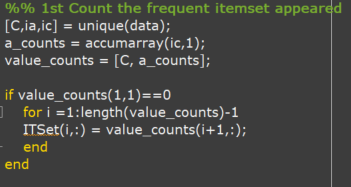
**Association Rule**

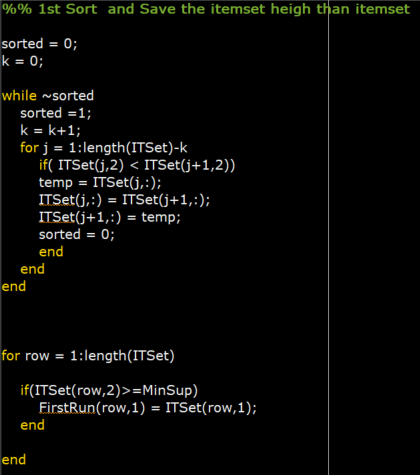
Data I used from IBM Quest Synthetic Data Generator (describe…)

**Apriori Algorithm**

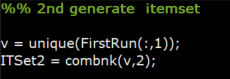
1. Count the frequent of all single items appeared



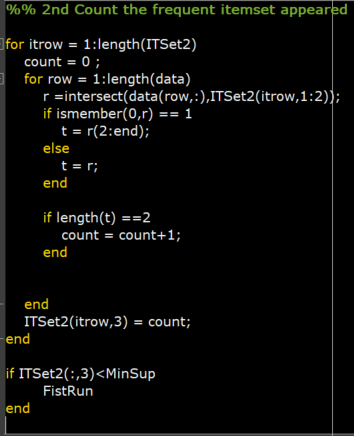
1. Compare with original data, save the itemset frequent higher than minimum support.



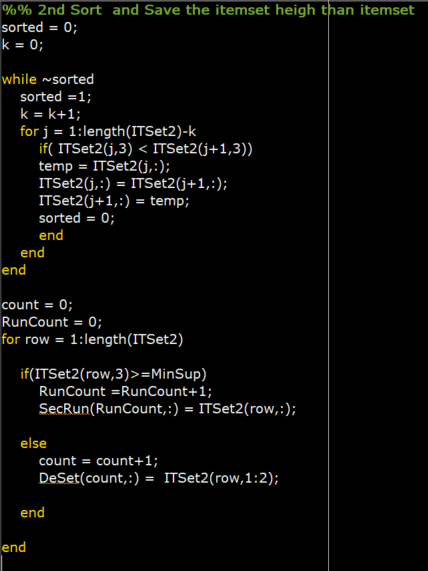
1. According the items of step2 appeared we generate 2 combination itemsets



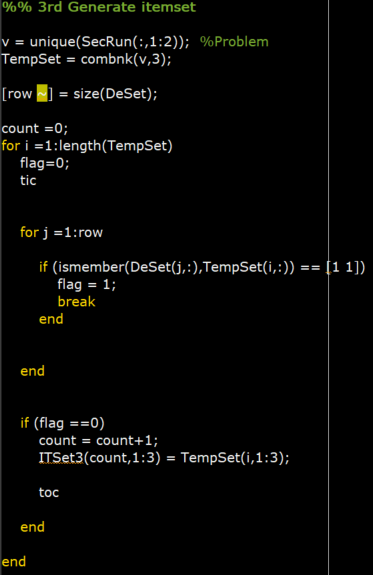
1. Count the frequent of 2 combination itemsets in dataset.



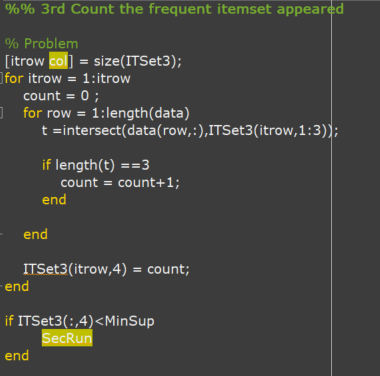
1. Save the itemsets higher than minimum support.



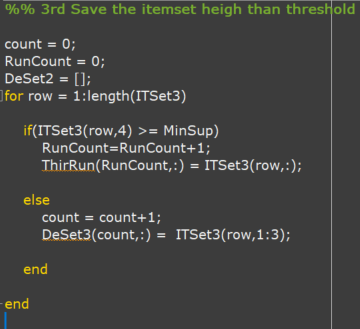
1. According the items of step5 appeared we generate 3 combination itemsets



1. Count the frequent of 3 combination itemsets in dataset.



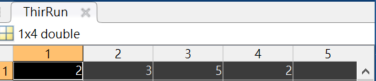
1. Save the itemsets higher than minimum support.



Result

Slide Example

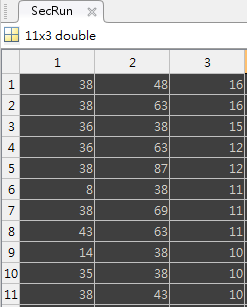
MinSup = 2



The column 1 to 3 is the itemset, column 4 is the frequent.

IBM

MinSup = 10

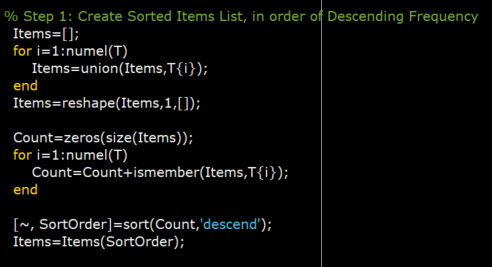


The column 1 to 2 is the itemset, column 3 is the frequent.

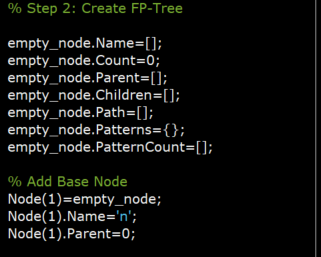
**Frequency Pattern Tree**

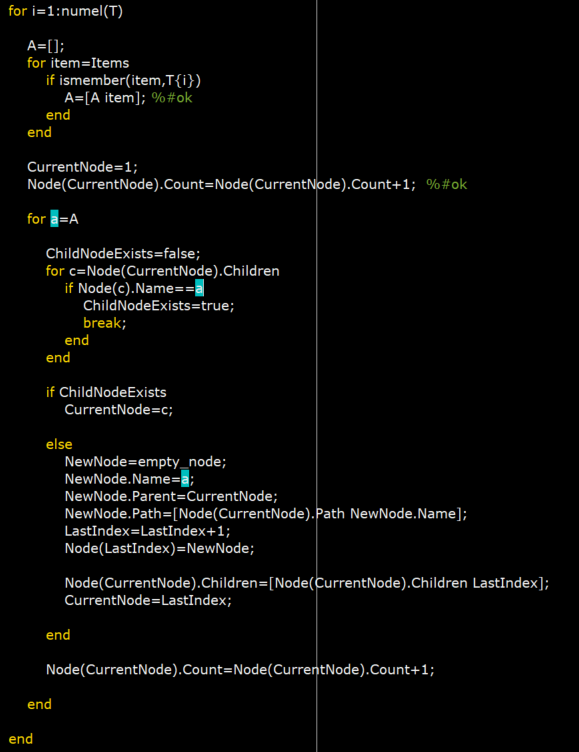
I reference from this website: https://www.mathworks.com/matlabcentral/fileexchange/52868-fp-growth-association-rule-mining

1. Count the frequent of all the items, then sort each itemset according the frequent, the most frequency item will be the first one in itemset.

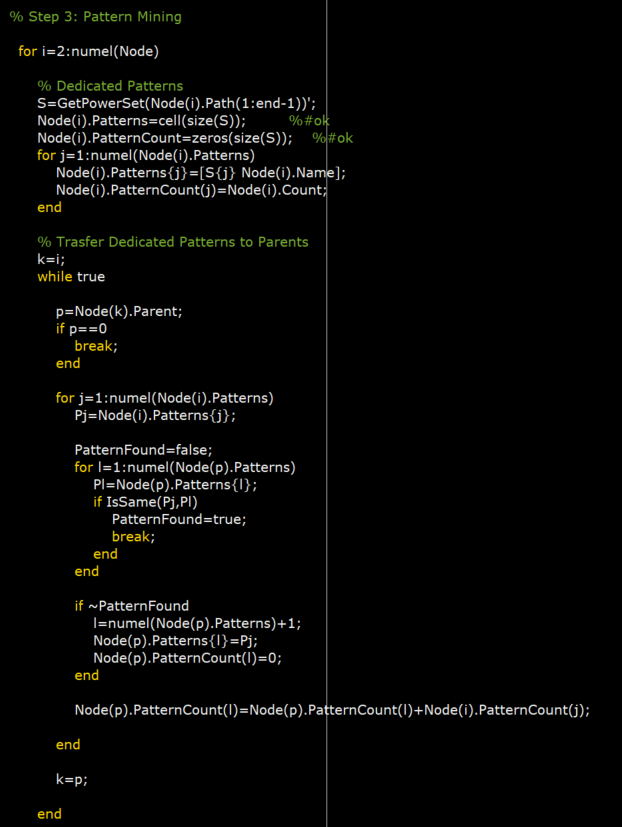


1. Generate the tree, most frequency item will be the first node, the fist itemset will be the first branch.





1. Find the suffix tree from the low frequent item, and scan the itemset that frequent is lowest, record the path, and generate another subtree



Result

Slide Example

