學號:P77101032

姓名:林瑋荏

環境: Windows 10下安裝Anaconda3 及python3.7

使用程式語言: python

主程式: poters\_stopword.py 計算文字出現頻率，並將計算結果儲存於cvs檔。

副程式: pubmed\_query.py，從pubmed網站抓取50000個COVID-19關鍵字abstract，並將每100個文件存成1個txt檔。

副程式: show\_plot.py，讀取計算結果cvs檔並加以排序繪圖顯示。

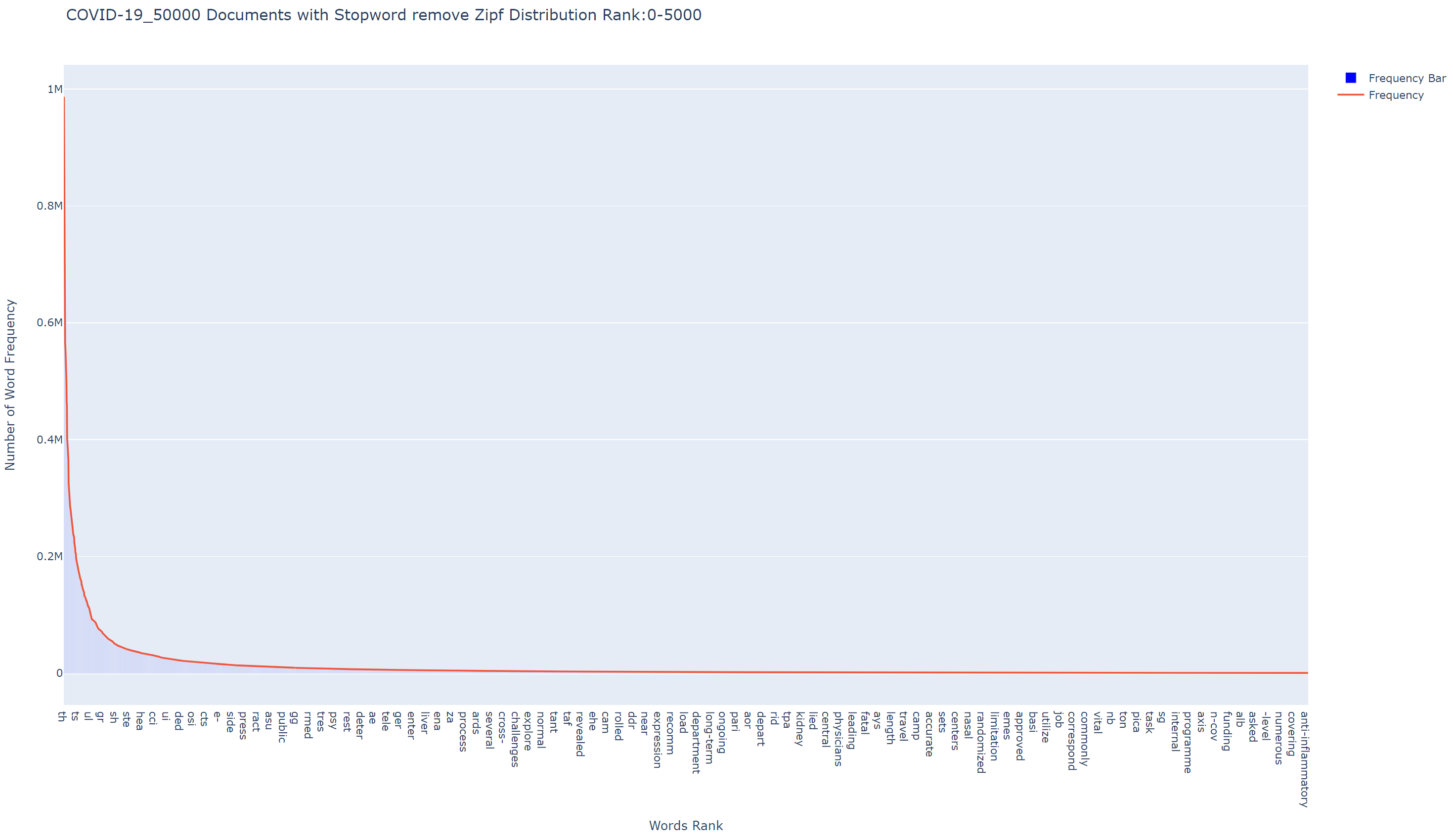
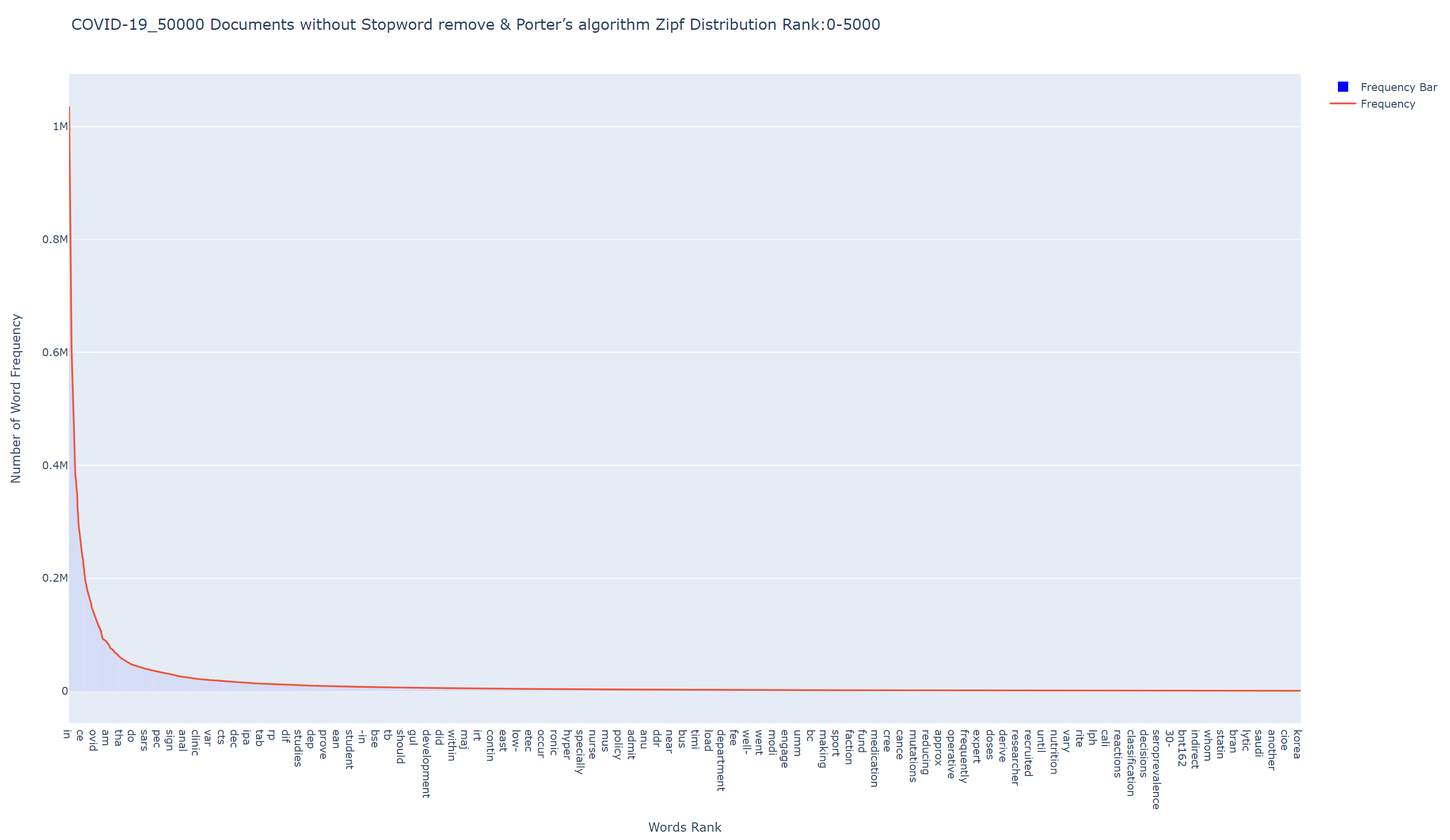
各程式使用library:

PyQt5、sys、os、time、re、nltk、Plotly、pandas

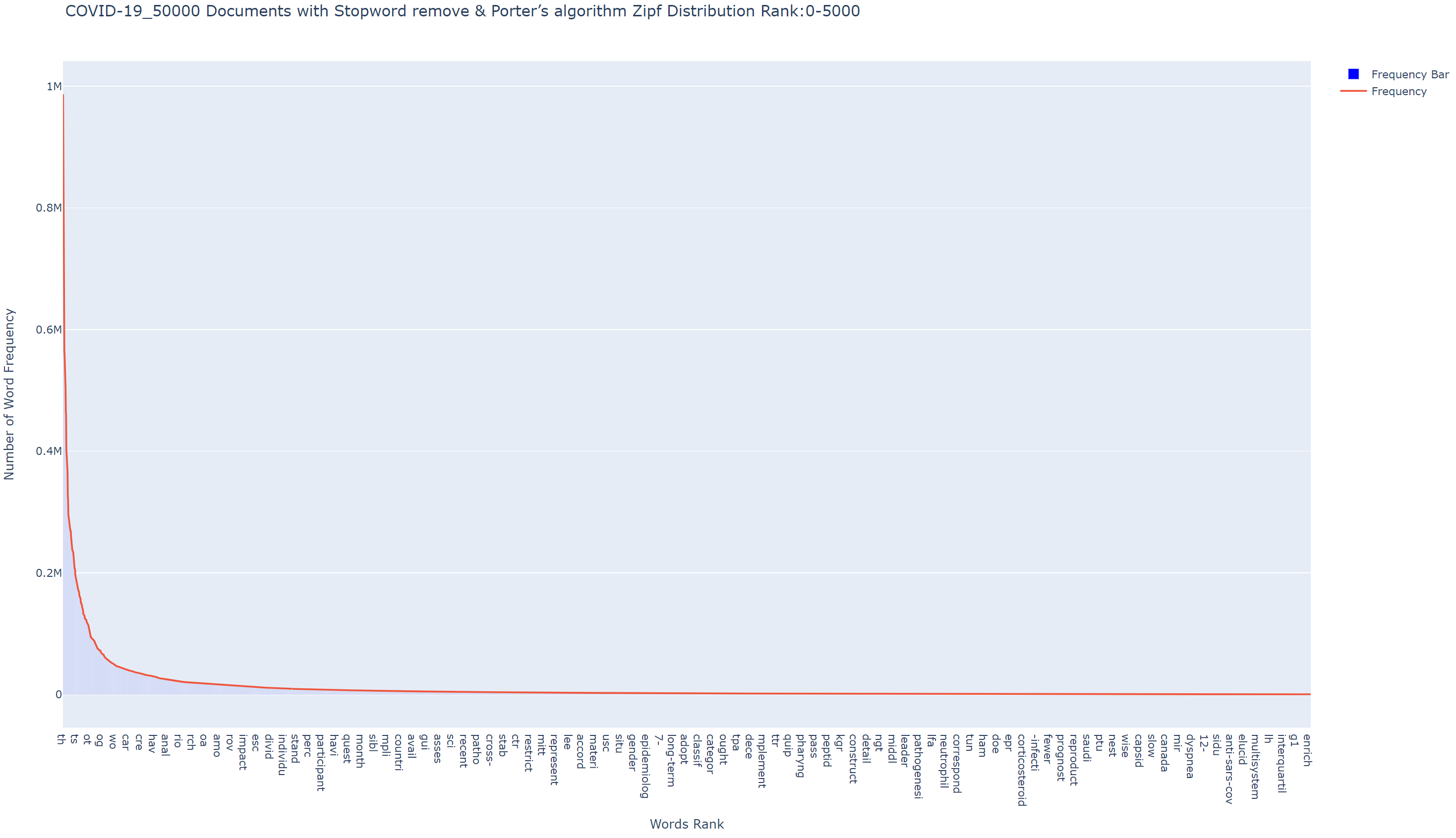
執行結果:



Rank :0 - 5000

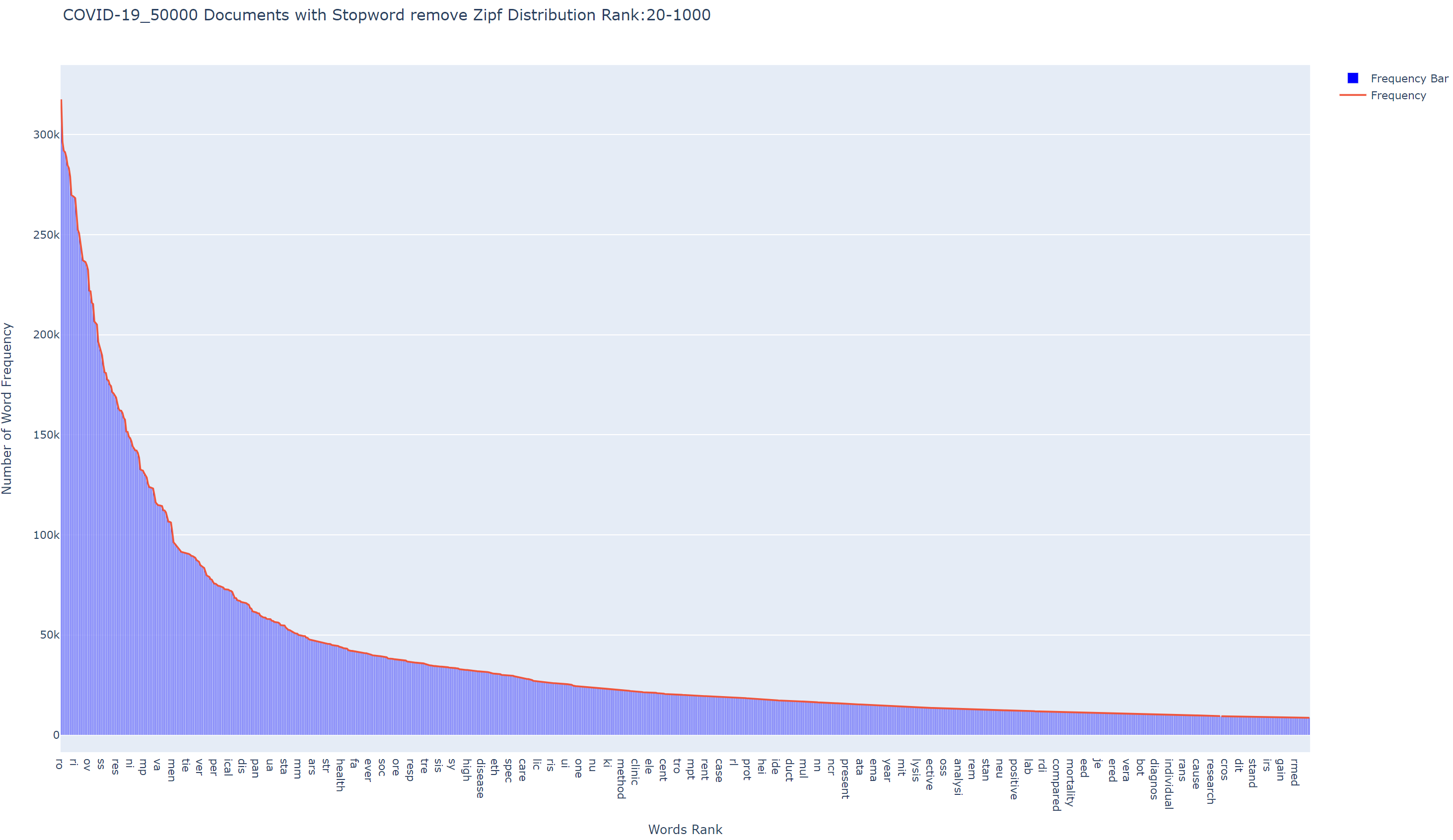
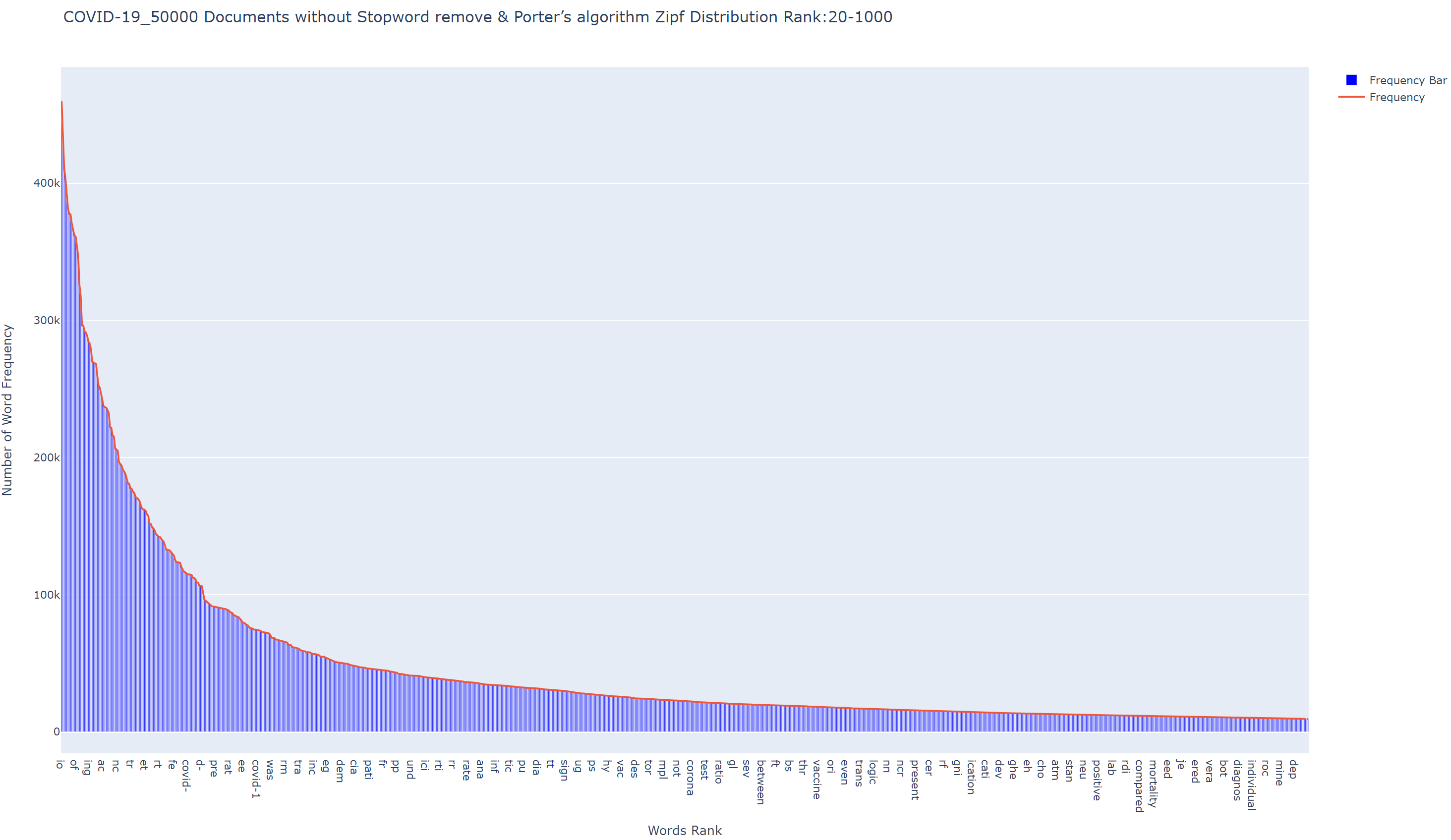


5萬筆不去除Stop word及不使用Porter’s algorithm 5萬筆去除Stop word及不使用Porter’s algorithm

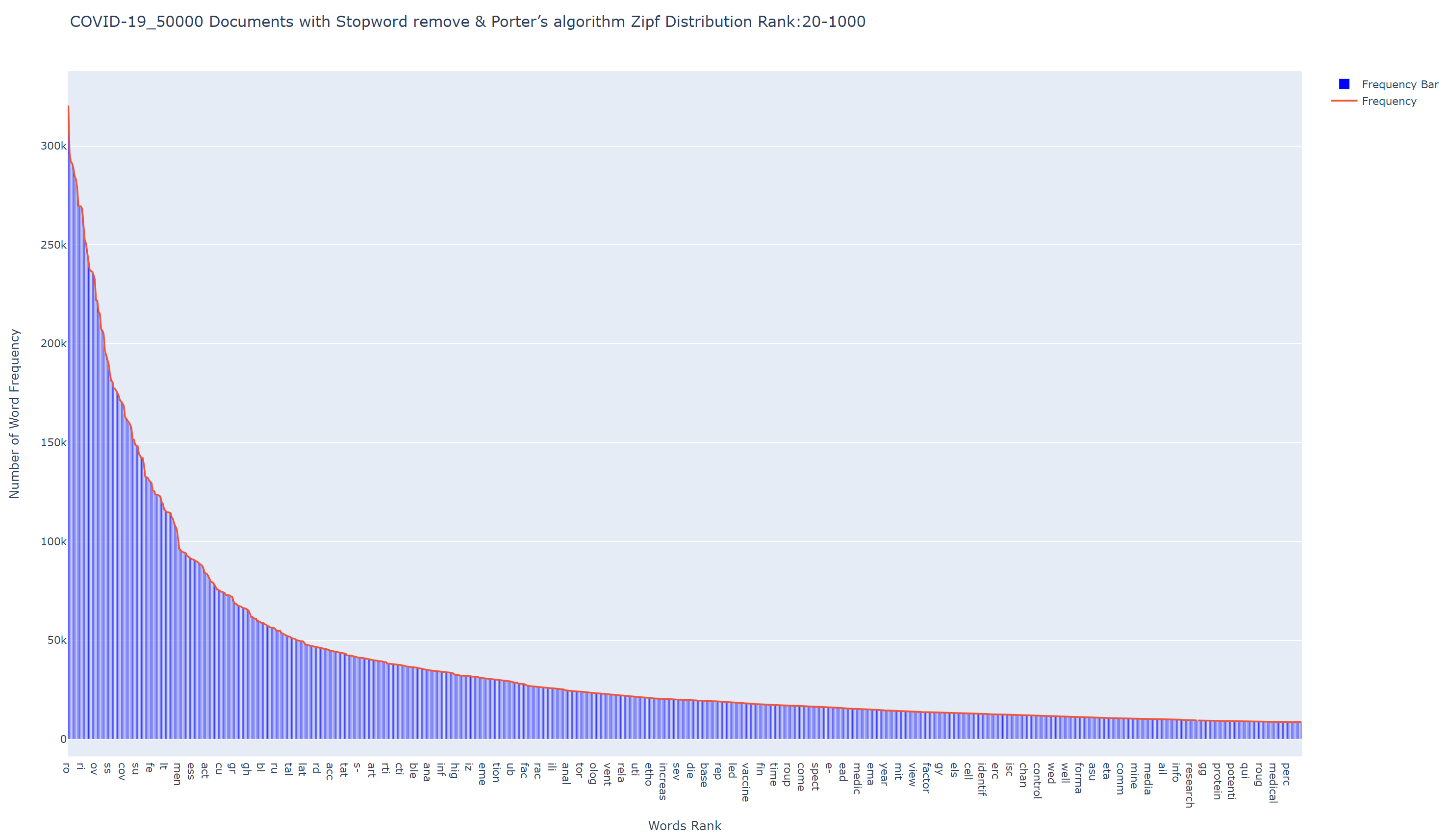


5萬筆去除Stop word並使用Porter’s algorithm

Rank :20 - 1000

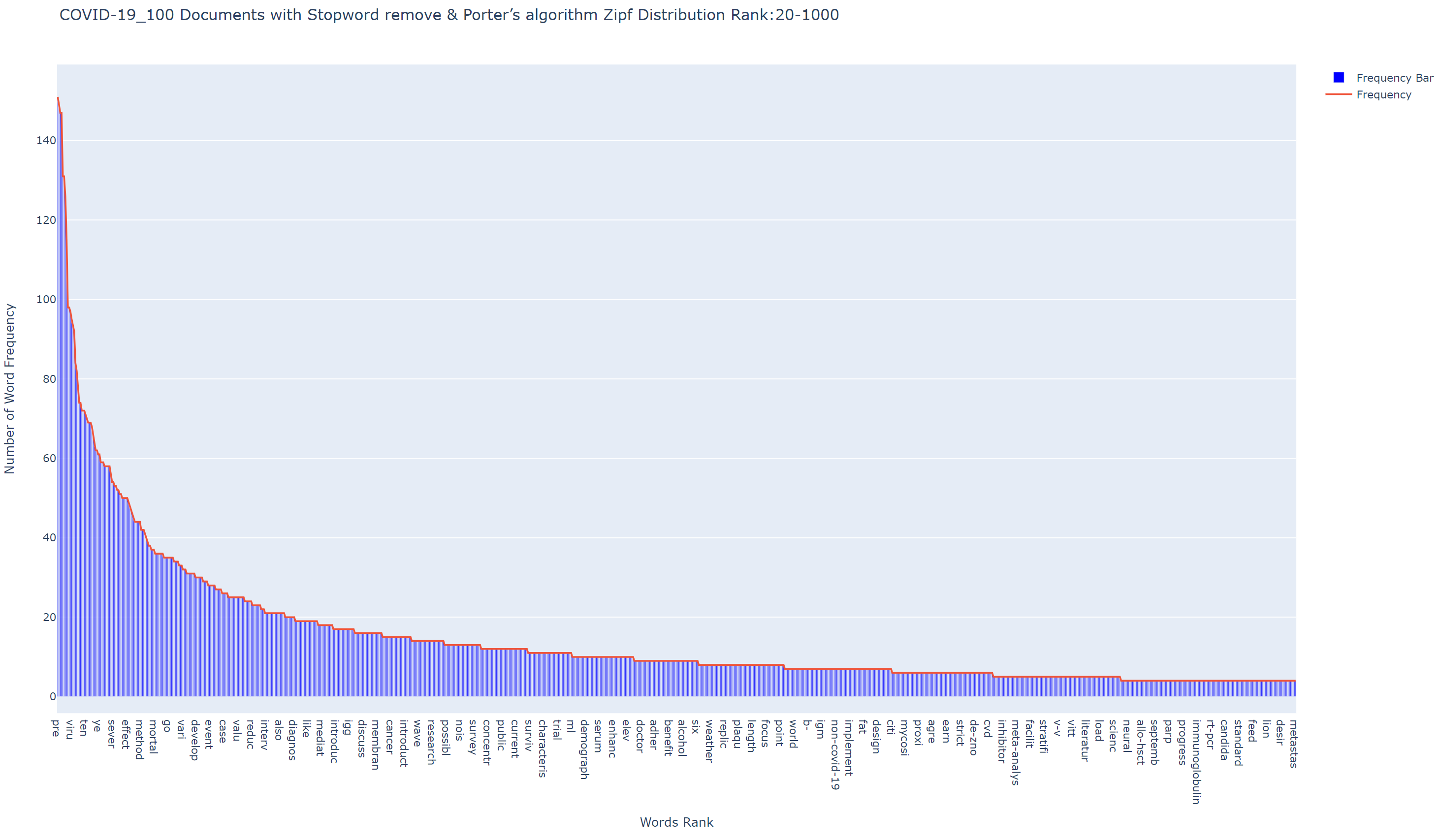
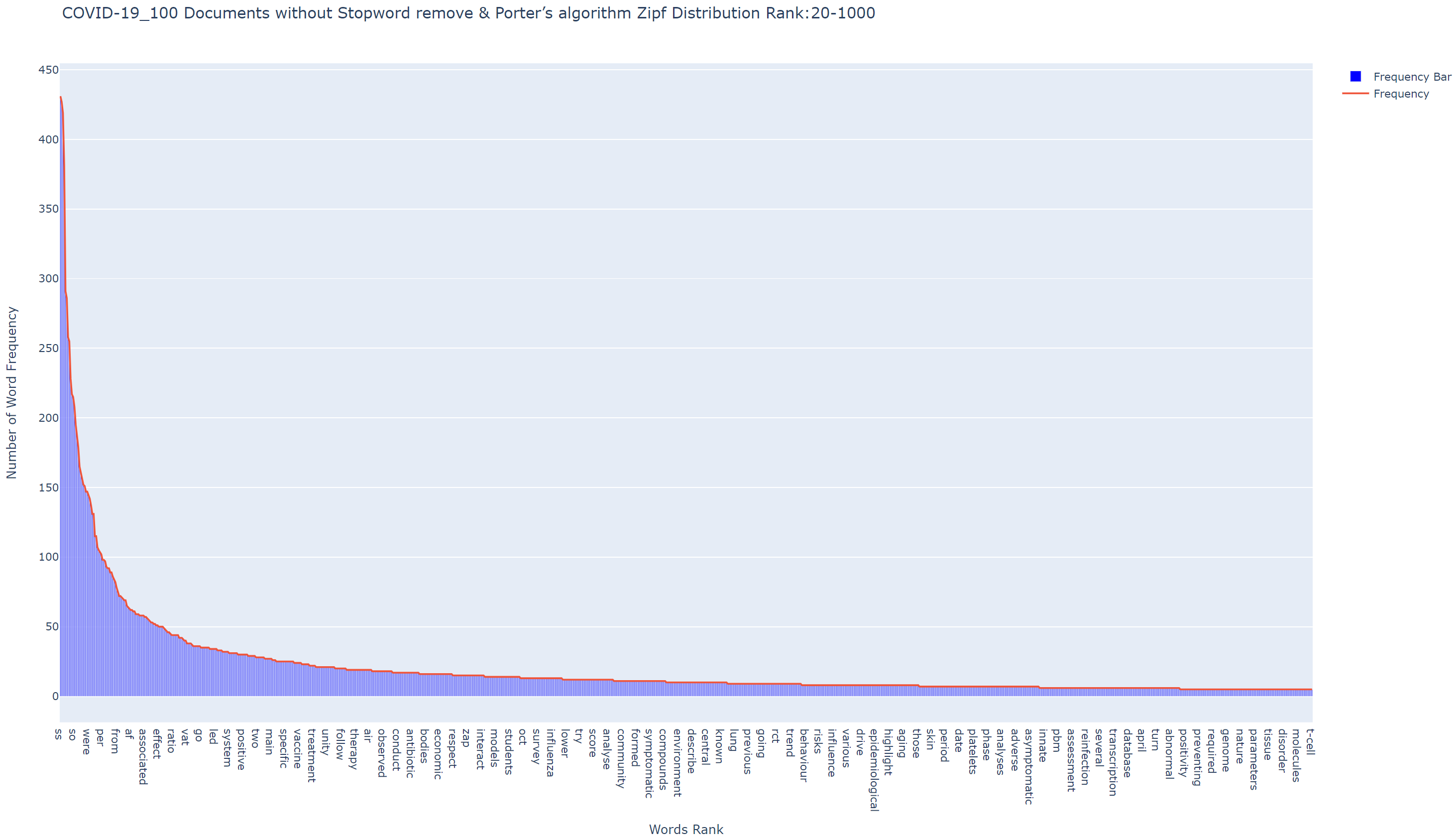


5萬筆不去除Stop word及不使用Porter’s algorithm 5萬筆去除Stop word及不使用Porter’s algorithm



5萬筆去除Stop word並使用Porter’s algorithm

Rank :20 - 1000



100筆不去除Stop word及不使用Porter’s algorithm 100筆去除Stop word並使用Porter’s algorithm