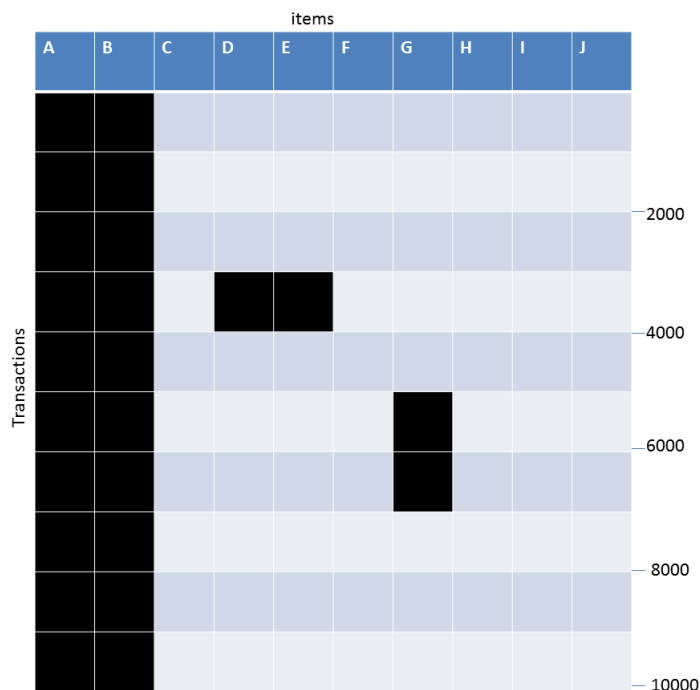
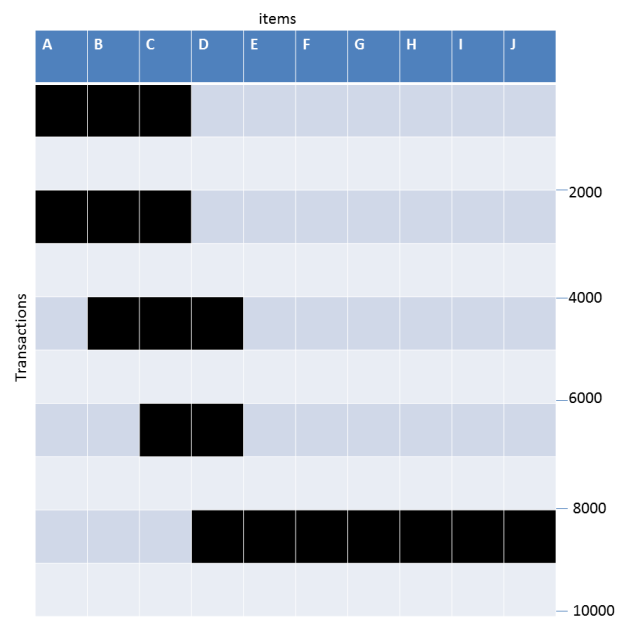
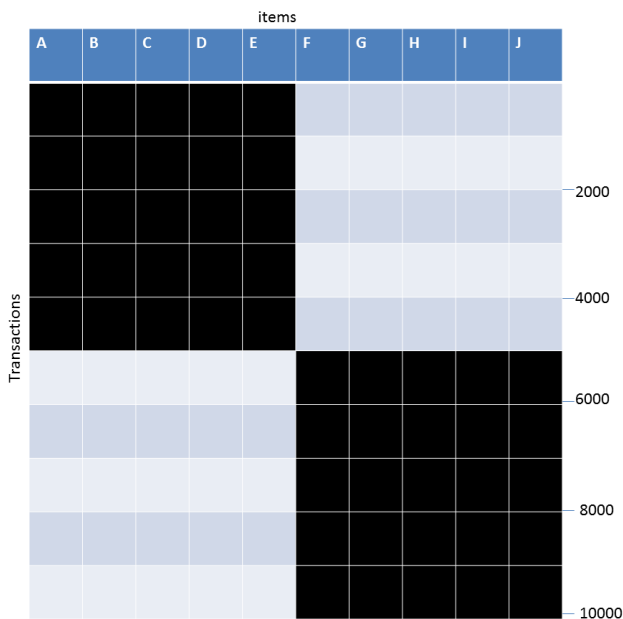


Assignment 2

- Submit the assignment by **20th November by 6:00 PM**.
- There will be a penalty for late submissions.
- Answers should be solved individually (this is not a group assignment).
- Please be careful, do not copy the assignments. **Zero marks** will be awarded to both persons.
- Try to submit a soft copy (using a word file) since a handwritten softcopy is hard to read.

1. Given the following transaction data sets (dark cells indicate the presence of an item in a transaction) and a support threshold of 20%, answer the following questions.



- a. What is the number of frequent itemsets for each dataset? Which dataset will produce the most number of frequent itemsets?
 - b. Which dataset will produce the longest frequent itemset?
 - c. Which dataset will produce frequent itemsets with highest maximum support?
 - d. Which dataset will produce frequent itemsets containing items with widely varying support levels (i.e., itemsets containing items with mixed support, ranging from 20% to more than 70%)?
 - e. What is the number of maximal frequent itemsets for each dataset? Which dataset will produce the most number of maximal frequent itemsets?
 - f. What is the number of closed frequent itemsets for each dataset? Which dataset will produce the most number of closed frequent itemsets?
2. Solve the following exercise questions as an assignment from the book "Introduction to data mining".

6.10 Exercises --- Question Numbers: 2, 3, 6, 7, 12

7.8 Exercises --- Question Numbers: 1, 10, 11

8.7 Exercises --- Question Numbers: 7, 10

Note: Copying assignments is strictly prohibited.