

INFO247 Final Project

Presence vs. Representation

The Portrayal of LGBTQ+ Communities in Movies

[Aoi Furukawa](#), [Han Yang](#)

1 Project Goals

Despite an increasing number of movies included LGBTQ+ characters over years, there has always been criticism on how these characters are portrayed.¹ As one of the most influential modern media types, movies should represent the LGBTQ+ community properly. There have been positive trends regarding the LGBTQ+ representation in movies,² possibly due to changing social, cultural, and political conditions. By analyzing the quantity, ratings, box office, plots, and characters of the movies with LGBTQ+ representation, we aim to measure the true representation of the LGBTQ+ communities in movies of the past decade.

Specifically, this visualization project will accomplish the following goals:

- Present the facts, including the quantity, categorization, and trends, of LGBTQ+ presence in movies of the past decade.
- Analyze the true representation of LGBTQ+ population from diverse perspectives, using various metrics.
- Educate the general public about LGBTQ+ representation in movies through compelling storytelling and visualization.

2 Related Work

¹ Baldwin, Grant (2021) "The Effect of LGBT Film Exposure on Policy Preference," Sigma: Journal of Political and International Studies: Vol. 38 , Article 6.

Available at: <https://scholarsarchive.byu.edu/sigma/vol38/iss1/6>

² Andrew Limbong, "LGBTQ Characters Got More Movie Screen Time in 2020, GLAAD Study Finds," NPR (NPR, July 16, 2021), <https://www.npr.org/2021/07/16/1016801763/lgbtq-characters-got-more-movie-screen-time-in-2020-glaad-study-finds>.

2.1 [“Data Mining Reveals the Trends in LGBT-Related Filmmaking.”](#) Medium. Medium, December 12, 2018.

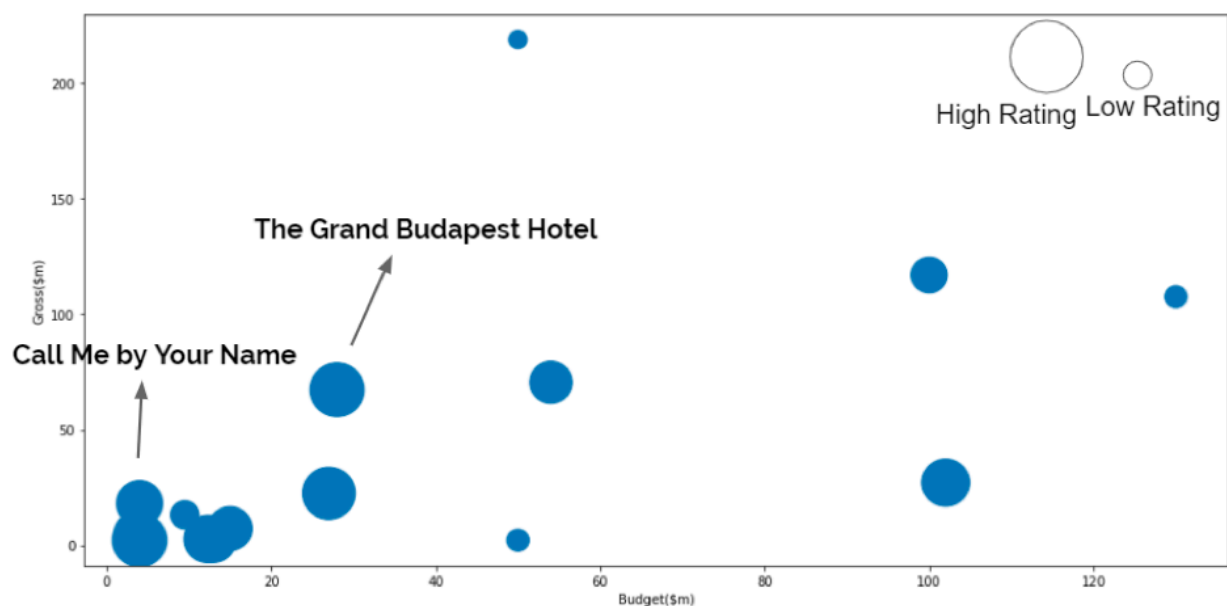


Fig. 1 Loss and Gain in Top-voters LGBT movies by Yice


Yice, the author of this Medium article, used IMDb movie data to analyze LGBTQ+ movie trends from 1998 to 2018. Yice inspired us to use IMDb LGBTQ+ plot keywords to scrape IMDb database to obtain movie entries with LGBTQ+ elements, but it also made us realize that this method was far from perfect, which we have noted in the project limitation below. For example, in Fig. 1, *The Grand Budapest Hotel* showed up in Yice’s dataset, which is only suspected to have LGBTQ+ characters. We learned a lot from Yice’s work, but also decided to focus the majority of our visualizations on LGBTQ+ focused movies to get more accurate conclusions.

2.2 [“Creating the next Bechdel Test.”](#) FiveThirtyEight, December 21, 2017.

Inspired by the original Bechdel Test, FiveThirtyEight attempted to measure the representation of women in the movie industry from a more holistic and diverse perspective. They created a variety of hypothetical metrics to inform readers that representation did not have a single definition. Our team appreciated FiveThirtyEight’s innovative approach and formulated a series of Rainbow Bechdel Tests, aiming to qualitatively analyze the representation of LGBTQ+ communities from various angles.

THE WAITHE TEST


Lena Waithe: Emmy Award-winning writer on “Master of None”



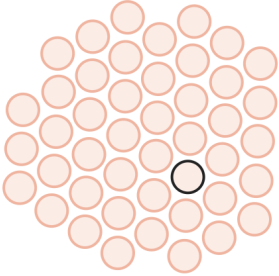
A movie passes if:

- There’s a black woman in the work
- Who’s in a position of power
- And she’s in a healthy relationship

○ **GHOSTBUSTERS**
THE PEANUT GALLERY SAYS:
“Why did it have to be the black character who wasn’t the scientist?”



5 passed



45 failed

> SHOW TEST NOTES

Fig. 2 One example of FiveThirtyEight’s Next Bechdel Test

2.3 “[Hollywood Equality: All Talk, Little Action.](#)” USC Annenberg School for Communication and Journalism, September 6, 2016.

Academic researchers in the media field have accomplished relevant studies as well. Professor Stacy L. Smith’s research is considered “the largest intersectional analysis of characters in motion picture content to date”. Her research has shown that, in Hollywood, the equality of genders, races, disabilities, sexual orientations, as well as identities, is still long way to go. Her research inspired some of the criteria in our Rainbow Bechdel Tests and our definition of “LGBTQ+ focused movies”.

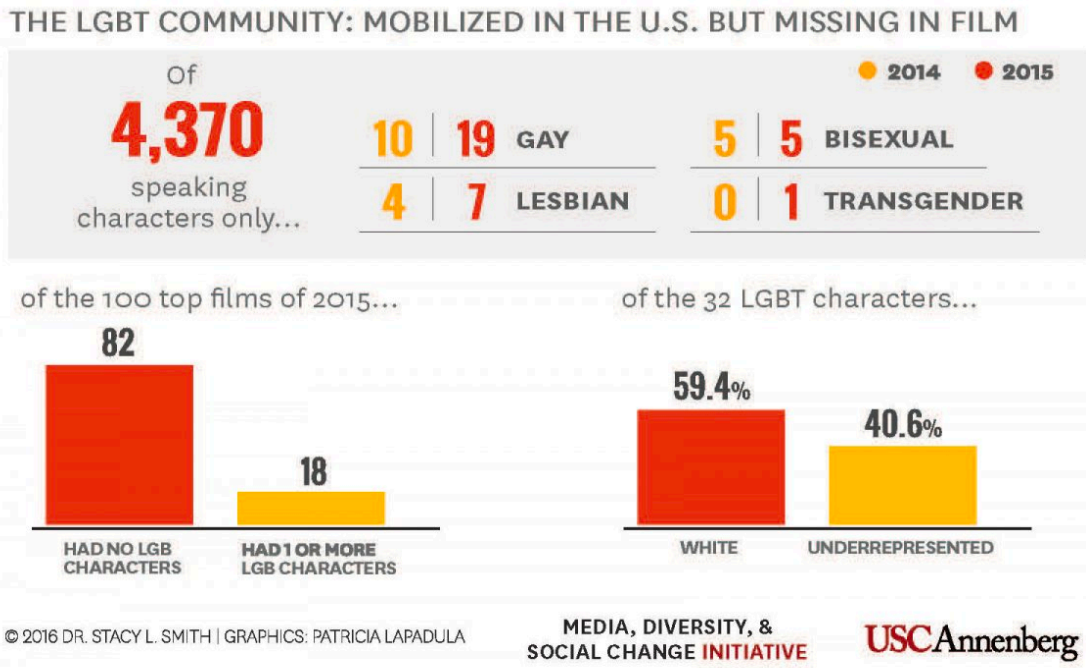


Fig. 3 Stacy Smith’s intersectional analysis of characters in motion picture

2.4 “[2020 GLAAD Studio Responsibility Index.](#)” GLAAD, July 15, 2021.

GLAAD, a non-governmental media monitoring organization, has been publishing the annual Studio Responsibility Index (SRI). Its annual publishing schedule allowed us to see the trends that GLAAD summarized. Similar to the Bechdel Test, GLAAD utilizes its “Vito Russo Test” to determine the true LGBTQ+ representation in movies with LGBTQ+ elements. The test inspired our Rainbow Bechdel Tests, but rather than using the Vito Russo Test as the basis shown on our webpage, we opted for the Bechdel Test due to its wider publicity.

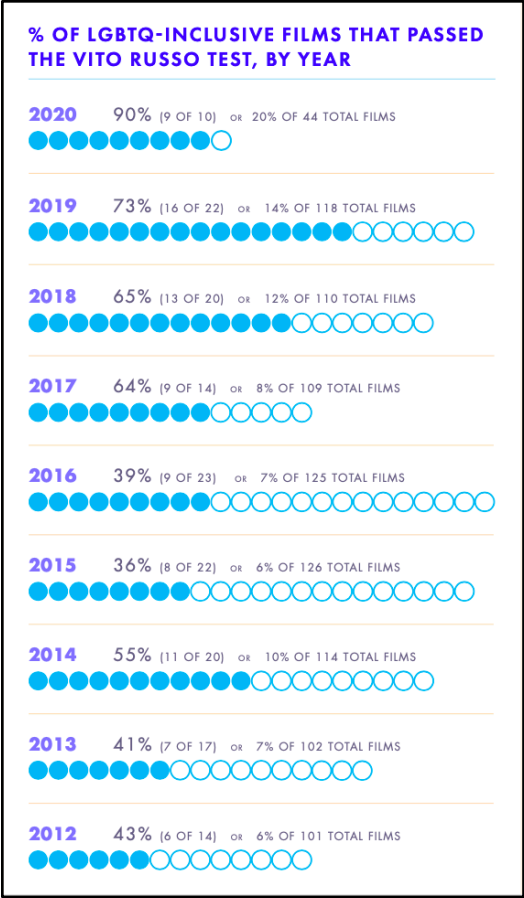


Fig. 4 The Vito Russo Test analysis by GLAAD

2.5 [“The Last 60 Years of Oscar Nominations in One Chart.”](#) FiveThirtyEight. FiveThirtyEight, March 3, 2018.

FiveThirtyEight’s Oscar nomination visualization (Fig. 5) is an example of how to visualize complex movie information in a relatively simple format. Even though the visualization itself is not interactive, it still communicates the information effectively. Inspired by the article, our teams opted for squarified icons to represent movies in some of the visualization prototypes. We also ranked the importance of interactive components in each visualization and decided to implement the most critical interactivity first, given the limited time. Visualizations with low interactivity needs will be implemented in a static format similar to this example.

Which Oscar category is the least open to newcomers?

Looking at the 9,942 Academy Award nominations awarded to 5,385 nominees, 1958-2017

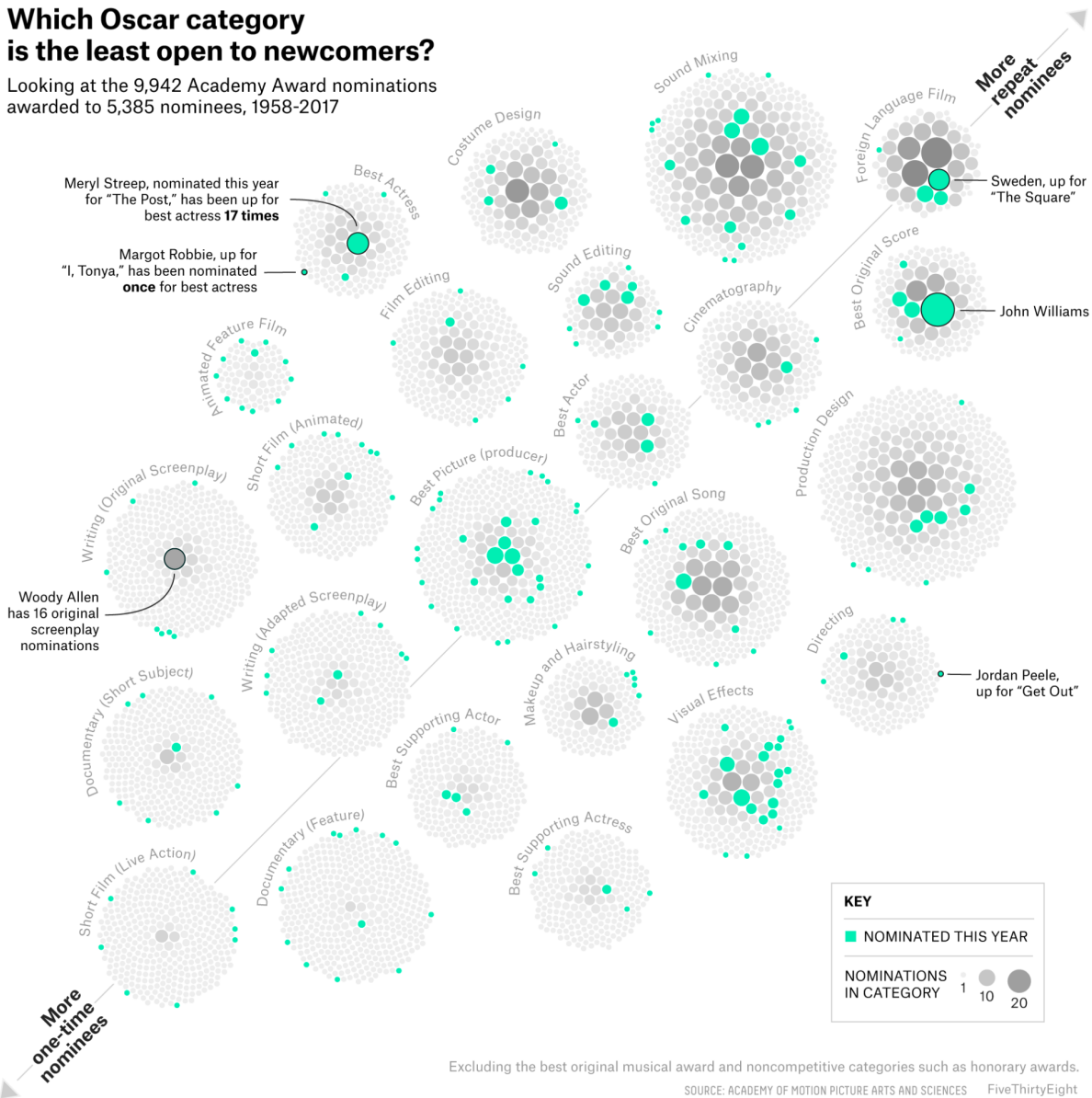


Fig. 5 FiveThirtyEight's Oscar nomination visualization

2.6 ["Ghostbusters' Is a Perfect Example of How Internet Movie Ratings Are Broken."](#) FiveThirtyEight. FiveThirtyEight, July 14, 2016.

FiveThirtyEight's review analysis for *Ghostbusters* inspired us to dive deeper into the review discrepancy across different sites. We obtained the box office data, IMDb ratings, and Metacritic ratings (metascores) for movies with LGBTQ+ elements and aimed to explore the correlations between them. However, the results of this analysis were proven to be insignificant and readers were unable to get meaningful information out of the visualization during the

usability study. Even though we decided to remove the visualization from the final design, this example and our exploration helped us understand what was important vs. unimportant for our final design.

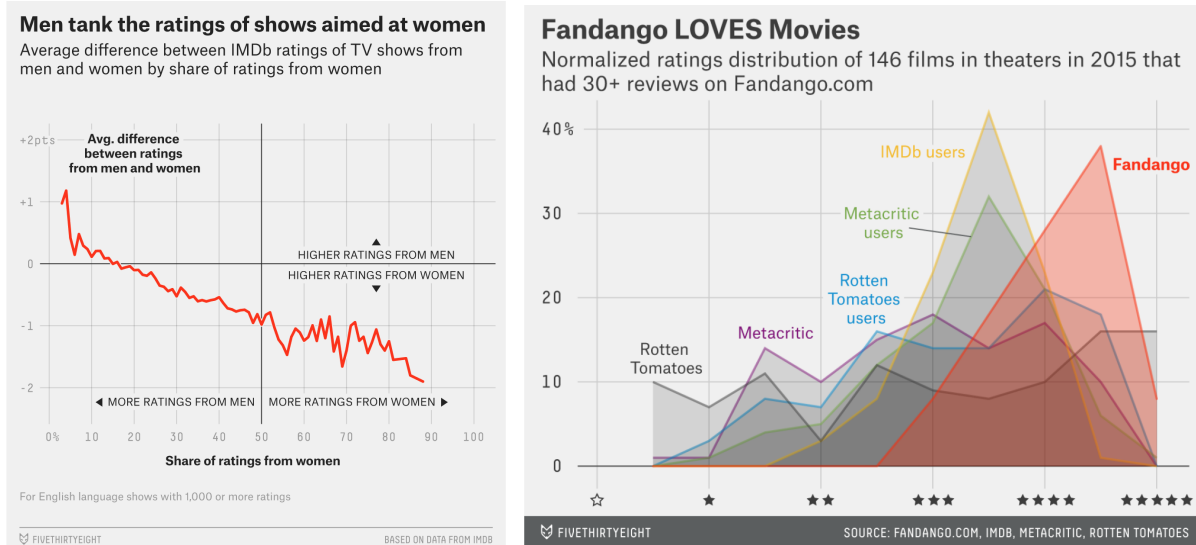


Fig. 6 *Ghostbusters'* review visualization by FiveThirtyEight

3 Visualization Description

To achieve the goal of educating the general public, the key of our visualization flow is storytelling (see detailed approach in section 4.3). Texts and individual visualizations take advantage of the “slow in” and “slow out” animation effects (where applicable) to improve the overall interactivity. The visualization webpage is optimized for desktop and tablet (landscape) viewing.

3.1 Introduction and the Persona Story

We begin the story with a “hook” – a persona and the severity of the problem of representation (Fig. 7).

To grab the attention of the readers, we challenge the perception of “presence” vs. “representation” right away. By focusing on the most influential movies (“top 50”) with LGBTQ+ elements first, we attempt to introduce the concept and the impact of top movies early on. The squarified icons are used throughout the webpage and are introduced right at the beginning as well. We use a trans persona to guide readers interacting with the visualization. Readers are asked to find the only trans-focused movie among the defined top 50 movies by hovering over the squarified icons. No matter whether readers are successful in finding the movie or not, they can learn the focuses of the top 50 movies gradually during the discovery process. The difficulty in finding a single movie also reinforces the central problem that the visualization is trying to address – presence is not equal to representation.

Presence vs. Representation

The Portrayal of LGBTQ+ Communities in Movies



First, let's imagine we are part of the transgender community...

Among the 50 most popular movies with LGBTQ+ elements released in the last decade*, can you find the movies that include our community as part of the main plot?

Hover over the matrix to reveal answers.

Captain Marvel	I Care a Lot	Bohemian Rhapsody	A Star Is Born	Solo: A Star Wars Story	Bird Box	Birds of Prey	Blade Runner 2049	It Chapter Two	Alien: Covenant
Get Out	Guardians of the Galaxy	Deadpool	Green Book	The Old Guard	Fantastic Beasts: The Crimes of Grindelwald	Deadpool 2	Beauty and the Beast	The Shape of Water	Guardians of the Galaxy Vol. 2
The Favourite	The Grand Budapest Hotel	Ghost in the Shell	Call Me by Your Name	The New Mutants	Moonlight	Star Trek Beyond	The Imitation Game	Atomic Blonde	The Lobster
A Simple Favor	Spotlight	Ted	The Perks of Being a Wallflower	Vice	Onward	The Intouchables	Booksmart	Suspria	The Mule
The World's End	The Neon Demon	Sausage Party	A Million Ways to Die in the West	Happy Death Day	Power Rangers	The Invitation	Love, Simon	Dallas Buyers Club	Bombshell

*Popularity based on the number of IMDb reviews
 *LGBTQ+ elements are defined as any LGBTQ+ characters or scenes

Fig. 7 Introduction, persona story and the flipped card visualization



Fig. 8 Introduction, persona story and the flipped card visualization (continued)

3.2 The Timeline

After the introduction, the timeline (Fig. 9) attempts to educate readers about a series of key events related to LGBTQ+ representation in movies. The timeline also transforms the tone of the visualization into the more positive side – even though the current status is not perfect, we are making progress.

The timeline was highly rated as well as effective during the usability testing. Readers appreciated the simplicity and the mental break between interactive and information-intensive visualizations. We further improved the visualization by emphasizing on information hierarchy and concise text descriptions.

Looking at LGBTQ+ representation in the history of movie production, we are making progress, albeit imperfect...



1934

"The Moral Guidelines"

The Hays Code, effective from 1934 to 1968, placed a series of censorship guidelines on movie production. Homosexuality was considered not acceptable.

1970



A Milestone in Queer Cinema

"The Boys in the Band", centers on a gay friend group, is among the first major American motion pictures to revolve around gay characters.



1993

Starring Tom Hanks

A-list actor Tom Hanks won an Oscar for his portrayal of a gay lawyer in "Philadelphia". The performance and the fame of Hanks helped broaden the understanding of AIDS.

Fig. 9 Timeline – key events related to LGBTQ+ representation in movies

According to IMDb, in the last decade, about 150–200 movies with LGBTQ+ elements were released every year...

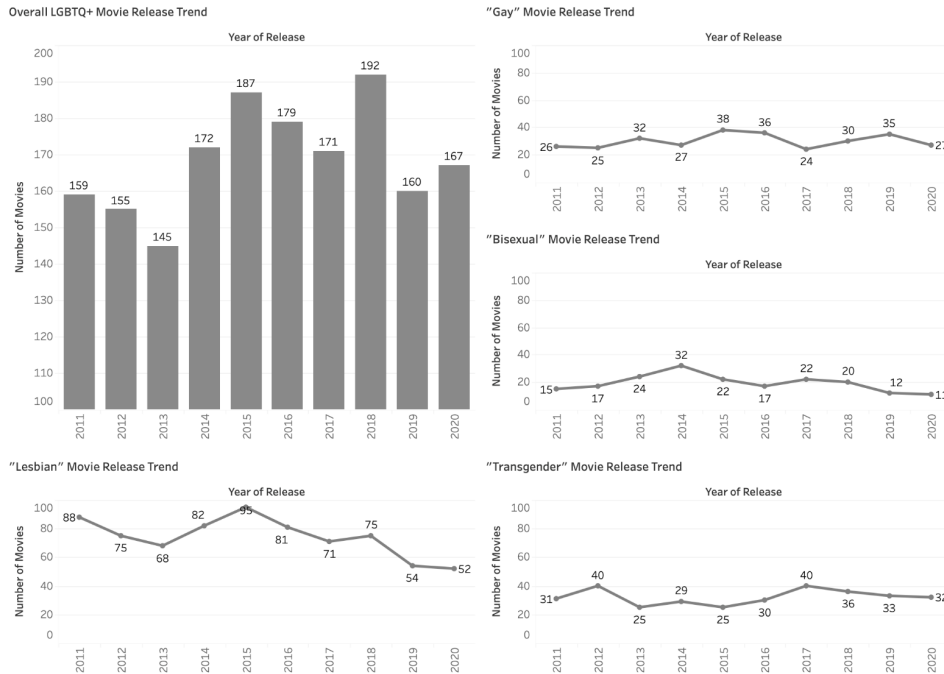


Fig. 10 Overall LGBTQ+ movie release dashboard

3.3 Overall Movie Release Dashboard

The timeline in the previous step introduces readers to the progress that we've made regarding representation over the years. After a brief introduction to the historical context, readers are guided to understand the progress in detail and the current trends (2011–2020) (Fig. 10).

Taking advantage of the functionality of a Tableau dashboard, we aim to show the overall trend as well as the trends by category. Readers can interact with the charts on the dashboard and dive into each year through brushing and linking. The overall release dashboard serves as the foundation for detailed-level visualizations later.

However, the majority of them only have LGBTQ+ characters that are peripheral...

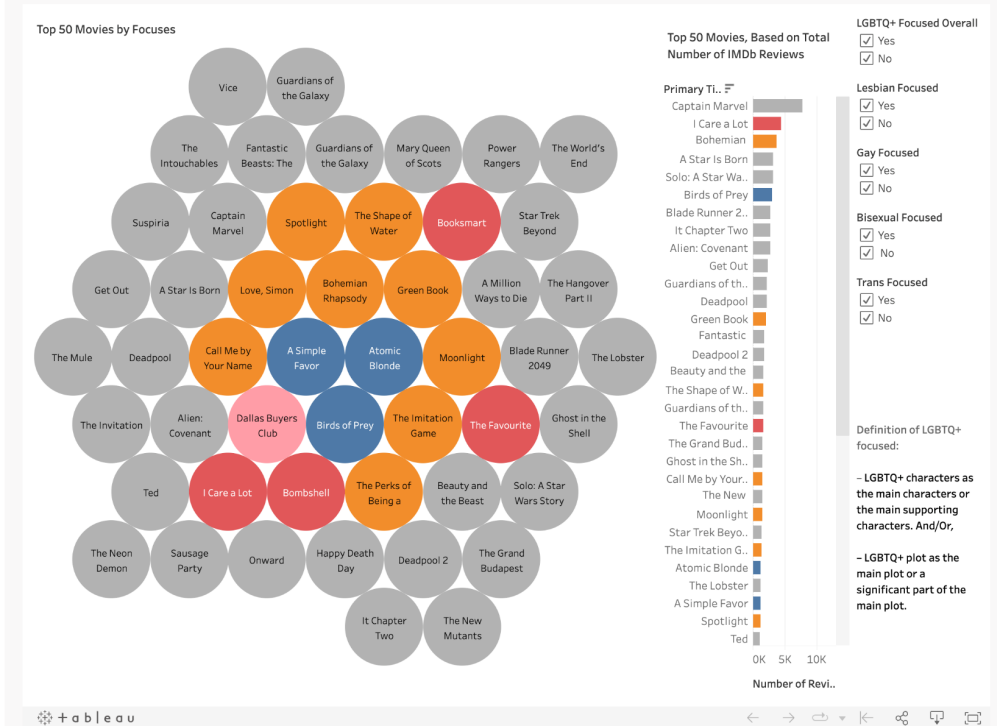


Fig. 11 Categorization of top movies by types

3.4 Top 50 Movie Categorization Dashboard

After the overall movie release dashboard, readers are brought to focus on the top movies again and the story turns back to the less positive side. Top movies are color coded to represent categories (“Lesbian”, “Gay”, “Bisexual”, “Trans”, or non-LGBTQ+ focused), and readers can hover over the icon to reveal tooltips with additional information (Fig. 11).

To visually reinforce our definition of “LGBTQ+ focused”, a bar chart ranking movies based on the number of IMDb reviews is shown. Readers can explore the relationship between charts through brushing and linking. Through the categorization and the ranking, readers are being communicated with the inequality within LGBTQ+ movies.

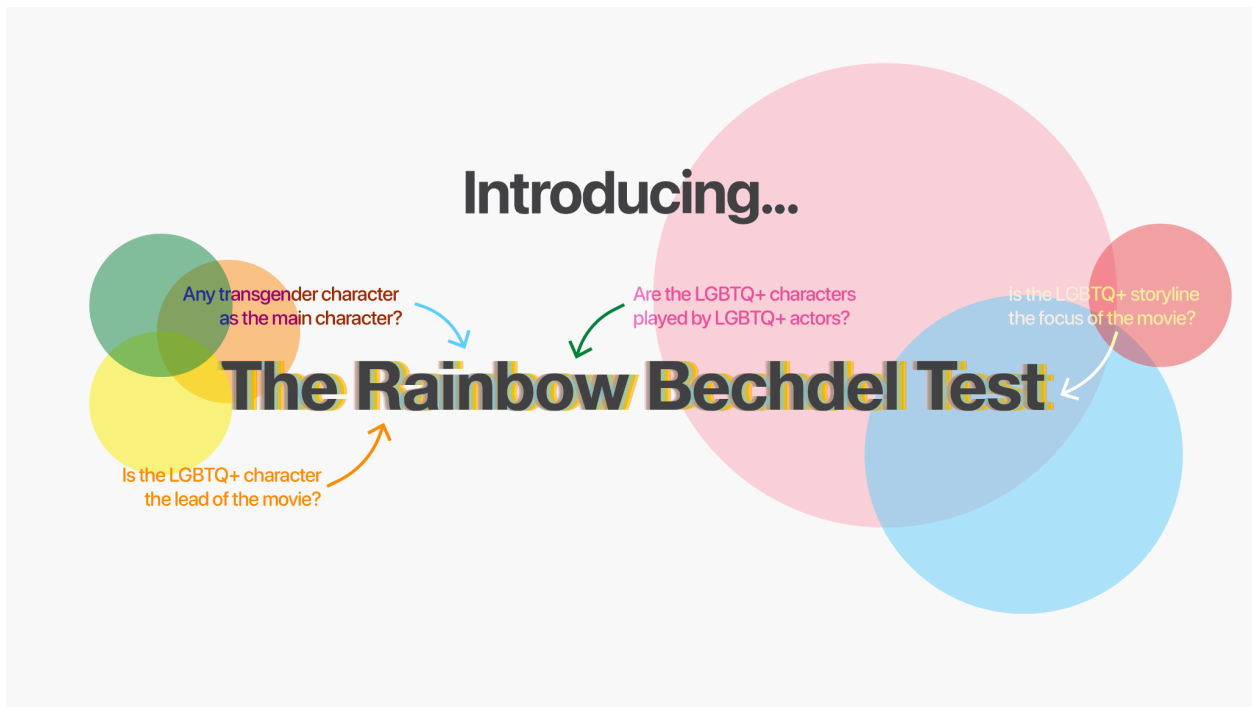


Fig. 12 Introduction to the Rainbow Bechdel Tests

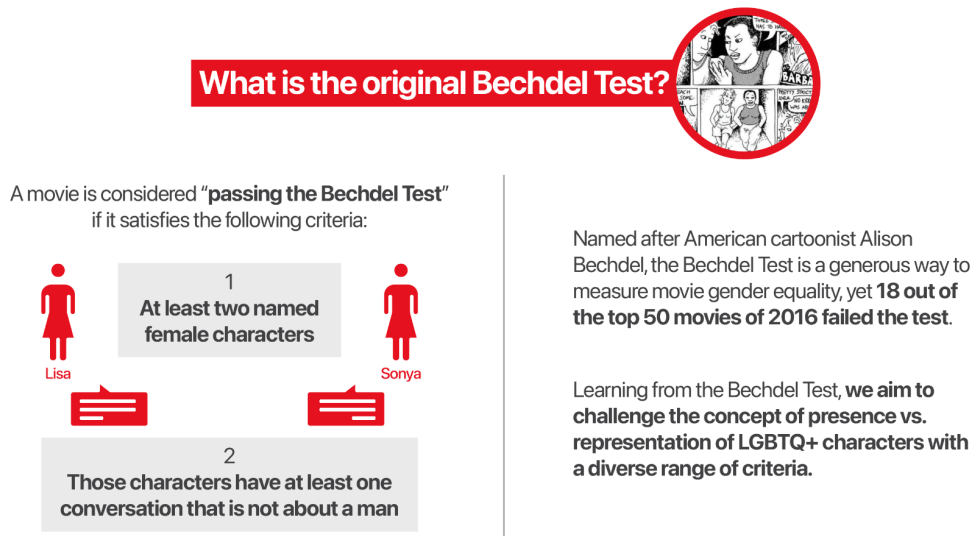


Fig. 13 The original Bechdel Tests

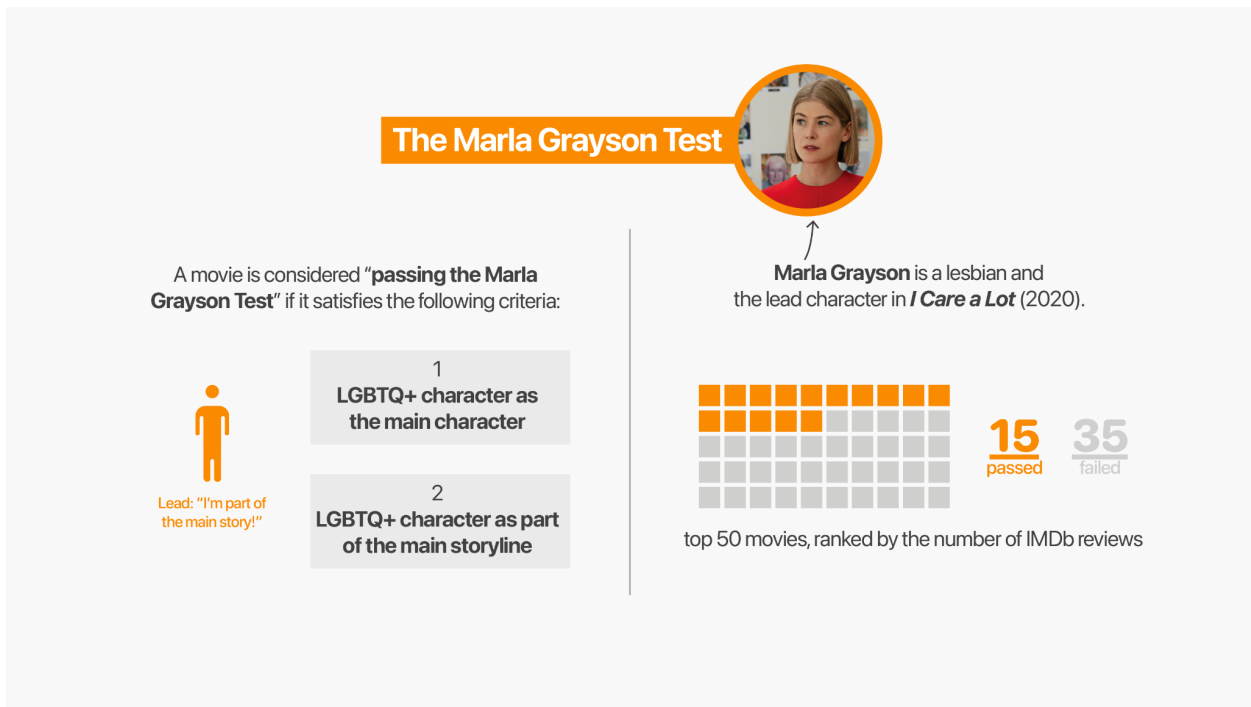


Fig. 14 Rainbow Bechdel Test 1: the Marla Grayson Test

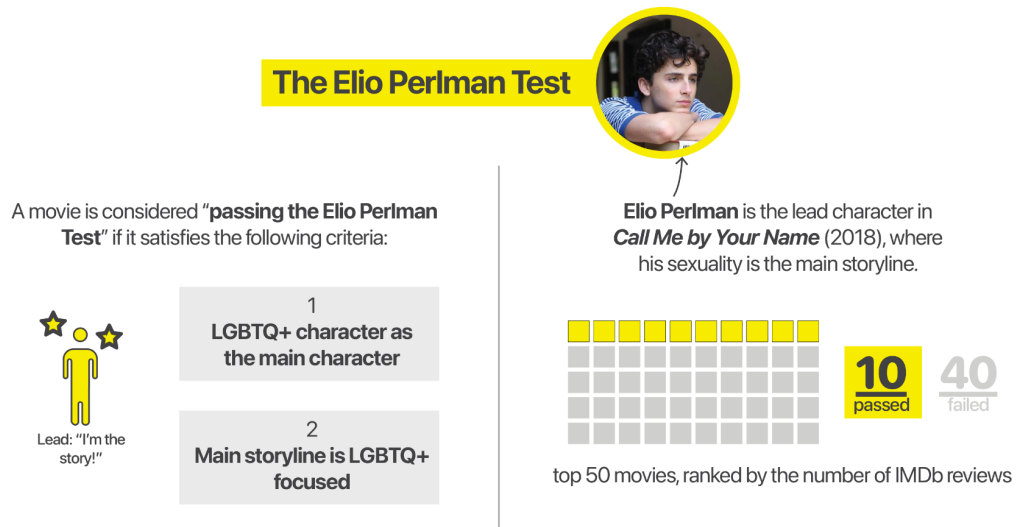


Fig. 15 Rainbow Bechdel Test 2: the Elio Perlman Test

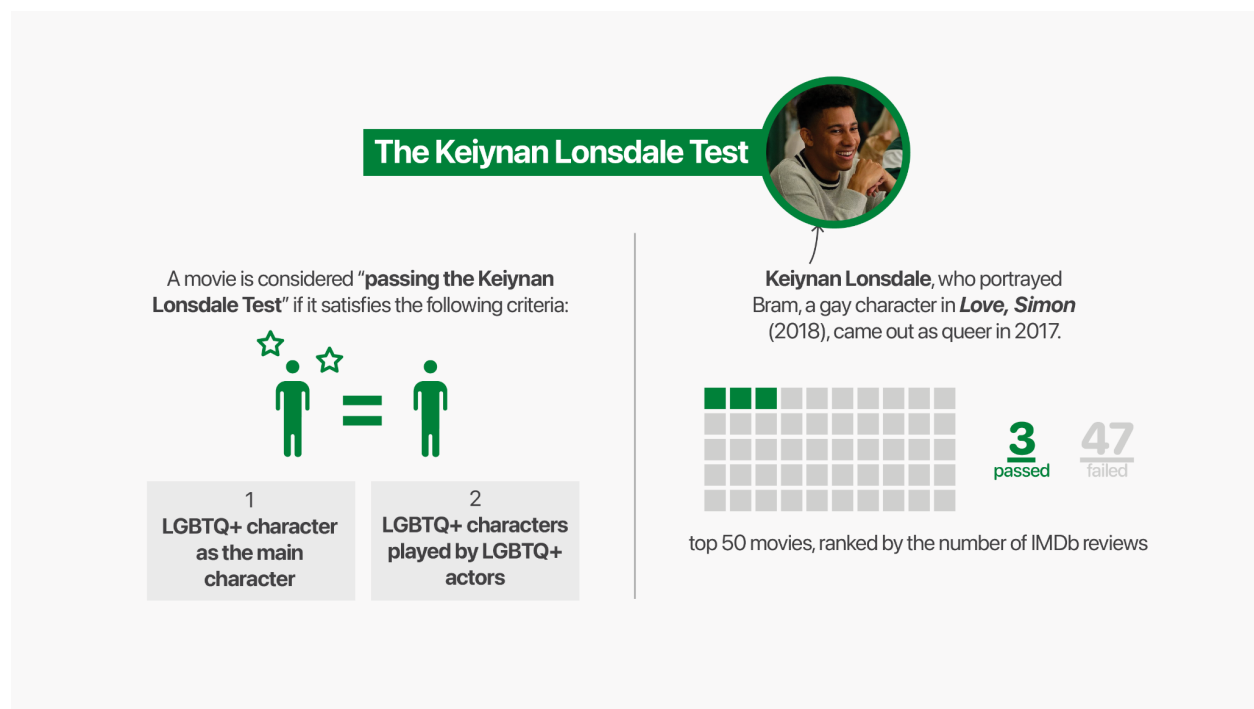


Fig. 16 Rainbow Bechdel Test 3: the Keiynan Lonsdale Test

3.5 The Rainbow Bechdel Tests

The visualizations so far have been focusing on the issue of LGBTQ+ movie representation through the lens of quantities and categories. It is possible that readers are unable to resonate with the idea in its entirety, since LGBTQ+ population is relatively small and readers might think the population's presence is adequate. Starting from the Rainbow Bechdel Tests (Fig. 12–16), readers are brought to the content level and dissect what true representations mean.

The Rainbow Bechdel Tests are a series of infographics that introduces readers to the concept of representation measurement. Similar to the timeline, these infographics emphasize on communicating some knowledge points in simple and effective ways, rather than the interactivity. Here, readers are also able to take another short break after exploring the information-intensive dashboards. With a set of gradually more demanding representation criteria, readers are being challenged with the concept of presence vs. representation.

However, even for LGBTQ+ movies that have passed the Rainbow Bechdel Tests, their characters and plots can be stereotypical...

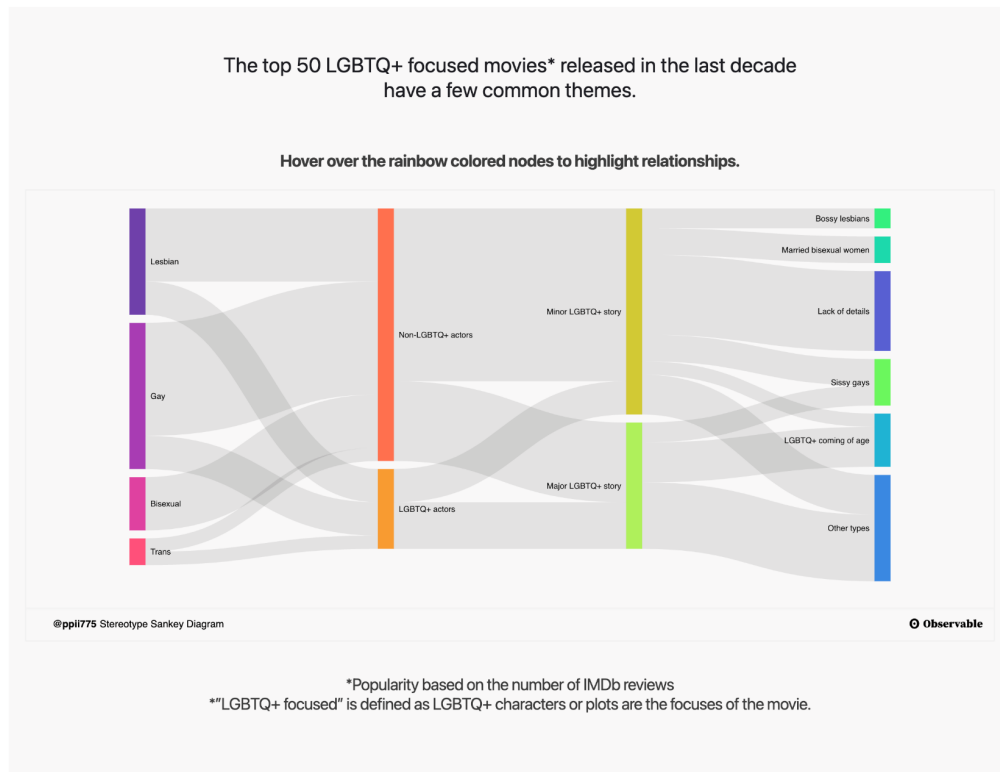
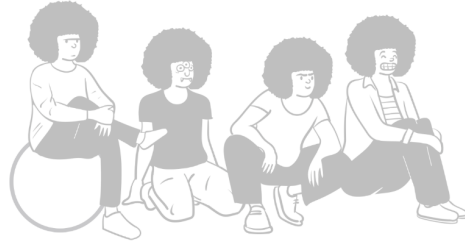


Fig. 17 Stereotype Sankey diagram

3.6 Stereotype Sankey Diagram

Even if the movie passed multiple Rainbow Bechdel Tests, the representation could still be too stereotypical. A Sankey diagram (Fig. 17) is presented to readers after the Bechdel Tests to dive even deeper into the movie plots and characters. Revised based on readers' feedback from the usability testing, only four sets of nodes are used to reduce the complexity and categorize the most important traits related to representation and stereotype. Readers can hover over the node to trigger the link highlights between adjacent nodes. Even though the

relative sizes of the nodes can communicate the problem, the added layer of interactivity was well received during the interviews.

3.7 Colors, Fonts, and Illustrations

In general, the webpage uses colors sparingly. Rainbow themed colors are used to represent critical categorical information and are incorporated into important visualizations to improve legibility. Alternating background colors are used for different sections of the visualization.

We used SF Pro sans serif font as the primary font for its excellent readability. In addition, we set up Inter sans serif font as the backup font to maximize consistency, in case readers' computers don't support Apple fonts. Different font weights and sizes are used to facilitate information hierarchy. Summary and title texts of each section are provided in bold and big format so that readers can easily grasp the central message. Readers can also dive into the detailed texts for additional content if they choose to do so.

Thanks to the free illustrations by Pablo Stanley, a series of monochromatic persona characters are placed throughout the page, which correspond to the message of each section. These illustrations were highly rated for their relevance and their aesthetic quality during the usability testing.

3.8 Video Walkthrough

Link: <https://youtu.be/b4IXPqPFUvU>

4 Approach

4.1 Data Preparation Approach

Our team primarily utilized the IMDb database ("the world's most popular and authoritative source for movie, TV and celebrity content") to obtain data. Learning from Yice's example mentioned in the Related Work section, we used plot keywords as criteria to scrape movie entries that included these specific keywords.

The final dataset contains movies with LGBTQ+ plot keywords ("gay", "lesbian", "transgender", "bisexual", "LGBT", or "LGBTQ") from 2011 to 2020. These movies have at least one LGBTQ+ character or scene, and the dataset is not only about LGBTQ+ focused movies (e.g., LGBTQ+ as the main plot). "Queer" has been omitted due to high percentage of overlap and its difficulty in being categorized. The dataset was generated by scraping the IMDb database in February 2022. Even though the accuracy is high, it is important to note that the original movie data can be crowdsourced by IMDb users. The dataset has 1687 unique entries.

Important column information is as follows:

Type	Column Name(s)	Description
Identifier	tconst	Unique identifier used for all productions on IMDb.
Title	primaryTitle originalTitle	For movie identification only. Original title is included for non-English movies.
Release year	startYear	The movie release year can be different from the movie production year. This dataset prioritizes release year, since movie production can be a long time.
Genres	genres	Movie genres. Each movie can have up to three genres.
Rating	Rate	IMDb ratings.
Metascore (metacritic.com)	metascore	Ratings from metacritic.com, another influential rating website, are included to cross-check IMDB ratings.
Plot keywords	keywords	These keywords are used to scrape the IMDb database to include movies with at least one LGBTQ+ character or scene.
Box office sales	Gross worldwide	Budget and box office data are incomplete in the IMDb database. "Gross worldwide" is determined to be the most complete.
Total number of reviews	review_total	The number of total user and critic reviews on IMDb. This number is used to determine movie popularity for this analysis and it is not related to Metascore.

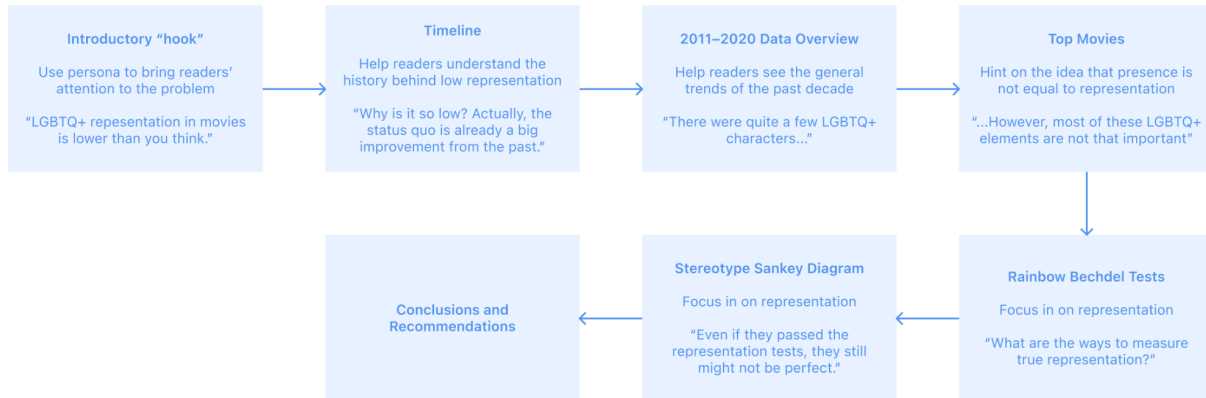
While the dataset was instrumental in identifying overall trends and forming the basis for the Tableau dashboards, we transformed the data for more granular analyses. By sorting the movie entries based on the total number of reviews (“review_total”) and qualitatively evaluating each movie plot, we identified the top 50 movies with LGBTQ+ elements and the top 50 LGBTQ+ focused movies. The subsequent grounded coding categorized these movies into multiple buckets. The top movie datasets were the basis for the intro flipped card visualization, the Rainbow Bechdel Tests, and the stereotype Sankey diagram.

4.2 Tools

- Python
We used Python library BeautifulSoup and requests for scraping detailed movie data from IMDb database. After scraping the data, we used pandas for data cleaning.
- Tableau
Tableau was used not only to create the final dashboards embedded onto the webpage, but also to conduct overall EDA, which helped us define the focus and storyline of our project.
- Figma
Our team utilized Figma to iterate through the overall webpage designs. The Figma prototype helped us design transitions between visualization. It was used as the primary prototype for our usability testing.
- Illustrator
For our Rainbow Bechdel Tests, Illustrator was used to design high-fidelity narrative infographics. These static visualizations have low dependence on interactivity and were directed embedded onto the final webpage.
- D3.js
Our team utilized D3.js to create visualizations that have high dependence on interactivity. A Sankey diagram was adopted for the final webpage to illustrate the categorical relationships.
- HTML, CSS, and Javascript
For developing the website, we used HTML, CSS, Javascript. We used the templates from [CodyHouse](#) for the vertical timeline, and from [scrolltrigger](#) for implementing the scroll triggered events.
- Github
We used Github pages for hosting the website we created.

4.3 Storytelling Approach

To achieve the goal of educating the general public about the topic, our team took storytelling as a key approach in organizing the final visualization flow. Our storytelling approach can be summarized as follows.



4.4 Steps

1. Data preparation

As mentioned above, we used plot keywords as criteria to scrape the IMDb database.

The final dataset contains movies with LGBTQ+ plot keywords (“gay”, “lesbian”, “transgender”, “bisexual”, “LGBT”, or “LGBTQ”) from 2011 to 2020.

Since the original dataset provided by IMDb did not contain the keyword information, we scraped the keyword data from each movie’s page. Also since each page contains a lot of information that is not necessarily related to this research, we cleaned the data and extracted relevant data.

Tools: Python (pandas, BeautifulSoup, requests), Jupyter notebook

2. Overall EDA

Utilizing the overall dataset, we performed EDA to validate hypotheses and discover insights. These insights helped us identify the story and focus, which guided our exploration of the individual visualizations in the next step. We identified the top 50 movies with LGBTQ+ elements and the top 50 LGBTQ+ focused movies based on the number of IMDb reviews.

Tools: Tableau

3. Individual visualizations

Through our own efforts as well as through completing other INFO247 assignments, we iterated on various visualization options, which were assembled during the prototyping phase to form a coherent story.

Tools: Tableau, Illustrator, Figma, D3.js

4. Prototypes and usability study

The primary goals of the prototyping phase were to test various implementation strategies, validate the storyline, and create functional prototypes for the usability study. Our usability study (see the next section for details) involved three participants and we obtained meaningful feedback and priority items that have been incorporated into the final submission.

Tools: Figma, HTML, CSS, Javascript

5. Final submission

The final project submission included a comprehensive writeup and a webpage. The webpage has been improved based on the results of the usability study. We incorporated all the priority revision items identified in the study.

Tools: HTML, CSS, Javascript, Github

5 Usability Testing & Results

5.1 Introduction

Our usability study was conducted two to three weeks before the final submission of the project. The primary goal of the usability study was to identify issues before the final implementation ([V1 prototype](#) used for the usability study, see [INFO247 Usability Testing Report](#) for more details).

The usability study focuses on both the aesthetics and the understanding of the current design. Specifically, we would like to find out:

1. For the overall storytelling and webpage design, whether or not readers enjoy them visually and/or have improvement in understanding of the topic;
2. For individual visualizations, whether or not they are aesthetically pleasing and/or successfully contribute to the overall storytelling;
3. The most important areas to be improved before the final submission, given the limited time.

5.2 Method

Due to continued disruptions caused by the pandemic, the usability study is conducted remotely via Zoom. As permitted by the participants, all Zoom sessions are recorded. Surveys are distributed, completed, and recorded digitally during the Zoom sessions. Observations and semi-structured interviews are conducted with video and screen sharings. Participants can control the prototypes freely during the screen sharings. For each participant, our usability test utilizes a predetermined sequence of methods. The entire process takes approximately 45min per participant. Two team members swap roles between sessions (interviewer vs. notetaker/facilitator). We used the Figma prototype as a primary testing interface due to

efficiency. Since some features are unable to be implemented on Figma, users are asked to switch between Figma and Tableau/Observable/webpage.

Test sequence for each participant

1. **Pretest question survey.** A multiple choice pretest short survey is distributed to each participant before the test to quantitatively assess the participant’s existing knowledge of the topic. The pretest survey includes introductory questions that measure the participants’ exposure in the topic area, as well as questions explored in the prototypes. The survey focuses on areas that participants can answer conveniently through multiple choice. More detailed usability questions are addressed in the semi-structured interviews. The result of the pretest survey serves as the baseline for each participant.
2. **Observation.** After completing the pretest survey, each participant is asked to browse the overall Figma webpage prototype freely. Following the thinking aloud method, each participant is asked to verbally communicate their thoughts while reading the page.
3. **Posttest question survey.** After the observation session, a posttest survey is distributed to each participant, which includes the same set of questions used in the pretest survey, minus the introductory questions. The result of the posttest survey is used to compare with the pretest survey baseline result to measure participants’ improvement in the knowledge of the topic.
4. **Summary survey.** An exit summary survey is distributed to participants after the posttest survey. Utilizing Likert scales and open-ended questions, the survey obtains qualitative data toward the aesthetics and understanding of the webpage and visualizations. Participants are informed that the summary survey results would be reviewed live.
5. **Semi-structured interviews.** Finally, participants and the design team walk through the webpage and visualization prototypes together. Based on the live summary survey results and the observations, impromptu questions are formulated to address participants’ immediate concerns. Afterward, with a set of predetermined interview questions, participants are asked to express their opinions on individual visualizations.

5.3 Results of the Quantitative Measures

The results of the correct response rate for survey questions are shown as the table below. While the rate increased for Question 3 asking the reason why there weren’t LGBTQ+ characters before 1970’s, it didn’t change for Question 1 and even decreased for Question 2.

	Pre-test	Post-test
1 – During the last decade (2011–2020), roughly how many movies with LGBTQ+ elements were released each year?	0%	0%

2 – Among these movies, which LGBTQ+ sub-community has the highest percentage?	33%	0%
3 – What was the primary reason that LGBTQ+ characters were almost nonexistent before the 1970s?	0%	100%
4 – What is the Bechdel Test?	100%	100%

5.4 Results of the Qualitative Measures

The summary of responses for qualitative summary survey questions are shown as the table below. Users appreciate the aesthetics of the visualization slightly more than its contents. Also, the score distribution is similar in both questions: The Tableau dashboards scored the lowest from both aesthetics and content perspectives, while the timeline and the Rainbow Bechdel Test series scored the highest.

Overall design

Question	Average score (1 poor – 5 great)
1 – From an aesthetic perspective, what’s your rating for the overall design?(1-5)	4.3
From a content perspective, how would you rate your understanding of the overall design?(1-5)	3.6

Aesthetics

Question:	Average score (1 least - 6 most effective)
From an aesthetic perspective, please rank the following visualizations from the most effective to the least effective.	
The flipped card visualization at the intro	4
The timeline	5.3
The overall LGBTQ+ movie release dashboard	4.3
The LGBTQ+ box office and review dashboard	2
The Rainbow Bechdel Test series	4.3
The Sankey diagram for LGBTQ+ movie stereotypes	4
Aesthetics average overall	3.98

Contents

Question: From a content perspective, please rank the following visualizations from the most clear to the least clear (6 being the most clear).	Average score (1 least – 6 most clear)
The flipped card visualization at the intro	3
The timeline	6
The overall LGBTQ+ movie release dashboard	3
The LGBTQ+ box office and review dashboard	2.3
The Rainbow Bechdel Test series	4.7
The Sankey diagram for LGBTQ+ movie stereotypes	3.7
Content average overall	3.78

5.5 Observation and Semi-structured Interview Results

Overall

- The testing took more time than we expected. Users spent fair amounts of time understanding the text description and the dashboards.
- Users' testing flow was complicated by prototype switching.
- Users understood the intention of the visualization, but not all the details.
- Users enjoyed the design overall, but their interests in the topic varied.

The introductory “hook” and persona

- Some terms used in the texts were confusing (e.g., “50 most popular movies”).
- Gray squarified icons didn't represent movies well. Most movie posters are rectangle-shaped.
- Users gave suggestions on the squarified icon flipping interaction:
 - What would the interaction look like if users successfully picked the correct movie?
 - Should the squarified icon stay flipped after the interaction?
- Users didn't get the meaning of the colored squarified icons until seeing multiple screens. The gradual exposure didn't work well.

The timeline

- Users liked the information and the simplicity of the design.
- The text description took a while for users to read through.

The overall LGBTQ+ movie release dashboard

- Users felt lost or simply didn't spend much time engaging with the dashboard – too many visualizations in one screen.
- Users needed concise information of what this visualization was about.
- Users thought the dashboard was too technical. Tooltips didn't contain much more useful information.

The LGBTQ+ box office and review dashboard

- The comments for the previous dashboard also applied here.
- The visualizations were too hard to understand. Specially, users couldn't grasp the meaning of the scatterplot and the log scale.

The Rainbow Bechdel Test series

- Users enjoyed the design and the information overall.
- Users were confused by the names of the tests, as they weren't sure of the relationship between the new tests and the original Bechdel tests. More information was needed to bridge the gap.
- Users thought these tests already existed – they didn't get the intent of the visualizations.

The Sankey diagram for LGBTQ+ movie stereotypes

- Users thought the design and the interaction were cool.
- It was hard to understand what this visualization meant.
- Users preferred to see relationships across nodes. Different users got different information from this visualization.

5.6 Analysis

The overall test results showed that the visualization was harder to understand than we expected. While we aimed to provide a holistic picture and detailed information to help them understand the topic, users enjoyed simpler visualizations (e.g., the timeline) than a group of charts (e.g., the dashboards). When users felt overwhelmed by the information, they would simply ignore certain information, which led to lower-than-expected understanding of the content. Besides visualizations, users would like to see clear explanations and summary texts to assist with their information processing. Too much text or inappropriate summaries can lead to further confusion.

Based on user feedback, we have prioritized the most important changes as follows, which haven't been addressed for the final submission.

- Revise all text descriptions and add appropriate summary texts to all visualizations. Texts need to reflect the intents and the key takeaways of the visualizations. Long texts would be redesigned to highlight information hierarchy.

- Reduce the number of the visualizations on the dashboards and revise the dashboards to reflect the key intents. Tooltips should include additional and appropriate information to help users understand the dashboards.
- Reduce the number of scrolls and visualizations used for the intro. The long intro was distracting and the gradual exposure was too long to be effective.
- Improve the timeline. Users highly rated the effectiveness of the timeline. Additional efforts will be spent on further improving the information hierarchy.
- Improve the understanding of the Rainbow Bechdel Tests series by adding descriptions and transition information.
- Reduce the complexity of the Sankey diagram. Add additional functionality (e.g., capability of rearranging nodes) to better communicate the intents.

6 Links to Demos

Link to webpage: https://ei-blue.github.io/lgbt_film/

Link to GitHub repository: https://github.com/ei-blue/lgbt_film

7 Collaboration

Overall, team members have taken on equal amounts of tasks. Table below includes a detailed contribution breakdown for each task.

Task	Component	Aoi	Han
Data preparation	Raw data scraping and organization	90%	10%
	Data processing and grounded coding	10%	90%
Overall design	Storytelling and design strategies	50%	50%
	Figma design prototype	0%	100%
	Webpage implementation prototype	100%	0%
Overall content	Text	30%	70%
	Webpage information architecture	70%	30%

Visualizations	Intro (flipped cards)	70%	30%
	Timeline	70%	30%
	Tableau dashboards	30%	70%
	Rainbow Bechdel Tests	30%	70%
	Stereotype Sankey diagram	30%	70%
	Flows between visualizations	70%	30%
Usability study	Strategies and interviews	50%	50%
	Analyses and reports	50%	50%
Final writeup	Final writeup	30%	70%

8 Appendices

8.1 Code Reference

- ScrollTrigger
<https://greensock.com/scrolltrigger/>
- CodyHouse vertical timeline
<https://codyhouse.co/gem/vertical-timeline/>

8.2 Illustrations

- Open Peeps by Pablo Stanly
“All Illustrations can be downloaded and used completely free.”

8.3 Content Reference

- Monteil, Abby. “A History of LGBTQ+ Representation in Film.” *Stacker*, 14 Oct. 2021,
<https://stacker.com/stories/4331/history-lgbtq-representation-film>.

8.4 Usability Study Tasks

Usability Test Consent Form

Participants sign consent forms.

Usability Test Background Introduction (3min)

- Thank you for agreeing to participate in this usability test. We are assessing an information visualization design.
- As mentioned in the consent form, we will be collecting data and recording the session. All information will be kept confidential. Nonetheless, feel free to communicate your concerns along the way.
- We will be showing you a collection of information visualizations. You will be asked some questions before and after reading the visualizations.

Pretest Questions (5min)

1. Introductory question 1: If you can recall, among all the movies that you have watched, roughly how many of them/what percentage of them have **LGBTQ+ elements** (These can be any LGBTQ+ characters or plot elements)?
2. Introductory question 2: What do you think about the representation of LGBTQ+ communities in movies?
3. During the last decade (2011–2020), roughly how many movies with **LGBTQ+ elements** were released each year?
4. Among these movies, which LGBTQ+ sub-community has the highest percentage?
5. What was the primary reason that LGBTQ+ characters were almost nonexistent before the 1970s?
6. What is the Bechdel Test?

Observation Tasks (10min)

- Please feel free to browse this webpage prototype at your own pace;
- Please tell us what you are thinking while you are browsing it – it can be about the look and feel, your understanding, etc.

Posttest Questions (5min)

Posttest survey includes pretest questions 3–6.

Summary Survey (7min)

1. From an aesthetic perspective, what's your rating for the overall design? (Scale 1–5)
2. From a content perspective, how would you rate your understanding of the overall design? (Scale 1–5)
3. From an aesthetic perspective, please rank the following visualizations from the most effective to the least effective. (Six individual visualizations used in the prototype)

4. From a content perspective, please rank the following visualizations from the most clear to the least clear. (Six individual visualizations used in the prototype)
5. Did any part of the prototype/visualization stand out for you (This could be in both good and bad ways)? Why?
6. Do you have any other comments?

Semi-structured Interview Questions (15min)

1. After taking the surveys, is there anything you would like to bring up first?
2. (Unstructured questions related to observing participants.)
3. (Participants are asked about the overall design and each of the following visualizations.) What do you think about it? Anything confusing, missing, or unnecessary?
 - a. The introductory “hook” and persona
 - b. The timeline
 - c. The overall LGBTQ+ movie release dashboard
 - d. The LGBTQ+ box office and review dashboard
 - e. The Rainbow Bechdel Test series
 - f. The Sankey diagram for LGBTQ+ movie stereotypes
 - g. The overall design

8.4 Usability Study Surveys

Pretest and Posttest Questions

[\(Link to Google Forms\)](#)

1. Introductory question 1: If you can recall, among all the movies that you have watched, roughly how many of them/what percentage of them have **LGBTQ+ elements** (These can be any LGBTQ+ characters or plot elements)?
(Open-ended question; no correct answer.)
2. Introductory question 2: What do you think about the representation of LGBTQ+ communities in movies?
(Multiple choice question; no correct answer.)
 - a. Overrepresented
 - b. Represented adequately
 - c. Represented inadequately
 - d. I don't know
3. During the last decade (2011–2020), roughly how many movies with **LGBTQ+ elements** were released each year?
 - a. Below 50
 - b. 50–100
 - c. 100–150
 - d. 150–200

- e. Above 200
- 4. Among these movies, which LGBTQ+ sub-community has the highest percentage?
 - a. Lesbian
 - b. Gay
 - c. Bisexual
 - d. Trans
 - e. Roughly the same
- 5. What was the primary reason that LGBTQ+ characters were almost nonexistent before the 1970s?
 - a. Because the society was too conservative.
 - b. Because of the movie industry self-imposed censure/guidelines.
 - c. Because no actors expressed their sexual orientation openly.
 - d. Because there was no such market.
- 6. What is the Bechdel Test?
 - a. It is a test used to identify whether or not the movie production had the freedom of expression.
 - b. It is a test used to measure whether or not the director did a good job.
 - c. It is a test used to measure the movie's popularity.
 - d. It is a test used to measure the representation of women in movies.

Summary Survey

[\(Link to Google Forms\)](#)