

社群媒體分析第二次讀書會作業

指導老師：黃三益 教授

組別：第六組

組員：

N104020001 李采容

N104020002 廖英捷

N104020007 郭育雯

N104020008 游淑媛

N104020009 蔡雨臻

N104020010 盧貴聰

B096060020 黃湘安

M106020015 林猷盛

目錄

一、分析議題說明	3
二、工作流程設計	3
三、資料清理	4
四、詞性篩選	8
五、情緒分析 2	12
六、分群彙總與視覺化圖表	15
七、心得與結論	32

一、分析議題說明

- 主題：社會大眾對於政府「普發六千元」政策的網路情緒分析

- 議題發想：

去年全國稅收超徵 4500 億，時任行政院長蘇貞昌表示，預計將其中的 1400 億「還稅於民」，每人可以領取六千元現金。此政策造成民眾諸多關注，網路上也開始有了無排富政策、詐騙問題、未繳稅也可領取等等的話題討論。故本組將運用社群媒體分析課程所學，探討以下議題：

1. 對於普發六千有哪些情緒詞彙？
2. 網路上對於此政策是正面態度還是負面態度？
3. 網路上的情緒趨勢走向為何？

- 使用平台：文字探勘工作流程設計平台

- 第二次讀書會補充：因三月底至四月已陸續登記並發放六千元，故可進一步探討發放之後的正負面留言趨勢是否有異動。

二、工作流程設計

- 工作流程：情緒分析 2_copy_N010

- 視覺化儀表板：0404_copy_N010

- 資料來源：PTT 留言萃取(Lifeismoney 省錢版、HatePolitics 政黑版、Gossiping 八卦版)

- 分析期間：112/03/01 ~ 112/04/07

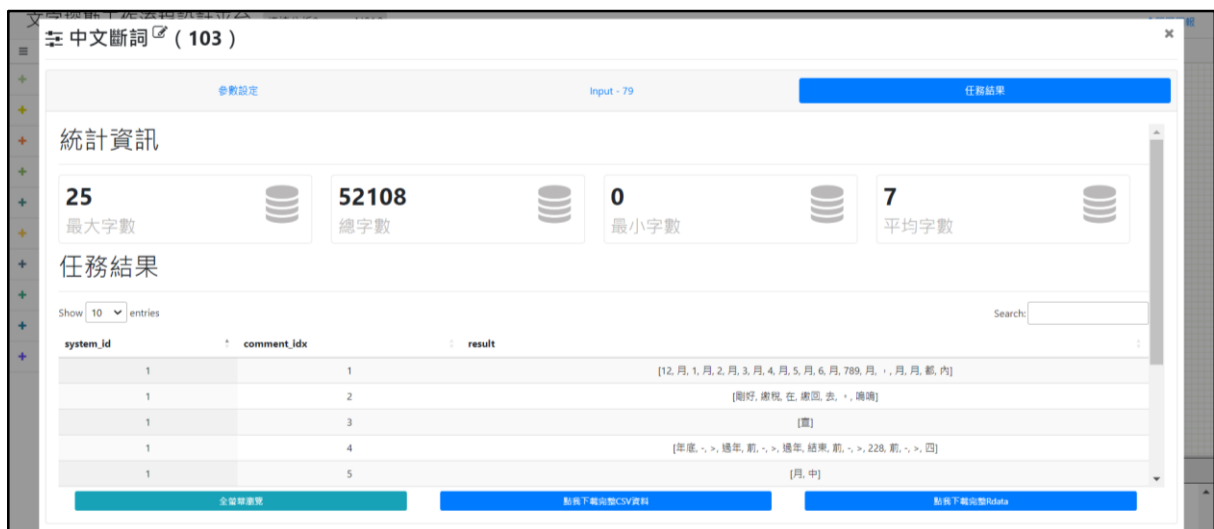
- 流程概述：

1. 指定六千元相關的關鍵字(超徵、還稅於民、還稅、普發 6000 元、普發 6 千、普發 6000、全民共享普發現金)，讀入 PTT 爬蟲資料，並使用留言萃取讀入 PTT 底下留言。
2. 將網址、PTT 平台發文文字和/n 等字詞替換成空白。
3. 用中文斷詞加大詞彙權重。
4. 將不具分析效益的字詞進行停用，但保留單字元。

(3) 以「中文斷詞」元件將文句分段，並在檢視輸出結果後增加部分字詞權重，使其能夠被篩選出來。除了第一次讀書會的「一次性」以外，增加兆豐卡、永豐銀行等銀行名稱，另外 PTT 上網友會故意用簡體字「党」做戲稱，但被字典分割，故新增民進黨、國民黨、雙標党等詞的權重，以確保結果符合語意。



斷詞後獲得的總字數為 52108 字



(4) 清除停用字：本來、這種、超過、出生、發個、拿去、網站、發個、一年、ooo、東西、反正、這幾年、準備好、好像等等，此外，為了後續的進階情緒分析，在是否清除單字元的部分選擇否。清除總數為 25881 個字詞。

清除停用詞 (106)

參數設定

Input - 103

任務結果

語言 *

Chinese

是否清除單字元 ⓘ

否

清除英文字母 *

否

清除換行符號 *

是

清除html tag *

是

使用預設停止詞

是

是否轉為小寫英文

是

清除數字 *

是

清除特殊標點符號 *

是

自定義停止詞

本來
這種
超過
出生
發價

儲存更改

清除停用詞 (106)

參數設定

Input - 103

任務結果

統計資訊

25881

清除總數

任務結果

Show 10 entries

Search:

system_id	comment_idx	result
1	1	[月, 月, 月, 月, 月, 月, 月, 月]
1	2	[慶祝, 慶回, 慶禮]
1	3	[雷]
1	4	[年底, 過年, 前, 過年, 結束, 前, 前]
1	5	[月, 中]

全部單字

點選下載完整CSV資料

點選下載完整data

四、詞性篩選

(1)詞性篩選(名詞)：選擇保留名詞 Noun 及專有名詞 Proper Noun

詞性篩選 (名詞) (255)

參數有做更動，建議重新執行

參數設定

Input - 106

任務結果

語言 *
Chinese

選擇保留詞性 *
Noun
Proper Noun
Verb
Adjective
Adverb

儲存更改

詞性篩選 (名詞) (255)

參數有做更動，建議重新執行

參數設定

Input - 106

任務結果

任務結果

Show 10 entries

Search:

system_id	comment_idx	result
1	1	[月, 月, 月, 月, 月, 月, 月, 月, 月]
1	2	[繳稅, 嗚嗚]
1	4	[年底]
1	5	[月]
1	7	[錢, 政績]
1	8	[稅金, 納稅人]
1	9	[很棒, 樣子]
1	10	[錢, 五月]

全螢幕瀏覽

點我下載完整CSV資料

點我下載完整Rdata

(2)詞頻計算(名詞)：自定義停用詞輸入「還」、「水」、「電」和「還不是」，採用預設篩選數量為 200，獲得出現頻率較高的名詞依序為「錢」、「政府」、「民進黨」、「王道銀行」、「台灣」、「銀行」、「垃圾」、「繳稅」、「笑死」、「詐騙」，大量針對執政黨、政府和對於提前發放的銀行進行討論。

詞性篩選 (形容詞) (252)

參數設定 Input - 106 任務結果

任務結果

Show 10 entries Search:

system_id	comment_idx	result
1	32	[常聞]
1	41	[金]
2	20	[小費]
2	22	[數人, 聰明]
2	26	[超微]
2	36	[超微]
3	7	[便宜]
3	16	[穩定]
3	18	[穩定]
3	21	[便宜]

全螢幕瀏覽 點我下載完整CSV資料 點我下載完整data

(4)詞頻計算(形容詞)：自定義停用詞輸入「找」、「早」、「這間」、「剪綵」、「重點」、「原」、「擋」，採用預設篩選數量為 200，獲得出現頻率較高的形容詞依序為「超微」、「笑死」、「年」、「死」、「低」、「高」、「爛」、「八卦」、「免費」、「特別」，大量出現超微、笑死等型從詞針對普發六千元進行討論。



詞頻計算 形容詞 (259)		
<div> <div>參數設定</div> <div>Input - 252</div> <div>任務結果</div> </div>		
n	Term	
21	超微	
18	笑死	
18	年	
16	死	
14	低	
13	高	
13	爛	
12	八卦	
11	免費	
<div> <div>全螢幕瀏覽</div> <div>點我下載完整CSV資料</div> <div>點我下載完整Rdata</div> </div>		

五、情緒分析 2

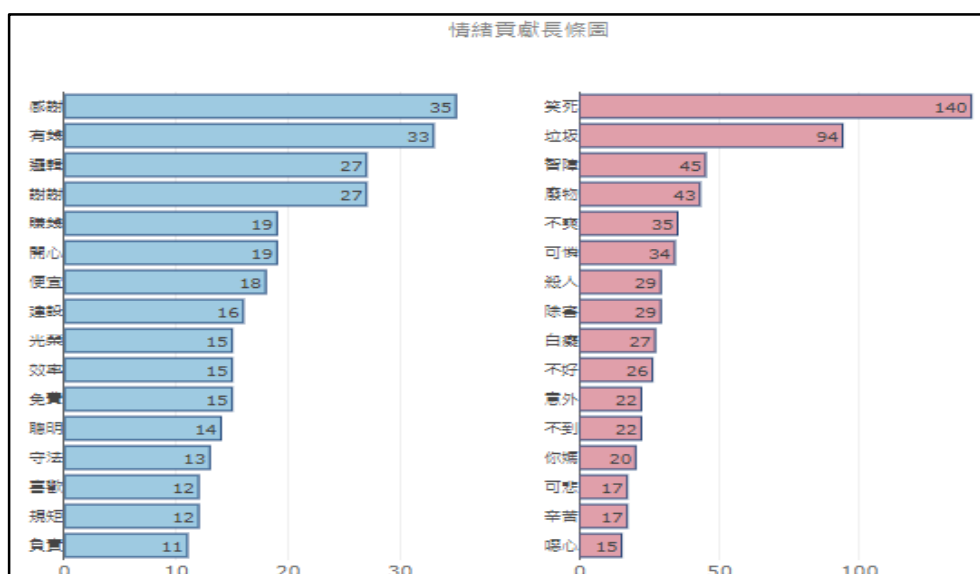
本組將「清除停用詞」後所得每則留言字詞透過「進階字典法情緒分析」、「Corpus Based 情緒分析」進行分析，目的主要是希望可以得知每則留言編屬正負面情緒以及各則留言中出現的正負面情緒字詞，並以此來做為後續彙整分析之基礎。

Ø 進階字典法情緒分析

首先本組進階字典法，採用台大情緒字典（NTUSD，National Taiwan University Sentiment Dictionary）來做為情緒詞彙判斷依據之基準，並自定義【笑死】為負面詞彙。考量 PTT 留言者多有否定詞、加強詞的使用，因此，啟用「否定詞」、「加強詞」來獲得更精準的計算情緒分數。

加強字的部分有包含「超」、「有夠」、「挺」、「太」等，強化各正負向情緒字詞的分數，較好的做到區別化。

從「情緒貢獻長條圖」，我們可發現【笑死】被普遍使用，情緒分數多達 140；其次，【垃圾】情緒分數則有 94。而正面情緒詞彙則有【感謝】、【有錢】等。這邊不難看出，民眾對於政府此「普發 6000 元」的政策常以較為偏激的負面言論來評論，對於此政策感到正向的反而相對保守。





後續將結果依情緒分數 0 分作為分水嶺，篩選看一下正負向字詞各自的數量：

資料篩選(正向) (379)

參數設定 Input - 243 任務結果

統計資訊

6655 移除數量

664 保留數量

任務結果

Show 10 entries Search:

system_id	comment_idx	positive_count	positive_words	negative_count	negative_words	booster_count	booster_words	negator_count	negator_words	sentiment_val
1	9	1.000	[很棒]	0.0	[]	0.000	[]	0	[]	1.000
1	19	1.000	[有錢]	0.0	[]	0.000	[]	0	[]	1.000
1	27	1.000	[溫暖]	0.0	[]	0.000	[]	0	[]	1.000
2	4	1.000	[努力]	0.0	[]	0.000	[]	0	[]	1.000

資料篩選(負向) (375)

參數設定 Input - 243 任務結果

統計資訊

5965 移除數量

1354 保留數量

任務結果

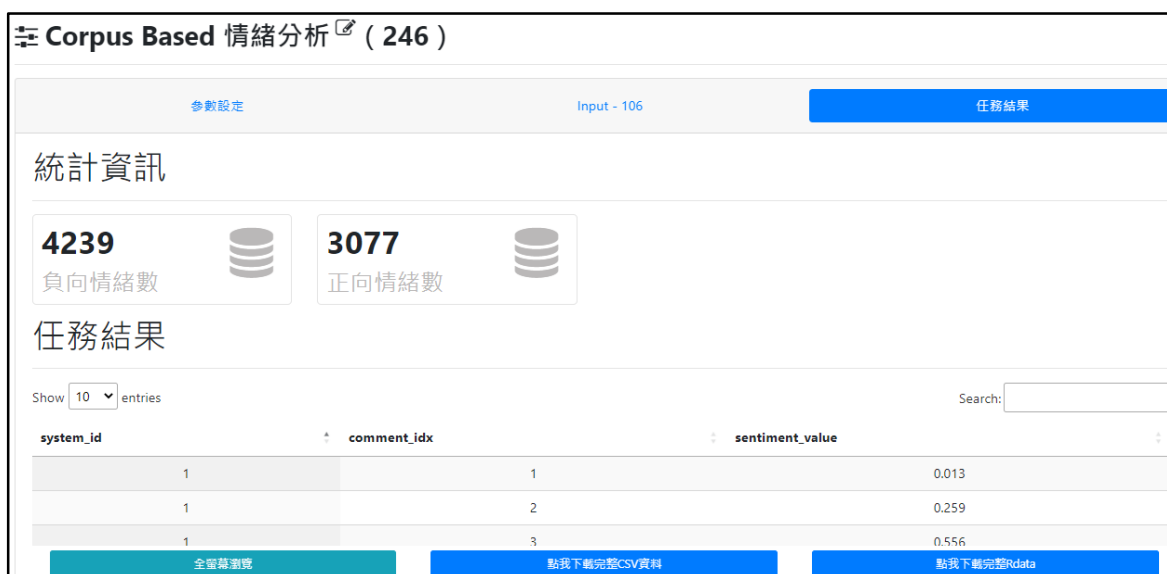
Show 10 entries Search:

system_id	comment_idx	positive_count	positive_words	negative_count	negative_words	booster_count	booster_words	negator_count	negator_words	sentiment_val
1	15	0.0	[]	-1.0000	[垃圾]	0.0000	[]	0	[]	-1.0000
1	17	0.0	[]	-1.0000	[乞丐]	0.0000	[]	0	[]	-1.0000
1	21	0.0	[]	-1.0000	[死掉]	0.0000	[]	0	[]	-1.0000
1	29	0.0	[]	-1.0000	[辛苦]	0.0000	[]	0	[]	-1.0000

從正負向情緒長條圖以及實際資料篩選後結果(正向字詞有 664 個、負向字詞有 1354 個)，進階字典法分析結果很明顯呈現負面詞彙多於正面詞彙，由此推論以詞彙來看，PPT 鄉民對於「普發 6000 元」的政策多持負向態度。

Ø Corpus Based 情緒分析

另外再依據 Corpus Based 模型作驗證分析各筆留言之情緒，依模型中預設情緒分數 0.5 分為分界，共分類出 4239 個負向情緒留言，以及 3077 個正向情緒留言，其中也可以看出負向情緒數多於正向情緒數，因此我們以 Corpus Based 分析留言句來看，初步推斷出本議題「普發 6000 元」於 PTT 鄉民討論中，較為偏向負面情緒。



依據上述兩項情緒分析模組(進階字典法與 Corpus Based 情緒分析)，得出結論皆為負向情緒居多，因結論類似，本組後續擇一選擇 Corpus Based 留言情緒分類結果來做進一步的分析。

六、分群彙總與視覺化圖表

Ø 合併資料 (290)

分群彙總前先進行「資料合併」，由「Corpus Based 情緒分析」資料與留言萃取之資料依 system_id 和 comment_idx 進行合併：

合併資料 (290)

參數設定Input - 76Input - 246任務結果

JOIN規則

新增規則刪除規則

left_key	right_key
comment_idx	comment_idx
system_id	system_id
-----請選擇-----	-----請選擇-----

儲存更改

任務結果							
Show 10 entries				Search:			
comment_idx	system_id	cmtStatus	cmtPoster	cmtContent	cmtDate	sentiment_value	
1	1	推	boy1031	12月1月2月3月4月5月6月789月，月月都內	2023-03-03 07:40:00	0.013	
1	2	推	LawLawDer	均富其實才奇怪吧	2023-03-03 14:29:00	0.334	
1	3	噓	jil	失敗的能源政策全民買單	2023-03-03 14:49:00	0.549	
2	1	推	link4517	剛好繳稅在繳回去。嗚嗚	2023-03-03 07:41:00	0.259	
2	2	→	LawLawDer	我躺平文組憑啥領和輪班GG一樣多?	2023-03-03 14:30:00	0.416	
2	3	推	WeGoStyle	喜迎電價上漲的時代來臨	2023-03-03 14:49:00	0.286	
3	1	→	boy1031	宣	2023-03-03 07:41:00	0.556	
3	2	→	LawLawDer	這就我當初放棄數學物理化學的代價啊	2023-03-03 14:30:00	0.398	
3	3	→	s142857	垃圾民進黨	2023-03-03 14:49:00	0.523	

Ø 每日留言數量 (297)

合併資料後，以日期作分群，並將計算欄位設定為 comment_idx_x，彙總函數設定 Max，在此保留日期欄位以利看出每日的留言數量趨勢。

每日留言數量 分群彙總(日期) (297)
參數有做更動，建議重新執行

參數設定
Input - 290
任務結果

選擇日期欄位 *
cmtDate

選擇日期類型 *
日

計算欄位(按住ctrl(Windows)或command(MAC)可以複選) *
system_id
comment_idx_x
cmtStatus
cmtPoster
cmtContent
cmtDate
comment_idx_y
sentiment_value

日期格式 *
%Y-%m-%d %H:%M:%S%z

匯總函數 *
max

保留欄位(按住ctrl(Windows)或command(MAC)可以複選) *
system_id
comment_idx_x
cmtStatus
cmtPoster
cmtContent
cmtDate
comment_idx_y
sentiment_value

儲存更改

統計資訊

35
群組數量

任務結果

Show 10 entries
Search:

cmtDate	comment_idx
2023-03-03	90
2023-03-04	43
2023-03-05	33
2023-03-06	83

Ø 每日情緒平均 分群彙總（日期）（299）

將合併後之資料使用「分群彙總(日期)」模組，取每一天所有留言的情緒數值平均，所得結果為每日情緒的平均，用來瞭解普發 6000 元的政策，隨著時間的推移，所獲得的正負面情緒趨勢。

每日情緒平均 分群彙總(日期) (299)
參數有做更動，建議重新執行

參數設定
Input - 290
任務結果

選擇日期欄位 *
cmtDate

選擇日期類型 *
日

計算欄位(按住ctrl(Windows)或command(MAC)可以複選) *
comment_idx
system_id
cmtStatus
cmtPoster
cmtContent
cmtDate
sentiment_value

日期格式 *
%Y-%m-%d %H:%M:%S%z

匯總函數 *
mean

保留欄位(按住ctrl(Windows)或command(MAC)可以複選) *
comment_idx
system_id
cmtStatus
cmtPoster
cmtContent
cmtDate
sentiment_value

統計資訊

35
群組數量

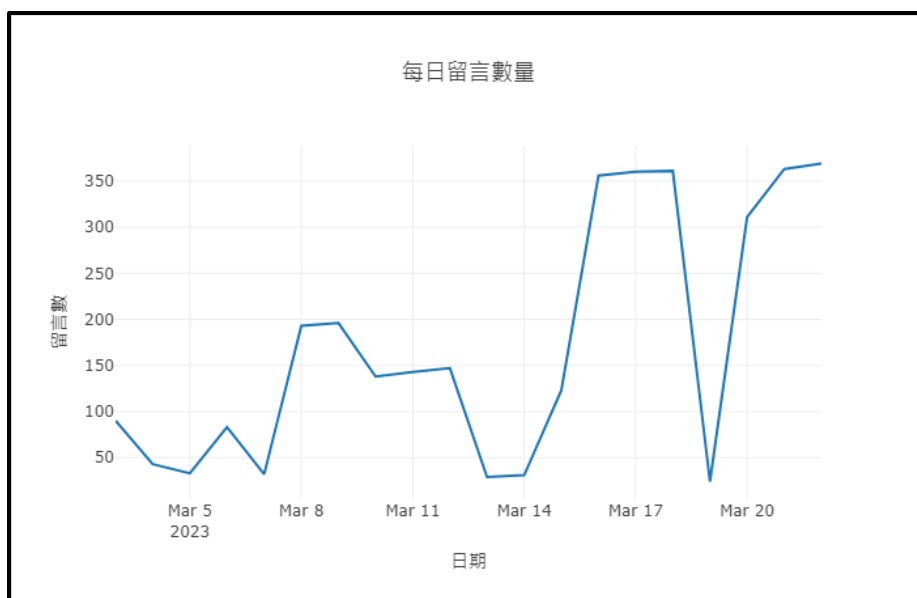
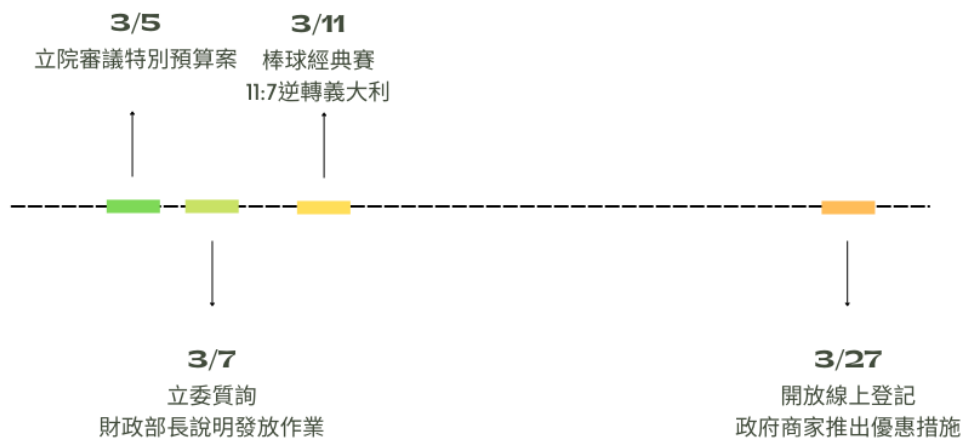
任務結果

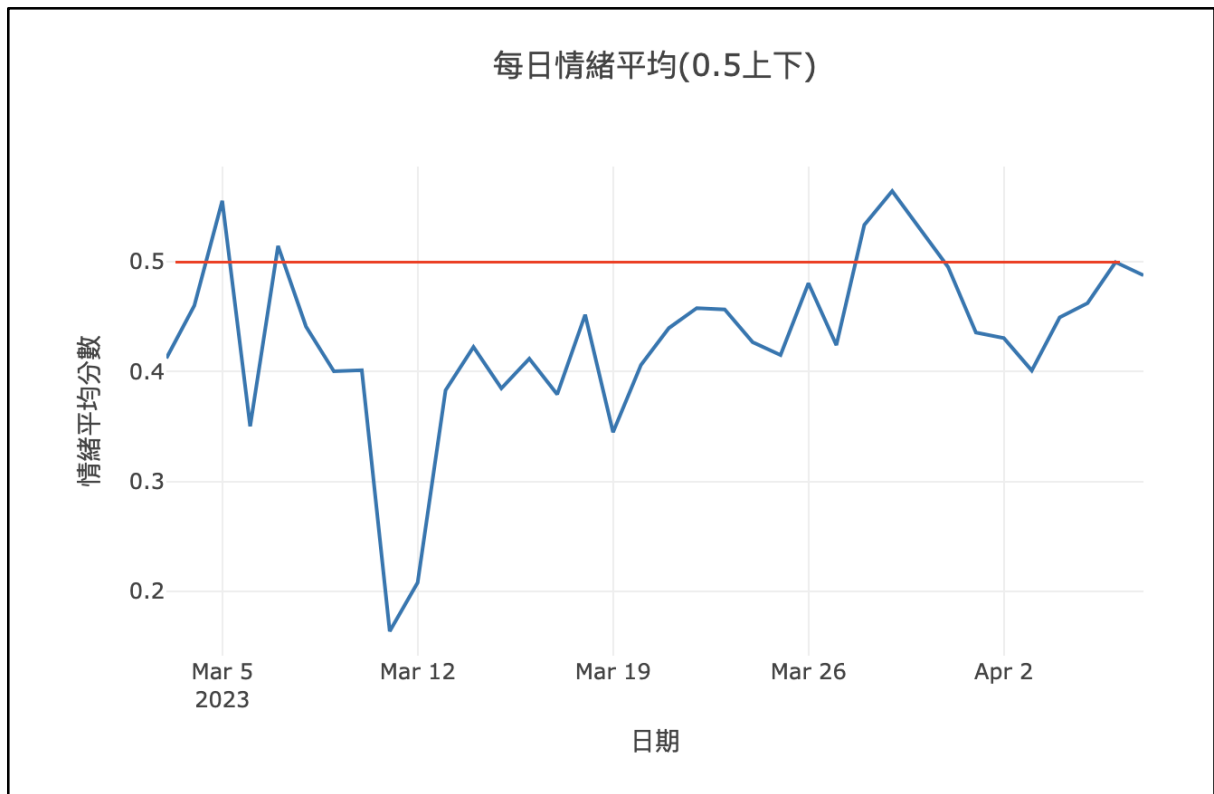
Show 10 entries
Search:

cmtDate	sentiment_value
2023-03-03	0.412087
2023-03-04	0.459764
2023-03-05	0.555000
2023-03-06	0.350043

每日留言數量與每日情緒平均的視覺化圖表

由上述兩項計算的每日留言數量與每日情緒平均可繪製成以下圖表，可發現時間區段內大多的平均情緒分數都低於原本分析設定的正向定義 0.5，更詳細的來看 35 天內只有 4 天的平均情緒分數大於 0.5。3/5 是因立院將審查普發現金的疫後特別預算案，其中財政部書面報告指出普發現金需編列新台幣 1416 多億元。3/7 則是在立委質詢下財政部長有明確表示有望在總統公布後一週內開始進行發放作業。3/27 起開放線上登記入帳以及公布可領 6000 元的 ATM 機台位置，各銀行、地方政府與商街也趁勢推出優惠活動，因此可推測隨著政策有新進展，往「領錢」又更邁向一步，留言者的情緒便較為正面。反之，3/11 適逢棒球經典賽台灣 11:7 逆轉義大利迎來首勝，因此對於普發六千元議題討論度遠低於其他日期，平均的情緒分數因為留言數量極少而被少數負面情緒強烈的留言拉低。





Ø 欄位篩選 (365)

下列分析將著重於日期、留言者與其留言情緒的統計，因而以「欄位篩選」功能保留「system_id」、「comment_idx_x」、「cmtPoster」、「cmDate」與「sentiment_value」欄位來進行分析。

欄位篩選 (365)

參數設定 Input - 290 任務結果

選擇要保留的欄位(按住ctrl(Windows)或command(MAC)可以複選) *

comment_idx
system_id
cmtStatus
cmtPoster
cmtContent
cmtDate
sentiment_value

儲存更改

關於正向情緒

Ø 資料篩選 (正向) (301)

使用資料篩選將 Corpus Based 情緒分析後情緒分數大於等於 0.5 的留言獨立出來分析。

資料篩選 (正向) (301)

參數有做更動，建議重新執行

參數設定

Input - 365

任務結果

條件式 *

'\$sentiment_value>=0.5'

篩選結果移除了 4239 筆留言，並保留 3077 個正向留言。

統計資訊

4239
移除數量

3077
保留數量

任務結果

Show 10 entries

Search:

comment_idx	system_id	cmtPoster	cmtDate	sentiment_value	result
1	3	jil	2023-03-03 14:49:00	0.549	_selected
1	4	saltlake	2023-03-03 17:51:00	0.674	_selected
1	5	PSptt	2023-03-03 21:33:00	0.642	_selected
1	6	sion1993	2023-03-03 23:44:00	0.547	_selected

Ø 每日正向言論數 分群彙總 (日期) (305)

依日期為分群單位，以「count」計算統計每日情緒分數大於等於 0.5 的留言數量，以取得每日的正向留言總數。

每日正向言論數 分群彙總(日期) (305)

參數設定

Input - 301

任務結果

選擇日期欄位 *

cmtDate

選擇日期類型 *

日

計算欄位(按住ctrl(Windows)或command(MAC)可以複選) *

comment_idx
system_id
cmtPoster
cmtDate
sentiment_value
result

日期格式 *

%Y-%m-%d %H:%M:%S%z

匯總函數 *

count

保留欄位(按住ctrl(Windows)或command(MAC)可以複選) *

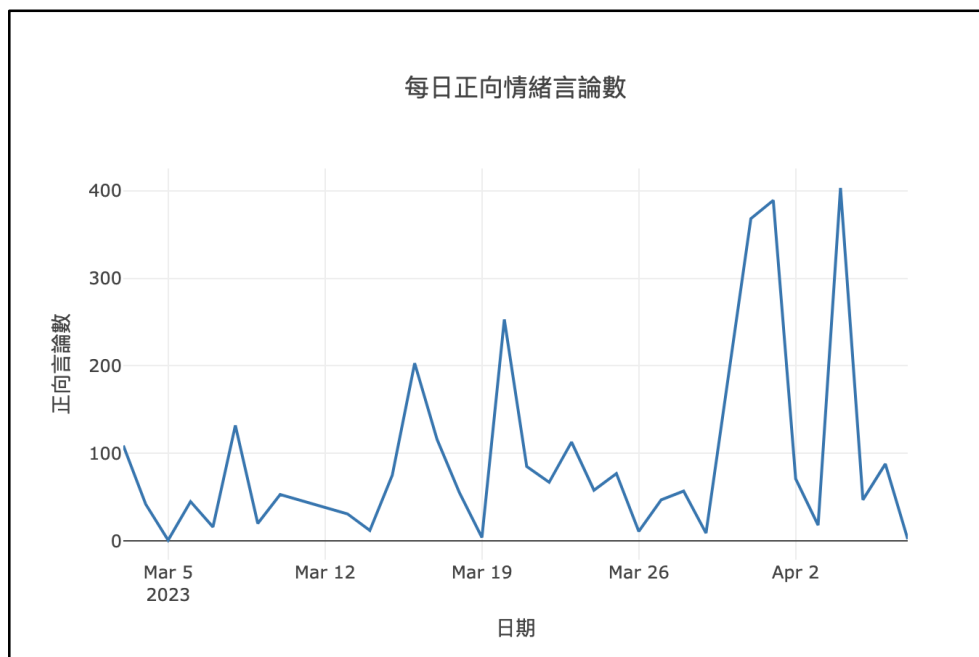
comment_idx
system_id
cmtPoster
cmtDate
sentiment_value
result

此正向分群數量為 33 與每日總留言群組數 35 不相等，推測可能有幾天的情緒分數未大於等於 0.5。可以此數字繪製視覺化圖表，觀察每日的留言情緒分數趨勢為何。



每日正向言論數的視覺化圖表

以上述統計的每日正向情緒言論數量繪製。正向情緒最高的相對高峰為 3/31、4/1 與 4/4。3/31 是王道銀行無預期的提前撥款，隔日財政部則宣布八大公股銀行的款項已入帳，只要各銀行作業完成即可提前撥付，4/3 為發放開始的首個週一，民營銀行陸續完成撥款，因此推測這幾個日期關於普發 6000 元的討論度較高，情緒分數因而在相對高點。



Ø 留言者的正向言論數 分群彙總（非數值）（333）

以留言者 ID 欄位進行分群，以「count」計算欄位數量的函數來計算「cmDate」留言時間欄位的總數，以取得各留言者之正向留言數。

留言者正向言論數 分群彙總 (非數值) (371)

參數設定

Input - 301

任務結果

使用...欄位進行分群(按住ctrl(Windows)或command(MAC)可以複選) *

comment_idx

system_id

cmtPoster

cmtDate

sentiment_value

result

匯總函數 *

count

nunique

min

max

first

last

sum

計算欄位(按住ctrl(Windows)或command(MAC)可以複選) *

comment_idx

system_id

cmtPoster

cmtDate

sentiment_value

result

任務結果共有 1925 個 ID 有情緒數值大於等於 0.5 的正向留言

留言者正向言論數 分群彙總 (非數值) (371)

參數設定

Input - 301

任務結果

統計資訊

1925

群組數量

任務結果

Show 10 entries

Search:

cmtPoster

cmtDate@count

Anvec7

AustinRivers4

PANAMIDA4

Ø 資料篩選（正面前 20）（341）

使用「資料篩選」功能篩選正向留言數前 20 的發文者 ID，多次嘗試後將 cmtDate@count（正向留言數總和）> 7 的發文者 ID 留下。

資料篩選 (正面前20) (341)

參數設定
Input - 371

條件式 * ⓘ

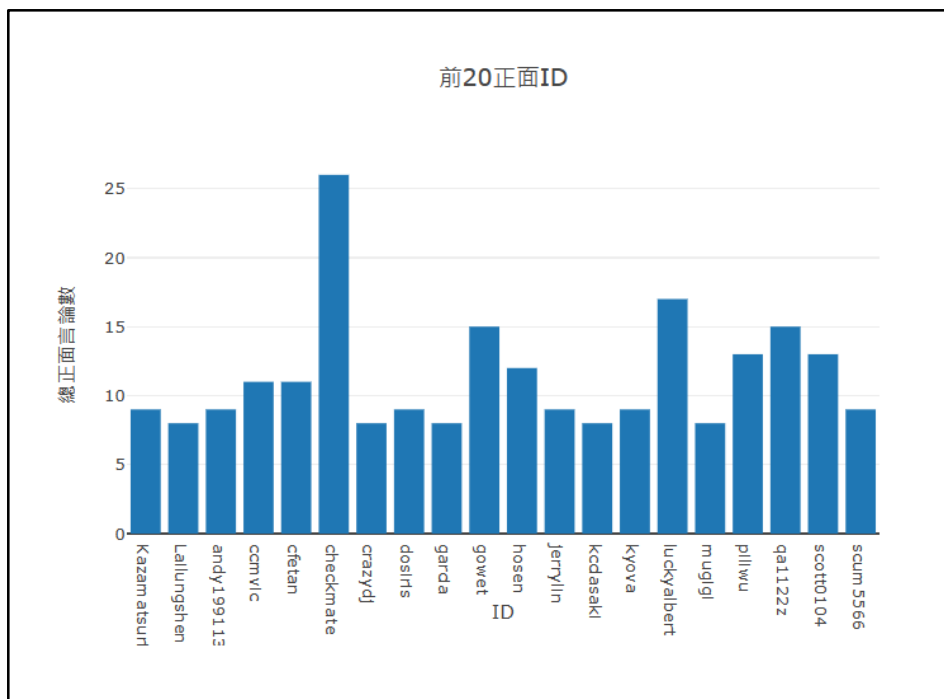
'\$cmtDate@count>7'

來回篩選後得出情緒分數大於 7 的留言者共有 27 位，後續以視覺化圖表方式將 ID 與數值羅列出來。



Ø 正向情緒最高的前 20 名 ID

因為上述「資料篩選」結果為 23 個群組，因此在使用視覺化繪圖時設定資料數量為 20 個，以求得正向情緒最高的前 20 名留言者。由圖表結果可清楚地看出名為「checkmate」的留言者總正面言論數最多，且明顯高於其他 ID，次多的 ID 為「luckyalbert」，其他的 ID 正面言論數皆不及 15。



關於負向情緒

Ø 資料篩選（負向）（324）

條件設定以情緒分數 < 0.5 為符合負面資料，一共移除 3,077 筆，最終保留 4,242 筆留言數量，其任務結果如下：

資料篩選 (負向) (324)

參數設定

Input - 365

任務結果

統計資訊

3077
移除數量

4242
保留數量

任務結果

Show 10 entries

Search:

comment_idx	system_id	cmtPoster	cmtDate	sentiment_value	result
1	1	boy1031	2023-03-03 07:40:00	0.013	_selected
1	2	LawLawDer	2023-03-03 14:29:00	0.334	_selected
2	1	link4517	2023-03-03 07:41:00	0.259	_selected
2	2	LawLawDer	2023-03-03 14:30:00	0.416	_selected
2	3	WeGoStyle	2023-03-03 14:49:00	0.286	_selected
3	2	LawLawDer	2023-03-03 14:30:00	0.398	_selected
4	1	kcclasaki	2023-03-03 07:41:00	0.179	_selected

全螢幕瀏覽

點我下載完整CSV資料

點我下載完整Rdata

Ø 每日負向言論數 分群彙總（日期）（327）

依日期為分群單位，以「count」的匯總函數計算出單位日期內的留言數量，並保留 cmtDate 欄位以利查看。

每日負向言論數 分群彙總(日期) (327)

參數有做更動，建議重新執行

參數設定

Input - 324

任務結果

選擇日期欄位 *

cmtDate

選擇日期類型 *

日

計算欄位(按住ctrl(Windows)或command(MAC)可以複選) *

comment_idx

system_id

cmtPoster

cmtDate

sentiment_value

result

日期格式 * ⓘ

%Y-%m-%d %H:%M:%S%z

匯總函數 * ⓘ

count

保留欄位(按住ctrl(Windows)或command(MAC)可以複選) *

comment_idx

system_id

cmtPoster

cmtDate

sentiment_value

result

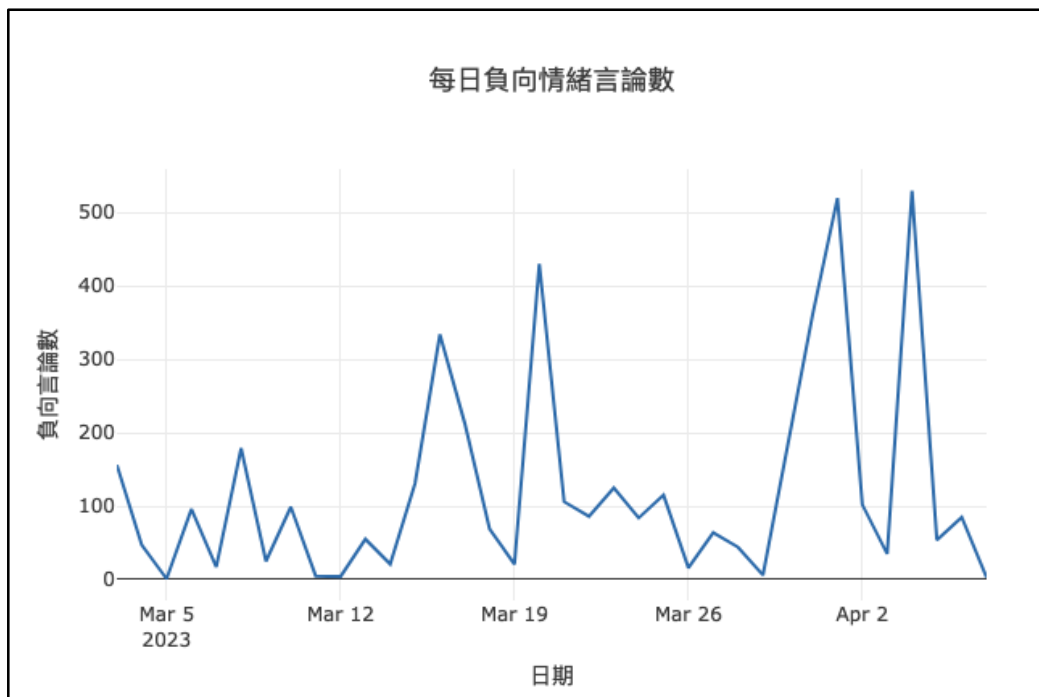
儲存更改

觀察每日的「負面」情緒留言數量的變化並確認某幾天負面高峰之原因：



Ø 每日負向情緒言論數視覺化圖表

可看到「4/1」與「4/4」達負面情緒高峰，清明連假第一天，銀行 ATM 發放 6000 元的評論數都大於 2,000。



Ø 留言者的負向言論數 分群彙總（非數值）（336）

以留言者欄位進行分群，以「count」匯總函數來計算「cmtDate」欄位，從而看出哪些發言者的負向言論特別多，並從中分析這些負面聲浪是對於哪方面不滿或覺得有不夠周全的地方。設定與結果如下(這邊使用留言時間來計算的原因為單一留言者不太可能同時發布多則訊息)：

留言者的負向言論數 分群彙總 (非數值) (336)

參數設定

Input - 324

任務結果

使用...欄位進行分群(按住ctrl(Windows)或command(MAC)可以複選) *

comment_idx

system_id

cmtPoster

cmtDate

sentiment_value

result

匯總函數 *

count

nunique

min

max

first

last

sum

計算欄位(按住ctrl(Windows)或command(MAC)可以複選) *

comment_idx

system_id

cmtPoster

cmtDate

sentiment_value

result

依照留言者的負向分群彙總，可得出 2,436 筆的留言者 ID，與其負向留言總數。

留言者的負向言論數 分群彙總 (非數值) (336)

參數設定 Input - 324 任務結果

統計資訊

2436
群組數量

任務結果

Show 10 entries Search:

cmtPoster	cmtDate@count
Anvec	15
AustinRivers	9

Ø 資料篩選 (負面前 20) (339)

使用「資料篩選」功能將情緒分數計數，cmtDate@count(負向留言數總和) > 9 的發文者留下，移除 2,406 筆留言者資料。



資料篩選 (負面前20) (339) 參數有做更動，建議重新執行

參數設定 Input - 336

條件式 * ⓘ

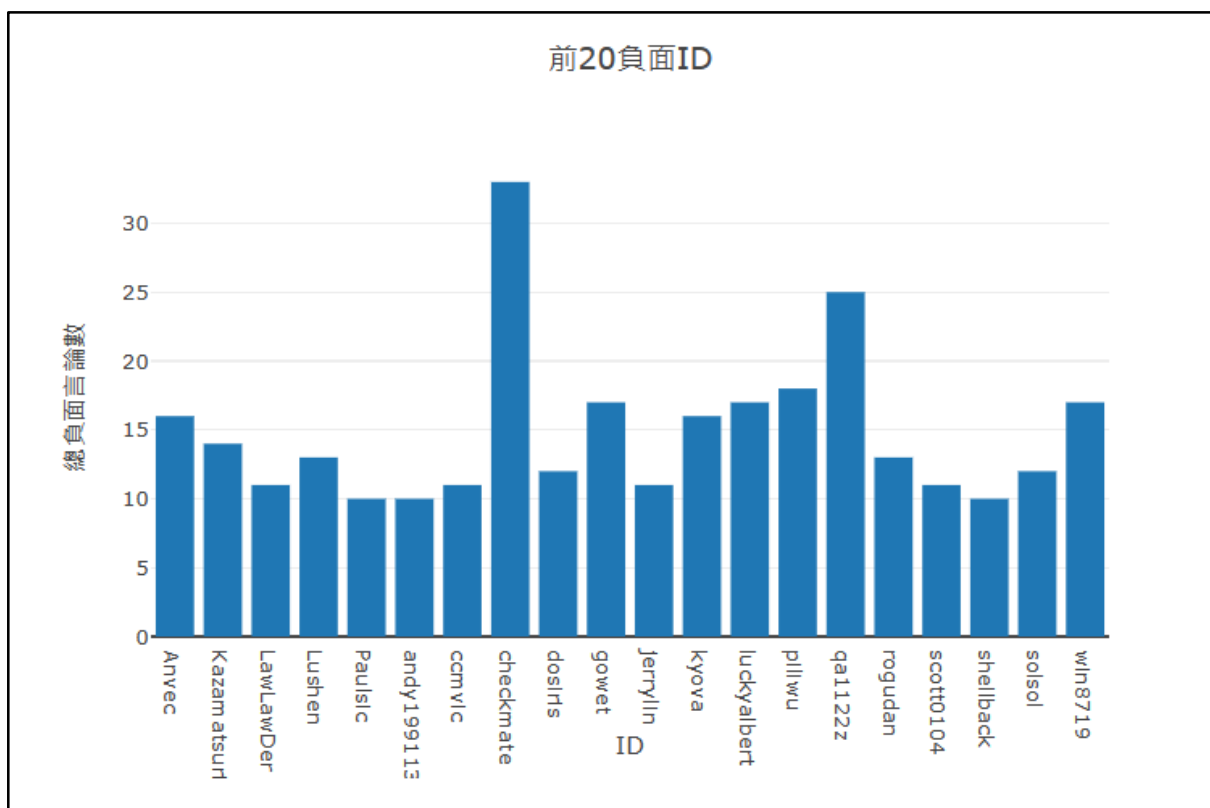
'\$cmtDate@count>9'

經反覆驗證後，條件以負向留言數總和 >9 做設定，會保留最終 24 筆留言者資料，故結果如下：

資料篩選 (負面前20) (339)			
參數設定		Input - 336	任務結果
2412 移除數量		24 保留數量	
任務結果			
Show 10 entries		Search: <input type="text"/>	
cmtPoster	cmtDate@count	result	
checkmate	33	_selected	
qa1122z	23	_selected	
gowet	18	_selected	
luckyalbert	18	_selected	

Ø 負向情緒最高的前 20 名 ID

因為上述「資料篩選」結果為 24 個群組，因此在繪圖時設定資料數量為 20 個。負面分數排序由高至低的留言者 ID：checkmate > qa1122z > piliwu ...等。



Ø 文字雲 - 詞頻計算 (名詞)

A word cloud featuring various Chinese characters. The most prominent words include "無限" (Infinite), "新" (New), "廢物" (Waste), "可悲" (Pitiful), "前" (Before), "不爽" (Not satisfied), "綠色" (Green), "正義" (Justice), "天才" (Genius), "無腦" (Brainless), "噁心" (Disgusting), "不多" (Not many), "不" (Not), "出" (Out), and "非" (Non-). The characters are in different colors like red, green, blue, and black.



不爽領 免費廣告 政府錢
金融機構 漲電價 天氣之子 特別預算
操你媽 給我 綠能
八卦仔 繳電費 下個月 為民除害 可憐 拖久
繳 不守規矩 王道銀行 銀行 不早說
做事 笑死 王道銀行
物價漲 繳稅 月月 幹你娘 守規矩 政府做
三小 撈錢 騙集團 債留子孫
明年選舉 政府有錢 繳稅 大撒幣 詐騙 買票
錢政府 劫富濟貧 垃圾 民進黨 錢積怨已久
謝主隆恩 發萬 拍廣告

七、心得與結論

本次主題延續前一次讀書會所定之「社會大眾對於政府「普發六千元」政策的網路情緒分析」，由於時間的推移，得以讓此主題有前後的觀察，已完成一個較為詳盡的研究也是一個完整的主題。

在這次獲得的資料當中能很清楚的看到討論方向的不同，從原先的發放前問題如：詐騙、宣導、如何登記、政府稅收等，轉為現在的銀行偷跑、繳稅等議題。

使用進階字典法情緒分析後，透過更精準地計算得到的情緒分數可知，本次普發 6000 議題仍然是負面情緒大於正面情緒的；在 Corpus Based 情緒分析上，透過將預設分數作為分界分類出正負向情緒 7000 多筆，其中正向為 3077，負向為 4239，由此也可推得本次議題即便來到發放的時間點，在 PTT 裡也是偏向負面情緒的。其中值得一提的是，透過正向負向分數排名篩選出的前 20 名留言者 ID，這有助於關注議題時能分辨該留言者之身分是否為公關操作所為，惟目前使用情緒分析尚無法看出全貌，需搭配原文和留言了解分析，但也是值得探討的方向。

本次觀察 3/31、4/1、4/4 的討論最高，不論是正向或負向的言論數，都於這三天呈現高峰，透過回顧新聞事件發現到普發 6000 議題在 3/31 因王道銀行偷跑、八大公股銀行決定提前發放等事件發生後討論熱絡。按理說發放現金後民眾在討論度上因該會趨向正向言論，透過新聞和文字雲的分析可以發現，因銀行偷跑事件，公股銀行和部分銀行陸續宣布提早發放，然其餘銀行皆須等到原先發放日，因此針對不同銀行間會有負面言論出現。除銀行以外，也因發放日近繳稅季，發放日前又有漲電價的議題，因此在討論中可以看到討論執政黨、退稅繳稅、發錢繳電費等問題取代，在討論中看到的能源議題等都可歸類於上述而非離題，其中有趣的是「天氣之子」一詞，小組成員不理解為何會出現於此，深入了解後發現這是在諷刺總統任內缺水缺電，以表達普發 6000 都拿去繳水費電費的不滿。

最後值得注意的是，這次有許多的髒話出現，回到原文觀看可以發現，部分的髒話作為語助詞出現，整句話的語義可能是正向的，如「X！拿到錢了好爽。」之類的言論會被歸類到負面，故本組認為，再遇到這類單詞或髒話的場合需考慮臺灣人的日常用語習慣，若做結合改善的話，或許本次的正負向言論數量會更加接近。