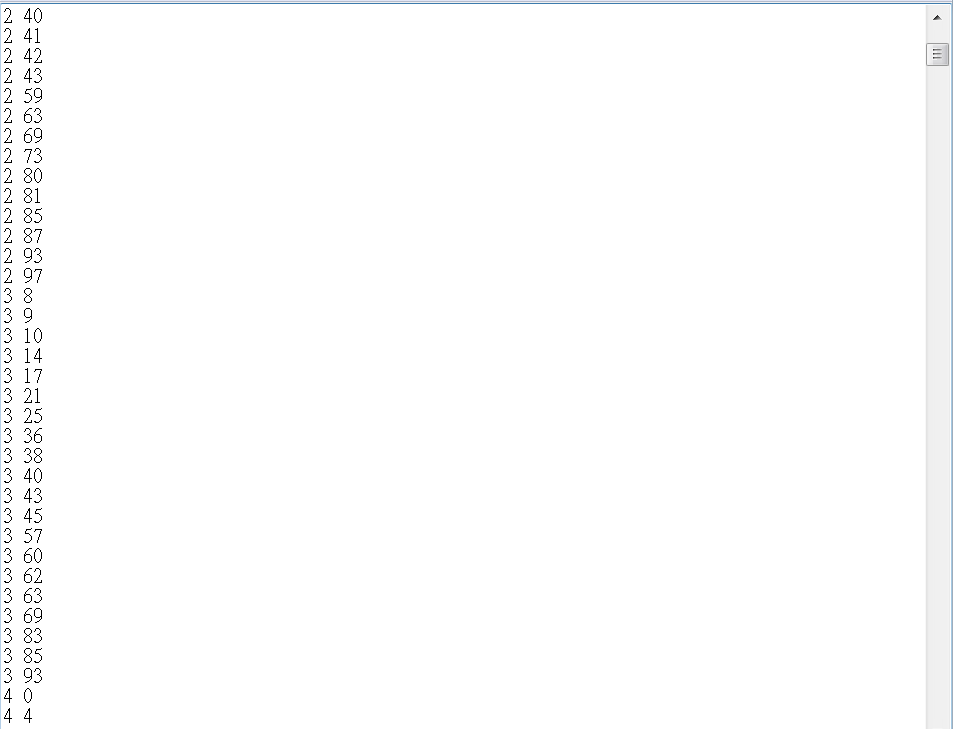
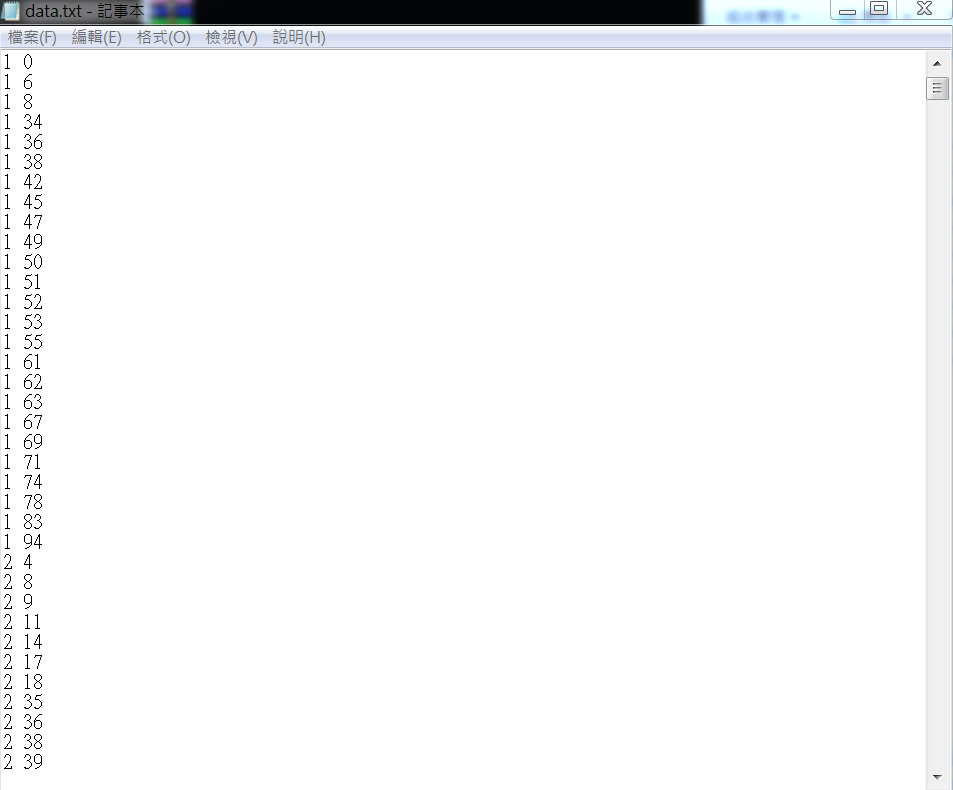
**Association Analysis**

**E94056178**

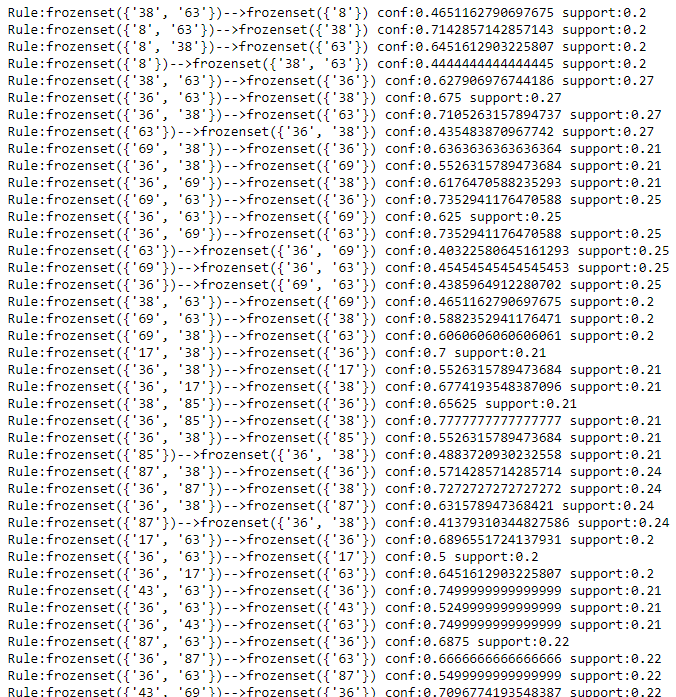
Dataset: data.txt (Use IBM Quest Synthetic Data Generator)

Apriori code : ” Apriori.py” or “Apriori.ipynb” (Python3執行產生結果)

FP-growth code : “FP-growth.py” or “FP-growth.ipynb” (Python3執行產生結果)

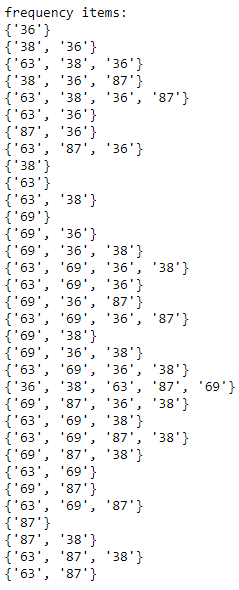
****Dataset(data.txt)

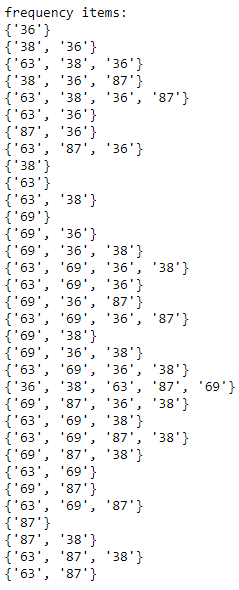
…………未列完

****Apriori result (min support = 0.2 confidence = 0.4)

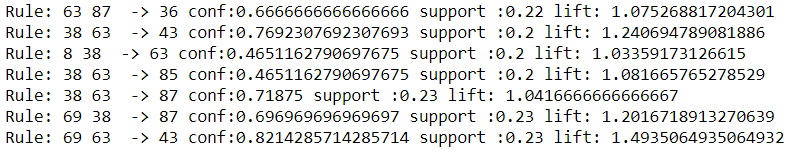
…………未列完

FP-growth result

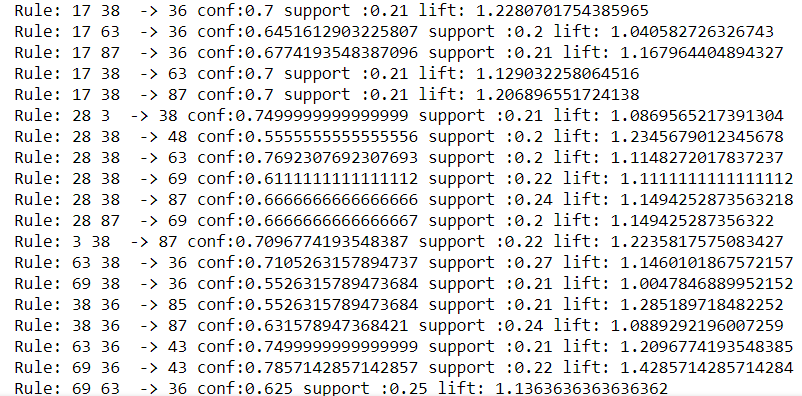
****

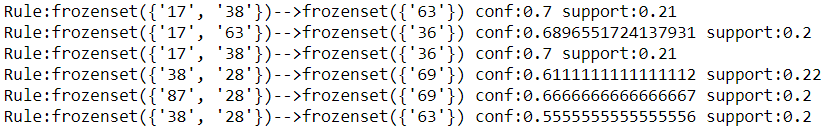
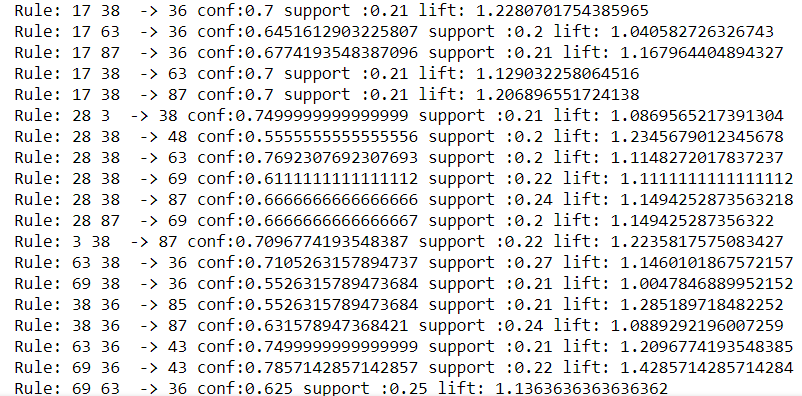
****

…………未列完

****¨ Use association analysis tools (沒使用WEKA 用python的apyori套件比較)

…………未列完

相關code: “Apriori\_tool.py” or “Apriori\_tool.ipynb”

****Compare results

「皆節錄部分結果」同rule的confidence 與 support一致

Apriori

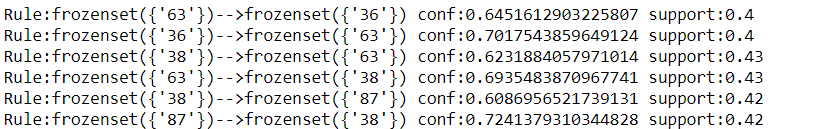
implementationn

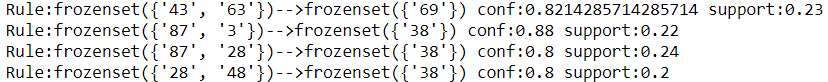
association analysis tool未列完

Find and Answer

對data.txt做APriori關聯分析(自己實作),不同support與confidence結果如下:

C:\Users\USER\Desktop\未命名100.pngRules with High support, high confidence

Rules with High support(0.4), Low confidence(0.2)

Rules with Low support(0.2), high confidence(0.8)

Rules with Low support(0.2), Low confidence(0.2)

…………共162列

此資料均為純數字，當作字串處理，每個元素對整體的support沒有太大，結果的support代表該rule整體出現在data的頻率，confidence代表該rule前面關聯後面的可能性多大。以市場交易為例：高support高confidence可能代表事物性質相近，一起購買的機率高，交易比重也高；高support低confidence交易比重高，但是一起購買機率不高；低support高confidence雖然事物一起買機率高，但是在市場交易比重占太少；低support低confidence則代表兩事物關聯性低。