BECHDEL TEST

A RELATIONAL APPROACH TO GENDER INEQUALITY

Alice Anna Maria Brunazzi, Alessandro Della Beffa, Daniele Lepre aa. 2023/2024

Abstract

The Bechdel Test is a sequence of three questions designed to assess the presence of women in movies. It tries to highlight the under representation and the overall gender asymmetry in fiction, using easily detectable metrics to assess the gender bias.

This study investigates the application of the Bechdel test to a selection of 20 Oscar-nominated films, aiming to automate the evaluation of gender representation within their scripts. The research methodology involved the creation and systematic analysis of a comprehensive dataset, encompassing casts and dialogues.

CONTENTS

INTRODUCTION

The film industry stands as a powerful cultural force, shaping perceptions and influencing social interactions across the globe. It serves not only as a source of entertainment but also as a mirror reflecting the complexities of the human experience. In recent years, the call for greater representation has intensified, collecting different opinions and representation- based metrics of evaluation. The Bechdel Test, devised by Alison Bechdel, provides a set of criteria for gender portrayal in films, by assessing if they feature at least two named female characters who engage in a conversation about something other than a male character. The objective of this study was to automate the analysis of 20 films, spanning multiple genres to ascertain the prevalence and characteristics of those conforming to Bechdel Test criteria. The list of movies analyzed can be found in Appendix 1.

Films were selected among the ones nominated for an Oscar, as this was considered an appropriate metric to indicate that a film was well-received by both the public and critics. Additionally, it is noted that none of the selected films have female directors.

Utilizing a systematic data management approach, various sources were leveraged to create a dataset, encompassing various facets of each film, including but not limited to cast, dialogue, and thematic content. Methodologically, the research involved the creation, categorization, and analysis of this dataset, employing both qualitative and quantitative methodologies to derive meaningful insights.

DATA COLLECTION

RESEARCH QUESTIONS

The research question that shaped this analysis adheres to the criteria outlined by the Bechdel test, which served as a guiding framework. The whole concept of this test is enclosed in the following text of a Bechdel's comic:

'I only go to a movie if it satisfies three basic requirements. One: it has to have at least two women in it, who, two, talk to each other about, three, something besides a man'[1] *(Figure 1)*

The analysis was therefore conducted in order to select the films that positively met the following criteria:

i. Does it contain two women?
ii. Do the (at least) two women alone interact with each other in al least a scene?
iii. Is the interaction about something else than a man?

The primary intent of this project was to automate the response process.

DATA SOURCES- DATA EXTRACTION

The data collected for this project encompasses various components sourced from multiple online platforms. The components were categorized into script data, gender, actors, and additional pertinent details. In details:

- Scripts: crucial data for the analysis of dialogue and scene content, the scripts were retrieved from the *Internet Movie Script Database (IMSD)*[2], using a custom extraction function, developed to obtain the script text for each title, then parsed and divided into different scenes, *The Beautiful Soup* [3] library was used for web scraping.
- Cast: information about the cast helped identify the actors and their respective roles and was sourced from *Internet Movie Database (IMD)*[4]. Utilizing web scraping techniques, the full credits page of each film was accessed to extract the cast list, including data inherent to the names and the corresponding roles.
- Gender: the determination of the gender of actors was performed utilizing the *Names Dataset Library*[5], and NLP techniques.

DATA PROFILING

STATISTICS

At the end of the importing phase, the dataset presents:

- 3012 scenes
- For each film, the average number of scenes is 150,6

The total amount of columns is, precisely:

- MOVIE table; 3 columns, 20 rows
- SCRIPT table: 9 columns, 3012 rows
- CAST table: 4 columns, and 876 rows

DATA SOURCES- DATA EXTRACTION

Given the nature of the analysis, missing data was not a problem, but checks on data integrity and quality were required. This involved cross-referencing data across multiple tables to ensure coherence and accuracy, thorough scrutiny of data entries, and validation against

[1] Dykes to Watch Out For' (1985, Bechdel Alison)

[2] https://imsdb.com/

[3] https://tedboy.github.io/bs4_doc/index.html

[4] https://www.imdb.com/

[5] https://pypi.org/project/names-dataset/

external sources (see footnote 2 and 4) where applicable.

Any discrepancies or irregularities discovered during this process were meticulously resolved, as explained in the *Operational Guide for Reproducibility*, to maintain the integrity and reliability of the dataset.

Outliers were not a concern.

DATA STORAGE

An *SQLite database* [6] *engine*[7] is created to store the data collected during script execution.

The choice of a relational SQL database for this project stems from several reasons. Primarily, the need of a flexible, powerful, and user-friendly language, was a key consideration. Additionally, it was chosen for the efficiency of the queries and the small size of the database.

The process initiates by creating tables within the database to accommodate diverse data types. This tables, namely 'MOVIE', 'SCRIPT' and 'CAST' were designed to capture the corresponding attributes, ensuring data integrity and efficient querying through the keys included in every table.

The SCRIPT table was designed to allow a scene-based analysis, segregating each scene using indicators of interior or exterior settings. The method was implemented with the idea of facilitating subsequent data insertion.

The CAST table contains details about the actors, the roles and the gender.

INTERLINKING BETWEEN TABLES

The database tables are connected via primary keys and foreign keys to establish relationships between the different data. The MOVIE table, that contains general information about the film, such as the title and the release year, is linked to the other tables using the 'film_id' field as a foreign key.

The 'CAST' dataset, retrieved raw by the IMD, gave information about the role of each actor, allowing us to determine the gender of the role using the first name of the actor, validated by the name's dataset library.

SCRIPT table- COUNT columns

To achieve the project's goals, it was necessary to find a way to create columns to facilitate the script analysis.

Once the data from each movie's script was loaded and cross-linked with the corresponding cast (and, obviously, the gender of each actor), each scene underwent a double analysis:

- Uppercase analysis: the names written in uppercase correspond to the character who is talking, so, by collecting and counting the unique uppercase world, the column 'personaggi' was created, and then used to compute the sum of female and male characters speaking (second level of the Bechdel test.)
- Lowercase analysis: to understand the object and subject of the conversation, it was also necessary to do the same analysis using a lowercase-dictionary, to detect if a known name was mentioned in the scene, but only if male (third level of the Bechdel test just require the presence of a male).

It is important to have a comprehensive dictionary of womanly and manly terms that the script might use, so, in order for both test to be exhaustive, other words were taken into account. The complete list of words used in the analysis can be found in Appendix 2.

DATA PREPARATION

The collected data underwent a preparation and cleaning process to ensure quality and consistency.

The data obtained at this point presents discrepancies, mainly because each script was written differently, not conformed to a specific standard, and some roles presented titles, terms of endearment and other distorting characteristics.

To resolve some of the issues:

– The script text underwent text preprocessing techniques to enhance suitability for analysis, such as removal of HTML tags, special characters, and not-needed white spaces. The scene descriptions were standardized to facilitate uniformity and comparability across the dataset. The only common thing that each script presented was the location indicator (i.e., INT/EXT), so the whole scene segregation was made exploiting this similarity.
– Cast information underwent standardization to rectify inconsistencies in actor names and role description.
The employed techniques are mainly string cleaning related, minimizing discrepancies in the dataset.

[6] https://www.sqlite.org/index.html,

[7] https://www.sqlalchemy.org/

– Gender inference involved verification and validation to mitigate inaccuracies. Each actor's first name is subjected to probabilistic analysis using the names dataset, which contains historical data[8] on name-gender associations, statistical algorithms analyze the probability distribution of gender associations for each name, determining the most likely gender for each actor. The inferred gender information is then associated with the corresponding actor in the dataset, enriching the dataset with gender-related insights for subsequent analysis.

DATA ANALYSIS

Following the strictly ordered framework provided by the Bechdel Test, the analysis delved into characterizing the nature of the female interactions within the films.

First Requirement

To satisfy the first test parameter, the answering process involved querying the database to identify films that met the 'at least two women' criterion. This step allows to select the films that will be further analyzed in the following steps.

To rule out the films that do not have at least two women in the cast the following query is performed:

```
SELECT m.titolo, 'PASS' AS at_least_2_woman
FROM MOVIE m
JOIN (
   SELECT film_id
   FROM CAST
   WHERE LOWER(gender) = 'female'
   GROUP BY film_id
   HAVING COUNT(*) >= 2
) c ON m.film_id = c.film_id;
```

Second Requirement

Once the films without two (at least) women were eliminated in the analysis, the second requirement can be tested.

To rule out which film does not have a scene in which two women are speaking alone, the required query is:

```
SELECT m.film_id, m.titolo, m.anno
FROM MOVIE m
JOIN SCRIPT s ON m.film_id = s.film_id
WHERE s.count_female >= 2
AND s.count_male=0;
```

[8] Provides information about

Third Requirement

Finally, if a movie passed the first and second requirement, for it to be a suitable candidate to pass the Bechdel test is necessary that, in the scene in which two women are speaking to each other, no man is mentioned in the dialog.

The process to assess this requisite is similar to the aforementioned, but the focus this time will be on the number of men involved in the scene. It is important to note that this requirement also provides an ultimate filter to select the scenes in which no man is spoken to, speaks, is mentioned or, finally, is present.

The following query is performed:

```
SELECT m.film_id, m.titolo, m.anno
FROM MOVIE m
WHERE EXISTS (
   SELECT 1
   FROM SCRIPT s
   WHERE s.film_id = m.film_id
     AND s.count_male = 0
     AND s.count_female >= 2
     AND s.count_male_nom=0)
```

More Analysis

It is possible to further explore and expand the analysis performed, to gain other insight on the topic.

For example, we can investigate the female presence within the cast.

```
SELECT
   (SELECT COUNT(*) FROM CAST WHERE gender = 'Male') AS male_count,
   (SELECT COUNT(*) FROM CAST WHERE gender = 'Female') AS female_count;
```

It might be also interesting to have an account, for each film, of how many scenes satisfy the requisites. One of the limitations of the Bechdel test is that, for it to be passed, it requires at least one scene. This kind of criteria does not allow us to perform a comparative analysis between different films. One way to perform this kind of analysis is to compare the percentage of Bechdel-friendly scenes on the total of scenes in the whole script.

```
SELECT m.film_id, m.titolo,
    ROUND(COUNT(
        CASE WHEN s.count_male = 0 AND s.count_female >= 2 THEN 1 END) * 100.0
        / COUNT(s.scene_id), 2) AS percentage
FROM MOVIE m
JOIN SCRIPT s ON m.film_id = s.film_id
GROUP BY m.film_id, m.titolo;
```

It might also be interesting to discover the number of times in which a male and a female appear in the movie:

```
SELECT film_id,
SUM(count_male),
SUM(count_female),
ROUND(SUM(count_male) * 100.0 / SUM(count_male + count_female)) AS male_percentage,
ROUND(SUM(count_female) * 100.0 / SUM(count_male + count_female)) AS female_percentage
FROM SCRIPT
GROUP BY film_id;
```

Another interesting analysis could be the investigation on the location of the scene where the film pass all the 3 level of the Bechdel test

```
SELECT s.place, COUNT(*) AS num_scenes_passed_test
FROM SCRIPT s
WHERE s.count_male = 0
  AND s.count_female >= 2
  AND s.count_male_nom = 0
GROUP BY s.place;
```

Other aspect is how many times a man is nominated o how many times a man Is the topic of a conversation

```
SELECT COUNT(*)
FROM SCRIPT
WHERE count_male_nom > 0;
```

DATA QUALITY

ACCURACY

To assess the quality of the information provided by the dataset we chose to confront the entries with external sources:

- Wikipedia, to assess if the cast of each film was correct
- Bechdel test website, to check if the evaluation provided by the model corresponded with the one provided by the forum.

It is important to note that the results taken from the Bechdel Test website are not entirely accurate, as they are based on evaluations made by film viewers. Many films (such as The Great Gatsby) have generated very divided opinions, with results not fully agreed upon by the community. Furthermore, we observe that the scenes we analyzed are much longer than those referenced by the site (which looks at micro-dialogues), making our test more stringent.

The main problem that the analysis has to overcome are recognize at the names dictionary: without the possibility to understand the context, a lot of names are not by the algorithm, making difficult the extraction and the count from the scene.

COMPLETENESS

The database does provide NULL values:

- In the ‘count’ columns, but in these cases the absence of values is informative and does not check as missing.
- In the ‘gender’ column, archived as *unknown*, the missing values were treated as None to avoid polluting the analysis.

RELEVANCE

The information in the database are divided into relevant columns, and, for the research goals, are not redundant.

To ensure that only relevant information was analyzed and extracted, in the importing phase all the roles that included:

- Articles: such as ‘the’, ‘a’
- Uncredited characters
- Character with ‘s
- Redundant actor for the same role

were treated differently, eliminating the confounding term. In this way, only the first name was analyzed.

READABILITY AND UNDERSTANDABILITY

The columns were named after the content, and the documentation describing the database structure and fields meaning is updated to the latest date.

CONSISTENCY

The gender is consistent across all the different cast analyzed, and it is consistent with both the gender of the

actor and the role interpreted. The films in which these two values were different were not taken into account to provide a more consistent dataset. To provide some examples, in Dallas Buyers Club Jared Leto, a male actor, has a female role, and in the movies Juno and Inception, a female role was recited by a now male actor. To protect the consistency of the dataset, these films were discharged. Other rules were enforced, such as the consistency between the scripts and the validity of the relationships between data (referential integrity). There were no conflicts between completeness and consistency.

OTHER METRICS

Due to the research needs, the database was created in a static timeline, so we do not consider timeliness and currency. It is possible to update the database when needed.

RESULT

Of the 20 films analyzed (found in Appendix 1).

- 8 pass only the first level
- 4 pass the first two levels
- 8 pass all three levels

Obviously the films that pass all three levels also pass the first two, so on a cumulative level we have

- 20 pass the first level
- 12 pass the second level
- 8 pass all the three levels

More specifically, the films that pass only the first level are:

Rocky (1976), Blade Runner (1982), Dead poets society (1989), Good will hunting (1997), The Truman show (1998), The gladiator (2000), Memento (2000) and The great Gatsby (2013)

Films that stop at the second level are:

Fargo (1996), American Psyco (2000), Big Fish (2003) and 12 years slave (2013).

Films that pass all levels are:

American Beauty (1999), Eternal Sunshine of the Spotless Mind (2004), Devil wears Prada (2006), No country for old man (2007), Anna Karenina (2012), Interstellar (2014), American Sniper (2014), La La Land (2016)

In percentage the results are divided as follows (figure2)

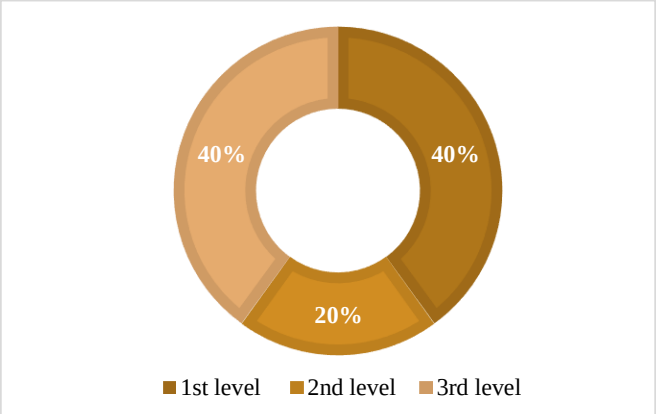

Figure 2

Looking at the distributions we see that there is a good balance especially between films that pass the first level and the third level, while those films that stop at the second level are in a clear minority.

In addition to whether or not they pass the Bechdel test, we analyzed other aspects, one of which is the female presence within the casts (figure 3)

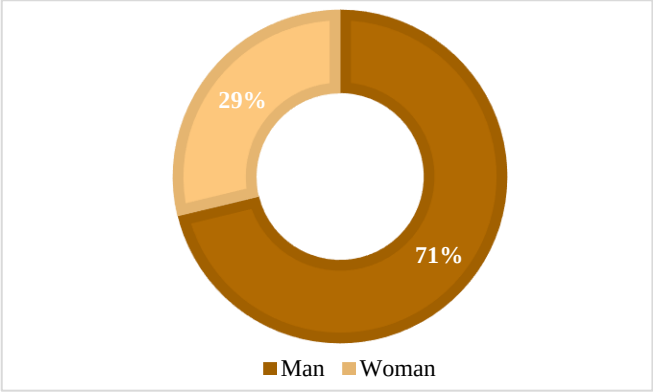

Figure 3

It is observed that female characters strongly outnumber men. More precisely, the average presence of men within the casts is around 32.2 men per film, while for women 13.05

This disparity is also confirmed on how many times women are in scene compared to men(figure 4)

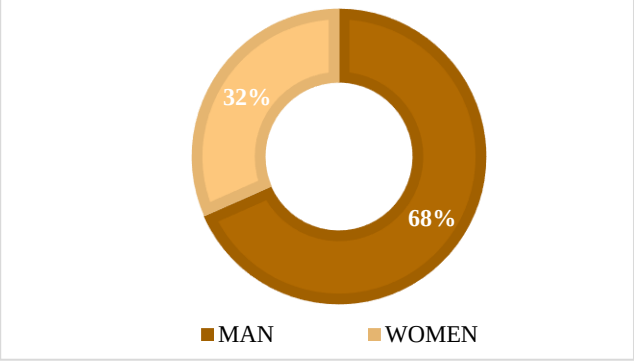

Figure 4

Moving on to additional analyses we have:

From the analysis of the locations where scenes passing Level 3 of the Bechdel Test occur, we have (tab1):

place	num_scenes_passed_test
EXT	13
INT	31

Tab 1

it's observed that 70% of these scenes take place indoors (INT). This highlights how the outdated stereotype of women as homemakers in these films has not been abandoned.

This doesn't change even for scene passing level 2 of the Bechdel test where over 78% of the scene are indoor (tab 2).

place	num_scenes_passed_test
EXT	19
INT	69

Tab 2

Moving to the percentage of Bechdel-friendly scenes, we observe that, on the total of scenes in the whole script are (tab 3)

titolo	percentage
Rocky	0.0
Blade Runner	0.0
Dead Poets Society	0.0
Fargo	4.35
Good Will Hunting	0.0
Truman Show, The	0.0
American Beauty	11.66
Gladiator	0.0
Memento	0.0
American Psycho	5.0
Big Fish	2.07
Eternal Sunshine of the Spotless Mind	4.27
Devil Wears Prada, The	14.22
No Country for Old Men	2.53
Anna Karenina	1.88
12 Years a Slave	0.65
Great Gatsby, The	0.0
Interstellar	0.4
American Sniper	1.6
La La Land	9.17

Tab 3

It's immediately evident how the percentages are quite low. This could be due to the film's topic, but at the same time, it's undeniable how the numbers are consistently too low in any case.

Analyzing how many times a man is nominated or how many times a man is the topic of a conversation we have that over the 3012 scene analyzed, 2319 have a man nominated in the scene. (figure 5)

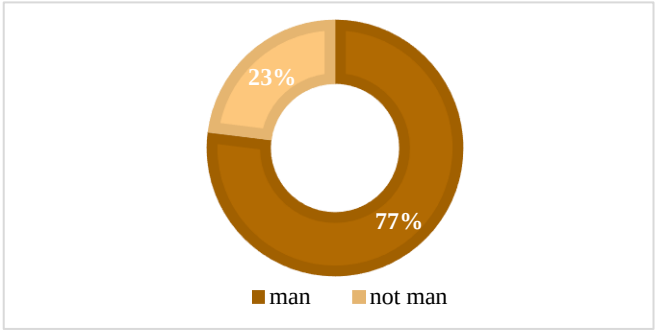

Figure 5

The value, as show in the graph, is equal to the 77% of the scene confirming the inequalities noted in our analysis.

CONCLUSIONS

This study underscores the importance of gender representation, and, applying the Bechdel test to the selection of films, how biased it can be. One of the most relevant observation one can make is the fact that, even if important for the plot and the number of lines is relevantly high, a lot of female characters were not referred to as their role's names, but as someone else's wife, mother or other relational terms. This phenomenon does not occur as frequently in reverse.

The process presented overall good results in terms of script analysis. Future improvements could expand on these methods, to include other parameters to refine the analysis.

It is quite interesting how, without the ability to read the context, gender inequality in movies' plots is quite frequent. Some of the film passed the test for a single scene, non-important for the plot, and are object of discussion to this day between the users of online forums. (e.g. *American Sniper* passes the test for a single line, in which a the wife of the protagonist speaks to a receptionist), so, even if the Test encompasses various scenarios, it is still not enough to provide a critical point of view.

For this reason, we think that this database, if expanded, might provide a useful tool to analyze other scripts and casts.

APPENDIX 1- MOVIES

The movies taken in account for our analysis are 20 Oscar's candidates movie:

- Rocky (John G. Avildsen, 1976)
- Blade Runner (Ridley Scott, 1982)
- Dead Poets Society (Peter Weir, 1989)
- Fargo (Joel ed Ethan Coen, 1996)
- Good Will Hunting (Gus Van Sant, 1997)
- The Truman Show (Peter Weir, 1998)
- American Beauty (Sam Mendes, 1999)
- Gladiator (Ridley Scott, 2000)
- Memento (Christopher Nolan, 2000)
- American Psycho (Mary Harron, 2000)
- Big Fish (Tim Burton, 2003)
- Eternal Sunshine of the Spotless Mind (Michel Gondry, 2004)
- The Devil wears Prada (David Frankel, 2006)
- No Country for old men (Ethan and Joel Coen, 2007)
- Anna Karenina (Joe Wright, 2012)
- 12 years a Slave (Steve McQueen, 2013)
- The Great Gatsby (Baz Luhrmann, 2013)
- Interstellar (Christopher Nolan, 2014)
- American Sniper (Clint Eastwood, 2014)
- La La Land (Demien Chazel, 2016)

Others movies participated in the creation process but were discharged. Other than the ones that are already nominated in previous sections, we also recall:

- The Green Mile
- A beautiful Mind
- The Lord of Rings
- Inglorious Bastards
- Inception

APPENDIX 2 – KEYWORDS

FEMALE- related keywords:

- Lady, miss, ms.
- Woman, female, girl
- Mom, mother
- Aunt
- Wife, Girlfriend
- Grandmother, Grandma
- Assistant, nurse, actress
- Princess, queen

Assistant and nurse might be considered gender neutral terms, but, due to the nature of the analysis and the scripts considered, it made sense to associate these terms with a feminine meaning. This necessity is also quite meaningful, as it offers insight on how the female role in film can be easily distorted.

Key words- MALE:

- Sir, mister
- Man, boy, male
- Dad, father
- Uncle
- Husband
- Grandad, grandfather
- Prince, king