Associate Based Rule Learning

It makes product recommendations according to the rules learned through association analysis. It is a rule-based machine learning technique used to find patterns (pattern, relationship, structure) in data. It very often obtains the probabilities of the products bought together and provides some suggestions based on these probabilities. One of the best examples of this is the work of American retail giant Walmart.

When Studies were conducted, it was concluded that As a result of these studies, it was noticed that there is a relationship between diapers and beer. It has been observed that a very large proportion of those who buy diapers in a certain period of time also buy beer. After this observation, the rows were arranged accordingly and the sales rates increased.

**Apriori Algorithm:**

To identify relationships between the products, and there are 3 metrics:

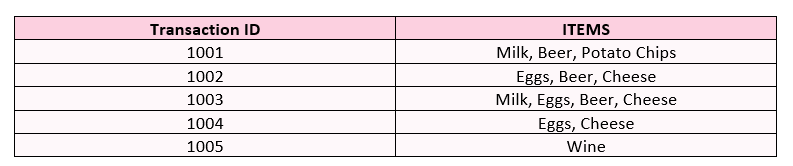
Here Support is the probability of X and Y occurring Together.

Confidence is the probability of selling Y when X is brought.

The coefficient of increase in the probability of purchasing product Y when product X is purchased.

**Working:**

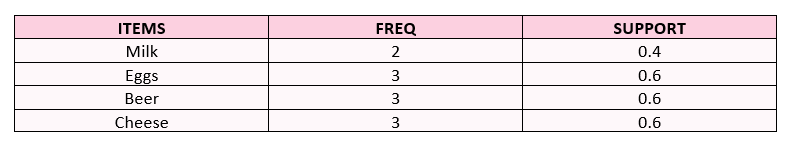
The apriori algorithm calculates the possible product pairs according to the support threshold value that is determined at the beginning of the process.



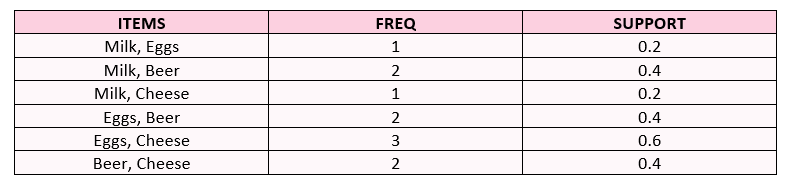
Consider the above table as a dataset. First step is to identify number of products and the occurring frequency. In the below table, we can see that we have 6 items in total. Now calculating each item’s Support using above formula. For milk: OR 2 / 5 = 0.2



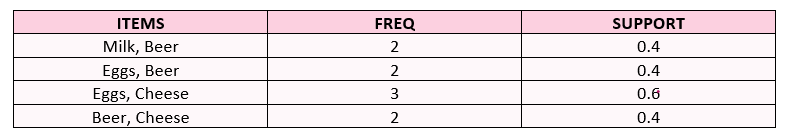
As some items occur only once, They are removed as they don’t show relationships. We set support values at 0.2, so less are eliminated. They updated table will be:



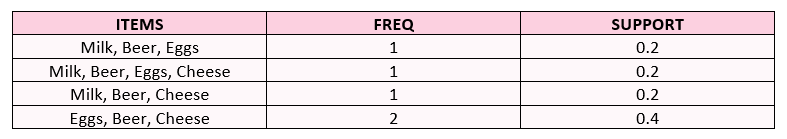
Now we have to identify possible combinations. Here we are checking combination of two items. Items whose combination exists are counted, and their support is also extracted.



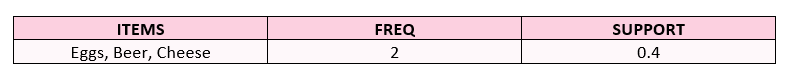
Once we get the combinations, we will eliminate the product pairs that have higher support values as initially identified.

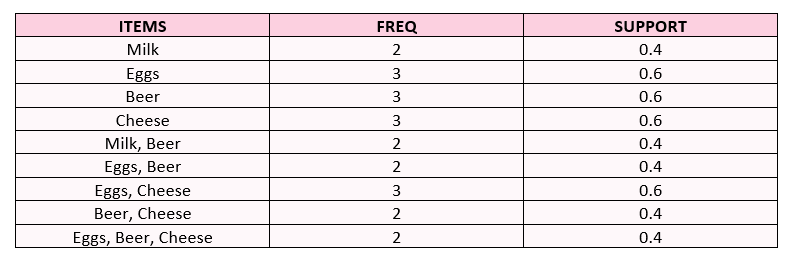


Now Identifying more product pairs, and calculating respective support values.



Eliminate according to the determined support threshold.

  
Now combining all the previous support tables.



Eggs and beer are observed together in 40% of all purchases. 67% of customers who buy eggs also buy beer. Beer sales increase 1.11 times in egg purchases. According to the final table, these kinds of comments can be made.