In [1]:

```
import pandas as pd
import numpy as np
from mlxtend.frequent_patterns import apriori, association_rules
```

In [2]:

```
df=pd.read_csv(r"C:\Users\Ritvik\Downloads\GroceryStoreDataSet.csv",names = ['products'
```

In [3]:

```
df.head()
```

Out[3]:

	products
0	MILK,BREAD,BISCUIT
1	BREAD,MILK,BISCUIT,CORNFLAKES
2	BREAD,TEA,BOURNVITA
3	JAM,MAGGI,BREAD,MILK
4	MAGGI,TEA,BISCUIT

In [4]:

```
df
```

Out[4]:

	products
0	MILK,BREAD,BISCUIT
1	BREAD,MILK,BISCUIT,CORNFLAKES
2	BREAD,TEA,BOURNVITA
3	JAM,MAGGI,BREAD,MILK
4	MAGGI,TEA,BISCUIT
5	BREAD,TEA,BOURNVITA
6	MAGGI,TEA,CORNFLAKES
7	MAGGI,BREAD,TEA,BISCUIT
8	JAM,MAGGI,BREAD,TEA
9	BREAD,MILK
10	COFFEE,COCK,BISCUIT,CORNFLAKES
11	COFFEE,COCK,BISCUIT,CORNFLAKES
12	COFFEE,SUGER,BOURNVITA
13	BREAD,COFFEE,COCK
14	BREAD,SUGER,BISCUIT
15	COFFEE,SUGER,CORNFLAKES
16	BREAD,SUGER,BOURNVITA
17	BREAD,COFFEE,SUGER

	products
18	BREAD,COFFEE,SUGER
19	TEA,MILK,COFFEE,CORNFLAKES

In [5]:

```
df.shape
```

Out[5]:

```
(20, 1)
```

In [6]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20 entries, 0 to 19
Data columns (total 1 columns):
 #   Column    Non-Null Count  Dtype 
---  ------    --------------  ----- 
 0   products  20 non-null     object
dtypes: object(1)
memory usage: 288.0+ bytes
```

In [7]:

```
data= list(df["products"].apply(lambda x:x.split(",") ))
```

In [8]:

```
data
```

Out[8]:

```
[['MILK', 'BREAD', 'BISCUIT'],
 ['BREAD', 'MILK', 'BISCUIT', 'CORNFLAKES'],
 ['BREAD', 'TEA', 'BOURNVITA'],
 ['JAM', 'MAGGI', 'BREAD', 'MILK'],
 ['MAGGI', 'TEA', 'BISCUIT'],
 ['BREAD', 'TEA', 'BOURNVITA'],
 ['MAGGI', 'TEA', 'CORNFLAKES'],
 ['MAGGI', 'BREAD', 'TEA', 'BISCUIT'],
 ['JAM', 'MAGGI', 'BREAD', 'TEA'],
 ['BREAD', 'MILK'],
 ['COFFEE', 'COCK', 'BISCUIT', 'CORNFLAKES'],
 ['COFFEE', 'COCK', 'BISCUIT', 'CORNFLAKES'],
 ['COFFEE', 'SUGER', 'BOURNVITA'],
 ['BREAD', 'COFFEE', 'COCK'],
 ['BREAD', 'SUGER', 'BISCUIT'],
 ['COFFEE', 'SUGER', 'CORNFLAKES'],
 ['BREAD', 'SUGER', 'BOURNVITA'],
 ['BREAD', 'COFFEE', 'SUGER'],
 ['BREAD', 'COFFEE', 'SUGER'],
 ['TEA', 'MILK', 'COFFEE', 'CORNFLAKES']]
```

In [9]:

```
from mlxtend.frequent_patterns import apriori, association_rules
from mlxtend.preprocessing import TransactionEncoder
a = TransactionEncoder()
a_data = a.fit(data).transform(data)
df = pd.DataFrame(a_data,columns=a.columns_)
df = df.replace(False,0)
df= df.replace(True,1)
df
```

Out[9]:

	BISCUIT	BOURNVITA	BREAD	COCK	COFFEE	CORNFLAKES	JAM	MAGGI	MILK	SUGER	TEA
0	1	0	1	0	0	0	0	0	1	0	0
1	1	0	1	0	0	1	0	0	1	0	0
2	0	1	1	0	0	0	0	0	0	0	1
3	0	0	1	0	0	0	1	1	1	0	0
4	1	0	0	0	0	0	0	1	0	0	1
5	0	1	1	0	0	0	0	0	0	0	1
6	0	0	0	0	0	1	0	1	0	0	1
7	1	0	1	0	0	0	0	1	0	0	1
8	0	0	1	0	0	0	1	1	0	0	1
9	0	0	1	0	0	0	0	0	1	0	0
10	1	0	0	1	1	1	0	0	0	0	0
11	1	0	0	1	1	1	0	0	0	0	0
12	0	1	0	0	1	0	0	0	0	1	0
13	0	0	1	1	1	0	0	0	0	0	0
14	1	0	1	0	0	0	0	0	0	1	0
15	0	0	0	0	1	1	0	0	0	1	0
16	0	1	1	0	0	0	0	0	0	1	0
17	0	0	1	0	1	0	0	0	0	1	0
18	0	0	1	0	1	0	0	0	0	1	0
19	0	0	0	0	1	1	0	0	1	0	1

In [10]:

```
df.sum()
```

Out[10]:

```
BISCUIT        7
BOURNVITA      4
BREAD         13
COCK           3
COFFEE         8
CORNFLAKES     6
JAM            2
MAGGI          5
MILK           5
SUGER          6
TEA            7
dtype: int64
```

In [11]:

```
df = apriori(df, min_support = 0.2, use_colnames = True, verbose = 1 )
df
```

```
Processing 42 combinations | Sampling itemset size 3
```

Out[11]:

	support	itemsets

	support	itemsets
0	0.35	(BISCUIT)
1	0.20	(BOURNVITA)
2	0.65	(BREAD)
3	0.40	(COFFEE)
4	0.30	(CORNFLAKES)
5	0.25	(MAGGI)
6	0.25	(MILK)
7	0.30	(SUGER)
8	0.35	(TEA)
9	0.20	(BREAD, BISCUIT)
10	0.20	(MILK, BREAD)
11	0.20	(SUGER, BREAD)
12	0.20	(BREAD, TEA)
13	0.20	(COFFEE, CORNFLAKES)
14	0.20	(SUGER, COFFEE)
15	0.20	(TEA, MAGGI)

In [15]:

```
df_ar = association_rules(df, metric = "confidence", min_threshold = 0.6)
df_ar
```

Out[15]:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	convi
0	(MILK)	(BREAD)	0.25	0.65	0.2	0.800000	1.230769	0.0375	
1	(SUGER)	(BREAD)	0.30	0.65	0.2	0.666667	1.025641	0.0050	
2	(CORNFLAKES)	(COFFEE)	0.30	0.40	0.2	0.666667	1.666667	0.0800	
3	(SUGER)	(COFFEE)	0.30	0.40	0.2	0.666667	1.666667	0.0800	
4	(MAGGI)	(TEA)	0.25	0.35	0.2	0.800000	2.285714	0.1125	

In [ ]: