
MOVIES

Domestic box office
prediction



INTRODUCTION

\$42.2 BILLION

- Global box office in 2019
-

METHODOLOGY

PREDICT TOTAL DOMESTIC BOX OFFICE GROSS

- Features:
 - Budget
 - Running time
 - Number of theaters
 - Rating
 - Genres
-

DATA

- Web-scraped from BoxOfficeMojo.com
- Top grossing movies for 2012 – 2019 years
- 917 movies

Box Office Mojo

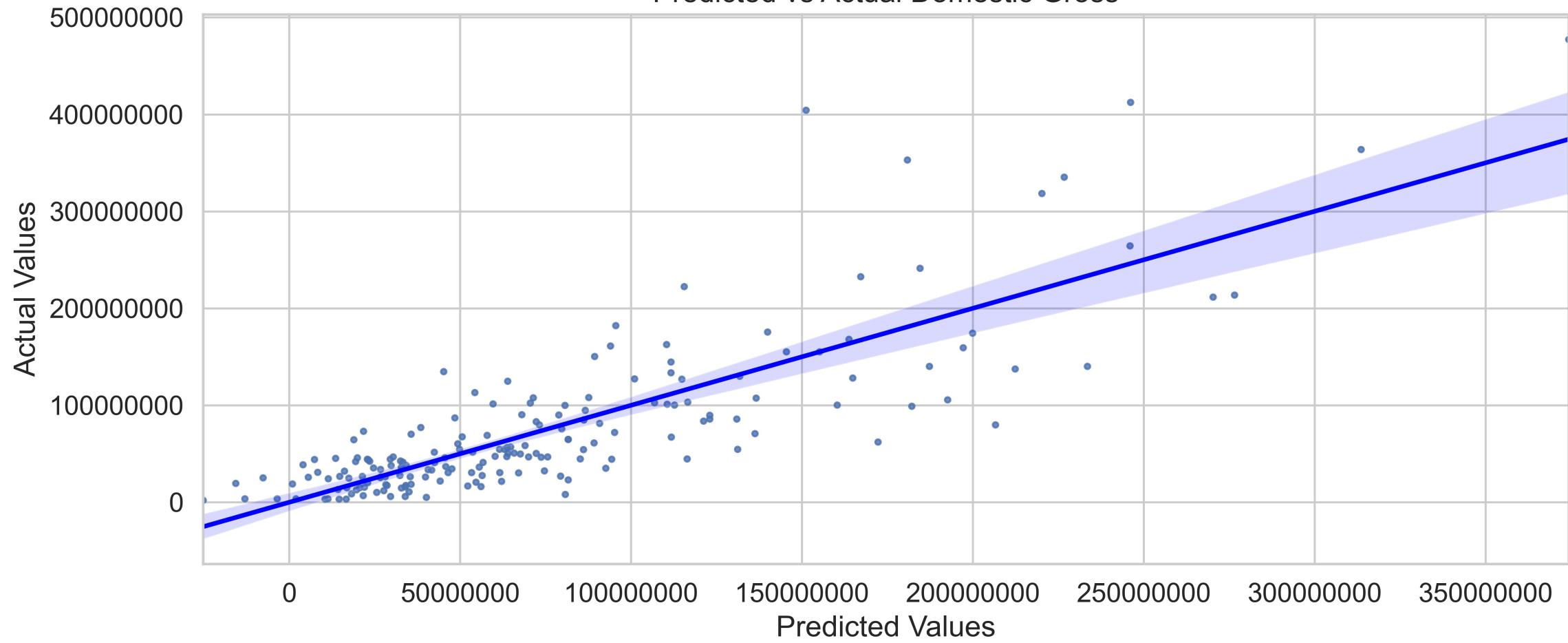
RESULTS

KEY FEATURES

- Budget (invest more money into your movie)
 - Theaters (get your movie to as many theaters as you can)
 - PG – 13
 - Make thrillers, comedies, or romances
-

R squared = 0.679

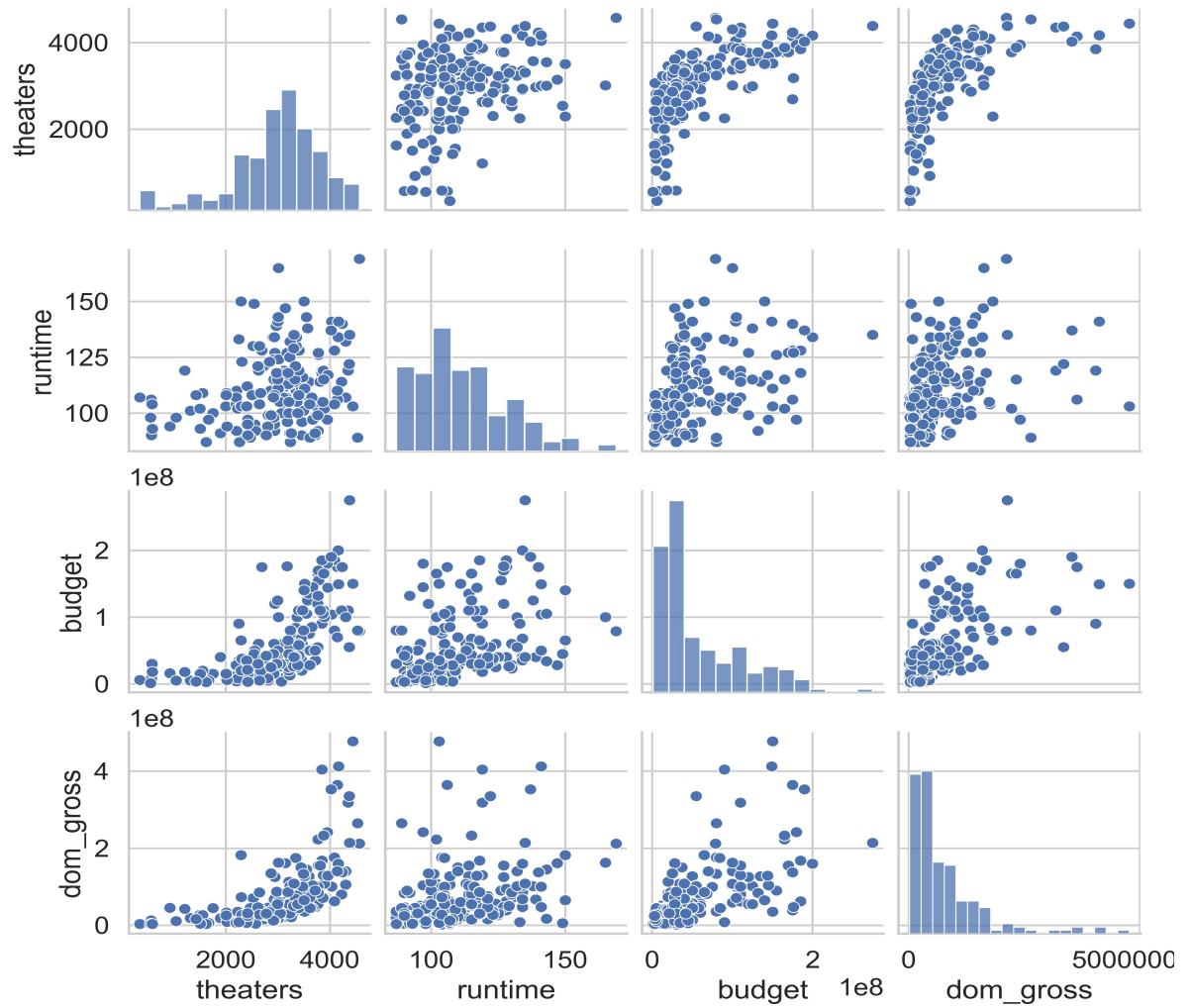
Predicted vs Actual Domestic Gross



FUTURE WORK

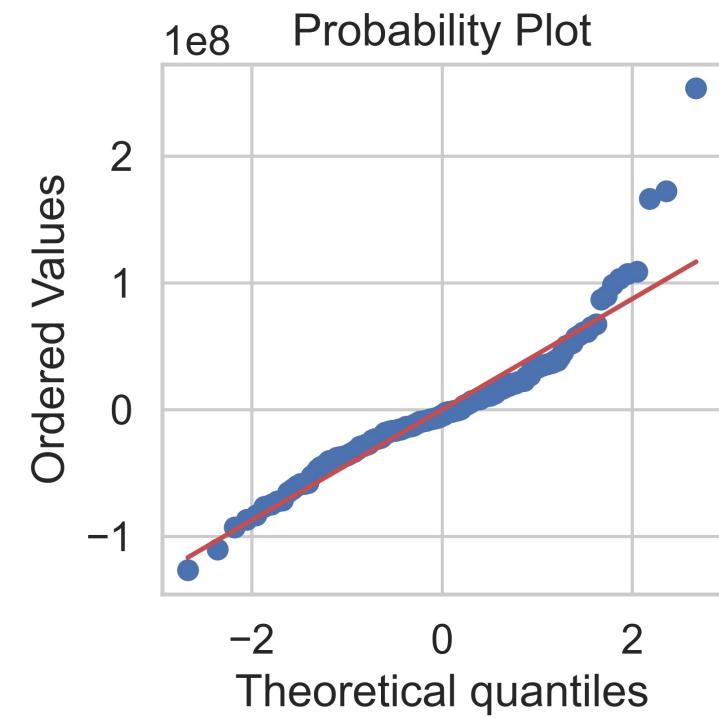
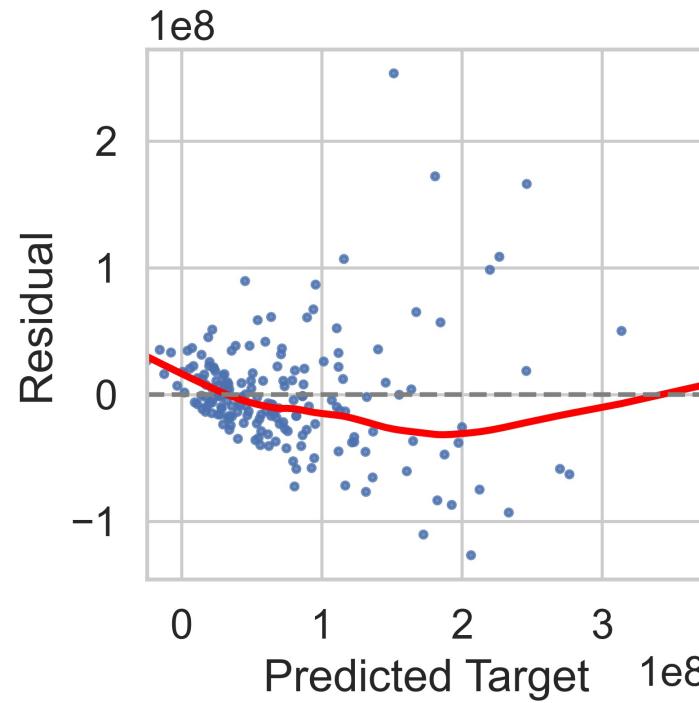
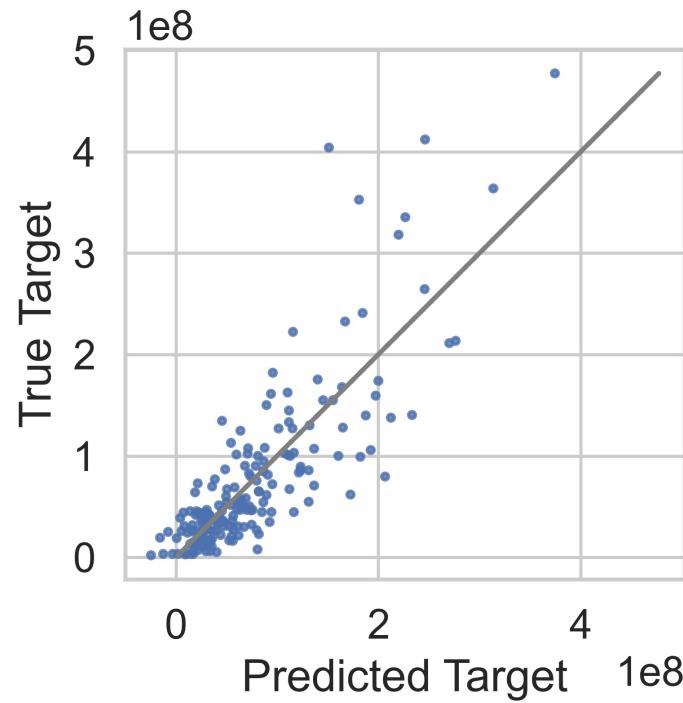
- Add more features to the model to investigate what other factors will play a role in box office gross
 - Add time series analysis to identify how changes in release date affects the target
 - Add data from other resources to the analysis (imdb.com, rottentomatoes.com, etc)
-

APPENDIX A

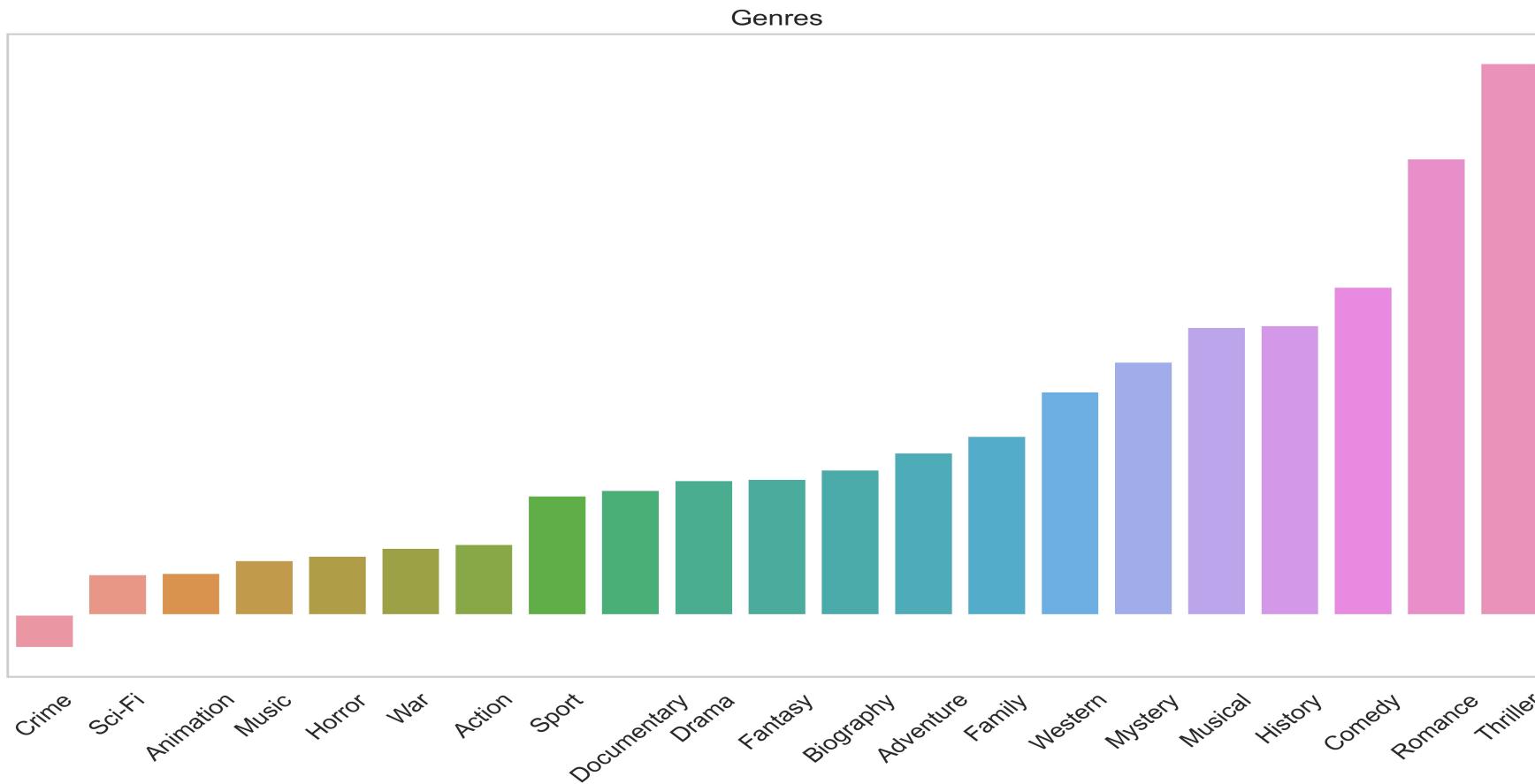


APPENDIX B

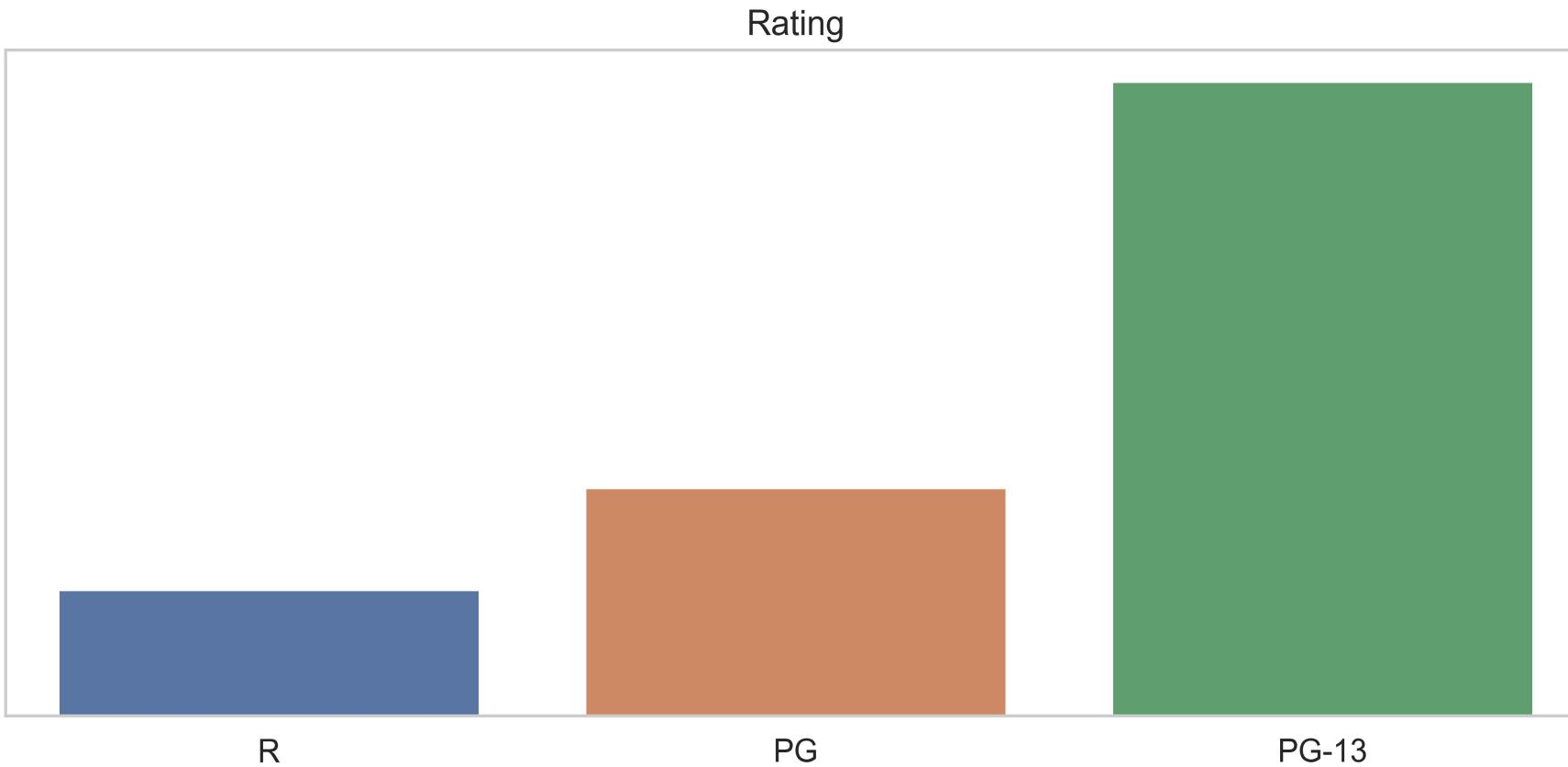
Diagnostic Plots



APPENDIX C



APPENDIX D



APPENDIX E

- Lasso Regularization - alpha = 0.01 R squared = 0.529
 - Ridge Regularization - alpha = 0.01 R squared = 0.533
-

QUESTIONS?
