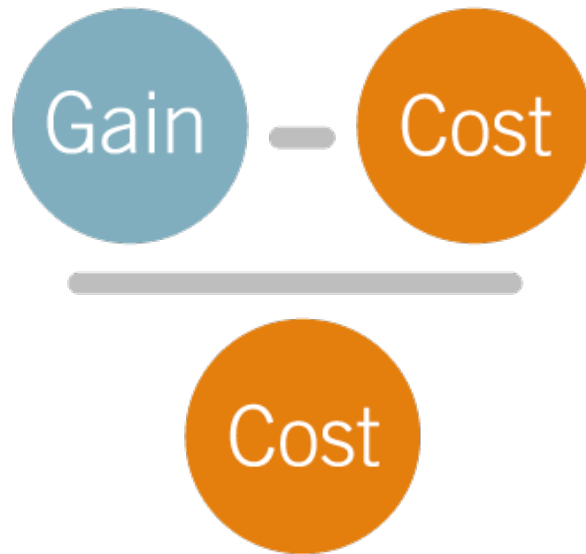


Low-Budget Movies with High ROIs

Project Luther

Definitions

- Low-Budget : Movies with budgets < \$2M
- Return on Investment (ROI) :



The diagram illustrates the Return on Investment (ROI) formula. It features a blue circle labeled 'Gain' on the left, a minus sign in the center, and an orange circle labeled 'Cost' on the right. A horizontal line is drawn below these three elements. Below the line is a single orange circle labeled 'Cost'.

$$\frac{\text{Gain} - \text{Cost}}{\text{Cost}}$$

Data

- Let's look at movies with enormous ROI
- What makes these movies special?
 - Possible Features:
 - Critic Scores
 - Runtime
 - Widest Release
 - Genre
 - MPAA Rating
 - Distributor

Analysis

- Let's try a linear regression model to associate ROI with these features
- Potential Model:

ROI ~ Critic Scores + Runtime + Release Size + Genre + MPAA + Distributor

- Data Acquired via Webscraping:
 - ROI : The-Numbers.com
 - Critic Scores : RottenTomatoes.com
 - Others: BoxOfficeMojo.com

Results

- Most viable model:

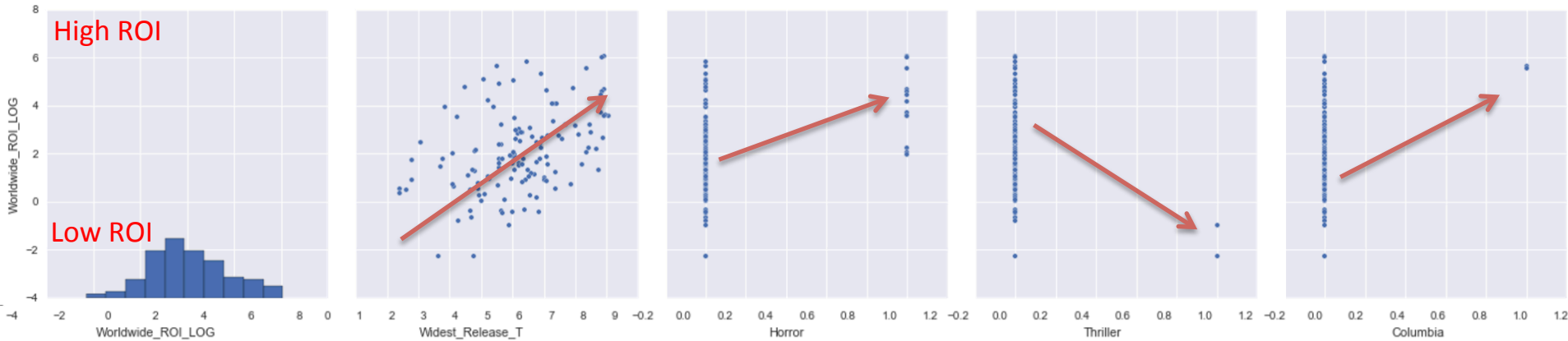
ROI ~ Release Size + Genre + Distributor

Worldwide_ROI_LOG ~ Widest_Release_T + Horror + Thriller + Columbia
(continuous) (continuous) (categorical)

- Achieved an Adjusted R² : 0.31
- All variables in the final model are statistically significant (significance level: 0.03)
- But what does this mean?

Associations

Worldwide_ROI_Log ~ Widest_Release_T + Horror + Thriller + Columbia



	coef	std err	t	P> t 	[95.0% Conf. Int.]
Intercept	0.2639	0.477	0.553	0.581	-0.680 1.208
Widest_Release_T	0.3292	0.092	3.562	0.001	0.146 0.512
Horror	1.0630	0.470	2.263	0.025	0.134 1.992
Thriller	-3.3137	1.002	-3.308	0.001	-5.296 -1.332
Columbia	2.8778	1.012	2.842	0.005	0.874 4.881

Sample Films

Release_Date	Movie_Title	Worldwide_ROI	Horror
9/25/09	Paranormal Activity	430.5178533	1
7/14/99	The Blair Witch Project	412.8333333	1
10/1/68	Night of the Living Dead	262.1578947	1
8/6/04	Open Water	110.037282	1
1/6/12	The Devil Inside	100.758489	1
10/29/04	Saw	85.56668917	1
4/1/11	Insidious	65.58059067	1
5/25/12	Chernobyl Diaries	41.411721	1
8/27/10	The Last Exorcism	37.98105556	1
4/20/79	Dawn of the Dead	35.66666667	1
4/10/53	House of Wax	35.17021277	1
3/13/15	It Follows	9.324887	1
9/25/15	The Green Inferno	7.801867	1
2/22/08	The Signal	7.12598	1

Release_Date	Movie_Title	Worldwide_ROI	Columbia
2/26/93	El Mariachi	290.704	1
10/1/68	Night of the Living Dead	262.157	1

Release_Date	Movie_Title	Worldwide_ROI	Thriller
9/18/15	Captive	0.378888	1
8/9/02	Pandora's Box	0.1024375	1

Movie_Title	Worldwide_ROI	Widest_Release
House of Wax	35.17	3111
The Last Exorcism	37.98	2874
Dawn of the Dead	35.67	2748
Paranormal Activity	430.52	2712
Open Water	110.04	2709
Nancy Drew	14.35	2612
The Devil Inside	100.76	2551
The Blair Witch Project	412.83	2538
Saw	85.57	2467
Chernobyl Diaries	41.41	2433

Summary

- Should you invest based on this model?
 - No
- But if you want to, aim at making a Horror film without any Thrill, that is distributed by Columbia with the widest release possible
- In the meantime, let's dig for more data or re-evaluate how we approach our question

Thank You

Additional Material

- Utilized Variable Selection Techniques to eliminate variables and optimize Adjusted R^2
 - Stepwise Variable Selection (SAS)
- Running a CVLasso (SciKit Learn) on this Model
 - Yields a $\lambda = 0.001$
 - Our variables do not warrant a penalty