

In [1]:

```
import pandas as pd
import numpy as np
import pyfpgrowth
from mlxtend.frequent_patterns import fpgrowth
from mlxtend.frequent_patterns import association_rules
from sklearn.model_selection import train_test_split
```

In [2]:

```
data = pd.read_csv(r"I:\Last Semester\477\mushroom.csv")
from mlxtend.preprocessing import TransactionEncoder
te = TransactionEncoder()
te_ary = te.fit(data).transform(data)
df = pd.DataFrame(te_ary, columns=te.columns_)
df
```

Out[2]:

	0	1	2	3	4	5	6	7	8	9
0	False	True	False	False	False	False	False	False	False	False
1	False	False	False	True	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	True
3	False	True	False	True	False	False	False	False	False	False
4	False	False	True	True	False	False	False	False	False	False
...
8118	False	False	False	False	False	False	False	False	False	False
8119	False	False	False	False	False	False	False	False	False	False
8120	False	False	False	False	False	False	False	False	False	False
8121	False	False	False	False	False	False	False	False	False	False
8122	False	False	False	False	False	False	False	False	False	False

8123 rows × 10 columns

In [3]:

```
data.head(100)
```

Out[3]:

	1	3	9	13	23	25	34	36	38	40	...	63	67	76	85	86	90	93	98	107	113
0	2	3	9	14	23	26	34	36	39	40	...	63	67	76	85	86	90	93	99	108	114
1	2	4	9	15	23	27	34	36	39	41	...	63	67	76	85	86	90	93	99	108	115
2	1	3	10	15	23	25	34	36	38	41	...	63	67	76	85	86	90	93	98	107	113
3	2	3	9	16	24	28	34	37	39	40	...	63	67	76	85	86	90	94	99	109	114
4	2	3	10	14	23	26	34	36	39	41	...	63	67	76	85	86	90	93	98	108	114
...
95	2	6	10	13	23	27	34	36	39	43	...	65	67	76	85	86	90	93	99	107	114
96	2	3	9	14	23	26	34	36	39	41	...	63	67	76	85	86	90	93	98	108	114
97	2	4	9	15	23	26	34	36	39	42	...	63	67	76	85	86	90	93	99	107	114
98	2	3	10	15	23	26	34	36	39	42	...	63	67	76	85	86	90	93	98	107	114
99	2	3	11	13	24	28	34	37	39	43	...	63	67	76	85	86	90	94	98	107	114

100 rows × 23 columns

In [4]:

```
df
```

Out[4]:

	0	1	2	3	4	5	6	7	8	9
0	False	True	False	False	False	False	False	False	False	False
1	False	False	False	True	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	True
3	False	True	False	True	False	False	False	False	False	False
4	False	False	True	True	False	False	False	False	False	False
...
8118	False	False	False	False	False	False	False	False	False	False
8119	False	False	False	False	False	False	False	False	False	False
8120	False	False	False	False	False	False	False	False	False	False
8121	False	False	False	False	False	False	False	False	False	False
8122	False	False	False	False	False	False	False	False	False	False

8123 rows × 10 columns

In [5]:

```
from mlxtend.frequent_patterns import fpgrowth  
fpgrowth(df, min_support=0.0006)
```

Out[5]:

	support	itemsets
0	0.001108	(3)
1	0.000616	(9)
2	0.000616	(5)
3	0.000616	(6)

In [6]:

```
fpgrowth(df, min_support=0.0006, use_colnames=True)
```

Out[6]:

	support	itemsets
0	0.001108	(3)
1	0.000616	(9)
2	0.000616	(5)
3	0.000616	(6)

In [7]:

```
from mlxtend.frequent_patterns import fpgrowth  
%timeit fpgrowth(df, min_support=0.0005)
```

27.6 ms \pm 7.64 ms per loop (mean \pm std. dev. of 7 runs, 10 loops each)

In [8]:

```
from mlxtend.frequent_patterns import fpgrowth  
%timeit fpgrowth(df, min_support=0.0006)
```

23.6 ms \pm 1.44 ms per loop (mean \pm std. dev. of 7 runs, 10 loops each)

In [9]:

```
from mlxtend.frequent_patterns import fpgrowth  
%timeit fpgrowth(df, min_support=0.0007)
```

27 ms \pm 6.48 ms per loop (mean \pm std. dev. of 7 runs, 10 loops each)

In [10]:

```
from mlxtend.frequent_patterns import fpgrowth  
%timeit fpgrowth(df, min_support=0.0008)
```

24.4 ms \pm 5.73 ms per loop (mean \pm std. dev. of 7 runs, 10 loops each)

In []: