

---

# 美國總統大選情緒分析

## TARFLOW

### 第九組

組員名單:

N124320001 沈世英

N124320005 趙千霽

N124320008 邱懷英

N124320014 陳柏豪

N124320018 邱翔

N124320021 永瑞華

N124320026 施泳呈

N084020030 黃緯翔

---

---

# 專案動機

現在的人越來越習慣在網路上討論政治，PTT 這類匿名論壇更是許多台灣網友發表意見的重要地方。特別是在美國總統大選期間，PTT「美國總統大選」政黑板（HatePolitics）上會有大量的討論，這些留言可能反映網友對候選人的看法，甚至影響其他人的意見。

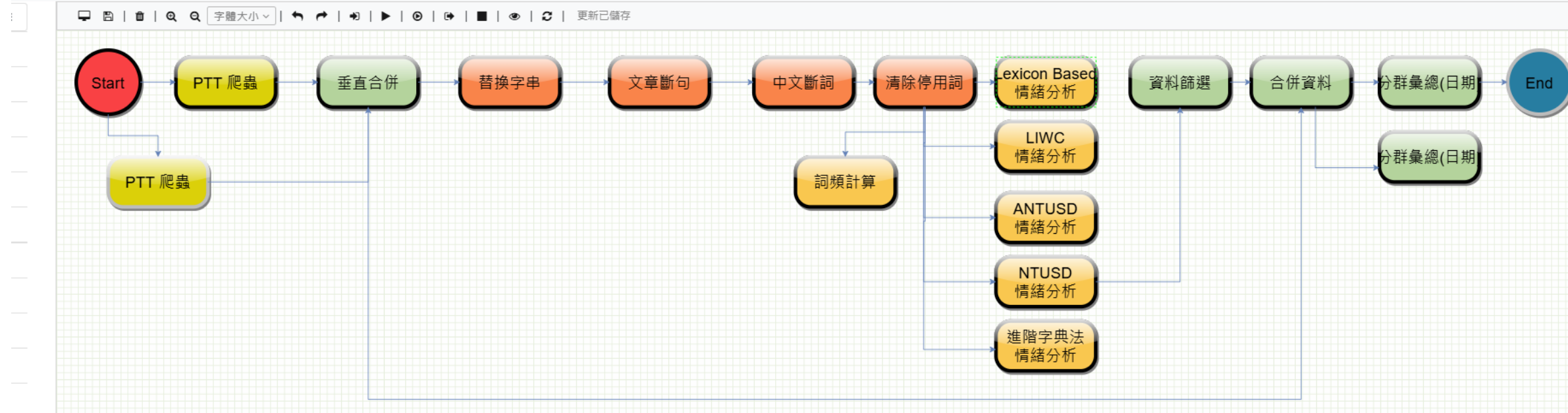
但因為這些討論內容多是文字，意見也非常多元，很難直接用肉眼判斷整體的情緒趨勢。因此，本研究希望分析 PTT 上的留言情緒，看看大家對不同候選人的態度是偏向支持、反對，還是中立。進一步來說，我們也想了解這些情緒變化是否與選舉過程中的重大事件、候選人的表現或媒體報導有關。這樣的研究不但能幫助我們更了解網路輿論的變化，甚至對未來的政治情緒分析或選舉預測也可能有所幫助。

# TARFLOW 流程說明

文字探勘工作流程設計平台

美國大選備份(世英)

★問!



# PTT爬蟲參數設定

☰ PTT 爬蟲  ( 20 )



參數設定	任務結果
<p><b>選擇看板 *</b></p> <div><div>GetMarry(結婚)</div><div>Gossiping(八卦)</div><div>graduate(研究所)</div><div>GuineaPig(天竺鼠)</div><div>Hamster(倉鼠)</div><div>HardwareSale(硬體二手)</div><div>HatePolitics(政黑)</div></div>	<p><b>搜尋關鍵字 ❶</b></p> <div><div>美選</div><div>美國大選</div><div>川普槍擊案</div><div>美國總統</div><div>川普當選</div></div>
<p><b>搜尋起始日期</b></p> <div><div>2024/03/01</div><div></div></div>	<p><b>排除關鍵字 ❶</b></p> <div><div>柯文哲</div><div>張安樂</div><div>竹聯幫</div><div>中國漁船</div><div>巴爾的摩</div></div>
<p><b>搜尋結束日期</b></p> <div><div>2024/12/31</div><div></div></div>	
<div>儲存更改</div>	

# PTT爬蟲結果

## 統計資訊

10

欄位數



1397

資料筆數



## 任務結果

Show 

10

 entries

Search:

system_id	artUrl	artTitle	artDate	artPoster	artCatagory	artContent	artComment	e_ip	insertedDate	dataSource
1	https://www.ptt.cc/bbs/HatePolitics/M.1709306853.A.20C.html	[討論]今年美國大選都是棄台論？	2024-03-01 23:27:31	taxuan	HatePolitics	拜登看起來也陷入烏克蘭以色列泥沼 根本沒力顧台灣 川普更不用說了 還說要課台積電重稅 看起來兩黨都對台灣很不友善啊 根本就是棄台論當道 想問問政黑版眾 今年是不是Team Biden	[{"cmtStatus": "推", "cmtPoster": "lono", "cmtContent": "當然支持跳脫藍紅的那一個", "cmtDate": "2024-03-01 23:2..."}]	223.137.59.58	2024-03-02 01:39:15	ptt

全螢幕瀏覽

點我下載完整CSV資料

點我下載完整Rdata

點我下載完整json資料

# 遇到的困難1:垂直合併(2024年&2025年無法一起跑)

≡ 垂直合併 ( 38 )

參數設定

Input - 20

Input - 27

任務結果

統計資訊

10  
欄位數

1515  
資料筆數

任務結果

Show 10 entries

Search:

system_id	artTitle	artUrl	artContent	artPoster	e_ip	artCatagory	dataSource	insertedDate	artComm
1	Fw: [NEWS]台 酒無法銷 售電子菸 流失業務 量莊翠	https://www.ptt.cc/bbs/HatePolitics/M.1735808422.A.449.htm	作者: tobetwob (WHO FCTC 菸草減害專家) 看板: tobacco 標題: [NEWS] 台酒無法銷售電子菸流失業務量莊翠 時間: Thu Jan 2 16:59:25 ...	tobetwob	None	HatePolitics	ptt	2025-02-23 09:49:03	[[{"cmtSta "噓", "cmtPost "rtwodtv "cmtCont ":沒有白藍 了的法 "cmtDat

全螢幕瀏覽

點我下載完整CSV資料

點我下載完整Rdata


點我下載完整json資料

## 統計資訊

取代數量



## 任務結果

Show 10  entries

Search:

system_id	result
1	<p>作者: tobetwob (WHO FCTC 菸草減害專家) 看板: Tobacco 標題: [NEWS] 台灣無法銷售電子菸流失業務量 莊翠雲： 時間: Thu Jan 2 16:59:25 2025 台灣無法銷售電子菸流失業務量 莊翠雲：將加強和衛福部溝通 2025-01-02 13:32 聯合報 / 記者 朱漢崙 / 台北即時報導 台灣菸酒董座湯期安在今日財委會上答詢時指出，配合政策明年將要求所有投放商，若沒有實名制，就避免去投放廣告，他也坦言，菸害防制法上路之後，很多民眾轉抽電子煙、加熱菸，但台灣還無法銷售加熱菸，流失很多營運量，對此，莊翠雲在答詢時指出，會加強和衛福部的溝通。而民眾黨籍立委黃珊珊質疑，大陸正進行低價啤酒傾銷，國人若喝到劣質啤酒但不自知，莊翠雲指出，已透過釀酒商協會會員的身分，向經濟部提出申請市場調查，亦即調查中國大陸啤酒是否有傾銷事實。民進黨籍立委吳秉叡指出，加熱菸遭衛福部擋關14、15個月，不但影響台灣公司的銷量，也造成稅收流失。湯期安坦言，台灣菸酒還無法銷售加熱菸，的確影響到40、50億元的銷量。湯期安也表示，菸害防制法上路之後，台灣的香菸銷量從289萬箱減少到168萬箱，這些減少的部分，大多已流向持有非合法的電子菸及加熱菸。湯期安在答覆李坤城質詢指出，目前菸酒公司大部分的廣告仍在傳統的電視與實體通路，去年整個行銷預算為3748萬元，在line、google、Fb的廣告才下1千多萬元。李坤城也指出，現在中華職棒大聯盟，已有簽約金信託，還有職棒球員防賄基金信託，一方面讓球員避免簽賄，另一方面也讓球員的收入有信託保障，他要求台啤與台銀向職籃聯盟談判加入這個信託，對此莊翠雲答覆，會讓公股的球隊也加入職籃防賄信託。為了維持物價穩定，湯期安在答覆立委羅明才指出，包括米酒等民生用酒售價都不會調高，至於紹興、紅露酒是否調整？他表示，會調整產品結構，但售價是否調整，則會再評估。心得：美國總統川普支持電子菸封閉</p>

# 替換字串/文章斷句/中文斷詞/清除停用詞

文章斷句 ( 24 )

參數設定

Input - 2

統計資訊

1515

原始數量

1514

切割總數

1

最大切割數

1

最小切割數

1.00

平均切割數

任務結果

Show 10 entries

Search:

system_id	sentence_idx	sentence
1	1	作者: tobetwob (WHO FCTC 菸草減害專家) 看板: Tobacco 標題: [NEWS] 台酒無法銷售電子菸流失業務量 莊翠雲： 時間: Thu Jan 2 16:59:25 20...
2	1	1.政權交接：川普總統回歸，人事大調動，美國駐日代表已更換 2.阿北大嘴巴：2024選舉時，就扯一堆AIT，大嘴巴，誰都怕跟他有關係 3.川黑民眾堂：阿北罵川普總統是瘋子，把川普總統看細漢，黃國昌嘲笑...

全螢幕瀏覽

點我下載完整CSV資料

點我下載完整Rdata

點我下載完整json資料

選擇處理欄位 \*

result

斷句符號

/n

使用默認符號進行斷句 \*

否



# 替換字串/文章斷句/中文斷詞/清除停用詞

中文斷詞 (3)

參數設定

## 統計資訊

4687

最大字數



510888

總字數



20

最小字數



337

平均字數



## 任務結果

Show 10 entries

Search:

system_id	sentence_idx	result
1	1	[作者, :, tobetwob, (, WHO, FCTC, 菸草, 減害, 專家, ), 看板, :, Tobacco, 標題, :, [, NEWS, ], 台酒, 無法, 銷售, 電子菸, 流失, 業務量, 莊翠雲, :, 時間, :, Thu, Jan, 2, 16, :, 59, :, 25, 2025, 台酒, 無法, 銷售, 電子菸, 流失, 業務量, 莊翠雲, :, 將, 加強, 和, 衛福部, 溝通, 2025, -, 01, -, 02, 13, :, 32, 聯合報, /, 記者, 朱漢, 崙, /, 台北, 即時, 報導, 台灣, 菸酒, 董座, 湯期, 安在, 今日, 財委會, 上, 答詢, 時, 指出, ,, 配合, 政策, 明年, 將, 要求, 所有, 投放, 商, ,, 若, 沒, 有, 實名制, ,, 就, 避免, 去, 投放, 廣告, ,, 他, ...]
2	1	[1, 政權, 交接, :, 川普, 總統, 回歸, ,, 人事, 大, 調動, ,, 美國, 駐日, 代表, 已, 更換, 2, 阿北, 大嘴巴, :, 2024, 選舉, 時, ,, 就, 扯, 一堆, AIT, ,, 大嘴巴, ,, 誰, 都, 怕, 跟, 他, 有, 關係, 3, 川, 黑, 民眾, 堂, :, 阿北, 罵, 川普, 總統, 是, 瘋子, ,, 把, 川普, 總統, 看, 細漢, ,, 黃國昌, 嘲笑, 2020, 美選, 開票, ,, 雖然, 針對, 政黑, 前, 版主, Kero, ,, 但是, 誰, 看, 了, 都, 知道, 現任, 民眾, 堂, 主席, 黃國昌, 先生, 是, 個川, 黑, ,, 這些, AIT, 都, 紀錄, 在, 案, 還有, 很多, 例子, ,, 反正, 阿北, 7414, ...]

參數設定

Input - 24

選擇處理欄位 \*

sentence

定義詞彙 ⓘ

美國總統 1000  
美國選舉 1000  
美國大選 1000  
澤倫斯基 500  
哈瑪斯 500

# 替換字串/文章斷句/中文斷詞/清除停用詞

## result

[作者, tobetwob, who, fctc, 菸草, 減害, 專家, 看板, tobacco, 標題, news, 台酒, 無法, 銷售, 電子菸, 流失, 業務量, 莊翠雲, 時間, thu, jan, 台酒, 菸, 流失, 業務量, 莊翠雲, 加強, 衛福部, 溝通, 聯合報, 記者, 朱漢, 台北, 即時, 報導, 台灣, 菸酒, 董座, 湯期, 安在, 今日, 財委會, 答詢, 指出, 配, 求, 投放, 實名制, 避免, 投放, 廣告, 坦言, 菸害, 防制法, 上路, 之後, 民眾, 轉抽, 電子, 加熱, 台酒, 無法, 銷售, 加熱, 流失, 營運量, 對此, 莊翠雲, 強和, 衛福部, 溝通, 民眾, 黨籍, 立委, 黃珊珊, 質疑, 大陸, 進行, 低價, 啤酒, 傾銷, 國人, 喝到, 劣質, 啤酒, 自知, 莊翠雲, 指出, 透過, 釀酒, 協會]

[政權, 交接, 川普, 總統, 回歸, 人事, 調動, 美國, 駐日, 代表, 更換, 阿北, 大嘴巴, 選舉, 一堆, ait, 大嘴巴, 關係, 民眾, 阿北, 川普, 總統, 瘋子, 川, 國昌, 嘲笑, 美選, 開票, 政黑, 版主, kero, 知道, 現任, 民眾, 主席, 黃國昌, 先生, 個川, ait, 紀錄, 例子, 反正, 阿北, sent, from, meowptt, on,

[柯建銘, 先衝, 雙罷, 黨政, 高層, 微詞, 記者, 總統, 賴清德, 元旦, 表態, 一月, 將邀, 立法院長, 韓國, 總統府, 和解, 咖啡, 立法院, 民進黨團, 總召, 出, 罷免, 立法院, 副院長, 國民黨, 四十一席, 區域, 立委, 指出, 府院, 立院, 黨團, 討論, 形成, 共識, 主張, 黨政, 高層, 感佩, 柯建銘, 奮戰, 苦功, 理解, 一線, 壓力, 柯未經, 黨內, 高層, 會議, 黨團, 會議, 共識, 決就, 先衝, 黨內, 異音, 黨政, 高層, 十分, 為難, 微詞, 據了解, 國民黨, 強勢, 主導, 三大, 爭議, 修法, 黨政, 高層, 考量, 啟動, 協調, 行政院長, 卓榮泰, 處理, 法案, 主動, 聯繫, 溝通, 最後, 藍營, 強行, 法案, 黨政, 高層, 認為, 做不, 礎, 這項, 方案, ...]

[新聞, 網址, 新聞, 來源, 自由時報, 完整, 新聞, 標題, 風向, 川普, 就職, 中國, 網路, 對美, 義和團, 言論, 一夕, 退場, 完整, 新聞, 內容, 編譯, 孫宇, 普, 就任, 美國總統, 日期, 逼近, 中國, 風向, 急遽, 轉變, 自由, 亞洲, 電台, rfa, 報導, 去年, 底美, 國駐, 使館, 微博, 發文, 充斥, 美國, 評論, 元旦, 關係, 內容, 學者, 分析, 中國, 政府, 在內, 外交, 困下, 擺出, 和緩, 姿態, 報導, 美國, 駐中, 大使館, 元旦, 微博, 發出, 回顧, 美中, 建交, 的雙, 關係, 貼, 友, 評論, 言論, 一掃, 往日, 戾氣, 出現, 中美, 關係, 人類史, 偉大, 雙邊, 關係, 美利堅, 中華民族, 帶來, 史性, 福祉, 世界, 帶來, 前所未有, 深遠, ...]

### 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

億元

### 語言 \*

Chinese

### 是否清除單字元 ⓘ

是

### 清除英文字母 \*

否

### 清除換行符號 \*

是

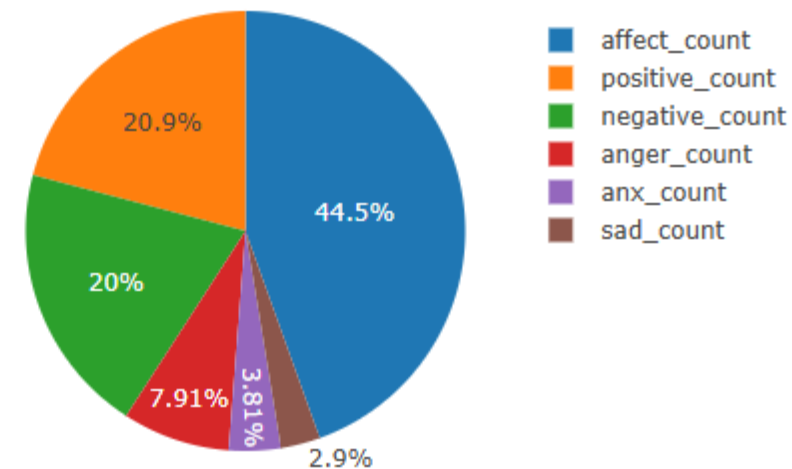
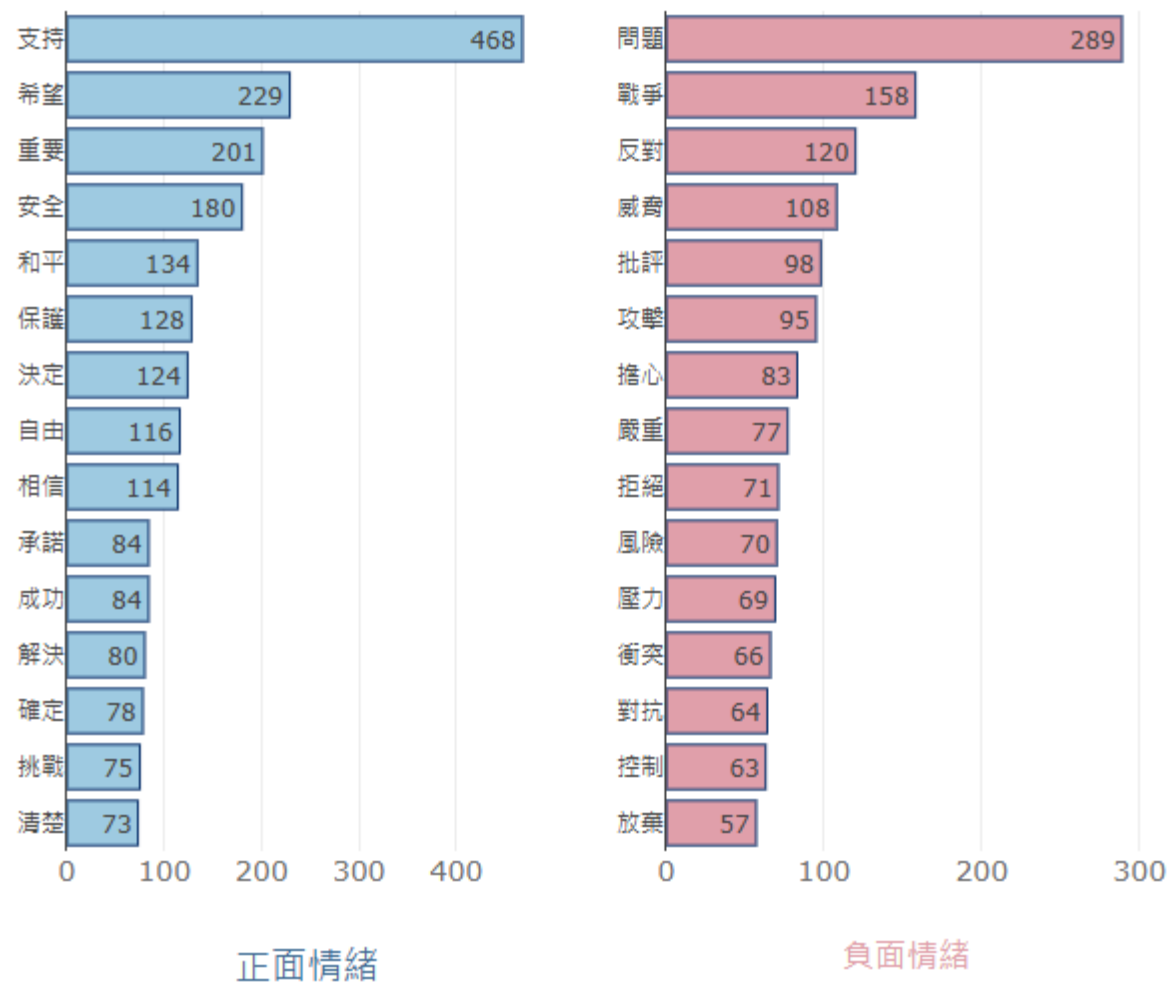
### 清除html tag \*

是

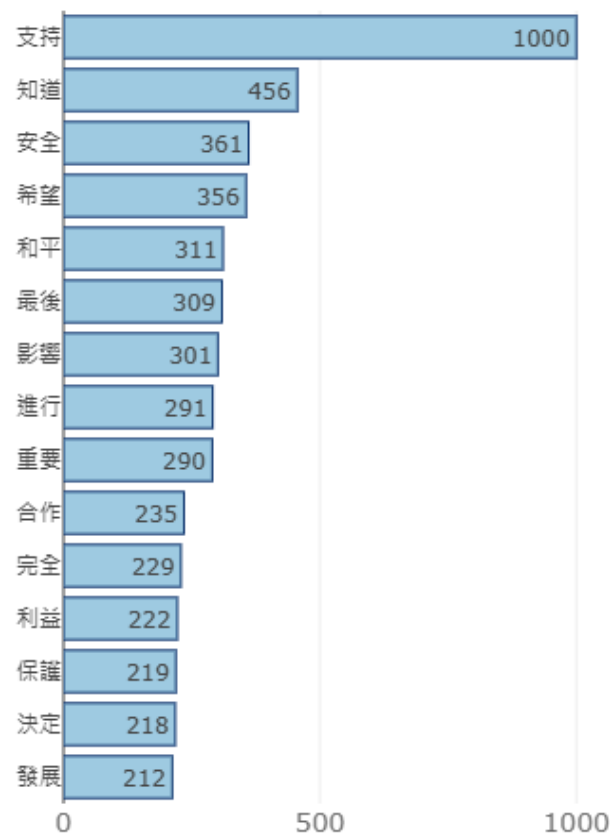
# 詞頻分析



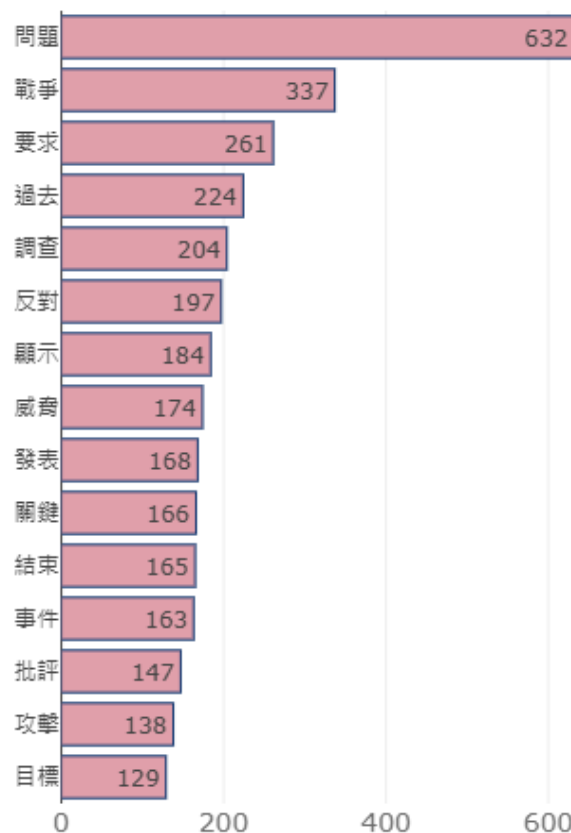
# 情緒分析:LIWC情緒分析



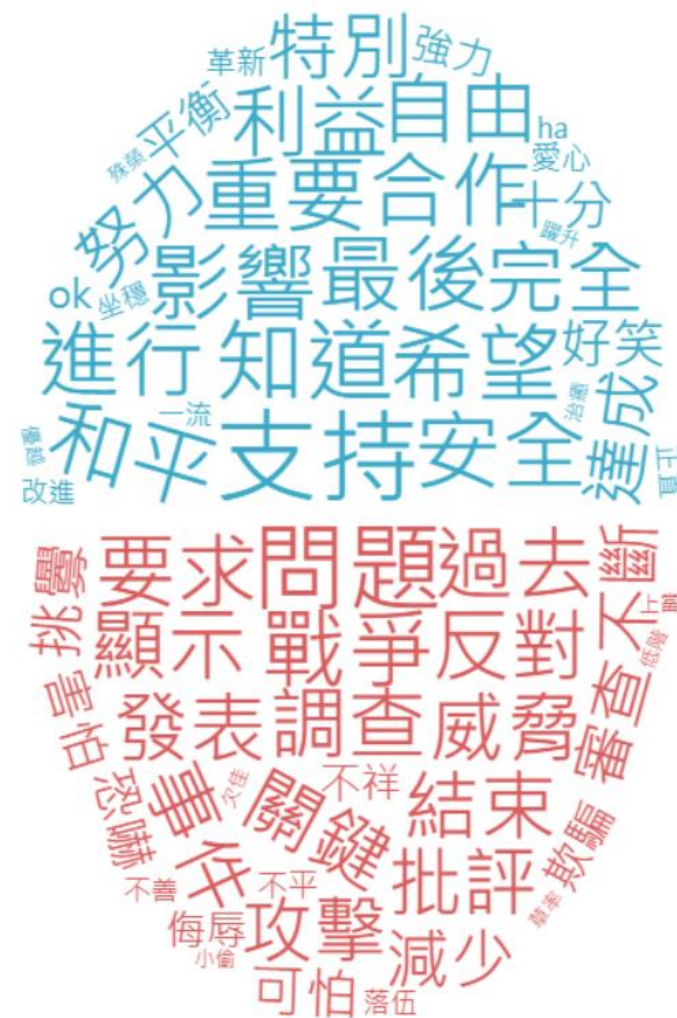
# 情緒分析: LEXICON BASED



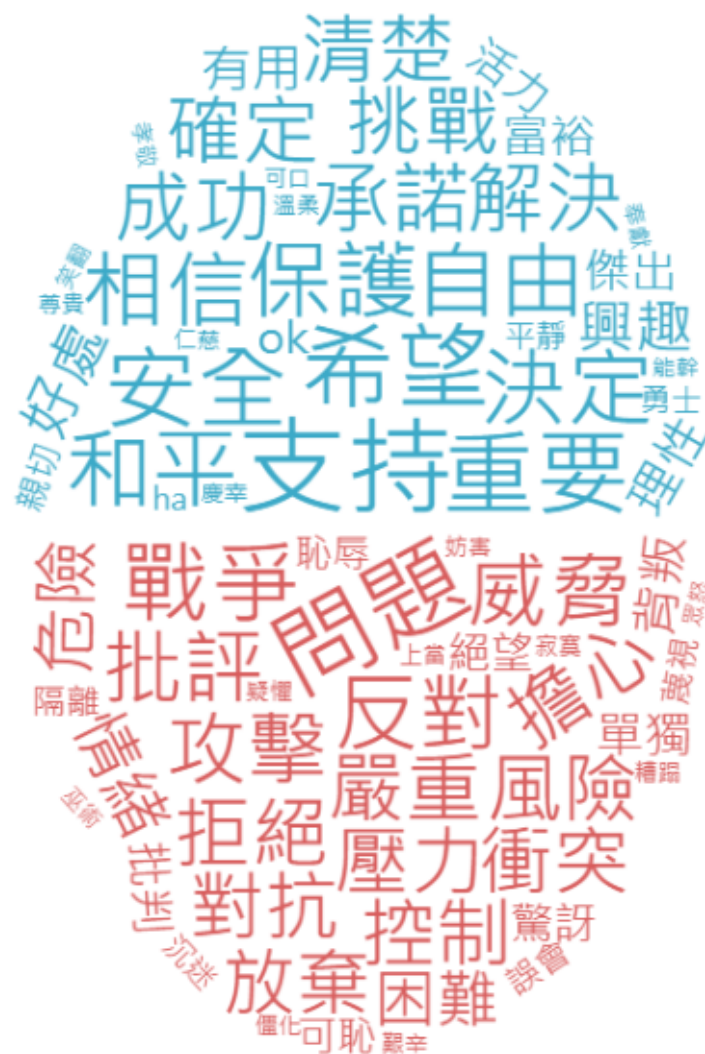
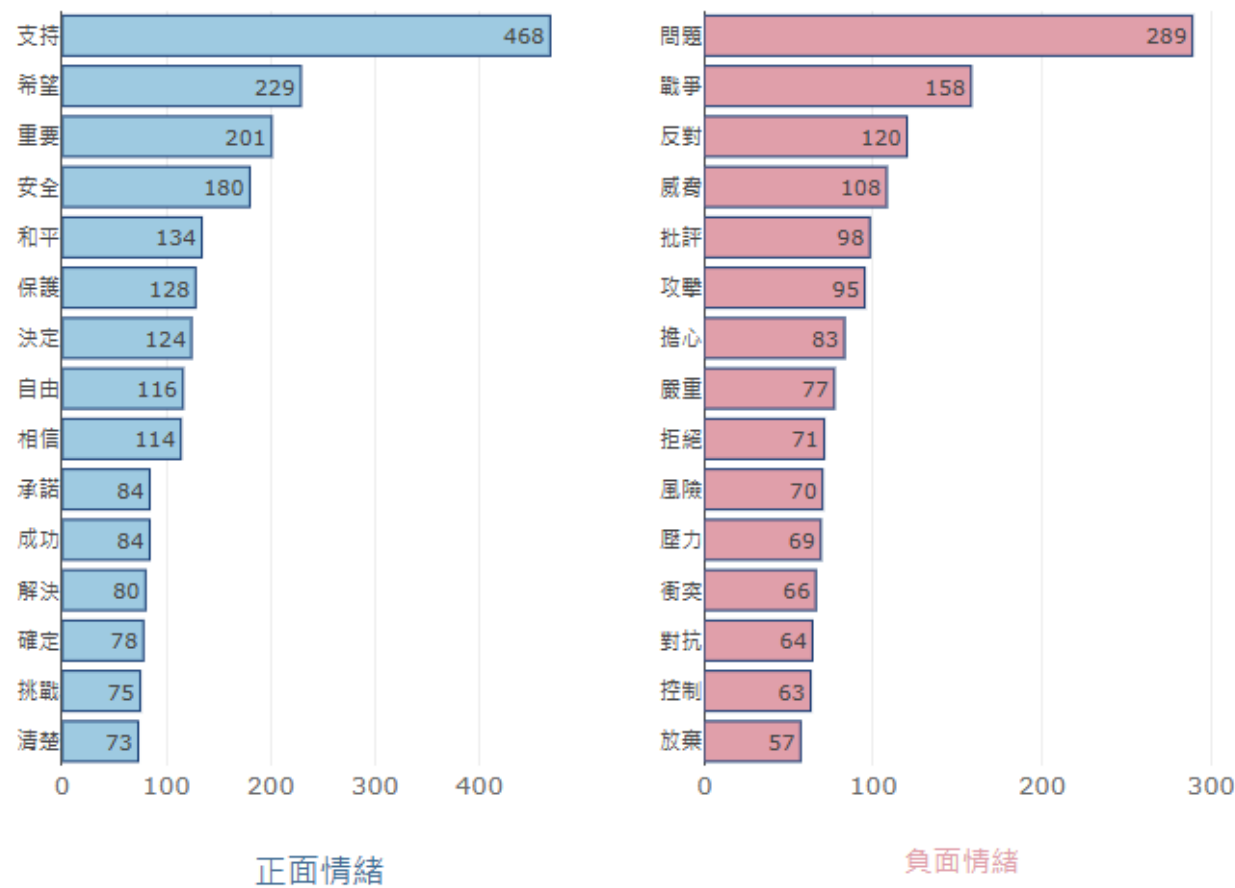
正面情緒



負面情緒

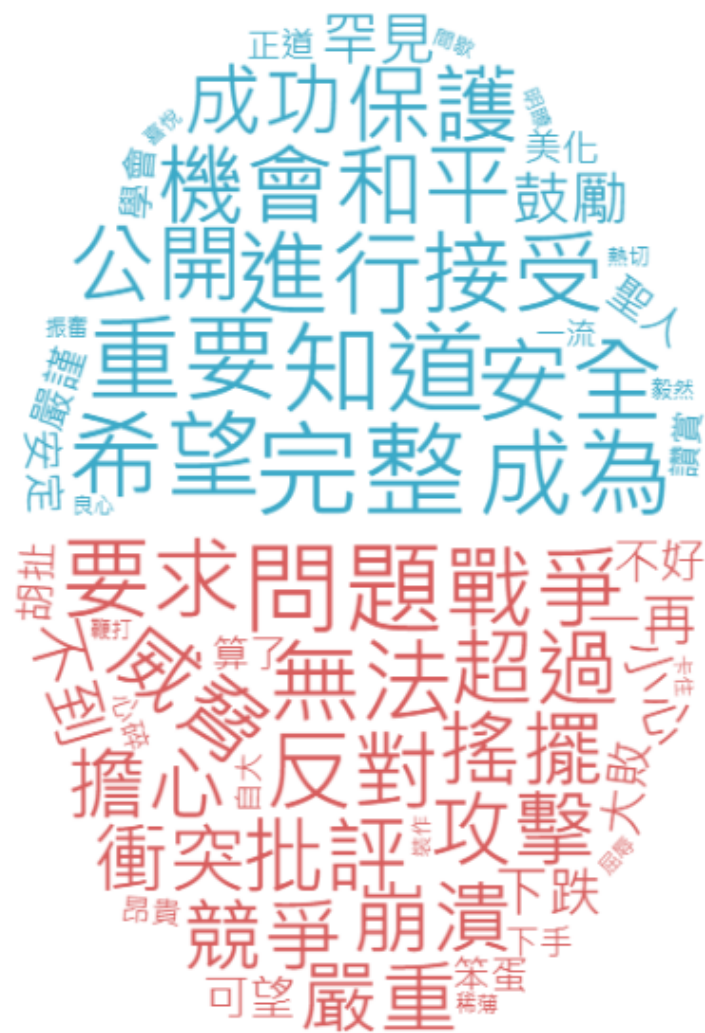
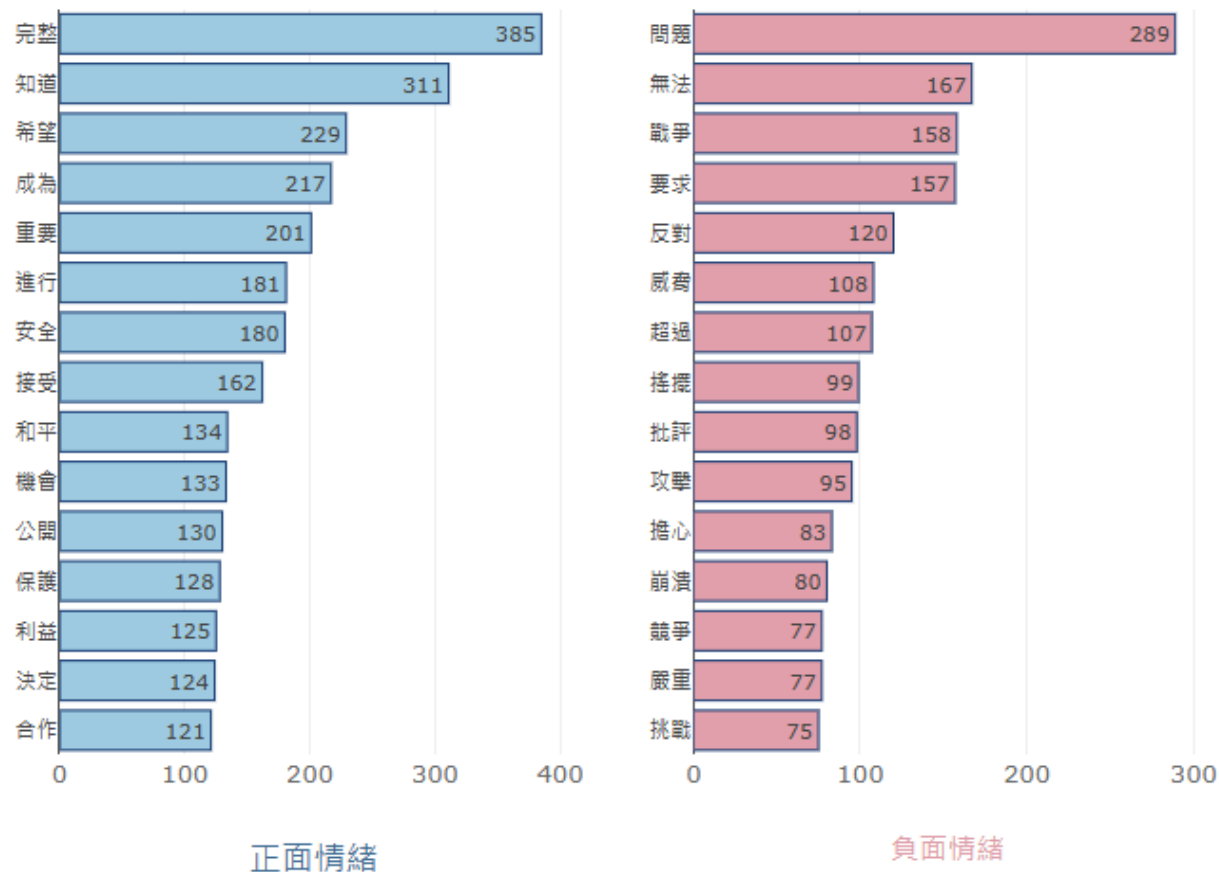


# 情緒分析:進階字典法

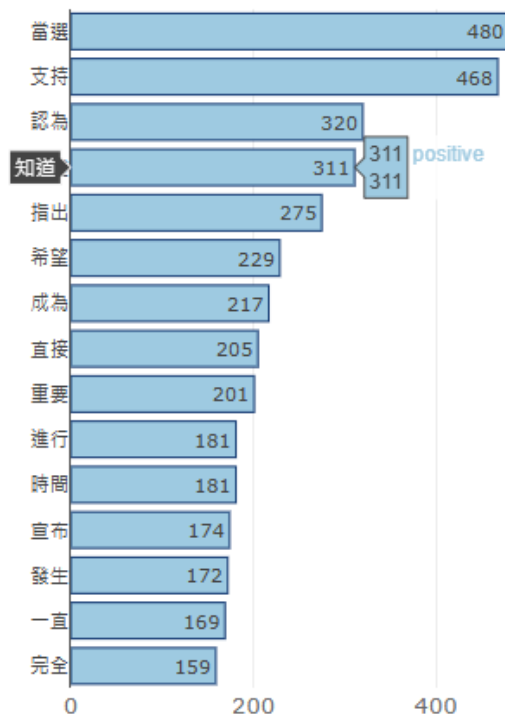




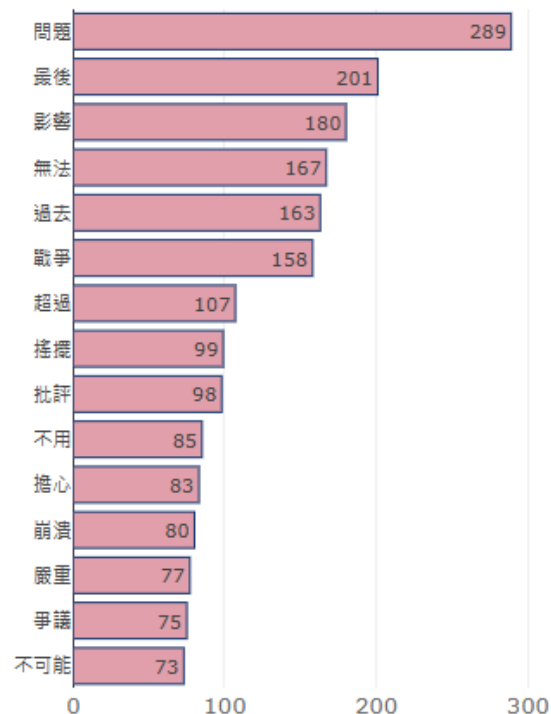
# 情緒分析:NTUSD



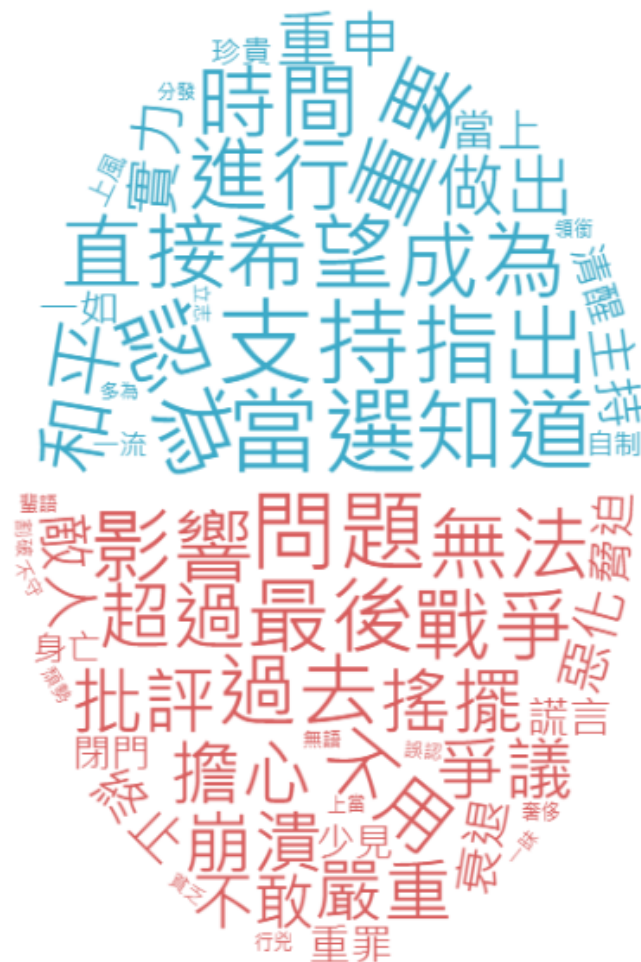
# 情緒分析:ANTUSD



正面情緒



負面情緒





## 遇到的困難2:沒有獨立思考的資料篩選

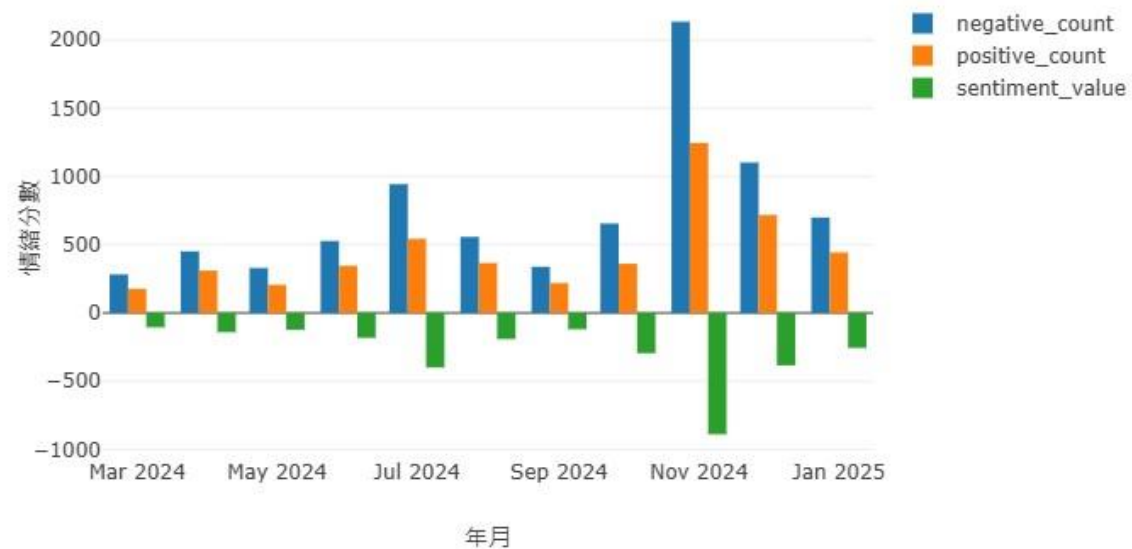
資料篩選 (21) 參數有做更動，建議重新執行

參數設定

條件式 \* ⓘ  
sentiment\_value<0

儲存更改

彙總結果分群長條圖



# 合併資料/欄位重構/分群彙總

JOIN規則

新增規則

刪除規則

任務一欄位

system\_id

任務二欄位

system\_id

欄位重構(Melt) ( 63 )

參數設定

要轉換的欄位(按住ctrl(Windows)或command(MAC)可以複選) \*

sentence\_idx

positive\_count

positive\_words

negative\_count

negative\_words

sentiment\_value

轉換的欄位新的欄位名稱 \*

variable

分群彙總(日期) ( 8 )

參數設定

選擇日期欄位 \*

artDate

選擇日期類型 \*

月

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

system\_id

artDate

variable

value

日期格式 \*

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

匯總函數 \*

sum

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

system\_id

artDate

variable

value

Input - 7

任務結果

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

e\_ip

artCategory

dataSource

insertedDate

artComment

artDate

值的欄位名稱 \*

value

# 視覺儀錶板

## 2024 美國總統大選重大事件時間軸

日期	事件描述
2024/3/12	川普鎖定提名：獲共和黨總統候選人資格。
2024/6/27	首場辯論：拜登表現受批評，健康受質疑。
2024/7/13	川普遭槍擊未遂：賓州集會遇襲，右耳受傷。
2024/7/15	共和黨提名：川普獲提名，J.D.萬斯任副手。
2024/7/21	拜登退選：支持賀錦麗接棒競選。
2024/7/22	賀錦麗獲提名：民主黨正式提名為總統候選人。
2024/9/10	第二場辯論：川普與賀錦麗於費城交鋒。
2024/9/15	川普再遭刺殺未遂：佛州遇襲，嫌犯被捕。
2024/9/24	賀錦麗辦公室遭槍擊：亞利桑那州發生槍擊，無人傷亡。
2024/11/5	大選日：川普勝出，獲 312 張選舉人票。
2025/1/20	川普就職：華盛頓宣誓，展開第二任期。



彙總結果分群長條圖

