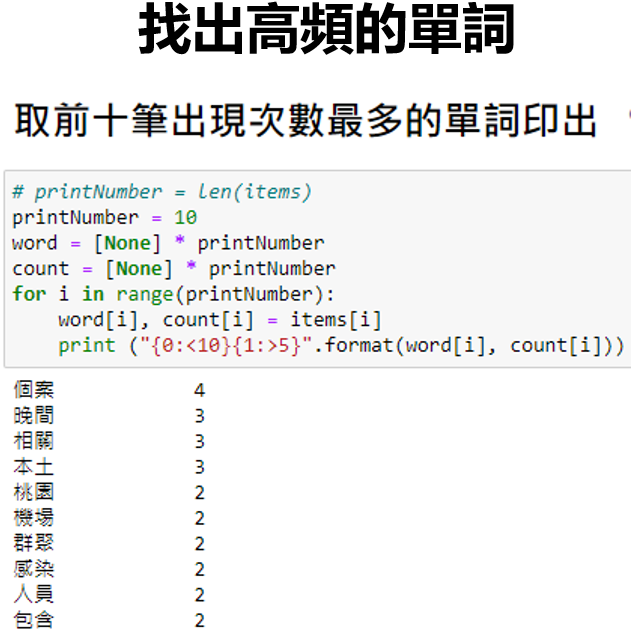
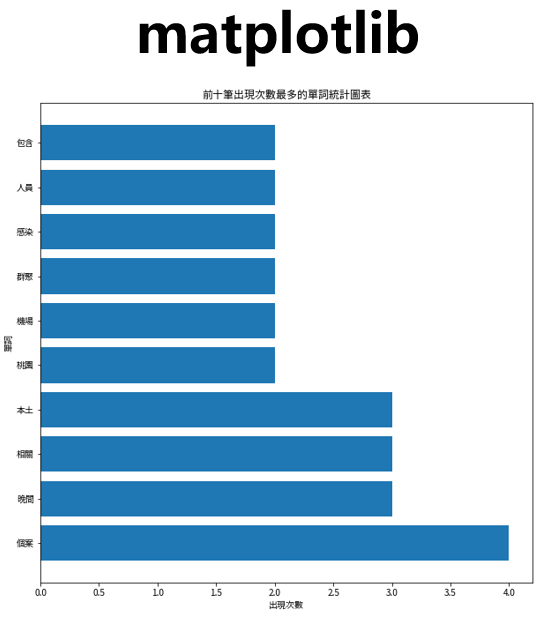
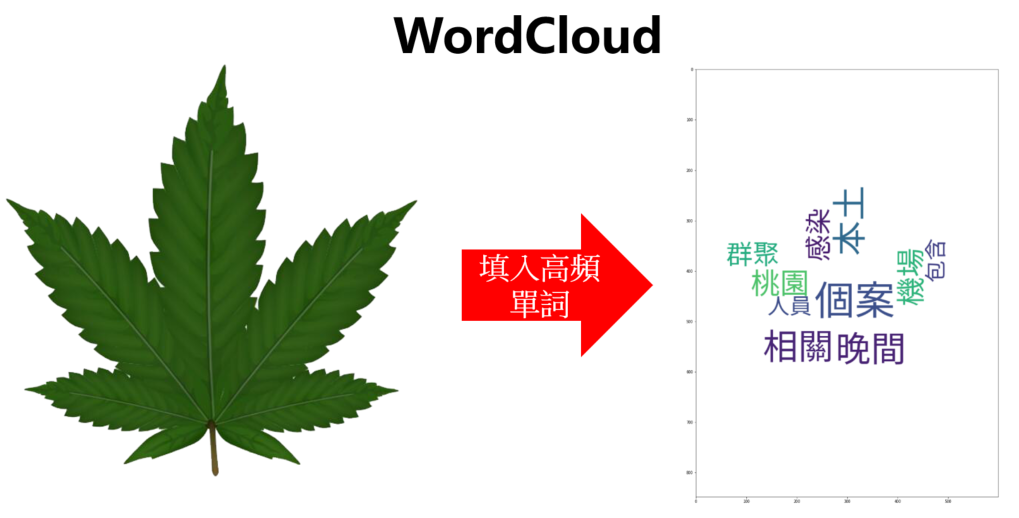
大數據分析期末報告-文檔字詞頻率分析

M11056002 林泓均

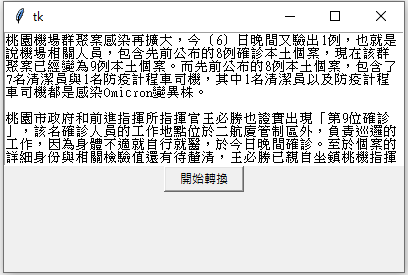
1. 應用套件  
   jieba : 最常使用到的 Python 中文斷詞套件  
   matplotlib : Python 的視覺化套件  
   WordCloud : 統計詞頻後視覺化成文字雲的套件  
   Tkinter: Python中的GUI圖形化介面開發套件
2. 文檔處理  
   可以將新聞或是論文內容、社群平台上的內容放進data.txt中，之後就可以利用程式讀檔進行分析  
   



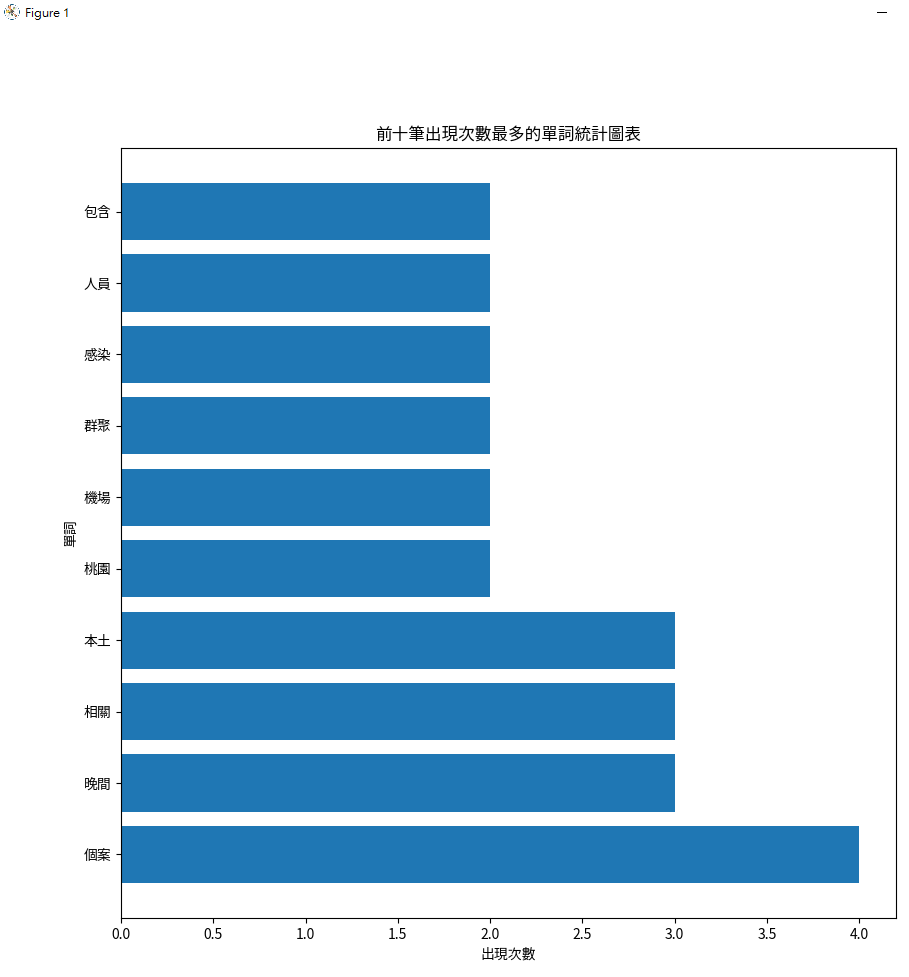


1. MapReduce  
   
2. 視覺化  
     
   

  
最後利用Tkinter 將程式流程轉換成使用者GUI  
以下將說明操作流程  
1.執行程式

2. 貼入要分析的文字後按開始轉換按鈕

3.程式將會產生斷詞後word count的視覺化圖形

4.並將這兩張分析圖自動儲存進資料夾中完成分析

