Quiz 3: Frequent Itemsets

Name: _____ ID: _____

1) Consider the following input file of basket data and a support threshold **s = 2**, answer the following questions.

$B_1 = \{m, c, b\}$	$B_2 = \{m, p, j\}$
$B_3 = \{m, c, b, n\}$	$B_4 = \{c, j\}$
$B_5 = \{m, p, b\}$	$B_6 = \{m, c, b, j\}$
$B_7 = \{c, b, j\}$	$B_8 = \{b, c\}$

Write down two association rules and their **confidence** and **interest** numbers. One of your association rule should be derived from a frequent pair (i.e., X->Y), and the other one should be derived from a frequent triplet (i.e., X, Y->Z) (3pts)

$\{m\} \rightarrow \{c\}$ (0.5 Points) confidence = $3/5 = 0.6$ (0.5 points) Interest = $|3/5 - 6/8|$ (0.5 points) $= 0.15$

$\{m, c\} \rightarrow \{b\}$ (0.5 points) confidence = $3/3 = 1$ (0.5 points) Interest = $|1 - 6/8|$ $= 0.25$ (0.5 points)

2) If you have n types of items and x baskets, at most how different many pairs you need to count for finding all frequent pairs (1pts)?

$$nC_2 = \frac{n!}{(n-2)!\,2!} = \frac{n(n-1)(n-2)!}{(n-2)!\,2} = \frac{n(n-1)}{2}$$ (1 point)

4) When should you use the table of triples approach instead of the triangular matrix and why? (2 pts)

[0.5 marks] if fewer than 1/3 of possible pairs actually occur in the market basket data (# pairs with count > 0) then we use table [0.5 points] of triples as we just need to store the pairs with count > 0 so it can save space. For reason

3) Given items apple, beer, tea, and coke if we need to count all possible pairs of them, we can use the one-dimensional vector to store the counters (triangular matrix). Show how you 1. assign an index to each item, 2. Use the lexicographic order to order the vector component, 3. an example of how to find the counter for (apple, tea) (e.g., the counter of (apple,beer) is at position 0 of your vector) (4 pts)

1. Assign 1, 2, 3, 4 to apple, beer, tea, coke respectively [1 point]

2. keep pair counts in lexicographical order [1 point]

   {1,2} {1,3} {1,4} {2,3} {2,4} {3,4}

3. pair $\{i, j\}$ is at position $(i-1)(n - i/2) + j - i$ [1 point]

   eg:- {apple, tea} => {1, 3} is at position 2 [1 point]