

網路爬蟲與文字分析

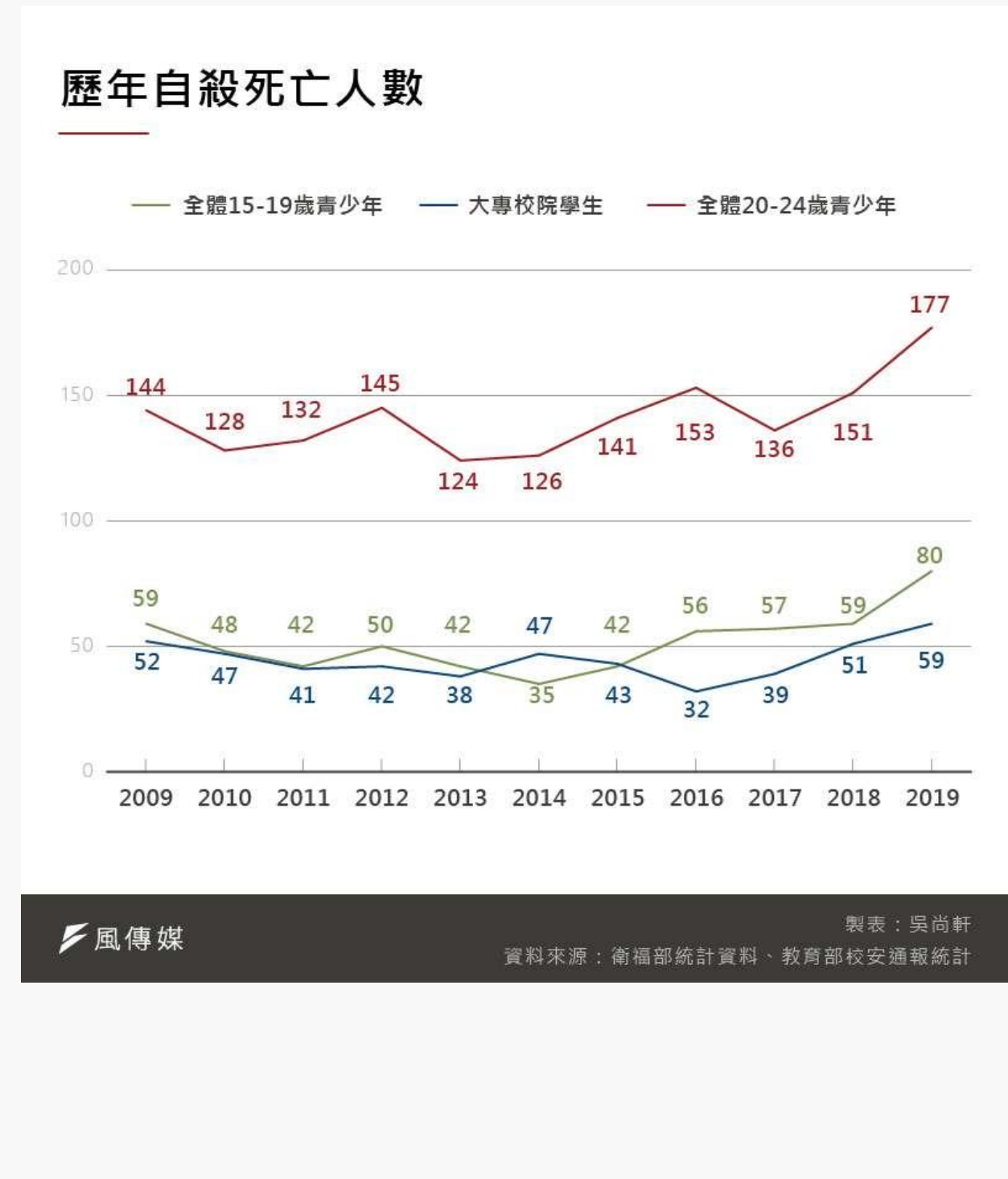
以大學生自殺事件產生之維特效應為例

林芊卉、黃亭瑜、許慈芸、郭思怡
經濟三、經濟四

背景

2020年已有約76名各級學校學生自殺身亡，意圖自殺學生更高達約1350名，相較於2019年，意圖自殺人數年增68%，台灣青年自殺問題已越顯嚴重。

2020年11月中旬，台灣大學五天內連續有三名學生自殺身亡，經媒體爭相報導後，繼台大後，短時間內，其他大專院校發生多起大學生自殺事件，11月份已發生六起大學生自殺案件。許多教育、心理專家分析自殺原因，多因學業、家庭等因素所引起。然而，大眾媒體過於關注報導自殺事件，也可能導致「維特效應」現象。



動機

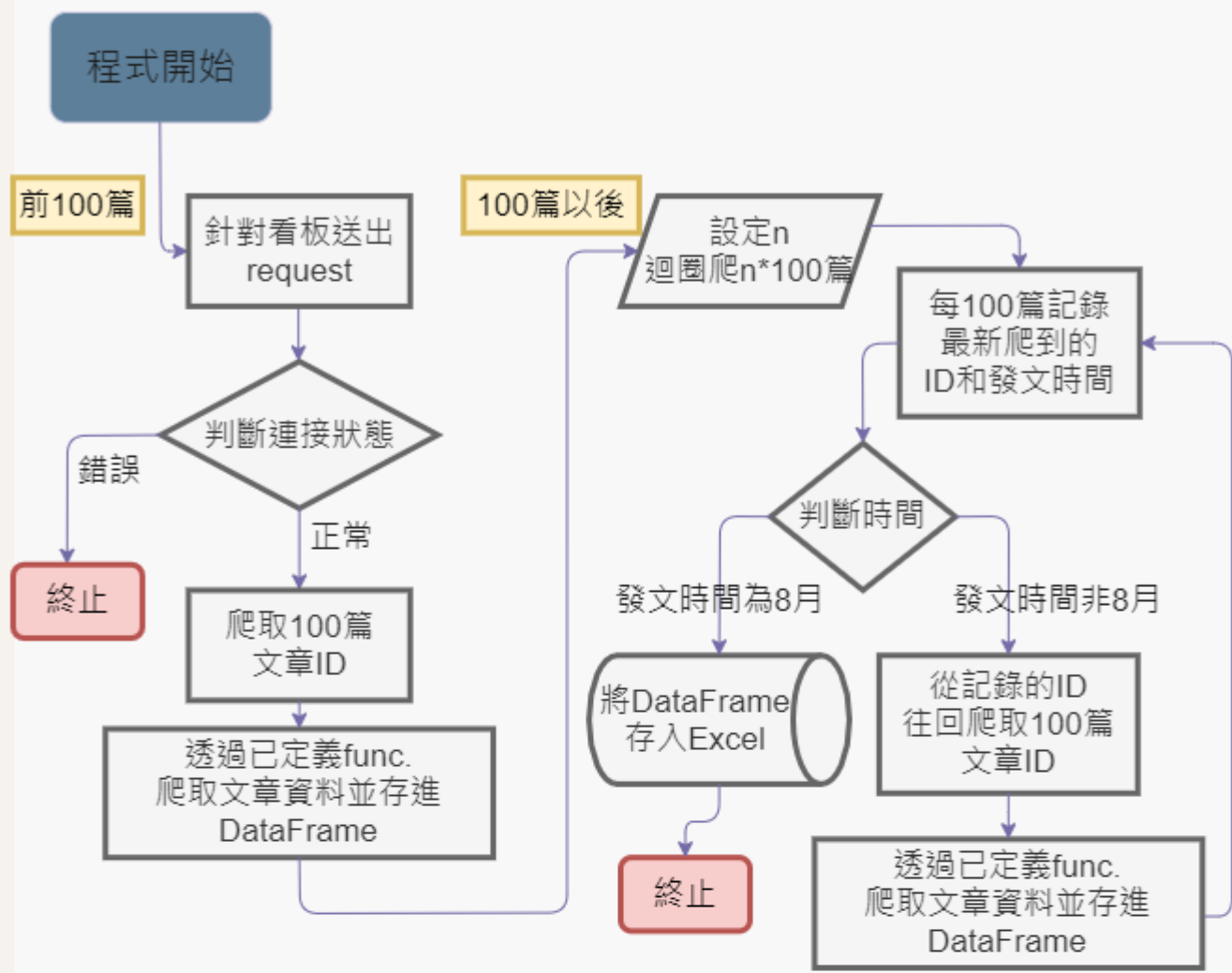
1. 探討台大學生自殺事件，經由媒體大肆報導傳播後，是否使大學生心理受到負面影響，進而引起維特效應。
2. 希望呈現媒體過度關注自殺事件可能對社會造成負面影響，並冀望能約束媒體報導自殺事件時的用詞及敘述內容。
3. 希望藉由本次研究，能使學校心理輔導單位、自殺防治團體建立提前預警機制，防止自殺模仿現象的發生。

方法

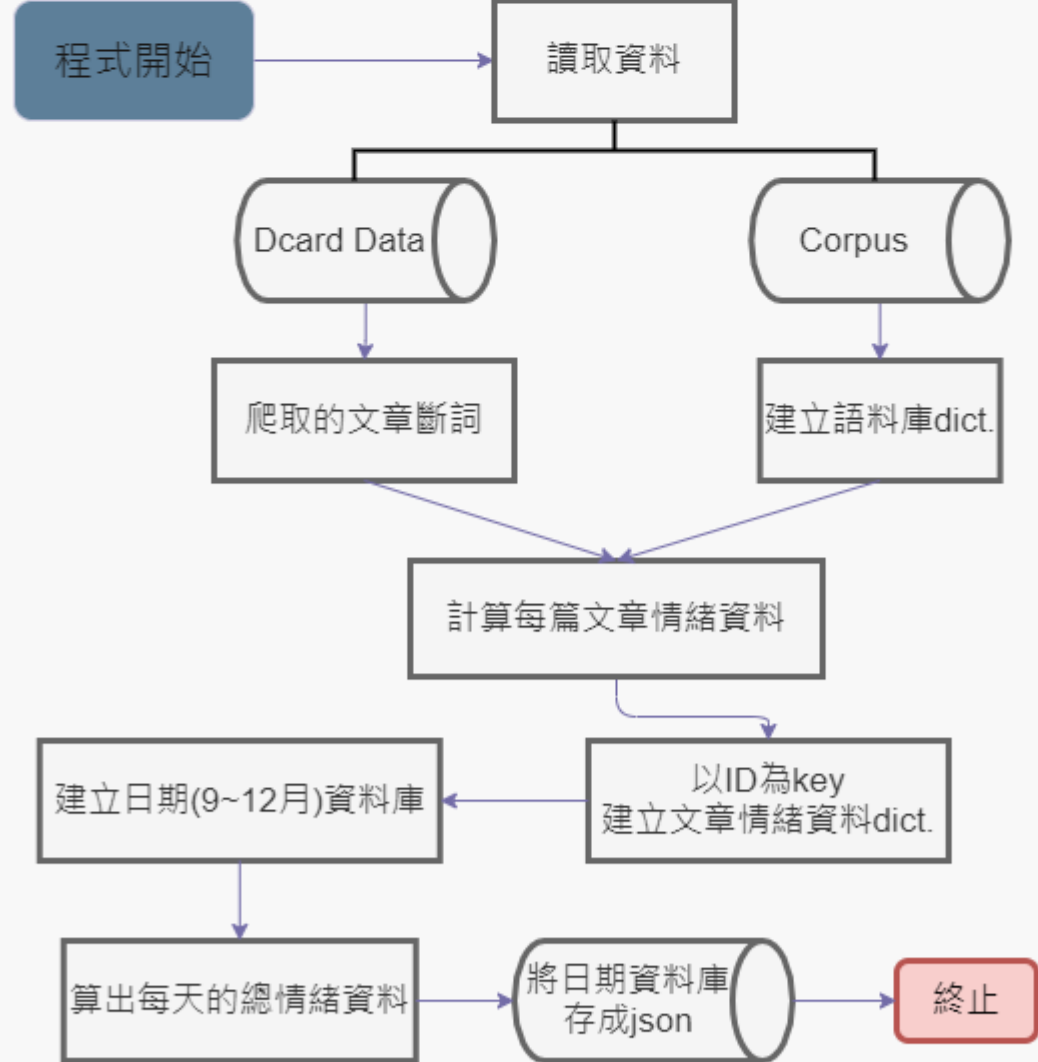
Part1 爬蟲

前置處理 { 設定header
定義透過ID爬取單篇文章寫入DataFrame的Function

- api一次最多爬100篇，所以以100篇為單位。
- 過程使用time.sleep()避免被擋下。
- 程式結束後手動刪除2021年1月及2020年8月的多餘資料。(筆數極少故手動處理)



Part2 資料處理



Part3 GUI

- 輸入計算日期區間(YYYY-MM-DD)
- 選擇欲計算的學校
- 得到 1.區間每天數據 2.區間平均數據 3.區間情緒折線圖

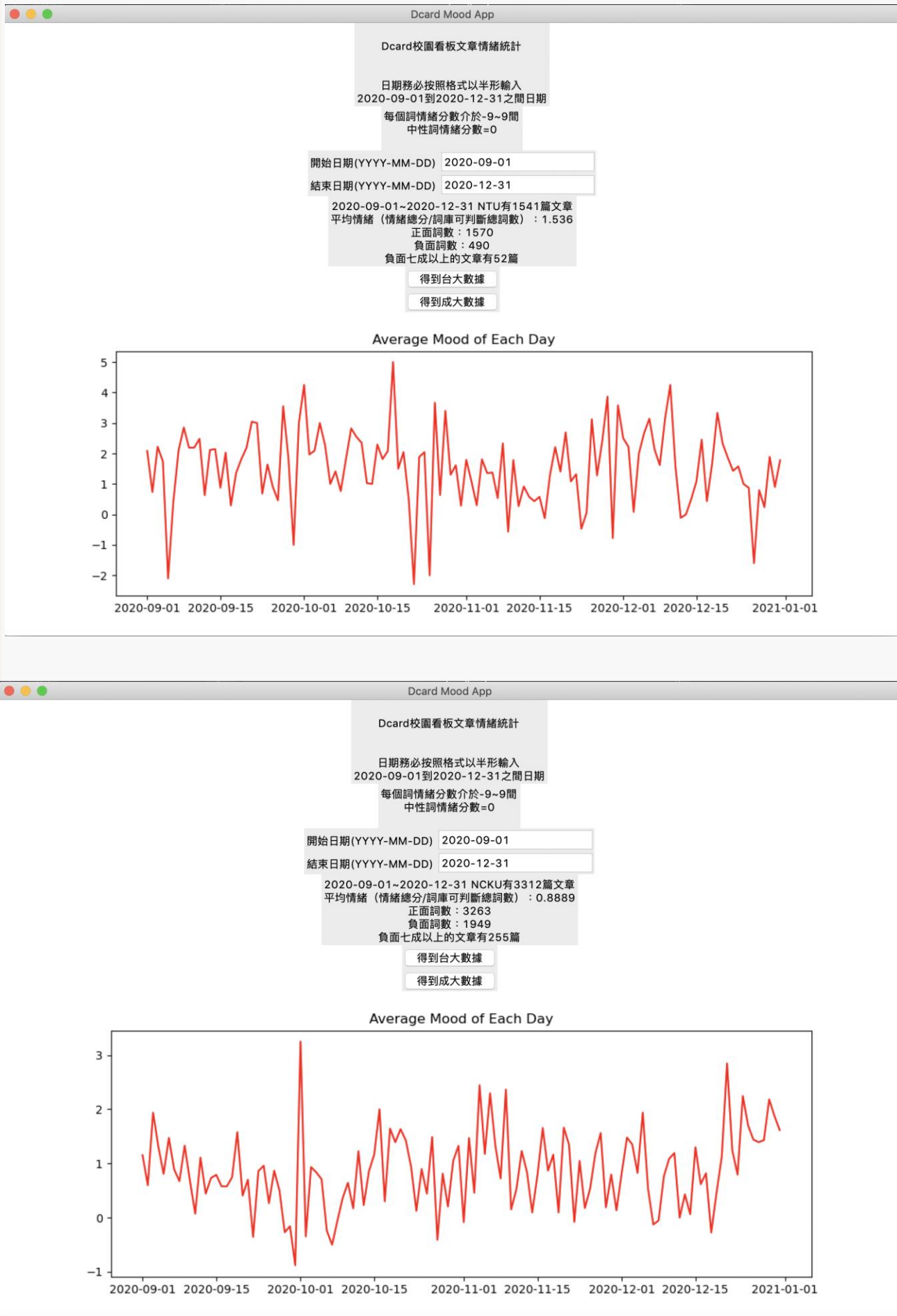
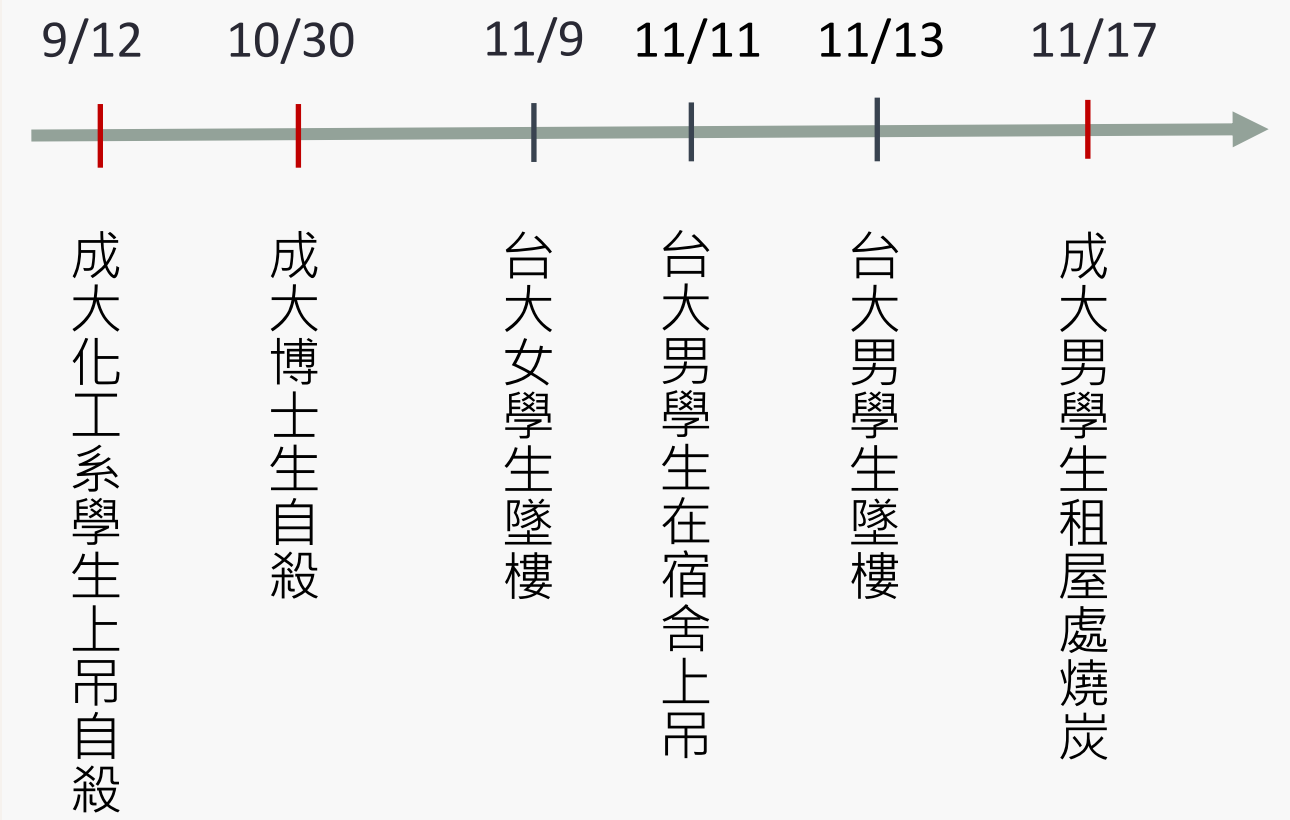


結果

- 台大與成大的情緒強度大致都分佈在正向，與預期中呈現負向的結果不符。
- 台大正面詞數大於負面詞數出現次數約3倍；成大正面詞數則是大於負面詞數出現次數約1.6倍。總體來看，正面詞數仍是大於負面詞數出現次數，與預期中呈現負面詞數大於正面詞數出現次數的結果不符。
- 台大負面情緒比例七成以上文章約有52篇，佔全部文章比例約3%；成大負面情緒比例七成以上文章約有255篇，佔全部文章比例約7%。總體來看，文章呈現負面情緒的比例並不多，與預期中負面情緒比例大於正面情緒比例的結果不符。

結果

- 9月-12月台大與成大大自殺事件時間表：



結論

- 根據圖表顯示，11月份連續自殺事件發生後，台大與成大的情緒強度多呈現正向，其原因可能是大學生普遍在自殺事件發生後，傾向發表正向情緒的文章，間接鼓勵與幫助受負面情緒影響的同儕。
- 10月中旬至11月自殺事件發生前，台大的情緒強度多次呈現負向；成大則是在10月初自殺事件發生前，情緒強度逐漸下滑至負向。由此可知，當一段期間的情緒強度呈現負向趨勢時，自殺事件發生的可能性或許會增高。因此，本文建議學校心理輔導單位及自殺防治團體可於這段期間，適時投放相關心理輔導管道與正向訊息，有效防治自殺事件的發生。
- 雖然，呈現的結果與預期中不符合，維特效應所帶來負面影響的程度不如預期的嚴重，但透過本次研究，可了解大學生情緒強度在不同時間點的趨勢，並在負面情緒強度特高的期間，建立提前預警機制，有效防止自殺模仿現象的發生。