

## Summary of Data Cleaning Changes

### 1. Missing Values Treated:

- Filled 'director' with "Unknown".
- Filled 'cast' with "Not Specified".
- Filled 'country' with "Unknown".
- Forward-filled 'date\_added' using .ffill().
- Replaced missing 'rating' with "Not Rated".
- Filled missing 'duration' with "Unknown".

### 2. Duplicates Removed:

- Removed exact duplicate rows using drop\_duplicates().

### 3. Standardized Text Formats:

- Title-cased 'type' and 'country' columns.
- Renamed all column headers to lowercase with underscores.

### 4. Date Formatting:

- Converted 'date\_added' to datetime format.

### 5. Data Types Corrected:

- Ensured 'release\_year' is integer.
- Created 'year\_added' column from 'date\_added'.