



# Behind the Box Office: A Data Analysis of Film Success Factors





# Meet The Team



*Britney Aparicio*



*Abdurrahman Elkhadrawy*



*Allen Kwan*



*Abraham Cordova*



*Darien Miller*

# Table of Contents

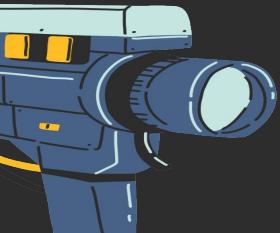
- 01 Introduction
- 02 Overview of Data
- 03 Analysis
- 04 Recommendations Sources



# Introduction



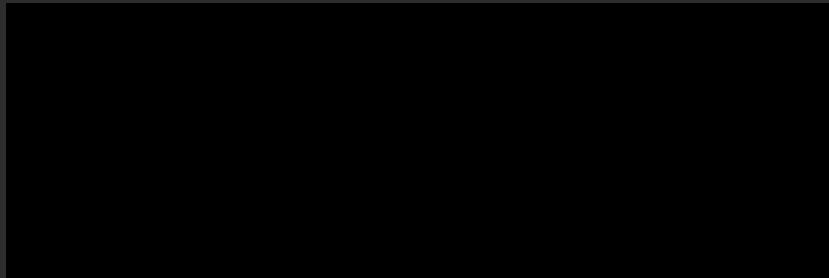
Movies are timeless ...

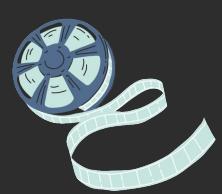




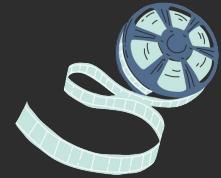
## ***Main Objective:***

Recommend the optimal movie investment  
for our stakeholders, focusing on  
maximizing returns by analyzing factors  
such as genre, release date, and  
classification.





# Basic Data Measurements



**Box Office Revenue (\$):** Money earned by films through ticket sales

**Budget (\$):** The grand total of actual spending to produce the film

**Return On Investment - ROI (%):** Gains or losses of an investment

$$\frac{\text{Revenues from Investment} - \text{Cost of Investment}}{\text{Cost of Investment}} \times 100 = \text{ROI (\%)}$$



# Data Overview for all Movies



## Number of Movies Analysed

### Count of Movie Data

508

### Number of Genres

17

## Median Values for Movies Analysed

### Median Box Office Revenue (\$)

\$79,350,000.00

### Median Budget (\$)

\$30,000,000.00

### Median ROI

170.99%

# Margin of Error



## Calculate your margin of error

Population Size ⓘ

680165

Confidence Level (%) ⓘ

95

Sample size ⓘ

508

Margin of error

**4%**

$$\text{MoE} = z \times \sqrt{\frac{p(1-p)}{n}}$$

$n$  = sample size

$z$  = z-score

$p$  = sample proportion

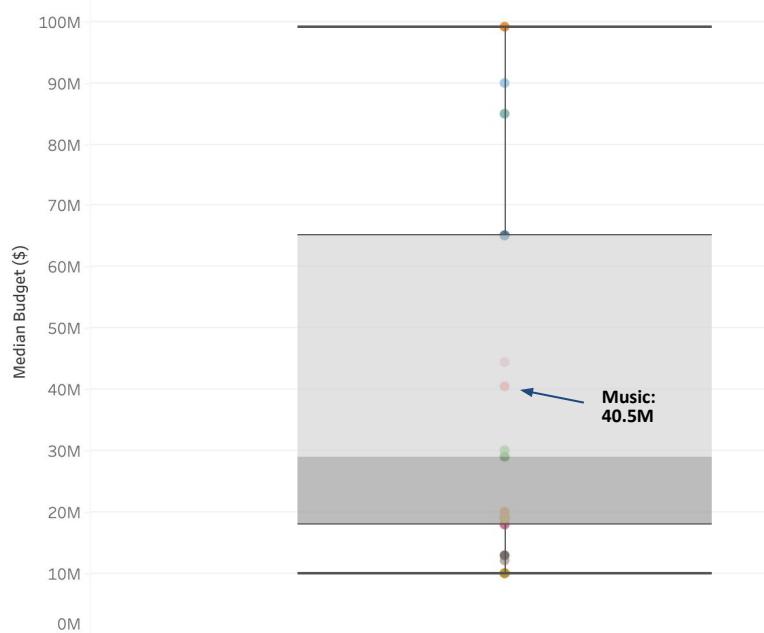
According to IMDb, there are **680,165 movies** in existence



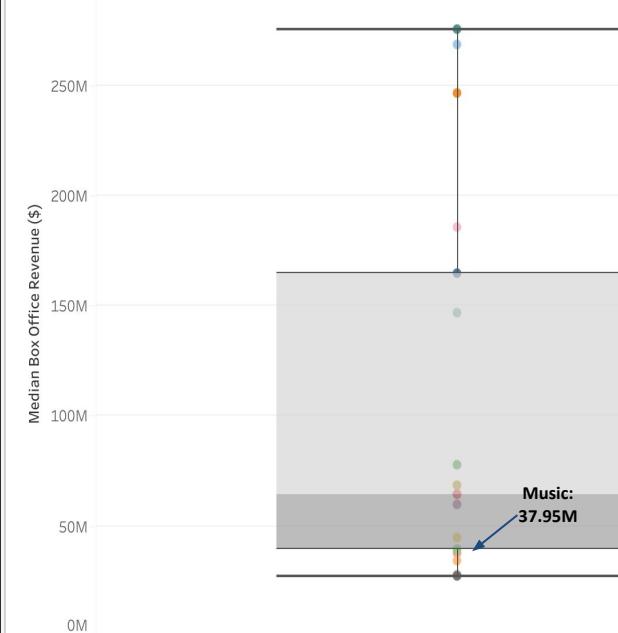
# Overall Insights: Median Revenue & Budget Breakdown



Median Budget



Median Revenue



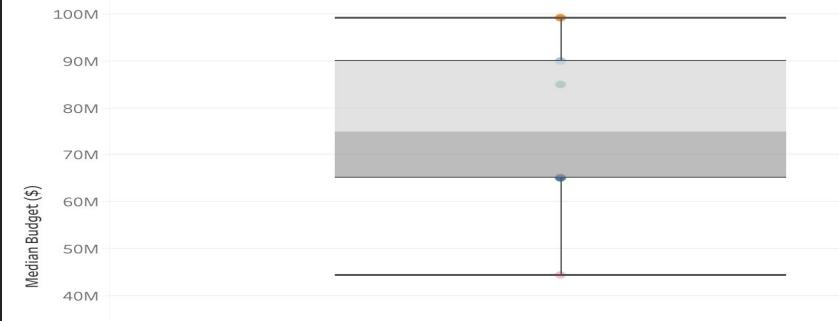
- Genre (1)
- Action
  - Adventure
  - Animation
  - Biography
  - Comedy
  - Crime
  - Documentary
  - Drama
  - Family
  - Fantasy
  - Horror
  - Musical
  - Mystery
  - Religious
  - Romance
  - Sci-Fi
  - Thriller



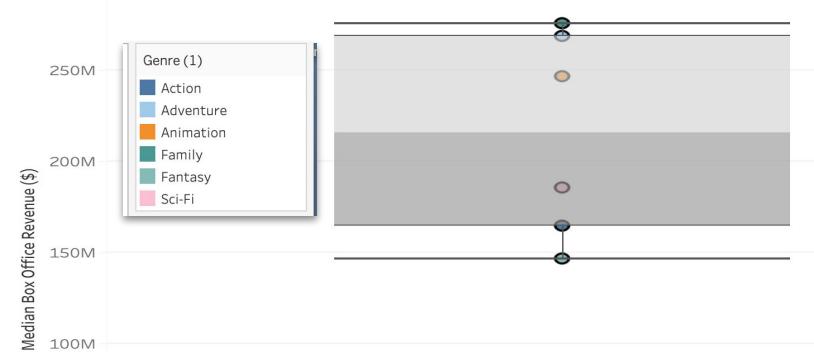
# Overall Insights: Median Revenue & Budget Breakdown



Median Budget



Median Revenue



**Top 5 genres Action, Adventure, Animation, Family, Fantasy and Sci- Fi**



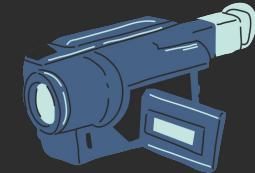
**Revenue:** Upper Whisker: \$275M Upper Hinge: \$164.5M



**Budget:** Upper Whisker: \$99M  
Upper Hinge: \$65M

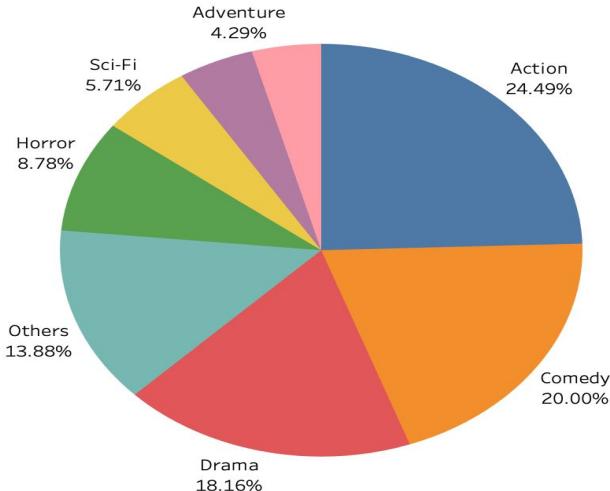


# Genre Popularity



- 💡 Action, Comedy and Drama comprise the biggest share of movies
- 🚀 They comprise 20%+ each, making up 60% of the dataset.
- 🔍 Crime and Adventure are the least popular sitting around 4% each.

What genre of Movies are the most Popular?





# Genre Profitability



**Horror is the most profitable genre by a whopping 460%!**

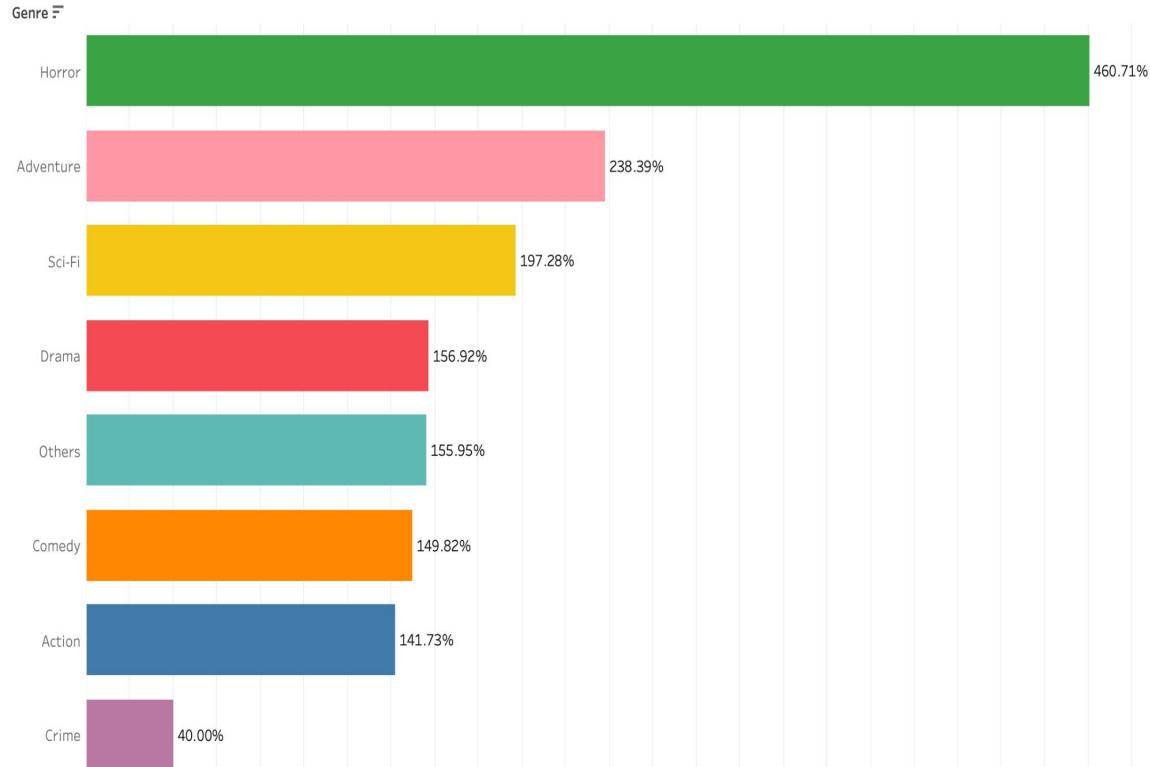


**Adventure is also very profitable but requires a higher budget.**



**Crime unfortunately while profitable, but not by much**

Which Genre gives us the best return on Investment?



# What's fueling Horror's Profitability?

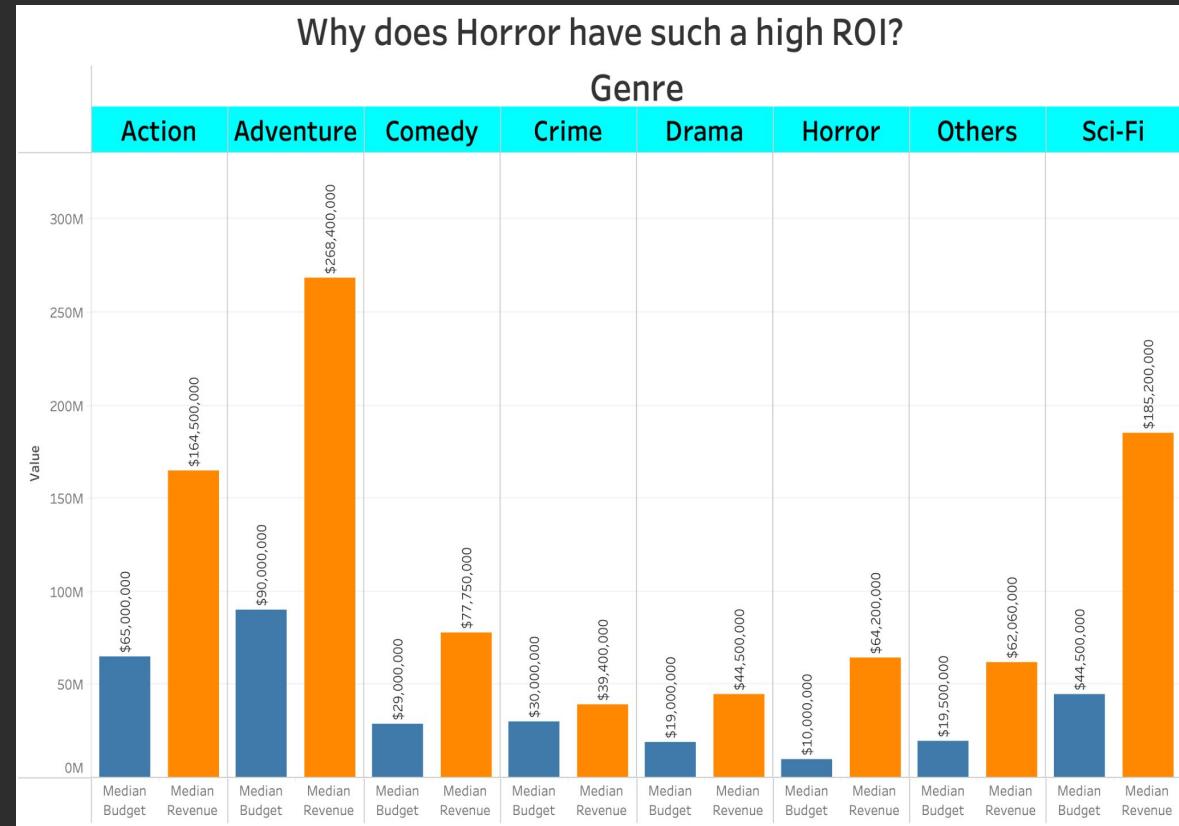


Genres like Adventure, Sci-Fi and Action generate a lot of revenue but aren't the most profitable.



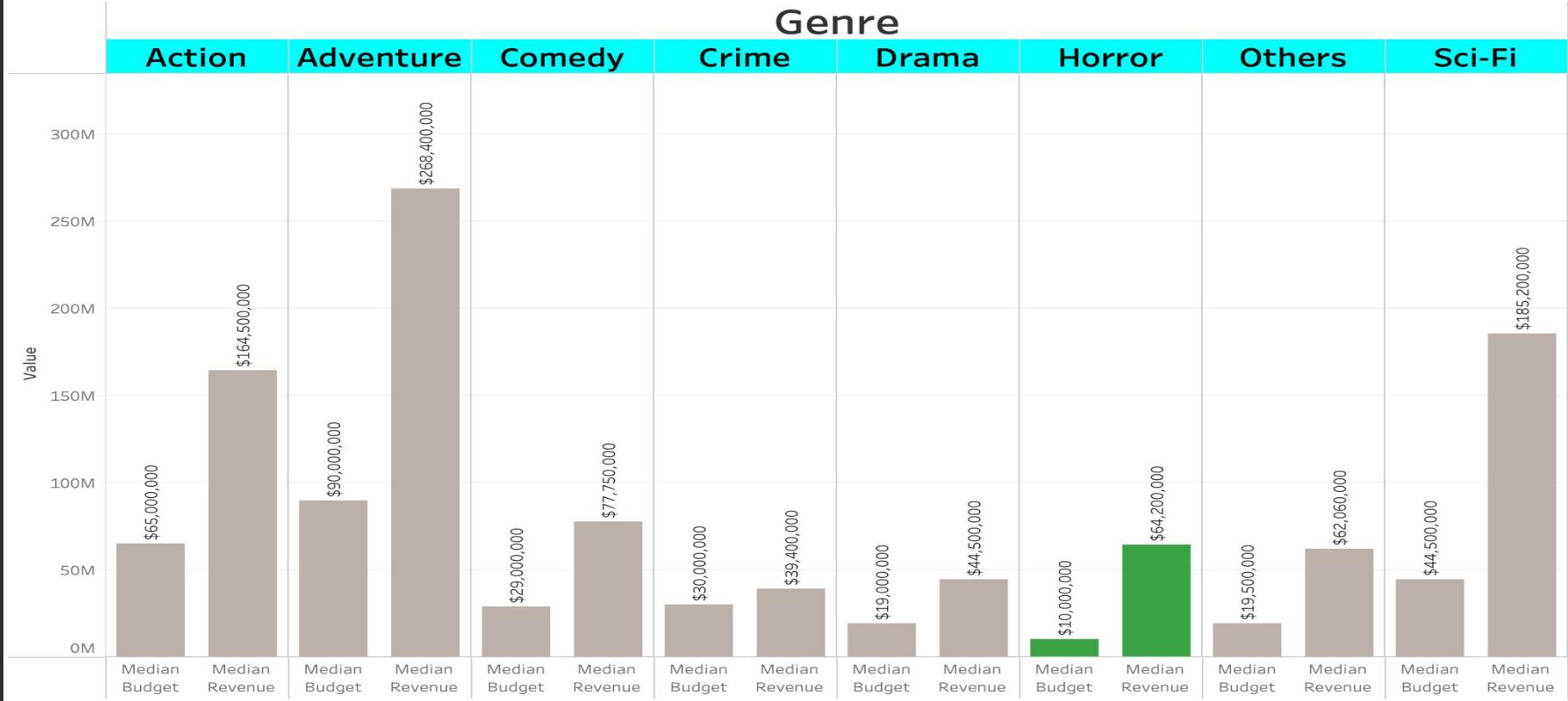
Crime generates a very slim revenue over it's budget.

Why does Horror have such a high ROI?



# Seems like Someone Is Standing Out

Why does Horror have such a high ROI?





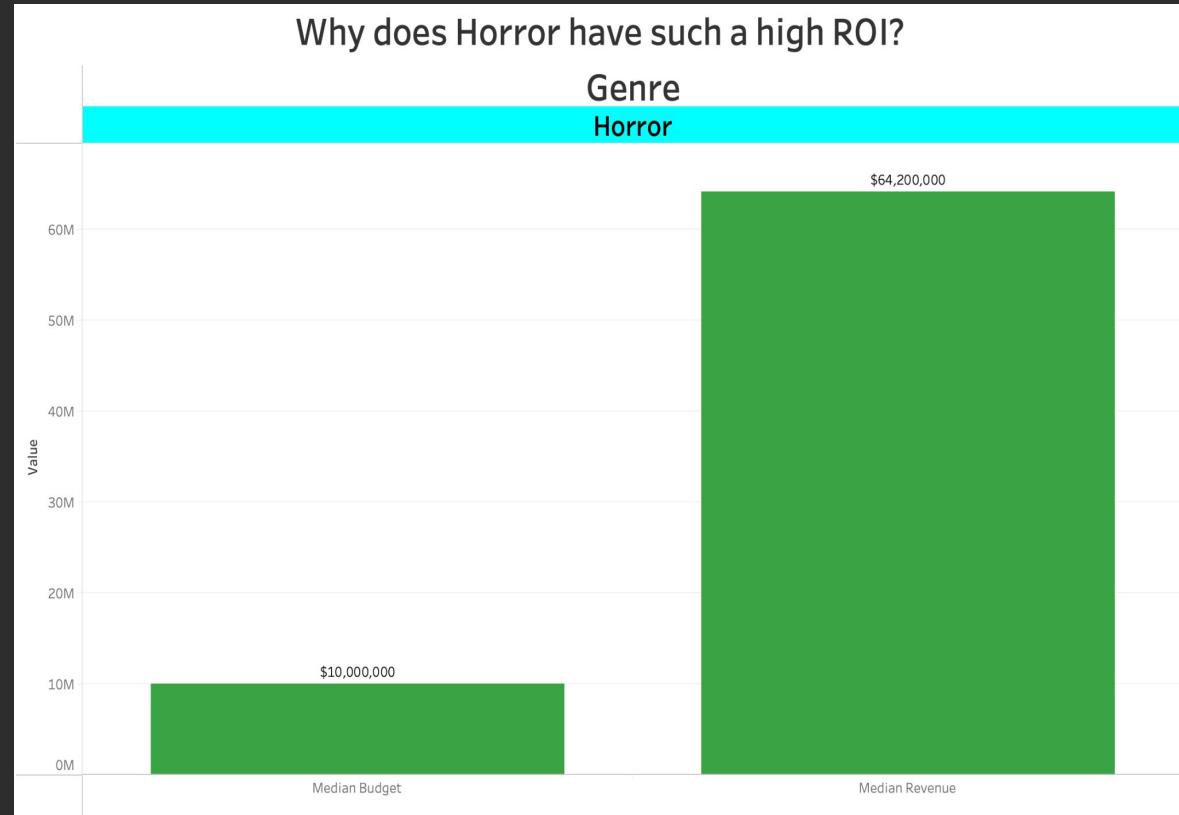
# Buy Low Sell High!



**The biggest reason for Horror's amazing ROI is its very below average low budget while also maintaining a average revenue**



**This is likely due to its creative storytelling instead of expensive movie set technology and actors**

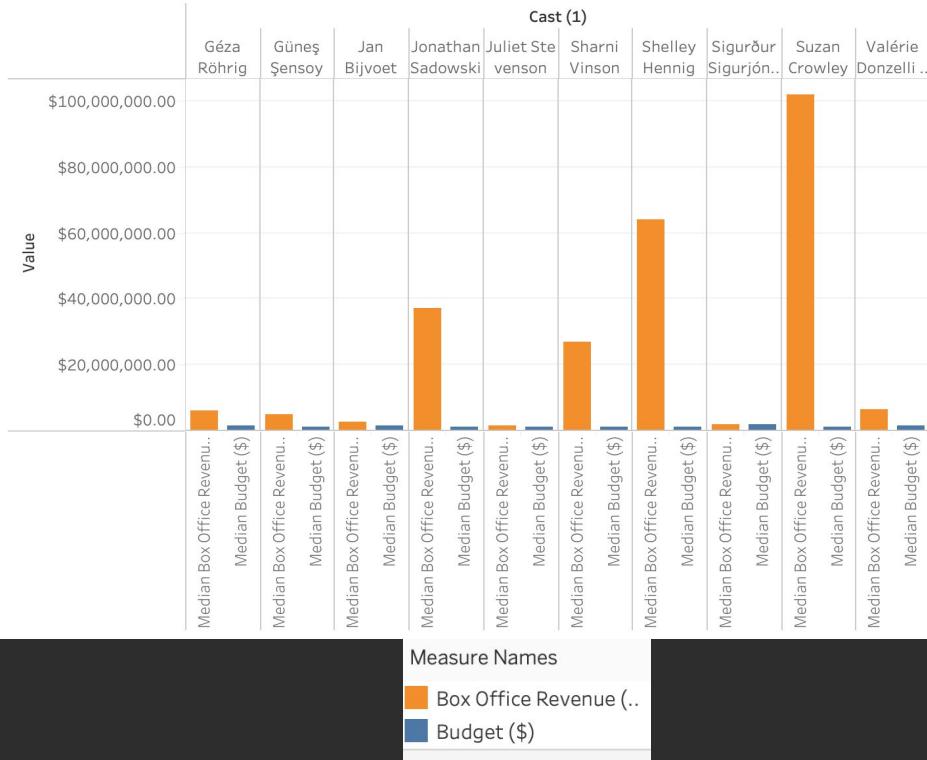




# Star Power vs. Dollar Power



Median Budget/ Revenue by Cast (LOW)



**Suzan Crowley has the lowest median movie investment of \$1M but the movie revedened \$101.8M of revenue**

💡 Horror Movie (only one movie)

💡 ROI: 10080%

💡 Franchise



**Sigurour Sigurjonsson has one of the lowest median movie investment of \$1.75M but created a revenue of \$1.74M**

💡 Drama Movie (only one movie)

💡 ROI: -0.57%

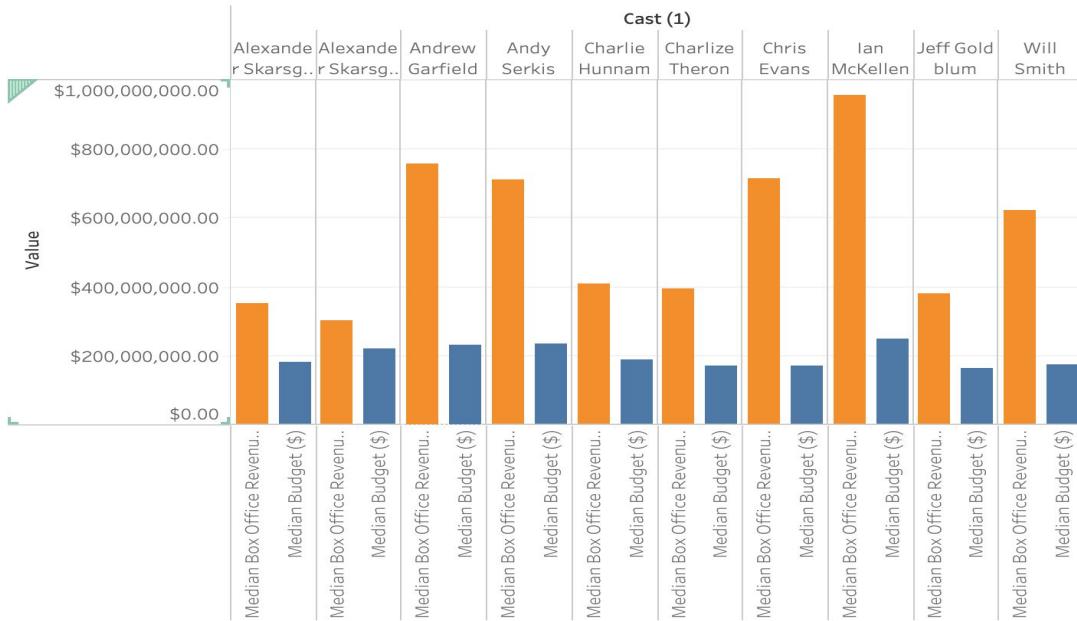
💡 New Movie



# Star Power vs. Dollar Power



Median Budget/ Revenue by Cast(Top)



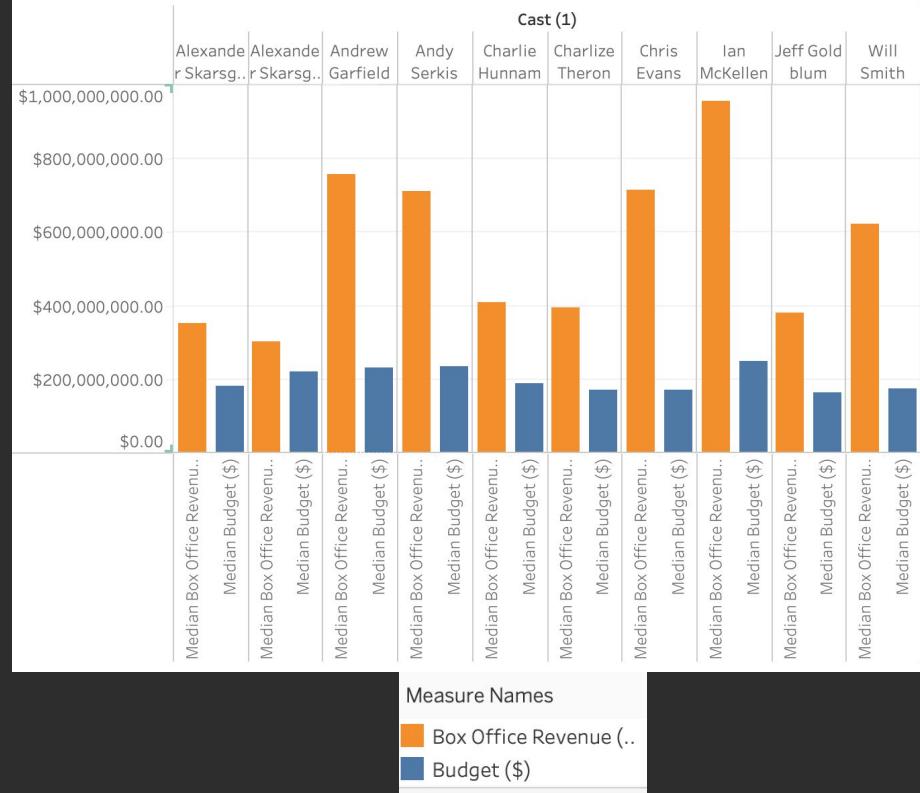
Is there a correlation  
between cast and a  
movie's profit?



# Star Power vs. Dollar Power



Median Budget/ Revenue by Cast(Top)



Ian McKellen w/ the highest median revenue at **\$956M but with a budget at \$250M**

💡 Adventure Movie (only one movie)

💡 ROI: 282.4%

💡 Release Date: December 1, 2014

💡 New Movie

**A Budget and its inclusion of an actor  
DOES NOT provide us with a successful  
movie**



# Pivot Tables



Horror Filter

Filter for ROI  
above 171%  
(median overall)

Patrick Wilson  
Honorable Mention



Cast (1)	COUNTA of Movie Title	Cast (2)	COUNTA of Movie Title	Cast (3)	COUNTA of Movie Title	Cast (4)	COUNTA of Movie Title	Cast (5)	COUNTA of Movie Title
Allison Miller	1	Alfre Woodard	1	Adam Scott	1	Adrian Rawlins	1	Alisa Khazanova	1
Annabelle Wallis	1	Allison Tolman	1	AJ Bowen	1	Aimee Carrero	1	Barbara Crampton	1
Anya Taylor-Joy	1	Betty Gabriel	1	Alexander DiPersia	1	Ana Coto	1	Elizabeth Blackmore	1
Ashley Bell	1	Ciarán Hinds	1	Angus Sampson	1	Brian Howe	1	Erin Moriarty	1
Blake Lively	1	Devin Kelley	1	Dakota Goyo	1	David Fox	1	Joseph Julian Soria	1
Daniel Radcliffe	1	Ed Oxenbould	1	Daniel Kash	1	Diana Hardcastle	1	Judd Lormand	1
Dermot Mulroney	1	Édgar Ramírez	1	David Mazouz	1	Evan Helmuth	1	Lin Shaye	1
Douglas Smith	1	Fernanda Andrade	1	Deanna Dunagan	1	Jessica Lucas	1	Nathan Phillips	1
Emjay Anthony	1	Gabriel Bateman	1	Donald Glover	1	Joel McHale	1	Oaklee Pendergast	1
Eric Bana	1	Jared Harris	1	Edwin Hodge	1	Kyle Secor	1	Olivia DeJonge	1
Ethan Hawke	1	Jeremy Irvine	1	Eric Ladin	1	Leigh Whannell	1	Vanessa Ray	1
Frank Grillo	1	Joe Swanberg	1	Fred Thompson	1	Liz White	1	Grand Total	11
James Ransone	1	Josh Hamilton	1	Helen McCrory	1	Louis Herthum	1		
Jane Levy	1	Juliet Rylance	1	Jacob Wysocki	1	Lucy Fry	1		
Jeffrey Dean Morgan	1	Kyra Sedgwick	1	Janet McTeer	1	Nicholas Tucci	1		
Jessica Chastain	1	Mark Duplass	1	Jesse McCartney	1	Olivia Taylor Dudley	1		
Jonathan Sadowski	1	Nikolaj Coster-Waldau	1	Jim Norton	1	Peter McRobbie	1		
Kathryn Hahn	1	Olivia Cooke	1	Kate Dickie	1	Saxon Sharbino	1		
Katie Featherston	1	Óscar Jaénada	1	Lin Shaye	1	Ty Simpkins	1		
Keri Russell	1	Radha Mitchell	1	Lou Taylor Pucci	1	Vincent D'Onofrio	1		
Kevin Bacon	1	Ralph Ineson	1	Muse Watson	1	Will Peltz	1		
Lauren Cohan	1	Renee Ostead	1	Olivia Munn	1	Grand Total	21		
Natalie Dormer	1	Rose Byrne	1	Rosemarie DeWitt	1				
Olivia Wilde	1	Rupert Evans	1	Sam Anderson	1				
Patrick Wilson	2	Shannyn Sossamon	1	Simon Quarterman	1				
Phoebe Fox	1	Shiloh Fernandez	1	Vivis Colombetti	1				
Sam Rockwell	1	Spencer Treat Clark	1	Grand Total	26				
Sharni Vinson	1	Stefanie Scott	1						
Shelley Hennig	1	Taylor Kinney	1						
Suzan Crowley	1	Tyler Craig	1						
Teresa Palmer	1	Vera Farmiga	1						
Grand Total	32	Zach Gilford	1						
		Grand Total	32						



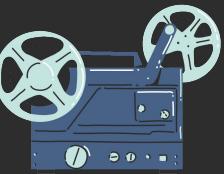
# Pivot Tables cont'd



**Horror performed well throughout the years**

GENRE BY ROI	Year						
		2012	2013	2014	2015	2016	Median
Horror	17696.53%	7994.30%	12943.78%	5478.64%	6809.54%	7994.30%	
Action	8457.61%	6306.69%	4656.94%	3561.50%	3992.40%	4656.94%	
Comedy	7232.09%	5497.06%	5777.97%	4335.42%	1707.01%	5497.06%	
Drama	4663.32%	1586.29%	9780.43%	5996.43%	1977.80%	4663.32%	
Sci-Fi	1165.31%	3171.24%	3416.23%	1989.43%	893.45%	1989.43%	
Thriller	2770.23%	1490.00%	868.57%	1722.62%	1202.41%	1490.00%	
Adventure	1792.29%	879.18%	695.88%	1469.41%	1121.60%	1121.60%	
Religious			3974.83%		370.00%	21.72416667	
Romance	768.62%		108.57%	1492.32%	501.89%	635.25%	
Family	303.79%	251.67%	2093.67%			303.79%	
Biography	-58.50%	342.92%	917.67%	1110.32%	-24.66%	342.92%	
Crime	228.22%	968.18%	331.98%	288.13%	23.75%	288.13%	
Fantasy	1106.86%	0.17%	20.21%	476.85%	-10.87%	20.21%	
Mystery		427.69%	432.38%	149.09%		4 276923077	
Animation			461.14%	334.61%		3.978778259	
Documentary		585.00%				5.85	
Musical	603.46%	-17.50%		-54.00%		-17.50%	

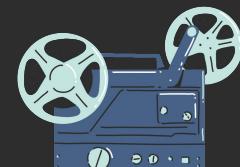
SUM of ROI	Year	2012	2013	2014	2015	2016	Median
Month							
January	10080.00%	1012.00%	427.14%		2064.67%	1538.33%	
July			6503.00%	3.85%	3425.39%	3425.39%	
September		3138.00%	3815.64%	1870.00%		3138.00%	
October	2651.50%		1972.00%	716.82%		1972.00%	
May	3620.00%	-65.52%		172.57%	172.50%	172.54%	
August	460.71%	2580.00%	0.00%	429.00%		444.86%	
February	756.67%	654.29%		1063.64%		756.67%	
June	16.98%			927.27%	1146.99%	927.27%	
April	121.67%	473.53%				297.60%	
December			226.00%	310.00%		268.00%	
March		202.00%				202.00%	
November	-11.00%			-14.50%		-12.75%	



# Process of movie classification

Questions I had to ask before starting:

- How many movie types are really out there?
- What can I do to add a new column to this table containing this information?
- Will this process be by hand? Automated?
- Where will this information come from?



```
13 // Declare constants using iota for auto-incremented values
14 const (
15     REMAKE string = "Remake"
16     SERIES string = "Series"
17     SEQUEL string = "Sequel"
18     PREQUEL string = "Prequel"
19     FRANCHISE string = "Franchise"
20     STAND_ALONE string = "Stand Alone"
21 )
22 func getMovieType(paragraphs []string) string{
23     for _, paragraph := range paragraphs{
24         if strings.Contains(strings.ToLower(paragraph), strings.ToLower(REMAKE)) {
25             return REMAKE
26         }
27
28         if strings.Contains(strings.ToLower(paragraph), strings.ToLower(SERIES)) ||
29             strings.Contains(strings.ToLower(paragraph), strings.ToLower(FRANCHISE)) ||
30             strings.Contains(strings.ToLower(paragraph), strings.ToLower(PREQUEL)) ||
31             strings.Contains(strings.ToLower(paragraph), strings.ToLower(SEQUEL)) {
32             return FRANCHISE
33         }
34     }
35
36     return STAND_ALONE
37 }
```





# Movie Type Pivot Tables



Movie Type	Budget	Box Office	Budget per Movie	Box Office Revenue per movie	SUM of ROI
Franchise	\$137,500,000.00	\$1,103,300,000.00	13750000	110330000	17906.34%
Remake	\$35,000,000.00	\$95,400,000.00	35000000	95400000	172.57%
Stand Alone	\$459,100,000.00	\$2,054,100,000.00	14346875	64190625	32843.89%
<b>Grand Total</b>	<b>\$631,600,000.00</b>	<b>\$3,252,800,000.00</b>	<b>14688372.09</b>	<b>75646511.63</b>	<b>50922.80%</b>

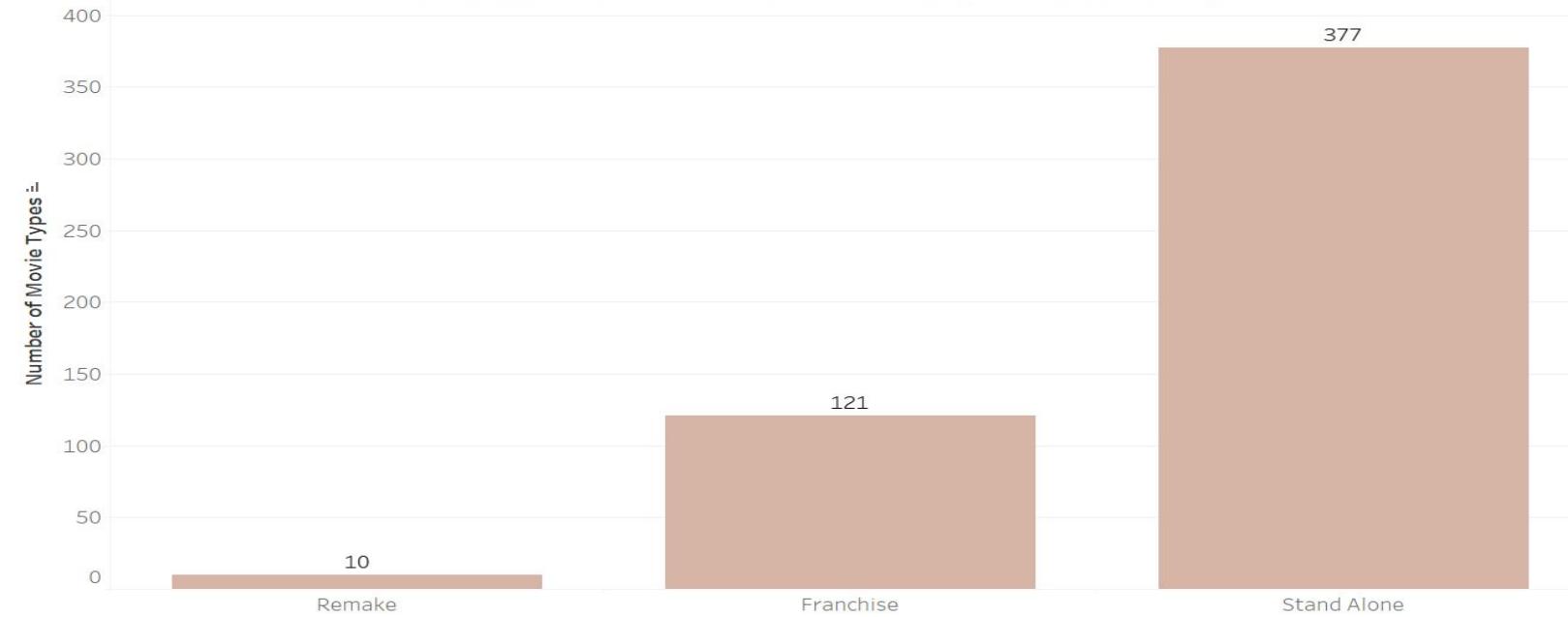
Movie Type	SUM of Budget (\$)	SUM of Box Office Revenue (\$)	SUM of ROI
Franchise	\$137,500,000.00	\$1,103,300,000.00	17906.34%
Remake	\$35,000,000.00	\$95,400,000.00	172.57%
Stand Alone	\$459,100,000.00	\$2,054,100,000.00	32843.89%
<b>Grand Total</b>	<b>\$631,600,000.00</b>	<b>\$3,252,800,000.00</b>	<b>50922.80%</b>



# \*Approximate\* # of Movie types

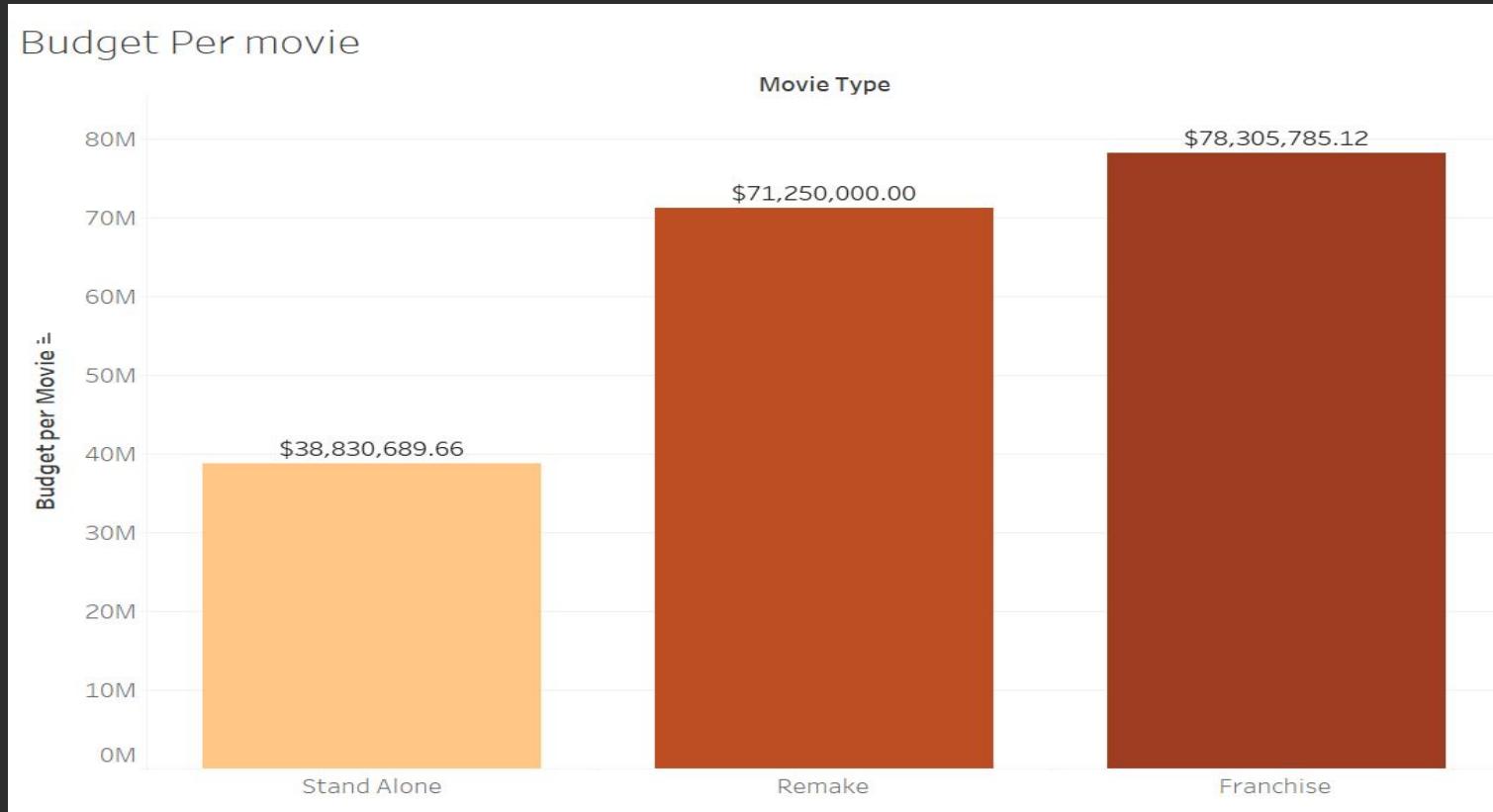
Number of Movie Types

Movie Type (Data Visualization Movie Data Starter Project - Number of movie type.csv)





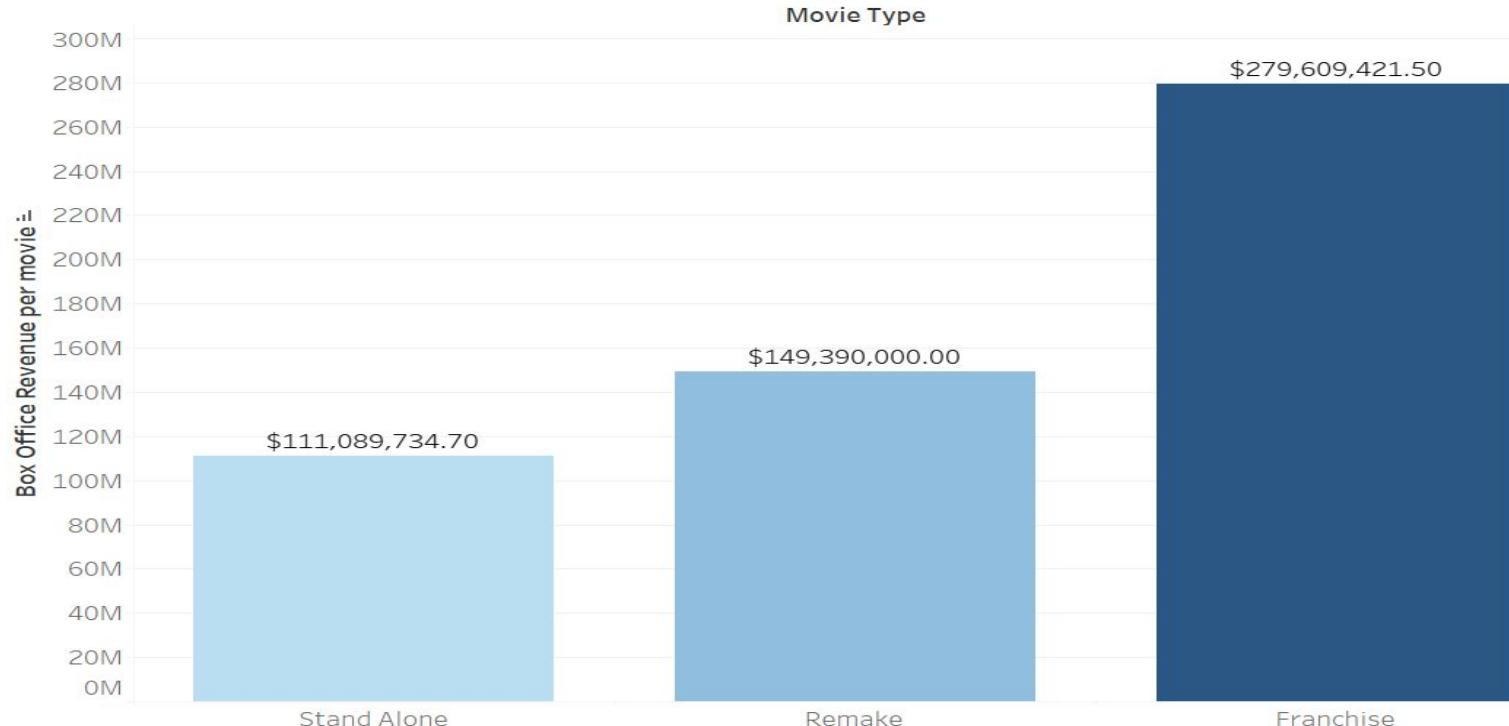
# Budget Per Movie (\$)





# Box Office Per Movie (\$)

Box Office per Movie

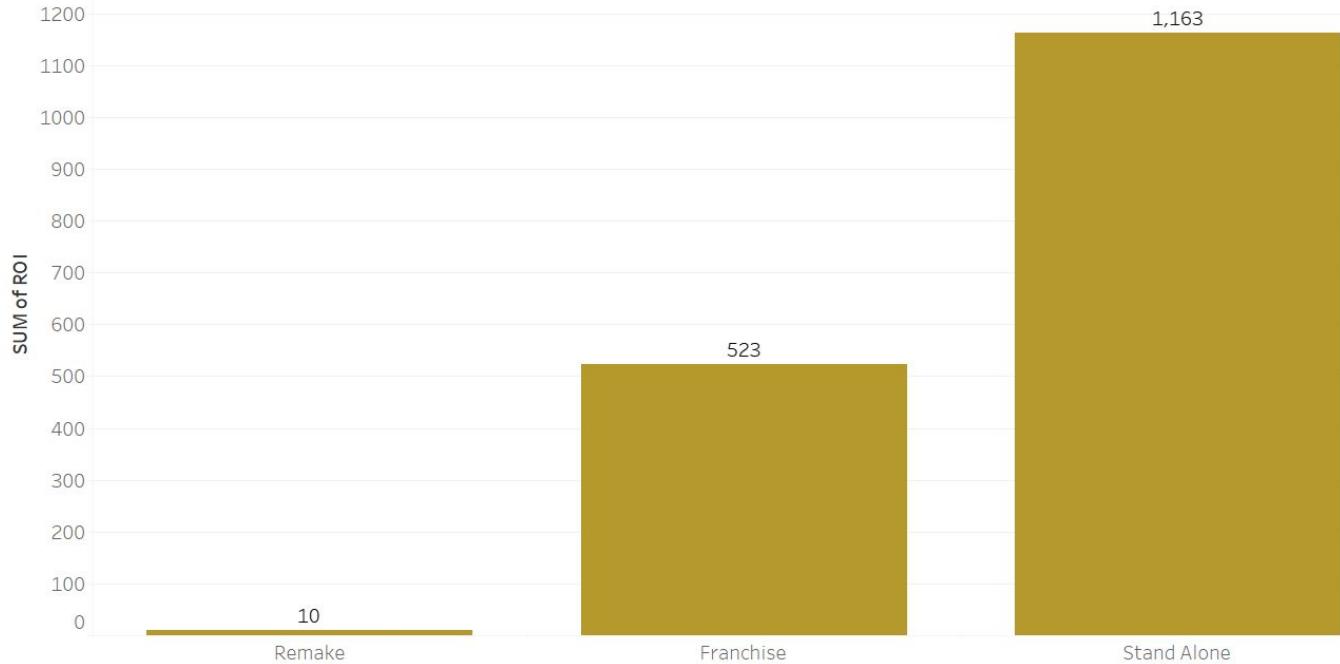




# Best performing movie type

Box office and ROI by Movie type

Movie Type (Data Visualization Movie Data Starter Project - Budget and ROI by movie type.csv)



💡 Stand alone movie perform the best.

⭐️ Stand Alone Movie Recommended

# Recommendations



## Genre Recommendations

**Horror:** High ROI with low budgets and unsaturated market; ideal for big returns with minimal investment.

**Other Genres:** Consistent and stable, but less explosive potential

**Crime:** Average budget, below-average returns—best to avoid.

## Release date Recommendation

**Months:** July, Sept/Oct, January

## Classification Recommendation

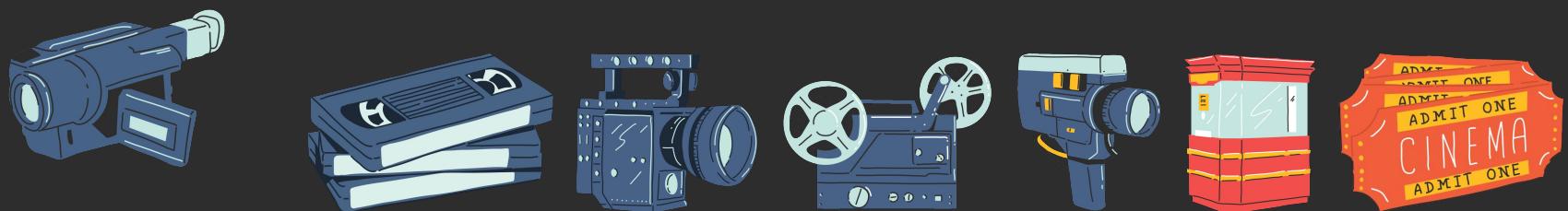
Horror **Stand Alone** movies have generated the highest total ROI, making them the safest pick.

**Stand Alone** movies across the board performed the best as well.



THANK YOU  
FOR  
LISTENING!

# RESOURCE



PAGE

# Icons (continued)

