

## 我國侵害名譽權慰撫金損賠裁判文本分析

公衛一 陳昱君、法律四 陳宛瑩、  
政治四 鄭亦君、職醫所三 黃宇丞

### 壹、研究動機

日常生活時有耳聞民眾因細故在公開場所口出穢語，衍生後續一連串爭訟相關問題之社會事件，在該種情況下，過去網路上流傳一張號稱經法官認證的「罵人價目表」，在新聞媒體廣泛報導、推波助瀾下，漸廣為人知，惟仔細檢閱相關該價目表及新聞報導的內容，其存在許多根本性的問題，如資料的取用來源以及法官的認證真實性皆無從考證，在這個數據科學發展迅速，講究實事求是的世代裡，我們認為可以透過本學期教授的文本分析技巧，進一步探究該資料與新聞報導的真實性。

### 貳、研究目的

本組想藉由本次報告，驗證網路上「罵人價目表」相關新聞等資訊的真偽，並藉以瞭解民事上因穢語涉訟的裁判分布概況、判賠比率；我國人最常使用哪些穢語致生侵害名譽之疑慮，是否有因地區不同而有所差異；最後，試圖以文本分析方法，初探法院再酌定判賠金額時，可能考慮的相關因素。

### 參、文本分析流程

#### 一、資料搜集

我國法院裁判書資料可自司法院法學檢索系統等網站取得，又由於本次研究需要預先收錄相關裁判完成資料庫建置，本組在報告前預先取得已使用網路爬蟲取得之 2008-2018 年間我國各地方法院裁判資料，共計 53 萬餘筆，經過多次嘗試後設定相關關鍵字，從中取出 4132 件符合本次研究標的之樣本。關鍵字的設定，整理如下表所示：

欄位	內容
案件類型	各地方法院+民事+裁判
期間	2008 / 1 / 1 - 2018 / 12 / 31
類別	裁判
案由	損害賠償
裁判全文	公然侮辱
條文	民法第 195 條

在案件類型上，選擇各地方法院裁判的原因是相較高等或最高等其他審級裁判，地院的裁判具有普遍性，能夠反映整體社會使用穢語涉及損害他人名譽權時法院裁判的普遍樣貌；又該類型裁判雖與刑法之公然侮辱有涉，惟本文探討的項目應與民事侵害名譽權非財產上損害賠償有關，故設定為民事裁判類型。經學生觀察判決案由，雖各法院在設定時未有統一規範，惟至少會存在「損害賠償」一詞，故本組將其作為文本選取關鍵字之一；在裁判全文處，為了使侵害名譽權之相關裁判類型包含「穢語」的使用，我們納入了刑法上「公然侮辱」作為關鍵字（經觀察結果所得）；條文部分，由於名譽權屬於人格權之範疇，就人格權的非財產上損賠主要規範在民法第 195 條中，故將其設為關鍵字的一部分。

設定好關鍵字後，我們將符合條件的裁判 json 檔整理成一個資料集，方便後續資料處理使用。

## 二、 資料清洗

### (一) 切分基本欄位

在資料清洗環節，我們先將全文切分成幾個主要欄位，如裁判全文、法院意見、日期、法院名稱、裁判案號等。

### (二) 清洗資料以合乎分析需求

進一步以正則表達式、if-else 條件設定清洗資料使其合乎分析需求，如

設置判賠金額、裁判結果（是否判賠）、各案件涉及髒話內容欄位，儲存初步清洗好的資訊

### (三) 組建語料庫

上一步驟麟話欄位的切分方法是，使用裁判中出現「」（上下引號）部分，將其內文資訊進行提取後，納入一個新的欄位，惟在此過程中會收錄到不屬於穢語的語彙，加上結巴斷詞無法判斷如閩南語等音譯麟話的使用，故我們先行查閱各欄位中提及的麟話，建制語料庫。

#### (四) 完成資料清洗並檢查

完成上述步驟後，檢視資料庫內容是否存在雜訊或其他問題，進行細部修正，例如：清洗穢語欄位的雜訊（如下圖），處理後抓取到的相關穢語語彙數明顯高於未處理時。

```
[ ] import re
f = open(ppp+"期末報告/"+'髒話語料庫.txt')
d = f.read()
d = d.split('\n')

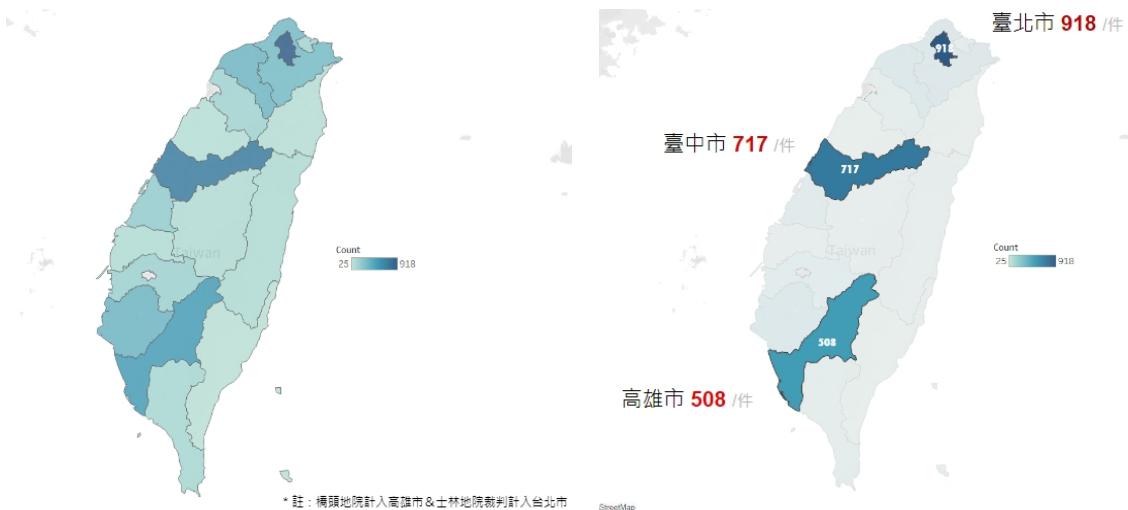
def clean(a):
    a=re.sub('0000000000|' , '' , a)
    return a

print('原髒話: ', len(dirty_word))
dirty=[]
for i in dirty_word:
    a = clean(i)
    for j in re.split(';|_|。|\n|\r|\t|!|\r\n|_|、|；', a):
        if j != '':
            if j in d:
                dirty.append(j)
print('處理後髒話: ', len(dirty))
```

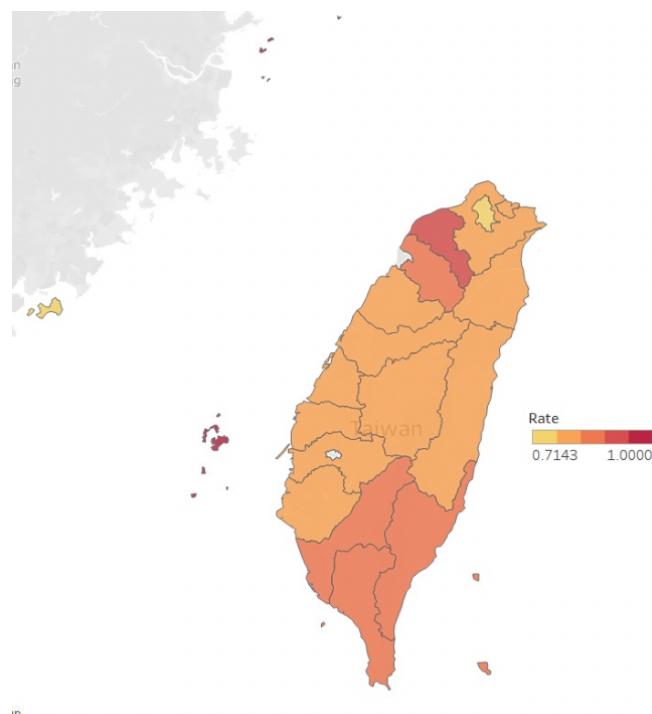
## 肆、 資料詮釋

## 一、 案件分布敘述統計

基於民事訴訟的特性，我們按照我國 22 個地方法院所在地區<sup>1</sup>，加總裁判數量後，可以看出台北、台中、高雄三區涉訟的件數最多，呈現如下圖：



又我們初步將是否判賠的資料予以統計，依區域差異以分層地圖呈現如下，可以看出判賠率多介於 7 成至 9 成間，澎湖地院由於涉訟件數少，判賠率達到 100%，屬於例外情況；我們也可以觀察到，台北地區的涉訟件數雖然最多，然而判賠率卻最低，未來或可進一步按區域不同探討涉訟與判賠的具體原因。



<sup>1</sup> 除高雄地區有加計橋頭地院裁判；台北加計士林地院裁判資料外，其餘地區為一地一法院。

## 二、穢語使用情況視覺化

我們以建置好的髒話語料庫將裁判中提及的穢語製作文字雲圖，可以發現因穢語侵害名譽權的相關裁判中，以「幹你娘」提及的次數最多，共計 1021 件裁判，佔整體裁判 24.7%。



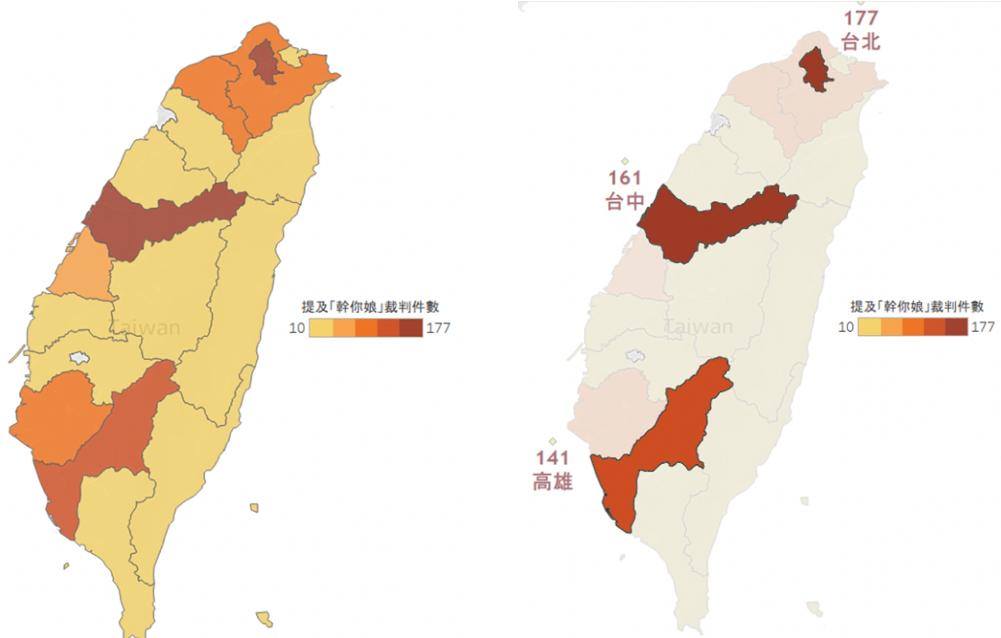
我們繼續進一步檢視各地區所涉髒話類型。<sup>2</sup>各地區髒話類型的統計結果可以看到各地區的涉訟的髒話類型極為相似。不過，從各地區的髒話類型文字雲圖可以發現各地區涉訟的髒話類型還是有些許差異。細看南北相關詞彙，南部有較多以台語發音髒話，且相較其他各區，東部常見涉訟髒話多了「死番仔」、「番仔」這類詞彙。

<sup>2</sup> 北中南東與離島的定義為：北部包含「臺北、新北、士林、新竹、基隆、桃園」地院裁判、中部包含「苗栗、臺中、彰化、南投」地院裁判、南部包含「高雄、臺南、雲林、嘉義、屏東、澎湖」地院裁判、東部包含「花蓮、臺東」地院裁判、離島包含「金門、連江」地院裁判。

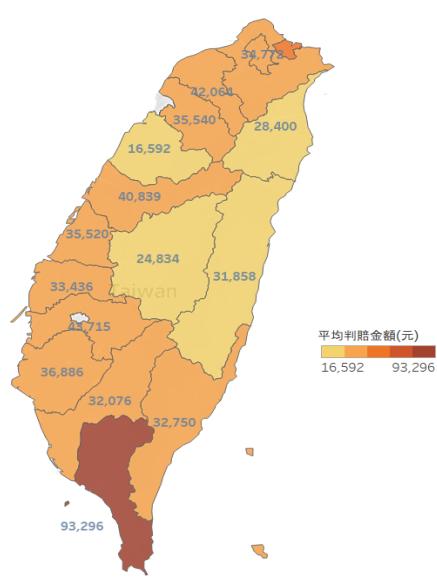


	全部		北部		中部		南部		東部		離島	
第一名	幹你娘	1870	幹你娘	706	幹你娘	419	幹你娘	606	幹你娘	67	幹你娘	9
第二名	幹	857	不要臉	470	幹	211	幹	241	垃圾	19	爛貨	6
第三名	不要臉	833	幹	360	不要臉	176	不要臉	161	幹	17	賤貨	5
第四名	垃圾	355	王八蛋	144	神經病	56	垃圾	133	幹你娘雞巴	16	笨蛋	3
第五名	神經病	273	垃圾	139	垃圾	49	瘋女人	86	神經病	12	神經病	3
原諱話樣本數	4132		1797		1010		1125		96		8	
處理後 諱話樣本數	8935		3775		2002		2661		230		29	

為了驗證諱話價目表的真實性，我們選擇最多裁判提及的「幹你娘」一詞進行分析，在案件數上，仍以案件總數最多的台北、台中、高雄地區涉訟案件最多。

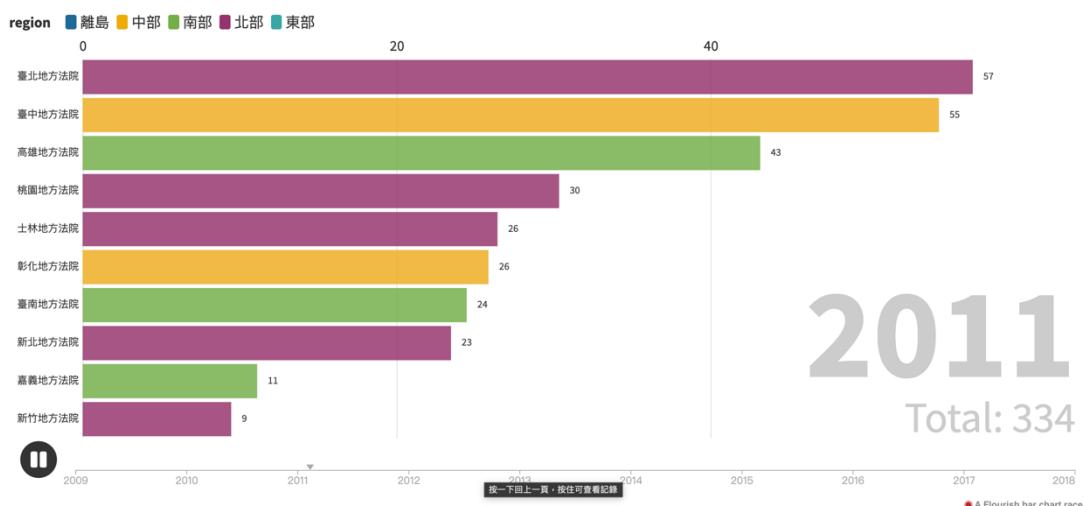


進一步觀察各地方法院判賠金額，可以發現各地區判賠金額平均介於 3 萬至 4 萬之間，屏東地院平均判賠金額最高，高達 93,296 元，而苗栗地院平均判賠金額最低，為 16,592 元。屏東地院判賠金額高的可能原因在於該院裁判由於多非屬單一穢語使用，加上一定的研究限制所致，而無從得到屏東地院偏好判賠金額最高的結論。而苗栗地院判賠金額低的可能原因整體裁判法院尚有觀察個案情境、被告資力、原告自身行為而做裁判。從結果上來看，價目表上「幹你娘」的價格與實際平均金額差異甚大，因此，網路上流傳的髒話價目表的可信度有待釐清。



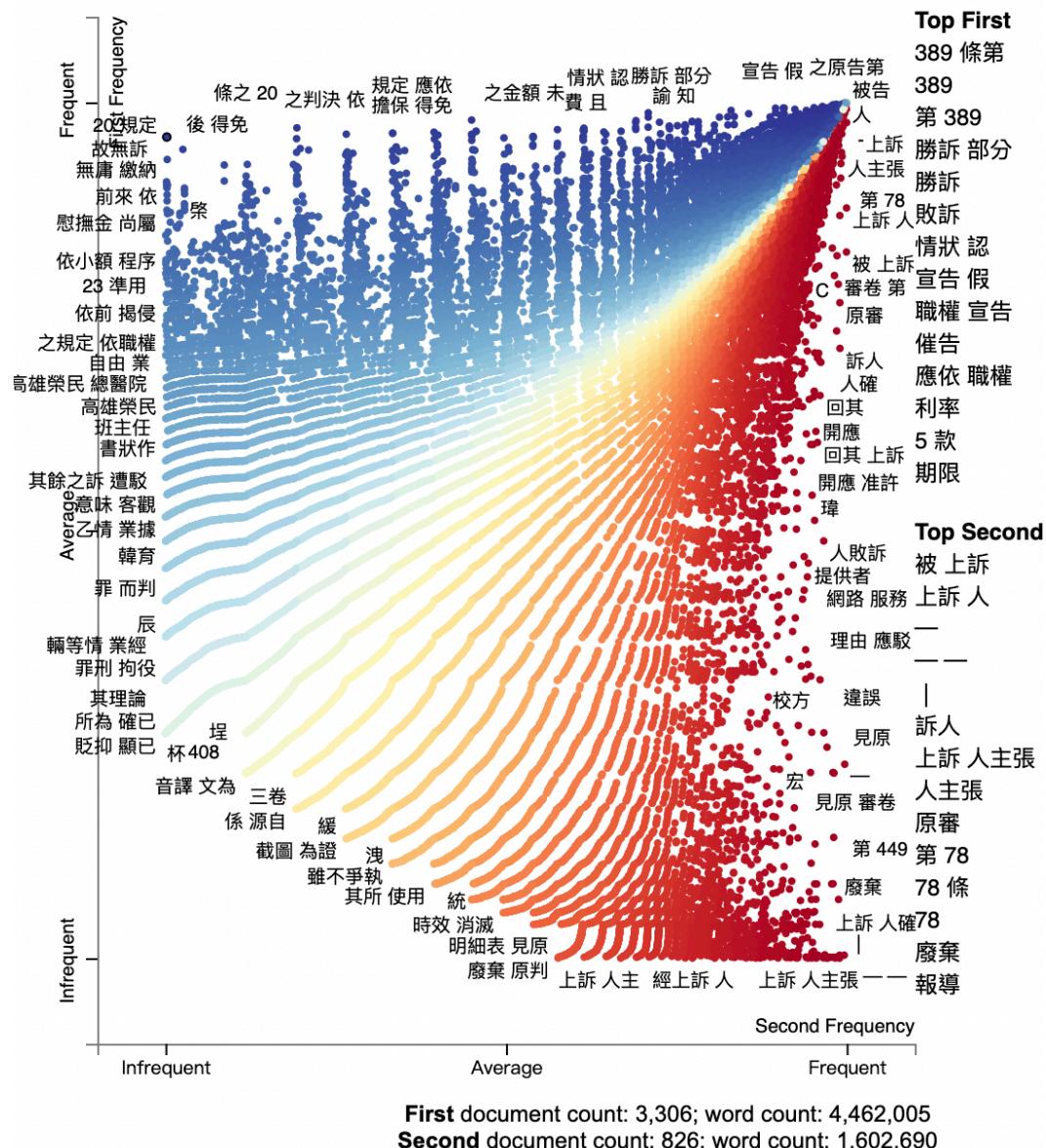
台灣罵人價目表	
用詞	罰款金額
1 幹	判賠 0 元
2 更年期到了	判賠 2 千
3 賤貨	判賠 3 千
5 米蟲	判賠 6 千
6 王八蛋	判賠 1 萬
7 不要臉的髒東西	判賠 2 萬
8 你去吃原啦	判賠 2 萬 5 千
4 敗類	判賠 3 萬
9 智障	判賠 5 萬
10 賤人就是嬌情	判賠 5 萬
11 頭殼裝屎	判賠 5 萬
12 白痴	判賠 5 萬 5 千
13 娘子	判賠 6 萬
14 下流	判賠 6 萬
15 幹 X 娘	判賠 8 萬
16 死番仔	判賠 10 萬
17 人渣公務員	判賠 30 萬
18 神經病	判賠 30 萬
19 特殊性關係	判賠 100 萬

為了呈現各地院因「幹你娘」一詞涉訟的變化，我們也製作了動態圖表（<https://wayne-growing.github.io/>），呈現 2008-2018 年間地院裁判件數的變化。



### 三、 裁判書文字散布圖

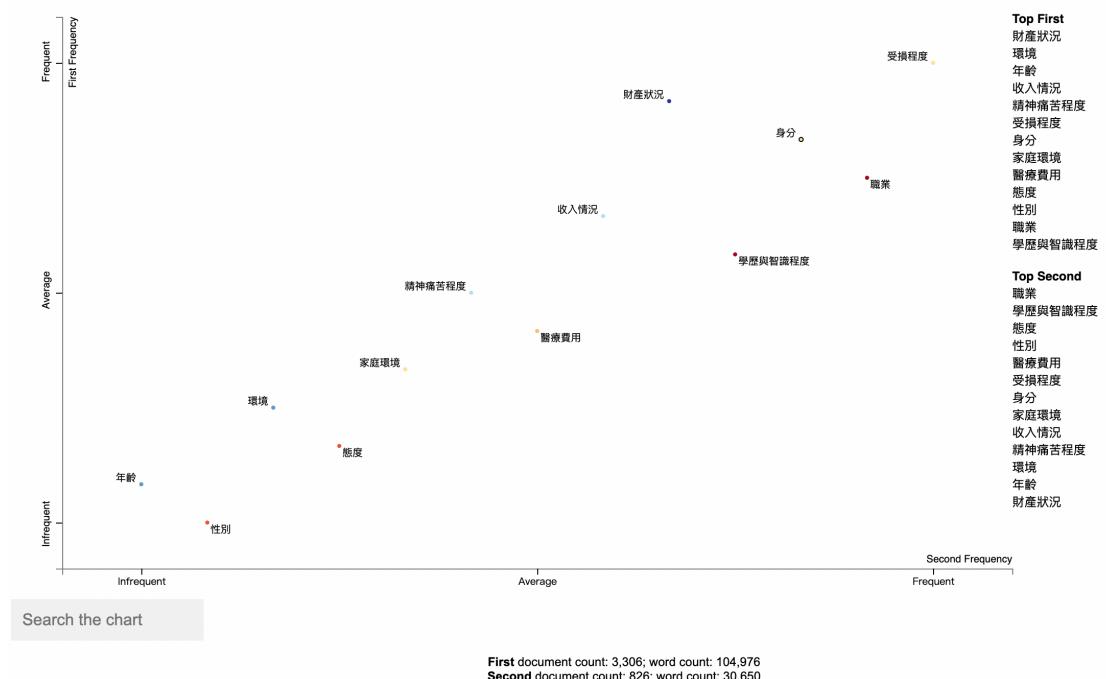
為進一步了解裁判事由，我們製作了文字散布圖，依有無判賠的裁判設定 y,x 軸（first 和 second 區域），將資料庫中所有裁判全文匯入觀察，或許是受到法律用語與一般文字使用的差異，並不能精準地看出相關因素，不過我們可以透過這個文字散布圖搜尋相關因素的用語，以利建置因素的字典。



經過檢視法院裁判結果，我們將其分為財產狀況、學歷與智識程度、身分、精神痛苦程度、職業、家庭環境、收入情況、醫療費用、態度、受損程

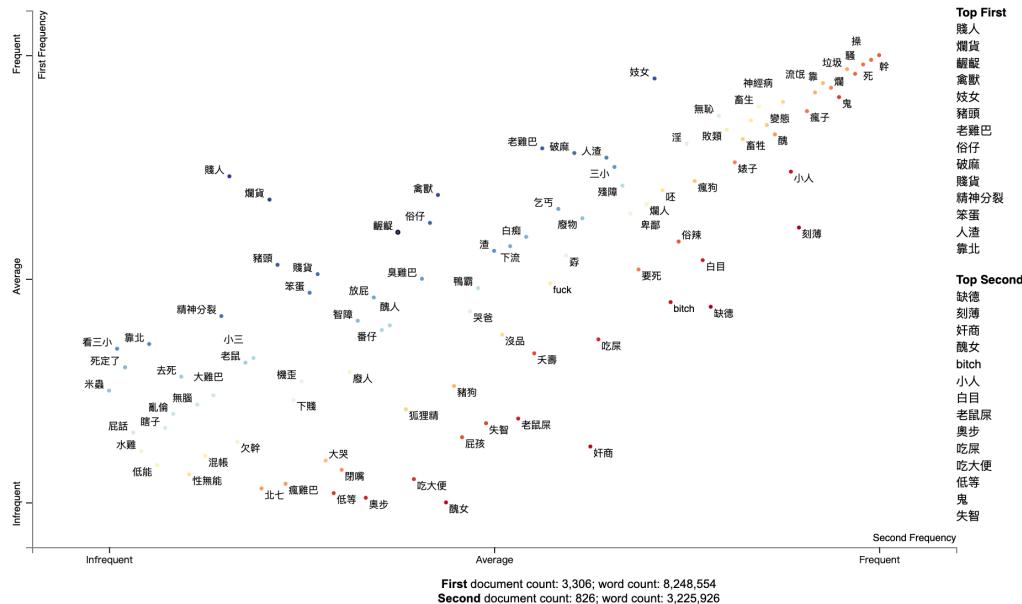
度、環境、性別、年齡等 13 大類，並就擷取的有關詞彙組合成字典，以法院心證理由欄位為標的，製作群組分類圖。由於法院應該會先做成是否構成名譽權損害的心證，再按照相關因素酌定判賠金額，根據圖表顯示的資訊，我們可以發現有判賠的裁判關鍵字前幾名是財產狀況、環境、年齡、收入情況等，其對法院判賠或有一定關聯，後續可以透過統計或機器學習的方法進一步驗證各因素的影響性，該群組散布圖也提供應納入哪些因素建置模型的參考。

```
topic_model = {
    '財產狀況': ['資力', '財產狀況', '財產情況', '經濟狀況', '經濟', '資產', '無財產', '經濟能力', '財產資料', '名下', '經濟情況'],
    '學歷與智識程度': ['畢業', '肄業', '學歷', '國小', '國中', '高中', '高職', '大專', '大學', '專科', '研究所', '五專', '教育', '程度', '教育程度', '知識程度', '智識', '識字'],
    '身分': ['公眾人物', '地位', '身分', '社會上地位', '社會地位', '社會經驗', '人格地位'],
    '精神痛苦程度': ['精神痛苦', '痛苦程度', '痛苦'],
    '職業': ['任職', '失業', '無業', '承包', '職務', '待業', '從事', '工作', '工', '職員', '擔任'],
    '家庭環境': ['家庭狀況', '婚', '子女', '扶養', '家庭生活', '家庭'],
    '收入情況': ['月薪', '年收', '薪資', '收入', '月收', '所得'],
    '醫療費用': ['精神科', '診所', '就診費用', '醫療費用', '看診', '身心科', '診斷證明', '就醫'],
    '態度': ['態度'],
    '受損程度': ['受損之程度', '名譽受損程度', '受損情形', '受損情節', '貶抑程度', '加害情形', '貶損', '受損之狀況', '損害', '加害情形'],
    '環境': ['客觀情境', '動機'],
    '性別': ['性別'],
    '年齡': ['歲', '年次', '高齡']
}
```

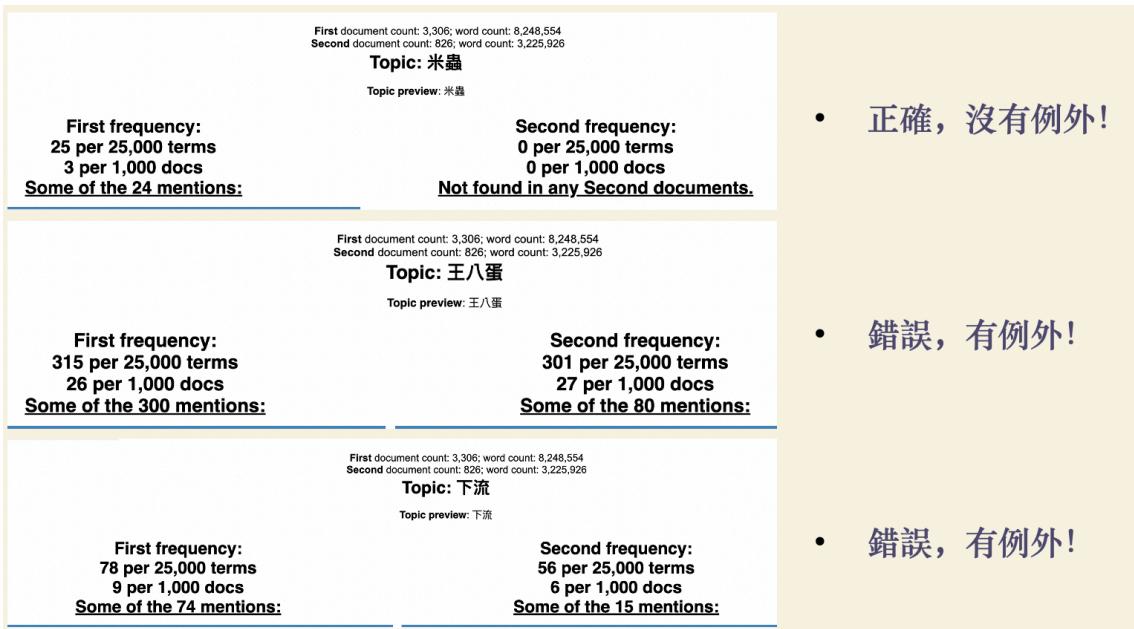


我們也想要知道普遍上法院會認為用何種穢語指涉別人可能構成名譽權侵害，因此，將建置好的髒話語料庫修改為字典形式後，繪製散布圖，從下

表可以看到髒話語料的分佈情況，可以看到法院有判賠的裁判多與性、疾患有關，相較之下，未判賠的用詞似乎比較符合社會通念中的客觀評價，如：缺德、醜女等。（但也有例外，如：bitch一詞似乎與性有關，卻被分類在第二類）



最後，我們進一步驗證罵人價目表中相關新聞的資訊是否正確，以新聞中提及的這段話為例「.....但如罵人「米蟲」、「王八蛋」、「下流」等，雖然不是髒話，卻構成公然侮辱、名譽受損等罪，可判賠新台幣2千到6萬元不等....」（雖然新聞內容與法學實際用語有意，不太精確），驗證文中的關鍵資訊：「米蟲」、「王八蛋」、「下流」是否可判賠？從散布圖搜尋結果中可以看到，「米蟲」與新聞所提的資訊吻合，皆有判賠；惟「王八蛋」和「下流」則有例外。



## 伍、小結

### 一、與報章雜記輿論差異探討

回到一開始的研究目的，我們先說明本次報告與罵人價目表之差異，並驗證、歸納新聞媒體所報導的內容真偽。網路上流傳的「罵人價目表」未能說明其來源並說明出處、樣本母數，同時找不到所謂「經過法官核准」的背書人選，相較之下，本次的報告具體提供了樣本的選擇標準、母數，同時亦有提供各案件裁判的法官人選，能夠初步的認為該研究結果是經過普遍心證後的結論。又報章雜誌所提及：「『米蟲』、『王八蛋』、『下流』等用語皆有構成名譽權損害」，經過驗證，除「米蟲」一詞找不到反例，其餘皆有未判賠的樣本存在，報章雜誌所撰寫的內文顯有可議之處。

而本次報告結果的小發現有以下四點：首先，整體而言，最常涉訟的穢語態樣為「幹你娘」，且不分區域，皆為常見的涉訟事由。第二，髒話價目表認定「幹你娘」判賠金額為八萬，惟實際上多介於3-4萬間，顯有落差，故表格上的資料正確性確有具體研議之處，非全然可信。第三，較顯見構成名譽權侵害態樣的髒話類型或與性、疾患、他人私領域親屬關係相關。最

後，法院在酌定判賠金額時，應尚有考量個案相關因素，未將名譽權受損事由（如：穢語種類）作為單一判賠認定基準。

## 二、未來展望

本次報告僅是簡單地針對使用穢語致他人受名譽權損害民事慰撫金損賠裁判進行初步文本分析，就構成法院心證事由的相關因素所占重要與影響性，未來或可透過統計分析、機器學習等技法，進一步探討法院的判賠標準。此外，在判賠金額上，未來或可利用上述方法，完成相關模型的建置，或可達到預測裁判之成效。