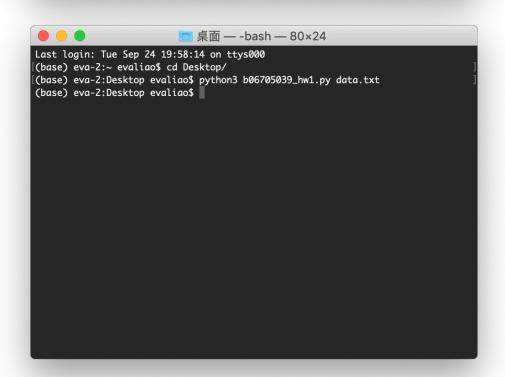
- 1. 執行環境
  Mac terminal
- 2. 程式語言 python 3.7.3
- 3. 執行方式 Install Package nltk

```
n evaliao — python3 - python3 — 80×24
Last login: Tue Sep 24 19:57:03 on ttys000
[(base) eva-2:~ evaliao$ python3
Python 3.7.3 (default, Mar 27 2019, 16:54:48)
[Clang 4.0.1 (tags/RELEASE_401/final)] :: Anaconda, Inc. on darwin
Type "help", "copyright", "credits" or "license" for more information.
[>>> import nltk
>>> nltk.download('punkt')
[nltk_data] Downloading package punkt to /Users/evaliao/nltk_data...
[nltk_data] Package punkt is already up-to-date!
True
[>>> nltk.download('stopwords')
 [nltk_data] Downloading package stopwords to
 [nltk_data]
                   /Users/evaliao/nltk_data...
 [nltk_data]
                Package stopwords is already up-to-date!
True
>>>
```



在terminal上打:

python3 b06705039\_hw1.py 要輸入的文件名.txt 結束後會在同個資料夾產生一個result.txt的文件

## 4. 作業處理邏輯說明

吃進檔名後,讀檔,用nltk的function, word\_tokenize來把資料切開,把有出現在 string.punctuation裡的字元刪除後,用lowercase的function變小寫,再用porterStemmer() function來把每個字變成最原始的樣子,最後把不是stopword的東西存進result.txt裡。