

心理實驗法作業
公衛三 梁嫚芳 b07801003

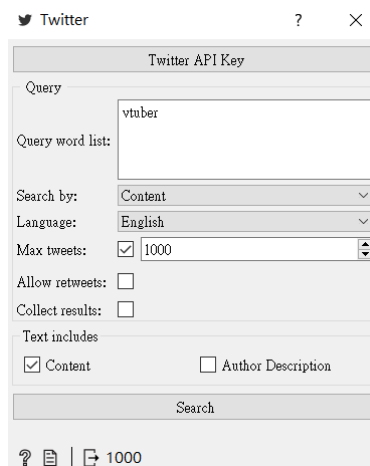
1 任選一與台灣, COVID-19, 或其它自己有興趣的小探索性研究主題 (1 分)

研究主題：英語使用者的”VTuber”文本情感分析

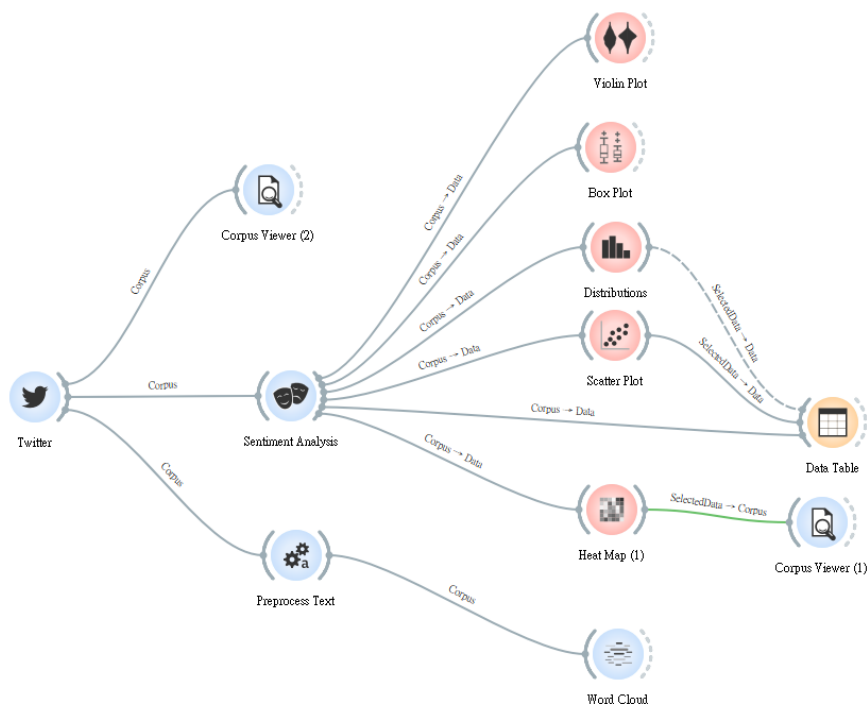
選擇此主題是由於近來 VTuber 浪潮席捲全球，根據我對身邊朋友的觀察，臺灣一部分人們對 VTuber 認為很新穎有趣，甚至有需多忠實粉絲，卻也有另一些人認為 VTuber 很宅、油，而帶有貶義。於是我便好奇，歐美生態的人們究竟對 VTuber 的情緒感知為何呢？以下我為的研究方法與結果。

2. 使用 Orange 擷取並分析如 Wikipedia 或 Twitter 等個人偏好之(社群)媒體資料 (3 分)

使用 Twitter 分析包含 VTuber 的 Tweets，語言限定為英文，並搜尋 1000 則。



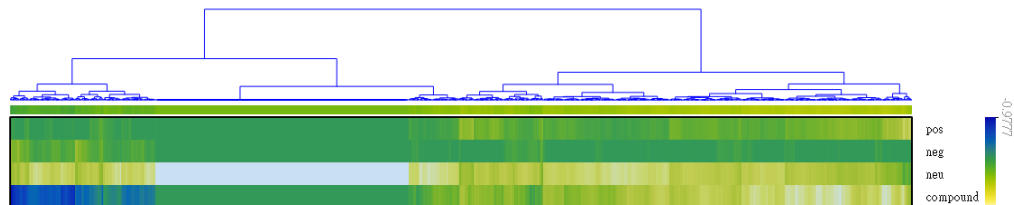
下圖為我的 work flow。



3. 呈現至少一圖或表並討論小研究發現(3 分)

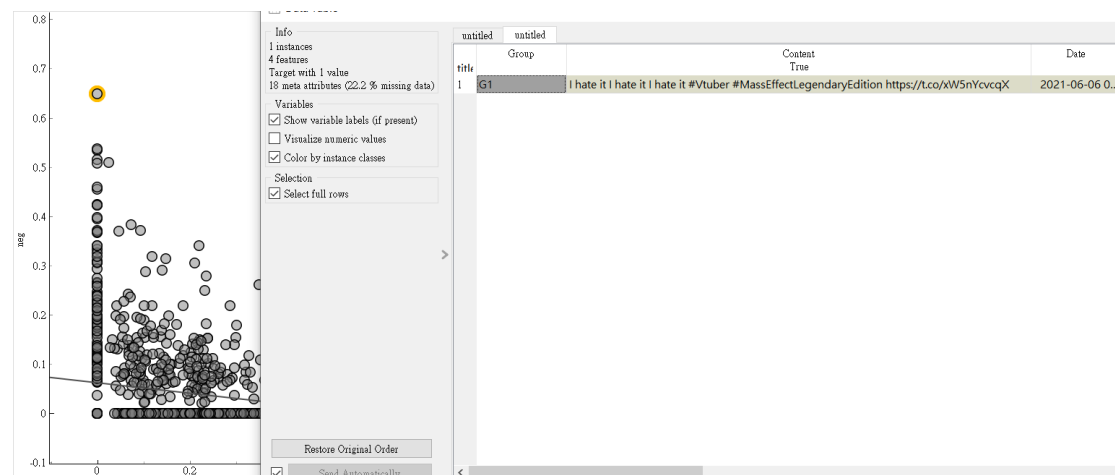
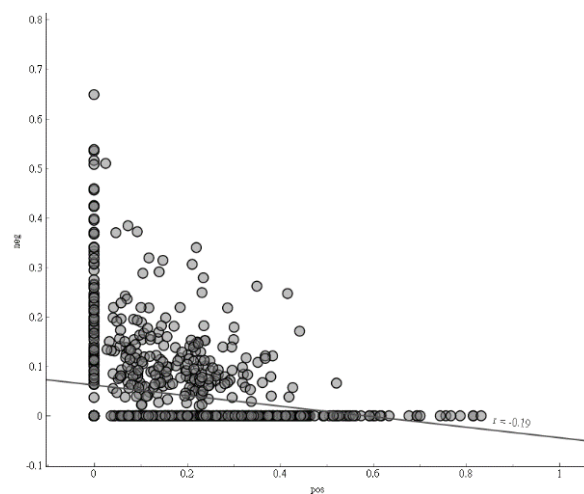
(1) Heat map

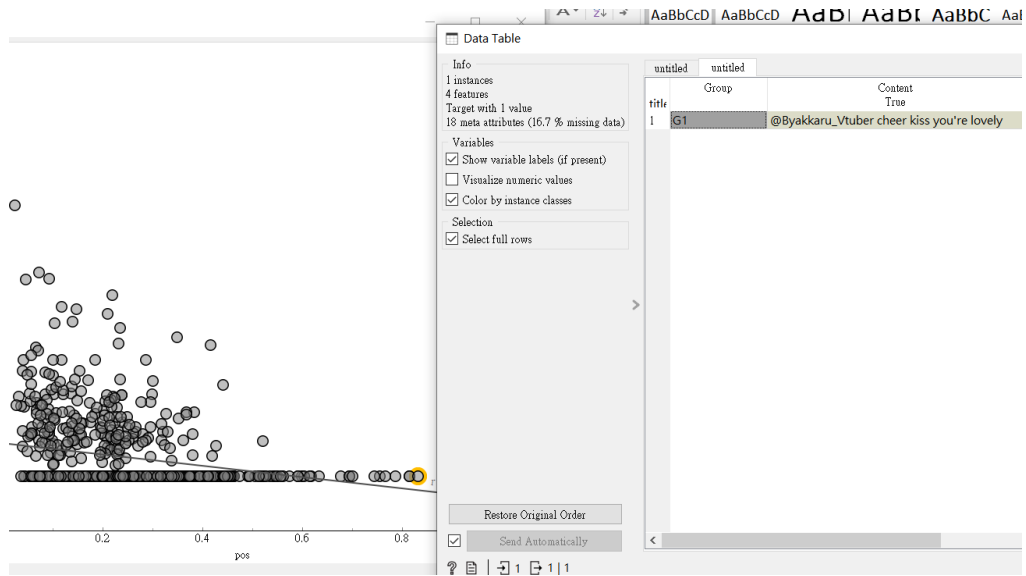
由此圖可見，英語使用者對於 VTuber 的情感大多為中性偏正性，於 compound 分布可見僅有少部分為負性（即藍色）。我認為此為可以預見的結果，畢竟以歐美生態且字串爬取的社群媒體為 Twitter，大多為年輕世代，即與 VTuber 收看觀眾群重疊度高。



(2) Scatterplot: positive vs negative

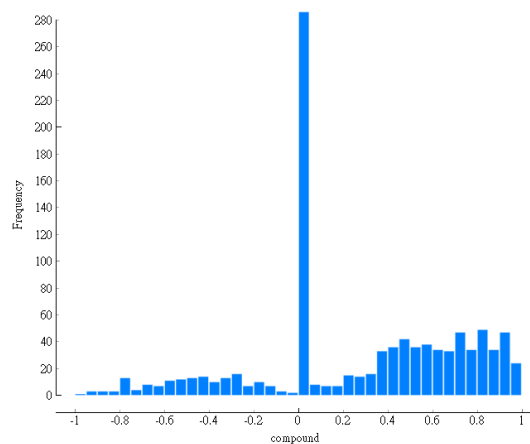
由此圖可見，大致的分布範圍為一以 0, 1, 1 為頂點的直角三角形，推測為大多推文內不會同時包含正性與負性的詞彙。檢視負性分數最高、正性分數最低（最左上角的點）的，果然包含了三次仇恨言論 (hate)。檢視正性分數最高、負性分數最低（最右下角的點）的，包含許多如 kiss, lovely 的言論。





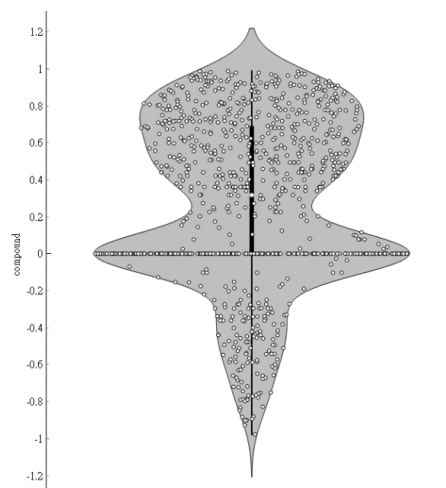
(3) Distribution: compound

由此圖可見，0 的偏正向的第一區間為最多 (280/1000)，此外，於負性與正性分數範圍中分別有兩個小高峰，約為-0.5 與 0.6。仍然可見正性分數的數量較多。



(4) Violin plot

由此圖可見，資料大多分布於 0 與正性區域，與上一張圖大同小異。



(5) Word cloud：從文字雲可見字串頻率的多寡，大致可看多為正性詞彙，如 like, good, want, hope, love 等。

