

Data Analysis Project on

Amazon Prime Movies and TV Shows

About Dataset:

Amazon Prime is another one of the most popular media and video streaming platforms. They have close to 10000 movies or tv shows available on their platform, as of mid-2021, they have over 200M Subscribers globally. This tabular dataset consists of listings of all the movies and tv shows available on Amazon Prime, along with details such as - cast, directors, ratings, release year, duration, etc.*

- Here we have 12 columns and 1000 rows
- We have blank values in few column values which you need to focus and replace by 0 in power query
- We have null values too, replace these null values to 0

Dataset: [Amazon Prime Movies and TV Shows | Kaggle](#)

Dataset Analyzation:

- Show Id: Given unique IDs for each title
- Type: Category types
- Title: Name of the movie, series etc...
- Director: List of director's names
- Cast: Names of cast
- Country: List of country name which belongs to
- Date added: the day show has been added to Amazon Prime
- Release Year: Year of movie, show, series released

- Rating: falls under categories
- Duration: Timing(duration) of particular title
- Listed In: Each show falls under which list(category)
- Description: A brief info about the shows

Interesting task ideas:

1. Understanding what content is available in different countries
2. Identifying similar content by matching text-based features
3. Network analysis of Actors / Directors and find interesting insights
4. Does Amazon Prime have more focus on TV Shows than movies in recent years?
5. Title having maximum watch time
6. Analyze which year been more release year
7. Understanding shows for kids based to increase children attention

Note: Make sure to use relatable charts and use as much as information out of it to show the entire details about data.