Revision and Preparation for the Exam_Solution

Q1.

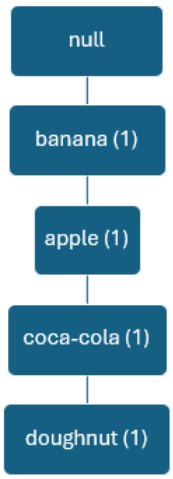
1. Calculate the frequency of each item across all transactions:

   banana: 5
   apple: 4
   coca-cola: 4
   doughnut: 3

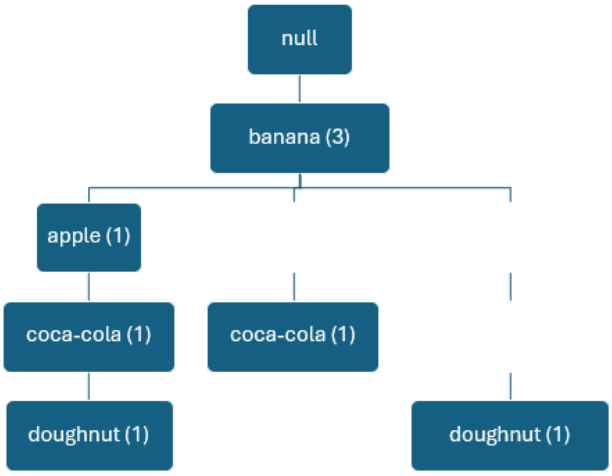
2. Sort the items in each transaction based on frequency:

   Transaction 1: banana, apple, coca-cola, doughnut
   Transaction 2: banana, coca-cola
   Transaction 3: banana, doughnut
   Transaction 4: apple, coca-cola
   Transaction 5: banana, apple, doughnut
   Transaction 6: banana, apple, coca-cola

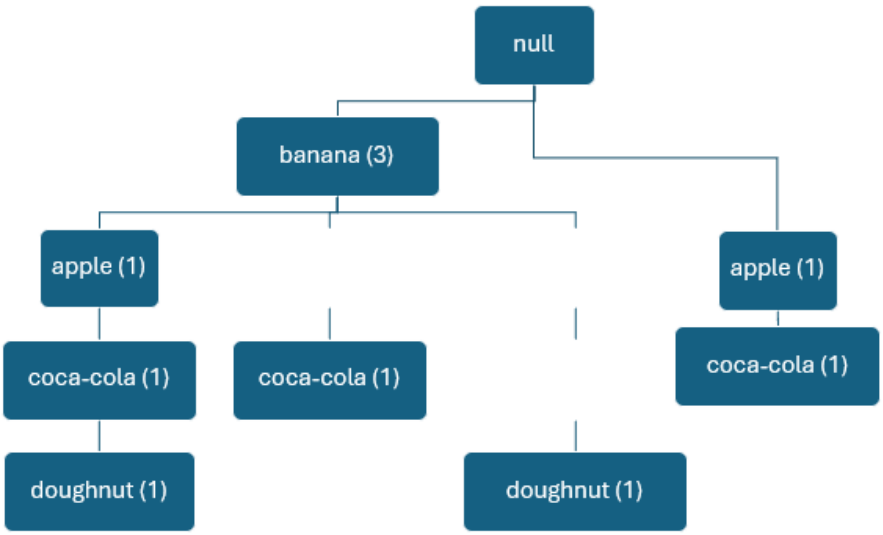
3. Adding Transaction 1:



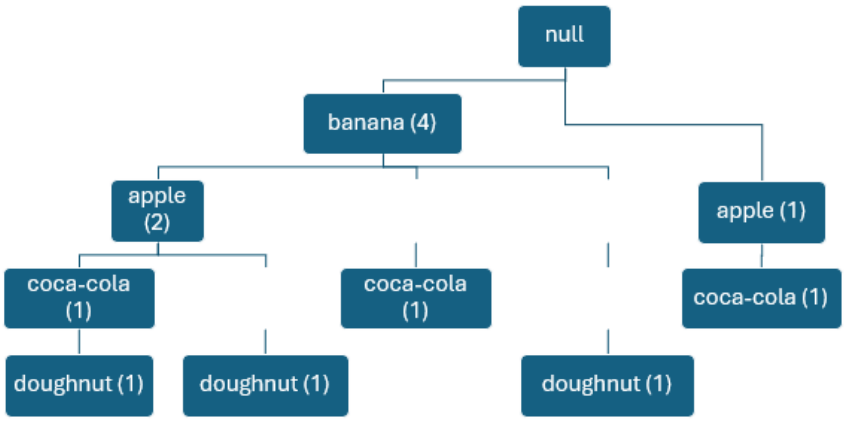
Adding Transaction 2 and then Adding Transaction 3:



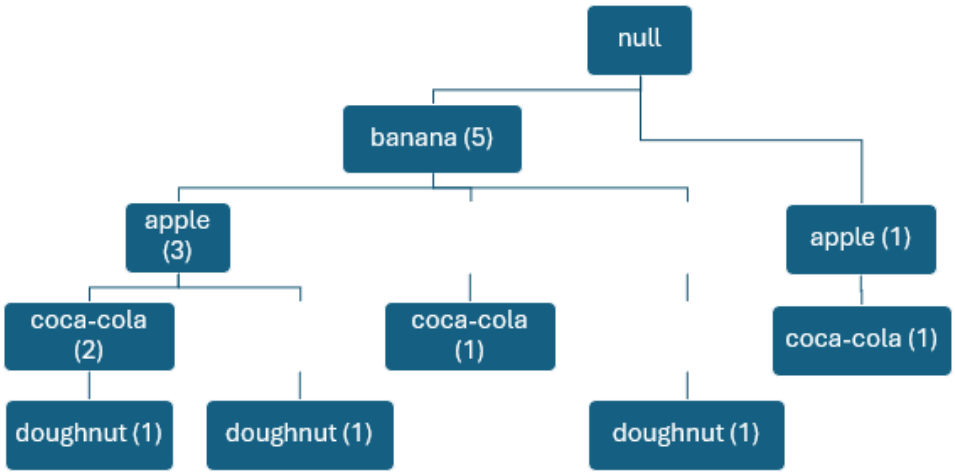
Adding Transaction 4:



Adding Transaction 5:



Adding Transaction 6:



2.

We'll apply the Apriori algorithm with a minimum support (min_sup = 2) to find the frequent itemsets.

1. Generate Frequent 1-Itemsets:

{banana}: 5
{apple}: 4
{coca-cola}: 4
{doughnut}: 3

2. Generate Frequent 2-Itemsets:

{banana, apple}: 3
{banana, coca-cola}: 3
{banana, doughnut}: 3
{apple, coca-cola}: 3
{apple, doughnut}: 2

3. Generate Frequent 3-Itemsets:

{banana, apple, coca-cola}: 2
{banana, apple, doughnut}: 2

4. Generate Frequent 4-Itemsets:

There are no 4-itemsets that meet the minimum support threshold.

5. Frequent Itemsets using Apriori:

{banana}: 5
{apple}: 4
{coca-cola}: 4
{doughnut}: 3
{banana, apple}: 3
{banana, coca-cola}: 3
{banana, doughnut}: 3
{apple, coca-cola}: 3
{apple, doughnut}: 2
{banana, apple, coca-cola}: 2
{banana, apple, doughnut}: 2

3.

For the frequent itemset {apple, banana, doughnut}, the possible association rules are:

{apple} → {banana, doughnut}
{banana} → {apple, doughnut}
{doughnut} → {apple, banana}
{apple, banana} → {doughnut}

{apple, doughnut} → {banana}
{banana, doughnut} → {apple}

Now, we calculate the confidence for each rule:

$$confidence(\{apple\} \rightarrow \{banana, doughnut\}) = \frac{\{apple, banana, doughnut\}}{\{apple\}} = \frac{2}{4} = 0.50$$

$$confidence(\{banana\} \rightarrow \{apple, doughnut\}) = \frac{\{apple, banana, doughnut\}}{\{banana\}} = \frac{2}{5} = 0.40$$

$$confidence(\{doughnut\} \rightarrow \{apple, banana\}) = \frac{\{apple, banana, doughnut\}}{\{doughnut\}} = \frac{2}{3} \approx 0.67$$

$$confidence(\{apple, banana\} \rightarrow \{doughnut\}) = \frac{\{apple, banana, doughnut\}}{\{apple, banana\}} = \frac{2}{3} \approx 0.67$$

$$confidence(\{apple, doughnut\} \rightarrow \{banana\}) = \frac{\{apple, banana, doughnut\}}{\{apple, doughnut\}} = \frac{2}{2} = 1$$

$$confidence(\{banana, doughnut\} \rightarrow \{apple\}) = \frac{\{apple, banana, doughnut\}}{\{banana, doughnut\}} = \frac{2}{3} \approx 0.67$$

Only the following rules meet the minimum confidence threshold of 70%:

{apple, doughnut} → {banana} with 100% confidence.

Q2.

```
SELECT employees.first_name, departments.department_name
FROM employees
INNER JOIN departments ON employees.department_id = departments.department_id;
```

Q3.

```
SELECT e.employee_name, d.department_name AS employee_department
FROM employees e
JOIN departments d ON e.department_ID = d.department_ID
JOIN assignments a ON e.employee_ID = a.employee_ID
JOIN projects p ON a.project_ID = p.project_ID
WHERE e.department_ID != p.department_ID;
```

Q4

E1 is head=narrow, E2 is eyes=elliptical, E3 is size=Large

P(yes) = 4/9

P(E1|yes) = ¼

P(E2|yes) = ¾

P(E3|yes)=1/4

$$P(yes|E) = \frac{\frac{1}{4}*\frac{3}{4}*\frac{1}{4}*\frac{4}{9}}{P(E)} = 0.021 / P(E)$$

P(no)=5/9

P(E1|no) = 3/5

P(E2|no) = 2/5

P(E3|no)=3/5

$$P(no|E) = \frac{\frac{3}{5}*\frac{2}{5}*\frac{3}{5}*\frac{5}{9}}{P(E)} = 0.08 / P(E)$$

⇨ The prediction will be that the example is not venomous