**Task 1 – Prime**

This dataset consists mainly of categorical data of the movies and TV shows available in Prime Video over the years, along with the director, cast, country and more fields. The dataset contains 12 columns and 8808 rows of data, out of which none were duplicate rows. However, there were null values and after removing them, the rows reduced to 5332. Python libraries such as pandas, numpy, matplotlib and seaborn can be utilized to visualize the data to understand it better.

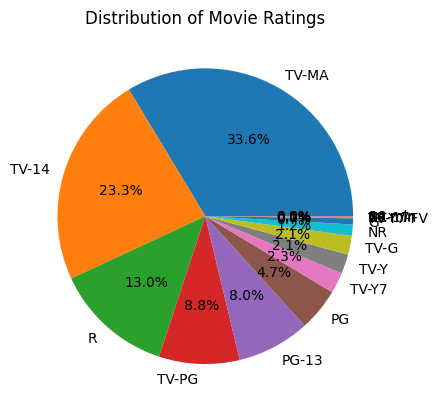
* **Year with the greatest number of movies**

Using the groupby() function we can group the data by release year and by selecting movie from the type field. And the count() function can be used to count the number of movies in each year and calculate the one with the maximum count to be the year with the most number of movies, which is 2017.

* **Year with the greatest number of shows**

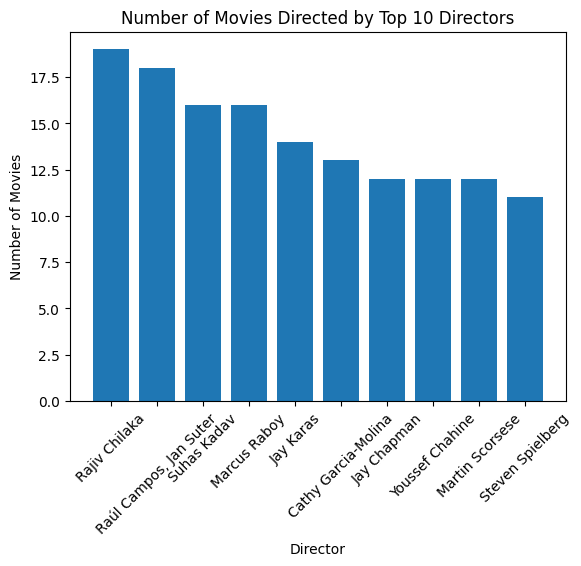
The same functions described above can be used to get the year that had the most number of TV Shows, and that is 2020.

* **Distribution of movie ratings in Prime Video**



It can be seen that the streaming platform, Prime Video has most of its content falling under the rating category of TV-MA, following which is TV-14, R and TV-PG.

* **Number of movies directed by top 10 directors**



From the bar chart above, we can get the total number of movies directed by the top 10 directors. The director with the highest number of films is Rajiv Chilaka.