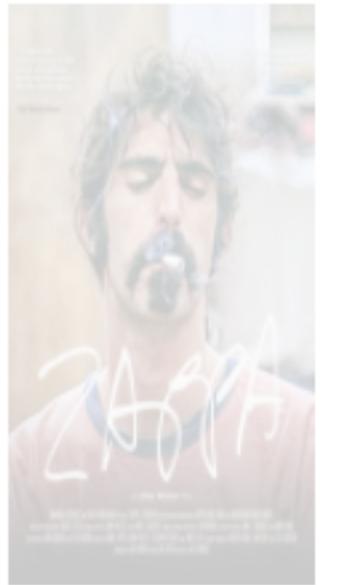


Liam-o-meter: 7.4

IMDB: 6.7



Zappa

Liam-o-meter: 7.4

IMDB: 7.6

Liam-o-meter: 7.4

IMDB: 8.0



Liam-o-meter: 7.4

IMDB: 4.1



Mantra of Rock

Liam-o-meter: 7.4

IMDB: 6.8



Sunset Blvd

Liam-o-meter: 7.4

IMDB: 8.4

MacGowan

Liam-o-meter: 7.4

IMDB: 8.0



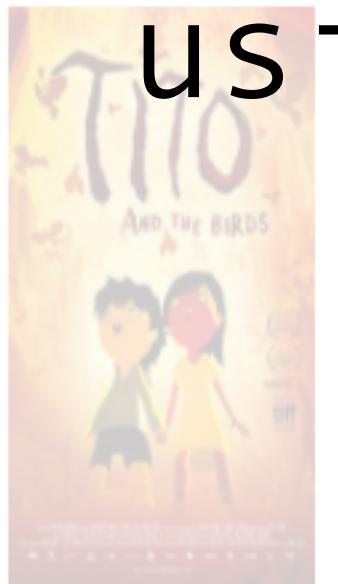
The Dissident

Liam-o-meter: 7.4

IMDB: 8.2

# the liamometer

using machine learning to simulate opinion



Tito and the Birds

Liam-o-meter: 7.4

IMDB: 6.5



Promare

Liam-o-meter: 7.4

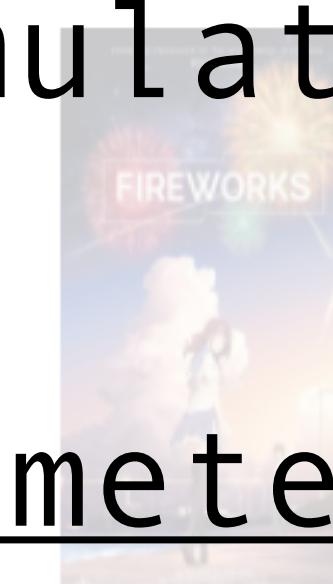
IMDB: 7.2



Liz and the Blue Bird

Liam-o-meter: 7.4

IMDB: 7.2



Fireworks

Liam-o-meter: 7.4

IMDB: 7.8



Burn the Stage: The Movie

Liam-o-meter: 7.4

IMDB: 8.8



22-Jan-2021 (Revised: 31-Mar-2021)

How do we want to rate movies?

# How do we want to rate movies?

---

Active choice  
we make off  
of  
preferences

# Can we simulate opinion?

Review systems as ways of manufacturing opinion, based off of opinions

Goal: rate based off of genre from  
internationally-popular  
distributors



# Roadmap

## How do we want to rate movies?

### EDA

- BoxOfficeMojo
- Review sources:  
IMDB,  
RottenTomatoes,  
Metacritic

Which reviews to  
use as a basis?

### Cleaning



- Remove budget,  
MPAA

Which data to  
ignore?

### Modeling

- Try: 4  
approaches
- LASSO/Ridge CV,  
Q-Q plots,  $R^2$

Which data to  
consider, and how  
to consider it?

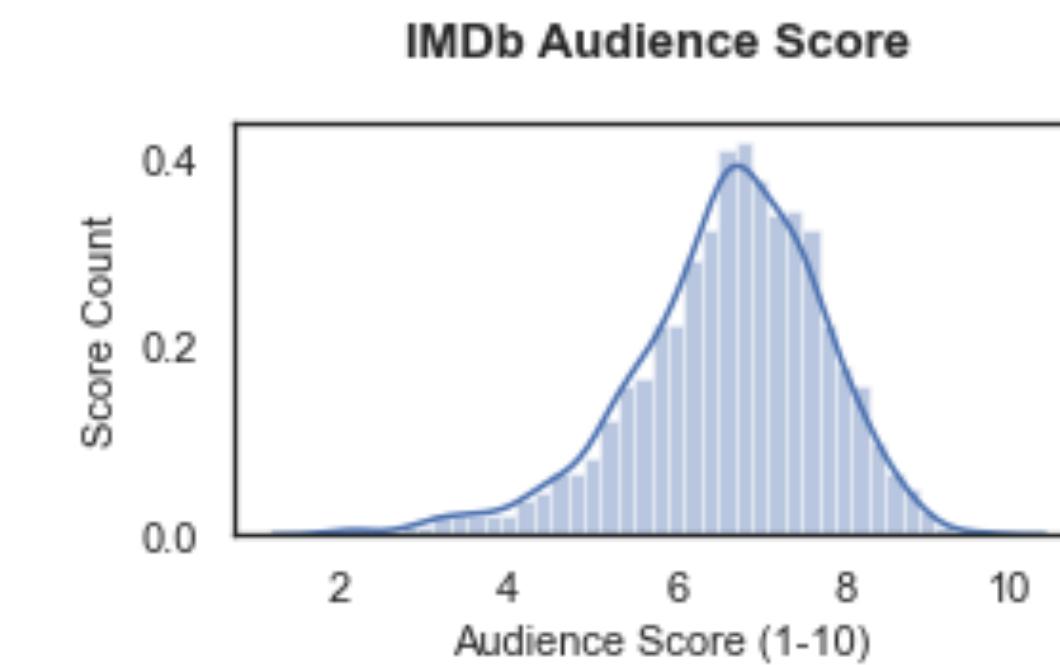
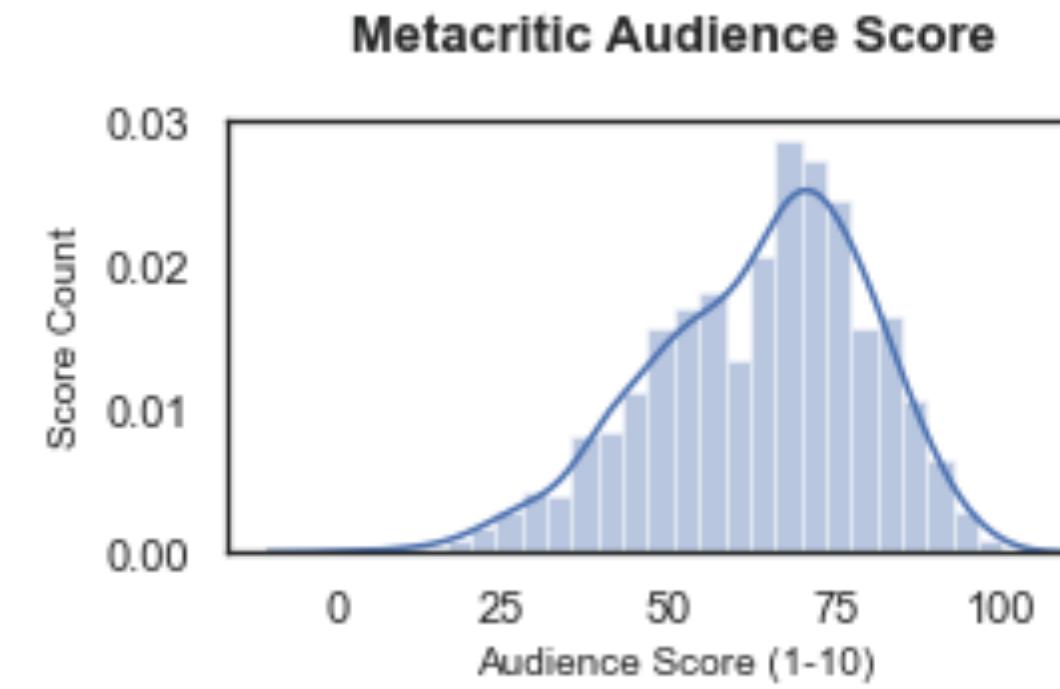
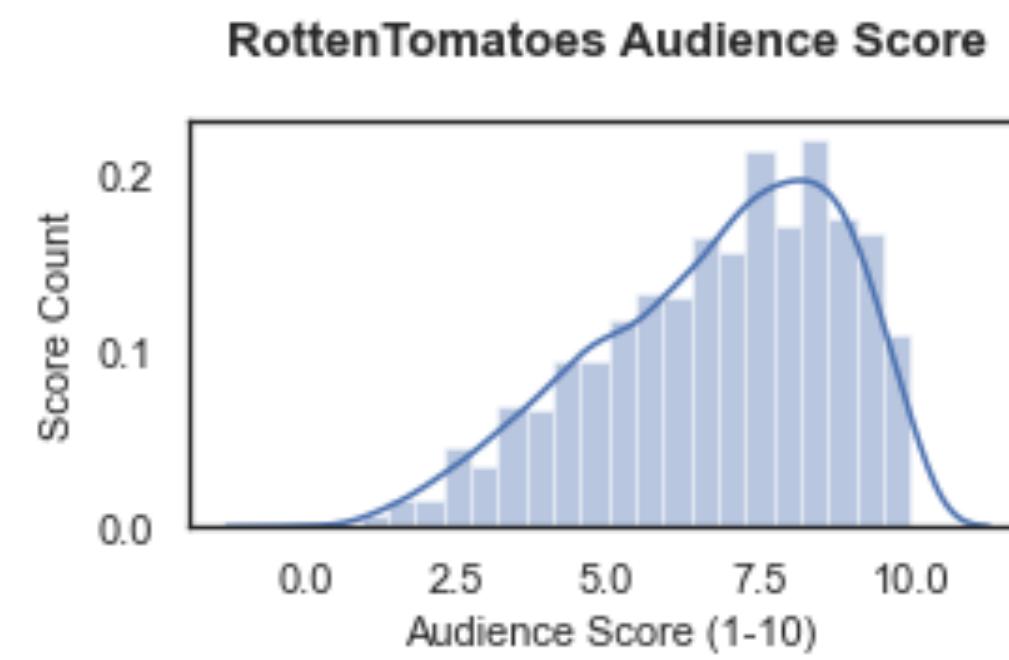
### Visualization

- Flask WebApp

How to use  
manufactured  
opinion?

# EDA

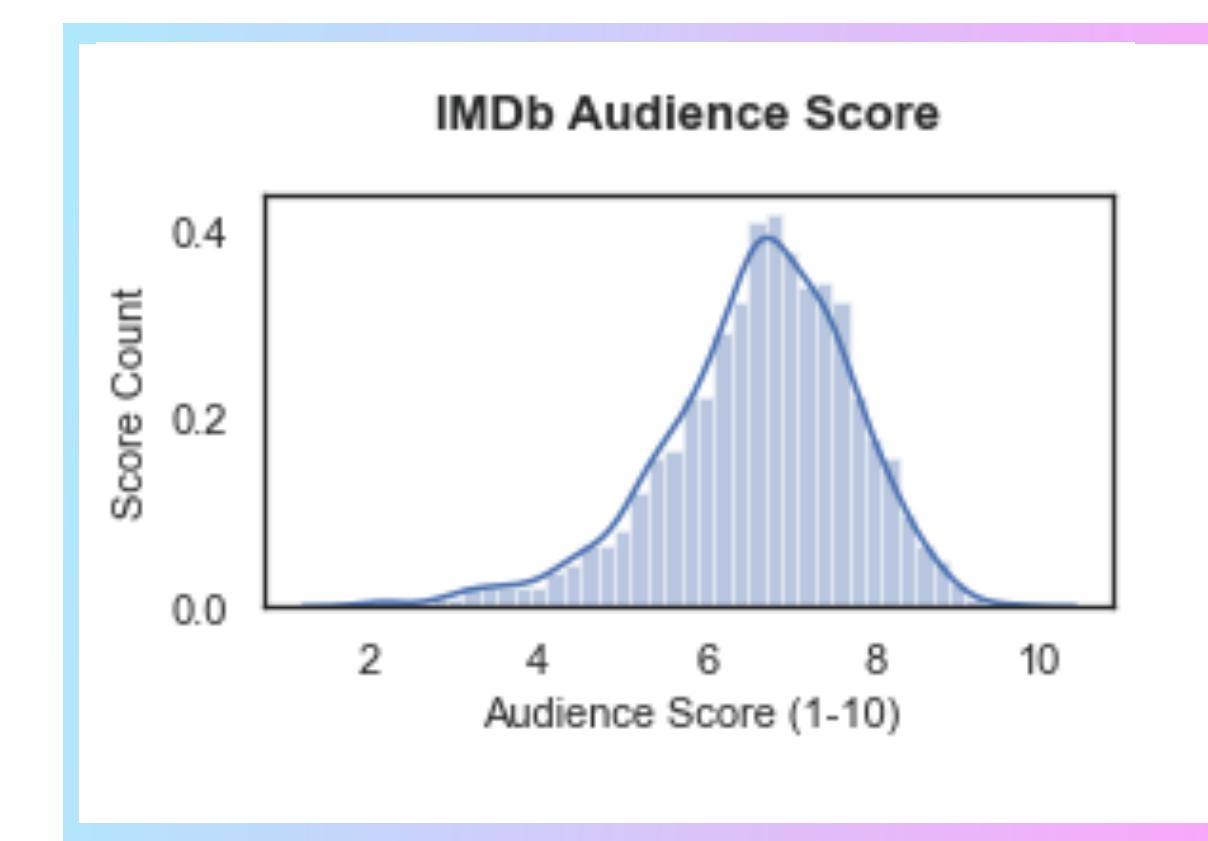
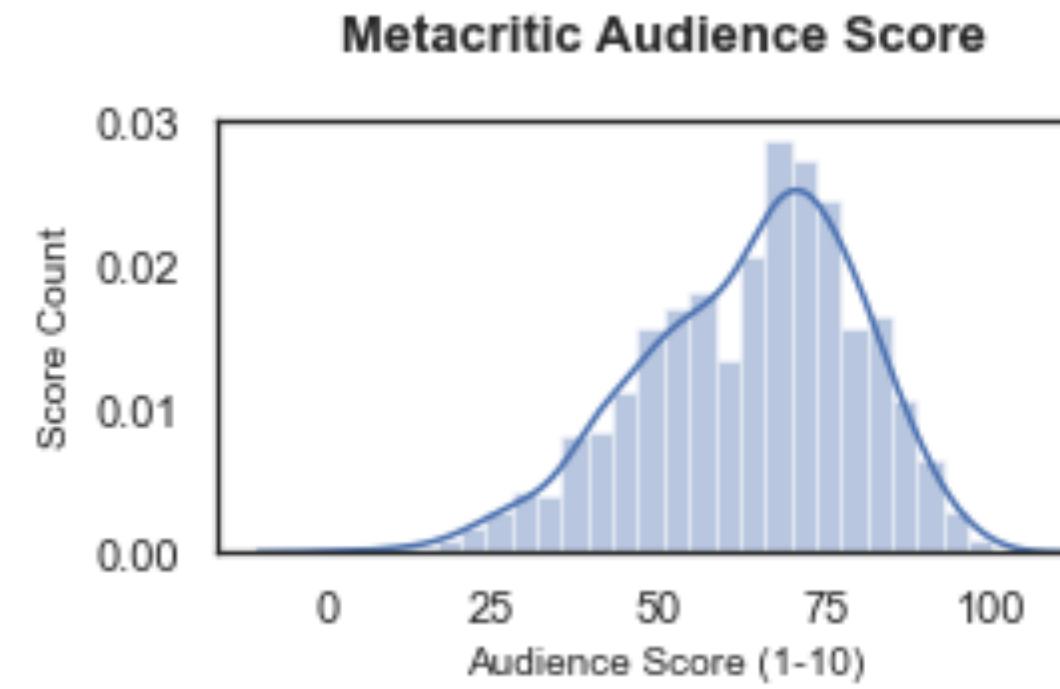
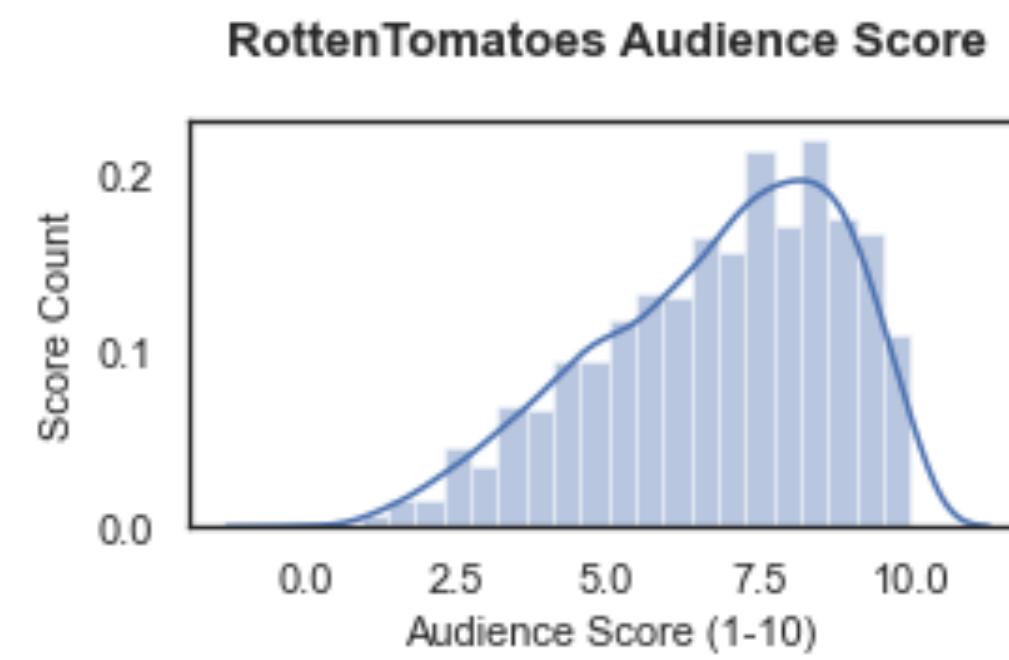
Data scraped using [Scrapy](#) web crawling framework ([tutorial](#))



Tend to be skewed due to “7 is average” mentality

# EDA

Data scraped using [Scrapy](#) web crawling framework ([tutorial](#))

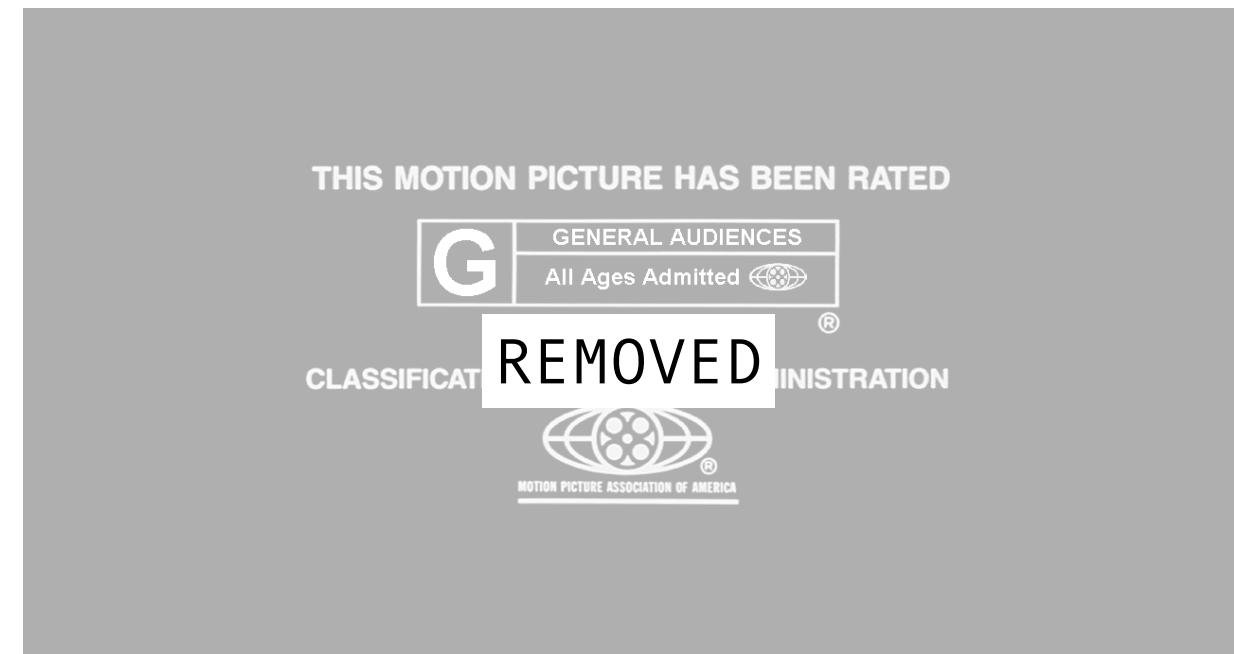


- Rating system is a “% positive” system, i.e. did you like it yes/no

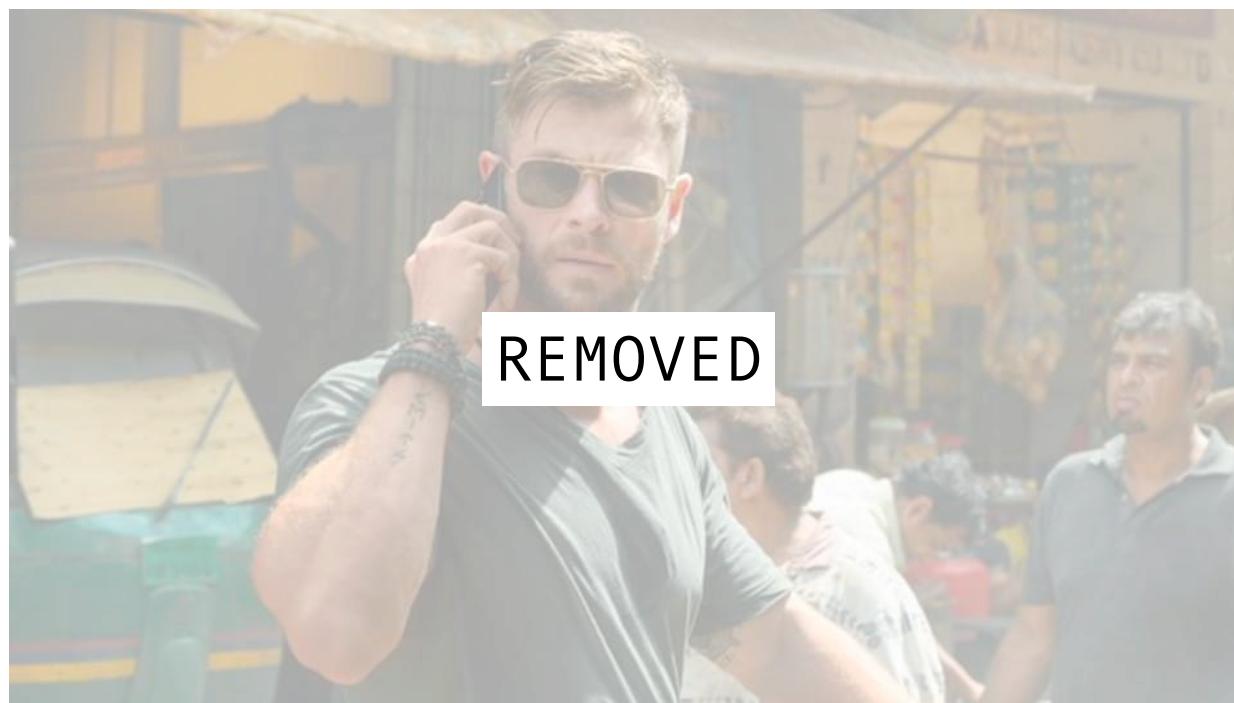
- Most normal
- Rating system is 1-10, not % positive

# EDA -> Cleaning

MPAA removed: we do not care about “PG-13” or “R” these are not interesting ways of reviewing movies (while they still might be valuable in making an accurate prediction model, we are more interested in interpretation)



Budget removed: we do not want to bias towards wealthy studios (fame is boring and with the way the movie industry is, money tends to be concentrated in the production of dumb films)



# EDA -> Cleaning -> Modeling

## Technical terms

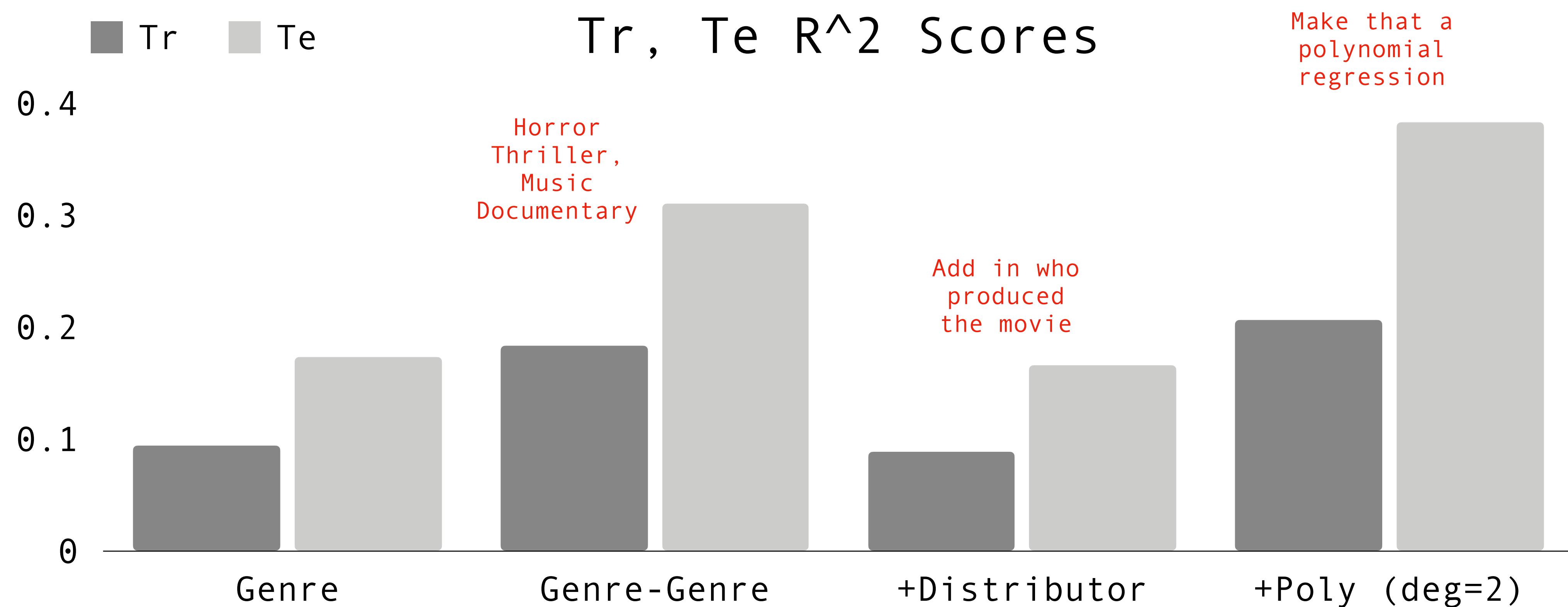
Add-in “domestic-international ratio”  
One-hot-encode genre  
Genre-genre interactions  
Over-represent smaller studios by replacing  
<40 movies a year with “Other”



## Non-technical translation

Favor internationally-successful movies  
Take into account genre  
Capture “anime+war” not just “anime”  
Bias towards smaller studios

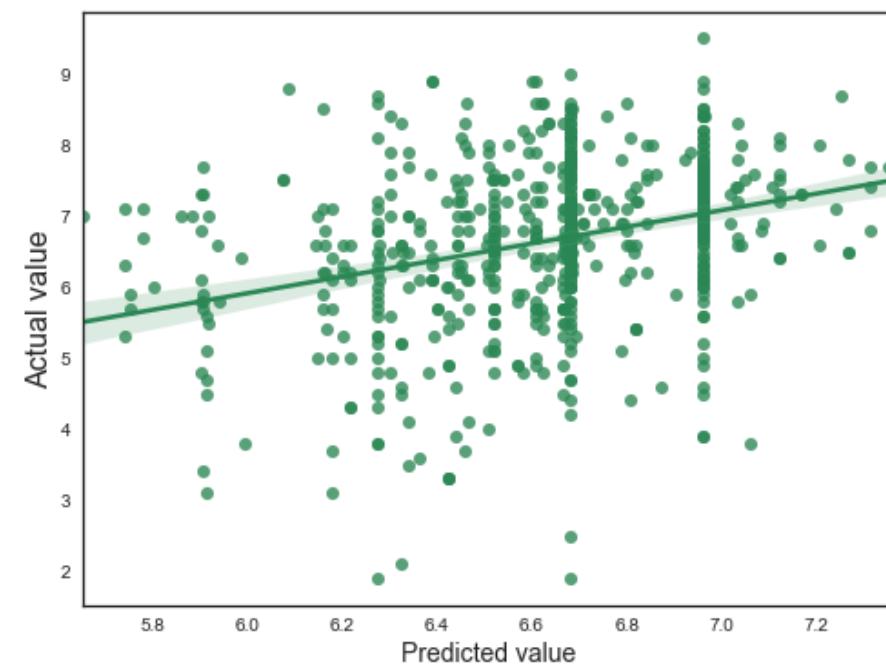
# EDA -> Cleaning -> Modeling



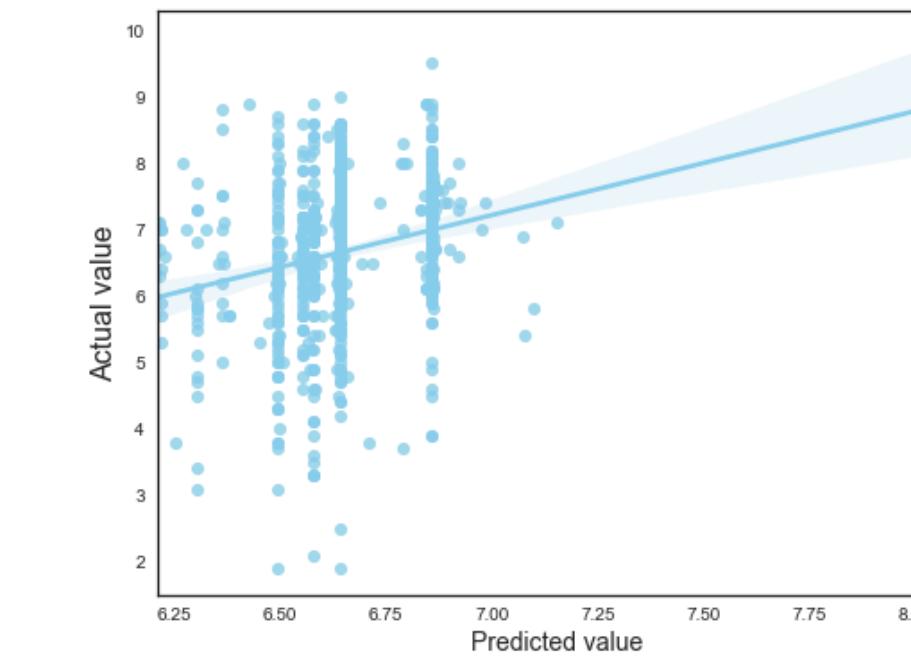
$R^2$  is just one metric. How can we look into the question: linear or polynomial?

# EDA → Cleaning → Modeling

Residual plot of LASSO 5-fold CV model

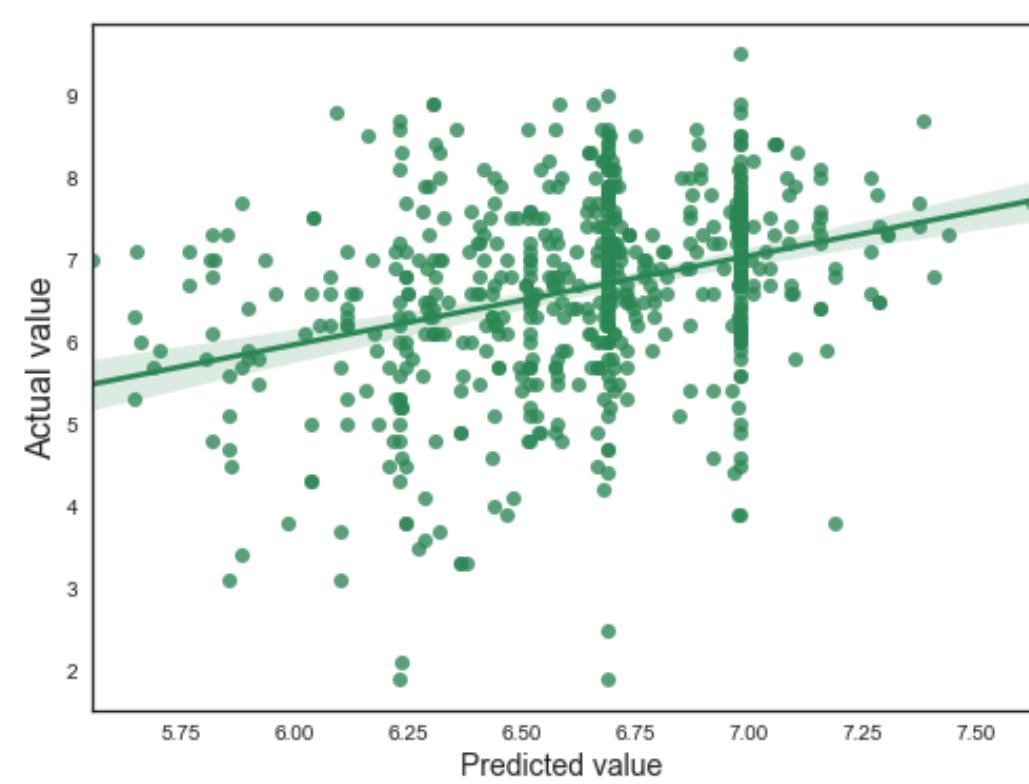


Residual plot of LASSO 5-fold CV model (polynomial)

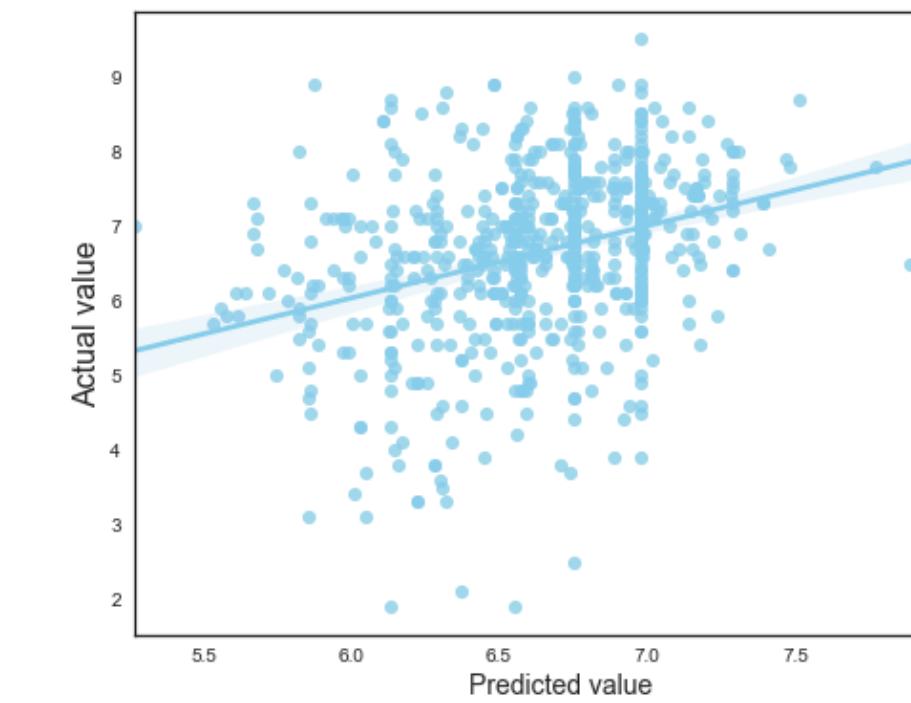


Residuals  
should be  
normally  
distributed  
with zero  
mean

Residual plot of Ridge 5-fold CV model

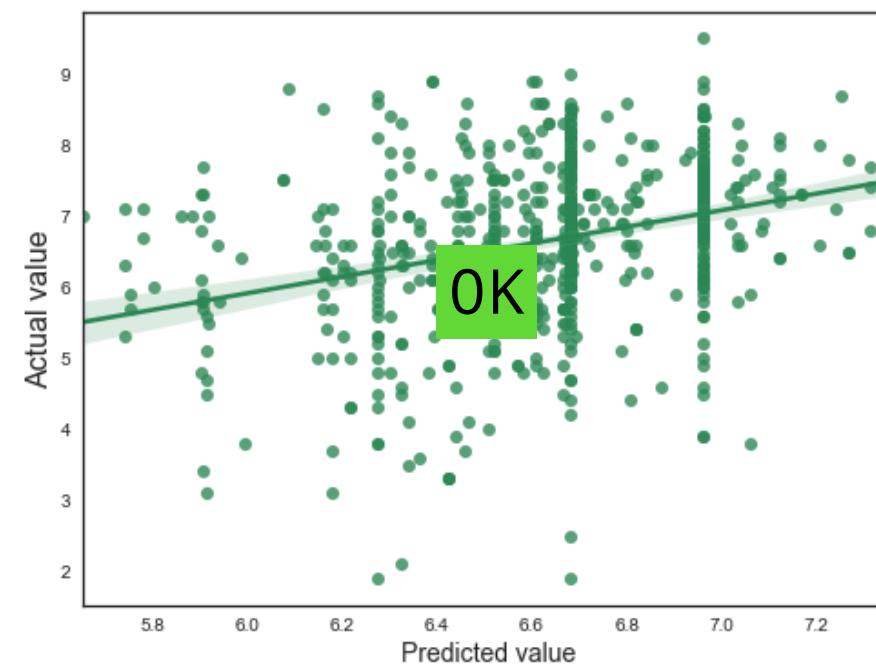


Residual plot of Ridge 5-fold CV model (polynomial)

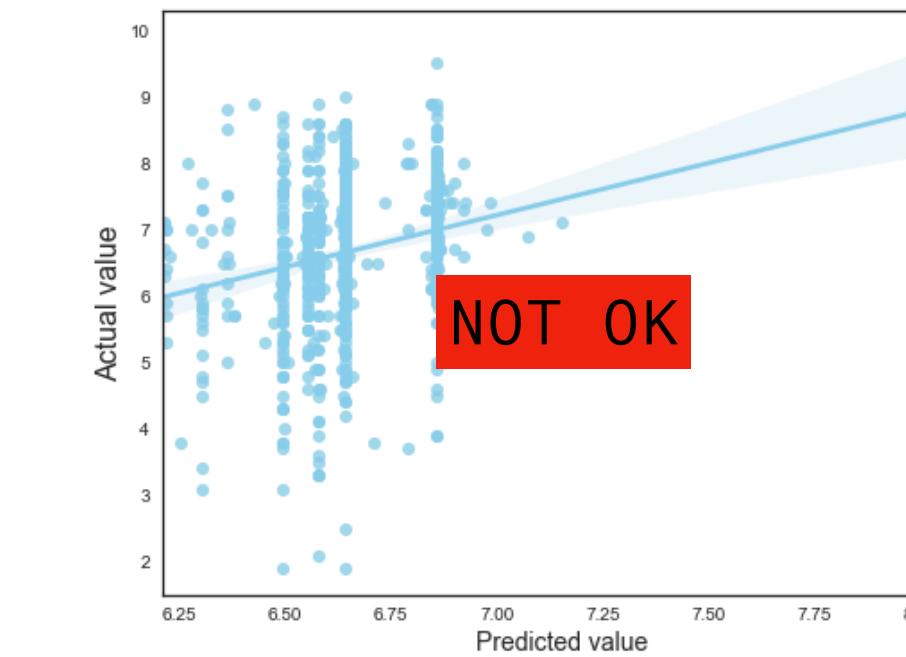


# EDA → Cleaning → Modeling

Residual plot of LASSO 5-fold CV model

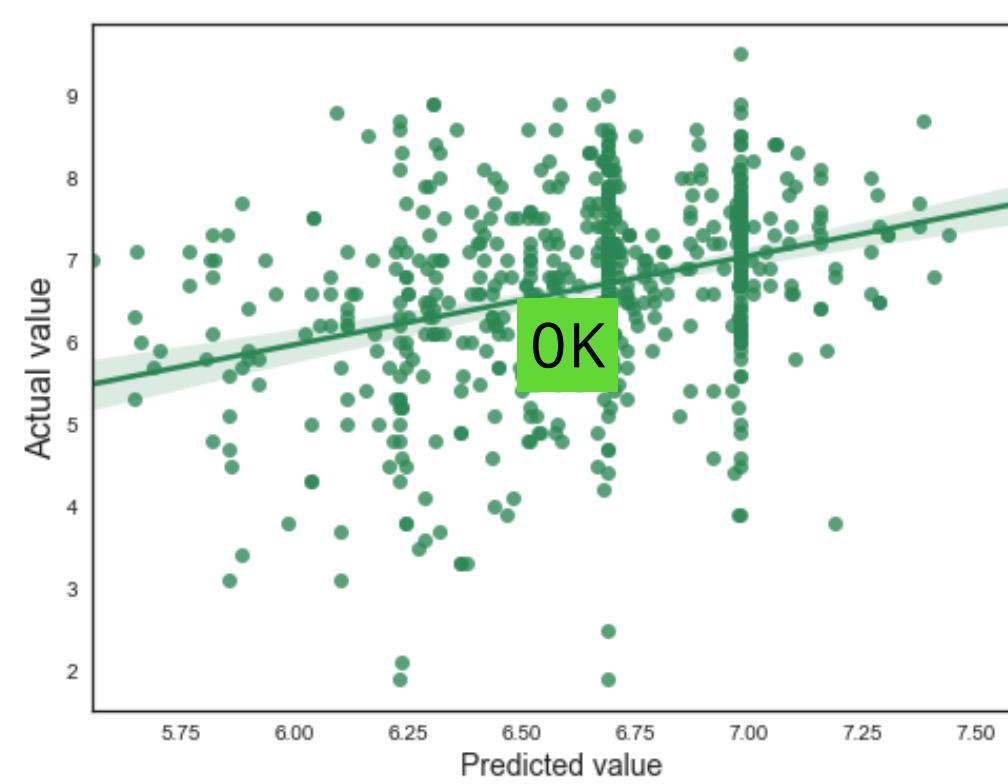


Residual plot of LASSO 5-fold CV model (polynomial)

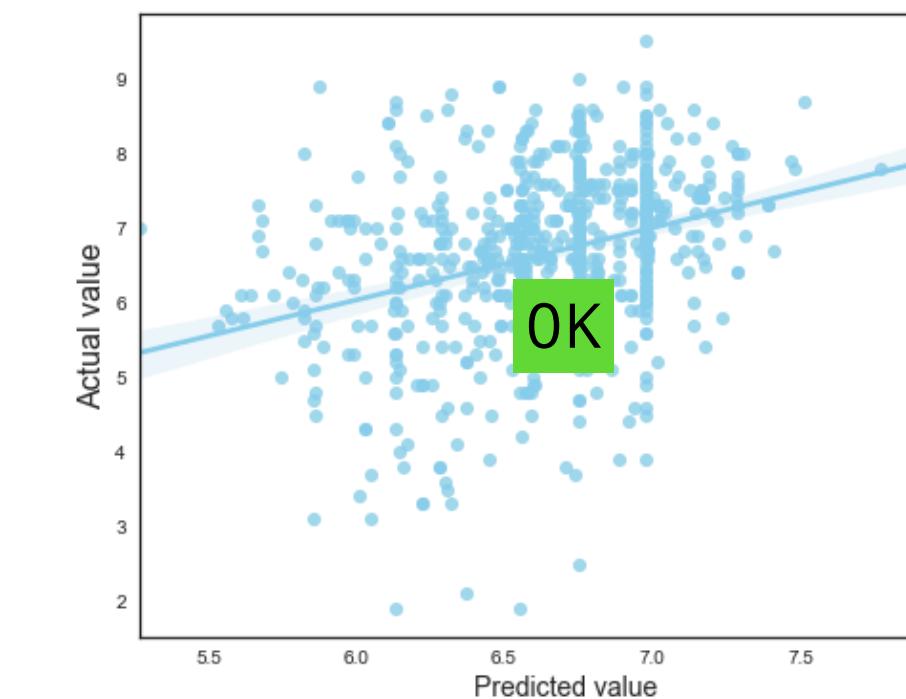


Residuals  
should be  
normally  
distributed  
with zero  
mean

Residual plot of Ridge 5-fold CV model

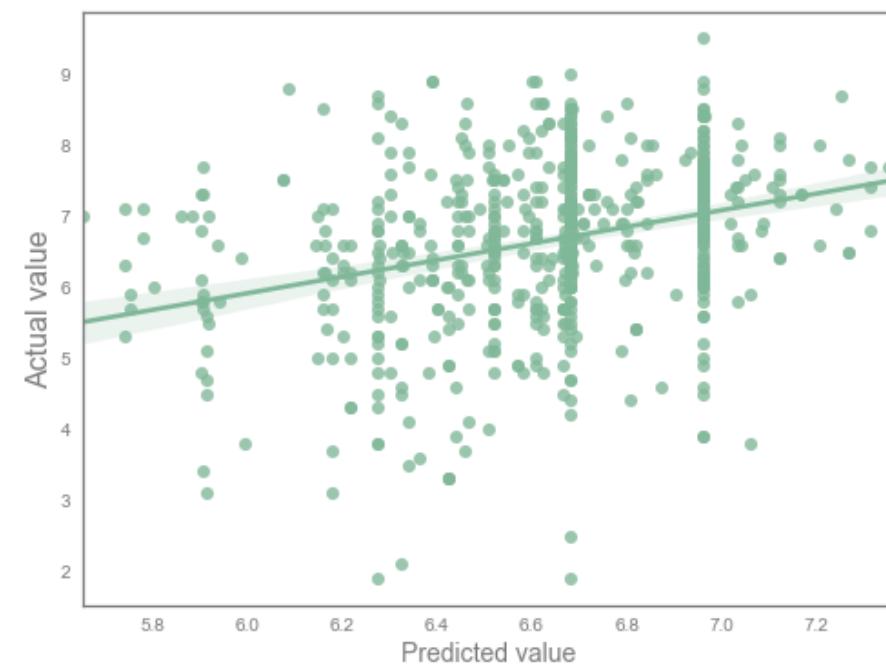


Residual plot of Ridge 5-fold CV model (polynomial)

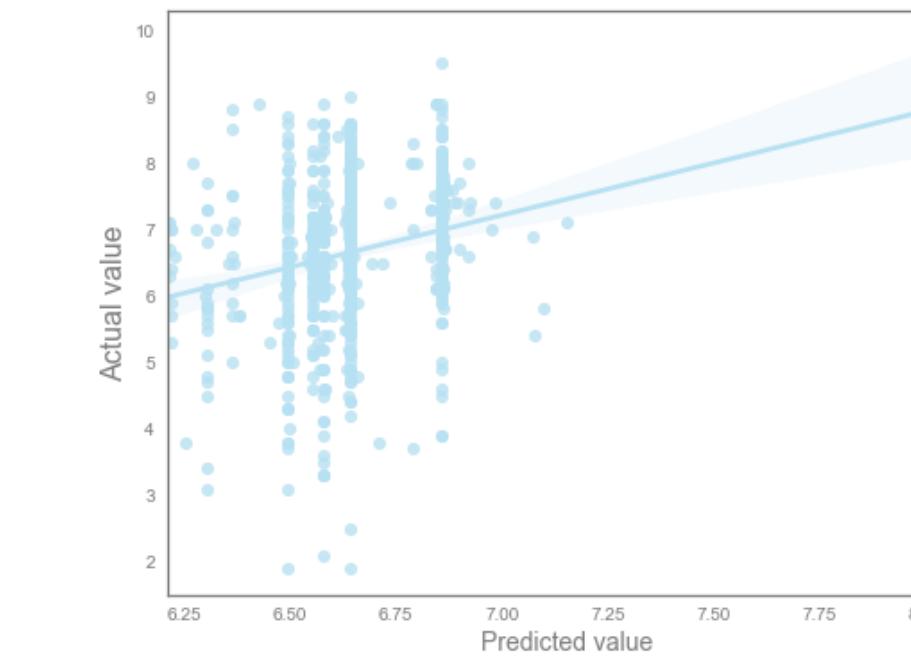


# EDA → Cleaning → Modeling

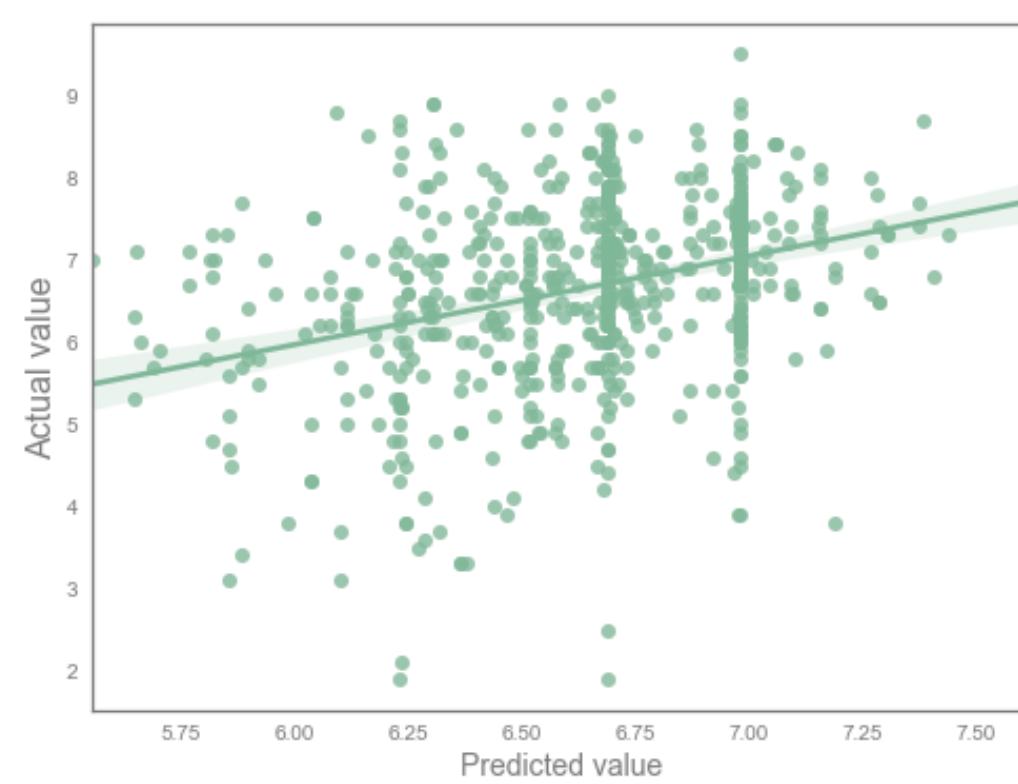
Residual plot of LASSO 5-fold CV model



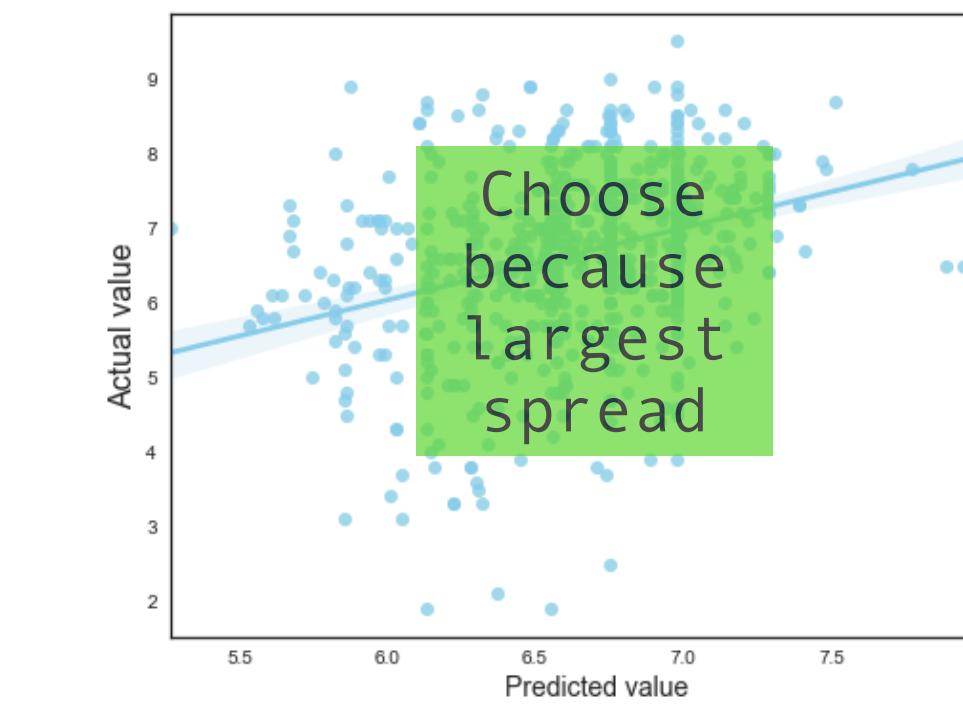
Residual plot of LASSO 5-fold CV model (polynomial)



Residual plot of Ridge 5-fold CV model



Residual plot of Ridge 5-fold CV model (polynomial)

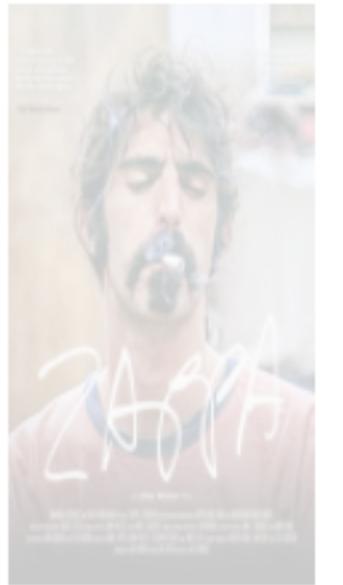


EDA -> Cleaning -> Modeling -> Visualizing

How do we make people care?

Liam-o-meter: 7.4

IMDB: 6.7



Zappa

Liam-o-meter: 7.4

IMDB: 7.6

Liam-o-meter: 7.4

IMDB: 8.0



Ruben Brandt,  
Collector

Liam-o-meter: 7.4

IMDB: 7.5

Liam-o-meter: 7.4

IMDB: 4.1



Beauty and the  
Beast

Liam-o-meter: 7.4

IMDB: 7.1

Mantra of Rock

Liam-o-meter: 7.4

IMDB: 6.8



Sunset Blvd.

Liam-o-meter: 7.4

IMDB: 8.4

MacGowan

Liam-o-meter: 7.4

IMDB: 8.0



The Dissident

Liam-o-meter: 7.4

IMDB: 8.2

By showing them movie pictures!

Tito and the  
Birds

Liam-o-meter: 7.4

IMDB: 6.5



Promare

Liam-o-meter: 7.4

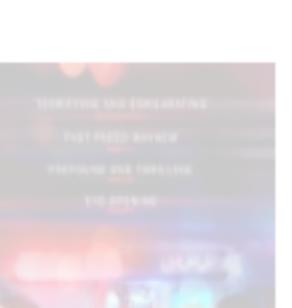
IMDB: 7.2



Liz and the  
Blue Bird

Liam-o-meter: 7.4

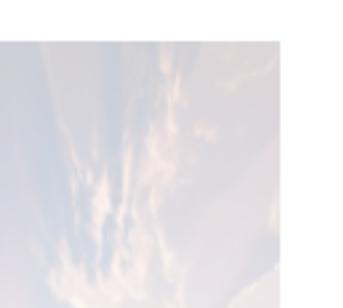
IMDB: 7.2



Fireworks

Liam-o-meter: 7.4

IMDB: 7.8



Burn the  
Stage: The  
Movie

Liam-o-meter: 7.4

IMDB: 8.8



# Demo

[liamisaacs.com/liamometer](http://liamisaacs.com/liamometer)

# Conclusions

- Now you know, if all you cared about was internationally successful movies from lesser-known studios, what you would care about.