HOW DOES

HOLLYWOOD

MAKE THE BIG BUCKS?

FT. ZUCK(R)BERG

Research Question:
Which factors
contribute most to a
more profitable movie?

Data

Objectives

"TMDB 5000 Movies"

Collected via "Movie Database API"

```
## Observations: 4,505
## Variables: 20
## $ budget
## $ genres
## $ homepage
## $ id
## $ keywords
## $ original_language
## $ original_title
## $ overview
## $ popularity
## $ production_companies
## $ production_countries
## $ release_date
## $ revenue
## $ runtime
## $ spoken_languages
## $ status
## $ tagline
## $ title
## $ vote_average
## $ vote_count
```

1. Define profitability

2. Eliminate unnecessary variables

3. Determine which factors are most influential on profitability

4. Model the influence of our variables

Data Problems

Budget

```
## # A tibble: 1 x 2
##   title        budget
##   <chr>         <dbl>
## 1 Bran Nue Dae      7
```
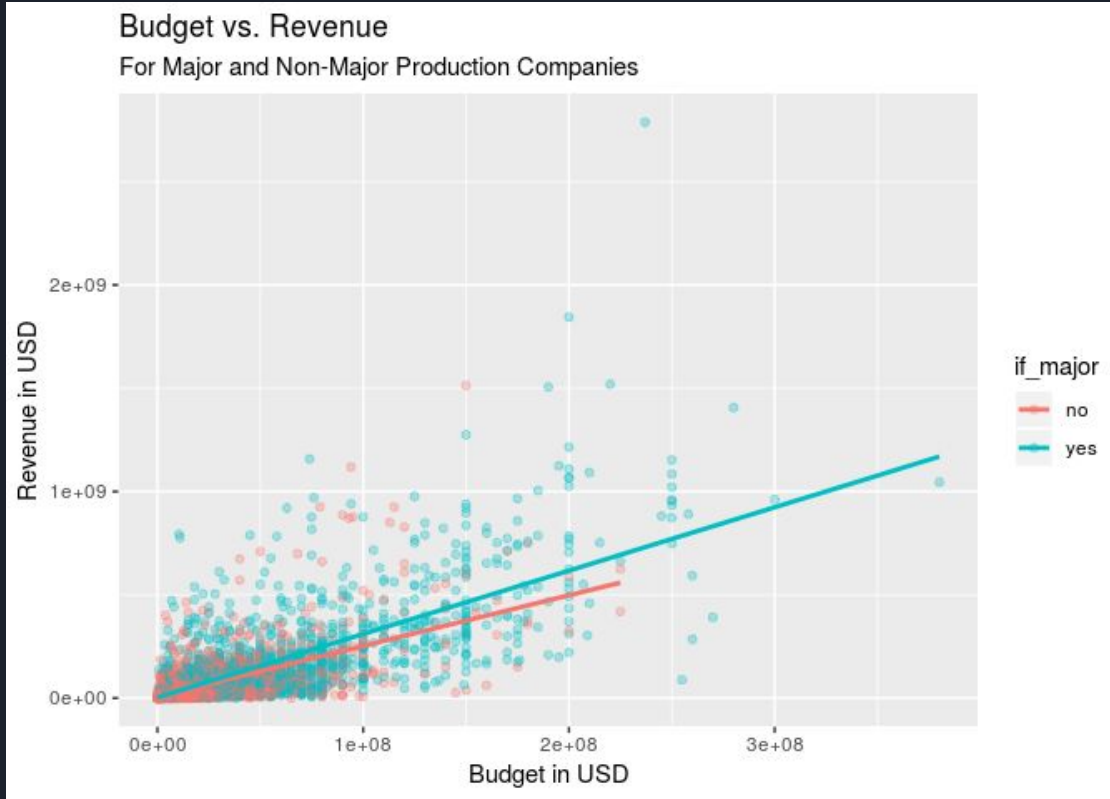
Bran Nue Dae / Budget

7 million AUD

- Some budgets are not in USD
- Inconsistent Units

Budget and Revenue

Budget vs. Revenue
For Major and Non-Major Production Companies

Revenue =
143491
+ 2.98 (Budget)

R-squared:

0.497329

Variable Creation - Profitability

prof·it·a·bil·i·ty
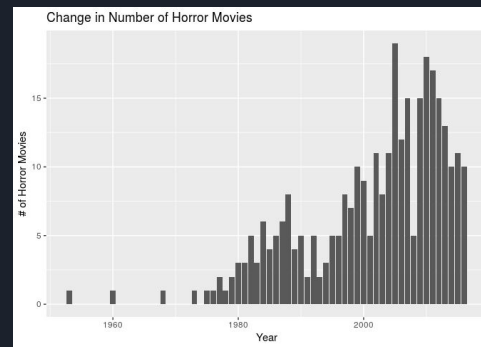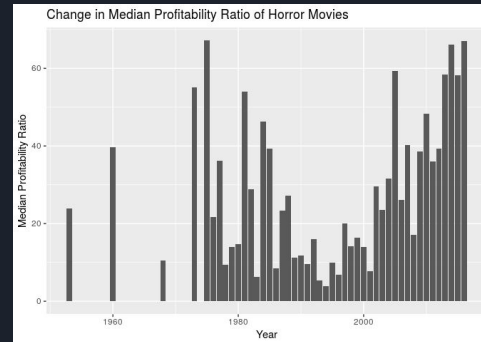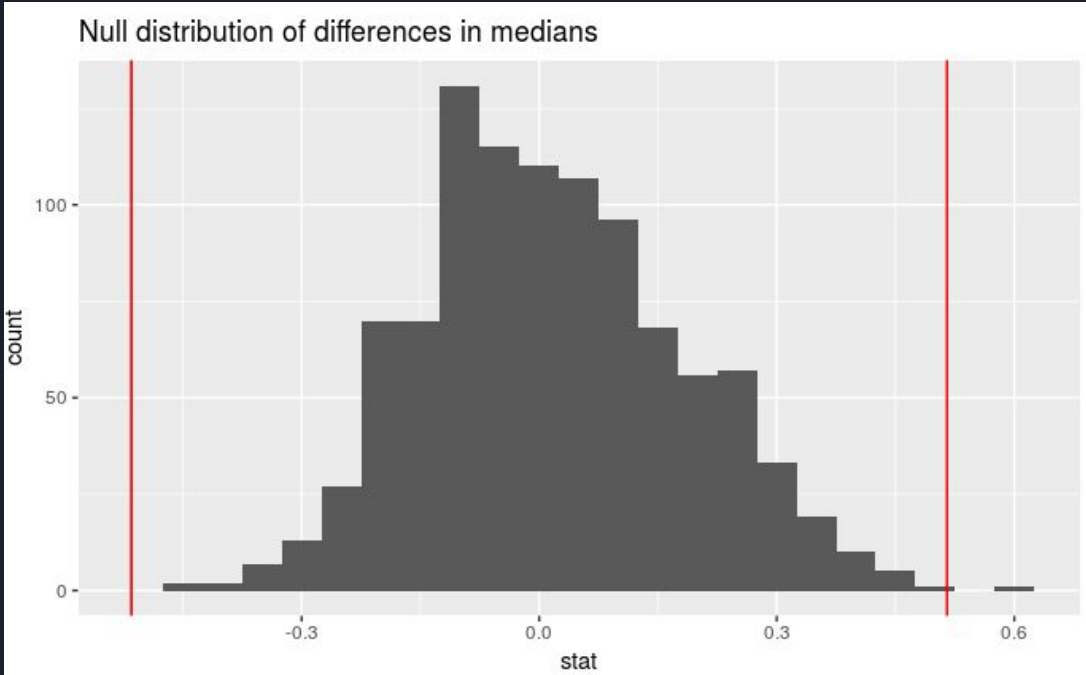
/ˌpräfədəˈbilədē/

noun

the degree to which a business or activity yields profit or financial gain.
"profitability may not improve until well into next year"

- the state of yielding profit or financial gain.
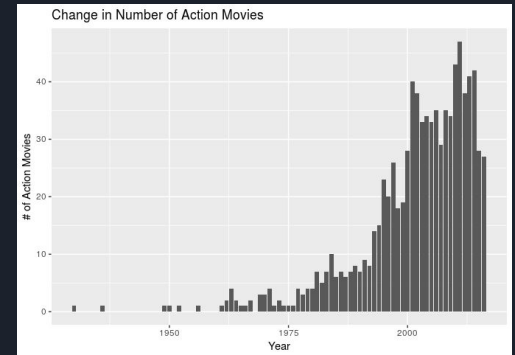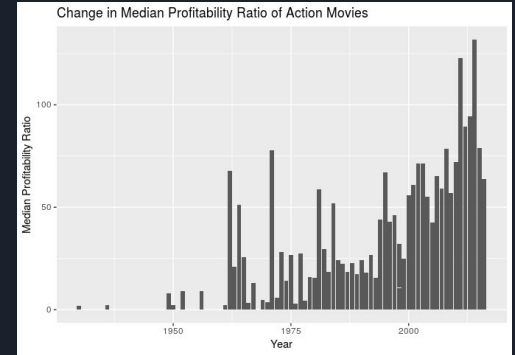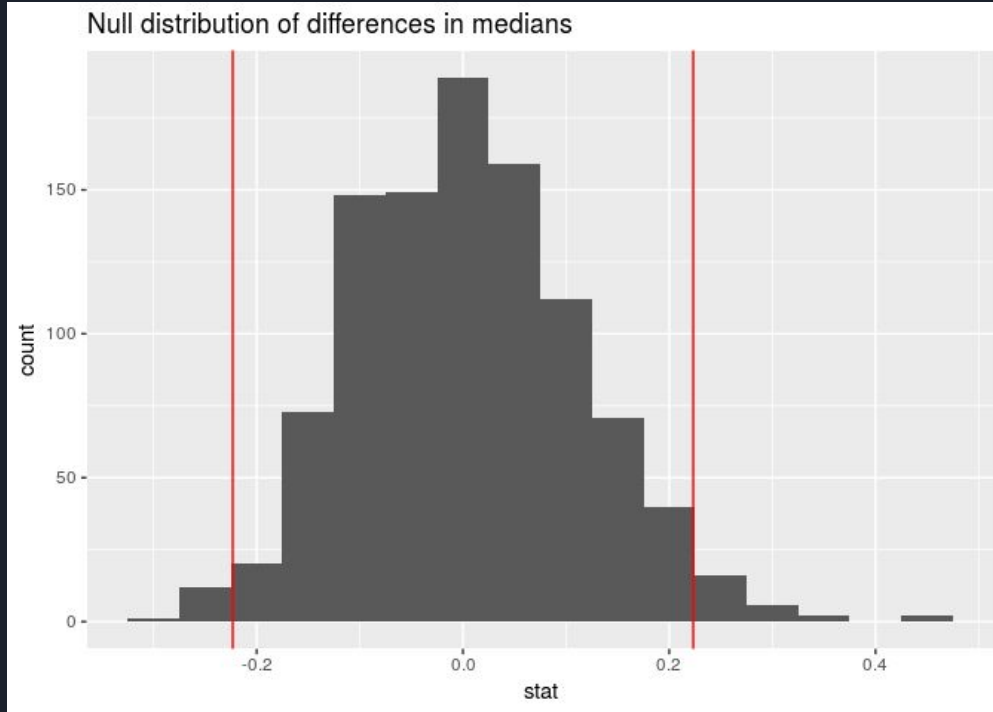"growing sales and a return to profitability"

```
## # A tibble: 1 x 2
##    lower uppper
##    <dbl>  <dbl>
## 1  2.20   2.39
```

- Numerical

- "pratio"

- Filtering out erroneous entries

- 95% Confidence Interval
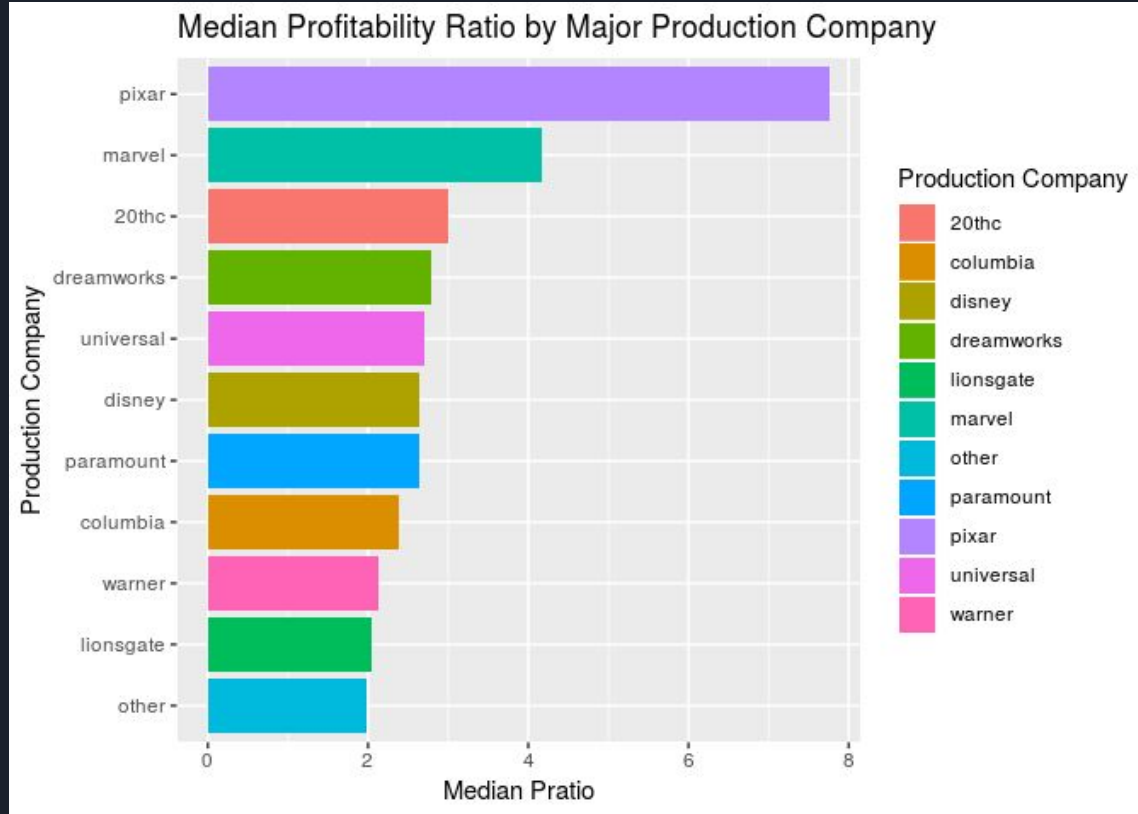
Horror Movie Trends

Action Movie Trends

Null distribution of differences in medians

Change in Median Profitability Ratio of Action Movies

Change in Number of Action Movies

Major Production Companies and Profitability

Hypothesis Testing
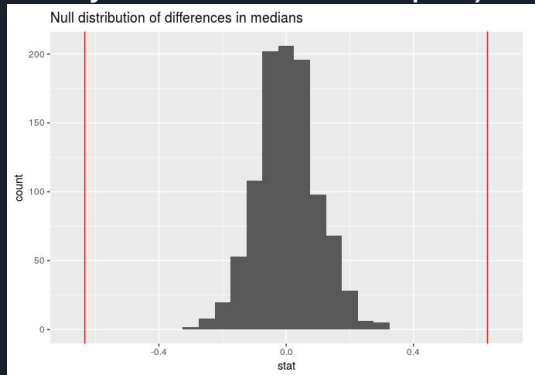
Questions for simulation:

Is there a difference in median profitability ratios for films that were made under major production companies and those that weren't?
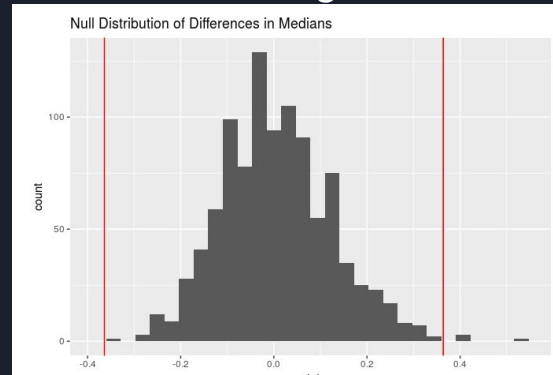
Is profitability ratio independent of title length?

Results:
All factors significantly affected pratios.
(All had p-values < .05)

Major Production Company

Title Length

Predicting Profitability Ratio: Initial Linear Model

- Holiday Release (-)
- Likely Sequel (-)
- Budget (-)
- If Major Production Co. (-)
- Production Co. Categories:
  - Pixar (+)   Lionsgate (-)
  - Columbia (-)   Marvel (+)
  - Disney (+)   Paramount (+)
  - DreamWorks (+)   Universal (+)
  - 20th Century Fox (+)
- Tagline Length (-)
- Spoken Language English (+)
- Runtime (+)
- Action (-)
- Horror (+)
- Few Word Title (+)
- Budget * Major Production Company (+)
- Likely Sequel * Budget (+)

Adjusted R-squared:

0.0407478

Predicting Profitability Ratio: Final Linear Model

- Holiday Release (-)
- Likely Sequel (-)
- Budget (-)
- If Major Production Co. (-)
- ~~Production Co. Categories:~~
  - ~~Pixar (+)~~   ~~Lionsgate (-)~~
  - ~~Columbia (-)~~   ~~Marvel (+)~~
  - ~~Disney (+)~~   ~~Paramount (+)~~
  - ~~DreamWorks (+)~~   ~~Universal (+)~~
  - ~~20th Century Fox (+)~~
- ~~Tagline Length (-)~~
- Spoken Language English (+)
- Runtime (+)
- Action (-)
- Horror (+)
- ~~Few Word Title (+)~~
- Budget * Major Production Company (+)
- Likely Sequel * Budget (+)

Adjusted R-squared:

0.0425732*

*Increase of 0.0018254

Profitability Ratio =

$$2.95$$
$$-\ 0.92 \text{ (Holiday Release)}$$
$$-\ 1.46 \text{ (Likely Sequel)}$$
$$-\ 6.47^{-8} \text{ (Budget)}$$
$$+\ 0.22 \text{ (Major Production Co.)}$$
$$+\ 0.69 \text{ (Spoken Language English)}$$
$$+\ 0.03 \text{ (Runtime)}$$
$$-\ 0.65 \text{ (Action)}$$
$$+\ 1.89 \text{ (Horror)}$$
$$-\ 1.16 \text{ (Budget*Major Production Co.)}$$
$$+\ 2.49 \text{ (Likely Sequel*Budget)}$$

Conclusion

Most prominent factors that affected pratio:

- Production company
- Genre
- Title length
- Budget

Surprising Takeaways:

- Trends only visible when looking at individual variables
- The effect of certain variables against others
- Movie success is unpredictable