

零售業顧客分群分析

卓宛俞

目錄

01 分析目的

02 資料前處理、特徵工程

03 建立模型、模型結果

04 結論與建議

分析背景與目的

背景

面對新零售的浪潮，在消費者的喜好隨市場環境迅速變動的趨勢下，無論企業欲採取何種營運策略。應先釐清現有顧客的消費習慣與消費者輪廓，並以此為基礎發展相應的行銷活動、制定營運策略，藉此提升企業營收、顧客黏著度

目的

- 1.藉由過去的消費紀錄，了解顧客的屬性、產品偏好、消費行為
- 2.通過顧客分群來進行更精準化的運營，根據不同的顧客群體執行不同經營策略、給予相應的行銷方案，以提升企業資源運用效率、營收

檢查缺失值

僅有Income有缺失值，將缺失值刪除

```
RangeIndex: 2240 entries, 0 to 2239
Data columns (total 29 columns):
#   Column                Non-Null Count  Dtype
---  -
0   ID                     2240 non-null   int64
1   Year_Birth             2240 non-null   int64
2   Education              2240 non-null   object
3   Marital Status        2240 non-null   object
4   Income                 2216 non-null   float64
5   Kidhome                2240 non-null   int64
6   Teenhome              2240 non-null   int64
7   Dt_Customer            2240 non-null   object
8   Recency                2240 non-null   int64
9   MntWines               2240 non-null   int64
10  MntFruits              2240 non-null   int64
11  MntMeatProducts        2240 non-null   int64
12  MntFishProducts        2240 non-null   int64
13  MntSweetProducts       2240 non-null   int64
14  MntGoldProds           2240 non-null   int64
15  NumDealsPurchases      2240 non-null   int64
16  NumWebPurchases        2240 non-null   int64
17  NumCatalogPurchases   2240 non-null   int64
18  NumStorePurchases      2240 non-null   int64
19  NumWebVisitsMonth      2240 non-null   int64
20  AcceptedCmp3           2240 non-null   int64
21  AcceptedCmp4           2240 non-null   int64
22  AcceptedCmp5           2240 non-null   int64
23  AcceptedCmp1           2240 non-null   int64
24  AcceptedCmp2           2240 non-null   int64
25  Complain               2240 non-null   int64
26  Z_CostContact           2240 non-null   int64
27  Z_Revenue              2240 non-null   int64
28  Response               2240 non-null   int64
dtypes: float64(1), int64(25), object(3)
memory usage: 507.6+ KB
```

特徵工程

創建新欄位

年齡

家庭人口數

總消費支出

會員年資

有無小孩

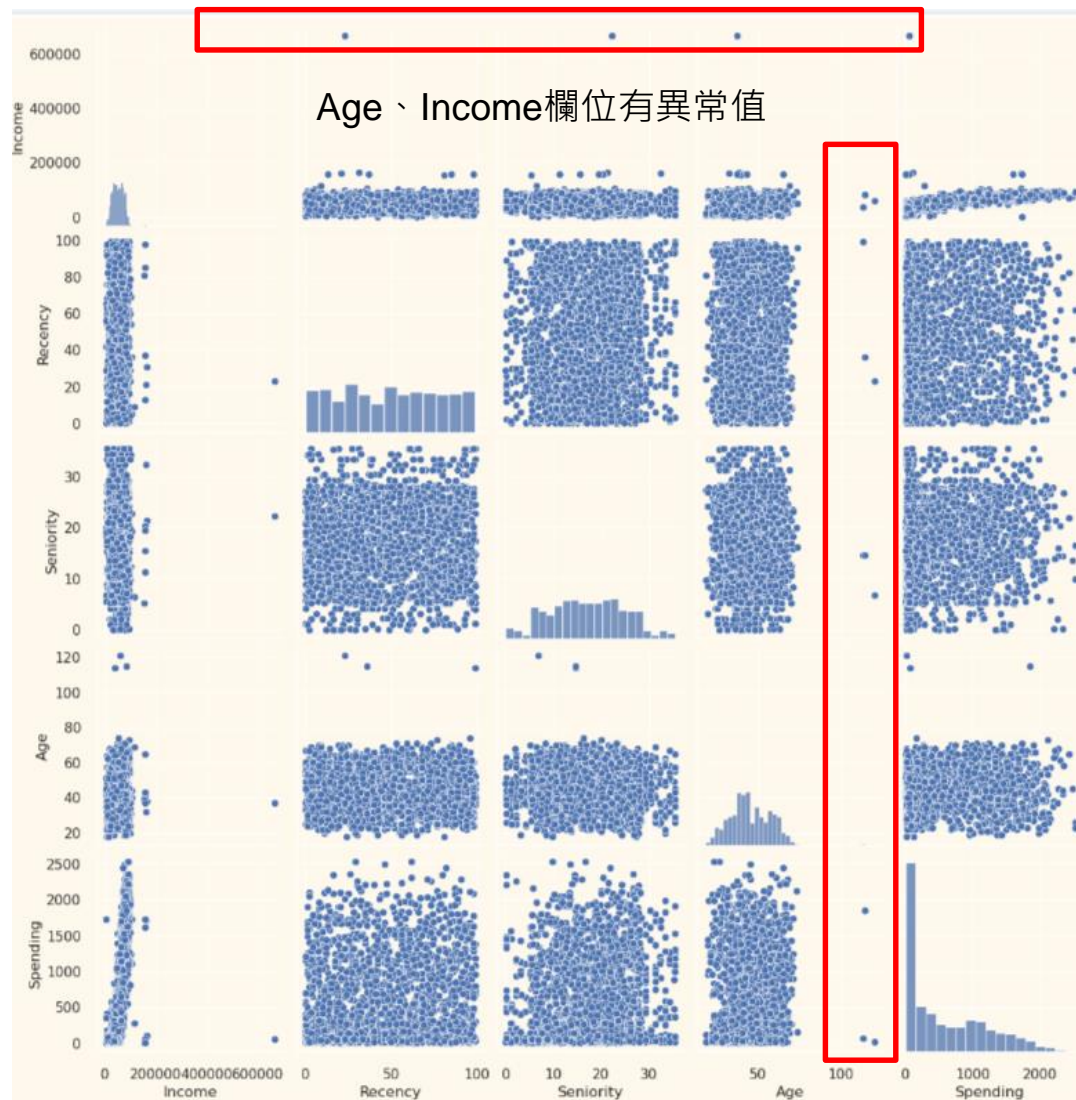
接受宣傳次數

資料欄位

顧客屬性	購買金額
Education	Wines
Age	Fruits
Marital_Status	Meat
Children	Fish
Family Size	Sweets
Has Child	Gold
Seniority	消費渠道
Campaign	Web
Income	Catalog
Spending	Store

折扣購買數量
Deals

了解資料相關性、偵測異常值

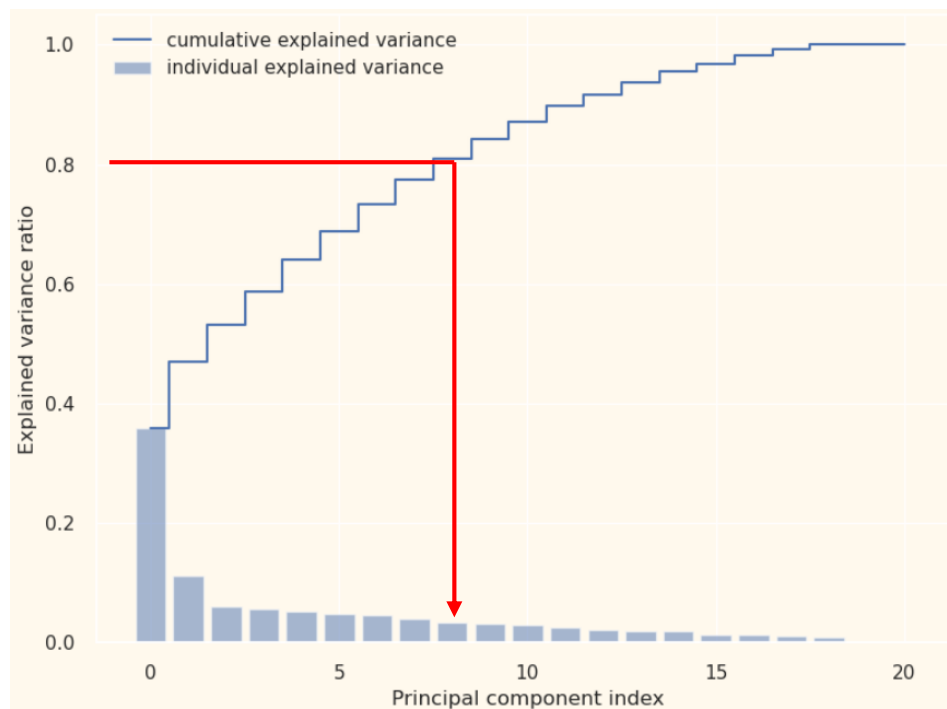


資料前處理

- Lable Encoding
- 特徵縮放 Standardization
- 主成分分析

1.主成分依序保留1~20特徵
之疊加變異解釋率

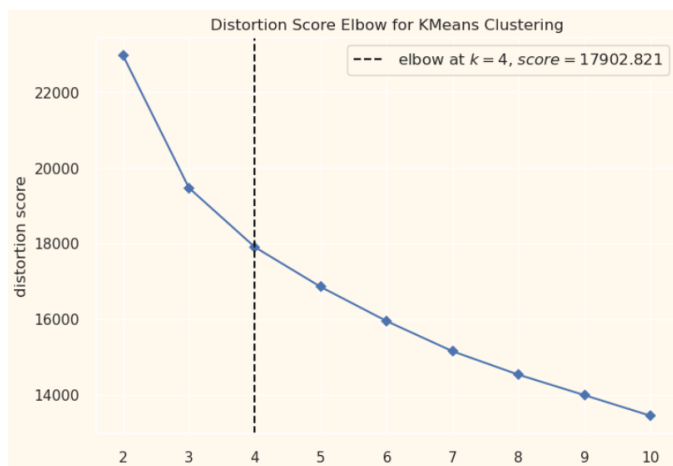
2.主成分保留8維建立模型



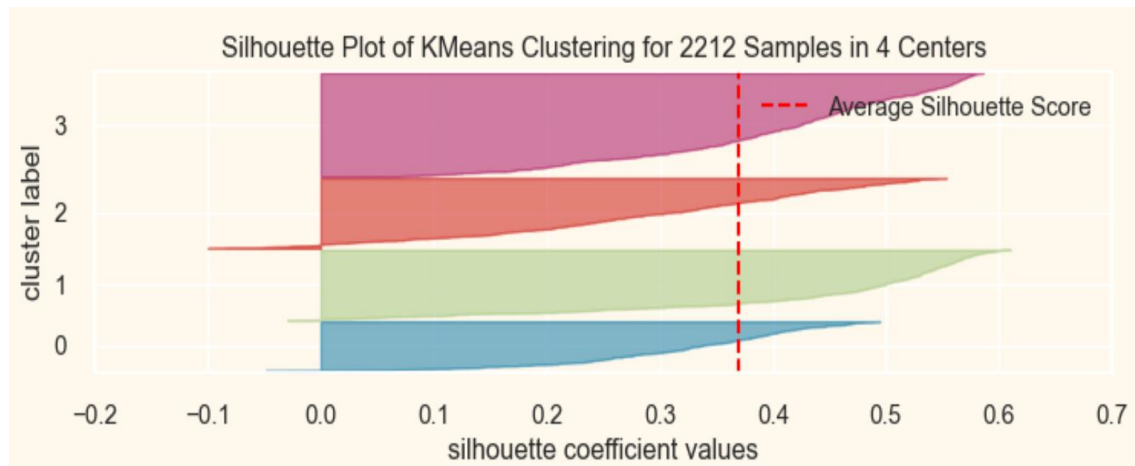
建置模型_K-means

- 使用手肘圖和輪廓分析設定群數

手肘圖



輪廓係數

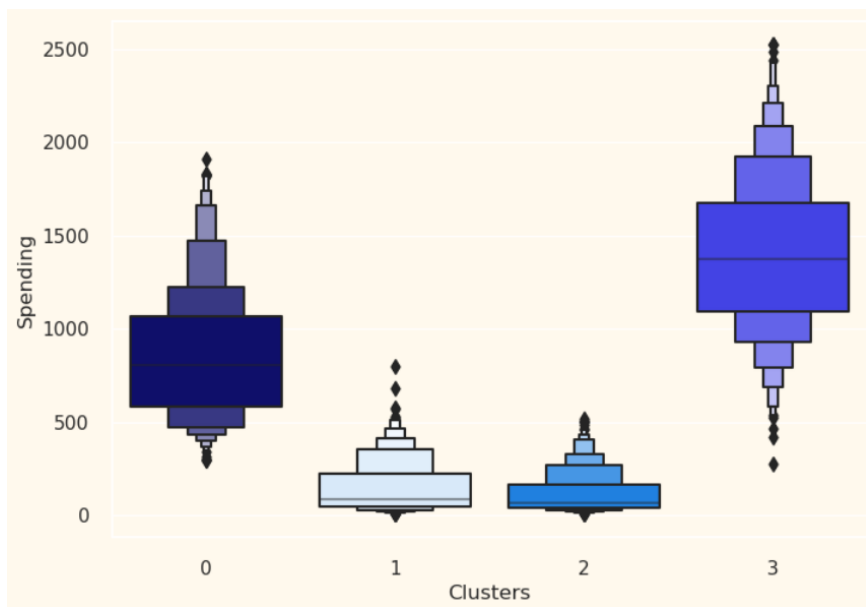


從輪廓係數圖可以看出 4 群波動性一致、均超過平均輪廓係數
搭配手肘圖，故選擇設定群數為 4

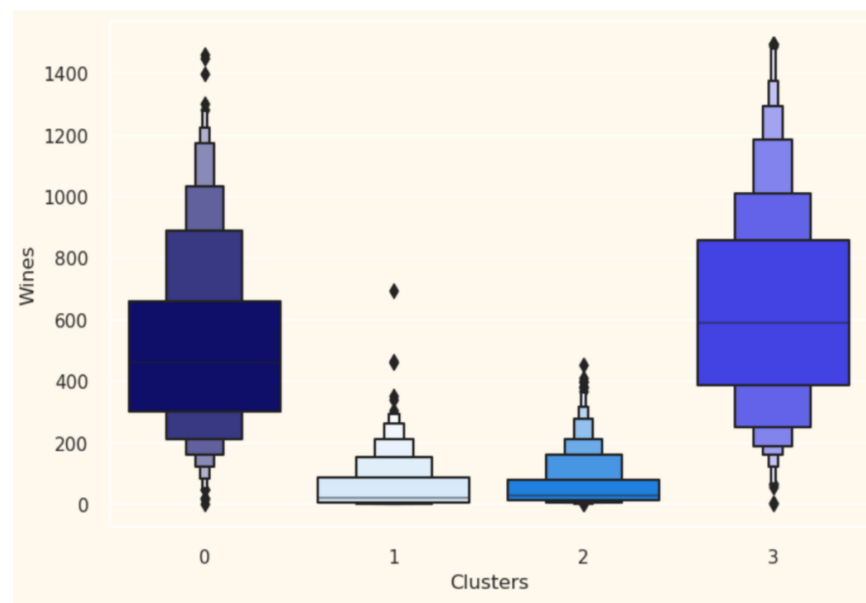
分析模型結果

- 分析順序為單變數(連續變數)、單變數(類別變數)、雙變數

總消費金額 群 3>0>1>2



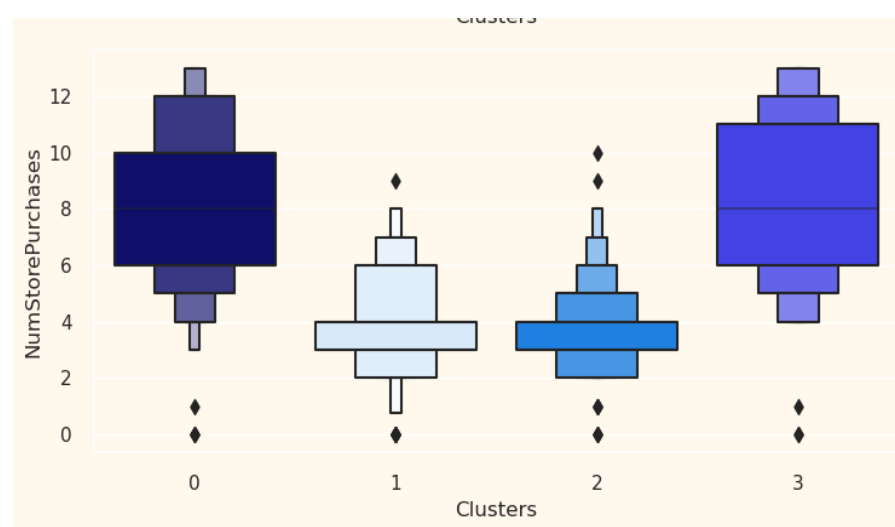
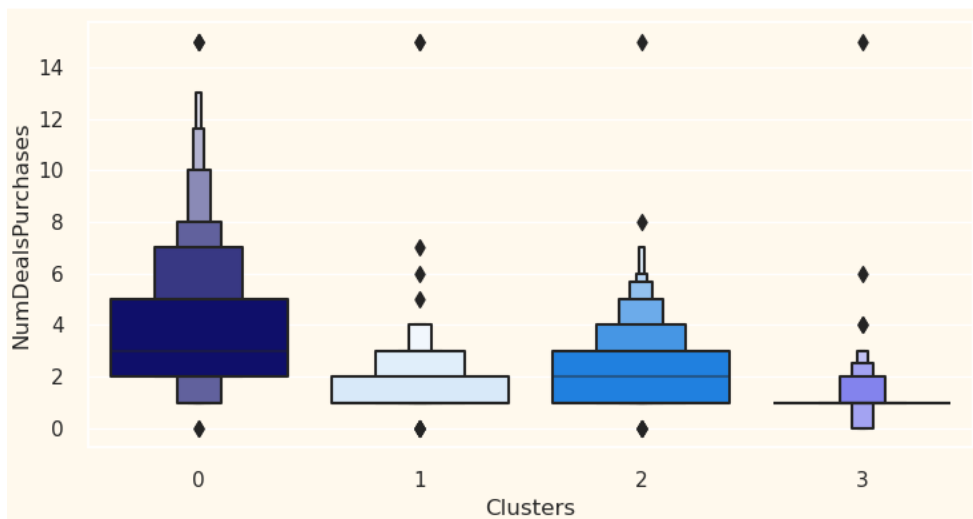
酒(黃金) 消費金額 群 3>0>1=2



分析模型結果

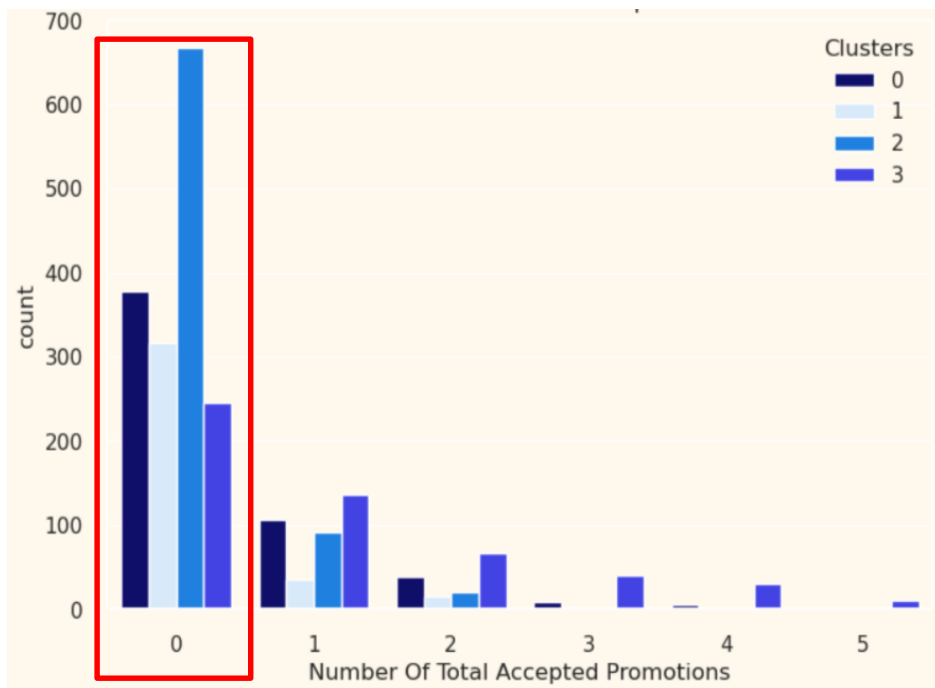
第0群的顧客喜愛打折商品，第3群則相反

大部分的顧客依然偏好實體購物而非網路

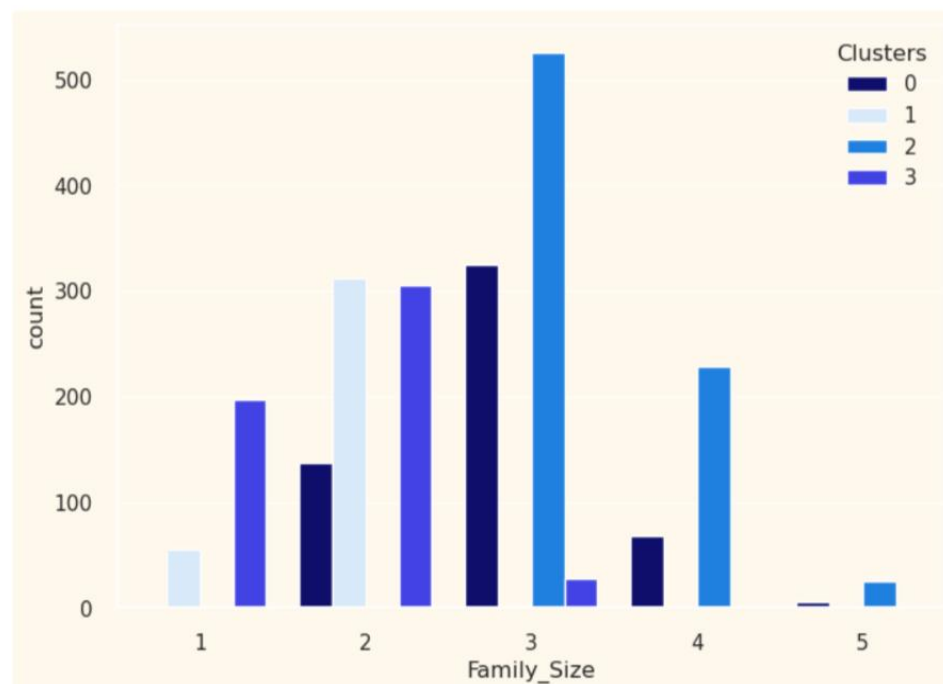


分析模型結果

過去行銷活動成效有限

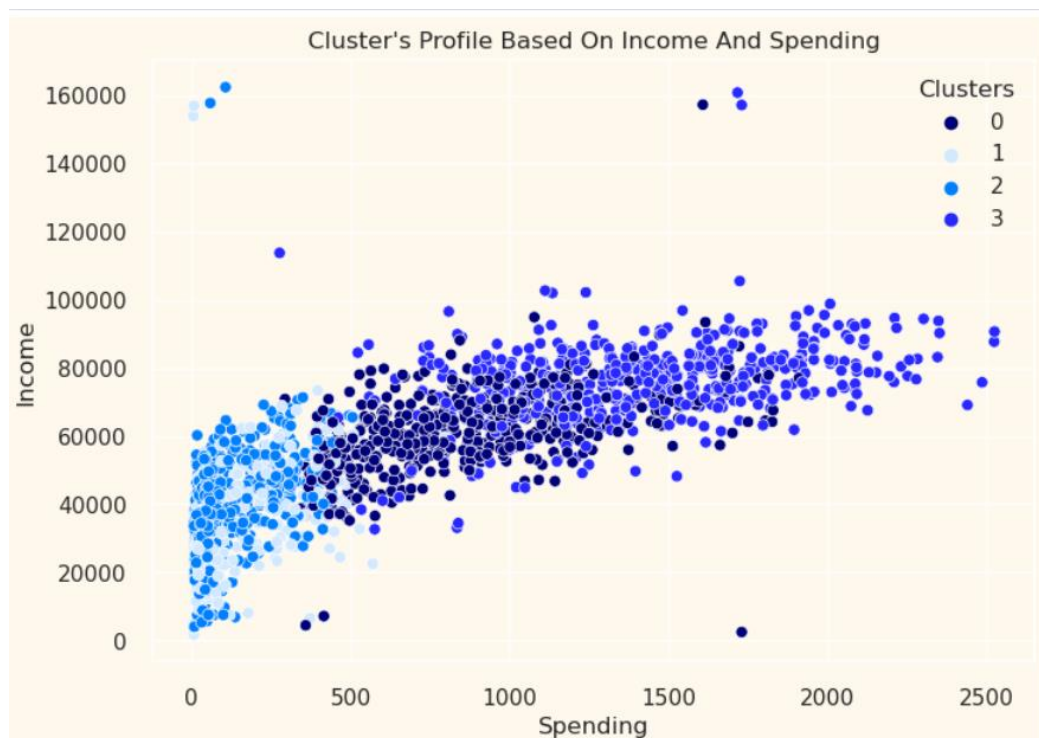


觀察婚姻、家庭數，了解第2群為家庭消費族群



雙變量_連續

消費金額與所得之關係呈現正相關，隨著所得提升，消費支出提高



第1、2群:低所得、低消費

第0群:介於中間

第3群:高所得、高消費

結論與建議

各群屬性歸納

第0群:喜愛奢侈品，在商品打折時會增加消費量，所得和支出中間偏高

第1群:顧客年齡較年輕、家庭人口數少(單身、伴侶)

第2群:家庭客，佔會員人數4成

第3群:高所得、高消費，所有商品皆購買

共同屬性:偏好實體購物、過去行銷活動效果不彰

建議

- 1.對第一次的行銷活動進行檢討，找出僅接受一次的原因
- 2.選擇價格彈性高的商品，對第0群做促銷活動
- 3.第3群為高價值客戶，需定期追蹤，避免顧客流失