CSCD 429/529 Data Mining Lab 7 (15 points)

Step 0: Download the data file data.csv from Canvas. Note that the TRUE value means the corresponding item is in the transaction, and the FALSE value means the corresponding item is not in the transaction.

Step 1: **Manually** (means not using any data mining libraries) choose **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 **Orange** to re-do this assignment. First you need to install Orange add-on "Associate", then you can use "Frequent Itemsets" and "Association Rules" two widgets from this add-on. In addition, you need to change the input data into a format that can be properly used by these two widgets. You can find the related document from this link:

https://orangedatamining.com/widget-catalog/associate/associationrules/

 Step 3: Use **R** to re-do this assignment. In R, Package "arules" supports the generation of frequent itemsets and associations rules. You need to change the input data into a format that can be properly used by "arules" library. 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 Orange.
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.