

Programming Assignment 1

Deadline: 11/13

此次作業主要目的在讓同學學習運用 Python 由 Foodmart Database 超級市場的顧客及其交易資料中，探勘 Frequent Itemsets、Association Rules、Quantitative Association Rules。

- 作業給定的 Foodmart Data 已經附在 Moodle 平台上。
 - 作業每人繳交一份報告，檔案類型以 pdf 為限。
上傳檔名格式為學號_P1，例如：112753XXX_P1.pdf
 - 此次作業可以使用現有套件執行運算。
1. 從 Foodmart Data 的交易資料中，探勘符合 Minimum Support = 0.00015 且 Minimum Confidence = 0.8 的 Association Rules，並列出 Confidence 最高的前 10 條 Rules 以及 lift 最高的前 10 條，並比較這兩者的異同。
 2. 有時候我們有興趣的資料不只有產品間的資訊，也會想要由 User Profile 探勘顧客的基本資料。在給定 Minimum Support = 0.05 且 Minimum Confidence = 0.9 的條件下，探勘 Foodmart 顧客基本資料的屬性 {customer_state_province, yearly_income, gender, total_children, num_children_at_home, education, occupation, homeowner} 間的 Association Rules，並列出10條。
 3. 請探勘 Foodmart Data 中，顧客背景資料與其交易資料之間的關係 (Quantitative Association Rules)。例如 80% 女性顧客常買保養品。請自行設定 Minimum Support、Minimum Confidence，找出 10 條你覺得有意義的 Rules。請說明你的作法及相關參數設定。
 4. 在美國由於聖誕節，12月是購物的旺季。請探勘分析比較 12 月與 1 ~ 11 月的顧客購物行為。有哪些相似的地方，有哪些差異的地方？