Project Name: Netflix Content Strategy Analyzer

1. Data Cleaning Steps

The dataset was first imported using pandas' read\_csv() function. Several cleaning steps were performed to ensure data quality:

• Missing values were identified using df.isnull().sum() and handled appropriately by either filling them with default values (e.g., 'Unknown' for missing director names) or dropping rows where critical data was missing.

• Duplicate rows were removed using df.drop\_duplicates() to avoid data repetition.

• Column names were standardized for readability and consistency.

• Data was filtered to create separate subsets for Movies and TV Shows when necessary.

2. Key Metrics and Insights

Several useful metrics were generated from the dataset to analyze Netflix's catalog:

• Total number of titles available on Netflix.

• Distribution of Movies vs. TV Shows.

• Number of titles added per year to observe growth trends.

• Top 10 countries contributing to Netflix content.

• Most common ratings assigned to Netflix titles.

• Top 10 directors with the most titles available.

• Frequency of genres/categories to identify the most popular genres.