



Full TMDb Movies Dataset 2024 (1M Movies)

Mihailo Vukorep IN 40/2021

Marko Kolarski IN 60/2021

Rekonstrukcija šeme i dodati indeksi

Iz kolekcije movies nastala je nova kolekcija movies_optimized sa sledećim izmenama:

- Polje release_date je pojednostavljeno i “flattenovano” u release_info sa dodatim poljem decade.
- Dodato polje financial.budget_category za klasifikaciju budžeta i polje financial.quality_tier za klasifikaciju ocena.
- Izračunata polja: financial.profit, financial.roi, financial.is_profitable.
- Žanrovi su sortirani i čuvaju se u content_info.sorted_genres.
- Producenti i zemlje su dopunjeni brojačima: production.company_count, production.country_count.

Rekonstrukcija šeme i dodati indeksi

Dodati indeksi:

- { financial.budget_category: 1, financial.revenue: 1, production.companies: 1 }

(Query 1: prosečan prihod po produkcijskim kućama sa budžetom > 50M)

- { release_info.decade: 1, ratings.vote_average: 1, content_info.genres: 1 }

(Query 2: prosečna ocena po žanrovima kroz decenije)

- { financial.budget_category: 1, release_info.month: 1 }

(Query 3: blockbuster filmovi po mesecima)

- { financial.revenue: 1, financial.budget: 1, content_info.sorted_genres: 1, financial.profit: 1, financial.roi: 1 }

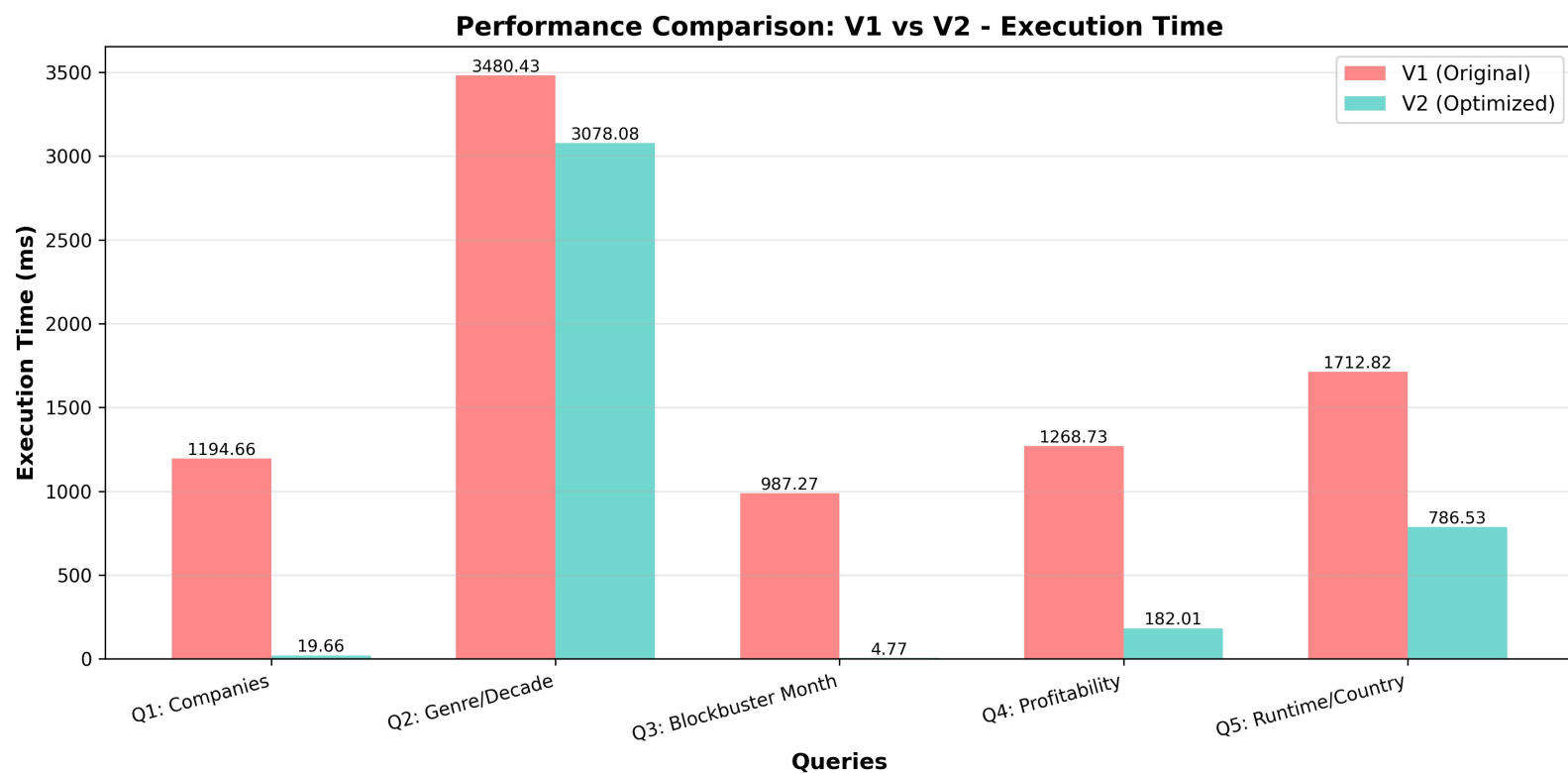
(Query 4: najprofitabilnije kombinacije žanrova)

- { production.countries: 1, content_info.runtime: 1, ratings.quality_tier: 1 }

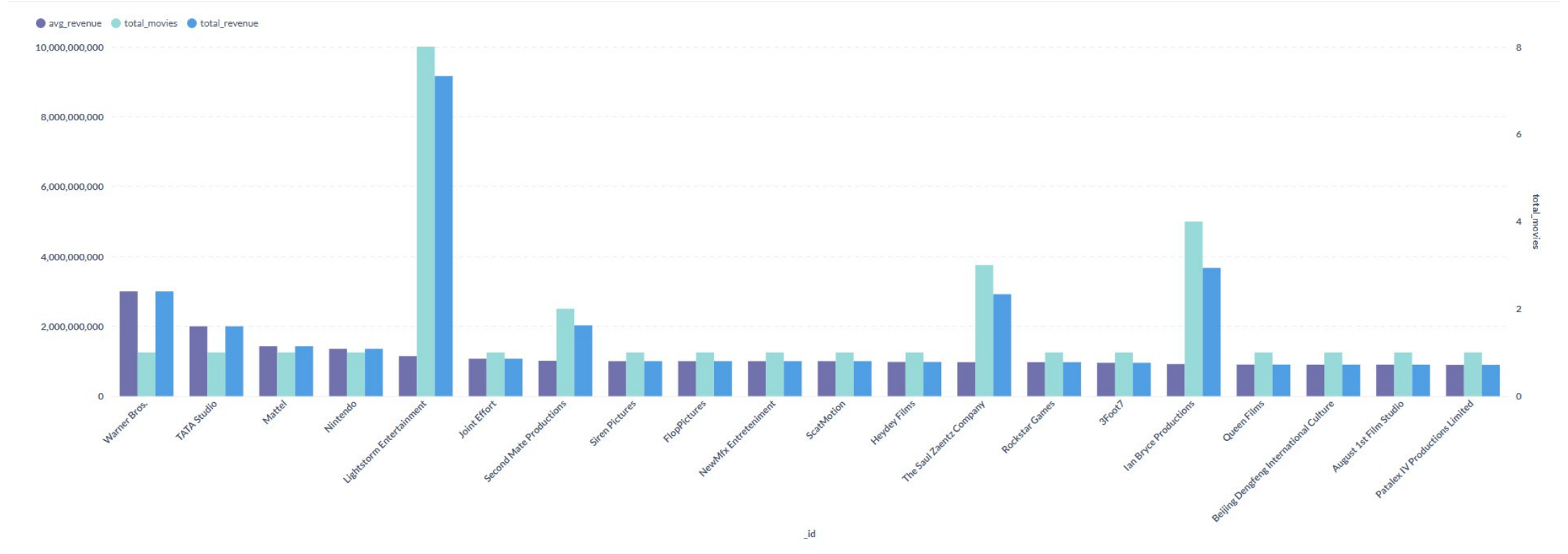
(Query 5: prosečno trajanje po zemlji sa ocenom > 7.0, min 100 filmova)

Uporedna analiza performansi pre i posle optimizacije

Uporedna analiza vremena izvršenja svih 5 upita

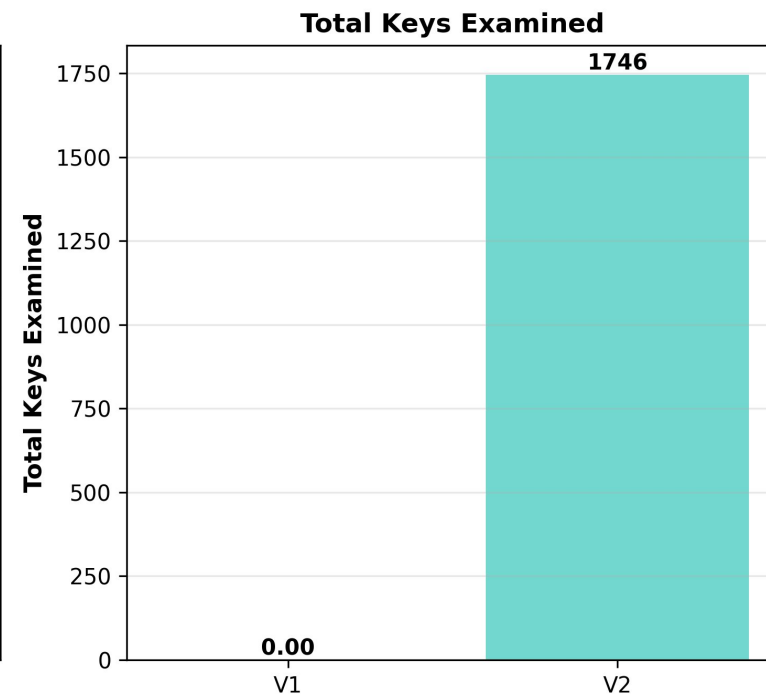
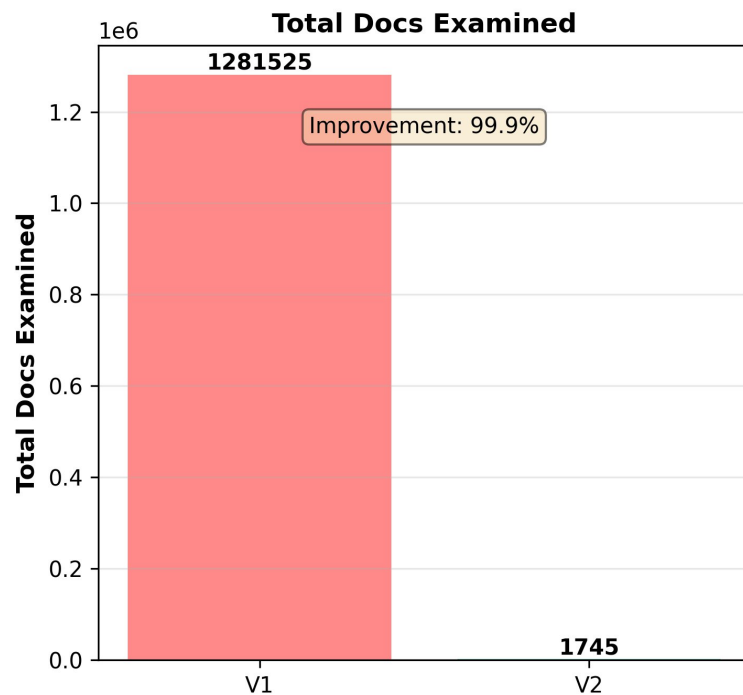
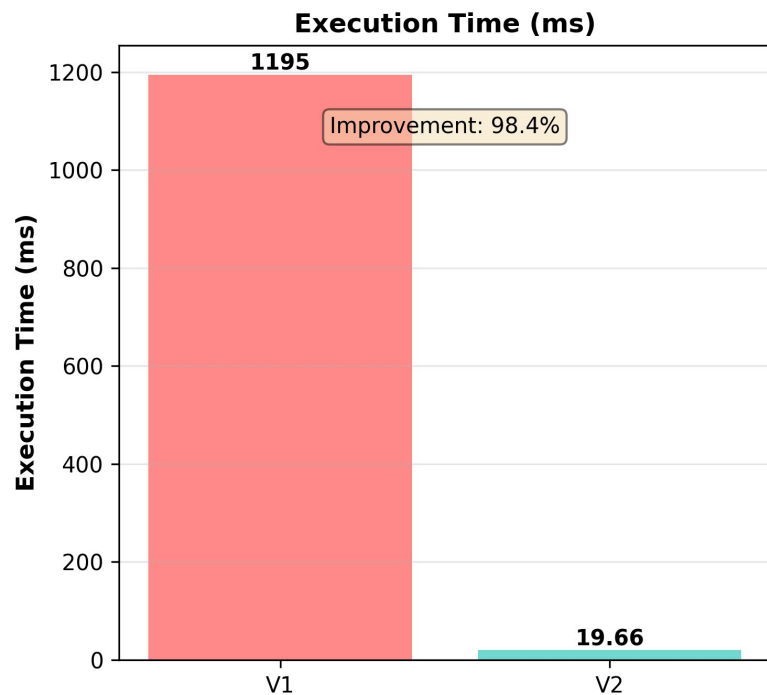


1. Koliki je prosečan prihod po filmu produkcijskih kuća čiji su filmovi imali budžet veći od 50 miliona dolara?

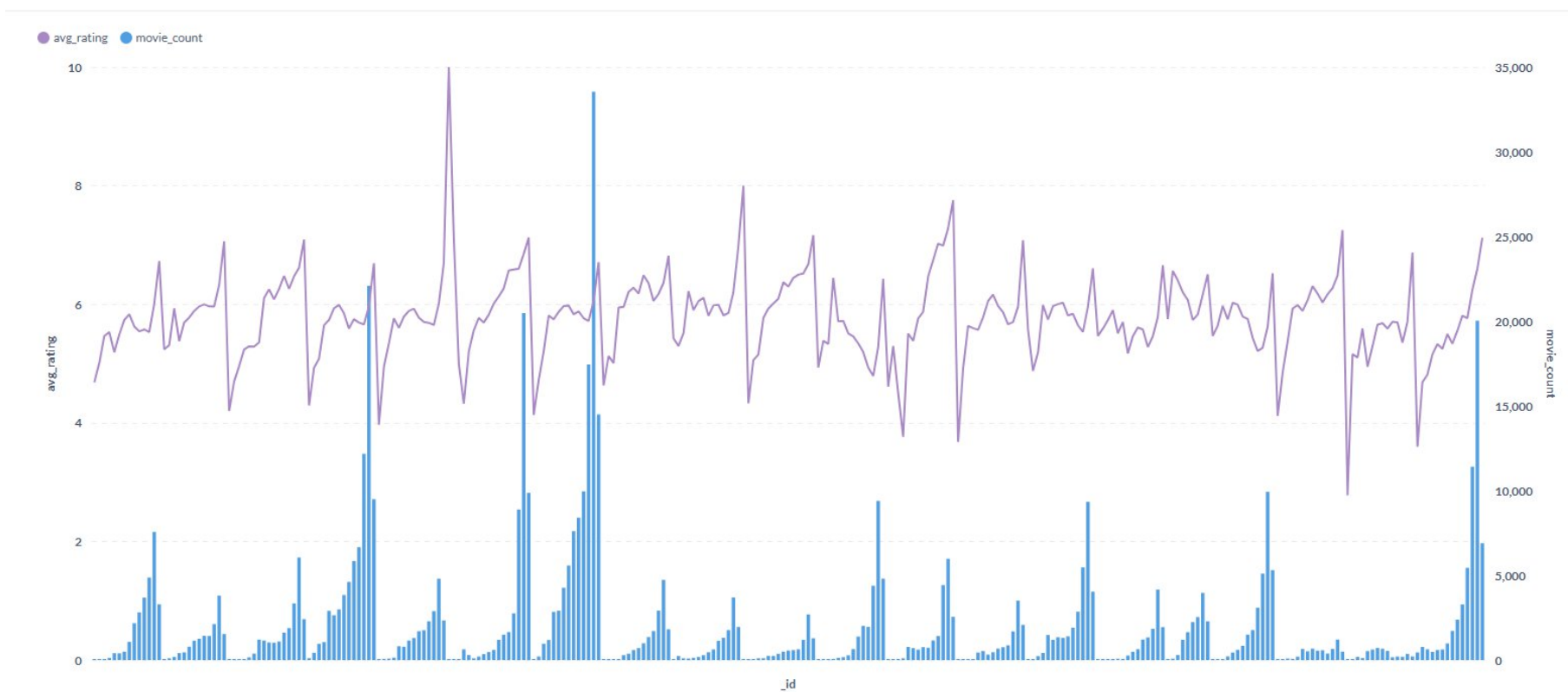


1. Koliki je prosečan prihod po filmu produkcijskih kuća čiji su filmovi imali budžet veći od 50 miliona dolara?

Q1: Companies - Performance Metrics (V1 vs V2)

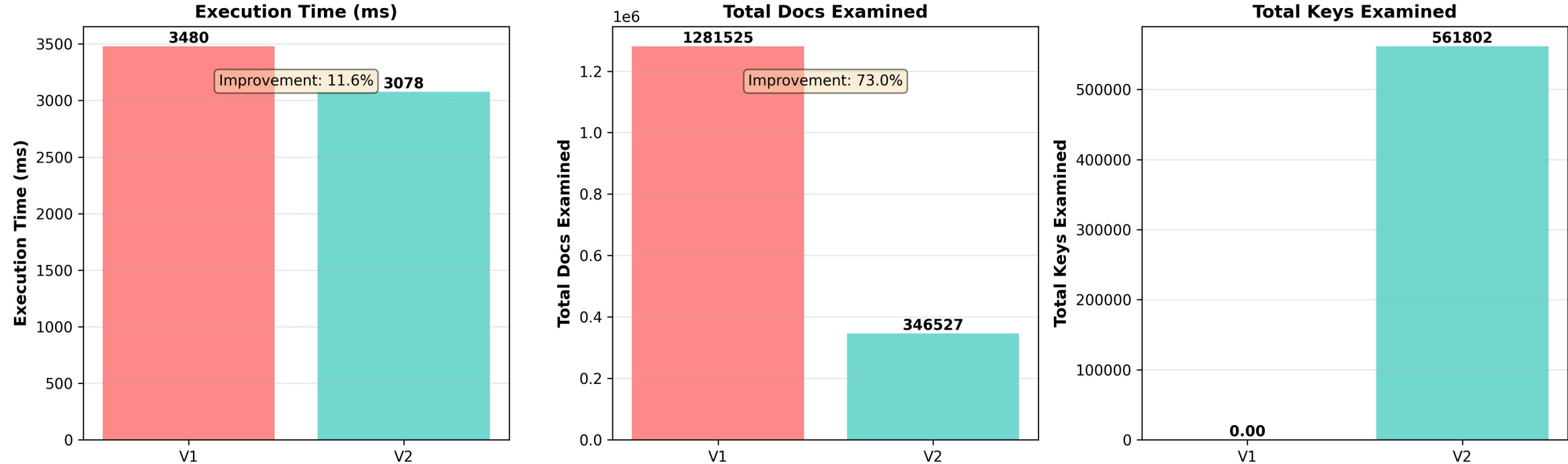


2. Kako se prosečna ocena filmova u različitim žanrovima menjala tokom decenija?

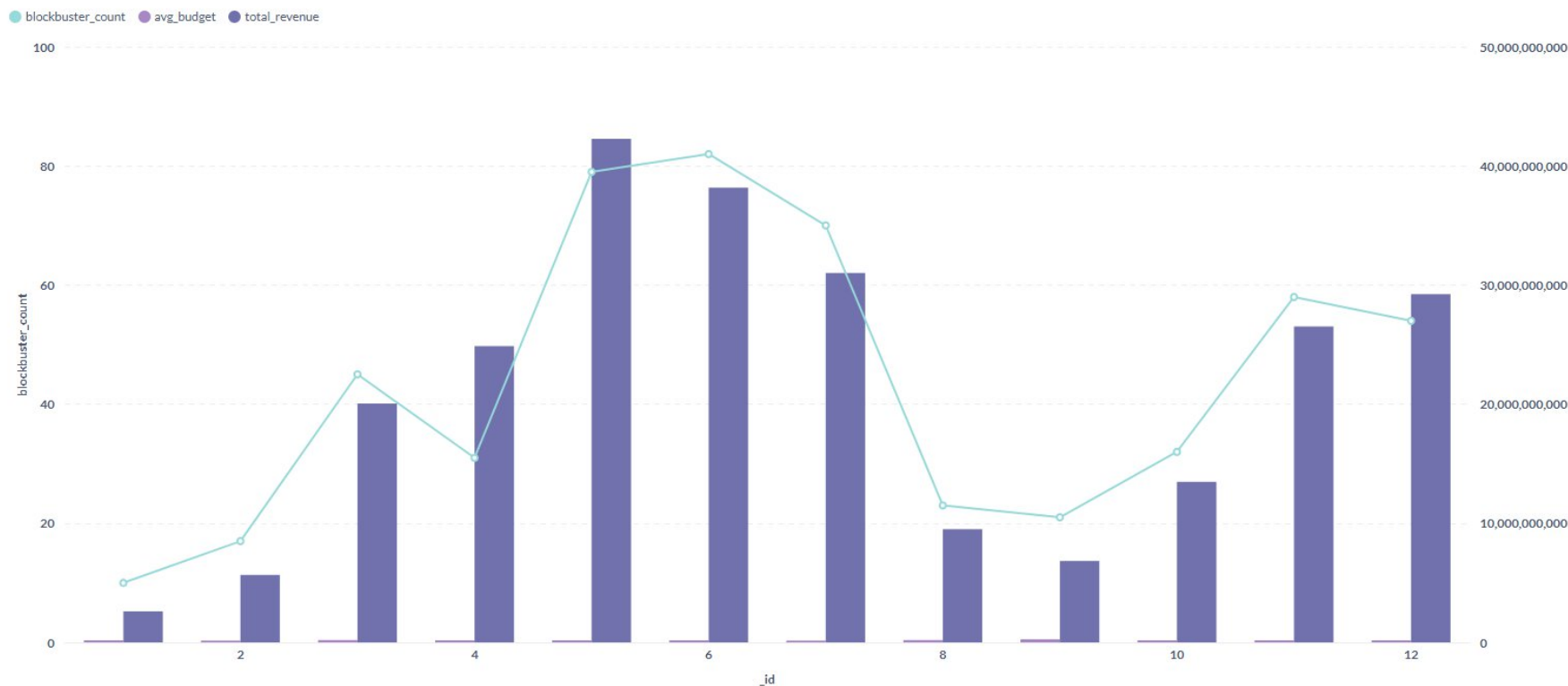


2. Kako se prosečna ocena filmova u različitim žanrovima menjala tokom decenija?

Q2: Genre/Decade - Performance Metrics (V1 vs V2)

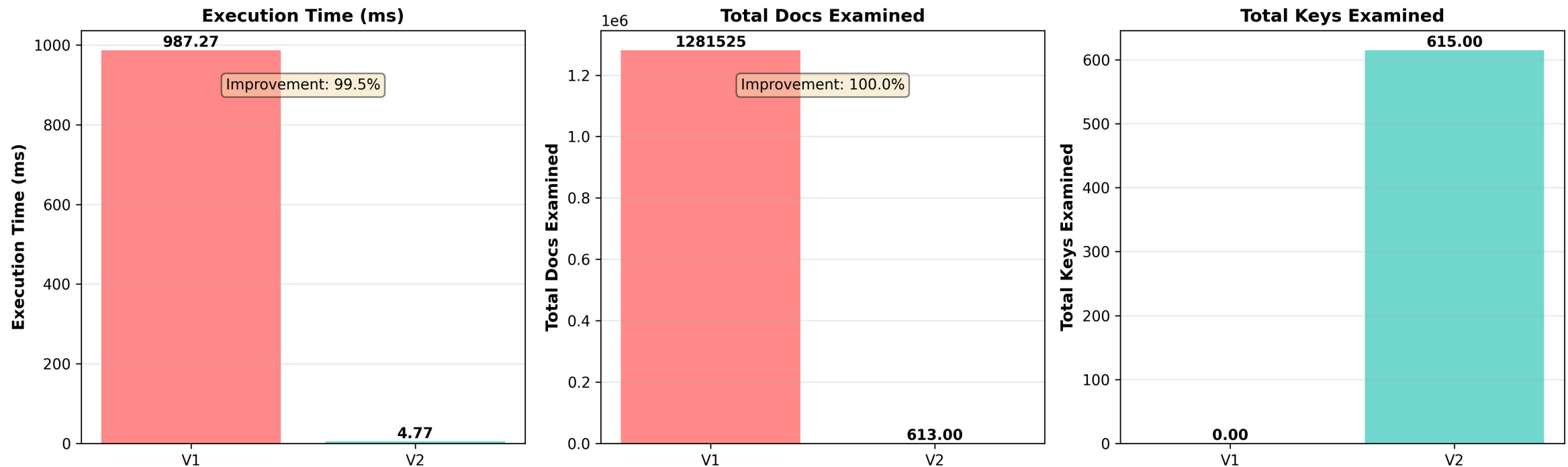


3. Koji meseci u godini imaju najveću stopu objavljivanja blockbuster (hit) filmova (budžet > 100M)?

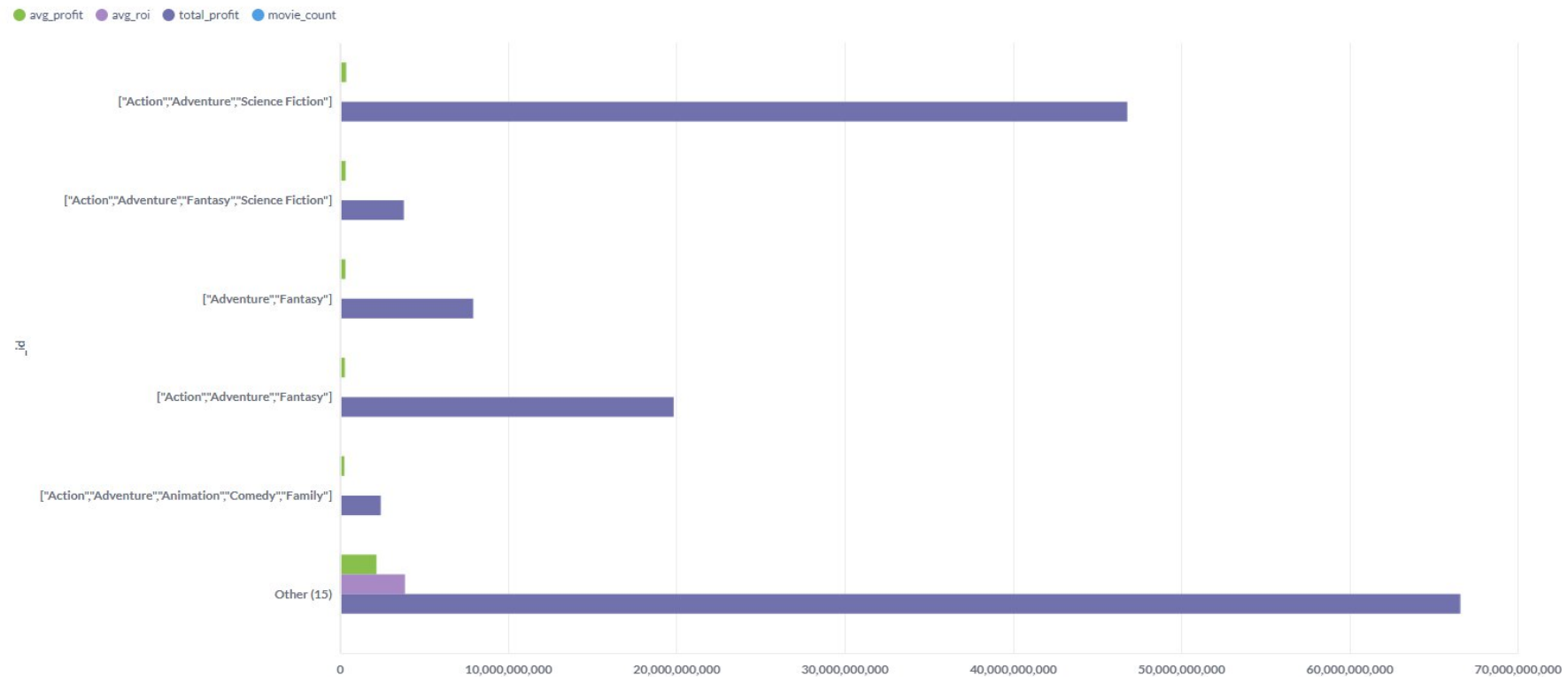


3. Koji meseci u godini imaju najveću stopu objavljivanja blockbuster (hit) filmova (budžet > 100M)?

Q3: Blockbuster Month - Performance Metrics (V1 vs V2)

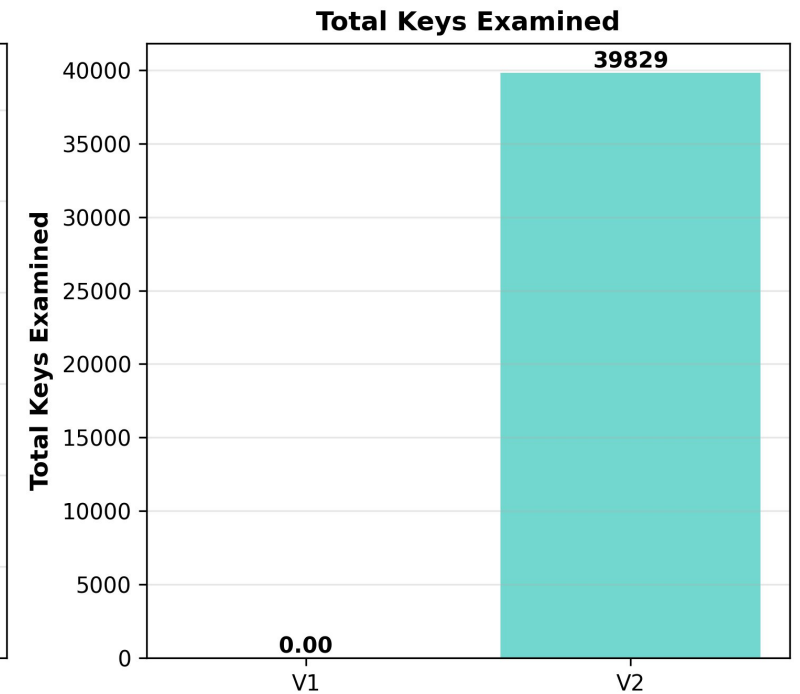
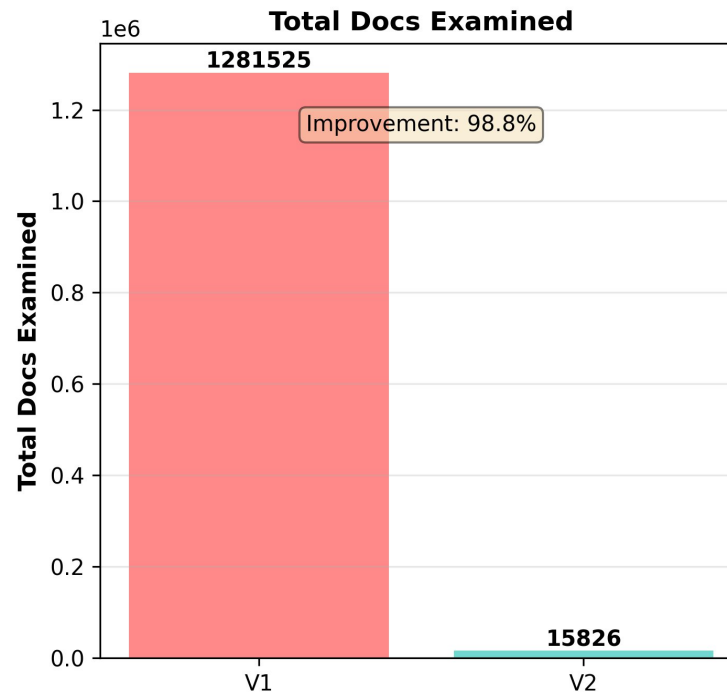
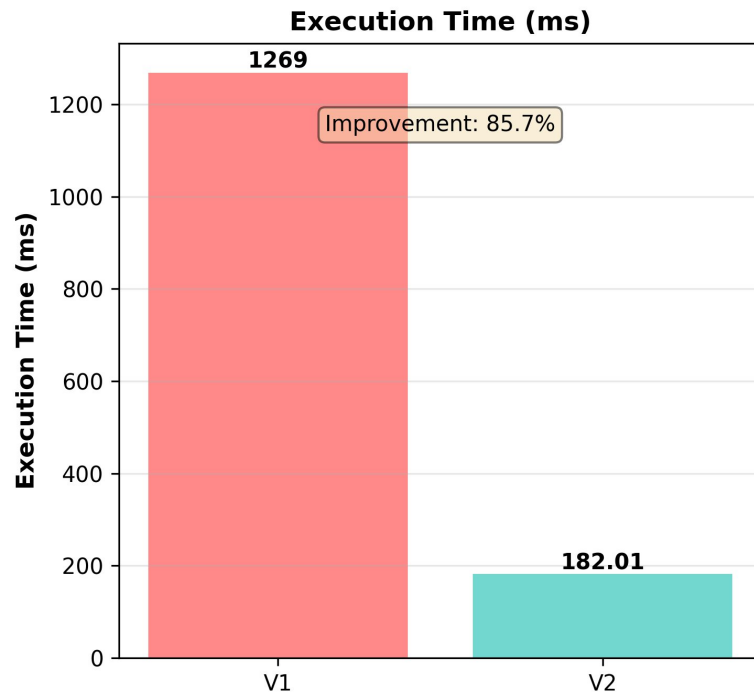


4. Koje su najprofitabilnije kombinacije žanrova?

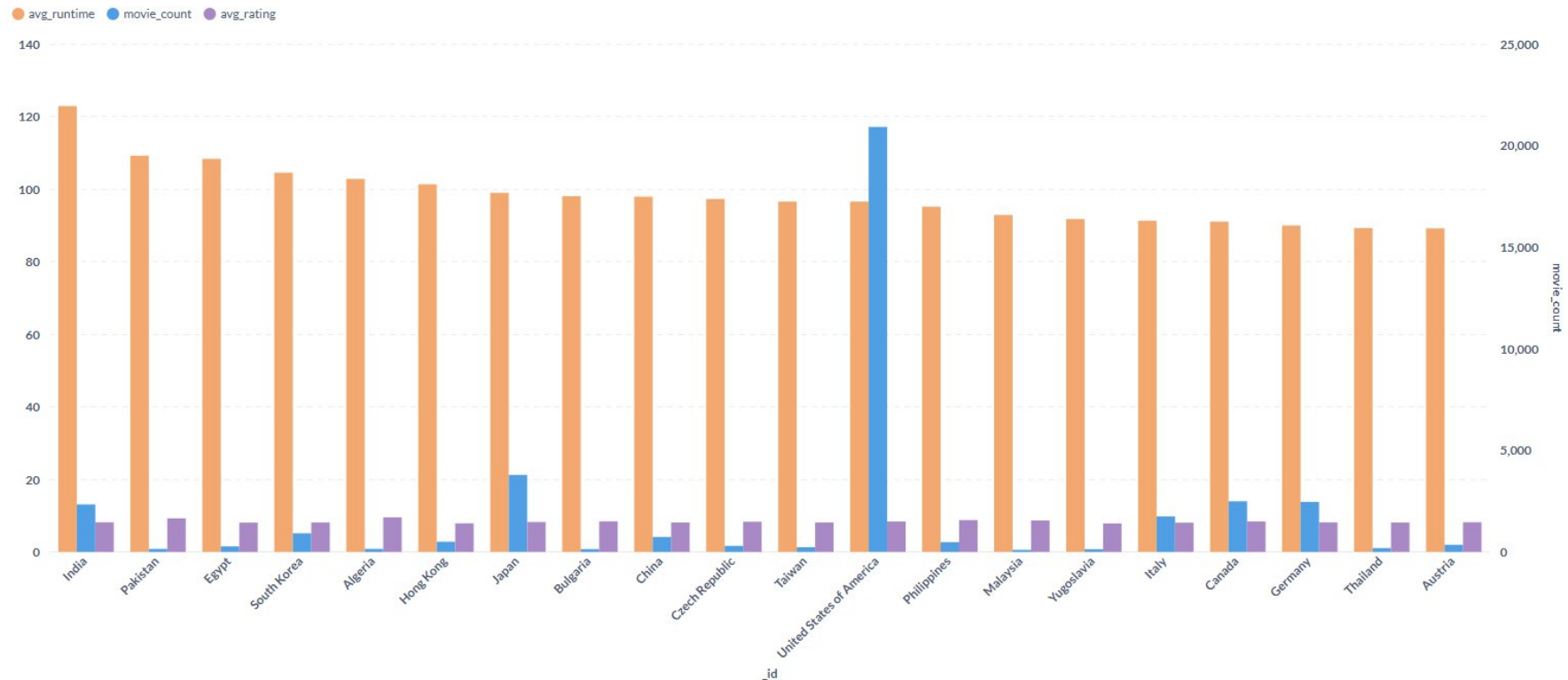


4. Koje su najprofitabilnije kombinacije žanrova?

Q4: Profitability - Performance Metrics (V1 vs V2)



5. Prosečno trajanje filma po zemlji produkcije sa ocenom iznad 7.0, uzimajući u obzir samo zemlje koje su proizvele više od 100 filmova?



5. Prosečno trajanje filma po zemlji produkcije sa ocenom iznad 7.0, uzimajući u obzir samo zemlje koje su proizvele više od 100 filmova?

Q5: Runtime/Country - Performance Metrics (V1 vs V2)

