

# Problem Statement

---

This dataset consists of TV shows and movies available on Netflix as of 2019. The dataset is collected from Flixable which is a third-party Netflix search engine.

In 2018, they released an interesting report which shows that the number of TV shows on Netflix has nearly tripled since 2010. The streaming service's number of movies has decreased by more than 2,000 titles since 2010, while its number of TV shows has nearly tripled. It will be interesting to explore what all other insights can be obtained from the same dataset.

Integrating this dataset with other external datasets such as IMDB ratings, rotten tomatoes can also provide many interesting findings.

---

## In this project, you are required to do

1. Exploratory Data Analysis
  2. Understanding what type content is available in different countries
  3. Is Netflix has increasingly focusing on TV rather than movies in recent years.
  4. Clustering similar content by matching text-based features
- 

## Attribute Information

---

1. show\_id : Unique ID for every Movie / TV Show
2. type : Identifier - A Movie or TV Show
3. title : Title of the Movie / TV Show
4. director : Director of the Movie
5. cast : Actors involved in the movie / show
6. country : Country where the movie / show was produced
7. date\_added : Date it was added on Netflix
8. release\_year : Actual Releaseyear of the movie / show
9. rating : TV Rating of the movie / show
10. duration : Total Duration - in minutes or number of seasons
11. listed\_in : Genre
12. description: The Summary description