

Data Support Specialist

Netflix Exercise

Please write up your responses according to this template.

To download the template, click File → Download

Part 1

1. In the 'Films (Non-English)' category, which film has spent the most weeks in the top 10? To estimate the number of users who watched this film, what assumptions would you make and what risks are involved?

Please limit your response to 150 words or less.

To identify the film with the most weeks in the top 10 under 'Films (Non-English),' I filtered the dataset by category and used the cumulative_weeks_in_top_10 column. The top title is *Through My Window* with 13 weeks.

To estimate the number of users, I divided weekly_hours_viewed by the film's runtime in hours, assuming each user watches the entire film once.

Risks include:

- Shared accounts or devices leading to undercounts per individual.
- Repeated or partial views affecting accuracy.
- Ignoring the week of May 22 (with missing viewership data) could skew total hours viewed and user estimates for that period.

Est Users:
Films_(Non-English)
112878333

2. What are the risks of ignoring the data from the week of May 22nd when calculating the metrics from the previous questions (first question above and prior pre-screen questions)?

Please limit your response to 150 words or less.

Excluding the week of *May 22, 2022*, removes one full week of viewership data due to a collection issue in the `weekly_hours_viewed` field.

The main risk is underestimating the cumulative viewing hours and user estimates for any film that appeared in the top 10 during that week, especially if it was a peak week for that title. However, this exclusion likely has minimal impact on rankings based on cumulative weeks in top 10, since that column is still accurate (the issue only affected hours viewed, not the presence in the list).

I might overestimate or underestimate by leaving out this week:

- *Churn rates*: The statistics would not demonstrate whether an interruption caused people to leave.
- *User engagement metrics*: If viewing hours drastically dropped during the outage, excluding this week could result in inaccurate weekly hours viewed.

Part 2

2. Please review the attached report and be prepared to answer live questions regarding its insights into NFLX's performance as a company.

← Estimated Users

0.00

33,50,10,000.00

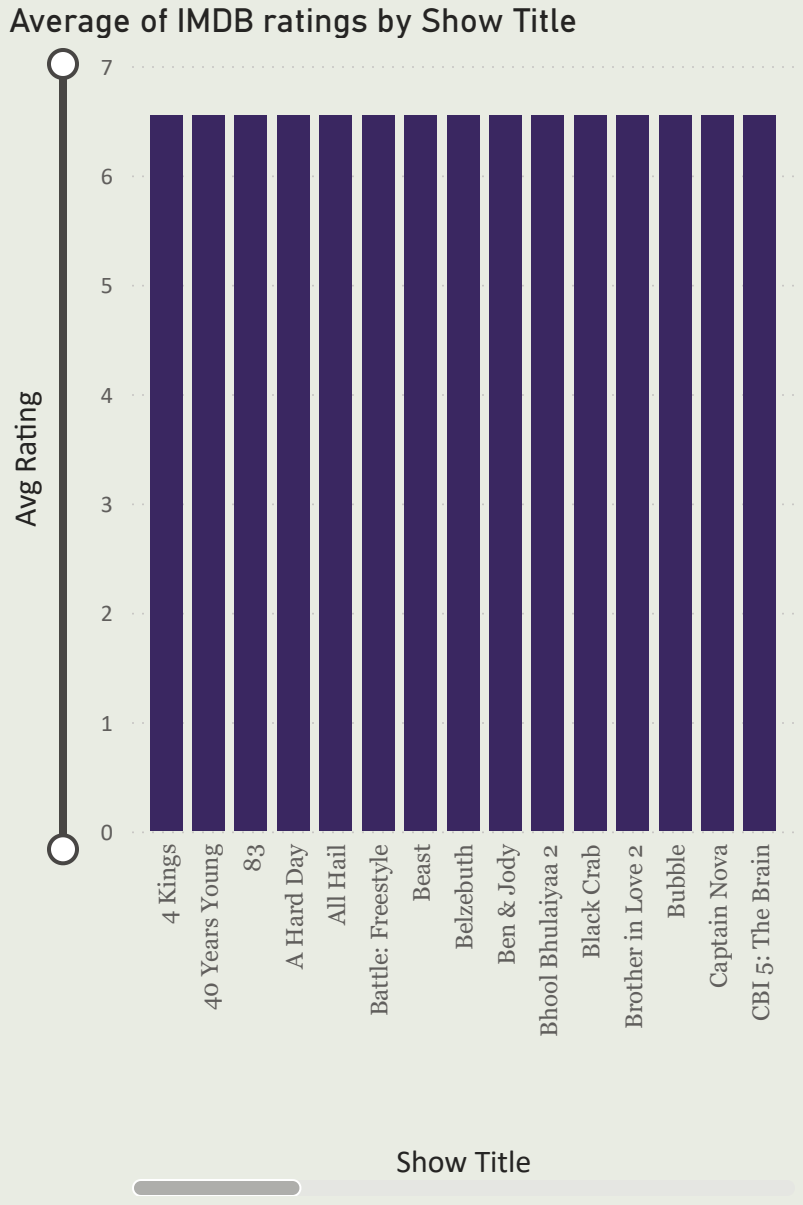
Category

Films (Non-English)

IMDB Rating

0.00

8.90



112.88M

Sum of Estimated Users

6.55

Avg Rating

13

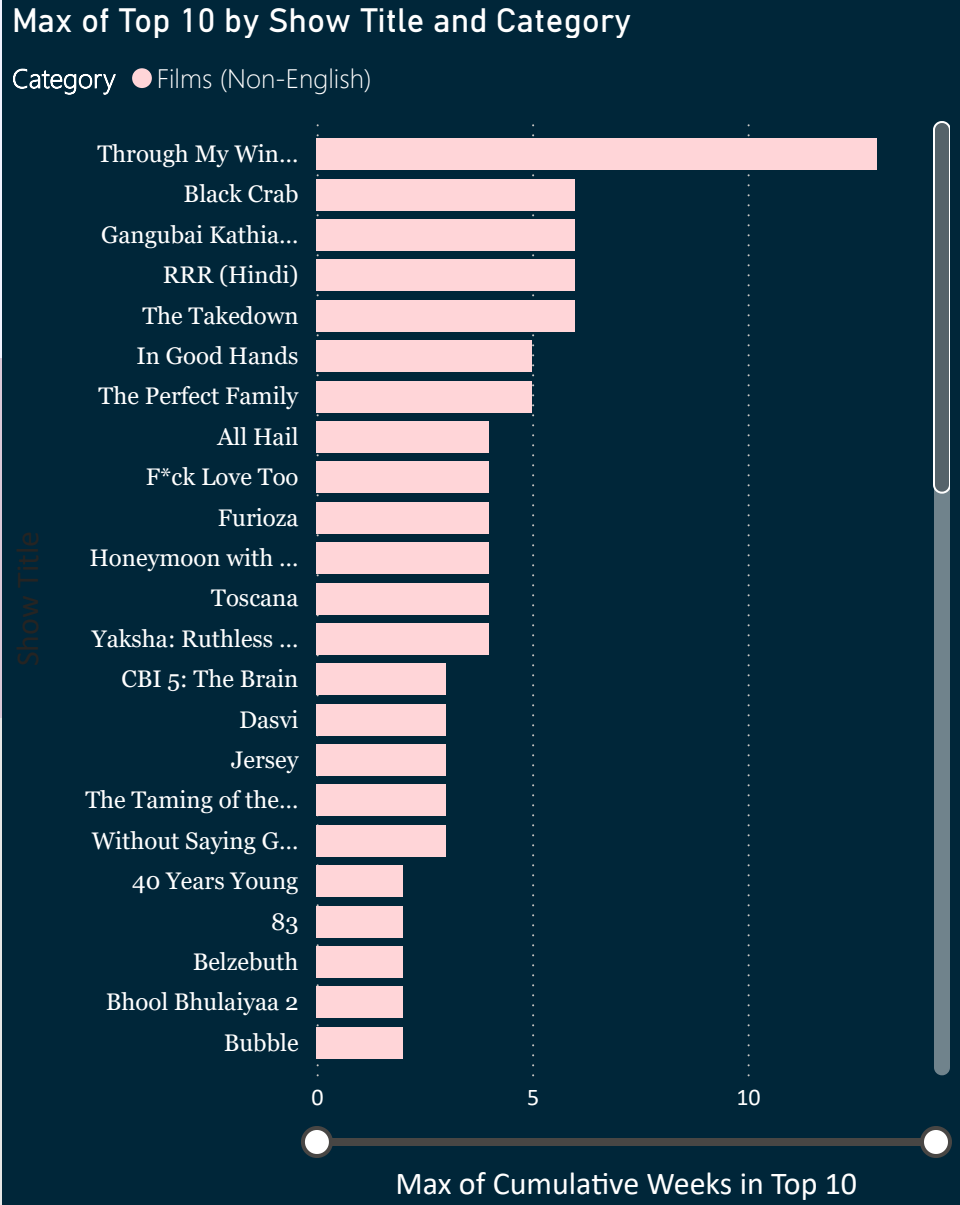
Max of Cumulative Weeks in Top 10

Through My Window

Top Show Title

Sum of Weekly Hours Viewed and Top Show Title by Category

758M (100%)



← Estimated Users

0.00

33,50,10,000.00

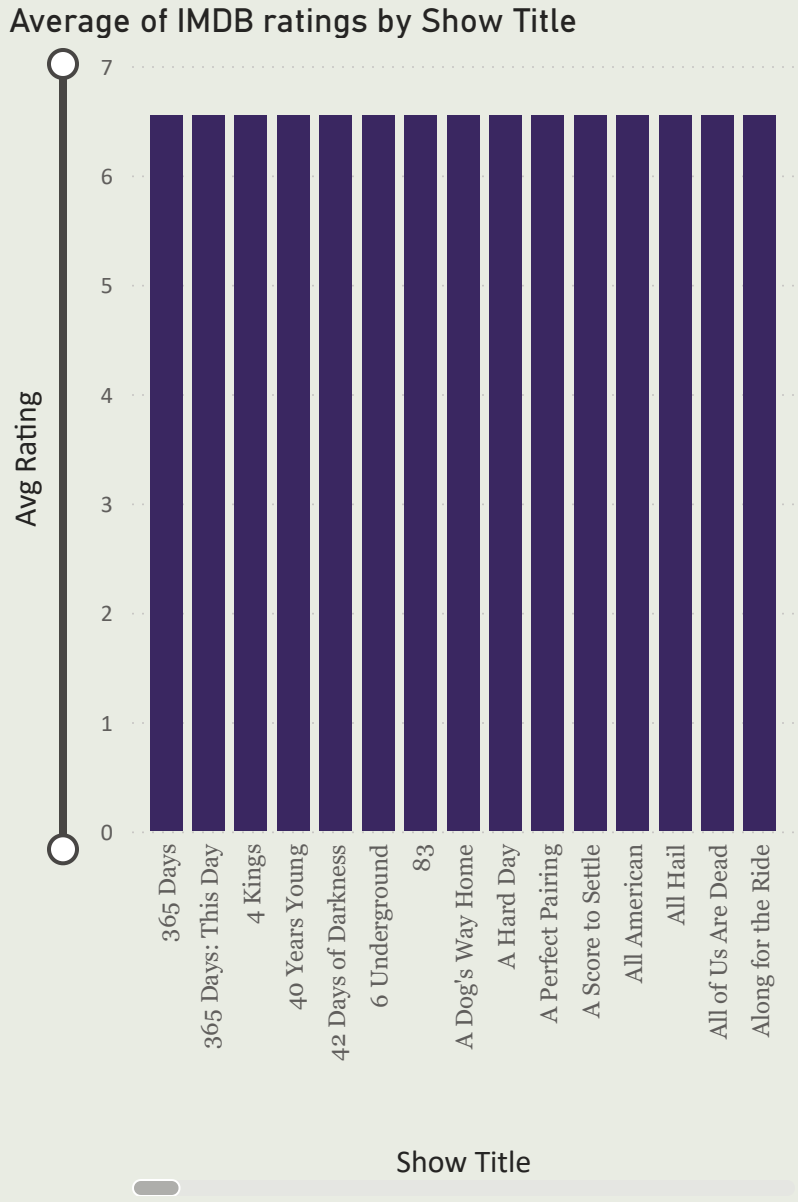
Category

All

IMDB Rating

0.00

8.90



5.52bn

Sum of Estimated Users

6.55

Avg Rating

28

Max of Cumulative Weeks in Top 10

Yo soy Betty, la fea

Top Show Title

Sum of Weekly Hours Viewed and Top Show Title by Category

Category	Hours Viewed	Percentage
TV (English)	4bn	50.21%
TV (Non-English)	2bn	22.7%
Films (English)	2bn	18.55%
Films (Non-English)	1bn	8.53%

