Input Dataset

```
import numpy as np
import pandas as pd
from apyori import apriori
```

```
basket_data = pd.read_csv('market-basket.csv')
basket = pd.DataFrame(basket_data)
```

```
basket
```

	Item(s)	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	...	Item 23	Item 24	Item 25	Item 26	Item 27	Item 28	Item 29	Item 30
0	4	citrus fruit	semi-finished bread	margarine	ready soups	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	3	tropical fruit	yogurt	coffee	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	1	whole milk	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	4	pip fruit	yogurt	cream cheese	meat spreads	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	4	other vegetables	whole milk	condensed milk	long life bakery product	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
9830	17	sausage	chicken	beef	hamburger meat	citrus fruit	grapes	root vegetables	whole milk	butter	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9831	1	cooking chocolate	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9832	10	chicken	citrus fruit	other vegetables	butter	yogurt	frozen dessert	domestic eggs	rolls/buns	rum	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9833	4	semi-finished bread	bottled water	soda	bottled beer	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9834	5	chicken	tropical fruit	other vegetables	vinegar	shopping bags	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

9835 rows × 33 columns

Determine association rules

```
basketContents = []
for i in range(0, 1000):
    basketContents.append([str(basket.values[i,j]) for j in range(1, 33)])
```

```
association_rules = apriori(basketContents, min_support=0.015, min_confidence=0.2, min_lift=3, min_length=2)
association_results = list(association_rules)
```

```
print(len(association_results))
```

```
10
```

Format results for proper display

```
df_results = pd.DataFrame(association_results)
```

```
first_values = []
second_values = []
third_values = []
fourth_value = []

for i in range(df_results.shape[0]):
    single_list = df_results['ordered_statistics'][i][0]
    first_values.append(list(single_list[0]))
    second_values.append(list(single_list[1]))
    third_values.append(single_list[2])
    fourth_value.append(single_list[3])

# support = list(df_results['support'])
support = pd.DataFrame(df_results['support'])
```

```
lhs = pd.DataFrame(first_values)
rhs= pd.DataFrame(second_values)
confidence=pd.DataFrame(third_values,columns=['confidence'])
lift=pd.DataFrame(fourth_value,columns=['lift'])
```

```
df_final = pd.concat([lhs,rhs,support,confidence,lift], axis=1)
```

```
df_final.fillna(value=' ', inplace=True)
df_final.columns = ['lhs',0,1,2,'rhs','support','confidence','lift']
```

```
df_final['lhs'] = df_final['lhs']+str(", ")+df_final[1]+str(", ")+df_final[2]
```

```
df_final.columns = ['items_base','items_add','support','confidence','lift']
```

```
df_final
```

	items_base	items_add	support	confidence	lift
0	berries, whipped/sour cream,		0.017	0.361702	4.887867
1	butter, root vegetables,		0.017	0.369565	3.359684
2	berries, nan, whipped/sour cream		0.017	0.361702	4.887867
3	butter, nan, root vegetables		0.017	0.369565	3.359684
4	curd, whole milk, yogurt		0.016	0.213333	3.809524
5	fruit/vegetable juice, whole milk, yogurt		0.015	0.223881	3.997868
6	whole milk, yogurt,		0.015	0.405405	3.192169
7	curd, nan, whole milk	yogurt	0.016	0.213333	3.809524
8	fruit/vegetable juice, nan, whole milk	yogurt	0.015	0.223881	3.997868
9	whole milk, nan, yogurt		0.015	0.405405	3.192169