Movie Data Linear Regression in R

A data science project executed in R Studio to build a linear regression model for predicting movie performance.

About

This project seeks to answer the question, "What constitutes a timeless movie that audiences love?" A new variable was created titled "ranked movie performance" from IMBD rating, meta score, and gross film revenue. The linear regression model seeks to identify correlation between the independent variables.

Link

Netflix Dataset: https://www.kaggle.com/shivamb/netflix-shows

Rotten Tomatoes Dataset: https://data.world/prasert/rotten-tomatoes-top-movies-by-genre

IMDB Movies Dataset: https://www.kaggle.com/harshitshankhdhar/imdb-dataset-of-top-1000-movies-and-tv-shows

Setup

This coding project was constructed in R Studio. Please import the following libraries:

- Ggplot2
- Readxl
- Purr
- Plyr
- Dplyr
- Magrittr
- Stringr
- Tidyverse
- Tidyr