EXP NO: **HYPOTHESIS GENERATION: DERIVE THE HYPOTHESIS FOR ASSOCIATION RULES TO DISCOVERY OF**

DATE: **STRONG ASSOCIATION RULES; USE CONFIDENCE AND SUPPORT THRESHOLDS.**

AIM:

BACKGROUND THEORY:

Hypothesis Generation:

1. **Define the Objective:**
   Determine what you aim to find with association rule mining.

   This could be identifying products frequently bought together, common sequences of events, etc.

2. **Select Data:**
   Choose the dataset that contains the items or events you want to analyze.

   This dataset should be formatted appropriately for association rule mining, typically in a transactional format.

3. **Identify Items and Transactions:**
   Define what constitutes an item and a transaction in your dataset. An item could be a product, an event, etc., and a transaction could be a purchase instance, a session, etc.

PROCEDURE:

1. **Load Data:**
   Import your dataset into Orange. You can do this by dragging the "File"widget and connecting it to the "Data Table" widget to view the data.
2. **Preprocess Data:**
   Ensure your data is in the correct format. If needed, use the "Preprocess" widget to transform your data.
3. **Apply Association Rule Mining:**
   Drag the "Association Rules" widget into the workspace.
   Connect it to the "File" widget to feed your dataset into the Association Rules widget
4. **Set Parameters:**
   In the "Association Rules" widget.
   Support Threshold: This parameter determines the minimum frequency at which At itemset appears in the dataset. Set a value that is suitable for your dataset size and the rarity of the combinations you are interested in. For example, a support threshold of 0.01 means that the itemset must appear in at least 1% of all transactions. Confidence Threshold: This parameter measures the likelihood that a rule is true for a transaction. Set a confidence level that reflects the strength of the rules you are interested in. For example, a confidence threshold of 0.7 means the rule must be true in at least 70% of the cases where the antecedent occurs.

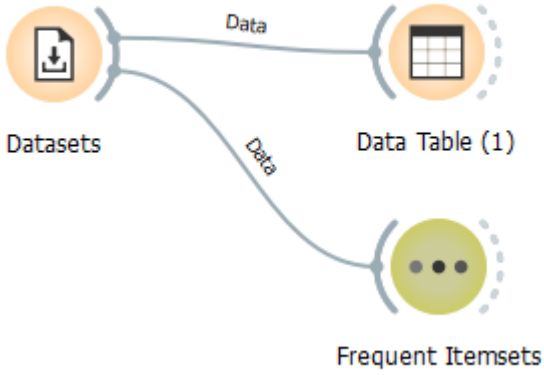
5.Generate Rules:

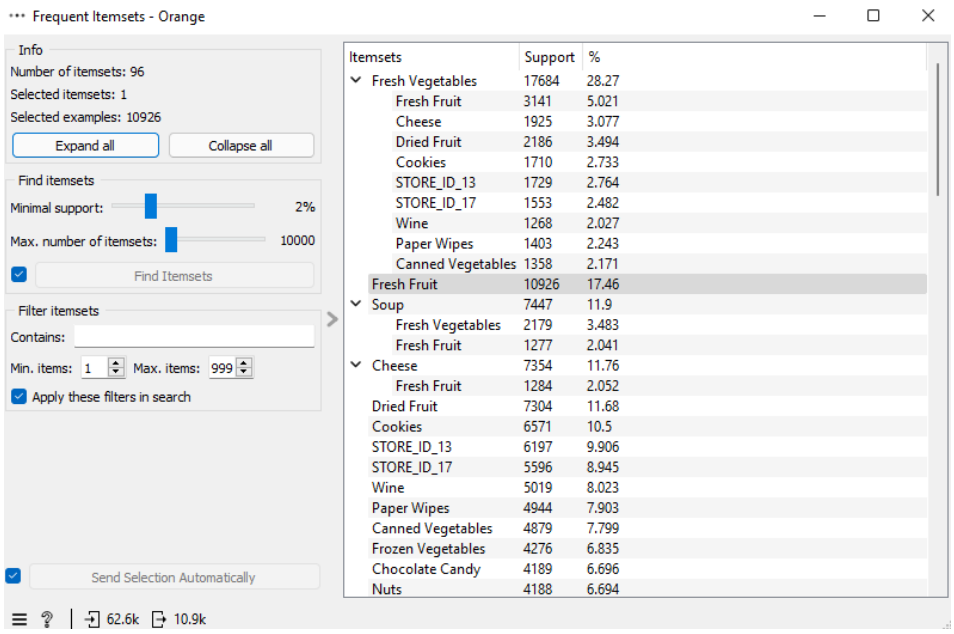
Run the algorithm to generate association rules. The output will be a list of rules that meet the specified support and confidence thresholds.

6. Evaluate and Filter Rules:

Examine the generated rules. You can further filter them by adjusting the thresholds or by using additional metrics such as lift or leverage.

OUTPUT:





Itemsets	Support	%
Fresh Vegetables	17684	28.27
Fresh Fruit	3141	5.021
Cheese	1925	3.077
Dried Fruit	2186	3.494
Cookies	1710	2.733
STORE_ID_13	1729	2.764
STORE_ID_17	1553	2.482
Wine	1268	2.027
Paper Wipes	1403	2.243
Canned Vegetables	1358	2.171
Fresh Fruit	10926	17.46
Soup	7447	11.9
Fresh Vegetables	2179	3.483
Fresh Fruit	1277	2.041
Cheese	7354	11.76
Fresh Fruit	1284	2.052
Dried Fruit	7304	11.68
Cookies	6571	10.5
STORE_ID_13	6197	9.906
STORE_ID_17	5596	8.945
Wine	5019	8.023
Paper Wipes	4944	7.903
Canned Vegetables	4879	7.799
Frozen Vegetables	4276	6.835
Chocolate Candy	4189	6.696
Nuts	4188	6.694

RESULT: