**NETFLIX DATA ANALYSIS**

**Problem Statement:**

To conduct a detailed analysis of Netflix movies to answer important business questions and help future decision making.

**Business Requirements:**

1. What is the most frequent genre of movies released on Netflix?
2. Which genres have highest votes?
3. Which movie is the most popular and what’s it’s genre?
4. Which movie got the lowest popularity and what’s it’s genre?
5. Which year got the highest number of movies filmed?

**Steps Followed:**

1. Jupyter Notebook was used to execute this project. Firstly, all necessary python libraries were imported necessary for the project.
2. Dataset was imported in CSV format.
3. Dataset was explored using various python functions and basic details about the dataset was understood like size, shape, some statistics etc.
4. Null and duplicate values were checked for in the dataset.
5. ‘Release\_Date’ column was changed to date-time format using pandas functionality.
6. Then only years part of the column was extracted and saved as necessary for the project.
7. Unnecessary columns were dropped.
8. ‘Vote\_Average’ column was categorized in 4 categories- popular, average, below average and not popular.
9. The value counts for all categories were checked.
10. Null values were removed if any.
11. ‘Genre’ column was split and formatted in such a way that each row contained only 1 genre.
12. Datatype of the ‘Genre’ column was changed to categorical data.
13. A ‘Genre vs Releases’ graph was plotted using seaborn library to visualize the most frequent genre of movies released on Netflix.
14. A new variable genre\_vote\_counts was created in which all the genres were stored alongside their total votes in descending order.
15. A barplot was plotted to visualize the genres with the highest no of votes.
16. The most and least popular movies and their genres were found out using python’s dataframes functionality.
17. A histogram was plotted to visualize the year with the most no of movie releases.

**Key Findings:**

1. Drama genre has the highest number of releases on Netflix.
2. Drama genre has the highest number of votes
3. The most popular movie is Spider-Man: No Way Home and it's genres are Action, Adventure and Science Fiction.
4. The least popular movie is The United States vs. Billie Holiday and it's genres are Music, Drama and History.
5. The year 2020 got the most number of releases