**烏俄戰爭輿論觀察**

### **動機**

近年來受到烏俄戰爭影響，全球經濟受到重挫，能源價格飆升是導致通膨的最大原因，各國央行透過激進的加息政策來壓低通膨，導致許多民眾的購買力嚴重下降，目前烏俄戰爭近期已滿一周年，且戰爭並沒有緩和的跡象，即使是離戰區很遠的台灣，也受到戰爭的影響，政府也將兵役延長，台灣民眾開始對於戰爭發展的情勢感到憂心，例如對岸是否會打過來？若發生戰爭的話美軍真的會過來協防嗎？各國對台灣的態度...等等。故本研究將以烏俄戰爭作為本次第一次讀書會報告主題，了解民眾對於烏俄戰爭的看法。

### **資料範圍、來源、關鍵字**

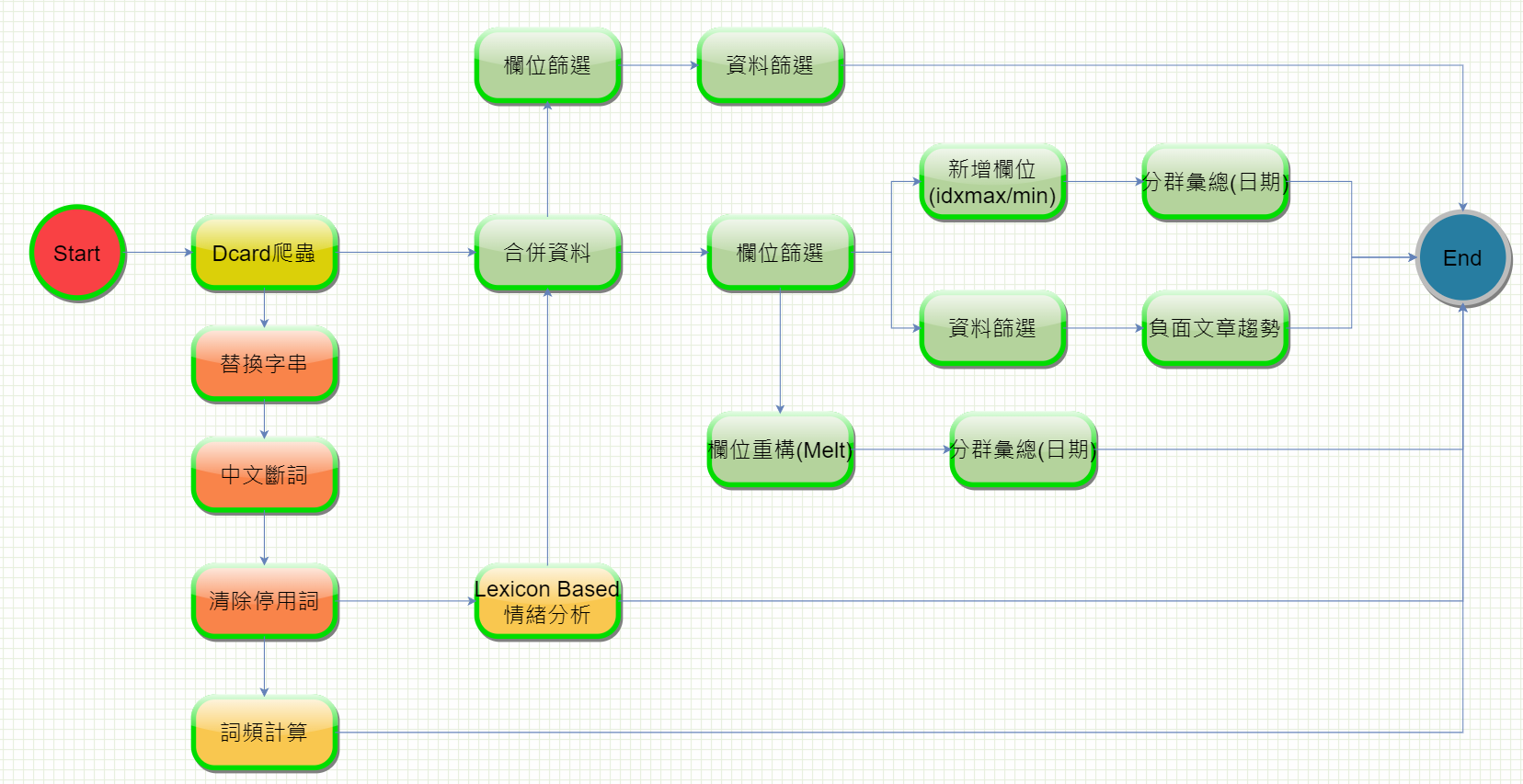
資料來源：Dcard時事版、閒聊版

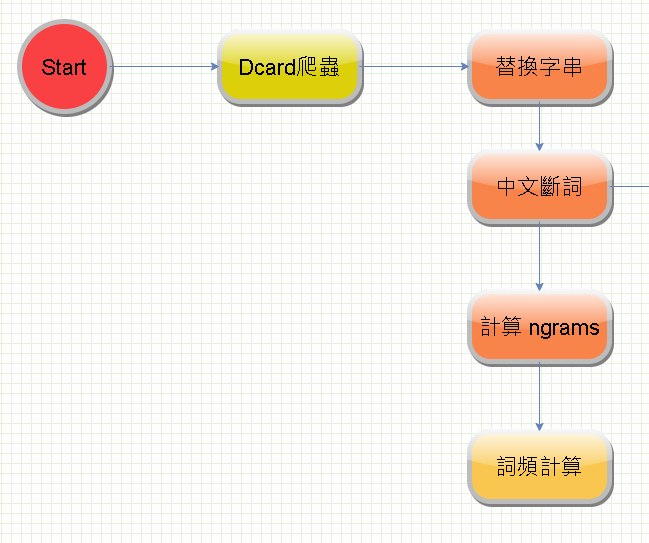
資料範圍：2022/2/24~2023/2/28

搜尋關鍵字：烏俄戰爭、俄烏戰爭、俄羅斯、烏克蘭、戰爭、基輔、澤倫斯基、普丁

### **系統流程**

流程名稱:Test\_01





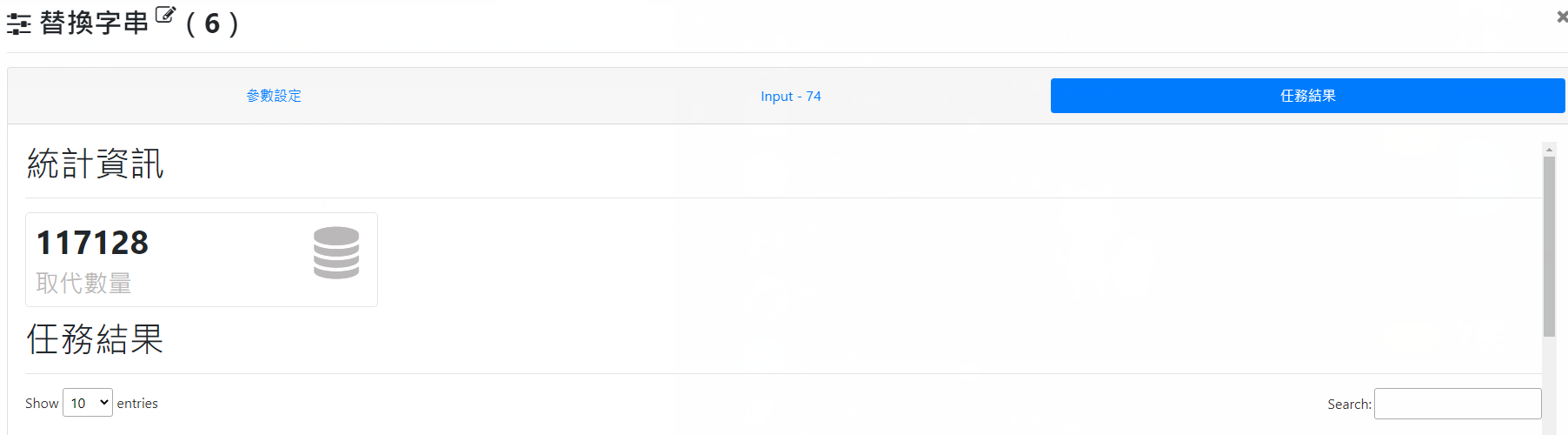
### **分析過程**

#### **替換字串**

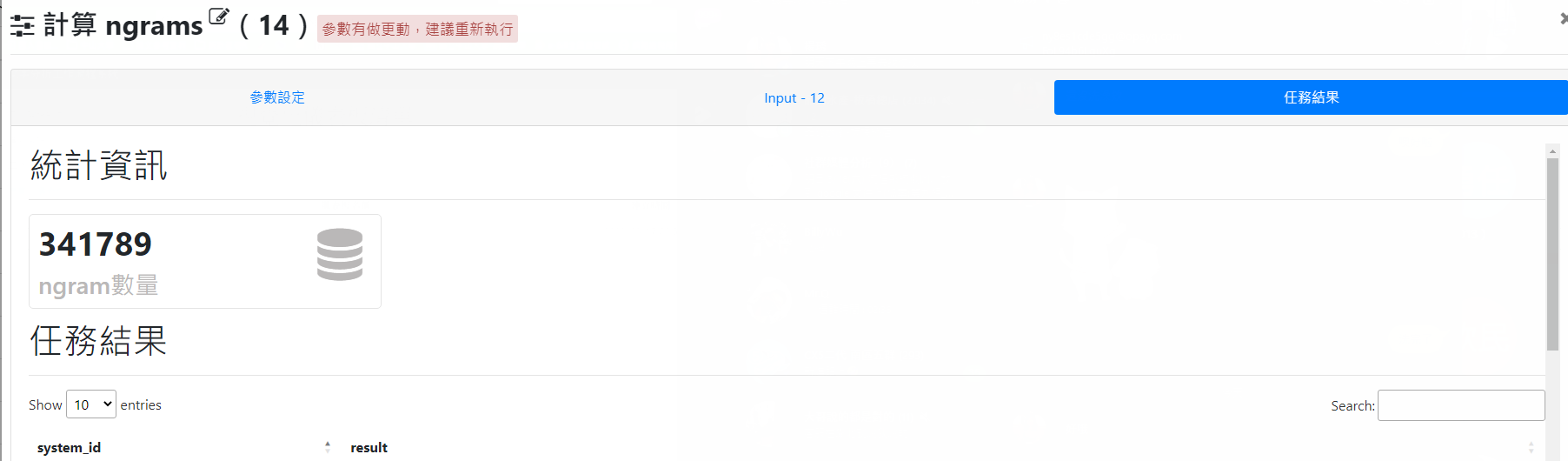
清理文章內的網址、HTML符號等，共清理了117128筆。

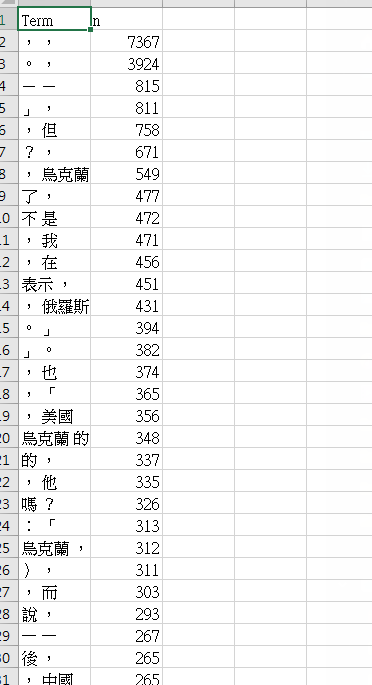


替換字串-任務結果



#### **計算 ngrams**

使用 ngrams 找出重複性高的字元並當作後續的中文斷詞。



#### **中文斷詞設定**

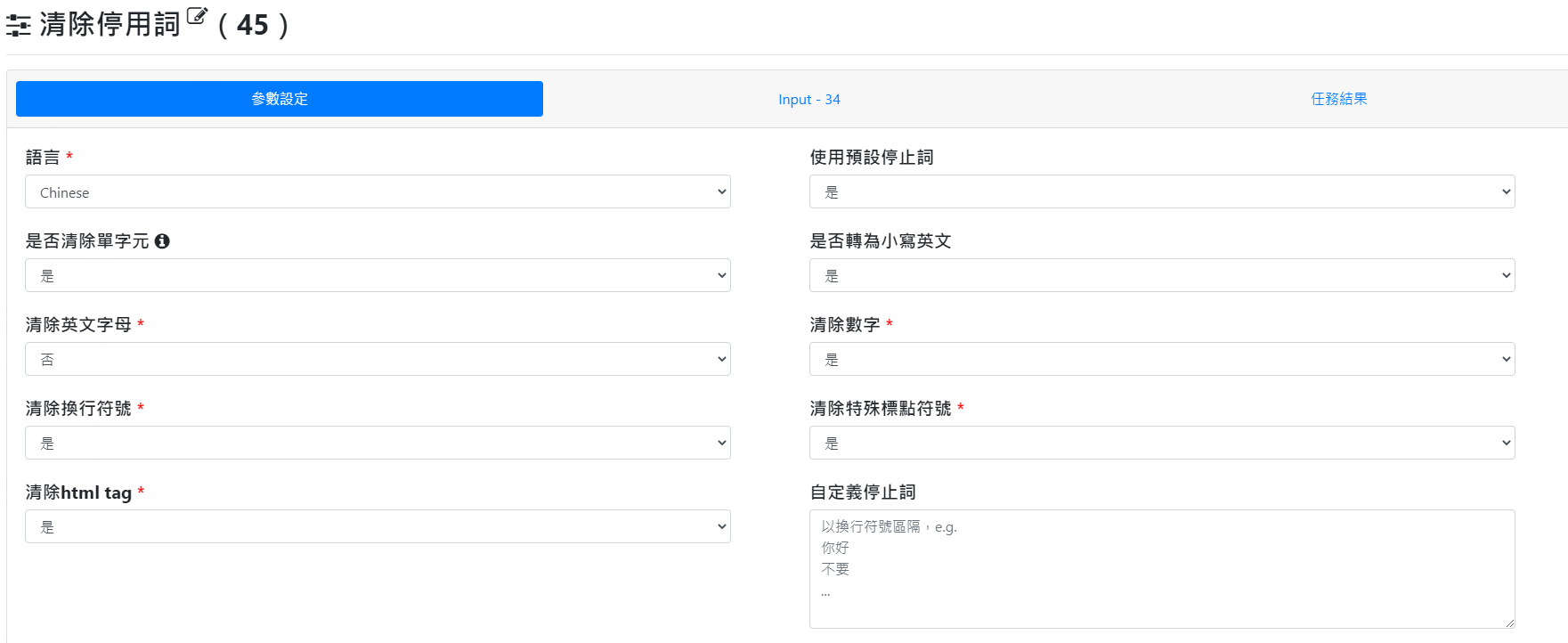
將上一步使用 ngrams 產出的結果輸入到中文斷詞詞彙，並把此次議題烏俄戰爭常用詞語設定為斷詞幫助分析。





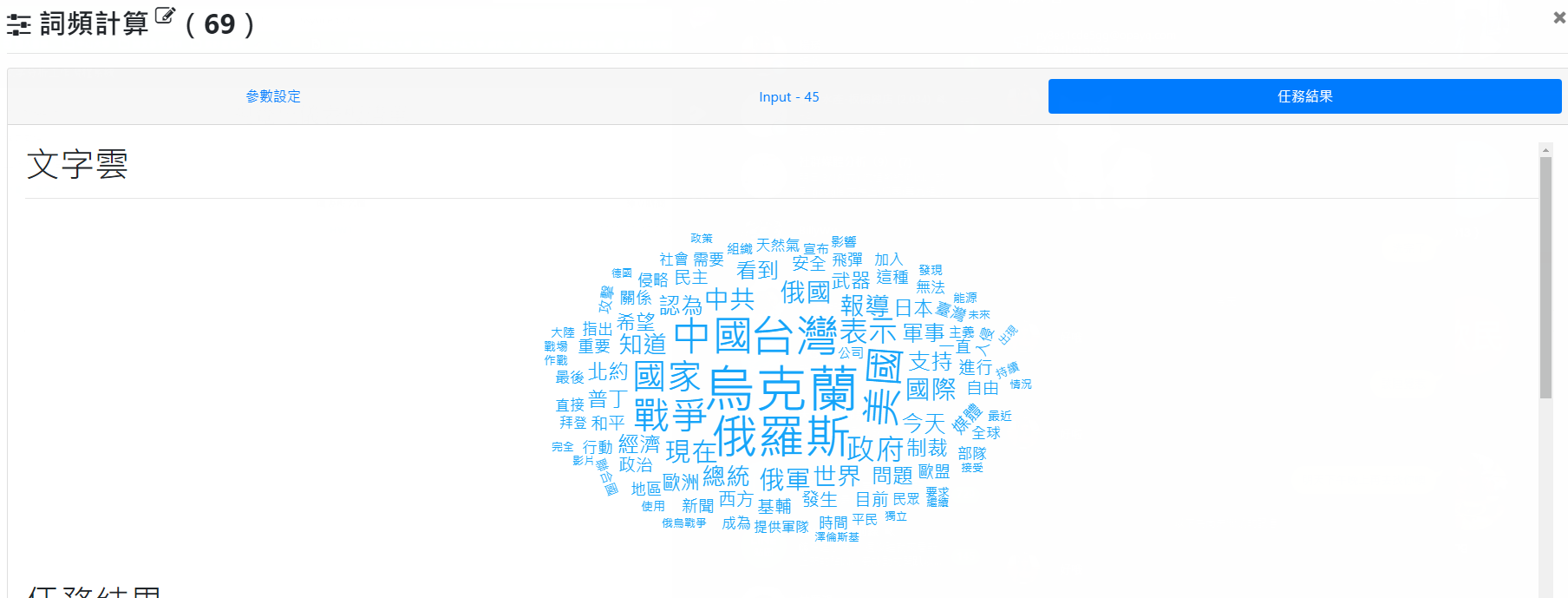
#### **清除停用字**

加入停用字清洗資料減少雜訊。



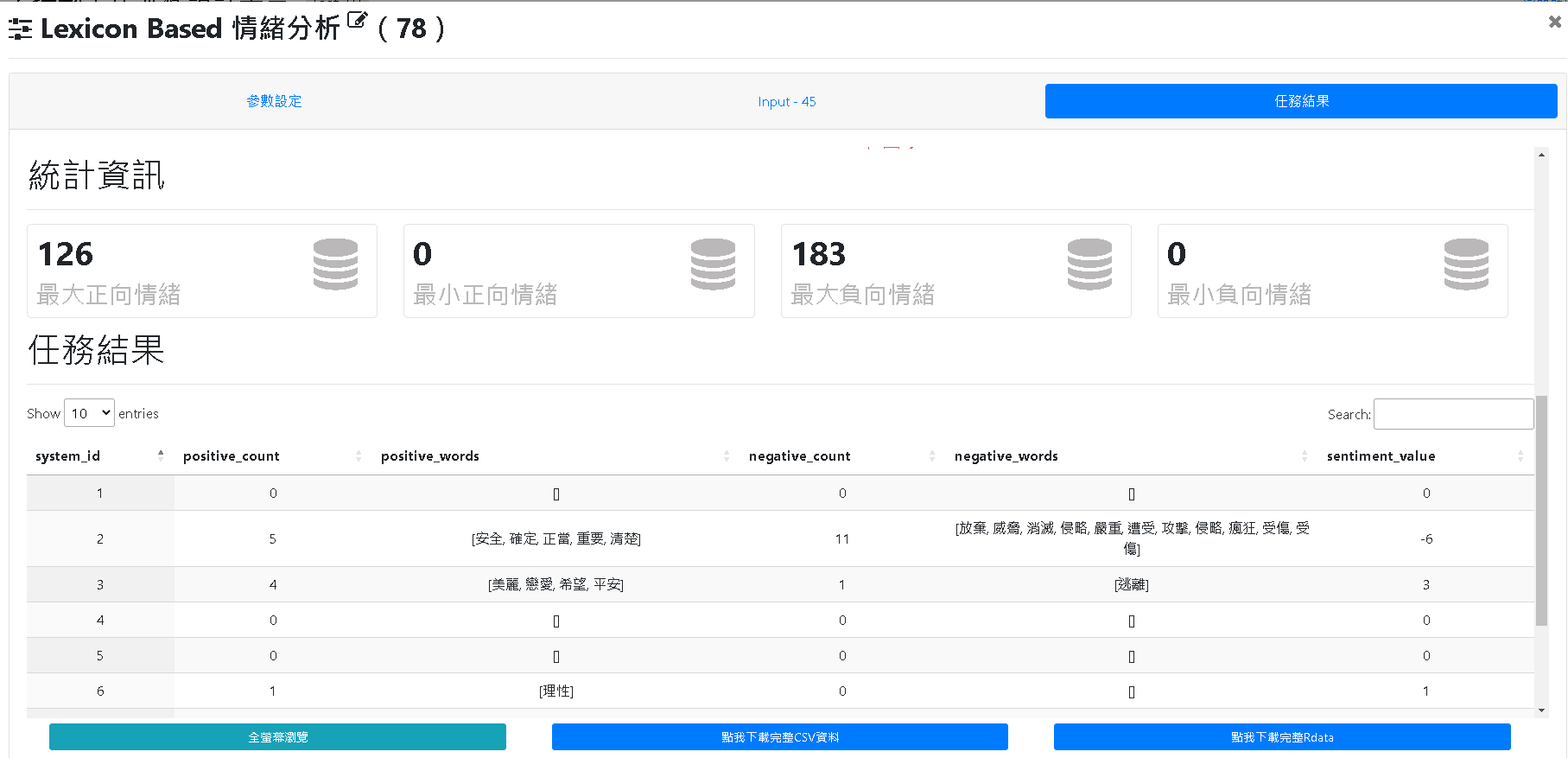
#### **詞頻計算**

產出文字雲觀看結果是否符合主題，若有不符合主題的字詞則重復上一步加入停用詞反覆清洗。

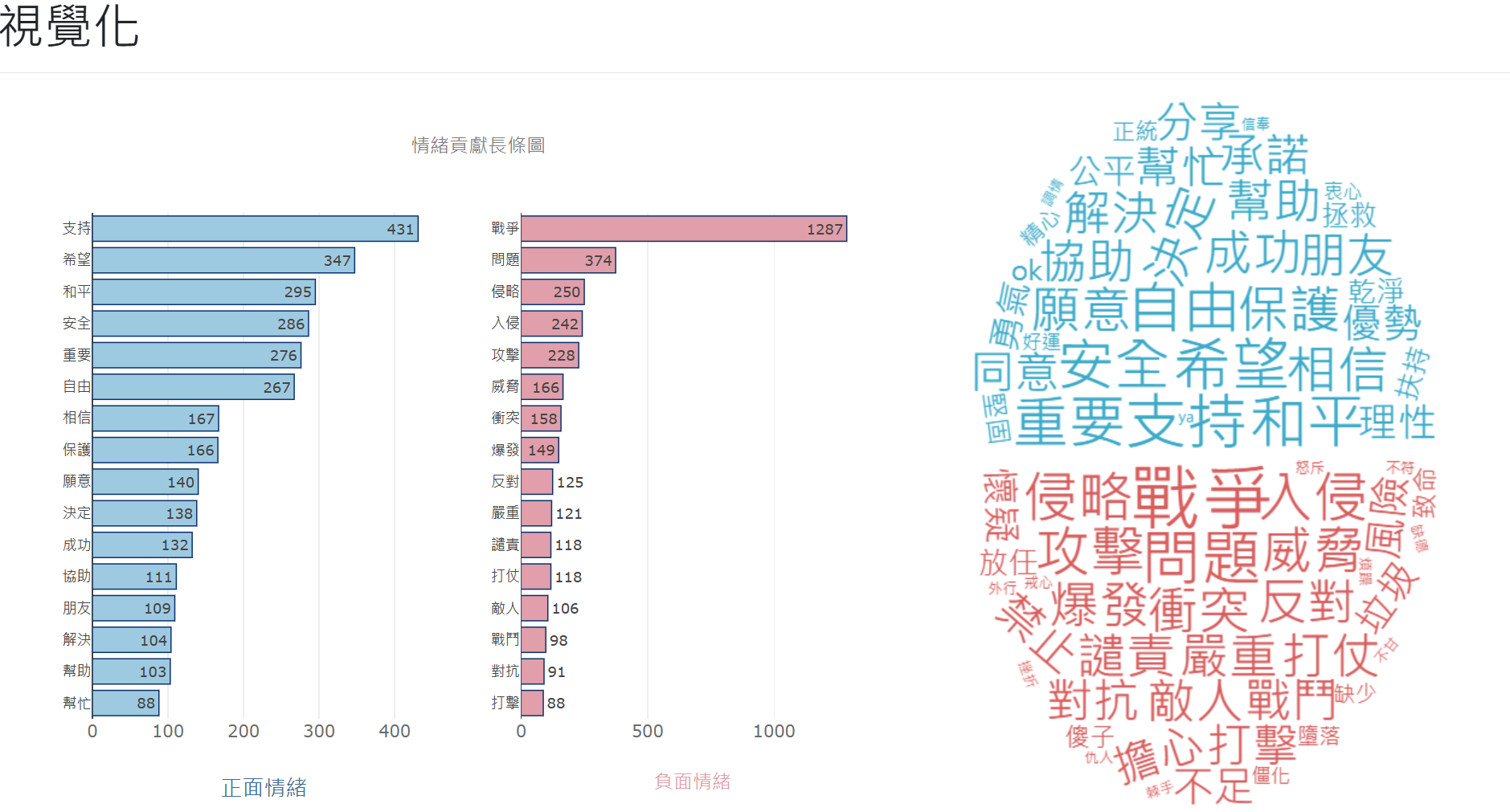


#### **Lexicon Based情緒分析**

一篇文章裡最大正向情緒為 126 個，負向情緒為 183 個。

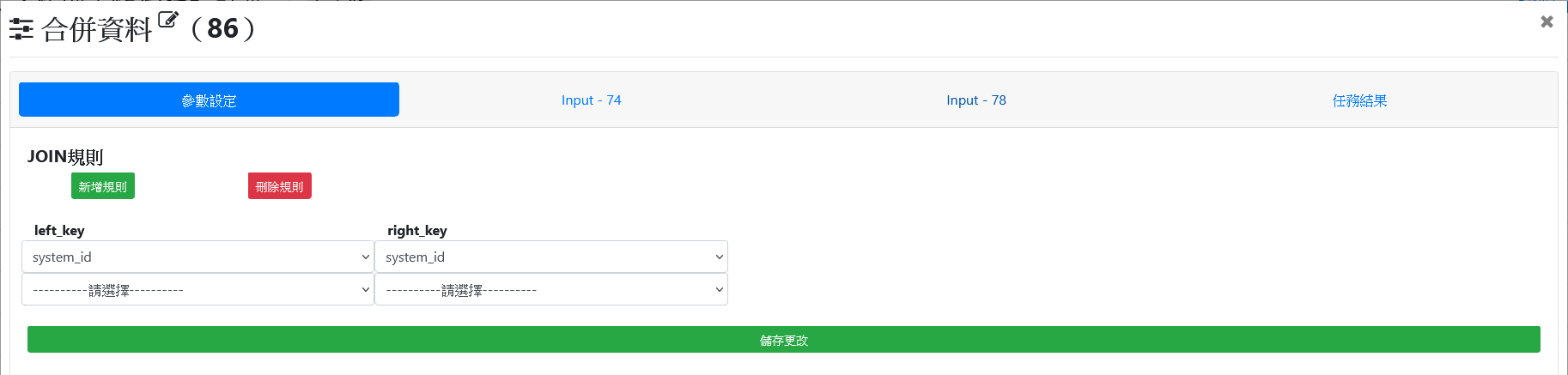


在經由視覺化呈現，正面情緒詞主要包含了支持、希望、安全、和平...等，而負面情緒詞主要包含了戰爭、問題、侵略、攻擊...等。

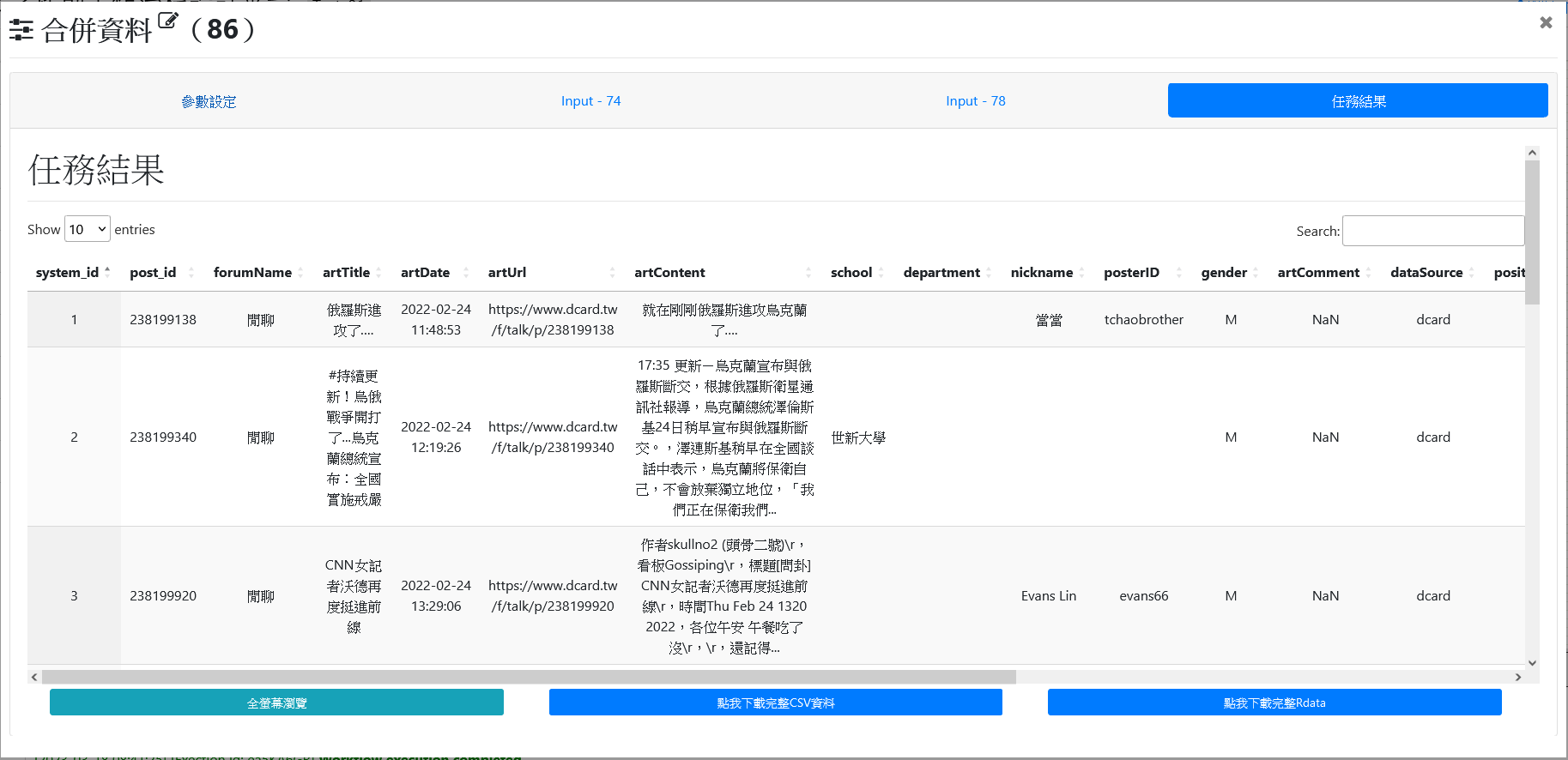


#### **合併資料**

欲進行正負面情緒分析，使用 system\_id 做為 key 值，將Lexicon Based情緒分析的結果與 Dcard 原始文章進行 join 來做後續分析使用。



合併資料任務結果



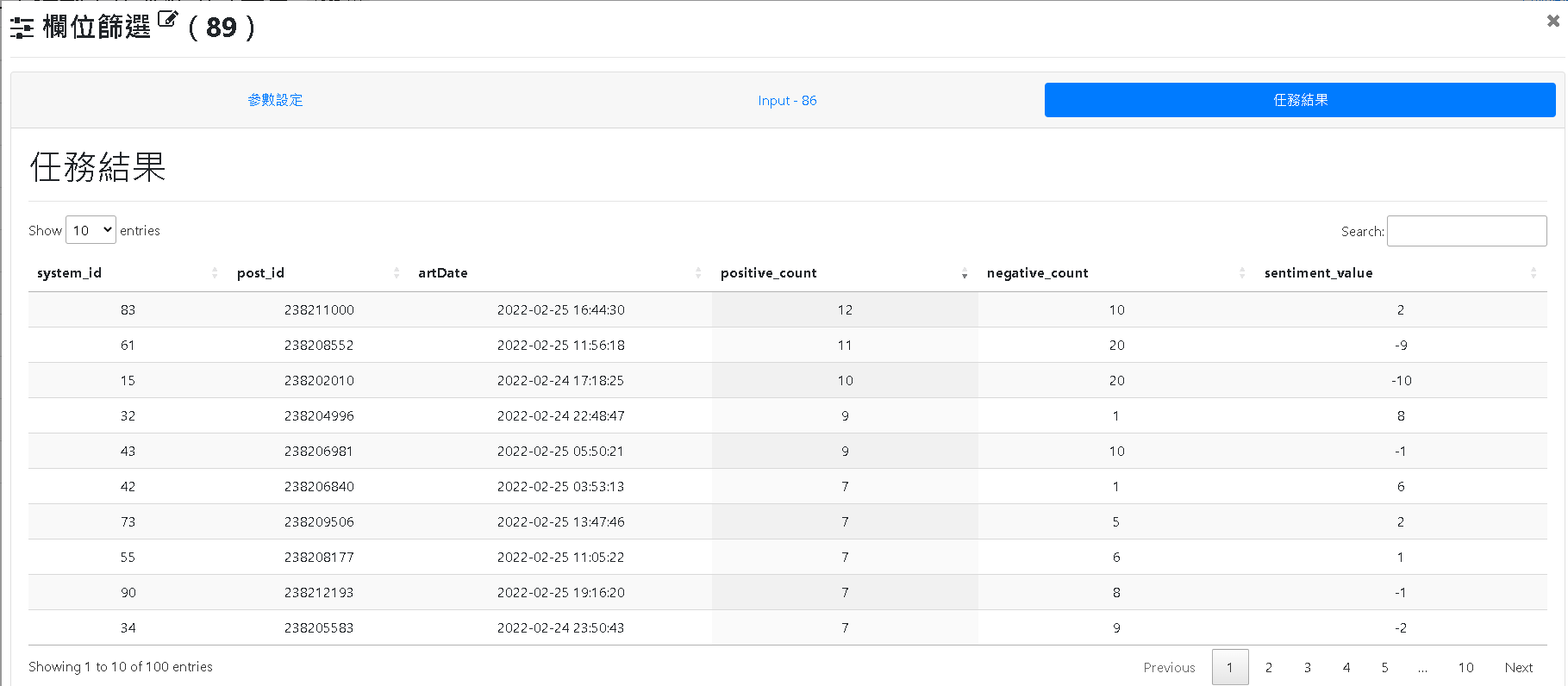
#### **欄位篩選**

選擇 artDate、positive\_count、negative\_count、sentiment\_valuen 四個欄位。



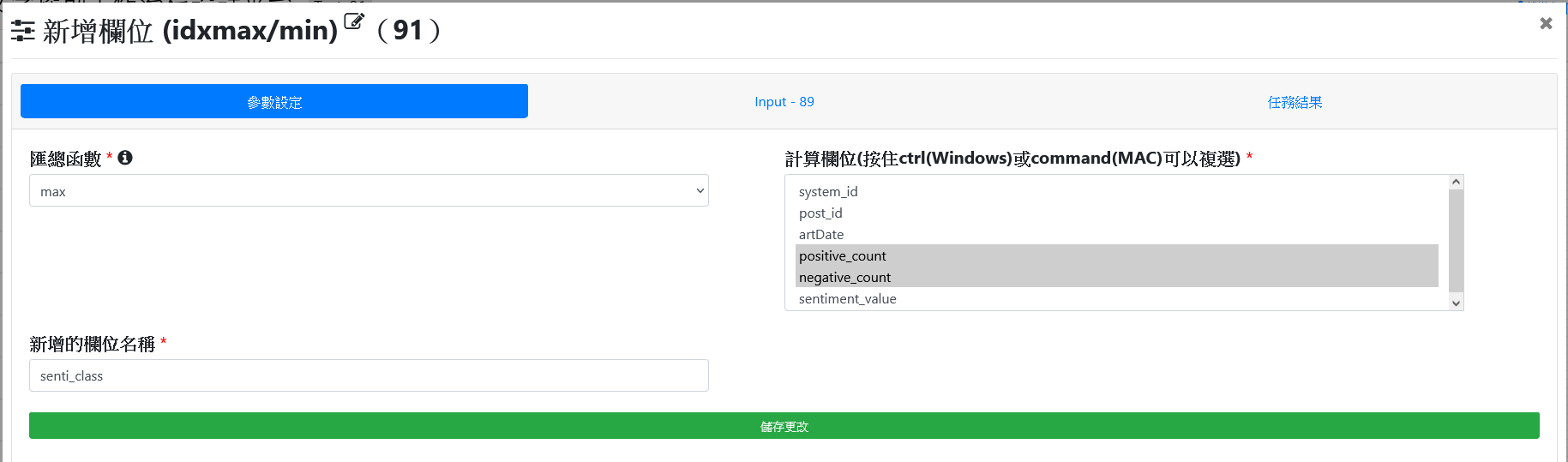
**欄位篩選-任務結果**

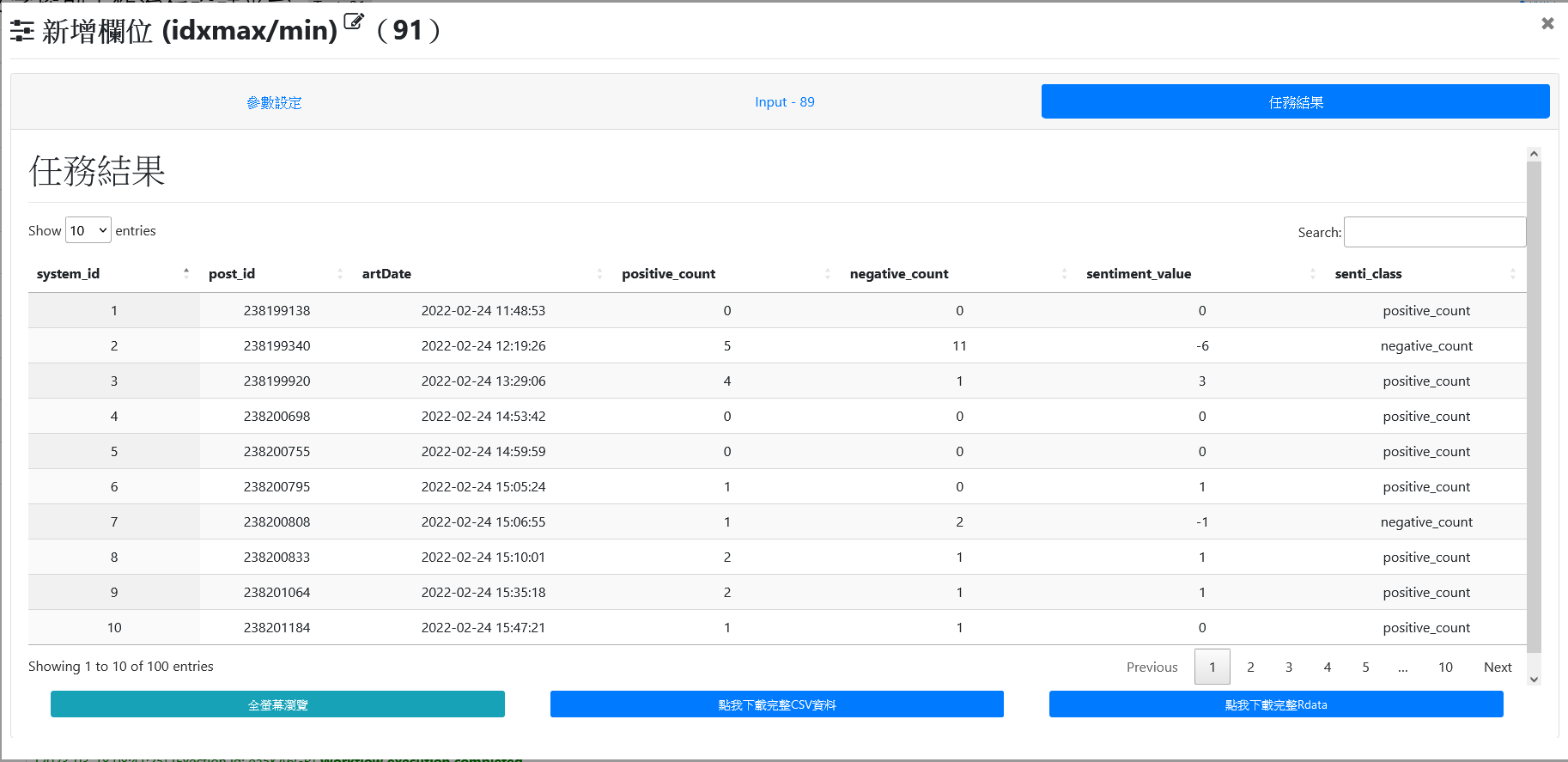
如圖為欄位篩選結果，若小於0則代表為負面情緒。



#### **新增欄位**

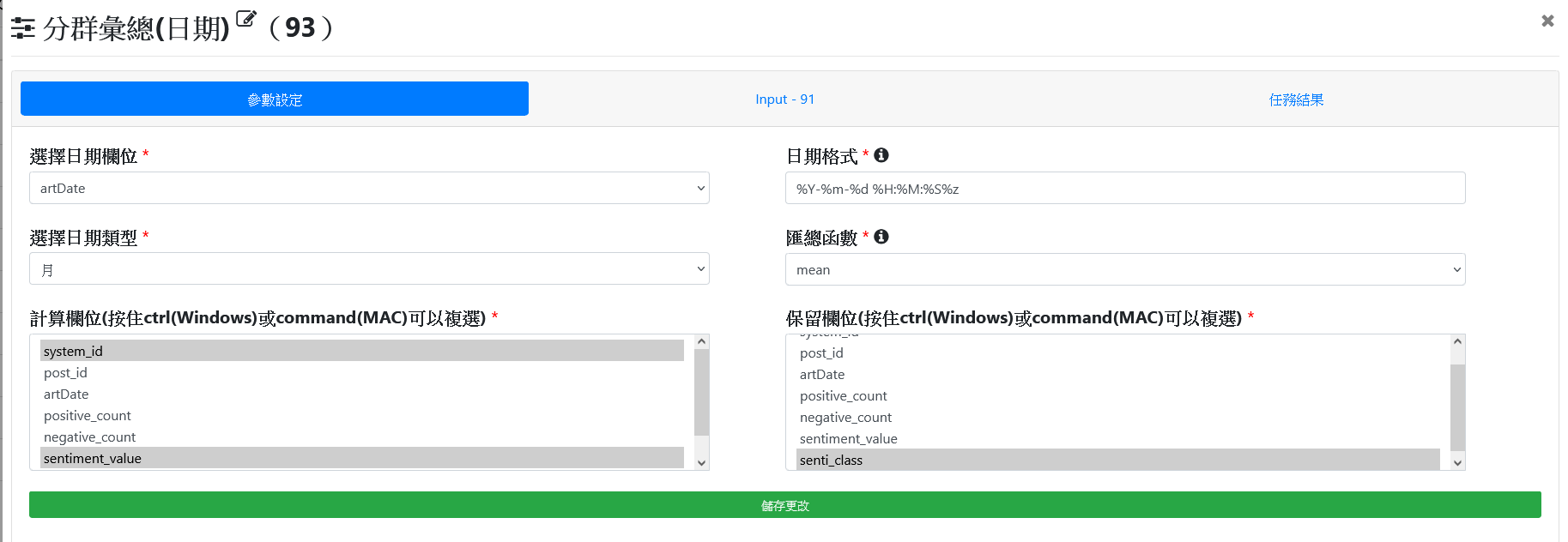
新增一欄位紀錄，將每篇文章的情緒歸類為正面或是反面情緒，再依此得出每月正反向各別的情緒分數。

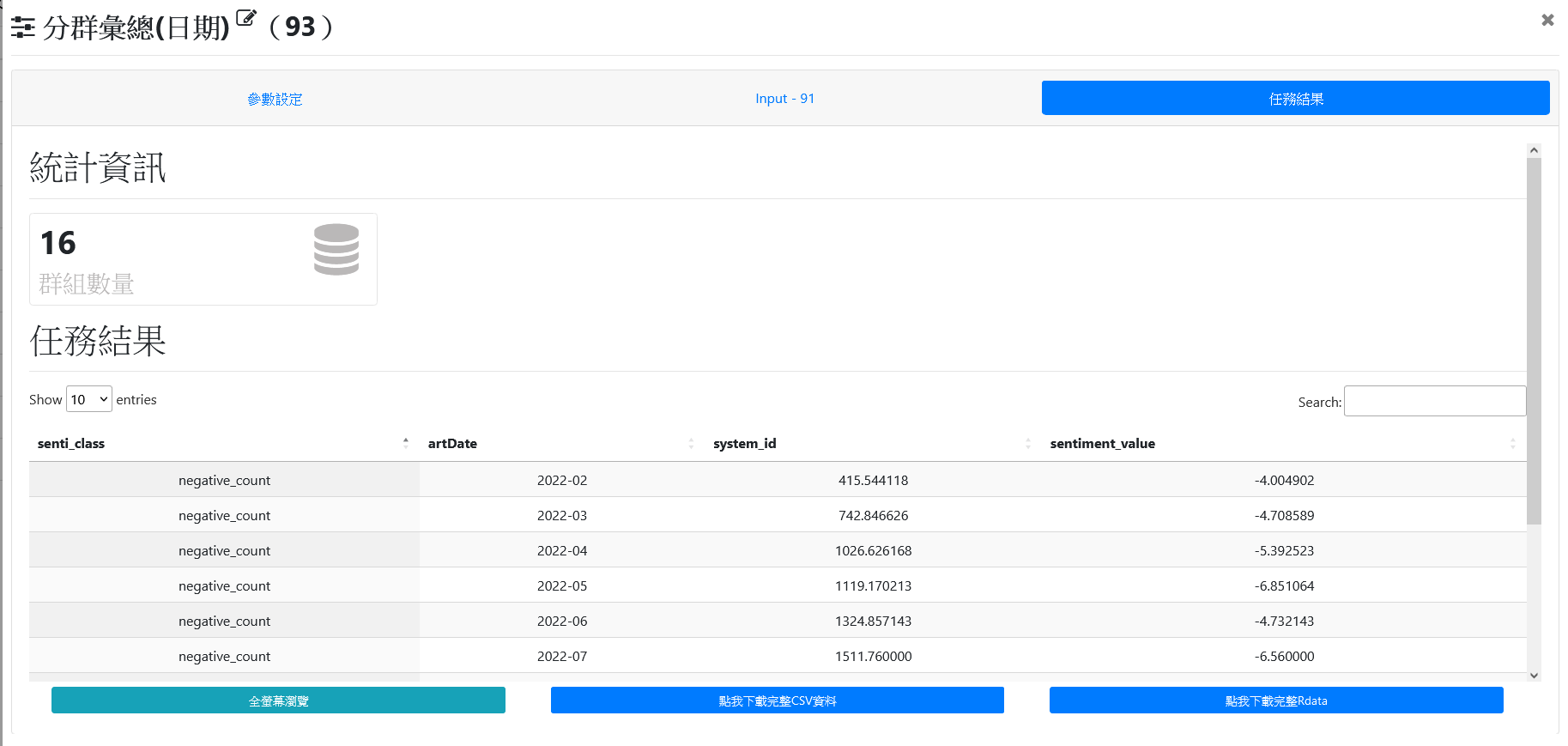




#### **分群彙總(日期)**

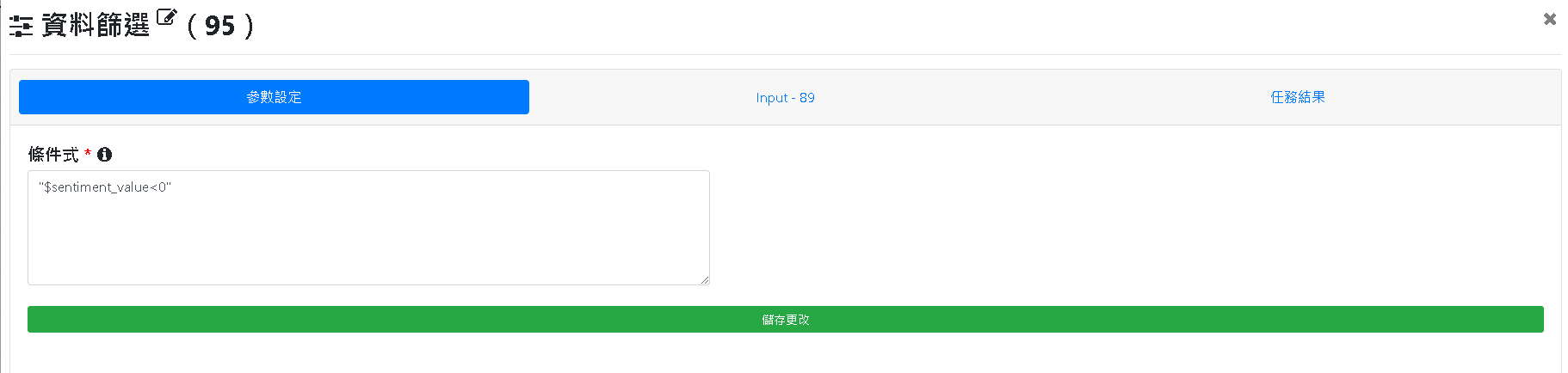
取得每月正反向各別的情緒分數後，以月為單位分群可以看出，每個月的情緒分數。

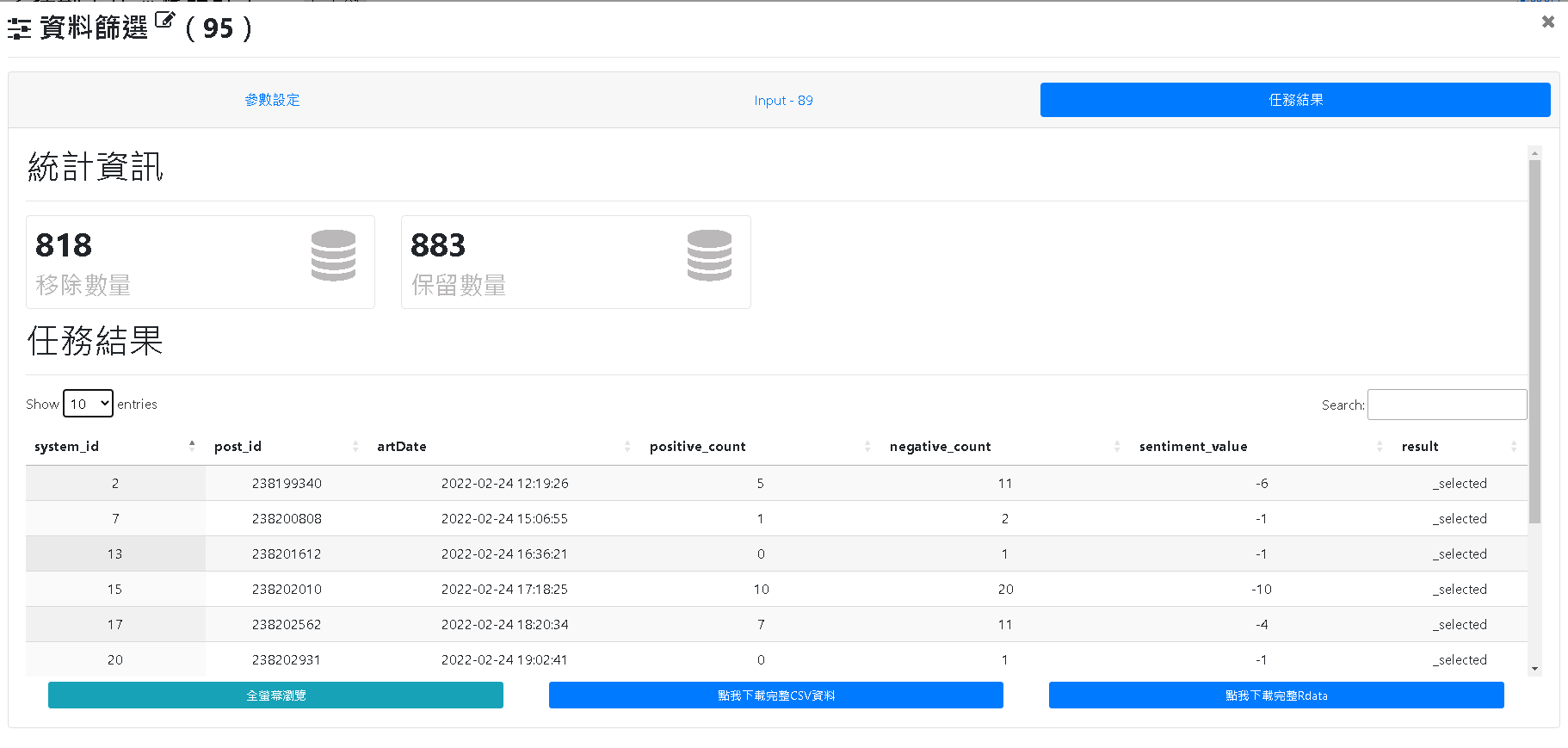




#### **資料篩選**

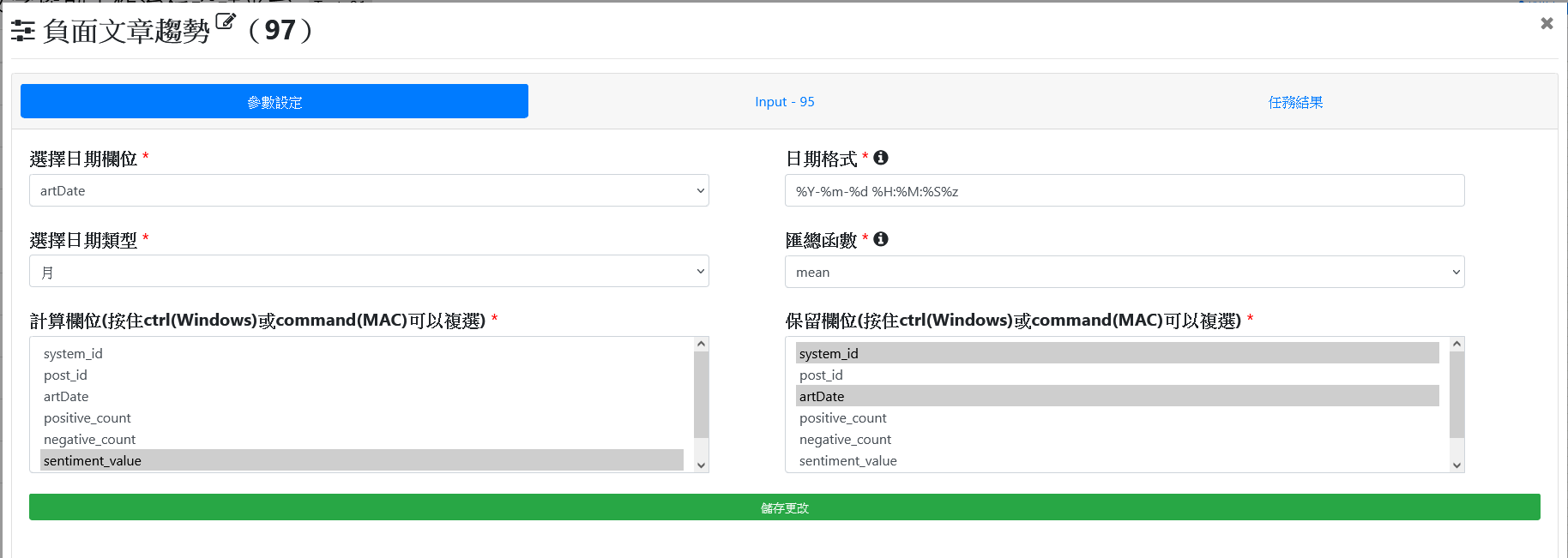
欲分析負面的情緒，故把情緒分數 < 0 的挑選出來。

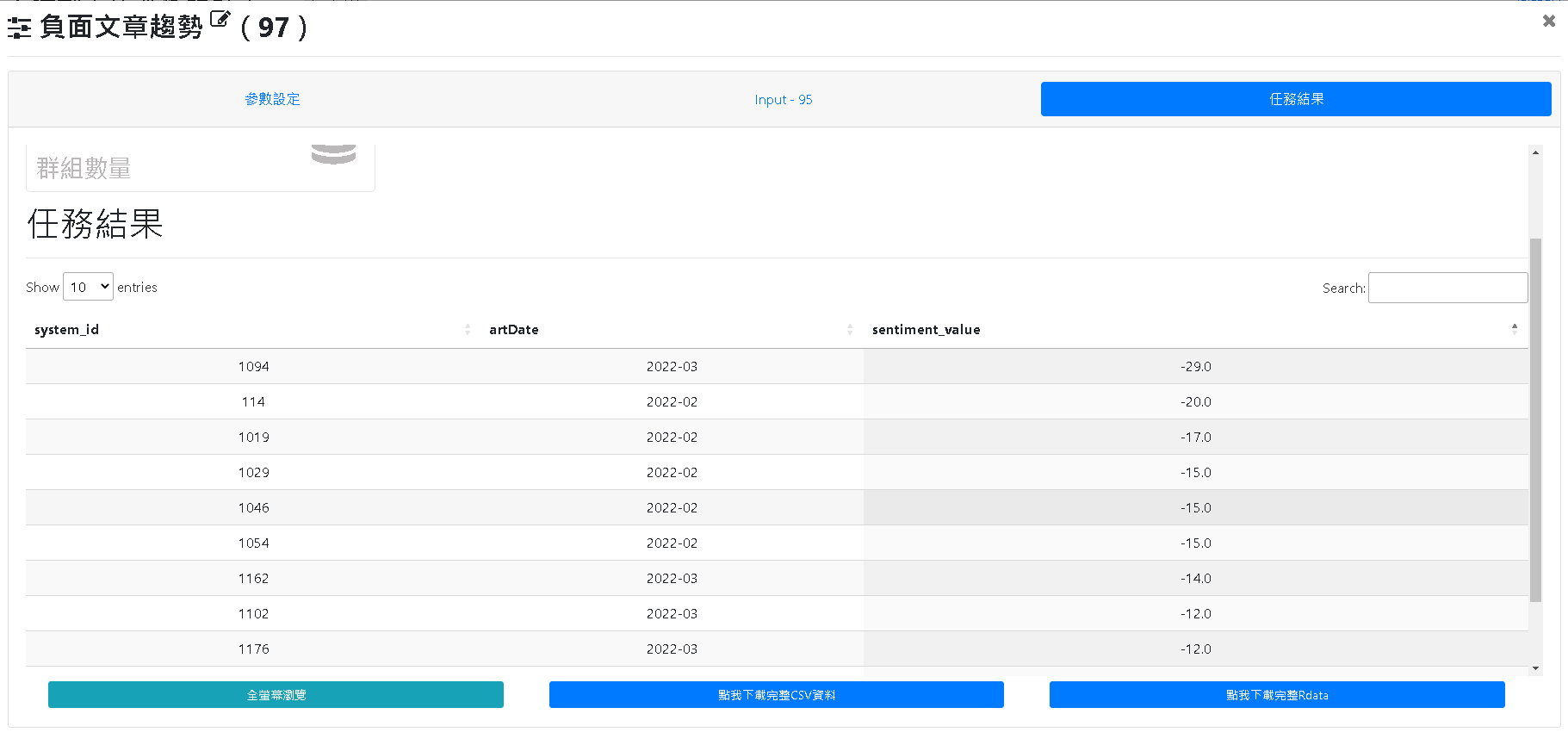




#### **負面文章趨勢**

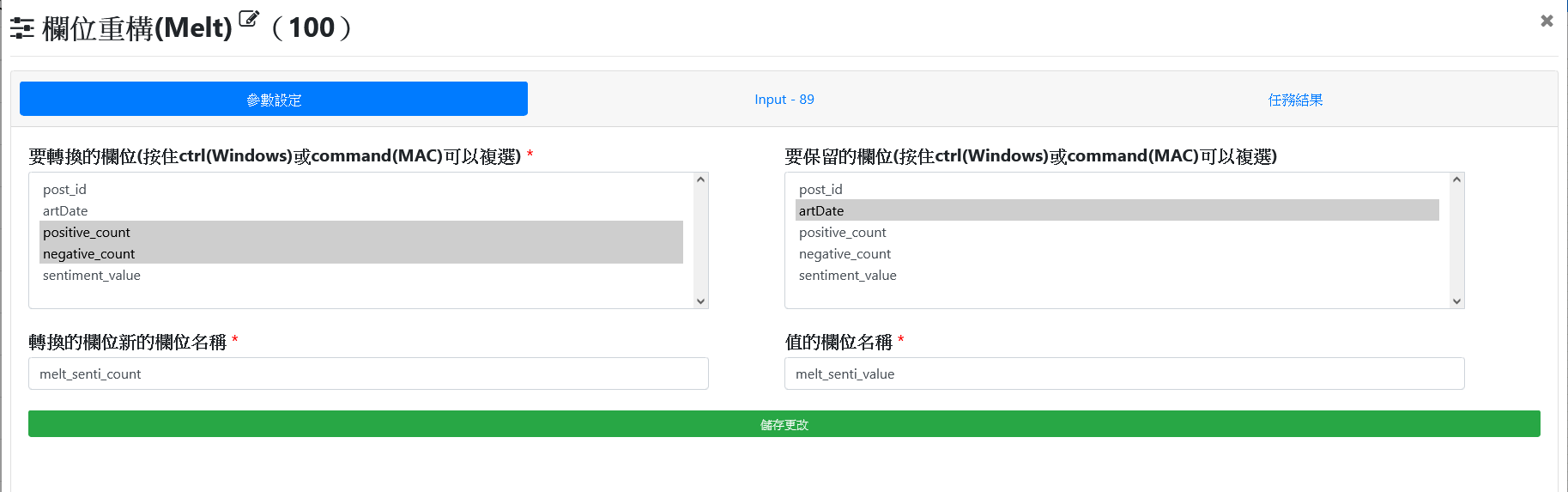
透過欄位篩選出計算欄位後再經由資料篩選過濾出情緒值小於0的文章。



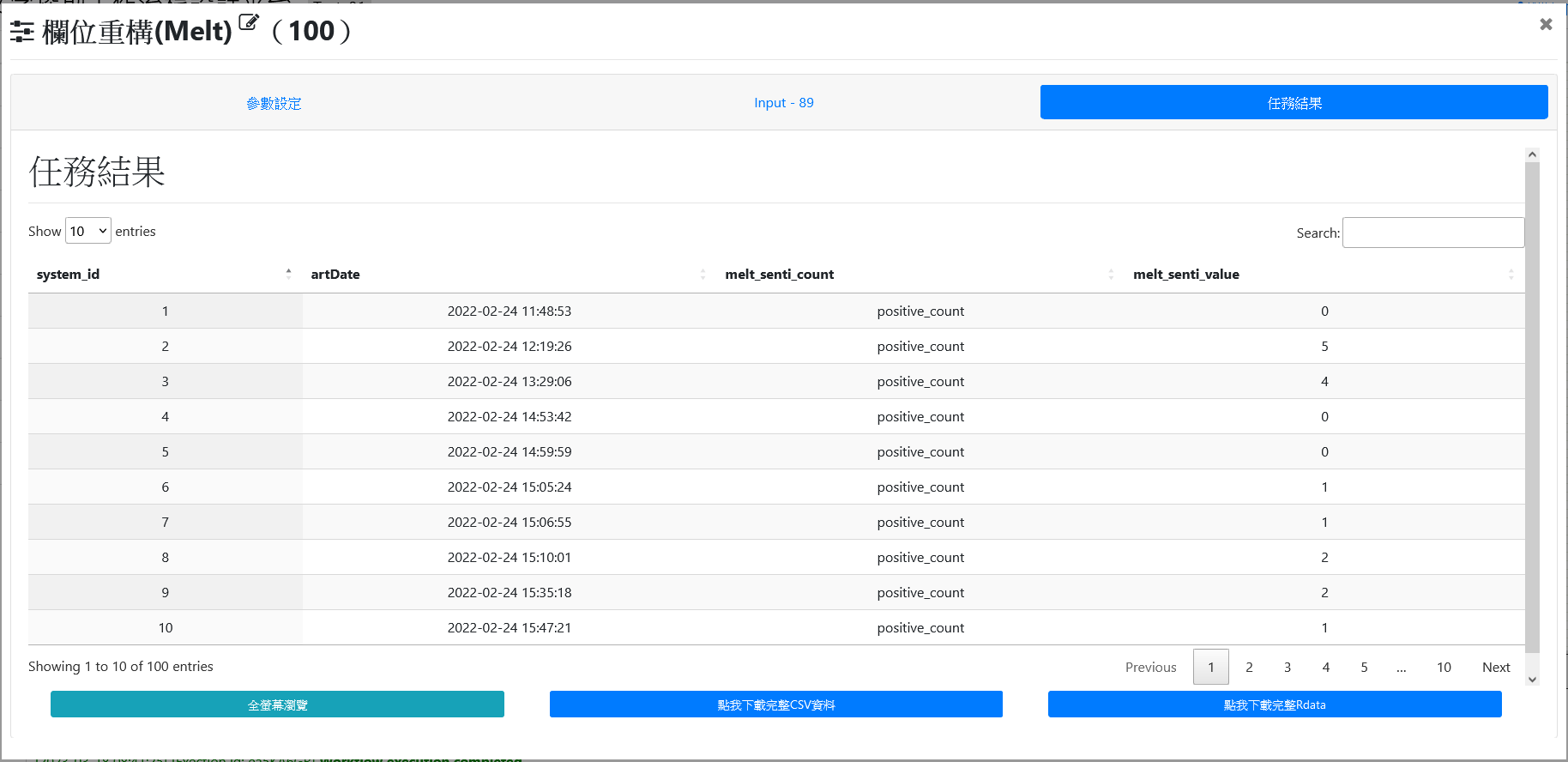


#### **欄位重構**

將每篇文章每日正反面的情緒數量轉成直的顯示方式。

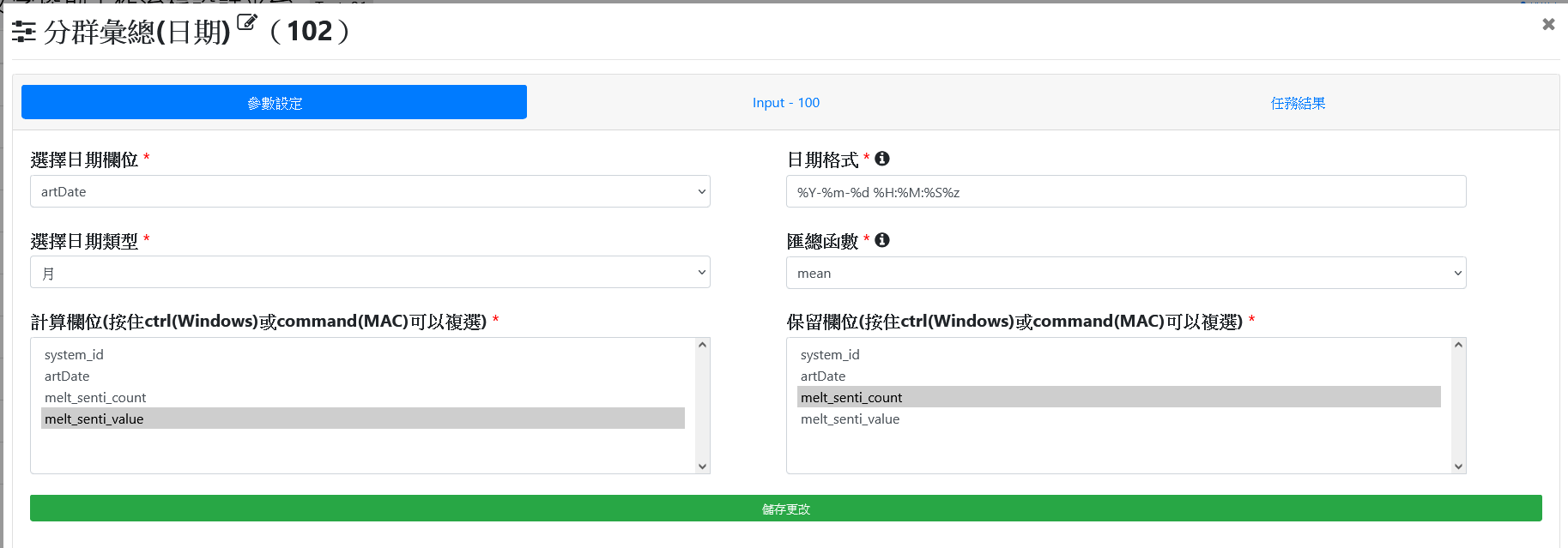


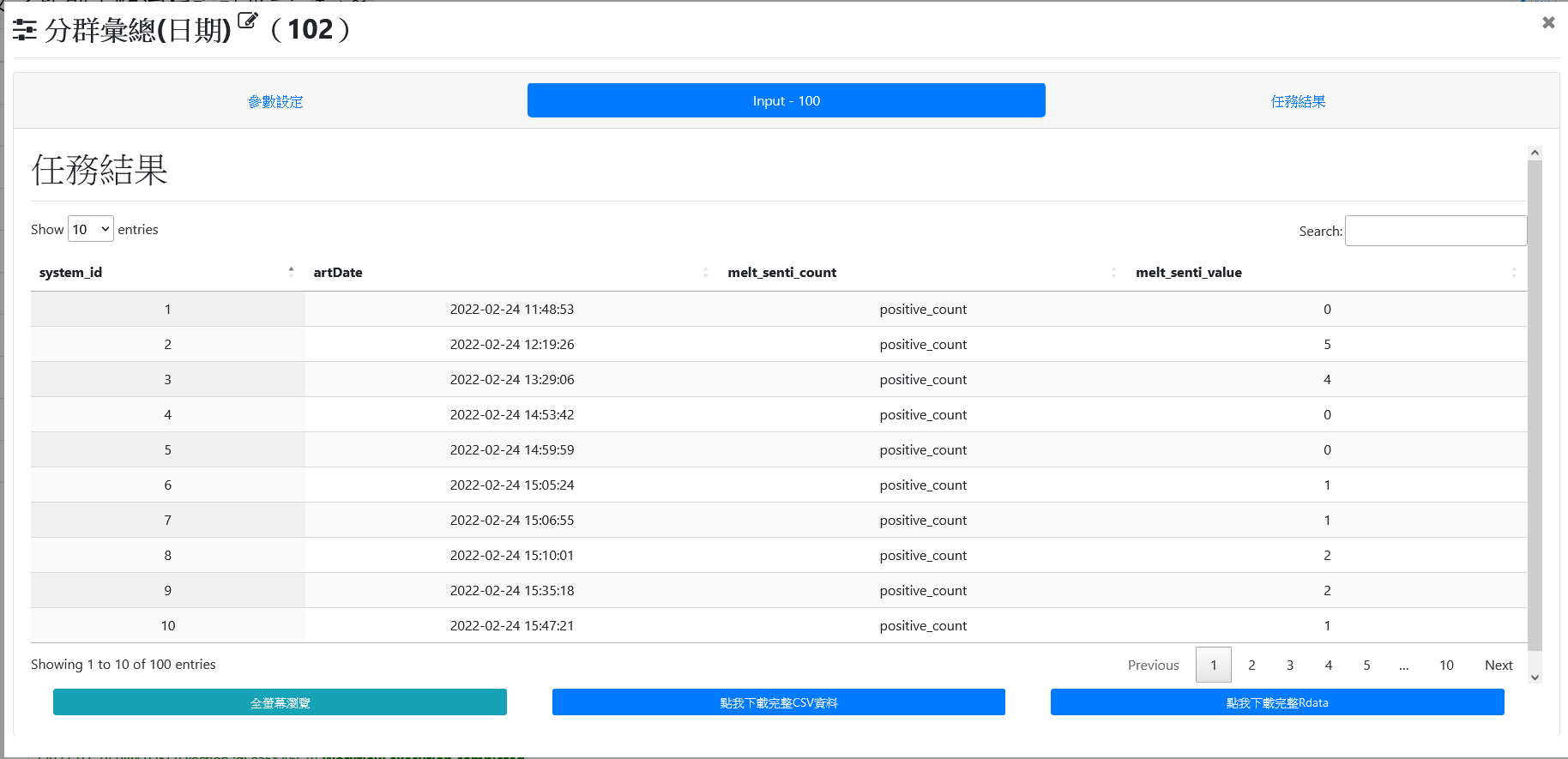
**欄位重構-任務結果**

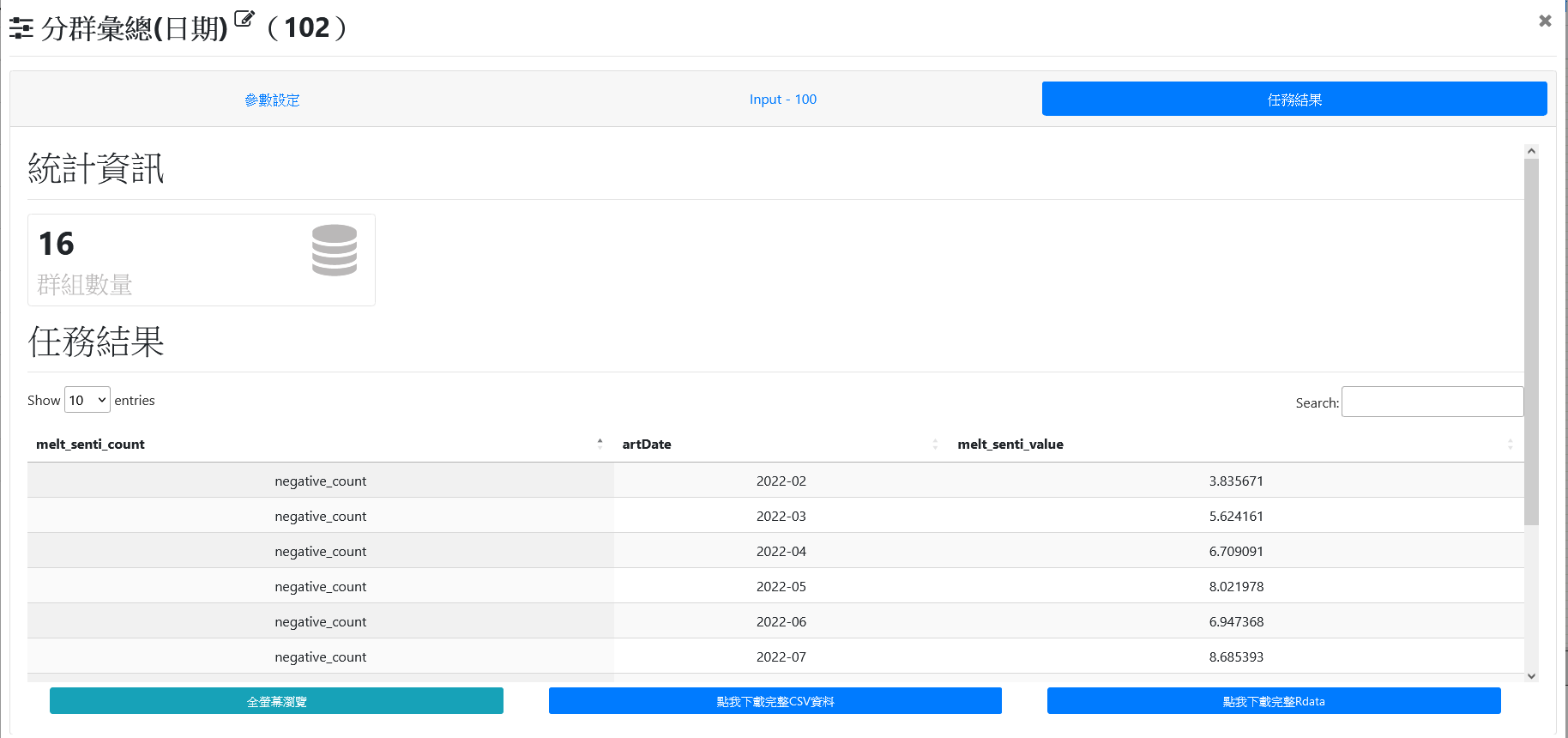


#### **分群彙總(日期)**

計算每月的各文章的正反面與情緒分數的平均值。



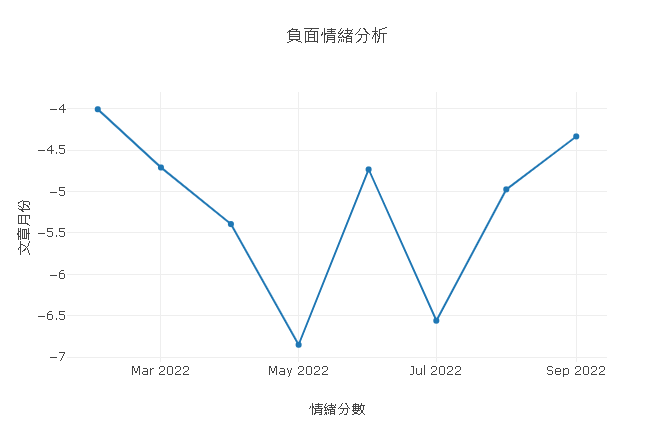




### **圖表分析**

#### **儀表板-折線圖**

可以看出在 2022 年的五月以及七月的文章負面情緒用詞最高，推估五月是因烏克蘭的抗俄象徵-亞速鋼鐵廠的最後一批守軍投降。七月則是一個烏克蘭在盧甘斯克地區最後的主要據點也淪陷。

****

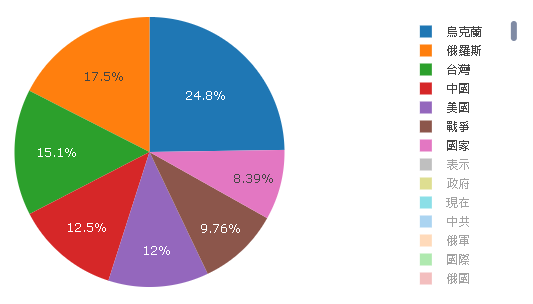
#### **儀表板-文字雲**

同樣是大國對上小國，加上因台海近年來局勢不穩定，即使是遠在歐洲的戰爭，民眾對於中國大陸來犯與否還是有疑慮。這一疑慮可以從文字雲看出遠在亞洲一端的中國、中共、台灣等字樣，也大量現在文字雲中。

****

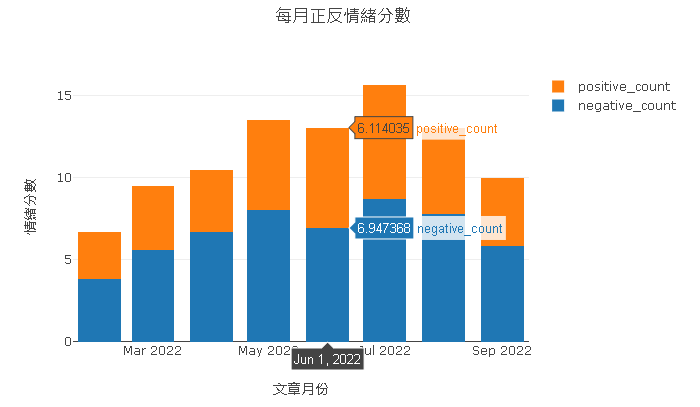
#### **儀表板-圓餅圖**

文章出現中字詞出現最多的頻率，以圓餅圖呈現，可以發現討論烏克蘭、俄羅斯、台灣、中國...等字詞出現頻率最多。

****

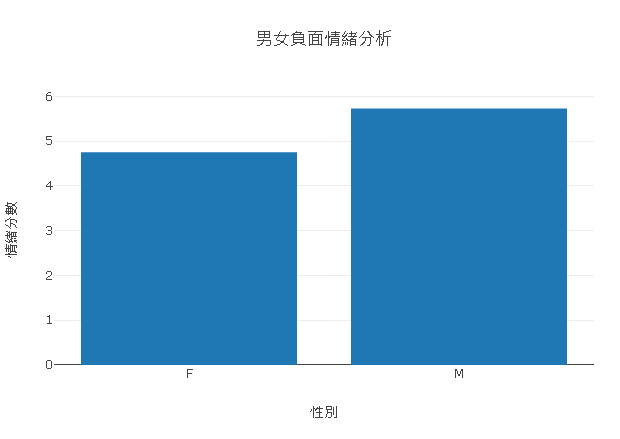
#### **儀表板-堆積長條圖**

可以看出在 2022 六月的正反情緒分數差距不大，當月烏克蘭擊沉俄羅斯軍艦，可能因此正面的情緒詞彙較多。

****

#### 儀表板-長條圖

從這張男女比例負面情緒長條圖來看，可以大概得知台灣男性相較女性有較多的負面情緒，推測可能是在台灣僅男性需要當兵，且Dcard上相較其他平台有更多的年輕人，故可推斷出年輕男性針對戰爭議題的危機感較為強烈。

****

### **結論**

烏克蘭戰爭雖然與台灣沒有直接關係，但台灣與烏克蘭在國內的政治立場同樣分為西方與非西方陣營，因此同樣的新聞內容透過不同媒體的報導後，呈現出了截然不同的立場觀點。

例如2022年7月，美國國防部前部長訪台的新聞，激起正反兩方激烈討論，當時資料出現了許多正、負面字詞，正面的部分提到了「支持」、「和平」、「承諾」、「自由」、「協助」等，而負面字詞則有「威脅」、「遠離」、「擔心」、「挑釁」、「沒用」等，這頁代表此次事件對台灣民眾也造成了一定程度上的影響。許多人擔心台灣是否會成為下一個烏克蘭，網路上也廣為流傳一句話：今日烏克蘭，明日台灣。

相較其他月份，2022年7月有資料期間內最高的情緒分數，代表了對於台灣人來說，比起烏克蘭現在正在發生的戰爭，我們更擔心大國間的角力是否會破壞台海和平。畢竟烏克蘭遠在歐洲的另一邊，而台海即在我們觸手可及之處。

第九組成員：

N104020003 王博玄

N104020004 葉子豪

N104020026 林冠宏

N104020028 姜孟嫻

N104020030 林欣民

N104020027 林欣穎

N104020029 陳勇任