**Data Cleaning -**

In the process of preparing our dataset for analysis, we conducted comprehensive data cleaning and transformation tasks. The main objectives were to handle null and NA values, create new columns for analysis, standardize date formats, remove unnecessary symbols, and convert runtime into a more readable format.

**1. Handling Null and NA Values:**

Identified and treated null and NA values in the release date columns.

Removed rows with null or NA values in specific columns to ensure data integrity.

**2. Date Transformation:**

Extracted Year and Month from the release date column, creating new columns for easy analysis.

Converted the date format to a more standardized representation for consistency in the dataset.

**3. Type Conversion:**

Ensured consistent data types across columns by converting them from one format to another.

Verified and adjusted data types to enhance analytical accuracy.

**4. Symbol Removal:**

Removed unnecessary symbols, such as dollar signs, from relevant columns.

Ensured uniformity for numerical columns, facilitating straightforward numerical analyses.

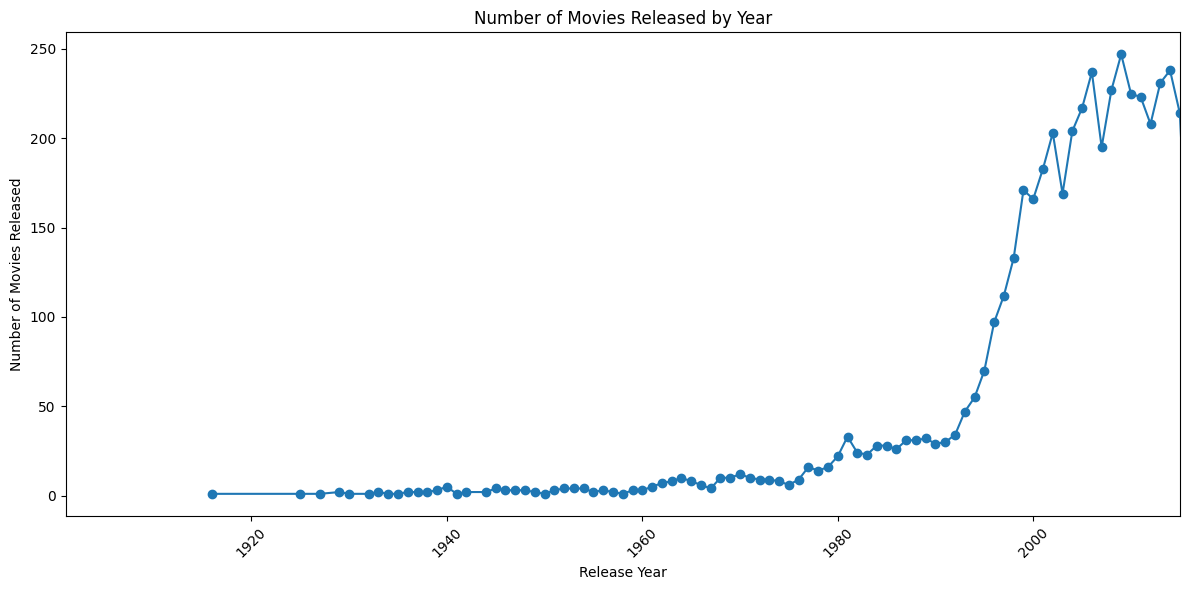
**5. Runtime Conversion:**

Standardized the runtime format into minutes for a more consistent and understandable measure.

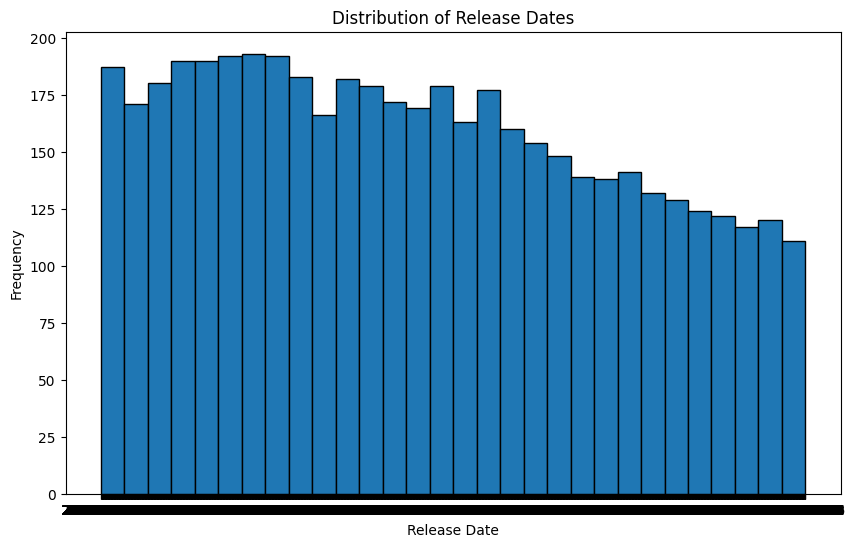
Ensured the runtime data is now in a readable format.

**EDA**

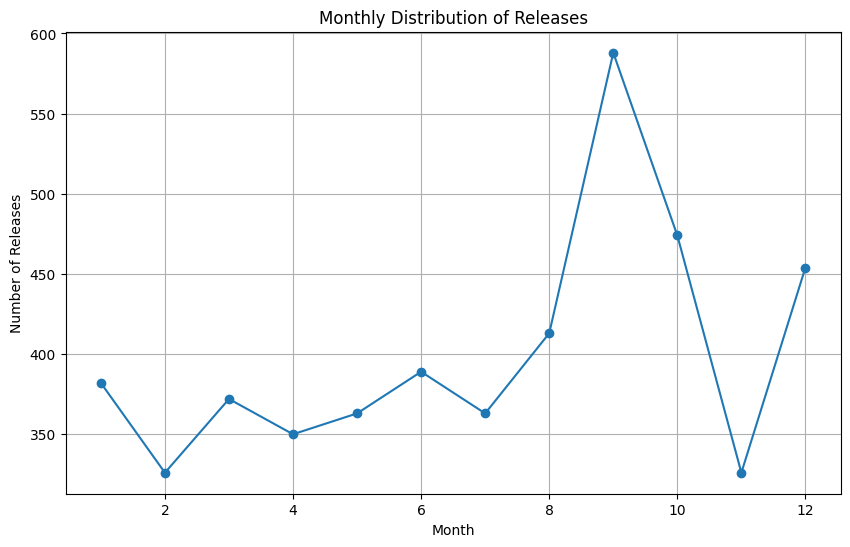
**Time series plot - Number of Movie releases - Year**



**Histogram and Distribution - Release Date frequency**

****

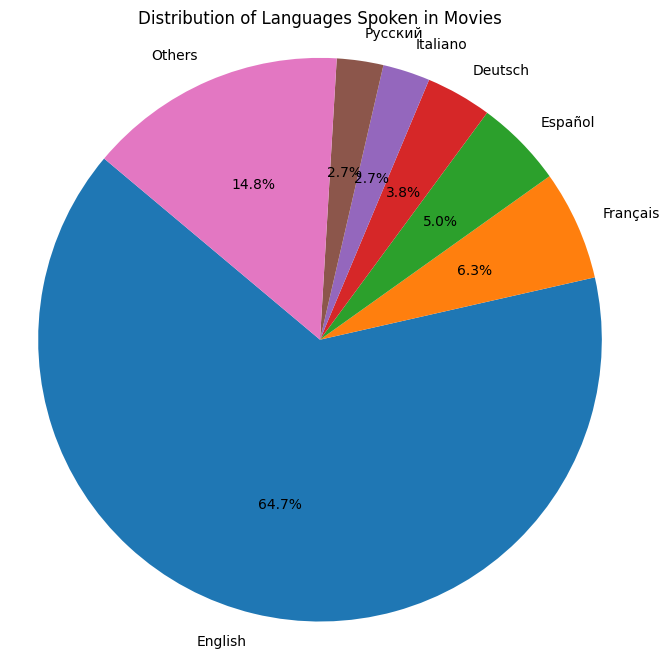
**Seasonality Analysis Monthly**

****

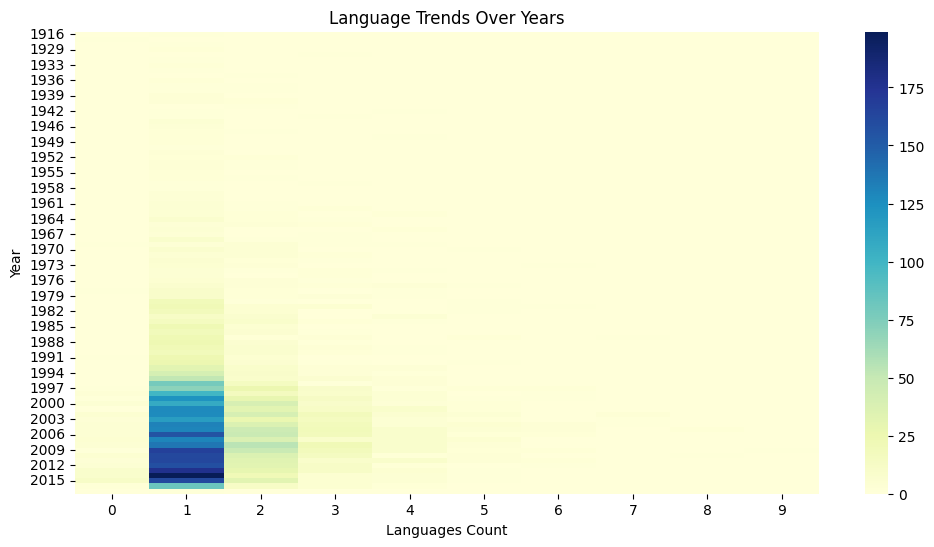
**Weekend vs. Weekday Releases**

(-3.139337589399419, 0.001703581856420019) - Based on the t-statistic and p-value, we can say that there is a significant difference in movie revenue between weekday and weekend releases, with weekend releases having a higher mean revenue.

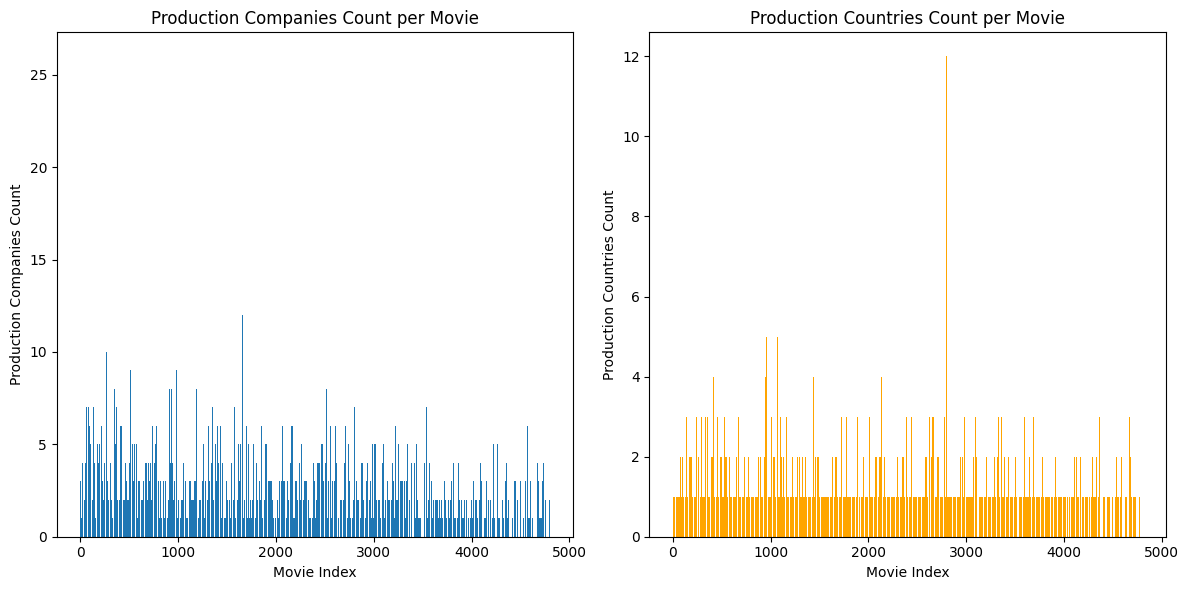
**Languages frequency - Movies**



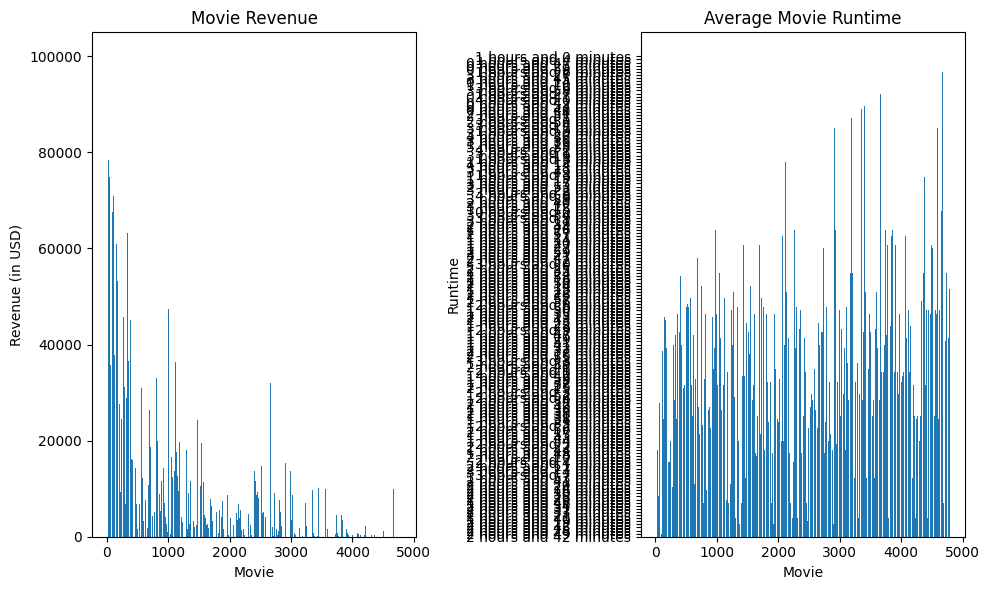
**Language Trends over Years**

****

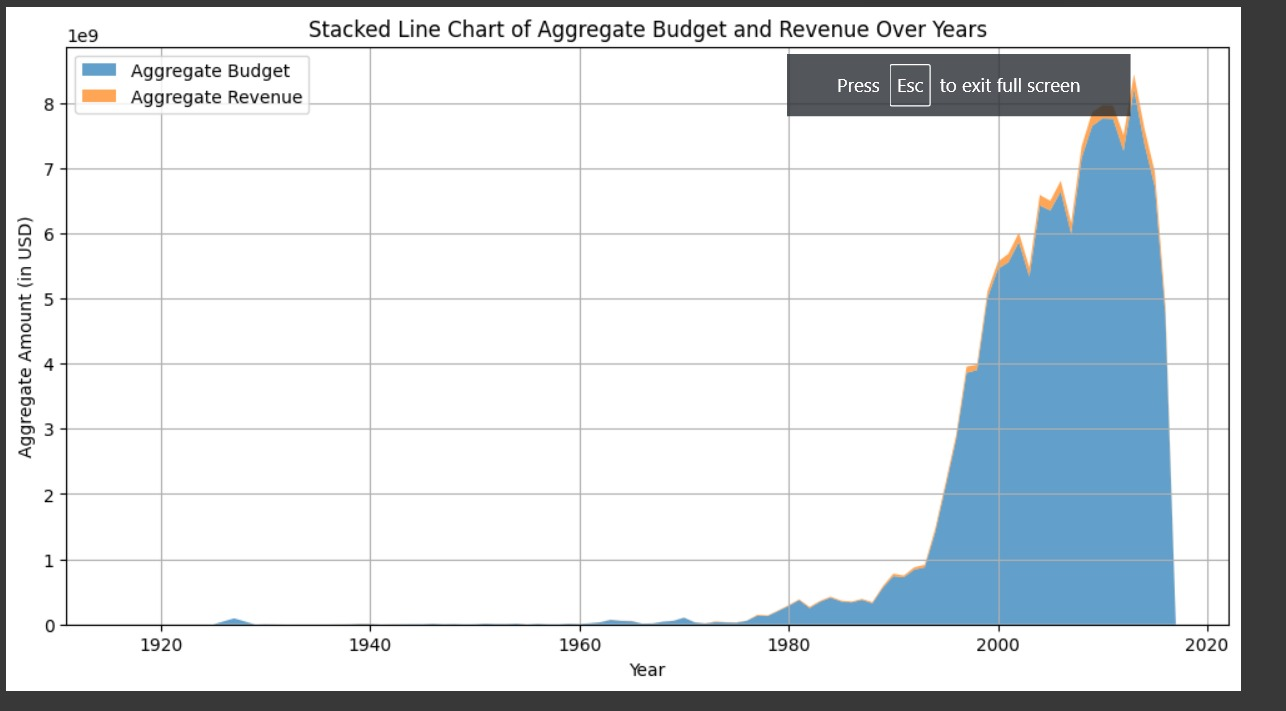
**Bar plots side by side to visualize the counts of production companies and production countries per movie**

****

**Movie runtime vs revenue**

****

**Revenue vs Budget stacked line chart**

****