Exercise 10: Apriori algorithm, solution

1.

We start by finding all the itemsets of size 1 and their support.

{bread} – 3/5, support 0.6
{butter} – 4/5, support 0.8
{beer} – 3/5, support 0.6
{milk} – 2/5, support 0.4
{water} – 2/5, support 0.4
{jam} – 2/5, support 0.4
{diapers} – 2/5, support 0.4
{juice} – 2/5, support 0.4

We then keep only the itemsets with support 0.6

{bread} – 3/5, support 0.6
{butter} – 4/5, support 0.8
{beer} – 3/5, support 0.6

Based on frequent itemsets of size 1 we generate itemsets of size 2 and compute their support.

{bread, butter} – 3/5, support 0.6
{butter, beer} – 3/5, support 0.6
{bread, beer} – 3/5, support 0.6

All of these sets have the minimal support of 0.6, they all become the basis for generating the itemsets of size 3. There is only one such set.

{bread, beer, butter} – 3/5, support 0.6

This ends the process of generating all frequent itemsets:

{bread} – 3/5, support 0.6
{butter} – 4/5, support 0.8
{beer} – 3/5, support 0.6
{bread, butter} – 3/5, support 0.6
{butter, beer} – 3/5, support 0.6
{bread, beer} – 3/5, support 0.6
Based on the frequent itemsets we found, we now need to generate association rules of the form:

\[ \text{item}_1 \Rightarrow \{\text{item}_2, \text{item}_3\} \]

Since there are three items in the rule we can only use frequent itemsets of size no less than three to generate the rule. The only such frequent itemset is \{bread, beer, butter\}. We generate all possible association rules for this itemset and compute their confidence:

\[ \text{bread} \Rightarrow \{\text{beer, butter}\} - \text{confidence } \frac{3}{3} = 1.0 \]
\[ \text{beer} \Rightarrow \{\text{bread, butter}\} - \text{confidence } \frac{3}{4} = 0.75 \]
\[ \text{butter} \Rightarrow \{\text{bread, beer}\} - \text{confidence } \frac{3}{4} = 0.75 \]

All of these rules satisfy the minimum confidence of 0.7.

3.

\[ \text{bread} \Rightarrow \{\text{beer, butter}\} - \text{confidence } \frac{3}{3} = 1.0 \]