

Controlling Bias in Online Movie Reviews

Reweighting IMDb Ratings to the U.S. Census

The Problem

The Problem

review bombing • an Internet phenomenon in which large groups of people leave negative user reviews for a product or work online in an attempt to harm their sales and/or popularity



'A Wrinkle in Time' Proves Rotten Tomatoes User Reviews and IMDb Ratings are Broken

Movie-goers should be allowed to express their opinions on movies, but what happens when fake reviews flood the web's biggest sites?

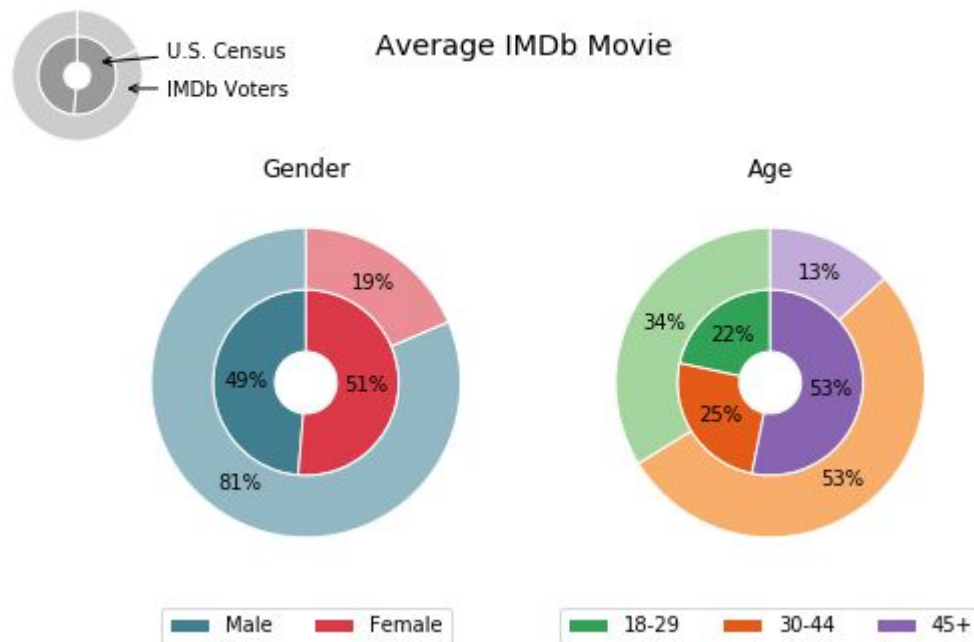


'Captain Marvel' Praised in Strong First Press Reactions, So Ignore Those Fake User Reviews

Listen to film critics and journalists. Not trolls on Rotten Tomatoes who have yet to see the movie.

The Problem

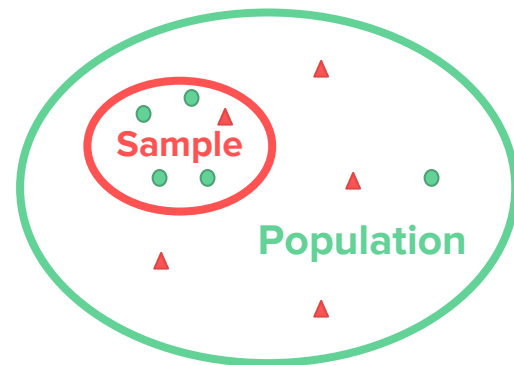
Essentially, the opinions shared on IMDb, Rotten Tomatoes, and other review aggregators are dominated by *young men*



The Problem

In academia and political polling, when the sample distribution does not reflect the overall population, it is referred to as ***sampling bias***

- Example:
 - The known population is 10 million men, 10 million women
 - The sample distribution is 800 men, 200 women
 - Therefore: the sample *overrepresents* men, *underrepresents* women



Specifically, for online reviews, we are dealing with *self-selection bias* — users actively choose to participate in IMDb or other review sites because they are invested in a topic or community

The Remedy

The Remedy

Survey Weighting

In academia, the correction for sampling bias is **survey weighting**, where underrepresented groups are valued higher and overrepresented groups are valued lower to align the known demographic distribution of the overall population

- Example:
 - Known population: 10 million men, 10 million women (50%-50%)
 - Sample population: 800 men, 200 women (80-20%)
 - Men will be weighted *down* by 0.625 (= 50% / 80%)
 - Women will be weighted *up* by 2.500 (= 50% / 20%)

Multiple demographic variables can be weighed simultaneously with the survey weighting method, *iterative proportional fitting*, otherwise known as **raking**

The Data

The Data


- **Known Population Distribution**

- 2017 American Community Survey, Age and Sex Distribution
- Collected and updated annually by the U.S. Census Bureau

Subject	United States											
	Total		Percent		Male		Percent Male		Female		Percent Female	
	Estimate	Margin of Error	Estimate	Margin of Error	Estimate	Margin of Error	Estimate	Margin of Error	Estimate	Margin of Error	Estimate	Margin of Error
Total population	325,719,178	*****	(X)	(X)	160,402,504	+/-26,371	(X)	(X)	165,316,674	+/-26,371	(X)	(X)
AGE												
Under 5 years	19,795,159	+/-20,249	6.1%	+/-0.1	10,138,447	+/-17,217	6.3%	+/-0.1	9,656,712	+/-16,582	5.8%	+/-0.1
5 to 9 years	20,095,947	+/-62,409	6.2%	+/-0.1	10,223,398	+/-42,853	6.4%	+/-0.1	9,872,549	+/-41,408	6.0%	+/-0.1
10 to 14 years	21,130,768	+/-56,442	6.5%	+/-0.1	10,854,851	+/-38,203	6.8%	+/-0.1	10,275,917	+/-41,176	6.2%	+/-0.1
15 to 19 years	21,497,166	+/-38,471	6.6%	+/-0.1	11,003,862	+/-26,744	6.9%	+/-0.1	10,493,304	+/-26,494	6.3%	+/-0.1
20 to 24 years	21,950,055	+/-33,648	6.7%	+/-0.1	11,280,143	+/-23,368	7.0%	+/-0.1	10,669,912	+/-22,910	6.5%	+/-0.1
25 to 29 years	23,085,789	+/-27,138	7.1%	+/-0.1	11,720,804	+/-18,237	7.3%	+/-0.1	11,364,985	+/-18,237	6.9%	+/-0.1

- **Sample Distribution**

- IMDb Rating by Demographic
- Collected for all 4,608 films eligible for the Top 250 (films with >25,000 votes)

 The Wizard of Oz ★★★★★ (5634) 101min 1939 PG					
Rating By Demographic					
	All Ages	<18	18-29	30-44	45+
All	8.0 347,829	7.9 383	7.9 79,878	8.0 125,972	8.4 50,546
Males	8.0 195,653	7.9 282	7.9 53,709	7.9 93,387	8.4 39,183
Females	8.2 71,333	7.8 86	8.0 25,356	8.2 31,283	8.6 10,729

The Data

Data Collection

- written in Python with the library IMDbPY, created by Davide Alberani
- Here's a taste of my code, more of which can be found on my Github:

```
#Import Libraries
!pip install git+https://github.com/alberanid/imdbpy
from imdb import IMDb, IMDbError

#Collect Demographic Info for each Movie
i = IMDb(accessSystem='http')
for m in movies:
    demographics=i.get_movie(m,'vote details').get('demographics')
    demo_clean = pd.DataFrame(demographics)[demo_list]

    #Reweigh Ratings to the Census Demographic Distribution
    demo_clean = pd.merge(demo_clean, us_targets[['gender', 'age', 'percent_census']], \
on = ['gender', 'age'])
    rating_census = demo_final['rating'].dot(demo_final['percent_census'])
```

The Results

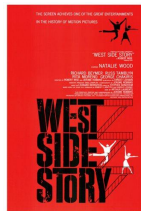
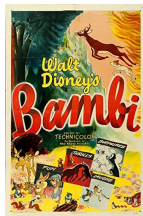
The Results

Title	Year	Original Rating	Weighted Rating	Original Rank	Weighted Rank
The Shawshank Redemption	1994	9.29	9.16	1	1
Schindler's List	1993	8.94	8.85	7	4
12 Angry Men	1957	8.94	8.83	6	5
...
Key Largo	1948	7.90	8.01	500	247
Spring, Summer, Fall, Winter, ... and Spring	2003	8.03	8.01	347	248
All About My Mother	1999	7.91	8.01	492	249
I, Daniel Blake	2016	7.89	8.01	507	250

The Results

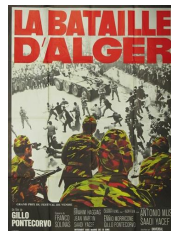
When looking at all 4,608 eligible films, classic family films or films based on racial/ethnic topics previously targeted by review bombers earned the greatest ranking boost, and foreign films suffered the greatest ranking decline

Most Positive Ranking Change



- The Promise (2016) 1778 ↑ 342
- 42 (2013) 1132 ↑ 638
- Radio Days (1987) 1007 ↑ 515
- Bambi (1942) 1321 ↑ 885
- West Side Story (1961) 968 ↑ 543

Most Negative Ranking Change



- The Battle of Algiers (1966) 279 ↓ 973
- Ikiru (1952) 157 ↓ 851
- Infernal Affairs (2012) 358 ↓ 971
- My Sassy Girl (2001) 232 ↓ 814
- Sin City (2005) 301 ↓ 849

The Results

Films with larger female and mature viewer bases typically increased in rating, while more younger, male-dominant films decreased in rating

Quantile	% Female Vote Range	Average Rank Change
1	5% - 11%	↓ 117
2	11% - 15%	↓ 29
3	15% - 19%	↑ 26
4	19% - 26%	↑ 56
5	26% - 71%	↑ 65

Correlation: 0.242

Quantile	% Age 45+ Vote Range	Average Rank Change
1	2% - 10%	↓ 169
2	10% - 13%	↓ 13
3	13% - 16%	↑ 18
4	16% - 21%	↑ 57
5	21% - 48%	↑ 107

Correlation: 0.354

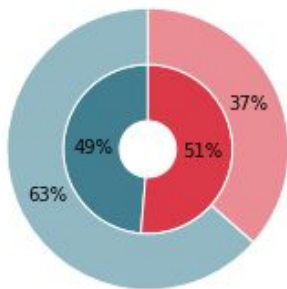
The Results

My Big Fat Greek Wedding (2002)

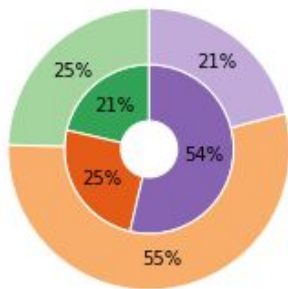


Mandy (2018)

Gender



Age

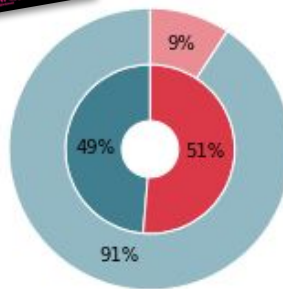


Male Female

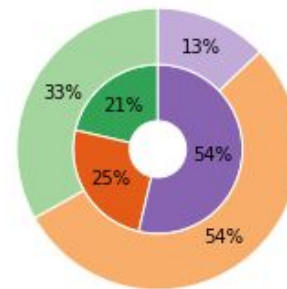
18-29 30-44 45+

- Unweighted Rating: 6.62
- Weighted Rating: 6.87
- **Rank** ↑ 499

Gender



Age



Male Female

18-29 30-44 45+

- Unweighted Rating: 6.58
- Weighted Rating: 6.00
- **Rank** ↓ 972

The Results

Foreign films received less reweighted love than English-language films. Particularly, films produced in Asian languages saw a more significant decline in reputation than foreign-films produced in western speaking countries.

Language	# Films	Average Rank Change
Hindi	73	↓ 762
Japanese	54	↓ 163
Korean	20	↓ 55
French	85	↓ 29
English	4,148	↑ 20
German	28	↑ 30
Spanish	45	↑ 50
Swedish	14	↑ 71
Norwegian	5	↑ 135

The Results

Genre	# Films	Average Rank Change
Sport	160	↓ 48
Comedy	1,681	↓ 37
Crime	959	↓ 26
Horror	586	↓ 2
Romance	966	↑ 9
Fantasy	635	↑ 50
Family	481	↑ 61
History	189	↑ 86
Film-Noir	15	↑ 116

Genres associated with younger audiences (family, fantasy, adventure, animation) and genres associated with older film eras (film noir), received the most significant rating boost

The Results

Generally speaking, the older the film, the greater the rating boost (with the exception of silent 1920s films which modern audience find difficult to connect with). This typically means that older films are undervalued by IMDb's current unweighted rating system.

Decade	# Films	Average Rank Change
1920s	10	↓ 150
1930s	23	↑ 86
1940s	33	↑ 169
1950s	68	↑ 52
1960s	94	↑ 75
1970s	156	↑ 7
1980s	383	↑ 16
1990s	757	↓ 8
2000s	1,571	↓ 5
2010s	1,513	↓ 6

The Results

74 Films were shuffled in and out of the IMDb Top 250

New Top 250 Entries

- The Lion in Winter (1968)
- Kind Hearts and Coronets (1949)
- Brief Encounter (1945)
- Philadelphia Story (1940)
- La Strada (1954)

Disappeared Top 250 Titles

- 3 Idiots (2009)
- The Wages of Fear (1953)
- La Haine (1995)
- Barry Lyndon (1975)
- The Thing (1982)

Demographic	Reweighted 250	Original 250
Avg % Female Votes	18.1%	16.3%
Avg % Age 45+ Votes	18.3%	14.9%
# of Foreign-Language Films	48	66
# of 1940s Films	25	12
# of 2000s Films	38	50

The Results



Unweighted: 8.11

Weighted: 8.26

Rank: 322 ↑ 86

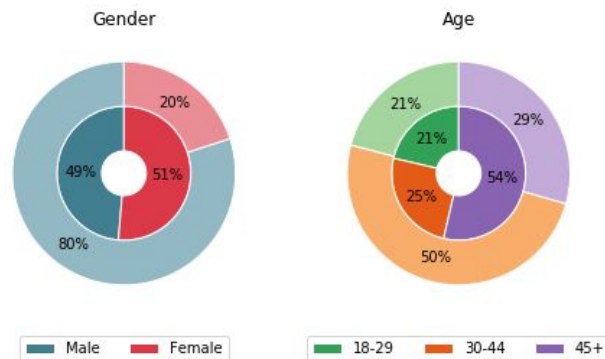


Unweighted: 8.05

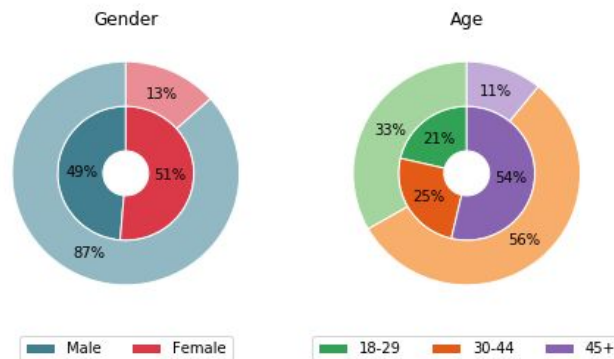
Weighted: 7.53

Rank: 229 ↓ 1009

Brief Encounter (1945)



La Haine (1995)



The Next Steps

The Next Steps

Takeaways:

- Be aware of the demographic bias embedded in simple statistics that we can take for granted
- Products that appeal to women and older audiences are undervalued by online review aggregators, like family films and classic films on IMDb
- Reweighing is accompanied with its own biases: older audiences are less open to foreign films than younger audiences


Next Steps:

- Work with more detailed user demographics: include nationality to limit to methodology U.S. users, adjust 'known' population to match internet-users, include race variable in user demographics to ensure minority voices are represented
- Monitor and reweigh ratings of new film releases in real-time, possibly partnering with [CherryPicks](#)
- Encourage more online review aggregators to collect user demographics and weigh user reviews appropriately

The Next Steps



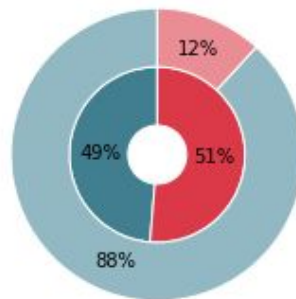
Captain Marvel

- Unweighted Rating: 6.48
- Weighted Rating: 6.92
- Rank:  835

Rating By Demographic

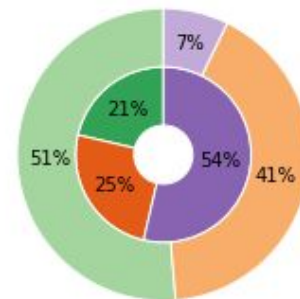
	All Ages	<18	18-29	30-44	45+
All	6.4 26,725	7.3 191	6.7 8,111	6.2 6,551	5.9 1,137
Males	6.2 16,141	7.1 158	6.5 6,752	6.0 5,735	5.8 990
Females	7.8 2,168	7.9 10	7.9 1,056	7.8 667	7.6 111

Gender



Male Female

Age



18-29 30-44 45+

Kerry Driscoll

Data Scientist

kerry.driscoll@claritas.com

[LinkedIn](#) | [Github](#)