第八組

鍾承玫N074220001  
羅育琦N094220002  
楊雯珺N094220006  
周昆穎N094220007  
陳冠汶N094220012  
 李明章 N094220023  
 鄭雅丹 N094220027

社群媒體分析

讀書會報告

目錄

[第１章 探討主題 2](#_Toc86681454)

[第２章 分析說明 3](#_Toc86681455)

[第１節 PTT爬蟲 4](#_Toc86681456)

[第２節 文章斷句 5](#_Toc86681457)

[第３節 中文斷詞 6](#_Toc86681458)

[第４節 清除停用詞 7](#_Toc86681459)

[第５節 詞頻計算 8](#_Toc86681460)

[第６節 Lexicon Based情緒分析 9](#_Toc86681461)

[第７節 每日情緒 10](#_Toc86681462)

[第１項 關聯分析7/25 10](#_Toc86681463)

[第２項 關聯分析9/15 11](#_Toc86681464)

[第８節 正向情緒文章(9/23最多文章) 13](#_Toc86681465)

[第１項 關聯分析9/23 13](#_Toc86681466)

[第９節 負向情緒文章(10/12最多文章) 15](#_Toc86681467)

[第１項 關聯分析10/12 15](#_Toc86681468)

# 探討主題

自從今年五、六月份開始，汽車開始缺晶片，民眾買不到浴缸，在美國購買一輛車，相較於一月，價格增加了22%，汽油的價格也增加了27%。其實那段時間的短缺是經濟復甦的訊號。

專家分析造成通貨膨脹的因素有三：

1. 因為消費者對汽車、家具、家用電器等商品的需求激增。消費者添購這些商品，是為了在新冠疫情期間，讓居家隔離生活更美好、過得更舒適。

2. 因為新冠疫情中斷了部分商品的全球供應鏈，如晶片的短缺就限制了汽車的生產製造。全球航運業的中斷，如停飛航線或塞港，也使得全球供應鏈的中斷更為嚴峻。

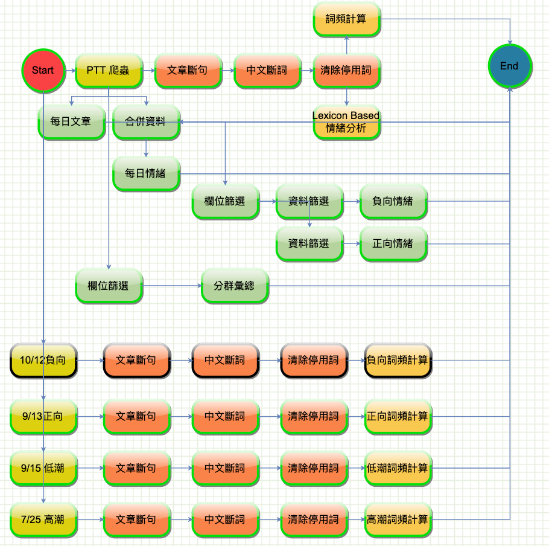
3. 因為服務業供不應求，導致價格開始反彈，這也可能是最重要的因素。隨著疫情慢慢降溫，消費者比就業者更快湧進餐廳、酒吧、美容理髮店或其他服務業。

因此本組將探討民眾對於通貨膨脹所帶來的正面及負面影響。

# 分析說明

工作流程名稱：test

建立時間：2021/10/20 23:41:48



## PTT爬蟲

日期：2021/06/01 ~ 2021/10/20

資料來源看板：Gossiping (八卦)

關鍵字：通貨膨漲、民生用品、漲價

共 **662** 筆貼文。



## 文章斷句

斷句符號：「， 。 \n ？ ：」

共 **20,696** 句。



## 中文斷詞

詞彙權重：漲價 1000、貴 1000、物價 1000

共 **129,913** 個詞。



## 清除停用詞

清除英文字母：設定為「是」。清除網址URL等英文…

自定義詞停止詞：「反正、之後、還要、而已、八卦」

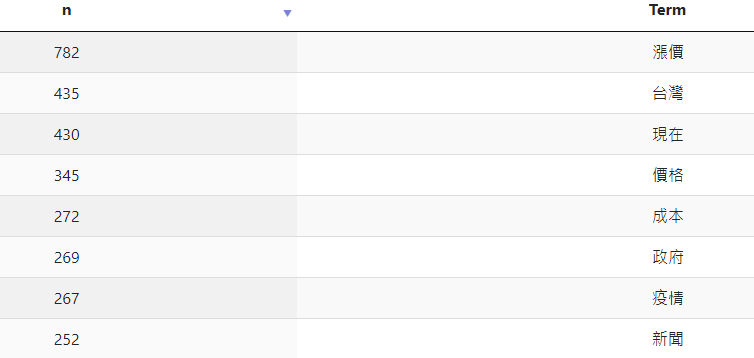
清除 **68,672** 個詞，保留 **61,241** 個詞。



## 詞頻計算

將「詞頻計算」的結果，透過「視覺化儀表板」中的「文字雲」進行呈現。

可以看出「漲價、台灣、價格、成本、房價、市場、物債、上漲、經濟、疫情等」，從這邊可以看出台灣物價確實是有上漲的趨勢。



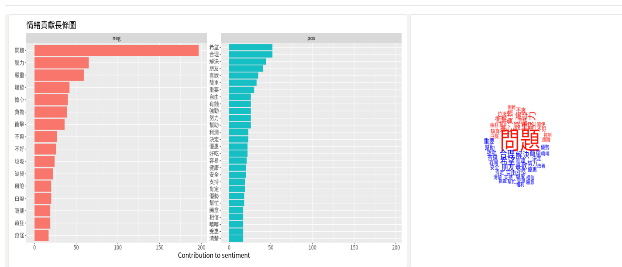


## Lexicon Based情緒分析

正面詞彙：「回饋金、利潤、好朋友、深愛、用心、做事、努力」

負面詞彙：「苦民、娘機、裝死、缺貨、落井下石、無關緊要、智障、白癡」

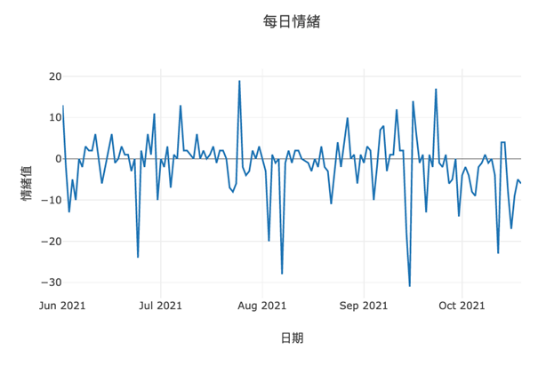




## 每日情緒

針對每日情緒做正向與負向情緒統計，並透過「視覺化儀表板」中的「折線圖」進行呈現。

正向情緒值最高在 **7/25** (值為19)、負向情緒值最高在 **9/15** (值為-31)



### 關聯分析7/25

發現跟通膨相關的文章，在7/25正向情緒值達最高峰，故欲了解7/24-7/25這兩日八卦版都在討論甚麼議題。

PTT爬蟲 🡪 文章斷句 🡪 中文斷詞 🡪 清除停用詞 🡪 詞頻計算，將「詞頻計算」的結果，透過「視覺化儀表板」中的「文字雲」進行呈現。

看到「奧運、選手、柔道、中國、日本、疫情等」，再上網調查一下，確實那天奧運有許多直撥及競賽。



### 關聯分析9/15

發現跟通膨相關的文章，在9/15負向情緒值達最高峰，故欲了解9/14-9/15這兩日八卦版都在討論甚麼議題。

PTT爬蟲 🡪 文章斷句 🡪 中文斷詞 🡪 清除停用詞 🡪 詞頻計算，將「詞頻計算」的結果，透過「視覺化儀表板」中的「文字雲」進行呈現。

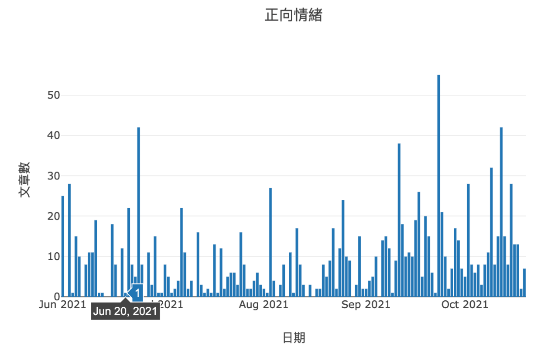
看到「疫苗、問題、高端、日本、小孩、政府、第二劑等」，在查網查看當日是否有相關性，確實9/14小孩本土確診、日本提供AZ疫苗50萬劑、長輩預約AZ第二劑。



## 正向情緒文章(9/23最多文章)

針對這段期間，每日出現正向情緒的文章數進行統計，並透過「視覺化儀表板」中的「長條圖」進行呈現。

最高峰文章數在 **9/23** (文章數為53)。



### 關聯分析9/23

發現跟通膨相關的文章，在9/23出現最多篇正向情緒的文章，故欲了解9/22-9/23這兩日八卦版都在討論甚麼議題。

PTT爬蟲 🡪 文章斷句 🡪 中文斷詞 🡪 清除停用詞 🡪 詞頻計算，將「詞頻計算」的結果，透過「視覺化儀表板」中的「文字雲」進行呈現。

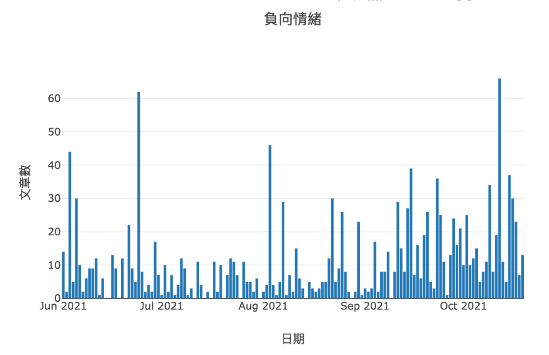
推敲該兩日，正向情緒文章，多為「疫苗」及「振興相關措施」相關議題而發表。



## 負向情緒文章(10/12最多文章)

針對這段期間，每日出現負向情緒的文章數進行統計，並透過「視覺化儀表板」中的「長條圖」進行呈現。

最高峰文章數在 **10/12** (文章數為67)。



### 關聯分析10/12

發現跟通膨相關的文章，在10/12出現最多篇負向情緒的文章，故欲了解10/11-10/12這兩日八卦版都在討論甚麼議題。

PTT爬蟲 🡪 文章斷句 🡪 中文斷詞 🡪 清除停用詞 🡪 詞頻計算，將「詞頻計算」的結果，透過「視覺化儀表板」中的「文字雲」進行呈現。

推敲該兩日，負向情緒文章，多為跟「中國」及「霸凌」相關議題而發表。或許跟國慶日總統發表的演說有關。

