

某女鞋品牌
之
大數據與商業分析

目錄

1. 分析不同年齡之消費行為
2. 分析會員等級之消費行為
3. 廣告來源分析
4. 新冠肺炎對於訂單的影響
5. 關鍵字搜尋分析

1. 分析不同年齡之消費行為

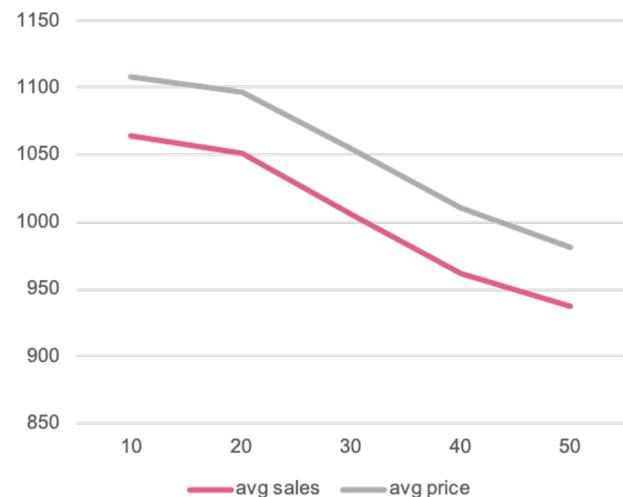
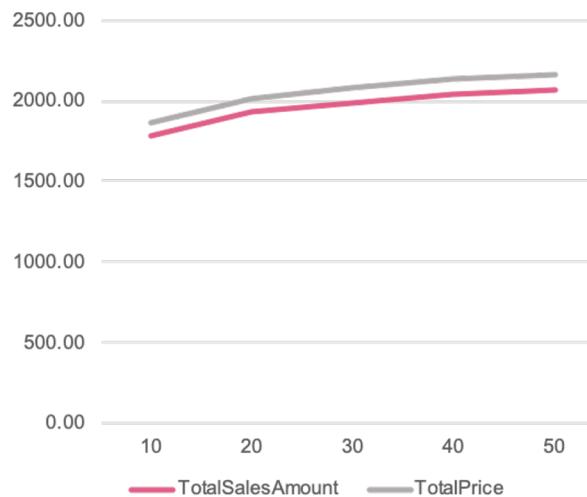
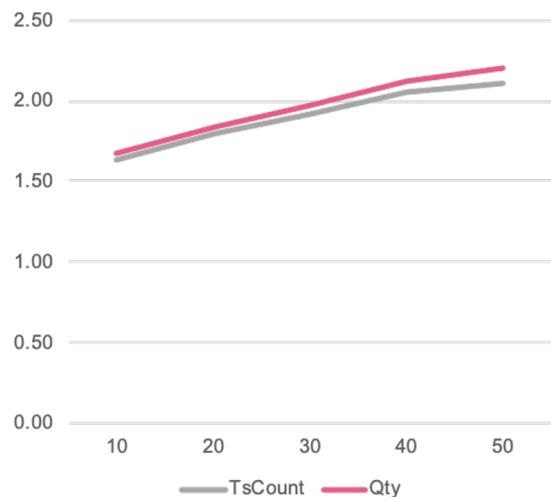
分析不同年齡層消費行為

1. 只留下Finish的訂單其他刪除
2. 將每筆訂單中的會員ID連結至MemberData去找尋相同的ID以取得會員生日
3. 利用訂單日期及生日去計算消費時的年齡
4. 將年齡過大或過小之不合裡的資料刪除
5. 以10歲為區間 將資料分為五類：10-20 \ 20-30 \ 30-40 \ 40-50 \ 50歲以上
6. 計算每一類之平均單筆消費總金額、平均單筆購買數量、購買商品之平均單價

發現不同年齡之消費行為不同

age	TsCount	Qty	TotalSalesAmount	TotalPrice	TotalDiscount	avg sales	avg price
10 - 20	1.63	1.68	1789.91	1863.81	-73.90	1064.38	1108.33
20 - 30	1.80	1.84	1930.56	2013.32	-82.75	1051.16	1096.22
30 - 40	1.92	1.97	1982.91	2080.10	-97.19	1005.61	1054.90
40 - 50	2.06	2.12	2036.26	2140.69	-104.42	961.04	1010.33
> 50	2.11	2.21	2073.42	2170.49	-97.07	937.75	981.65

不同年齡消費行為之圖表



建議針對不同年齡提出不同促銷方案

- 針對低年齡客群：提高總消費金額
Ex. 滿千送百
- 針對高年齡客群：提高單項之價格
Ex. 購買單價千元以上之商品可集點

2. 分析不同會員等級 之消費行為

分析不同會員等級之消費行為——資料前處理

STEP1：利用python中的pandas將會員資料中所需的資訊拿出來，分別為“MemberCardLevel”，“MemberID”，“Birthday”並將這些資料輸入成由會員等級、會員ID、年紀所組成的dataframe。

STEP2：將訂單資料中符合訂單狀態為“Finish”的訂單總金額、商品總數量、該筆的會員ID取出。

STEP3：將STEP2的資訊與STEP1對照，最後得出不同會員等級對應的總人數、平均消費總金額、平均商品價格。

	MemberCardLevel	...	age
0	10	...	36.0
1	10	...	30.0
3	10	...	43.0
4	10	...	37.0
5	10	...	36.0
...
337567	10	...	56.0
337568	10	...	24.0
337570	10	...	26.0
337573	10	...	39.0
337574	10	...	44.0

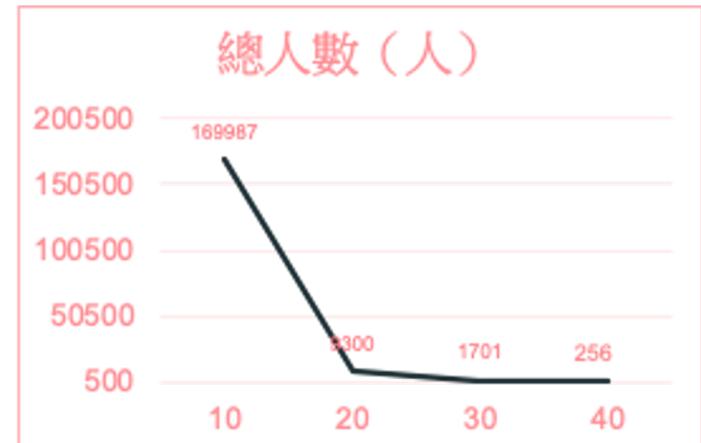
STEP1完成時的dataframe

結果

會員等級	10 (一般)	20 (中間)	30 (次高)	40 (最高)
總人數 (人)	169987	9300	1701	256
平均消費總金額 (NTD)	\$2866	\$12057	\$32971	\$7358
平均商品價格 (NTD)	\$1050	\$997	\$913	\$908
平均年齡 (歲)	32	33	33	28

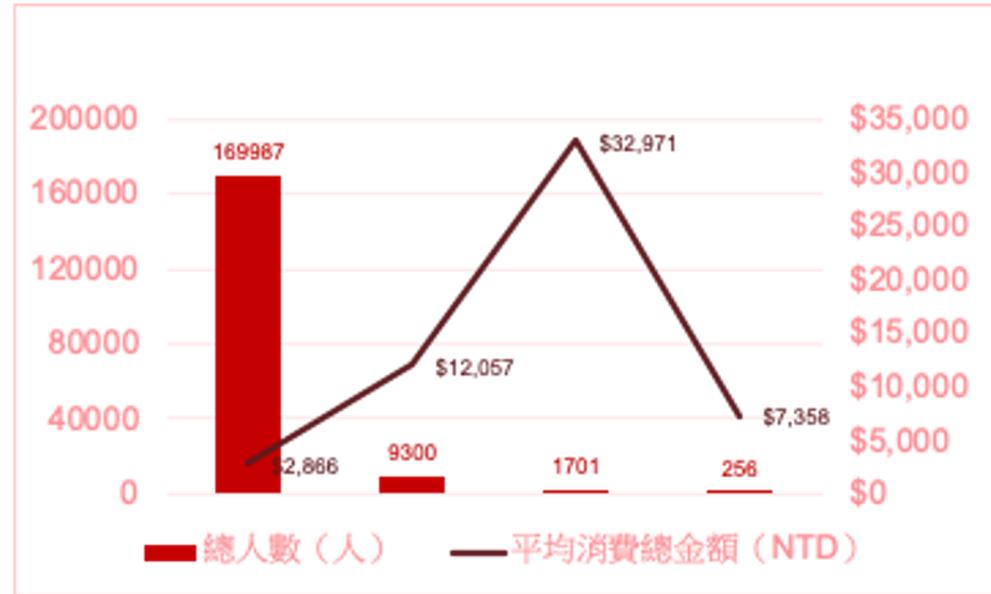
分析結果與建議 ①

- 由年齡可知各個會員等級的年齡大致為33歲，雖然最高會員等級的平均年齡較低，但是相對地，最高會員等級的人數明顯比其他會員等級少很多。因此，我們認為年齡與會員等級不具有交互作用。
- 利用會員卡等級分類所得出的平均商品價格，僅有一般會員等級有超過\$1000元，因此若將目標定為使每個會員卡等級的平均商品價格超過1000元，我們認為在廣告投放內容中，加強**金額較高**的商品投放與行銷或許能達成此目的。



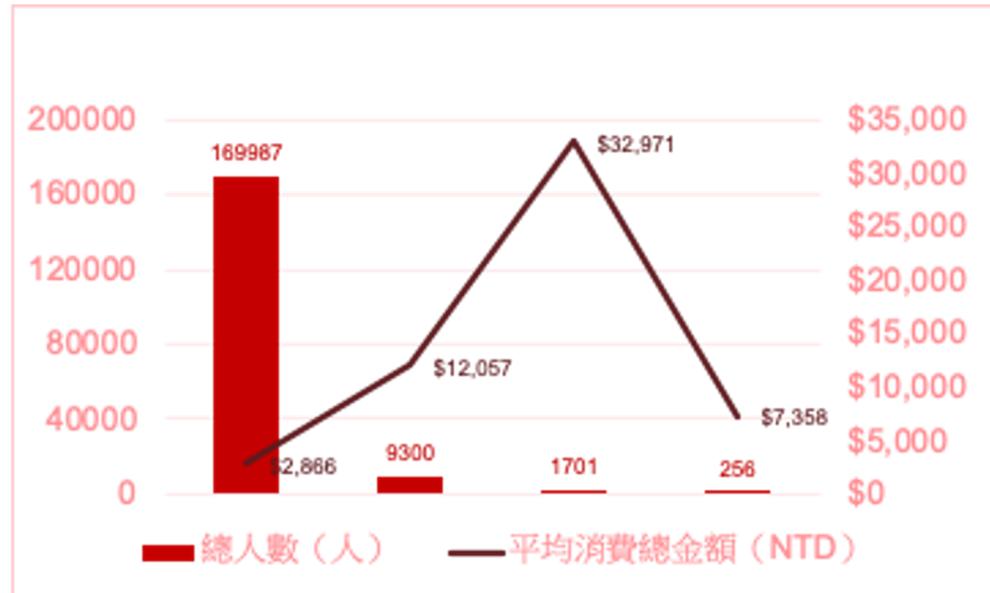
分析結果與建議 ②

- 雖然屬於一般會員人數最多，但平均消費總金額卻最少，因此為了使會員卡等級為10的人消費金額增加，除了投放更多廣告外，在廣告行銷內容上做更精心的著墨或許也能增加平均消費金額。



分析結果與建議 ③

- 由於消費者都是由一般等級的會員當起，因此我們認為，若放寬進入更高級會員級數的標準，使一般會員等級的會員移動至更高等級；以及在高級會員等級做更多的行銷，例如更加客製化廣告內容等等，使他們的平均消費總金額增加。



3. 廣告來源分析

1. 依據年齡
2. 依據會員等級

資料前處理

1. 合併2018-06至2020-02的trafficSource資料
2. 刪除非會員的行為資料
3. 刪除重複記錄的資料 (同一時間之行為會被記錄不只一次，造成重複欄位)
4. 利用trafficSource會員ID連結至MemberData，找尋相同的ID，取得會員生日
以及會員分級
5. 利用生日計算會員年齡
6. 將ht轉換成 datetime 的datatype

資料前處理

1	MemberID	ht	bh	cs	age	MemberCardLevel
2	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-08 13:06:40	traffic	google	38	10
3	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-09 19:40:00	traffic	FBAD	38	10
4	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-09 22:26:40	traffic	m.facebook	38	10
5	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-10 9:33:20	traffic	google	38	10
6	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-12 17:06:40	traffic	google	38	10
7	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-18 17:33:20	traffic	google	38	10
8	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-18 6:26:40	traffic	BRIDGE\	38	10
9	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-22 13:13:20	traffic	FBPO	38	10
10	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-22 18:46:40	traffic	FBLIVE	38	10
11	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-22 21:33:20	traffic	google	38	10
12	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-23 0:20:00	traffic	affiliate	38	10
13	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-23 8:40:00	traffic	BRIDGE\	38	10
14	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-04-29 9:06:40	traffic	BRIDGE\	38	10
15	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-05-03 21:26:40	traffic	google	38	10
16	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-05-04 19:40:00	traffic	google	38	10
17	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-05-05 15:06:40	traffic	google	38	10
18	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-06-03 19:06:40	traffic	FBPO	38	10
19	%2B%2B3XVjBiQPLoAXWc7gK%2BTdvTH	2019-06-06 19:20:00	traffic	FBPO	38	10
20	%2B%2Bc1c1PmW8BW4lP3bLm6163xTwILZ	2019-05-18 5:53:20	traffic	FBAD	33	10
21	%2B%2BNsfFBYC%2BjEi%2B6UABBcjFEsu	2019-06-04 9:00:00	traffic	google	30	10
22	%2B%2BNsfFBYC%2BjEi%2B6UABBcjFEsu	2019-06-04 17:20:00	traffic	google	30	10
23	%2B%2BvRWVI5IU2ccaBFhoZjhjOABLJdjxjF	2019-01-26 13:20:00	traffic	m.facebook	33	10
24	%2B%2BvRWVI5IU2ccaBFhoZjhjOABLJdjxjF	2019-01-28 12:33:20	traffic	BRIDGE\	33	10

計算方法

- 認列模式：First-In
- 會員分群：年齡、會員等級

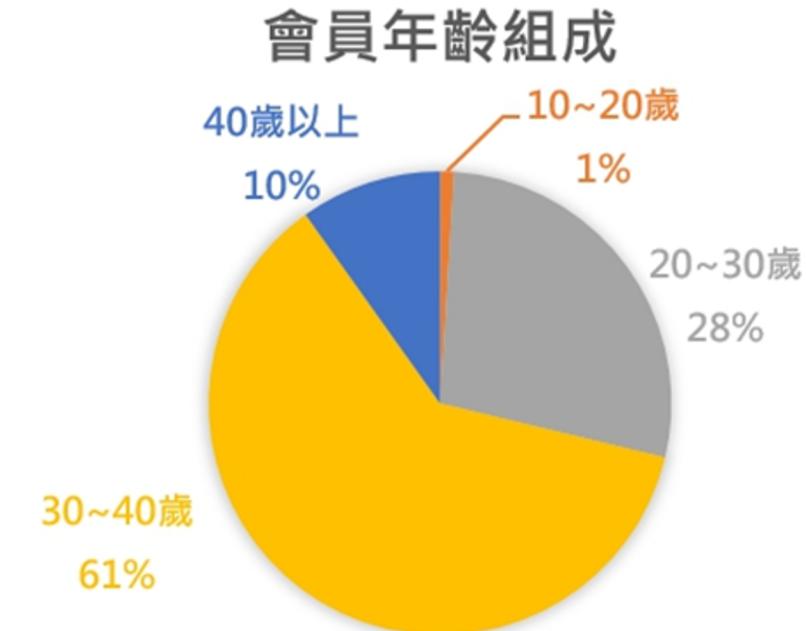
方法：

1. 將會員首次進入官網或app的流量來源加權 = 3
2. 其餘非首次進入之流量來源加權 = 1
3. 統計各來源之分數，並針對各個分群做後續分析及建議

3-1 依年齡將顧客分群

會員年齡組成

如右圖，會員年齡主要為30~40歲，其次是20~30歲，40歲以上的會員約占一成，而10~20歲的會員僅佔百分之一。



廣告來源排名(依加權後的分數)

年齡 排名	10~20歲	SCOR E	20~30歲	SCOR E	30~40歲	SCOR E	40歲以上	SCOR E
1	GOOGLE	313	GOOGLE	4922	GOOGLE	8875	GOOGLE	1350
2	LINE	78	FBPO	3005	FBPO	7062	FBAD	1345
3	BRIDGEWELL	70	FBAD	2437	FBAD	6729	FBPO	1167
4	System_inform	68	System_inform	2323	System_inform	4651	BRIDGEWELL	680
5	FBPO	61	LINE	1431	BRIDGEWELL	3811	System_inform	579
6	m.facebook.com	4	BRIDGEWELL	1232	LINE	3131	LINE	576

分析

- 由上表，我們發現每個年齡層的會員，廣告來源為GOOGLE皆佔最多數。
- 在SCORE較高的客群(20歲以上)中，FB PO文與FB AD的排名都佔據了來源中的第二、第三名。排名緊接在後的是System inform、Bridgewell及Line，這三種管道囊括了20歲以上顧客來源的第四到第六名，可見上述這些廣告的投放非常有效。
- 上表顯示，會員年齡越低，廣告來源為LINE之排名越前。可推知對年齡較低的會員投放LINE的廣告可能較有效益。

廣告來源分析(依加權後的分數)

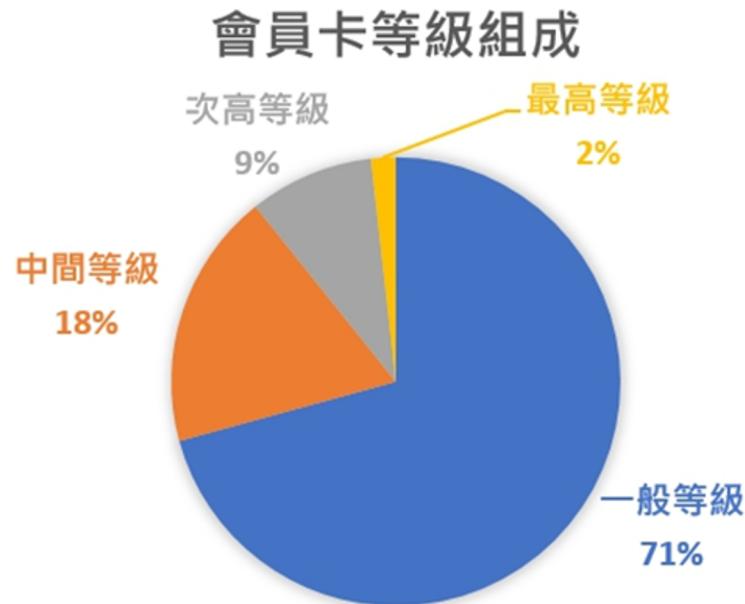
分類 \ 年齡	10~20歲	20~30歲	30~40歲	40歲以上
SCORE加總	716	20,622	45,551	7,332
廣告來源個數	15	57	69	35

- 30~40歲的會員廣告來源高達69種，比起其他年齡段的顧客更為多樣化，而10~20歲的顧客使用的管道較為固定，僅有15種。
- 由上表可知，廣告來源個數與SCORE加總呈正相關。

3-2 依會員卡等級將顧客分群

會員卡等級組成

如右圖，會員卡等級愈高的會員數愈少，其中高達七成的會員都是一般等級的會員，而最高等級的會員僅有2%。



廣告來源排名(依加權後的分數)

等級 排名	一般等級	SCOR E	中間等級	SCOR E	次高等級	SCOR E	最高等級	SCOR E
1	GOOGLE	11687	FBPO	2421	FBPO	1341	FBPO	313
2	FBAD	7753	GOOGLE	2333	GOOGLE	1139	l.facebook.com	233
3	FBPO	7182	FBAD	1931	FBAD	738	at-talk.line.me	160
4	System_inform	5777	System_inform	1336	LINE	541	GOOGLE	119
5	BRIDGEWELL	3993	BRIDGEWELL	1238	BRIDGEWELL	485	FBAD	96
6	LINE	3361	LINE	1213	System_inform	445	TAGTOO	78

分析

- 由上表可知，等級愈高的會員，廣告來源為FB PO文者愈多，可能表示FB PO文較能吸引消費力高、忠誠度高的顧客。而等級愈低的會員，廣告來源為GOOGLE者愈多。
- FB AD及System inform的效益在會員等級較低的客群中較有效，此店若要吸引新顧客，可以嘗試增加投放FB AD。

廣告來源分析(依加權後的分數)

分類 \ 等級	一般等級	中間等級	次高等級	最高等級
SCORE加總	52,213	13,665	6,621	1,310
廣告來源個數	64	41	45	37

- 一般等級的會員廣告來源最廣，共計64種，而最高等級的顧客之廣告來源較窄，共37種。

結論

- 若要發展較低年齡層(10~20歲)的顧客，可廣泛投放LINE及BRIDGEWELL的廣告以吸引年輕族群；而20歲以上的客群，以GOOGLE及FACEBOOK投放廣告應會有較佳的成效。
- 10~30歲的客群也可找網路紅人以業配、用YOUTUBE幫忙宣傳的方式，再統計顧客是被哪個網紅推薦來的，最後依據流量與消費金額多寡給薪。
- 從會員等級觀察，廣告來源主要為FB PO文者，會員等級普遍較高，可推知其忠誠度或消費力較高。品牌應持續積極經營FB PO文以鞏固忠實客戶。此外，品牌若希望能發展新客群，可增加FB AD的投放量。

4. 新冠肺炎對於訂單的影響

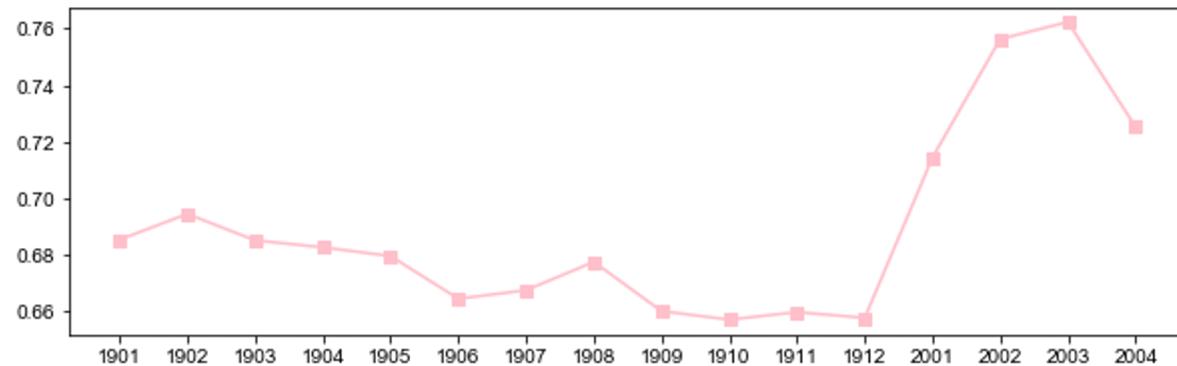
訂單數量

- 預期：消費者無法出門消費，造成網購次數增加
- 結果：訂單數量大幅下降



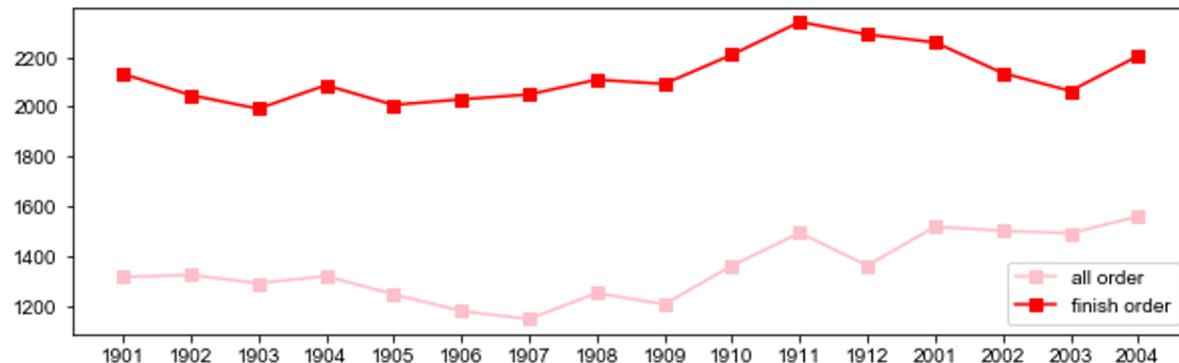
訂單完成率

- 預期：與平時應該不會相差太大
- 結果：訂單完成率明顯提高



訂單價格

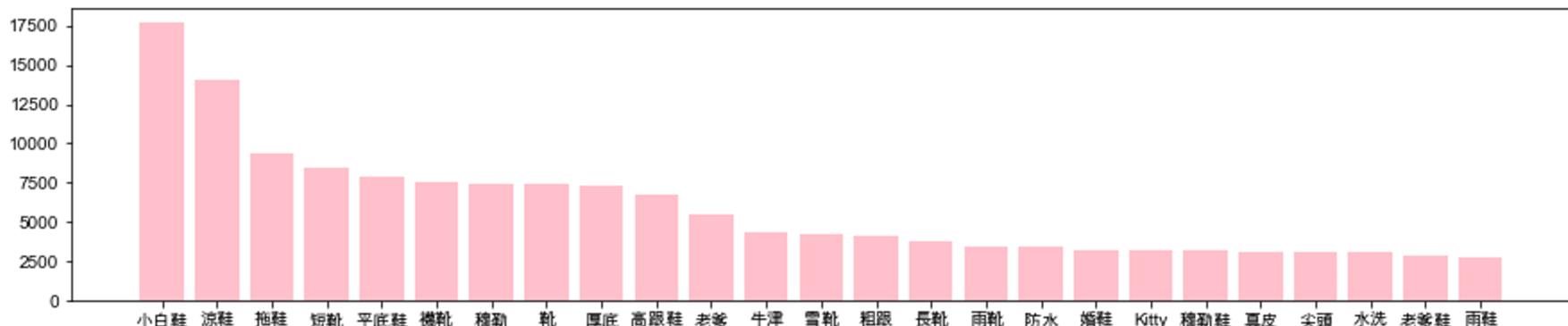
- 雖然訂單量下降、訂單完成率上升，每筆訂單價格在今年1-4月並無明顯變化



5. 關鍵字搜尋分析

搜尋關鍵字排行

- 小白鞋與涼鞋搜尋次數明顯高於其他 – 使用Google投放此關鍵字
- 搜尋方向：鞋子種類 > 形容詞 > 鞋子全名



遇到的問題

- Search行為資料中有 37% 的 User ID 為 NaN，或許存在搜尋次數非常多的潛在顧客，可以針對該客群行銷，但我們無法從資料中找出這些人是誰。

搜尋次數與訂單關聯分析

- 資料處理：找出搜尋次數>1000的會員ID，並分析訂單狀態
 搜尋VIP平均訂單價格：1220.575
 - 分析結果：**搜尋VIP訂單完成率：0.275**
 - 平均訂單價格與整體訂單相差不大，但訂單完成率<<平均完成率(0.6949)
 - 以uid = "jmPin5xB18Y61zGiw9xqSO3v5CW4eZqpq6oypBX%2F9BI%3D" 為例
 - 搜尋次數高達41837次
 - 訂單狀態幾乎全為Cancel

謝謝聆聽～

影片網址

<https://youtu.be/DXNxWuDSYM8>