Cairo University
Faculty of Engineering
Computer Engineering Department

CMPN451
Spring 2020

Data Mining, Big Data and Analytics.

Lab 3 – Association Rule Mining

Requirement (1):

Hint: Use the help to know what you don't know! It is called help for a reason.

1.	First of all, start by cleaning the workspace and setting the working directory.
2.	Load the libraries **arules** and **arulesViz**
3.	Load the transactions in the file **AssociationRules.csv** using the function **read.transactions.** **Make sure you don't include the header line in the dataset.**
4.	Display the transactions in a readable format using the function **inspect. Display only the first 100 transactions.**
5.	What are the most frequent two items in the dataset? What are their frequencies? **Hint: use the function itemFrequency or use the function summary.**
6.	Plot the 5 most frequent items of the transactions using the function **itemFrequencyPlot**
7.	Generate the association rules from the transactions using the apriori algorithm. Set the minimum support = 0.01, minimum confidence = 0.5, minimum cardinality (number of items in the rule) = 2. Use the function **apriori**
8.	Now, sort the generated rules by support. Search the function sort found in the arules package. Show only the first 6 rules.
9.	Sort the generated rules by confidence. Show only the first 6 rules.
10.	Sort the generated rules by lift. Show only the first 6 rules.
11.	Plot the generated rules with **support** as x-axis, **confidence** as y-axis and lift as shading. **Use the function plot** in arules package.
12.	Based on (8-11), Can you tell now what are the most interesting rules that are really useful and provide a real business value and an insight to the concerned corporate?