

機器學習觀念與應用 作業二
112703003 資訊二 黃柏淵

第一題：

- Recency = 當前日期 - 客戶最近一次購買的日期
= 客戶最近一次購買的時間

```
df_trans["trans_days"] = [(date(int(d.split("/")[2]),int(d.split("/")[0]),int(d.split("/")[1]))-date(1998,12,31)).days for d in df_trans["transaction_date"]]
```

為 df_trans DataFrame 新增一個欄位 "trans_days"，計算每筆交易的日期與固定參考日期 1998 年 12 月 31 日 之間的天數差異。

- Frequency = 客戶在特定時間範圍內的購買次數。

```
f_value = df_trans.groupby(["customer_id"])["transaction_date"].nunique()
```

把交易數據按客戶 ID 分組，然後使用 nunique() 計算每個客戶的獨立交易日期數量。

- Monetary = 客戶在特定時間範圍內的總消費金額。

```
# look up the price of the product by its product_id
def lookup_price(product_id, df=df_prod):
    return df_prod[df_prod["product_id"]==product_id]["product_retail_price"].values[0]
0.0s

df_trans["price"] = [lookup_price(pid) for pid in df_trans["product_id"]]

df_trans["amount"] = df_trans["quantity"]*df_trans["price"]
m_value = df_trans.groupby(["customer_id"])['amount'].apply(sum)
```

先為每筆交易添加產品價格，再計算每筆交易的金額 (quantity * price)，按客戶分組，計算每位客戶的總消費金額 (M)。

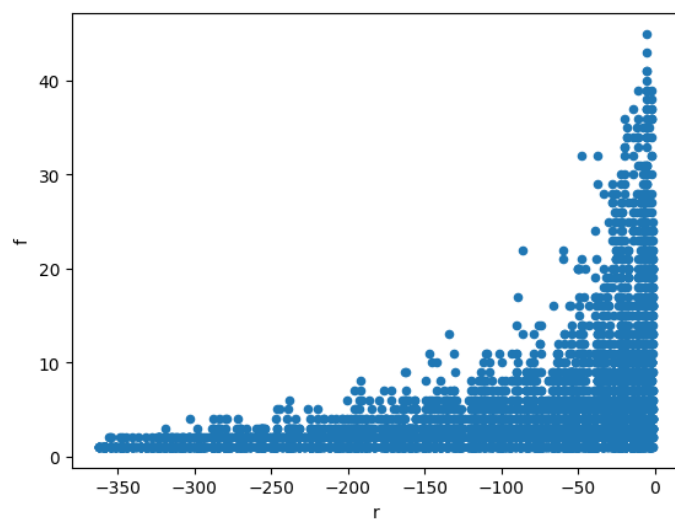
第二題：請針對你所設計的 RFM 計算方式，根據顧客在 1 ~ 10 月的購物行為資料，進行 RFM 值轉換

- 根據所有顧客的 RFM 值分佈，提供各別的最小值和最大值

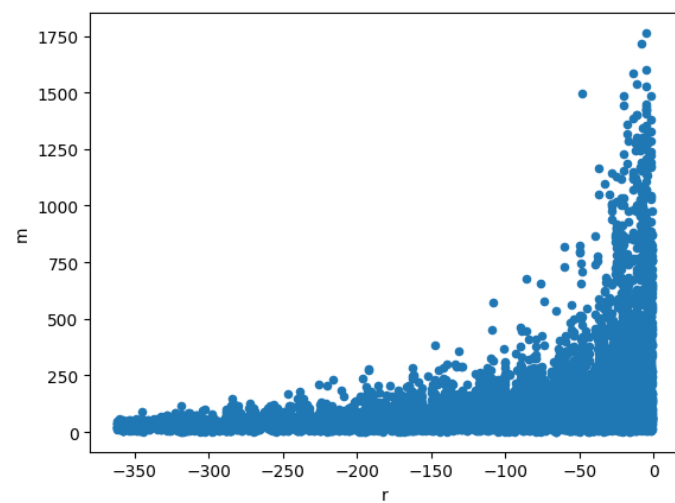
	Metric	Min	Max
0	Recency (R)	-362.00	-1.00
1	Frequency (F)	1.00	45.00
2	Monetary (M)	1.06	1763.58

ii. 分別畫出 RF、RM、FM 二維散布圖

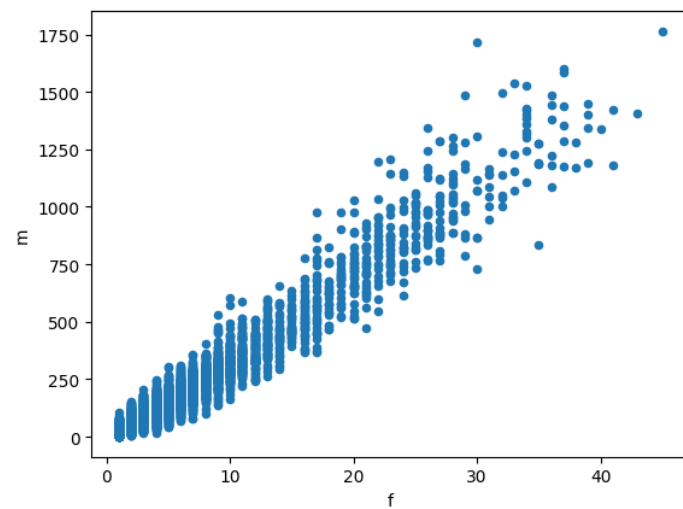
(1)RF



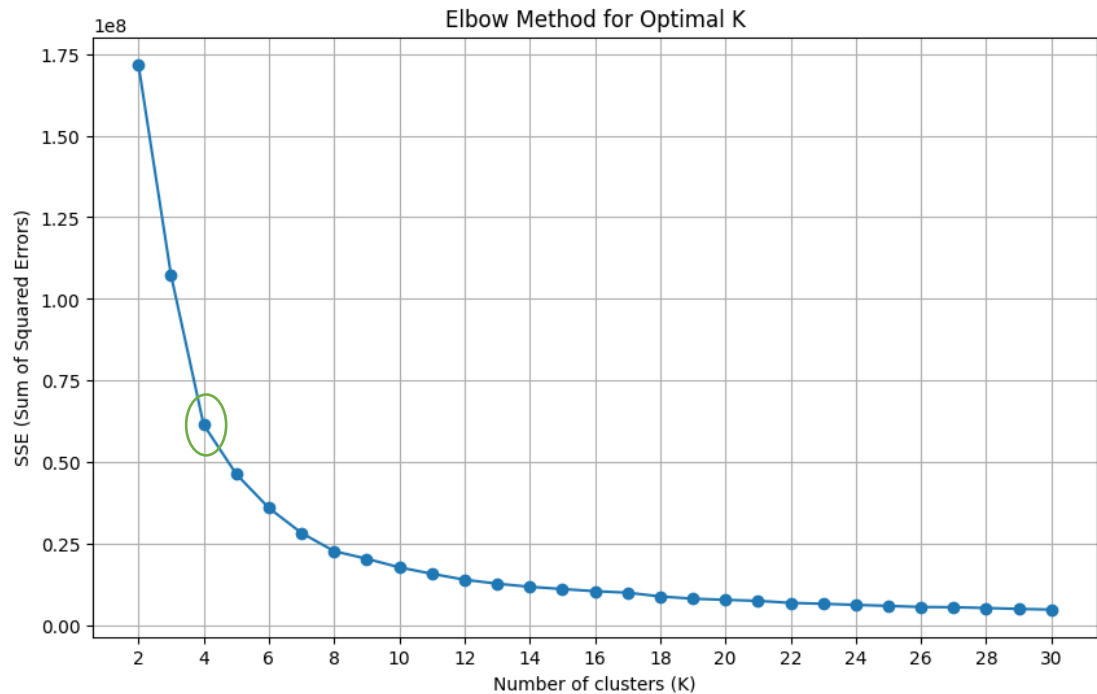
(2)RM



(3)FM



第三題：承第二題，根據轉換後資料，每一位顧客包含三個屬性資料(即 R 值, F 值, M 值)，請根據此資料，透過 K-means 分群演算法，將顧客分群，並觀察不同參數 K 對分群結果 SSE 的影響，請設置 K = 2 ~ 30，利用 Elbow 方法找出 K 應該設置多少？



K 應該設置為 4。

第四題：承第三題，使用 Elbow 方法選定一個 K 值，再執行 K-means 分群演算法，請列出每一群的中心點，例如中心點 C1: R=70、F=25、M=800。K 個中心點的列表請根據中心點的 M 數值降冪排序。

	F	M	R
1	25.288820	940.794907	-12.770186
3	10.424670	344.674714	-28.936564
0	3.211346	91.545168	-53.590434
2	1.601518	44.435968	-230.483397

第五題：承第四題，請你從分群結果挑選出其中一群顧客，為他們設計專屬的行銷活動，以增加客戶留存或增加營收。

i. 此群中心點的 R、F、M 值分別是多少？

我選群組 1:

	F	M	R
1	25.288820	940.794907	-12.770186

ii. 此群的顧客總人數？

```
Group 1 count: 322
```

iii. 請針對這群顧客的 RFM 值分佈，分別說明最小值和最大值的範圍

```
R range: [-86, -1]
F range: [14, 45]
M range: [643.8000000000001, 1763.58]
```

iv. 請說明你的具體行銷活動內容與目的

群組 1 特性分析：

- Recency (-86 到 -1)：這群客戶最近都有購買行為，最遠不超過 86 天，平均約 12.77 天前，顯示他們是活躍用戶。
- Frequency (14 到 45)：購買頻率高，平均 25.29 次，最低也有 14 次，顯示他們是忠實客戶。
- Monetary (643.80 到 1763.58)：消費金額範圍廣，平均 940.79，最低 643.80，最高可達 1763.58，顯示他們願意花費，且部分客戶有更高消費潛力。
- 產品偏好

```
=== Cluster 1 產品偏好 ===
product_name
High Top New Potatos      48
Nationeel Fudge Brownies  47
Super Chunky Peanut Butter 46
Better Noodle Soup        45
Moms Roasted Chicken      45
..
Plato Sesame Oil          15
Colony White Bread         14
Washington Diet Soda      14
Washington Berry Juice    12
CDR Apple Preserves       9
Name: count, Length: 1559, dtype: int64
```

此 322 位客戶是高價值、活躍且忠誠的群體，偏好多樣化的日常飲食商品，適合透過產品組合促銷以及節慶活動來維持他們的消費行為並進一步提升營收。

A. 產品組合促銷、季節性與節慶行銷

- 家庭採購優惠包：
 1. High Top New Potatoes (馬鈴薯) + Moms Roasted Chicken (烤雞) + Better Noodle Soup (湯麵)
 2. Super Chunky Peanut Butter (花生醬) + Colony White Bread (白麵包) + CDR Apple Preserves (蘋果果醬)

3. 提供組合折扣，例如「買 3 件享 85 折」，鼓勵顧客一次購買多樣商品，增加客單價。

- 節慶採購促銷：

1. 主打 Moms Roasted Chicken、High Top New Potatoes 和 Nazioneel Fudge Brownies，推出「節慶家庭採購包」，提供折扣或贈品（例如免費小份 Washington Berry Juice）。

- 健康飲食推廣：

1. Washington Diet Soda 和 Washington Berry Juice 顯示顧客可能注重健康選擇。推出「健康飲食週」，主打低脂、低糖商品（如 Washington Diet Soda 和 Better Noodle Soup 的低鈉版本），並提供試吃或小份試用裝。

B. 累積消費滿額贈禮

- 活動內容：

1. 「消費滿 1000 元送 100 元優惠券」，優惠券可在下次消費時使用，無最低使用限制，有效期 60 天。
2. 累積期間：1 個月

- 預期效果：

1. 提升金額：平均消費 940.79 的客戶只需增加約 60 元即可達標，誘因高。
2. 增加留存：優惠券鼓勵下次消費，延長客戶生命週期。
3. 針對高消費客戶：M > 1000 的客戶（最高 1763.58）可輕鬆達標，進一步鞏固消費者忠誠度。