

[introduction](#) [samples](#) [compare](#) [non-English](#) [related sites](#) [get data](#)

1. In any materials that you develop with the data, end users cannot see the exact frequency of a word (e.g. it occurs 823 times in the corpus) or the exact rank order (e.g. it is the 2,920th most common word). But you can group words into frequency bands (e.g. the word is in the band of words from rank order 3000-5000), although the number of frequency bands must be limited to 20 or less.
2. In no case can the word frequency data be distributed beyond the university or company that purchased the data. A small, unique change has been made to each dataset that is sold, and this can serve as a "fingerprint" to identify you as the unique source of the data. Automated queries run every night to search the web (including the "dark web") to find copies of the data.
3. If you are in violation of the licensing terms, you will be contacted by our IP (Intellectual Property) lawyers (Sumsion Law, in Provo, Utah, USA), and will be liable for damages for lost sales, depending on how many people have downloaded the data that you have illegally made available to others.
4. If you have questions about possible (re-)use of the data, contact us *before* you purchase the data.