

MA615 Text Mining Task II Report

Yifeng He

2021/12/07

Task I Book Choice

The book I chose is ‘The Origin of The Family, Private Property and the State’ by Friedrich Engels.

TASK II

In this task, I will first use three methods to analyze ‘The Origin of The Family, Private Property and the State’. The sentiment analysis includes “AFINN”, “BING”, “NRC”. For the extra credit, I will use “Loughran-McDonald” for another sentiment analysis.

The content for this book is about early human history and reveal the destruction of old social rules; the development of new society based on private property. Engels first discusses the development of family, from blood-related family to monogamy family. Then he talks about the foundation of Germany and Rome; next he states specialization and social class and the start and end of countries. Engels discussed marriage, slaves, social progress, and human rights. As a result, more positive words appear in the book than negative words.

As shown in the below code chunks and graphs, more positive words appear in the book than negative ones. However, in some chapters, Engels talks about the dissemble of nations and the situation of slaves. These result in the increase of negative words and the downward bars in the sentiment analysis.

The code I use refer to Jin Yuli’s code. But I made some major changes to his code so that the graphs are visually adaptable to my book. The code is also simplified.

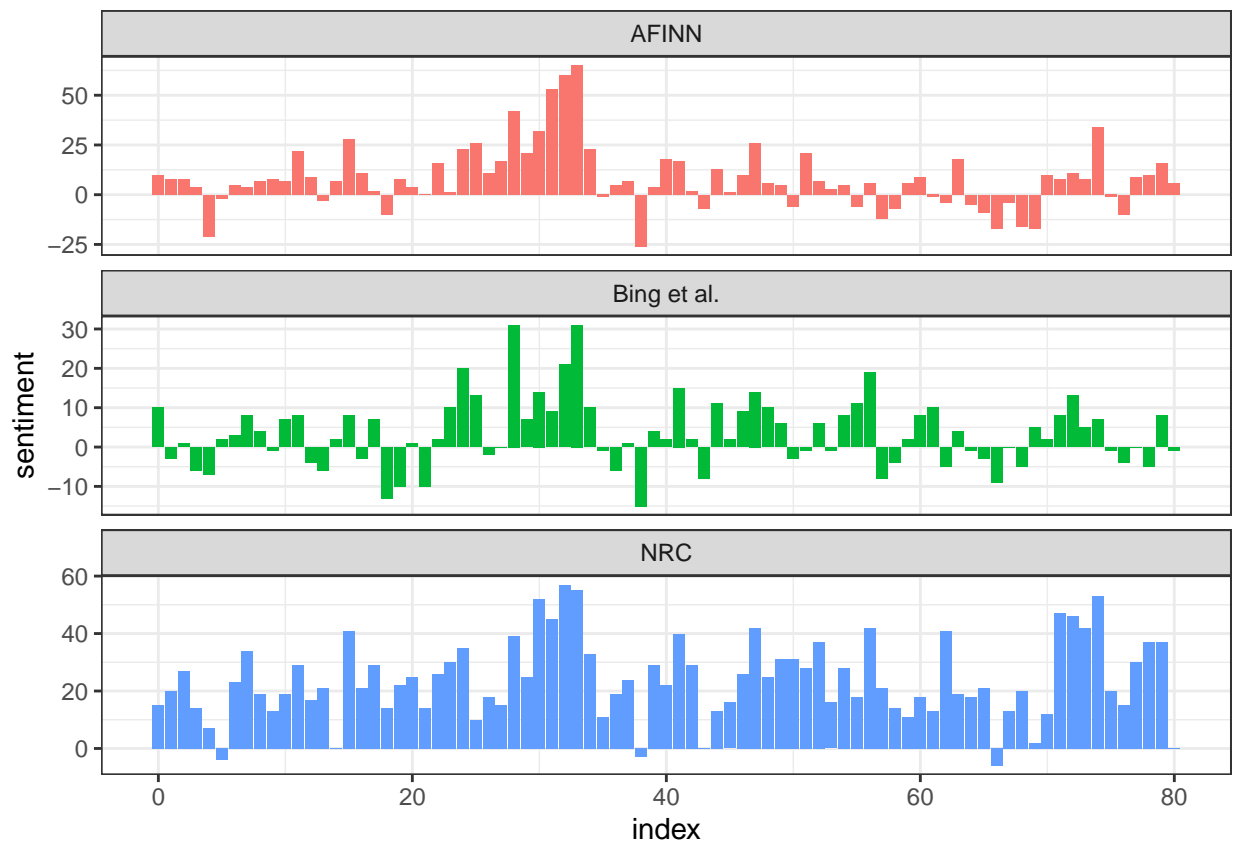


Figure 1: sentiment plot

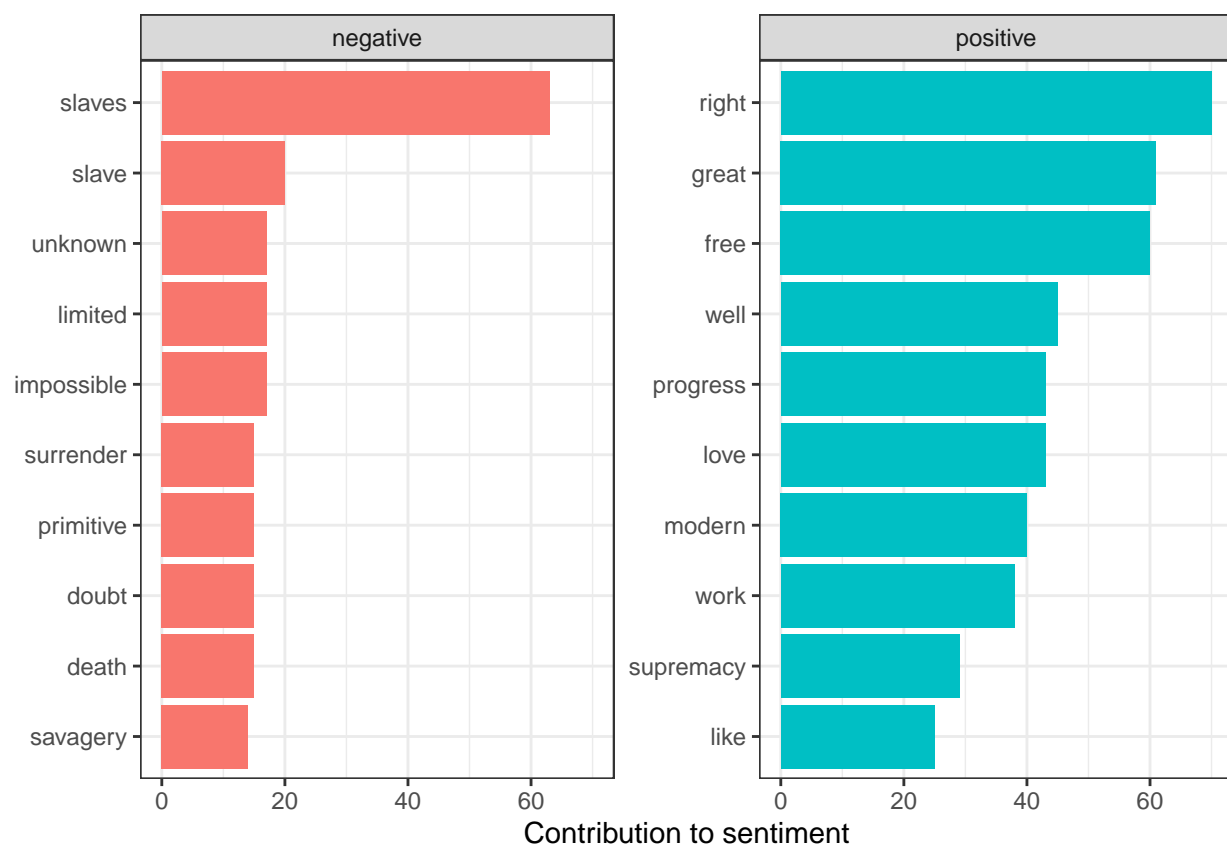


Figure 2: nega-posi words count



Figure 3: Word Cloud Graph



Figure 4: sentiment word cloud

This is the part for extra credit:

