

Feminism

Saumya Pandey

2022-09-20

In 2022, the topic, feminism, creates divides. There is one side of people (including females) that don't think of themselves as feminists. Their reasons include that it is a man-hating group that exists solely for the benefit of women. They believe that the feminist-group believes in them being superior than males and that they deserve more. While, there are some feminists who certainly believe in that and preach that, for the remainder of the people believing in feminism, it is further from the truth. For them, feminism is a movement to support them in getting the same opportunities as men - whether it be in terms of career, salary, etc. So, as just described, there is a lot of controversies and mixed opinions surrounding feminism at the present. This is what feminism is like in 2022. But what was it like in the earlier years and centuries? What were people striving for back then? What were the issues and concerns raised at that time?

The answers to these questions can be found through a data set that was compiled for the Philosophy Data Project. The data set contains 51 texts written by different authors. It contains attributes like "school" (categories include rationalism, feminism, empiricism, etc.), "author", "title", "sentence_spacy" (sentences from a text), "original_publication_date", etc. By looking at the data set, there are three authors that fall under the "school" of feminism. The authors were Mary Wollstonecraft, Simone de Beauvoir and Angela Davis. The data set includes one book for each author. Mary Wollstonecraft wrote the book called "Vindication of the Rights of Woman" in 1792. Simone de Beauvoir wrote "The Second Sex" in 1949 and Angela Davis wrote the book "Women, Race, And Class" in 1981. Through data analysis techniques on this subset of the entire data set, I was able to get a sense of what feminism was about earlier through the general ideas that were presented in each text, the way the authors expressed themselves in the text, etc.

The first question is what were the emotions of the author in a text? Did they express with positive or negative sentiments? If negative, then what kind of emotions were expressed? Did different authors tend to have more negative (or positive) sentiments than other authors? It was of interest of me to gauge the overall sentiment of the author.

Sentiment analysis accomplishes this. Here, I used the `get_nrc_sentiment` function on each text. This function uses the NRC sentiment dictionary to calculate the presence of eight different emotions (anger, fear, joy, anticipation, etc.), as well as a positive and negative sentiment valence. To showcase the emotions expressed by each author, as well as the positive and negative sentiment valences, I made some visualizations as well which are down below.

To get more detailed steps of the implementation, please do take the time to read my comments in the code which is in the RMD file.

Positive vs. Negative Sentiments Expressed from each Author

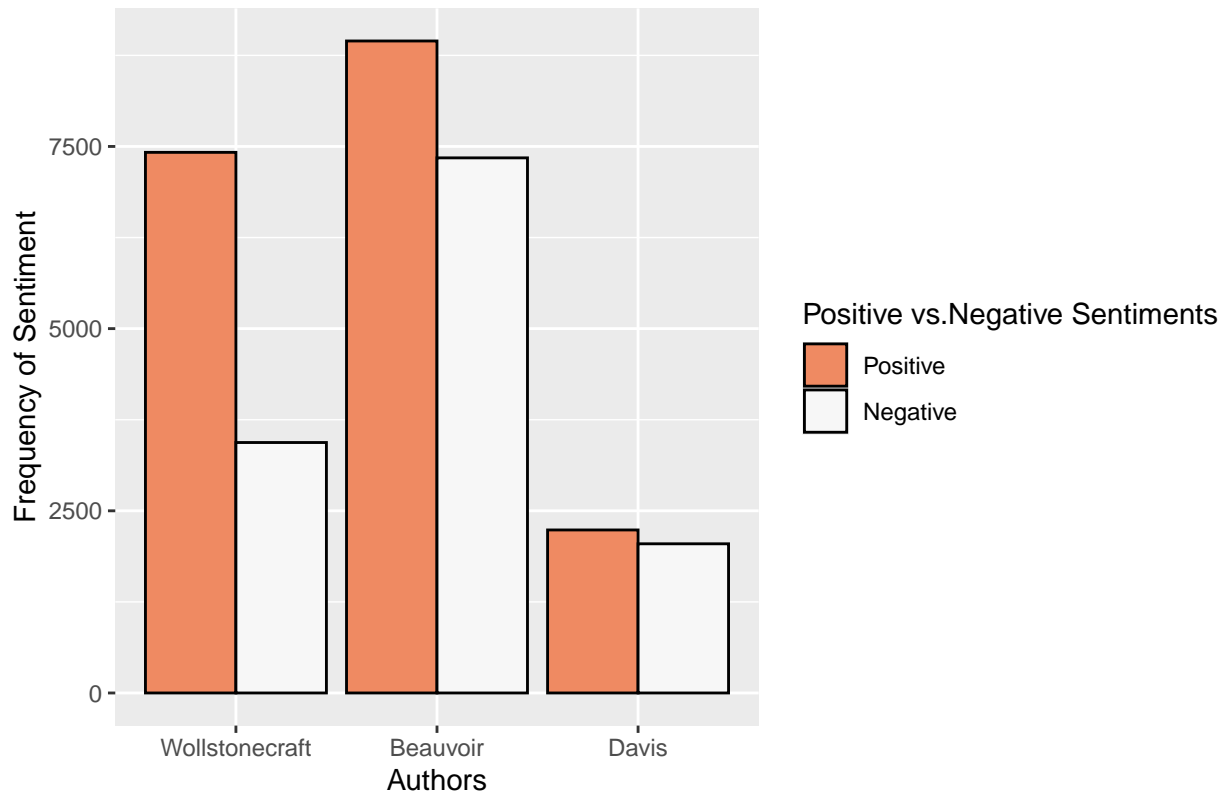
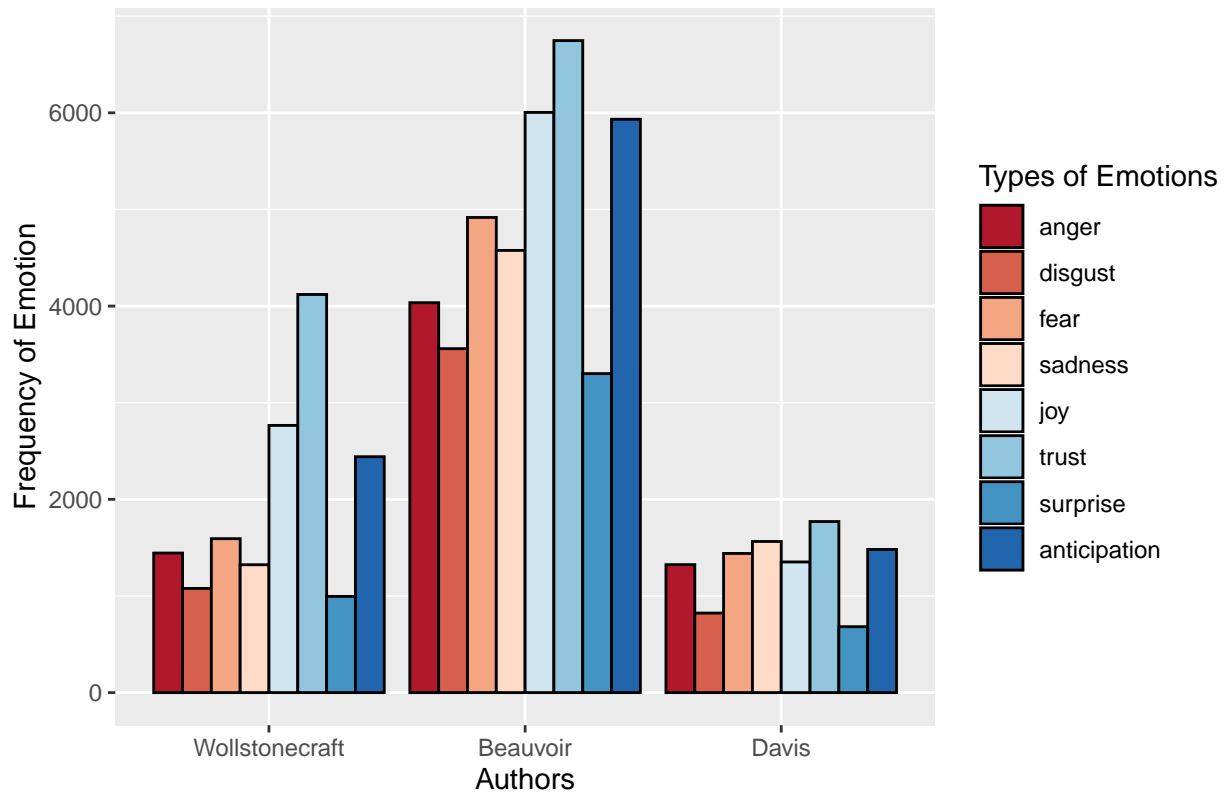


Chart of Types of Emotions Expressed by each Author



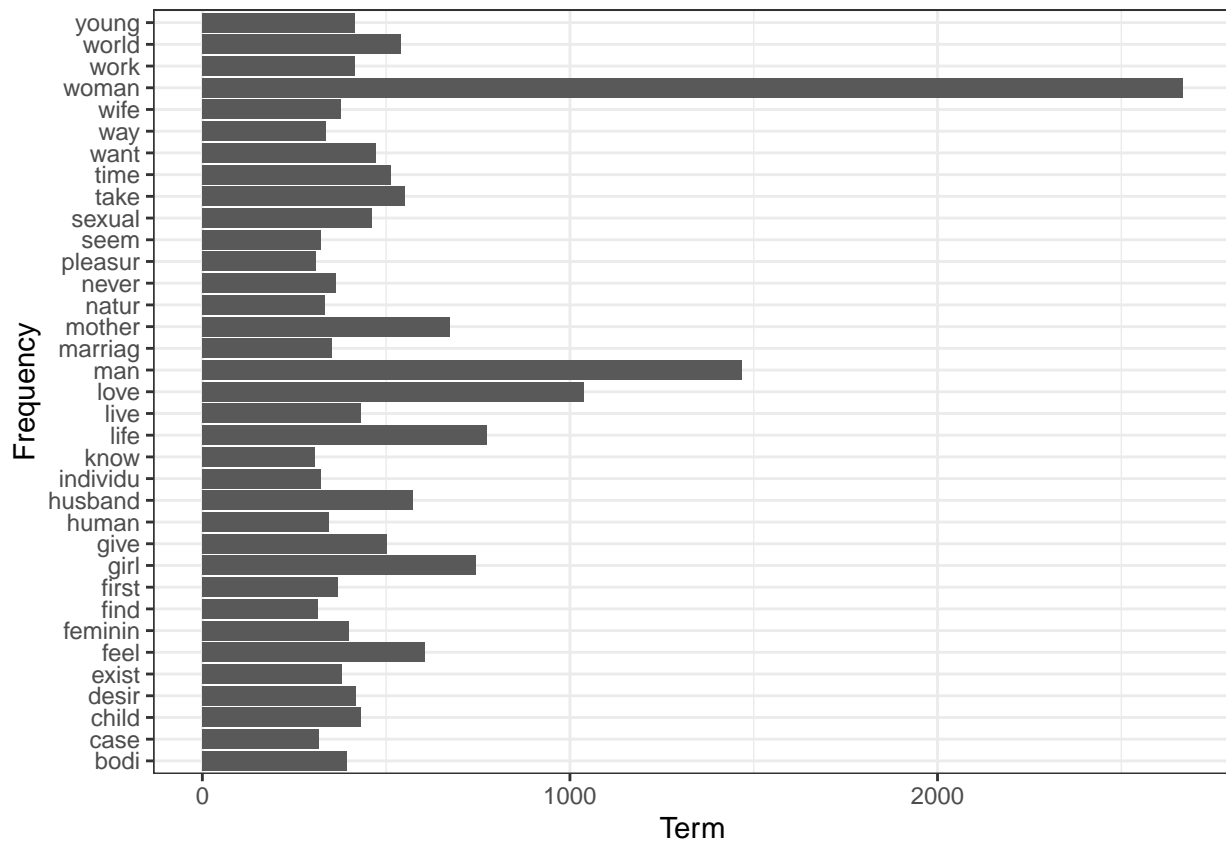
Looking at the first plot, it can be seen the amount of positive and negative sentiments for Beauvoir are a lot

higher than the other authors. The reason is that there are a lot more sentences present in Beauvoir's case than there are for the remaining two authors. Looking at Beauvoir's example in the first plot, it can be seen that there are a lot more positive sentiments expressed. However, the margin between negative and positive sentiments is not too great which means that there were some serious topics discussed as well that could have led to negative sentiments. Looking at the first plot for Wollstonecraft's instance, it can be seen that positive sentiments are hugely expressed by the author. The margin between negative and positive sentiments is also wide, compared to Beauvoir's case. Looking at the first plot for Davis's case, this was an interesting one. The margin is quite low between negative and positive sentiments expressed which means the level of positive and negative sentiments expressed was almost the same. It means that there were several serious topics discussed in the text which could have led to negative sentiments. To get a better idea of what kind of negative and positive sentiments were expressed, the second plot can be looked at. In Wollstonecraft's case, since positive sentiments are expressed hugely, the top positive emotions that are expressed are trust and joy. In Beauvoir's case, the margin between negative and positive sentiments is not by a wide margin. So, the top positive emotions expressed were anticipation and trust while the top negative emotions that were expressed were fear and sadness. For Davis's case, the margin between negative sentiments and positive sentiments is even lower. Here, the top negative emotions expressed were fear and sadness while trust and anticipation were the top positive emotions expressed.

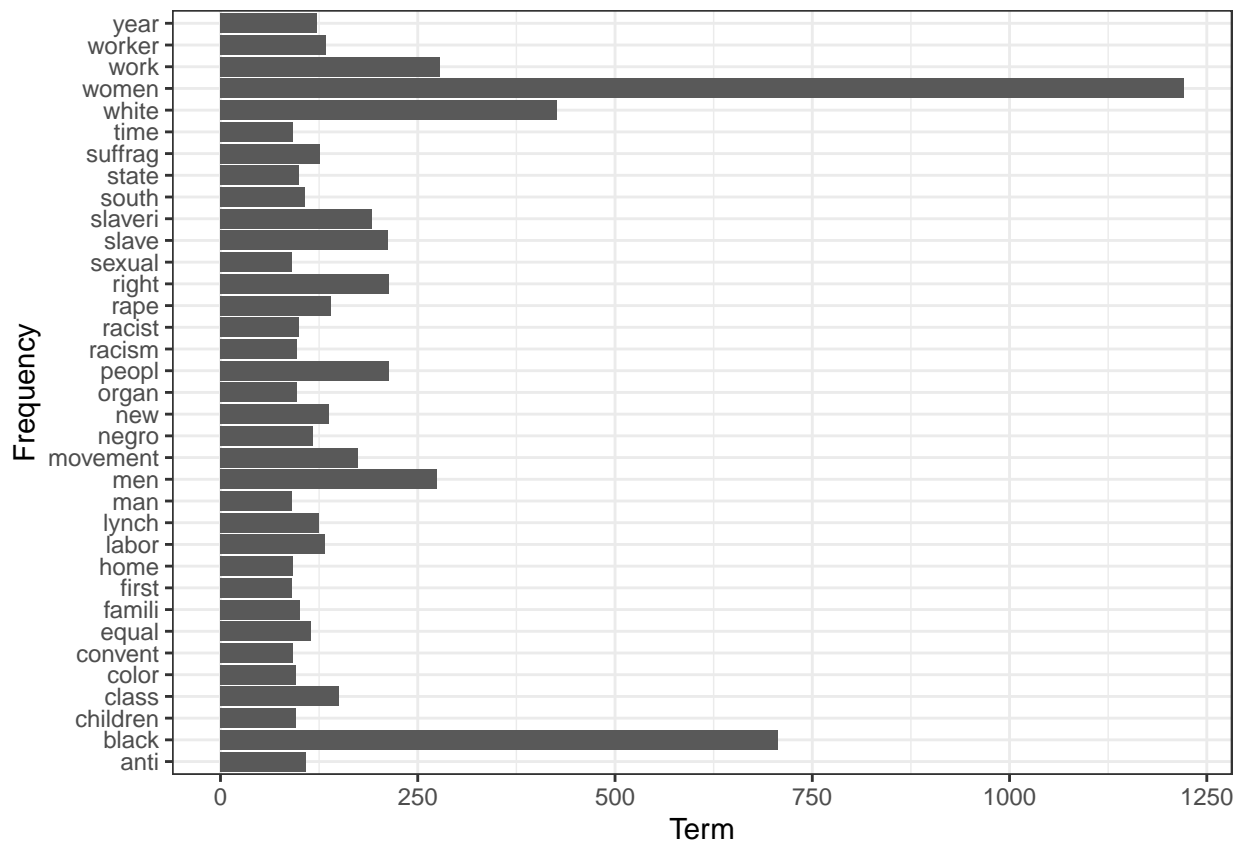
So through this analysis, I can conclude that the text written by Wollstonecraft was expressed with a lot more positive sentiments (trust and joy) than other texts since the margin between positive and negative sentiments was quite huge compared to the margins of the other texts. I believe Davis's text has more negative sentiments (fear and sadness) than the rest of the authors since again in this case, the margin between the positive and negative sentiments is very very low compared to the margins of the other texts.

Davis's case is what led to thinking what ideas was she talking about that led to negative sentiments being expressed - and that too, at a frequency where almost the positive and negative sentiments being expressed is equal? What ideas was she discussing that led to sadness, anger? The same type of questions can be asked regarding the other two authors. What is so different in the ideas of Wollstonecraft vs. the rest of the authors that makes her express positive sentiments hugely and that too by a wide margin? This leads to the second analysis which is finding out the general and popular ideas behind the text. One approach is to find the words with the most occurrences in the text. A word cloud is included for visualization purposes.

women towards her children, in society, etc.



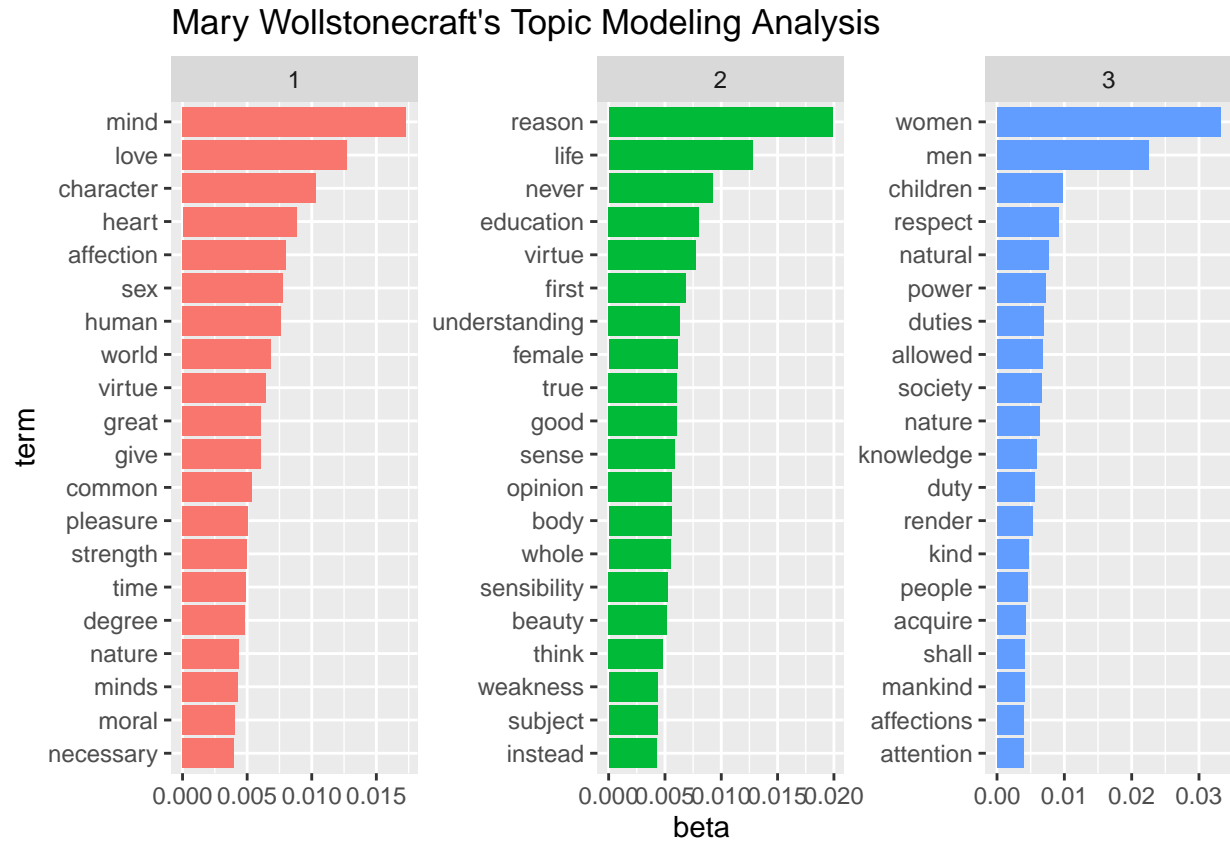
As it can be seen through the frequency plot and word cloud, besides from woman and man, terms like world, work, sexual, object, feminine, mother, marriage, husband, etc. are at a high frequency. From this, it can be deferred that some main ideas or topics in the book could be the topic of marriage, the topic of their role and responsibilities as a mother and to her husband, the topic of being feminine, the topic of objectifying women.



As it can be seen through the frequency plot and word cloud, besides from women, terms like white, black, work, slave, movement, suffrage, rape, labour, racism, oppress, campaign are at a high frequency. Some main ideas or topics in the book could be the topic of women's movement and campaigning for it, the topic of the suffering that women went through - rape, slavery, lynching, the topic of race and inequality (white vs. black women and men).

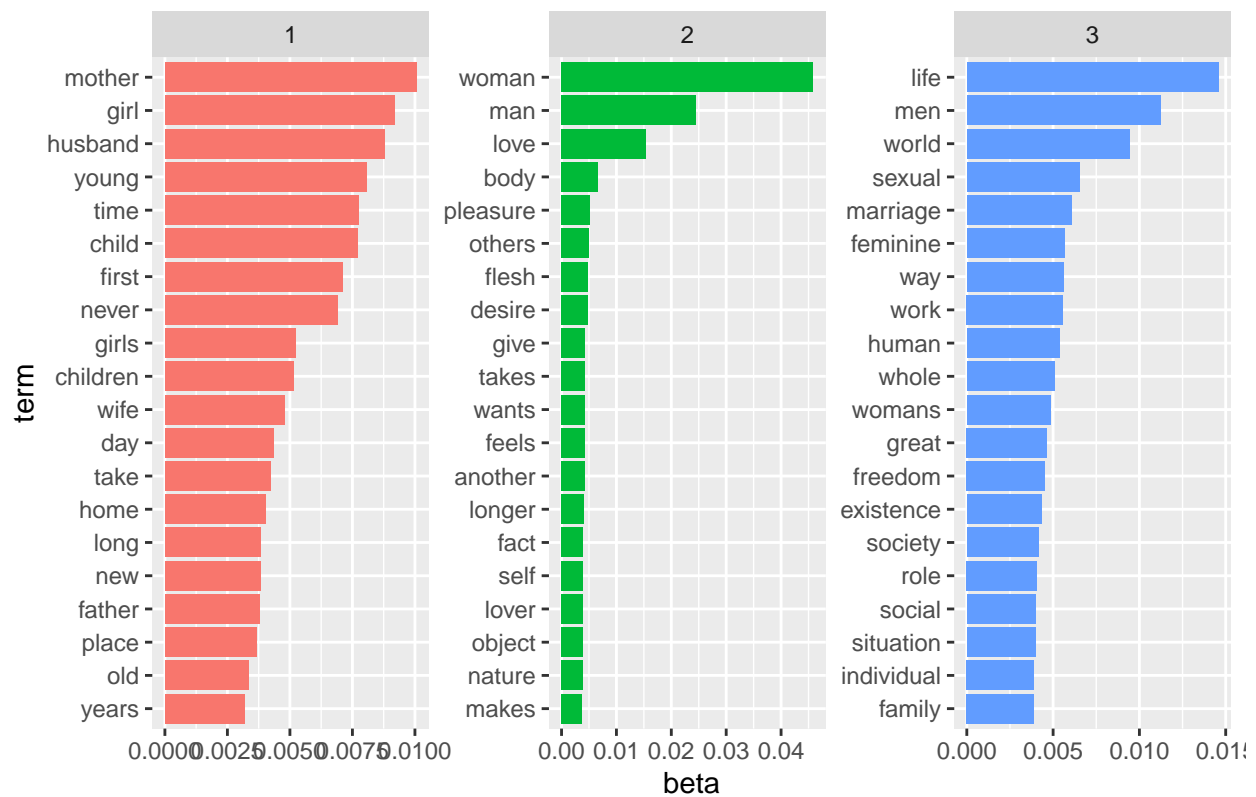
While the simple word frequency metric gave good, general ideas about each text, there are better approaches to understanding the context behind the words better, to understand even more about what the author was thinking while writing down this words. One approach would be topic modeling. Topic modeling looks at groups of words that often occurred together in documents (in this case, sentences), instead of counting each word individually. This approach helps in capturing the broader context of the text. If we can understand

what ideas the groups represent based on the most popular terms, then we can be more definite about the main ideas talked about in the text. In this analysis, the LDA model using Gibbs sampling was used through the LDA function on each text individually to figure out the three topics/ideas for each text. When the model was created, there are some specific elements of interest - the beta values. Beta values basically show the probability of the word being associated in the topic. Since three topics are to be evaluated for each text - for each topic, the top 20 words with the highest beta probability are shown in a bar chart. The higher the beta value, the more frequent the term appears in the topic.

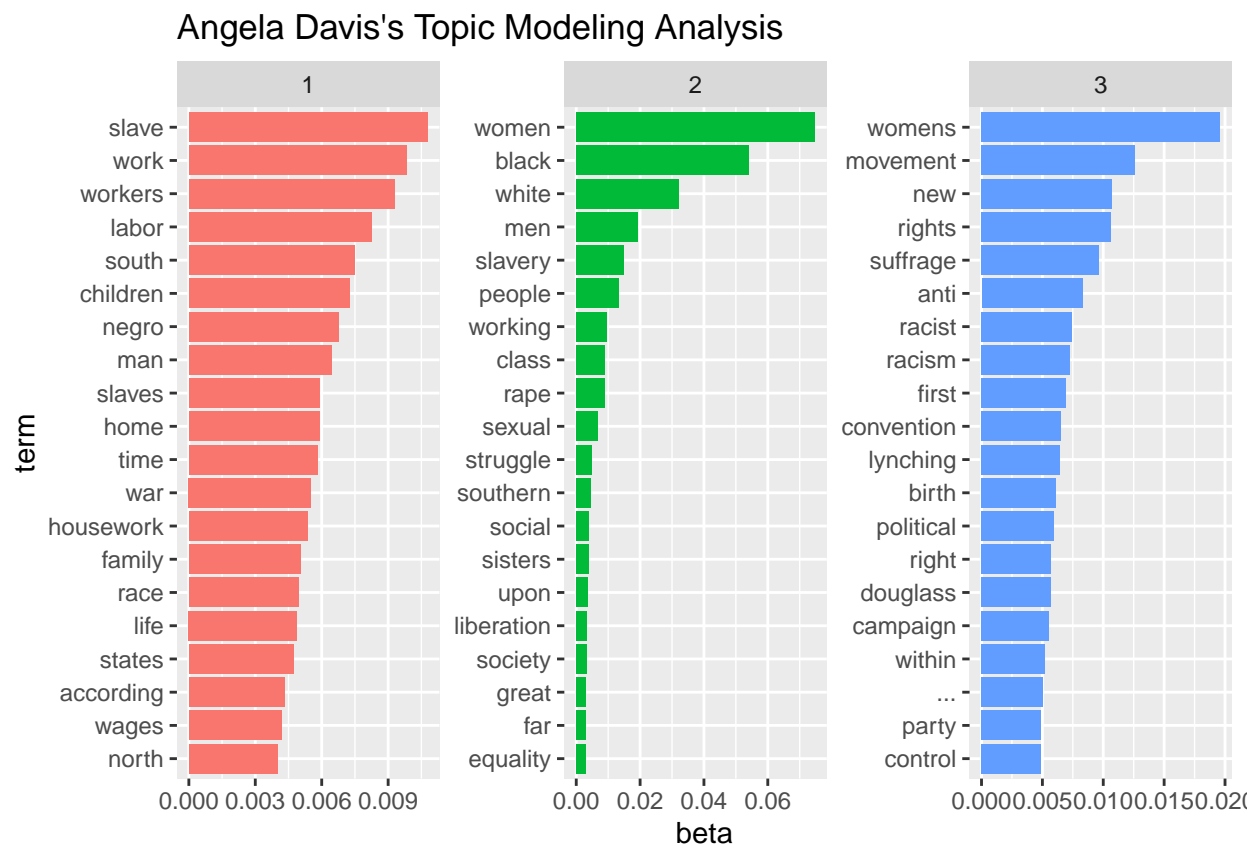


Looking at Mary Wollstonecraft's topic modeling results, the third topic has women, men, children with the highest beta probabilities which means that in this topic, these words occurred the most. Following that, words like society, duties, allowed, respect, power occur. Hence, like mentioned in the discussion regarding the word frequency plot and word cloud for this author, a general topic would be of the role and responsibilities of women towards her children, towards her husband. In the second topic, words like reason, life, education, virtue occur the most. Words like weakness, beauty, occur too. This could be highlighting the importance of education and virtue so to not seem as a weakness and that the focus should not be just on beauty instead. The topic of education was also mentioned in my discussion of the word frequency plot and word cloud. In the first topic, mind, love, character, heart, affection, virtue are the top words with the highest beta values. Followed are the terms like strength, necessary, nature, etc. The topic could be on developing morals and great virtues, character.

Simone de Beauvoir's Topic Modeling Analysis



Looking at Simone de Beauvoir's topic modeling results, the third topic has life, men, marriage, world as the top words with the highest beta probabilities. These words occur the most in the topic. Followed are terms like existence, society, role, family, etc. This could be highlighting what the role of a female, her existence is - her world is her marriage, her family, etc. This was mentioned in my discussion regarding the word frequency plot and word cloud as well. Looking at the second topic, words like man, woman, love exists the most in the topic. Followed are terms like body, pleasure, flesh, object, desire. The topic could be what women could be perceived or objectified as - as an object for pleasure, desire, etc. This was also mentioned in my discussion regarding the word frequency plot and word cloud. The first topic has words like mother, girl, husband occur the most in the topic. Followed are terms like day, home, children, first. An idea in the text could be of the routine of the female - staying at home all day, child is her first priority, etc.



Looking at Angela Davis's topic modeling results, the third topic has womens, movement, rights, suffrage as the top words with the highest beta probabilities. These words occur the most in the topic. Followed are terms like anti, racism, douglass, campaign, etc. This could be highlighting new changes for female - women's movement regarding their rights, suffrage, campaign against anti- racism. It also mentions the term douglass, which I believe refers to Frederick Douglass who was one of the prominent advocates of women's rights in the nineteenth century. This topic was mentioned in my discussion regarding the word frequency plot and word cloud too. Looking at the second topic, words like women, black, white exists the most in the topic. Followed are terms like men, slavery, class, rape, struggle. The topic could be the inequality in class and race (white vs. black) and the suffering that women went through. This was also mentioned in my discussion regarding the word frequency plot and word cloud. The first topic has words like slave, work, children, negro, man occur the most in the topic. Followed are terms like war, race, slaves. An idea in the text could be of the exploitation of children and men of black african ancestry as slaves, for war, for labor, etc.

To conclude the analysis, it makes sense now as to why the margin between negative and positive sentiments is so low in Davis's text. It is because such serious topics were discussed - regarding inequality in class and race, regarding the many types of suffering the women went through like rape, racism, oppression. However, since the margin was low, this also means there were a series of topics expressing positive sentiments (anticipation and trust) which there were - women's movement regarding their rights, discussion regarding Frederick Douglass, suffrage, campaigning, etc. On the other hand, Wollstonecraft expressed positive sentiments by a huge margin, compared to other authors. This could be because it might have been a text to educate women on the importance of education and how it can serve women well. There's a difference there.

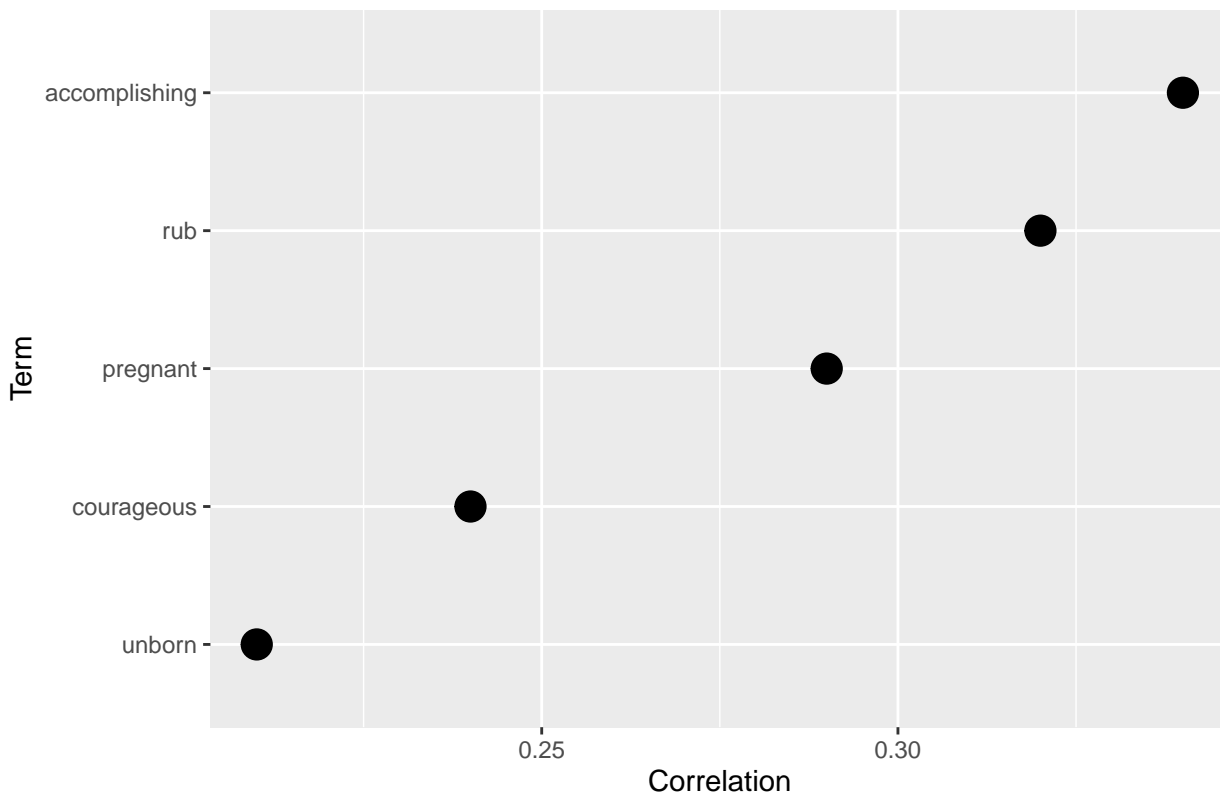
Another insightful question to explore was regarding finding what the opinion of females in other schools were - how were they perceived as, were they seen in a negative manner or were they respected? To get a broader context regarding females in other schools, I used a function called findAssocs. This helps in finding associations between terms - so when the text discusses females, what are the other terms used? For any given word, this function calculates the correlation with every word in the term document matrix or document term matrix. Though, the function is done at document level. For every document that contains

the word asked for, the other terms in those same documents are associated. Documents that don't have the word are ignored. In this case, a document is each sentence of all texts from a school. I want to find the broader context when the term, female or woman or women, etc. is present in a sentence so finding good associations in the same document or sentence is what I am looking for. Documents or sentences that don't contain the term, female or women, etc. are simply ignored which is fine. The scores range from 0 to 1. A score of 1 means that the two words always appear together in a document/sentence. A score near 0 means that the words do not really appear together in a document/sentence.

I tried this function for all schools where the terms to search for were: female, male, man, woman, girl and boy. The correlation/association between the mentioned terms above and other terms in the document/sentence has to be more than 0.2. Not all the associations between each mentioned term above and every other word is shown here as there were many that did not result in anything insightful. Also, not all schools are mentioned due to the same reason. For instance, the texts from the schools, Plato and Aristotle, did not have anything related to females, women, woman, etc. So, only the interesting and insightful associations are mentioned down below.

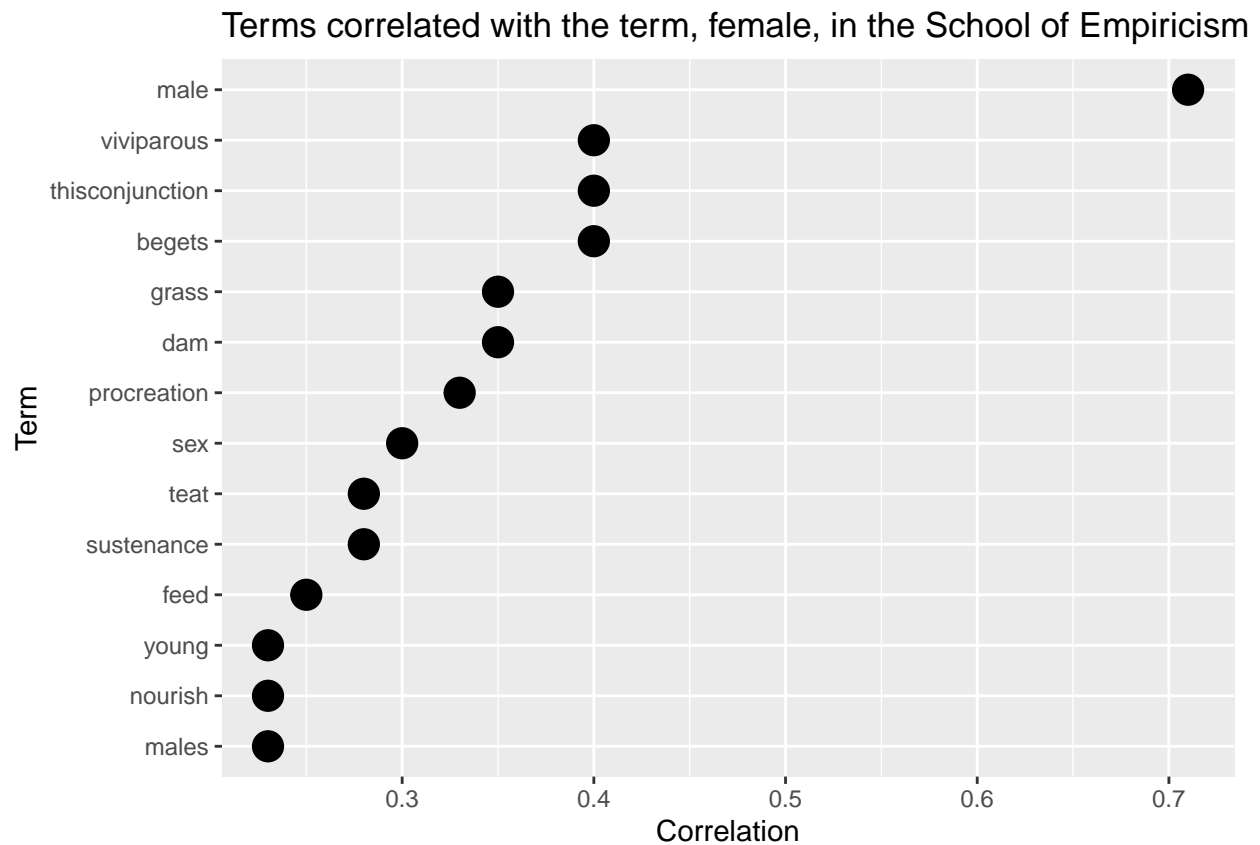
To know more about the detailed steps, please do take the time to read the comments in the code which is in the

Terms correlated with the term, women, in the School of Rationalism

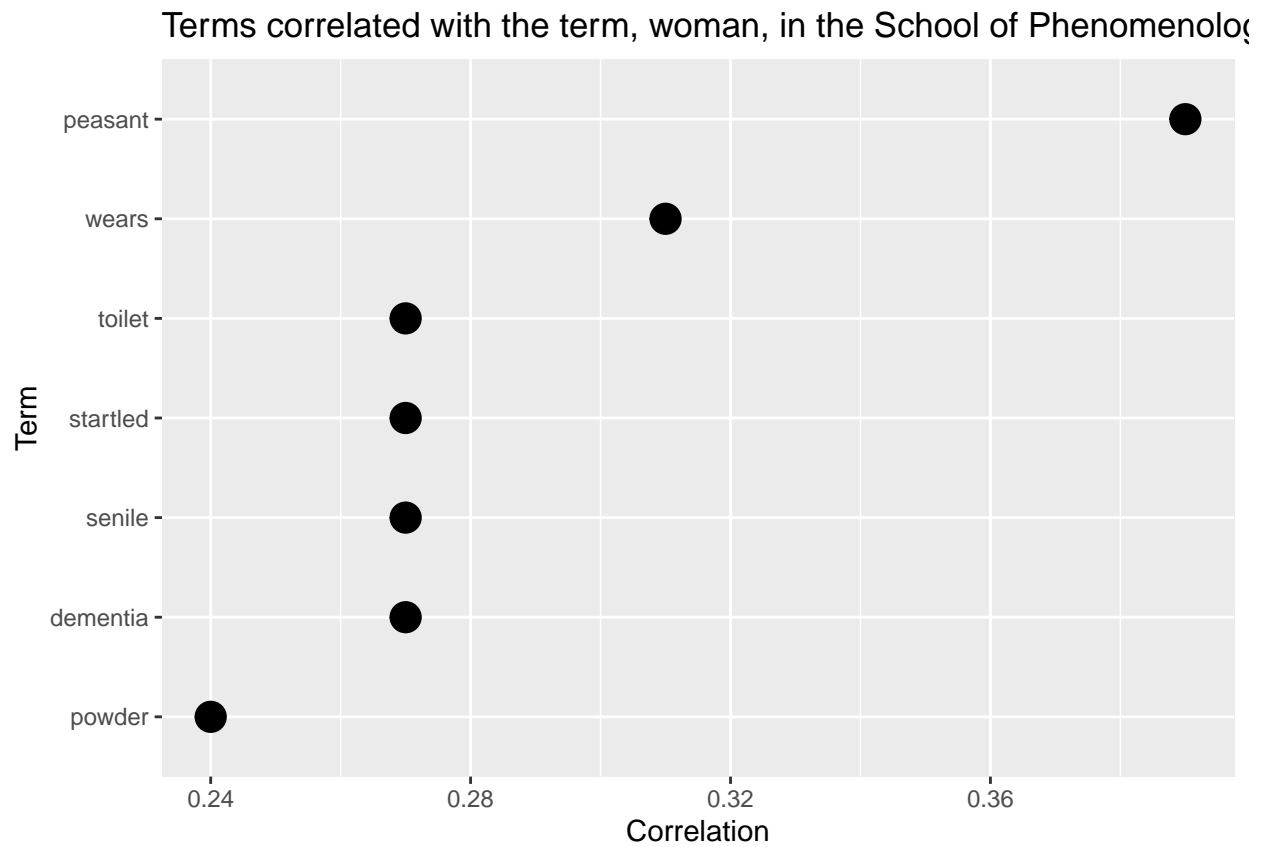


RMD file.

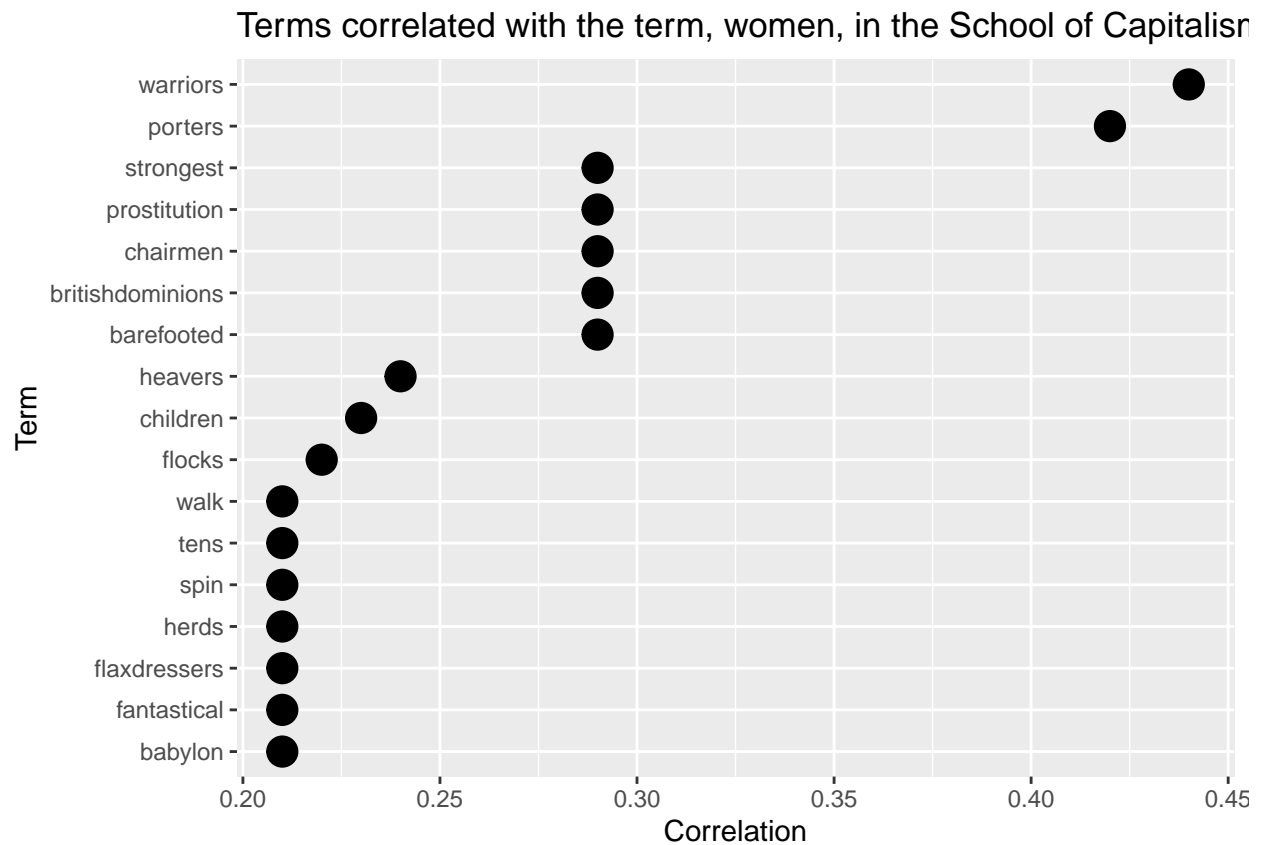
In the school of rationalism, when the text discussed “women”, then other terms associated were like accomplishing, courageous which is quite surprising. It was quite interesting to see good connotations being associated with women and that too at high correlations (e.g. accomplishing having a correlation of more than 0.35).



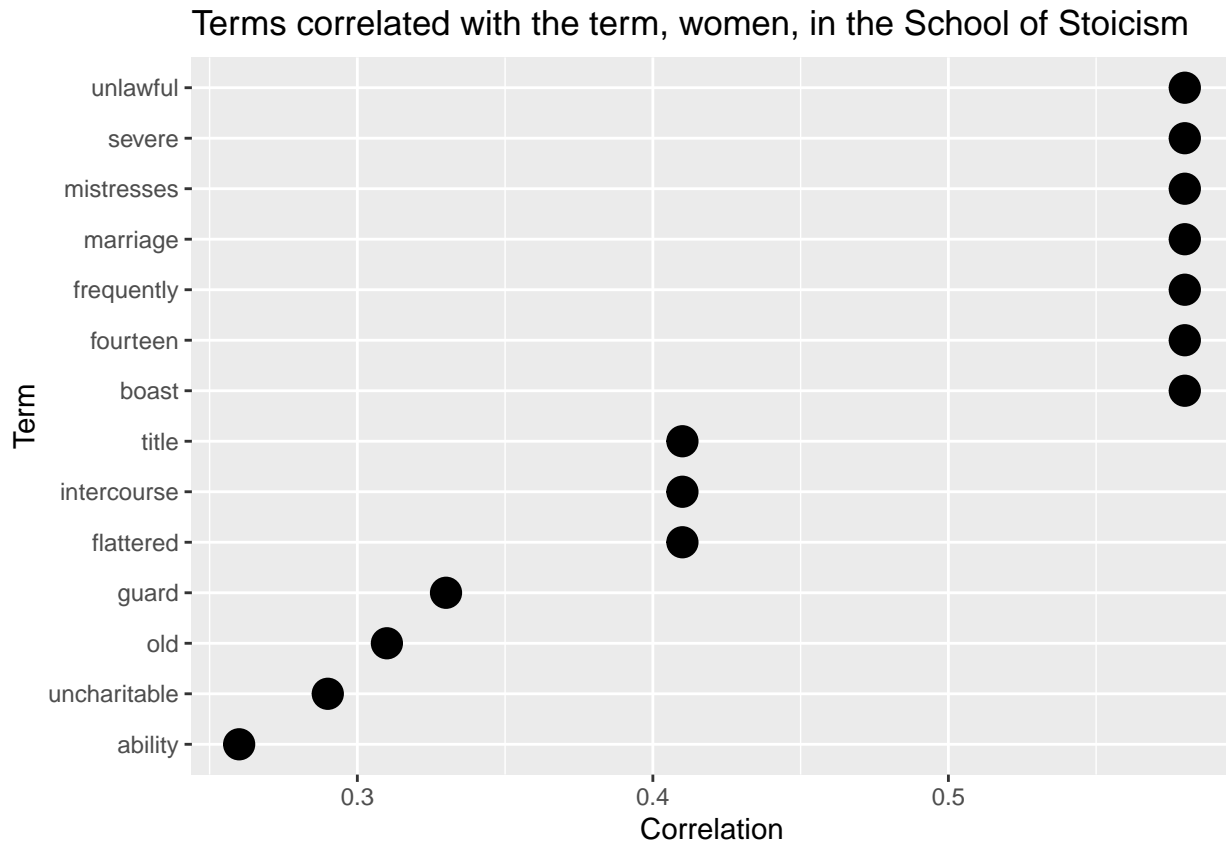
In the school of empiricism, the words associated with female are viparious (definition is producing living young), begets (definition is reproduction), procreation, sustenance (definition is nourishment), etc. So, females here were associated with conceiving a child and nurturing them. This gives a broader context as to the role of women and the opinion of them from the texts in the school of empiricism.



In the school of phenomenology, the words associated with female are peasant, toilet, senile. The highest correlation was to do with peasant. Again, compared to the other schools, this association is quite different with women.



The strongest associations here with the term, woman, has to be warriors and porters. This is followed by the terms strongest, prostitution, chairmen, heavers (definition is a person employed to carry luggage or other loads), children. This gives a broader context from texts as here it seems the texts discussed as to what a woman did to earn a living.



The strongest associations with the term, women, has to be mistresses and marriage. This is followed by the terms intercourse, uncharitable, old. This gives a broader context as to how women were perceived as objects in the texts from the school of stoicism.

To conclude this analysis, different schools looked at females in different ways. Some schools like the school of rationalism associated with women with terms like courageous and accomplishing. Some schools like empiricism looked at women as people who took care of children and conceived them. So, different schools saw females in different ways.

In conclusion, the role of females back then was generally to take care of their family, their marriage. The issues or concerns back then were slavery, oppression, rape, being seen and treated as objects, class inequality, etc. The feminism movement back then was access to basic human rights, access to education, was regarding suffrage, etc. Even though, with time, the topics of the movement have changed, the core of it remains the same - to give them the same opportunities as everyone else.