

關聯規則分析流程

原始資料

Item.csv
同筆訂單可能
有多個商品

purchaseDate	purchaseTime	providerID	productID	productName	fridgeName	orderID	customerID	category	price
2020/9/10	18:18:44	SW004	30200013	楓糖蔓越莓蛋糕捲	內湖長虹	A105040	L_pEwdm	甜點	40
2020/9/10	18:18:44	MD048	10100054	擋不住熱情的蔥抱牛	內湖長虹	A105040	L_pEwdm	主餐	94
2020/9/10	18:18:44	MD048	10100054	擋不住熱情的蔥抱牛	內湖長虹	A105040	L_pEwdm	主餐	94
2020/9/10	18:18:44	SD029	20100011	肉包	內湖長虹	A105040	L_pEwdm	輕食	31
2020/9/10	18:18:44	SD029	20100011	肉包	內湖長虹	A105040	L_pEwdm	輕食	31
2020/9/10	18:18:44	SD029	20100011	肉包	內湖長虹	A105040	L_pEwdm	輕食	31

(1) 同筆訂單多個商品合併

orderID	prod_item
A105040	'擋不住熱情的蔥抱牛', '肉包', '楓糖蔓越莓蛋糕捲'

資料處理

(2) 轉為0/1矩陣

orderID	二吉軒 原味豆乳	肉包	招牌 杏仁茶	美?廢人 專門品 玻妞優格飲	楓糖蔓越莓 蛋糕捲	綜合水果盤	輕焙凍頂 烏龍	擋不住熱情 的蔥抱牛	藍莓奶酪
A105040	0	1	0	0	1	0	0	1	0
A527386	0	0	1	1	0	0	0	1	0
B230468	0	1	1	0	0	0	0	0	0
B298505	1	0	0	0	1	0	0	0	0
B454146	1	1	0	0	1	1	0	0	0
B988453	0	1	1	0	0	0	1	0	0
C111127	1	1	0	0	0	0	0	0	1

資料分析

- 分析方法：關聯規則(Apriori)
- 說明：(1) 用來找出經常一起出現的商品組合
(2) 計算支持度、信賴度、提升度等指標，評估商品之間的關聯強度
- 評估指標門檻：提升度 > 1，大於1表示關聯規則有效

關聯規則分析結果

註1：總訂單數22,833筆
註2：因個人電腦受限無法運算到支持度0.15%以下的規則

					支持度	信賴度	提升度
規則	A商品	B商品	買A比例 (%)	買B比例 (%)	同時買AB 的比例 (%)	買A再買B 的機率 (%)	提升度
1	佩拉琪巧克力蛋糕捲	蜂蜜鮮奶油蛋糕捲	1.35	1.16	0.17	12.62	10.92
2	不知春	輕焙凍頂烏龍	3.38	2.80	0.24	7.13	2.54
3	綜合水果盤	麻辣茶葉蛋二入	11.07	1.26	0.15	1.39	1.10
4	二吉軒原味豆乳	二吉軒微糖豆乳	6.86	5.13	0.35	5.17	1.01

指標	定義	說明
支持度	同時買AB商品的訂單佔總訂單的比例	支持度高→常出現的商品組合
信賴度	買A商品再買B商品的機率	信賴度高→買完A越可能買B
提升度	買B商品的比例因商品組合提高的程度，大於1表示關聯規則有效	提升度高→越適合做商品組合

分析結果表格說明：以規則1為例

- 買A比例：有買「佩拉琪巧克力蛋糕捲」的訂單佔總訂單1.35%
- 買B比例：有買「蜂蜜鮮奶油蛋糕捲」的訂單佔總訂單1.16%
- 支持度：同時買「佩拉琪巧克力蛋糕捲」和「蜂蜜鮮奶油蛋糕捲」的訂單佔總訂單0.17%
- 信賴度：買「佩拉琪巧克力蛋糕捲」再買「蜂蜜鮮奶油蛋糕捲」的機率為12.62%
- 提升度：買「蜂蜜鮮奶油蛋糕捲」的比例1.35%因商品組合提高為12.62%，提升度10.92，表示關聯規則很有效

結論與建議

註1：總訂單數22,833筆

註2：因個人電腦受限無法運算到支持度0.15%以下的規則

					支持度	信賴度	提升度
規則	A商品	B商品	買A比例 (%)	買B比例 (%)	同時買AB 的比例 (%)	買A再買B 的機率 (%)	提升度
1	佩拉琪巧克力蛋糕捲	蜂蜜鮮奶油蛋糕捲	1.35	1.16	0.17	12.62	10.92
2	不知春	輕焙凍頂烏龍	3.38	2.80	0.24	7.13	2.54
3	綜合水果盤	麻辣茶葉蛋二入	11.07	1.26	0.15	1.39	1.10
4	二吉軒原味豆乳	二吉軒微糖豆乳	6.86	5.13	0.35	5.17	1.01

結論

- 在**(1) 提升度 > 1**、**(2) 支持度 > 0.15%** (註2) 的條件下，僅能找出4項關聯規則，若能排除(2)的限制，應能找出更多規則
- 4項關聯規則的支持度相差不大，可考慮都作為商品組合
- 若只需一組最佳商品組合，可選擇提升度最高的「規則1」

建議

- 由規則1、2、4看出，消費者可能喜歡購買不同口味的商品(巧克力vs.鮮奶油 蛋糕捲、不知春vs.烏龍茶飲、原味vs.微糖 豆乳)，若要找新的商品組合可考慮**搭配不同口味的商品組合**
- 由規則3看出商品組合偏向健康餐，且有買「綜合水果盤」的訂單佔總訂單11.07%(2,527/22,833)，隨著國人健康觀念提升，若要找新的商品組合也可考慮**健康餐商品組合**