Movie Runtime Regression

Introduction

- Does a movie's runtime have any relationship to its box office performance?
- Should the studio make longer or shorter movies?
- What other factors might change this relationship?

Methodology

- Collect data from boxofficemojo.com
- Top 200 highest grossing movies in each year from 2015-2019
- Build regression model to predict box office
- Test model
- Report coefficients for runtime, and interactions with runtime

Model

Lasso Regression Target

Worldwide Gross Box Office

Independent Variables

Budget

$Budget^2$

Genres (Action & Comedy)

Distributor (Disney & Universal)

Independent Variables (Runtime)

Runtime

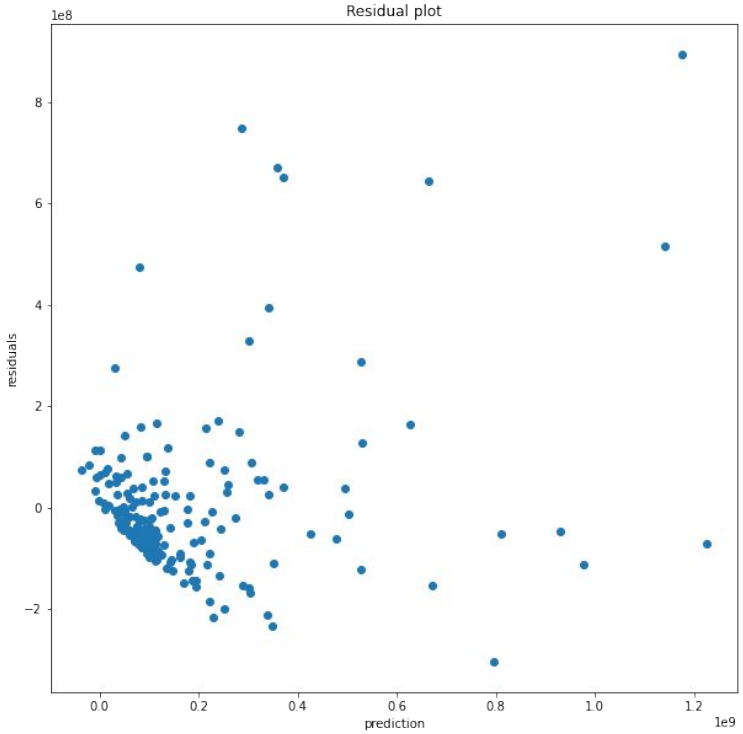
Comedy Runtime

Action Runtime

Performance

Mean Absolute Error
$95,724,303

Performance



Results

Relationship between Runtime and Box Office Gross

$2,463,203

Increase in Action Movie Gross for each marginal minute of runtime

$879,826

Increase in Comedy Gross for each marginal minute of runtime

$1,637,976

ncrease in other genre's Gross for each marginal minute of runtime

Conclusion

- Movies of all genres have a positive correlation between runtime and worldwide box office
- This correlation is weakest in comedies
- We don’t know if there is causation

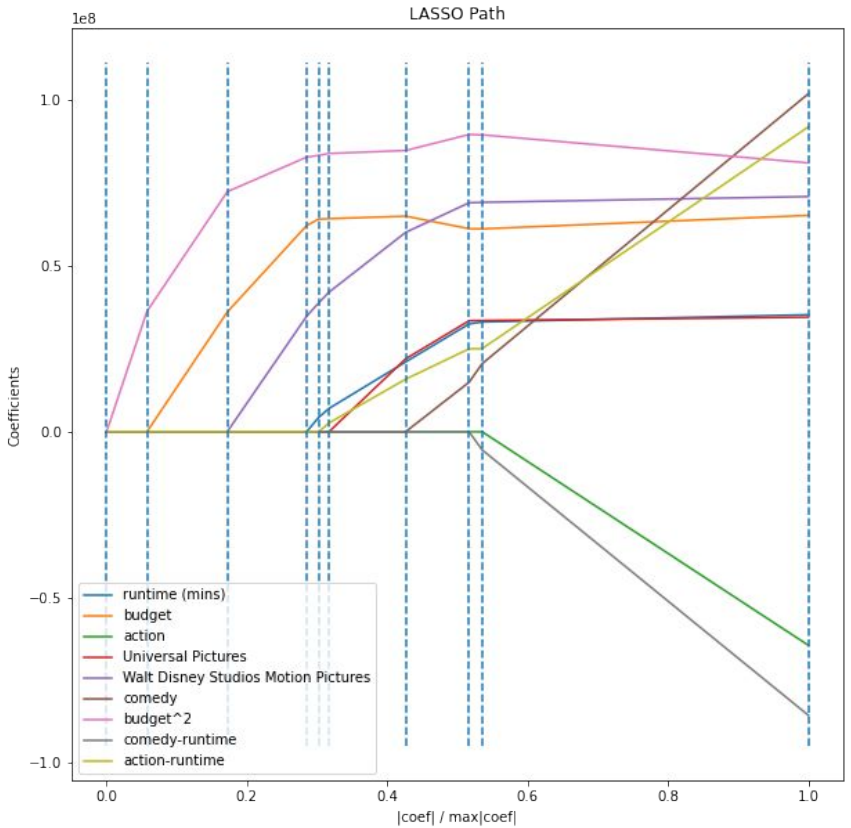
Appendix (Regression Coefficients)

Coefficients scaled to one standard deviation:

Runtime 33820364
Budget 62502833
Genre Action -20191724
Genre Comedy 45945968
Distributor Universal Pictures 33898144
Distributor Walt Disney Studios Motion Pictures 69807308
Budget² 87026541
Comedy x runtime -30435923
Action x runtime 46077369

R^2 on test data: .7058

Appendix (Lasso Path)



Appendix

Data from boxofficemojo.com

Slide template from SlidesCarnival