```
import pandas as pd
import numpy as np
import random

num_transactions = 10000

items = ['I1', 'I2', 'I3', 'I4', 'I5', 'I6', 'I7', 'I8', 'I9']
data = {'Transaction Id': ['T' + str(i) for i in range(1,
num_transactions + 1)]}
df = pd.DataFrame(data)

def generate_random_itemset():
    num_items = random.randint(3, len(items))
    return sorted(random.sample(items, num_items))

df['Items'] = [generate_random_itemset() for _ in
range(num_transactions)]

print(df)
```

```
     Transaction Id                                 Items
0                T1  [I1, I2, I3, I4, I5, I6, I7, I8, I9]
1                T2                  [I1, I2, I3, I7, I8]
2                T3      [I1, I2, I3, I4, I5, I6, I8, I9]
3                T4      [I1, I2, I4, I5, I6, I7, I8, I9]
4                T5      [I1, I3, I4, I5, I6, I7, I8, I9]
...             ...                                   ...
9995          T9996  [I1, I2, I3, I4, I5, I6, I7, I8, I9]
9996          T9997      [I1, I2, I3, I4, I5, I6, I8, I9]
9997          T9998                  [I2, I3, I5, I6, I9]
9998          T9999                          [I1, I2, I5]
9999         T10000  [I1, I2, I3, I4, I5, I6, I7, I8, I9]

[10000 rows x 2 columns]
```

```
df.to_csv("NIT_Transaction1.csv",index=False)
from itertools import combinations
df.rename(columns={'Transaction
Id':'TId','Items':'Item_Ids'},inplace=True)
msp=1000

flattened_list = [item for sublist in df['Item_Ids'].values.tolist()
for item in sublist]
ids = sorted(list(set(flattened_list)))
itemset=pd.DataFrame(columns=['Itemset','Count'])
for i in range(len(ids)):
    itemset.loc[i] = [ids[i],df['Item_Ids'].apply(lambda x: ids[i] in
x).sum()]
print('C1: ')
print(itemset)
print('L1: ')
itemset[itemset['Count']>=msp]

C1:
  Itemset  Count
0      I1   6623
1      I2   6625
2      I3   6614
3      I4   6738
4      I5   6649
5      I6   6632
6      I7   6673
7      I8   6672
8      I9   6636
L1:

  Itemset  Count
0      I1   6623
1      I2   6625
2      I3   6614
3      I4   6738
4      I5   6649
5      I6   6632
6      I7   6673
7      I8   6672
8      I9   6636

i=2
while True:
    ids=list(combinations(ids, i))
    itemset1=pd.DataFrame(columns=['Itemset','Count'])
    for j in range(len(ids)):
        itemset1.loc[j] = [ids[j],df['Item_Ids'].apply(lambda x:
all(element in x for element in ids[j])).sum()]
    print(f'C{i}\n',itemset1)
    itemset1=itemset1[itemset1['Count']>=2]
    print(f'L{i}\n',itemset1)
    ids=list(set([element for tup in
itemset1['Itemset'].values.tolist() for element in tup]))
    if itemset1.empty:
        break
    itemset=itemset1
    i+=1

C2
      Itemset  Count
0   (I1, I2)   4684
1   (I1, I3)   4655
2   (I1, I4)   4722
3   (I1, I5)   4687
```

```
4   (I1, I6)   4715
5   (I1, I7)   4668
6   (I1, I8)   4731
7   (I1, I9)   4673
8   (I2, I3)   4652
9   (I2, I4)   4733
10  (I2, I5)   4709
11  (I2, I6)   4641
12  (I2, I7)   4717
13  (I2, I8)   4717
14  (I2, I9)   4694
15  (I3, I4)   4748
16  (I3, I5)   4663
17  (I3, I6)   4705
18  (I3, I7)   4695
19  (I3, I8)   4720
20  (I3, I9)   4656
21  (I4, I5)   4759
22  (I4, I6)   4722
23  (I4, I7)   4751
24  (I4, I8)   4782
25  (I4, I9)   4730
26  (I5, I6)   4650
27  (I5, I7)   4705
28  (I5, I8)   4722
29  (I5, I9)   4629
30  (I6, I7)   4696
31  (I6, I8)   4716
32  (I6, I9)   4707
33  (I7, I8)   4710
34  (I7, I9)   4690
35  (I8, I9)   4749
L2
      Itemset  Count
0   (I1, I2)   4684
1   (I1, I3)   4655
2   (I1, I4)   4722
3   (I1, I5)   4687
4   (I1, I6)   4715
5   (I1, I7)   4668
6   (I1, I8)   4731
7   (I1, I9)   4673
8   (I2, I3)   4652
9   (I2, I4)   4733
10  (I2, I5)   4709
11  (I2, I6)   4641
12  (I2, I7)   4717
13  (I2, I8)   4717
14  (I2, I9)   4694
```

```
15  (I3, I4)   4748
16  (I3, I5)   4663
17  (I3, I6)   4705
18  (I3, I7)   4695
19  (I3, I8)   4720
20  (I3, I9)   4656
21  (I4, I5)   4759
22  (I4, I6)   4722
23  (I4, I7)   4751
24  (I4, I8)   4782
25  (I4, I9)   4730
26  (I5, I6)   4650
27  (I5, I7)   4705
28  (I5, I8)   4722
29  (I5, I9)   4629
30  (I6, I7)   4696
31  (I6, I8)   4716
32  (I6, I9)   4707
33  (I7, I8)   4710
34  (I7, I9)   4690
35  (I8, I9)   4749
C3
          Itemset  Count
0   (I6, I9, I3)   3555
1   (I6, I9, I5)   3475
2   (I6, I9, I1)   3595
3   (I6, I9, I4)   3570
4   (I6, I9, I8)   3594
..           ...    ...
79  (I1, I2, I7)   3533
80  (I4, I8, I2)   3602
81  (I4, I8, I7)   3585
82  (I4, I2, I7)   3591
83  (I8, I2, I7)   3573

[84 rows x 2 columns]
L3
          Itemset  Count
0   (I6, I9, I3)   3555
1   (I6, I9, I5)   3475
2   (I6, I9, I1)   3595
3   (I6, I9, I4)   3570
4   (I6, I9, I8)   3594
..           ...    ...
79  (I1, I2, I7)   3533
80  (I4, I8, I2)   3602
81  (I4, I8, I7)   3585
82  (I4, I2, I7)   3591
83  (I8, I2, I7)   3573
```

```
[84 rows x 2 columns]
C4
               Itemset  Count
0    (I6, I9, I3, I5)   2782
1    (I6, I9, I3, I1)   2851
2    (I6, I9, I3, I4)   2856
3    (I6, I9, I3, I8)   2884
4    (I6, I9, I3, I2)   2841
..                ...    ...
121  (I1, I4, I8, I2)   2882
122  (I1, I4, I8, I7)   2858
123  (I1, I4, I2, I7)   2851
124  (I1, I8, I2, I7)   2841
125  (I4, I8, I2, I7)   2879

[126 rows x 2 columns]
L4
               Itemset  Count
0    (I6, I9, I3, I5)   2782
1    (I6, I9, I3, I1)   2851
2    (I6, I9, I3, I4)   2856
3    (I6, I9, I3, I8)   2884
4    (I6, I9, I3, I2)   2841
..                ...    ...
121  (I1, I4, I8, I2)   2882
122  (I1, I4, I8, I7)   2858
123  (I1, I4, I2, I7)   2851
124  (I1, I8, I2, I7)   2841
125  (I4, I8, I2, I7)   2879

[126 rows x 2 columns]
C5
                   Itemset  Count
0    (I6, I9, I3, I5, I1)   2327
1    (I6, I9, I3, I5, I4)   2335
2    (I6, I9, I3, I5, I8)   2368
3    (I6, I9, I3, I5, I2)   2329
4    (I6, I9, I3, I5, I7)   2327
..                    ...    ...
121  (I5, I1, I4, I8, I7)   2368
122  (I5, I1, I4, I2, I7)   2346
123  (I5, I1, I8, I2, I7)   2349
124  (I5, I4, I8, I2, I7)   2395
125  (I1, I4, I8, I2, I7)   2383

[126 rows x 2 columns]
L5
                   Itemset  Count
0    (I6, I9, I3, I5, I1)   2327
```

```
1    (I6, I9, I3, I5, I4)   2335
2    (I6, I9, I3, I5, I8)   2368
3    (I6, I9, I3, I5, I2)   2329
4    (I6, I9, I3, I5, I7)   2327
..                    ...    ...
121  (I5, I1, I4, I8, I7)   2368
122  (I5, I1, I4, I2, I7)   2346
123  (I5, I1, I8, I2, I7)   2349
124  (I5, I4, I8, I2, I7)   2395
125  (I1, I4, I8, I2, I7)   2383

[126 rows x 2 columns]
C6
                      Itemset  Count
0   (I6, I9, I3, I5, I1, I4)   1997
1   (I6, I9, I3, I5, I1, I8)   2035
2   (I6, I9, I3, I5, I1, I2)   1996
3   (I6, I9, I3, I5, I1, I7)   1988
4   (I6, I9, I3, I5, I4, I8)   2045
..                       ...    ...
79  (I3, I5, I1, I4, I2, I7)   1992
80  (I3, I5, I1, I8, I2, I7)   2003
81  (I3, I5, I4, I8, I2, I7)   2038
82  (I3, I1, I4, I8, I2, I7)   2030
83  (I5, I1, I4, I8, I2, I7)   2025

[84 rows x 2 columns]
L6
                      Itemset  Count
0   (I6, I9, I3, I5, I1, I4)   1997
1   (I6, I9, I3, I5, I1, I8)   2035
2   (I6, I9, I3, I5, I1, I2)   1996
3   (I6, I9, I3, I5, I1, I7)   1988
4   (I6, I9, I3, I5, I4, I8)   2045
..                       ...    ...
79  (I3, I5, I1, I4, I2, I7)   1992
80  (I3, I5, I1, I8, I2, I7)   2003
81  (I3, I5, I4, I8, I2, I7)   2038
82  (I3, I1, I4, I8, I2, I7)   2030
83  (I5, I1, I4, I8, I2, I7)   2025

[84 rows x 2 columns]
C7
                          Itemset  Count
0   (I6, I9, I3, I5, I1, I4, I8)   1783
1   (I6, I9, I3, I5, I1, I4, I2)   1757
2   (I6, I9, I3, I5, I1, I4, I7)   1733
3   (I6, I9, I3, I5, I1, I8, I2)   1776
4   (I6, I9, I3, I5, I1, I8, I7)   1775
5   (I6, I9, I3, I5, I1, I2, I7)   1747
```

```
6   (I6, I9, I3, I5, I4, I8, I2)   1808
7   (I6, I9, I3, I5, I4, I8, I7)   1778
8   (I6, I9, I3, I5, I4, I2, I7)   1766
9   (I6, I9, I3, I5, I8, I2, I7)   1795
10  (I6, I9, I3, I1, I4, I8, I2)   1817
11  (I6, I9, I3, I1, I4, I8, I7)   1798
12  (I6, I9, I3, I1, I4, I2, I7)   1790
13  (I6, I9, I3, I1, I8, I2, I7)   1805
14  (I6, I9, I3, I4, I8, I2, I7)   1817
15  (I6, I9, I5, I1, I4, I8, I2)   1787
16  (I6, I9, I5, I1, I4, I8, I7)   1764
17  (I6, I9, I5, I1, I4, I2, I7)   1747
18  (I6, I9, I5, I1, I8, I2, I7)   1778
19  (I6, I9, I5, I4, I8, I2, I7)   1795
20  (I6, I9, I1, I4, I8, I2, I7)   1801
21  (I6, I3, I5, I1, I4, I8, I2)   1779
22  (I6, I3, I5, I1, I4, I8, I7)   1774
23  (I6, I3, I5, I1, I4, I2, I7)   1737
24  (I6, I3, I5, I1, I8, I2, I7)   1757
25  (I6, I3, I5, I4, I8, I2, I7)   1784
26  (I6, I3, I1, I4, I8, I2, I7)   1789
27  (I6, I5, I1, I4, I8, I2, I7)   1763
28  (I9, I3, I5, I1, I4, I8, I2)   1775
29  (I9, I3, I5, I1, I4, I8, I7)   1739
30  (I9, I3, I5, I1, I4, I2, I7)   1731
31  (I9, I3, I5, I1, I8, I2, I7)   1756
32  (I9, I3, I5, I4, I8, I2, I7)   1781
33  (I9, I3, I1, I4, I8, I2, I7)   1775
34  (I9, I5, I1, I4, I8, I2, I7)   1763
35  (I3, I5, I1, I4, I8, I2, I7)   1760
L7
                          Itemset  Count
0   (I6, I9, I3, I5, I1, I4, I8)   1783
1   (I6, I9, I3, I5, I1, I4, I2)   1757
2   (I6, I9, I3, I5, I1, I4, I7)   1733
3   (I6, I9, I3, I5, I1, I8, I2)   1776
4   (I6, I9, I3, I5, I1, I8, I7)   1775
5   (I6, I9, I3, I5, I1, I2, I7)   1747
6   (I6, I9, I3, I5, I4, I8, I2)   1808
7   (I6, I9, I3, I5, I4, I8, I7)   1778
8   (I6, I9, I3, I5, I4, I2, I7)   1766
9   (I6, I9, I3, I5, I8, I2, I7)   1795
10  (I6, I9, I3, I1, I4, I8, I2)   1817
11  (I6, I9, I3, I1, I4, I8, I7)   1798
12  (I6, I9, I3, I1, I4, I2, I7)   1790
13  (I6, I9, I3, I1, I8, I2, I7)   1805
14  (I6, I9, I3, I4, I8, I2, I7)   1817
15  (I6, I9, I5, I1, I4, I8, I2)   1787
16  (I6, I9, I5, I1, I4, I8, I7)   1764
```

```
17  (I6, I9, I5, I1, I4, I2, I7)   1747
18  (I6, I9, I5, I1, I8, I2, I7)   1778
19  (I6, I9, I5, I4, I8, I2, I7)   1795
20  (I6, I9, I1, I4, I8, I2, I7)   1801
21  (I6, I3, I5, I1, I4, I8, I2)   1779
22  (I6, I3, I5, I1, I4, I8, I7)   1774
23  (I6, I3, I5, I1, I4, I2, I7)   1737
24  (I6, I3, I5, I1, I8, I2, I7)   1757
25  (I6, I3, I5, I4, I8, I2, I7)   1784
26  (I6, I3, I1, I4, I8, I2, I7)   1789
27  (I6, I5, I1, I4, I8, I2, I7)   1763
28  (I9, I3, I5, I1, I4, I8, I2)   1775
29  (I9, I3, I5, I1, I4, I8, I7)   1739
30  (I9, I3, I5, I1, I4, I2, I7)   1731
31  (I9, I3, I5, I1, I8, I2, I7)   1756
32  (I9, I3, I5, I4, I8, I2, I7)   1781
33  (I9, I3, I1, I4, I8, I2, I7)   1775
34  (I9, I5, I1, I4, I8, I2, I7)   1763
35  (I3, I5, I1, I4, I8, I2, I7)   1760
C8
                             Itemset  Count
0  (I6, I9, I3, I5, I1, I4, I8, I2)   1589
1  (I6, I9, I3, I5, I1, I4, I8, I7)   1566
2  (I6, I9, I3, I5, I1, I4, I2, I7)   1554
3  (I6, I9, I3, I5, I1, I8, I2, I7)   1578
4  (I6, I9, I3, I5, I4, I8, I2, I7)   1600
5  (I6, I9, I3, I1, I4, I8, I2, I7)   1606
6  (I6, I9, I5, I1, I4, I8, I2, I7)   1574
7  (I6, I3, I5, I1, I4, I8, I2, I7)   1567
8  (I9, I3, I5, I1, I4, I8, I2, I7)   1558
L8
                             Itemset  Count
0  (I6, I9, I3, I5, I1, I4, I8, I2)   1589
1  (I6, I9, I3, I5, I1, I4, I8, I7)   1566
2  (I6, I9, I3, I5, I1, I4, I2, I7)   1554
3  (I6, I9, I3, I5, I1, I8, I2, I7)   1578
4  (I6, I9, I3, I5, I4, I8, I2, I7)   1600
5  (I6, I9, I3, I1, I4, I8, I2, I7)   1606
6  (I6, I9, I5, I1, I4, I8, I2, I7)   1574
7  (I6, I3, I5, I1, I4, I8, I2, I7)   1567
8  (I9, I3, I5, I1, I4, I8, I2, I7)   1558
C9
                                 Itemset  Count
0  (I6, I9, I3, I5, I1, I4, I8, I2, I7)   1419
L9
                                 Itemset  Count
0  (I6, I9, I3, I5, I1, I4, I8, I2, I7)   1419
C10
 Empty DataFrame
```

```
Columns: [Itemset, Count]
Index: []
L10
 Empty DataFrame
Columns: [Itemset, Count]
Index: []
```

```
implied_combinations = []
for tup in itemset['Itemset'].values.tolist():
    for a, b in combinations(tup, 2):
        for c in tup:
            if c != a and c != b:
                implied_combinations.append((a, b, c))
association_df=pd.DataFrame(columns=['Rule','Confidence'])
for i in range(len(implied_combinations)):
    tuple=implied_combinations[i]
    association_df.loc[len(association_df.index)]=[f'{tuple[:2]} ->
{tuple[-1]}',
    round(df['Item_Ids'].apply(lambda x: all(element in x for element
in tuple)).sum()*100/df['Item_Ids'].apply(lambda x: all(element in x
for element in tuple[:2])).sum(),2)]
    association_df.loc[len(association_df.index)]=[f'{tuple[-1]} ->
{tuple[:2]}',
    round(df['Item_Ids'].apply(lambda x: all(element in x for element
in tuple)).sum()*100/df['Item_Ids'].apply(lambda x: tuple[-1] in
x).sum(),2)]
print("All Possible Association rules: ")
association_df
```

```
All Possible Association rules:

                   Rule  Confidence
0    ('I6', 'I9') -> I3       75.53
1    I3 -> ('I6', 'I9')       53.75
2    ('I6', 'I9') -> I5       73.83
3    I5 -> ('I6', 'I9')       52.26
4    ('I6', 'I9') -> I1       76.38
..                  ...         ...
499  I1 -> ('I2', 'I7')       53.34
500  ('I2', 'I7') -> I4       76.13
501  I4 -> ('I2', 'I7')       53.29
502  ('I2', 'I7') -> I8       75.75
503  I8 -> ('I2', 'I7')       53.55

[504 rows x 2 columns]
```

```
print(association_df)
association_df.to_csv("Allassocitation.csv",index=False)
```

```
                   Rule  Confidence
0    ('I6', 'I9') -> I3       75.53
1    I3 -> ('I6', 'I9')       53.75
2    ('I6', 'I9') -> I5       73.83
3    I5 -> ('I6', 'I9')       52.26
4    ('I6', 'I9') -> I1       76.38
..                  ...         ...
499  I1 -> ('I2', 'I7')       53.34
500  ('I2', 'I7') -> I4       76.13
501  I4 -> ('I2', 'I7')       53.29
502  ('I2', 'I7') -> I8       75.75
503  I8 -> ('I2', 'I7')       53.55

[504 rows x 2 columns]
```

```
print("Final Association Rules: ")
print(association_df[association_df['Confidence']>70])
association_df[association_df['Confidence']>70].to_csv('AcceptedAssoci
ated.csv',index=False)
```

```
Final Association Rules:
                   Rule  Confidence
0    ('I6', 'I9') -> I3       75.53
2    ('I6', 'I9') -> I5       73.83
4    ('I6', 'I9') -> I1       76.38
6    ('I6', 'I9') -> I4       75.84
8    ('I6', 'I9') -> I8       76.35
..                  ...         ...
494  ('I2', 'I7') -> I3       74.79
496  ('I2', 'I7') -> I5       75.20
498  ('I2', 'I7') -> I1       74.90
500  ('I2', 'I7') -> I4       76.13
502  ('I2', 'I7') -> I8       75.75

[252 rows x 2 columns]
```