

Note: This homework is worth a total of 15 points

Dataset: Convenience store transactions

ID	Bread	Milk	Sugar	Eggs	Shampoo	Conditioner
1	True	True	True	True	True	False
2	True	False	False	True	False	False
3	True	True	False	False	False	False
4	True	True	True	True	True	True
5	True	True	False	False	True	False
6	False	False	False	True	True	False
7	True	True	False	True	False	False
8	True	False	True	True	False	False
9	True	True	True	False	True	True
10	True	True	False	False	True	False

Q1 (5pts): Find all frequent item-sets using Support ≥ 0.45

{Bread} : 90%

{Milk} : 70%

{Eggs} : 60%

{Shampoo} : 60%

{Bread, Milk} : 70%

{Bread, Eggs} : 50%

{Bread, Shampoo} : 50%

{Milk, Shampoo} : 50%

{Bread, Milk, Shampoo} : 50%

Q2 (5pts): Find all strong rules using Confidence ≥ 0.75

Only create rules for item-sets that passed the Support threshold

{Bread} \rightarrow {Milk} : 78%

{Milk} \rightarrow {Bread} : 100%

{Eggs} \rightarrow {Bread} : 83%

{Shampoo} \rightarrow {Bread} : 83%

{Shampoo} \rightarrow {Milk} : 83%

{Bread, Shampoo} \rightarrow {Milk} : 100%

{Milk, Shampoo} \rightarrow {Bread} : 100%

{Shampoo} \rightarrow {Bread, Milk} : 83%

Q3 (5pts): Rank your rules by how informative they are according to their lift values

{Bread, Shampoo} -> {Milk} : 1.43

{Shampoo} -> {Bread, Milk} : 1.19

{Shampoo} -> {Milk} : 1.19

{Milk, Shampoo} -> {Bread} : 1.11

{Bread} -> {Milk} : 1.11

{Milk} -> {Bread} : 1.11

{Eggs} -> {Bread} : 0.92

{Shampoo} -> {Bread} : 0.92