NRC Bible Sentiment Analysis
NLP Analysis of bible books and chapters

This project analyses the bible (*World English Bible - WEB*) and gives different insights using methods from NLP. The problem to solve is to give information about the tendencies of the sentiments covered in chapters and books without having to read the whole bible.

Data & Cleaning

The dataset (Bible Corpus) includes many different versions of the bible. For this analysis the *World English Bible* was chosen. It consists of 31,102 rows and 5 columns. The columns have been renamed for better understanding. They consist of *id, book, chapter, verse* and *text*. As the goal was to analyse the complete chapters and not their single verses, the text has been grouped by books and chapters to concatenate the single verses in the text column. After that, contractions of the English language (e.g.: won't → will not) have been fixed, special characters have been removed and the text was converted to lowercase, before creating a tokenized data frame to get insights into word frequencies.

Exploratory Analysis

The most frequently used words of the bible are the following:

To give more significance to the single words TF-IDF was used:

NRC Sentiment Analysis

By using the [NRC Word-Emotion Association Lexicon](#) the sentiment of the bible as a whole was analysed.
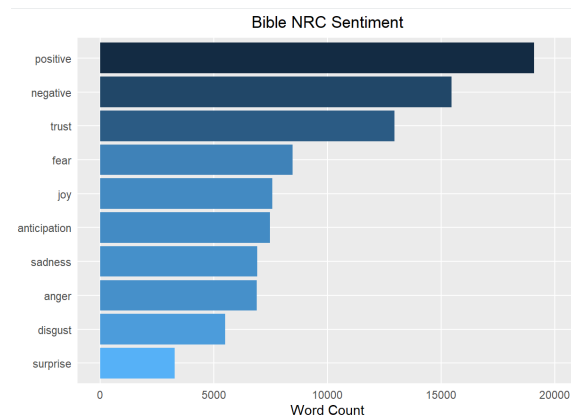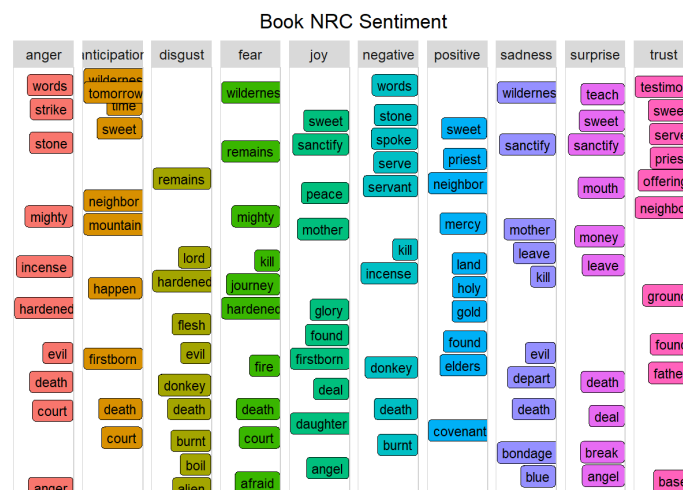

This graph shows that the positively and negatively associated words are fairly balanced. To get insights into the single chapters, a mood ring was created that shows the mood between the first 30 chapters.
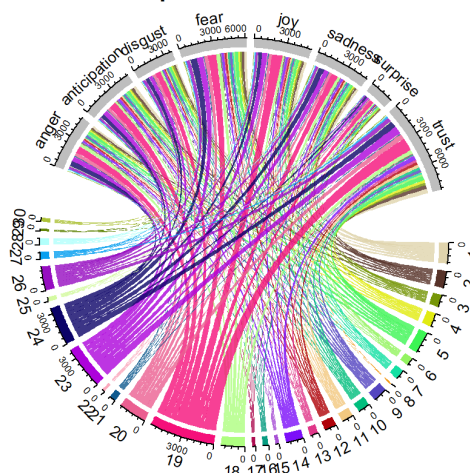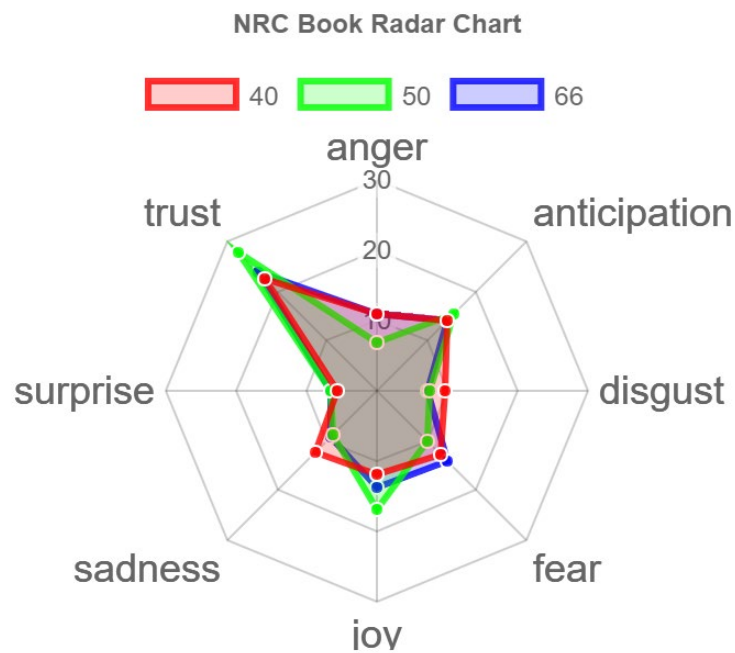

To zoom in even more, a facet grid of a single book can be plotted. In this instance the second book is being displayed:

And last but not least, radar charts have been implemented to make it easy for the books of your choice (*40, 50, 66* in this case) to be compared:

NRC Book Radar Chart

These graphs and analysis give the user the opportunity to explore the sentiments presented in the bible and its individual chapters. The gathered information shows that the main sentiment being conveyed is trust, fear and joy.

The link to the Github account can be found here.