# TASK – 1

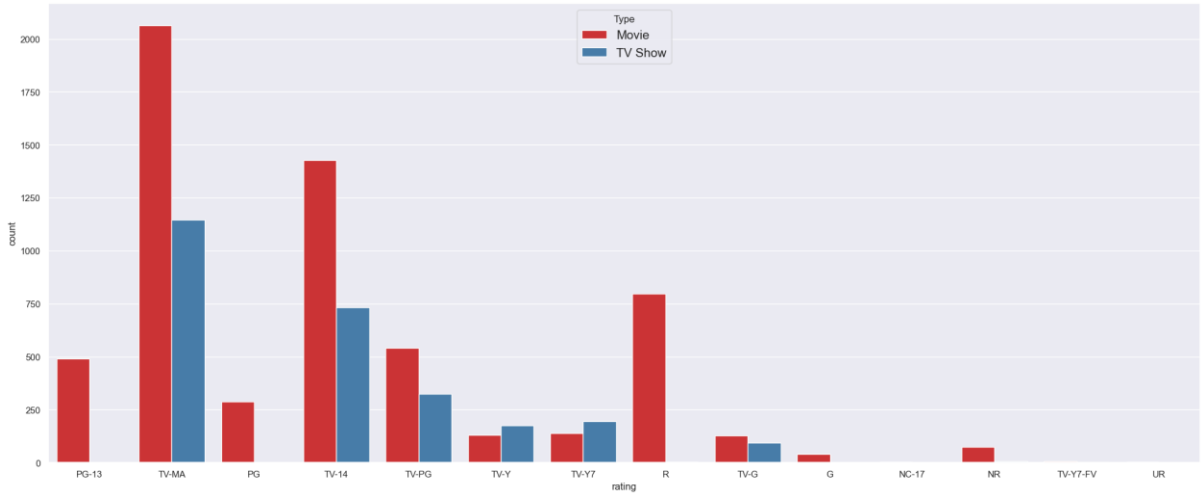
# AMAZON PRIME DATA VISUALIZATION

**Dataset Information :**

```
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 12 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   show_id       8807 non-null   object
 1   type          8807 non-null   object
 2   title         8807 non-null   object
 3   director      6173 non-null   object
 4   cast          7982 non-null   object
 5   country       7976 non-null   object
 6   date_added    8797 non-null   object
 7   release_year  8807 non-null   int64
 8   rating        8803 non-null   object
 9   duration      8804 non-null   object
 10  listed_in     8807 non-null   object
 11  description   8807 non-null   object
dtypes: int64(1), object(11)
```

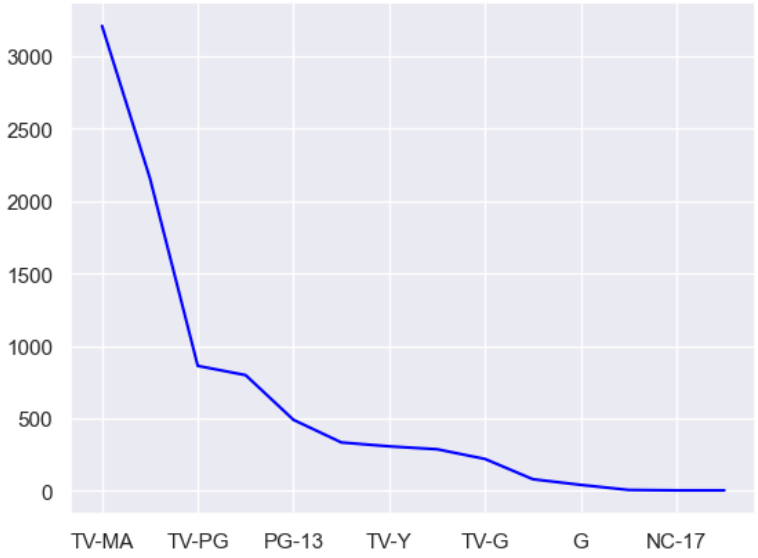
**Ratings Count:**

We are able to see that **Movies** are a major source of content for users to explore in Amazon Prime videos. There are many movies around **TV-MA category.**



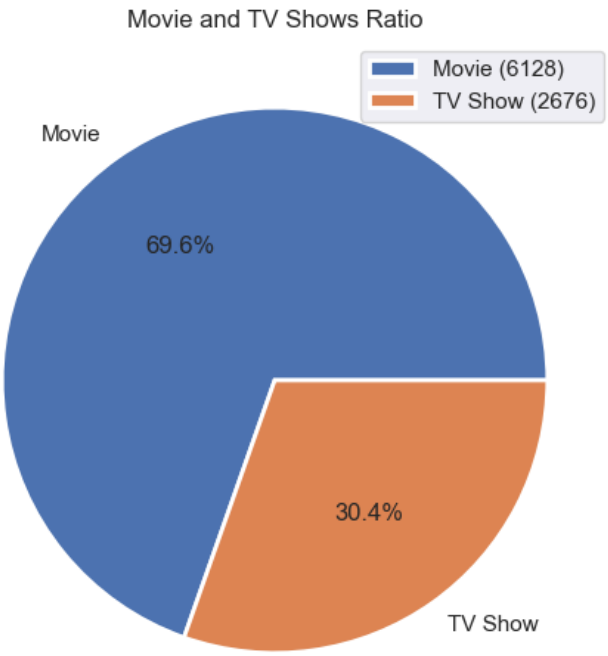
**Popular Ratings Category:**

According to the dataset, **TV-MA** is the most popularly rated category liked by the users.




**Movie and TV Shows Ratio:**

Movies contribute to nearly **70%** of the content, whereas TV Show contribute to around **30%**.

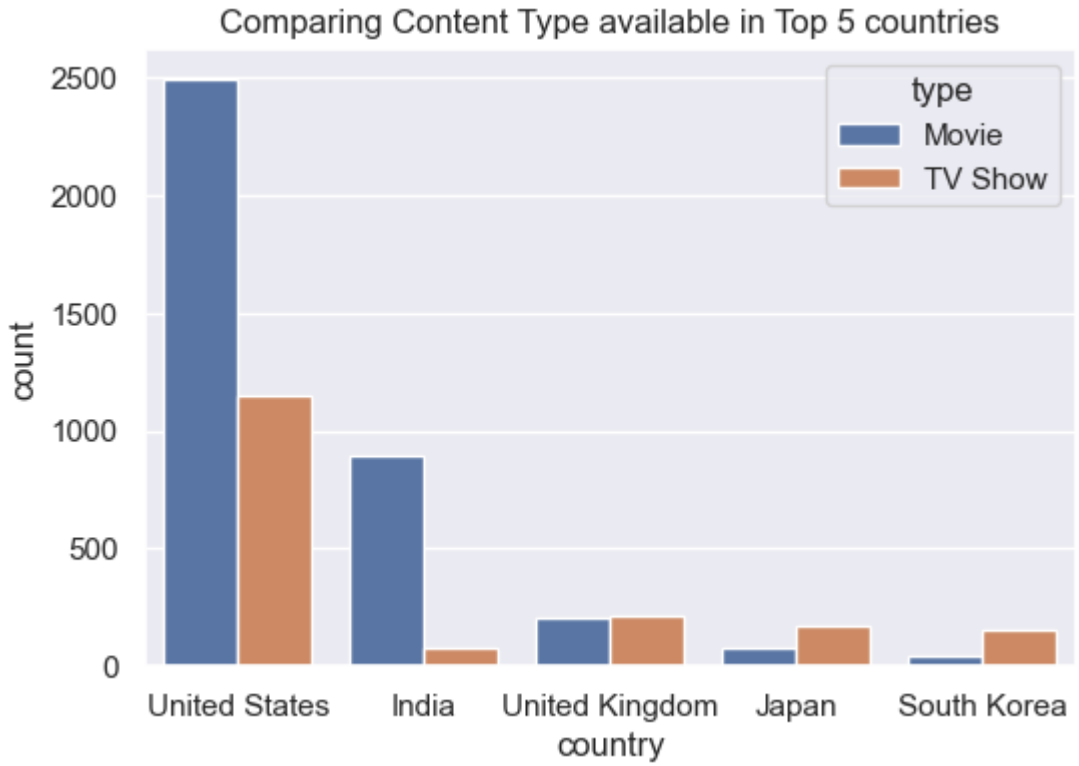


**Top 5 countries :**

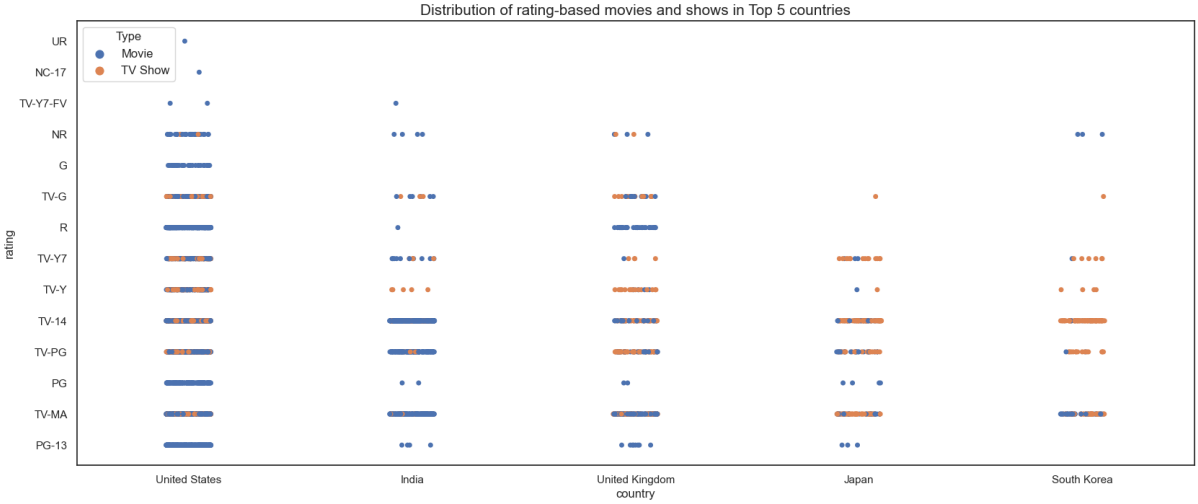
1. United States
2. India
3. United Kingdom
4. Japan
5. South Korea

**Comparing Content Type available in Top 5 countries :**

We are able to visualise and observe that movies are highly available in 3 of the top 5 countries.



## Distribution of rating-based movies and shows :




## Show categories and its count :

**International Movies** category occupies the top place in **Show Categories**.