社群媒體分析_MIS581_第10組(WAKA組)



期末Project

超商聯名商品的輿論分析

指導教授: 黃三益 教授

組員: 周冠呈 N114320001

洪緯哲 N114320013

郭玟慧 N114320024

陳昱維 N114320026

目 錄

壹、分析議題簡介	3
一、研究動機	3
二、研究目的	4
貳、工作流程設計	5
一、資料前處理	5
二、LEXICON-BASE 情緒分析	9
三、LIWC 依日期分群彙總	11
四、WORD2VEC	14
參、詞彙相關性觀察	15
一、相關性矩陣熱圖	15
二、關聯式文字雲 (SHINY)	16
肆、GUIDEDLDA 主題模型	18
伍、社會網路圖 (SHINY)	23
陸、總結	26
一、情緒分析結果	26
二、主題模型分析結果	27
三、社會網絡分析結果	27
柒、影片連結	28

圖目錄

圖	1 :	分析資訊 TarFlow 截圖	3
昌	2	TarFlow 截圖	3
昌	3	PPT 爬蟲截圖	5
昌	4	PPT 爬蟲任務結果截圖	5
圖	5	替換字串截圖	6
昌	6	中文斷詞截圖	6
昌	7	清除停用字截圖	7
昌	8	詞頻計算橫向長條圖	8
昌	9	文字雲	8
昌	10	情緒分析圖	10
昌	11	LIWC 情緒分析	11
昌	12	依日期分群彙總長條圖	12
昌	13	情緒計數圓餅圖	12
昌	14	情緒分析折線圖	13
昌	15	情緒分析盒狀圖	14
昌	16	Word2Vec	14
昌	17	詞彙相關性熱圖	15
圖	18	相關性 DTM 矩陣	16

圖 19	GuidedLDA 主題模型	18
圖 20	GuidedLDA 主題模型第一群	19
圖 21	GuidedLDA 主題模型第二群	20
圖 22	GuidedLDA 主題模型第三群	21
圖 23	GuidedLDA 主題模型第四群	22
圖 24	GuidedLDA 主題模型第五群	23
圖 25	社會網路圖第一群	24
圖 26	社會網路圖第二群	24
圖 27	社會網路圖第三群	25
圖 28	社會網路圖第四群	26

壹、 分析議題簡介

一、研究動機

本次研究的主題著重於對PTT CVS(超商)版的文字探勘與情緒分析,進行超商聯名食物輿 論分析可以幫助我們瞭解消費者對聯名食品、商品的看法、喜好、意見和需求,進而提供有 價值的市場洞察,未來或許可協助超商了解消費者需求,並持續改進其產品和服務。

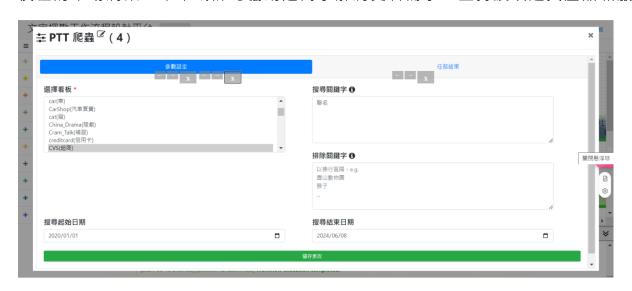


圖 1 分析資訊 TarFlow 截圖

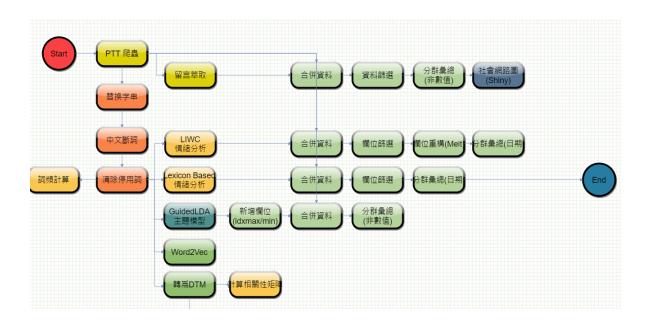


圖 2 TarFlow 截圖

以下為分析資訊說明:

● 分析期間:2020.01.01-2024.06.08

■ 資料來源: CVS(超商)

■ 使用工具: TarFlow 工作流程系統

■ 搜尋關鍵字:聯名

二、研究目的

(一) 瞭解消費者對不同類型聯名產品的整體滿意度和評價:

- 1. 包括食物、飲品、文具、生活用品等。
- 2. 探索消費者對聯名產品的接受度及購買行為。

(二) 分析消費者對不同聯名主題的評價因素:

- 1. 例如老牌口味、創新開發、品牌影響力、設計風格等。
- 2. 研究這些評價因素對消費者滿意度的影響。

(三) 研究消費者對聯名產品的改進建議和期望:

- 1. 瞭解消費者希望超商在哪些方面進行改進,以提升聯名產品的吸引力。
- 2. 探索消費者對未來可能聯名主題的期望和需求。

貳、 工作流程設計

一、資料前處理

- 1. 資料匯入
 - ■指定「聯名」關鍵字與設定排除字,讀入「PTT爬蟲」的資料。

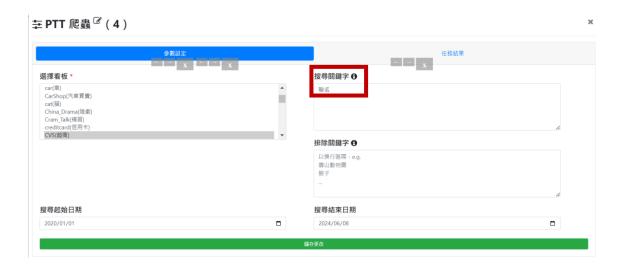


圖 3 PPT 爬蟲截圖

■ 我們針對2020年1月1日至2024年6月8日的PTT中 CSV(超商)版進行聯名關鍵字的爬蟲·總共有1,053筆資料。

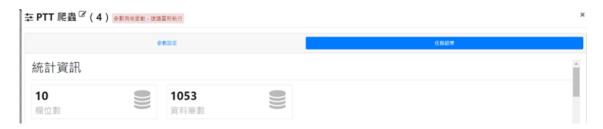


圖 4 PPT 爬蟲任務結果截圖

2. 資料清理

● 使用「替換字串」·將「com」、「from」、「Sent from JPTT on my

\+」、「https」、「imgur」、「jpeg」、「jpg」、「my」、「sent」、

「xd」、「news」、「tw」替換成「空白」、替換成「空白」、替換成

「空白」。



圖 5 替換字串截圖

■ 設定「中文斷詞」定義詞彙。



圖 6 中文斷詞截圖

● 使用「清除停用詞」設定停用的詞彙,移除不具分析效益的詞彙。在清除停用字的部分,因為超商有7-11、FamilyMart等數字及英文字,故無法直接清除英文及數字,透過情緒分析、LDA、詞頻計算等方式交叉比對,清除不要之字眼。

幸清除停用詞 ^② (9) _{参数有做更動,建議重新執行}

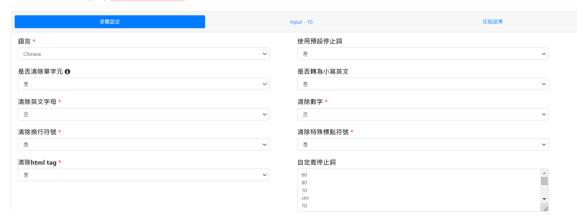


圖 7 清除停用字截圖

● 詞頻計算-長條圖

從橫向長條圖可以觀察到幾個現象:

- A. 全家:這個字詞出現頻率最高,接近 1500 次,可解讀為消費者在評論中經常提到全家便利商店。
- B. 內容和口味:這個字詞出現的頻率也比較高,均超過 500 次。消費者可能 對聯名產品的內容和口味非常關注。
- C. 價錢:這個字詞出現的頻率也較高·接近 1000 次·可見消費者對產品價格的敏感度較高。

消費者經常在評論中提到全家便利商店,這可能是因為其推出了多種聯名產品,吸引了廣泛關注。消費者特別關注這些聯名產品的內容和口味,這表明產品的內在品質和味道對消費者滿意度有重要影響。此外,價格也是消費者關注的焦點,合理的定價策略對於提升消費者滿意度至關重要。

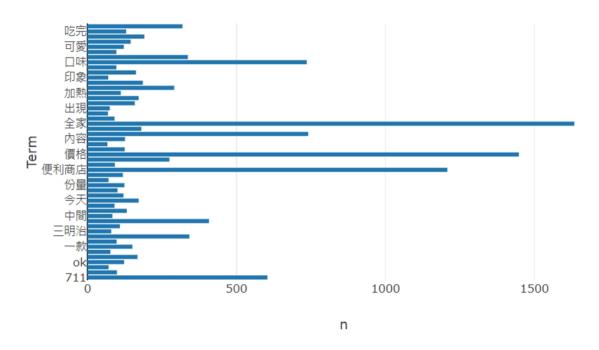


圖 8 詞頻計算橫向長條圖

• 繪製文字雲

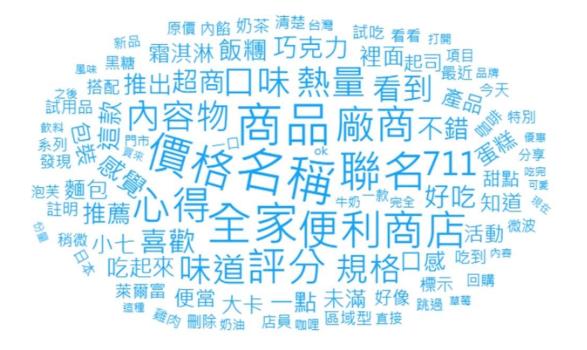


圖 9 文字雲

二、Lexicon-base 情緒分析

- 設定「Lexicon Based」和「LIWC」六種情緒分析字典,以瞭解情緒詞彙、正面與負面詞彙的發生情形。
 - 在我們的研究中,透過採用進階字典法對PTT生活版進行情緒分析,此過程 基於台灣大學情緒字典(NTUSD)進行的定制化調整,挑選了與超商聯名 商品輿論情緒密切相關的正面及負面詞彙。為了更精確地分析聯名商品的 情緒,我們特意去除了一些與超商情緒分析無直接關聯的情緒詞,如「謝 謝」、「考慮」等,並針對聯名商品的特性加入了更多相關詞彙。
 - 正面情緒詞彙:「喜歡」、「美味」、「推薦」、「滿意」以及與聯名商品 特性關聯的詞彙包括「創新」、「高品質」、「方便」和「受歡迎」等。
 - 在負面情緒詞彙:「不滿」、「失望」、「劣質」、「價格高」、「難吃」、「退貨」和「差評」等,這些詞彙通常在商品評價中表達負面情緒,反映 消費者的不滿和抱怨。另外,我們還特別關注了一些關鍵的負面情緒詞彙,如「過期」、「偽劣」、「不新鮮」和「服務差」,這些詞彙對於理解消費者的負面情緒至關重要,也涉及了潛在的商品質量問題和服務質量的問題。

運用上述方法,我們不僅識別了與超商聯名商品相關的積極與負面情緒,還能夠深入理解特定時期(例如:新品上市期間),哪些情緒詞彙特別頻繁地被提及。這種分析不僅可以讓我們對超商聯名商品輿論情緒傾向有更深入的洞察,也可提供超商及品牌商關於市場情緒的見解,作為商品開發

與改進的參考。

情緒	詞彙
正面	喜歡、美味、推薦、滿意、創新、高品質、方便、受歡迎
負面	不滿、失望、劣質、價格高、難吃、退貨、差評、過期、
	偽劣、不新鮮、服務差

- 在我們的研究中,分別使用Lexicon Based與LIWC這兩個情緒字典來分析 發文內容的情緒文字。首先,運用語言詢問詞庫(LIWC)來進行更為廣泛 的情緒分析,LIWC能識別多種不同的情緒維度。
 - ✓ 正面情緒詞彙:排名最高的詞彙是「喜歡」,出現了458次。其他高頻 詞彙包括「好吃」(421次)、「不錯」(408次)和「清楚」(183 次)。
 - ✓ 負面情緒詞彙:排名最高的詞彙是「問題」,出現了54次。其他高頻 詞彙包括「缺點」(52次)、「失望」(49次)和「難吃」(48次)。

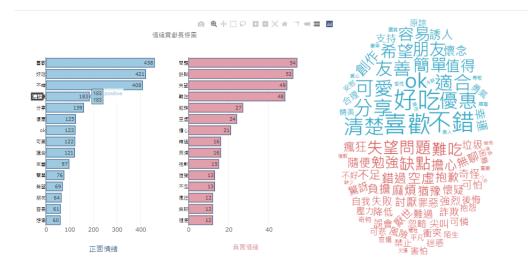


圖 10 情緒分析圖

從這張圖中可以看出,消費者對超商聯名商品的正面情緒佔比最大,達到35.3%,而負面情緒則佔9.89%。情感總數(affect_count)佔比48.5%,表明消費者對聯名商品有相當多的情感反應。此外,負面情緒中的具體分類如悲傷、焦慮和憤怒佔比相對較小,這些分析結果可以幫助企業了解消費者對產品的總體反應,進而進行產品改進和市場策略調整。

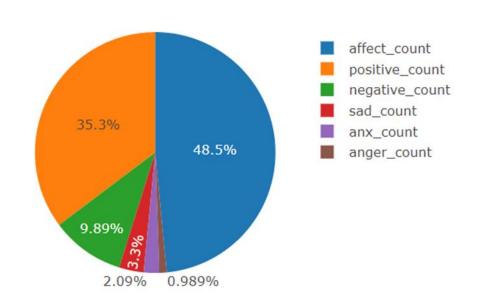


圖 11 LIWC 情緒分析

三、LIWC依日期分群彙總

這張圖表展示了從2020年1月到2021年7月,不同情緒計數(affect_count、anger_count、anx_count、negative_count、positive_count、sad_count)隨時間變化的情況。在某些月份(如2020年4月),情緒表達特別活躍,可能與超商推出特定聯名商品有關,吸引了消費者的廣泛關注和討論。正面情緒和負面情緒的波動顯示了消費者對不同聯名商品的反應,這有助於分析消費者對聯名活動的整體滿意度和反應。

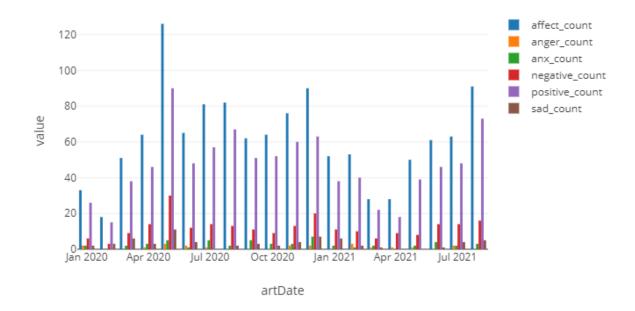


圖 12 依日期分群彙總長條圖

這張圓餅圖展示了不同情緒計數在整體中的比例分佈。情緒總體計數(29.7%)占比最大,表明超商聯名商品引發了大量情緒反應。正面情緒計數占比第二大(25.5%),顯示出消費者對聯名商品普遍持正面態度。然而,負面情緒(16.4%)和其他情緒也不容忽視,反映了聯名商品在推出過程中存在的一些問題或挑戰。

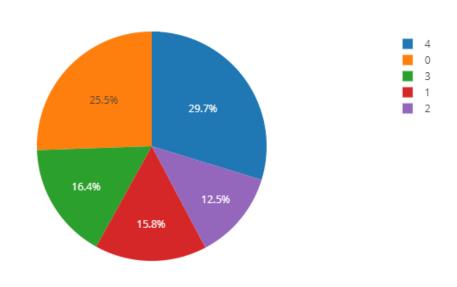


圖 13 情緒計數圓餅圖

這張折線圖展示了系統ID隨時間變化的情況,從2020年1月到2021年7月。某些月份

(如2020年4月和2020年8月)的系統ID峰值,可能反映了超商在這些時期推出的重要聯名商品或活動,這些商品或活動引發了大量的消費者評論和互動,表明這些聯名活動在市場上產生了顯著影響。

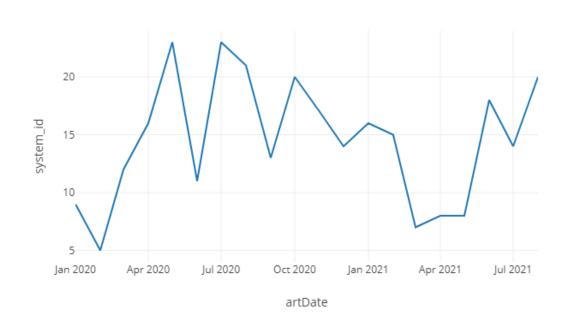


圖 14 情緒分析折線圖

這張盒狀圖展示了不同情緒計數的分佈情況。情緒總體計數的分佈範圍廣,有幾個顯著的離群點,這表明某些聯名商品引發了特別強烈的情緒反應。正面和負面情緒的分佈顯示出聯名商品在市場上的反響不一,某些商品可能受到消費者熱烈歡迎,而另一些則引起了爭議或不滿。了解這些情緒分佈有助於超商在未來策劃聯名活動時,考慮消費者的多樣化反應和需求。

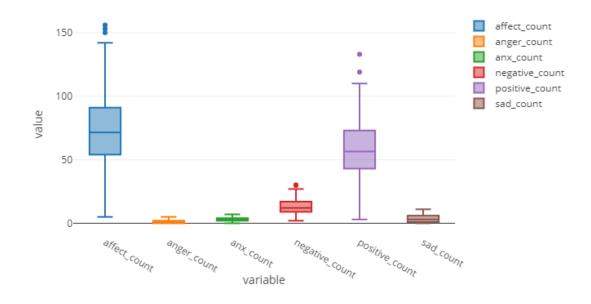


圖 15 情緒分析盒狀圖

四、Word2Vec

1. 下載後的csv資料轉分別轉為兩份TSV檔分別為Embedding及metadata

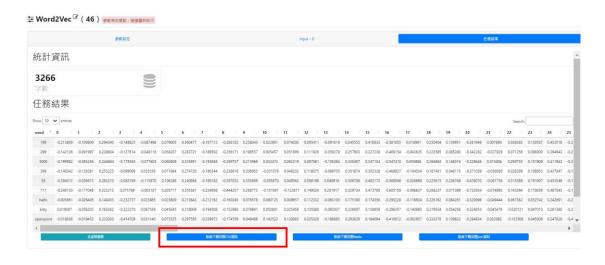


圖 16 Word2Vec

參、 詞彙相關性觀察

一、相關性矩陣熱圖

這張熱圖幫助我們理解不同情緒詞彙在消費者評論或討論中的關聯性。透過觀察顏色深 淺,可以發現哪些詞彙之間有較強的關聯。橫軸和縱軸列出主要情緒詞彙,如「便利商店」、 「品牌」、「商品」和「價格」等。顏色深淺代表關聯強度,從淺藍色(低關聯度)到深藍 色(高關聯度)。對角線上的深色顯示詞彙自身的高關聯度。

從圖中可以觀察到哪些詞彙之間具有較高的關聯度。例如,如果「便利商店」和「品牌」 之間的關聯度較高,這表示在評論或討論中,這兩個詞彙經常被一起提及。

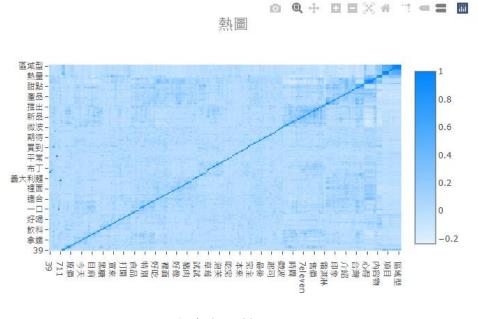


圖 17 詞彙相關性熱圖

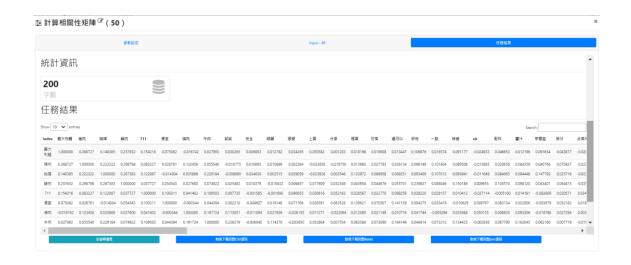
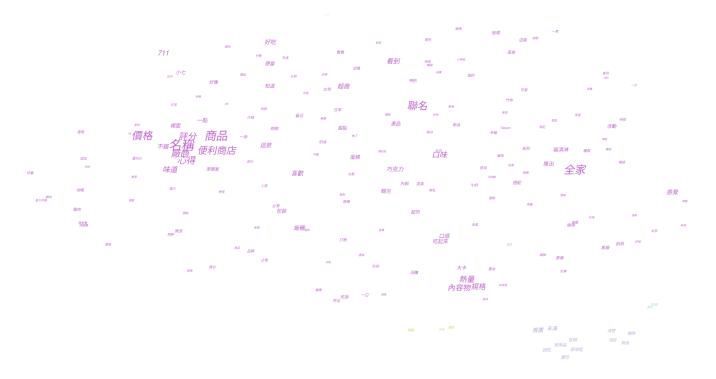


圖 18 相關性 DTM 矩陣

二、關聯式文字雲 (shiny)

這張關聯式文字雲圖將消費者對超商聯名商品的討論分成了五個群組,每個群組代表了不同的情緒或關鍵詞的關聯情況,顯示消費者對超商聯名商品在品牌、商品質量、價格、口味、優惠、新品和包裝設計等多方面的關注和討論。



以下為五個群組說明:

(一)左側群組:

這個群組的關鍵詞包括「商品」、「價格」、「味道」、「便利商店」等。這個群組集中在討論商品的基本特性,包括價格、味道和在便利商店中的可獲得性。消費者關注的是商品的性價比及其感官體驗。

(二)中心群組:

這個群組的關鍵詞如「口味」、「聯名」、「巧克力」、「推」等,這個群組主要 涉及聯名活動和促銷推廣。消費者對於聯名商品的口味和特色,以及推廣活動的效 果有較多討論,顯示出對創新和特別活動的興趣。

(三)右上角群組:

關鍵詞包括「全家」、「霜淇淋」、「感覺」,這個群組反映了消費者對品牌(如全家)的偏好和情感反應。討論中涉及消費者的喜好、滿意度以及品牌給他們帶來的感覺。

(四)中下方群組:

這個群組主要集中在「口感」、「內容物規格」、「口感吃起來」等關鍵詞上,這個群組集中在產品的具體規格和詳細評價,特別是口感和內容物的詳細特性。消費者對產品的具體品質有深入的探討,反映出他們對產品體驗的重視。

(五)右下方群組:

關鍵詞如「試用品」、「區域型」、「試吃」、「推薦」,這個群組的關鍵詞如「試用品」、「區域型」、「試吃」等,顯示了消費者對便利商店新推出的試用品和區域型商品的興趣和關注。特別是試吃活動,這些詞語表明消費者對能夠親自體驗產品的重要性,以及區域性產品在不同地區的影響和評價。

建、 GuidedLDA 主題模型

我們根據字詞內容先抓出並設定便當、飯糰、甜點、霜淇淋4種主題種子字,並把主題數 設為5,以此觀察主題分類的效果。



 幸 GuidedLDA 主題模型 ^② (42)

 ●數有做更動・建議重新執行

圖 19 GuidedLDA 主題模型

下圖為GuidedLDA模型的主題視覺化結果。左側的主題間距地圖(Intertopic Distance Map)顯示了五個主題的相互關係,圓圈的大小表示每個主題在整體文本中的比例,而圓圈之間的距離則代表主題之間的相似度。可以看到,主題2和3之間距離較近,表明它們的內容較為相似。

右側的圖表則展示了主題1的最相關詞彙,列出了與主題1關聯性最高的30個詞彙。這些 詞彙中,像便當、飯、甜點等詞彙在主題1中具有高度相關性,顯示出主題1可能集中在食品 相關的內容上。紅色和藍色條形圖分別代表詞彙的頻率和特異性,數值越高表示詞彙對主題 的貢獻越大。這表明,消費者對於超商的食品產品特別關注,而這些產品的內在品質和味道 是消費者滿意度的重要因素。

■ 根據下圖的主題1的關鍵詞,我們可以看到一些頻繁出現的詞彙,如「便當」、「飯糰」、「小七」、「萊爾富」、「咖哩」等。這些詞彙都與便當主食相關。因此,我們可以將

主題1定義為"便當主食群"。這個群組主要集中在描述和討論7-11和萊爾富的各種主食的產品口味,顯示消費者對這些產品的內容和口味特別關注。

GuidedLDA Vis

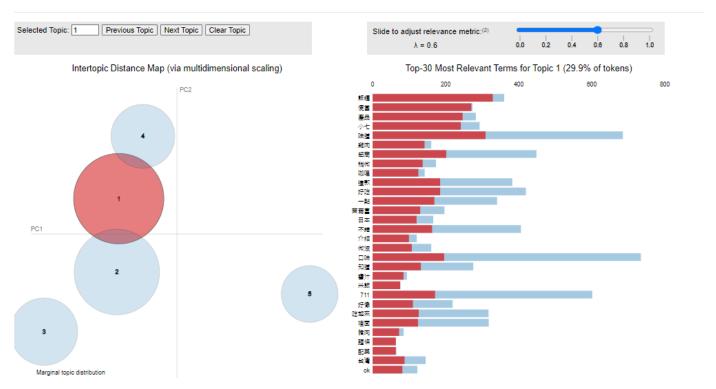


圖 20 GuidedLDA 主題模型第一群

根據下圖的主題2的關鍵詞,我們可以看到關鍵詞如「奶茶」、「黑糖」、「711」、「好喝」等頻繁出現。因此,我們可以將主題2定義為「飲料群」。這個群組主要集中在描述和討論超商的各種飲品的甜度和口味策略。



圖 21 GuidedLDA 主題模型第二群

根據下圖的主題3及其最相關詞彙,根據關鍵詞分析,主題3中的高頻詞包括「試吃」、「區域型」、「試用品」、「圖文」、「嘗鮮」等,這些詞彙顯示主題3主要集中於討論 超商新品的嘗鮮和區域型的試用活動居多。因此,我們可以將主題3定義為「嘗鮮與消費 體驗群」。這表明消費者對超商的服務質量、顧客服務和整體購物體驗有高度關注。

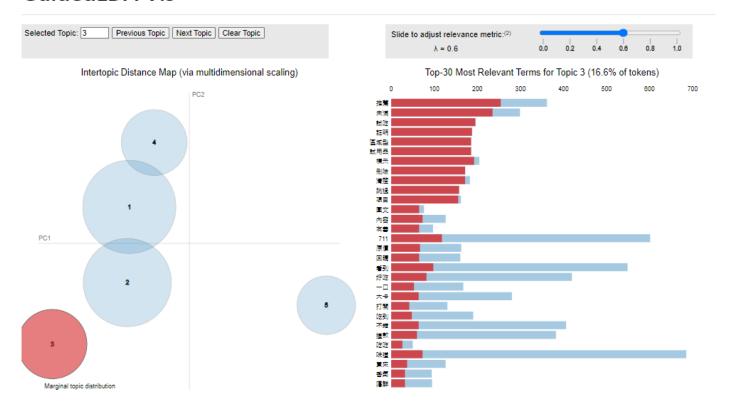


圖 22 GuidedLDA 主題模型第三群

根據下圖的主題4的高頻詞包括「巧克力」、「麵包」、「蛋糕」、「甜點」、「泡芙」、「牛奶」等,顯示主題5主要集中在甜點、麵包和飲料上。因此,我們可以將主題5定義為「麵包甜點群」。這表明消費者對於便利商店中各種麵包、甜點,如「巧克力」、「麵包」、「餅乾」等,有高度的關注和興趣。

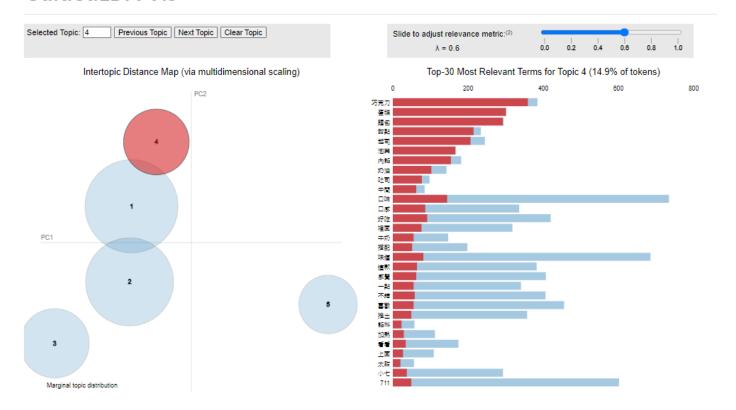


圖 23 GuidedLDA 主題模型第四群

■ 根據下圖的主題5的高頻詞包括「霜淇淋」、「活動」、「優惠」、「app」、「門市」、「推出」等,顯示主題5主要集中在超商推出的活動和優惠上。因此,我們可以將主題5 定義為「活動優惠群」。這表明消費者對於便利商店中各種活動、優惠的關注和興趣。

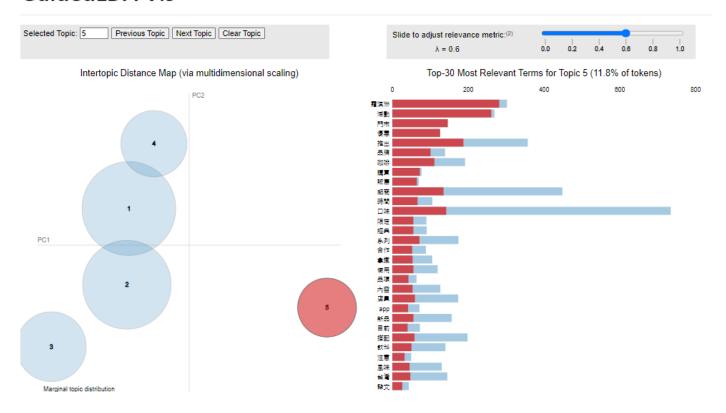


圖 24 GuidedLDA 主題模型第五群

透過GuidedLDA模型的主題分析,我們能夠了解消費者對超商聯名產品和服務的具體關注點。從圖表結果顯示,消費者對超商產品的品質、促銷活動、服務體驗、會員制度以及日常食品的供應都有著不同層次的需求和期望。這些結果有助於超商制定更精確的市場策略和提升服務質量,從而增強消費者滿意度和忠誠度。

伍、 社會網路圖 (Shiny)

透過社會網路圖可以看出節點之間的互動和連接方式更加明確,整個網絡可以分成四個主要群體,每個群體都有明確的核心節點和其相關聯的用戶。

■ 第一個群體以missheti為核心節點,該節點與多個用戶有直接連接,顯示missheti在這個 群體中具有很高的影響力。這些用戶之間也存在一定的互動連接,形成了一個緊密聯繫 的社交網絡。

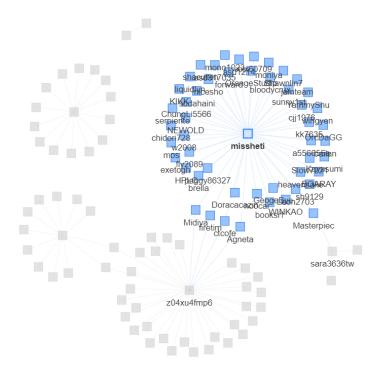


圖 25 社會網路圖第一群

第二個群體以z04xu4fmp6為核心節點,並與多個用戶直接連接。這些用戶之間也有一些 互動,顯示出z04xu4fmp6在這個群體中扮演著重要的橋樑角色,將多個用戶聯繫在一起。

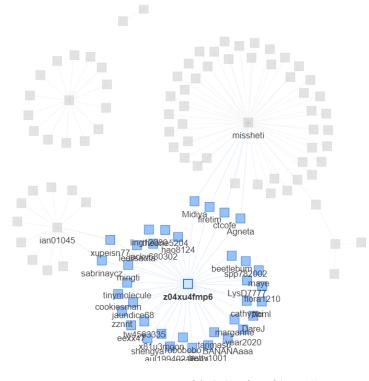


圖 26 社會網路圖第二群

第三個群體的核心節點是ian01045,這個節點與其他用戶有直接連接,這些用戶之間也存在一定的互動連接,形成了一個相對緊密的社交網絡,顯示出ian01045在這個群體中的重要地位。

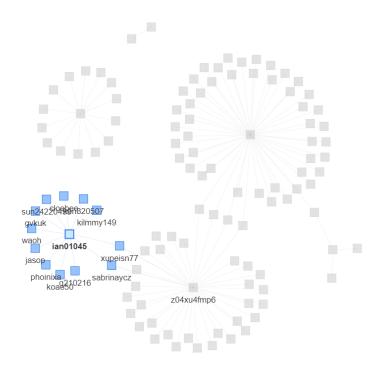


圖 27 社會網路圖第三群

■ 第四個群體以xuxu為核心節點,並與多個用戶直接連接。群體內的用戶之間也有一些互動,顯示出這是一個緊密聯繫的群體,核心節點xuxu在其中起到關鍵作用。

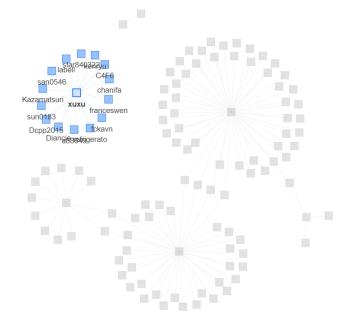


圖 28 社會網路圖第四群

這些社會網絡圖顯示了四個主要的互動群體,每個群體都有一個核心節點,這些核心節點與其他用戶有密切的互動,形成了清晰的社交結構。這些圖表為我們提供了用戶之間互動模式的可視化展示,有助於理解用戶在社交網絡中的角色和影響力,並可以用於進一步分析和制定有效的社交網絡策略。

陸、總結

本研究結果顯示,消費者對超商聯名商品的關注點主要集中在食品飲品的品質和口味、促銷活動、服務體驗、會員制度和日常食品等方面。超商可以根據這些結果,針對不同的消費者需求制定精確的市場策略和促銷活動,進一步提升消費者滿意度和品牌忠誠度。透過持續的情緒和主題分析,超商可以更好地了解消費者的需求和期望,從而在市場競爭中取得更大的優勢。

一、情緒分析結果

- (一) 正面情緒佔優勢:根據情緒分析結果,消費者對超商聯名商品的正面情緒佔比較高, 詞彙如「喜歡」、「美味」、「推薦」、「滿意」等頻繁出現。這表明聯名商品普 遍受到消費者的好評,且這些正面情緒詞彙的出現頻率顯示出消費者對商品的高度 滿意度。
- (二) 負面情緒關注問題:負面情緒詞彙如「失望」、「問題」、「難吃」、「拉肚子」 等顯示出消費者在某些聯名商品上存在不滿。這些詞彙反映了產品在品質和口味上 的問題,這些負面反饋對未來產品改進有重要參考價值。

二、主題模型分析結果

- (一) 便當主食群:消費者對於便當、飯糰、咖哩、雞肉等主食關注度最高,這些產品的 內容和口味對消費者滿意度影響顯著。
- (二) 飲料群:消費者對於超商飲品的口味、甜度也都有相當高的討論度。
- (三) 嘗鮮與消費體驗群:消費者對超商新品的嘗鮮和區域型的試用活動也有高度關注。
- (四)麵包甜點群:巧克力、麵包、蛋糕、泡芙、牛奶等日常食品受到消費者的高度關注,這類產品的質量和新鮮度是消費者滿意度的重要因素。
- (五) 活動優惠群:優惠、活動、app、活動等是消費者關注的另一重點,合理的促銷活動策略能有效提升消費者的購買意願。

三、社會網絡分析結果

社會網絡圖顯示了四個主要的互動群體,每個群體都有明確的核心節點,如: missheti、z04xu4fmp6、ian01045和xuxu,這些核心節點在輿論傳播中扮演重要角色。透過這些核心節點,超商可以更有效地傳遞訊息和促銷活動。

透過上述總結分析,超商聯名商品在市場上引起了廣泛的消費者關注和討論,情緒反應 具有多樣性。正面情緒詞彙如「推薦」、「優惠」、「方便」等反映了消費者對聯名商品的 普遍好評及其便捷性和實惠性。負面情緒詞彙如「失望」、「問題」等則揭示了部分商品在 品質和口味上的不足,這些問題需要引起企業的重視。

透過分析消費者對不同類型聯名產品的滿意度和評價,超商可以更加精確地定位市場需求,針對不同的消費者群體制定相應的市場策略。進一步了解消費者對服務體驗、促銷活動、

會員制度和日常食品的關注點,有助於超商提升服務質量和產品供應,增強消費者滿意度和 品牌忠誠度。

未來,超商應持續進行情緒和主題分析,以捕捉市場趨勢和消費者需求的變化,從而在 競爭激烈的市場中保持優勢。這些深入洞察為企業提供了寶貴的市場資訊,有助於超商在聯 名商品的開發和推廣中取得更大的成功。

柒、 影片連結

https://www.youtube.com/watch?v=INouBhBt5Ys