

Work by Kinyanjui.

**MOVIES ANALYSIS  
PROJECT.**

**By Kinyanjui .C. Kamau.**

# PROBLEM STATEMENT

I am to explore the types of films doing the best on the box office.

My hypothesis will be ;

- The International market will be more profitable than the domestic.
- The English language is by far the most available language for movies.
- There are consistent times of the month every year which yields lower profits when movies are released on, termed as 'Dump Months' and also there is a preferred length of time for movies which the audience can make time.
- There are genre types which are preferred by the majority of movie watchers.

## ***DATA UNDERSTANDING.***

The data was collected from **github** . There are 6 datasets provided by ;

- Internet Movie Database (**IMDB**)
- The Movie Database (**TMBD**)
- Box Office Mojo (**BOM**)
- Rotten Tomatoes website (**RT**)
- The Numbers (**TN**)

Work by Kinyanjui.

## ***DESCRIBING THE QUESTION***

As a hired Data Scientist, I am to find the types of films which are currently doing the best and translate them into actionable insights.

Work by Kinyanjui.

# DEFINING THE METRIC OF SUCCESS

Using the visualizations to show that;

- There exist the worst months to release films.
- There are preferred genres.
- More money is to be made Internationally.

Work by Kinyanjui.

## ***DATA PREPARATION***

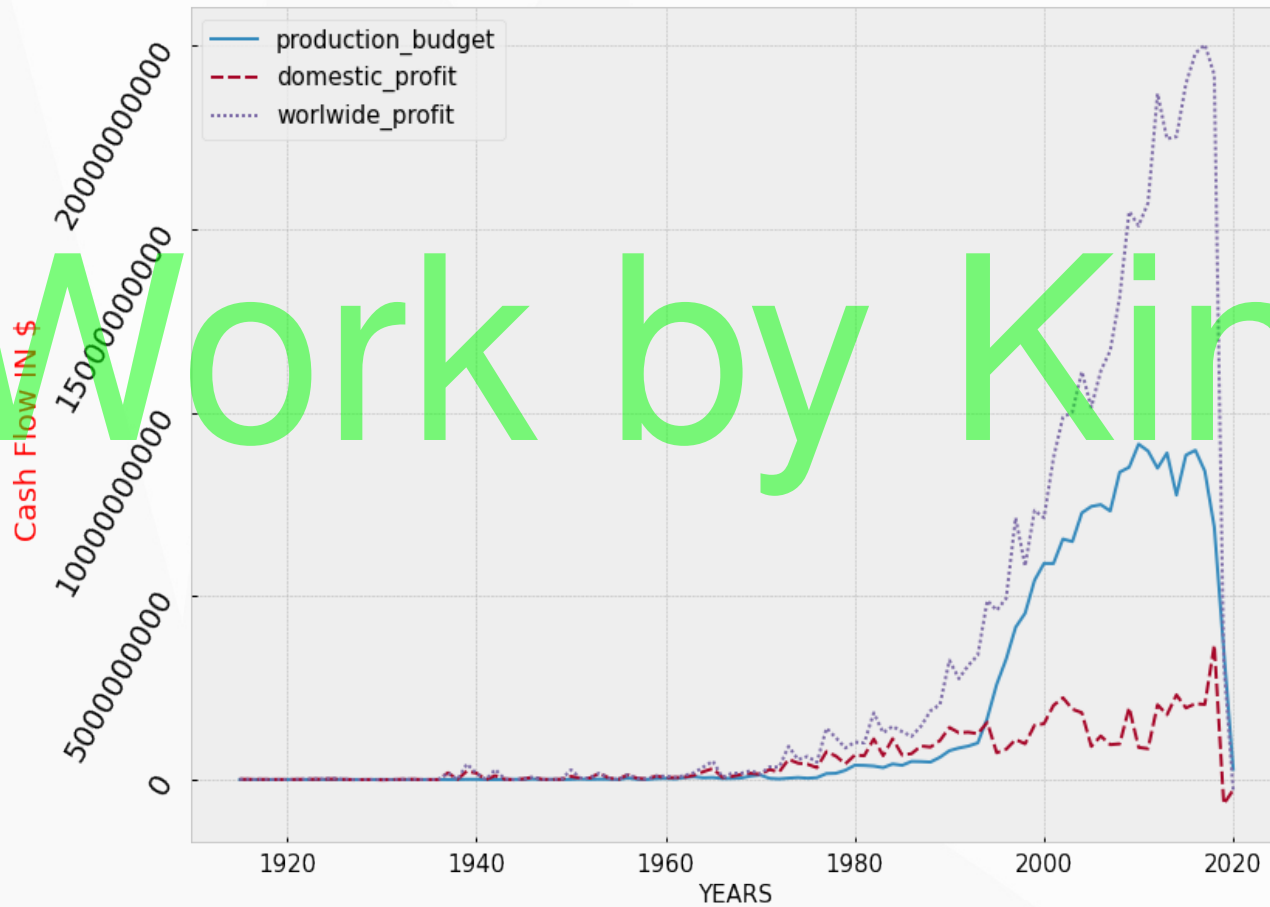
I will select the following tables to work with;

- tn.movie\_budgets.csv.gz database.
- tmdb.movies.csv database.
- *imdb.sql* database.

Work by Kinyanjui.

# DATA VISUALIZATION

FINANCIAL TRENDS OVER THE YEARS



Work by Kinnyanjui.

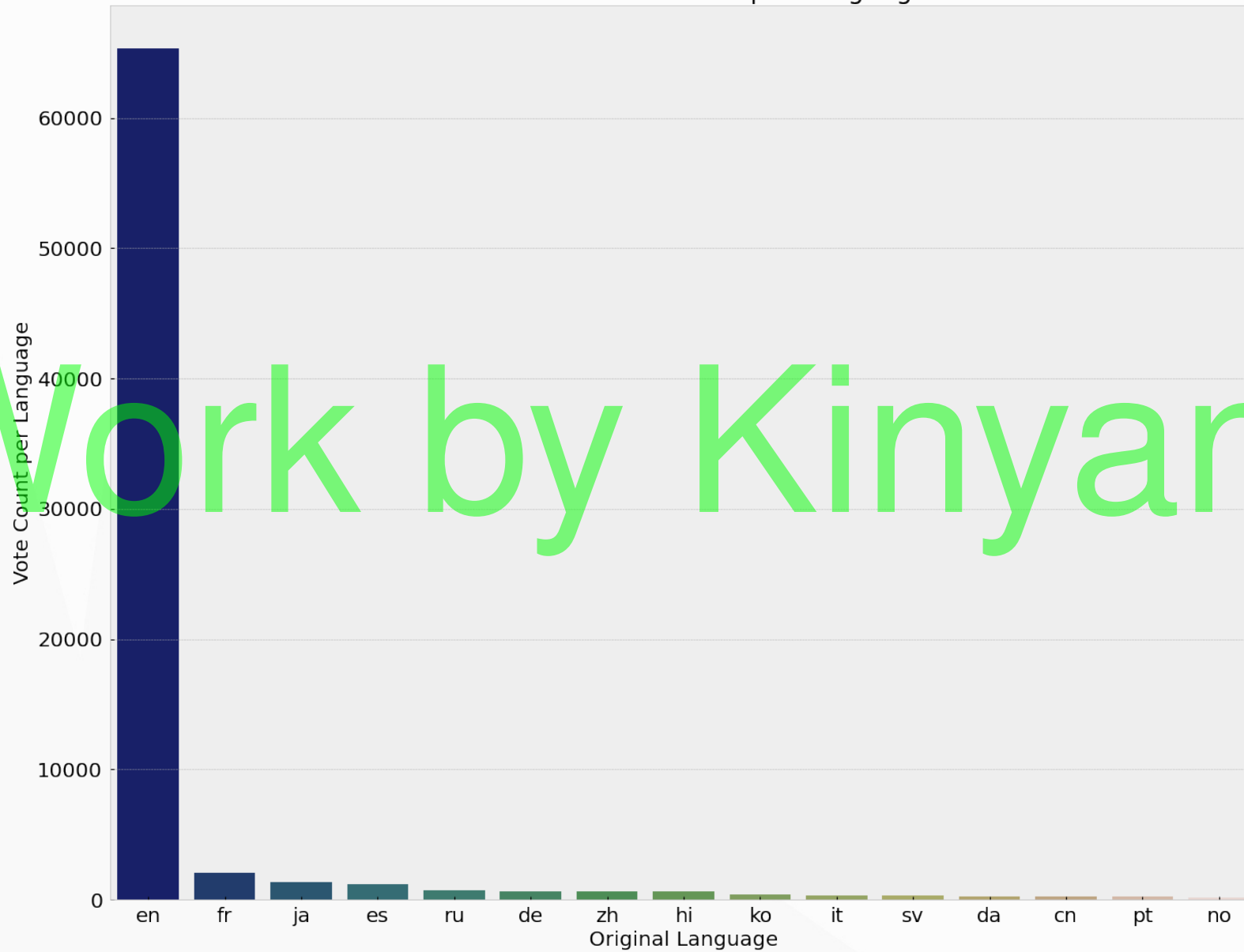
The previous line plot follows the trend of cash flows based on the included movies. Since the 1920's to 2020. It was grouped by year of release and summed to make the Data Frame "*Budget\_yearly*".

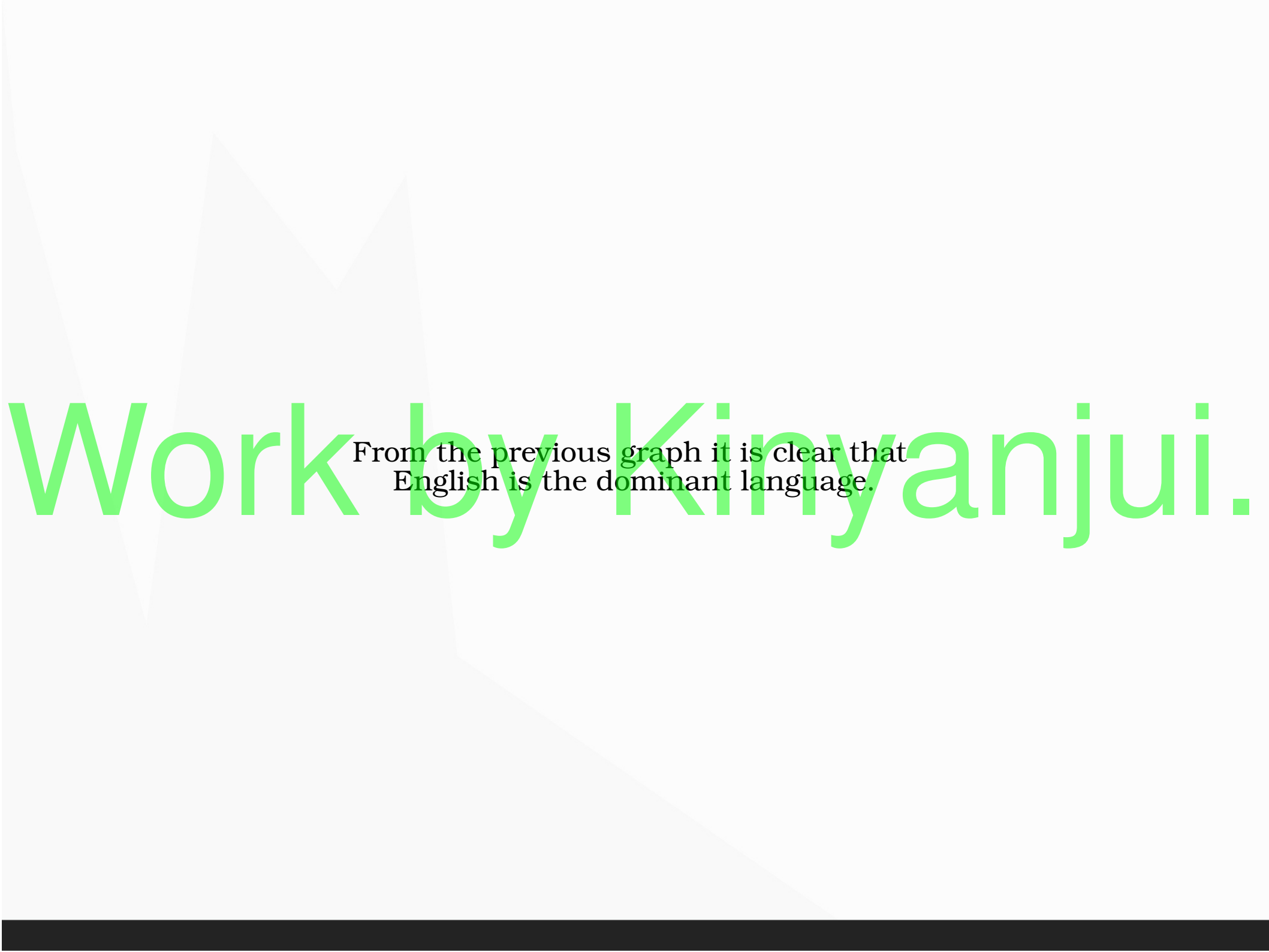
It is clear that the industry has been growing steadily and shot up between 1990 and 2000, especially the worldwide\_profit, basically the international audience.

The worldwide\_profit is clearly more attractive than the domestic profit.



Vote Count Didtribution per Language

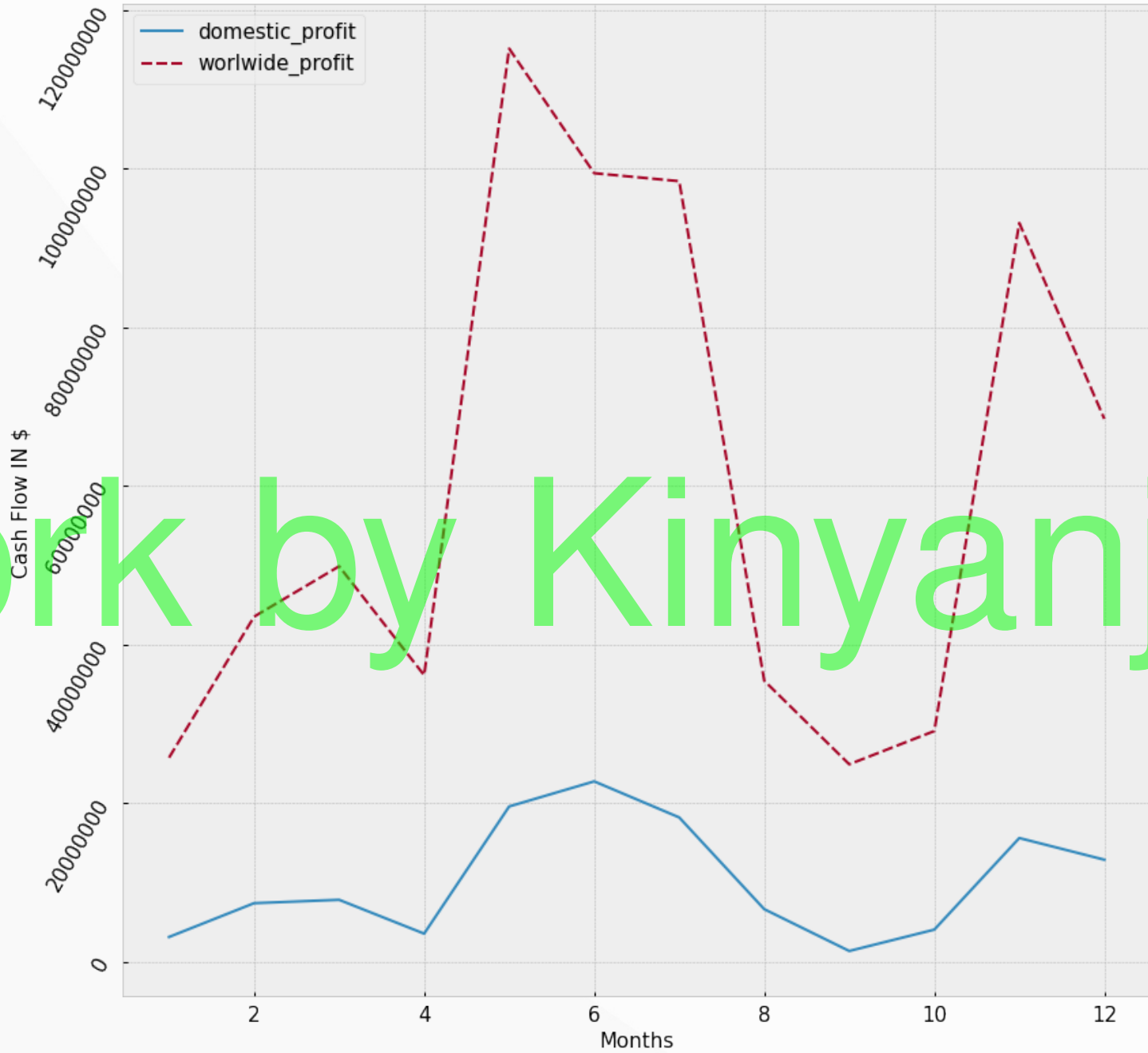




# Work by Kinnyanjui.

From the previous graph it is clear that  
English is the dominant language.

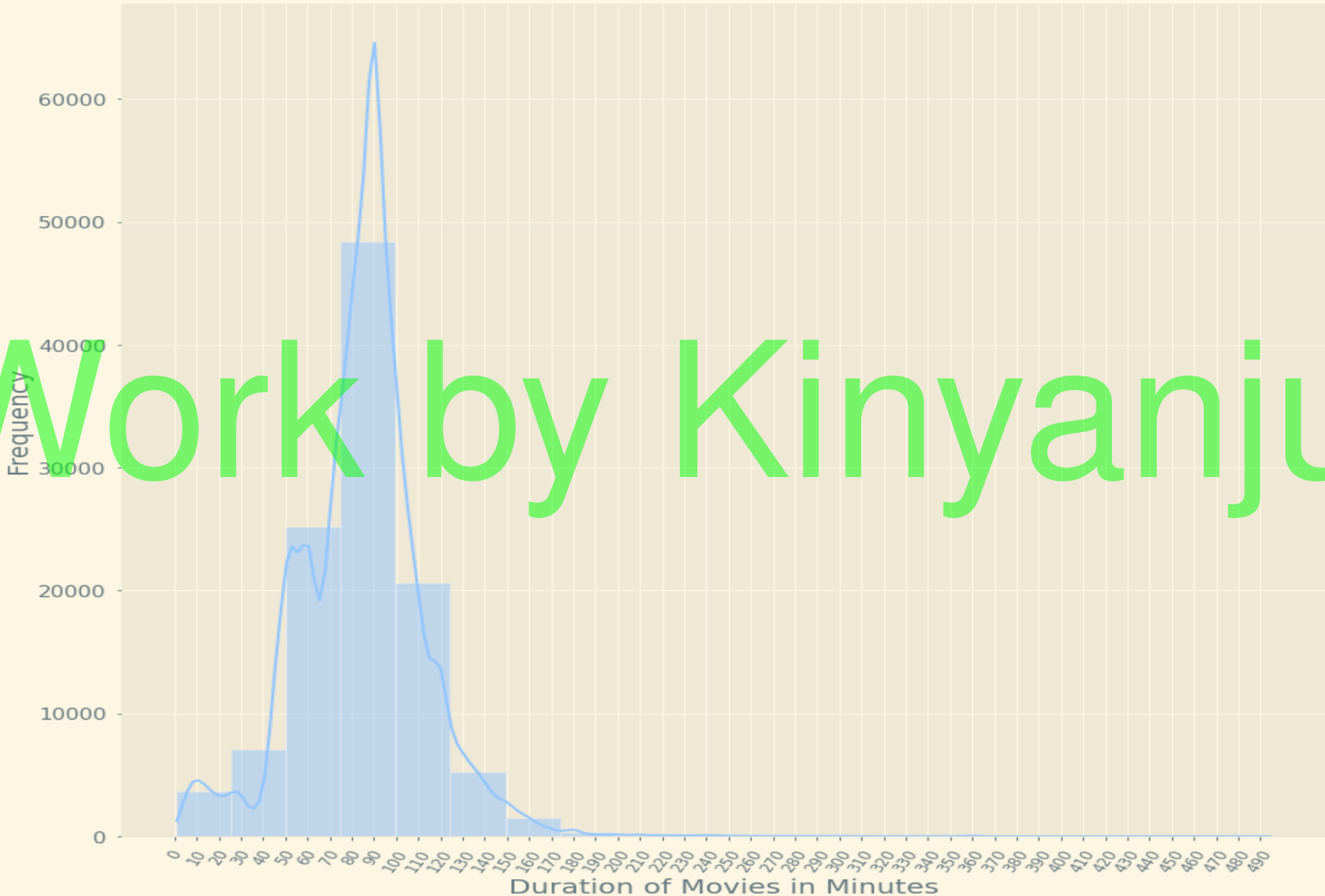
Cash Flow On Diff Release Months



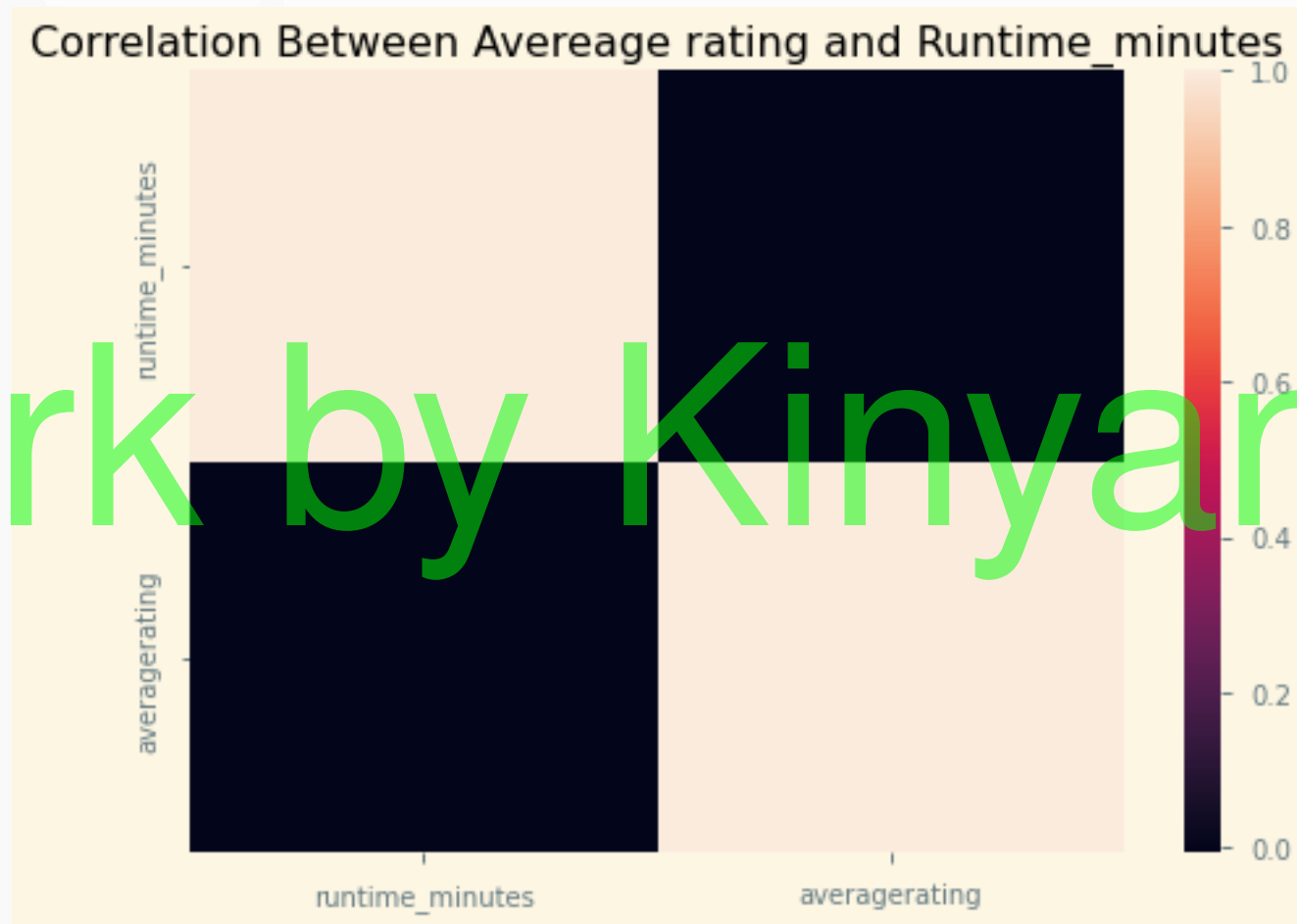
Work by Kinyanjui.

There are consistent times of the month every year which yields lower profits when movies are released on, termed as 'Dump Months' and also shows a goldilocks period with high profits.

Distribution Plot



Work by Kinyanjui.

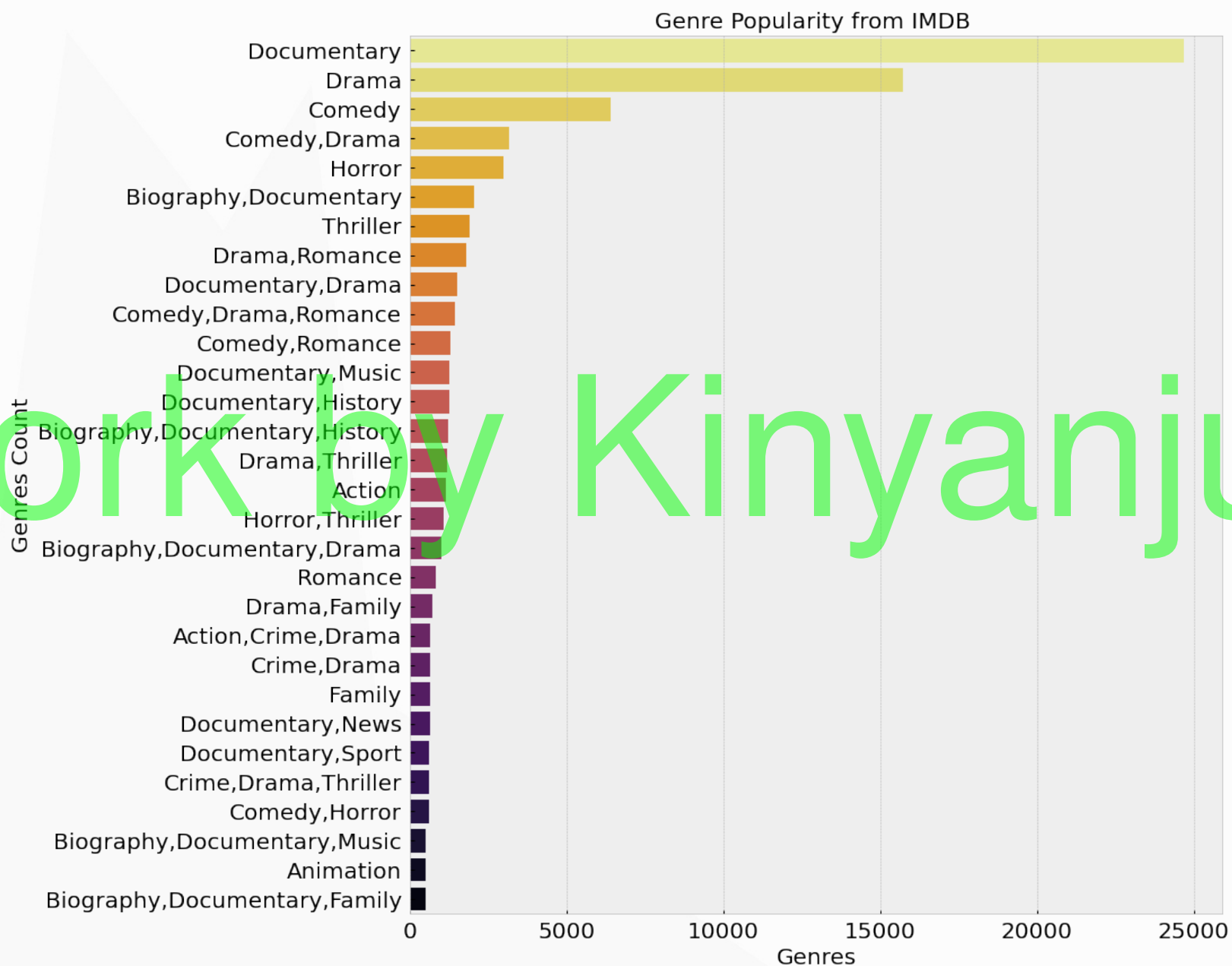


Work by Kinyanjui.

RUNTIME TRENDS OVER THE YEARS



Work by Kinyanjui.





Genres Popularity from AVG ratings from IMDB

