

Ratings, Reviews, and Where to Watch
the Best Movies & TV Shows

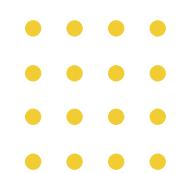
IMDb is the world's most popular and authoritative
source for movie, TV and celebrity content. Find...



IMDb Insights:

Analisis dan Visualisasi Data Film dari IMDb

Kelompok 2 - P2



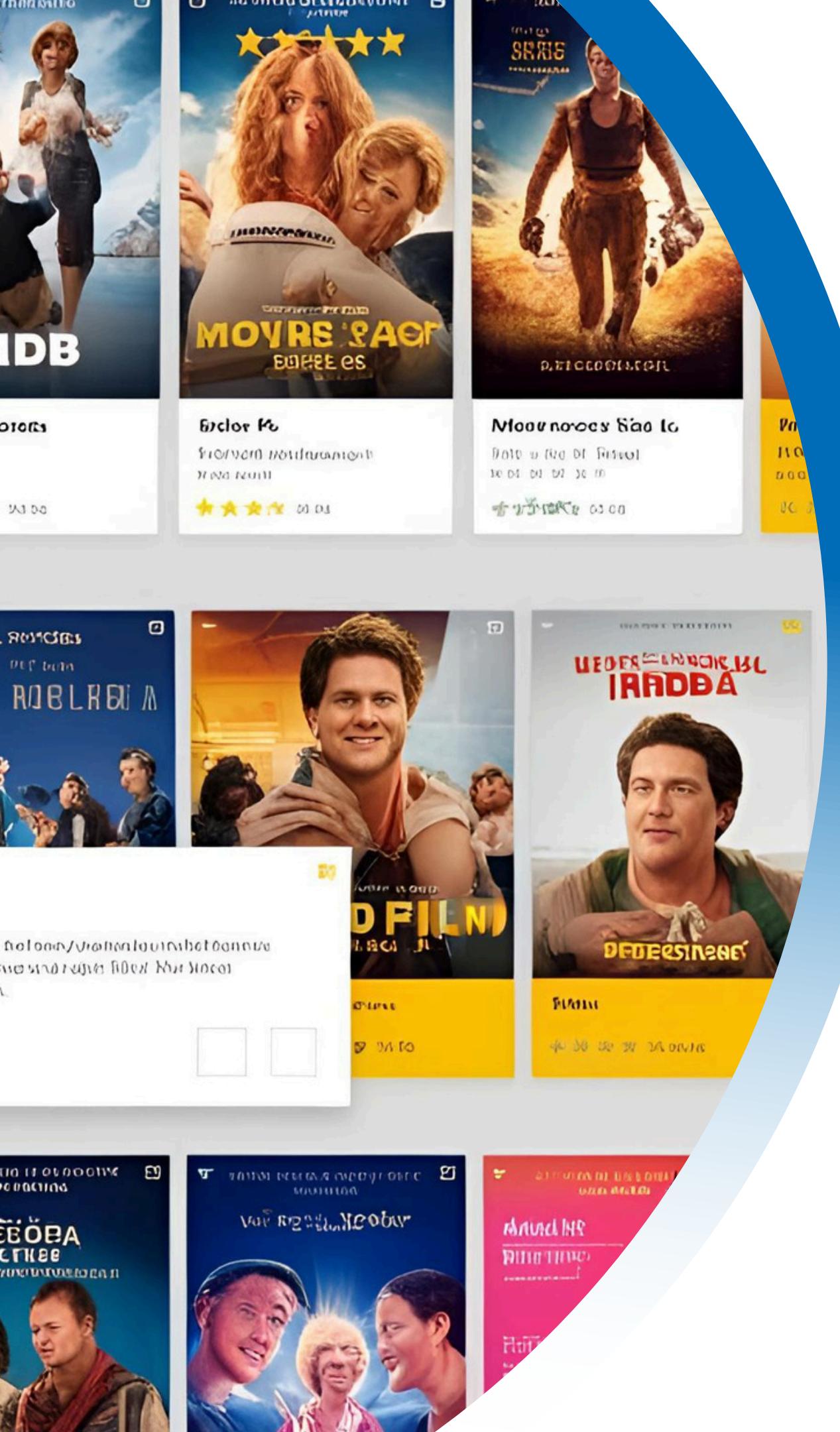
Meet Our Team

Dela
Gustiara

Rizka Annisa
Mingka

Wildatul
Maulidiyah

Windi
Pangesti

**IMDb***Internet Movie Database*

Merupakan situs web yang **menyediakan informasi lengkap tentang film**, serial televisi, video game, dan konten hiburan lainnya.

Didirikan pada tahun 1990 dan kini dimiliki oleh Amazon, IMDb digunakan oleh **jutaan pengguna** untuk mencari informasi, menilai, dan membahas film.

Situs ini memuat data seperti judul, rating, ulasan, genre, pemeran, serta kru produksi.

Tahapan yang dilakukan

Scraping data
pada web <https://www.imdb.com/>
dengan bantuan web ekstensi *data miner*



Cleaning Data Hasil Scrapping
Membersihkan data hasil scrapping
pada setiap variabel



Penyimpanan data ke Mongo DB
untuk pengelolaan yang fleksibel
dan terstruktur.



Visualisasi
Melakukan analisis dan
visualisasi menggunakan Rstudio



Data Film yang Diambil dari IMDb

Data yang diambil merupakan data film skala global 5 tahun terakhir (2020 - 2025)

Variabel	Deskripsi
Judul	Nama atau judul resmi dari film seperti yang tercantum di IMDb.
Tahun rilis	Tahun pertama kali film tersebut dirilis atau ditayangkan.
Durasi film	Lama waktu penayangan film dalam satuan menit.
Rating	Skor penilaian film berdasarkan pengguna IMDb dalam skala 1-10.
Jumlah vote	Jumlah total suara atau penilaian yang diberikan oleh pengguna IMDb.
Kategori	Klasifikasi usia penonton berdasarkan isi film, seperti R, PG-13.



Insight yang Diharapkanya



Ratings, Reviews, and Where to Watch
the Best Movies & TV Shows

IMDb is the world's most popular and authoritative
source for movie, TV and celebrity content. Find...

IMDb

1

Profil Film Berdasarkan Tahun dan Kategori

Profil Film Berdasarkan Tahun dan Kategori serta distribusi berdasarkan rating usia (PG, PG-13, R, dll.),

2

Tren jumlah film rilis 5 tahun terakhir

Melihat bagaimana jumlah film yang rilis dari waktu ke waktu

3

Analisis Popularitas Film:

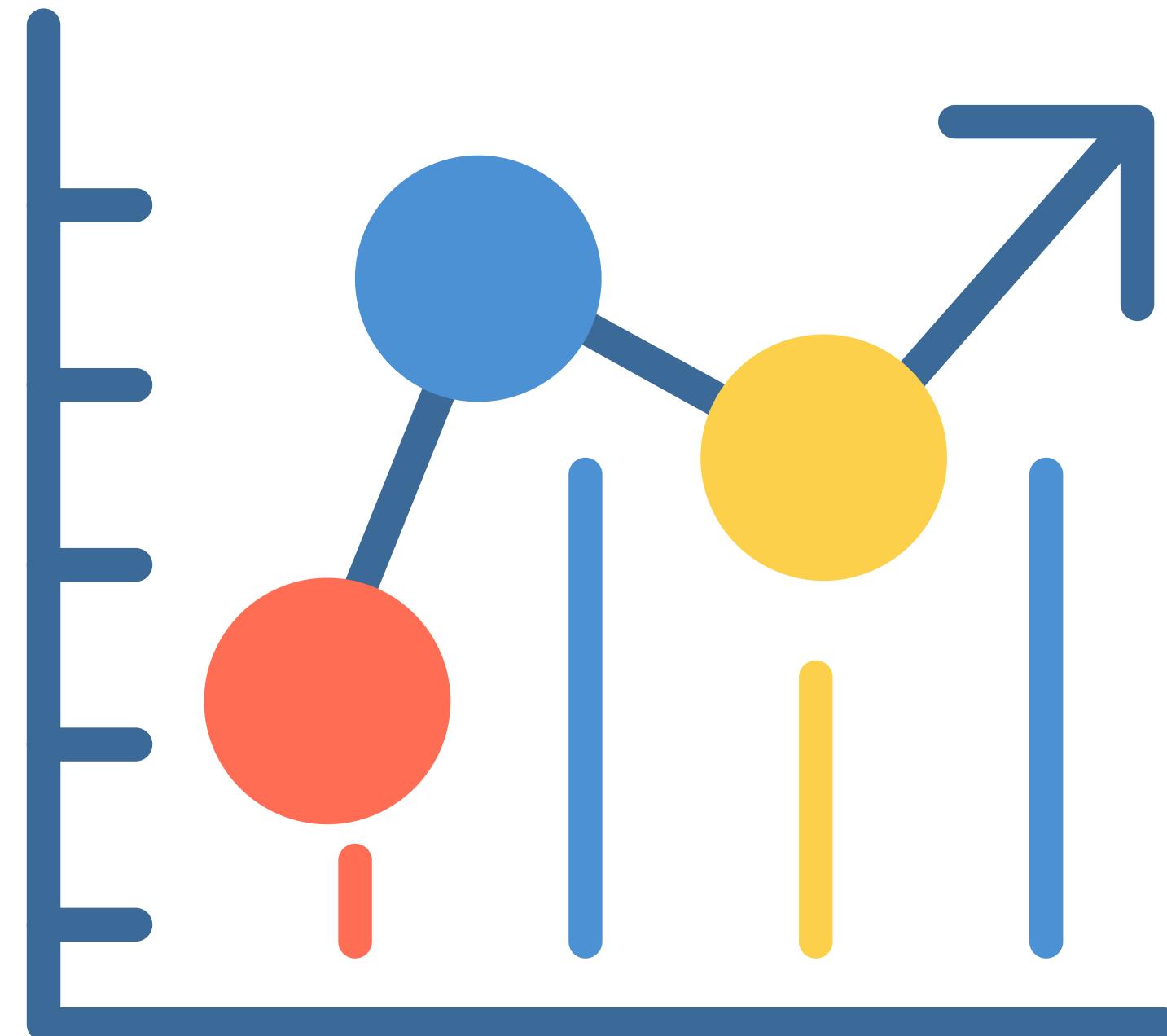
melihat film berdasarkan rating tertinggi dan jumlah vote terbanyak.

4

Pengaruh durasi dan jumlah vote terhadap rating

Mengkaji hubungan durasi dan jumlah vote terhadap rating film

HASIL **VISUALISASI**



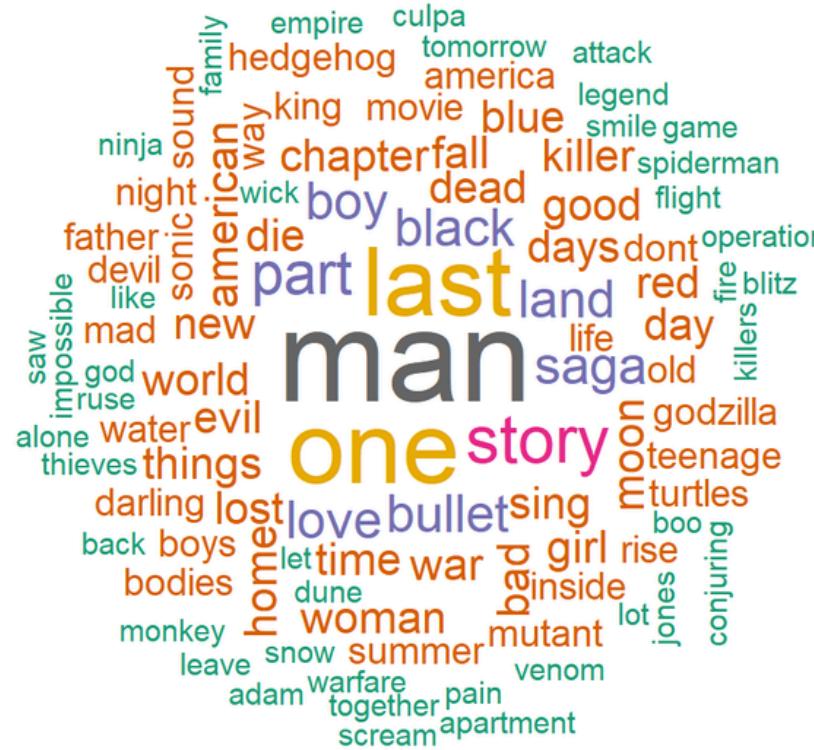
Informasi Umum Dataset

TOTAL FILM	JENIS KATEGORI	RATA-RATA RATING	RATA-RATA VOTE	RATA-RATA DURASI
645	10	6,46	92167	115,28

VISUALISASI

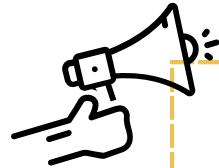
WORDCLOUD

JUDUL FILM



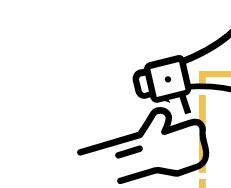
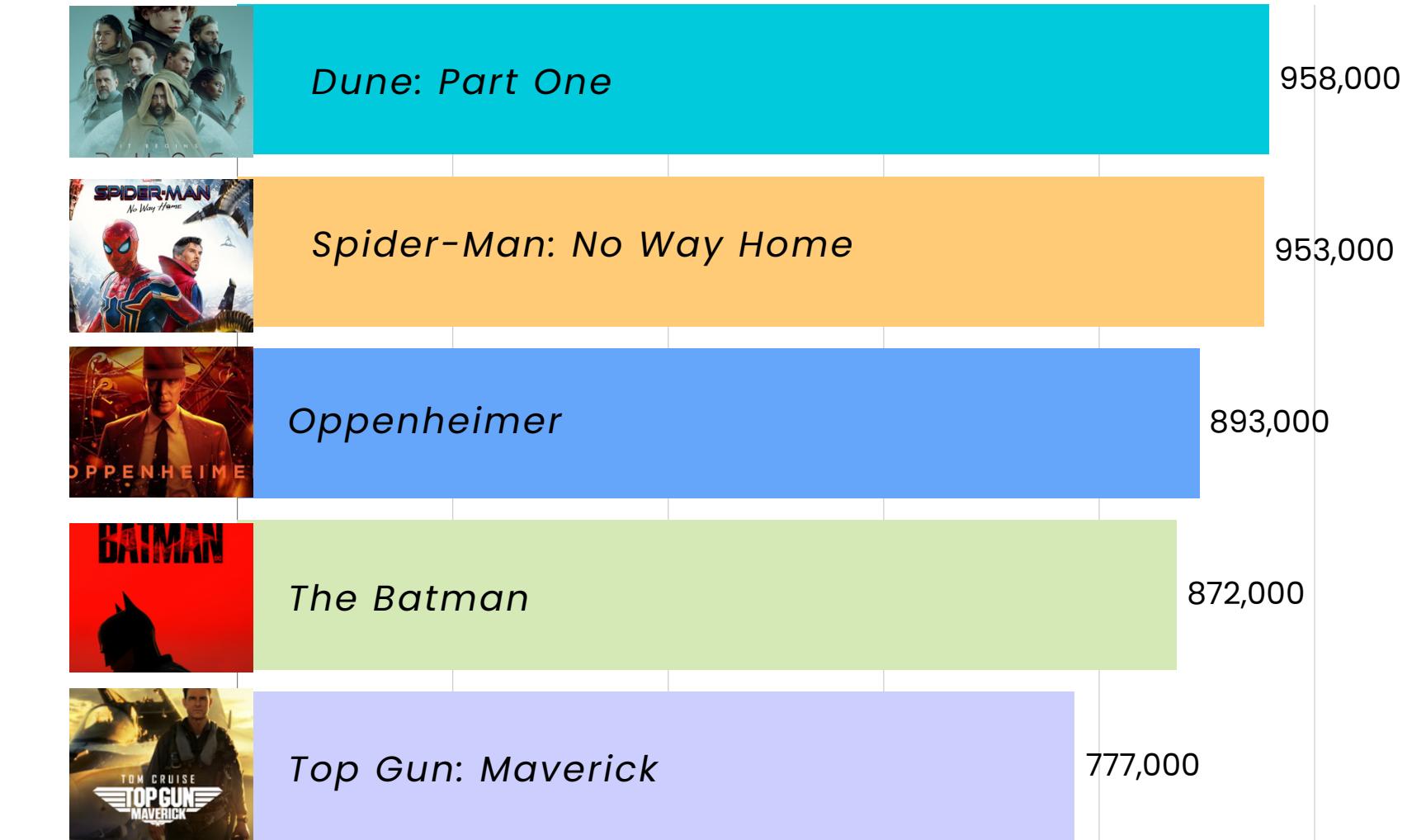
Visualisasi word cloud ini menunjukkan kata-kata yang paling sering digunakan dalam judul film selama lima tahun terakhir. Kata seperti "**man**", "**one**", "**last**", dan "**story**" mendominasi

TOP 5 FILM BERDASARKAN RATING



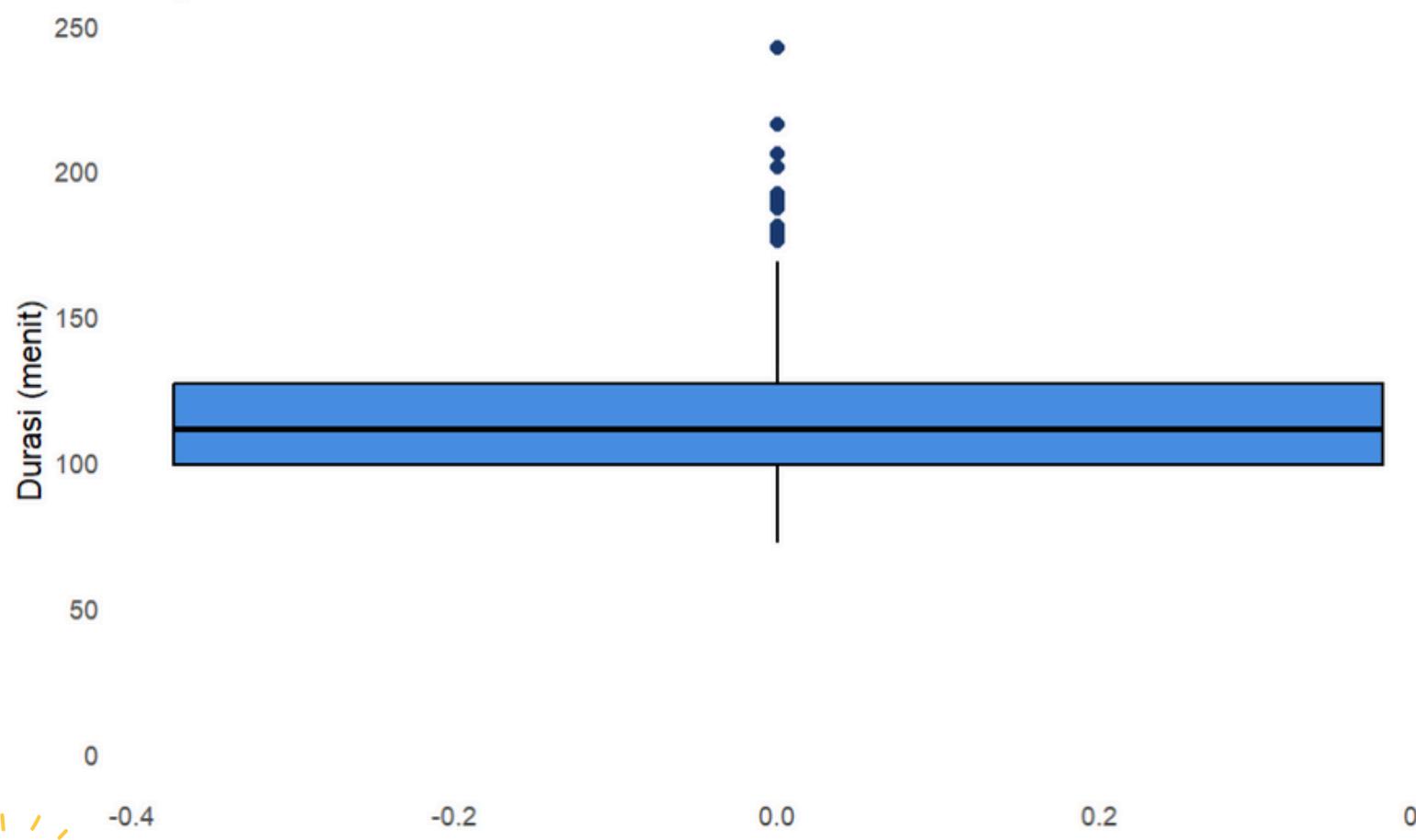
- **Guru Nanak Jahaz** mendapat rating tertinggi yaitu 9.4.
- Film-film dalam daftar ini memiliki rating **minimal 8.7**, menunjukkan **apresiasi tinggi** dari penonton.

TOP 5 FILM BERDASARKAN JUMLAH VOTE



- **Dune: Part One dan Spider-Man: No Way Home** mendapat lebih dari 950.000 vote.
- Film dengan jumlah vote tinggi menunjukkan **tingkat popularitas** yang besar secara global.

BOXPLOT SEBARAN DURASI FILM



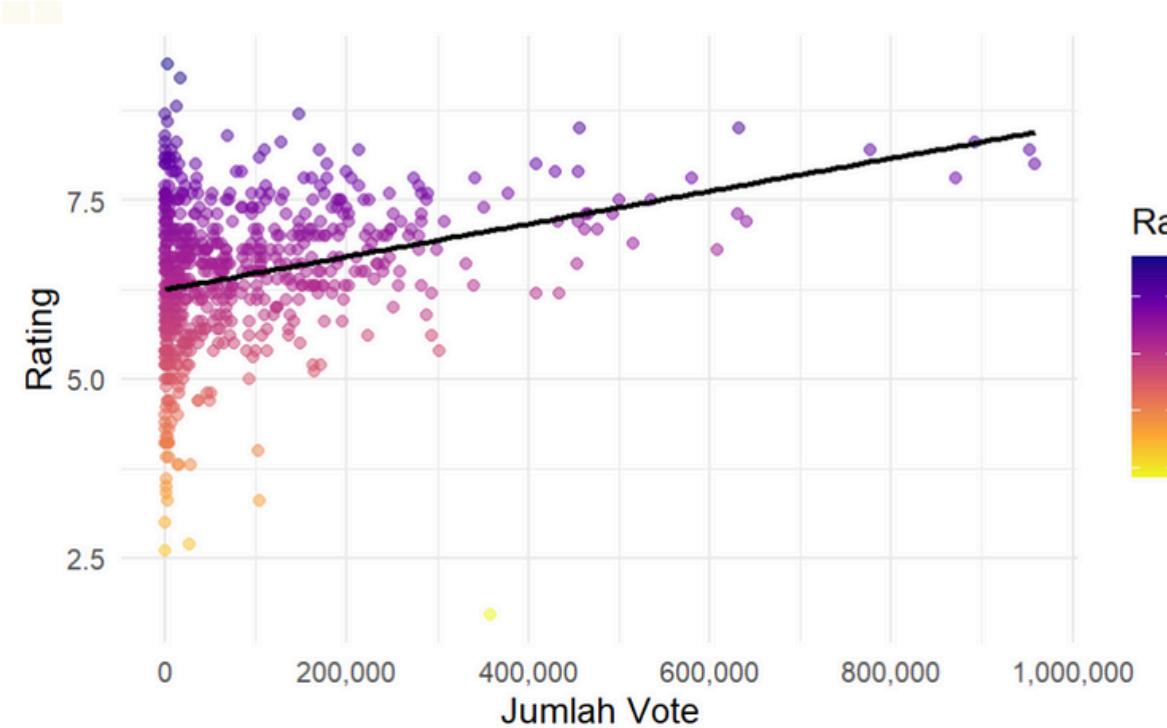
- Mayoritas durasi film berkisar **antara 90 hingga 140 menit**
- Terdapat outlier berupa film berdurasi sangat panjang, **melebihi 200 menit**

TOP 3 FILM DURASI TERLAMA

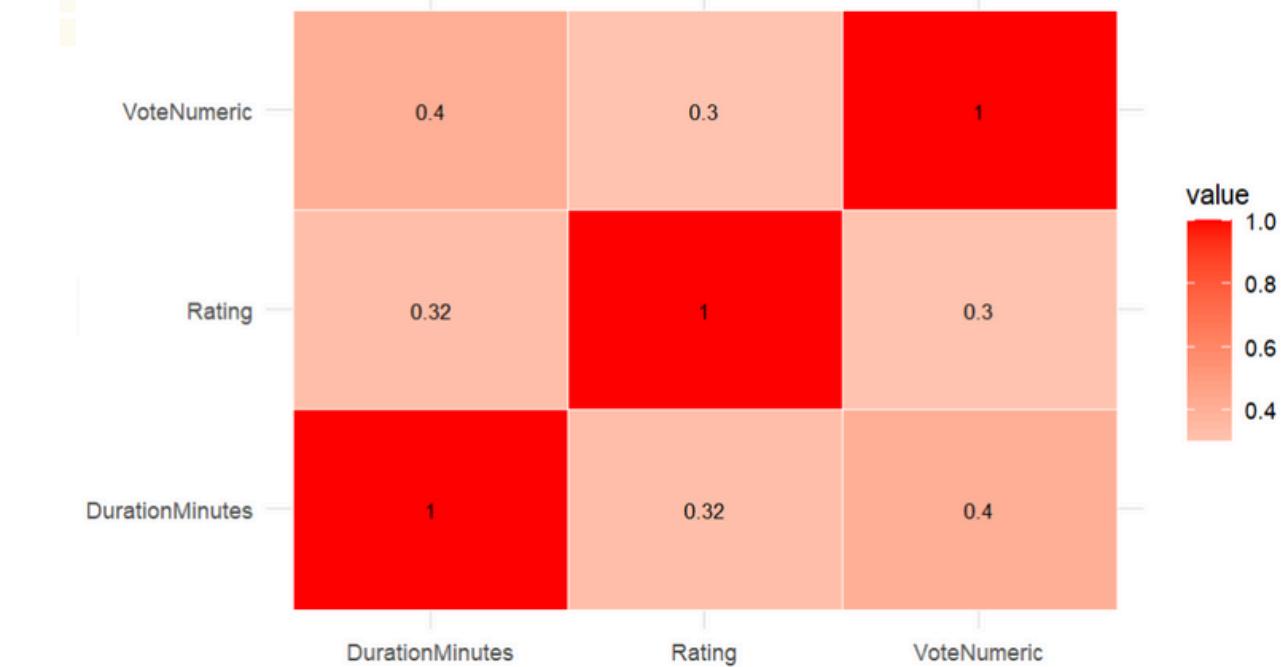


- **Zack Snyder's Justice League** menjadi film terlama dengan durasi 242 menit
- Ketiga film tersebut jauh di atas **rata-rata durasi umum** film lainnya

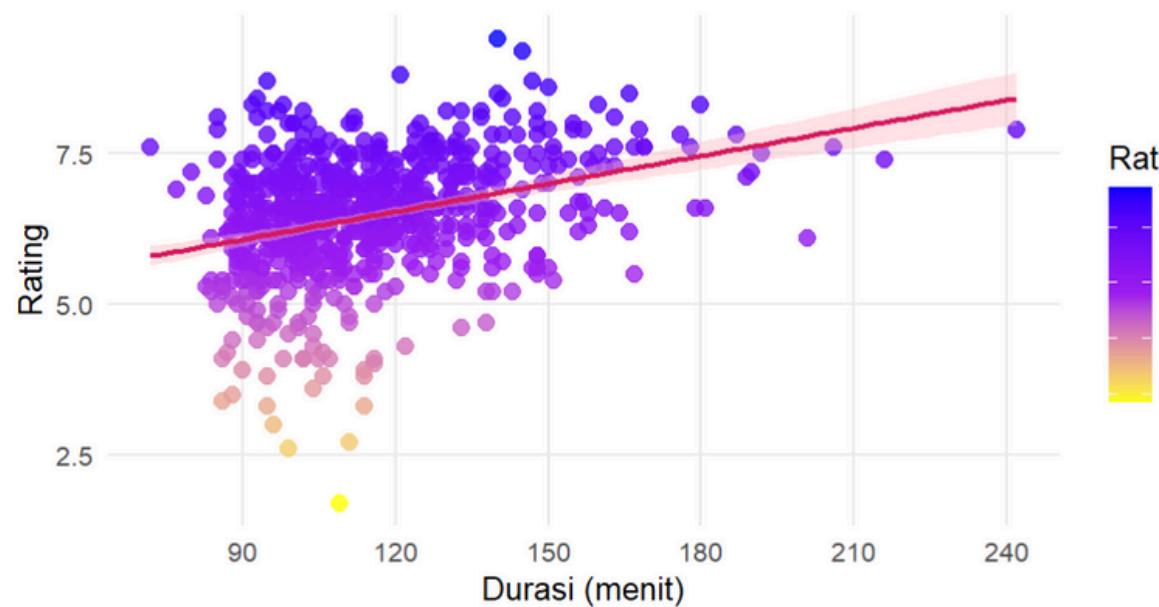
KORELASI JUMLAH VOTE VS RATING



HEAT MAP VARIABEL NUMERIK



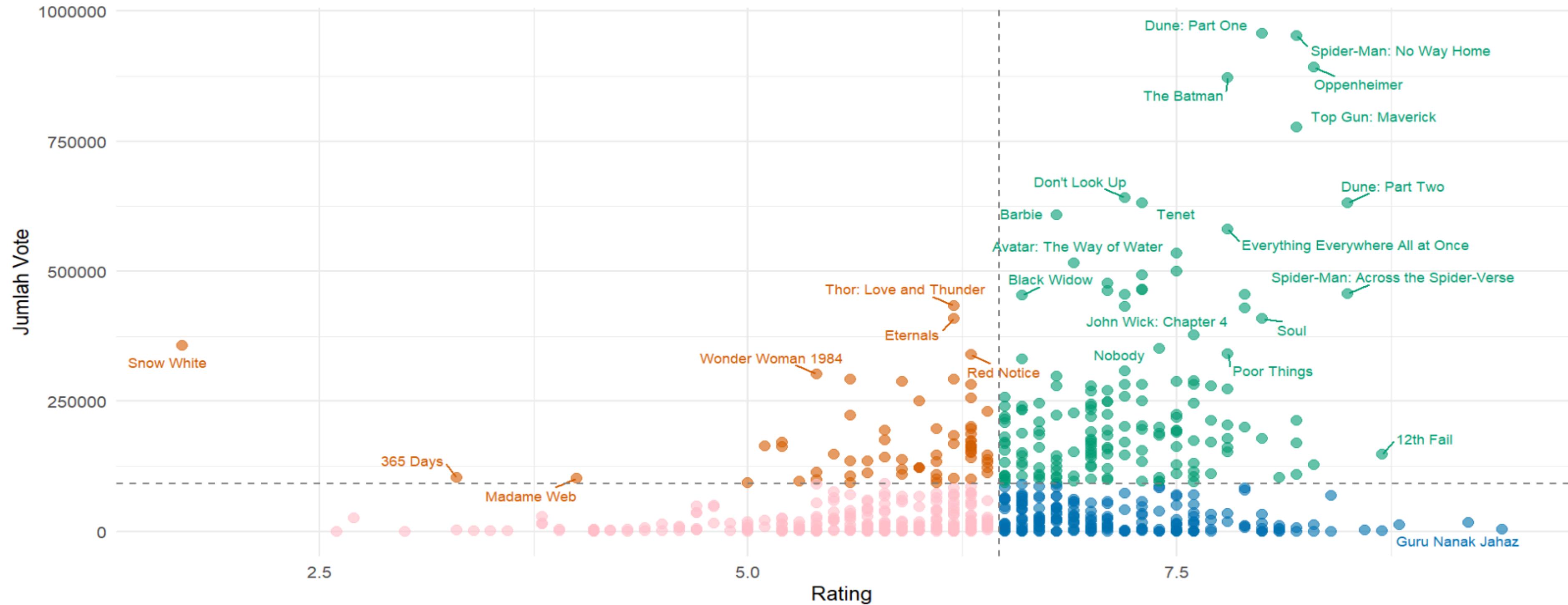
KORELASI DURASI VS RATING



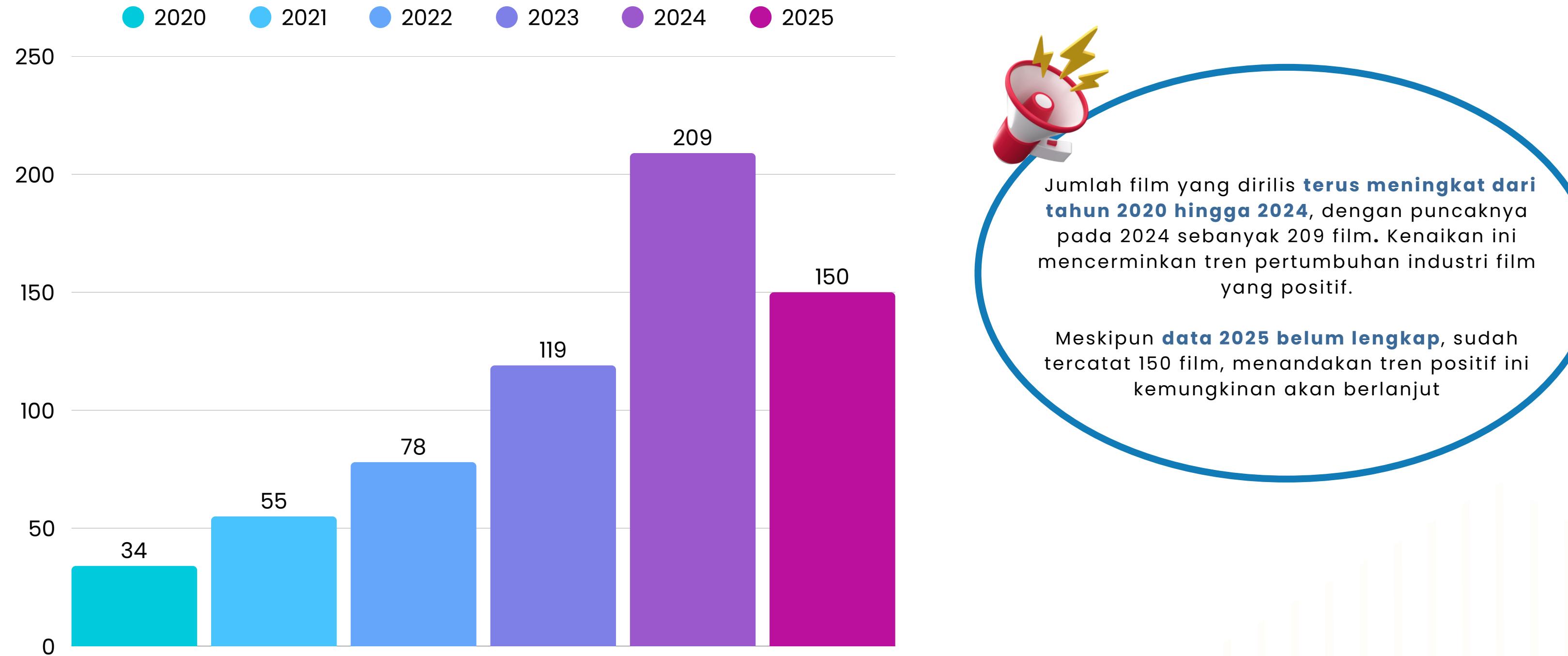
- Semakin banyak jumlah vote, biasanya rating film semakin tinggi.
- Film yang lebih panjang durasinya cenderung memiliki rating lebih tinggi.
- Artinya, **ada hubungan positif jumlah vote dan durasi film terhadap jumlah vote**
- Namun, hubungan tersebut tergolong **lemah**, karena berdasarkan heatmap korelasi hanya sekitar 0.3 (jumlah vote) dan 0.3 (durasi).
- Kedua variabel memang berpengaruh, tetapi sepertinya bukan faktor utama yang menentukan rating film..

PEMETAAN KUADRAN

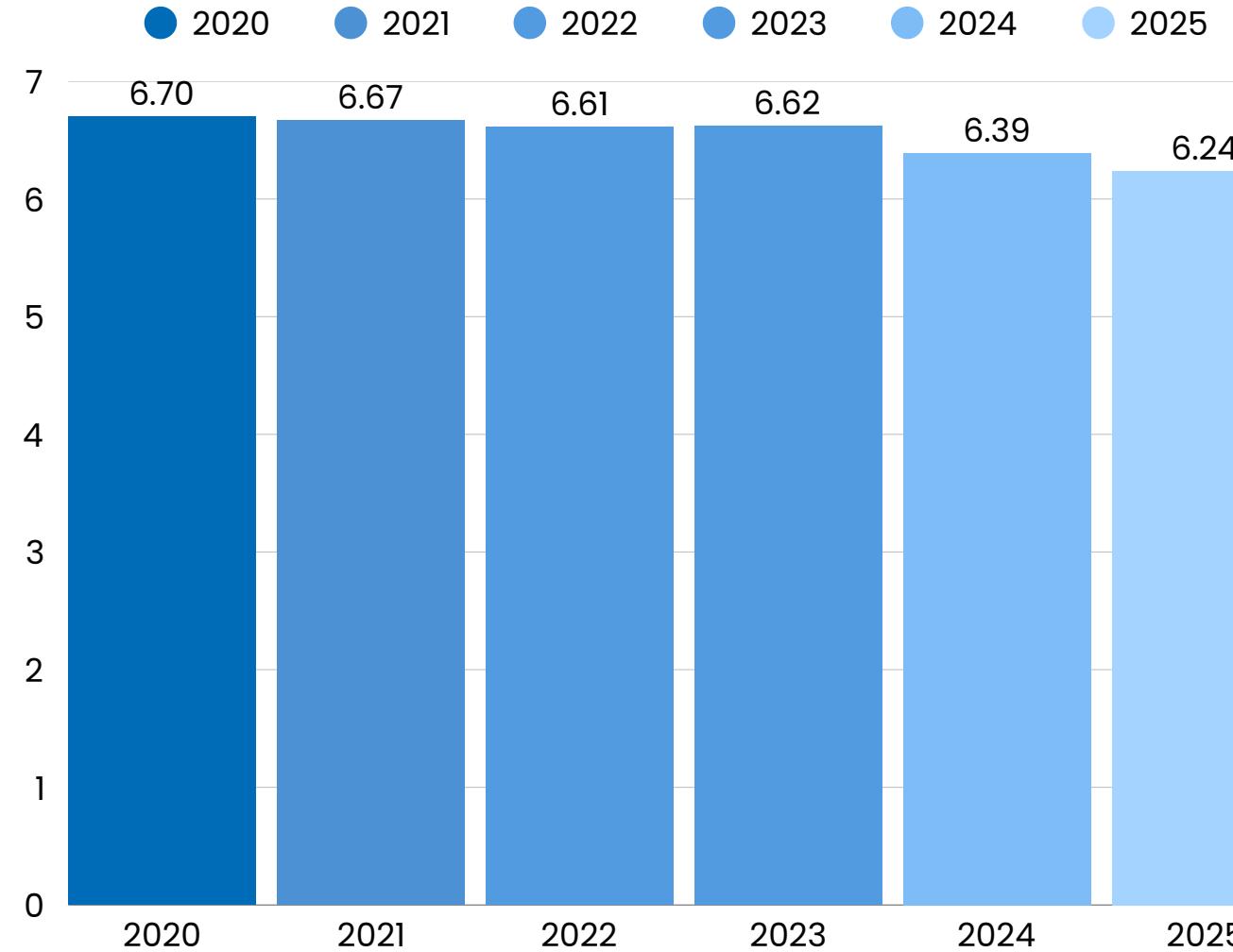
POPULARITAS DAN KUALITAS



JUMLAH FILM RILIS PER TAHUN

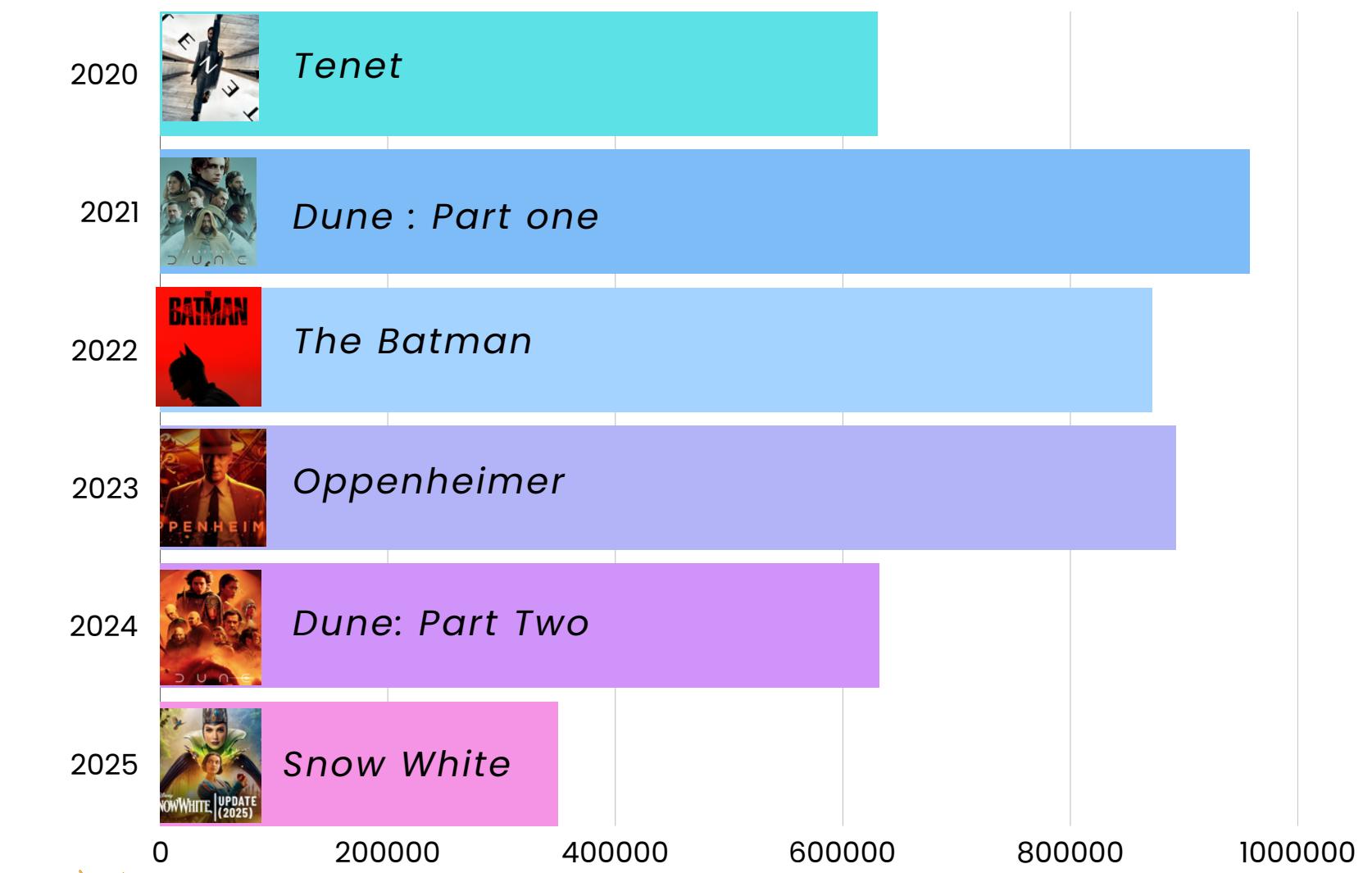


RATA-RATA RATING FILM PER TAHUN



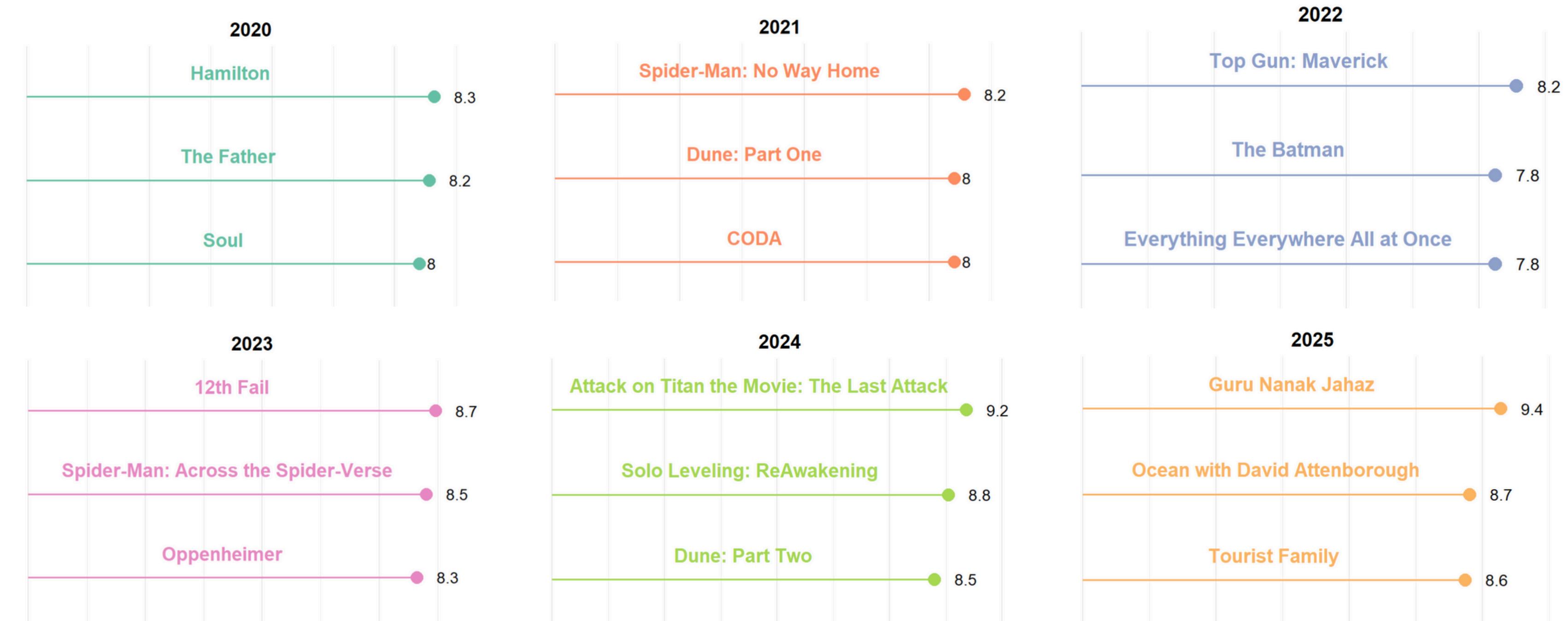
- Rata-rata rating film mengalami tren **penurunan dari tahun 2020 hingga 2025**.
- Pada tahun 2020, **rating tertinggi tercatat sebesar 6.7**, namun secara bertahap **menurun hingga mencapai 6.24** pada tahun 2025.

FILM DENGAN VOTE TERTINGGI PER TAHUN

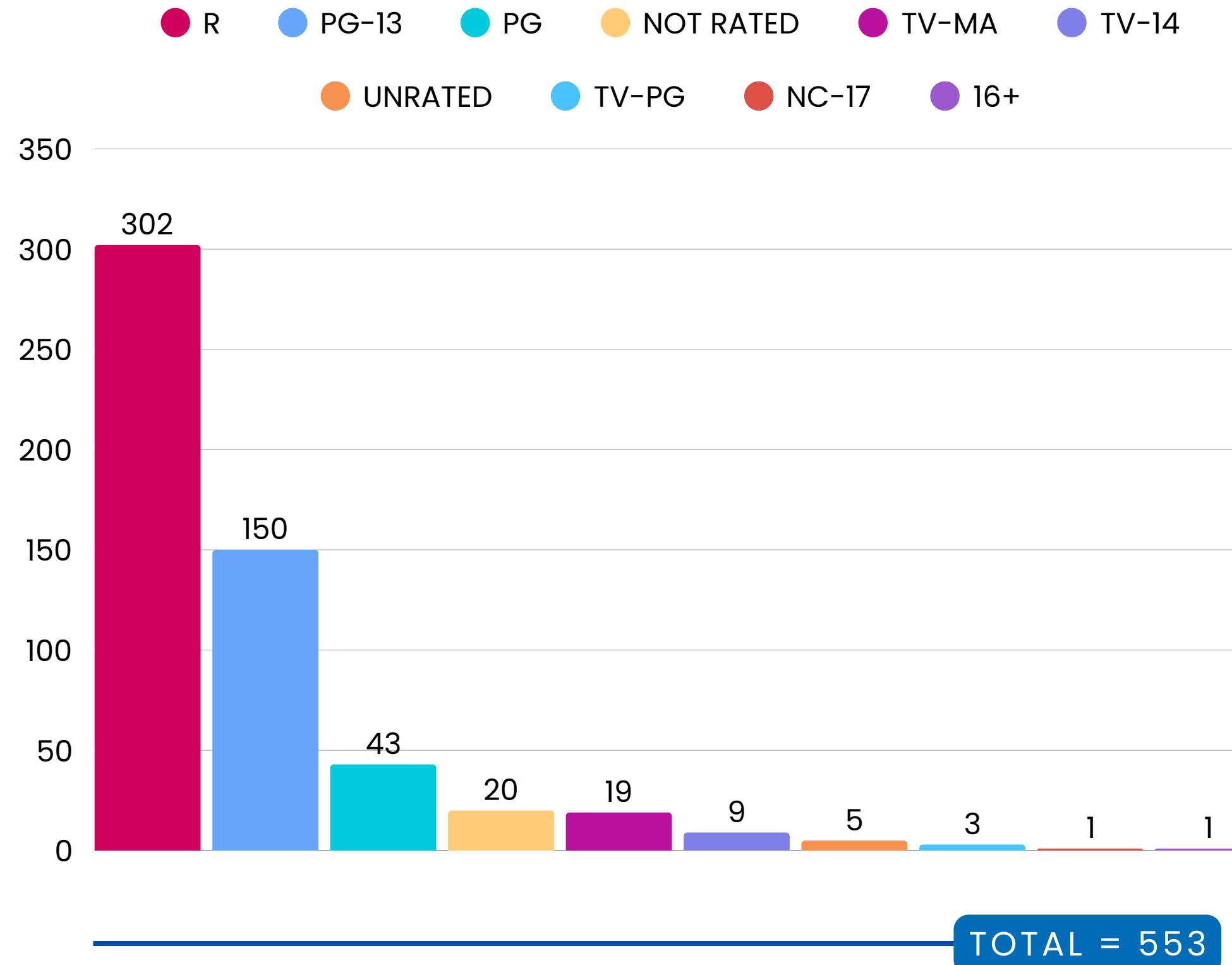


- Setiap tahun, terdapat **satu film yang paling banyak ditonton** dengan jumlah vote tertinggi.
- Film **Dune: Part One (2021) dan Oppenheimer (2023)** menjadi yang paling banyak dipilih oleh penonton.
- **Snow White (2025)** memiliki jumlah vote **paling rendah** dibanding tahun-tahun sebelumnya.

TOP 3 FILM DENGAN RATE TERTINGGI PER TAHUN



JUMLAH FILM RILIS UNTUK TIAP KATEGORI



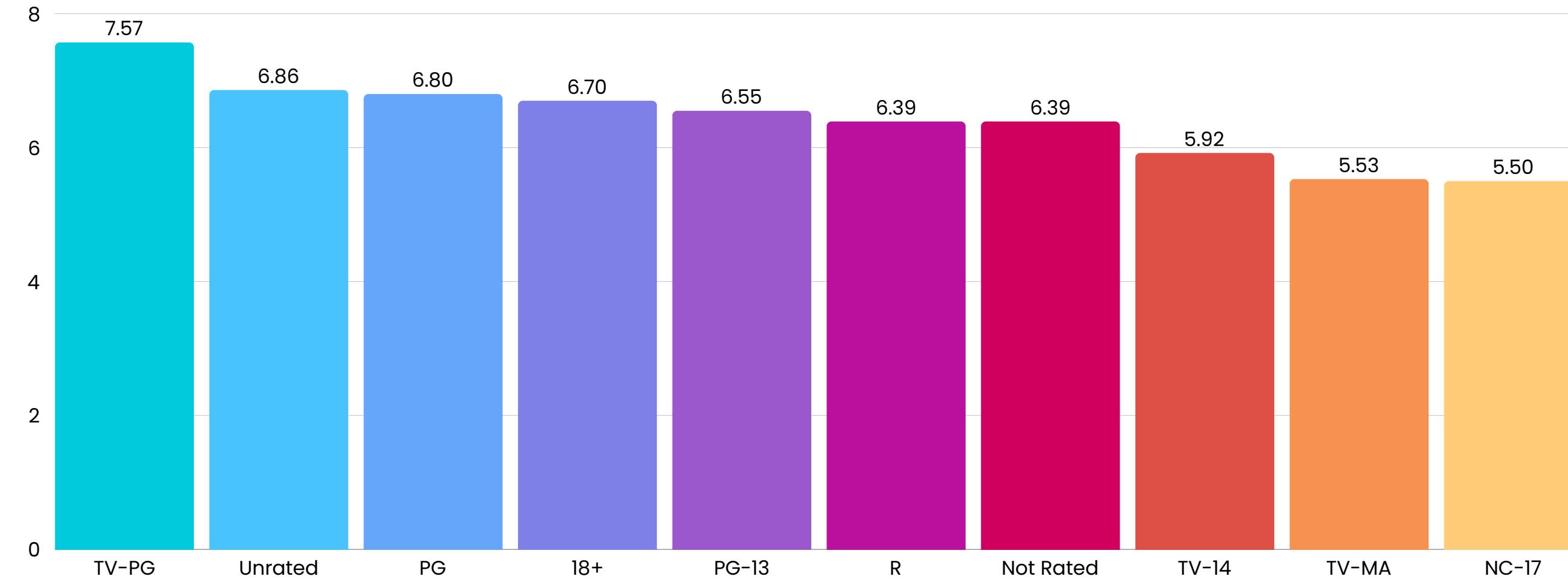
Jumlah film yang dirilis **terbanyak** ada pada kategori **R (Restricted)** atau **"hanya untuk usia 17 tahun ke atas"**.

Dibanding film dengan kategori Parental Guidance (PG, TV-PG, PG-13, TV-14) dengan target audiens anak dibawah umur hingga remaja, **R adalah kategori paling umum pada data**, menandakan film-film dewasa **mendominasi** koleksi IMDb

Dari total film 645, terdapat **92 film yang tidak memiliki kategori** pada situs IMDb



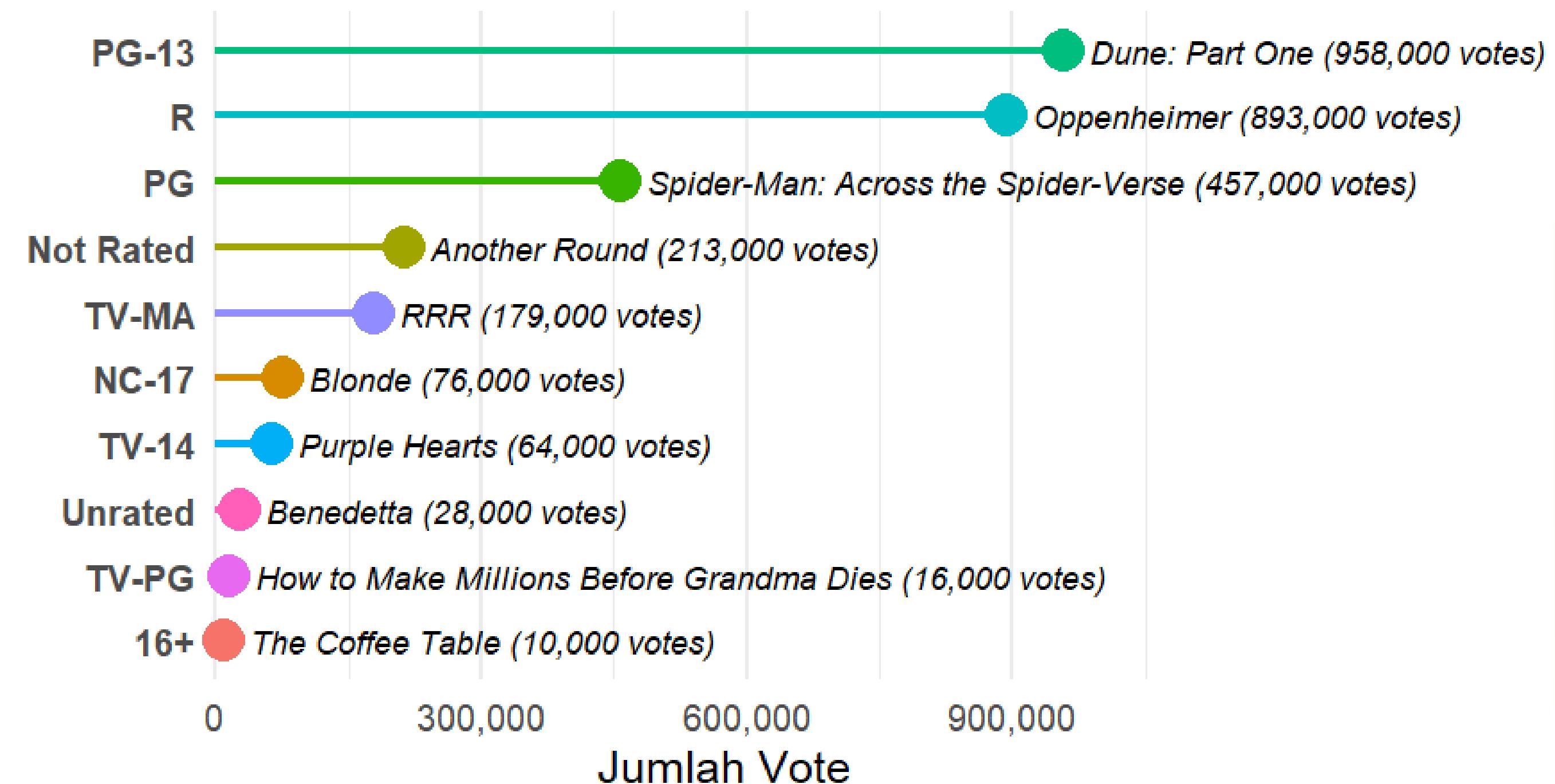
RATA-RATA RATING FILM PER KATEGORI

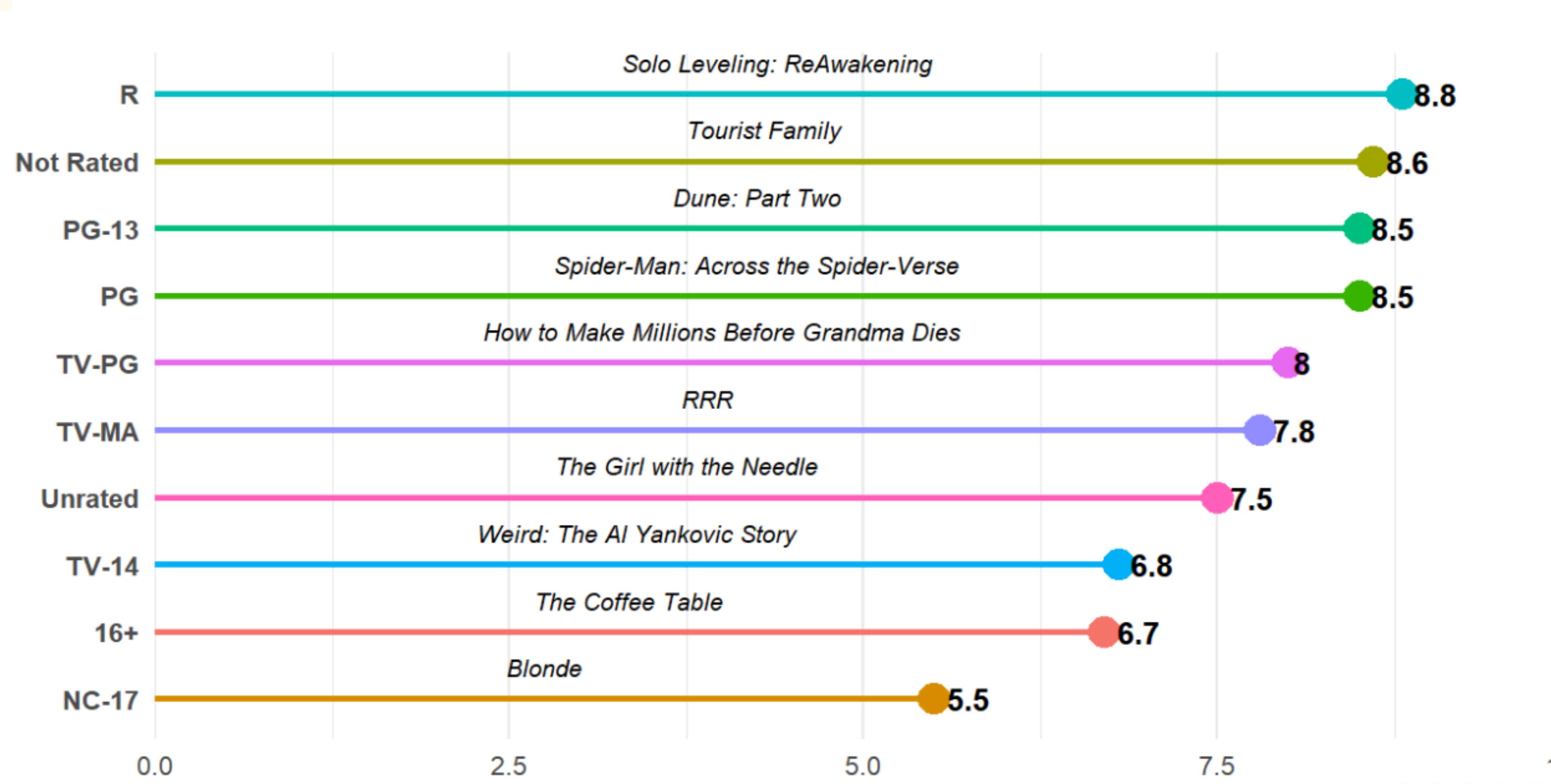


- Kategori **TV-PG** memiliki **rating tertinggi (7.57)**
- Menunjukkan bahwa **konten yang aman** untuk **anak-anak di bawah pengawasan orang tua** cenderung mendapatkan **apresiasi yang tinggi** dari penonton.



Film dengan Jumlah Vote Tertinggi per Kategori





Thank You!

