Exploratory Tools



Word Counter

- https://databasic.io/en/wordcounter/
- A user-friendly basic word counting tool
- It allows you to count words, bigrams, and trigrams in plain text files and to download spreadsheets with your results
- The max file upload is 10MB
- Can be run with and without **stopwords**



Word Counter Examples

TOP WORDS ⊕		
Word	Frequency	
could	356	
one	314	
would	308	
us	292	
little	269	
time	247	
life	218	
father	210	

Shows the top words in the text.

Stopwords aren't removed for the bigrams and trigrams because they need context.

BIGRAMS ①		
bigram [©]	Frequency	
of the	885	
in the	726	
i was	497	
it was	371	
to the	319	
on the	309	
and the	296	
i had	272	

TRIGRAMS ①		
trigram [©]	Frequency	
i did not	89	
that i was	61	
i could not	47	
i do not	45	
it was a	40	
and i was	40	
there was no	38	
it was not	35	

It is interesting how many of the trigrams are negations!



Word Trees

- https://www.jasondavies.com/wordtree/
- A word tree depicts multiple parallel sequences of words
- This is a good way to see patterns in word usage, based on words that appear before and after a term or terms of interest.
- There are some restrictions in size: fewer than 1 million words should work



Word Tree Examples

Reflects the focus of the book as on the Jewish community, while Gentiles are more likely neighbors.

Even though written by a woman, the book references Jewish and Gentile boys more than girls.

The punctuation following Gentile suggests it is often the end of the sentence. Jewish

Shift-click to make that word the root

word tree

ewish

might accept; and th

children

boy

, your heart heavy al
. The church door ge
for one day, I thought, he would k
was allowed to grow up without at
that ever was born, but he hated to
across the aisle trod on my foot to
that feeling. The

child nad to answer que

But little it matters now what

boys who were kidnapped by the Czar's
perused their Hebrew books in per
recruit was likely to be treated with se
dreaded, indeed, brutality and
neighbors . They were always resp
and one for the Gentilies
and one for the Gentilies

neighbors They were always resp and one for the Gentile household where the first generat community thiouties challenged without its scholars, claimed kinship with

street, and received alms from Jewish people, they cr people, they crossed themselves and mumbled Christ families lived scattered, by special permission of the settlements, searching for his family; hiding the scars rites, he was subjected to the severest punishment. M laws of daily life while in the service. A soldier often h law. When he returned home, at the end of his term of little girl in Polotzk. "It is a false world," you heard, a candidates, of course; a nine-year-old Jewish child ha practices. The public school interrupted the boy's sac wife; for, of course, every girl hoped to be a wife. A gi man and woman had a part in the fulfilment of the ar cemetery, through miles of desolate country, no livin society for years at a time, to live among pork-eaters, restrictions. Finally he reached the position of assista life in the Pale offered no opening to any other mode and Russian; solemn songs, and jolly songs, and song scholarship, as we have seen, was confined to a know enterprise. In the vocabulary of the more intelligent practice. That which had been taught them as the tru world with the familiar fear. The wholesale expulsion persecution. Passover was celebrated in tears that yes

word tree

gentile

Shift-click to make that word the root.

of his own free will. And David knows if he failed in Polotzk, could go els she kept a poodle, and she had no the women who have been to the rexisting to harass the Jews, while a though I can still prove that I was it was the symbol of tyranny; on the

neighbors

, they had as a differ

The Gentiles do as they like with u
" The third time, when Vanka spat
Piety alone was enough to make th

house there was what they ca
. I was afraid of the cro
boys ran away soon, or ran awa
. The way I dodge a snowl

family which was very friendly. There was a families, till they were old enough to enter t comrades. He might resort to all sorts of tri what he would call a sin against a fellow Jeckept store side by side, the Gentile could co could content himself with smaller profits. Was hardly expected to understand. But tha courts. The Gentiles were false, judges and

Gentile

Northeastern University
NULab for Texts, Maps, and Networks

Word Tree: Reverse Trees

When words are commonly followed by punctuation, it is worth reversing the tree to see the words that often precede it. To do this, click "reverse tree" next to the search bar.



