Age Gaps in Hollywood Films

Project **T**eddy
Harper
Andres
Naman

Our Topic

- Explores the age difference between actors playing couples in Hollywood movies
  - Demonstrates the male-female power dynamic in the movie industry
  - Will tell us if criticism within the industry and from viewers has actually been effective in making a change in the age gaps portrayed in film
- We examine:
  - Whether the male or female member of the couple tends to be older
  - Whether the uncomfortable age gaps between male and female romantic interests in movies has been waning over time.
  - Whether the proportion of films with large age gaps has decreased over time

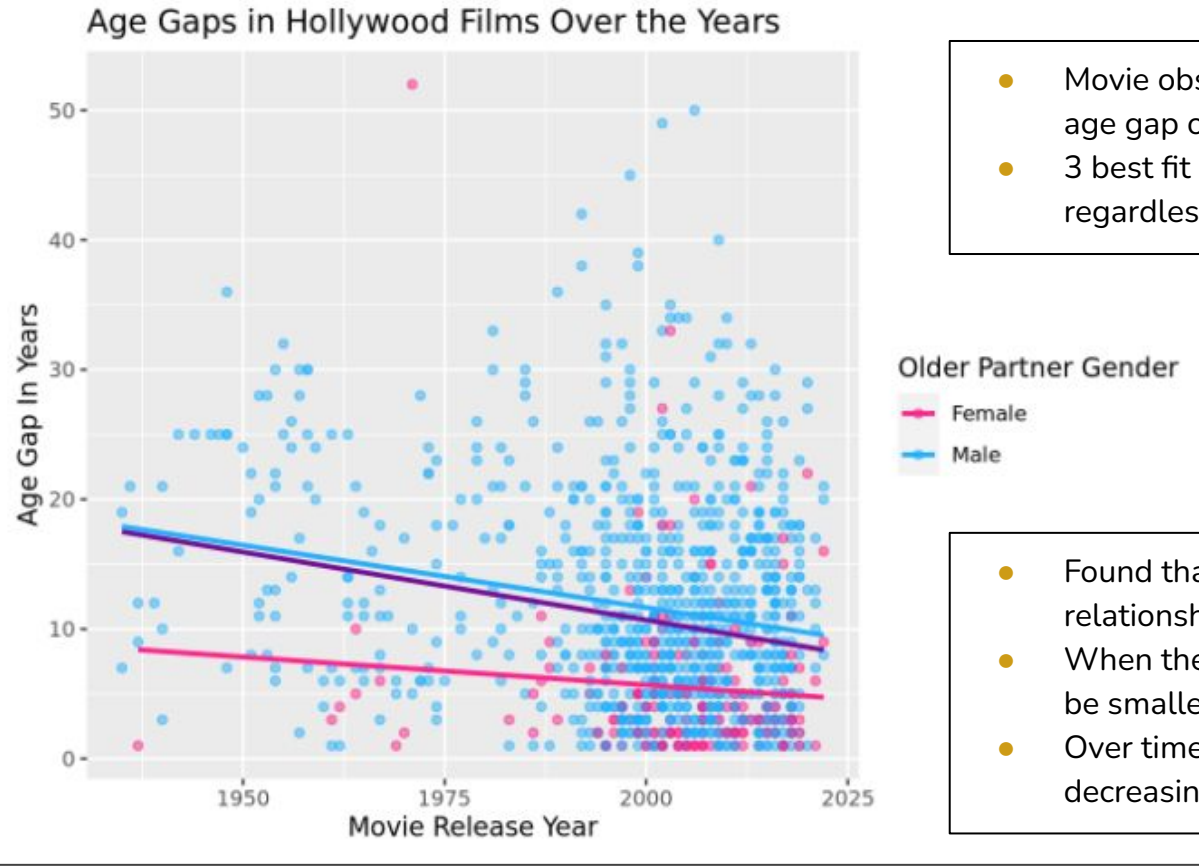
Research Question: Does the age difference between actors in relationships in movies decrease as the release date becomes more modern?

The Data Set

- From Tidy Tuesday repositories on GitHub via the site "Data is Plural."
- **Hollywood Age Gap**
  - Has information from upwards of 630 movies
  - Films from 1935 to 2022
  - Difference between the love interests' ages
  - Collected and published around February 7th, 2018
    - Most recently updated on February 25th, 2023
  - Observations of the data set include:
    - The movie name, director, release year, and then the age difference, the actors' names, their characters' names, their birthdates, and their ages

Methodology & Predictions

- Filtering the data
  - Same sex couples
- Scatter-plot used to plot the age difference in movies over time
  - Prediction: The gap will lessen a bit over time due to the criticism of this issue in the movie industry.
- Hypothesis testing
  - Prediction: Older partner tends to be male
  - Prediction: Proportion of films with large age gaps decreased over time
- Bootstrapping and Central Limit Theorem to compare the true mean age gap in each film era
- Fitting a Linear Regression model to see how a films release year correlates to the age gap in the film



- Movie observations over time where the couple has age gap of at least one year, colored by gender.
- 3 best fit lines: purple represents all couples regardless of older partner gender

- Found that older partners in Hollywood movie relationships tend to be disproportionately male
- When the older partner is female, the age gap tends to be smaller.
- Over time it appears that the mean age gaps are decreasing.

Hypothesis Testing: Gender

Null Hypothesis: the true mean age gap for couples where the older partner is male is the same as the true mean age gap for couples where the older partner is female.

$$H_o : \mu_m - \mu_f = 0$$

Alternative Hypothesis: the true mean age gap for couples where the older partner is male is greater than the true mean age gap for couples where the older partner is female.

$$H_o : \mu_m - \mu_f > 0$$

```
set.seed(0)
null_dist_hetero <- agegaps_hetero |>
  specify(response = age_difference, explanatory = older_partner) |>
  hypothesize(null = "independence") |>
  generate(reps = 1000, type = "permute") |>
  calculate(stat = "diff in means", order = c("Female", "Male"))
```

P-value < 0.001

Decide: reject the null hypothesis

Conclude: strong evidence that the true mean age gap for couples where the older partner is male is greater than it is in couples where the older partner is female.

Bootstrapping/CLT

Pre-1950: 12.9 to 21.1 years
1950-1975: 12.3 to 16.4 years
1975-2000: 13.3 to 15.4 years
Post-2000: 8.9 to 10.0 years

```r
#label: boot-1925
set.seed(1)

boot_1925 <- agegaps_1925 |>
  specify(response = age_difference) |>
  generate(reps = 1000, type = "bootstrap") |>
  calculate(stat = "mean")

boot_1925 |>
  summarize(lower = quantile(stat, 0.025),
            upper = quantile(stat, 0.975))
```

```r
#label: CLT-1950
set.seed(2)

est_mu_1950 <- 14.35632
est_sigma_1950 <- sd(agegaps_1950$age_difference , na.rm = T) / sqrt(87)

qnorm(c(0.025, 0.975), est_mu_1950, est_sigma_1950)
```

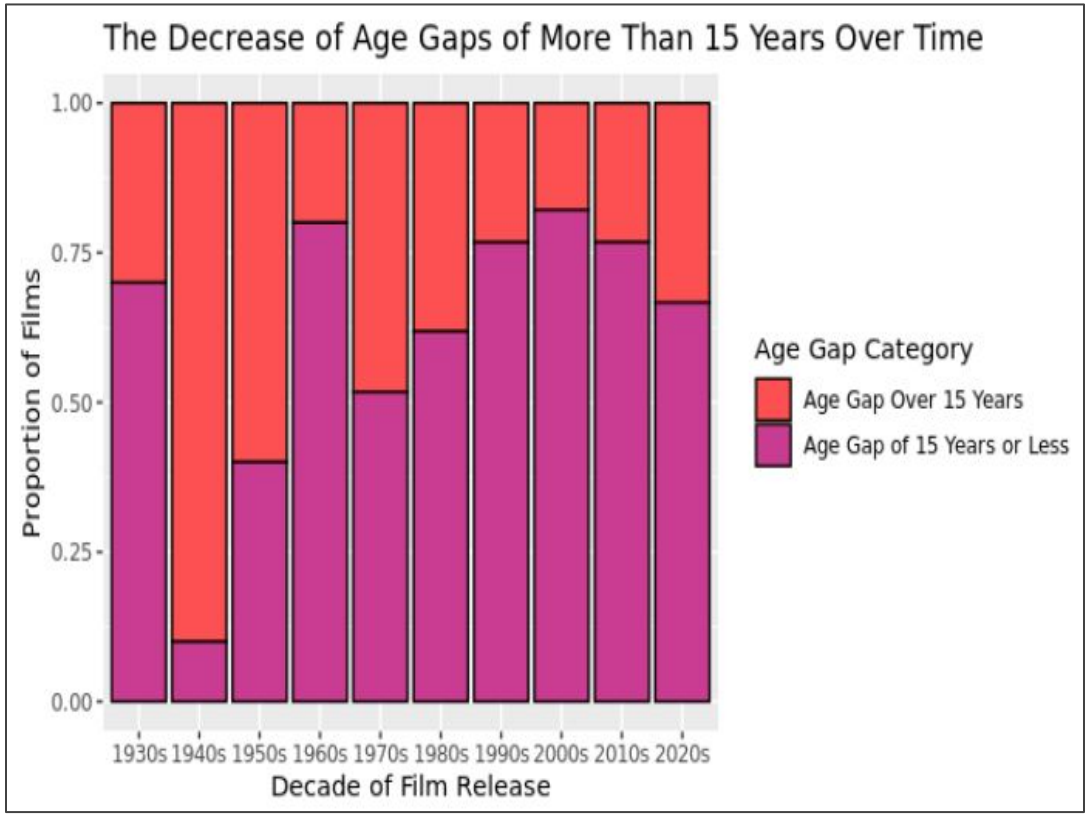
Linear Regression Model

$$\widehat{AgeDifference} = 231.0808 - 0.1103 * release\_year$$

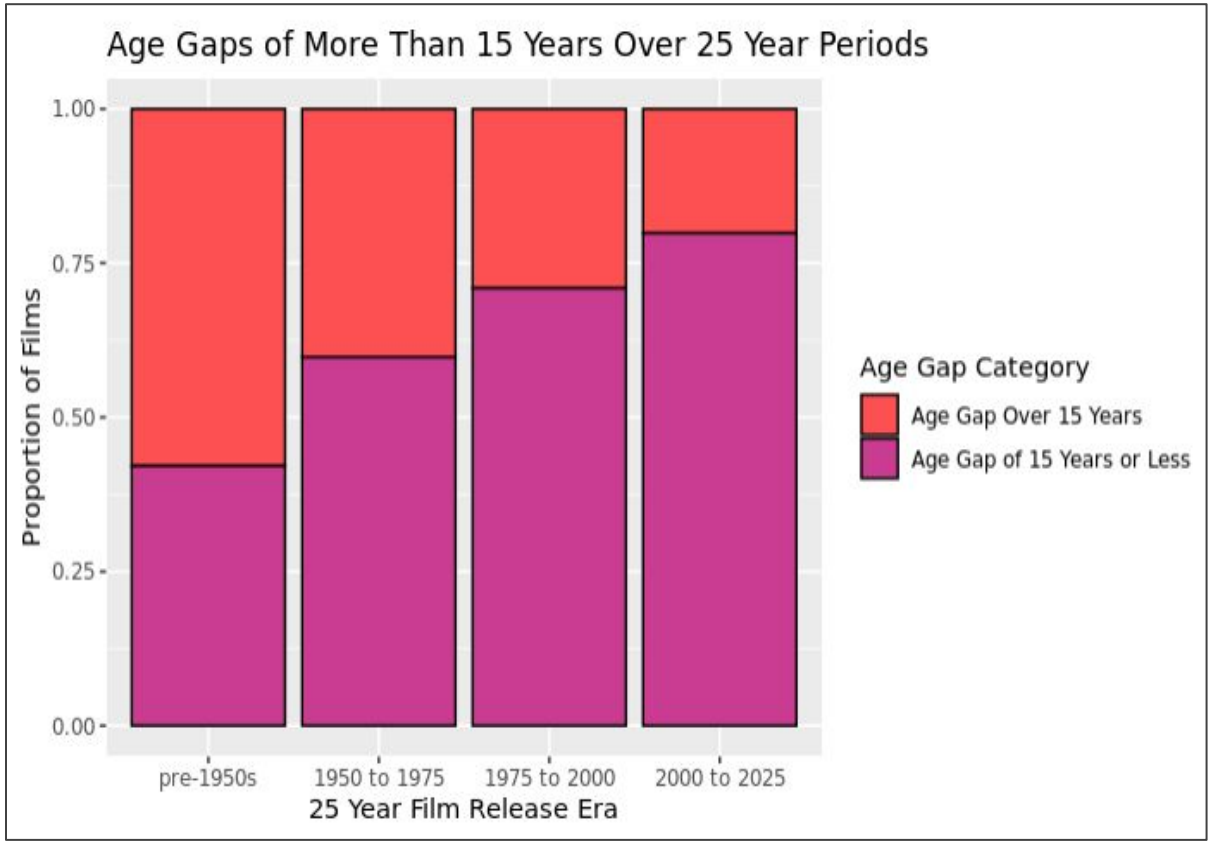
R-squared: 0.04524371

Model with release year explains roughly 4.5% of the variation in age gaps

Year has a negative coefficient

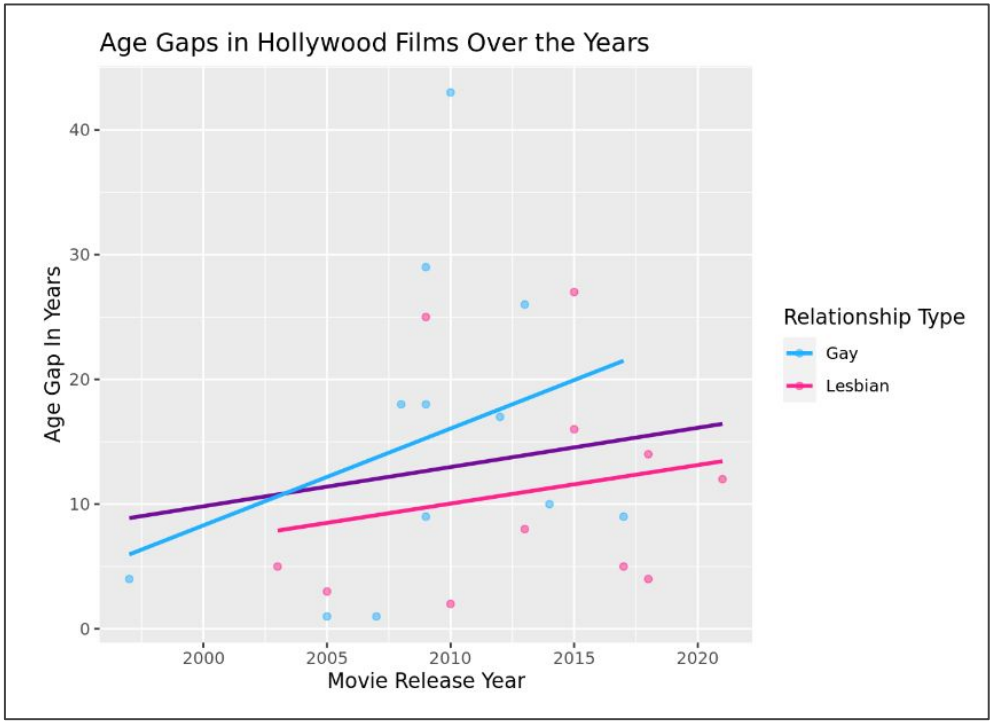


- Movie observations into the decade that they were released
- Over time, proportion of films containing age gaps of over 15 years decreases
- 1940s: over 80% of films had a large age gap of over 15 years
    - Notable that World War II occurred during most of this decade
    - Age gaps over 15 years could've been due to younger men fighting in war rather than acting
- Not many data points for the 1930s decade
    - Movies were just becoming more prevalent during this time period
- 2020s: around 30% of films contain a couple age gap of over 15 years



- 25 year eras
- Hypothesis Testing:
  - Over each 25-year period, the proportion of films containing age gaps of over 15 years decreases.
- Pre-1950: Over 50% of films having large age gaps of over 15 years
- 2000-2025: less than 25% of films contain a couple with an age gap of over 15 years

Same-sex couples



- Age gaps appear to have increased over time
- Only 23 observations (small sample size)
- Can see that overtime more films with same sex couples have been made

Conclusions

- Age difference between actors whose characters are in relationships decreases as time progresses
- As time progresses, the proportion of films with age gaps of over 15 years has decreased consistently.
- It is significantly more common for the older partner in these couples to be male but in cases where the older partner is female, the age gap tends to be smaller.

THANk you for your time