

# 結合文字探勘與量化工具從線上留言挖掘旅館業者的競爭優勢

- 線上留言對於了解消費者需求提供一個異於過去的重要管道，也成為旅館業者制定品牌策略的重要資訊來源。
- 對於如何利用線上評論來了解消費者需求並進一步找出各旅館的競爭優勢，在文獻上仍然存在有缺口。
- 利用文字探勘工具找出消費者在乎的旅館屬性，各旅館在各屬性上的表現，並利用對應分析找出消費者心中各旅館的群組歸屬
- 最後結合消費者購買行為與邏輯斯迴歸找出各旅館的競爭優勢。

## 研究背景

- 入境的旅遊人數增加但是星級旅館的住房率反而減少，主要原因之一歸諸於 Airbnb 提供房間數的增加，造成旅館業直接的財務影響包括住房率
- Airbnb 的崛起對於台灣旅館產業星級飯店的營運已形成一個重大的挑戰
- 產生一個獨特且具差異化的品牌定位
- 旅館產業過去的相關文獻，通常都是用問卷調查來探索旅館的品牌定位及其競爭優勢
- 問卷題項侷限在之前定義好的問題上，也常因為樣本數不大，造成樣本代表性受到質疑
- 大部分的研究聚焦在品牌的建構，顯示出對於社群媒體與品牌策略的相關研究仍然存在有很大的缺口。
- TripAdvisor 上的線上留言作為研究對象來說明如何利用線上評論來做旅館品牌的建立與競爭優勢的探索。

## 文獻探討

- 品牌定位
- 旅館的品牌屬性

- 包括形象、價格/價值、旅館、服務、房間、行銷、食物與飲料、安全、地點
- 消費者的線上評論
  - 線上評論的平台讓消費者可以用文字、照片、甚或影音來表達他們對產品、服務、品牌的知識、意見、感受與看法等。

## 研究方法

- 應於品牌建立的三階段
- TF-IDF 值計算來決定術語的重要性
  - 情緒分析  
AFINN 情緒字典
  - 對應分析(分群)  
邏輯斯迴歸
- TripAdvisor
- 關鍵字分析
  - 字詞擷取方法多以 TF-IDF 為主。TF 指的是詞頻 (term frequency)，是依字詞在某一文件中的出現頻率代表該字詞的重要性，頻率越高則表示越重要，而 IDF (Inverse Document Frequency) 是依據字詞在文件中出現的比例，依一定公式計算其權重。
- 情緒分析的目的是檢測和分類評論者所表達的情感
  - 第一種  
是基於詞彙的比較方法，根據網路口碑（eWOM）中已經確定是正面或是負面情緒的句子來計算評論的語義方向，這種方法通常需要依照不同情境下的情緒詞典來作比較分析；
- AFINN 情緒辭典較常用在旅遊業領域

## 結論

情緒值為正的前三項，為晚餐(547)、房間整潔(333)與早餐(321)

在品牌建立的三個步驟(認知、情感與意動)中，過去的研究在認知與情感前兩個步驟的操作上有比較完整的說明，然而在意動部分可能由於牽涉到消費者的購買行為資訊的不容易取得，使得在文獻上存在有很大的缺口，

只挑選台中西屯區五

間有名奢華旅館的 500 筆資料來分析