**It’s a man’s canvas**

# Motivation:

Movies are present in our lives just about anywhere and everywhere these days. In the cinema after a nice dinner, on the go with our phones or as a procrastination tool for the next EPFL homework. There's nothing quite like escaping into a good movie, no matter where or when we choose to do it!



As something that reaches so many people in the world, it is bound to influence our society. But not only can it influence us, there exist a reciprocal relationship where it simultaneously reflects the current state of the world.

With the subject of gender gap being hotter than ever, the question of how female representation in the movie industry is arises.

We want to analyse gender equality not because it is a buzzword or a trendy topic. We believe that on-screen representation has an impact on how we see ourselves and others. Underrepresentation or poor representation can contribute to negative stereotypes. An authentic representation is therefore vital to avoid reinforcement of gender inequalities in society.

Not only on-screen but also off-screen representation is important. When men dominate the film industry, it could lead to some bias for the movie content. It could also perpetuate the idea that men are more suitable or valuable for certain positions, which has serious consequences for women's careers and opportunities.

# Goal:

In this project we aim to analyse the on and off-screen gender representation in the movie industry to understand and address the gender gap. This is done by comparing data from different genres, geographical areas and searching for trends.

The results were found based on:

Movie metadata (2012, Freebase)

81’741 Movies

147 Countries

363 Genres.

**CMU Movie Summary Corpus**

Character metadata (2012, Freebase)

405’060 Movies

270’257 Male characters

134’803 Female characters

Plot summary (2012, English-language Wikipedia)

42’303 Plots

Character tropes (TV tropes)

72 Character types

501 Instances of those types

**IMDB website**

Director (2022, IMDb)

47’253 Movies

186’751 Male directors

34’153 Female directors

Writers (2022, IMDb)

49’730 Movies

279’186 Male writers

73’026 Female writers

With this we are ready to analyse the data!

# Which gender dominates the movie canvas?

The simple and short answer is: Men!

But let’s first talk about why this question is important. The ratio of male to female characters can be used as a metric to reflect gender equality in the world of fiction. It goes beyond just counting bodies on a screen. We believe that it speaks to the values and beliefs of the current society. So, let’s dig deeper into that question

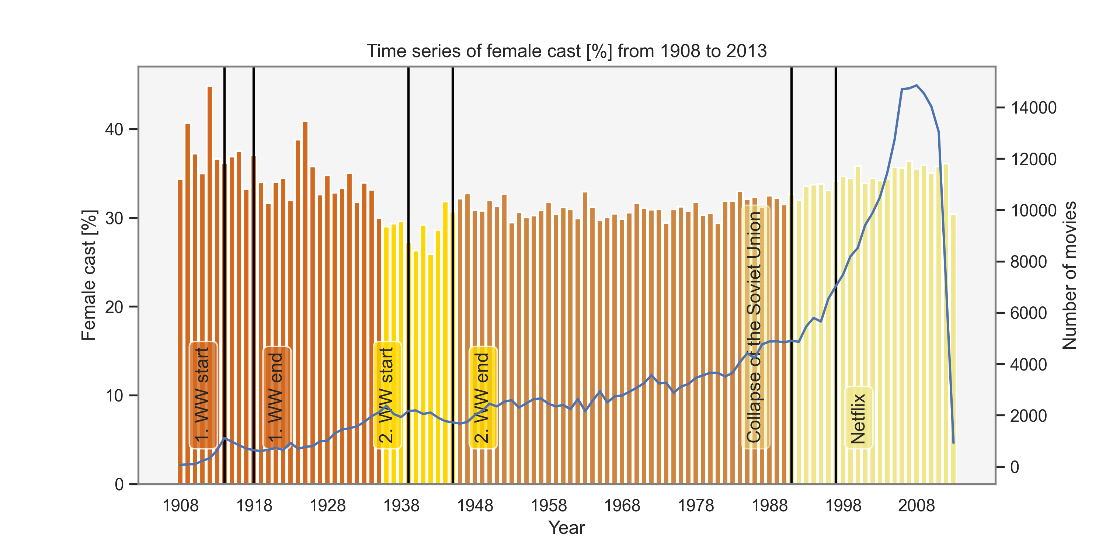
## Time series

In the time series below, we can see that men have always dominated on screen. We broke down the time series into four periods that show interesting behaviour. In the beginning of the dataset the number of movies is rather low, which causes the variance to be high. Therefore, the dataset should spawn over a longer time, namely from 1908 to 1935 (mean: 0.35, variance: 0.01).

Right around the second world war we see a decrease in female cast members, from 1935 to 1945. The variance as well as the mean is lower compared to the previous period (mean: 0.29, variance: 0.003).

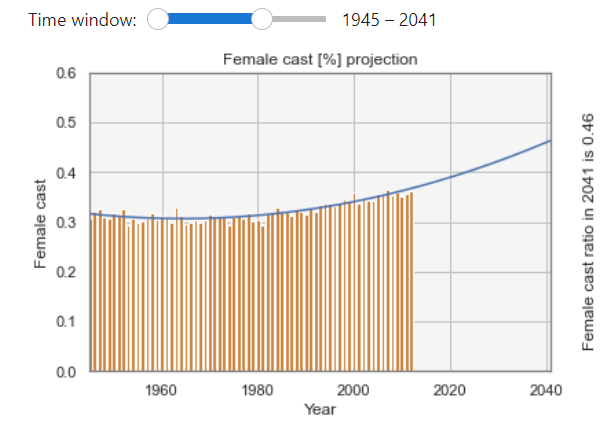
The third period was chosen over a long time between 1945 to 1990 because the variance is rather low (mean: 0.31 variance: 0.0001).

The last period should be representable for the current situation, so we choose the years between 1990 to 2013- the last year available in our dataset with the necessary information (mean: 0.34, variance: 0.0002). During this period the number of movies sky rocks, which might be due to the introduction of streaming services. In 2013 the value of female cast member percentage and number of movies within the dataset is unexpected small. Therefore, this value was taken out from the dataset for further analysis.



Now that we’ve seen how the time series looks like it makes us think how the future will evolve.

We believe that the dataset of the last period is most suitable to predict how the world of movies will progress since it is most similar politically and technology wise to today. By fitting a least squares model with polynomial degree of 2 we found out that in year 2050 the female cast ratio will be equivalent to the male cast ratio. So, if the growth rate is equal to 1900 to 2013, we will reach gender equality on screen in almost 15 years!

 or

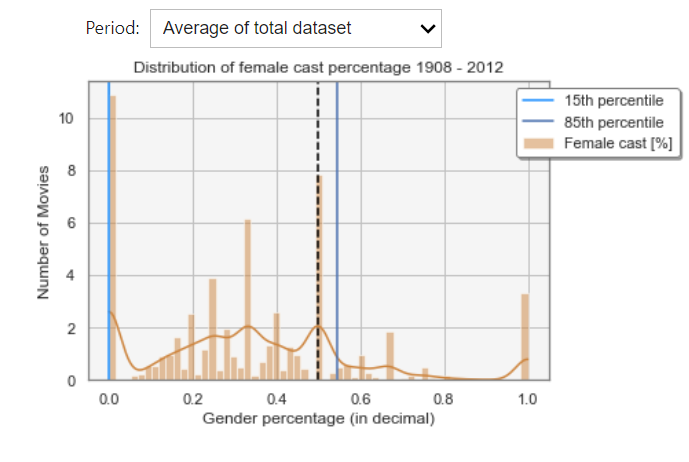
To see how the model prediction is performing we will compare it to the study ‘It’s a Man’s (Celluloid) World' (https://womenintvfilm.sdsu.edu/wp-content/uploads/2022/03/2021-Its-a-Mans-Celluloid-World-Report.pdf).

Feel free to try it for yourself, on the left side of the figure you see the predicted value of the last year.

We chose to compare the year 2020 instead of 2021 since 2021 data is impacted by the Covid pandemic. In the study female characters accounted for 38 % of major characters compared to our 39 % projection.

In order to see why our prediction is higher than what reality show we take a closer look of the dataset. The study has monitored portrayals of girls and women in the top 100 domestic grossing films since 2002 and found that 85% of films featured more male than female characters.

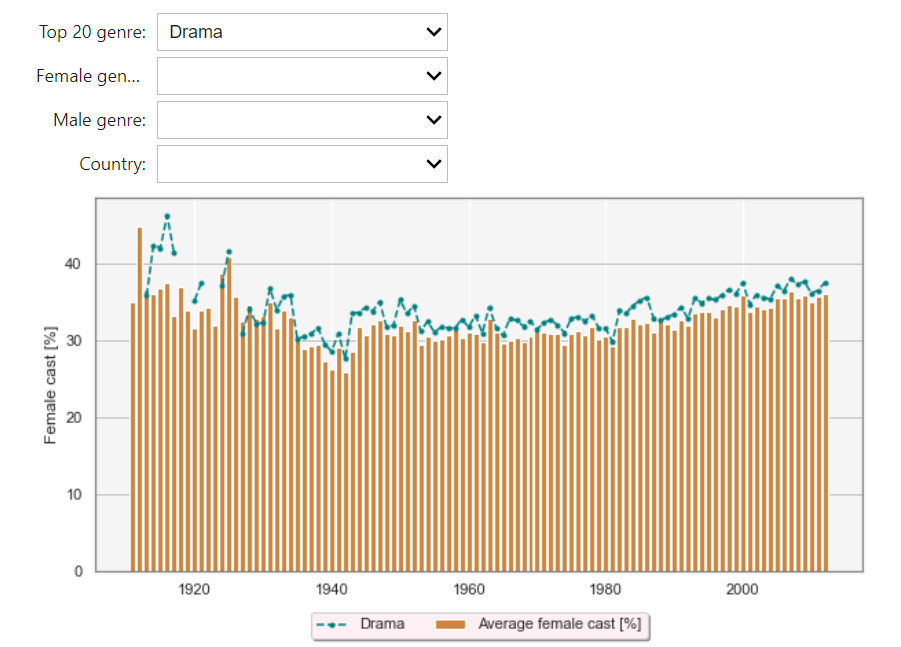
What about our dataset?



In our dataset this is only true for the two periods 1935-1945 and 1945-1990. For the other two datasets the percentage of films featuring more male than female is lower. We could therefore hypnotise that because the dataset used for training the model had a lower percentage of male dominant movies the prediction is lower than the real situation.

We have now seen how the cast gender distribution looks like over time. In the plot below you can explore the respective female cast percentage for: the recent 20 most popular genres, the genres that have the highest female respective male percentage at some point in the time and the countries that produce most movies currently.

Try to see if you notice anything interesting. We will talk take a deeper dive in the next few chapters.



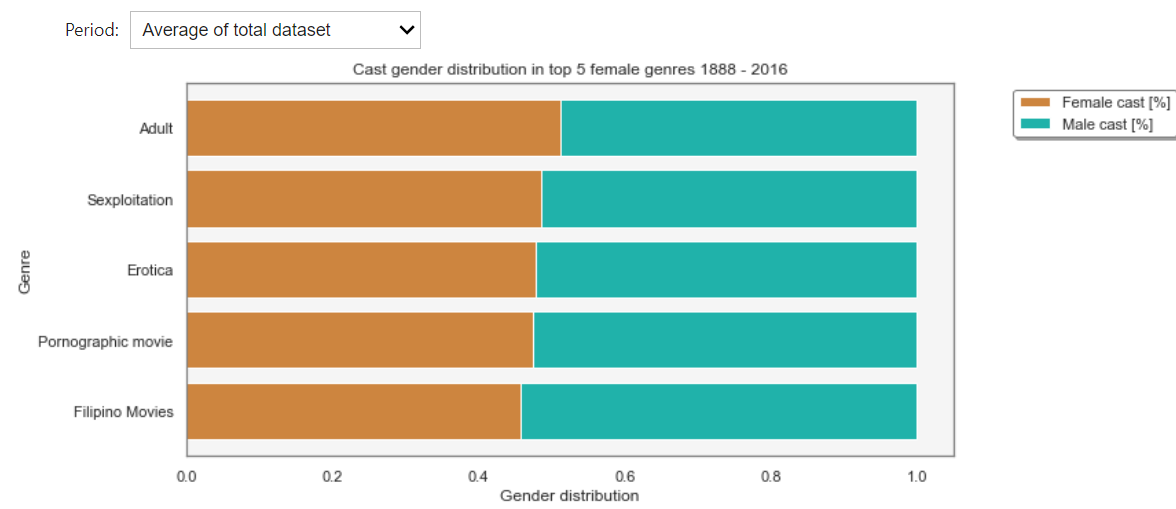
## Gender gap regarding movie genre

We all have encountered a similar situation: Boys get superheroes comics and girls receive princess books.

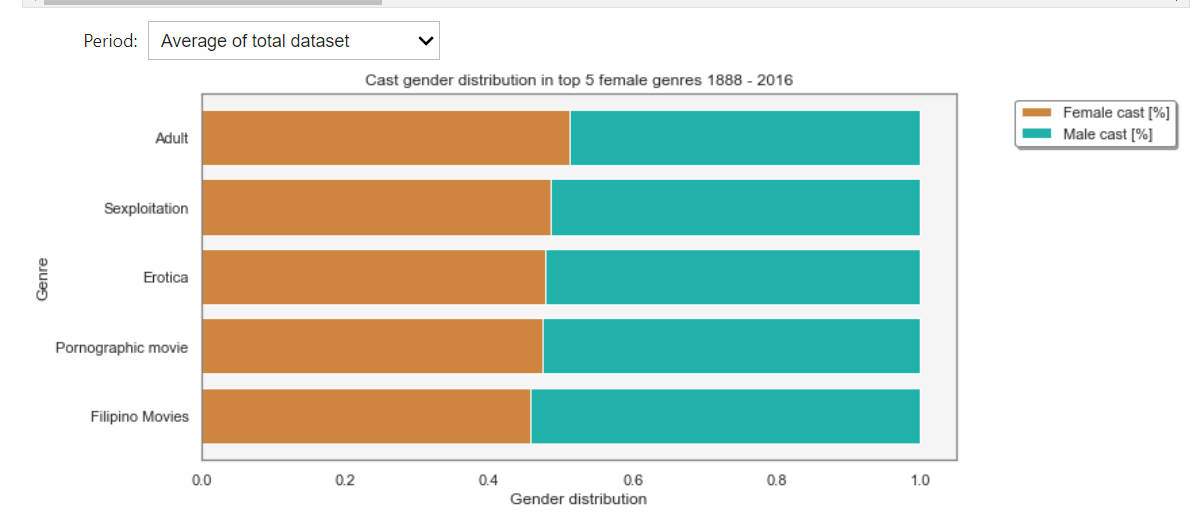
This made us think if there is a similar separation between gender in the movie domain. Do the topics vary between movies with higher female lead? By examining the gender distribution in different genres, we can gain a better understanding of the challenges and obstacles that women in these fields may face.

While analysing gender regarding the gender gap there is the trade-off between searching for the most extreme gender distribution while still choosing representable genres not analysing niche genres. We chose to set a threshold 100 movies. Any genre that has less movies is discarded in this analysis.

Not sure if we should include the most popular genres or if this is already covered with the plot above.



Let’s begin to look at the genre with the highest female cast percentage.



We can see that the genre with the highest female cast ratio changes in each period. Until 1945 women appear most often in comedy, romance and musical. Those genres often cover topics such as relationships, slice of life and family. At first glance it appears that women are more likely to play in light-hearted, feel good movies. Upon reflection they often are emotionally powerful, make us shed some tears while also being uplifting and hopeful. Based on these findings we can say that the top 5 female genres up to 1945 are often tied with emotions.

Some examples are: (We can provide a link or insert video playing the trailer”)

“The Wizard of Oz" A musical about a young girl's journey through a magical land to find her way home

“Gone with the Wind” About a strong-willed woman's tumultuous relationships and survival during the Civil War.

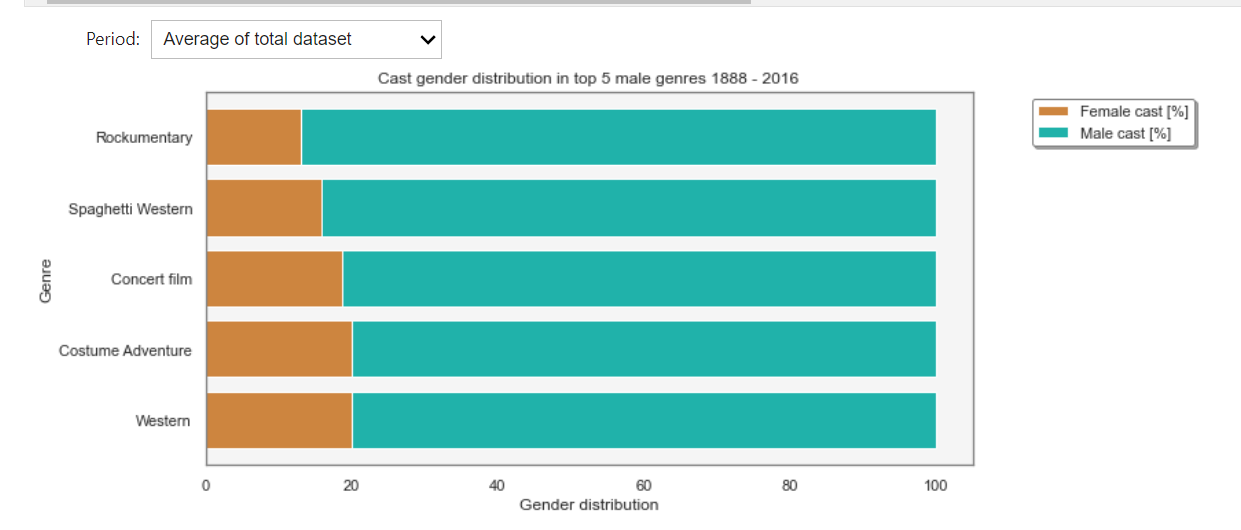
Between 1945-1990, sexual genres began to dominate the actress’s world. During this time the public attitude towards sexuality began to change. The new liberality towards sexuality grew with the counterculture movement of 1960s that caught momentum with the civil right movement. You might recognize this time with “Make love not war”. In addition, technology become more accessible, allowing easier access of explicit content at home.

Movies such as "Deep Throat" pave the way for the mainstream acceptance of erotic movie content.

In the latest period the two genres comedy and melodrama appear again with addition to Filipino movies. The genre ‘Filipino movies’ signifies movies that were produced in the Philippines. Therefore, it is difficult to check what topic this genre covers. We checked with which other genres it appears often. Three from the 5 most co-occurring genres were previously classified with high female percentage such as: Comedy, Romance Film, Romantic drama.

Even for the genres with the highest female cast ratio, the percentage of male cast is high. We discovered that actresses are more likely to play in sexual content which associates women with appearance.

Now let’s look at the male dominant genres.



We can quickly see that a lot of genres are related to violence and adrenaline, which fits into the stereotypical interest of men. It might appear apparent that war and western film mostly are dominant by men. But if we take a step back, we need to ask ourselves why we feel this way. While the frontline and western world of cowboys are dominated by the male gender, women also lived through this time. They served as nurses and administrators during the first world war, cared for the animals and were involved with ranch sales. So, to represent all sides of the story, both genders should be taken into account.

For action and adventure movies there is even less a reason why women do not have an equal appearance on screen. This finding might reflect how the society sees the role of a woman and

In both the first and last period war is listed as a genre with the dominant male cast percentage. In the time between those periods women have started taking more active roles on the frontline such as ‘Women Airforce Service Pilots’. Now the movie industry would have been able to solely focus on the frontline, while also including female cast. But compared to the first period, the percentage of women playing in war movies sank.

The fact that documentary, culture & society have such a significant higher male cast percentage might indicate that society evaluate them as more trustworthy and serious or that they prioritize their interest.

Chinese movies are listed as a male dominant genre. Upon further research we found out that under the 5 most co-occurring genres are two male dominant genres such as ‘Action’ and ‘Martial Arts Film’

After looking at the genre data we can conclude that this sentence reflects the findings: “Men have to be strong, and women have to be nurturing.” This biased view can be harmful for both genders, preventing men to express emotion or vulnerability and not enabling women to take leading positions.

Seeing the two genres related to country, the Filipino movie genre with high female cast percentage and Chinese movies with high male cast percentage makes us wonder if this means that the corresponding countries produce movies with higher, respective lower female cast percentage. Check out the next chapter if you’re also curious about this.

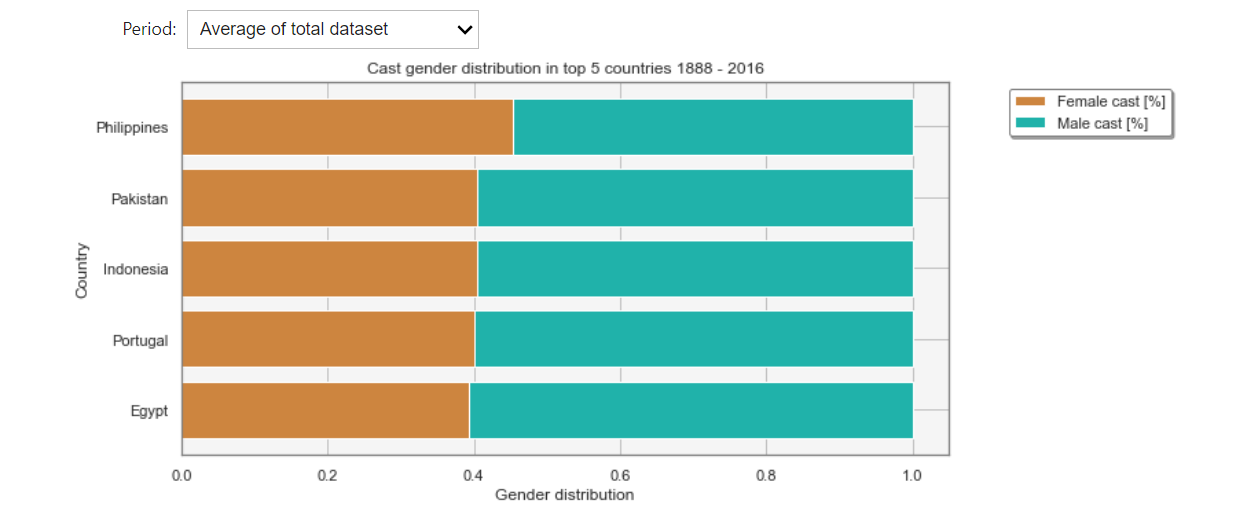
## Gender gap regarding countries

We remember that the genre Filipino movies had one of the highest female cast percentage in the last period. According to the Global Gender Gap Report 2021 Philippines is ranked 17th regarding the ‘Global Gender Gap Index’ which makes it the highest-ranking Asian country.

(<https://www3.weforum.org/docs/WEF_GGGR_2021.pdf>)

China with very low female cast percentage is ranked below average placed 107th which corresponded with the low female cast ratio in the genre analysis. This makes us wonder if the portrayed female cast percentage in movies reflect gender equality of the corresponding country.

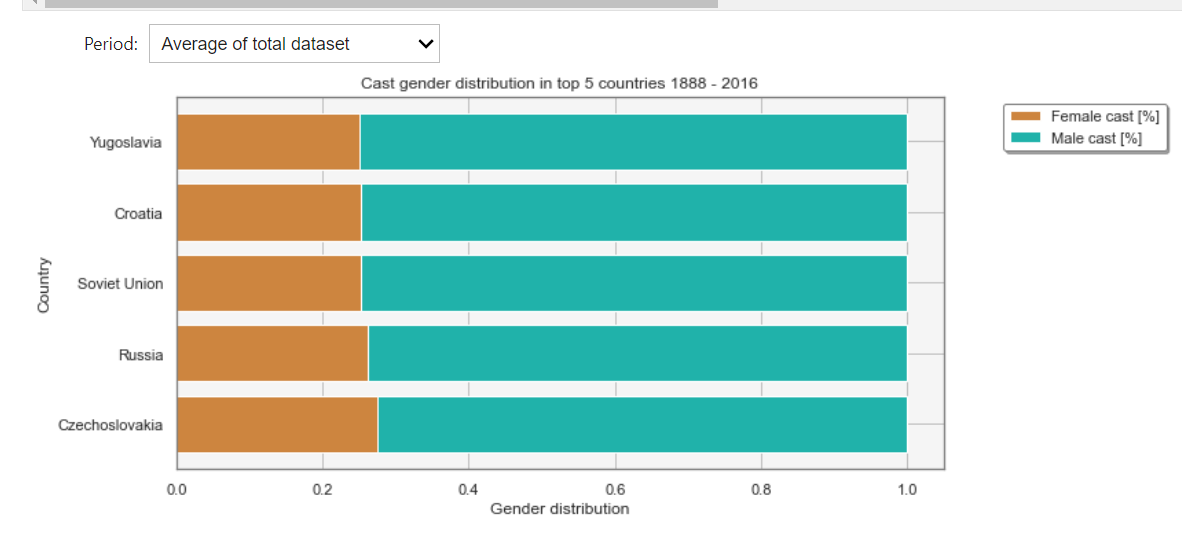
For that we computed the countries with highest female cast percentage across the periods. For this we again set a threshold of 100 movies to consider a country. So only countries that have at least produced 100 movies are taken into account.



Comparing the latest period of the dataset we can see that it does not reflect the finding of the Global Gender Gap Report 2021. Iran performs below average according to the study but is ranked 2nd according to the female cast percentage. Through that we could conclude that there is no linear correlation between female cast percentage and the finding of the World Economic Forum which measures the gender gap with following indexes:

Economic Participation and Opportunity, Educational Attainment, Health and Survival and Political Empowerment.

Same goes for the analysis of men cast percentage were Croatia performs above average regarding gender equality but has a highly skewed gender distribution.



In the plot below we can see that almost all countries are male dominant.

Ein Bild, das Karte enthält.

Automatisch generierte Beschreibung

We believe that the gender distribution in movies does not show an immediate and direct causal effect to indexes such as the ones used for world economic forum since other factors such as economy wealth and cultural tradition play a stronger role. Nevertheless, we think that it does subconsciously affect our thinking.

This is supported by David Gauntlett's Gender and Identity Theory where he states that media can influence our establishment of our own identity – there is always some representations of gender that will appeal to our own sensibilities.

But instead of only analysing this quantitative feature it would be beneficial to check a qualitative feature that shows how the characters are portrayed. This is what we’ve analysed in the next chapter.

# On-screen representation - Is it always the helpless princess waiting for her dream prince?

The world is full of stereotypes. For example, some say that women are superficial, only interested in outside appearance and primarily interested in romance and relationships.

It is important to recognize and challenge stereotypes to promote equality, respect, and mutual understanding. We therefore want to see if there is a difference between the portrayal of men and women.

## Bechtel test: Show how it behaves for time and country



In a 1985 comic strip called "Dykes to Watch Out For." Shows two characters discussing the criteria for a movie they would both want to see. The response led to the so called Bechdel test, which tests following three criteria:

https://bechdeltest.com/

This test checks if following three criteria are met

1. The work must have at least two named female characters.
2. These characters must talk to each other.
3. They must talk to each other about something other than a man.

Since then, the Bechdel test has become commonly used tool for evaluating the representation of women in media.

Everyone can participate in the data collection, so next time you watch a movie, consider contributing to the dataset here: <https://bechdeltest.com/add/>

## Sentiment analysis

The Bechdel test tell us more about qualitative female representation but there is just so much more that play into gender equality. For example, there could be a movie that that passes all three Bechdel test but still contain sexist content. How can we check harmful stereotypes?

### Based on TV tropes (Do we want to keep this? TA said it could be interesting)

The paper “Learning Latent Personas of Film Characters» chose 72-character tropes that they seem suitable as character types. Those 72 characters are connected to 501 individual characters. We assume that those 72-character correspond were picked so that they are representable for the movie industry. After manual evaluation of the described character traits, we gave assigned each trope a sentinel

- 1 represents a simple, passive, character

0 represents a neutral character

+ 1 represents an active, assertive and complex character

The sentinel score of female types resulted in -0.54 compared to -0.007 of men. We see that with our dataset, female characters more often represent simple and passive roles.

To give those numbers a bit more dimensions the top three most appearing tropes for each gender is listed below:

dumb\_blonde

brainless\_beauty

chanteuse

crazy\_jealous\_

corrupt\_corporate\_executive

byronic\_hero

Within the dataset 30 % of female character tropes are unique compared to 15 % for male tropes. So, while there are more character types for female characters, they more often represent passive characters. This can be detrimental for society. Based on the study "The Effects of Gender Stereotyping in Media on Children's Socialization" children exposed to media with traditional gender show gender-stereotyped behaviours.

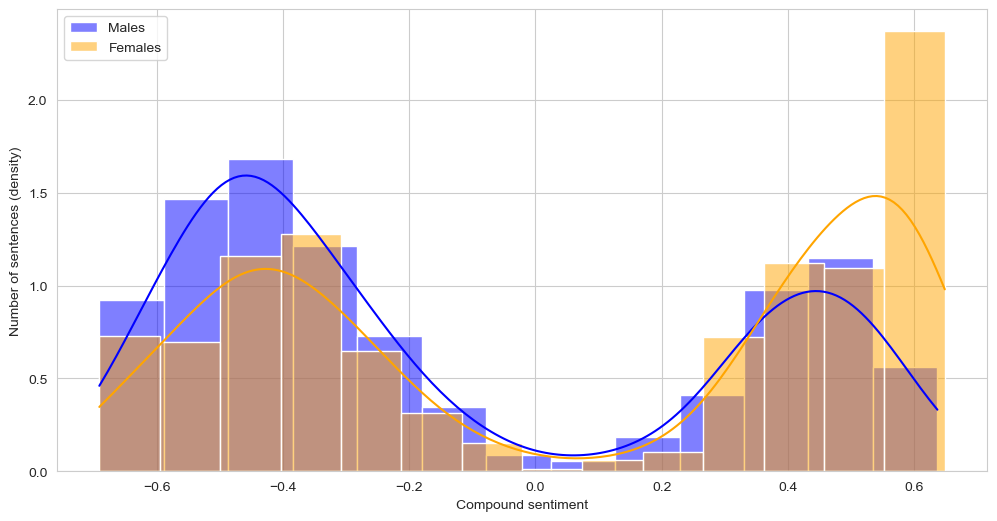
While those 72-character types give us a good first representation of how characters are portrayed we would like to go a bit deeper. Therefore, a sentiment analysis based on the movie summary was done.

### Based on movie summary

Analysing the movies summary can help us understand how male and female characters are depicted. In particular are females portrayed more positively or more negatively than males? To do so, we have compiled the most commonly used adjectives to describe the characters.



Those words give us a good idea of the adjectives used by each gender, but it is still be difficult to fully answer the question of whether females are depicted more positively or negatively compared to males. To further analyse this, we have assigned a sentiment score to every adjective on a scale from -1 to 1, with 1 being positive, -1 being negative, and 0 being neutral. Then we create a histogram of these scores, without the neutral score since lots of words are neutral.

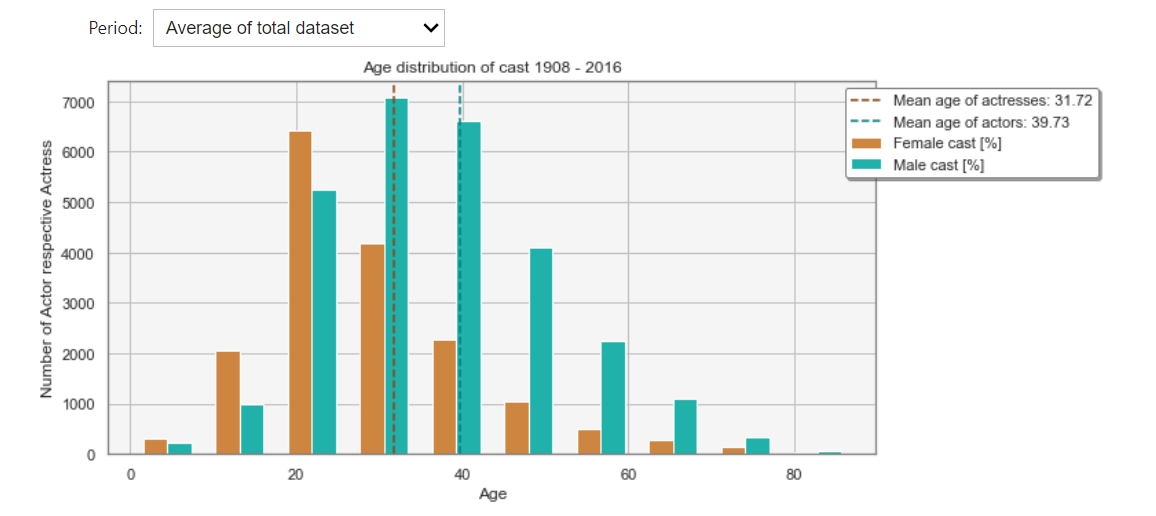


We clearly see that the females are generally described more positively, especially with extreme positive adjectives. This may be because females are frequently highlighted using positive adjectives such as beautiful, attractive, young or charming, since appearance is often particularly important for female characters. This is in contrast to male characters, who may not be as frequently described with such emphasis on appearance.

## Age of actor

It's a tale as old as time: men age like fine wine, while women are left to wilt on the vine. In society there tend to be a stigma that older women are no longer wanted. Have you ever heard of “Shèngnǚ”? It literally means “Leftover Woman” and describes unmarried women in their late twenties and beyond.

Let’s see if the age distribution in the movie industry is skewed.



For the last period representing, women are in average 8.5 years younger than their counterpart. The age distribution of women is right skewed, showing a dramatic decline once the “best age of 20” is reached. This could imply that looking young and fitting into today’s beauty standard is more important for actresses than actors in movie industry.

The same values are reflected through the results of the other periods.

# Behind the scenes power - Has the world in his hands

The feminist film critic Laura Mulvey believes that the movie production is in control of men who make decisions that appeal to their own values and interests.

(https://media-studies.com/male-gaze/#:~:text=The%20male%20gaze%20refers%20to,to%20accept%20this%20narrow%20representation. )

She describes a term called ‘the male gaze’ where female characters are often depicted as attractive rather than having a fully rounded, complex character and directs the root of the problem of men being dominant in the powerful behind the scenes roles. We want to see if this is true, therefore we analysed data about writers and directors.

## Time series

Let’s start off again with a time series. To put the numbers into relation we also show the female employment rate. Compared to the growth rate of female employment rate, the female writer and director growth rate is much lower. Fitting a 2-degree least squares model shows that in 2127 the female writer percentage will be equivalent to the male percentage and for the female director equal employment rate will be reached in 2133.

The growth rate of the behind-the-scenes roles are much slower compared to the women cast percentage.

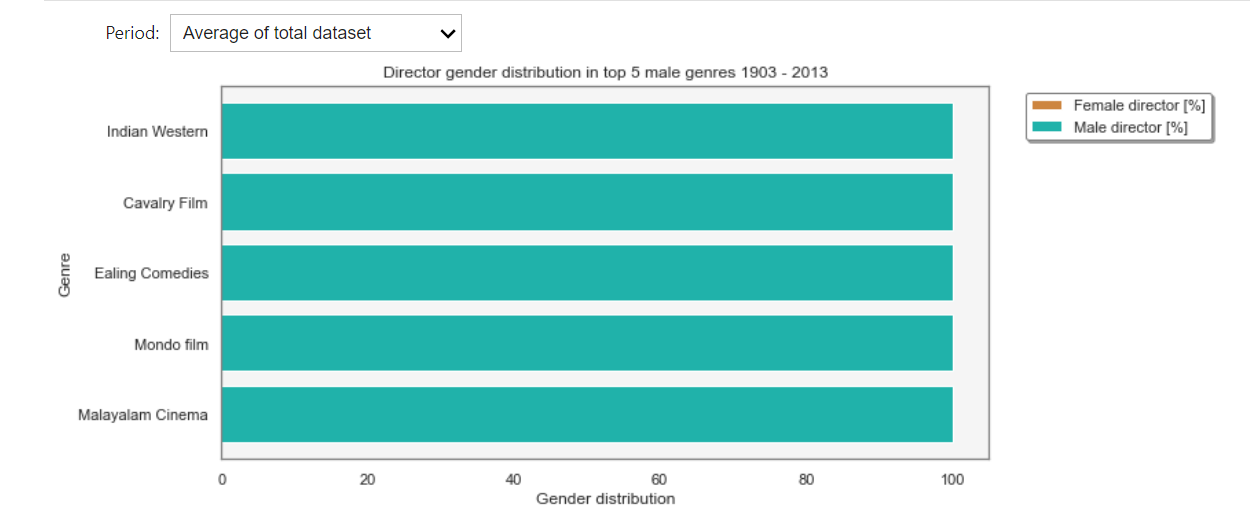
Ein Bild, das Text, Screenshot, Schreibgerät, Briefpapier enthält.

Automatisch generierte Beschreibung

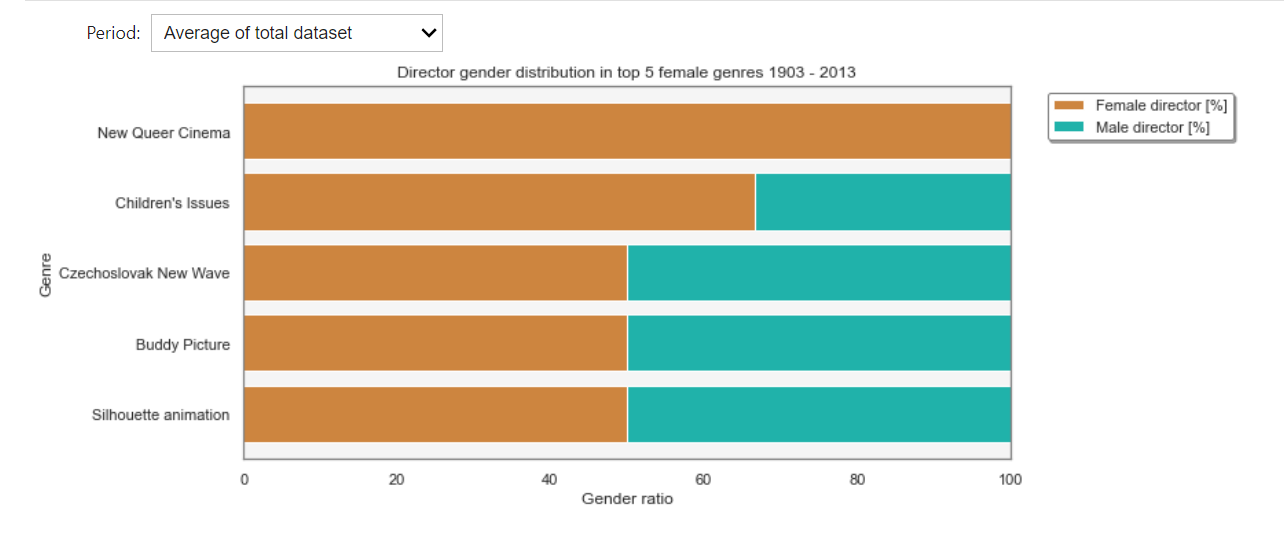
Having fewer female voices and perspectives might lead to a more inaccurate portrayal of women and their experiences which can lead to one dimensional portrayal of female characters.

## Gender gap for directors

First, let’s look at directors. Overall, there are 29 from 357 genres that solely have male directors. We show you the most popular genres within those in the figure below:



The genres cover a wide range of topics. There are many comedic genres such as “Bloopers & Candid Camera” and “Courtroom Comedy”. Others are action related such as “Prison escape” and “Sci-Fi Thriller”. Lastly, there are some cultural content such as “Journalsism”, “Indian Western”. All in all, we can clearly see that during each period you can find genres where only men director participates. Let’s see how it looks like for female directors:



For the genres dominated by female directors’ societal issues are often in the centre such as queerness, children’s or gender issues. There are also political topics such patriotic films, avant-garde, which is a political and artistic movements towards the idea of progress) or the Czechoslovak new wave which consists of socially critical films after the liberalization of the socialist regime and the lifting of censorship. Interestingly production in the former British Indish area is often a top female director genre with Bollywood, Tollywood and Bengali movies in the first periods. Lastly, there are two emotional touching movies such as ‘Buddy picture’ and ‘Coming of age’ movies.

Based on these findings there seems to be a trend where women are more likely to be in charge of emotionally charged content compared to men with wilder topics such as action and comedy.

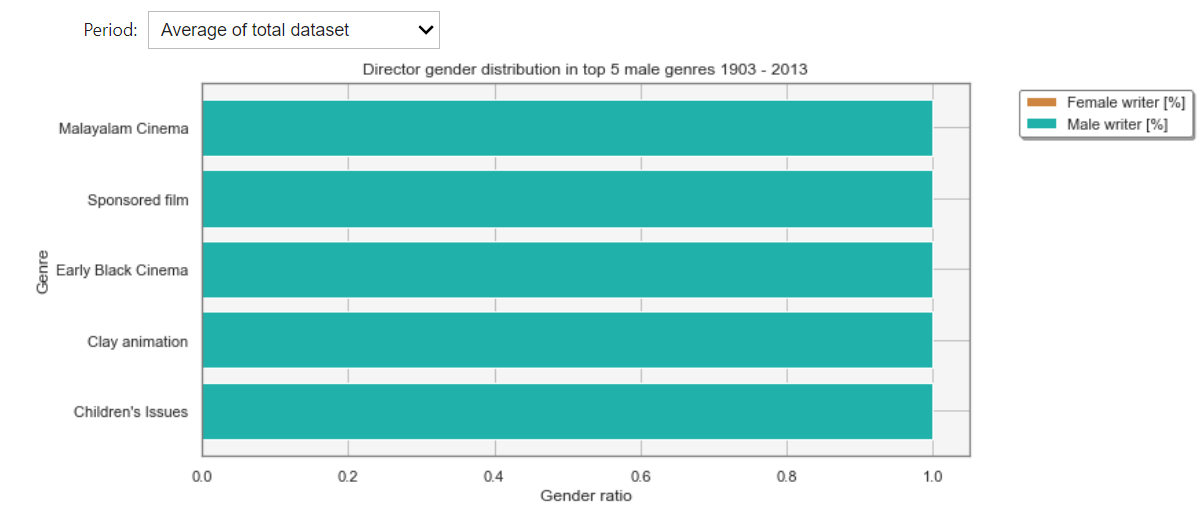
Let’s circle back to the phenomena of “male gaze”:

The genres with higher female director percentage are not typically known for this phenomenon. Male dominant genres such as action have been criticized wildly for sexualizing women. An example are the various studies about male gaze in superhero films.

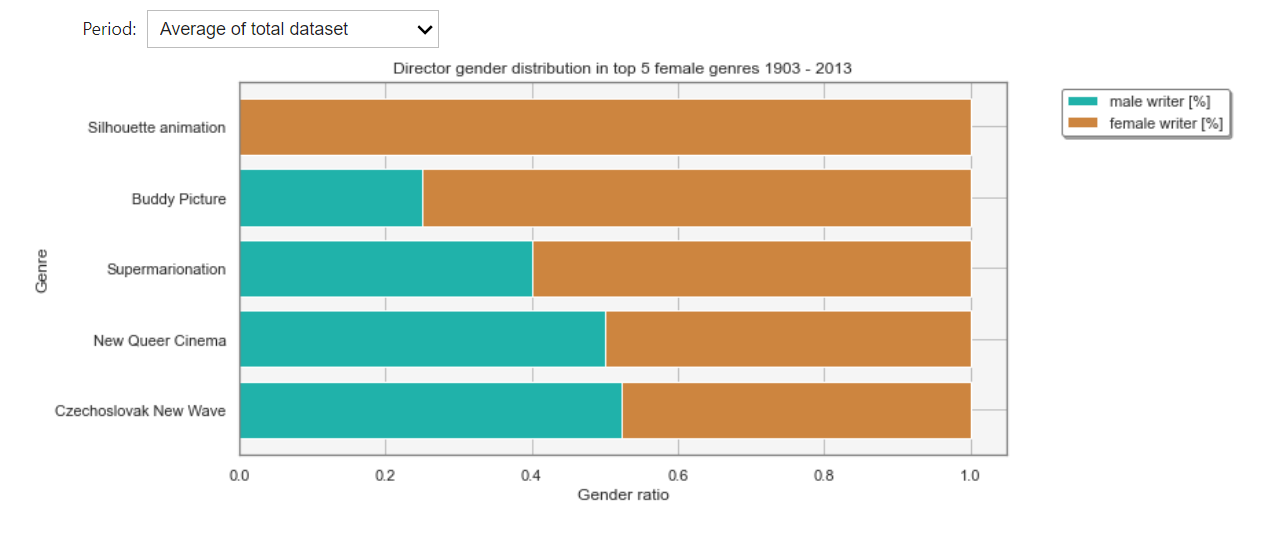
We believe that having more men director does not by default corresponds to ‘male gaze’ but that compared to women, men might not be as aware about this problem. With an improvement of gender representation on-screen, the phenomena of “male gaze” might decrease for all genres.

## Gender gap for writers

A total of 16 from 358 genres solely have male writers. The topics are very widespread and there is no clear pattern that can be seen.



The same goes for the female writers.



Only genres were considered where at least 100 movies belong to. Nevertheless, the genres with most extreme gender distribution belong mostly to niche genres. One could conclude that for writers, the topic is less biased towards gender stereotypes than directors.

# Conclusion

The analysis of the gender gap in the movie industry shows that while there is an improvement to be seen, the gender gap is still present many areas. The percentage of male cast, male director and writer is greatly higher compared to women with some variation between genres and countries.

Female on-screen representation based on tv tropes is poor.....Sentiment analysis .. Bechdel

The cultivation theory says that the more we see a particular representation, the more we will believe it is important and true. We should strive for good representation, both for men and women to break down society’s inequalities.

Not just society profits from a more diverse and well represented movie industry. Studies such as the one from In the media research agency Shift7 in collaboration with leading agency CAA (<https://shift7.com/media-research>) shows that female-led films did best, at every budget level.

Lastly, it is important to mention that the quality of the findings depend heavily on how representative the data is.

# Notes (Just for me while writing)

Studies:

"Gender Bias Without Borders: An Investigation of Female Characters in Popular Films Across 11 Countries" by the Geena Davis Institute on Gender in Media: This report analyzed the representation of female characters in popular films across 11 countries and found that female characters were less likely to be depicted in leading roles and were more likely to be sexualized or depicted in stereotypical roles.

"The Celluloid Ceiling: Behind-the-Scenes Employment of Women on the Top 100, 250, and 500 Films of 2019" by the Center for the Study of Women in Television and Film: This report analyzed the employment of women in behind-the-scenes roles on the top 100, 250, and 500 films of 2019 and found that women were underrepresented in many key roles, including directing, writing, and producing.

"It's a Man's (Celluloid) World: Portrayals of Male and Female Characters in Top Grossing G-Rated Films" by the Geena Davis Institute on Gender in Media: This study analyzed the representation of male and female characters in top grossing G-rated films and found that male characters were more likely to have leading roles and to be depicted in more active and assertive ways, while female characters were more likely to be depicted in supportive or passive roles.

**Inequalities in the film industry**: With the recent Zurich film festival the subject of gender gap in the film industry was brought up more. Not only are men heavily dominant in the behind-the-scenes roles as directors, writers, and critics but the on-screen roles are also given to men more often. In the study “It's a Man's (Celluloid) World” it was shown that 85% of films featured more male than female characters.

To tackle the problem of inequality it first needs to be understood properly. Therefore, a deeper analysis about diversity and inclusion will be done especially with respect to traditional gender roles, racial or sexual stereotypes and to which kind of people roles are given most to. Different genres will be compared to each other and the change of inequality over time will be studied. To find out if a certain part of the world is more advanced regarding equality movies with different languages will be evaluated. During this step a representative pick of movie samples is important to avoid any artificial skews.

gender equality is not just a buzzword or a trendy topic