Wikipedia is an open source online encyclopedia of knowledge, and one can find information on anything there. However, our male dominated society has led to an imbalance in sources that provide information to Wikipedia, as male contributors seem to vastly outnumber female contributors. This paper analyses the data from 20 'male identified' movies and 20 'female identified' movies, and compares them to show that this hypothesis is a real phenomenon. Our source was a movie rating website called movielens.org, which is owned and run by GroupLens, which is a research lab in the University of Minnesota.

Firstly, we tackled the problem of the criteria regarding the segregation of movies into the categories of male identified and female identified. Our initial approach was to sort the movies according to which gender rated the movie more often. However, this approach would skew our analysis as it is possible that raters can rate movies that are not particularly aligned with their gender. Our second approach was to sort by which movies had the highest number of male raters and the highest number of female raters. However, this also didn't work as the most popular movies had hundreds of ratings from both genders, and this would lead to popular movies dominating both categories. In the end, we decided to use the difference between the number of male and female raters to find out whether a movie was male identified or female identified, as if more males rated a movie as compared to females, it can reasonably be classified as male identified, and vice versa. By doing this, we were able to get a comprehensive list of the top twenty male identified and female identified movies.

Our next step was to get the length of the Wikipedia articles associated with these movies. The movies in the movies.dat file gotten from MovieLens did not match up with the titles of their corresponding Wikipedia articles, and so we had to change the names in the file to match the ones in the Wikipedia article. Once this was done, we ran our program and got some very interesting results.

We found that the average length of the Wikipedia articles of the top twenty male identified movies is 316,966, and the average length of the Wikipedia articles of the top twenty female identified movies is 54,067. This shows us the vast disparity between male identified movies and female identified movies, where the articles for female identified movies are one order of magnitude smaller than the articles for male identified movies. This is a huge difference, and there are many possibilities as to why this is occuring.

One possibility is that women have less access to movies and the internet than men, and as such, they do not rate movies as often as men do. This can be explained by the patriarchal conditioning that women go through in our heteronormative society, where women are told to take care of the house and have kids and devote all their time to making the home a better place. There is not much time for entertainment for women under these conditions, and as such, it is possible that women either do not watch a lot of movies of their liking, or do not have equal access to the internet.

Another possibility is that men are more prevalent in the movie industry than women are. Wikipedia is a public sourced encyclopedia, and people who contribute to the articles of movies are often those who work within the movie industry. It is known that women face a lot of sexist behaviour within the industry, and this leads to a lack of female raters who can provide their opinion online. It is also a trait of heteronormative behaviour for women to be submissive and not speak out and voice their opinions. Thus, they are generally not inclined to rate movies or voice their beliefs on Wikipedia. This can explain the disparity in the article lengths, as more men are able to add to the articles of male identified movies than women are to the articles of female identified movies.

In conclusion, the heteronormative standards of society that women are forced to comply with can be identified as one of the explanations for this disparity in the lengths of Wikipedia articles, and this disparity will hopefully lessen in size over time as feminist movements progress and grow, leading to the formation of a society with more equality.