Exercice 1

- Give a characterization of the set M_θ with $\theta = 1$ on any given dataset.
- Suppose we have $\forall\theta : F_\theta = M_\theta = C_\theta$. What does the according dataset look like ?

Exercice 2

Let us have :

case 1 : $M_\theta = \{ABC^3, DE^2, EF^5\}$

case 2 : $C_\theta = \{ABC^3, ABE^5, DE^2, EF^5\}$

- Give the set of FIs and their respective frequencies of the two cases.
  Let $\mathcal{D}$ a transaction dataset with the following horizontal representation $\mathcal{H}_\mathcal{D}$:

t	Items
t_1	A C D
t_2	B C E
t_3	A B C E
t_4	B E
t_5	A B C E
t_6	B C E

- Give the frequent / maximal / closed itemsets with $\theta = 3$.

Exercice 3

Let us consider the following queries where a user asks for particular frequent itemsets :

— $Q_1 : frequent_\theta(P) \wedge minSize_{lb}(P)$

— $Q_2 : frequent_\theta(P) \wedge mand_{set}(P)$

where :

— P is the itemset that we are looking for ;

— $frequent_\theta$: the user asks for frequent itemsets ;

— $minSize_{lb}$: the user asks for itemsets of at least a given size lb ;

— $mand_{set}$: the user asks for itemsets that involve a particular set set of mandatory items.

- Revise the Apriori algorithm to deal with Q_1 and Q_2.