

Project proposal

Intro

Netflix is one of the most popular media and video streaming platforms. It offers subscription-based video on demand from a library of films and television series, 40% of which is Netflix original programming produced in-house. They have over 8000 movies or tv shows available on their platform, as of mid-2021, they have over 200M Subscribers globally. This dataset consists of listings of all the movies and tv shows available on Netflix, along with details such as - cast, directors, ratings, release year, duration, etc.

The goal of this study is to analyze the types of listings, and whether Netflix is focused more on streaming TV-shows or movies based on how many listings there are and their ratings by users.

Dataset:

The data set [Netflix Data](#) obtained from Kaggle has about 8807 rows. Columns of the dataset include(*show_id, type, title, director, cast, country, date_added, release_year, rating, duration listed_in description*)

Sample

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mababane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...

Tools:

For the purpose of this study, I obtained a netflix dataset for Kaggle [Netflix Data](#) as .csv file. I'll be using pandas, seaborn, and matplotlib to visualize the data and do some cleanup. Programming using python will be on a Jupyter notebook.

To Do:

Investigate the dataset, and perform exploratory data analysis to come up with a model that can help with interpreting Netflix focused areas.