An Exploratory Data Analysis on Movie Data

• • •

Flatiron School - Phase 1
June 3rd, 2022

Outline

- Overview
- The Business Problem
- Data Sources
- Methods
- Results
- Recommendations

Overview

- Microsoft Studios is up and coming

- Where do we start?
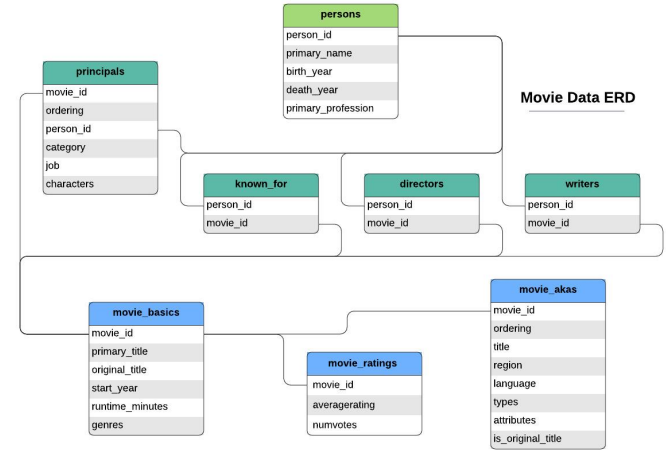 - Types of films?
  - Budgets?
  - Ratings?

The Business Problem:

- To explore the demographic of movies and analyze metrics for success in the box office.

- Translate these findings into actionable insights for Microsoft's new movie studio.

Data Sources

Data:

- IMDB Movie Database
  - Movie Information
- The Numbers (www.the-numbers.com)
  - Budgets
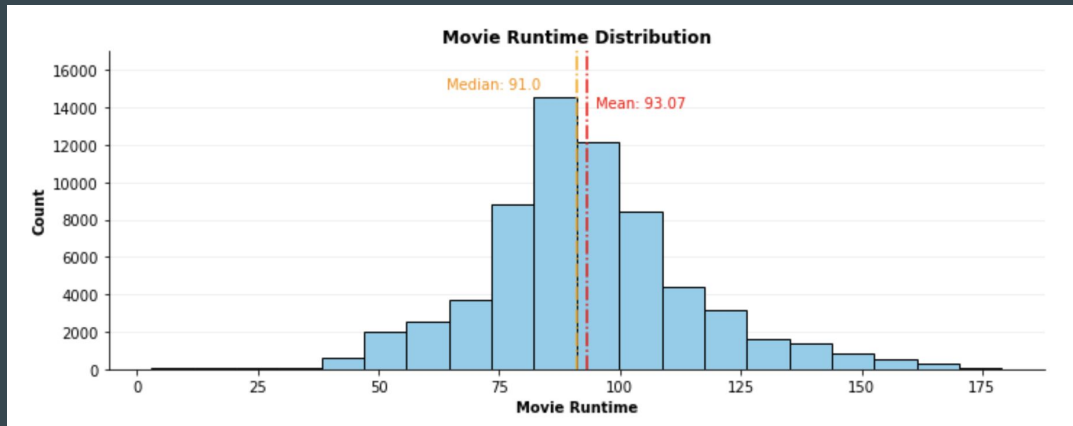

Movie Data ERD

Methods

Data Cleaning
- Cleaning across datasets
- Joining relevant information
  - Budgets, ratings, release dates

Results

Movie Runtime

Most movies are centered around a 90 minute runtime.

- Average rating for 90 minute movies range between 5 - 7 (on a scale of 1-10).
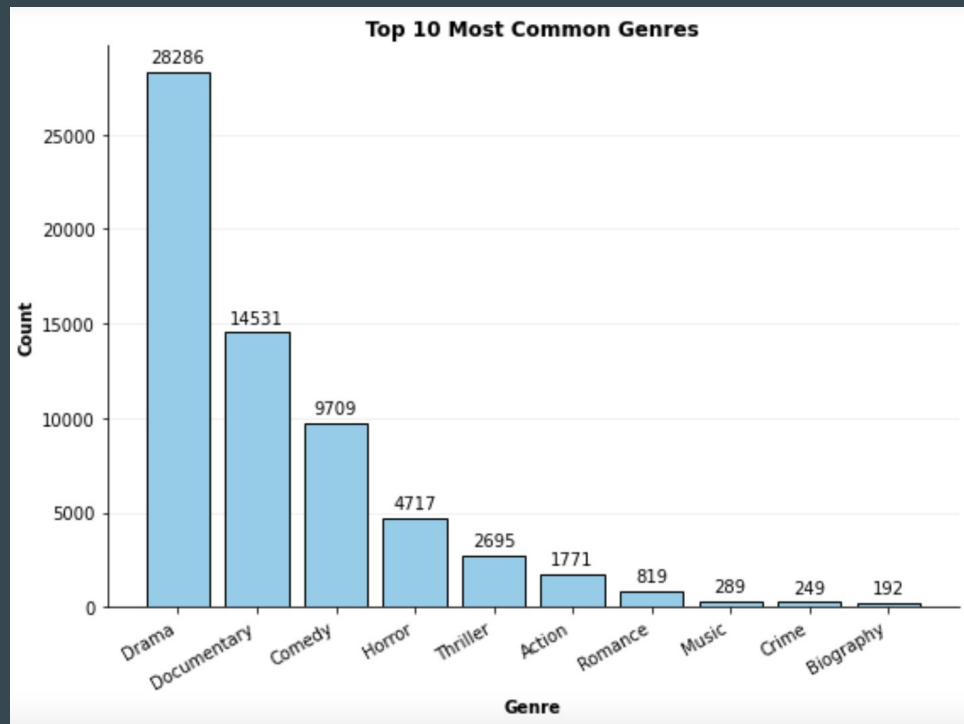- Generally, movies with less than a 90 minute runtime yielded higher ratings.

Results

Top Genres

Most frequently classified genre is the "Drama"

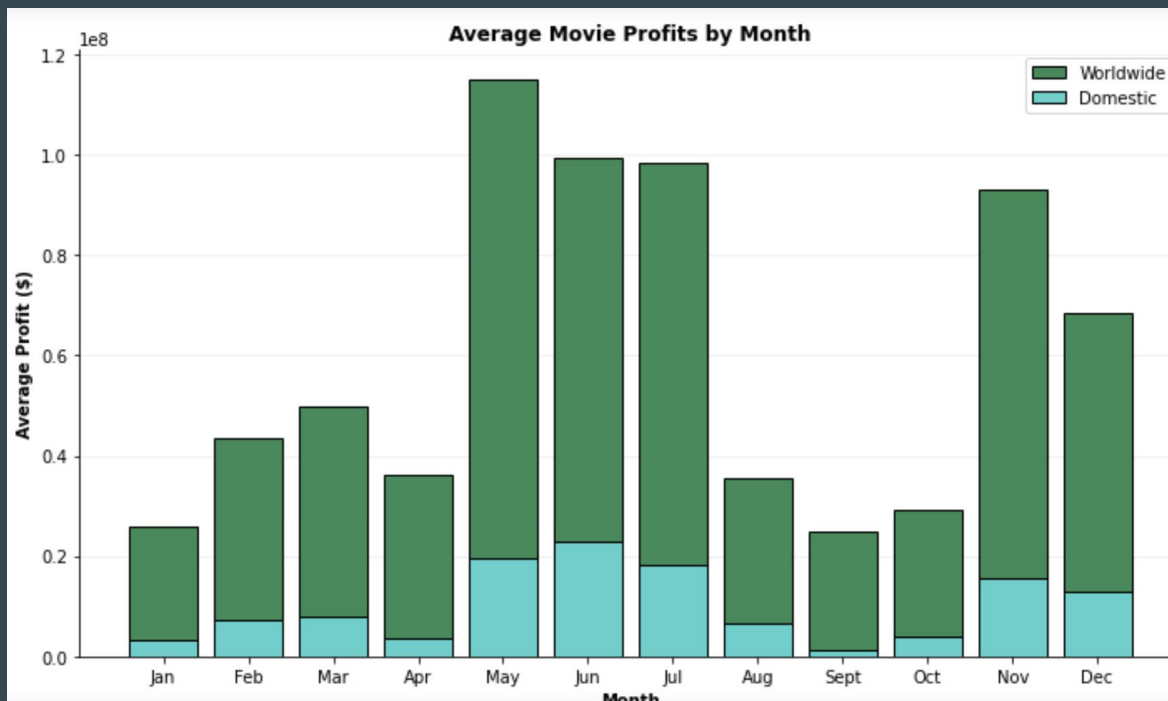- But does this necessarily mean the most *popular?*.

Top 10 Most Common Genres

Results

Profits - Month

More People = More **Profit**

Most popular times of the year

- Summer, holiday months
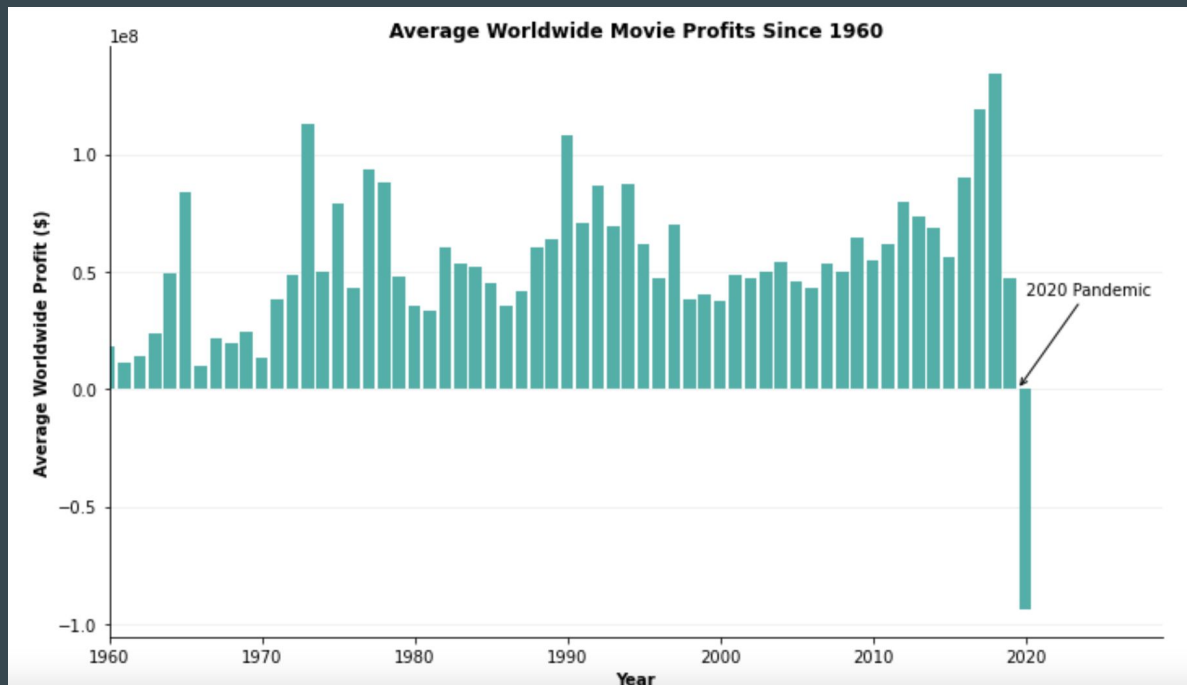- More kids out of school

Results

Profits (Since 1960)

Cyclical trend over decades.

Peaks in the worldwide profits

- Early 90s blockbusters
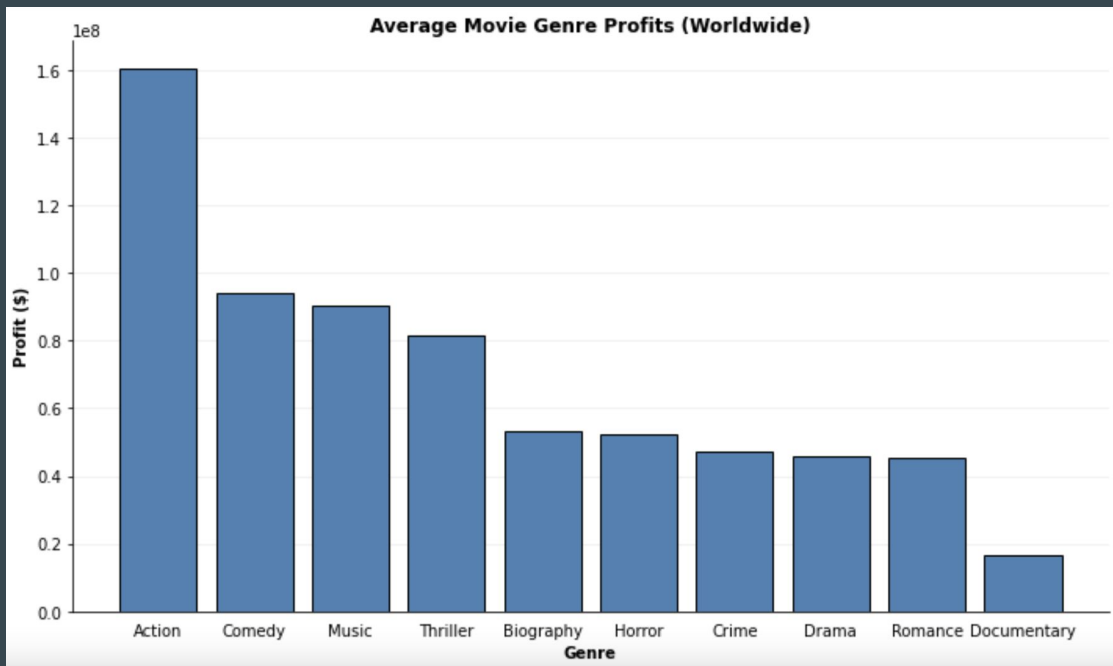- Mid-late 2010s superhero movies

2020 Pandemic

Results

Profits by Genre

Most profitable movies are action movies

Compare to most frequently classified genre

Average Movie Genre Profits (Worldwide)

Results

Top 20 Movies

Half of the top 20 movies are classified under the action, adventure, sci-fi genre.

- Trending towards kid friendly movies
- Adjust for inflation

Recommendations

Runtime & Ratings	• Most movies are 90 minutes • Generally, shorter movies are more highly rated.
Movie Genres	• Drama, documentaries, comedy most frequent • Action, comedy, music generate most profit (worldwide) • Focus on action, adventure, and sci-fi
Release Date	• Movies generate greatest profit in the summer and winter holiday months
Future data needed	• Monitor movie theater attendees post 2020

Questions?

Email: eric8395@gmail.com

GitHub: https://github.com/eric8395/Movie-EDA

LinkedIn: www.linkedin.com/in/eric-au8395/