**Purchase agreement**

between Mark Davies (seller) and

| Buyer (your name) | Brian Putnam |
| --- | --- |
| Name of company | N/A |
| URL / website for company | https://brian-putnam.com/fractograms/ |
| ID from the page you clicked on to see this document | [220408] |

For the sum of $ **295**, Mark Davies sells to the buyer listed above the following datasets from the Corpus of Contemporary American English (see https://www.wordfrequency.info/samples.asp).

* Top 60,000 lemmas with frequency in eight main genres, as well as nearly 100 sub-genres
* Frequency of word forms for each of these 60,000 lemmas
* Top 219,000 word forms (which occur at least 20 times in 5 different texts)

**Terms of license:**

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 should be limited to 20 or less.

2. In no case can the Data be distributed beyond the company listed above. 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.

| It will be used to generate puzzles for a simple word game where the user tries to reassemble a 5x5 word square (see link above). The game already works, but I need a better (and bigger) word list. I’ll produce a sub-list of common 5-letter words, and maybe other lengths too for future variations, and feed that to my puzzle-generating program.  Once I’ve created my 5-letter (or n-letter) lists I’d love to publicly redistribute them (with attribution and without frequency or rank data) if you’d be willing to allow that, but it’s not a dealbreaker for me if you’d prefer not to allow that. (If I’ve read the TOS correctly, by default that's not allowed without an exemption).  Note that the URL for the game may change in the future.  Happy to answer more questions if you’ve got them, just let me know! |
| --- |