ДЗ №1

1. Объем SpecC:

количество словоупотреблений - 5 431 157 слов

тематика / жанр корпуса - здоровье и медицина

источник корпуса - НКРЯ

1. Объем RefC

количество словоупотреблений - 122 411 245 слов

источник - НКРЯ

1. Метод 2: weirdness
2. Таблица 1.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| wi | Тип | CountSpecC | CountRefC | LogLikelihood | Ранг | Вес 2 | Ранг |
| болезнь | спец | 1611 | 4589 | 3471.78 | 2 | 7.91 | 2 |
| вирус | спец | 1408 | 169 | 7835.33 | 1 | 187.78 | 1 |
| ужин | общ | 78 | 3213 | 33.79 | 4 | 0.55 | 3 |
| стол | общ | 210 | 25 463 | 1100.60 | 3 | 0.19 | 4 |

1. Комментарий: Специфические слова при обоих методах получили более высокие ранги, чем общеупотребительные.