

# » 巨量資料分析與應用 (3)

楊立偉教授

台灣科技大學資管系

2013 Fall

# 楊立偉教授

## ◆ 專長領域為資料庫及語意分析技術、知識管理、數位行銷

現任	台科大資管系兼任助理教授 2008~
	台大工管系暨商研所兼任助理教授 2006~
	資訊及通信國家標準技術委員
	意藍資訊 董事總經理（創辦人） 1999~ 國內規模最大的網路情報與社群口碑自動分析平台
	龍捲風科技 董事總經理 國內企業搜尋引擎市佔率最高；國際檢索競賽第一名
經歷	智威湯遜數位行銷首席顧問、尚藍互動行銷共同創辦人
	2009年獲選100 MVP最有價值經理人，擁有超過20項語意分析專利
	2012年榮獲國家雲端創新獎、數位時代「創業之星」首獎

# 課程大綱

## ◆ 第一部份

- 巨量資料導論
- 巨量資料分析與管理架構
- 巨量資料分析技術

## ◆ 第二部份

- 應用案例與研討 – 企業個案 (1)
- 應用案例與研討 – 企業個案 (2)
- 應用案例與研討 – Open Data



# 應用案例與研討

---

## 運用社群資料分析對企業之影響研究

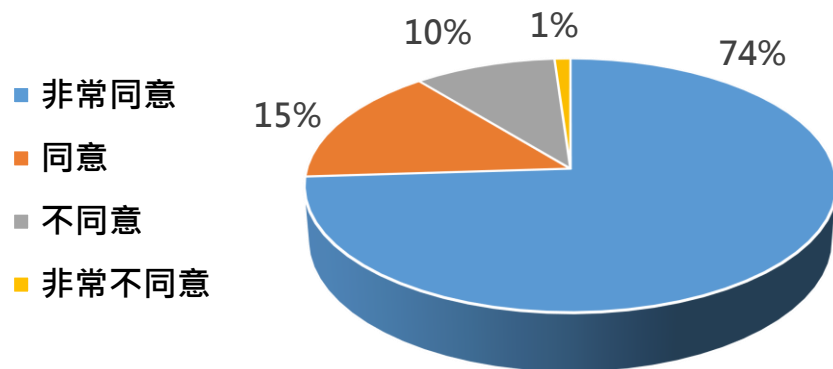
# 網路社群輿情的影響力快速增加

## ◆ 社群媒體影響力超越電視、雜誌

- 89.1% 台灣民眾受到網路社群的影響

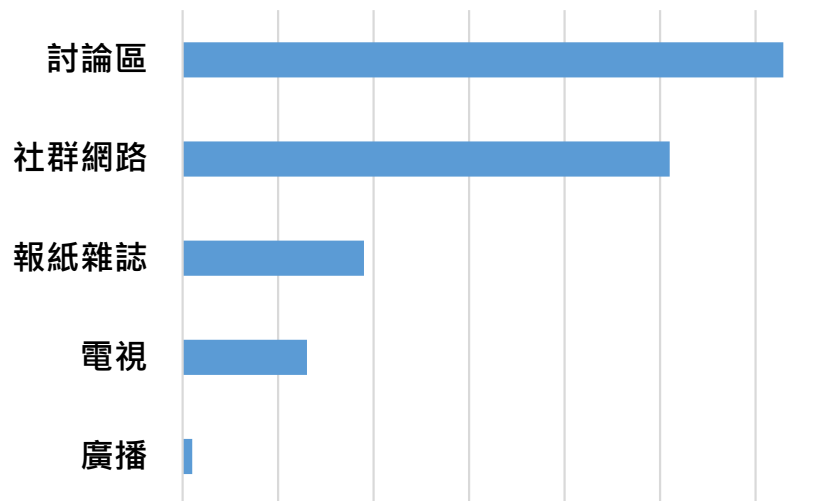
## ◆ 各產業皆受影響；是機會 也是威脅

是否受社群媒體影響



Source: Insightexplorer

受到媒體影響的民眾數



Source: MIC

# 2012~2013 社群輿情 發展趨勢觀察

## 一. 全面影響各產業

- 銀行、汽車、政府等也張手迎接

## 二. 社群輿情擴散的速度更快，層面更廣

- 平均在 48 小時內達到高峰，要與時間賽跑

## 三. 帶有情緒的輿論較去年增加 28.3%

- 網友更勇於在網路上表達自己的喜惡

## 四. 負面輿論成長的速度比正面輿論快 1.3 倍

- 網民意識抬頭，「壞事傳千里」更需戰戰兢兢

# 網路社群輿情對相關人員的影響

- ◆ 好比電視產業沒有了收視率調查與觀眾分析
- ◆ 缺少有力的資料支持
- ◆ 在迷霧中作戰



# Social Listening

- ◆ 又稱 Social analytics 或 Social intelligence
- ◆ 為社群輿情管理循環的起點
- ◆ 新一代市場調研的利器





# 用大數據為社群輿情大解密

---

使用《OpView社群口碑資料庫》內容觀察和自動語意技術  
二年有效討論量1億5618萬則，逐篇即時處理  
涵蓋台灣流量9成以上的社群、討論區、部落格、BBS等  
超過五千個網站頻道

# 掌握 品牌影響力 與 消費行為 的利器

## ◆ 國際趨勢 Social Listening & Intelligence

	革命性大數據新方法	民意或市場調查
主動性	主動	被動
準確性	非干擾式的內容觀察法 大數據下母體即抽樣	存在抽樣及干擾誤差
時效性	即時性高	低
資料量	龐大	1068份樣本
觀點形成	意見可深入分析	量化成選項，彈性低
執行難易	平台化工具，易執行	作業繁複

周變小時 快100倍

千到千萬份 大1萬倍

# 用Big Data站上制高點 – 社群口碑資料庫

## OpView Service

您的輿論觀測站

連上OpView · 品牌形象、輿論觀點監測好輕鬆



- ◆ 國內最大的雲端網路輿情觀測中心
  - 上百台雲端主機，涵蓋台灣九成以上社群流量
  - 超過二年度、2 億多筆的社群數據
- ◆ 多項Social media分析技術專利與獲獎
- ◆ 監測分析涵蓋 500大品牌、50個產業
  - 學術單位與知名媒體研究引用，權威性最高





### 網路影響力品牌7月月報：BURBERRY、CITIZEN 部落格發文帶動討論 經理人月刊 - 2013/8/18

再針對這些品牌，利用OpView雲端情報平台網路口碑資料庫，計算出2012年1月至7月間「品牌影響力指標」得分前200名者，為2012年影響力品牌TOP 200。從8月份...



### 慶祝父親節全台網路口碑大調查 中央日報 - 2013/7/29

國內網路口碑分析權威意藍資訊《OpView社群口碑資料庫》，公布了最「榜」，可供天下兒女表達孝心之參考。此次觀測...



### 七夕約會首選網友愛美食餐廳 中央通訊社 - 2013/8/7

中央社記者林紳旭台北7日電) OpView社群口碑資料庫今天公布七夕情物，美食餐廳和鮮花位居第1。約會地點排行... + 顯示更多

中央日報



### 情人節變不出新花樣餐廳放閃蟬聯TOP1 ETtoday - 2013/8/10

國內網路口碑分析權威意藍資訊《OpView社群口碑資料庫》統計及分析各大果發現今年「約會地點排行榜」的第一名與第二...



# OpView – 技術架構

Listening

Analyzing

Engaging



熱門議題分析

關鍵字詞擷取

競爭風暴圖

意見觀點分析

正負情緒分析

時間趨勢與分布

自動摘要分析



搜尋引擎



語意分析

索引

分析

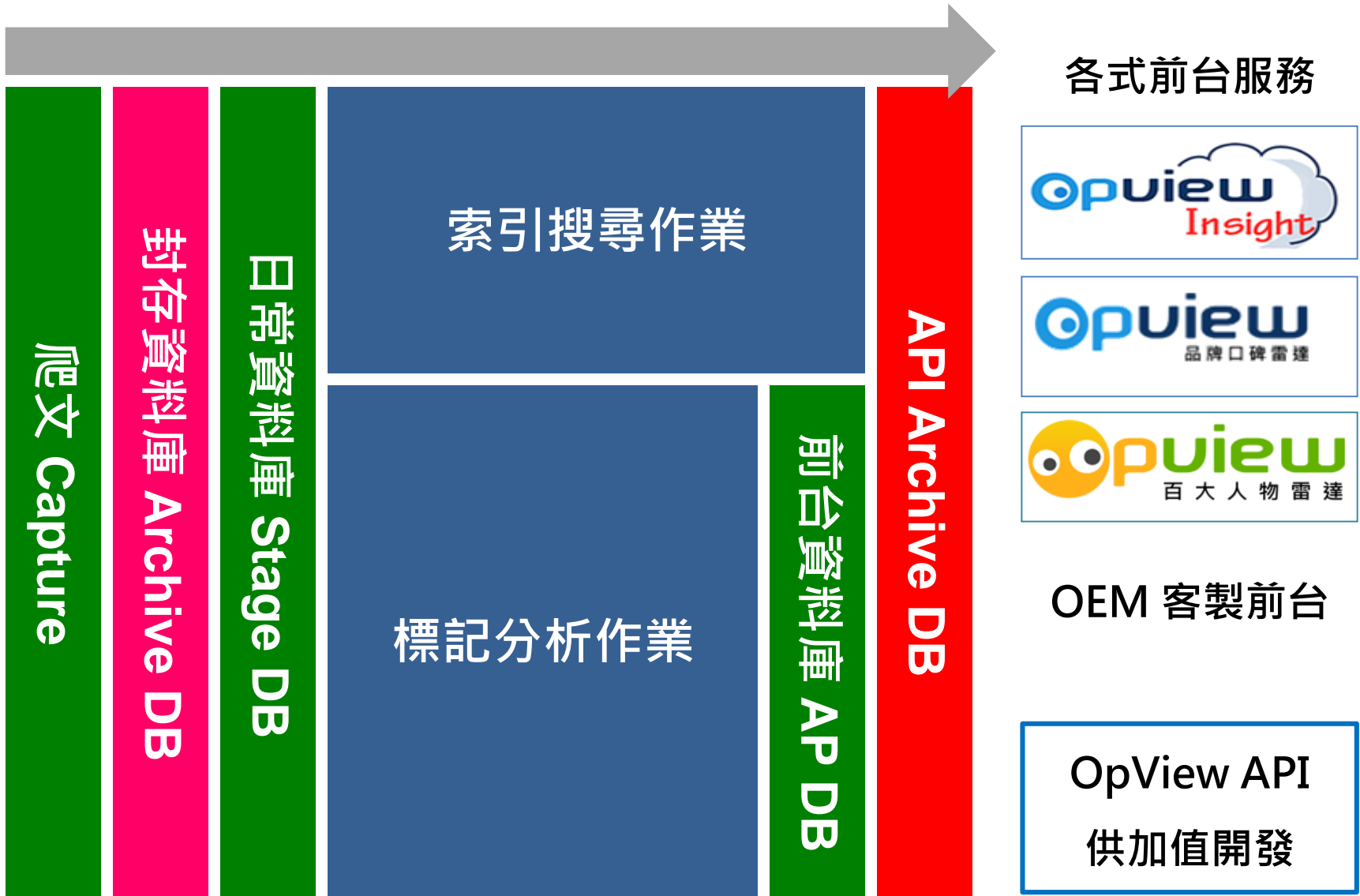
Opview  
Service (SaaS)

呈現

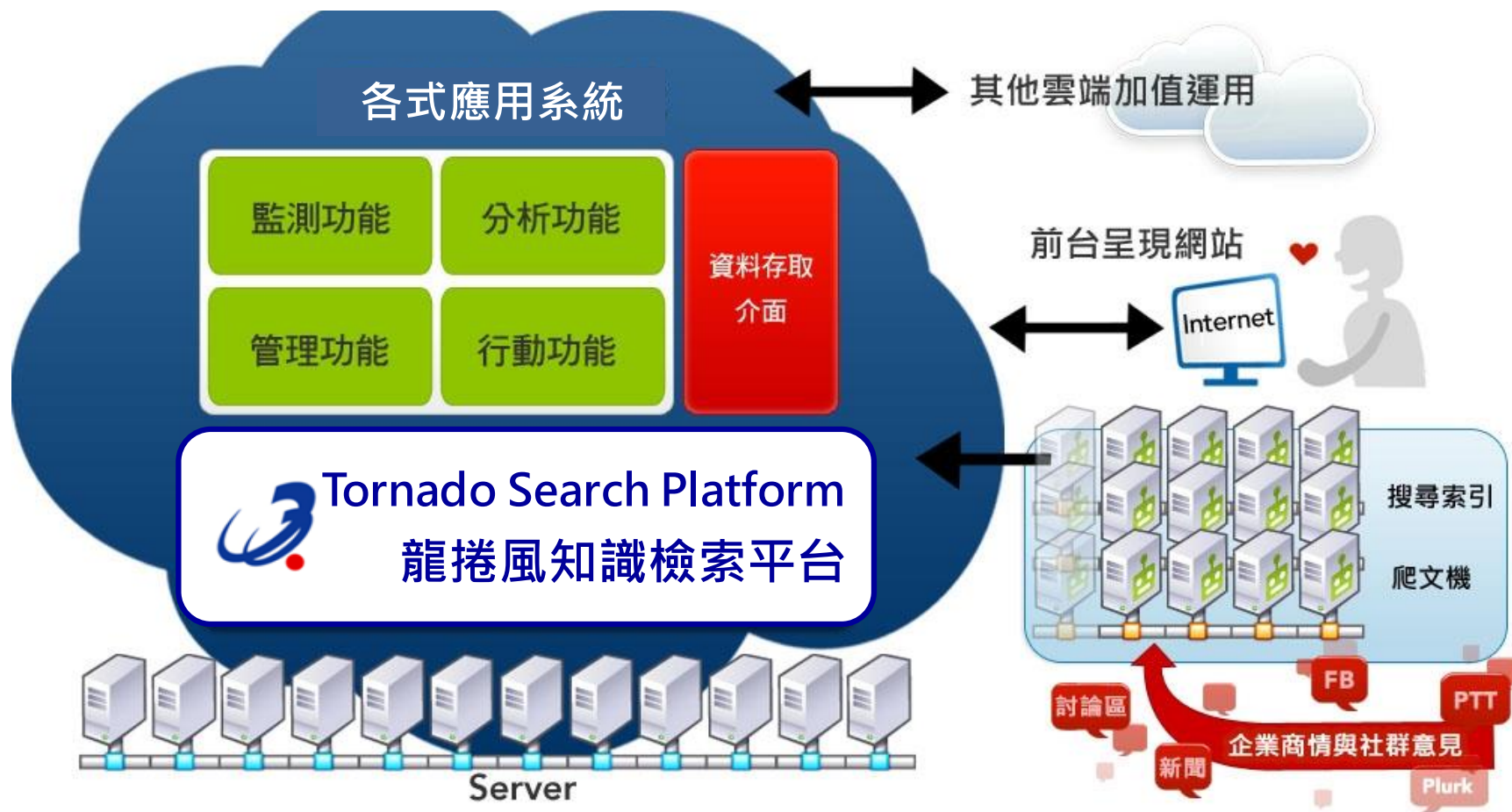
Opview  
Open API



# OpView – 多層式系統架構



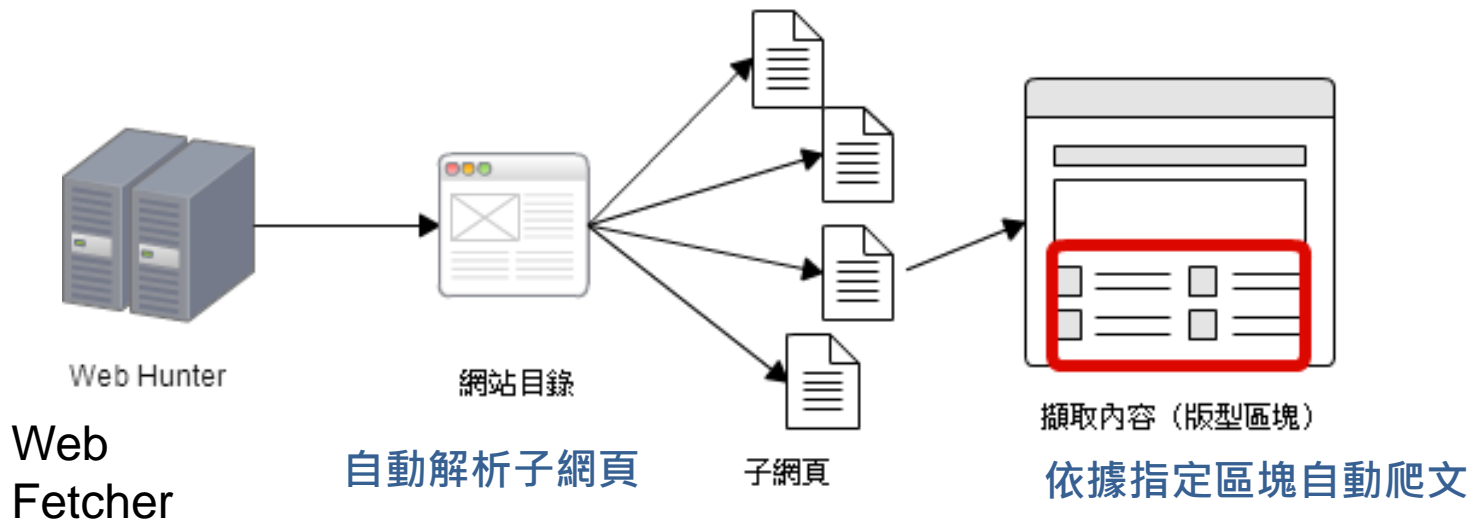
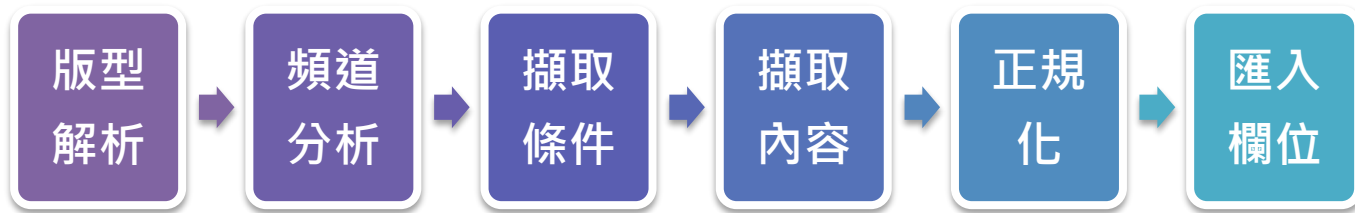
# 以雲端佈署方式提供服務



各服務功能具有完善的備援機制

# 資料源定期擷取更新，資料穩定、品質高

- ◆ 重要來源每小時更新
- ◆ 資料源每日至少更新三次





# OpView 輿情觀測平台 畫面範例

**查詢條件**

**期間：**

- 最近15天
- 最近30天
- 2013/06
- 設定開始日期
- 設定結束日期

**來源：**

- 討論區
- 社群網站
- 部落格
- 新聞

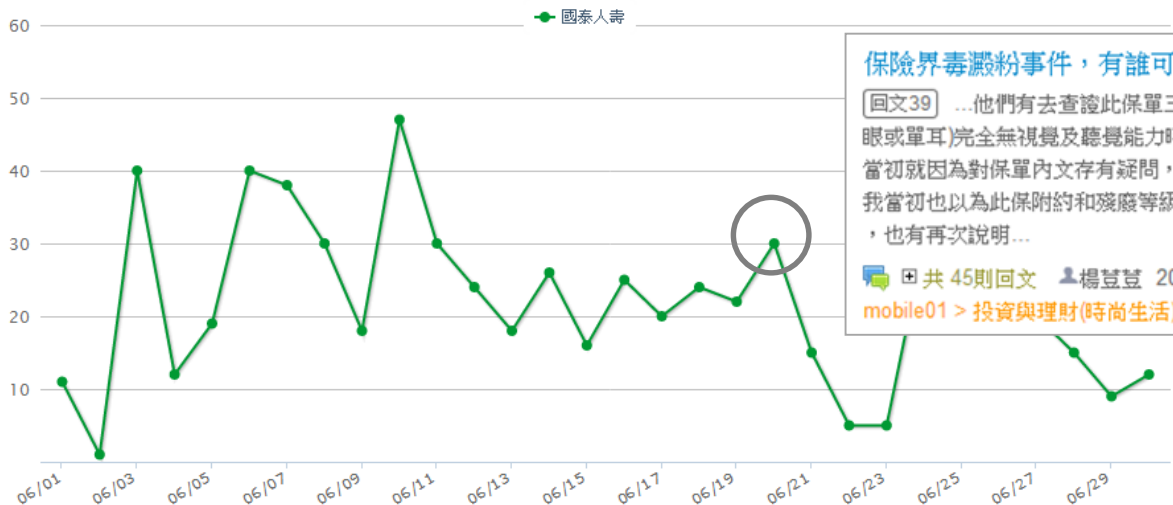
**主題：**

- 國泰人壽
- 南山人壽
- 富邦人壽
- 三商美邦
- 新光人壽
- 意外險
- 儲蓄險
- 投資型保單
- 保險業

**維度：** 不指定維度

**關鍵字篩選：** 請輸入搜尋關鍵字...

- 不指定維度
- 醫療險
- 意外險



**保險界毒澱粉事件，有誰可以幫幫我，我只是一個無助的媽媽！**

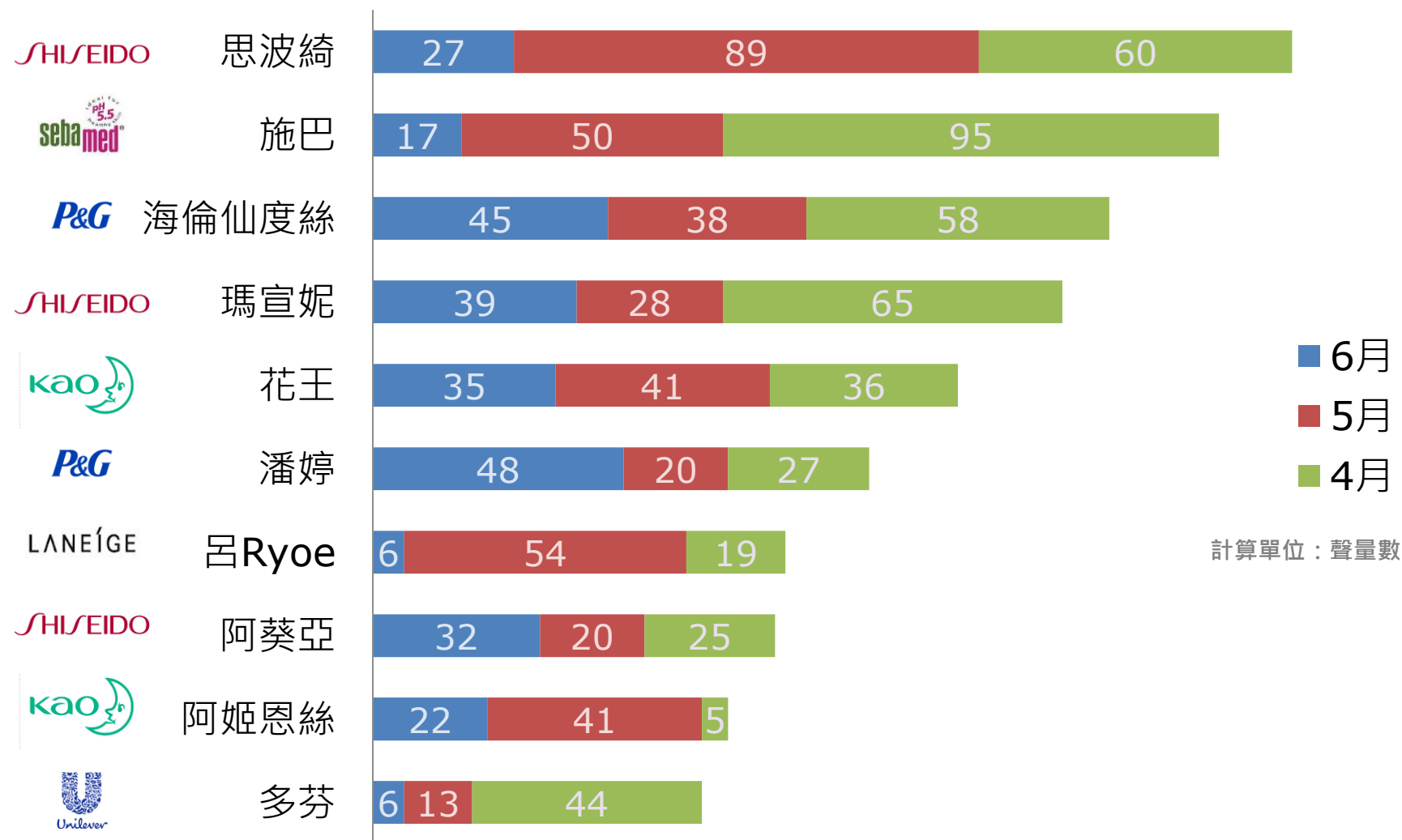
回文39 ...他們有去查這此保單三十七條(新生兒失明)及三十八條(新生兒失聰)這兩條的的內容都的確有(眼或單耳)完全無視覺及聽覺能力時，會讓消費者對於 **保險** 內容的解讀和 **保險** 公司產生不同的認知，當初就因為對保單內文存有疑問，所以才問專業的 **保險** 業務人員，但所得來的結果.....卻是各說各話，我當初也以為此保附約和殘廢等級有相關，一開始也有問及業務員，業務員也告知我保單設計和殘廢等是，也有再次說明...

共 45則回文 揚登登 2013-06-30 00:21:00 共 11726 次點閱 😊 0.226 😞 0.574

mobile01 > 投資與理財(時尚生活) 庫存內容 加入追蹤

# 解密一：生活消費品產業 (1)

## ◆ 從口碑聲量看品牌知名度



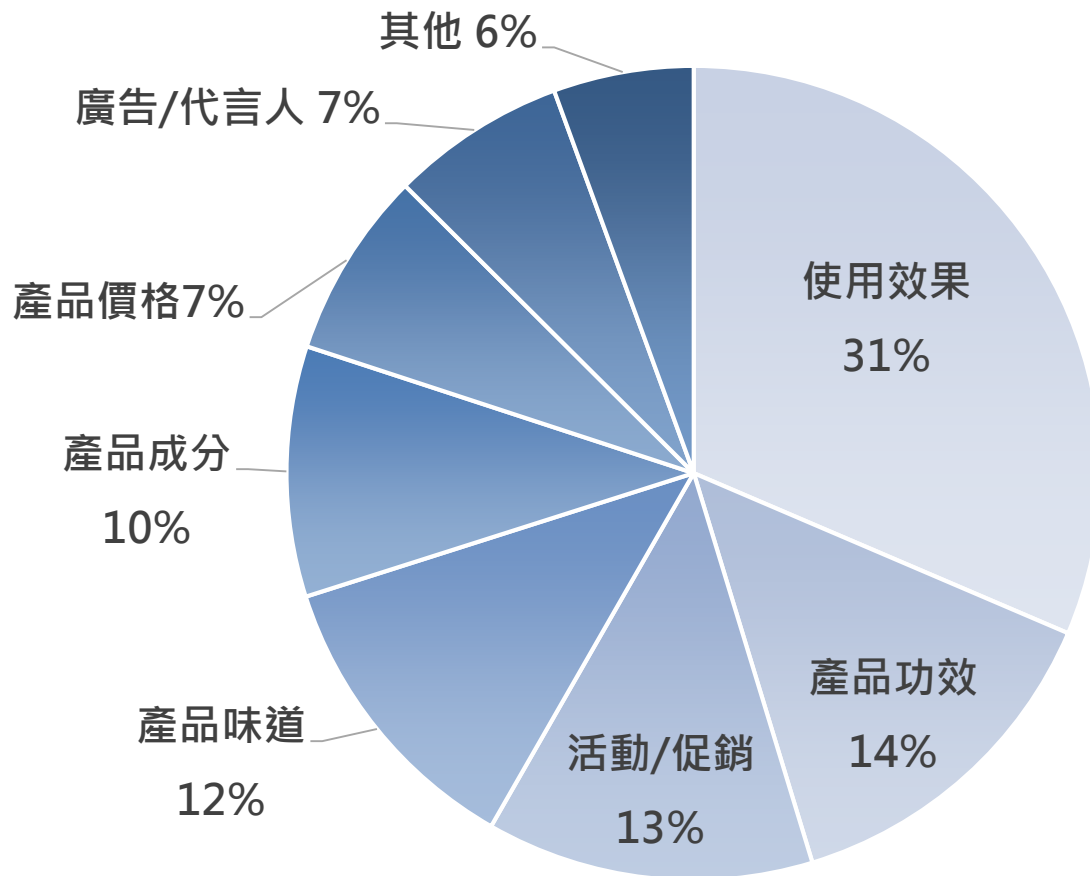
# 解密一：生活消費品產業 (2)

## ◆ 用情緒分析看品牌關係，掌握網友偏好

品牌	總聲量數	正面情緒 總聲量數	正面情緒 比例	負面情緒 總聲量數	負面情緒 比例	P/N比
阿姬恩絲	65	36	0.55	5	0.08	7.20
阿葵亞	69	55	0.80	12	0.17	4.58
瑪宣妮	125	96	0.77	22	0.18	4.36
呂Ryoe	75	60	0.80	14	0.19	4.29
思波綺	149	86	0.58	26	0.17	3.31
施巴	156	96	0.62	37	0.24	2.59
潘婷	80	42	0.53	19	0.24	2.21
花王	111	76	0.68	37	0.33	2.05

# 解密一：生活消費品產業 (3)

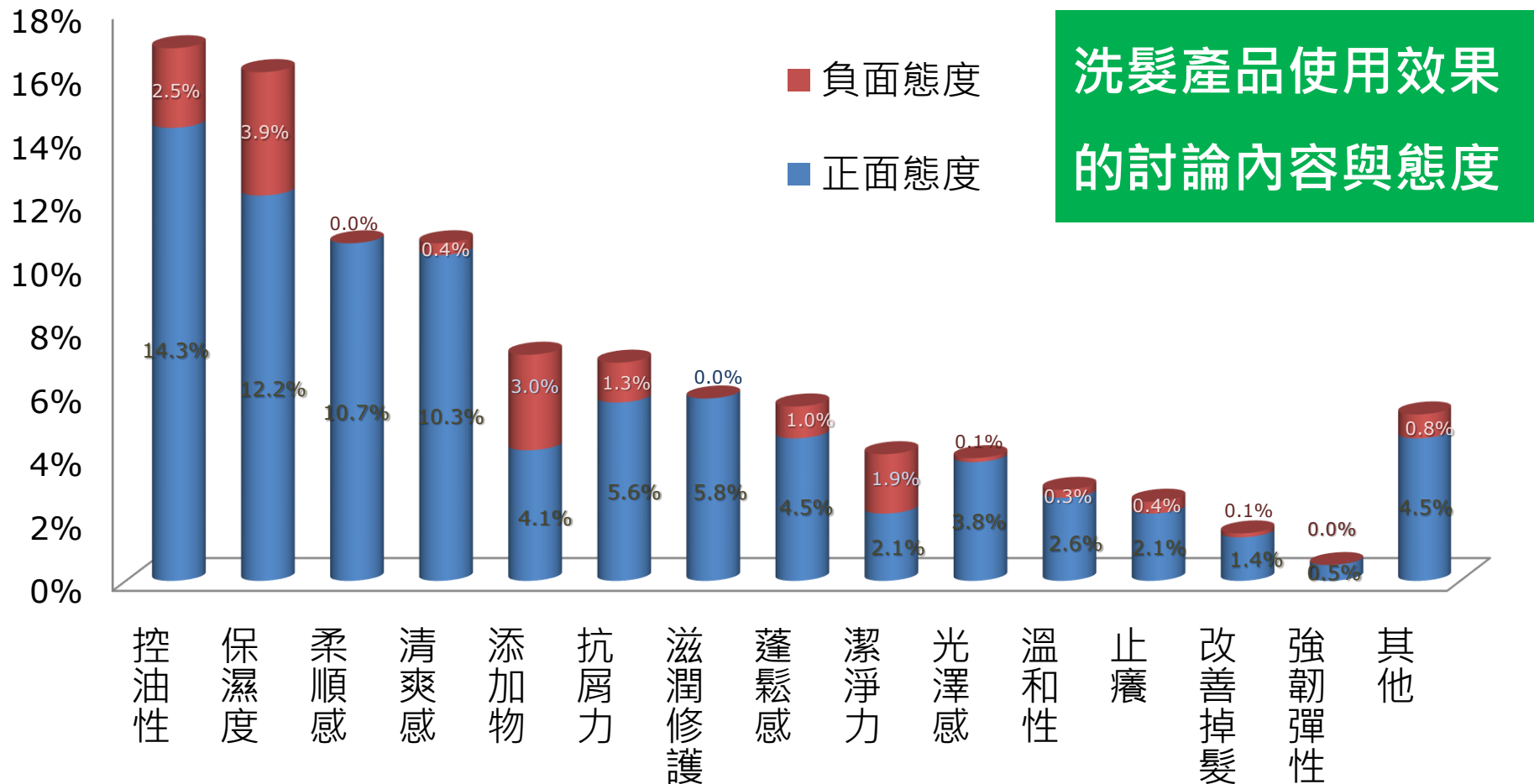
- ◆ 用語意技術分析消費行為，找出關鍵決策因素



消費者提及洗髮精  
時所關心的內容

# 解密一：生活消費品產業 (4)

◆ 用語意技術分析消費行為，找出關鍵決策因素



# 解密一：生活消費品產業 (5)

## ◆ 用語意技術分析消費行為，看心佔率

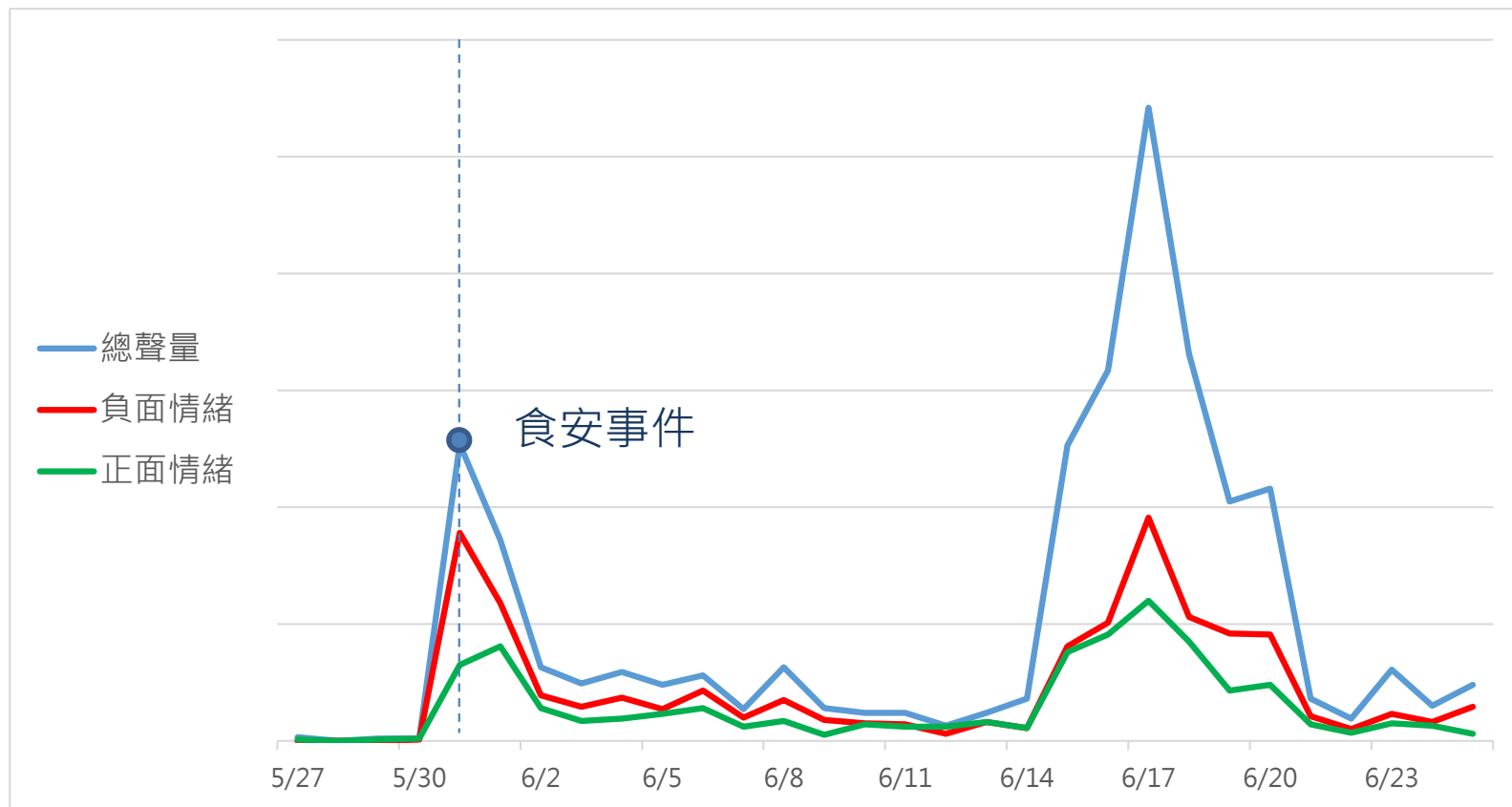
### ● 品牌於該項效果提及之比例

效果	品牌	占比	效果	品牌	占比
控油	施巴	31.6%	蓬鬆	瑪宣妮	75.0%
柔順	思波綺	31.8%	深層清潔	花王	41.2%
清爽	施巴	29.3%	光澤	思波綺	53.3%
不含有害添加物	花王	45.5%	溫和	施巴	57.1%
抗屑	海倫仙度絲	57.8%	強韌彈性	潘婷	50.0%

# 解密二：食品安全 從發現到處理 (1)

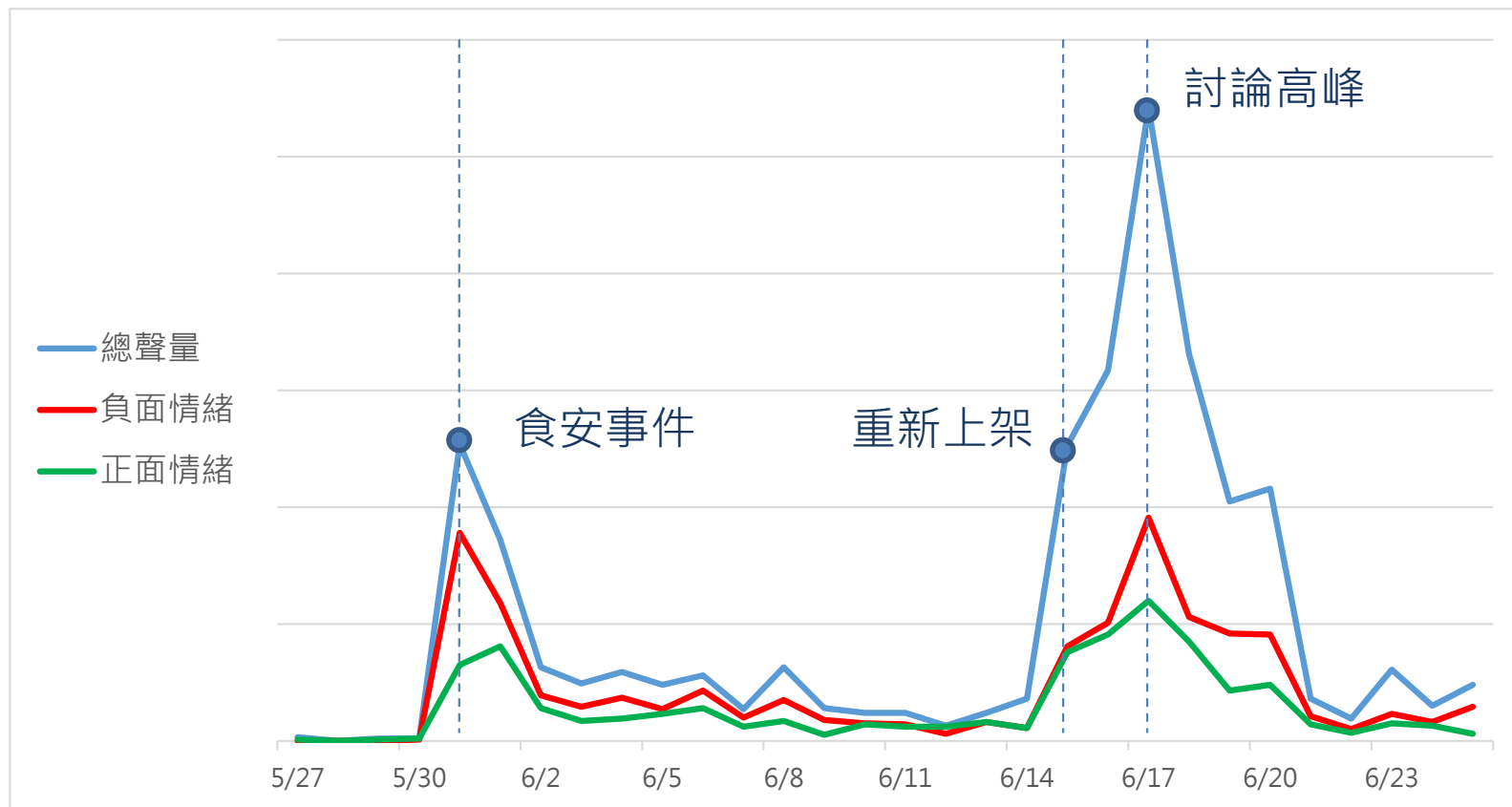
## ◆ 某品牌布丁 5月30日 當日各項指標急速惡化

- 品牌聲量竄達 5.9倍；單品聲量達 112倍；負評率 70.1%；P/N 0.37



## 解密二：食品安全 從發現到處理 (2)

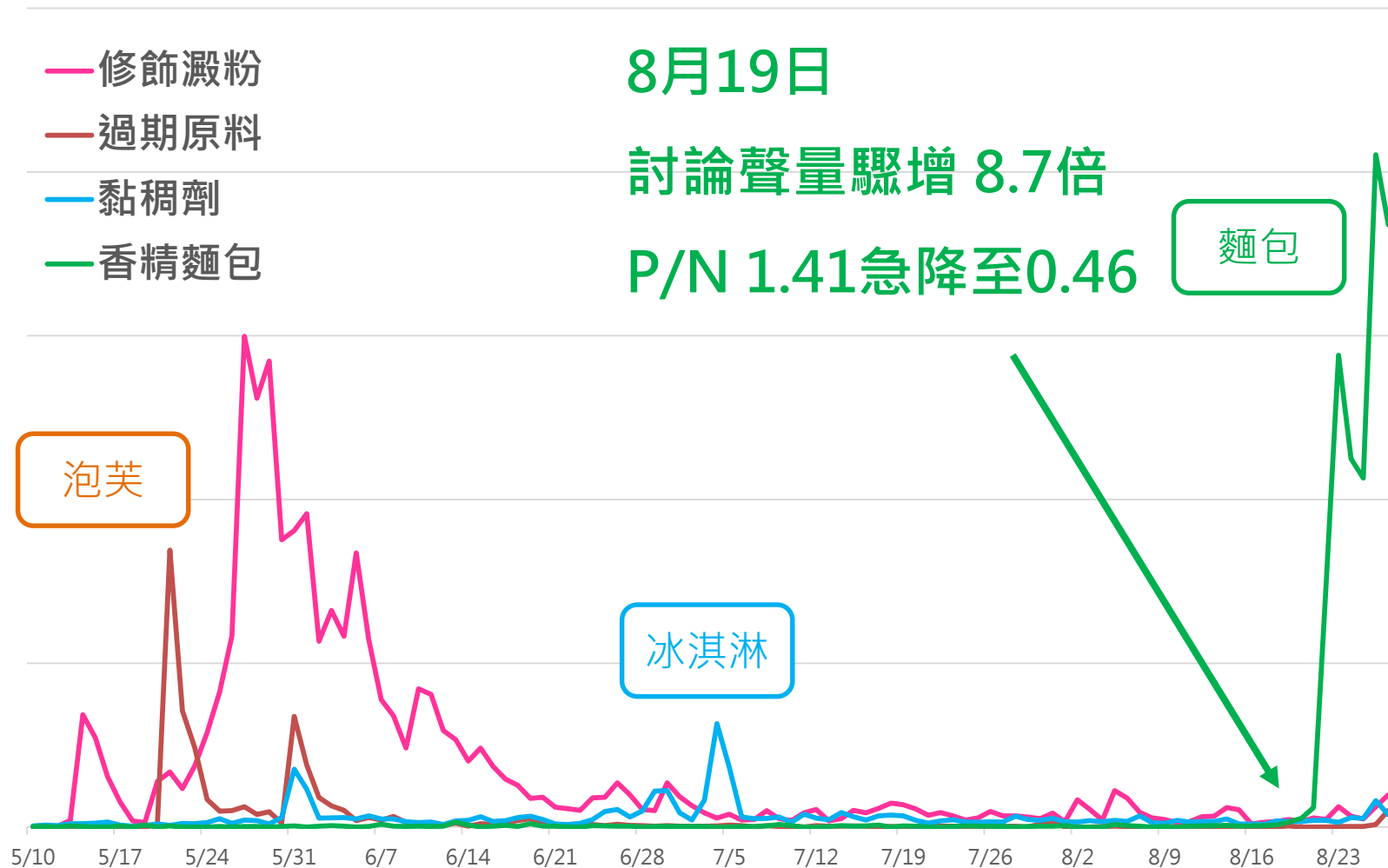
- ◆ 某品牌布丁 6月15日 重新上架，48小時內聲量達最高峰
  - 討論聲量倍增，負評率 70.1% → 32%；P/N 0.37 → 0.94 恢復水準





# 解密二：食品安全 輿論風暴比一比

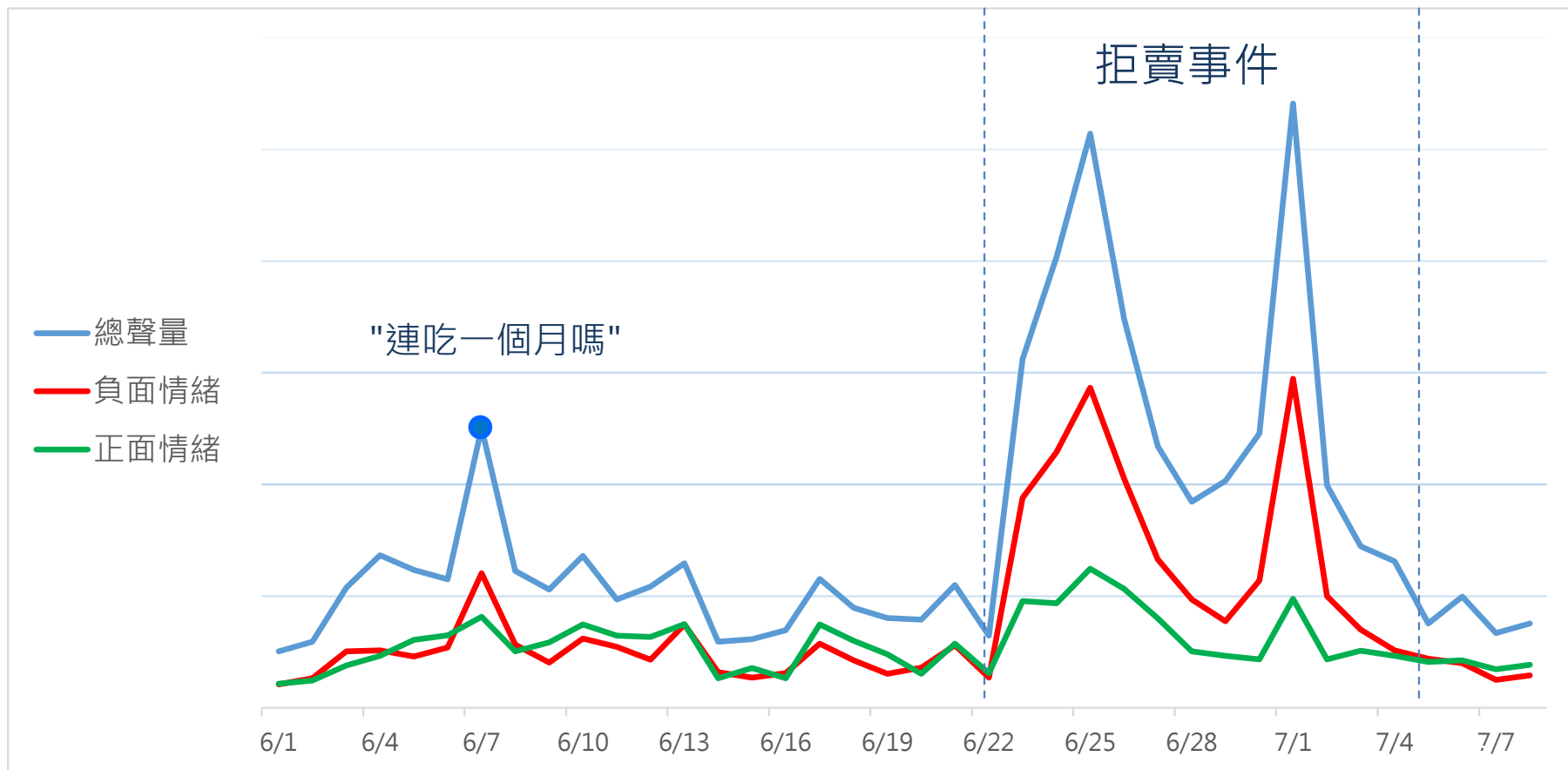
## ◆ 網友關注程度 與 擴散速度



# 解密三：從社群輿論看品牌形象

## ◆ 連鎖速食餐飲拒賣事件：黑暗的13天

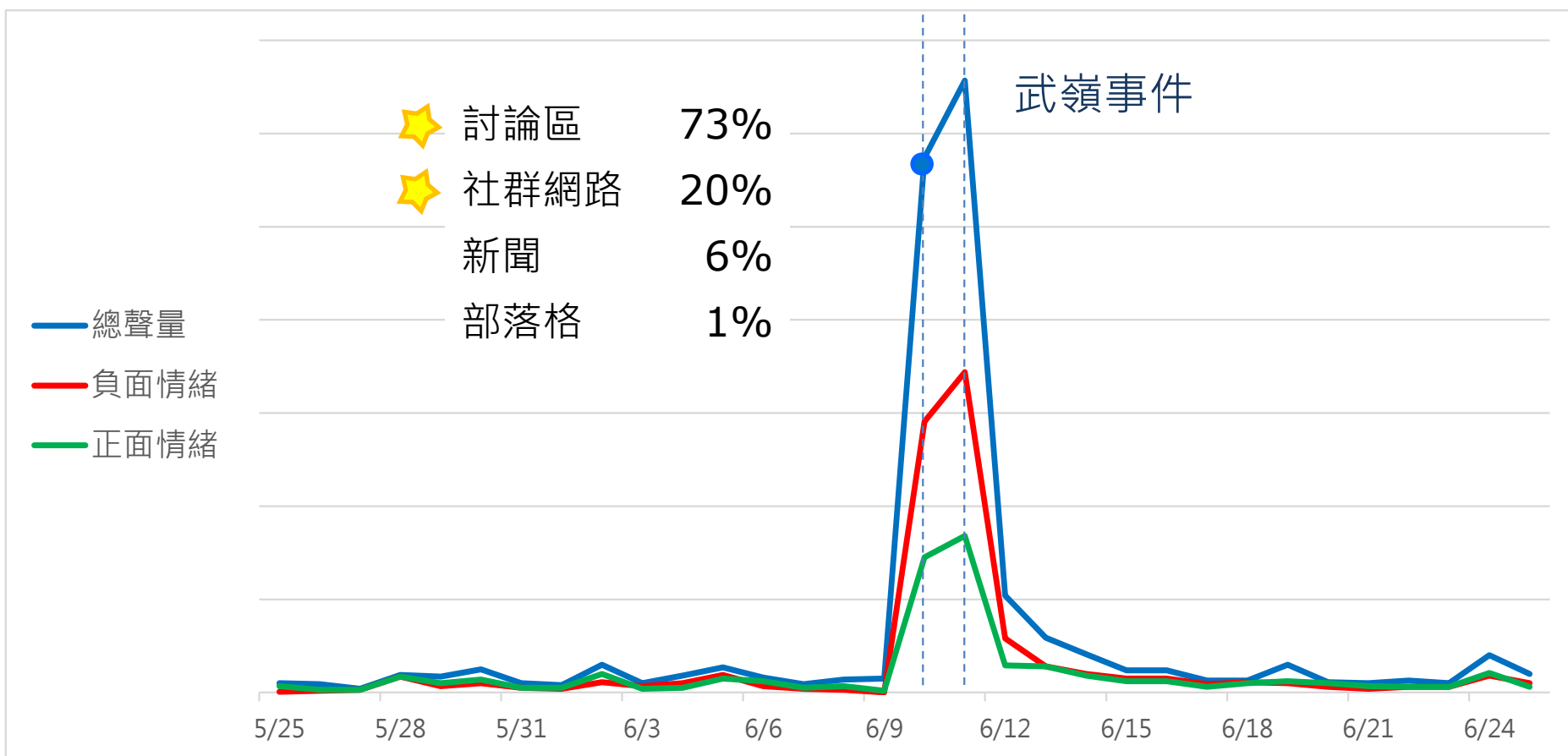
- P/N 逆轉，日均值 1.11 → 0.47 ( 負評率50±5%，正評率急降一半 )



# 解密四：從社群媒體看議題事件 (1)

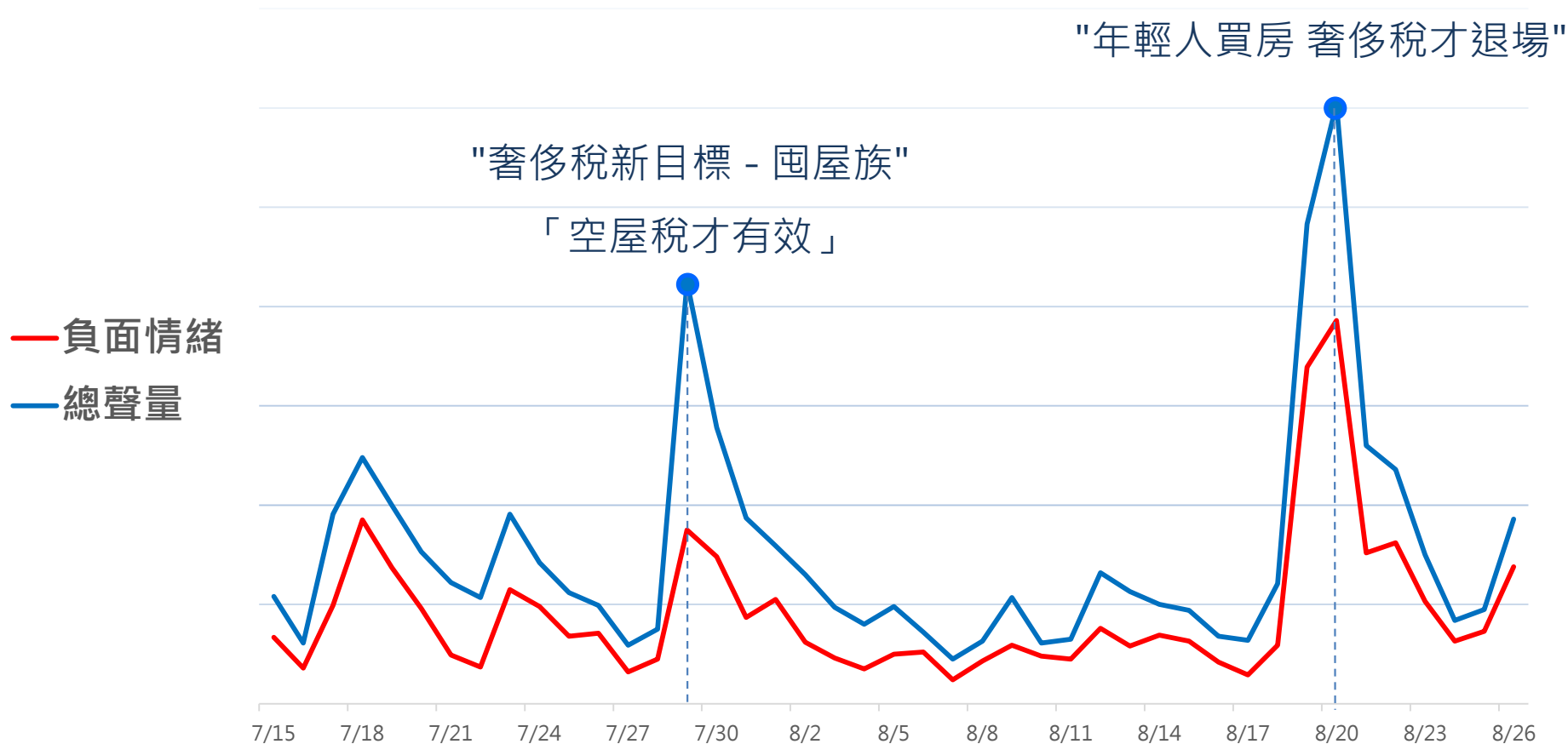
## ◆ 「微笑單車上武嶺」：機會或威脅？

- 聲量急拉 41.7倍；負評率 52% 不變，正評率 49%→25% 急降一半



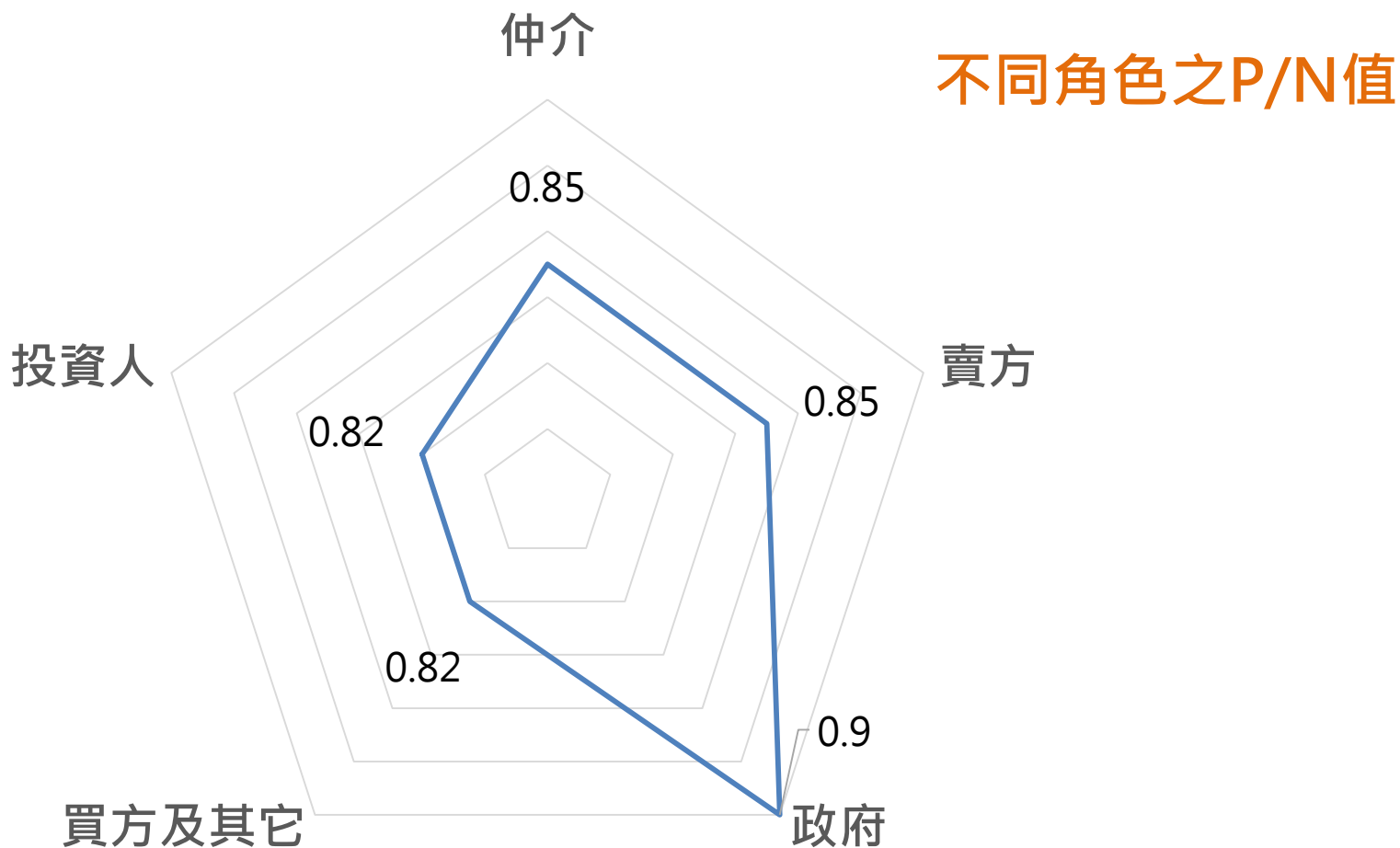
# 解密四：從社群媒體看議題事件 (2)

◆ 奢侈稅議題：15日均值 負評率64%▼ P/N值0.87▼

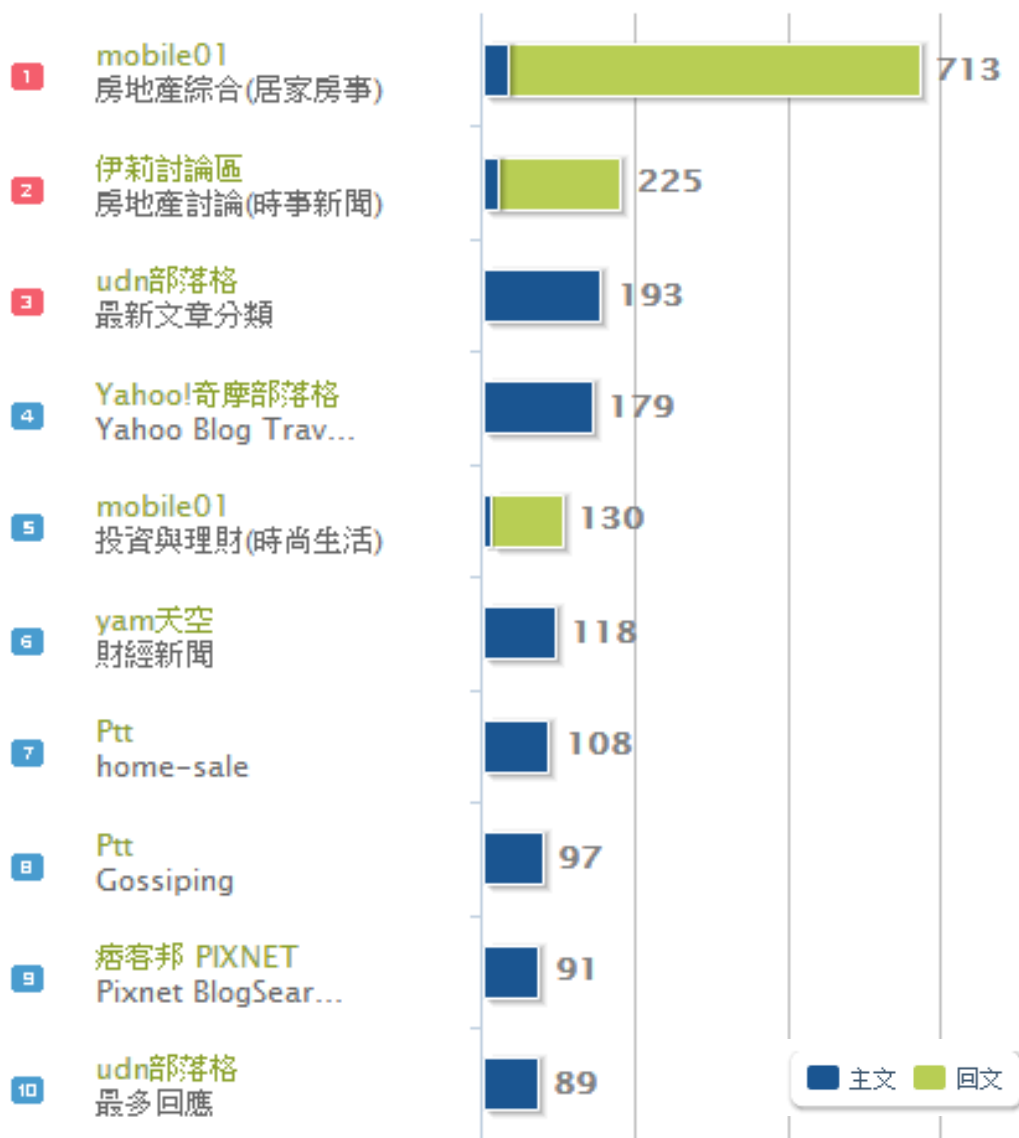


# 解密四：從社群媒體看議題事件 (3)

## ◆ 奢侈稅議題：用語意技術看關係人態度



# 解密四：從社群媒體看議題事件 (4)



## 奢侈稅議題

◆ 主要輿論來源

◆ 擴散分析

● 前十名集中度 43.4%

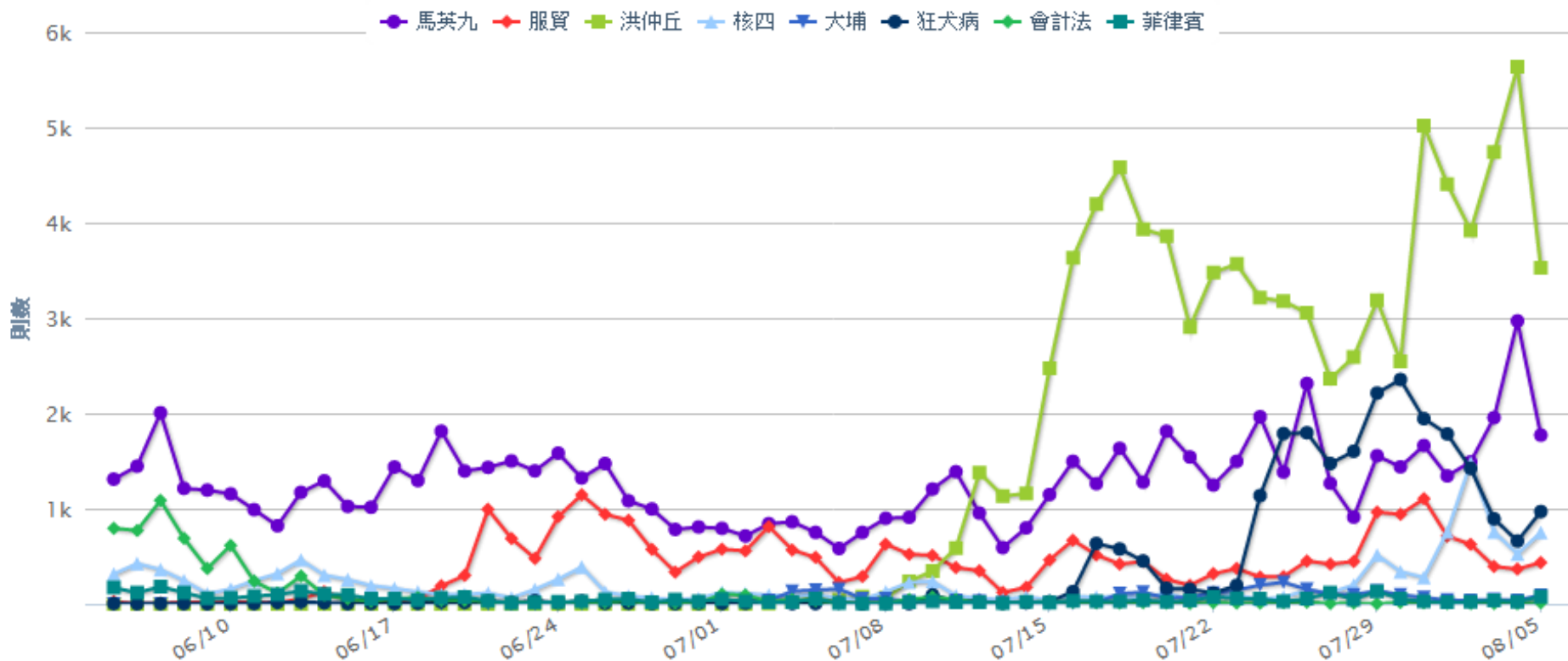
分散廣

● 回應比例低，無明顯

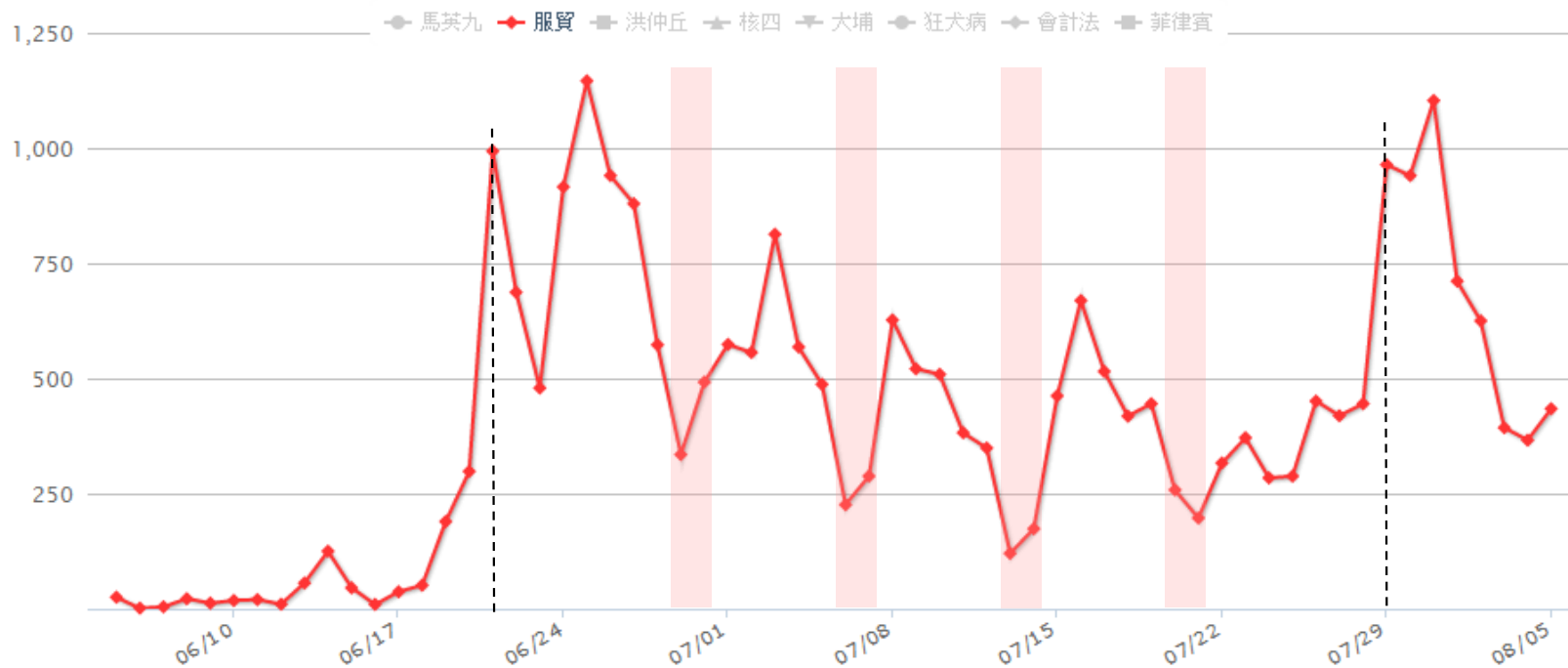
意見領袖

# 解密五：重大議題比一比 (1)

## ◆ 今年6月~7月政府重大議題觀測



# 解密五：重大議題比一比 (2) 服貿協議



■ 6/21簽訂創下第一波聲量高峰

■ 後續相關曝光起伏有周期性漸減；周末量少

■ 7/29 立院臨時會 聚焦【服貿協議】【核四公投】引發曝光高峰。



# 解密五：重大議題比一比 (3) 大埔事件

## ◆ 以風暴圖分析相關性

時間區間：2013/07/01~2013/08/31

資料來源：討論區，社群網站，部落格，新聞

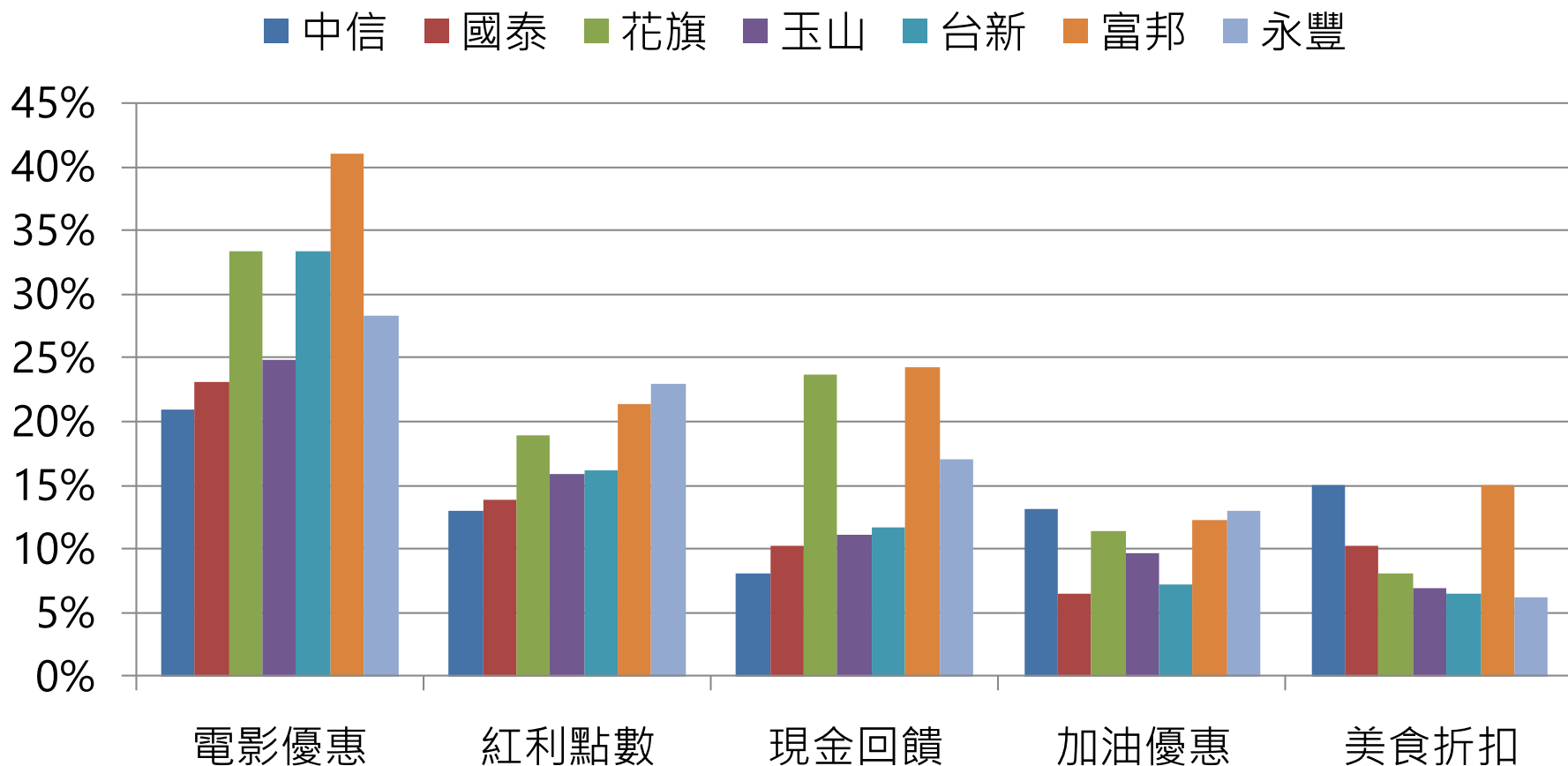
競爭中心	
大埔	
競爭主題	
吳敦義	輕度
江宜樺	中度
馬英九總統	超級強烈
郝龍斌	輕度
朱立倫	輕度



# 解密六：用口碑大數據 了解民眾看法 (1)

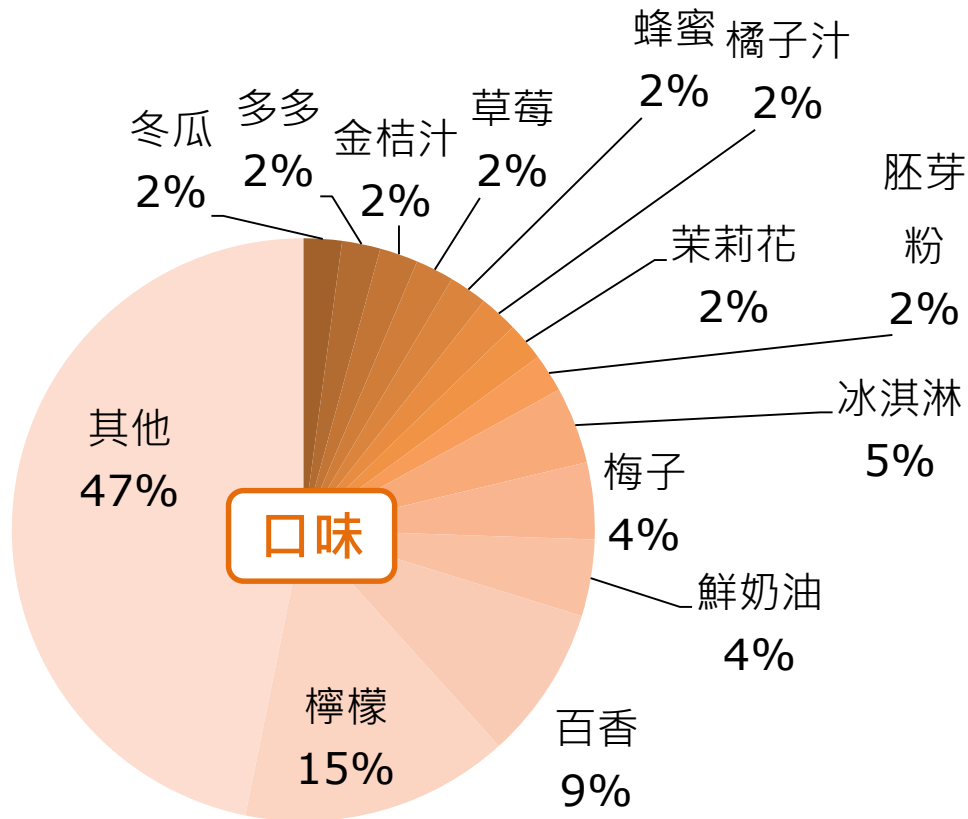
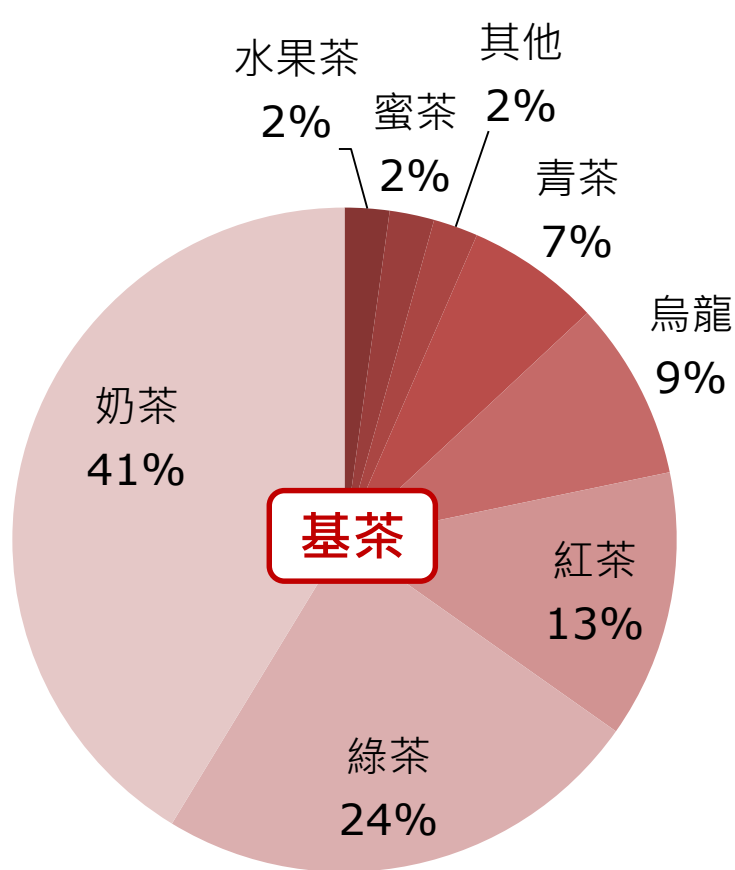
## ◆ 比傳統調查方法更為準確快速

- 以信用卡產業為例，分析用卡原因



# 解密六：用口碑大數據 了解民眾看法 (2)

## ◆ 以手搖飲料為例：民眾消費行為分析



# 解密六：用口碑大數據 了解民眾看法 (3)

## ◆ 以保險產業為例：品牌及產品偏好分析

數據：P/N 比

	醫療險	意外險	儲蓄險	投資型保單
南山人壽	1.17	1.24 	1.47	0.89
國泰人壽	1.32 	1.17	1.83 	0.67
富邦人壽	1.01	1.10	1.37	0.93
三商美邦	0.75	0.80	1.05	2.38 
新光人壽	0.98	0.86	1.00	1.00



# 用Big Data站上制高點 – 應用模式

## 觀測

- 大眾民意
- 議題事件
- 利害團體
- 意見領袖

## 分析

- 情緒走勢
- 內容變化
- 相關輿論

## 預測

- 民意選情
- 電影票房
- 流行疾病
- 股價走勢

# 對企業的應用 – 擁抱網路社群大數據

- ◆ 用聆聽與觀測踏出第一步
  - 各種量化與質化指標
- ◆ 建立社群輿情管理循環
- ◆ 走出迷霧中的作戰
  - 看清目標，量測成效

社群輿情管理循環

Listening

自動化網路輿情監測平台

- 覆蓋完整 即時性高

Analysis

語意分析技術

- 提供 數量/趨勢 與 議題/情緒
- 影響力分析 (媒體, 來源, 意見領袖)
- 競爭分析
- 擴散分析

Targeting

辨識及設定目標

Engaging

互動與溝通

Measuring

擴散與衡量

- 前後期成效 + 持續追蹤



# » 問題討論