

社群媒體分析 - 第13組 - 期中專案報告

PTT八卦版 薪資調漲討論分析

N094220012 陳冠汶
N094220019 林明禾
N094220027 鄭雅丹
N094220032 吳憶萱





CONTENT



01

資料基本介紹

02

前言
研究動機

03

分析資料
工作流程及結果

04

結論



PART 1

基本資料介紹



資料來源

- PTT八卦版

關鍵字

- 加薪
- 基本薪資
- 調薪



PART 2

前言 研究動機

前言

研究動機



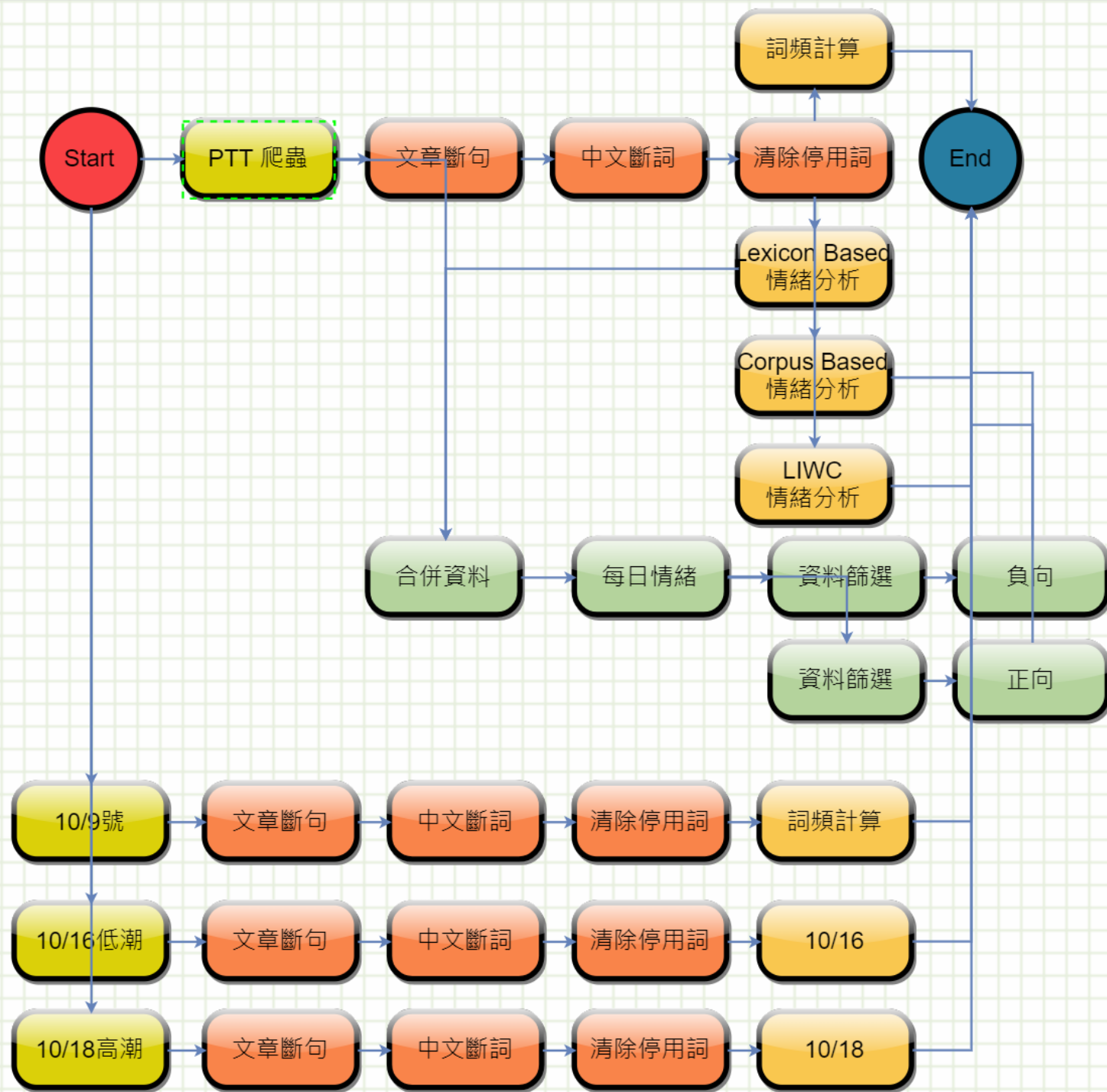
- 疫情嚴峻之下，物價也不斷上漲，唯有薪水漲幅微小，為了改善民衆生計，近期政府終於拍板定案調薪方案。
- 但對於調薪是否人人都看好呢？本組將就PTT八卦版上對於調薪相關的討論文章，進行情緒分析，探討正面與負面聲量的比重，藉此作為民衆對於政策的推行上，正反兩方論述的分析與討論。



PART 3

資料分析

工作流程與結果



參數設定

任務結果

選擇看板 *

Eseries
Exotic_Pet
Finance
Food
ForeignEX
GetMarry
Gossiping

搜尋關鍵字 ⓘ

加薪
基本薪資
調薪

排除關鍵字 ⓘ

中共
共同富裕

搜尋起始日期

2021/09/01

搜尋結束日期

2021/11/25

統計資訊

10

欄位數



533

資料筆數



本組篩選PTT的八卦版在09/01~11/25期間有關「加薪、基本薪資、調薪」等字詞篩選關鍵字，並排除「中共、共同富裕」之關鍵字，避免篩選到較不相關之文章。共篩選出533筆資料筆數。

參數設定

Input - 4

任務結果

選擇處理欄位 *

artContent

斷句符號

.

。

\n

\n\n

?

使用默認符號進行斷句 * ⓘ

是

儲存更改

統計資訊



任務結果

Show 10 entries

Search:

| system_id | sentence_idx | sentence |
|-----------|--------------|-----------------------|
| 1 | 1 | 不好意思 |
| 1 | 2 | 看到這篇一點同情都沒有 |
| 1 | 3 | 社工整天抱怨薪水沒調漲 |
| 1 | 4 | 來看看頭目說什麼 |
| 1 | 5 | #1WPHHD4G (Gossiping) |

中文斷詞 (6)



參數設定

Input - 5

任務結果

選擇處理欄位 *

sentence



選取字典 ⓘ

-----請選擇-----



定義詞彙 ⓘ

置裝費 100
一直漲 100
慣老闆 100
基本工資 100
勞動部 1000



置裝費 100
一直漲 100
慣老闆 100
基本工資 100
勞動部 1000
金融海嘯 100
明星高中 100

調整詞彙權重，將關鍵詞彙給予較高權重，以利斷詞分析。

統計資訊

54

最大字數



103121

總字數



0

最小字數



6

平均字數



任務結果

Show 10 ▾ entries

Search:

| system_id ▲ | sentence_idx ◆ | result ◆ |
|-------------|----------------|--|
| 1 | 1 | [不好意思] |
| 1 | 2 | [看到, 這篇, 一點, 同情, 都, 沒有] |
| 1 | 3 | [社工, 整天, 抱怨, 薪水, 沒, 調漲] |
| 1 | 4 | [來, 看看, 頭目, 說, 什麼] |
| 1 | 5 | [#, 1WPHHD4G, (, Gossiping,)] |
| 1 | 6 | [[, 新聞,], 蔡稱, 社工, 調薪, 不, 重要, 實踐, 才, 重要] |
| 1 | 7 | [只會, 說, 社會, 安全網, 、, 前朝, 遺毒] |
| 1 | 8 | [也不, 看, 現在, 誰, 執政] |

檢查一下中文斷詞結果，確認是否需要進行其他關鍵字權重的調整。

將字串轉換成詞彙之後，清除與主題無關詞彙，清除停用字參數設定大多為系統預設，在白訂義停止詞共選擇了12個詞彙進行清除動作(討論是否有影響情緒的詞彙)。

清除停用詞 (7)

| 參數設定 | Input - 6 | 任務結果 |
|-------------------------------------|---|------|
| 語言 * <div>Chinese</div> | 使用預設停止詞 <div>是</div> | |
| 是否清除單字元 ⓘ <div>是</div> | 是否轉為小寫英文 <div>是</div> | |
| 清除英文字母 * <div>是</div> | 清除數字 * <div>是</div> | |
| 清除換行符號 * <div>是</div> | 清除特殊標點符號 * <div>是</div> | |
| 清除html tag * <div>是</div> | 自定義停止詞 <div>一直 是不是 一下</div> | |

一直
是不是
一下
有人
一次
都是
也是
有沒有
一堆
一些
部分
一點

統計資訊

52826

清除總數



任務結果

Show 10 ▾ entries

Search:

| system_id | sentence_idx | result |
|-----------|--------------|------------------------------|
| 1 | 1 | [不好意思] |
| 1 | 2 | [看到, 這篇, 同情] |
| 1 | 3 | [社工, 整天, 抱怨, 薪水, 調漲] |
| 1 | 4 | [看看, 頭目] |
| 1 | 5 | [] |
| 1 | 6 | [新聞, 蔡稱, 社工, 調薪, 重要, 實踐, 重要] |
| 1 | 7 | [只會, 社會, 安全網, 前朝, 遺毒] |
| 1 | 8 | [也不, 現在, 執政] |



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。

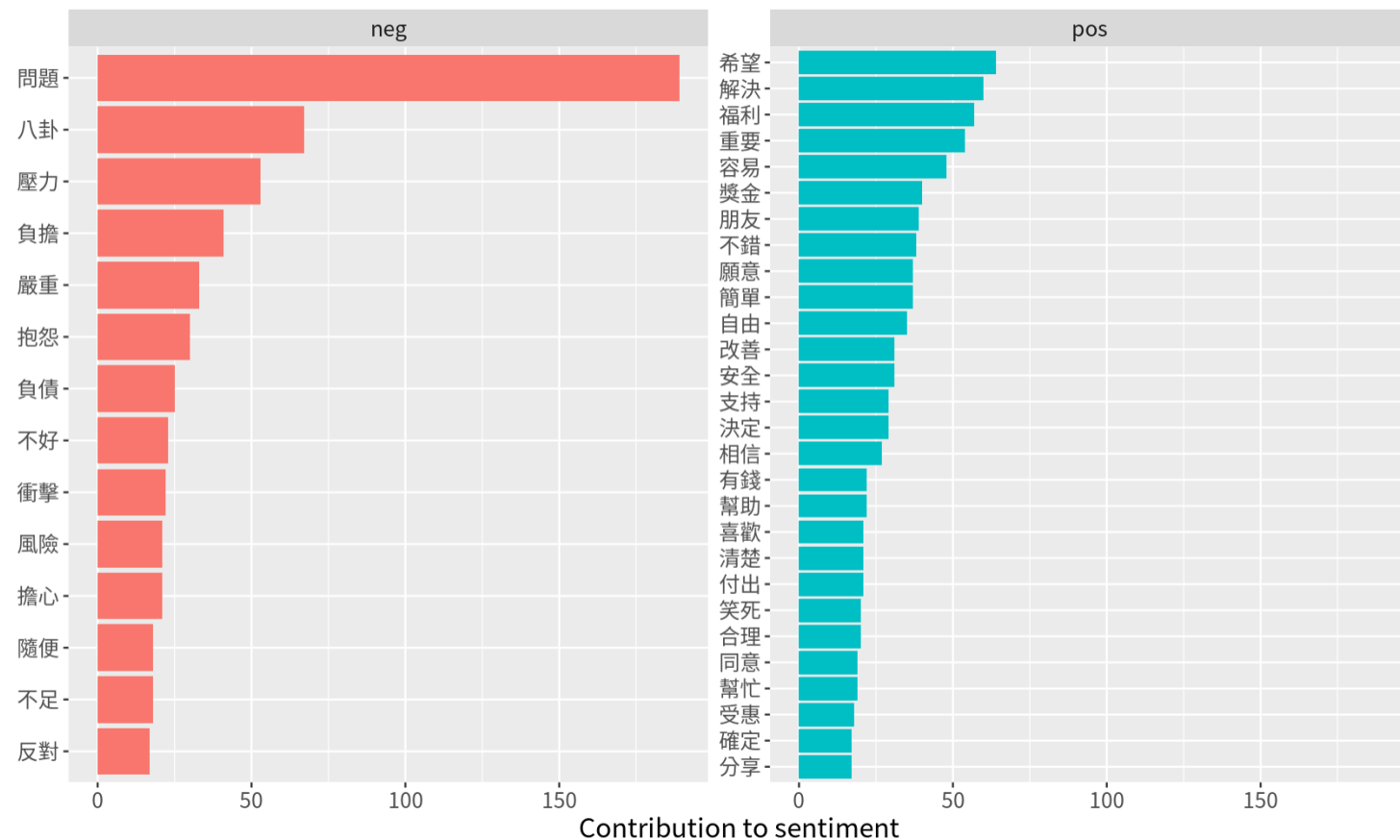
根據詞頻計算結果，對於薪資、薪水、工作等等的相關詞彙有相當高的討論。

Showing 1 to 10 of 100 entries Previous 1 2 3 4 5 ... 10 Next

Lexicon Based 情緒分析 (9)

視覺化

情緒貢獻長條圖



由圖表可看出，負面詞彙有「問題」(非常高)、「八卦」、「壓力」、「負擔」、「嚴重」……等等。
正面詞彙則有「希望」、「解決」、「福利」、「重要」、「容易」……等等。

☰ Corpus Based 情緒分析 (26)



參數設定

Input - 7

任務結果

統計資訊

14877

負向情緒數



0

正向情緒數



1570

總數



任務結果

Show entries

Search:

system_id

sentence_idx

sentiment_value

1

1

0.419

1

2

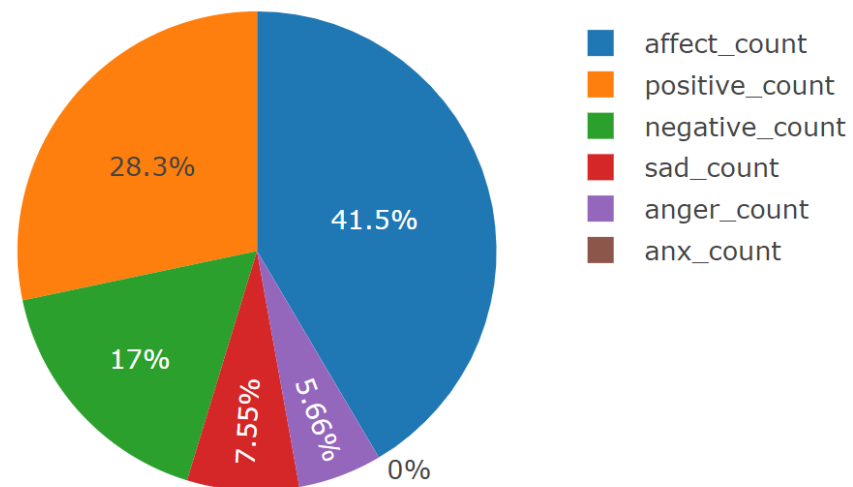
0.955

參數設定

Input - 7

任務結果

情緒分佈



全螢幕瀏覽

點我下載完整CSV資料

點我下載完整Rdata

每日情緒 (19)

參數設定

Input - 16

任務結果

統計資訊

533

群組數量



任務結果

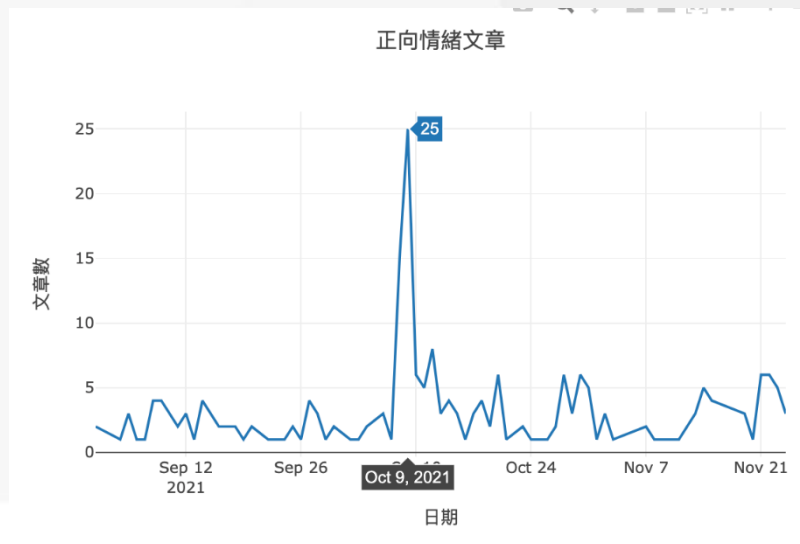
Show 10 entries

Search:

| system_id | artDate | sentiment_value |
|-----------|------------|-----------------|
| 1 | 2021-09-01 | -1 |
| 10 | 2021-09-02 | -3 |

使用視覺化儀表板，產出每日情緒折線圖。
可看出情緒值之最高峰與最低峰。





負向與正向文章數在10/9均達最高峰。

產出文字雲，探討關鍵字為「薪資」「基本工資」「物價」「台灣」……等等。

搜尋相關新聞，確實有關於工資調漲等的討論。



明年基本工資漲5.21%！月薪2萬5250元、時薪168元

2021-10-08 14:17 聯合報 / 記者葉冠妤／台北即時報導



PART 4

結論

結論

- 以上本組關於調薪相關的討論聲量分析，以初步情緒分析結果來說，負面情緒較高。其中最高的關鍵字為「問題」，但對於各篇文章中的「問題」，是否代表的都是負向的，本組認為應還有探討空間。可能後續透過字典的選擇或工具的使用上做調整，能夠更準確的判讀出情緒。
- 正向的部分關鍵字最高為「希望」、「解決」與「福利」，可知民衆對於調薪的正向討論可能是充滿希望、期待解決問題以及社會福利等等。



基本工資



— END —
THANK YOU

