MS Box Office Insights

Finding Opportunity Through Data

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

- According to data from The-Numbers.com, the US box office has increased from 5.3 billion USD in 1995 to 11.3 billion USD in 2019.
- Identifying opportunities with the highest probability for profit and ROI in this growing market is our strategic priority.

Data

- Data from Rotten Tomatoes and The-Numbers.com were combined to form a primary data frame of 4,776 films.
- The-Numbers.com and Internet Movie Database data were joined into a complimentary dataframe of 32,088 individuals in the film industry.

Methods

SQLite, Python, Pandas, and Seaborn were used for data preparation, analysis and visualization.

Budget and Cost Analysis

Results:
Cost Estimate for the First Year's Films

- Mean Production Budget per Film: 70,620,199 USD
- Projected Marketing Budget per Film: 35,310,100 USD
- Total Cost per Film: 105,930,299 USD
- Mean Number of Films per Year by a Major Studio: 15
- Total Estimated Cost for First Year: 1,588,954,479 USD

Results:
Budget vs Gross (Films Released Since 2009)

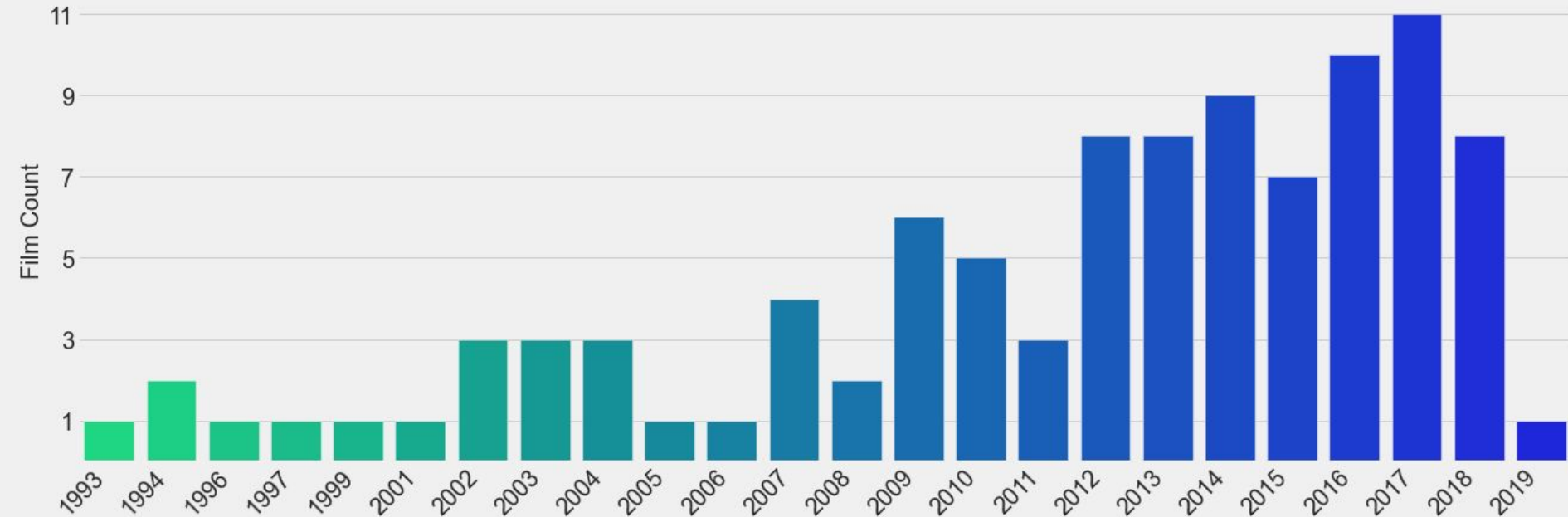Data Shows a Positive Correlation Between Budget and Gross
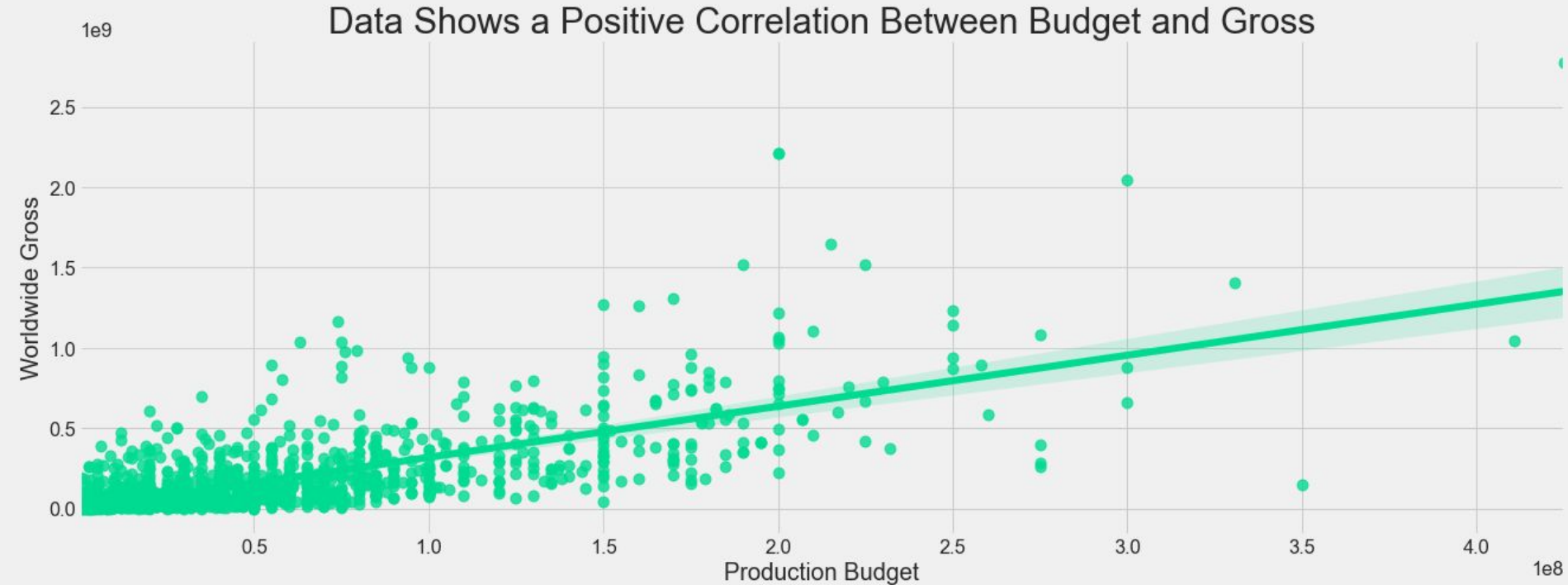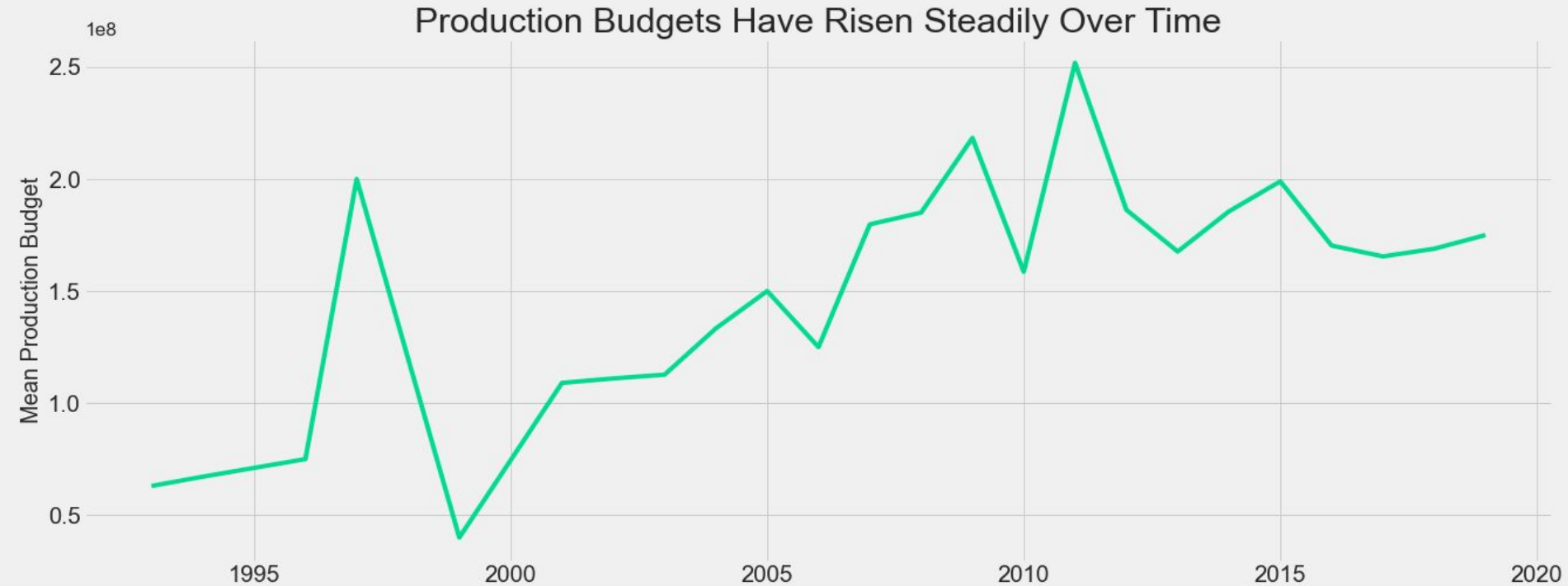
Top 100 Films by Gross

Results:
Distribution of the Top 100 Grossing Films

The Bulk of Top Grossing Films Were Released After 2008

Results:
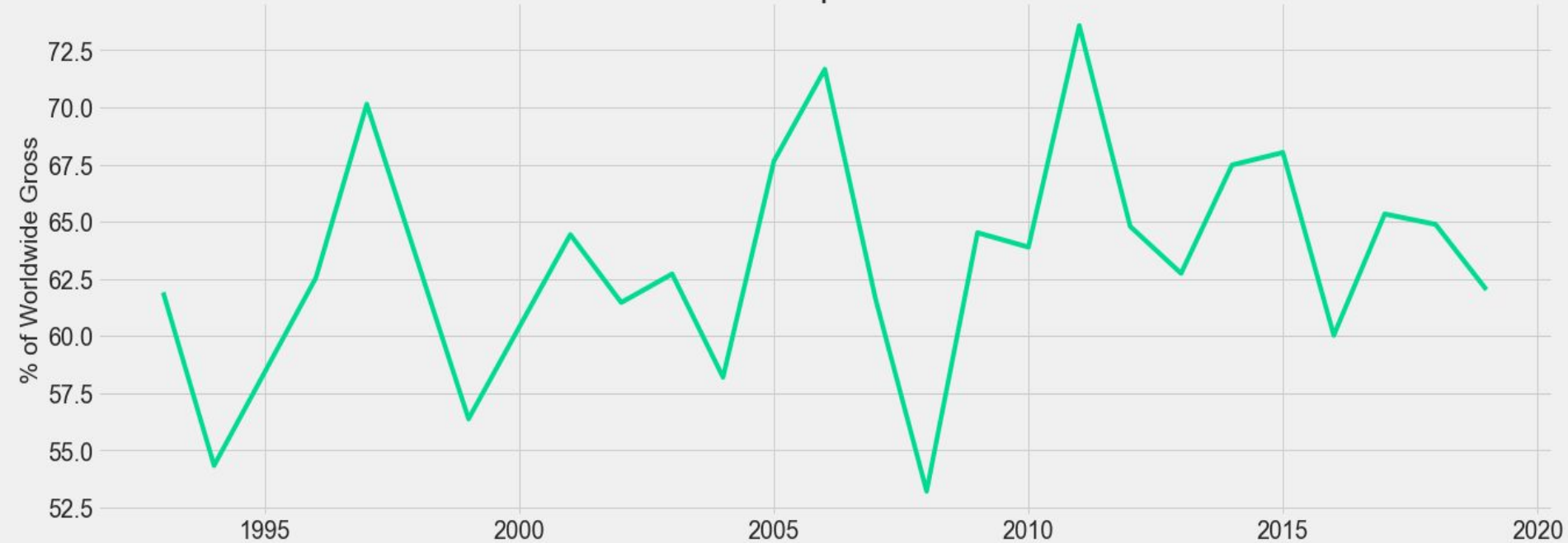# Mean Budget For a Top 100 Grossing Film

Production Budgets Have Risen Steadily Over Time

Results:
Mean Share of Gross From Outside US

The International Market Makes Up 50% to 75% of Worldwide Profits

Genre Analysis
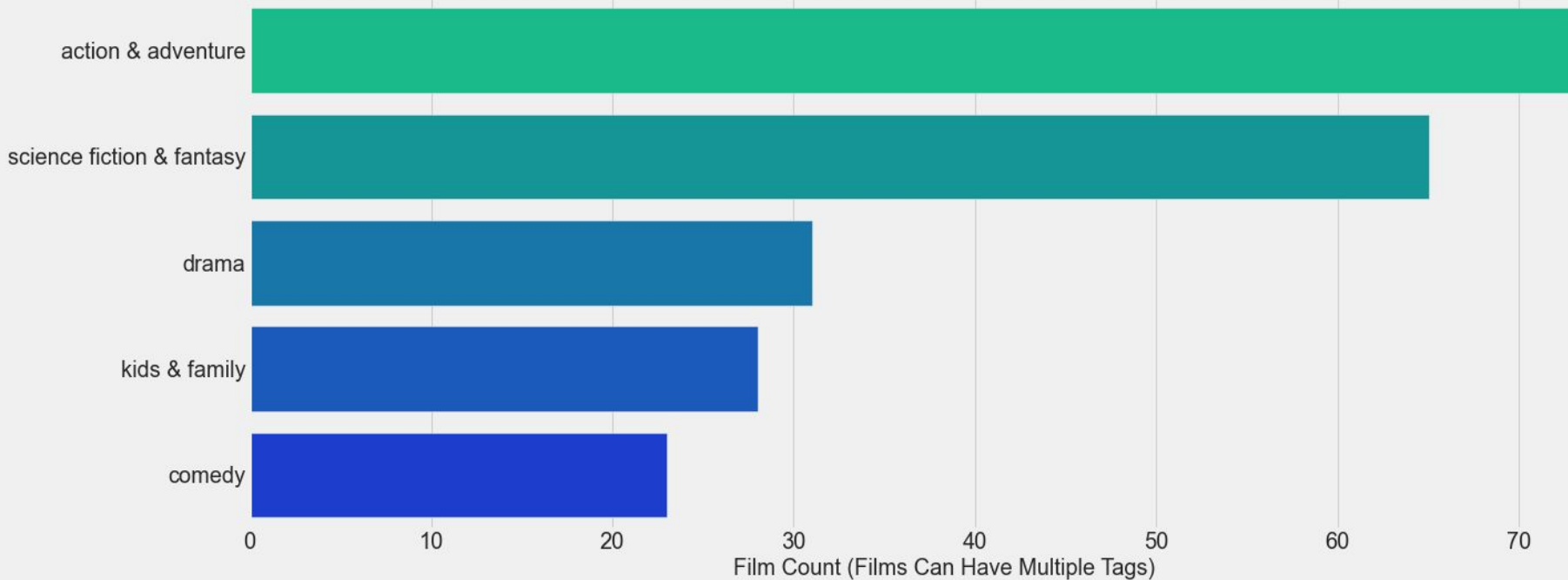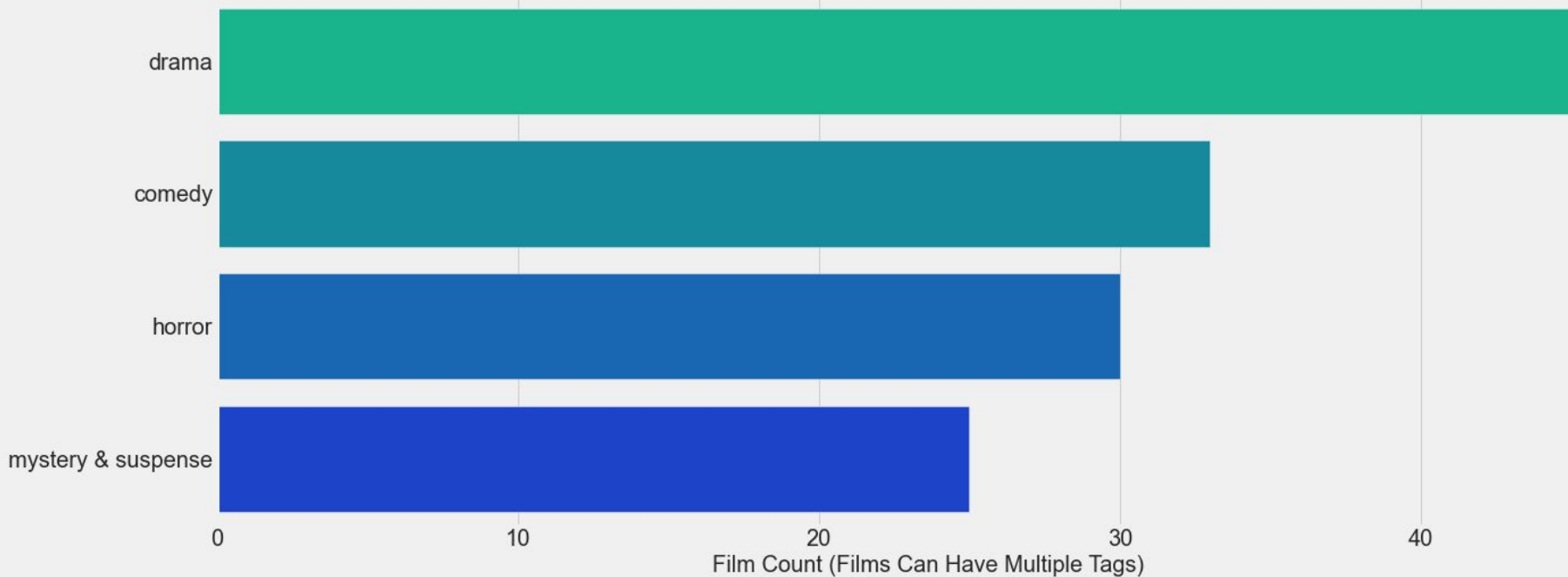
Results:
The Big Five Genres

Genre Tags with the Highest Counts in the Top 100 Films

Film Count (Films Can Have Multiple Tags)

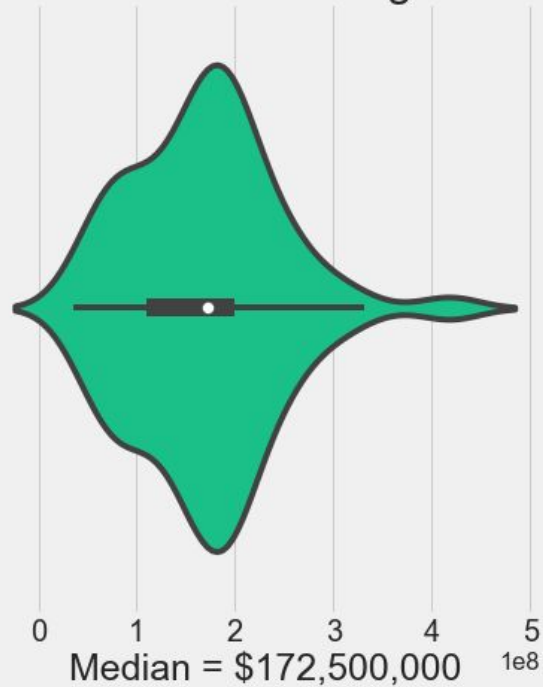# Results:
Hidden Value Genres

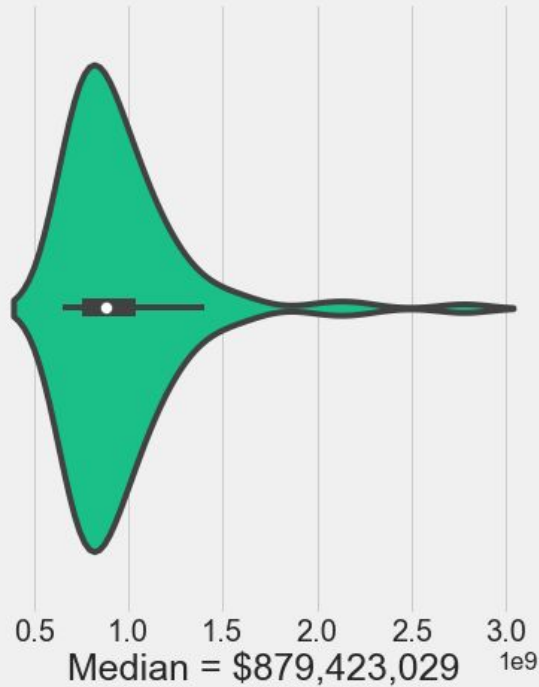Genre Tags with the Highest Counts in the Top 100 ROI Films
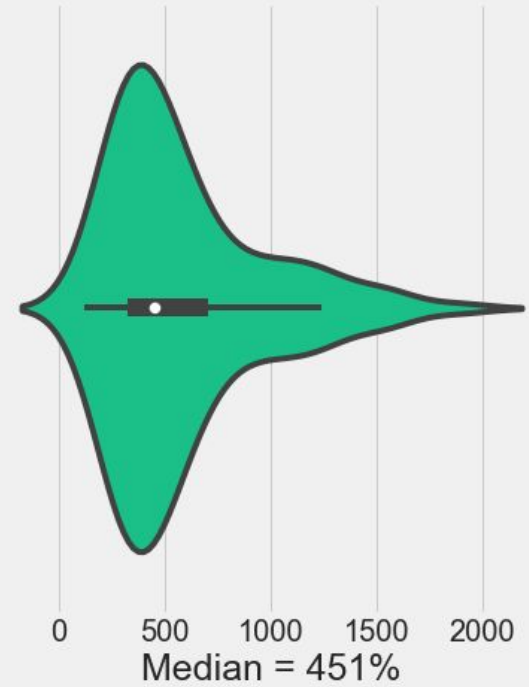
Results:
Statistics for the Top 100 Grossing Films
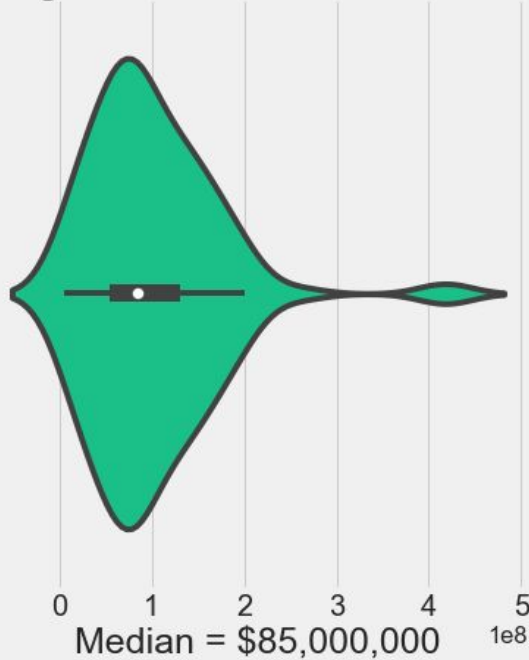

Production Budget	Worldwide Gross	Worldwide ROI
Median = $172,500,000	Median = $879,423,029	Median = 451%

Results:
Statistics For Comedy Films

Budget in Hundreds of Millions

0   1   2   3   4   5
Median = $85,000,000   1e8

Worldwide Gross in Billions

0   1   2   3
Median = $517,083,788   1e9

Worldwide ROI

0   2000   4000   6000   8000
Median = 515%
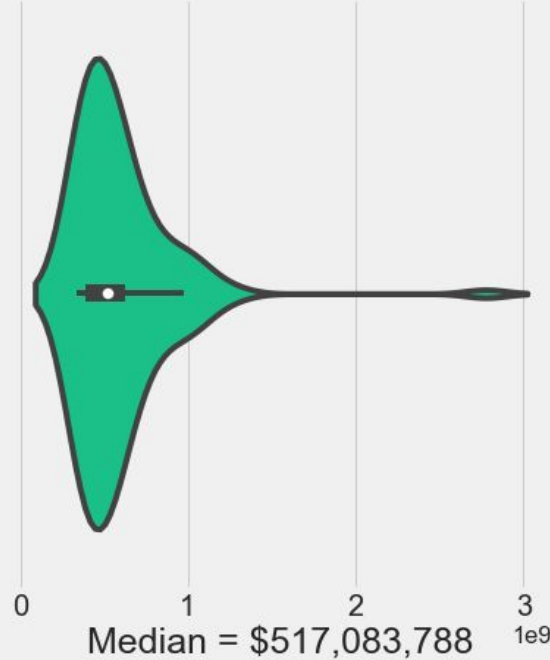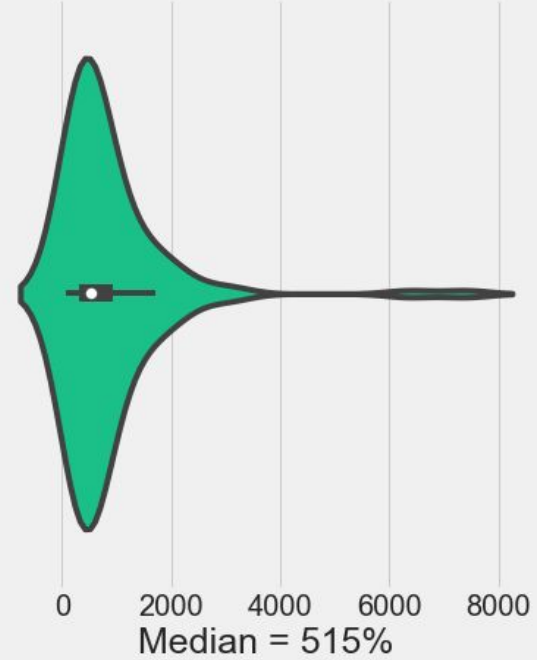
Conclusions

Summary

Descriptive analysis of film industry data reveals the following:

- The estimated cost for production and marketing in the first year: 1,588,954,479 USD.
- The bulk of film revenue will come from the international market.
- The Big Five genres of Action, SciFi, Drama, Family, and Comedy offer the best chances to maximize profits.
- The Hidden Value genres of Comedy, Horror, Mystery, and Drama have the highest probability for optimal return on investment.

Next Steps

Possible Extensions of the Analysis:

- Analysis of creative talent is available upon request.
- Incorporating streaming market data into the analysis.
- Researching the purchase of a production company.

Thank You!

Email: jeffrey.h.watson@protonmail.com
GitHub: @jeffrey.h.watson