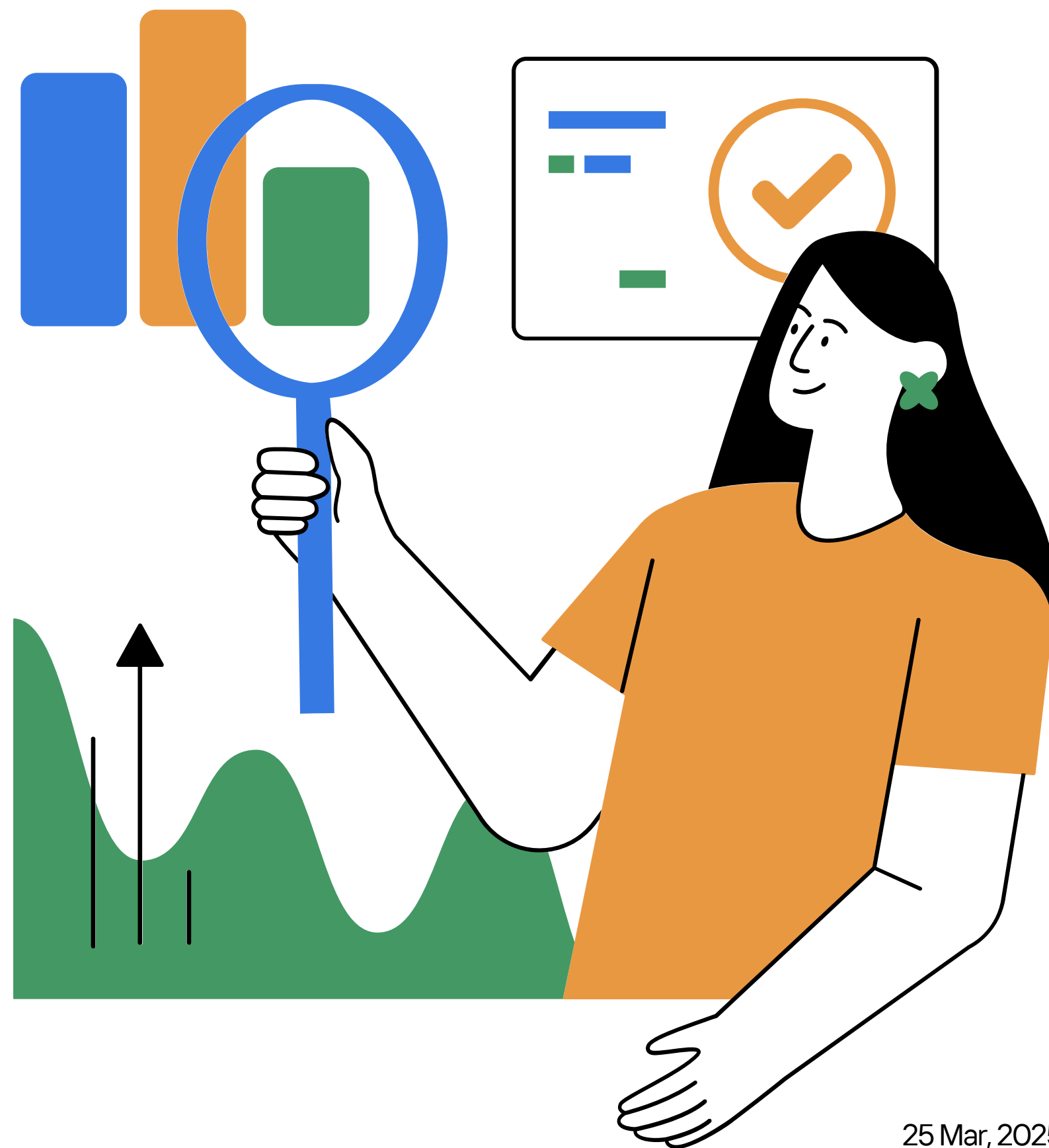


社群媒體分析 第一次讀書會報告



25 Mar, 2025



組員名單

N124320007 藍筱琦 N124320011 林紀吟 N124320012 鄭義璋
N124320017 郭良益 N124320019 吳邦齊 N124320020 陳軍弦
N124320024 何牧 N124320029 黃靖紋



專案動機

隨著中國大陸美食在台灣市場迅速崛起，例如螺螄粉、酸菜魚等正夯的美食，我們注意到，過去因食品安全考量而較為保守的台灣消費者，現在逐漸開始嘗試並接受這些來自大陸的風味。另一方面，社群媒體平台（像是Instagram）的廣泛應用，更使得大陸美食的討論與推廣迅速蔓延，形成一股顯著的美食熱潮。

因此，我們決定運用大數據及情緒分析技術，從Instagram等平台收集相關資料，深入了解台灣民眾對大陸美食的情緒反應。透過這個專案，我們希望能揭示台灣消費者對這些美食的真實態度，並從中發掘出市場趨勢與文化交流的契機。





TARFlow 流程說明

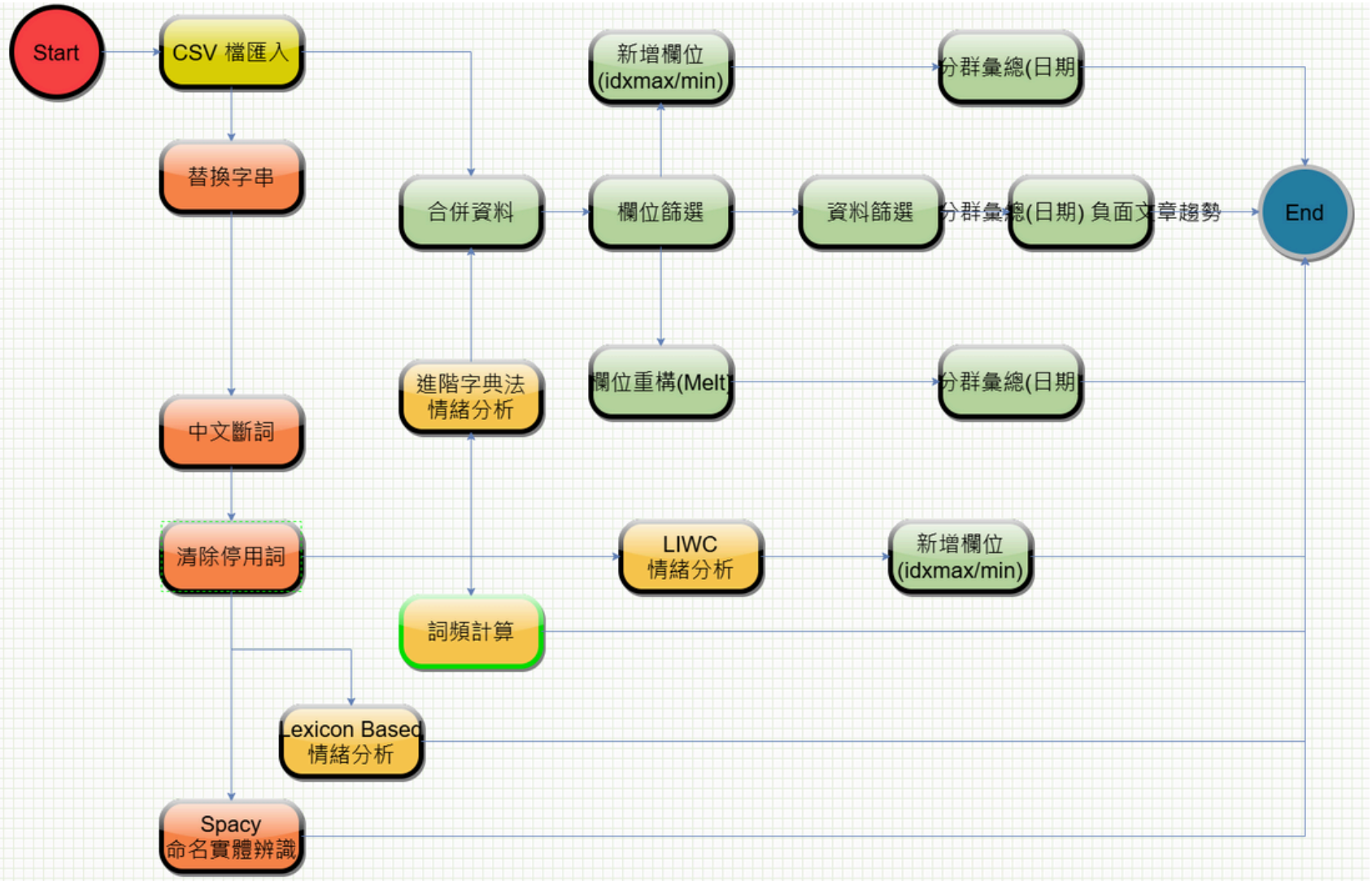
- 1.資料收集：

利用爬蟲抓取2024年全年的Instagram美食文章關鍵字設定：「中華美食」、「爆紅美食」、「特色小吃」等
- 2.資料前處理：

正規表達式清除網址與emoji
中文斷詞與停用字處理（自定義清洗無意義字元）
- 3.情緒分析：

LIWC與Lexicon-Based情緒分析（正向與負向情緒指標）
- 4.進階字典法調整：新增正向詞、移除誤判負面詞（如「錯過」）

以TARFlow平台搭配程式碼實現資料清洗、斷詞、情緒計算與視覺化



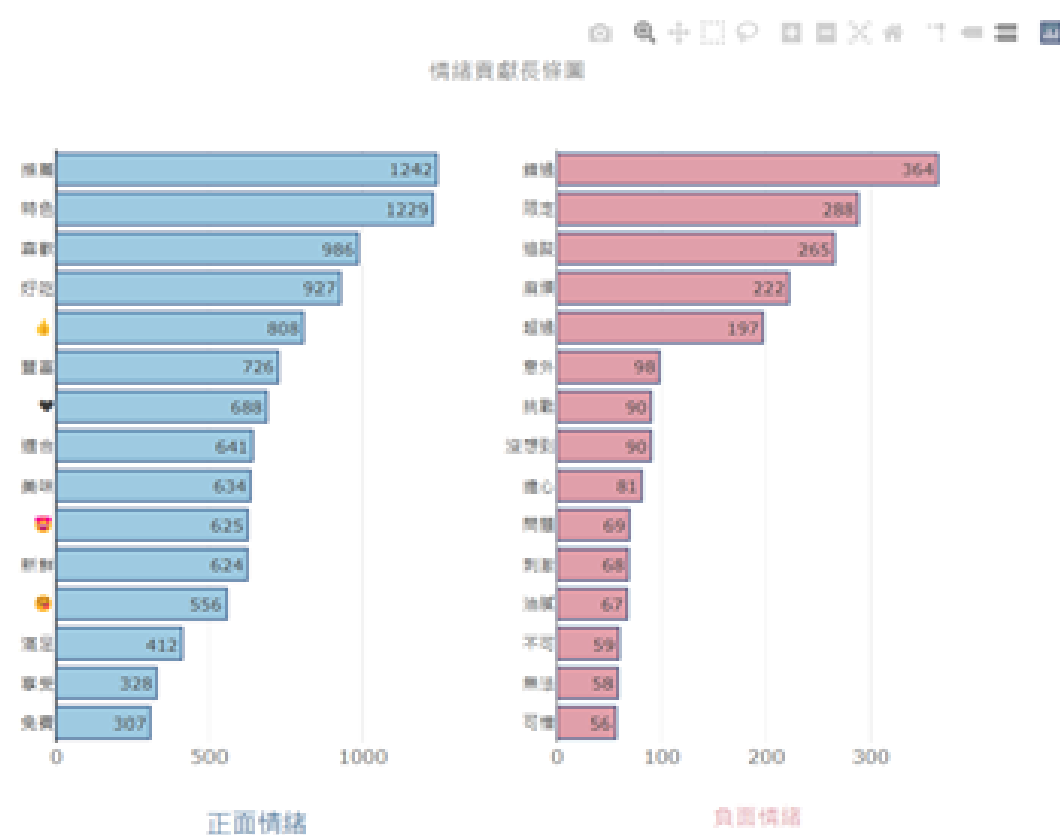


遇到的困難 與解決方式

資料清洗挑戰：

大量無意義的emoji及符號干擾詞頻統計

解決方式：增設停用字庫，精細正規表達式替換



情緒詞斷詞問題：

例如「油膩」與「不油膩」容易誤判情緒

解決方式：調整字典權重，增加反向情緒詞

資料來源限制：

Instagram一次僅能抓取1000筆，需分月份整合

解決方式：分批爬取後再合併處理



參數設定

38-分群彙總(日期) ▼

Group column:

senti_class ▼

X軸變數:

artDate ▼

Y軸變數: 獨立數據變數

system_id ▼

匯總函數:

sum ▼

圖片數量:

200

圖片標題設定:

請輸入圖片標題...

X軸標題設定:

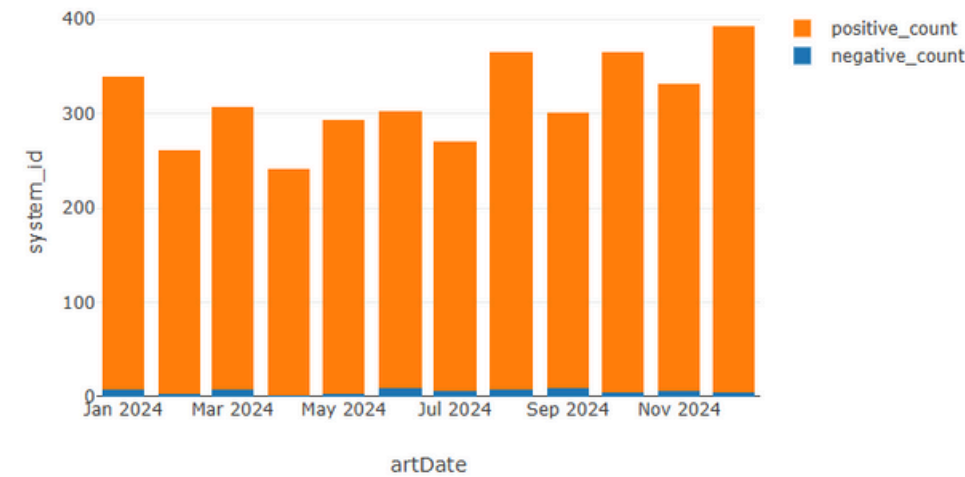
請輸入X座標軸名稱...

Y軸標題設定:

請輸入Y座標軸名稱...

預覽

圖片預覽



在專案成果方面，我們的數據分析結果顯示，台灣民眾對大陸美食的整體情緒呈現明顯的正面趨勢。從收集的Instagram文章中進行情緒計算後，正向情緒的關鍵詞出現頻率遠超過負向情緒，顯示消費者對這些美食普遍抱有高度認同和喜愛。此外，透過進階字典法的調整，我們有效避免了誤判情況，如「不能錯過」這類描述正面情緒的詞彙被準確歸類。

這一結果不僅反映出大陸美食在台灣市場的廣泛接受度，也暗示著台灣消費者對於新奇美食體驗的開放態度。整體正面情緒的數據支持了美食市場進一步多元化的趨勢，並為後續擴大大陸美食在台灣推廣與營銷提供了有力的數據依據。

針對未來的發展，我們計劃進一步篩選出中國美食的各個關鍵字，並觀察各個美食的聲量趨勢。透過更精細的分類與比較，不僅能夠瞭解哪些美食在台灣市場更受關注，也能夠預測市場潛力，進而協助相關業者制定更精準的行銷策略。



分析結果



Thank You

