Given the following transaction record

Transaction Records	
Transaction ID	**Items**
#1	apple, banana, coca-cola, doughnut
#2	banana, coco-cola
#3	banana, doughnut
#4	apple, coca-cola
#5	apple, banana, doughnut
#6	apple, banana, coca-cola

1. Build the FP-tree using a minimum support **min_sup** = 2. Show how the tree evolves for each transaction.

2. With the previous transaction record, Use the Apriori algorithm on this dataset and verify that it will generate the same set of frequent itemsets with **min_sup** = 2.

3. Suppose that { **Apple, Banana, Doughnut** } is a frequent item set, derive all its association rules with
*min_confidence = 70%*

1.

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

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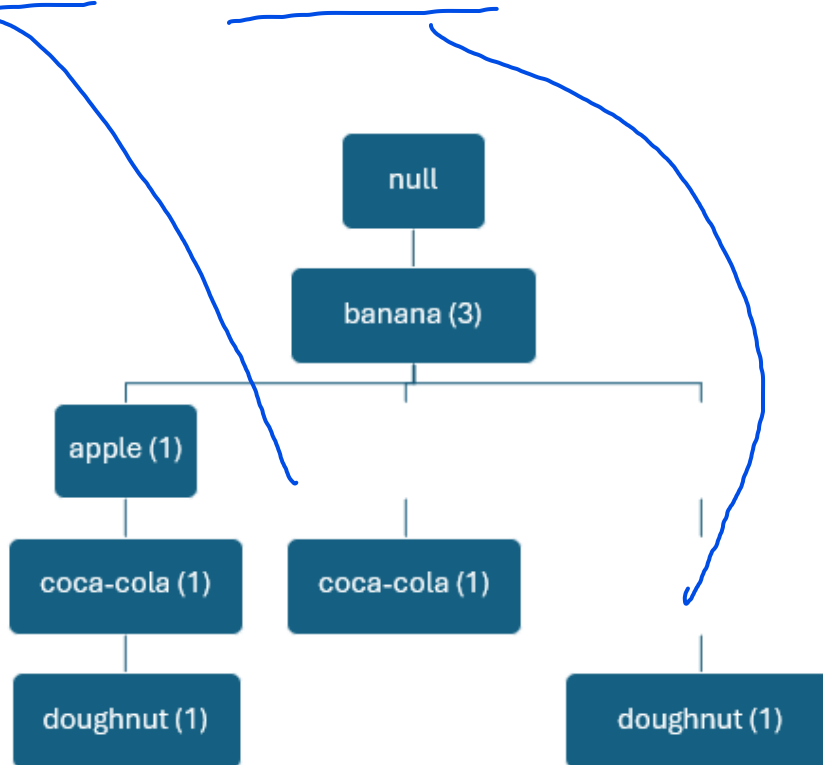apple: 4
coca-cola: 4
doughnut: 3

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
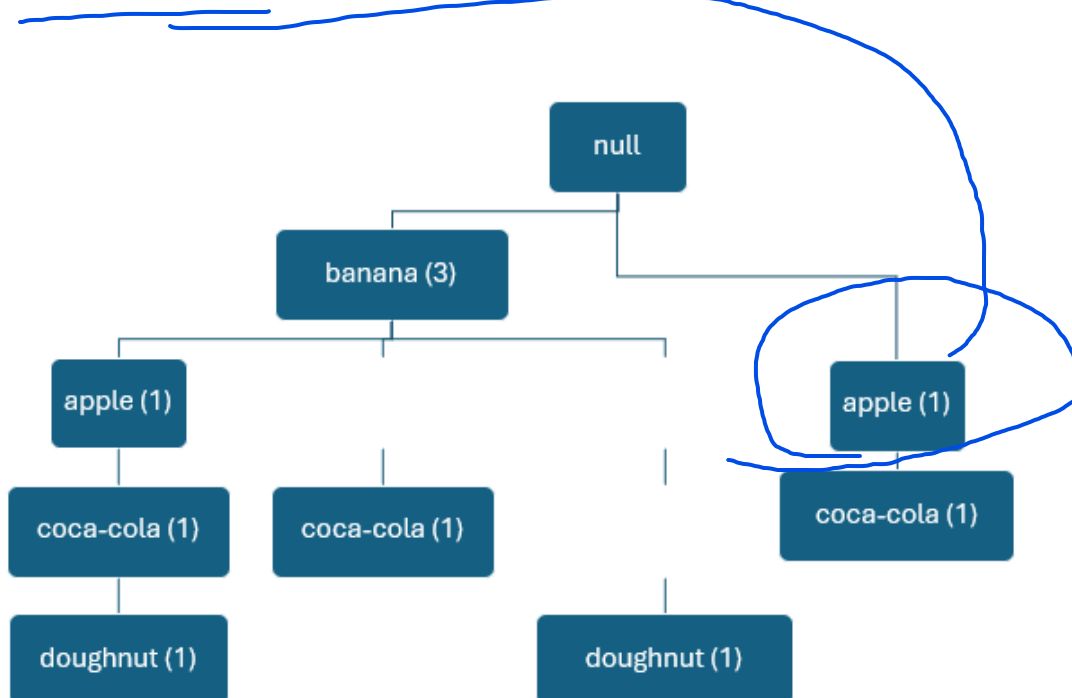   Transaction 6: banana, apple, coca-cola
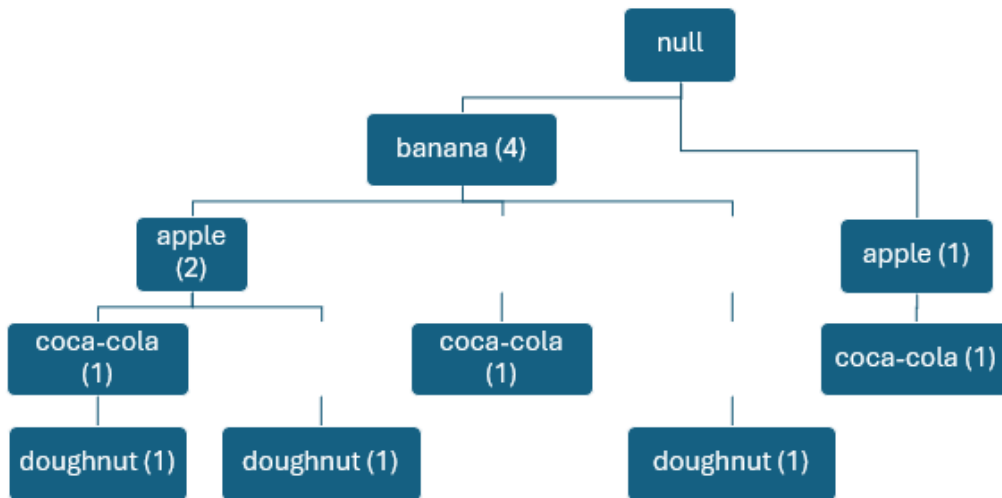
3. Adding Transaction 1:
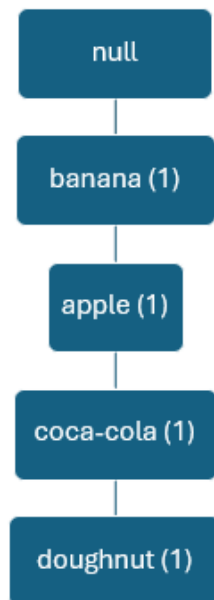
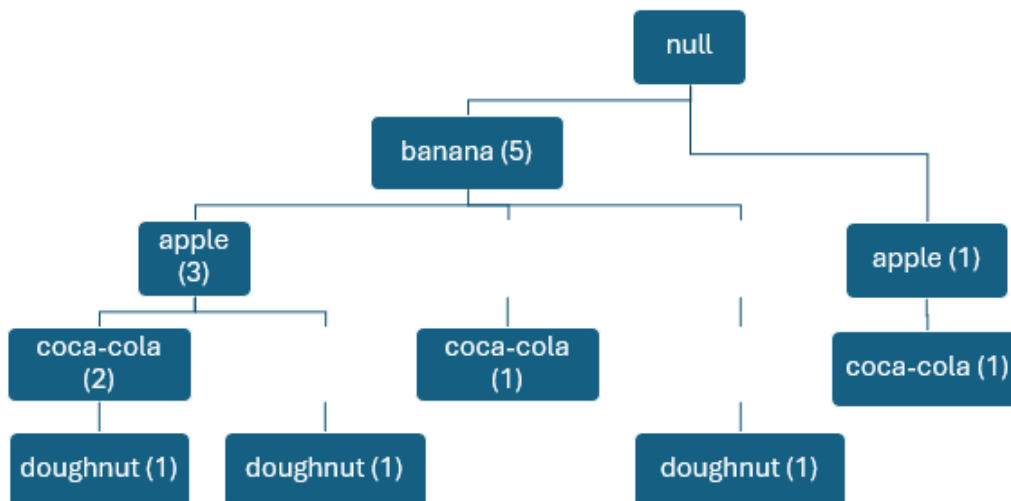Adding Transaction 2 and then Adding Transaction 3:

```
                          null
                            |
                       banana (3)
            ┌───────────────┼───────────────┐
        apple (1)           |                |
            |          coca-cola (1)         |
       coca-cola (1)                         |
            |                            doughnut (1)
       doughnut (1)
```

Adding Transaction 4:

```
                                 null
                    ┌──────────────┼──────────────┐
               banana (3)                      apple (1)
         ┌─────────┼─────────┐                     |
     apple (1)     |          |               coca-cola (1)
         |    coca-cola (1)   |
    coca-cola (1)             |
         |               doughnut (1)
    doughnut (1)
```

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}: 2
{banana, doughnut}: 2
{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}: 2
{banana, doughnut}: 2
{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:

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

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

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

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

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

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

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

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