# Predicting Bechdel Test Results through Statistical Modeling

Hailey Quintavalle

## Table of contents

## Abstract

The Bechdel Test is a simple measurement designed to analyze the representation of women in film. To pass, a movie must have two female characters who have a conversation that is not about a man. The present research aims to use genre, release year, movie budget, user ratings, and critics rating scores to predict the probability a movie will pass the test. Logistic regression analysis reveals more recent movies are predicted to have a higher probability of passing the test. For most years, genres such as Romance and Comedy are predicted to have a higher probability of passing, while genres such as Action, Sport, War, and Western are predicted to have a lower probability of passing the test. Although the Bechdel Test has its flaws, the test is a useful metric to bring attention to the roles women hold in film.

## Introduction

If you think of your all-time favorite movie, try to think about how many of the main characters were women. In 2023, 35% of speaking roles in movies belonged to women, and only 28% of the top grossing films contained female protagonists (Lauzen, 2024). Just one decade before, in 2013, women occupied 30% of all speaking roles and just 15% of protagonists were women (Lauzen, 2014). This is a small change for a span of 10 years, showcasing the underrepresentation of women in film, and the slow steps towards progress the film industry seems to be taking. As we examine the portrayal of women's roles in film, it is useful to explore various criteria used to measure their representation. One notable measure is the Bechdel Test, unintentionally introduced by American cartoonist Alison Bechdel in 1986, and the primary measure of women in film that is the focus of the present analysis.

The Bechdel Test originated from a comic strip titled "The Rule", a part of her comic *Dykes To Watch Out For*. The comic depicts two women discussing certain criteria necessary for them to watch a movie. The criteria for a movie included: two named women, who talked to each other, and held a conversation with each other that was not about a man. In the early 2000s, the test quickly gained popularity online and has since been used as a common tool for analyzing the role of women in film.

The test has only three basic requirements, meaning that a film can pass with just one line. Because of the simplicity of this tool, many have come forward with fair criticisms of the test. For example, the test does not take into account demographic factors, such as age, the voices of women of color, or those who do not speak English as their first language (O'Meara, 2016). Other flaws of the test include the oversight of conversations that are not directly about a man, but may be indirectly about them, where the conclusion of the test are unclear.

Regardless of its flaws, the Bechdel Test is a straightforward and easy tool that allows people to quickly make base assumptions about the presence of women in film. For this reason, this analysis will be examining the characteristics of thousands of movies in an attempt to use logistic regression analysis to model the likelihood of passing the Bechdel Test.

Read the comic here.

### Data

The data used for the present analysis comes from a combination of multiple online sources. A data set containing 10,183 movie titles was available through IMDb Non-Commercial Datasets. Release dates for these movies range from 1874 to 2023. Variables included in this data set can be seen in the preview of the data below.

IMDb Non-Commercial Data set:

| tconst | originalTitle | startYear | runtimeMinutes | genres |
|---|---|---|---|---|
| tt27502426 | Les filles d'Olfa | 2023 | 107 | Documentary |
| tt15398776 | Oppenheimer | 2023 | 180 | Biography,Drama,History |
| tt15326988 | Ghosted | 2023 | 116 | Action,Adventure,Comedy |
| tt8400584 | The Perfect Find | 2023 | 99 | Comedy,Drama,Romance |
| tt14230388 | Asteroid City | 2023 | 105 | Comedy,Drama,Romance |
| tt18257464 | Polite Society | 2023 | 104 | Action,Comedy |

The code manual for the data from IMDb:

- **tconst**: alphanumeric unique identifier of the title
- **originalTitle**: title of the movie
- **startYear**: release year of the movie
- **runtimeMinutes**: runtime in minutes
- **genres**: up to three genres associated with the title

Other data that was used in this analysis comes from the Bechdel Test Movie List, where users can submit movies with their Bechdel test rating through their online platform. The data set pulled from this website contains 10,251 movies with release dates ranging from 1874 to 2024. For the rating variable, a movie is given a rating from one to three, directly corresponding with the number of requirements of the Bechdel test that it passes. A preview of this data can be shown below.

| id | title | imdbid | year | rating |
|---|---|---|---|---|
| 11166 | Hunger Games: The Ballad of Songbirds & Snakes, The | 10545296 | 2023 | 3 |
| 11169 | Marvels, The | 10676048 | 2023 | 3 |
| 11171 | Royal Hotel, The | 18363072 | 2023 | 3 |
| 11172 | Nowhere | 15789472 | 2023 | 3 |
| 11173 | Leo | 15654328 | 2023 | 3 |
| 11230 | Book of Clarence , The | 22866358 | 2023 | 1 |

The code manual for the Bechdel Test Movie List data:

- **id**: ID number
- **title**: title of the movie
- **imdbid**: IMDb number ID
- **year**: release year

- **rating**: Bechdel Test rating (0-3)

The data from IMDb Non-Commercial Datasets was joined with the data set pulled from the Bechdel test movie list to be used in the final model concerning genres.

The final data set that was used in this analysis comes from the TidyTuesday social data project through GitHub. This data set contains 1,794 movies released from 1970 up to 2013.

| year | imdb | title | test | clean_test | binary | budget |
|---|---|---|---|---|---|---|
| 2013 | tt0770828 | Man of Steel | ok-disagree | ok | PASS | 2.25e+08 |
| 2013 | tt1821549 | Nebraska | ok-disagree | ok | PASS | 1.20e+07 |
| 2013 | tt1670345 | Now You See Me | notalk-disagree | notalk | FAIL | 7.50e+07 |

| domgross | intgross | code | budget_2013 | domgross_2013 | intgross_2013 |
|---|---|---|---|---|---|
| 291045518 | 687999518 | 2013PASS | 2.25e+08 | 291045518 | 687999518 |
| 17482517 | 17482517 | 2013PASS | 1.20e+07 | 17482517 | 17482517 |
| 117723989 | 351723989 | 2013FAIL | 7.50e+07 | 117723989 | 351723989 |

| period_code | decade_code | imdb_id | response | rated | language | runtime |
|---|---|---|---|---|---|---|
| 1 | 1 | 0770828 | TRUE | PG-13 | English | 143 min |
| 1 | 1 | 1821549 | TRUE | R | English, Spanish | 115 min |
| 1 | 1 | 1670345 | TRUE | PG-13 | English | 115 min |

| plot |
|---|
| A young itinerant worker is forced to confront his secret extrastellar origin when Earth is invaded by members of his own race. |
| An aging, booze-addled father makes the trip from Montana to Nebraska with his estranged son in order to claim a million-dollar Mega Sweepstakes Marketing prize. |
| An FBI agent and an Interpol detective track a team of illusionists who pull off bank heists during their performances and reward their audiences with the money. |

| writer |
|---|
| David S. Goyer (screenplay), David S. Goyer (story), Christopher Nolan (story), Jerry Siegel (Superman created by), Joe Shuster (Superman created by) |
| Bob Nelson |

| writer |
| --- |
| Ed Solomon (screenplay), Boaz Yakin (screenplay), Edward Ricourt (screenplay), Boaz Yakin (story), Edward Ricourt (story) |

| country | metascore | imdb_rating | director | released | imdb_votes |
| --- | --- | --- | --- | --- | --- |
| USA, Canada, UK | 55 | 7.4 | Zack Snyder | 14 Jun 2013 | 359556 |
| USA | 86 | 7.9 | Alexander Payne | 24 Jan 2014 | 33503 |
| France, USA | 50 | 7.3 | Louis Leterrier | 31 May 2013 | 280199 |

| actors | genre | awards |
| --- | --- | --- |
| Henry Cavill, Amy Adams, Michael Shannon, Diane Lane | Action, Adventure, Fantasy | 4 wins & 16 nominations. |
| Bruce Dern, Will Forte, June Squibb, Bob Odenkirk | Adventure, Drama | Nominated for 6 Oscars. Another 26 wins & 61 nominations. |
| Jesse Eisenberg, Mark Ruffalo, Woody Harrelson, Isla Fisher | Crime, Mystery, Thriller | 1 win & 3 nominations. |

The code manual for the TidyTuesday data:

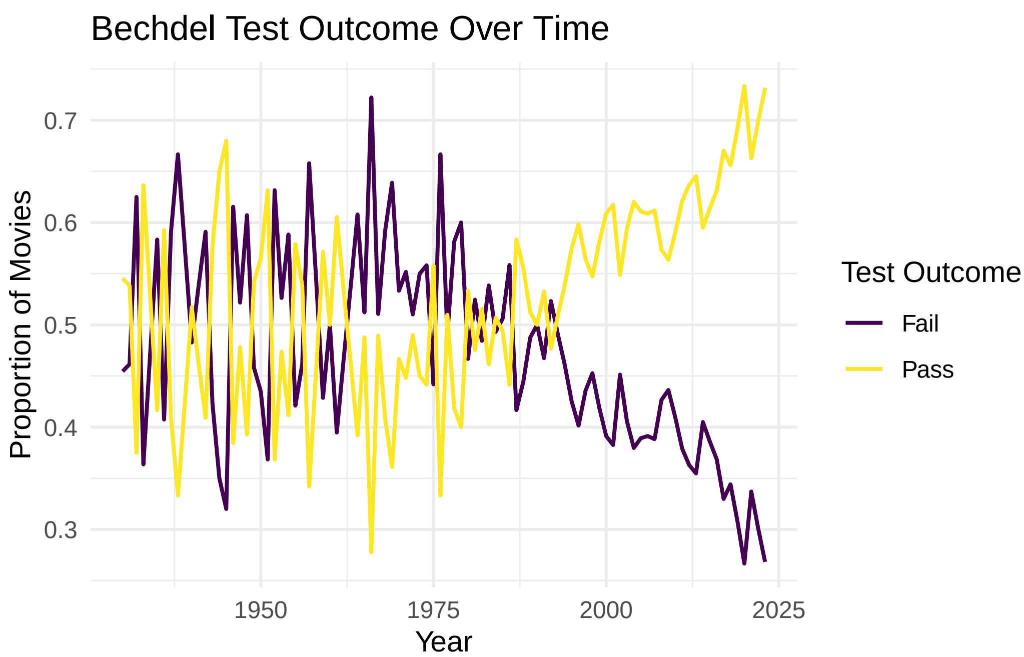
- **year**: release year
- **imdb**: IMDb id
- **title**: title of movie
- **test**: Bechdel Test outcome
- **clean_test**: Bechdel Test outcome (cleaned)
- **binary**: binary Pass or Fail of the Bechdel Test
- **budget**: budget as of release year
- **domgross**: domestic gross in release year
- **intgross**: international gross in release year
- **code**: code for movie ###

- **budget_2013**: budget normalized to 2013
- **domgross_2013**: domestic gross normalized to 2013
- **intgross_2013**: international gross normalized to 2013
- **period_code**: period code ###
- **decade_code**: decade code
- **imdb_id**: IMDb ID ###
- **plot**: plot summary of the movie
- **rated**: content rating of movie
- **response**: ??? ###
- **language**: language of movie
- **country**: country movie is produced in
- **writer**: writer of the film
- **metascore**: Metascore (critic) rating (0-100)
- **imdb_rating**: IMDb (user) rating (0-10)
- **director**: Director(s) of movie
- **released**: released date
- **actors**: main actors in movie
- **genre**: genre
- **awards**: awards won
- **runtime**: runtime in minutes
- **type**: type of film ###
- **imdb_votes**: number of IMDb votes

## Exploratory Analysis/Overview

Looking into the IMDb movie data set, the overall percentage of movies that pass the Bechdel Test is 57.0%, with the remaining 43.0% failing the test. The plot below demonstrates the proportion of movies each year that pass and fail the test each year, starting in 1930 up to 2023. The movies released in the years before 1930 have been excluded from this plot because of the amount of variation coming from a small number of movies each year, with an average of around 4 movies per year.

```
basics_movies |>
  mutate(binary = if_else(rating == 3, "Pass", "Fail")) |>
  select(year, binary) |>
  group_by(year, binary) |>
  summarise(n = n()) |>
  filter(year >= 1930) |>
  pivot_wider(names_from = binary, values_from = n) |>
  mutate(Pass = if_else(is.na(Pass), 0, Pass),
         Fail = if_else(is.na(Fail), 0, Fail)) |>
  mutate(total_movies = sum(Pass, Fail)) |>
  mutate(Fail = Fail / total_movies,
         Pass = Pass / total_movies) |>
  pivot_longer(c(Fail, Pass), names_to = "binary", values_to = "prop") |>
  ggplot(aes(x = year, y = prop)) +
  geom_line(aes(color = binary), linewidth = 0.7) +
  scale_color_viridis_d(name = "Test Outcome") +
  labs(x = "Year",
       y = "Proportion of Movies",
       title = "Bechdel Test Outcome Over Time") +
  theme_minimal()
```




As seen in this plot, the proportion of movies passing and failing the test has much variation from the 1930s up until the early 90s. After the early 90s, however, the difference between the proportion of those that pass and fail each year becomes much more prominent. The proportion of movies that are passing the Bechdel Test are increasing each year.

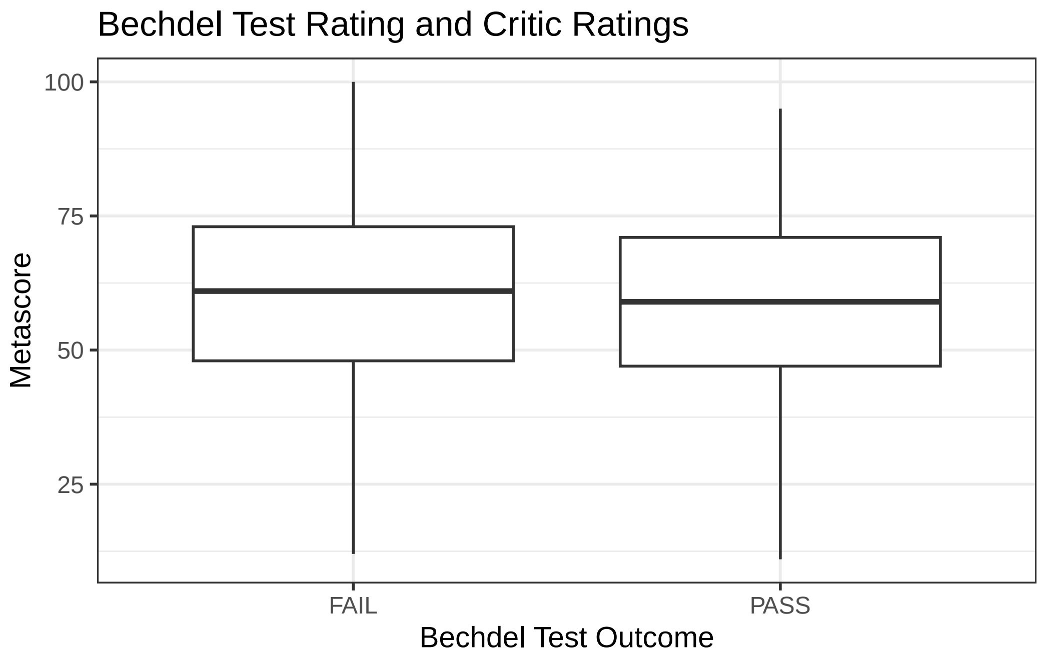
```
# movies appear multiple times for multiple listed genres
movies_model <- basics_movies |> separate_rows(genres, sep = ",")
movies_model <- movies_model |> mutate(binary = if_else(rating == 3, "1", "0")) |>
  mutate(binary = as.numeric(binary)) |>
  filter(!genres %in% c("News", "Adult", "Talk-Show", "\\N")) |>
  relocate(binary)
```

[enter some visual about genres here?]

Exploring the TidyTuesday data set allows us to examine other characteristics of movies, such as movie budget and different types of ratings. The plot below shpws the difference in critic ratings, or Metascore, between movies that pass and fail the Bechdel Test.

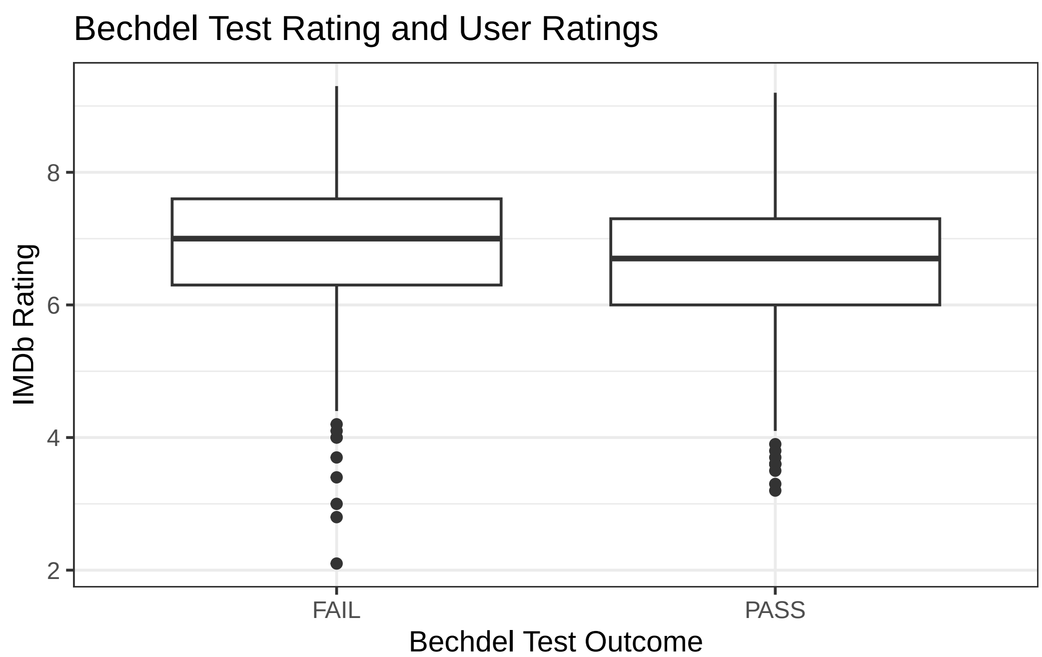
```
movies |>
  ggplot(aes(x = binary, y = metascore)) +
  geom_boxplot() +
  theme_bw() +
  labs(x = "Bechdel Test Outcome",
```

```
       y = "Metascore",
       title = "Bechdel Test Rating and Critic Ratings")
```



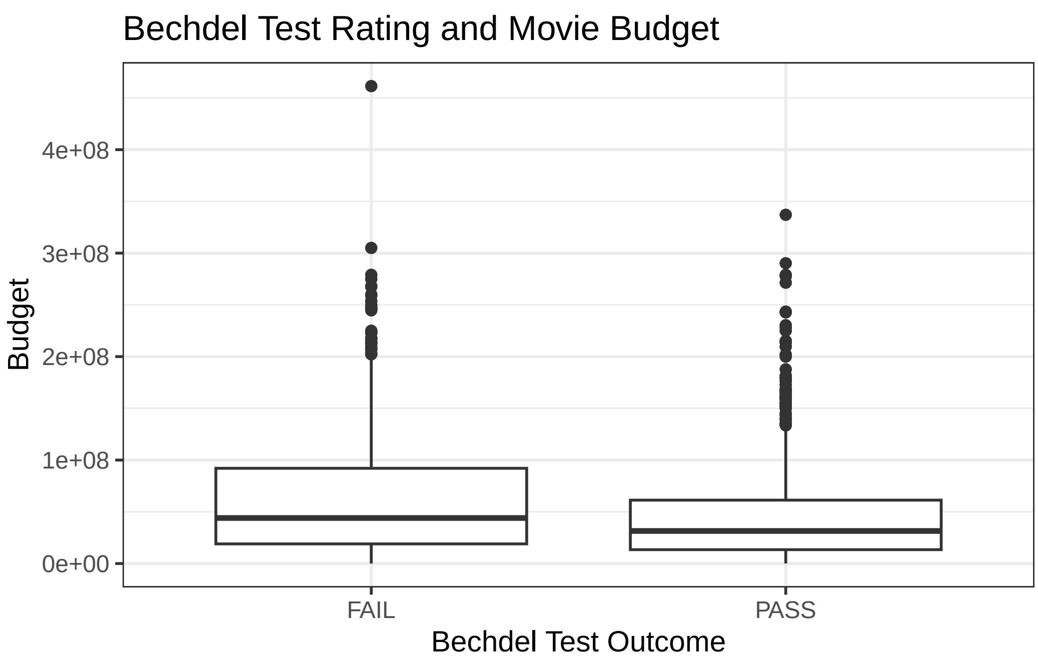

The plot above suggests a slight increase in average Metascore ratings for the movies that do not pass the Bechdel Test, but the difference appears to be too small to be notable. Demonstrated in the following plot, movies that do not pass the Bechdel Test also appear to have higher average user ratings than movies that pass.

```
movies |>
  ggplot(aes(x = binary, y = imdb_rating)) +
  geom_boxplot() +
  theme_bw() +
  labs(x = "Bechdel Test Outcome",
       y = "IMDb Rating",
       title = "Bechdel Test Rating and User Ratings")
```




In terms of movie budget, it appears that movies that fail tend to show larger movie bugdets than movies that pass the test, as shown in the following plot.

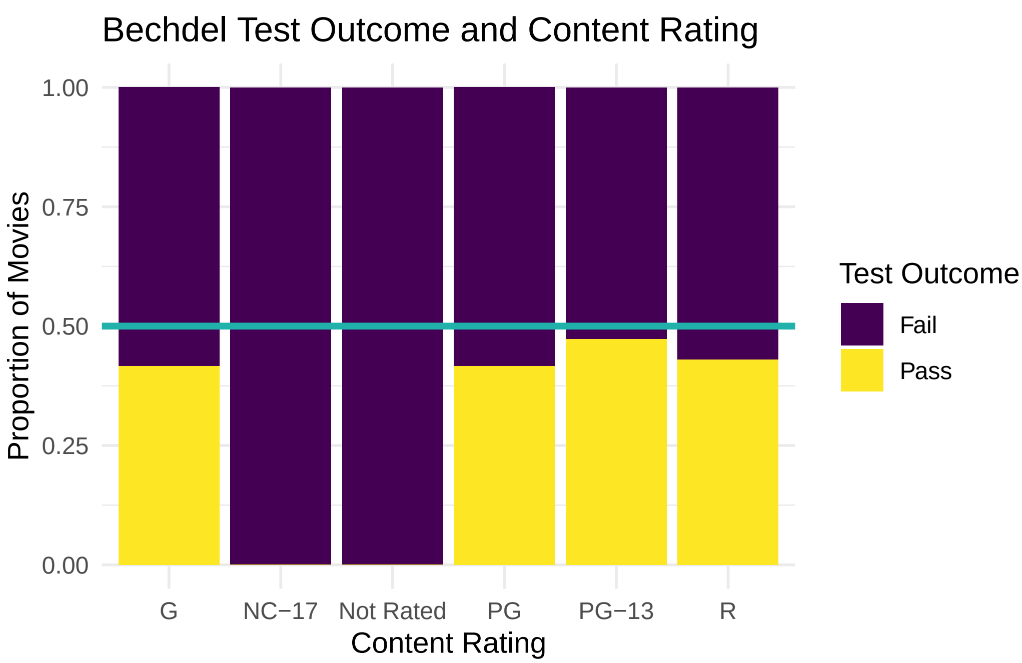
```
movies |>
  ggplot(aes(x = binary, y = budget_2013)) +
  geom_boxplot() +
  theme_bw() +
  labs(x = "Bechdel Test Outcome",
       y = "Budget",
       title = "Bechdel Test Rating and Movie Budget")
```




Another interesting variable in this data set is content rating. The content rating is designed to classify movies based on the audiences they are suitable for, based on factors like profanity, violence, sexual content, substance use, and any other topics that are not recommend for children to watch. The plot below contains five different content ratings with the proportion of movies that pass and fail the Bechdel Test in those categories.

```
movies |>
  select(rated, binary) |>
  group_by(rated, binary) |>
  summarise(n = n()) |>
  filter(n > 4) |>
  pivot_wider(names_from = binary, values_from = n) |>
  mutate(Pass = if_else(is.na(PASS), 0, PASS),
         Fail = if_else(is.na(FAIL), 0, FAIL)) |>
  mutate(total_movies = sum(Pass, Fail)) |>
  mutate(Fail = Fail / total_movies,
         Pass = Pass / total_movies) |>
  pivot_longer(c(Fail, Pass), names_to = "binary", values_to = "prop") |>
  filter(rated != "N/A") |>
  ggplot(aes(x = rated, y = prop)) +
  geom_col(aes(fill = binary), position = "stack") +
  geom_hline(yintercept = 0.5, linewidth = 1.2, color = "lightseagreen") +
  theme_minimal() +
```

```r
labs(x = "Content Rating",
     y = "Proportion of Movies",
     title = "Bechdel Test Outcome and Content Rating") +
scale_fill_viridis_d(name = "Test Outcome")
```




The blue line represents the point of indifference, where the proportion of movies that pass is equal to 50%. The proportion of movies that pass the test does not reach half for any of the content rating categories, revealing that the majority of movies in each of these categories do not seem to pass.

## Modeling

### Genre Analysis

Full model with all genres and all interactions:

$$\text{logit}(\pi) = \beta_0 + \beta_1 \cdot \text{Year} + \beta_2 \cdot \text{Animation} + \beta_3 \cdot \text{Action} + \beta_4 \cdot \text{Adult} + \beta_5 \cdot \text{Adventure} + \beta_6 \cdot \text{Biography} + \beta_7 \cdot \text{Comedy} + \beta_8 \cdot \text{Cri}$$

```r
## basics_movies with indicator variable for each genre
movies_indi <- read_csv(here::here("data/movies_indicator.csv"))

new_grid <- read_csv(here::here("data/movies_grid1.csv"))
library(modelr)
library(broom)

genre_model <- glm(factor(binary) ~ year + Animation + Action + Adult +
                     Adventure + Biography + Comedy + Crime +
                     Documentary + Drama + Family + Fantasy +
                     History + Horror + Music + Musical +
                     Mystery + Romance + Short + Sport +
                     Thriller + War + Western + year:Animation +
                     year:Action + year:Adult + year:Adventure +
                     year:Biography + year:Comedy + year:Crime +
                     year:Documentary + year:Drama + year:Family +
                     year:Fantasy + year:History + year:Horror +
                     year:Music + year:Musical + year:Mystery +
                     year:Romance + year:Short + year:Sport +
                     year:Thriller + year:War + year:Western,
                   family = "binomial", data = movies_indi)
```

```r
genre_model |> pander::pander(caption = "Model Coefficients")
```

Table 10: Model Coefficients

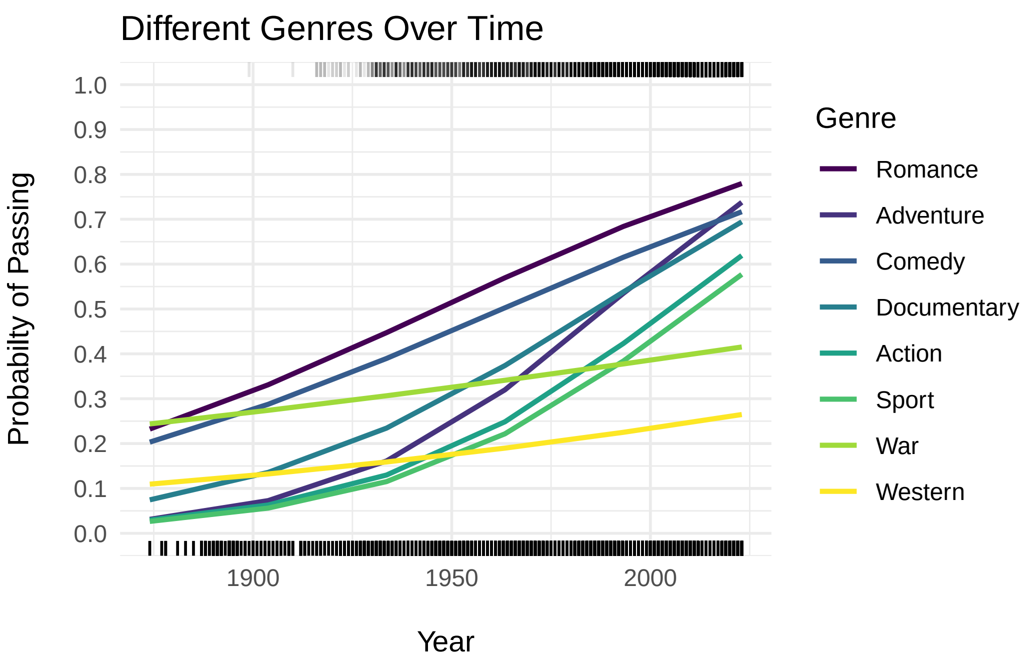
| | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| **(Intercept)** | -37.56 | 6.867 | -5.471 | 4.483e-08 |
| **year** | 0.01908 | 0.00344 | 5.546 | 2.916e-08 |
| **Animation** | -24.18 | 11 | -2.198 | 0.02798 |
| **Action** | -15.97 | 7.381 | -2.164 | 0.03046 |
| **Adult** | 31.48 | 18752 | 0.001679 | 0.9987 |
| **Adventure** | -22.12 | 6.641 | -3.331 | 0.0008664 |
| **Biography** | 5.874 | 9.814 | 0.5985 | 0.5495 |
| **Comedy** | 7.334 | 4.872 | 1.505 | 0.1322 |
| **Crime** | 13.73 | 5.736 | 2.393 | 0.0167 |
| **Documentary** | -6.955 | 19.51 | -0.3564 | 0.7215 |
| **Drama** | 10.29 | 5.026 | 2.048 | 0.04059 |
| **Family** | 1.427 | 8.118 | 0.1758 | 0.8604 |
| **Fantasy** | -2.324 | 7.531 | -0.3086 | 0.7576 |
| **History** | 2.483 | 10.03 | 0.2475 | 0.8046 |

| | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| **Horror** | -4.725 | 6.937 | -0.6811 | 0.4958 |
| **Music** | 2.424 | 12.14 | 0.1997 | 0.8417 |
| **Musical** | 4.9 | 10.15 | 0.4829 | 0.6292 |
| **Mystery** | 8.571 | 7.066 | 1.213 | 0.2251 |
| **Romance** | 5.439 | 4.928 | 1.104 | 0.2697 |
| **Short** | -26.7 | 13.9 | -1.921 | 0.05469 |
| **Sport** | -15.28 | 21.04 | -0.7263 | 0.4676 |
| **Thriller** | 10.38 | 6.921 | 1.5 | 0.1336 |
| **War** | 26.5 | 10.91 | 2.429 | 0.01514 |
| **Western** | 21.96 | 18.45 | 1.19 | 0.2339 |
| **year:Animation** | 0.01204 | 0.00549 | 2.192 | 0.02835 |
| **year:Action** | 0.007626 | 0.003685 | 2.069 | 0.03853 |
| **year:Adult** | -0.009716 | 9.432 | -0.00103 | 0.9992 |
| **year:Adventure** | 0.01093 | 0.003322 | 3.291 | 0.0009967 |
| **year:Biography** | -0.003053 | 0.0049 | -0.6229 | 0.5333 |
| **year:Comedy** | -0.003677 | 0.002439 | -1.508 | 0.1316 |
| **year:Crime** | -0.007114 | 0.002871 | -2.478 | 0.01321 |
| **year:Documentary** | 0.003333 | 0.00974 | 0.3422 | 0.7322 |
| **year:Drama** | -0.005127 | 0.002516 | -2.037 | 0.0416 |
| **year:Family** | -0.0005071 | 0.004068 | -0.1247 | 0.9008 |
| **year:Fantasy** | 0.001091 | 0.003769 | 0.2896 | 0.7721 |
| **year:History** | -0.001418 | 0.005022 | -0.2823 | 0.7777 |
| **year:Horror** | 0.002513 | 0.003472 | 0.7238 | 0.4692 |
| **year:Music** | -0.001224 | 0.006075 | -0.2015 | 0.8403 |
| **year:Musical** | -0.002443 | 0.005126 | -0.4766 | 0.6336 |
| **year:Mystery** | -0.00426 | 0.003536 | -1.205 | 0.2283 |
| **year:Romance** | -0.002575 | 0.002473 | -1.041 | 0.2977 |
| **year:Short** | 0.0127 | 0.006953 | 1.827 | 0.06775 |
| **year:Sport** | 0.007197 | 0.01051 | 0.6847 | 0.4935 |
| **year:Thriller** | -0.005301 | 0.003461 | -1.532 | 0.1256 |
| **year:War** | -0.01378 | 0.005508 | -2.502 | 0.01237 |
| **year:Western** | -0.01187 | 0.009348 | -1.27 | 0.2041 |

```r
aug_mod3 <- augment(genre_model, newdata = new_grid, se_fit = TRUE)

aug_mod3 <- aug_mod3 |> mutate(.pi = exp(.fitted) / (1 + exp(.fitted)))

list3 <- aug_mod3 |> pivot_longer(cols = c(Animation, Action, Adult,
                                           Adventure, Biography, Comedy,
                                           Crime, Documentary, Drama,
```

```
                                                      Family, Fantasy, History,
                                                      Horror, Music, Musical, Mystery,
                                                      Romance, Short,
                                                      Sport, Thriller,
                                                      War, Western), names_to = "genre",
                                           values_to = "values")

movies_pass <- movies_indi |> filter(binary == 1)
movies_fail <- movies_indi |> filter(binary == 0)
```

```
list3 |> filter(values == 1, genre %in%
                  c("Sport", "Comedy", "Romance", "Adventure",
                    "Action", "Documentary", "War", "Western")) |>
  mutate(genre = fct_reorder2(genre, .x = year, .y = .pi)) |>
  ggplot(aes(x = year, y = .pi)) +
  geom_line(aes(color = genre), linewidth = 0.9) +
  geom_rug(data = movies_pass,
           aes(x = year,
               y = as.numeric(binary)),
           sides = "t", alpha = 0.1) +
  geom_rug(data = movies_fail,
           aes(x = year,
               y = as.numeric(binary)),
           sides = "b") +
  scale_y_continuous(breaks = seq(0, 1, by = 0.1),
                     limits = c(0, 1)) +
  theme_minimal() +
  labs(x = "\nYear",
       y = "Probabilty of Passing\n",
       color = "Genre",
       title = "Different Genres Over Time") +
  scale_color_viridis_d()
```




## Further Analysis

```
movies <- movies %>% mutate(binary_0 = ifelse(binary == "PASS", 1, 0))

mod_year <- glm(binary_0 ~  budget_2013 +
                  year +
                  imdb_rating +
                  metascore,
                data = movies, family = "binomial")
```

```
mod_year |> pander::pander(caption = "Model Coefficients")
```

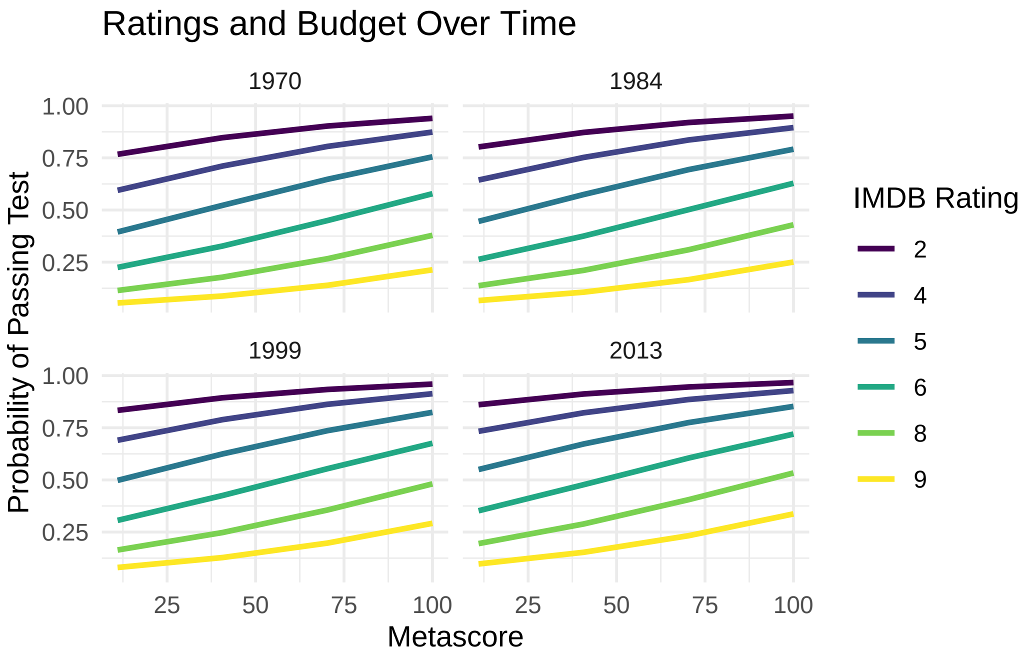
Table 11: Model Coefficients

| | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| **(Intercept)** | -26.36 | 15.38 | -1.714 | 0.08654 |
| **budget_2013** | -6.397e-09 | 1.078e-09 | -5.935 | 2.939e-09 |
| **year** | 0.01461 | 0.007634 | 1.914 | 0.05566 |
| **imdb_rating** | -0.5625 | 0.08984 | -6.261 | 3.821e-10 |
| **metascore** | 0.01745 | 0.004819 | 3.622 | 0.0002928 |

| | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|

```r
grid <- movies |>
  data_grid(
    year = seq_range(year, n = 4),
    binary_0 = c(0, 1),
    imdb_rating = seq_range(imdb_rating, n = 6),
    budget_2013 = median(movies$budget_2013, na.rm = T),
    metascore = seq_range(metascore, n = 4)
    )

aug_year <- augment(mod_year, newdata = grid, se_fit = TRUE)
aug_year <- aug_year |> mutate(.pi = exp(.fitted) / (1 + exp(.fitted))) |>
  mutate(year = round(year, 0))
```

```r
ggplot(data = aug_year, aes(x = metascore, y = .pi)) +
  geom_line(aes(color = as.factor(round(imdb_rating))), linewidth = 1) +
  facet_wrap(~ year) +
  labs(y = "Probability of Passing Test",
       x = "Metascore",
       title = "Ratings and Budget Over Time") +
  scale_color_viridis_d(name = "IMDB Rating") +
  theme_minimal()
```




## Conclusion

## References

1. Bechdel, A. (1986). Dykes to watch out for. *Firebrand Books.*

2. Lauzen, M. (2024). It's a Man's (Celluloid) World: Portrayals of female characters in the top grossing U.S. Films of 2023. *Center for the Study of Women in Television & Film.* https://womenintvfilm.sdsu.edu/its-a-mans-celluloid-world-portrayals-of-female-characters-in-the-top-grossing-u-s-films-of-2023/

3. Lauzen, M. (2014). It's a Man's (Celluloid) World: On-Screen Representations of Female Characters in the Top 100 Films of 2013. *Center for the Study of Women in Television & Film.* https://womenintvfilm.sdsu.edu/files/2013_It's_a_Man's_World_Report.pdf