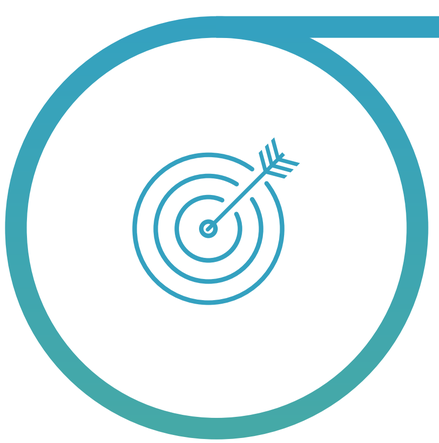
Data Science

The Movie Industry

Bao Tram Duong

Introduction

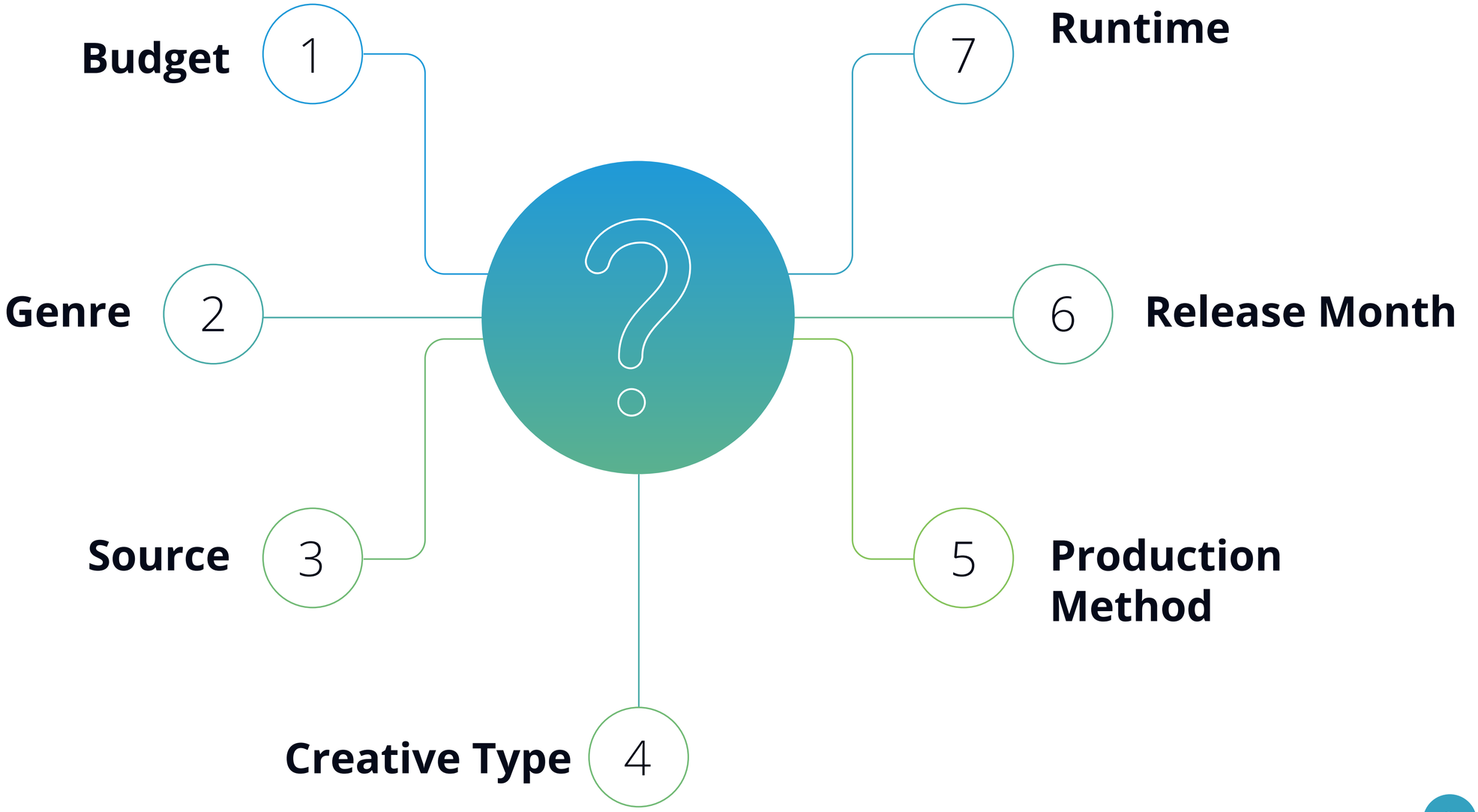


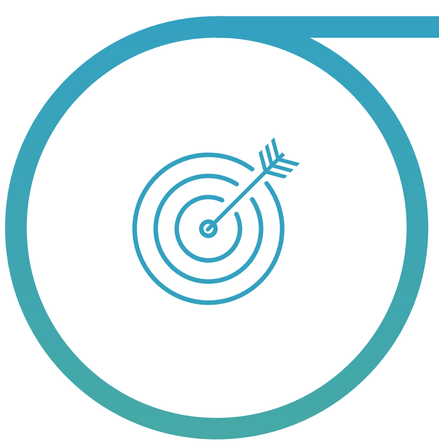
- The movie industry generated $35.3B in revenue in 2019

- The U.S is among the biggest film industries in the world

We are here to help **Microsoft** to become a powerful contributor in this industry.

Business Problems





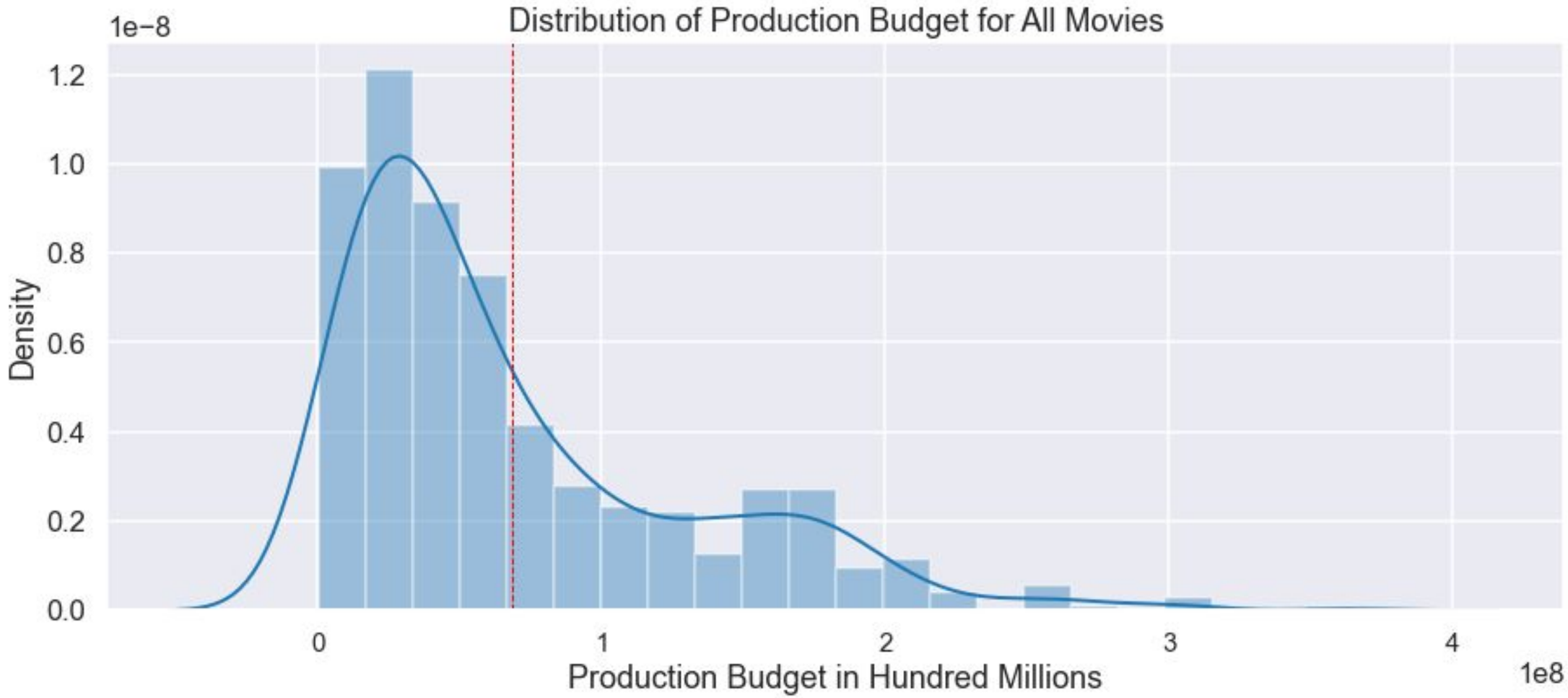
Business Problems

Q1. Is there a correlation between production budget and profit? If so, how much should Microsoft invest into production?

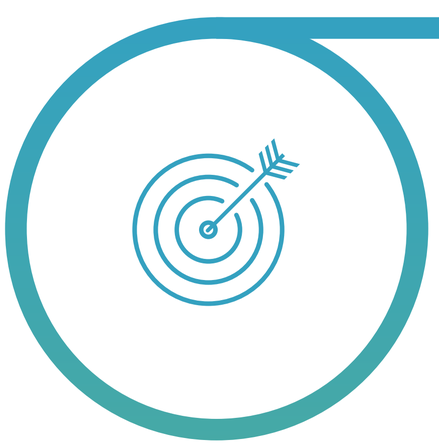
Q2. What kind of movie contents, in terms of genre, source, creative type, production method, perform the best?

Q3. When is the best time of year to release a movie?

Q4. Is there a correlation between runtime and profit? What is the best runtime?



- **$68M** to produce a major studio movie
- **$35M** for marketing and distributions

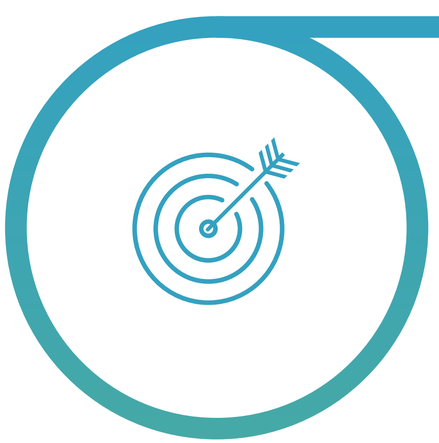


Methodology

We split our dataset into 2 groups:

- The **General Movies** group contains all 625 movies in the dataset.
- The **Top 100 Movies** group contains the top 100 movies that generate the highest worldwide profit.

Compare & observe what the Top 100 Movies are doing differently to set themselves apart



Data Sources

Predict the success of a movie based on its profitability by leveraging historical data from various sources:

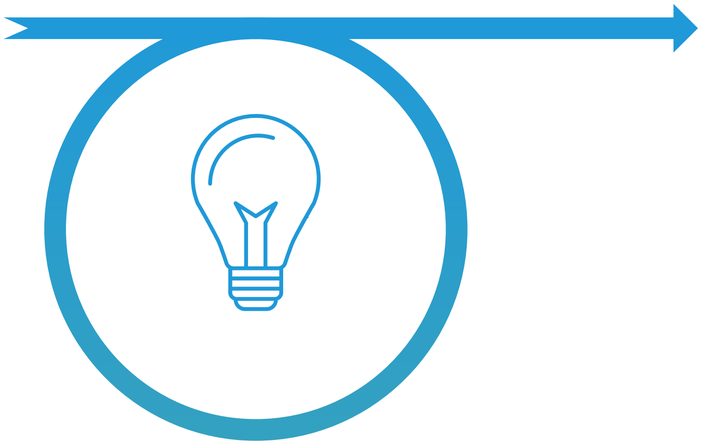
1. Box Office Mojo

2. IMDB

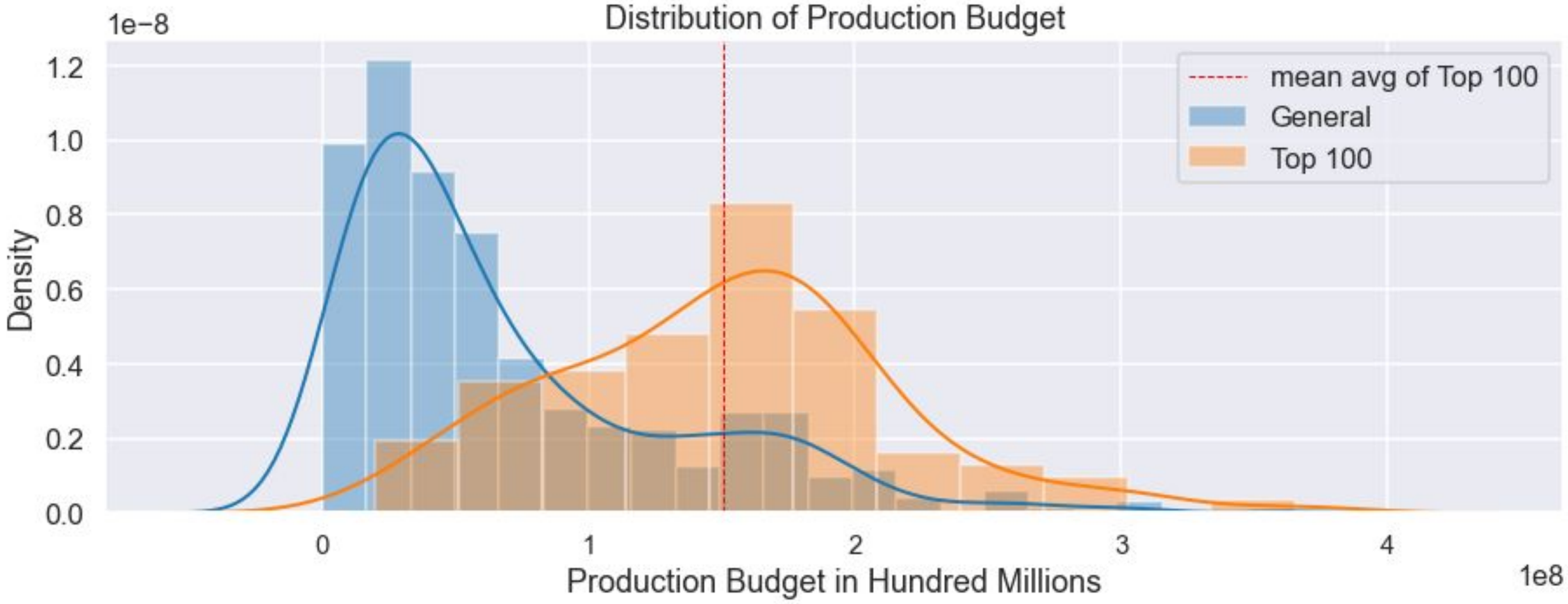
3. Rotten Tomatoes

4. TheMovieDB.org

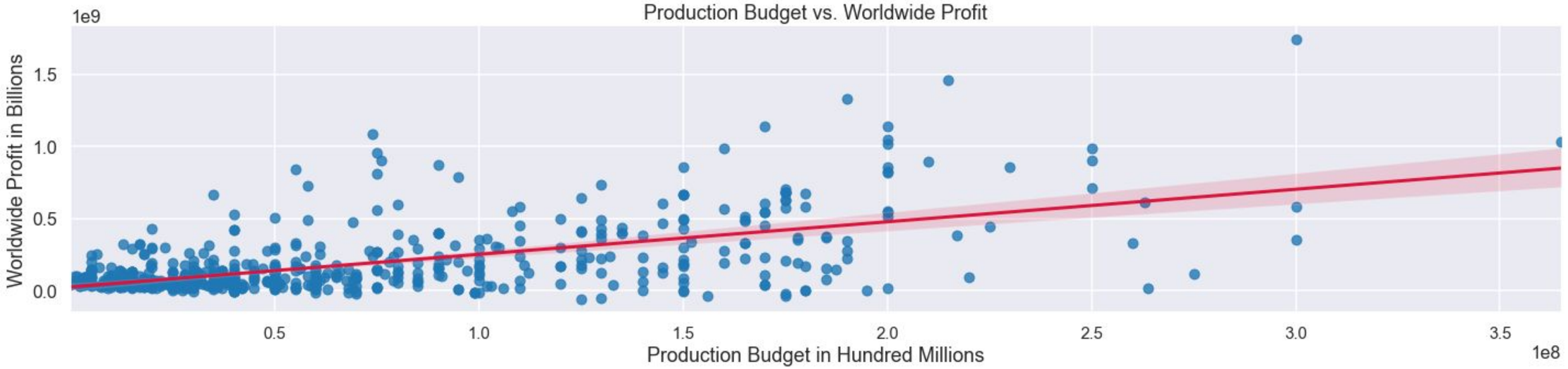
5.The-Numbers.com



Business Solution
Production Budget



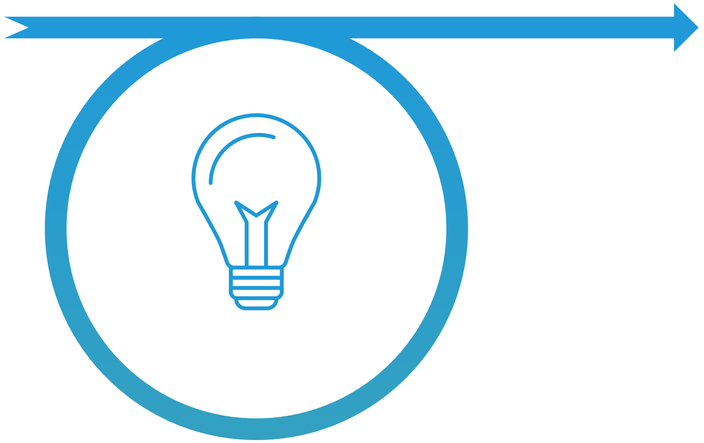
- The average production budget in the General Movies group is **$68M**.
- The average production budget in Top 100 Movies group is **$150M**.



- The average worldwide profit in General Movies group is **$179M**.
- The average worldwide profit of the Top 100 Movies group is **$624M**.

Conclusion

$150+ M



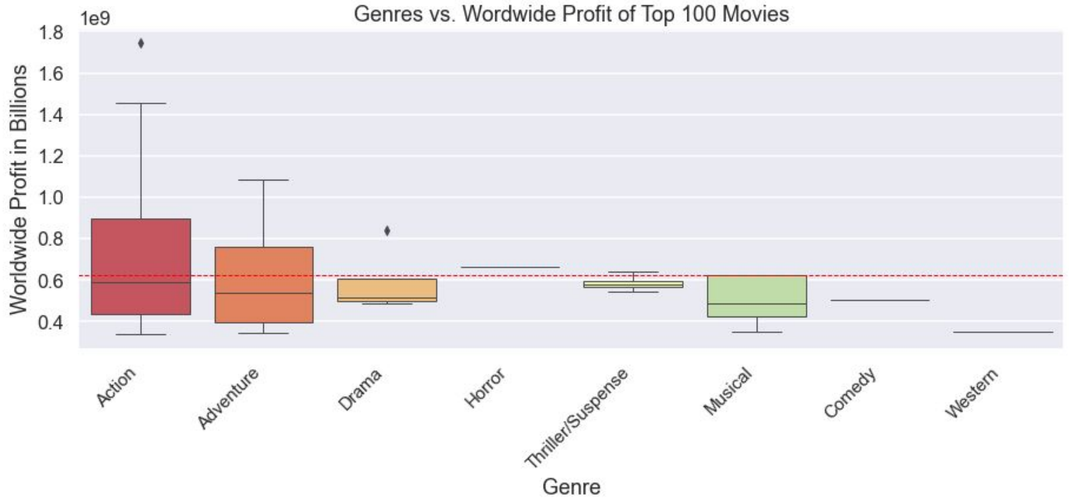
Business Solution

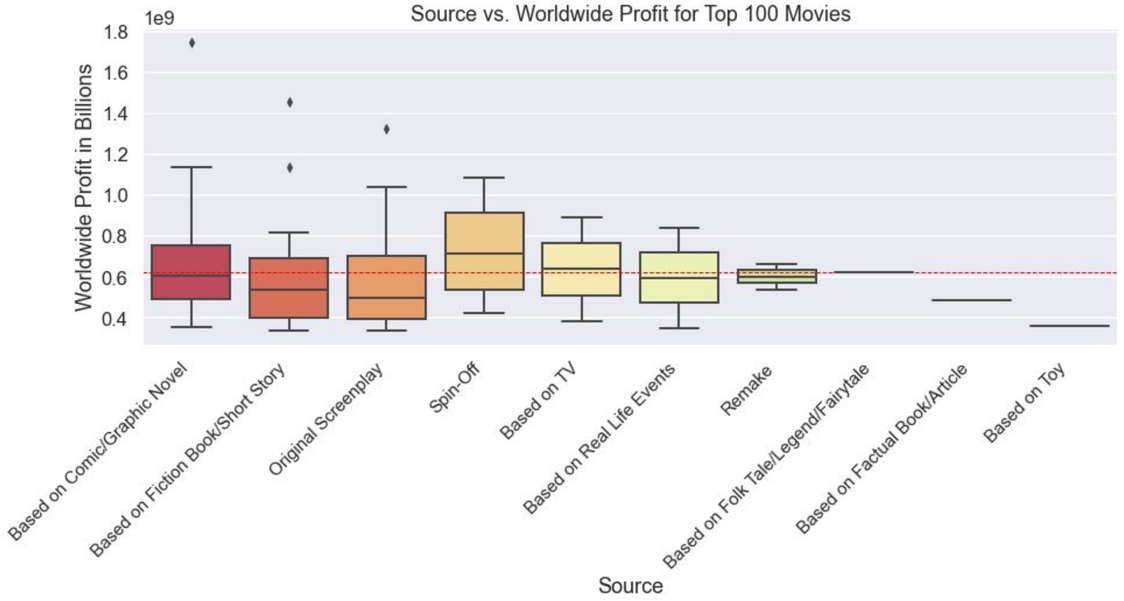
Genre

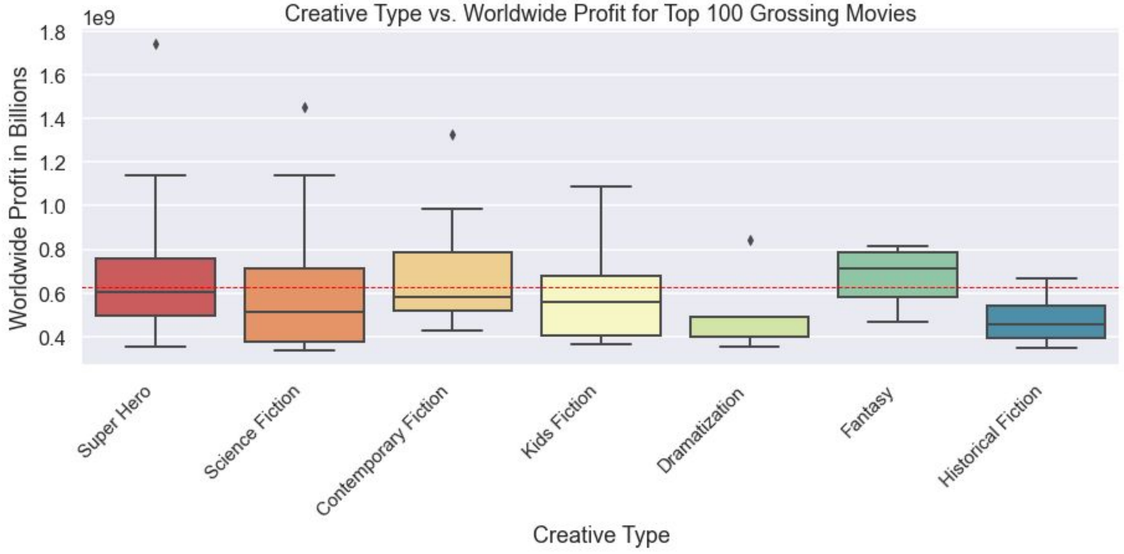
Source

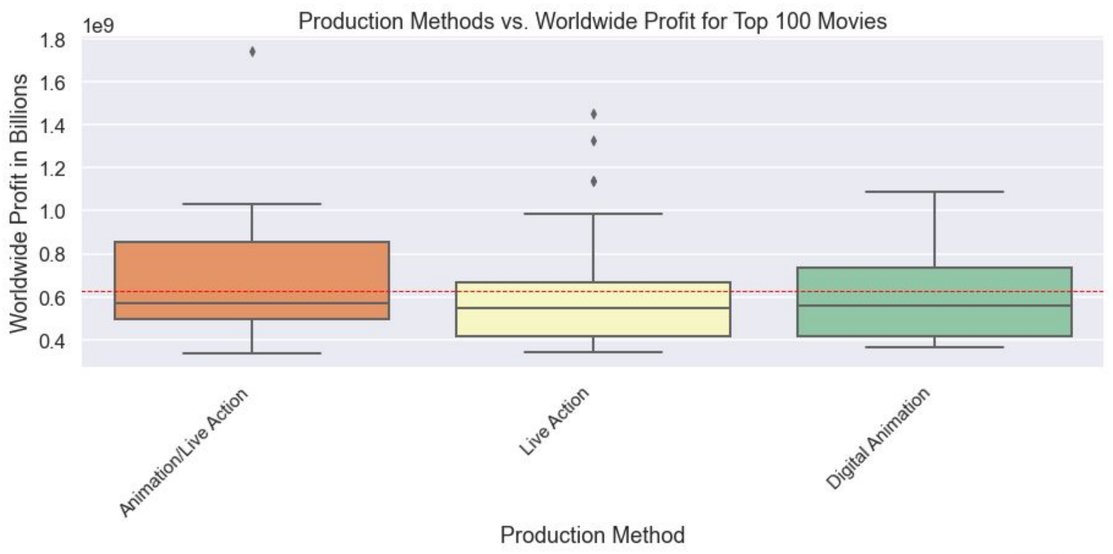
Creative Type

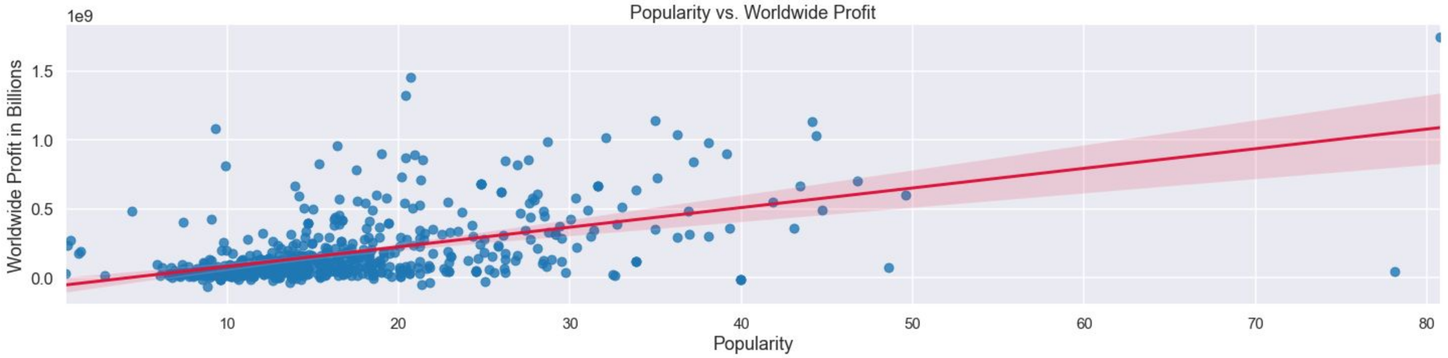
Production Method

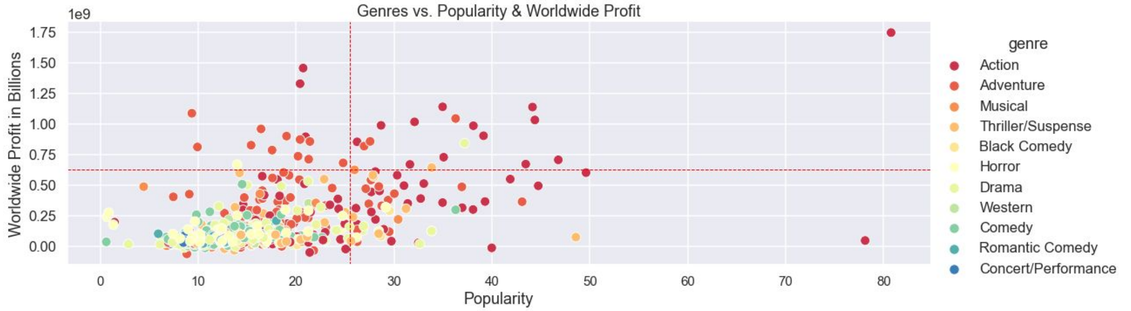

Genres vs. Wordwide Profit of Top 100 Movies
1e9
Worldwide Profit in Billions
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Action
Adventure
Drama
Horror
Thriller/Suspense
Musical
Comedy
Western
Genre

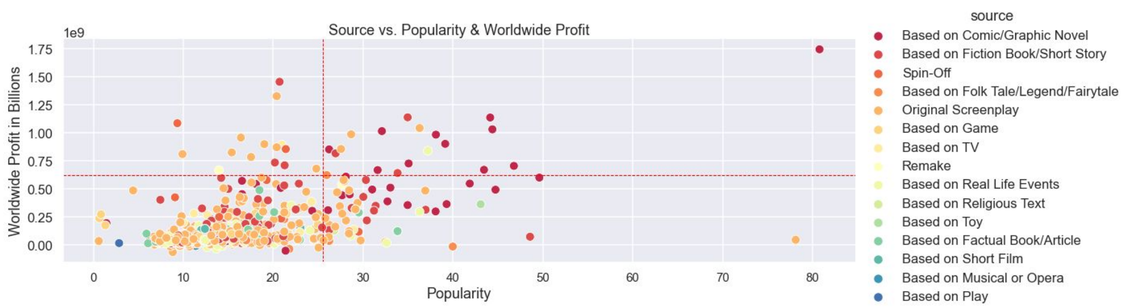

Source vs. Worldwide Profit for Top 100 Movies
1e9
Worldwide Profit in Billions
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Based on Comic/Graphic Novel
Based on Fiction Book/Short Story
Original Screenplay
Spin-Off
Based on TV
Based on Real Life Events
Remake
Based on Folk Tale/Legend/Fairytale
Based on Factual Book/Article
Based on Toy
Source

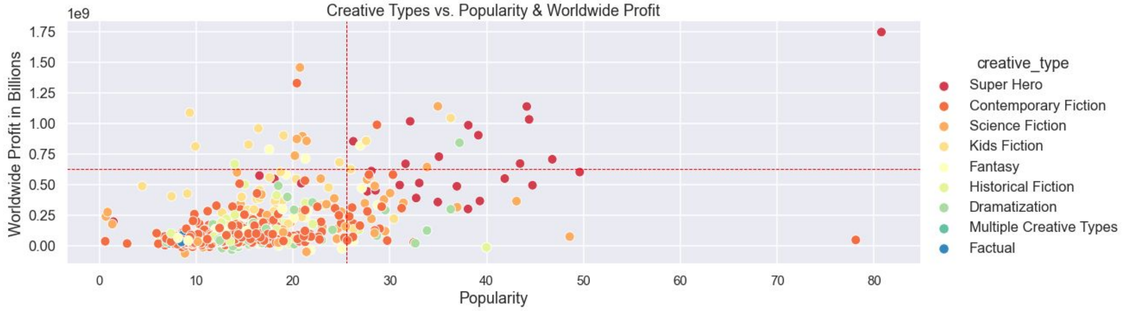

Creative Type vs. Worldwide Profit for Top 100 Grossing Movies
1e9
Worldwide Profit in Billions
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Super Hero
Science Fiction
Contemporary Fiction
Kids Fiction
Dramatization
Fantasy
Historical Fiction
Creative Type

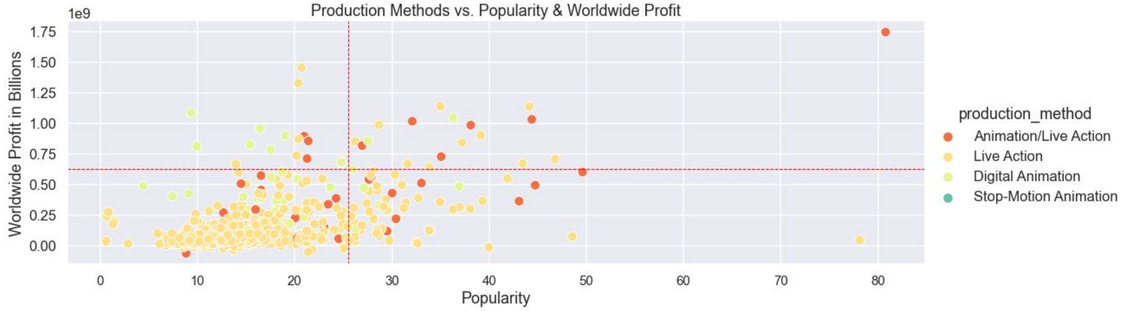

Production Methods vs. Worldwide Profit for Top 100 Movies
1e9
Worldwide Profit in Billions
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Animation/Live Action
Live Action
Digital Animation
Production Method


Popularity vs. Worldwide Profit
1e9
Worldwide Profit in Billions
Popularity


Genres vs. Popularity & Worldwide Profit
1e9
Worldwide Profit in Billions
Popularity
genre
Action
Adventure
Musical
Thriller/Suspense
Black Comedy
Horror
Drama
Western
Comedy
Romantic Comedy
Concert/Performance


Source vs. Popularity & Worldwide Profit
1e9
Worldwide Profit in Billions
Popularity
source
Based on Comic/Graphic Novel
Based on Fiction Book/Short Story
Spin-Off
Based on Folk Tale/Legend/Fairytale
Original Screenplay
Based on Game
Based on TV
Remake
Based on Real Life Events
Based on Religious Text
Based on Toy
Based on Factual Book/Article
Based on Short Film
Based on Musical or Opera
Based on Play


Creative Types vs. Popularity & Worldwide Profit
1e9
Worldwide Profit in Billions
Popularity
creative_type
Super Hero
Contemporary Fiction
Science Fiction
Kids Fiction
Fantasy
Historical Fiction
Dramatization
Multiple Creative Types
Factual


Production Methods vs. Popularity & Worldwide Profit
1e9
Worldwide Profit in Billions
Popularity
production_method
Animation/Live Action
Live Action
Digital Animation
Stop-Motion Animation

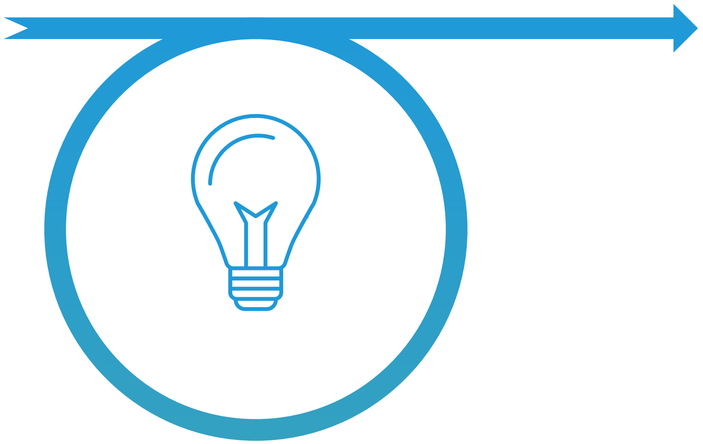
Conclusion

Action

Based on Comic Graphic/ Novel

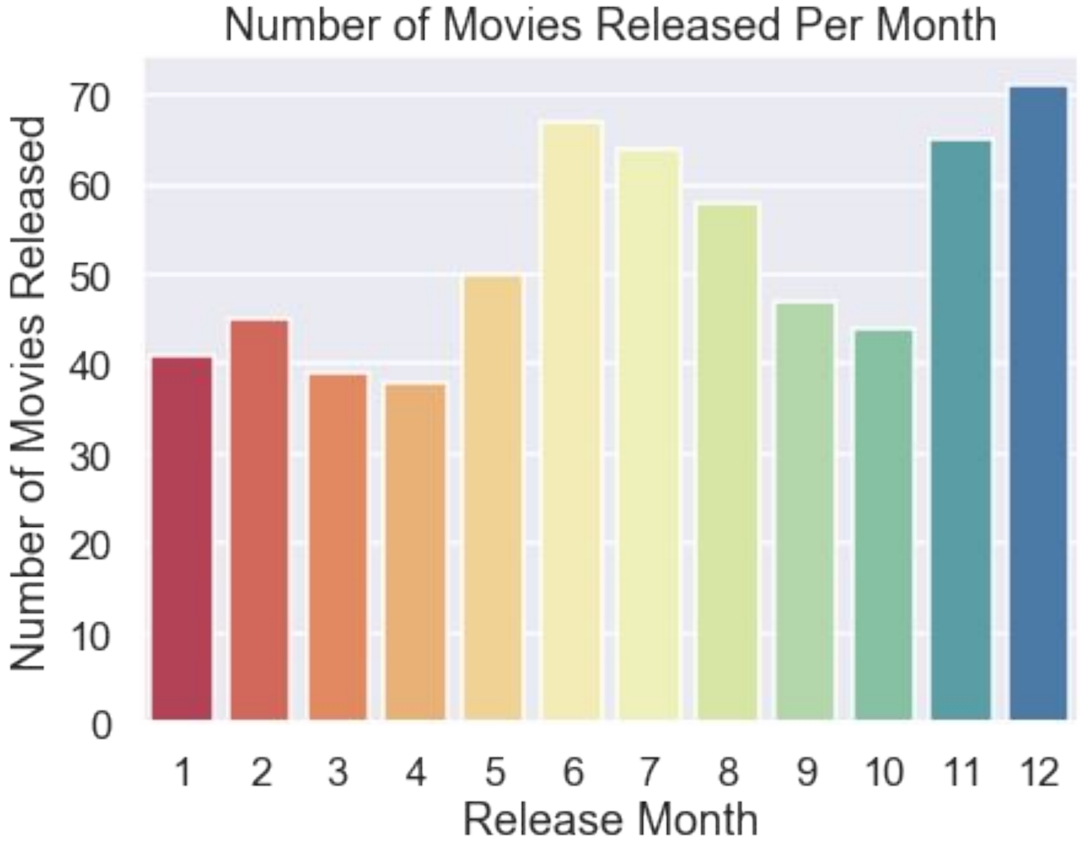
Super Hero

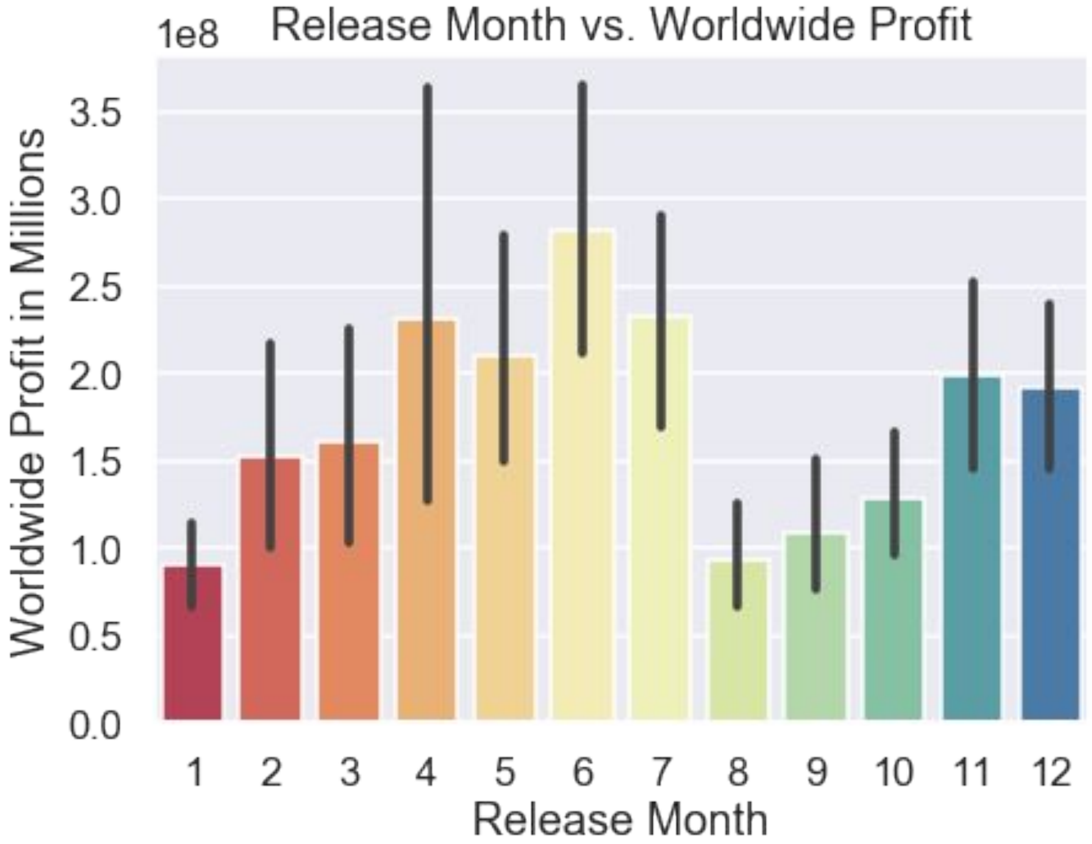
Animation/ Live Action



Business Solution

Release Month

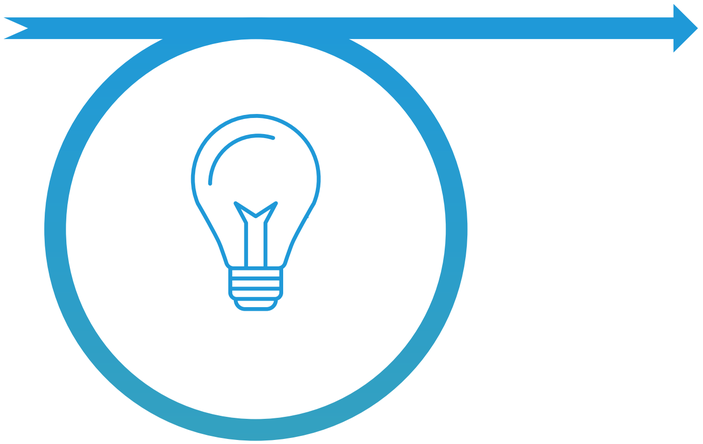




- Most release = June, December
- Most profit = June

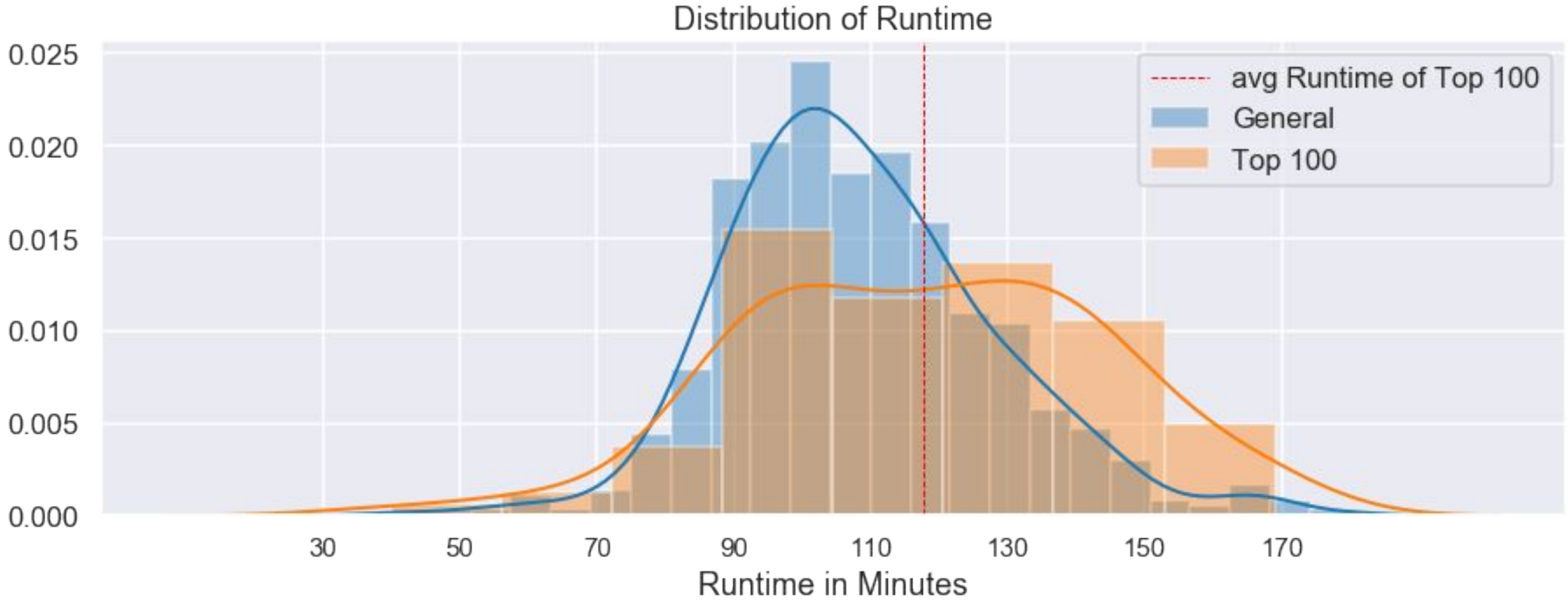
Conclusion

June



Business Solution

Runtime



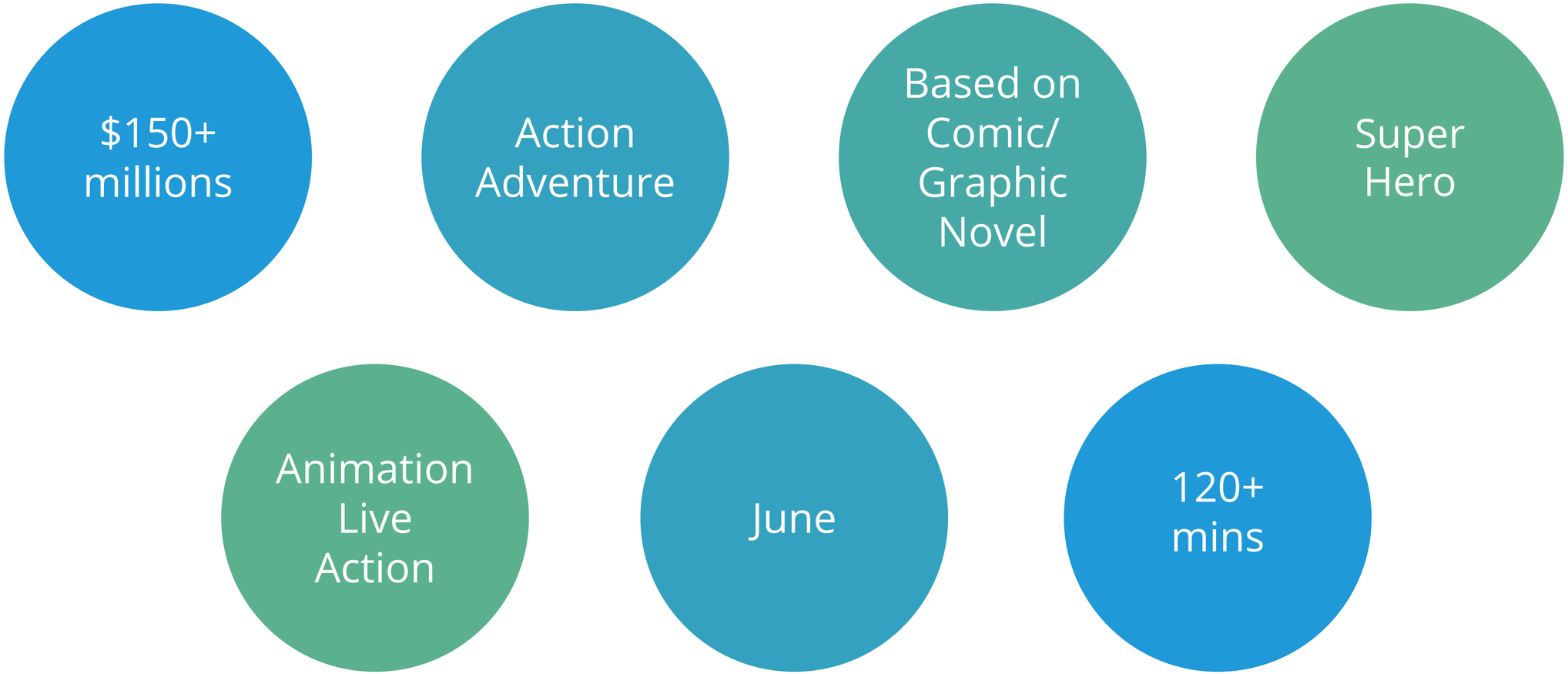
- The average runtime in General Movies is 108 minutes.
- The average runtime in Top 100 Movies is 118 minutes.

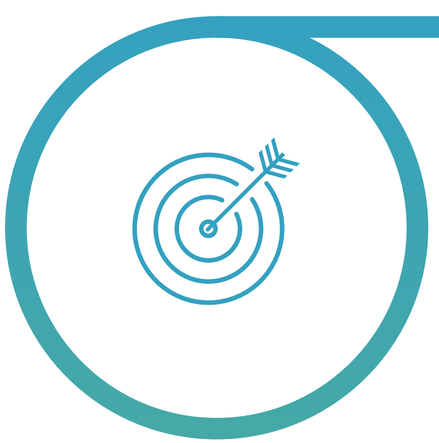
Conclusion

120+ MINS

Our **Recommendations**

What Makes a Movie Successful





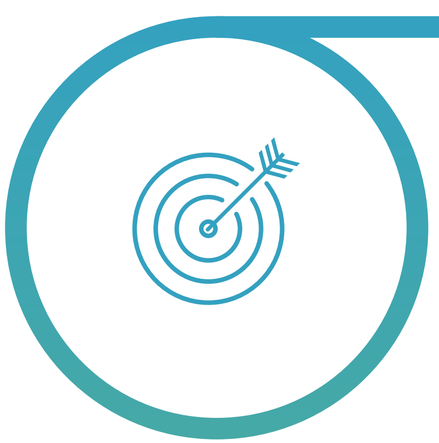
Future Works

- Do analysis in term of ROI
- Domestic market and international market
- MPAA rating
- Star quality
- Script quality
- Special effects, sound design, music
- Marketing campaign
- Popularity of the film preceding it (if it’s a sequel)
- Directors
- Writers
- Critics
- Diversity: female directors, women of color, Black Lives Matter movement, LGBTQ+, etc.

Thank You

Q & A

Appendix

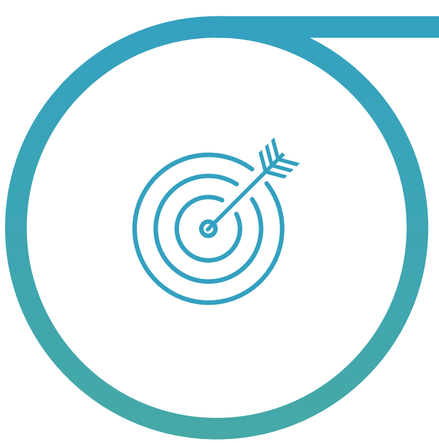


Examples

- ***Pirates of the Caribbean: On Stranger Tides*** (2011) - $379M budget, $1B earned

- ***Avenger: Age of Ultron*** (2015) - $365M budget, $1.3B earned

- ***Spectre*** (1997) - $300M budget, $879M earned

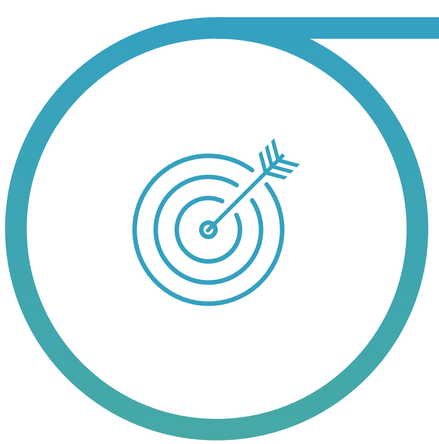


Examples

- ***Facing the Giants*** (2006) - $100,000 budget, $24.4M earned

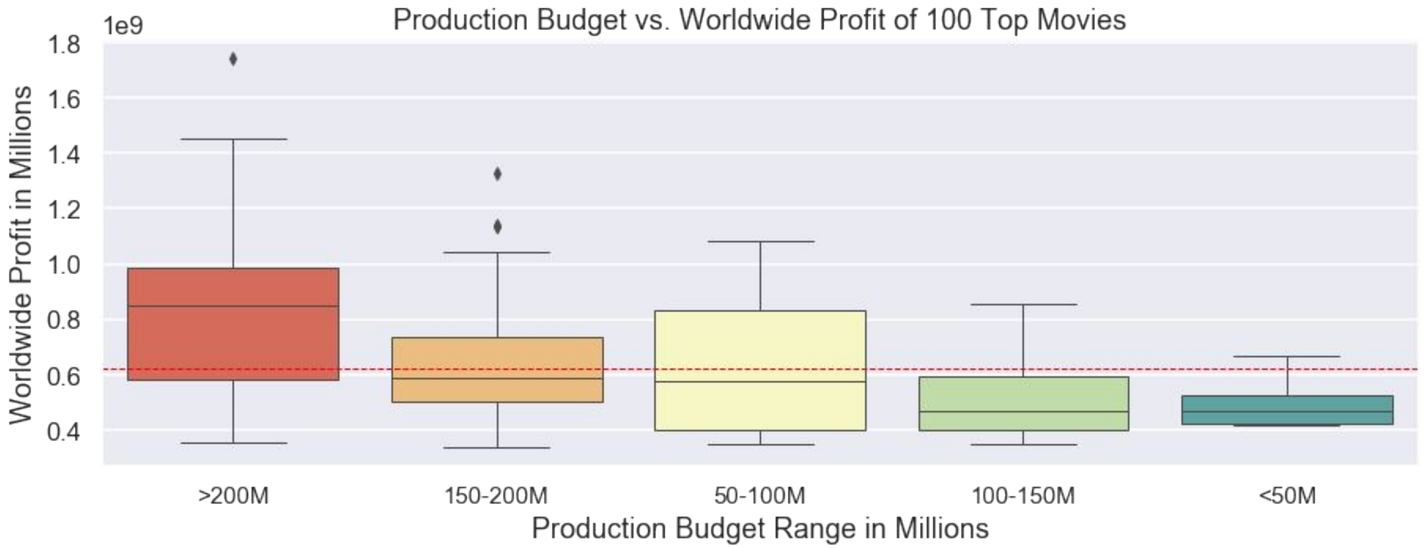
- ***Once*** (2007) - $150,000 budget, Oscar nominations, $19M earned

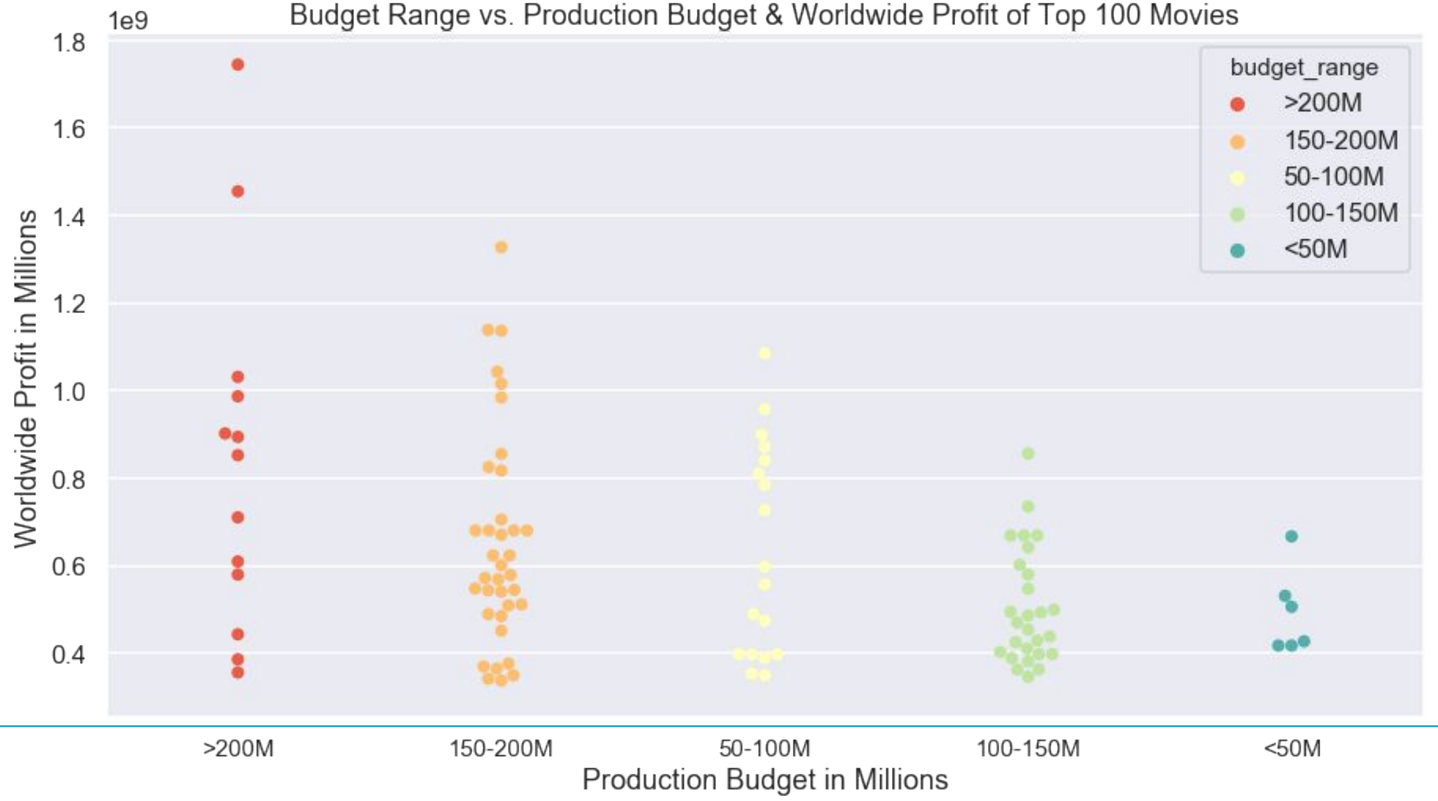
- ***Mad Max*** (1980), $200,000 budget, $49.6M earned

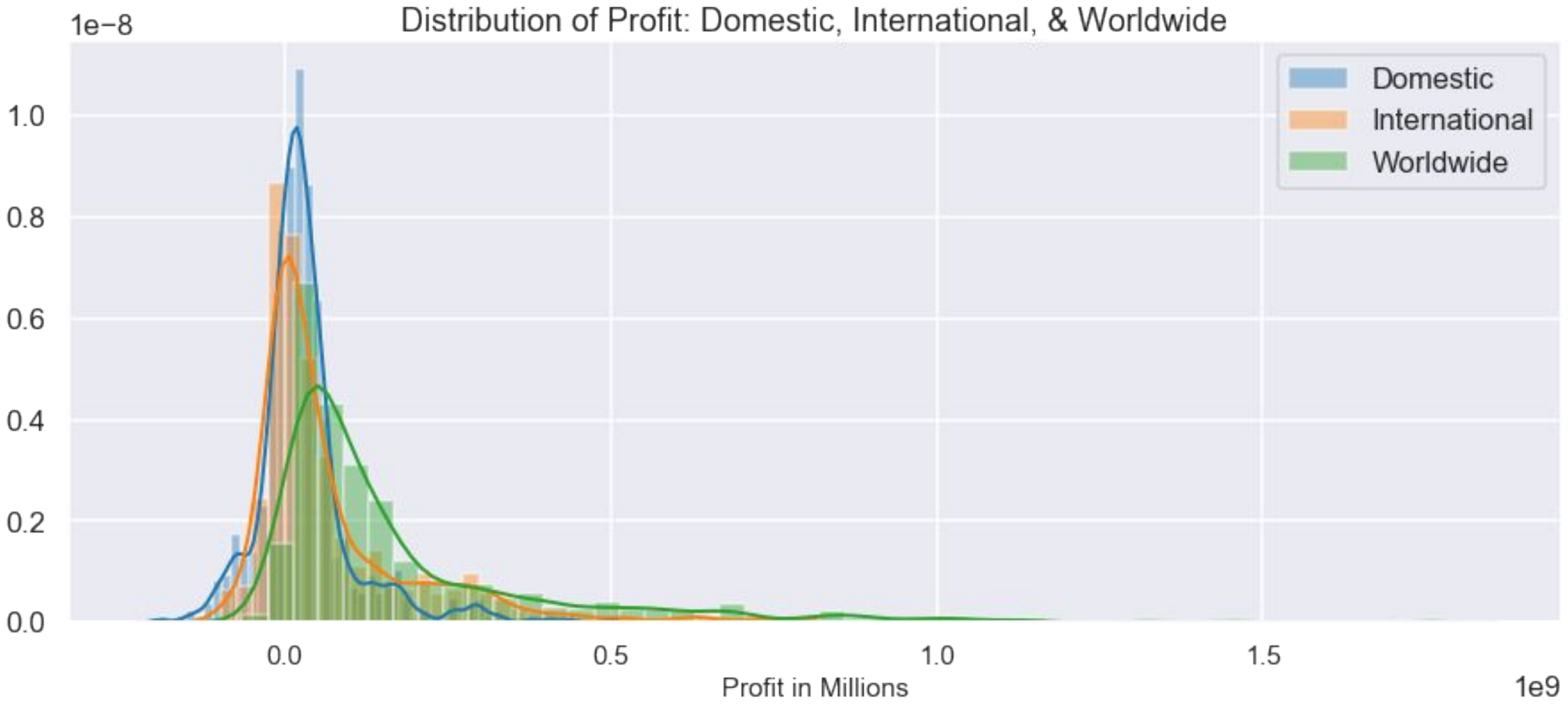


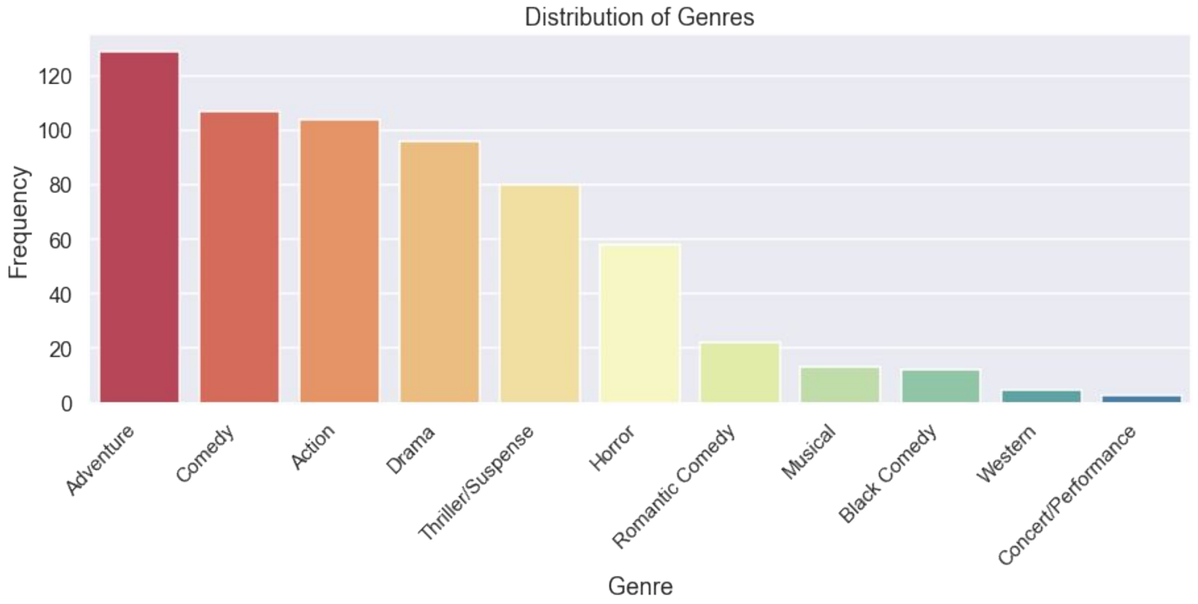
Notes

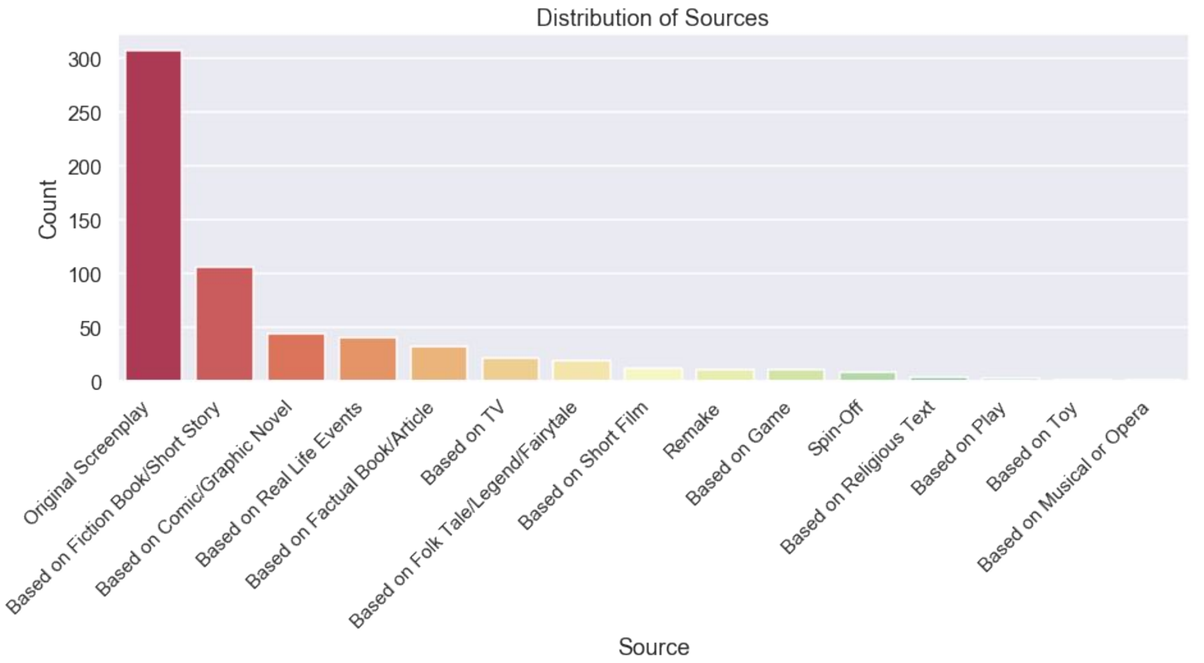
- **1e9** = in Billions
- **1e8** = in Hundred Millions

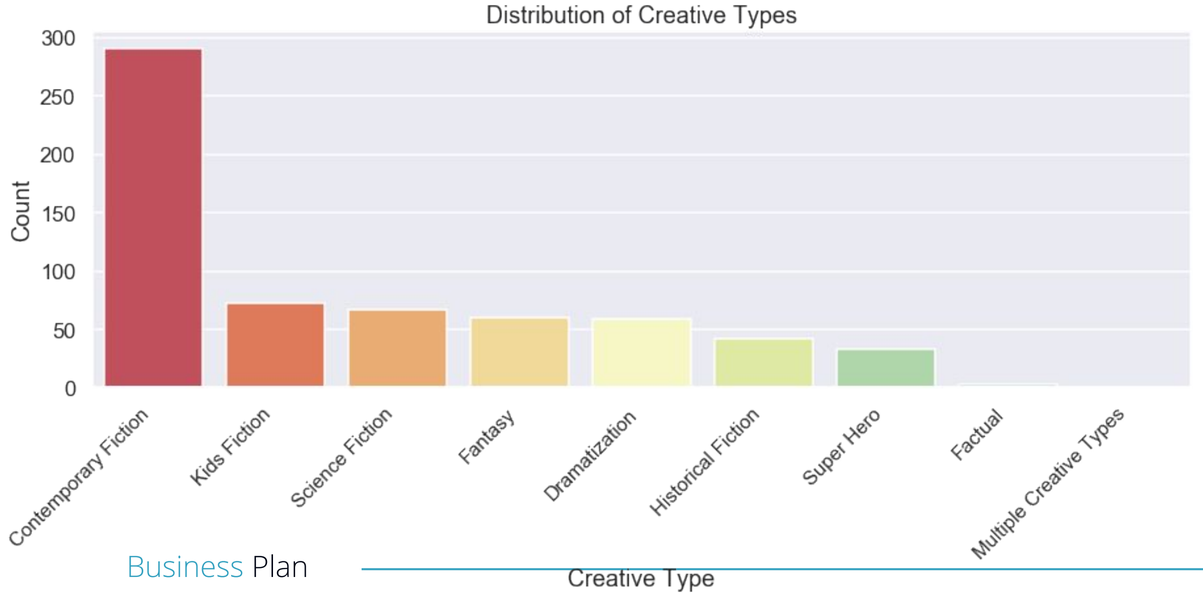

Production Budget vs. Worldwide Profit of 100 Top Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
>200M
150-200M
50-100M
100-150M
<50M
Production Budget Range in Millions

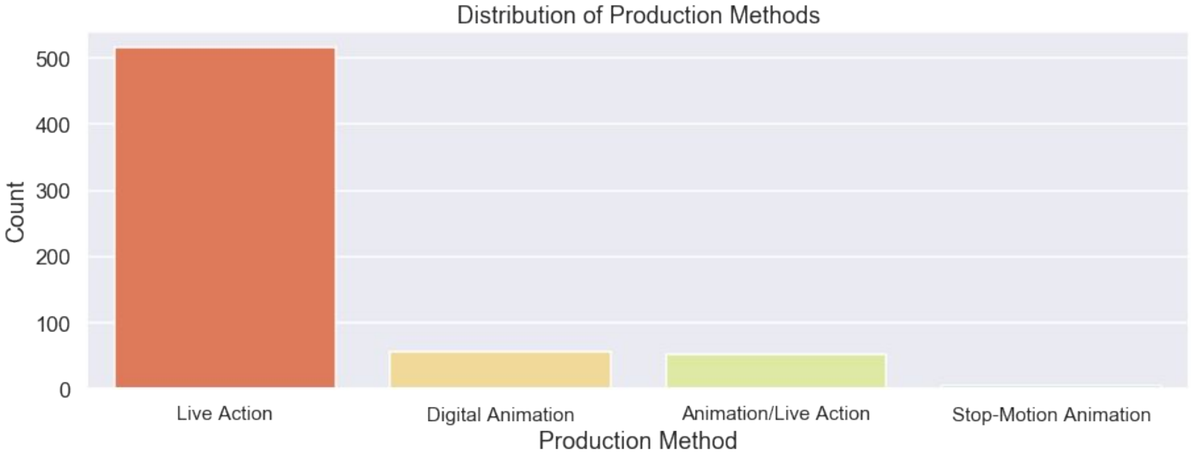

Budget Range vs. Production Budget & Worldwide Profit of Top 100 Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
budget_range
>200M
150-200M
50-100M
100-150M
<50M
>200M
150-200M
50-100M
100-150M
<50M
Production Budget in Millions

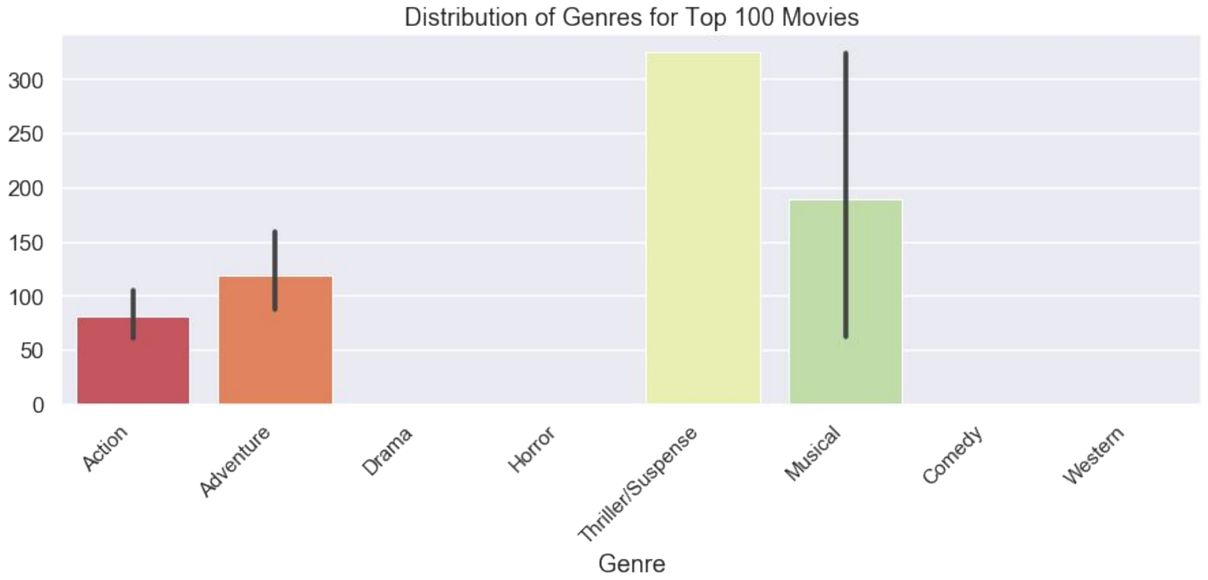

Distribution of Profit: Domestic, International, & Worldwide
1e−8
1.0
0.8
0.6
0.4
0.2
0.0
Domestic
International
Worldwide
0.0
0.5
1.0
1.5
Profit in Millions
1e9

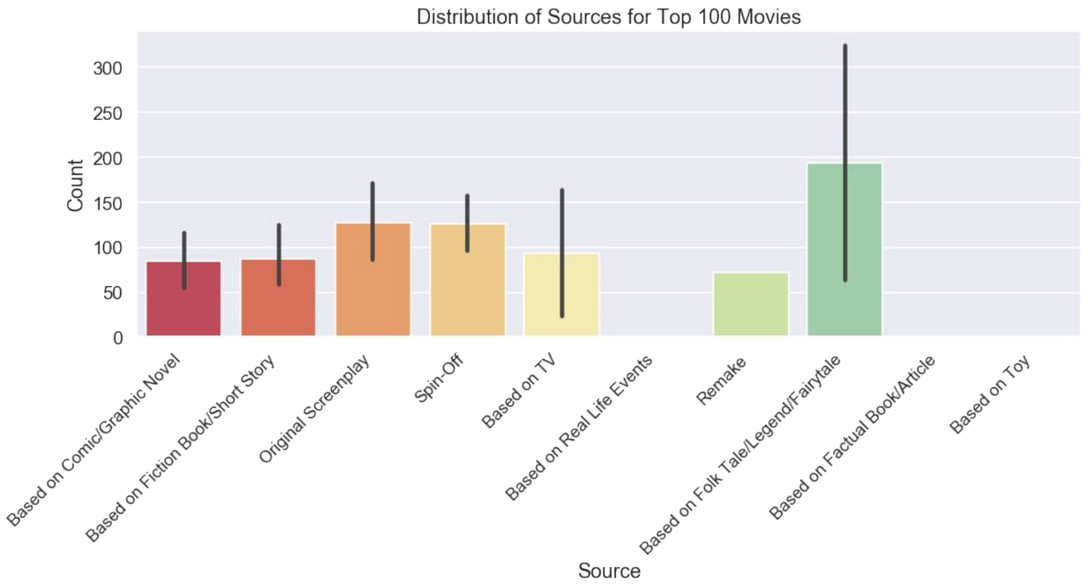

Distribution of Genres
Frequency
0
20
40
60
80
100
120
Adventure
Comedy
Action
Drama
Thriller/Suspense
Horror
Romantic Comedy
Musical
Black Comedy
Western
Concert/Performance
Genre

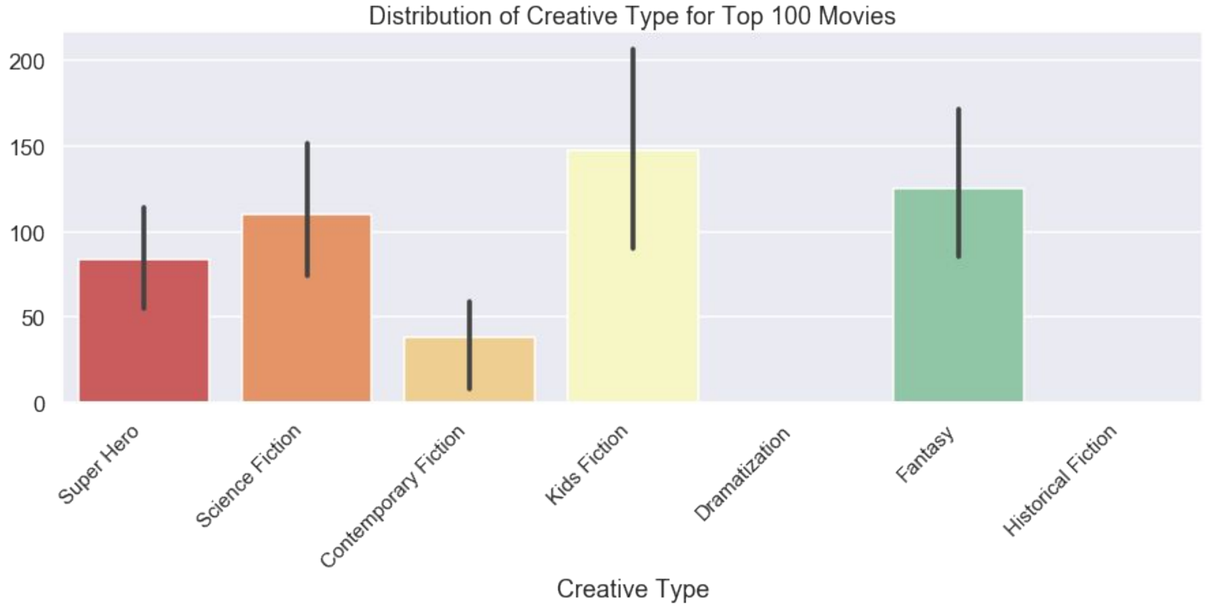

Distribution of Sources
Count
0
50
100
150
200
250
300
Original Screenplay
Based on Fiction Book/Short Story
Based on Comic/Graphic Novel
Based on Real Life Events
Based on Factual Book/Article
Based on TV
Based on Folk Tale/Legend/Fairytale
Based on Short Film
Remake
Based on Game
Spin-Off
Based on Religious Text
Based on Play
Based on Toy
Based on Musical or Opera
Source

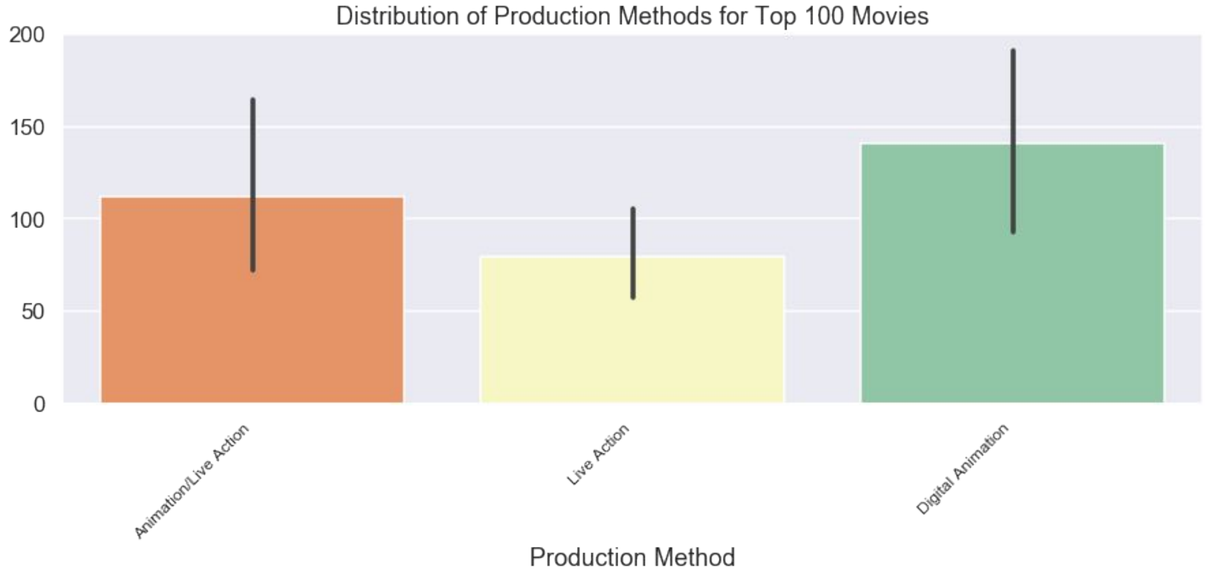

Distribution of Creative Types
Count
0
50
100
150
200
250
300
Contemporary Fiction
Kids Fiction
Science Fiction
Fantasy
Dramatization
Historical Fiction
Super Hero
Factual
Multiple Creative Types
Creative Type

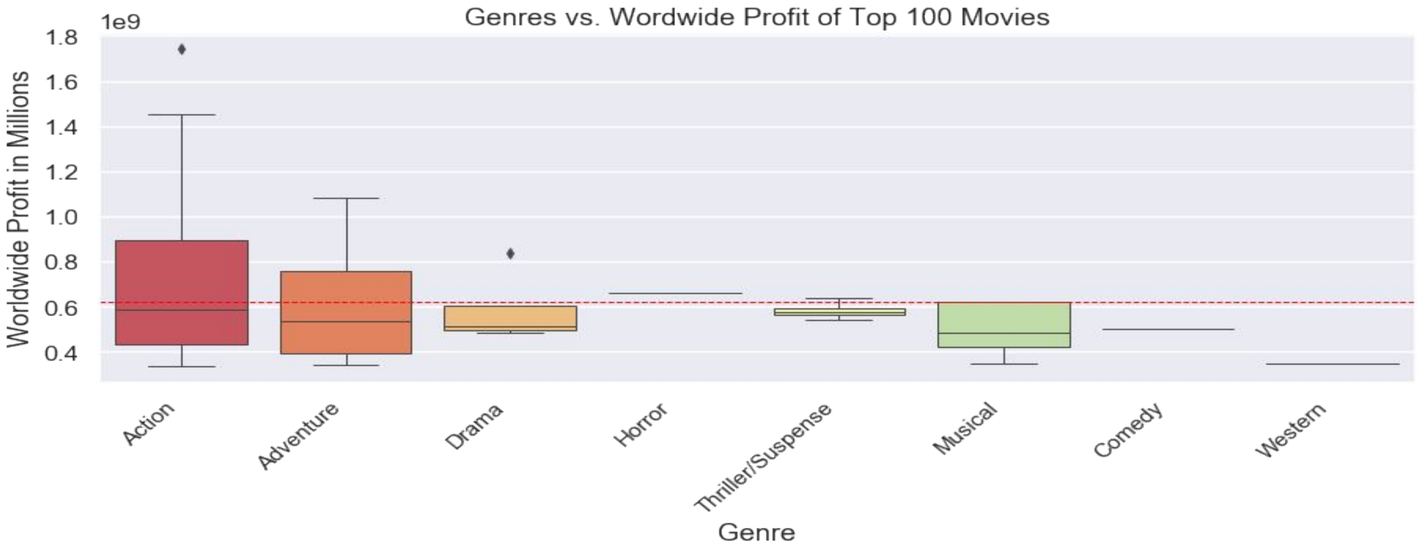

Distribution of Production Methods
Count
0
100
200
300
400
500
Live Action
Digital Animation
Animation/Live Action
Stop-Motion Animation
Production Method

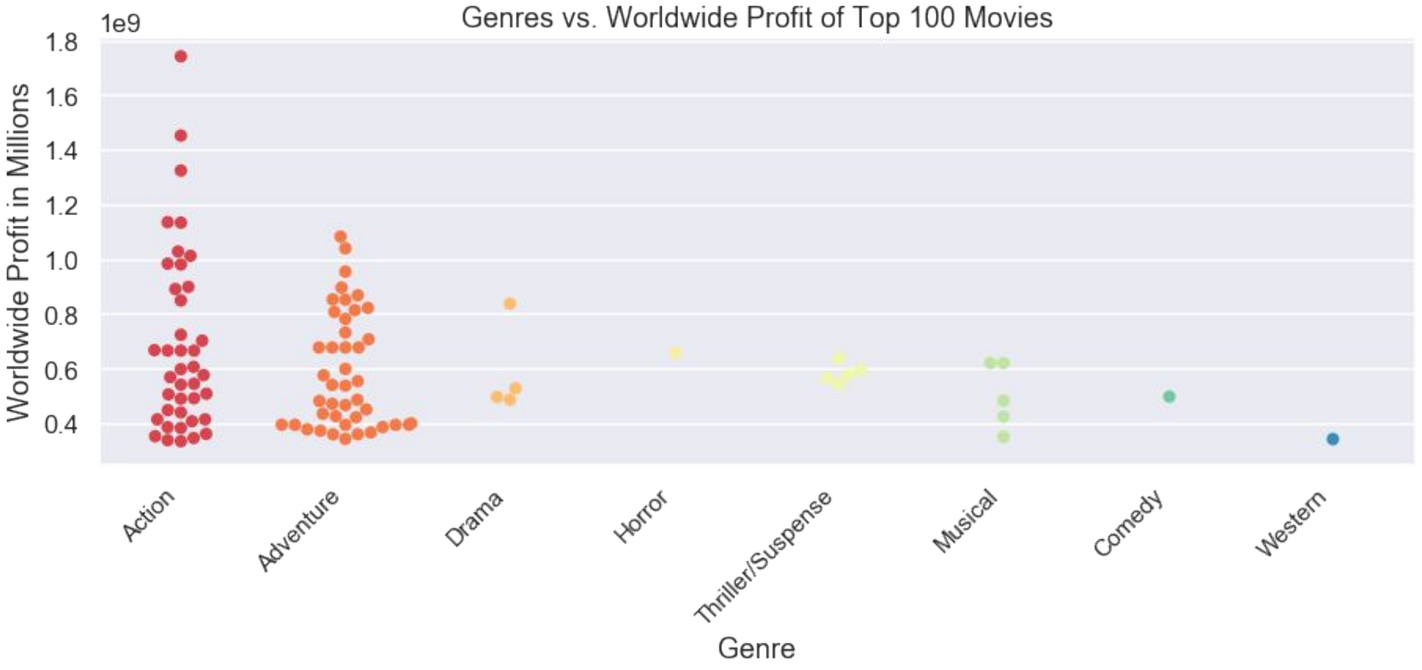

Distribution of Genres for Top 100 Movies
300
250
200
150
100
50
0
Action
Adventure
Drama
Horror
Thriller/Suspense
Musical
Comedy
Western
Genre

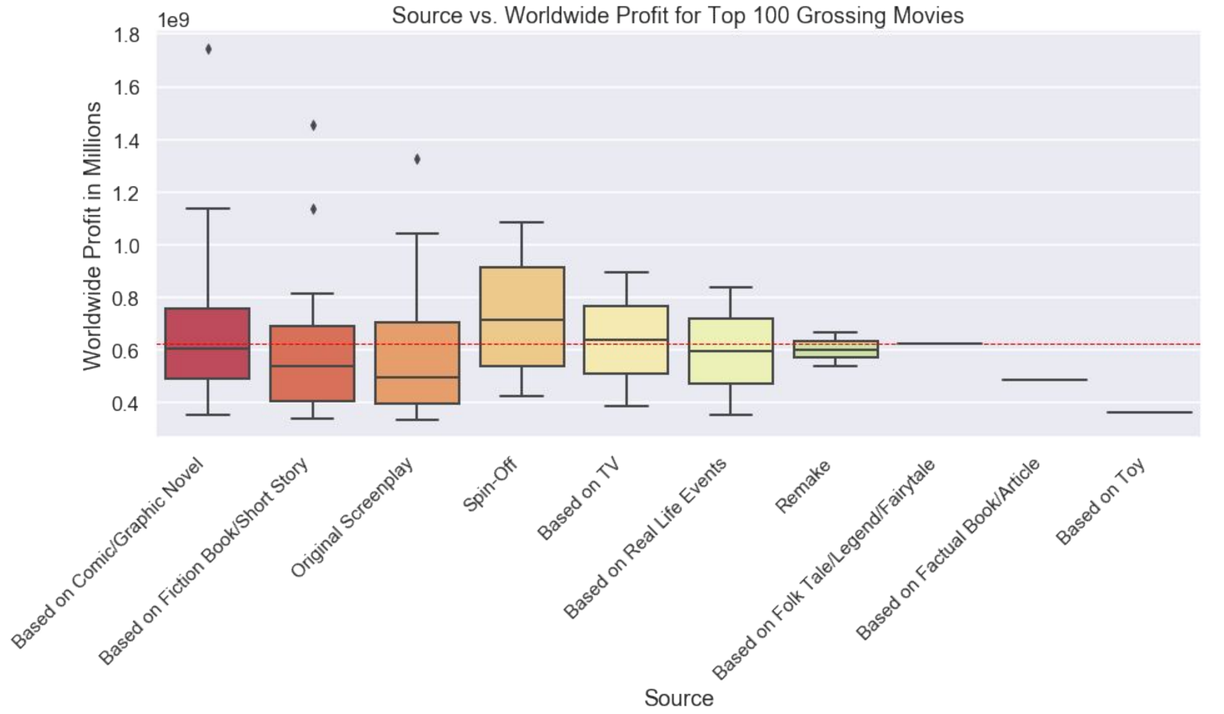

Distribution of Sources for Top 100 Movies
Count
300
250
200
150
100
50
0
Based on Comic/Graphic Novel
Based on Fiction Book/Short Story
Original Screenplay
Spin-Off
Based on TV
Based on Real Life Events
Remake
Based on Folk Tale/Legend/Fairytale
Based on Factual Book/Article
Based on Toy
Source

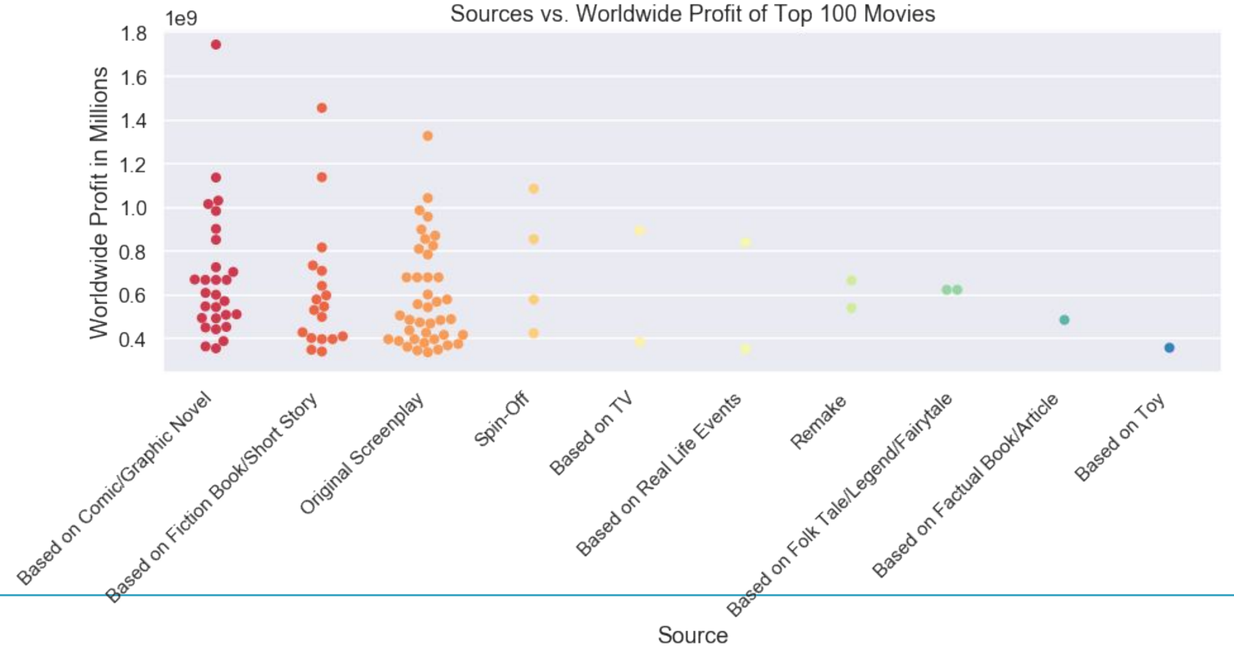

Distribution of Creative Type for Top 100 Movies
200
150
100
50
0
Super Hero
Science Fiction
Contemporary Fiction
Kids Fiction
Dramatization
Fantasy
Historical Fiction
Creative Type

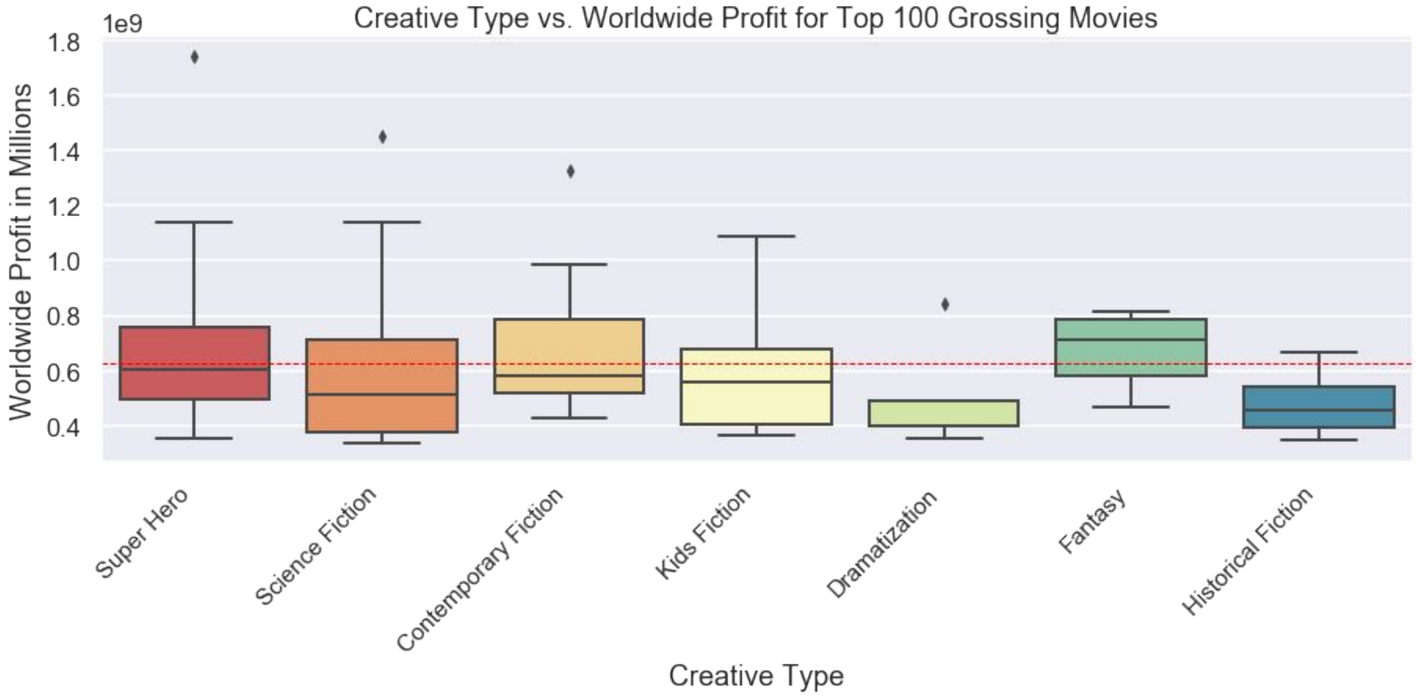

Distribution of Production Methods for Top 100 Movies
200
150
100
50
0
Animation/Live Action
Live Action
Digital Animation
Production Method

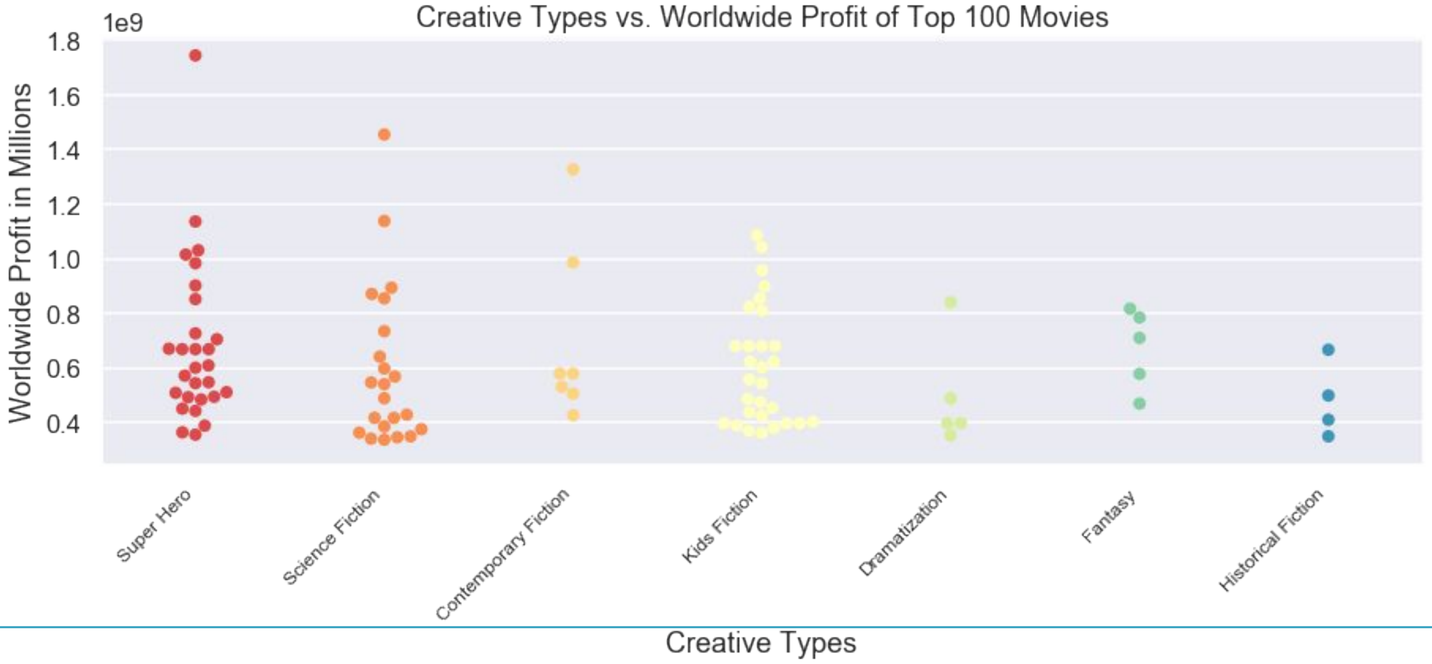

Genres vs. Wordwide Profit of Top 100 Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
Action
Adventure
Drama
Horror
Thriller/Suspense
Musical
Comedy
Western
Genre

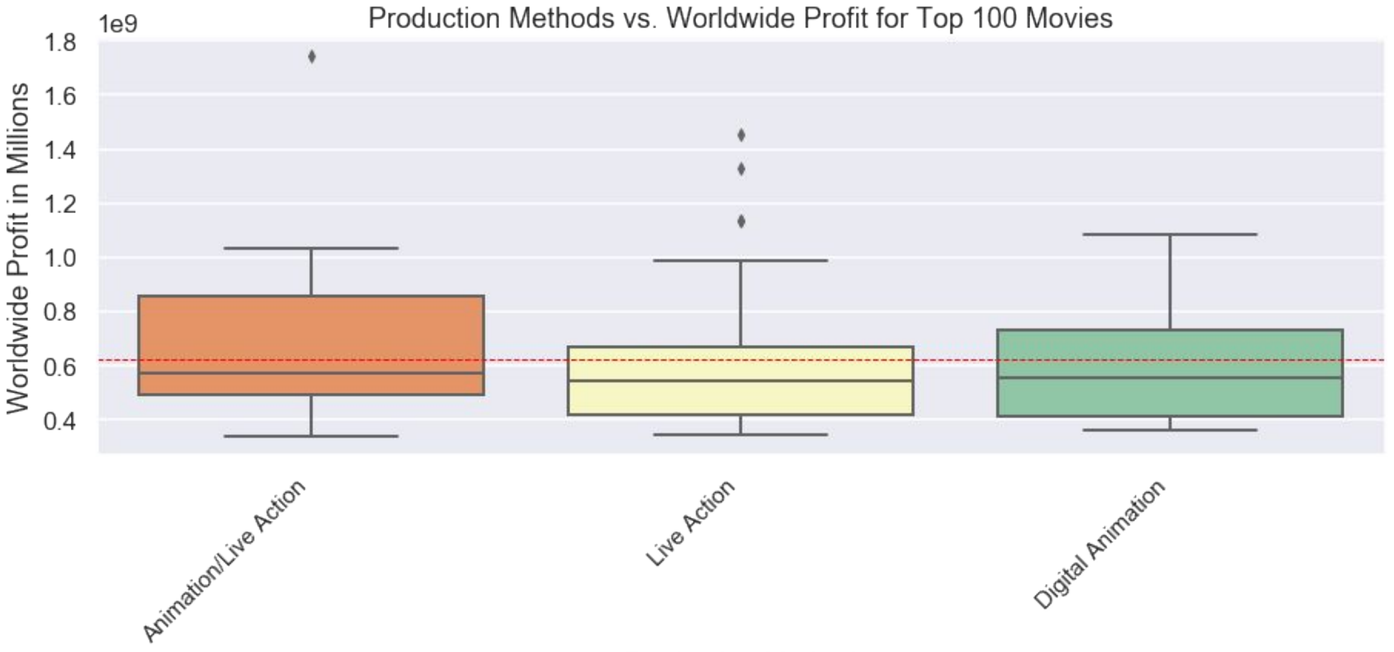

Genres vs. Worldwide Profit of Top 100 Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
Action
Adventure
Drama
Horror
Thriller/Suspense
Musical
Comedy
Western
Genre

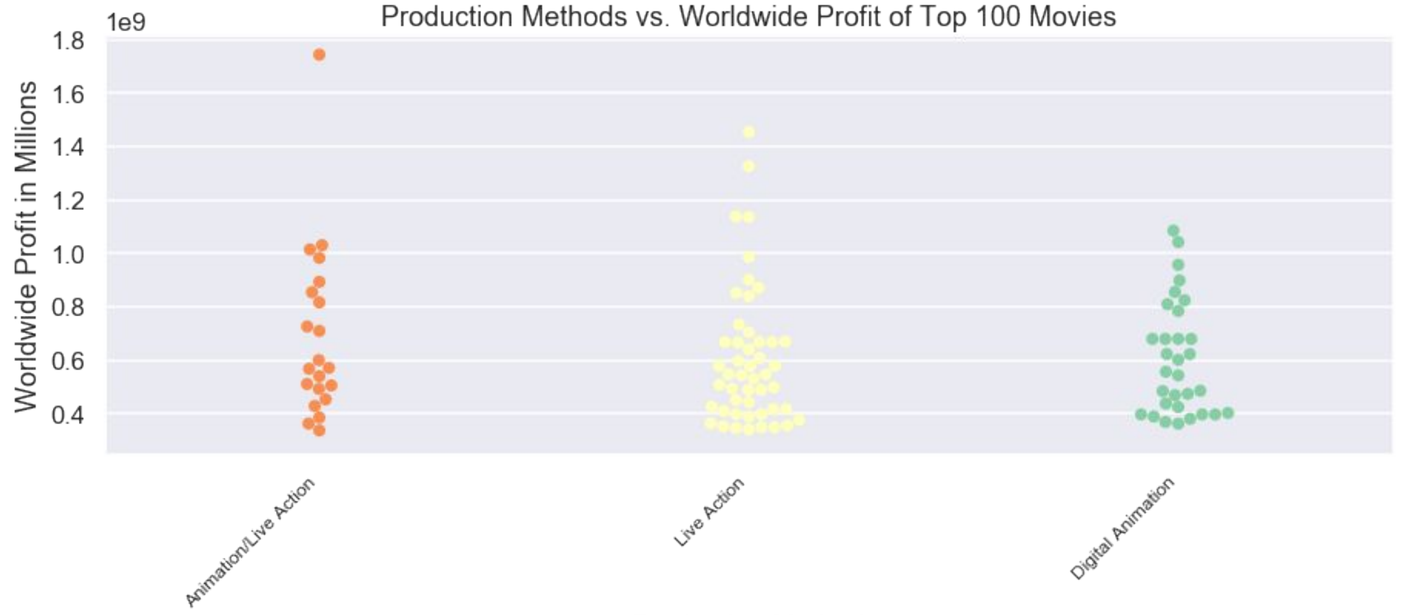

Source vs. Worldwide Profit for Top 100 Grossing Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
Based on Comic/Graphic Novel
Based on Fiction Book/Short Story
Original Screenplay
Spin-Off
Based on TV
Based on Real Life Events
Remake
Based on Folk Tale/Legend/Fairytale
Based on Factual Book/Article
Based on Toy
Source

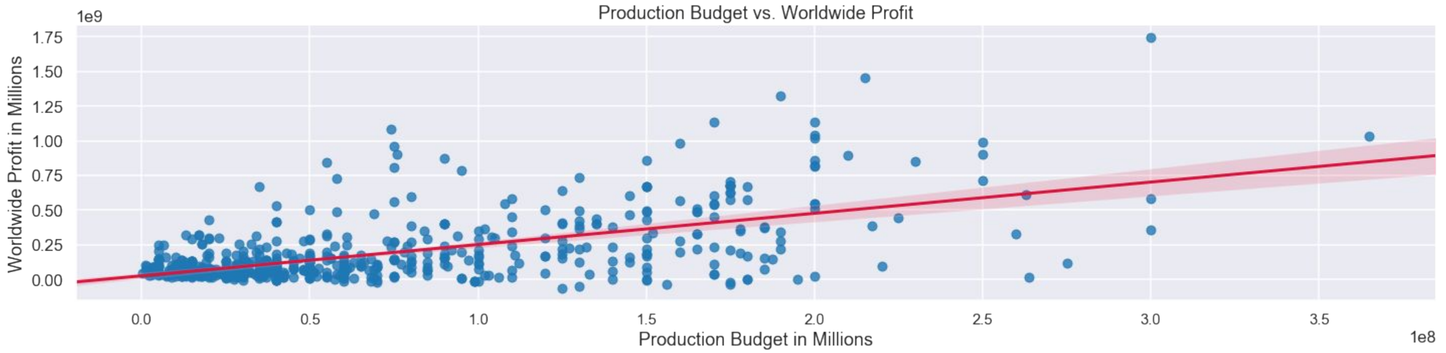

Sources vs. Worldwide Profit of Top 100 Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
Based on Comic/Graphic Novel
Based on Fiction Book/Short Story
Original Screenplay
Spin-Off
Based on TV
Based on Real Life Events
Remake
Based on Folk Tale/Legend/Fairytale
Based on Factual Book/Article
Based on Toy
Source

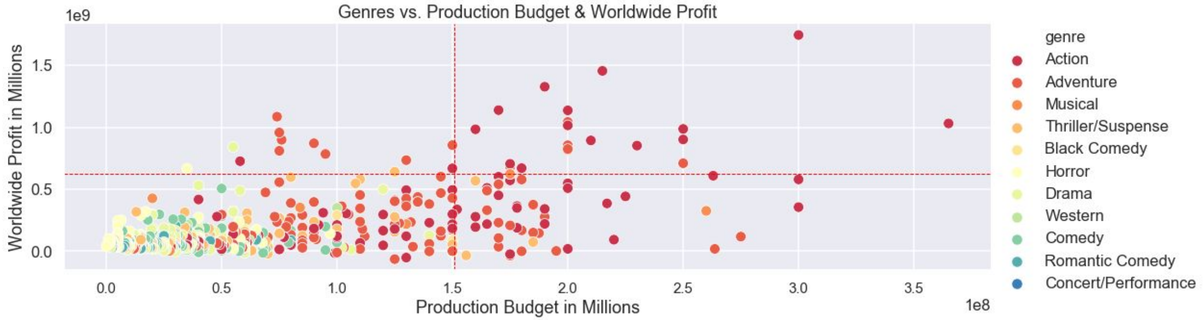

Creative Type vs. Worldwide Profit for Top 100 Grossing Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
Super Hero
Science Fiction
Contemporary Fiction
Kids Fiction
Dramatization
Fantasy
Historical Fiction
Creative Type

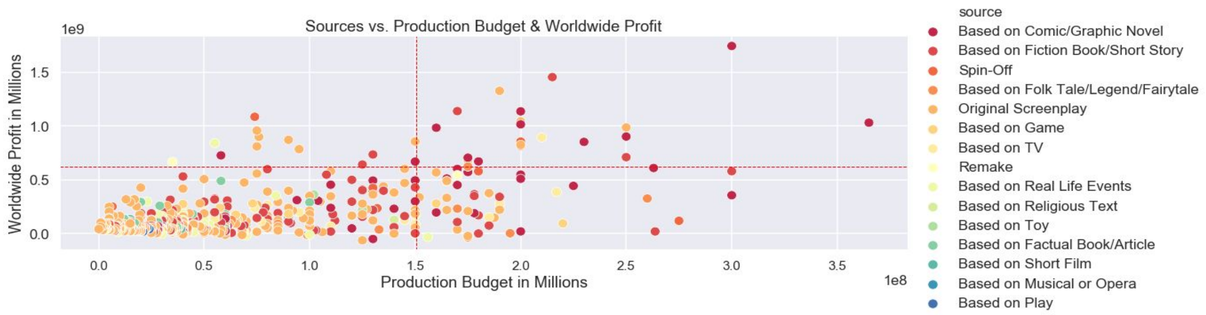

Creative Types vs. Worldwide Profit of Top 100 Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
Super Hero
Science Fiction
Contemporary Fiction
Kids Fiction
Dramatization
Fantasy
Historical Fiction
Creative Types

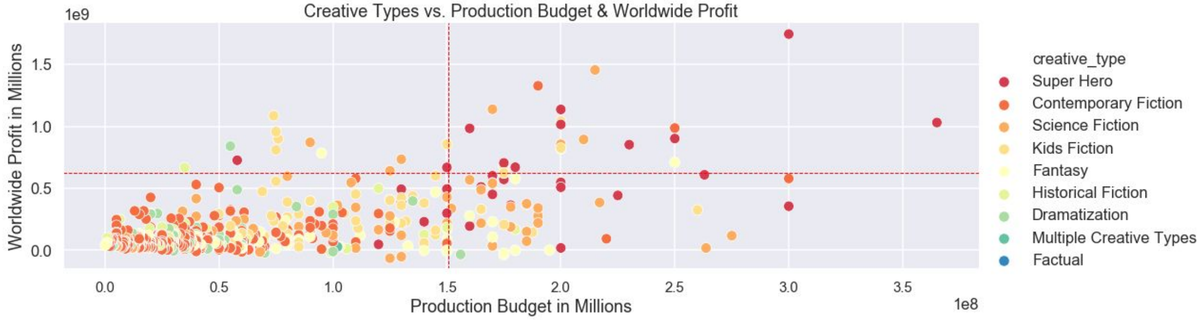

Production Methods vs. Worldwide Profit for Top 100 Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
Animation/Live Action
Live Action
Digital Animation
Production Method

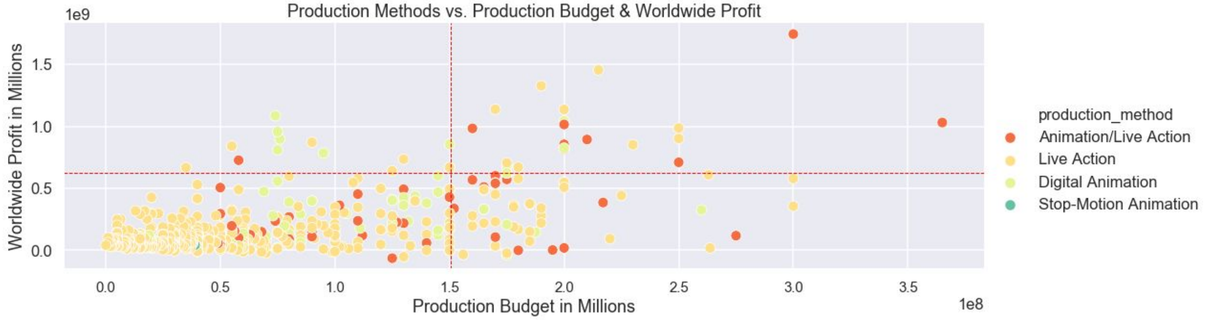

Production Methods vs. Worldwide Profit of Top 100 Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
Animation/Live Action
Live Action
Digital Animation
Production Method

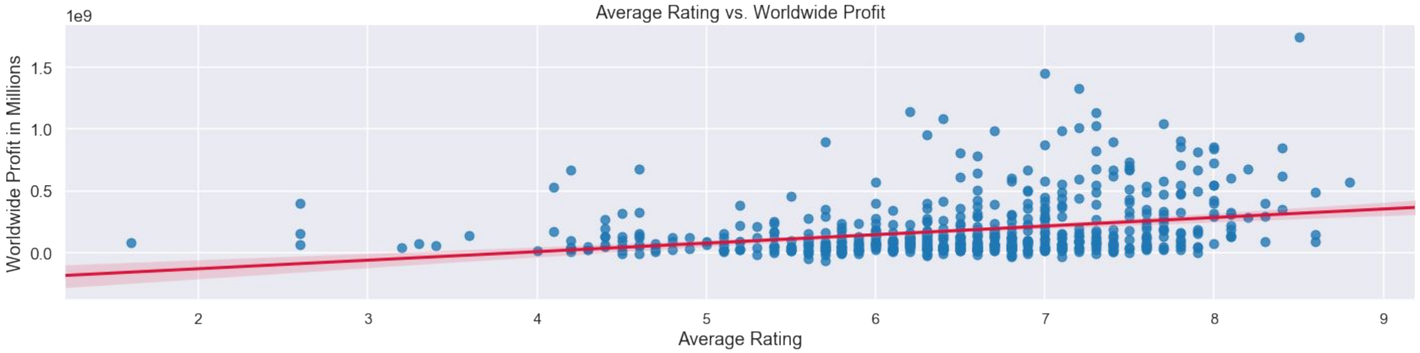

Production Budget vs. Worldwide Profit
1e9
Worldwide Profit in Millions
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Production Budget in Millions
1e8

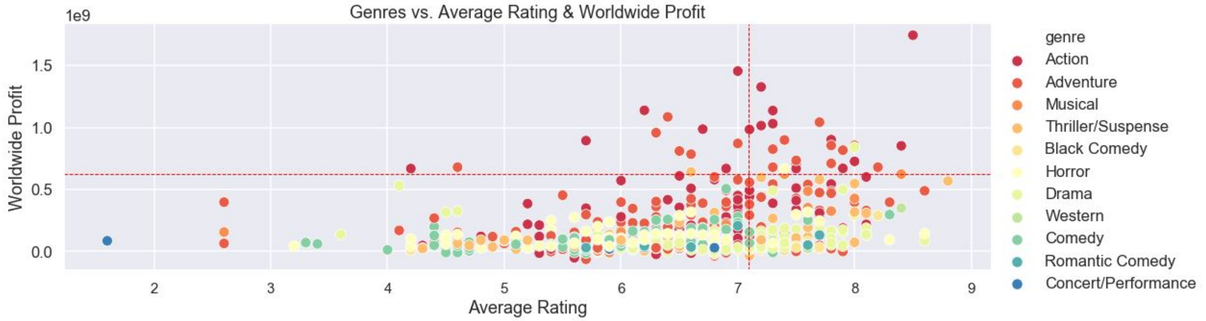

Genres vs. Production Budget & Worldwide Profit
1e9
Worldwide Profit in Millions
1.5
1.0
0.5
0.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Production Budget in Millions
1e8
genre
Action
Adventure
Musical
Thriller/Suspense
Black Comedy
Horror
Drama
Western
Comedy
Romantic Comedy
Concert/Performance

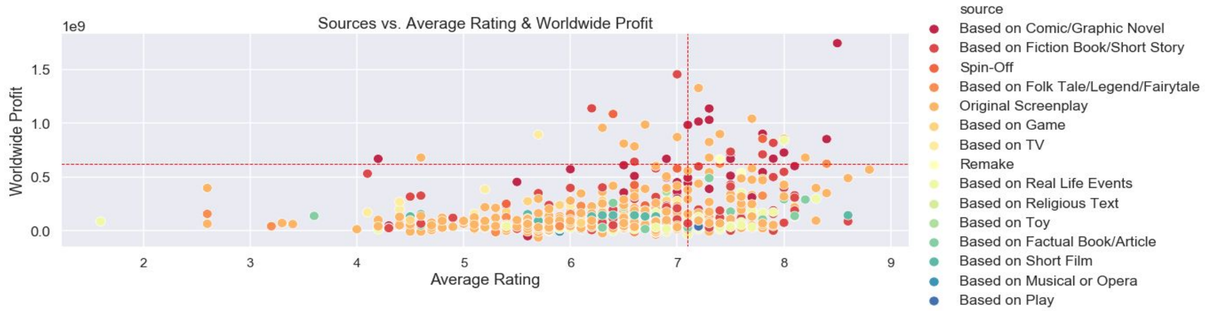

Sources vs. Production Budget & Worldwide Profit
1e9
Worldwide Profit in Millions
1.5
1.0
0.5
0.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Production Budget in Millions
1e8
source
Based on Comic/Graphic Novel
Based on Fiction Book/Short Story
Spin-Off
Based on Folk Tale/Legend/Fairytale
Original Screenplay
Based on Game
Based on TV
Remake
Based on Real Life Events
Based on Religious Text
Based on Toy
Based on Factual Book/Article
Based on Short Film
Based on Musical or Opera
Based on Play


Creative Types vs. Production Budget & Worldwide Profit
1e9
Worldwide Profit in Millions
1.5
1.0
0.5
0.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Production Budget in Millions
1e8
creative_type
Super Hero
Contemporary Fiction
Science Fiction
Kids Fiction
Fantasy
Historical Fiction
Dramatization
Multiple Creative Types
Factual

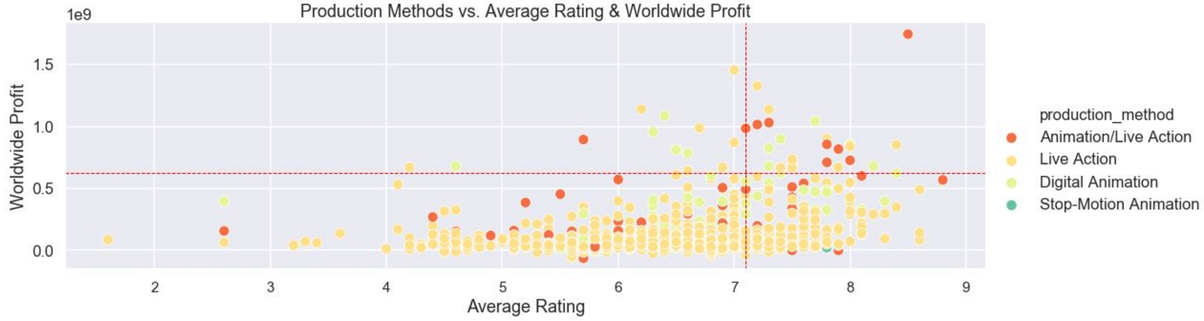

Production Methods vs. Production Budget & Worldwide Profit
1e9
Worldwide Profit in Millions
1.5
1.0
0.5
0.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Production Budget in Millions
1e8
production_method
Animation/Live Action
Live Action
Digital Animation
Stop-Motion Animation

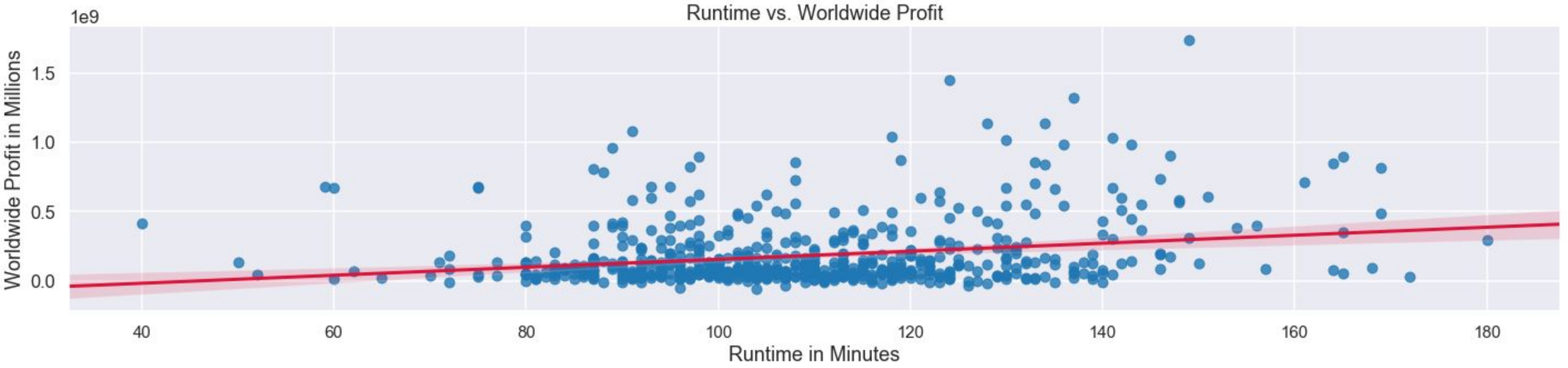

Average Rating vs. Worldwide Profit
1e9
Worldwide Profit in Millions
1.5
1.0
0.5
0.0
2
3
4
5
6
7
8
9
Average Rating

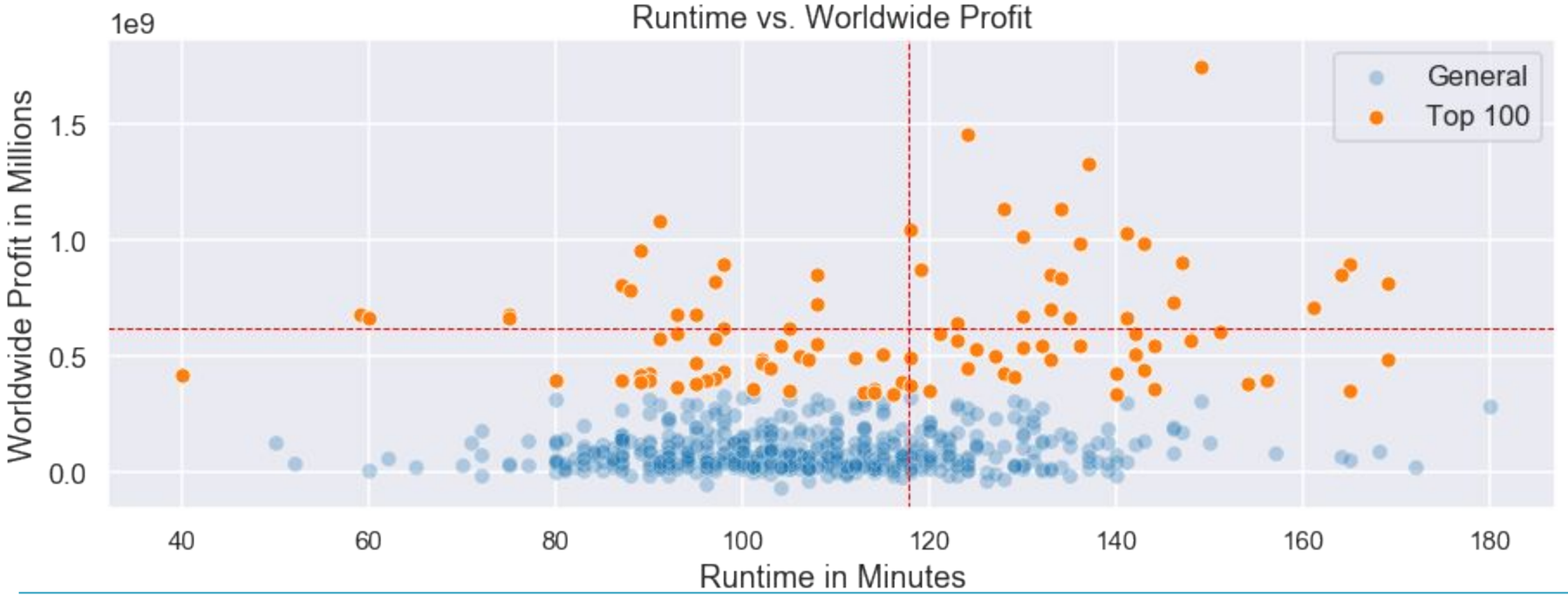

Genres vs. Average Rating & Worldwide Profit
1e9
Worldwide Profit
1.5
1.0
0.5
0.0
2
3
4
5
6
7
8
9
Average Rating
genre
Action
Adventure
Musical
Thriller/Suspense
Black Comedy
Horror
Drama
Western
Comedy
Romantic Comedy
Concert/Performance

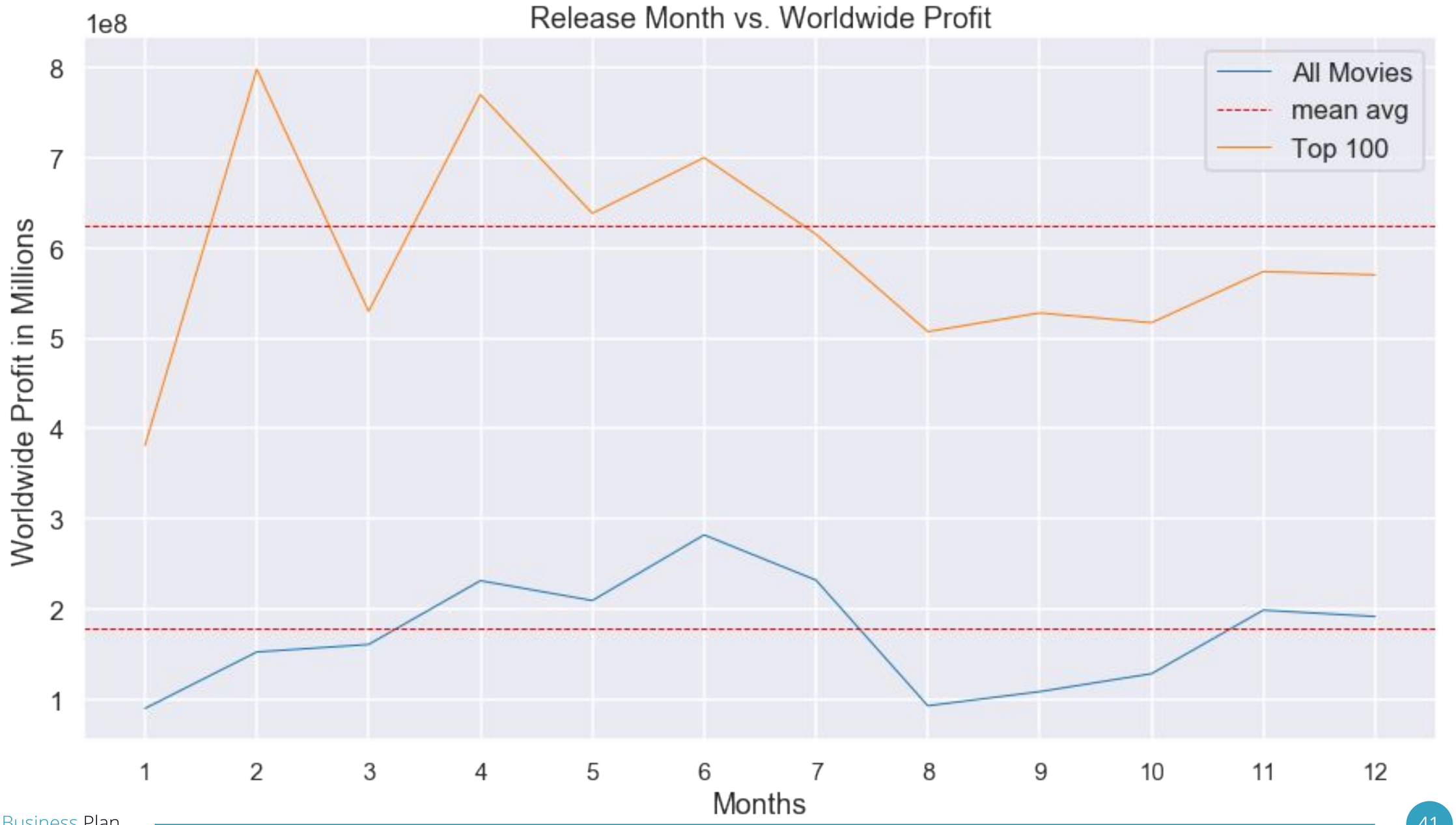

Sources vs. Average Rating & Worldwide Profit
1e9
Worldwide Profit
1.5
1.0
0.5
0.0
2
3
4
5
6
7
8
9
Average Rating
source
Based on Comic/Graphic Novel
Based on Fiction Book/Short Story
Spin-Off
Based on Folk Tale/Legend/Fairytale
Original Screenplay
Based on Game
Based on TV
Remake
Based on Real Life Events
Based on Religious Text
Based on Toy
Based on Factual Book/Article
Based on Short Film
Based on Musical or Opera
Based on Play

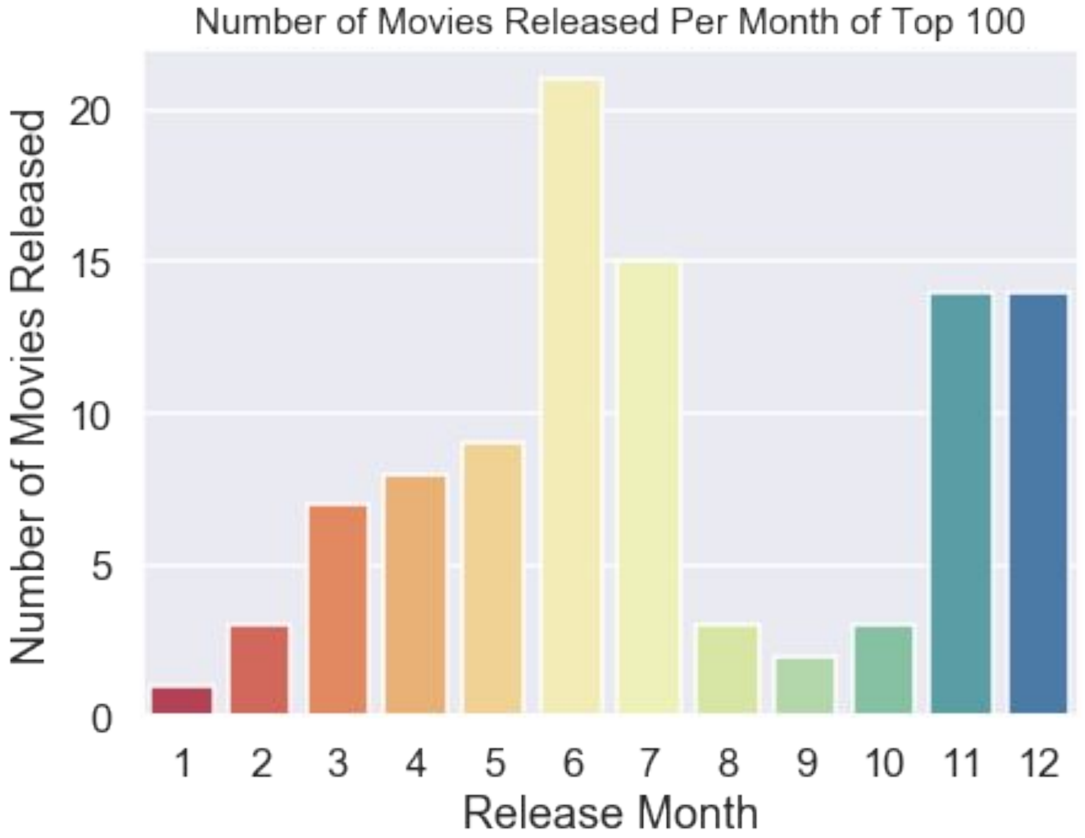
Creative Types vs. Average Rating & Worldwide Profit
1e9
Worldwide Profit
1.5
1.0
0.5
0.0
2
3
4
5
6
7
8
9
Average Rating
creative_type
Super Hero
Contemporary Fiction
Science Fiction
Kids Fiction
Fantasy
Historical Fiction
Dramatization
Multiple Creative Types
Factual

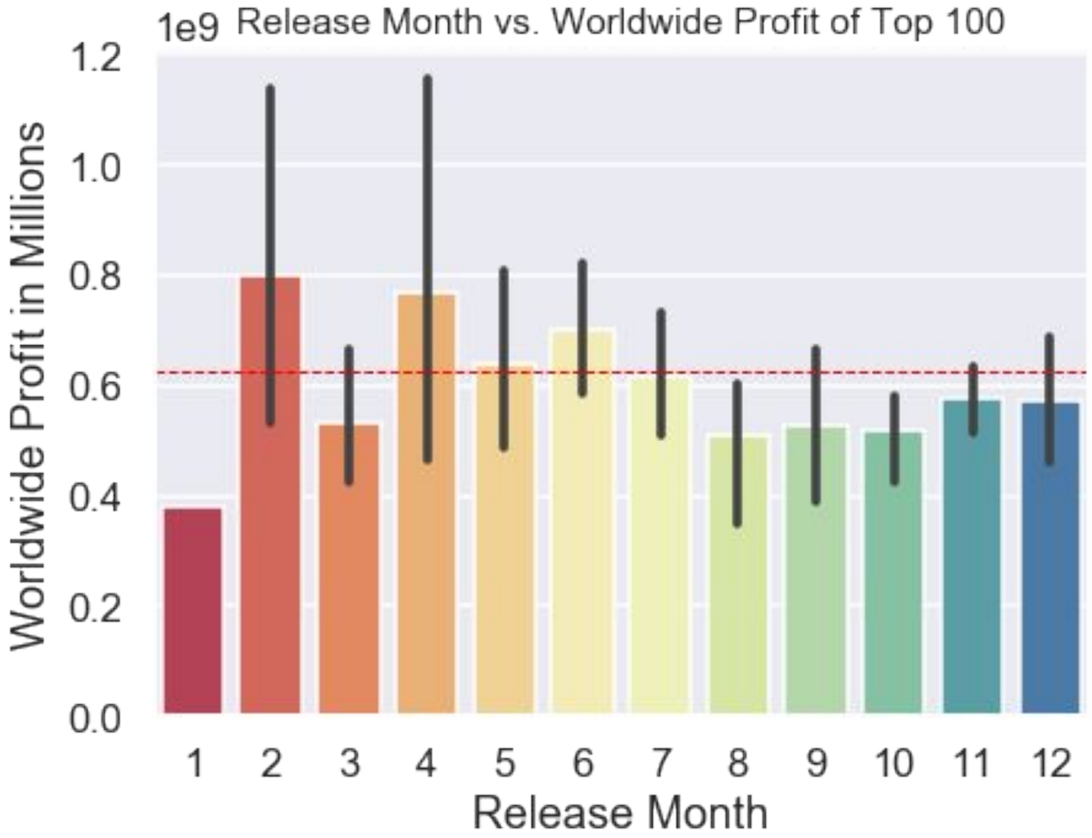

Production Methods vs. Average Rating & Worldwide Profit
1e9
Worldwide Profit
1.5
1.0
0.5
0.0
2
3
4
5
6
7
8
9
Average Rating
production_method
Animation/Live Action
Live Action
Digital Animation
Stop-Motion Animation


Runtime vs. Worldwide Profit
1e9
Worldwide Profit in Millions
1.5
1.0
0.5
0.0
40
60
80
100
120
140
160
180
Runtime in Minutes

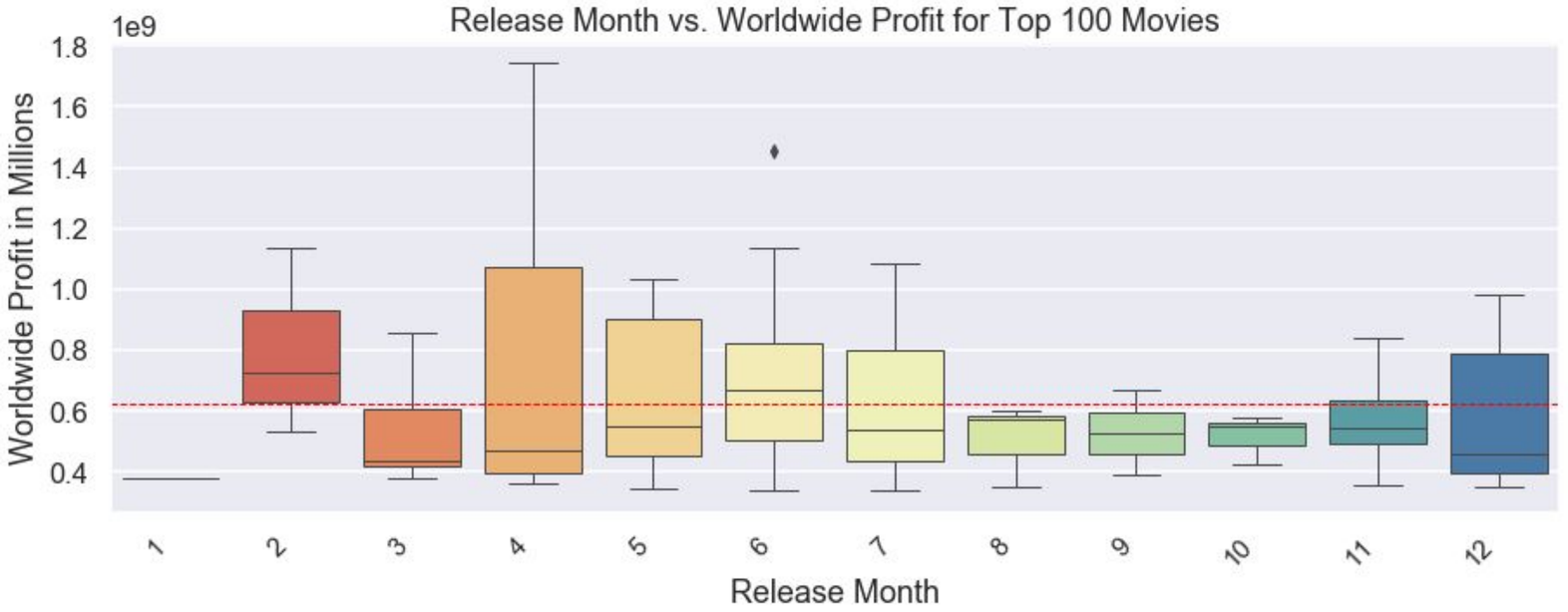

Runtime vs. Worldwide Profit
1e9
General
Top 100
Worldwide Profit in Millions
1.5
1.0
0.5
0.0
40
60
80
100
120
140
160
180
Runtime in Minutes

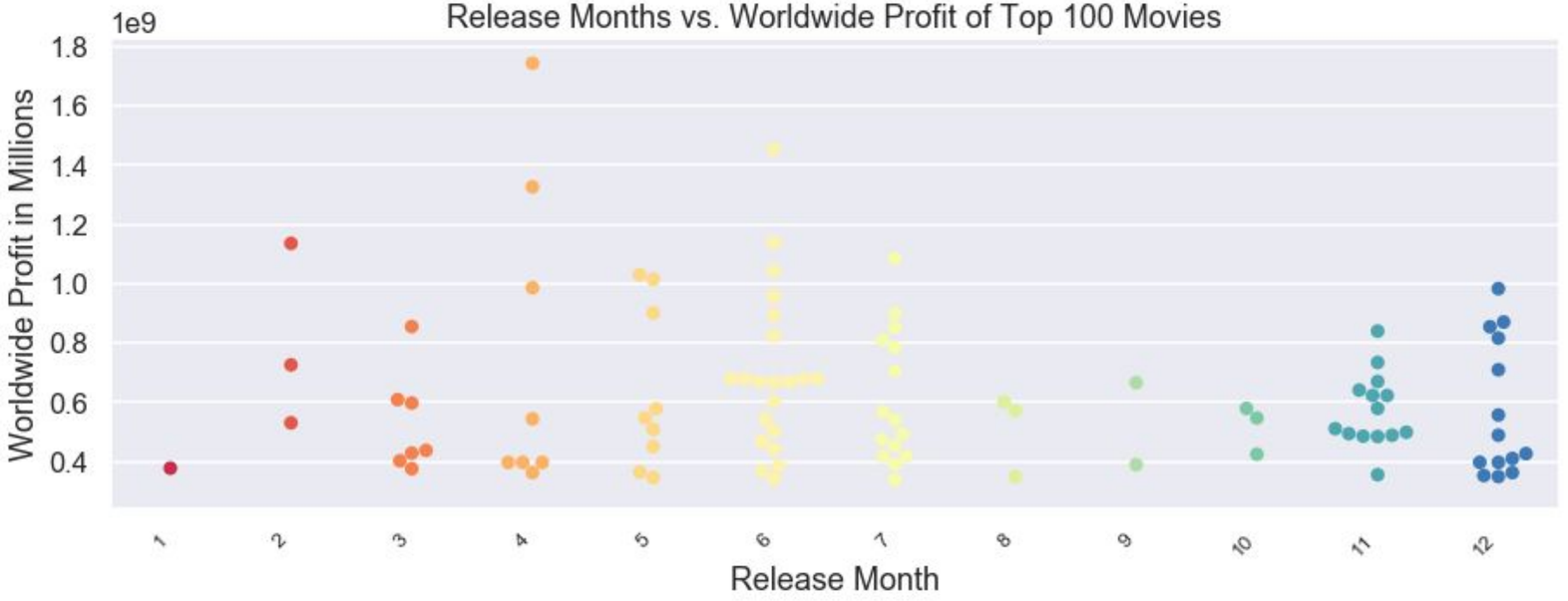

Release Month vs. Worldwide Profit
1e8
Worldwide Profit in Millions
Months
All Movies
mean avg
Top 100
1
2
3
4
5
6
7
8
9
10
11
12

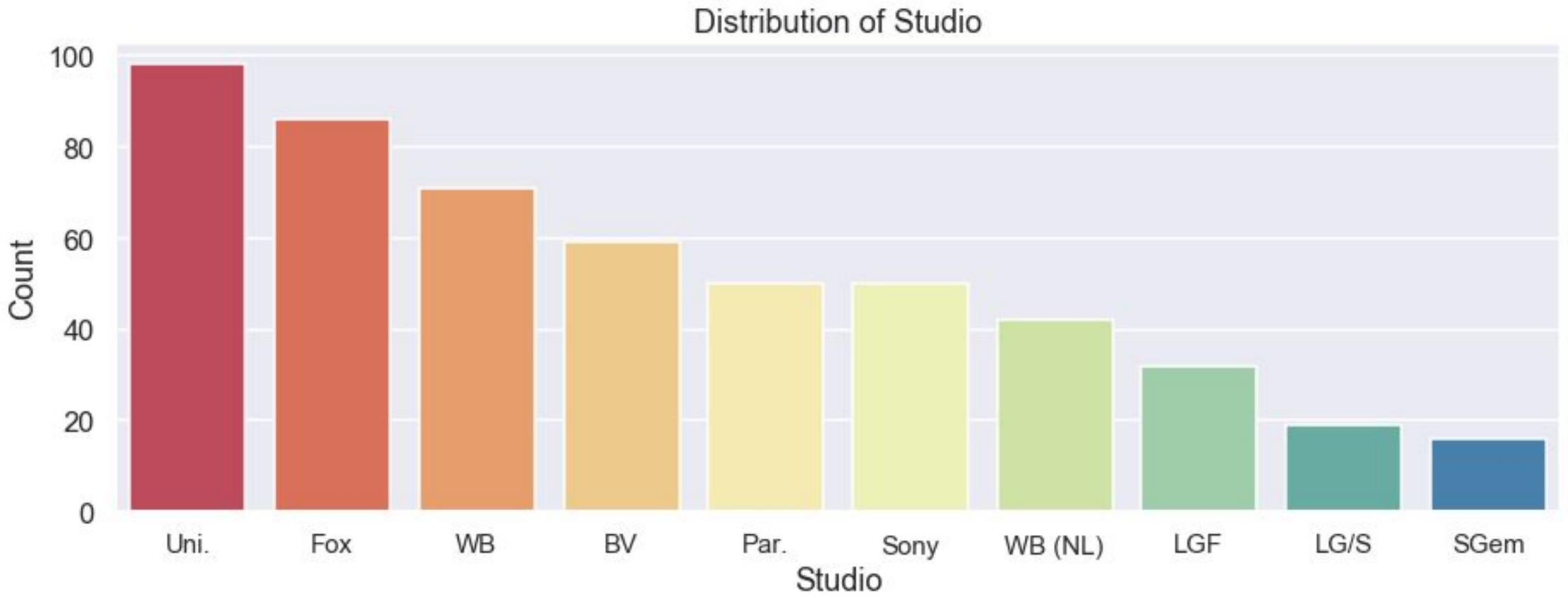


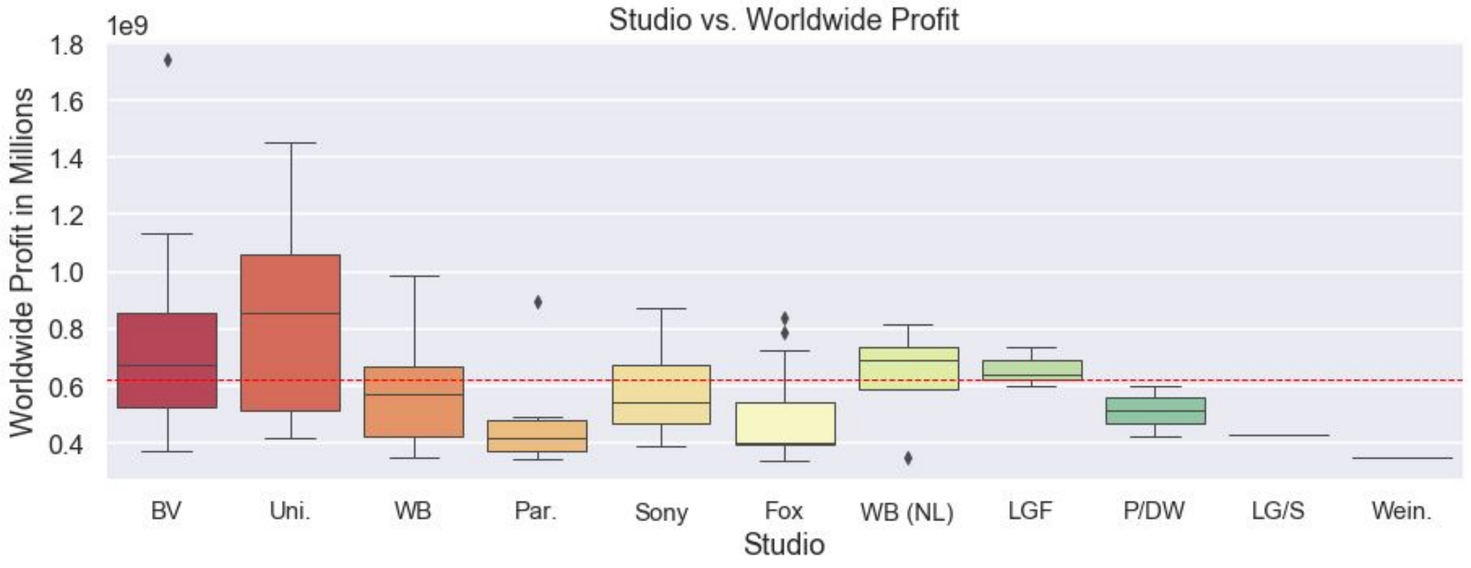


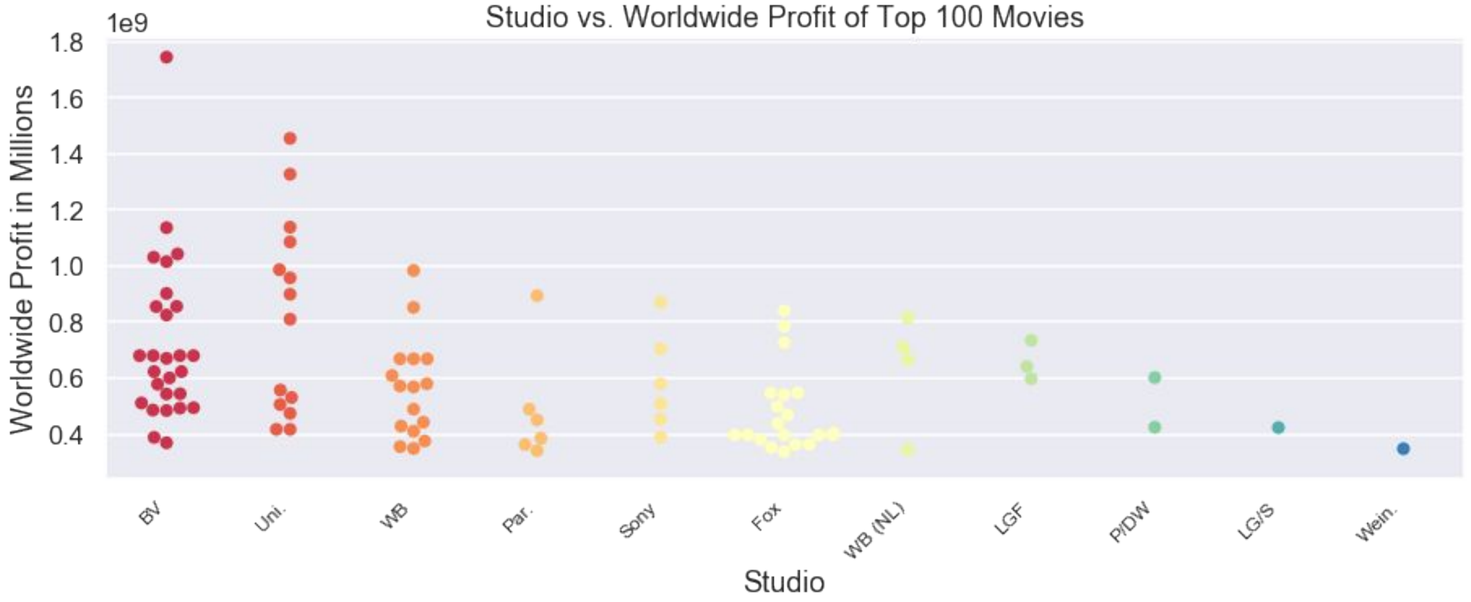
- Most release = June, December
- Most profit = Febuary, April but not by much

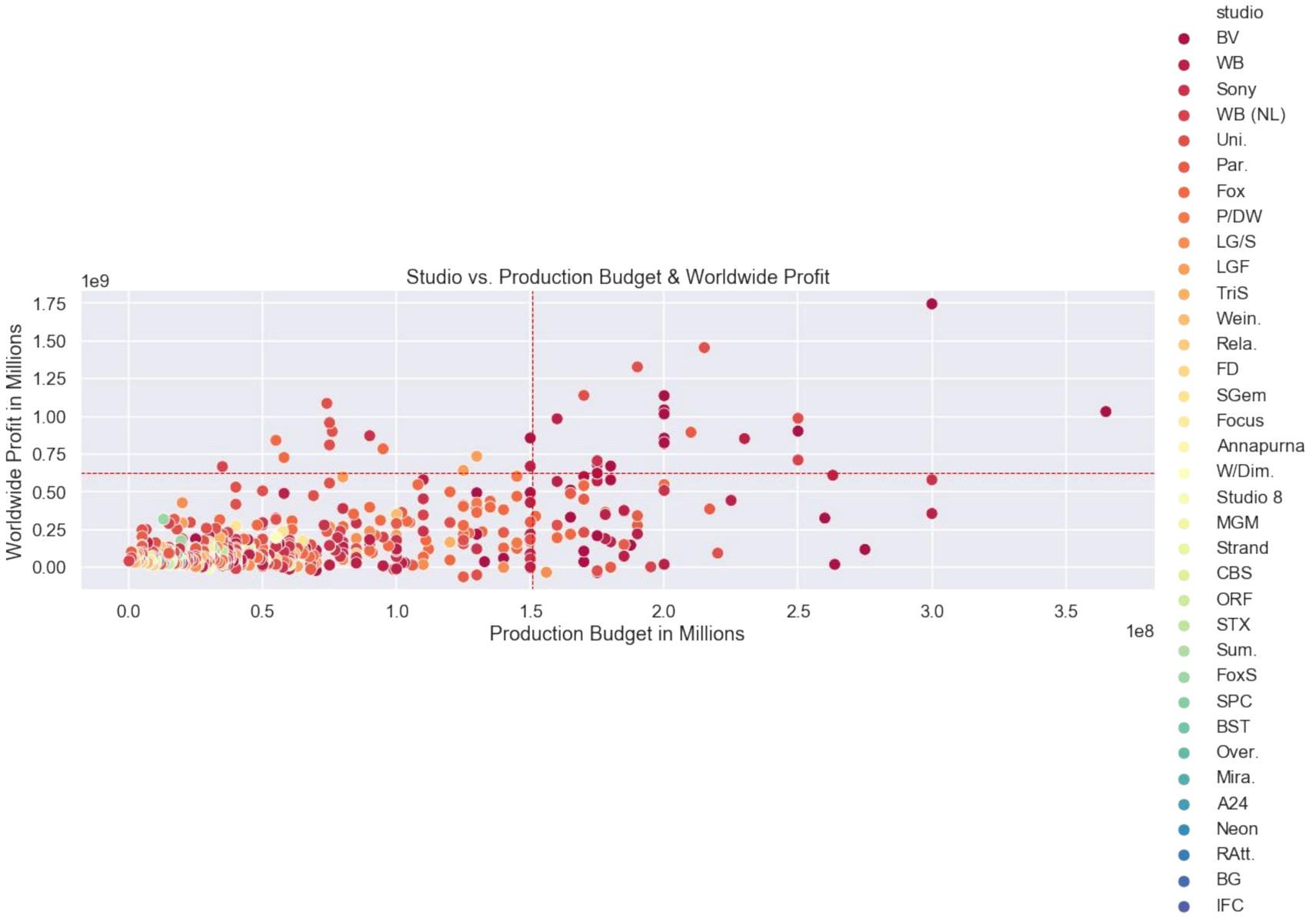

Release Month vs. Worldwide Profit for Top 100 Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
1
2
3
4
5
6
7
8
9
10
11
12
Release Month


Release Months vs. Worldwide Profit of Top 100 Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
1
2
3
4
5
6
7
8
9
10
11
12
Release Month


Distribution of Studio
Count
100
80
60
40
20
0
Uni.
Fox
WB
BV
Par.
Sony
WB (NL)
LGF
LG/S
SGem
Studio


Studio vs. Worldwide Profit
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
BV
Uni.
WB
Par.
Sony
Fox
WB (NL)
LGF
P/DW
LG/S
Wein.
Studio


Studio vs. Worldwide Profit of Top 100 Movies
1e9
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
Worldwide Profit in Millions
BV
Uni.
WB
Par.
Sony
Fox
WB (NL)
LGF
P/DW
LG/S
Wein.
Studio


Studio vs. Production Budget & Worldwide Profit
1e9
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
Worldwide Profit in Millions
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
1e8
Production Budget in Millions
studio
BV
WB
Sony
WB (NL)
Uni.
Par.
Fox
P/DW
LG/S
LGF
TriS
Wein.
Rela.
FD
SGem
Focus
Annapurna
W/Dim.
Studio 8
MGM
Strand
CBS
ORF
STX
Sum.
FoxS
SPC
BST
Over.
Mira.
A24
Neon
RAtt.
BG
IFC

Examples

- Action + Science Fiction
- Based on a popular book
- Star power: Jennifer Lawrence
- Underserved need for strong female protagonists