1. 執⾏環境 Mac terminal

2. 程式語⾔ Python 3.8.5

3. 執⾏⽅式

下載套件：nltk

pip install nltk

python3

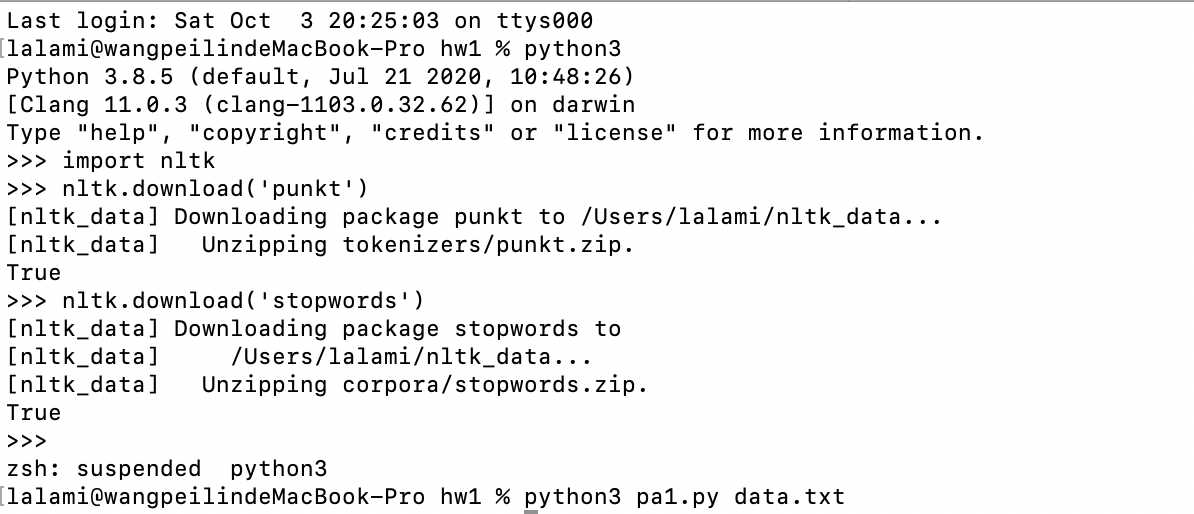
import nltk

nltk.download('punkt')

nltk.download('stopwords')

執行指令：

python3 pa1.py data.txt



結束後會在同個資料夾產⽣result.txt的⽂件

4. 作業處理邏輯說明：

以檔名為參數，讀檔後用replace移除標點符號和句點，用split作 tokenize，用lower()變小寫，再用PorterStemmer function作Stemming，最後把不在stopword（使用nltk.corpus）的結果寫入檔案。