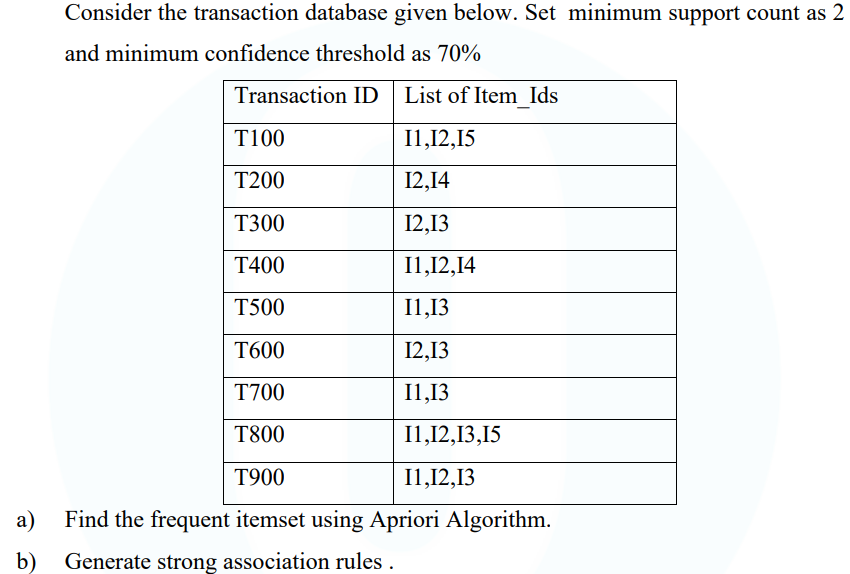
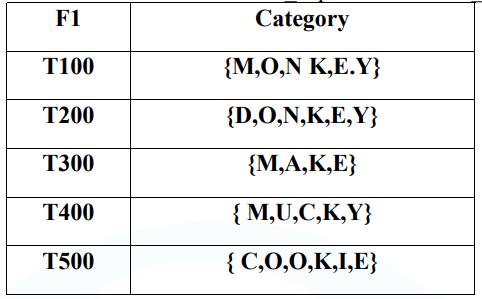
1. Problem related to missing data (Mean. Medium, most probable value, global constant]
2. Problem related to decimal normalization, min-max, z-score normalization
3. Problem related to remove noisy data [ Binning by mean, medium, boundaries]
4. Classification of DM (justify)
5. Box plot – problem
6. Sampling
7. Problem related to Aprior algorithm

|  |  |
| --- | --- |
| **Tid** | **Items** |
| 10 | A, C, D |
| 20 | B, C, E |
| 30 | A, B, C, E |
| 40 | B, E |

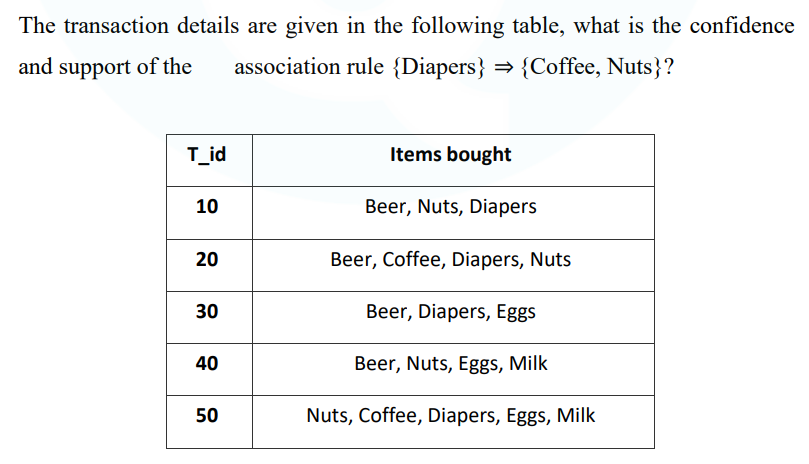


A database has five transactions. Let min\_sup=60% and min\_conf=80%**.**

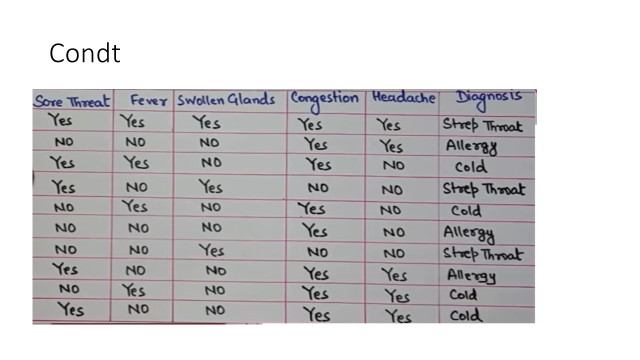


Construct the FP tree for following data set, show the trees separately after reading each transaction**.**

|  |  |
| --- | --- |
| **Tid** | **Item** |
| 1 | {a,b} |
| 2 | {b,c,d} |
| 3 | {a,c,d,e} |
| 4 | {a,d,e} |
| 5 | {a,b,c} |
| 6 | {a,b,c,d} |
| 7 | {a} |
| 8 | {a,b,c} |
| 9 | {a,b,d} |
| 10 | {b,c,e} |



8.



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Color** | **Type** | **Doors** | **Tires** | **Class** |
| Red | SUV | 2 | WhiteWall | + |
| Blue | Minivan | 4 | WhiteWall | - |
| Green | Car | 4 | WhiteWall | - |
| Red | Minivan | 4 | Blackwall | - |
| Green | Car | 2 | Blackwall | + |
| Green | SUV | 4 | Blackwall | - |
| Blue | SUV | 2 | Blackwall | - |
| Blue | Car | 2 | WhiteWall | + |
| Red | SUV | 2 | Blackwall | - |
| Blue | Car | 4 | Blackwall | - |
| Green | SUV | 4 | WhiteWall | + |
| Red | Car | 2 | Blackwall | + |
| Green | SUV | 2 | Blackwall | - |
| Green | Minivan | 4 | WhiteWall | - |