Exercices 3

google ngrams

- Read Michel, Jean-Baptiste, Yuan Kui Shen, Aviva P. Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, et al. 2010. "Quantitative Analysis of Culture Using Millions of Digitized Books." Science, December. `https://doi.org/10.1126/science.1199644`. (use scholar.google.com to find a pdf copy)

- Go to the ngram viewer `https://books.google.com/ngrams`.

- At the bottom of the page, there is a message "Raw data is available for download here"". Follow the"here" link. download the 1-grams file 'z' for the dataset "English Version 20120701"". Uncompress it (it is in the .gzip format). It is a text file listing words starting with a 'z'.

  Zipfian . . .