

# Netflix Data Visualization

## Describing My Dataset:

### Description

The two data sets that I will be working with will conjoin into one and will provide me with information regarding when Netflix titles went live on the platform, as well as when users watched the titles. The individual user streaming data encompasses data collected from 1/1/17 to 6/30/19. The Netflix live titles data encompasses data collected between 1/1/08 to 9/25/21.

### Usage

To measure traffic on the Netflix streaming service when certain “high-profile” titles are added.

### Format Individual User Data

A data frame with 671736 observations on 8 variables.

1. **View** numeric identification numbers (58928 - 730508)
2. **datetime** numeric date/time of view (date - time)
3. **duration** numeric Duration of watch (seconds)
4. **title** character Title of stream
5. **genres** character Genre of stream
6. **release\_date** numeric Day of release to theaters (date)
7. **movie\_id** numeric Individual movie ID number
8. **user\_id** numeric User ID number

### Details

Analysis of user stream times from users who volunteered their information. Data gathered from 1/1/17 to 6/30/19.

- View: The incremental number which is associated with the stream of a title. Netflix data for the view number on the platform.
- Datetime: The date and time for which the viewer opened and streamed a title.
- Duration: The length of time, in seconds, that a user viewed a title.
- Movie\_id: The identification number granted to a title or easy correlation accross data.
- user\_id: The identification number of the user who volunteered their information.

### Format for Netflix Titles

A data frame with 8807 observations on 12 variables.

1. **Show\_id** numeric Identification numbers of each title
2. **type** character Category of title, ex. Movie, Tv Show, etc...
3. **title** character Title of stream

4. **director** character Director of the title
5. **cast** character Lead actors and actresses
6. **country** character Country in which the title was available on Netflix
7. **date\_added** numeric Date the title went live on Netflix (date)
8. **release\_year** numeric Release year of title
9. **rating** character Viewers rating ex. R, TV-MA, etc...
10. **duration** numeric Length of title in minutes.
11. **listed\_in** numeric Subcategory of Netflix for which the title is found
12. **description** numeric Brief description of title as seen by viewers on netflix.

## Details

Analysis of titles added to the Netflix streaming repertoire between 1/1/08 and 9/25/21.

- Show\_id: The identificaiton number of the title added to netflix, can be correlated over different add and drops.
- Title: The title of the film which can be used to find the peak view times, as well as view traffic trends on the service.
- Date\_added: The date when the title was added to netflix, and will also be important in correlating traffic trends.

## Source

The data were obtained from the “Netflix Movies and TV Shows” data set created by SSHIVAM BANSAL on kaggle.com and the “Netflix audience behaviour - UK movies” by Stephen Follows and Eliel CM on kaggle.com.  
(<https://www.kaggle.com/datasets/shivamb/netflix-shows?rvi=1>)  
(<https://www.kaggle.com/datasets/shivamb/netflix-shows?rvi=1>)  
(<https://www.kaggle.com/datasets/vodclickstream/netflix-audience-behaviour-uk-movies?rvi=1>)  
(<https://www.kaggle.com/datasets/vodclickstream/netflix-audience-behaviour-uk-movies?rvi=1>)

## References

1. <https://www.kaggle.com/datasets/shivamb/netflix-shows?rvi=1>  
(<https://www.kaggle.com/datasets/shivamb/netflix-shows?rvi=1>)
2. <https://www.kaggle.com/datasets/vodclickstream/netflix-audience-behaviour-uk-movies?rvi=1>  
(<https://www.kaggle.com/datasets/vodclickstream/netflix-audience-behaviour-uk-movies?rvi=1>)