第八組

鍾承玫N074220001  
羅育琦N094220002  
楊雯珺N094220006  
周昆穎N094220007  
陳冠汶N094220012  
 李明章 N094220023  
 鄭雅丹 N094220027

社群媒體分析

第二次讀書會報告

目錄

[第１章 探討主題 2](#_Toc89563394)

[第２章 分析說明 3](#_Toc89563395)

[第１節 PTT爬蟲 4](#_Toc89563396)

[第２節 中文斷詞 5](#_Toc89563397)

[第３節 清除停用詞 6](#_Toc89563398)

[第４節 詞頻計算 7](#_Toc89563399)

[第１項 詞頻計算 7](#_Toc89563400)

[第２項 Ngram的詞頻計算 7](#_Toc89563401)

[第５節 情緒分析 7](#_Toc89563402)

[第１項 Lexicon based 7](#_Toc89563403)

[第２項 Corpus based 8](#_Toc89563404)

[第３項 LIWC 9](#_Toc89563405)

[第６節 每日情緒 9](#_Toc89563406)

[第１項 Lexicon based 9](#_Toc89563407)

[第２項 Corpus based 10](#_Toc89563408)

[第３項 LIWC 11](#_Toc89563409)

[第４項 關聯分析7/10 11](#_Toc89563410)

[第５項 關聯分析9/15 12](#_Toc89563411)

[第６項 關聯分析10/16 12](#_Toc89563412)

[第７項 關聯分析11/26 12](#_Toc89563413)

[第７節 文字關係分析 13](#_Toc89563414)

[第１項 計算共現矩陣及字詞網路圖 13](#_Toc89563415)

[第２項 計算相關性矩陣及字詞網路圖 14](#_Toc89563416)

# 探討主題

自從今年五、六月份開始，汽車開始缺晶片，民眾買不到浴缸，在美國購買一輛車，相較於一月，價格增加了22%，汽油的價格也增加了27%。其實那段時間的短缺是經濟復甦的訊號。

專家分析造成通貨膨脹的因素有三：

1. 因為消費者對汽車、家具、家用電器等商品的需求激增。消費者添購這些商品，是為了在新冠疫情期間，讓居家隔離生活更美好、過得更舒適。

2. 因為新冠疫情中斷了部分商品的全球供應鏈，如晶片的短缺就限制了汽車的生產製造。全球航運業的中斷，如停飛航線或塞港，也使得全球供應鏈的中斷更為嚴峻。

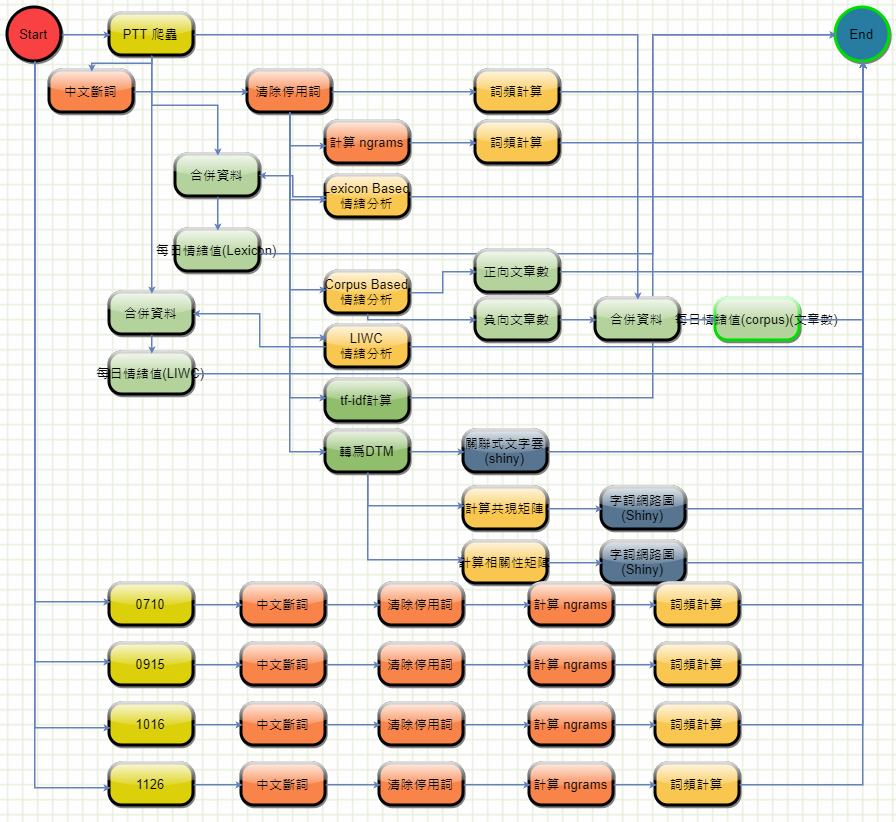
3. 因為服務業供不應求，導致價格開始反彈，這也可能是最重要的因素。隨著疫情慢慢降溫，消費者比就業者更快湧進餐廳、酒吧、美容理髮店或其他服務業。

因此本組將探討民眾對於通貨膨脹所帶來的正面及負面影響。

# 分析說明

工作流程名稱：tw2

建立時間：2021/12/04 17:06:26



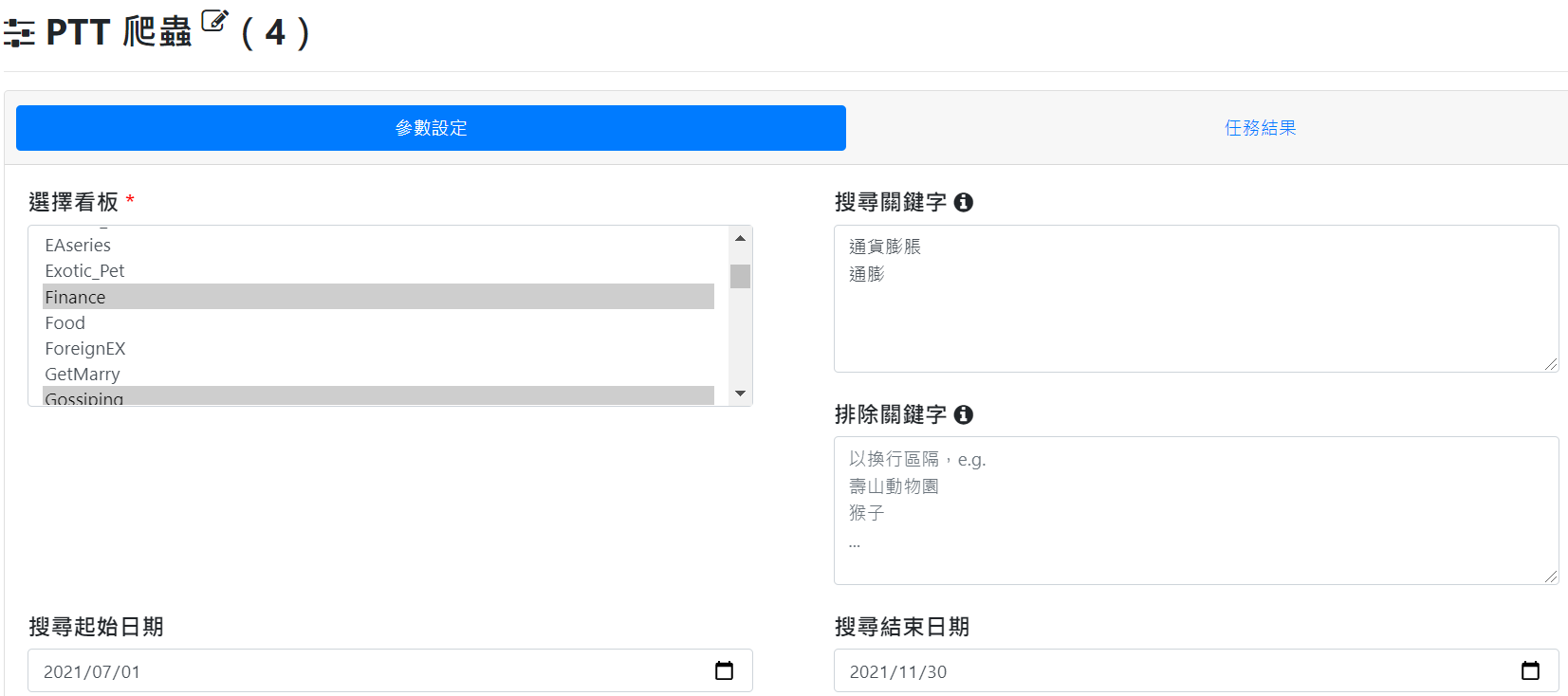
## PTT爬蟲

日期：2021/07/01 ~ 2021/11/30

資料來源看板：Finance(財經)、Gossiping (八卦)

關鍵字：通貨膨漲、通膨

共 **1101** 筆貼文。

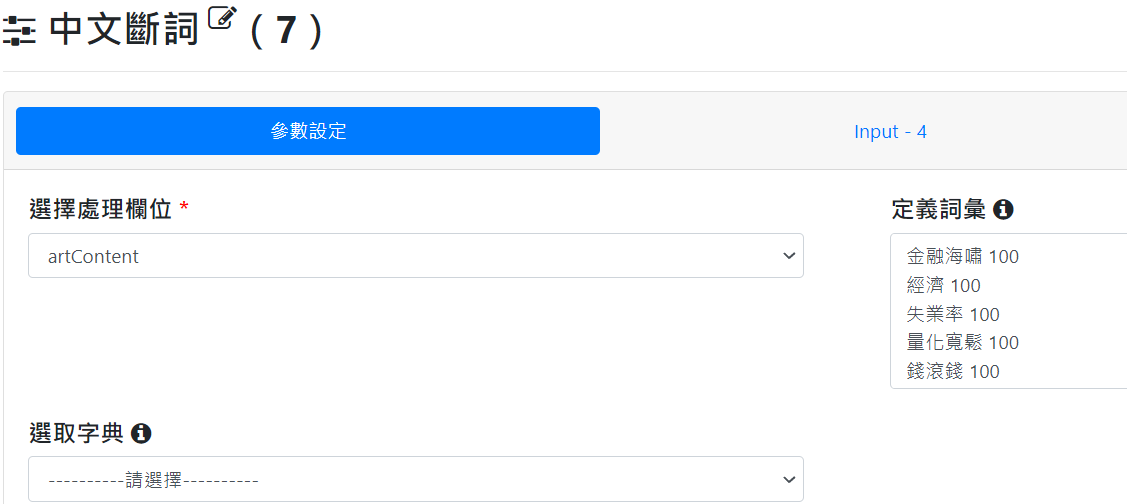


## 中文斷詞

詞彙權重：都是設定100，詞彙如下

金融海嘯、經濟、失業率、量化寬鬆、錢滾錢、實質購買力、惡性通膨、物價指數、國民所得、主計處、印鈔票、貿易協議、就業市場、理財專員、大盤指數、武漢肺炎、新冠肺炎、標準普爾、那斯達克、通貨膨脹、主計總處、聯合新聞網、工藤優作、房地合一、中央銀行、南韓央行

共 **259,645** 個詞。



## 清除停用詞

清除英文字母：設定為「是」。清除網址URL等英文…

自定義詞停止詞：

通膨、台灣、美國、中國、政府、一堆、一下、一直、新聞、記者、這種、有沒有、認為、造成、根本、表示、標題、出現、所以、連結、一點、備註、繼續、之後、越來越、部分、內文、主要、網址、最後、知道、完整、不用、署名、報導、媒體、來源、昨天、新增、即時、中央社、潘姿羽、請放、違者、死亡、復原、綜合、外電、聯合新聞網、八卦、問卦、工藤優作、文章、刪除、確診、日電、至少、看看、需要、好像、完全、導致、幾個、成為、看看

清除 **154,073** 個詞。



## 詞頻計算

### 詞頻計算

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。

可以看到「現在、物價、房價、經濟、問題、上漲、貨幣、漲幅…等」字詞出現的頻率較高，從這可觀察到通貨膨脹相關議題正在發生及發酵的趨勢。

一張含有 文字 的圖片

自動產生的描述

### Ngram的詞頻計算

一張含有 文字, 白板 的圖片

自動產生的描述

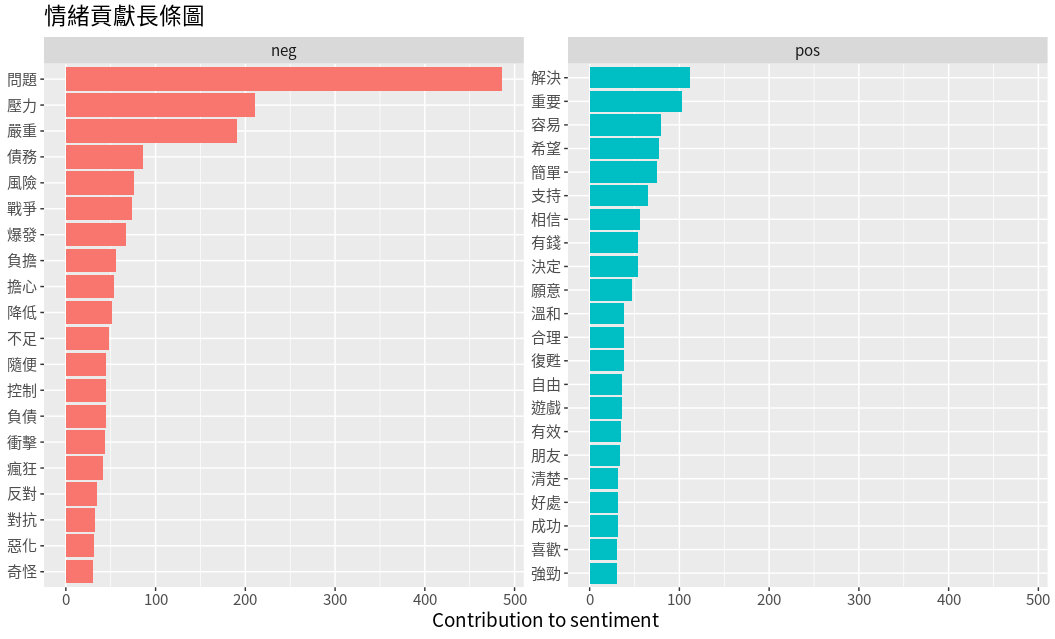
## 情緒分析

### Lexicon based

從情緒貢獻長條圖看出，負向情緒值明顯高於正向情緒值。

負向情緒詞：問題、壓力、嚴重、債務、風險…

正向情緒詞：解決、重要、容易、希望、簡單…



一張含有 文字 的圖片

自動產生的描述

### Corpus based

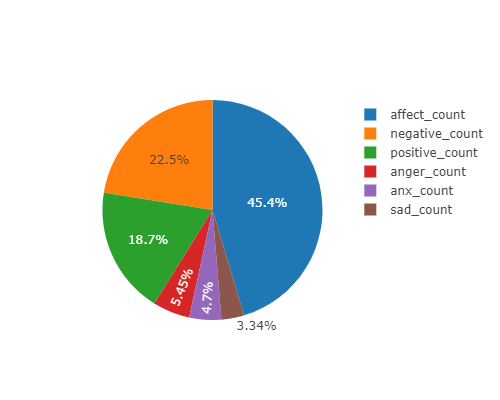
負向文章數(822筆)明顯高於正向文章數(279筆)。

sentiment\_value以0.5區分正負向。大於0.5表示正向，小於0.5表示負向。

一張含有 文字 的圖片

自動產生的描述

### LIWC

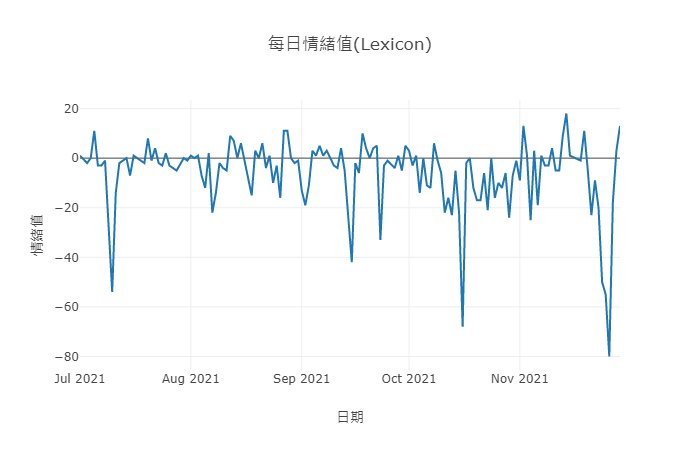


## 每日情緒

### Lexicon based

針對每日情緒做正向與負向情緒統計，並透過「視覺化儀表板」中的「折線圖」進行呈現。

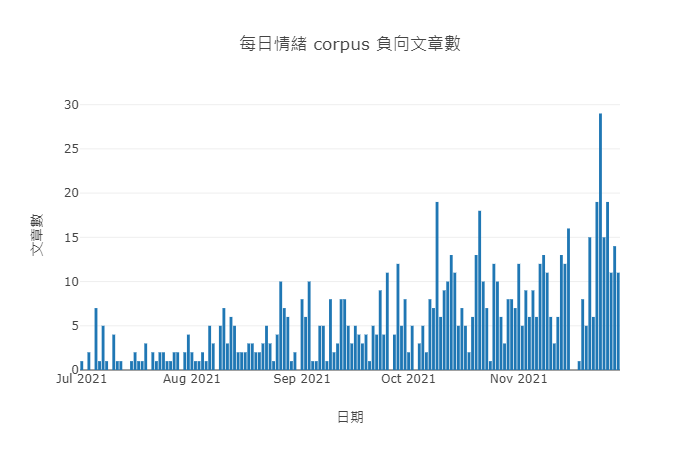
負向情緒值在**7/10、9/15、10/16、11/26**都偏高，最高落在**11/26(-80)**。



### Corpus based

針對每日負向文章數統計，並透過「視覺化儀表板」中的「長條圖」進行呈現。

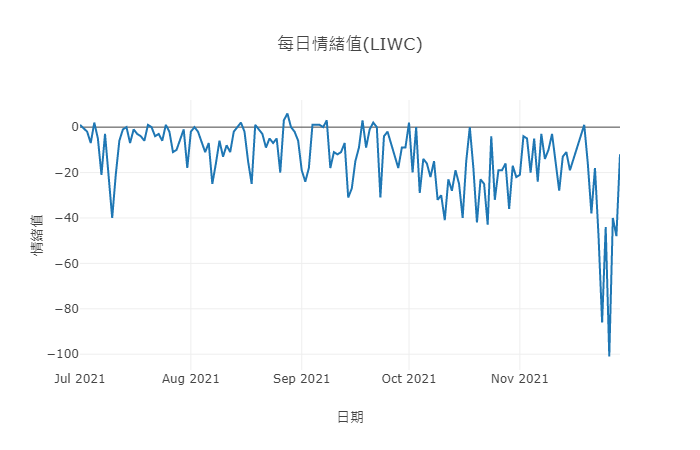
負向文章數在**10/9、10/21、11/24**都偏高，最高落在**11/24(29筆)**。



### LIWC

針對每日情緒做正向與負向情緒統計，並透過「視覺化儀表板」中的「折線圖」進行呈現。

負向情緒值在**7/10、10/16前後、11/26**都偏高，最高落在**11/26(-101)**。



### 關聯分析7/10

從上述每日情緒分析結果，探討7/6至7/10在通貨膨脹這個關鍵字下，負向情緒值偏高的可能原因。

Need to write something

一張含有 文字 的圖片

自動產生的描述

### 關聯分析9/15

從上述每日情緒分析結果，探討9/11至9/15在通貨膨脹這個關鍵字下，負向情緒值偏高的可能原因。

Need to write something

一張含有 文字, 白板 的圖片

自動產生的描述

### 關聯分析10/16

從上述每日情緒分析結果，探討10/12至10/16在通貨膨脹這個關鍵字下，負向情緒值偏高的可能原因。

Need to write something

一張含有 文字 的圖片

自動產生的描述

### 關聯分析11/26

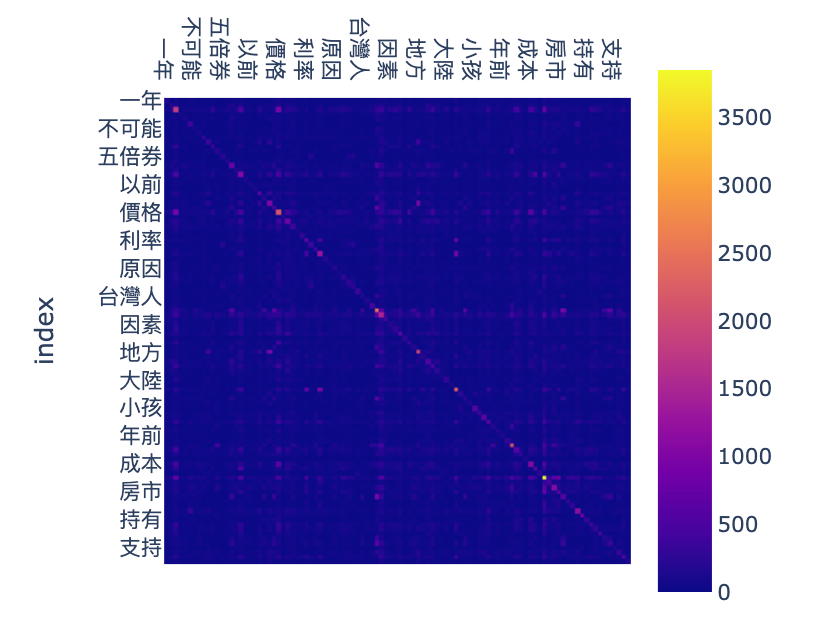
從上述每日情緒分析結果，探討11/22至11/26在通貨膨脹這個關鍵字下，負向情緒值偏高的可能原因。

Need to write something

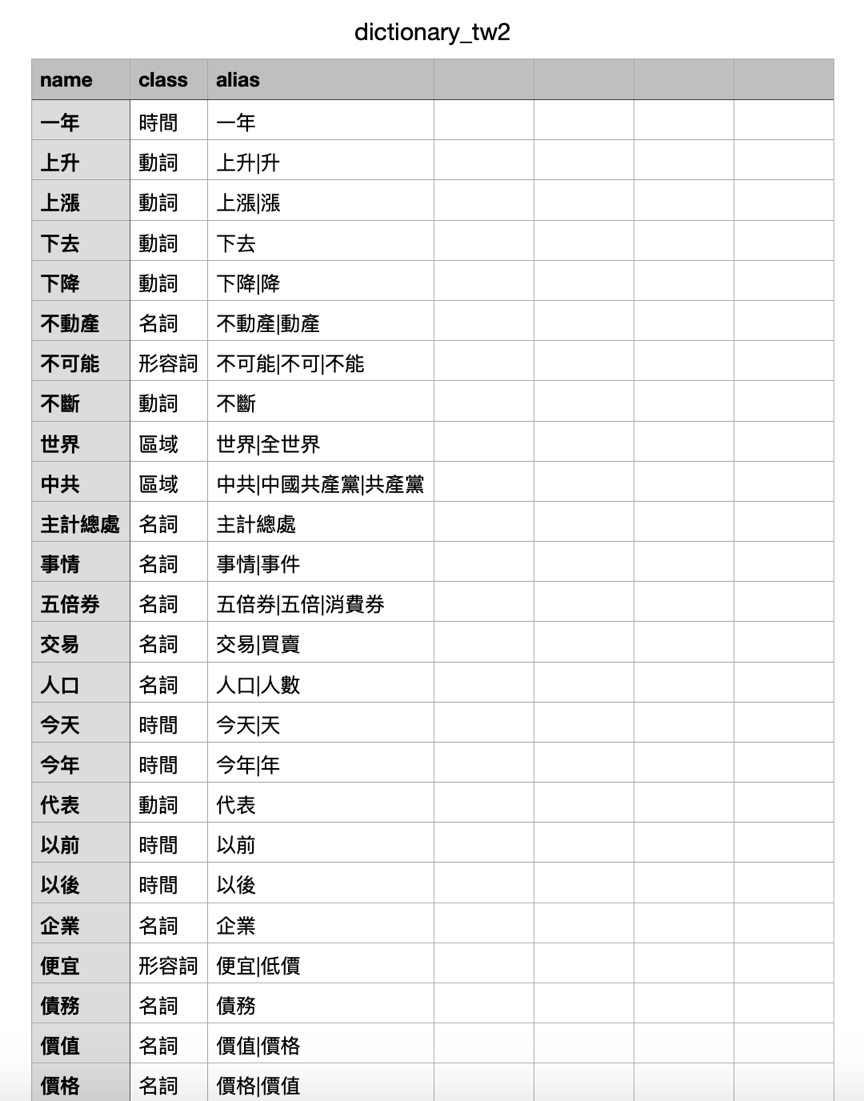


## 文字關係分析

### 計算共現矩陣及字詞網路圖

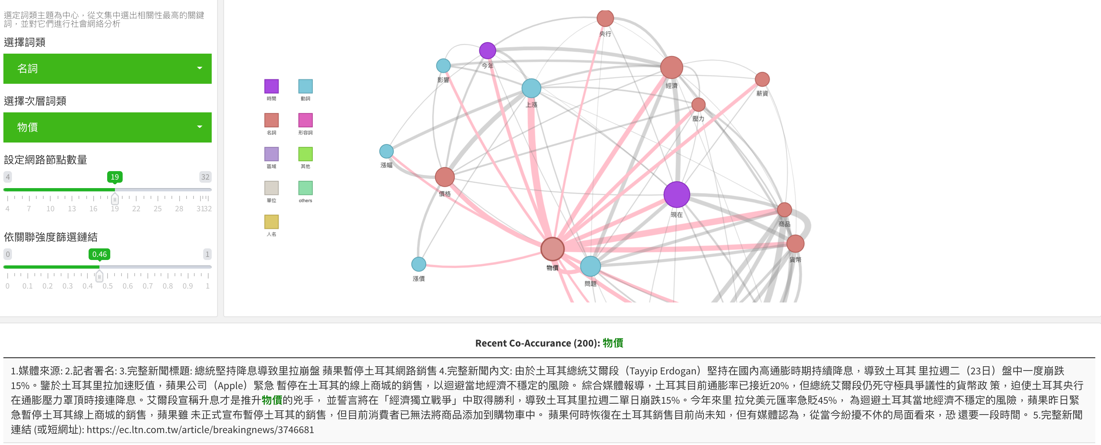


自製字典分類（共現矩陣匯出資料），分類：時間、動詞、名詞、形容詞、區域、單位、人名、其他。

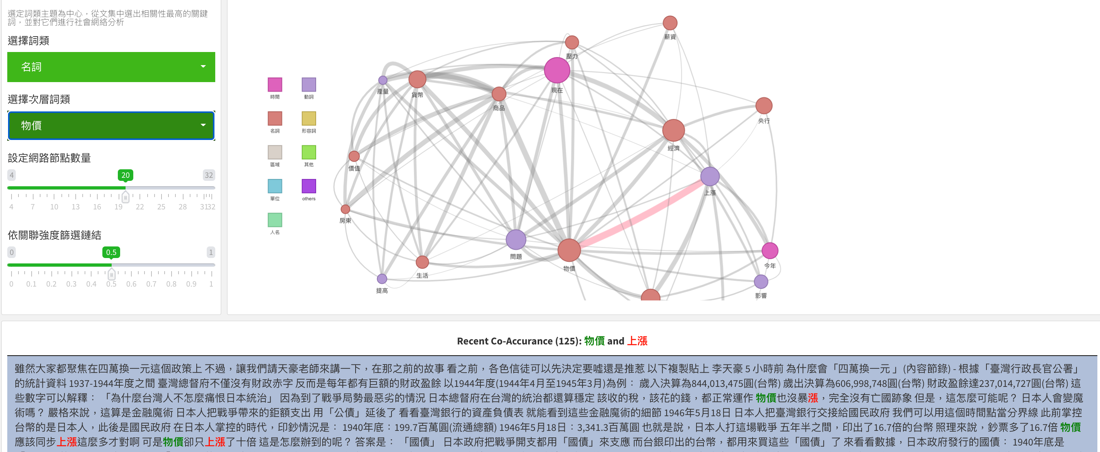


字詞網路圖shiny（選擇自製的字典）

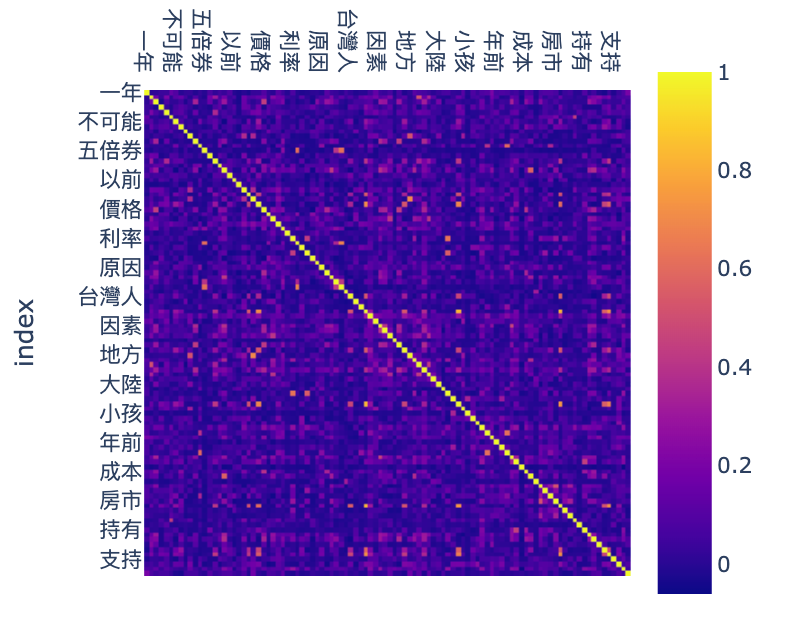
「物價」的字詞網路圖，從圖片可看出與物價關聯性強度的字詞，底下也可以看到相關性的文章。



點選節點間的連接鍊可以看出兩兩相關性的文章，例：「物價&上漲」

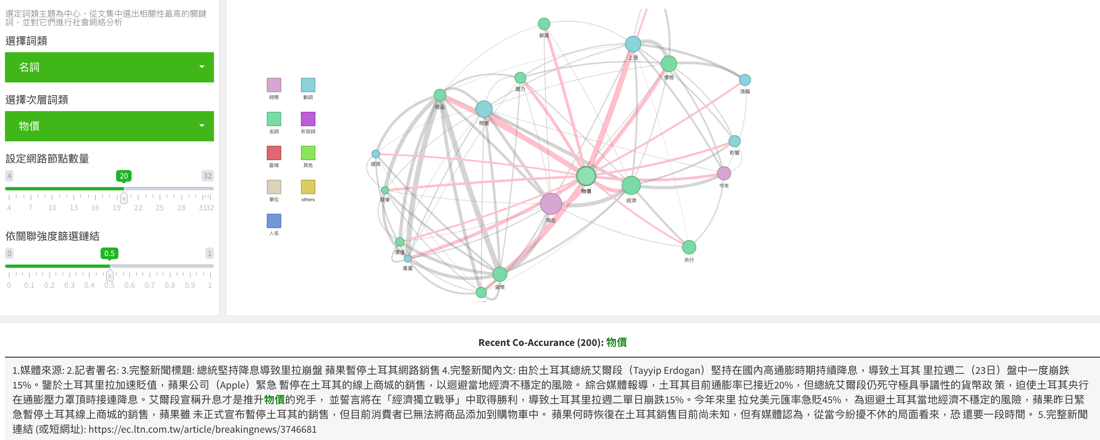


### 計算相關性矩陣及字詞網路圖

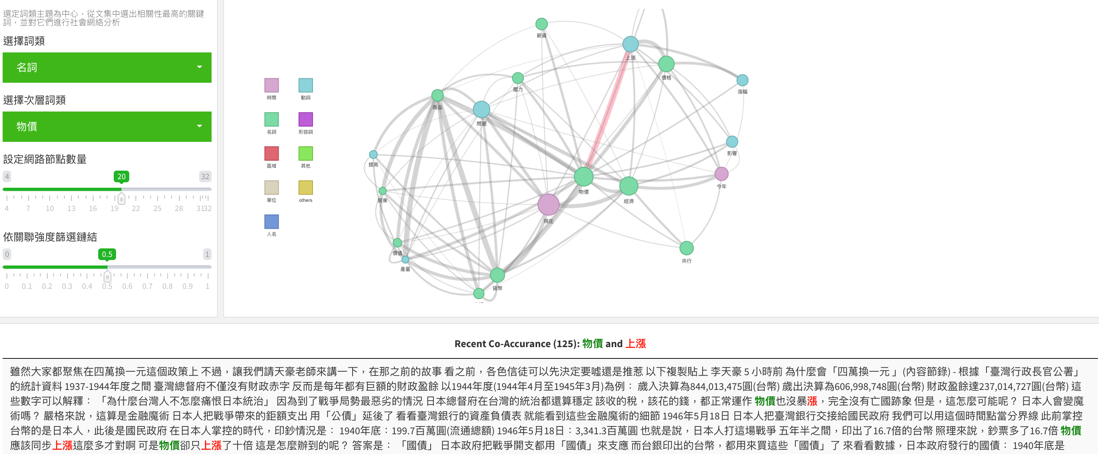


字詞網路圖shiny（選擇自製的字典）

「物價」的字詞網路圖，從圖片可看出與物價關聯性強度的字詞，底下也可以看到相關性的文章。



點選節點間的連接鍊可以看出兩兩相關性的文章，例：「物價&上漲」



**第3項 關聯性文字雲shiny**



從圈圈裡面可以看出大家在ppt裡討論的議題，如果把字一起唸會變成，「現在嚴重問題解決薪水」、「市場央行政策升息利率，房價房市打房」、「時代買房、房子買不起，現金貸款房貸」、「國際國內物價上漲，價格漲幅、成本原物料漲價」

