

## 以網頁資料分析台灣手搖飲品牌的語言定位與語意空間

### 【壹、摘要】

本研究利用 BERT 語言模型，分析台灣 11 個熱門手搖飲品牌在網路媒體上的語言形象。透過語意向量與降維（PCA），並結合 Cosine Similarity、K-means、階層聚類等方法，描繪品牌語意分佈地圖與策略關係（圖 1、圖 2）。結果顯示：多數品牌集中在語意空間左上角，競爭激烈；UG 與再睡五分鐘（2022）為流量明星，50 嵐則展現獨立的 brand 定位與高度認同，為強勢品牌代表。

### 【貳、研究架構】

一、資料來源與處理：透過搜尋引擎 API、爬蟲獲取 ETToday、經濟日報、ELLE 等新聞網站之內容，依據品牌市佔與聲量選出清心福全、50 嵐、COCO、可不可、迷客夏、麻古、UG、先喝道、再睡五分鐘、茶之魔手、春水堂等品牌，經資料清理後保留近 3700 篇文章。

二、分析方法：使用 BERT 提取語意向量，並進行品牌年度平均後降維（PCA），視覺化品牌語意定位（圖 1），並計算品牌間語意相似度（圖 2）。進一步以 K-means 聚類探討語意類型（圖 3、圖 4），並以階層聚類追蹤品牌語言隨年份變化（圖 5、圖 6）。

### 【參、研究發現】

一、語意地圖顯示大多品牌語意定位接近，集中於左上角，反映競爭高度同質化；UG、再睡五分鐘則位於語意空間邊角，呈現（基於流量的）利基定位（圖 1）。

二、品牌間語意距離：如麻古與迷客夏相似度 0.46，UG 與茶之魔手 0.37，麻古與再睡五分鐘為 0.76，顯示語言形象的差異。

三、文章語意聚類：K-means 雖受雜訊影響但仍分為三類：快速展店（UG、茶魔）、活動導向（COCO、麻古）、高互動度（再睡五分鐘、50 嵐），部分品牌如可不可、春水堂則呈語意中立（圖 2）。

四、品牌語意隨年份變化：UG 於 2024 - 2025 明顯差異，語意位置位於空間下方，展現獨特形象；50 嵐位置長期穩定，為經典品牌代表（圖 5）。

### 【肆、商業應用與建議】

一、模型應用：可開發品牌語意儀表板，提供品牌診斷、語言監測與競品分析服務。

二、策略建議：

- 欲模仿 UG 的品牌可強化視覺設計與社群話題，推升聲量。
- 若欲複製 50 嵐的情懷定位，可聚焦品牌故事與情感連結，建立深度認同。
- 新品牌進入競爭激烈的活動導向區需明確差異化，避免落入同質競爭。

### 【伍、結語】

本研究展示文字探勘於品牌語意分析之應用潛力，未來可結合社群語料與瀏覽量權重，建立更即時、動態的品牌追蹤工具，輔助商業決策與行銷策略制定。

### 【圖表與詳細圖說，見後續頁】

圖 1：品牌語意空間（PCA 降維後）、圖 2：品牌語意彼此的相似度關係

圖 3：品牌語意 K-means 分群結果

圖 4：K-means 分群後，各品牌語意與三個中心（centroid）的相似度

圖 5：品牌語意隨年份變化軌跡、圖 6：品牌一年份語意階層聚類圖

## 【參考資料】

以文字探勘分析品牌和代言人之品牌個性 知覺定位

<https://ndltd.ncl.edu.tw/cgi-bin/gs32/gswweb.cgi/login?o=dnclcdr&s=id=%22105NCHU5402011%22.&searchmode=basic>

**Leveraging user-generated social media content with text-mining examples**

<https://www.ibm.com/think/topics/text-mining-use-cases>

**LnData** 《激飲狂潮》

<https://www.lndata.com/explore/trending/article/taiwan-drink>

網路溫度計

<https://dailyview.tw/top100/topic/35>

圖 1：品牌語意空間（PCA 降維後）

- 每一個點代表一個品牌在網路上呈現的語意內容
- 品牌在網路上呈現的整體意象
- 每個品牌都有約 200 篇文章，取向量之平均

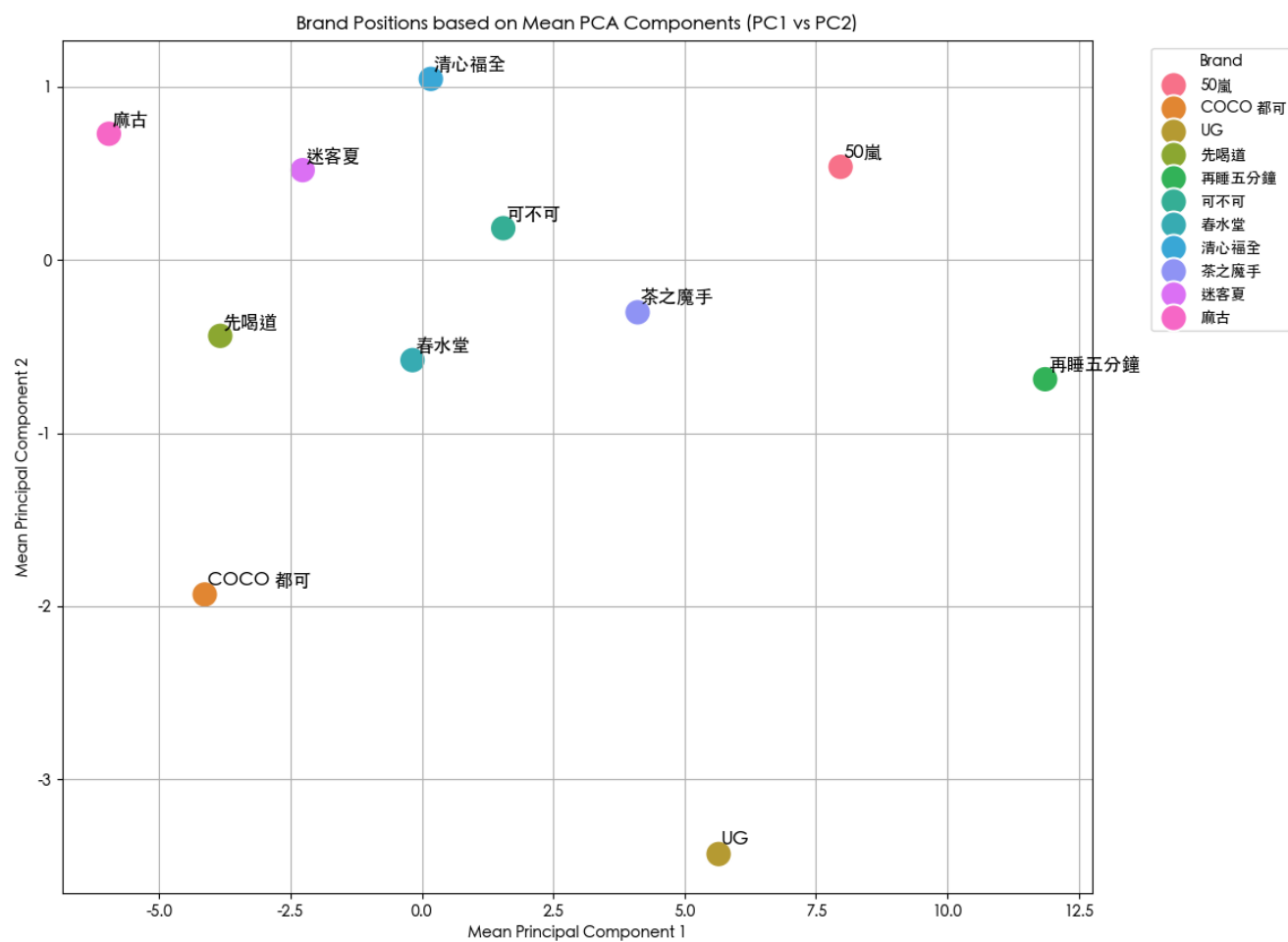


圖 2：品牌語意彼此的距離關係

最相近與最相異者

- 麻古 vs. 迷客夏 0.46
- 麻古 vs. 先喝道 0.39
- UG vs. 茶之魔手 0.37
- 麻古 vs. 再睡五分鐘 -0.76
- 可不可 vs. 春水堂 -0.29

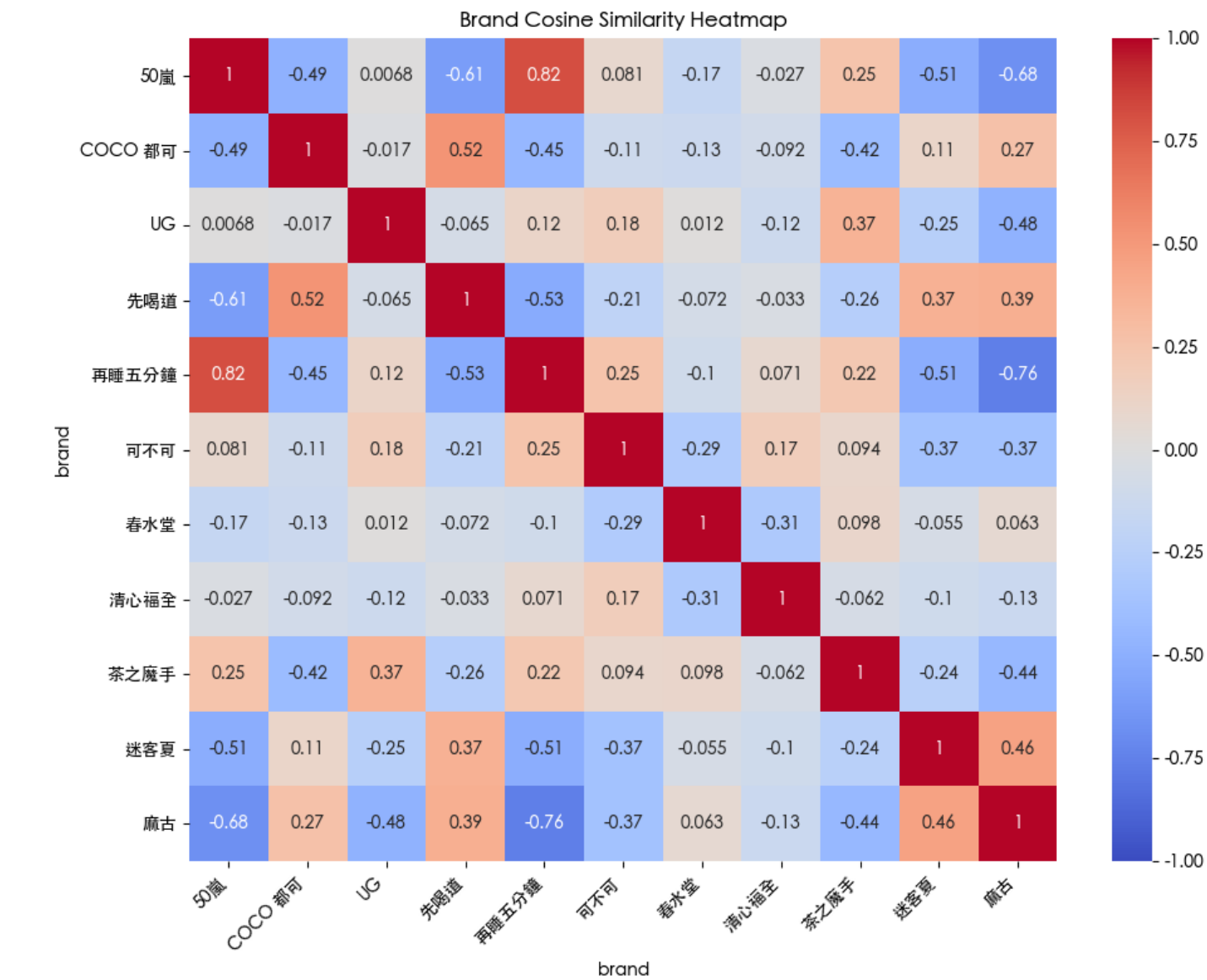


圖 3：品牌語意 K-means 分群結果

- k-means 分群成效不好，因為文章屬於雜訊非常多的資料，但我們還是能看到大致的趨勢。

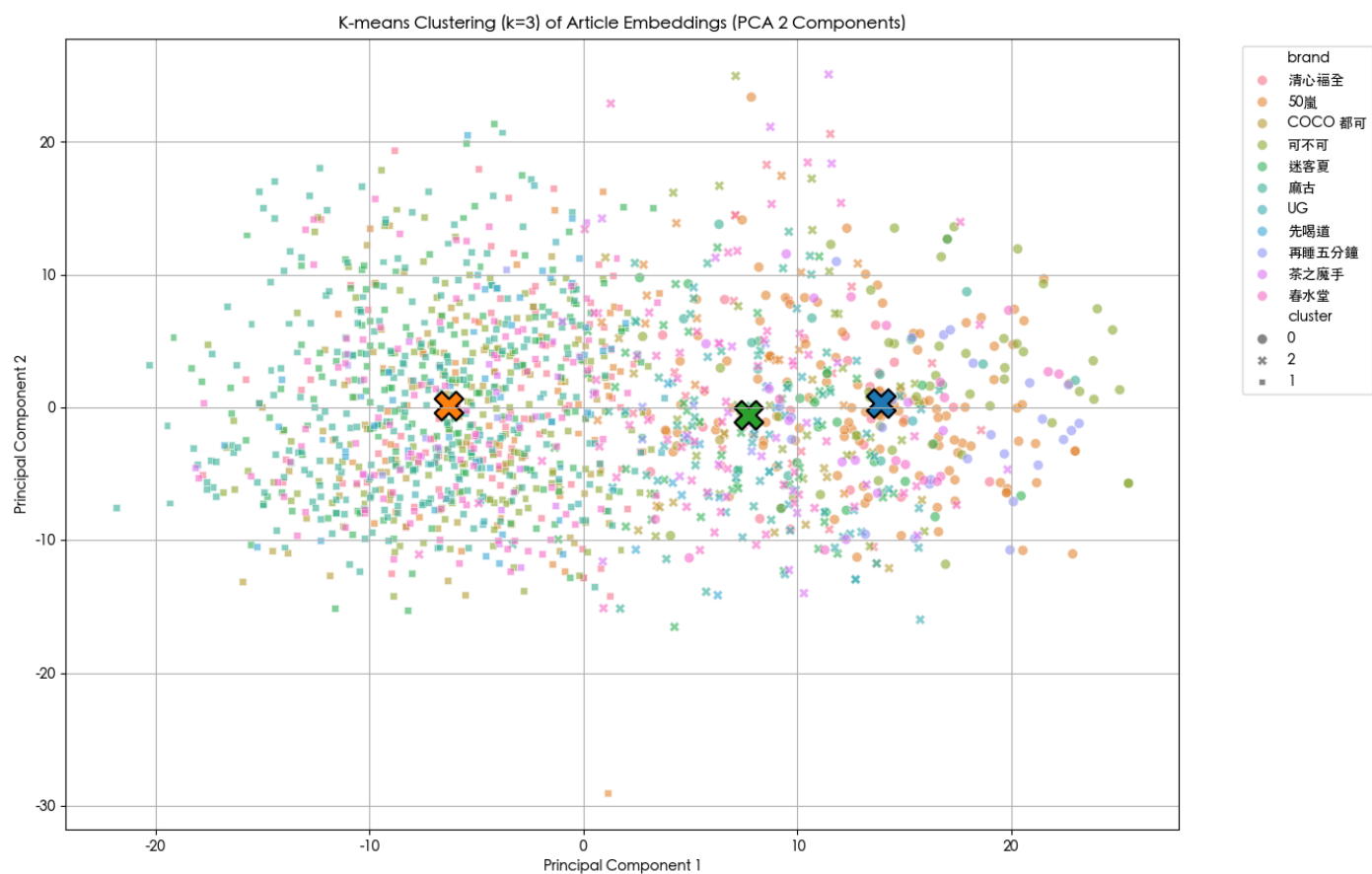


圖 4：K-means 分群後，各品牌語意與三個中心（centroid）的相似度

透過文章的檢索與分析，我們得知：

Cluster 1：快速展店、商業版圖

Cluster 2：新品與活動

Cluster 3：綜合討論、互動性。對於品牌的分群結果：

- 展店：UG、茶魔、（可不可）
- 新品與活動：COCO、先喝道、迷客夏、麻古
- 討論度：再睡五分鐘、50嵐、（可不可）
- 中立：春水堂、清心福全

以上結果，符合大多數人對於品牌的常識

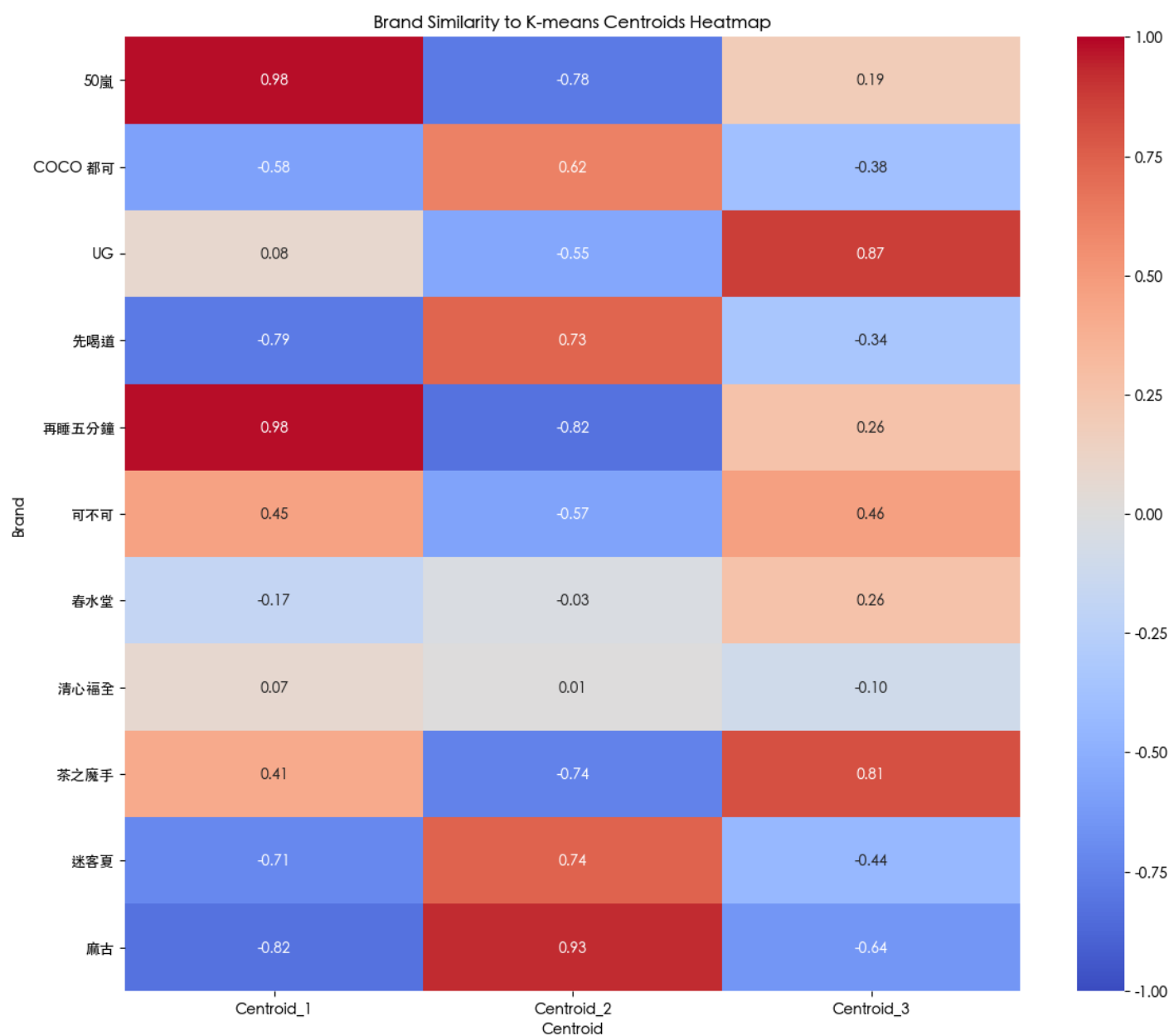


圖 5：品牌語意隨年份變化軌跡

- 大多品牌都聚集在左上角，可能表示競爭格外激烈在左上角：2022 年的再睡 5 分鐘、2024 的 UG、都是當下最高流量的品牌。是品牌地圖中的空缺。
- 50 嵐在空間相對獨立，可能反映消費者特殊的情懷與偏好。也佐證他們的競爭優勢（24%的市佔，來自參考資料《激飲狂潮》）。

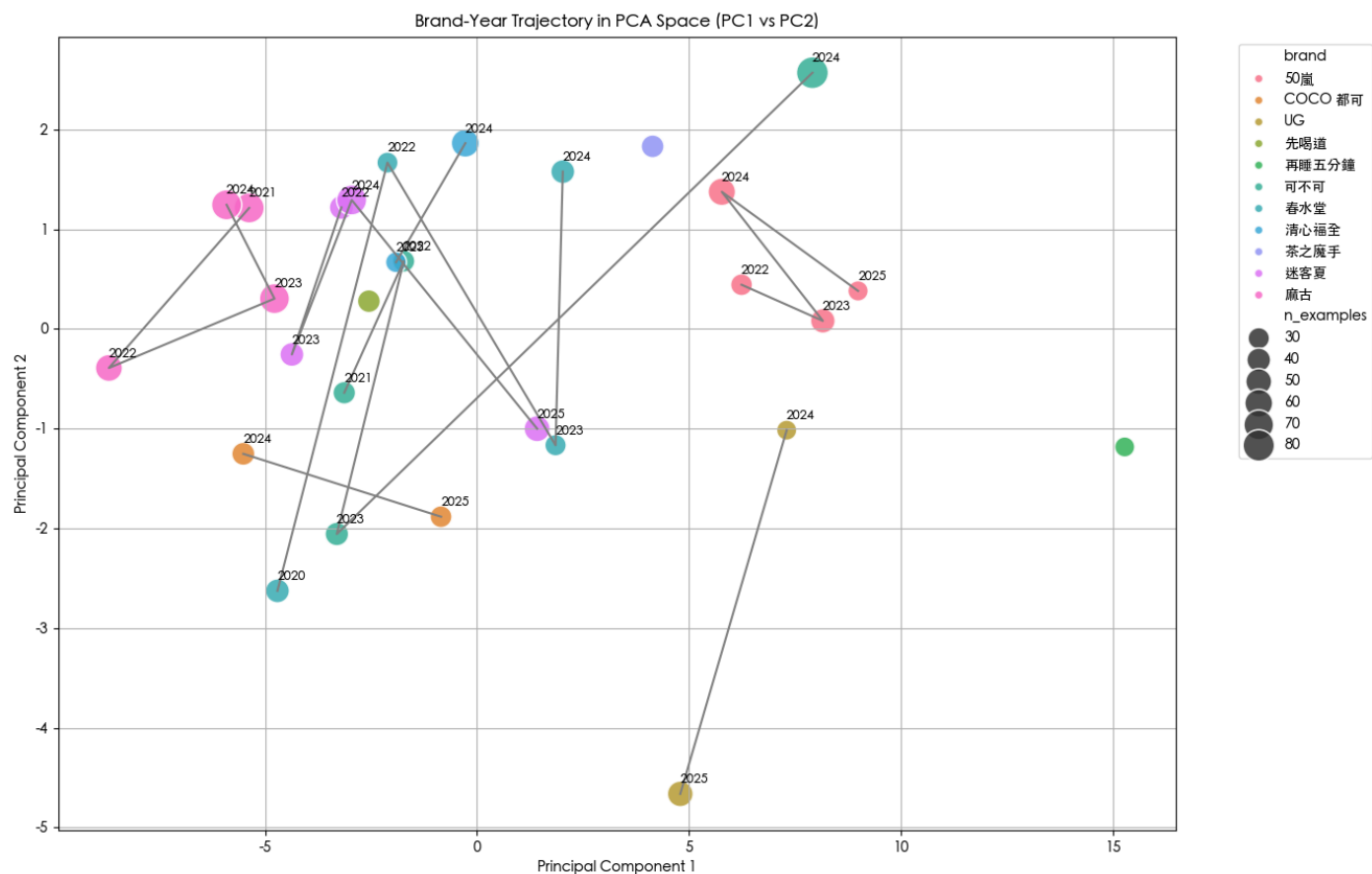


圖 6：品牌一年份語意階層聚類圖

