



行銷4P架構下的 購物籃分析應用

影片、程式碼與簡報連結:

https://drive.google.com/drive/folders/12TU89Zxt66jH21iPia38DogVux0LmUf9?usp=drive_link

<https://youtu.be/A50yzZxUDrk>

第21組

柯宥圻 B10703049 財金二

曾繁宸 B10703022 財金二

林品柔 B09303130 經濟三

陳柏宇 B07302249 政治五

陳瑄 B08208045 地理四



影片大綱

- 專案結合行銷4P架構策略
- 行銷4P之一：Place
- 行銷4P之二：Price
- 行銷4P之三：Promotion
- 行銷4P之四：Product
- 購物籃分析：介紹
- 購物籃分析：前處理
- 購物籃分析：規則創造
- 購物籃分析：個別分析
- 購物籃分析：規則介紹
- 購物籃分析：輸出解釋
- 購物籃分析：未來展望

專案結合行銷4P 架構策略

1. 第一步：資料分析找出通路、價格和促銷方式
2. 第二步：購物車分析找出要推播的產品
3. 以人事時地物等面向發展，最終收斂成一個可行的商業提案





行銷4P之一：Place

- 分析會員資料、主單資料和行為資料
- 依照不同廣告來源，進行年齡與會員等級的次數加總
- 最終取得容易促使不同年齡層、會員等級的會員完成購買的廣告通路

行銷4P之一：Place

排名	<= 20歲	Score	20~30歲	Score	30~40歲	Score	40~50歲	Score	>= 50歲	Score
1	app_sharing	8	affiliate	827	affiliate	2008	FB_CA	1625	FB_CA	1576
2	LINEOA	7	FB_CA	366	FB_CA	1303	affiliate	1368	affiliate	762
3	omnichat	5	gkw	246	omnichat	390	omnichat	529	omnichat	581
4	FB_CA	4	google	167	gkw	365	LINEOA	390	LINEOA	458
5	affiliate	4	omnichat	144	LINEOA	364	System_Inform	371	google	418
6	gkw	4	LINEOA	131	google	328	google	295	System_Inform	389



行銷4P之一：Place

排名	10級(一般)	Score	20(中等)	Score	30(高級)	Score
1	FB_CA	2808	affiliate	1153	affiliate	1507
2	affiliate	2309	FB_CA	1139	FB_CA	927
3	omnichat	920	omnichat	392	System_Inform	379
4	google	749	LINEOA	321	Audience	370
5	LINEOA	741	System_Inform	288	omnichat	337
6	gkw	619	google	284	LINEOA	288

行銷4P之一：Place

- 針對年輕族群：app_sharing和LINE廣告通路
- FB和affiliate廣告通路利用：網紅行銷作為一種管道的可能性





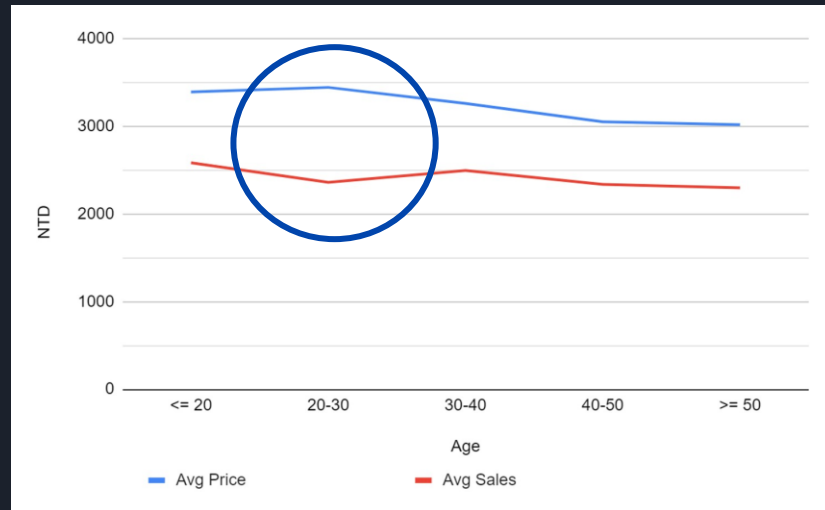
行銷4P之二：Price

age	Qty	TotalPrice (每單銷售總金額)	TotalDiscount (每單折扣總金額)	TotalSalesAmount (每單總金額)	Avg Price (每樣商品平均單價)	Avg Sales (每樣商品平均消費金額)
<= 20	2.41	7475.93	-1882.79	5593.14	3391.88	2583.73
20-30	2.20	6913.48	-1767.93	5154.55	3444.00	2361.64
30-40	2.23	6695.57	-1720.18	4975.39	3261.20	2496.32
40-50	2.24	6402.01	-1639.91	4762.10	3051.42	2337.97
>= 50	2.26	6475.91	-1700.95	4774.96	3018.28	2298.13

行銷4P之二：Price



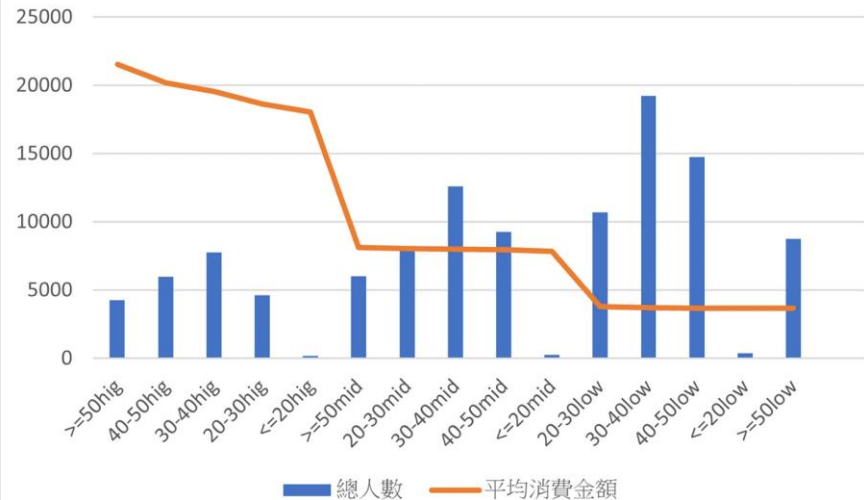
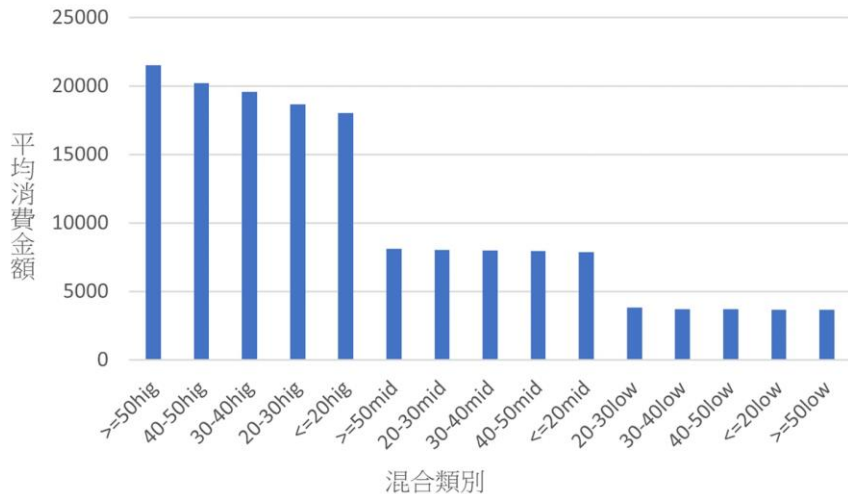
- 年齡層越高，每筆訂單消費總金額越低



- 折扣使用情形之比較

行銷4P之二：Price

- 年輕族群消費總額高、使用折扣的情形也較多
- 結合會員等級資料後仍為相同結果
- 中高齡會員人數較多，折扣策略主推中高年齡族群





行銷4P之三：Promotion

- 分析折扣使用的熱門時段
- 在最合適的時段寄出線上折扣

消費時間 (時段)					
index	TotalSalesAmount	TotalPrice	average_spending	order_num	discount_percentage(%)
df0_6	22123378	26467610	3422	6465	62.5
df6_12	52121894	66499461	3492.5	14924	68.1
df12_18	388458508	536151650	4939.6	78642	82.5
df18_24	479816297	640500484	5233.5	91681	78.8

行銷4P之三：Promotion

消費時間 (幾點)					
index	TotalSalesAmount	TotalPrice	average_spending	order_num	discount_percentage(%)
df0_	9552640	11473031	3355.3	2847	61.9
df1_	5682814	6745674	3507.9	1620	62.3
df2_	2891169	3399150	3409.4	848	63.4
df3_	1847957	2208737	3519.9	525	62.1
df4_	1067566	1289007	3432.7	311	64.6
df5_	1081232	1352011	3443.4	314	65.3
df6_	1861840	2238782	3295.3	565	61.4
df7_	3478924	4233270	3251.3	1070	58.8
df8_	5238997	6158659	2953.2	1774	58.2
df9_	6641731	7789769	3164.2	2099	61.6
df10_	11595871	14764293	3281.2	3534	69.6
df11_	23304531	31314688	3962	5882	74.9
df12_	36189750	49603766	4385.6	8252	77.2
df13_	48209889	66068609	4712.6	10230	80
df14_	62505336	86435405	4799.2	13024	82
df15_	75962589	105938554	4962	15309	84.1
df16_	82752242	114474775	5106.6	16205	84
df17_	82838702	113630541	5302.7	16222	84.1
df18_	78986411	107600256	5339.1	14194	82.9
df19_	91611069	123838316	5176.4	17698	80.4
df20_	120686964	161467117	5333.5	22628	80.6
df21_	132894543	176634780	5545.8	23963	79.3
df22_	38114045	49561724	4870.2	7826	70.6
df23_	17523265	21397871	3672.1	4772	62.7



行銷4P之四：Product

- 以上述通路、時間和折扣策略為基礎
- 根據購物籃分析，找出最適合用來推播的產品
- 購物籃分析系統：

輸入目標受眾 → 計算關聯規則 → 獲得購物籃推薦



行銷4P之四：Product

- 20歲以下推薦：印花、logo、白、有機棉
- 20-30歲推薦：全粒面、皮革、防水、黃、小麥
- 30-40歲推薦：磨砂革、防水、logo、有機棉
- 40-50歲推薦：皮革、全粒面、防水、磨沙革、休閒
- 50歲以上推薦：有機棉、磨沙革、logo、休閒

購物籃分析：介紹

Rule: $X \Rightarrow Y$

Support = $\frac{freq(X,Y)}{N}$

Confidence = $\frac{freq(X,Y)}{freq(X)}$

Lift = $\frac{Support}{Supp(X) \times Supp(Y)}$

購物籃分析藉由指標找到高關聯性的商品組合，找出不同的會員分群中，不一樣的關聯性組合。

我們可以針對消費者性質以及購物資料進行推薦，當進行瀏覽活動時，針對不同、潛在的需求，給予折扣券，或聯合促銷，拉高單筆消費金額。

購物籃分析：前處理

關鍵字	出現次數
休閒	2392
防水	2000
有機棉	1975
磨砂革	1854
logo	1439
連帽	1293
...120個形容詞，各出現至少45次	

利用商品品項進行購物籃分析已屢見不鮮，我們想要創造新的潛在需求規則：產品的不同特性，是不是會員購買的依據？當客人購買某特性的物品時，是否會有潛在需求？我們藉此將每個商品貼多個標籤。

舉例來說，如果發現，當人們購買「抗UV」商品時，也對「防水」有顯著需求，則可以推測說該使用者族群可能喜愛進行戶外親水活動。

購物籃分析：規則創造

	buy	maybuy	support	confidence	lift
0	[1]	[111]	0.047393	0.222222	2.131313
1	[111]	[1]	0.047393	0.454545	2.131313
2	[2]	[4]	0.094787	0.526316	2.019139
3	[4]	[2]	0.094787	0.363636	2.019139
4	[2]	[7]	0.037915	0.210526	2.613003
5	[7]	[2]	0.037915	0.470588	2.613003
6	[2]	[11]	0.037915	0.210526	2.776316
7	[11]	[2]	0.037915	0.5	2.776316
8	[104]	[2]	0.033175	0.466667	2.591228
9	[2]	[110]	0.061611	0.342105	2.489111
10	[110]	[2]	0.061611	0.448276	2.489111
11	[3]	[5]	0.056872	0.375	3.596591
12	[5]	[3]	0.056872	0.545455	3.596591
13	[3]	[8]	0.042654	0.28125	3.296875
14	[8]	[3]	0.042654	0.5	3.296875
15	[3]	[112]	0.066351	0.4375	2.977823

我們以2022年的消費資料創造規則：年齡5組，等級3組，以及交感項15組，並針對符合條件會員的購物籃進行**Apriori**演算法。

此處的篩選條件是：support 0.015, confidence 0.2-0.9, lift 1.5 以上。support篩選出現次數太少的組合，而confidence設定不大於**0.9**，是透過觀察發現，若大於0.9，則一定是商品名稱的重複，而非消費者主要的需求。

購物籃分析：個別分析

```
Cxjx1zdA==': ['皮革', '綠'], 'ka5mb2  
4WPQ==': ['皮革', '針織', '灰', '褐'  
nZXUpHUDf85YLg==': ['休閒', '皮革',  
'GW4MPaR9q0HyW52ejyhfpQ==': [], 'TE  
布', '卡其.1'], 'vk+8quPhai6H2vLogK/  
iKQ==': ['帆布', '藍'], 'hE1khqzoR06  
nJNSUWrvttA==': ['休閒', '黑'], 'eJh  
'白'], 'OCMI8D/v4k6FaFTP38Jw9g==': [  
==': ['網布', 'UV', '黑'], 'CoFWNwV9  
KdNlmA==': ['網布', 'UV', '白'], '6i  
EuKXplu1RZexd96viN0Q==': ['黑'], 'U  
3EqVCNH5FAosX7S/B5gg==': ['休閒', '男  
'織布', '黑'], 'Vhf+Qsap/ofIMXRExTxj  
7bRkaIlgw==': ['休閒', '帆布', '綠']
```

Step1: 針對**商品資料**，建立商品名稱-標籤dict，以及特徵名稱-編號系統。

Step2: 輸入消費者ID，**尋找他2022年的購物資料**，轉換成編號並裝入一個list中。

Step3: 利用會員資料**分入不同群組**。

Step4: 遍歷年齡規則表、等級規則表，以及年齡-等級會員表，**尋找購買紀錄中，和關聯規則有關的規則，並加權**。

購物籃分析：規則介紹

加權方式：**Confidence*Lift**，為了展現該商品組合的正相關強度(lift)，以及不同方向之間的貝式機率。

同時，如果**潛在購物籃中的商品有n個**，則每個商品的權重降低 $1/n$ ，以修正偏誤

此外，如果**客戶個別購買list內某商品出現過n次**，也會直接將權重提高n倍。

最後，**年齡-等級交叉**

購物籃權重*3

```
if set(row["buy"]).issubset(set_buylist):
    for i in row["maybuy"]:
        multiplier = 1/len(row["buy"])
        if i not in appear:
            appear.append(i)
            output_dict[i] = row["confidence"]*row["lift"]*multiplier
        else:
            output_dict[i] += row["confidence"]*row["lift"]*multiplier
    for j in row["buy"]:
        output_dict[i] = output_dict[i] * Counter(num_buylist)[j]
```

購物籃分析：輸出解釋

```
# 結果：LOG版本(大樣本降低權重)
for i, j in sorted_dict.items():
    sorted_dict[i] = sorted_dict[i]/math.log(item[name])
sorted_dict2 = dict(sorted(sorted_dict.items(), key=lambda x: x[1]))
for i, j in sorted_dict2.items():
    print(name_dict_reverse[i], j)

有機棉 6.979628122040438
logo 0.7405476242429689
印花 0.5413713047733338
藍 0.46595798486537454
休閒 0.14309134043387547
白 0.12789110439951112

print(Counter(d))

Counter({'logo': 2, '胸前': 2, '燈芯絨': 2, '白': 2,
'多功能': 1, '藍': 1, '有機棉': 1, '工裝': 1, '抽繩':
1, '粉': 1, '灰': 1})
```

將前頁輸出結果除以 \log (特徵總體出現次數)以平衡母體特徵出現次數帶來的影響。

顯性需求：已經重複購買許多次的特徵，如logo、胸紋，而Logo也在推薦系統中以第二順位被推薦。

隱性需求：藉由個人購買資料與群體進行對比，尋找新的潛在需求，如「有機棉」是他最有可能有潛在需求的商品，可以進行強力推薦。

購物籃分析：未來展望

1	市場	335.96	董事長	218.91	鴻海	15829	郭台銘	5349.8
2	公司	293.1	北報導	200.37	夏普	10658	鴻夏戀	2238.9
3	營收	282.32	台北報	196.95	社長	956.9	戴正吳	1970.1
4	董事	281.52	金管會	184.63	面板	689.56	富士康	1412.1
5	客戶	261.8	壽險公	157.9	權值	652.94	富智康	1022.3
6	億元	261.4	總經理	134.56	簽約	565.82	呂芳銘	850.89
7	產品	258.52	供應鏈	131.65	日圓	558.93	權值股	635.41
8	今年	243.28	月營收	129.91	蘋果	557.11	大立光	578.91
9	銀行	231.21	管會昨	124.81	群創	511.38	諾基亞	560.85
10	獲利	219.83	作者是	122.3	液晶	487.01	蘋概股	487.48
11	業者	205.49	物聯網	116.09	台股	481.78	台積電	477.09
12	事長	203.05	金融業	113.99	集團	462.89	金融股	389.29
13	產業	200.85	昨天股	111.59	組裝	424.64	機器人	361.82
14	北報	196.95	據了解	109.29	出資	393.16	電子股	326.04
15	營運	187.64	每股純	107.53	可成	312.97	副總裁	261.67
16	去年	187.46	股純益	107.08	盤中	288.28	成交值	260.43
17	成長	185.79	險公司	106.87	加權	281.25	概念股	214.14
18	集團	184.46	市占率	105.27	外資	269.1	顯示器	208.33
19	鴻海	182.15	業界認	103.11	大阪	254.04	聯發科	197.2
20	金管	176.22	壽險業	98.223	日經	245.46	董事會	164.37

1. **冷啟動**：目前至少要一筆購物資料才能分析，同時無法分析它需要的品項。藉由拆字和HW1學到的**TF-IDF**等技術，可以尋找某特定群組和其他群組之間的差異，預先推薦某個客戶所在的群組，他們的群體需求。
2. **協同過濾**：此報告並未使用到協同過濾，未來可加入並和此方法加權。
3. **尋找更好的估算方式**，並且消除掉關聯性出在商品名稱的問題。