社群媒體分析期末專案報告

題目：

從PTT論壇分析年終獎金及調薪議題

之情緒分析

指導老師：黃三益教授

第三組 郭耿耀N094020005

周昆穎N094220007

李昀瑾N094020022

王莉筑N094220029

# 大　綱

[大　綱 2](#_Toc93042055)

[**一、動機和分析目的** 3](#_Toc93042056)

[**二、資料集的描述** 3](#_Toc93042057)

[**三、資料清理過程** 5](#_Toc93042058)

[**四、視覺化的分析結果與解釋** 7](#_Toc93042059)

[**五、結論** 19](#_Toc93042060)

## **一、動機和分析目的**

2121.12.30這天最受到屬目的一筆新聞是長榮海運年終獎金，消息傳出發給員工最高40個月的年終獎金，很多報導指出有員工一早發現新台幣200萬元入帳！對此長榮只低調說明，年終獎金依營運績效及員工個人表現核發，每個人狀況不同。

每到等待年終獎金發放的季節，我們常常被問到，你拿到年終獎金了嗎？年終獎金是員工辛苦為公司賣命、任勞任怨的成果，也是公司展現良心與氣度的時候，除了年終獎金多寡常被拿來比較，不同產業能領到的金額也都不同。所以我們欲透過PTT論壇對於此議題做文字探勘、情緒分析等，作為本組的期末報告議題。

**二、資料集的描述**

本組從工作流程平台既有資料來源爬取PTT於2021.12.01~2021.01.07區間:八卦版討論版，找出與年終獎金、調薪、長榮相關討論文章。

整體流程架構圖

資料平台:**PTT八卦版討論版**

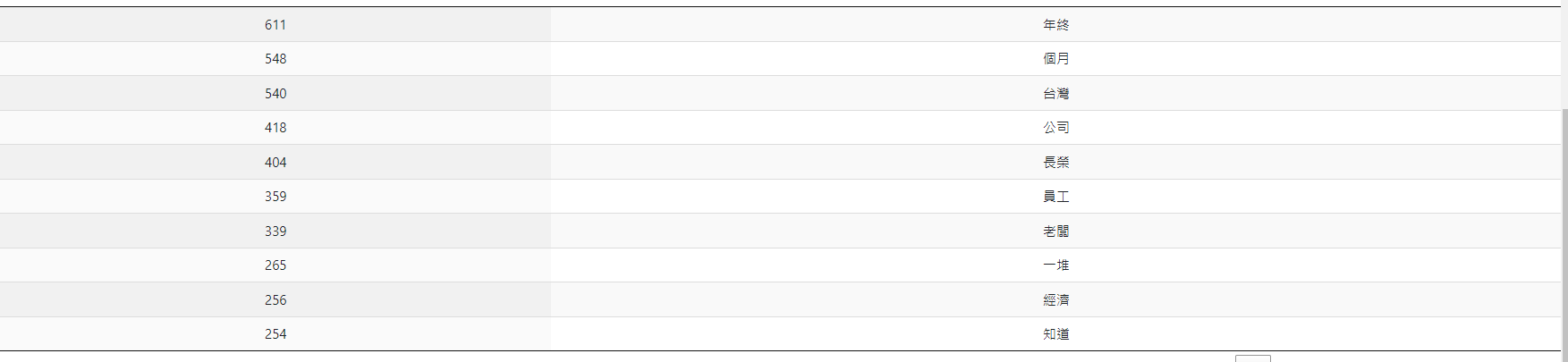
關鍵字:年終獎金、調薪、長榮等相關字詞從PTT兩個討論版共取得409筆數資料



**計算詞頻後文字雲(200字):**

熱門詞彙為:年終、(幾)個月、台灣、公司、長榮、員工、老闆、經濟

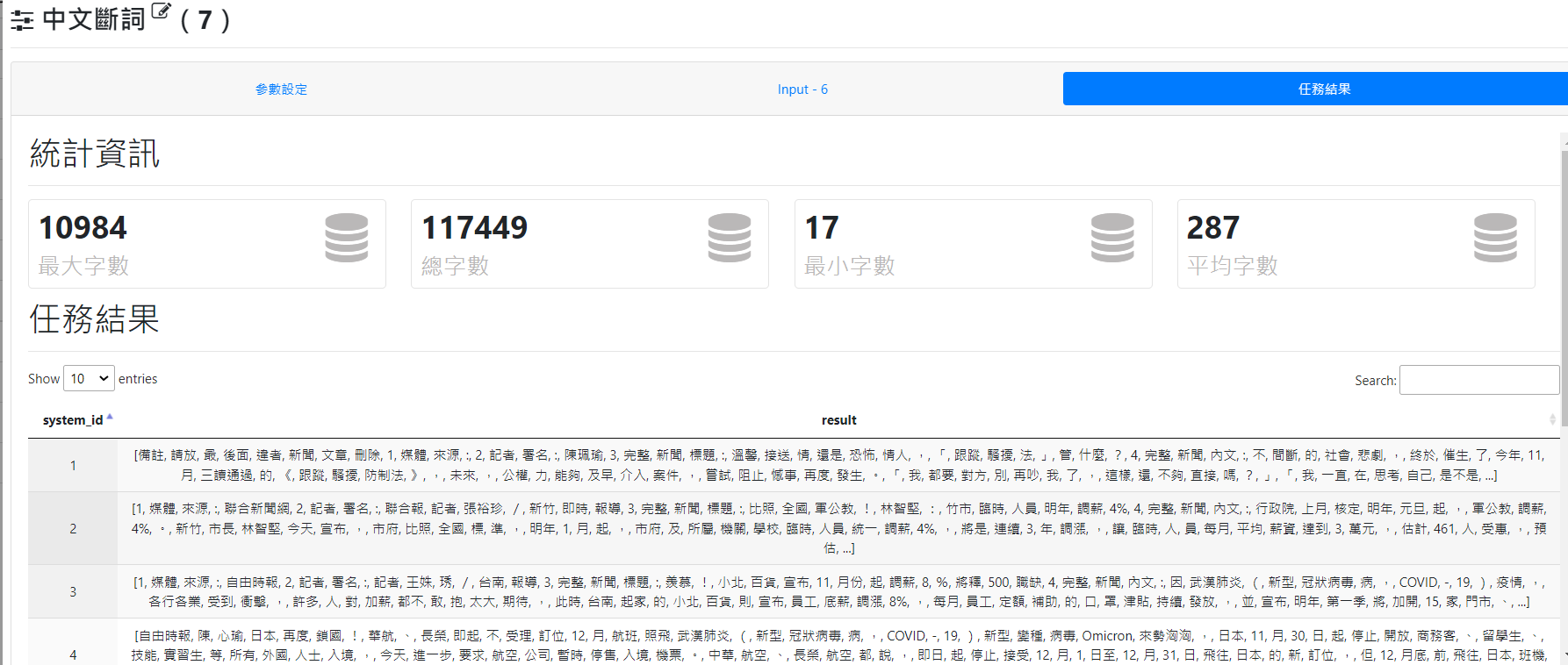




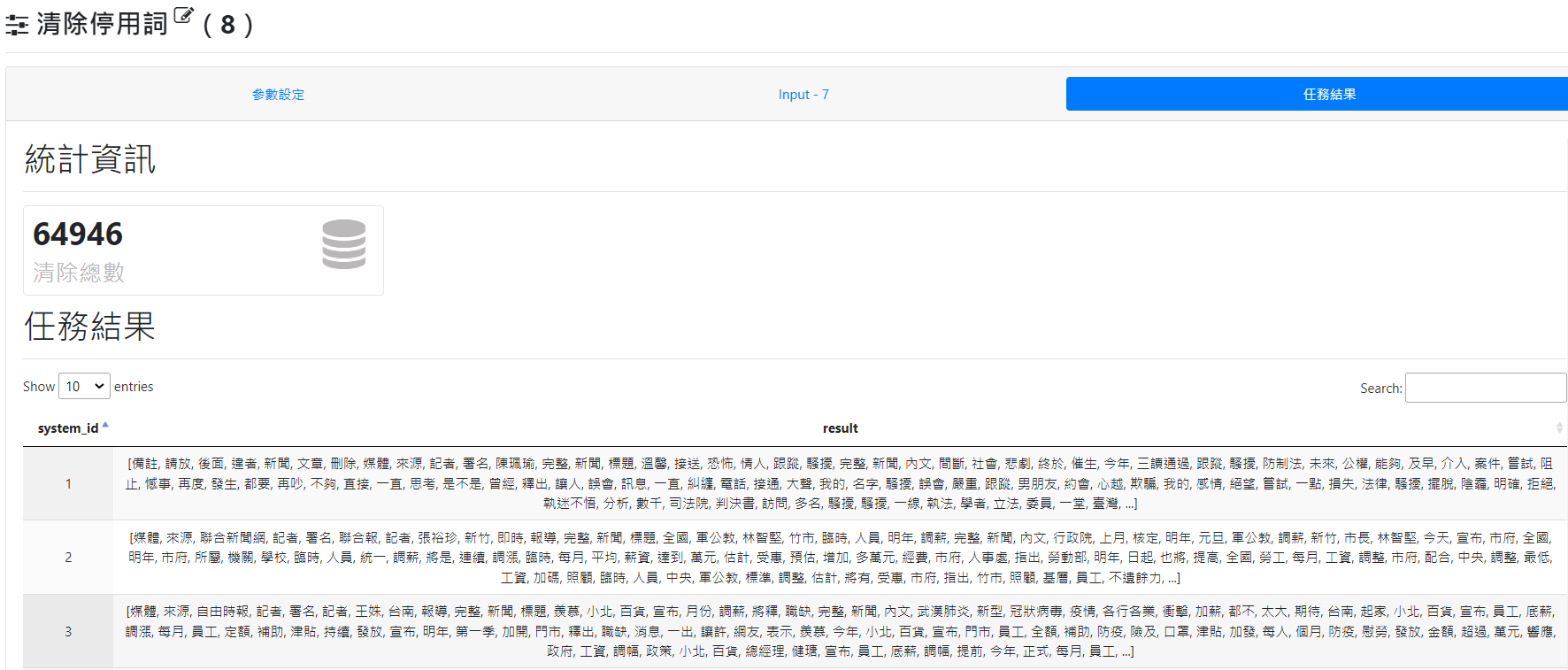
**三、資料清理過程**

替換字詞: 名言、表示、持續、超過、表示、萬元、護管員(指森林護管員)、今天、超過、山老鼠、蔡英文、人員、備註…等較無關字詞。

中文斷詞處理結果:10,984字數



**清除停用詞:**手動將網址及跟內容無關自詞做過濾，並啟用清除html lag\*、清除英文字元…等



**處理結果:**共清除64,946筆

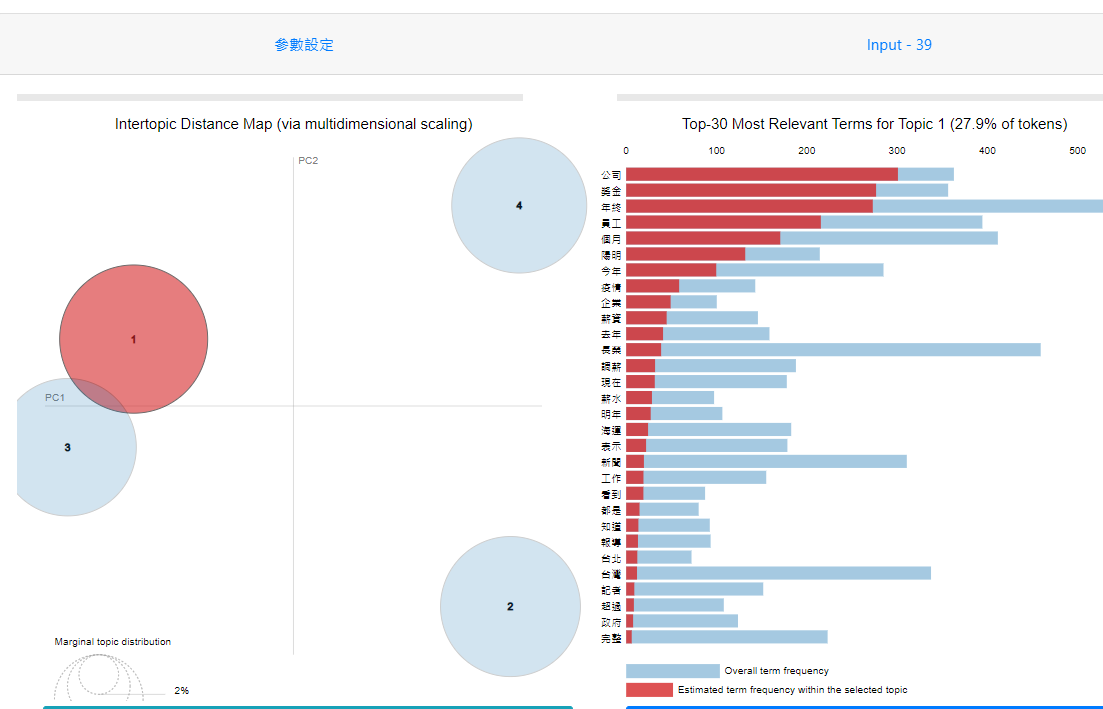
**四、視覺化的分析結果與解釋**

視覺化儀表板:

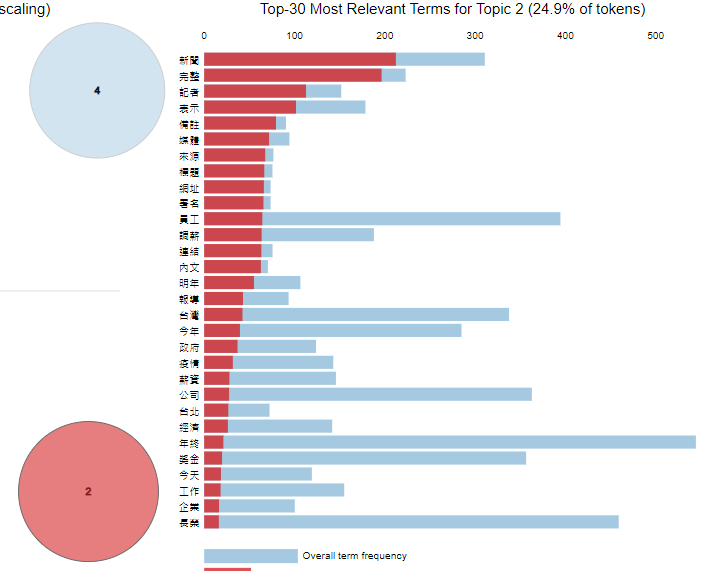
LDA主題模型: 本組分成4個分群



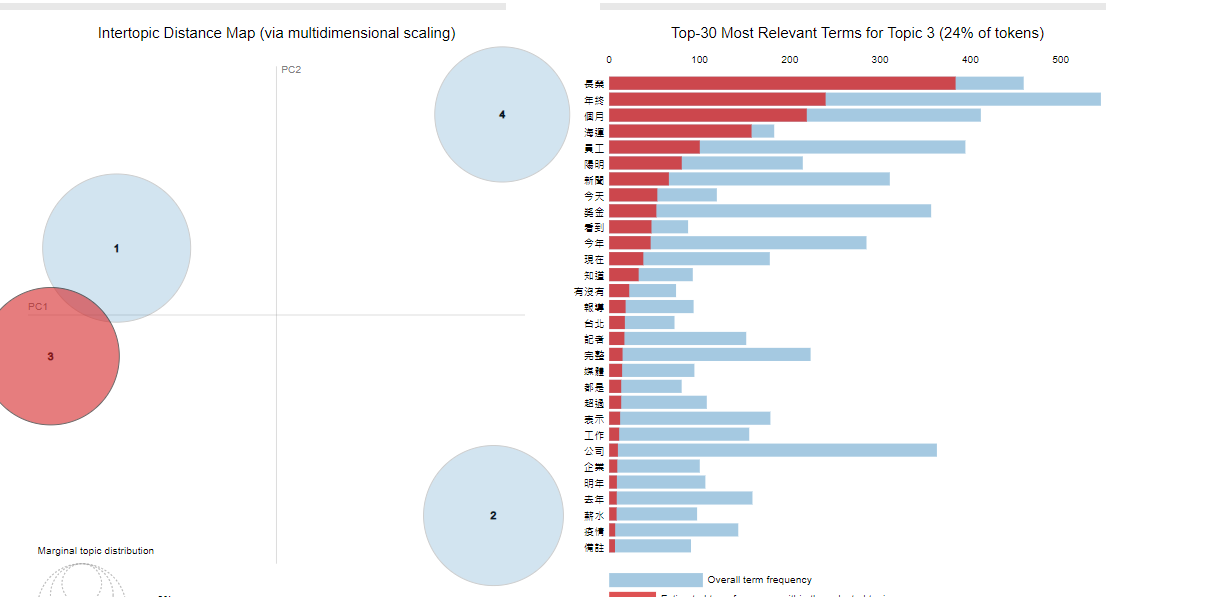
LDA主題一(公司、年終、獎金):公司；獎金、年終、幾個月、陽明



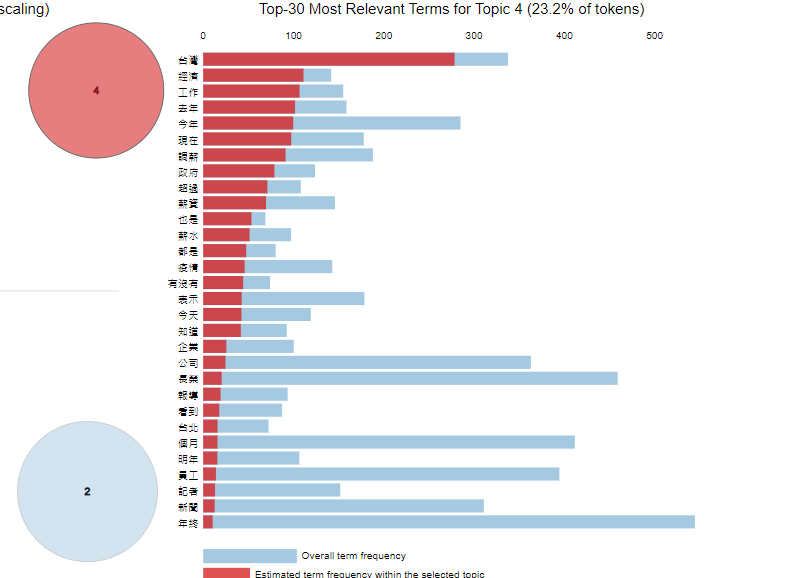
LDA主題二(媒體報導):新聞、記者、媒體



LDA主題三(長榮等海運年終為熱門探討):長榮、年終、員工、海運



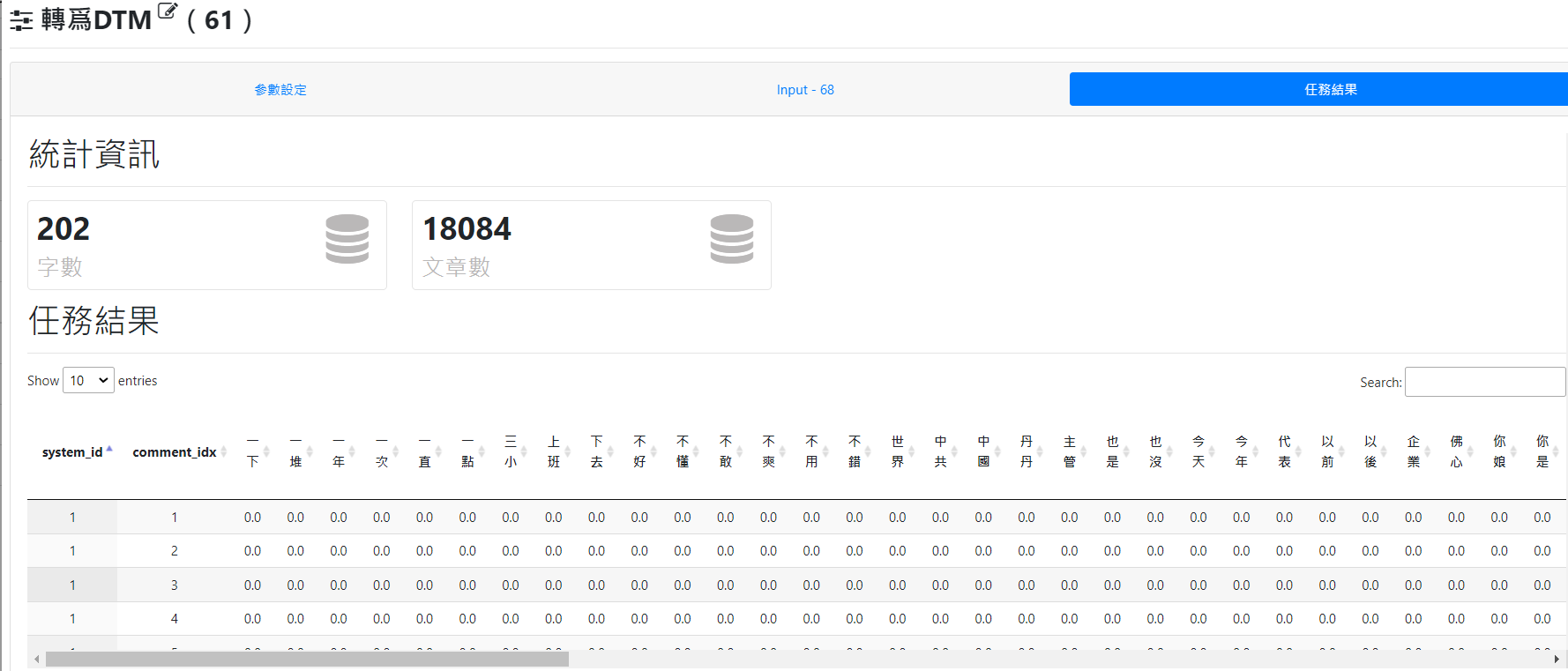
LDA主題四(時事類):台灣、經濟、去年、今年、工作、調薪



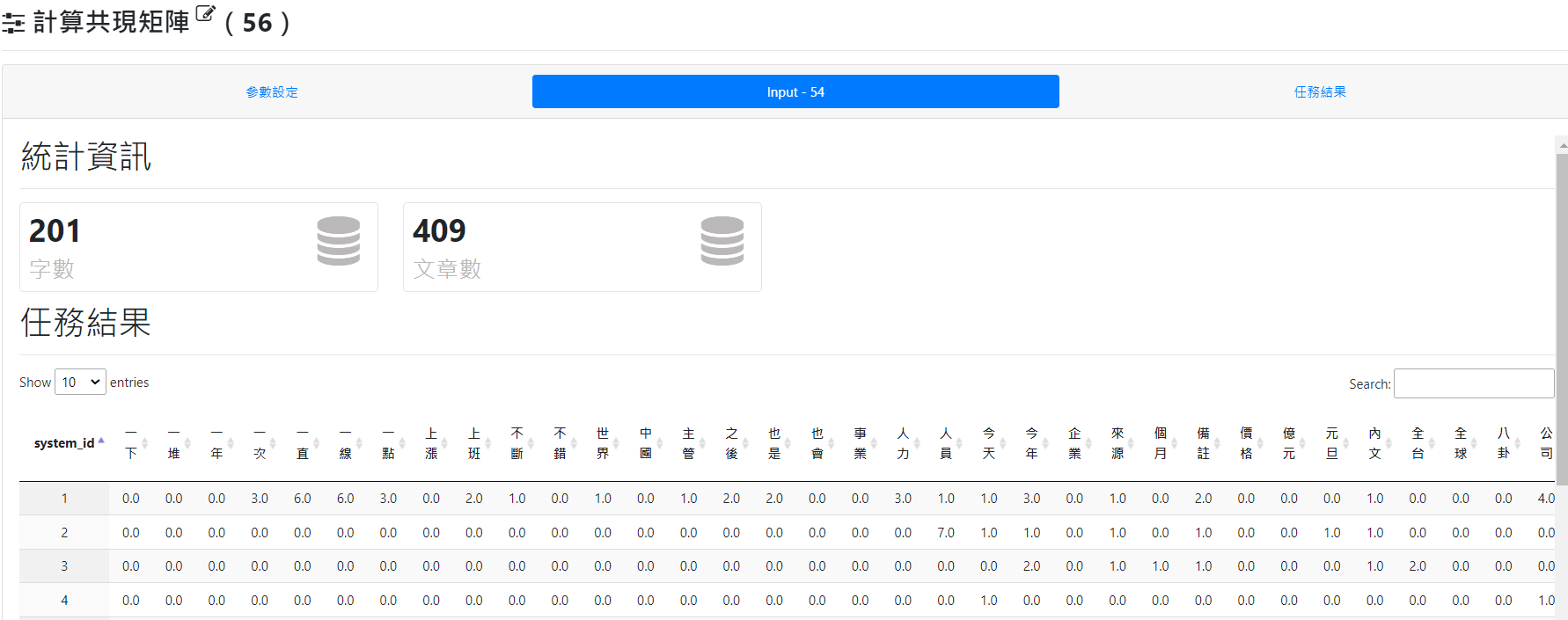
**字詞網路圖:**

接下來把計算相關性矩陣的結果進行文詞網路圖的詞彙分類，從計算相關性矩陣取得詞彙csv資料檔，並進行分類後匯入字詞網路圖，可於字詞網路圖選定詞類主題為中心，從文集中選出相關性最高的關鍵詞，並對它們進行社會網絡分析。

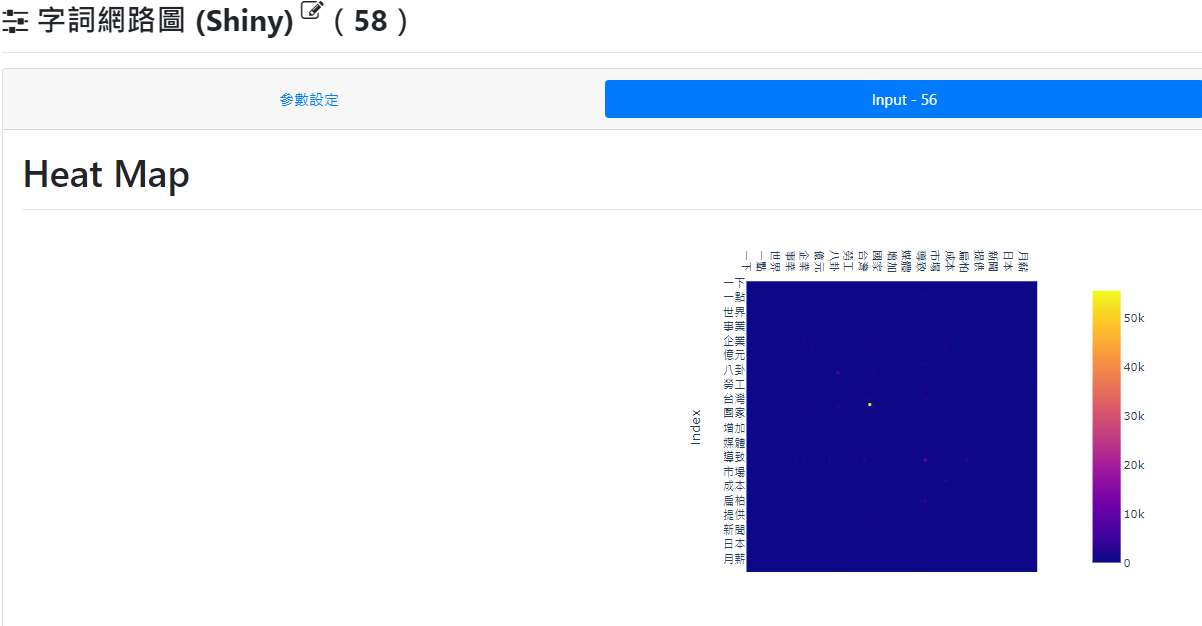
原PTT爬蟲字詞網路圖



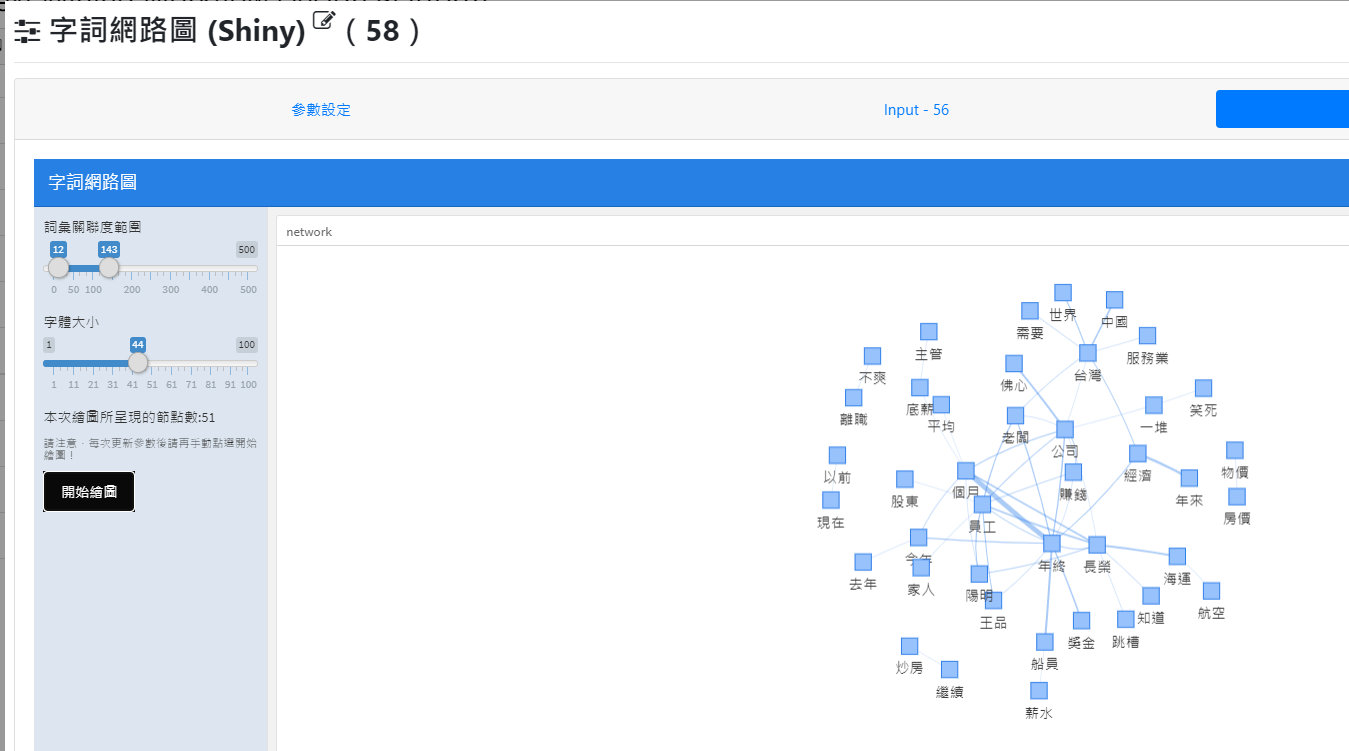
**計算共現矩陣:字數201筆、409篇文章**

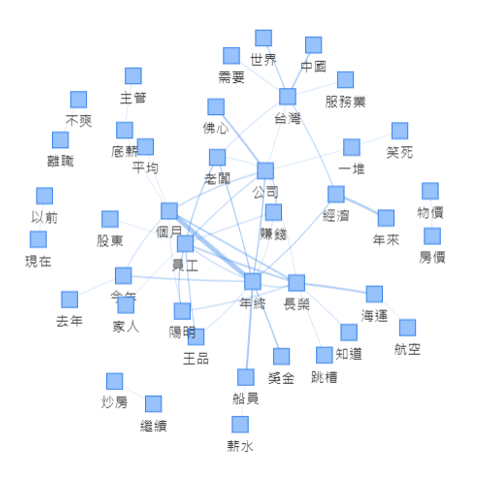


**字詞網路圖:**



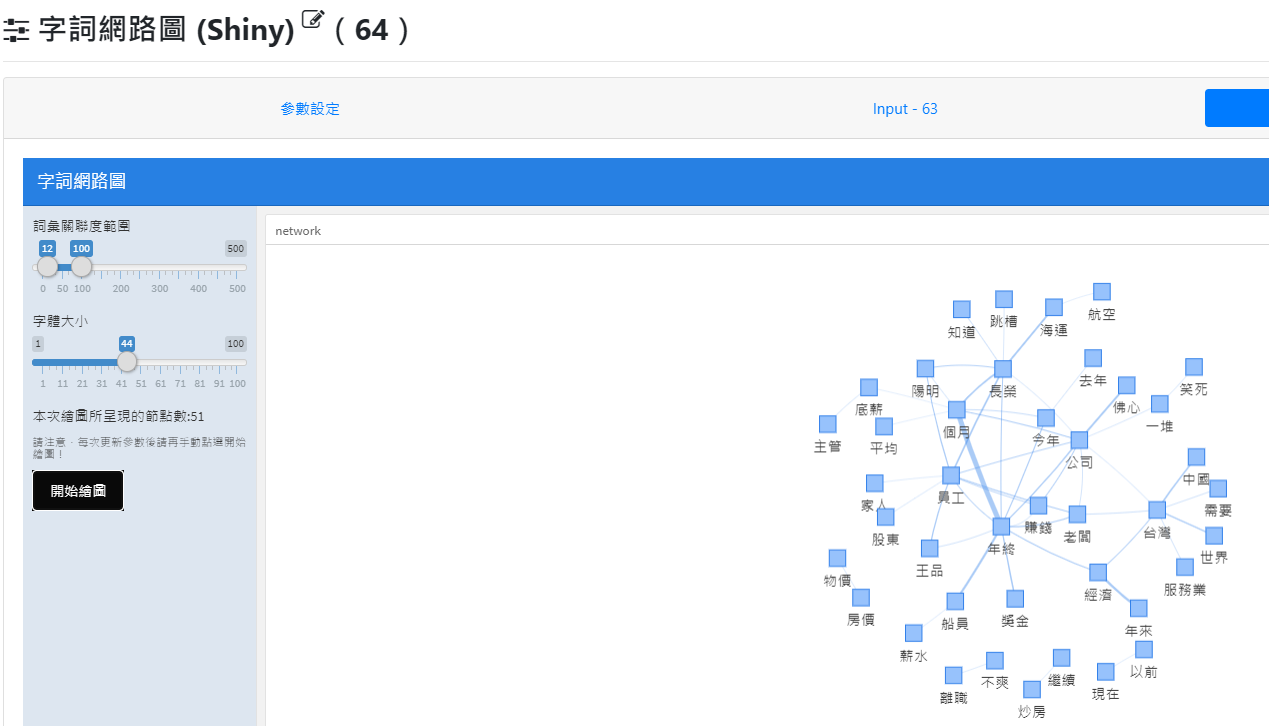
**之後產出原始PTT分析後字詞網路圖**





**網路節點分布分別為:公司、年終、台灣、長榮、幾個月等**

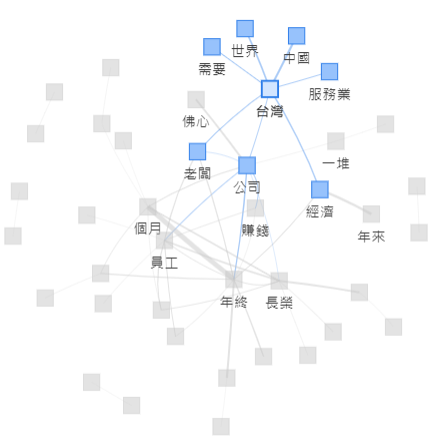
**留言萃取後產出之字詞網路圖**



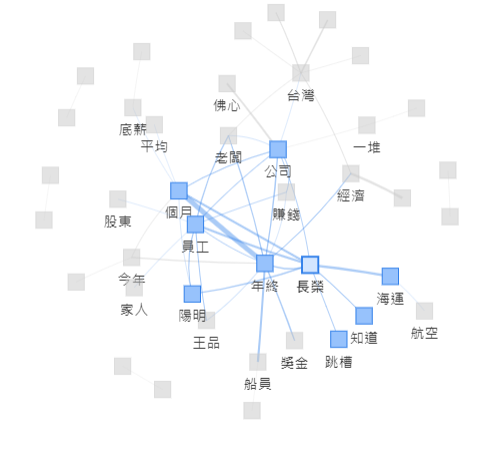
主要節點1:公司



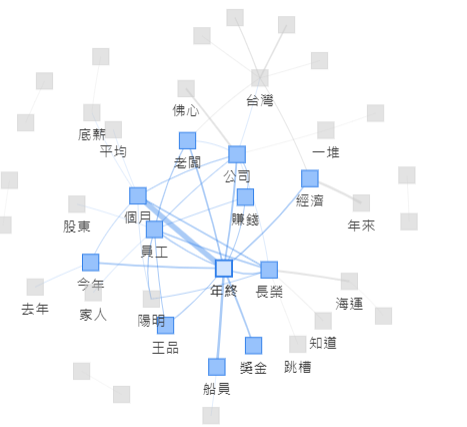
主要節點2:台灣



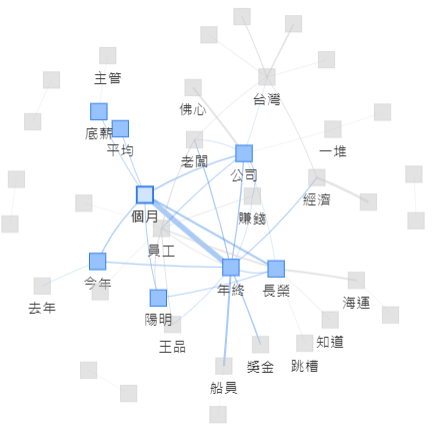
主要節點3:長榮



主要節點4:年終



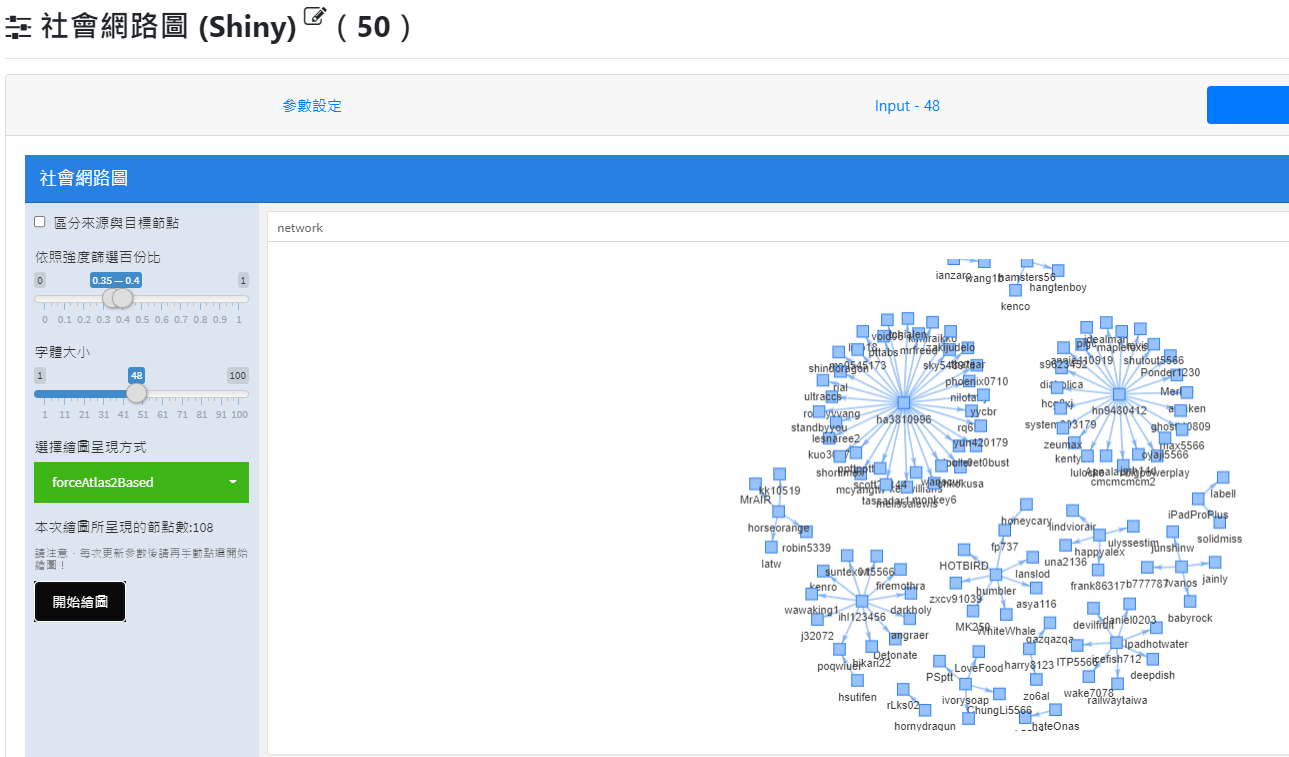
主要節點5:(幾)個月



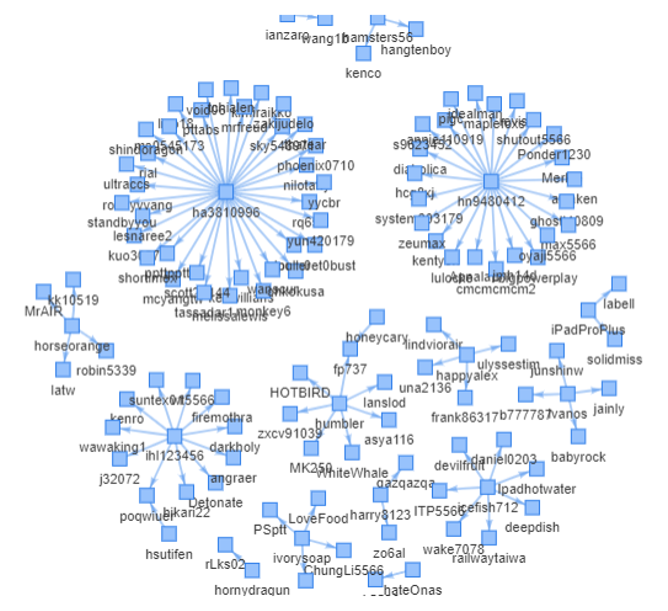
基本上與PTT原始網路圖分布內容相差不大

**社會網路圖分析:**

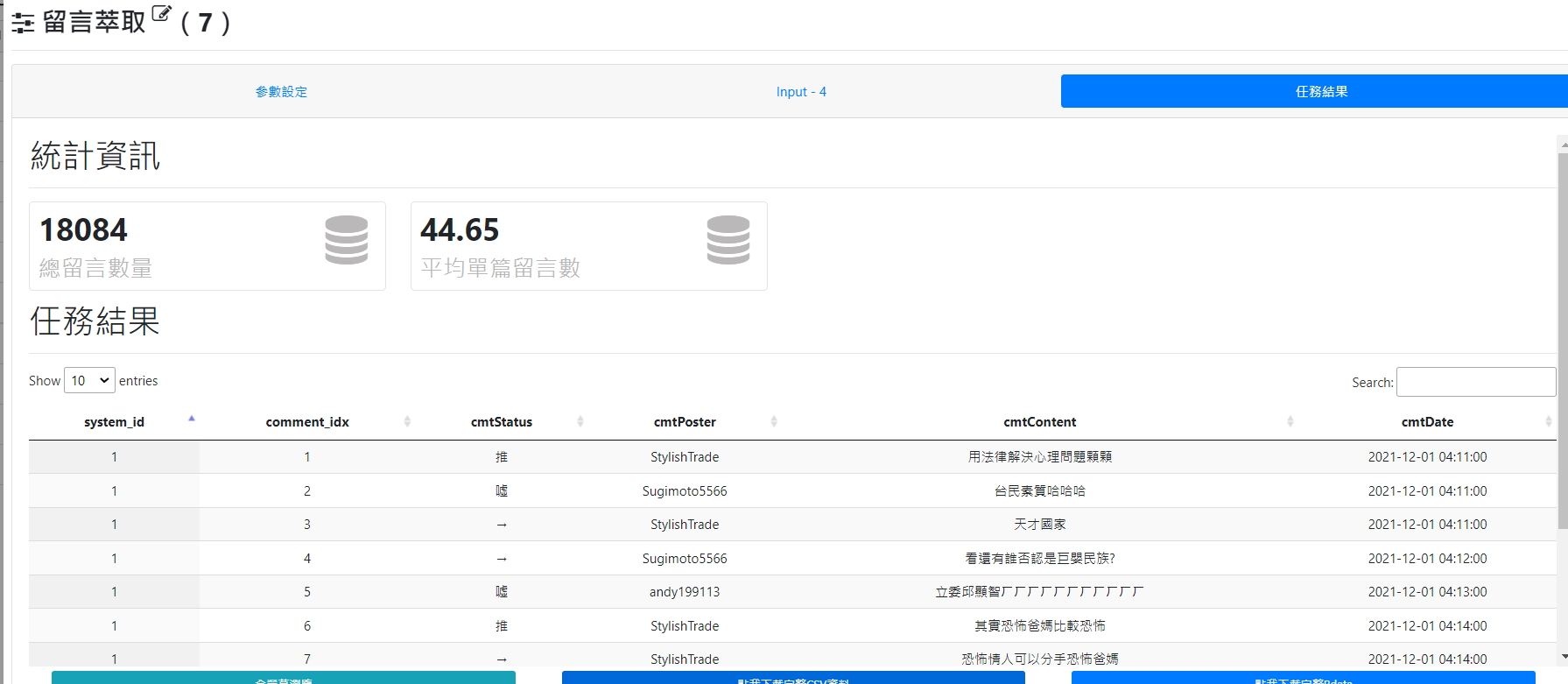
Topic分布社會網路圖



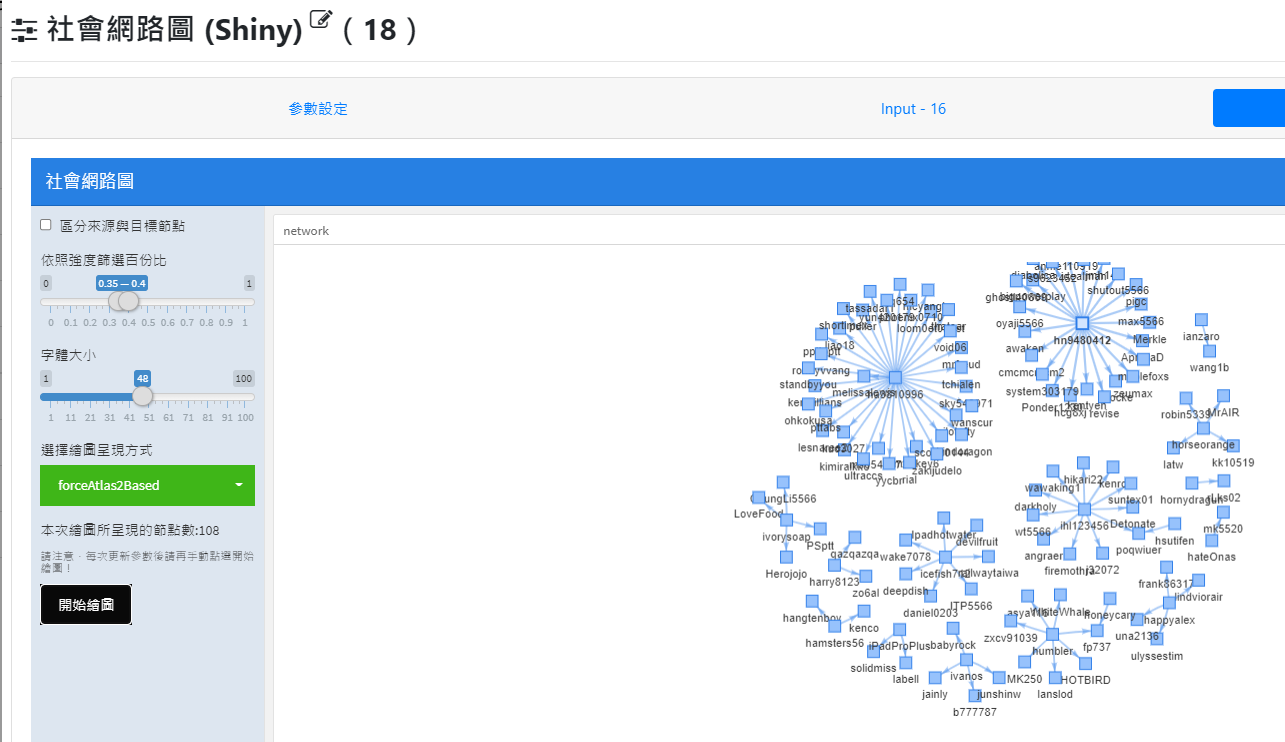
**主要意見領袖節點:ha3810996、hn9480412、及其他小群分布**

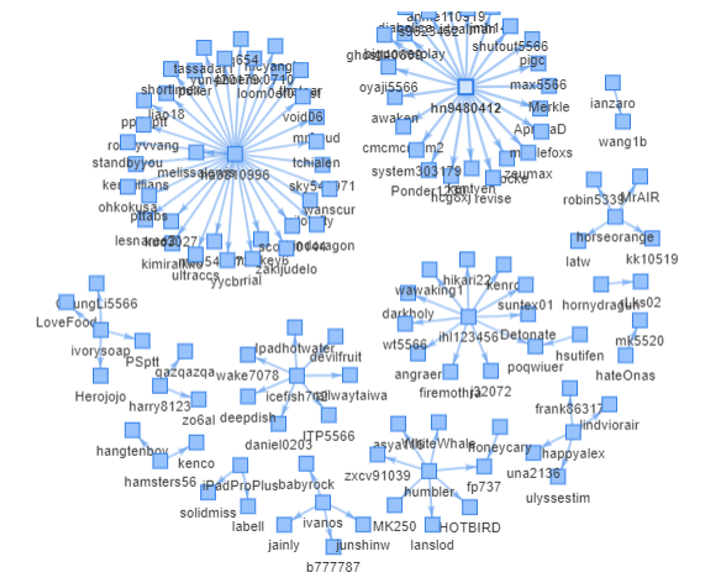
****

**留言萃取處理後:**



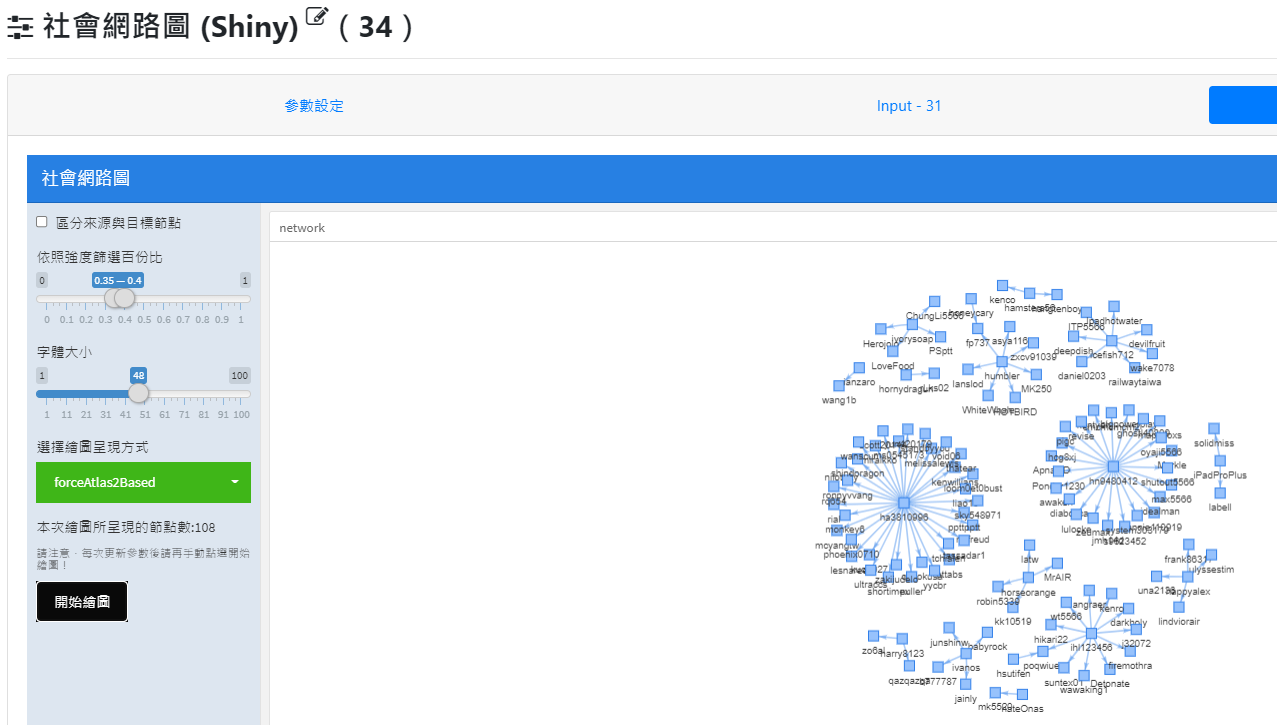
**以留言萃取產出之社會網路圖:**

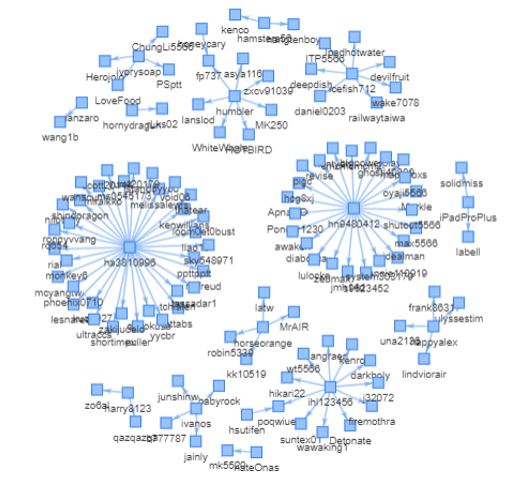




**意見領袖:xxx0996、hn9480412為最高，另外horseorande、ihl123456、icefish7xxx、humbler為小群分布**

**以情緒分析分析後產出社會網路圖**





**意見領袖為:humbler、ha3810996、hn9480412、ihl123456**

**五、結論**

1.關於主題分類分析:

LDA主題一:以公司的年終獎金為網友熱門討論議題。

LDA主題二:有關媒體報導之熱門新聞(如長榮40個月議題)。

LDA主題三:長榮、陽明等海運年終為熱門探討

LDA主題四:時事類，以台灣經濟、並比較去年、今年、及各產業工作、調薪、比較情形。

2.網路圖之意見領袖留言分析:

ptt社會網路圖:

ha3810996 :今天看到長榮海運年終40個月， 真的看到眼睛爆凸、口水快流出來， 不過沒關係， 之前考古題已經做到爛， 要削科技業工程師， 就看台積在哪設廠。

留言萃取後處理網路圖:

hn9480412: 食店員工「整年0請假」年終超大包　本人也嚇到：27萬。

情緒分析後處理網路圖:

Humbler:針對不同議題回應較為熱絡，如一銀年終6個月、炒房、房價，故節點分部較廣。