DATA MINING - ASSIGNMENT 3 [40 points]

Assignment Date: 27 October 2020
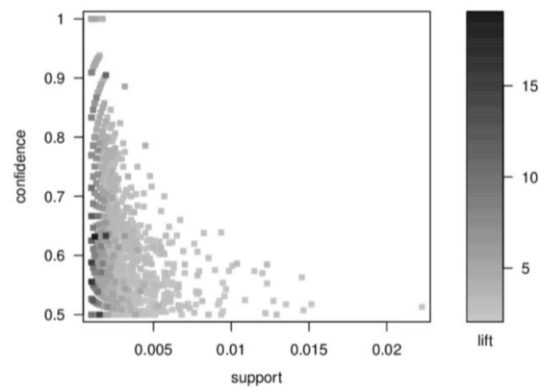
Due Date: 08 November 2020 Midnight

INSTRUCTIONS:

A.   Please do your own work.

B.   Submit your work well before time.

TASKS:

1.   Consider the following dataset: https://www.kaggle.com/irfanasrullah/groceries.

2.   You are required to find the following:

a.   The first five transactions. [1 Point]

b.   The most frequent item. [2 Points]

c.   The length of the longest transaction. [2 Points]

d.   Mine all frequent itemsets and report top-10 longest frequent itemsets. [10 Points]

e.   Mine all association rules and report top-10 longest association rules. [10 Points]

f.   Draw a scatter plot of the form shown below where x-axis shows the support and y-axis show

the confidence and the shading of a point shows the lift. [5 Points]

3.   Prepare a detailed report for tasks in point 2. [10 Points]

****