Women, Dialogue, and the Box Office

The Case for Hollywood to Make Films that Humanize Women

Project Luther
Metis NYC DS9

Ting Neo
October 7, 2016
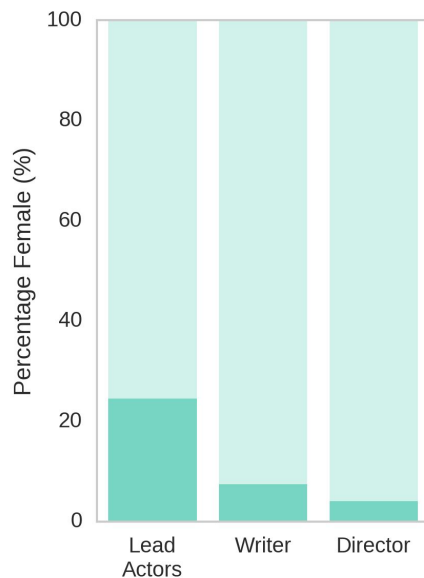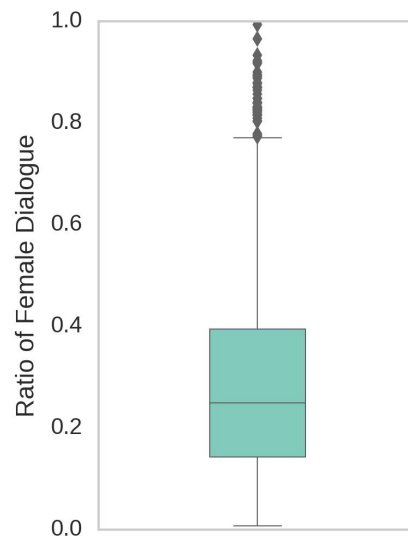
Hollywood has a gender problem

1. Does data prove the existence of this gender bias?

2. Is there a financial case for Hollywood to make more films starring, written by, and directed by women?

Defining gender features: measuring the presence (or lack) of women in the filmmaking process

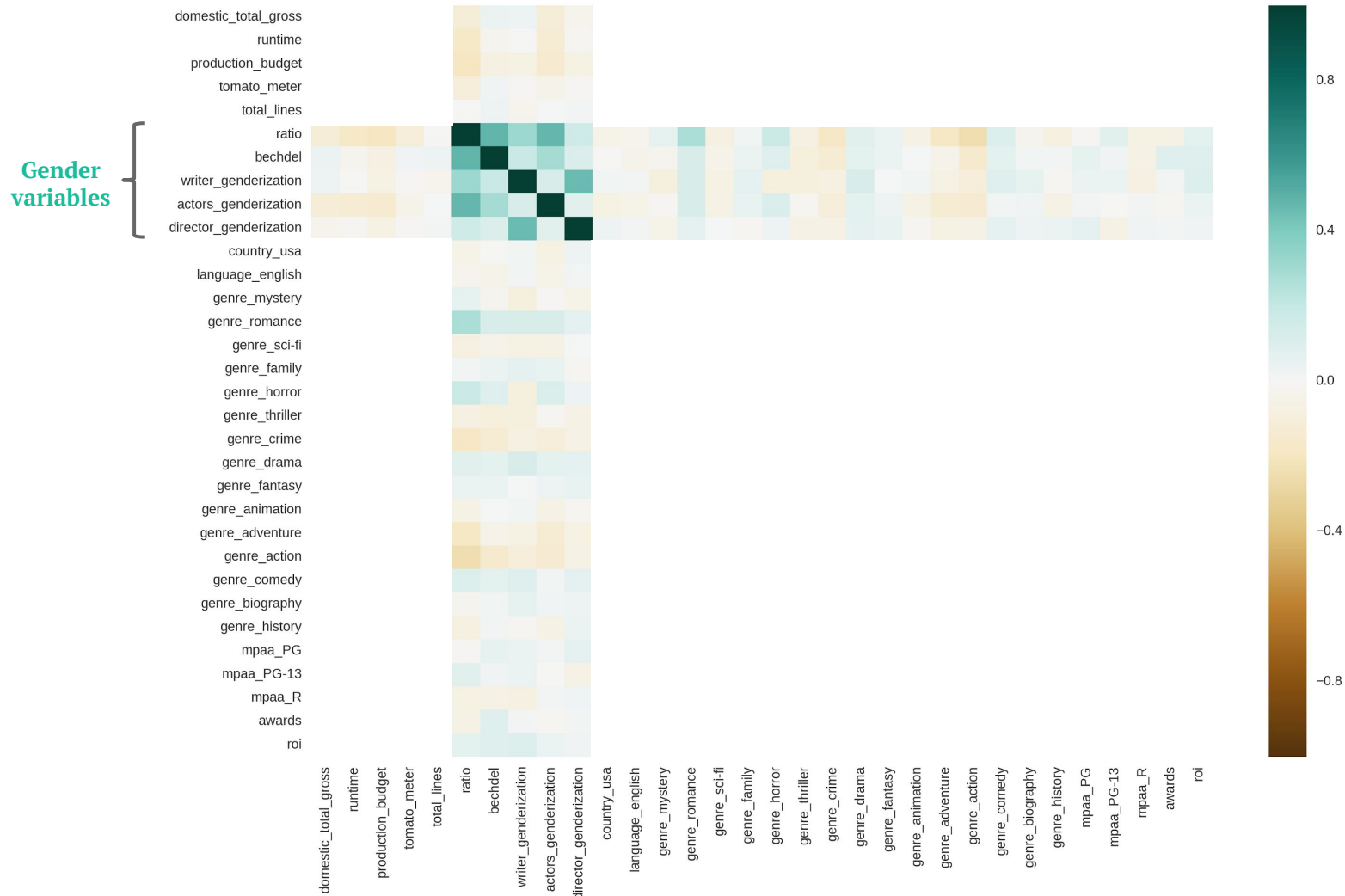## Gender Proportion

Dialogue by Gender

Bechdel Test
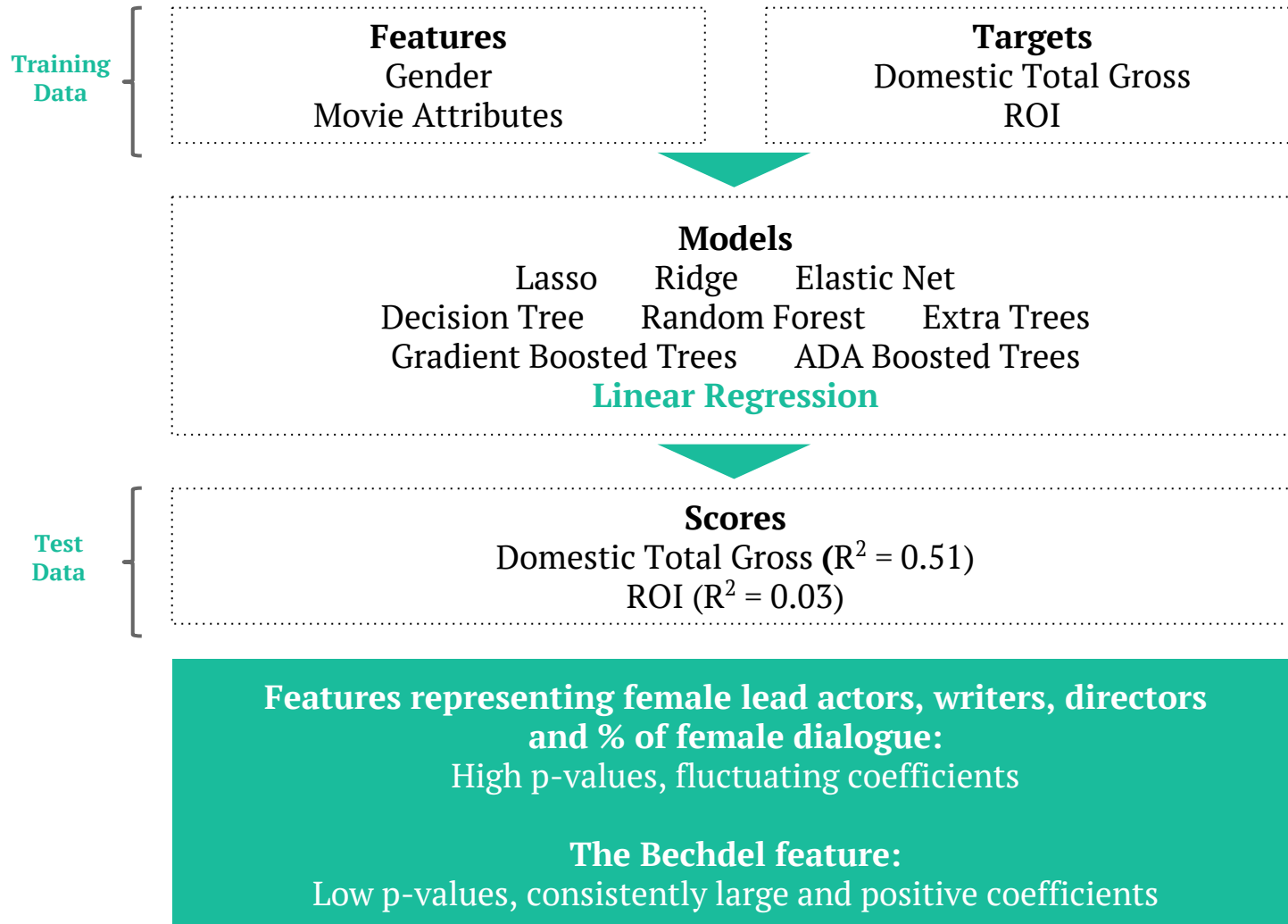
A film passes the Bechdel test if it:

1. Has at least two [named] women
2. Who talk to each other
3. About something besides a man

Only **50%** of films pass

Are other movie attributes correlated with gender features?

Gender variables

Testing for significance

Features
Gender
Movie Attributes

Targets
Domestic Total Gross
ROI

Models
Lasso   Ridge   Elastic Net
Decision Tree   Random Forest   Extra Trees
Gradient Boosted Trees   ADA Boosted Trees
Linear Regression

Test Data

Scores
Domestic Total Gross **(**R^2 = 0.51)
ROI (R^2 = 0.03)

Features representing female lead actors, writers, directors and % of female dialogue:
High p-values, fluctuating coefficients

The Bechdel feature:
Low p-values, consistently large and positive coefficients

No evidence to support that lead actresses, an increased proportion of female dialogue, or films made by women lead to lower returns

The cost of objectifying women:
$22 million* per film in domestic total gross

Image Credit: New York Times

thanks!

Any questions?

You can find me at

http://neokt.github.io/

neo.kaiting@gmail.com

Sources and Tools

- ▣ Data
  - ○ [Box Office Mojo](#)
  - ○ [Bechdel Test](#)
  - ○ [OMDB (Open Movie Database)](#)
  - ○ [Polygraph's Film Dialogue Dataset](#)

- ▣ Python Packages
  - ○ Pandas
  - ○ Matploblib
  - ○ Seaborn
  - ○ BeautifulSoup
  - ○ Genderizer
  - ○ Statsmodel
  - ○ Sklearn