



The background of the slide features a dark, moody aesthetic. Two large, metallic film reels are positioned in the upper half, slightly overlapping. A strip of film with visible sprocket holes is draped across the lower half, winding around the base of the reels. The lighting is soft, highlighting the textures of the reels and the film strip.

Movieland Production Company

Analysis to predict Gross revenue of a movie

Data from www.boxofficemojo.com

Top 200 ranked domestic movies from years 2015 to 2022

			<h3>The Batman</h3> <p>When the Riddler, a sadistic serial killer, begins murdering key political figures in Gotham, Batman is forced to investigate the city's hidden corruption and question his family's involvement.</p>		
 Title Summary			Original Release ▾ Domestic ▾		
Grosses DOMESTIC (50.2%) \$302,863,177 INTERNATIONAL (49.8%) \$300,400,000 WORLDWIDE \$603,263,177			Distributor	Warner Bros. See full company information	
			Opening	\$134,008,624 4,417 theaters	
			Release Date	Mar 4, 2022	
			MPAA	PG-13	
			Running Time	2 hr 56 min	
			Genres	Action Crime Drama Mystery	
			In Release	18 days/2 weeks	
			Widest Release	4,417 theaters	

Sample page

Features for Target Variable: Gross

Extracted features

- Genres
- MPAA rating
- Duration
- Number of theatres
- Number of days released
- Domestic release ratio
- Distributor
- Release date
- Opening week gross
- Movie title

Engineered features

- 5 Genres or less
- PG-13
- PG
- R
- Duration
- Number of theatres
- Number of days released
- Domestic release ratio
- Warner Bros.
- Other distributors
- Weekend release
- Opening

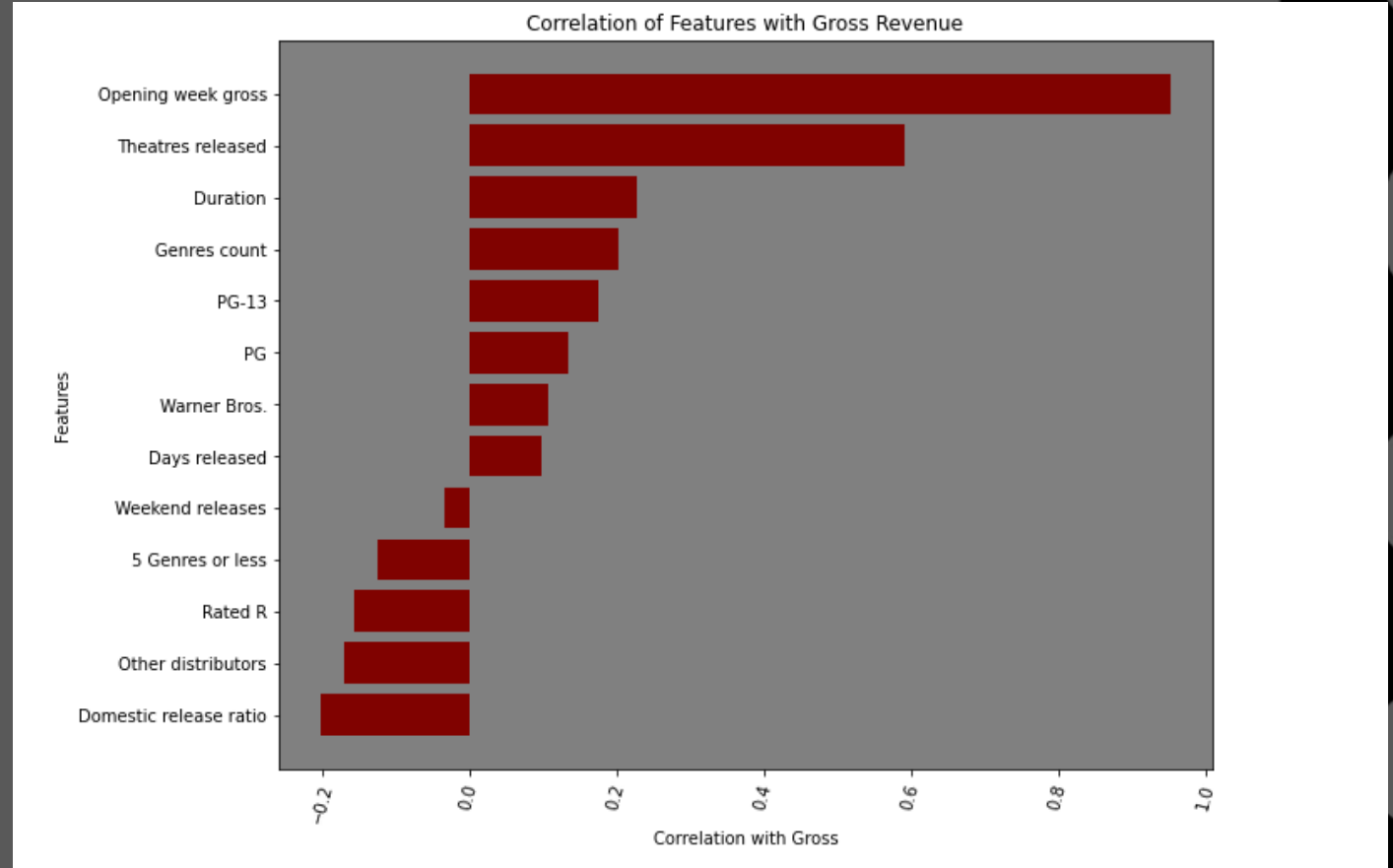
Correlation between features and Gross. revenue

Positively Correlated (top 3)

- Opening
- Theaters released
- Duration

Negatively correlated (bottom 3)

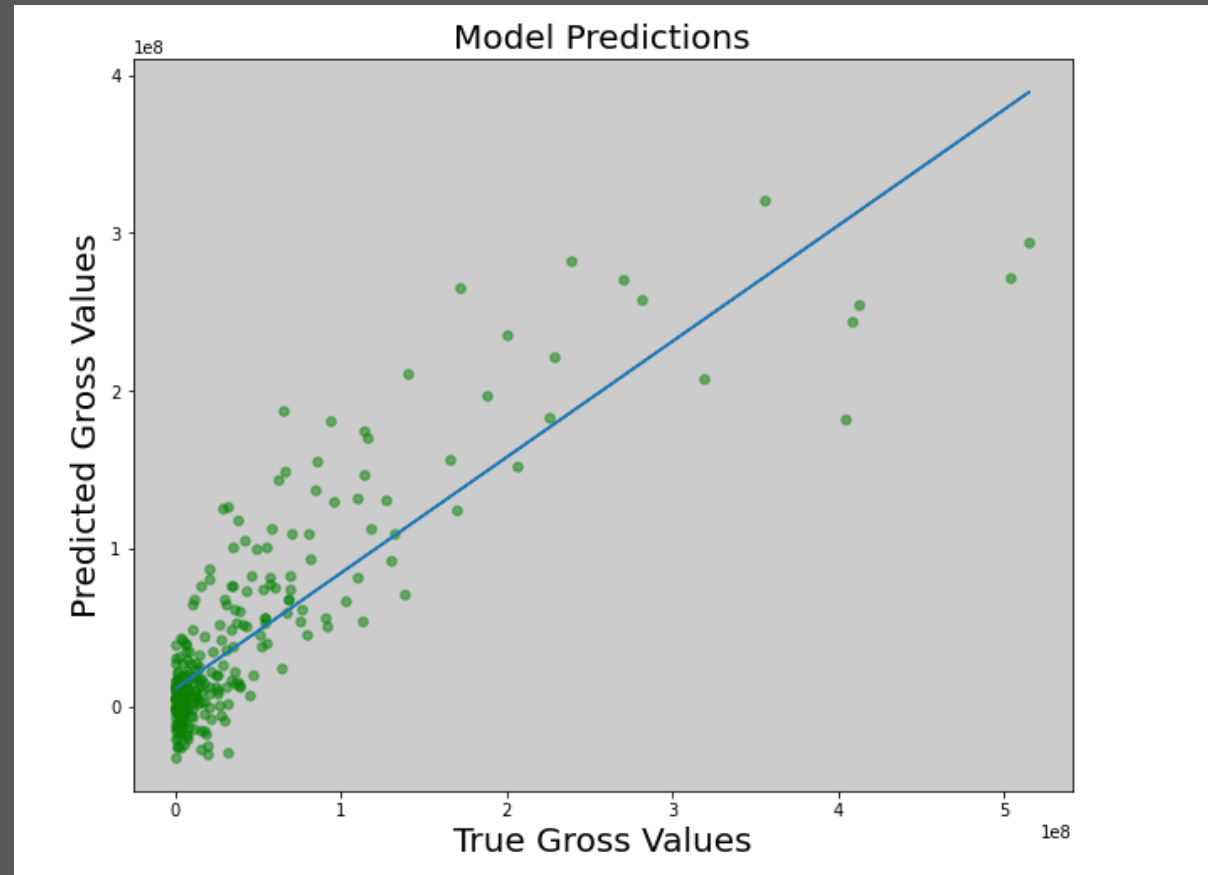
- Domestic release ratio
 - Other distributors
 - R Rated Movies
- * Dropped Opening week gross as it is a subset of Gross revenue and that is why the high corr.



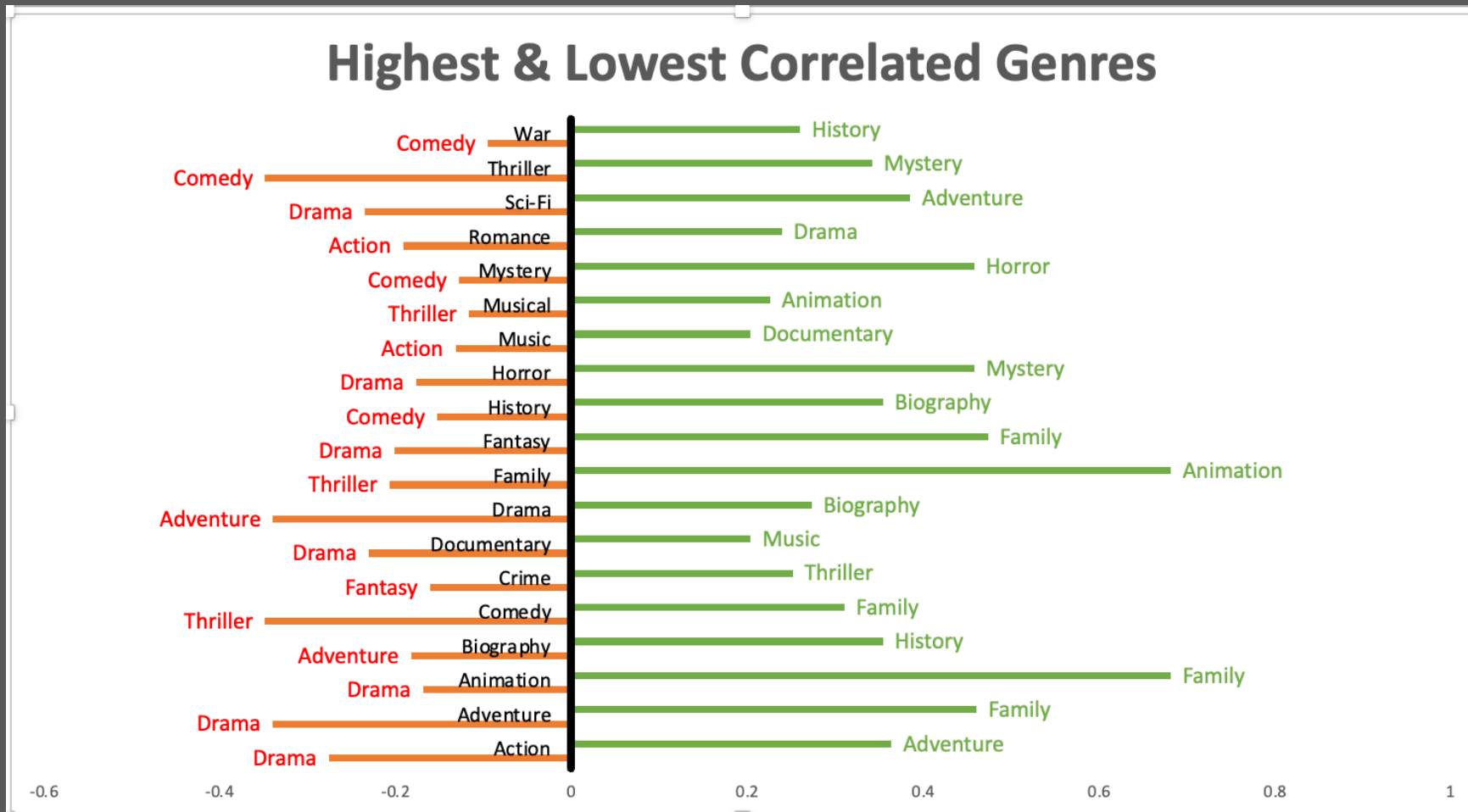
Model Predictions

Model Performance

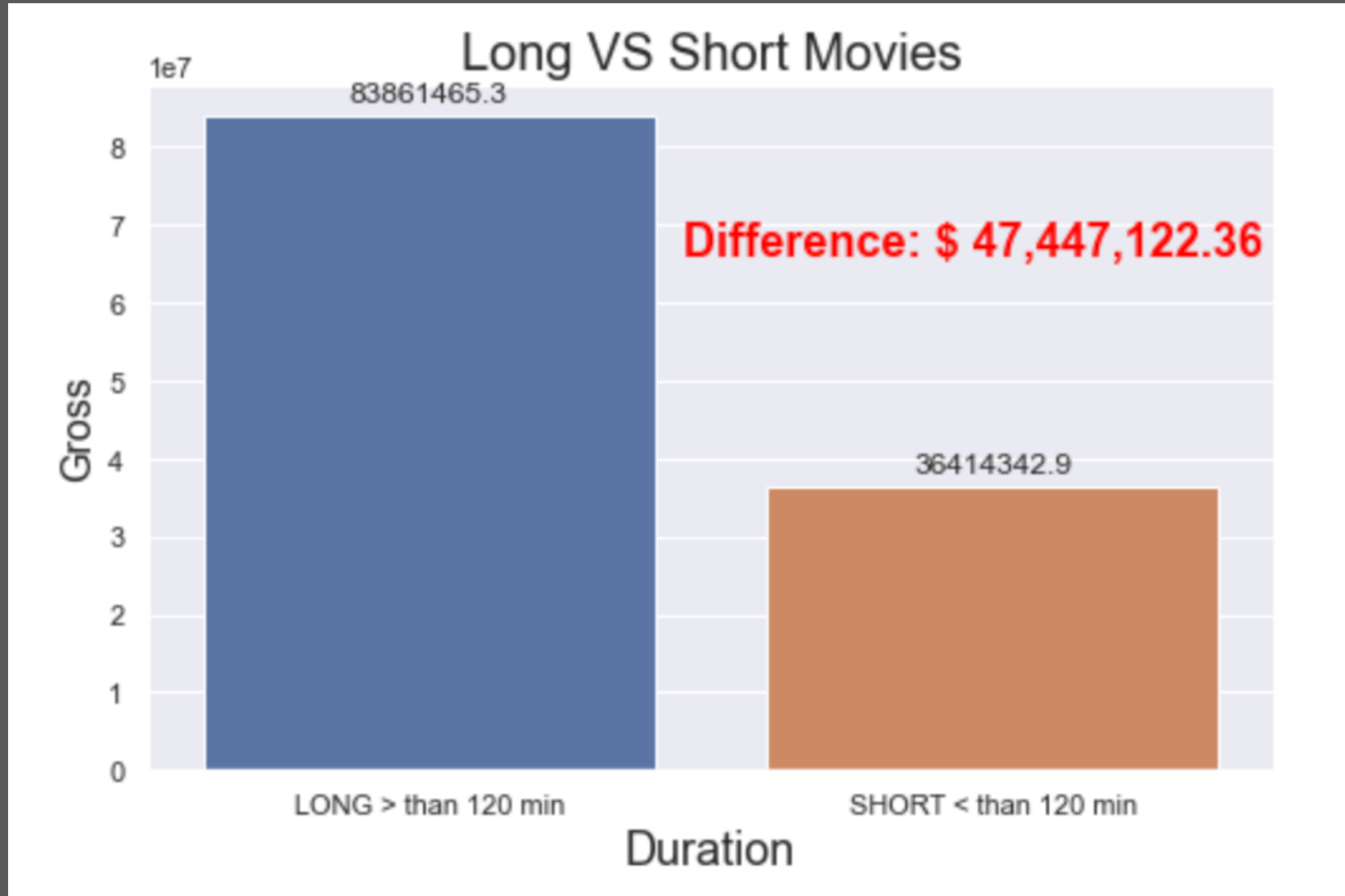
- Linear Regression with Polynomial transformation.
- R Square value of 0.73
- Predictions on 'Test data'



Interesting Insights



Interesting Insights



Future Work

- Incorporate Budget and do Profit analysis
- Movie recommendation based on Genre and other features.
- Add more features like IMDB Rating, Directors, Lead actors and Cast.

Features Explained

- 5 Genres or less: Each movie is associated with 1 or more genre. Some have up to 9 genres. So made a column that will display binary values of 1 or 0 based on the number of genres. If 5 or less, then it has a value of 1.
- Domestic release ratio: Proportion of released made with the country Vs made outside.
- Release date: is converted to know if it was released on a weekday or weekend and assigned binary values accordingly. This feature is now called weekend release.

The background of the slide is a photograph of a theater interior. In the foreground, several rows of plush red seats are visible, receding into the distance. The stage is at the far end, covered by a heavy, vertically pleated golden-brown curtain. The lighting is warm and focused on the stage area, creating a sense of anticipation.

Questions