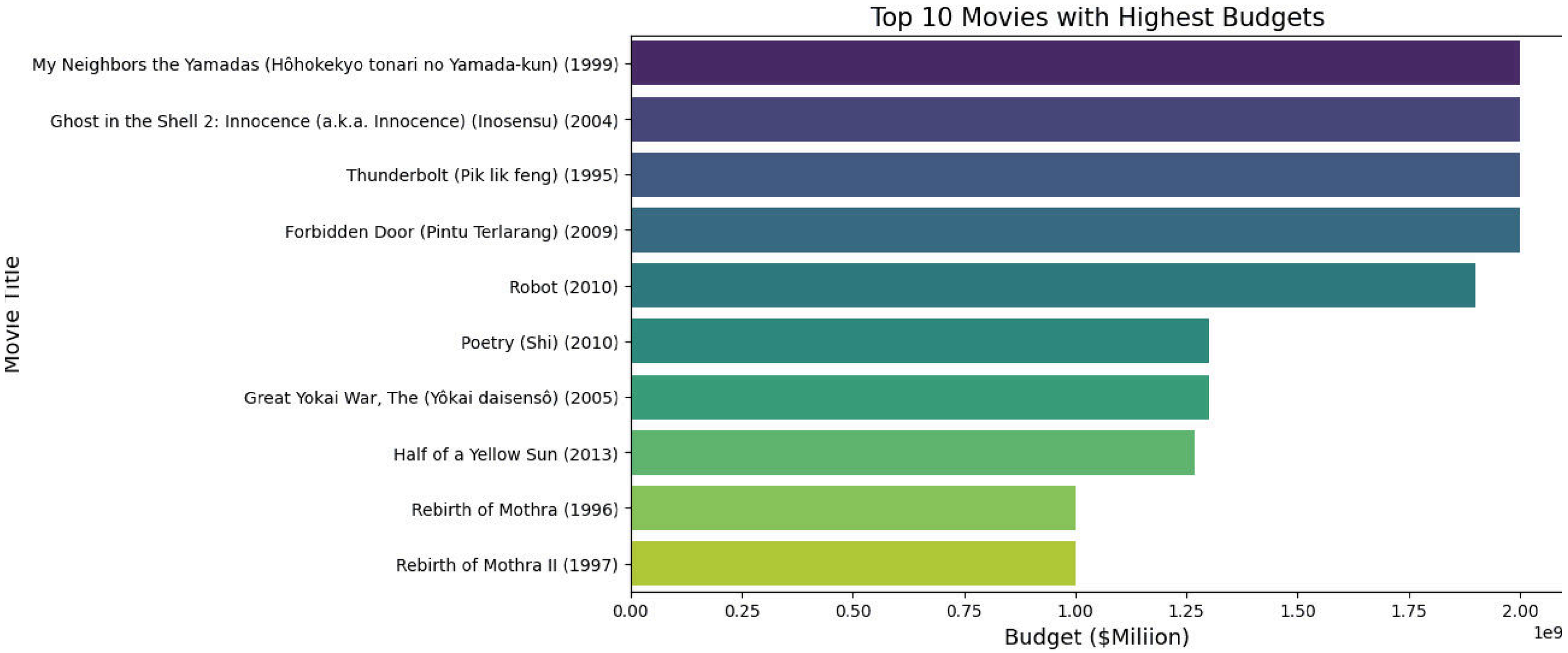
- Maintained by GroupLens research group at University of Minnesota.

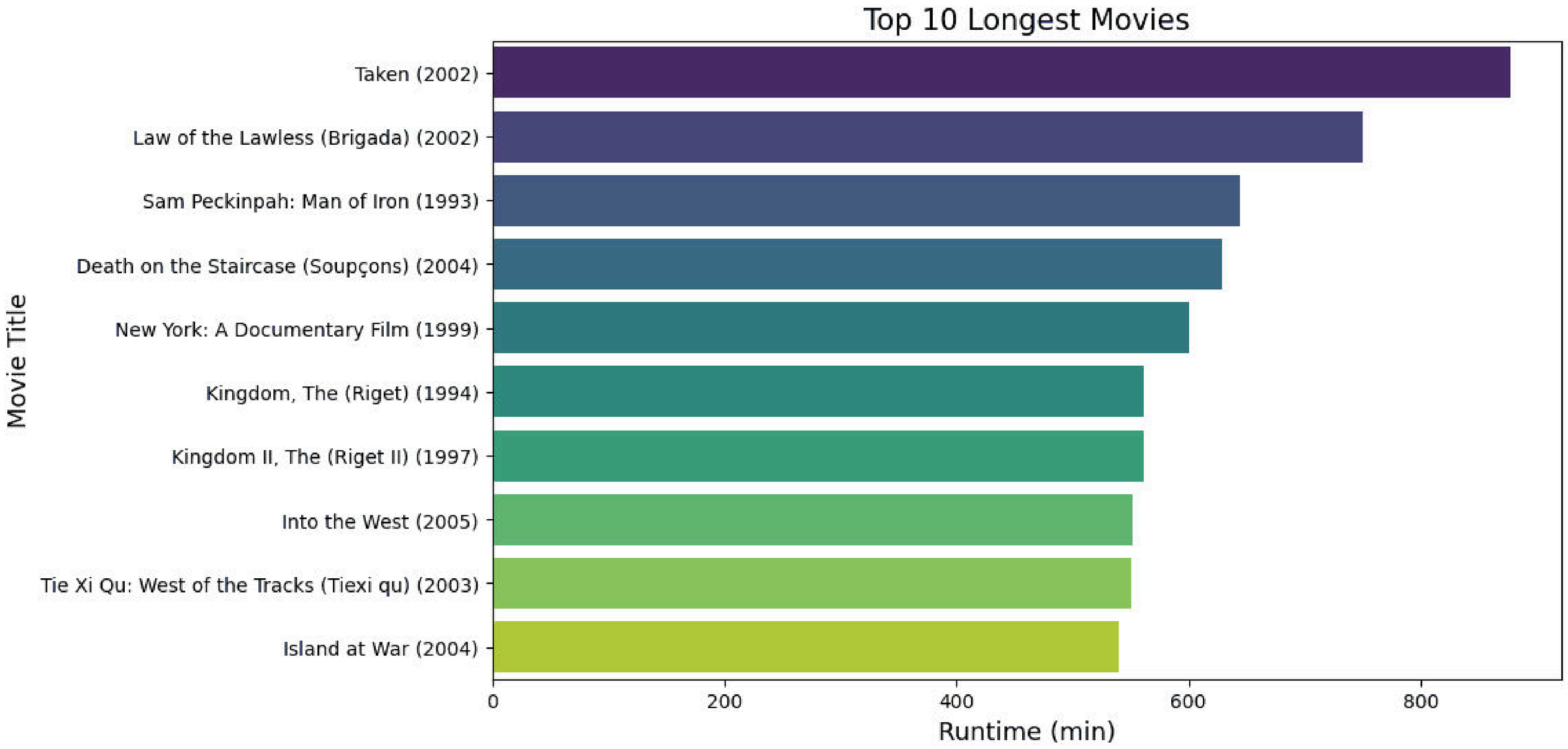
Data Preprocessing

- We checked for missing values and duplicates duplicates.
- We made and used a copy of the train data
- We merged datasets


Top 5 Most Occurring Genres
Count
25000
20000
15000
10000
5000
0
Drama
Comedy
Thriller
Romance
Action
Horror
Documentary
Crime
Genre


Number of Movies Published Each Year
Number of Movies
600000
500000
400000
300000
200000
100000
0
1891
1899
1909
1915
1920
1925
1930
1935
1940
1945
1950
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
Year

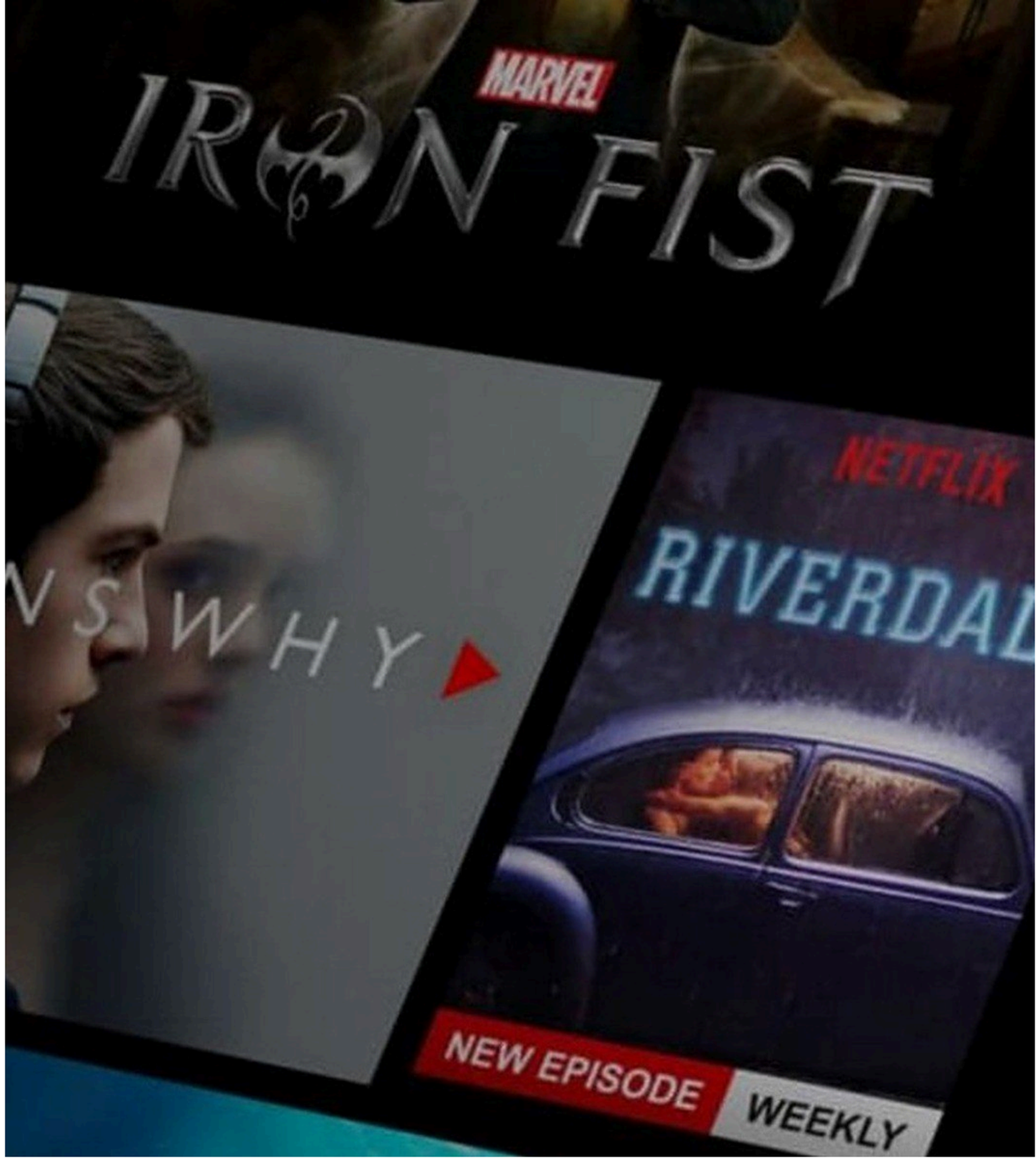

Top 10 Movies with Highest Budgets
My Neighbors the Yamadas (Hôhokekyo tonari no Yamada-kun) (1999)
Ghost in the Shell 2: Innocence (a.k.a. Innocence) (Inosensu) (2004)
Thunderbolt (Pik lik feng) (1995)
Forbidden Door (Pintu Terlarang) (2009)
Robot (2010)
Poetry (Shi) (2010)
Great Yokai War, The (Yôkai daisensô) (2005)
Half of a Yellow Sun (2013)
Rebirth of Mothra (1996)
Rebirth of Mothra II (1997)
Movie Title
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
1e9
Budget ($Miliion)

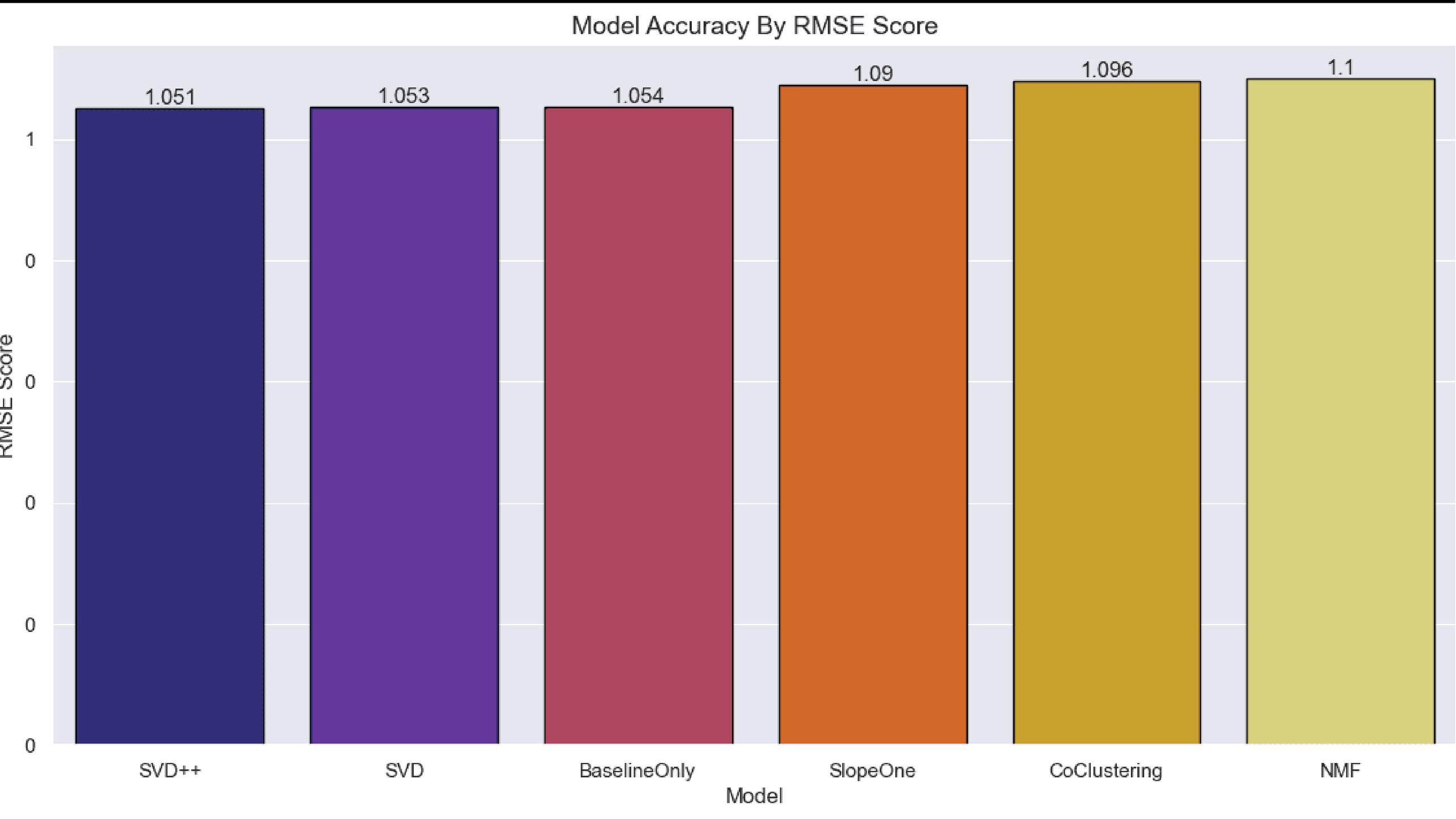

Top 10 Longest Movies
Taken (2002)
Law of the Lawless (Brigada) (2002)
Sam Peckinpah: Man of Iron (1993)
Death on the Staircase (Soupçons) (2004)
New York: A Documentary Film (1999)
Kingdom, The (Riget) (1994)
Kingdom II, The (Riget II) (1997)
Into the West (2005)
Tie Xi Qu: West of the Tracks (Tiexi qu) (2003)
Island at War (2004)
Movie Title
0
200
400
600
800
Runtime (min)

Stand-Out Words Present on Movie Titles

Collaborative-based filtering

- SVD (Singular Value Decomposition)
- SVD++
- BaselineOnly
- SlopeOne
- CoClustering
- NMF (Non-Negative Matrix Factorization)




Model Accuracy By RMSE Score
1.051
1.053
1.054
1.09
1.096
1.1
1
0
0
0
0
RMSE Score
SVD++
SVD
BaselineOnly
SlopeOne
CoClustering
NMF
Model

Live Demonstration

Conclusion

Sortify's groundbreaking movie recommendation application marks a significant leap forward in the realm of digital entertainment.