Association-Rules

In [1]:

```
!pip install mlxtend
```

...

In [2]:

```
# import libraries
import pandas as pd
from mlxtend.frequent_patterns import apriori,association_rules
from mlxtend.preprocessing import TransactionEncoder
```

In [3]:

```
# load dataset
titanic=pd.read_csv("C:/Users/Ashraf/Documents/Datafiles/Titanic.csv")
titanic.head()
```

Out[3]:

	Class	Gender	Age	Survived
0	3rd	Male	Child	No
1	3rd	Male	Child	No
2	3rd	Male	Child	No
3	3rd	Male	Child	No
4	3rd	Male	Child	No

Pre-Processing

As the data is not in transaction formation We are using transaction Encoder

In [4]:

```
df=pd.get_dummies(titanic)
df.head()
```

Out[4]:

	Class_1st	Class_2nd	Class_3rd	Class_Crew	Gender_Female	Gender_Male	Age_Adult	Ag
0	0	0	1	0	0	1	0	
1	0	0	1	0	0	1	0	
2	0	0	1	0	0	1	0	
3	0	0	1	0	0	1	0	
4	0	0	1	0	0	1	0	

Apriori Algorithm

In [6]:

```
frequent_itemsets = apriori(df, min_support=0.6, use_colnames=True)
frequent_itemsets
```

Out[6]:

	support	itemsets
0	0.786461	(Gender_Male)
1	0.950477	(Age_Adult)
2	0.676965	(Survived_No)
3	0.757383	(Gender_Male, Age_Adult)
4	0.619718	(Gender_Male, Survived_No)
5	0.653339	(Survived_No, Age_Adult)
6	0.603816	(Gender_Male, Survived_No, Age_Adult)

In [7]:

```
rules = association_rules(frequent_itemsets, metric="lift", min_threshold=0.7)
rules
```

Out[7]:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leve
0	(Gender_Male)	(Age_Adult)	0.786461	0.950477	0.757383	0.963027	1.013204	0.00
1	(Age_Adult)	(Gender_Male)	0.950477	0.786461	0.757383	0.796845	1.013204	0.00
2	(Gender_Male)	(Survived_No)	0.786461	0.676965	0.619718	0.787984	1.163995	0.08
3	(Survived_No)	(Gender_Male)	0.676965	0.786461	0.619718	0.915436	1.163995	0.08
4	(Survived_No)	(Age_Adult)	0.676965	0.950477	0.653339	0.965101	1.015386	0.00
5	(Age_Adult)	(Survived_No)	0.950477	0.676965	0.653339	0.687380	1.015386	0.00
6	(Gender_Male, Survived_No)	(Age_Adult)	0.619718	0.950477	0.603816	0.974340	1.025106	0.01
7	(Gender_Male, Age_Adult)	(Survived_No)	0.757383	0.676965	0.603816	0.797241	1.177669	0.09
8	(Survived_No, Age_Adult)	(Gender_Male)	0.653339	0.786461	0.603816	0.924200	1.175139	0.08
9	(Gender_Male)	(Survived_No, Age_Adult)	0.786461	0.653339	0.603816	0.767764	1.175139	0.08
10	(Survived_No)	(Gender_Male, Age_Adult)	0.676965	0.757383	0.603816	0.891946	1.177669	0.09
11	(Age_Adult)	(Gender_Male, Survived_No)	0.950477	0.619718	0.603816	0.635277	1.025106	0.01

An leverage value of 0 indicates independence. Range will be [-1 1]

A high conviction value means that the consequent is highly depending on the antecedent and range [0 inf]

In [8]:

```
rules.sort_values('lift',ascending = False)[0:20]
```

Out[8]:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leve
7	(Gender_Male, Age_Adult)	(Survived_No)	0.757383	0.676965	0.603816	0.797241	1.177669	0.09
10	(Survived_No)	(Gender_Male, Age_Adult)	0.676965	0.757383	0.603816	0.891946	1.177669	0.09
8	(Survived_No, Age_Adult)	(Gender_Male)	0.653339	0.786461	0.603816	0.924200	1.175139	0.08
9	(Gender_Male)	(Survived_No, Age_Adult)	0.786461	0.653339	0.603816	0.767764	1.175139	0.08
2	(Gender_Male)	(Survived_No)	0.786461	0.676965	0.619718	0.787984	1.163995	0.08
3	(Survived_No)	(Gender_Male)	0.676965	0.786461	0.619718	0.915436	1.163995	0.08
6	(Gender_Male, Survived_No)	(Age_Adult)	0.619718	0.950477	0.603816	0.974340	1.025106	0.01
11	(Age_Adult)	(Gender_Male, Survived_No)	0.950477	0.619718	0.603816	0.635277	1.025106	0.01
4	(Survived_No)	(Age_Adult)	0.676965	0.950477	0.653339	0.965101	1.015386	0.00
5	(Age_Adult)	(Survived_No)	0.950477	0.676965	0.653339	0.687380	1.015386	0.00
0	(Gender_Male)	(Age_Adult)	0.786461	0.950477	0.757383	0.963027	1.013204	0.00
1	(Age_Adult)	(Gender_Male)	0.950477	0.786461	0.757383	0.796845	1.013204	0.00

In [9]:

```
rules[rules.lift>1]
```

Out[9]:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leve
0	(Gender_Male)	(Age_Adult)	0.786461	0.950477	0.757383	0.963027	1.013204	0.00
1	(Age_Adult)	(Gender_Male)	0.950477	0.786461	0.757383	0.796845	1.013204	0.00
2	(Gender_Male)	(Survived_No)	0.786461	0.676965	0.619718	0.787984	1.163995	0.08
3	(Survived_No)	(Gender_Male)	0.676965	0.786461	0.619718	0.915436	1.163995	0.08
4	(Survived_No)	(Age_Adult)	0.676965	0.950477	0.653339	0.965101	1.015386	0.00
5	(Age_Adult)	(Survived_No)	0.950477	0.676965	0.653339	0.687380	1.015386	0.00
6	(Gender_Male, Survived_No)	(Age_Adult)	0.619718	0.950477	0.603816	0.974340	1.025106	0.01
7	(Gender_Male, Age_Adult)	(Survived_No)	0.757383	0.676965	0.603816	0.797241	1.177669	0.09
8	(Survived_No, Age_Adult)	(Gender_Male)	0.653339	0.786461	0.603816	0.924200	1.175139	0.08
9	(Gender_Male)	(Survived_No, Age_Adult)	0.786461	0.653339	0.603816	0.767764	1.175139	0.08
10	(Survived_No)	(Gender_Male, Age_Adult)	0.676965	0.757383	0.603816	0.891946	1.177669	0.09
11	(Age_Adult)	(Gender_Male, Survived_No)	0.950477	0.619718	0.603816	0.635277	1.025106	0.01

In [ ]:

```

```