

# MAKING MOVIE MAGIC

---

*By Natalya Doris*



## OVERVIEW

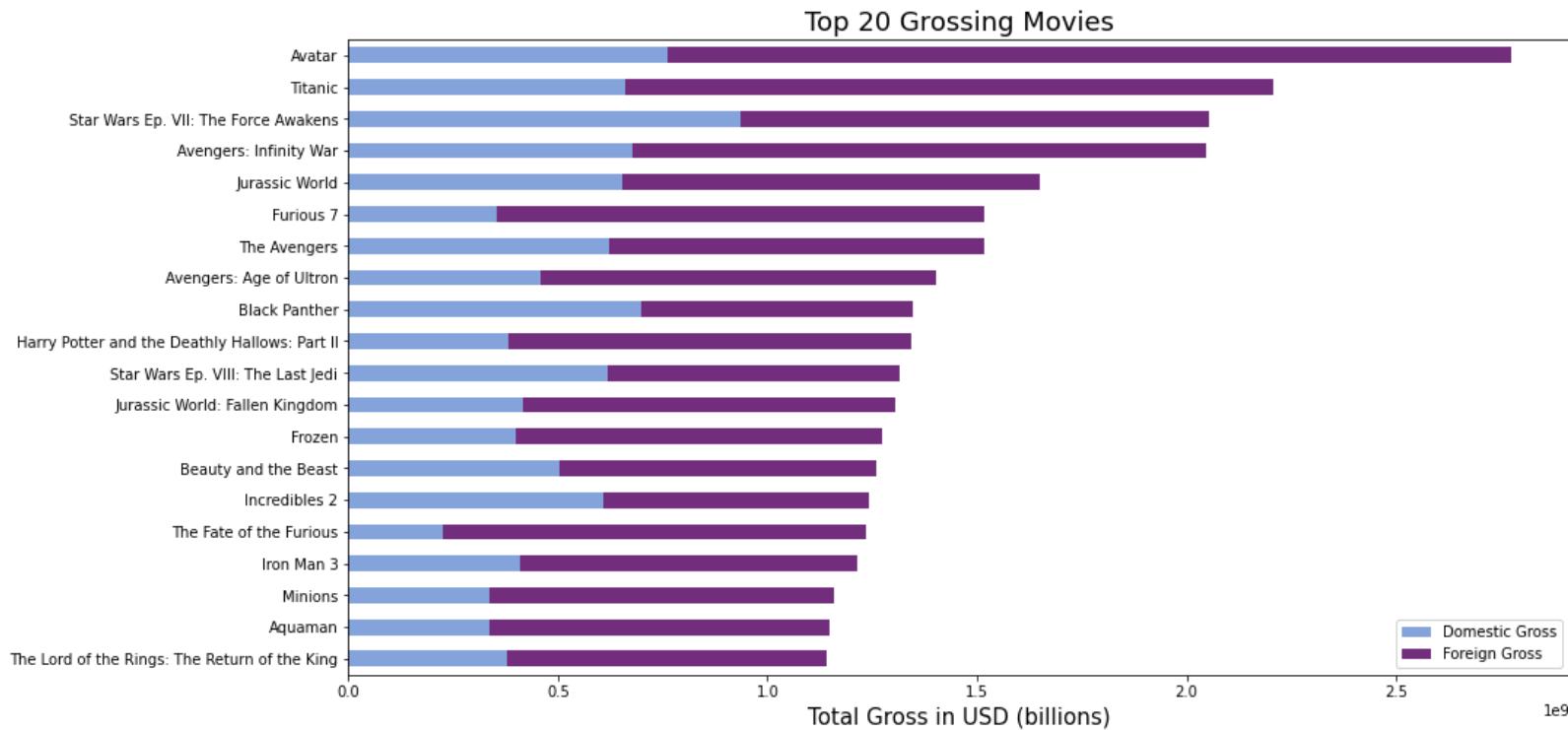
---

- This project analyzes movie data from Box Office Mojo, IMDB, Rotten Tomatoes, TheMovieDB and The Numbers to better understand what types of movies have been the most successful in the last decade, both from a profitability and popularity perspective. Analysis of recent historical budget, revenue, genre and review data can help Microsoft determine how to best strategize their impending entrance into the movie market.

# BUSINESS PROBLEM

---

- Microsoft is considering opening a movie studio, and faces a number of strategic choices, including genre, budget size, and timing. In this analysis, we will take a deep dive into the current landscape of the film industry to determine the most efficient use of their capital across these categories.

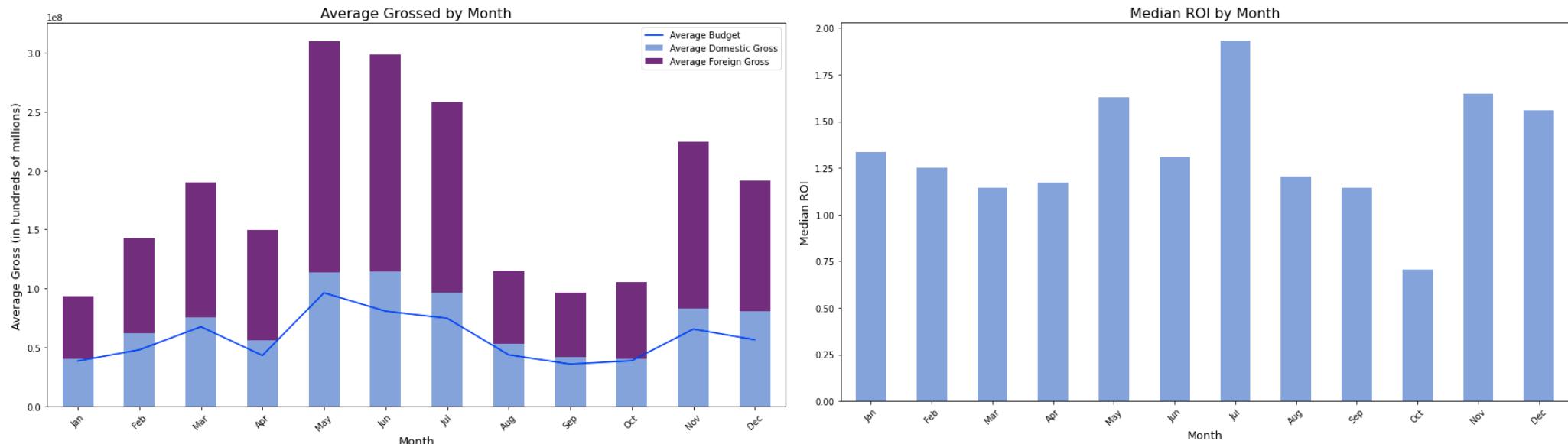


## DATA UNDERSTANDING

---

- The Movie Database (TMDB) dataset includes ~27,000 movies released from 1930 to 2020, with variables describing genre, popularity, vote average & vote count and release date.
- The movie budget data frame contains ~5,800 movies released between 1915 and 2020, and includes variables for domestic gross revenues, worldwide gross revenues, production budget, and release date.
- In this analysis, we combine the above datasets, linking genre and review data with revenue and budget data. The final dataset has a total of ~1,800 rows and spans the years 1946 to 2019. We focus primarily on movies released on or after 2010 with medium to very high budgets.

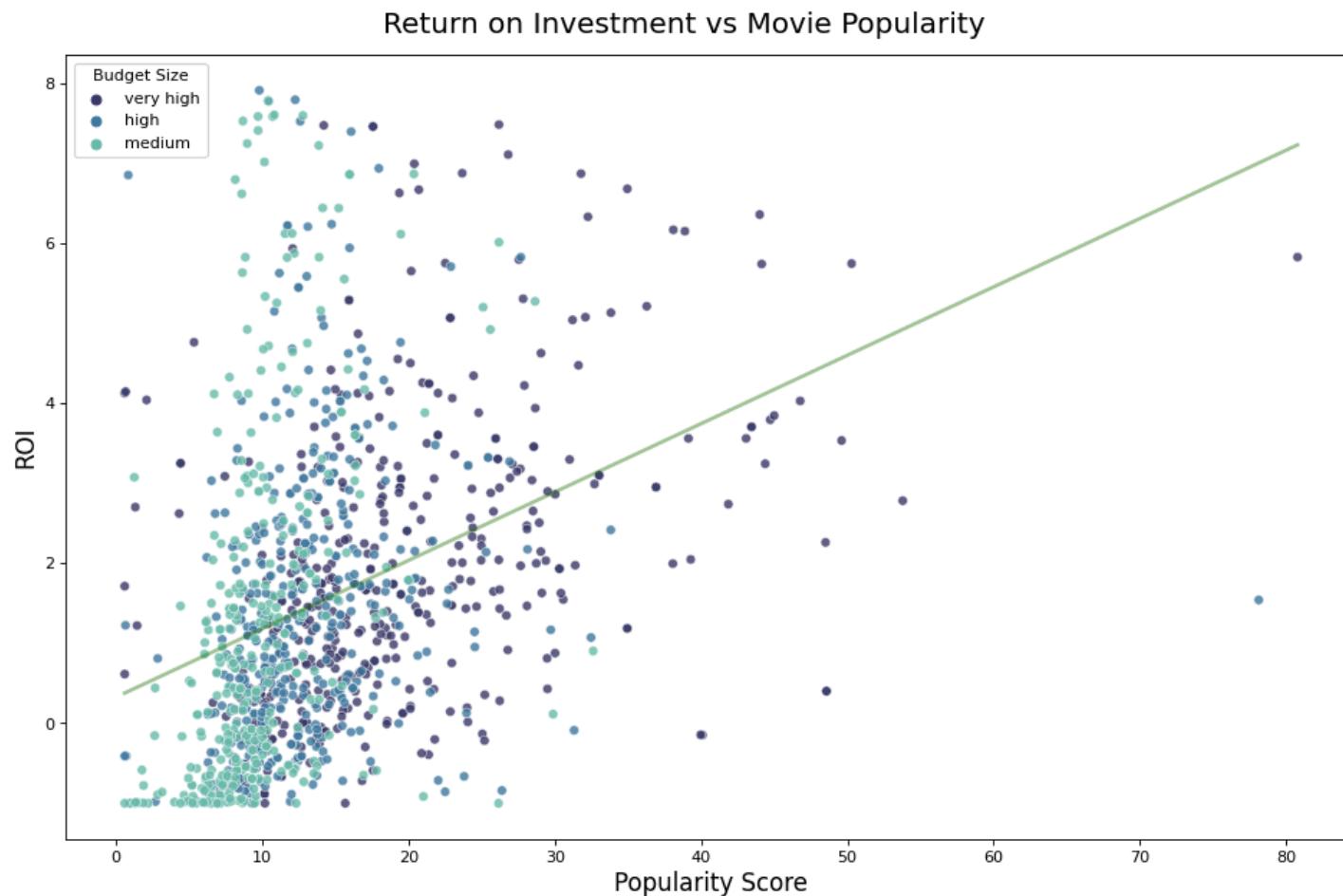
# CAPITALIZE ON HOLIDAYS



- Summer, winter holiday months see largest gross revenues and returns
- Movie budgets most elevated during summer months

# POPULARITY MATTERS

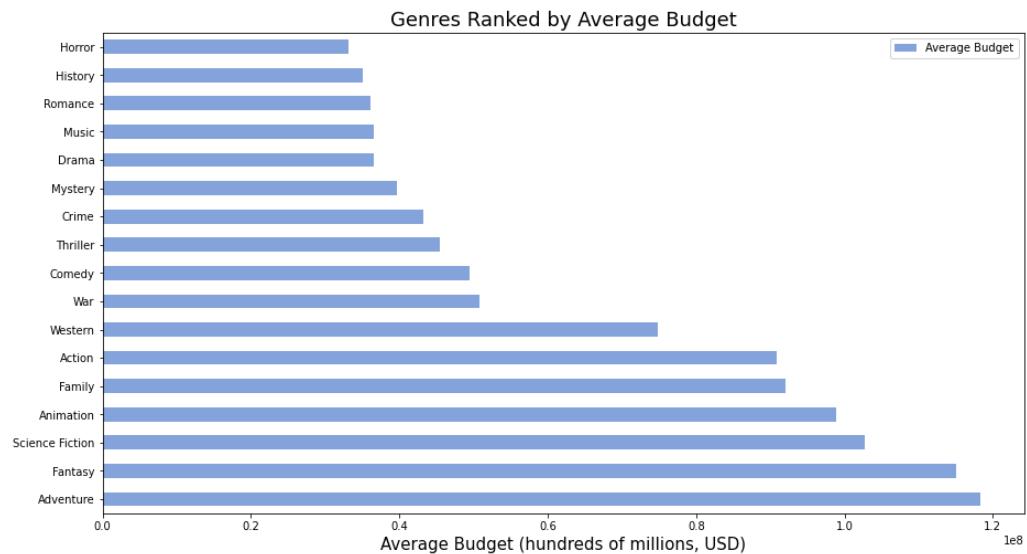
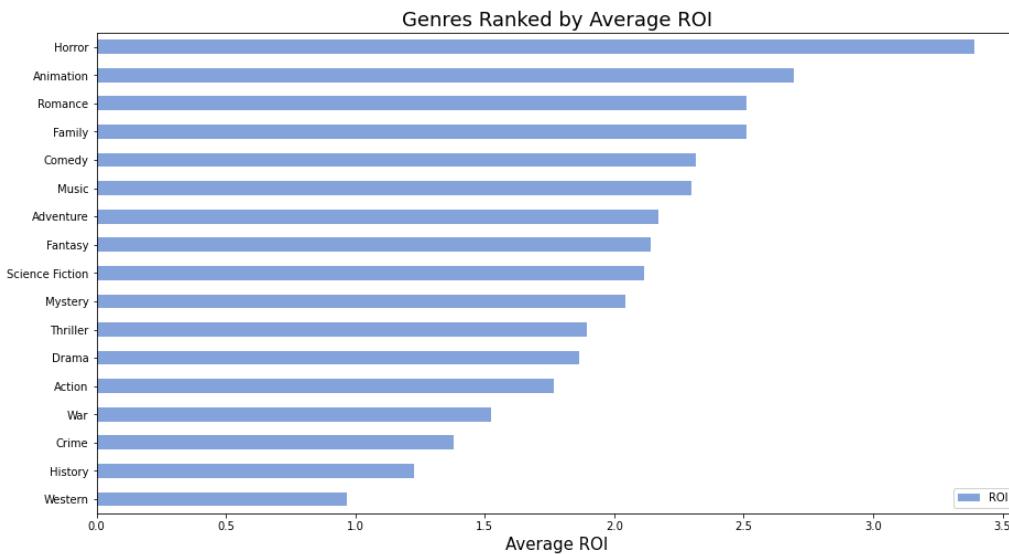
---



# GENRE

---

- Horror, the lowest budget genre, also has the highest ROI
- Despite higher budgets, animation is second highest ROI genre
- Romance, music among lowest budget & highest ROI genres



# SCORECARD - GENRE

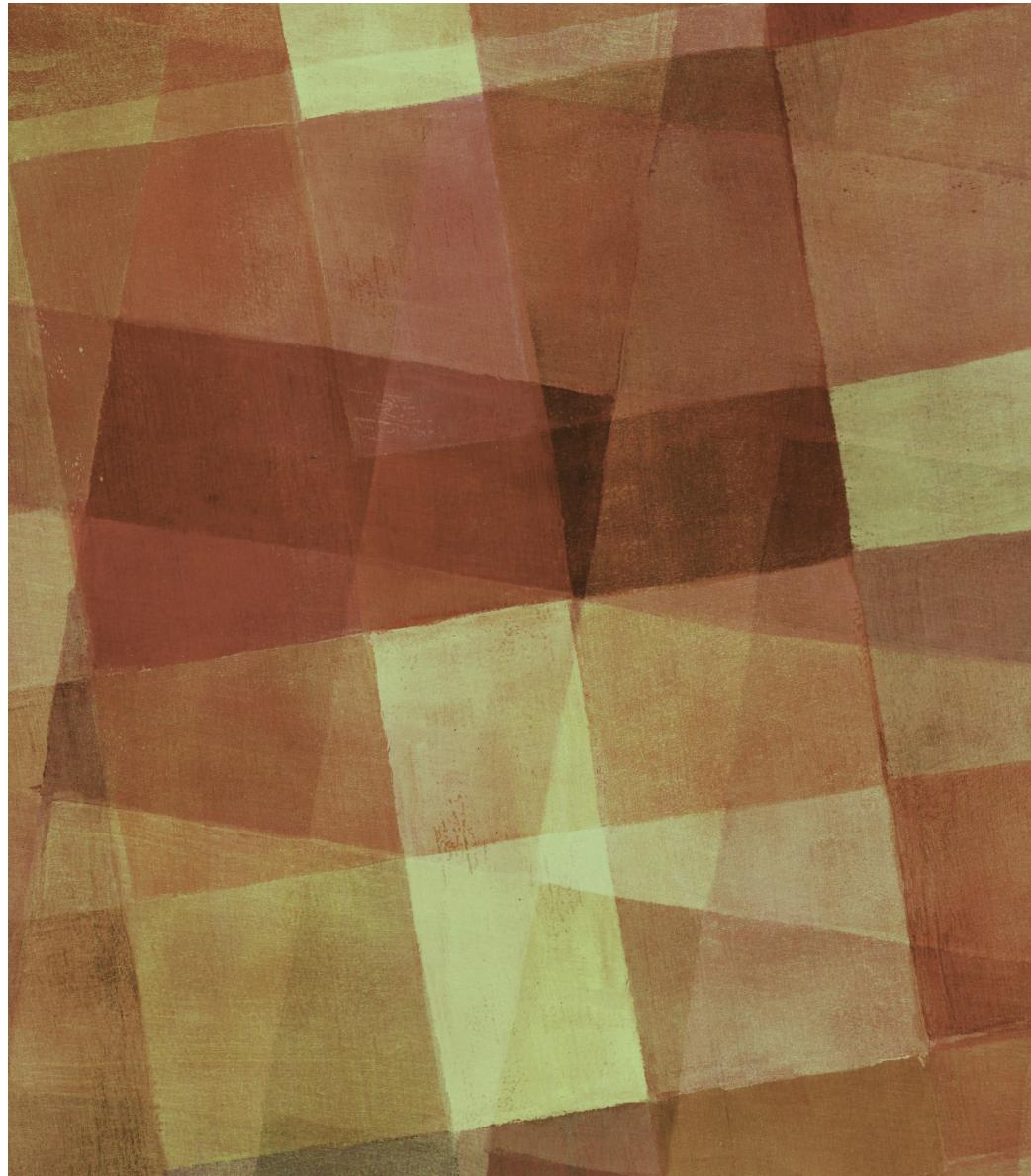
Genre	Count	Total Gross	Average Profit	Average Budget	Median ROI	% Foreign Revenues	Avg. Profit Margin	Avg. Vote	Popularity	Budget Category	Score
Animation	109	\$41,456,366,568	\$281,485,014	\$98,848,624	2.70	61.2%	0.74	6.54	15.73	very high	0.82
Adventure	287	\$114,859,571,883	\$281,931,958	\$118,275,610	2.17	62.0%	0.70	6.44	19.53	very high	0.68
Science Fiction	175	\$58,577,281,983	\$232,057,611	\$102,669,714	2.12	60.4%	0.69	6.39	19.70	very high	0.57
Fantasy	153	\$57,029,415,024	\$257,651,732	\$115,089,542	2.14	61.9%	0.69	6.33	19.84	very high	0.52
Family	170	\$55,355,999,945	\$233,568,235	\$92,055,294	2.51	57.9%	0.72	6.46	15.20	very high	0.50
Romance	156	\$18,251,426,642	\$80,894,402	\$36,101,923	2.51	50.2%	0.69	6.48	11.69	high	0.10
Action	393	\$110,840,804,013	\$191,218,076	\$90,819,593	1.77	57.8%	0.68	6.24	18.30	very high	0.06
Comedy	432	\$73,348,622,494	\$120,331,765	\$49,456,713	2.32	45.3%	0.71	6.26	13.19	high	-0.02
Horror	103	\$11,433,894,154	\$77,928,099	\$33,080,583	3.39	54.1%	0.70	5.81	13.17	high	-0.08
Drama	586	\$60,581,202,870	\$66,747,275	\$36,633,618	1.87	48.8%	0.65	6.64	12.58	high	-0.09
Music	35	\$4,227,519,643	\$84,229,133	\$36,557,143	2.30	35.9%	0.70	6.48	10.23	high	-0.10
War	42	\$5,865,804,654	\$88,809,635	\$50,852,381	1.52	48.3%	0.64	6.62	14.20	high	-0.17
Mystery	91	\$10,244,912,167	\$72,960,573	\$39,620,879	2.04	57.1%	0.65	6.23	14.20	high	-0.33
Thriller	340	\$43,050,542,607	\$81,095,125	\$45,524,118	1.89	55.4%	0.64	6.22	14.70	high	-0.39
Crime	194	\$22,965,852,528	\$75,099,240	\$43,281,443	1.38	51.5%	0.63	6.37	14.27	high	-0.50
History	70	\$4,821,672,962	\$33,781,042	\$35,100,000	1.22	51.5%	0.49	6.86	12.95	high	-0.72
Western	18	\$2,927,798,322	\$87,822,129	\$74,833,333	0.97	46.6%	0.54	6.54	14.64	very high	-0.85

# SCORECARD - STUDIOS

	Movie Count	Genre Count	Total Budget	Total Gross	ROI	Profit Margin	Avg. Vote	Avg. Vote Count	Popularity	Avg. Budget Size	Score
studio											
BV	79	17	\$10,677,300,000	\$38,248,690,075	2.58	0.72	6.94	5490	21.099658	very high	1.22
FoxS	32	12	\$488,100,000	\$2,369,639,346	3.85	0.79	6.88	2331	12.107469	medium	0.81
WB (NL)	37	13	\$2,245,000,000	\$9,049,872,986	3.03	0.75	6.46	3348	15.305514	very high	0.69
Wein.	36	12	\$818,100,000	\$2,767,132,410	2.38	0.70	6.96	2599	12.917417	high	0.67
Fox	117	17	\$8,535,000,000	\$29,734,509,875	2.48	0.71	6.45	3429	16.151769	very high	0.66
P/DW	10	8	\$1,334,000,000	\$5,078,027,601	2.81	0.74	6.52	3378	14.669800	very high	0.65
A24	16	9	\$183,500,000	\$460,260,186	1.51	0.60	6.94	2317	14.250125	medium	0.61
Uni.	113	16	\$6,749,200,000	\$27,445,930,630	3.07	0.75	6.26	2676	15.283752	very high	0.61
WB	97	17	\$8,773,000,000	\$23,527,640,561	1.68	0.63	6.45	3691	17.274979	very high	0.60
LGF	54	16	\$2,279,600,000	\$7,875,861,491	2.45	0.71	6.45	3016	15.176074	high	0.59
STX	17	11	\$604,100,000	\$1,453,150,394	1.41	0.58	6.51	1410	16.623000	high	0.55
Sony	74	16	\$5,377,500,000	\$19,380,490,279	2.60	0.72	6.16	2980	16.055770	very high	0.54
LG/S	27	14	\$1,608,000,000	\$3,567,604,679	1.22	0.55	6.39	3467	17.201852	very high	0.50
Par.	71	18	\$4,993,000,000	\$14,056,291,609	1.82	0.64	6.37	3372	15.041268	very high	0.45
TriS	13	11	\$454,800,000	\$1,383,485,283	2.04	0.67	6.55	2411	13.070846	high	0.44
Focus	33	14	\$772,800,000	\$2,064,112,073	1.67	0.63	6.70	1803	12.496848	high	0.41
Sum.	11	12	\$349,800,000	\$1,242,197,975	2.55	0.72	6.30	1728	11.134545	high	0.28
CBS	8	9	\$182,000,000	\$473,844,600	1.60	0.62	6.43	1414	10.728125	high	0.16
SPC	12	8	\$179,200,000	\$506,099,122	1.82	0.65	6.53	947	8.975833	medium	0.13
ORF	24	14	\$569,700,000	\$1,035,331,792	0.82	0.45	6.35	1502	11.853167	high	0.07
SGem	24	13	\$721,500,000	\$2,356,561,693	2.27	0.69	6.06	1435	9.854542	high	0.04
RAtt.	11	6	\$184,400,000	\$306,435,564	0.66	0.40	6.52	1045	10.074091	medium	0.01
BST	7	9	\$132,000,000	\$140,904,444	0.07	0.06	6.61	890	10.789286	medium	0.00
Rela.	23	15	\$679,000,000	\$1,328,915,479	0.96	0.49	6.04	1046	10.581043	high	-0.13
IFC	6	5	\$83,000,000	\$19,844,214	-0.76	-3.18	6.68	340	7.364500	medium	-0.34
BG	6	9	\$119,000,000	\$98,387,434	-0.17	-0.21	5.72	705	9.241167	medium	-0.54

## KEY TAKEAWAYS

- Aim to release in **summer or winter**, as movies that hit the theatre during this times yield the biggest revenues and returns on investment.
- If budget size isn't a barrier, **animation** is the most successful genre in terms of ROI and popularity/voted rating, followed by **adventure, sci-fi, fantasy** and **family**.
- On the lower end of the budget spectrum, focus on **horror** movies to yield the largest returns. **Romance** and **family** earned higher scores than horror, however, when taking into account both returns and popularity/rating.
- Successful studios tended to produce movies across a **variety of genres and budget categories**, implying that there are a number of potential strategies that yield success in the film industry





## NEXT STEPS

- While popularity tends to increase with budget size, a movie's average vote seems to be evenly distributed among budget sizes, implying other factors (movie quality, for example) not visible in this analysis impact voted rating. I would like to explore how rating (G / PG / PG-13 / R / Not-Rated), runtime, director quality, among other factors, impact average vote.
- I would like to better understand how genre success has evolved over time - are certain genres showing higher growth rates than others? Can we use this data to predict what genres show promise in the future?
- I would also like to see more detailed geographical revenue data - are there certain regions in the world that have produced more ROI & gross revenues than others?
- *If any further questions, please reach out to Natalya Doris at [ntdoris@gmail.com](mailto:ntdoris@gmail.com)*