CSCD 429 Data Mining Lab6 (15 points)

Step 0: Download the data file data.csv from Canvas.

Step 1: **Manually** use **any** one of the frequent itemset and association rule discovery algorithms introduced in class to find frequent itemsets and strong association rules in "**data.csv**". Assume the *min_support* = 40%, and the *min_confidence* = 100%.

Step 2: Use **RapidMiner** to re-do this assignment. RapidMiner has included FP-Growth frequent pattern discovery algorithm. Use it to find all the frequent itemsets in the data file. Please make sure to set up all the input parameters properly (FP-Growth DOES work correctly as long as you set the parameters right!!). Later add another operator to generate all the strong association rules.

Step 3: Use **R** to re-do this assignment. In R, Package 'arules' supports the generation of frequent itemsets and associations rules. You can find the document of this package from this link:

https://cran.r-project.org/web/packages/arules/arules.pdf

Submission:

You need to submit your report online via Canvas which should include:

(1) (5 points) All the frequent itemsets and strong association rules manually generated in Step 1. Please show your work.
(2) (3 points) A screen copy of all the frequent itemsets and association rules generated in Step 2 using RapidMiner.
(3) (5 points) The code you used in R, and the frequent itemsets and association rules generated in Step 3.
(4) (2 points) Comparison of all three steps. If you see different results among them, explain why.