

## Week – 10

### Write a Python program to generate frequent item sets / association rules using Apriori algorithm.

**AIM:** To write a Python program to generate frequent item sets / association rules using Apriori algorithm.

#### DESCRIPTION:

The algorithm is used to finding frequent itemset in a dataset for boolean association rule.




Name of the algorithm is Apriori because it uses prior knowledge of frequent itemset properties. We apply an iterative approach or level-wise search where k-frequent itemsets are used to find k+1 itemsets.

To improve the efficiency of level-wise generation of frequent itemsets, an important property is used called **Apriori property** which helps by reducing the search space.

**Apriori Property:** All non-empty subset of frequent itemset must be frequent. The key concept of Apriori algorithm is its anti-monotonicity of support measure.

**Frequent Itemset:** The itemset that occurs frequently is called frequent itemset.

#### PROGRAM:

 `pip install apyori` Looking in indexes: <https://pypi.org/simple>, <https://us-python.pkg.dev/colab-wheels/public/simple/>  
Collecting apyori  
 Downloading apyori-1.1.2.tar.gz (8.6 kB)  
Building wheels for collected packages: apyori  
 Building wheel for apyori (setup.py) ... done  
Created wheel for apyori: filename=apyori-1.1.2-py3-none-any.whl size=5975 sha256=15a6583e7917bfeb62bedbd6532e8c18b988a87105cfc6715453f07cc4ac0198  
Stored in directory: /root/.cache/pip/wheels/cb/f6/e1/57973c631d27efd1a2f375bd6a83b2a616c4021f24aab84080  
Successfully built apyori  
Installing collected packages: apyori  
Successfully installed apyori-1.1.2 `import numpy as np  
import pandas as pd  
from apyori import apriori`

```
data=pd.read_csv('/apri.csv')
```

```
print(data)
```

```
0      Wine  Chips  Bread  Butter  Milk  Apple
1      Wine  Chips  Bread  Butter  Milk   NaN
2      NaN   Chips   NaN    NaN    NaN  Apple
3      Wine  Chips  Bread  Butter  Milk  Apple
4      Wine  Chips   NaN    NaN    Milk   NaN
5      Wine  Chips  Bread  Butter   NaN  Apple
6      Wine  Chips   NaN    NaN    Milk   NaN
7      Wine   NaN   Bread   NaN    NaN  Apple
8      Wine   NaN   Bread  Butter  Milk   NaN
9      NaN   Chips  Bread  Butter   NaN  Apple
10     Wine   NaN   NaN    Butter  Milk  Apple
11     Wine  Chips  Bread  Butter  Milk   NaN
12     Wine   NaN   Bread   NaN    Milk  Apple
13     Wine   NaN   Bread  Butter  Milk  Apple
14     Wine  Chips  Bread  Butter  Milk  Apple
15     NaN   Chips  Bread  Butter  Milk  Apple
16     NaN   Chips   NaN    Butter  Milk  Apple
17     Wine  Chips  Bread  Butter  Milk  Apple
18     Wine   NaN   Bread  Butter  Milk  Apple
```

```
data.shape
```


```
(19, 6)
```

```
records=[]
for i in range(0,19):
    records.append([str(data.values[i,j]) for j in range(0,6)])
```

```
ass_rules=apriori(records,min_support=0.5,confidence=0.7)
result=list(ass_rules)
```

```
print(len(result))
```

```
21
```

 `print((result))`

## OUTPUT:

```
[RelationRecord(items=frozenset({'Apple'}), support=0.6842105263157895,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Apple'}), confidence=0.6842105263157895, lift=1.0)]),
RelationRecord(items=frozenset({'Bread'}), support=0.7368421052631579,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Bread'}), confidence=0.7368421052631579, lift=1.0)]),
RelationRecord(items=frozenset({'Butter'}), support=0.7368421052631579,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Butter'}), confidence=0.7368421052631579, lift=1.0)]),
RelationRecord(items=frozenset({'Chips'}), support=0.631578947368421,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Chips'}), confidence=0.631578947368421, lift=1.0)]),
RelationRecord(items=frozenset({'Milk'}), support=0.7894736842105263,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Milk'}), confidence=0.7894736842105263, lift=1.0)]),
RelationRecord(items=frozenset({'Wine'}), support=0.7368421052631579,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Wine'}), confidence=0.7368421052631579, lift=1.0)]),
RelationRecord(items=frozenset({'nan'}), support=0.8421052631578947,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'nan'}), confidence=0.8421052631578947, lift=1.0)]),
RelationRecord(items=frozenset({'Bread', 'Apple'}), support=0.5263157894736842,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Bread', 'Apple'}), confidence=0.5263157894736842, lift=1.0),
OrderedStatistic(items_base=frozenset({'Apple'}), items_add=frozenset({'Bread'}),
confidence=0.7692307692307692, lift=1.043956043956044),
OrderedStatistic(items_base=frozenset({'Bread'}), items_add=frozenset({'Apple'}),
confidence=0.7142857142857143, lift=1.043956043956044)]),
RelationRecord(items=frozenset({'Butter', 'Apple'}), support=0.5263157894736842,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Butter', 'Apple'}), confidence=0.5263157894736842, lift=1.0),
OrderedStatistic(items_base=frozenset({'Apple'}), items_add=frozenset({'Butter'}),
confidence=0.7692307692307692, lift=1.043956043956044),
OrderedStatistic(items_base=frozenset({'Butter'}), items_add=frozenset({'Apple'}),
```

```

confidence=0.7142857142857143, lift=1.043956043956044))),
RelationRecord(items=frozenset({'nan', 'Apple'}), support=0.5263157894736842,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'nan', 'Apple'}), confidence=0.5263157894736842, lift=1.0),
OrderedStatistic(items_base=frozenset({'Apple'}), items_add=frozenset({'nan'}),
confidence=0.7692307692307692, lift=0.9134615384615384),
OrderedStatistic(items_base=frozenset({'nan'}), items_add=frozenset({'Apple'}),
confidence=0.625, lift=0.9134615384615384)]),
RelationRecord(items=frozenset({'Bread', 'Butter'}), support=0.631578947368421,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Bread', 'Butter'}), confidence=0.631578947368421, lift=1.0),
OrderedStatistic(items_base=frozenset({'Bread'}), items_add=frozenset({'Butter'}),
confidence=0.8571428571428571, lift=1.163265306122449),
OrderedStatistic(items_base=frozenset({'Butter'}), items_add=frozenset({'Bread'}),
confidence=0.8571428571428571, lift=1.163265306122449)]),
RelationRecord(items=frozenset({'Bread', 'Milk'}), support=0.5789473684210527,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Bread', 'Milk'}), confidence=0.5789473684210527, lift=1.0),
OrderedStatistic(items_base=frozenset({'Bread'}), items_add=frozenset({'Milk'}),
confidence=0.7857142857142858, lift=0.9952380952380954),
OrderedStatistic(items_base=frozenset({'Milk'}), items_add=frozenset({'Bread'}),
confidence=0.7333333333333334, lift=0.9952380952380954)]),
RelationRecord(items=frozenset({'Bread', 'Wine'}), support=0.5789473684210527,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Bread', 'Wine'}), confidence=0.5789473684210527, lift=1.0),
OrderedStatistic(items_base=frozenset({'Bread'}), items_add=frozenset({'Wine'}),
confidence=0.7857142857142858, lift=1.0663265306122451),
OrderedStatistic(items_base=frozenset({'Wine'}), items_add=frozenset({'Bread'}),
confidence=0.7857142857142858, lift=1.0663265306122451)]),
RelationRecord(items=frozenset({'Bread', 'nan'}), support=0.5789473684210527,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Bread', 'nan'}), confidence=0.5789473684210527, lift=1.0),
OrderedStatistic(items_base=frozenset({'Bread'}), items_add=frozenset({'nan'}),
confidence=0.7857142857142858, lift=0.9330357142857144),
OrderedStatistic(items_base=frozenset({'nan'}), items_add=frozenset({'Bread'}),
confidence=0.6875000000000001, lift=0.9330357142857145)]),
RelationRecord(items=frozenset({'Butter', 'Milk'}), support=0.631578947368421,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Butter', 'Milk'}), confidence=0.631578947368421, lift=1.0),
OrderedStatistic(items_base=frozenset({'Butter'}), items_add=frozenset({'Milk'}),
confidence=0.8571428571428571, lift=1.0857142857142856),

```

```

OrderedStatistic(items_base=frozenset({'Milk'}), items_add=frozenset({'Butter'}),
confidence=0.7999999999999999, lift=1.0857142857142856)),
RelationRecord(items=frozenset({'Butter', 'Wine'}), support=0.5263157894736842,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Butter', 'Wine'}), confidence=0.5263157894736842, lift=1.0),
OrderedStatistic(items_base=frozenset({'Butter'}), items_add=frozenset({'Wine'}),
confidence=0.7142857142857143, lift=0.9693877551020409),
OrderedStatistic(items_base=frozenset({'Wine'}), items_add=frozenset({'Butter'}),
confidence=0.7142857142857143, lift=0.9693877551020409))],
RelationRecord(items=frozenset({'nan', 'Butter'}), support=0.5789473684210527,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'nan', 'Butter'}), confidence=0.5789473684210527, lift=1.0),
OrderedStatistic(items_base=frozenset({'Butter'}), items_add=frozenset({'nan'}),
confidence=0.7857142857142858, lift=0.9330357142857144),
OrderedStatistic(items_base=frozenset({'nan'}), items_add=frozenset({'Butter'}),
confidence=0.6875000000000001, lift=0.9330357142857145))],
RelationRecord(items=frozenset({'Milk', 'Wine'}), support=0.631578947368421,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Milk', 'Wine'}), confidence=0.631578947368421, lift=1.0),
OrderedStatistic(items_base=frozenset({'Milk'}), items_add=frozenset({'Wine'}),
confidence=0.7999999999999999, lift=1.0857142857142856),
OrderedStatistic(items_base=frozenset({'Wine'}), items_add=frozenset({'Milk'}),
confidence=0.8571428571428571, lift=1.0857142857142856))],
RelationRecord(items=frozenset({'nan', 'Milk'}), support=0.631578947368421,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'nan', 'Milk'}), confidence=0.631578947368421, lift=1.0),
OrderedStatistic(items_base=frozenset({'Milk'}), items_add=frozenset({'nan'}),
confidence=0.7999999999999999, lift=0.95),
OrderedStatistic(items_base=frozenset({'nan'}), items_add=frozenset({'Milk'}),
confidence=0.75, lift=0.95))], RelationRecord(items=frozenset({'nan', 'Wine'}),
support=0.5789473684210527,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'nan', 'Wine'}), confidence=0.5789473684210527, lift=1.0),
OrderedStatistic(items_base=frozenset({'Wine'}), items_add=frozenset({'nan'}),
confidence=0.7857142857142858, lift=0.9330357142857144),
OrderedStatistic(items_base=frozenset({'nan'}), items_add=frozenset({'Wine'}),
confidence=0.6875000000000001, lift=0.9330357142857145))],
RelationRecord(items=frozenset({'Bread', 'Butter', 'Milk'}),
support=0.5263157894736842,
ordered_statistics=[OrderedStatistic(items_base=frozenset(),
items_add=frozenset({'Bread', 'Butter', 'Milk'}), confidence=0.5263157894736842,
lift=1.0), OrderedStatistic(items_base=frozenset({'Bread'}),
items_add=frozenset({'Butter', 'Milk'}),

```

```
confidence=0.7142857142857143, lift=1.130952380952381),  
OrderedStatistic(items_base=frozenset({'Butter'}), items_add=frozenset({'Bread',  
'Milk'}), confidence=0.7142857142857143, lift=1.2337662337662338),  
OrderedStatistic(items_base=frozenset({'Milk'}), items_add=frozenset({'Bread',  
'Butter'}), confidence=0.6666666666666666, lift=1.0555555555555556),  
OrderedStatistic(items_base=frozenset({'Bread', 'Butter'}),  
items_add=frozenset({'Milk'}), confidence=0.8333333333333334,  
lift=1.0555555555555556),  
OrderedStatistic(items_base=frozenset({'Bread', 'Milk'}),  
items_add=frozenset({'Butter'}), confidence=0.909090909090909,  
lift=1.2337662337662336),  
OrderedStatistic(items_base=frozenset({'Butter', 'Milk'}),  
items_add=frozenset({'Bread'}), confidence=0.8333333333333334,  
lift=1.1309523809523812)]]]
```