1. **對 CreditCardPromotion 進行 Association Rule , 並使用 Apriori 演算法, 設定 confidence = 0.9、minimum support = 0.2 , 並回答以下問題：**

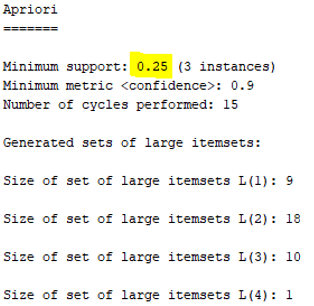
**使用 Weka 軟體：**

**(a) 請嘗試著修改 CreditCardPromotion.arff 的欄位與上圖相同,使其可以執 行 Association Rule, 請說明使用的方法以及解釋原來的檔案不能執行的原因? (10%)**

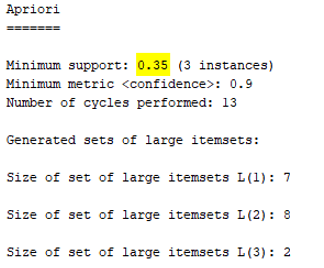
因為Sex屬性的型態應該是Nominal，而不是Numeric，而原本屬於Sex屬性的資料皆為0跟1，屬於Numeric而不是Nominal，所以無法執行。

**(b) 請將 numRule 設成 5 和 10,其各別執行後的 Minimum support 為何,請比較兩 者並說明造成其差異的原因。(15%)**

當numRule 設成 10，Minimum Support 為0.25

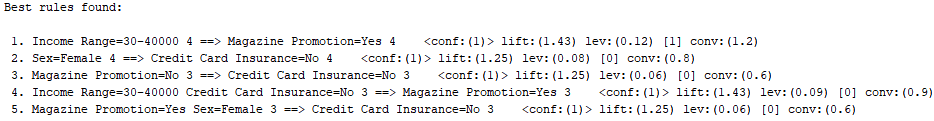


當numRule 設成 5，Minimum Support 為0.35



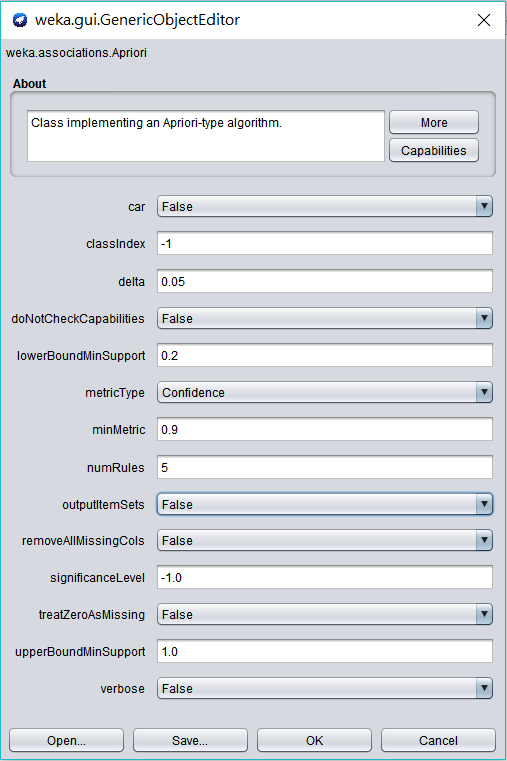
結果不同的原因：當要求較多的Rule時，資料之間的關聯相對會較不緊密，因此Minimum support會降低。

**(c) 將 numRule 設成 10, 列出前 5 條 rule(15%)**



**(d) 如何在 Associator output 產生 Itemset, 請截圖說明並附上 Itemset 結果。(15%)**

在參數設定的地方，將「outputItemSets」設為「True」，按下「OK」後「Start」



以下為Itemset的結果。



在Large Itemsets L(1)的部分：Weka找出各屬性值「單一」存在的個數。例如：Income Range = 20-3000的資料有3筆。

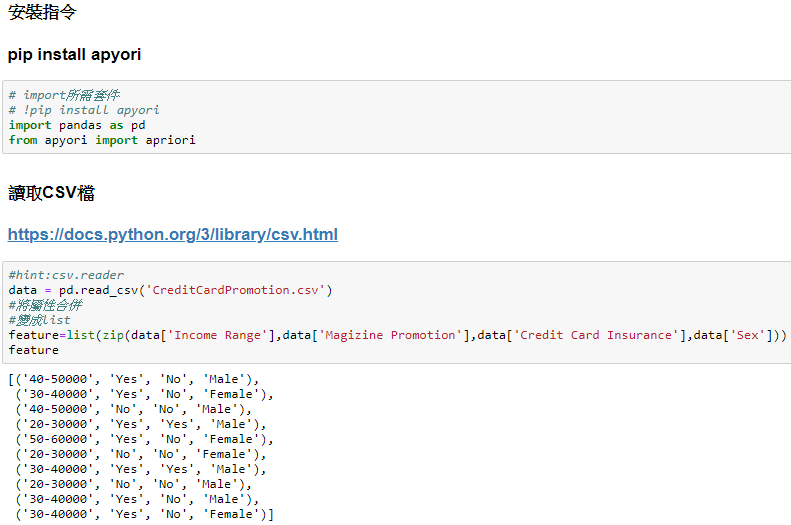
在Large Itemsets L(2)的部分：由Large Itemsets L(1)找出「任兩者」之間依賴關係的個數。例如：Income Range = 30-4000且Magazine Promotion= Yes的資料有4筆。

在Large Itemsets L(3)的部分：由Large Itemsets L(1)找出「任三者」之間依賴關係的個數。例如：Income Range = 30-4000且Magazine Promotion= Yes且Credit Card Insurance = No的資料有3筆。

**使用 Python ：**

**(e) 將已修改過的 CreditCardPromotion.arff 轉成 csv 檔，使用 Apriori 演算法進 行分析,設定 confidence = 0.9、minimum support = 0.2，過程中對所有重要程式 步驟進行截圖並加以說明，越詳盡越好。(15%)**

1. 首先，透過指令「pip install apyori」安裝與import引入所需套件，並讀取CreditCardPromotion.csv檔，將屬性變為list。



1. 使用aporiori()，將min\_support設為0.25，min\_confidence設為0.9，再將結果轉為list並直接印出。

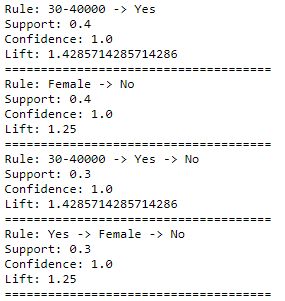


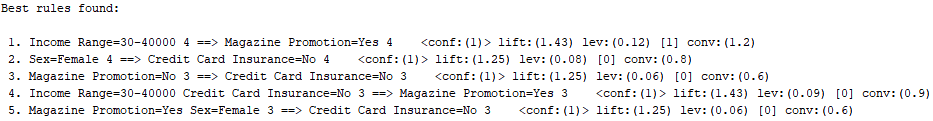
1. 排序：將結果轉為list後再轉成Dataframe並使用，sort\_values() 以support值進行排序(ascending = False 可讓排序由大至小)，再將結果轉為list，並印出。



**(f) 調整 apriori( )內的參數，產生與(c) 小題一樣的結果，截圖並加以說明(15%)**

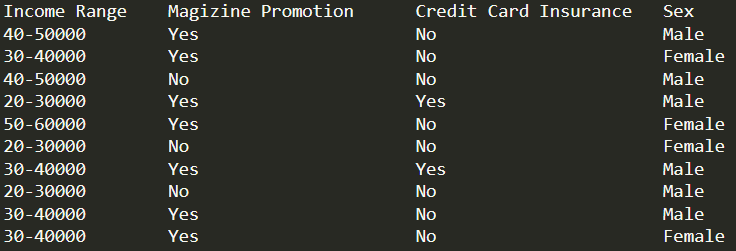
執行結果如下：



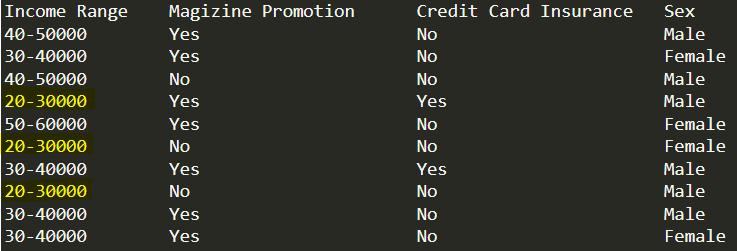
比較Weka執行結果（numRule = 10 , confidence = 0.9, minimum support = 0.2）  


說明：相較於Weka，Python的執行結果只有四筆，而Weka的第四筆在Python中是不存在的，其餘的Rule在兩者之間都是一樣的。

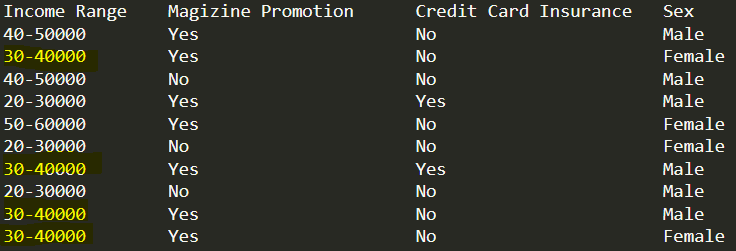
**(g) 請自己計算 (記錄在 Word 上或手算拍照附圖皆可),並與 (d) 小題結果做驗證。 (15%)**



1. 首先找出Large Itemsets L(1)

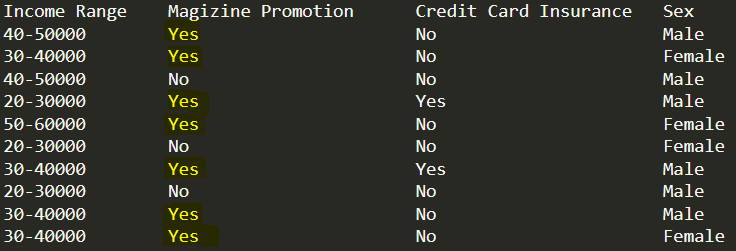
Income Range = 20-3000，3筆

Income Range = 30-4000，4筆

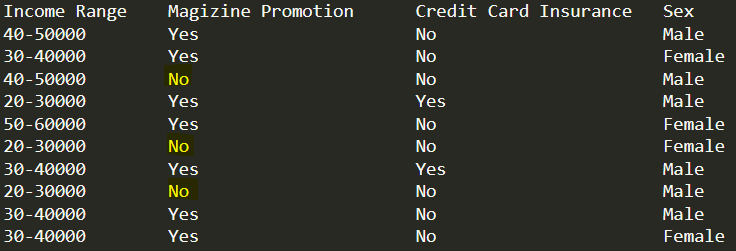


Income Range = 40-5000，2筆，數值太小不列入

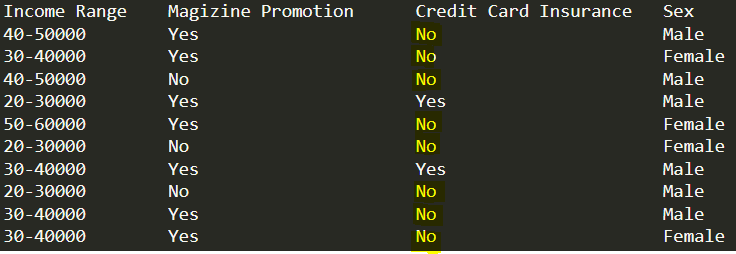
Magazine Promotion = Yes，7筆



Magazine Promotion = No，3筆

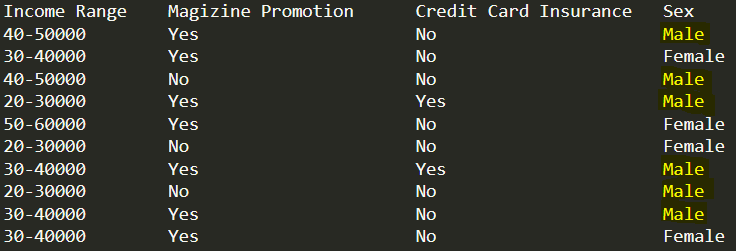


Credit Card Insurance = No，8筆

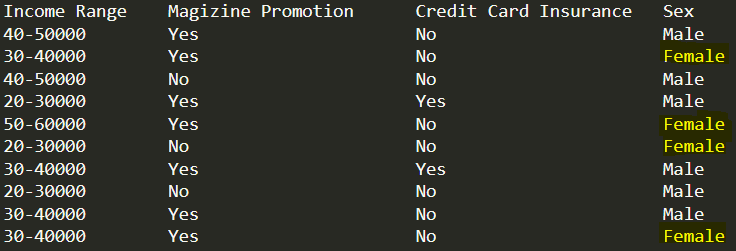


Credit Card Insurance = Yes，2筆，數值太小不列入

Sex = Male，6筆



Sex = Female，4筆



1. 由 i. 找出列出所有可能的Large Itemsets L(2)

Income Range = 20-3000，Magazine Promotion = Yes，1筆，不列入

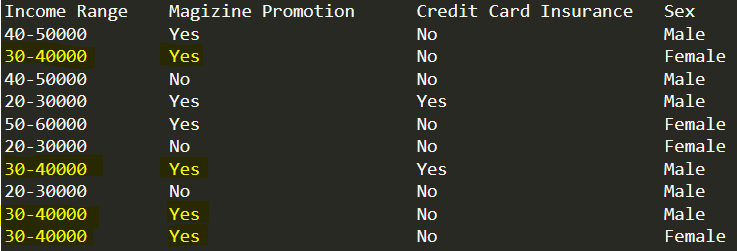
Income Range = 20-3000，Magazine Promotion = No，2筆，不列入

Income Range = 20-3000，Credit Card Insurance = No，2筆，不列入

Income Range = 20-3000，Sex = Male，2筆，不列入

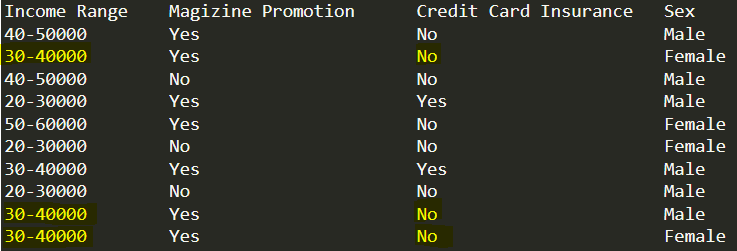
Income Range = 20-3000，Sex = Female，1筆，不列入

Income Range = 30-4000，Magazine Promotion = Yes，4筆



Income Range = 30-4000，Magazine Promotion = No，0筆，不列入

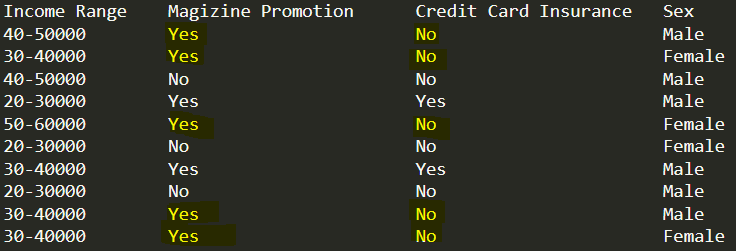
Income Range = 30-4000，Credit Card Insurance = No，3筆



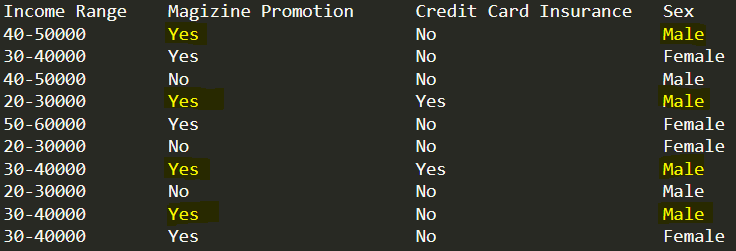
Income Range = 30-4000，Sex = Male，2筆，不列入

Income Range = 30-4000，Sex = Female，2筆，不列入

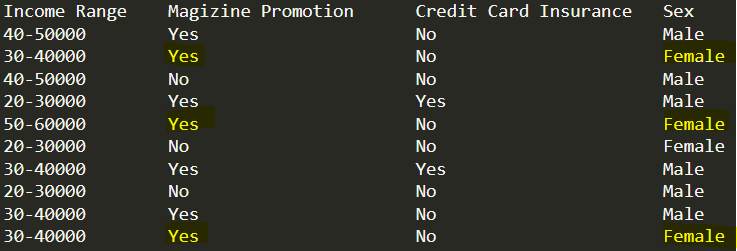
Magazine Promotion = Yes，Credit Card Insurance = No，5筆



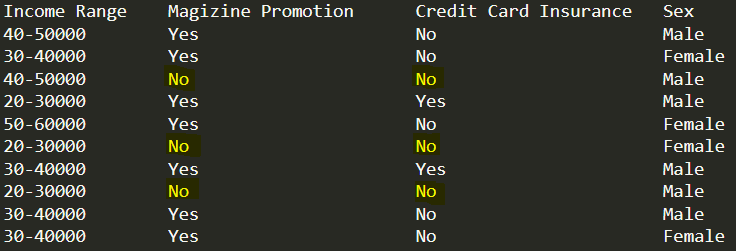
Magazine Promotion = Yes，Sex = Male，4筆



Magazine Promotion = Yes，Sex = Female，3筆



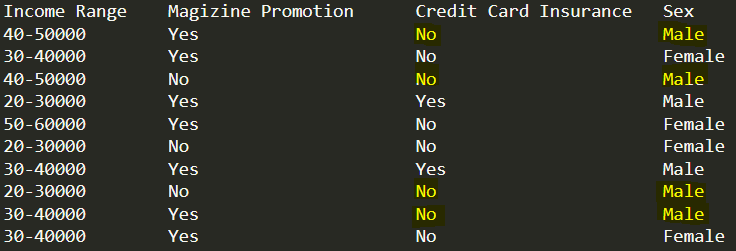
Magazine Promotion = No，Credit Card Insurance = No，3筆



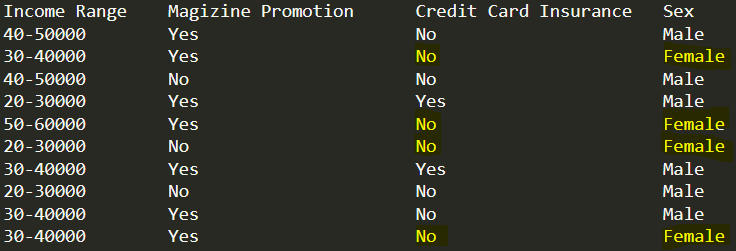
Magazine Promotion = No，Sex = Male，2筆，不列入

Magazine Promotion = No，Sex = Female，1筆，不列入

Credit Card Insurance = No，Sex = Male，4筆

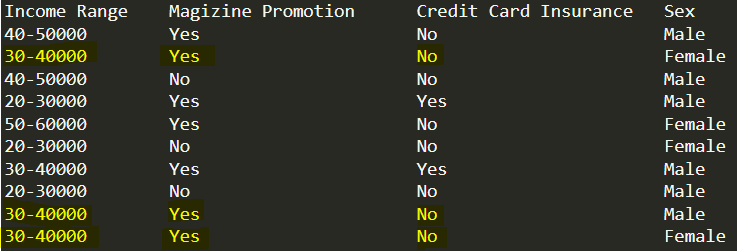


Credit Card Insurance = No，Sex = Female，4筆



1. 由 ii. 找出列出所有可能的Large Itemsets L(3)

Income Range = 30-4000，Magazine Promotion = Yes，Credit Card Insurance = No，3筆



Income Range = 30-4000，Magazine Promotion = Yes，Sex = Male，2筆，不列入

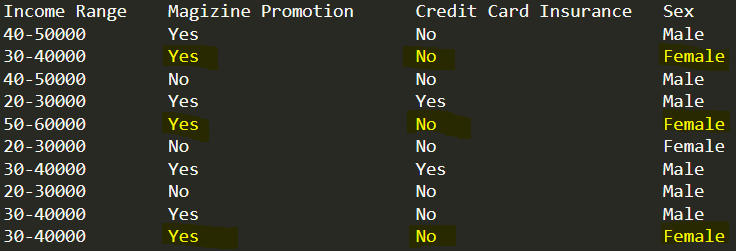
Income Range = 30-4000，Magazine Promotion = Yes，Sex = Female，2筆，不列入

Income Range = 30-4000，Credit Card Insurance = No，Sex = Male，1筆，不列入

Income Range = 30-4000，Credit Card Insurance = No，Sex = Female，2筆，不列入

Magazine Promotion = Yes，Credit Card Insurance = No，Sex = Male，2筆，不列入

Magazine Promotion = Yes，Credit Card Insurance = No，Sex = Female，3筆



Magazine Promotion = No，Credit Card Insurance = No，Sex = Male，2筆，不列入

Magazine Promotion = No，Credit Card Insurance = No，Sex = Female，1筆，不列入