

PROJECT REPORT

APRIORI ALGORITHM



Course: Data Mining

Prepared By: Vinay Basavaraja

GitHub link:

Introduction

Apriori Algorithm is the classical algorithm which is used in data mining. It is used for mining frequent item sets and relevant association rules for those item sets. It is devised to operate on a database containing more transactions, for instance, items brought by customers in a store.

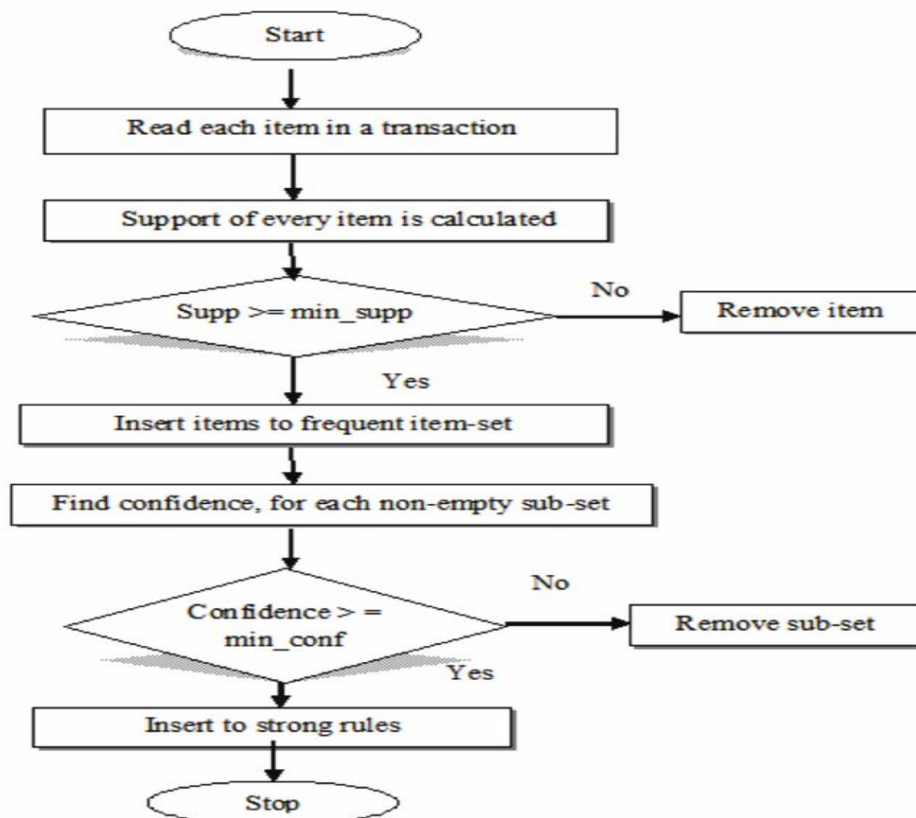
It is important for effective Market Basket Analysis and it helps the users in purchasing their products with more ease which provides more sales in the markets. It has been even used in the healthcare industry for the finding of adverse drug reactions. It produces association rules that shows what all combinations of medications and patient characteristics lead to ADRs.

Implementation

An implementation of Apriori Algorithm to generate Association rules with user input values for support and confidence. The support and confidence values are given in decimal values. It gives associations rules for different datasets: Amazon, BestBuy, KMART, Nike, Grocery.

High values of support and confidence gives an out of index value. Select a numerical value (input) for obtaining the respective data set's association rules. The csv files are uploaded too. The screenshots are attached in the file that is submitted. Check the file for report and for basic software and hardware used to implement the project. the Project is not software or hardware specific.

Implementation Architecture



Software Configuration

1. Environment: Anaconda-Navigator
2. IDE: Jupiter Notebook 6.1.4
3. Python version: 3.8
4. Libraries csv (for reader)

Source of Data:

The data used are the ones provided in the dataset examples in the midterm folder. I have also created a custom dataset for grocery store. All the data from these files are written into csv files. The datasets used are:

1. Amazon
2. BestBuy
3. KMART
4. Nike
5. Custom/Grocery

Hardware Configuration

1. Laptop: Apple MacBook Pro
2. Processor: M1 / Apple Silicon chip

Code

Coding is complete and these are the few functions used in the code

1. `append_data` : The purpose is to load the data set and return the dataset that contains a list of transactions and further, each transaction contains several items in it.
2. `gen_Candidate1`: Generate and return a set(dataset) which contains all frequent candidate 1-itemsets.
3. `apriori` : To check whether a frequent candidate k-itemset satisfy Apriori property and return true/false accordingly.
4. `gen_Cank` : Generate and return candidate k, a set which contains all frequent candidate k-itemsets.
5. `gen_Lisk` : Generate and return a list/set which contains all frequent k-itemsets.
6. `gen_L` : Generate all frequent itemsets by the value of minimum support's input.
7. `association_rules`: Generate and return association rules.
8. `Main` : Execution of functions/code and take user input for minimum support and minimum confidence.
9. Encoding used : windows-1252

Difficulties faced:

To implement the understood logic of the functions as this was my first python project.

Screenshots:

Screenshots are attached for different dataset and different user input values for support and confidence.

CSV files.:

CSV files are attached in the zip file.

GitHub Link:

Source Code:

```
from csv import reader
```

```
def append_data():
```

```
    #The purpose is to load the data set and return the dataset that contains a list of transactions  
    and further, each transaction contains several items in it.
```

```
    print("Hello, welcome to association rule mining\nPlease choose the dataset for which you  
    would like to get the Association rules")
```

```
    print("Please type \n1 for Amazon\n2 for BestBuy\n3 for KMART\n4 for Nike\n5 for  
    Custom")
```

```
    while True:
```

```
        datachoice = input()
```

```
        if (datachoice == '1'):
```

```
            data = '/Users/VINAY/midterm_datamining_vinay/amazon.csv'
```

```
            print('You selected Amazon Dataset')
```

```
            break
```

```
        elif(datachoice == '2'):
```

```
            data = '/Users/VINAY/midterm_datamining_vinay/bestbuy.csv'
```

```
            print('You selected BestBuy Dataset')
```

```
            break
```

```
        elif(datachoice == '3'):
```

```
            data = '/Users/VINAY/midterm_datamining_vinay/kmart.csv'
```

```
            print('You selected KMART Dataset')
```

```
            break
```

```
        elif(datachoice == '4'):
```

```
            data = '/Users/VINAY/midterm_datamining_vinay/nike.csv'
```

```
            print('You selected Nike Dataset')
```

```
            break
```

```
        elif(datachoice == '5'):
```

```
            data = '/Users/VINAY/midterm_datamining_vinay/grocery.csv'
```

```

        print('You selected Custom Dataset')
        break
    else:
        print("Please select a valid choice from the above list")
with open(data, 'r',encoding="windows-1252") as read_obj:
    # pass the file object to reader() to get the reader object
    csv_reader = reader(read_obj)

    # Pass reader object to list() to get a list of lists
    dataset = list(csv_reader)
    for items in dataset:
        for j in range(0, len(items)):
            for items1 in items:

                if items1 == "":
                    items.remove("")

    return dataset

def gen_Candidate1(dataset):
    # Generate and return a set(dataset) which contains all frequent candidate 1-itemsets.

    Can1 = set()
    for t in dataset:
        for item in t:
            item_set = frozenset([item])
            Can1.add(item_set)
    return Can1

```

```
def apriori(Cank_item, Lisk1):
```

```
    #To check whether a frequent candidate k-itemset satisfy Apriori property and return  
    true/false accordingly.
```

```
    for item in Cank_item:
```

```
        sub_Cank = Cank_item - frozenset([item])
```

```
        if sub_Cank not in Lisk1:
```

```
            return False
```

```
    return True
```

```
def gen_Cank(Lisk1, k):
```

```
    #Generate and return candidate k, a set which contains all frequent candidate k-itemsets.
```

```
    Cank = set()
```

```
    length_Lisk1 = len(Lisk1)
```

```
    listof_Lisk1 = list(Lisk1)
```

```
    for i in range(length_Lisk1):
```

```
        for j in range(1, length_Lisk1):
```

```
            list1 = list(listof_Lisk1[i])
```

```
            list2 = list(listof_Lisk1[j])
```

```
            list1.sort()
```

```
            list2.sort()
```

```
            if list1[0:k-2] == list2[0:k-2]:
```

```
                Cank_item = listof_Lisk1[i] | listof_Lisk1[j]
```

```
                if apriori(Cank_item, Lisk1):
```

```
                    Cank.add(Cank_item)
```

```
    return Cank
```

```

def gen_Lisk(dataset, Cank, minsup, sup_data):
    #Generate and return a list/set which contains all frequent k-itemsets.

    Lisk = set()
    item_count = {}
    for t in dataset:
        for item in Cank:
            if item.issubset(t):
                if item not in item_count:
                    item_count[item] = 1
                else:
                    item_count[item] += 1
    t_num = float(len(dataset))
    for item in item_count:
        if (item_count[item] / t_num) >= minsup:
            Lisk.add(item)
            sup_data[item] = item_count[item] / t_num
    return Lisk

```

```

def gen_L(dataset, k, minsup):
    #Generate all frequent itemsets by the value of minimum support's input.

    sup_data = {}
    Can1 = gen_Candidate1(dataset)
    List1 = gen_Lisk(dataset, Can1, minsup, sup_data)
    Lisksub1 = List1.copy()
    List = []
    List.append(Lisksub1)
    for i in range(2, k+1):

```



```

Ci = gen_Cank(Lisksub1, i)
Li = gen_Lisk(dataset, Ci, minsup, sup_data)
Lisksub1 = Li.copy()
List.append(Lisksub1)
return List, sup_data

```

```

def association_rules(List, sup_data, minconf):

```

```

    #Generate and return association rules.

```

```

    asso_rules = []

```

```

    sublist = []

```

```

    for i in range(0, len(List)):

```

```

        for freq_set in List[i]:

```

```

            for sub_set in sublist:

```

```

                if sub_set.issubset(freq_set):

```

```

                    conf = sup_data[freq_set] / sup_data[freq_set - sub_set]

```

```

                    rule = (freq_set - sub_set, sub_set, conf)

```

```

                    if conf >= minconf and rule not in asso_rules:

```

```

                        # print freq_set-sub_set, " => ", sub_set, "conf: ", conf

```

```

                        asso_rules.append(rule)

```

```

                    sublist.append(freq_set)

```

```

    return asso_rules

```

```

if __name__ == "__main__":

```

```

    #print("Please enter the min_support")

```

```

    #inp = input("Enter the minimum support: ")

```



```
Command Prompt
frequent 1-itemsets      support
-----
frozenset({'Java For Dummies'}) 0.6
frozenset({'Android Programming: The Big Nerd Ranch'}) 0.65
frozenset({'Java: The Complete Reference'}) 0.45
frozenset({'Java 8 Pocket Guide'}) 0.15
frozenset({'Head First Java 2nd Edition'}) 0.4
frozenset({'HTML and CSS: Design and Build Websites'}) 0.1
frozenset({'Beginning Programming with Java'}) 0.3
frozenset({'A Beginner's Guide'}) 0.55
frequent 2-itemsets      support
-----
frozenset({'A Beginner's Guide', 'Android Programming: The Big Nerd Ranch'}) 0.3
frozenset({'Java For Dummies', 'Head First Java 2nd Edition'}) 0.15
frozenset({'Beginning Programming with Java', 'Android Programming: The Big Nerd Ranch'}) 0.15
frozenset({'Java: The Complete Reference', 'A Beginner's Guide'}) 0.4
frozenset({'Java For Dummies', 'HTML and CSS: Design and Build Websites'}) 0.1
frozenset({'Head First Java 2nd Edition', 'Android Programming: The Big Nerd Ranch'}) 0.3
frozenset({'Java: The Complete Reference', 'Android Programming: The Big Nerd Ranch'}) 0.3
frozenset({'Java For Dummies', 'A Beginner's Guide'}) 0.4
frozenset({'HTML and CSS: Design and Build Websites', 'Java: The Complete Reference'}) 0.1
frozenset({'Beginning Programming with Java', 'Java 8 Pocket Guide'}) 0.1
frozenset({'Beginning Programming with Java', 'Head First Java 2nd Edition'}) 0.2
frozenset({'Java For Dummies', 'Android Programming: The Big Nerd Ranch'}) 0.45
frozenset({'Java For Dummies', 'Java: The Complete Reference'}) 0.45
frozenset({'HTML and CSS: Design and Build Websites', 'A Beginner's Guide'}) 0.1
frozenset({'Head First Java 2nd Edition', 'A Beginner's Guide'}) 0.15
frequent 3-itemsets      support
-----
frozenset({'Java For Dummies', 'Java: The Complete Reference', 'A Beginner's Guide'}) 0.4
frozenset({'Java For Dummies', 'A Beginner's Guide', 'Android Programming: The Big Nerd Ranch'}) 0.25
frozenset({'Java For Dummies', 'Head First Java 2nd Edition', 'Android Programming: The Big Nerd Ranch'}) 0.15
frozenset({'Java: The Complete Reference', 'A Beginner's Guide', 'Android Programming: The Big Nerd Ranch'}) 0.25
frozenset({'Java For Dummies', 'Java: The Complete Reference', 'Android Programming: The Big Nerd Ranch'}) 0.3
frozenset({'Java For Dummies', 'HTML and CSS: Design and Build Websites', 'A Beginner's Guide'}) 0.1
frozenset({'Java For Dummies', 'HTML and CSS: Design and Build Websites', 'Java: The Complete Reference'}) 0.1
frozenset({'HTML and CSS: Design and Build Websites', 'Java: The Complete Reference', 'A Beginner's Guide'}) 0.1
frozenset({'Head First Java 2nd Edition', 'A Beginner's Guide', 'Android Programming: The Big Nerd Ranch'}) 0.1
```

```
Command Prompt
Association Rules with Confidence
-----
frozenset({'A Beginner's Guide'}) => frozenset({'Android Programming: The Big Nerd Ranch'}) confidence: 0.5454545454545454
frozenset({'Android Programming: The Big Nerd Ranch'}) => frozenset({'A Beginner's Guide'}) confidence: 0.4615384615384615
frozenset({'Head First Java 2nd Edition'}) => frozenset({'Java For Dummies'}) confidence: 0.37499999999999994
frozenset({'Java For Dummies'}) => frozenset({'Head First Java 2nd Edition'}) confidence: 0.25
frozenset({'Beginning Programming with Java'}) => frozenset({'Android Programming: The Big Nerd Ranch'}) confidence: 0.5
frozenset({'Android Programming: The Big Nerd Ranch'}) => frozenset({'Beginning Programming with Java'}) confidence: 0.23076923076923075
frozenset({'A Beginner's Guide'}) => frozenset({'Java: The Complete Reference'}) confidence: 0.7272727272727273
frozenset({'Java: The Complete Reference'}) => frozenset({'A Beginner's Guide'}) confidence: 0.8888888888888889
frozenset({'HTML and CSS: Design and Build Websites'}) => frozenset({'Java For Dummies'}) confidence: 1.0
frozenset({'Head First Java 2nd Edition'}) => frozenset({'Android Programming: The Big Nerd Ranch'}) confidence: 0.7499999999999999
frozenset({'Android Programming: The Big Nerd Ranch'}) => frozenset({'Head First Java 2nd Edition'}) confidence: 0.4615384615384615
frozenset({'Java: The Complete Reference'}) => frozenset({'Android Programming: The Big Nerd Ranch'}) confidence: 0.6666666666666666
frozenset({'Android Programming: The Big Nerd Ranch'}) => frozenset({'Java: The Complete Reference'}) confidence: 0.4615384615384615
frozenset({'A Beginner's Guide'}) => frozenset({'Java For Dummies'}) confidence: 0.7272727272727273
frozenset({'Java For Dummies'}) => frozenset({'A Beginner's Guide'}) confidence: 0.6666666666666667
frozenset({'HTML and CSS: Design and Build Websites'}) => frozenset({'Java: The Complete Reference'}) confidence: 1.0
frozenset({'Java: The Complete Reference'}) => frozenset({'HTML and CSS: Design and Build Websites'}) confidence: 0.22222222222222224
frozenset({'Beginning Programming with Java'}) => frozenset({'Java 8 Pocket Guide'}) confidence: 0.3333333333333333
frozenset({'Java 8 Pocket Guide'}) => frozenset({'Beginning Programming with Java'}) confidence: 0.6666666666666667
frozenset({'Beginning Programming with Java'}) => frozenset({'Head First Java 2nd Edition'}) confidence: 0.6666666666666667
frozenset({'Head First Java 2nd Edition'}) => frozenset({'Beginning Programming with Java'}) confidence: 0.5
frozenset({'Android Programming: The Big Nerd Ranch'}) => frozenset({'Java For Dummies'}) confidence: 0.6923076923076923
frozenset({'Java For Dummies'}) => frozenset({'Android Programming: The Big Nerd Ranch'}) confidence: 0.75
frozenset({'Java: The Complete Reference'}) => frozenset({'Java For Dummies'}) confidence: 1.0
frozenset({'Java For Dummies'}) => frozenset({'Java: The Complete Reference'}) confidence: 0.75
frozenset({'HTML and CSS: Design and Build Websites'}) => frozenset({'A Beginner's Guide'}) confidence: 1.0
frozenset({'A Beginner's Guide'}) => frozenset({'Head First Java 2nd Edition'}) confidence: 0.7272727272727273
frozenset({'Head First Java 2nd Edition'}) => frozenset({'A Beginner's Guide'}) confidence: 0.37499999999999994
frozenset({'Java: The Complete Reference', 'A Beginner's Guide'}) => frozenset({'Java For Dummies'}) confidence: 1.0
frozenset({'Java For Dummies', 'A Beginner's Guide'}) => frozenset({'Java: The Complete Reference'}) confidence: 1.0
frozenset({'Java For Dummies', 'Java: The Complete Reference'}) => frozenset({'A Beginner's Guide'}) confidence: 0.8888888888888889
frozenset({'Java For Dummies'}) => frozenset({'Java: The Complete Reference', 'A Beginner's Guide'}) confidence: 0.6666666666666667
frozenset({'Java: The Complete Reference'}) => frozenset({'Java For Dummies', 'A Beginner's Guide'}) confidence: 0.8888888888888889
frozenset({'A Beginner's Guide'}) => frozenset({'Java For Dummies', 'Java: The Complete Reference'}) confidence: 0.7272727272727273
frozenset({'A Beginner's Guide', 'Android Programming: The Big Nerd Ranch'}) => frozenset({'Java For Dummies'}) confidence: 0.8333333333333333
frozenset({'Java For Dummies', 'A Beginner's Guide'}) => frozenset({'Android Programming: The Big Nerd Ranch'}) confidence: 0.625
frozenset({'Java For Dummies', 'Android Programming: The Big Nerd Ranch'}) => frozenset({'A Beginner's Guide'}) confidence: 0.5555555555555556
frozenset({'Java For Dummies'}) => frozenset({'A Beginner's Guide', 'Android Programming: The Big Nerd Ranch'}) confidence: 0.4166666666666667
```

2.BestBuy

```
Command Prompt
C:\Users\VINAY\midterm_datamining_vinay>python midterm_apriori.py
Hello, welcome to association rule mining
Please choose the dataset for which you would like to get the Association rules
Please type
1 for Amazon
2 for BestBuy
3 for KMART
4 for Nike
5 for Custom
2
You selected BestBuy Dataset
Enter the minimum support: 0.15
Enter the minimum Confidence: 0.2
*****
frequent 1-itemsets      support
*****
frozenset({'External Hard-Drive'}) 0.45
frozenset({'Anti-Virus'}) 0.7
frozenset({'Microsoft Office'}) 0.55
frozenset({'Speakers'}) 0.55
frozenset({'Lab Top Case'}) 0.7
frozenset({'Digital Camera'}) 0.45
frozenset({'Lab Top'}) 0.6
frozenset({'Flash Drive'}) 0.65
frozenset({'Desk Top'}) 0.3
frozenset({'Printer'}) 0.5
*****
frequent 2-itemsets      support
*****
frozenset({'Flash Drive', 'Speakers'}) 0.3
frozenset({'Speakers', 'Printer'}) 0.25
frozenset({'External Hard-Drive', 'Anti-Virus'}) 0.45
frozenset({'Lab Top', 'Speakers'}) 0.25
frozenset({'Desk Top', 'Microsoft Office'}) 0.25
frozenset({'Desk Top', 'Anti-Virus'}) 0.2
frozenset({'Desk Top', 'External Hard-Drive'}) 0.15
frozenset({'External Hard-Drive', 'Digital Camera'}) 0.15
frozenset({'Lab Top Case', 'Flash Drive'}) 0.45
frozenset({'Microsoft Office', 'Flash Drive'}) 0.55
frozenset({'Microsoft Office', 'Printer'}) 0.45
```

```
Command Prompt
frozenset({'Anti-Virus', 'Flash Drive'}) 0.5
frozenset({'Desk Top', 'Speakers'}) 0.2
frozenset({'Speakers', 'Digital Camera'}) 0.35
frozenset({'Anti-Virus', 'Printer'}) 0.35
frozenset({'Lab Top', 'Digital camera'}) 0.25
*****
frequent 3-itemsets      support
*****
frozenset({'Lab Top Case', 'Flash Drive', 'Speakers'}) 0.25
frozenset({'Lab Top Case', 'Speakers', 'Printer'}) 0.2
frozenset({'Desk Top', 'Microsoft Office', 'Speakers'}) 0.15
frozenset({'Desk Top', 'Lab Top Case', 'Speakers'}) 0.15
frozenset({'Lab Top Case', 'Speakers', 'Digital Camera'}) 0.3
frozenset({'Anti-Virus', 'Flash Drive', 'Lab Top'}) 0.3
frozenset({'Anti-Virus', 'Lab Top', 'Printer'}) 0.2
frozenset({'External Hard-Drive', 'Speakers', 'Printer'}) 0.15
frozenset({'External Hard-Drive', 'Flash Drive', 'Speakers'}) 0.2
frozenset({'Anti-Virus', 'Speakers', 'Printer'}) 0.25
frozenset({'Anti-Virus', 'Flash Drive', 'Speakers'}) 0.3
frozenset({'External Hard-Drive', 'Lab Top', 'Printer'}) 0.15
frozenset({'External Hard-Drive', 'Flash Drive', 'Lab Top'}) 0.15
frozenset({'Flash Drive', 'Speakers', 'Digital Camera'}) 0.15
frozenset({'Anti-Virus', 'Lab Top Case', 'Lab Top'}) 0.45
frozenset({'Microsoft Office', 'Anti-Virus', 'Lab Top'}) 0.2
frozenset({'Lab Top Case', 'Flash Drive', 'Digital Camera'}) 0.15
frozenset({'External Hard-Drive', 'Speakers', 'Lab Top Case'}) 0.3
frozenset({'External Hard-Drive', 'Lab Top Case', 'Lab Top'}) 0.25
frozenset({'Anti-Virus', 'Speakers', 'Lab Top Case'}) 0.4
frozenset({'Microsoft Office', 'External Hard-Drive', 'Speakers'}) 0.2
frozenset({'Lab Top', 'Flash Drive', 'Printer'}) 0.25
frozenset({'Microsoft Office', 'Lab Top', 'Lab Top Case'}) 0.2
frozenset({'External Hard-Drive', 'Anti-Virus', 'Speakers'}) 0.3
frozenset({'Printer', 'Flash Drive', 'Digital Camera'}) 0.15
frozenset({'Desk Top', 'External Hard-Drive', 'Anti-Virus'}) 0.15
frozenset({'Anti-Virus', 'Speakers', 'Digital Camera'}) 0.25
frozenset({'Microsoft Office', 'External Hard-Drive', 'Lab Top Case'}) 0.25
frozenset({'External Hard-Drive', 'Flash Drive', 'Printer'}) 0.25
frozenset({'Anti-Virus', 'Speakers', 'Lab Top'}) 0.25
frozenset({'Lab Top Case', 'Flash Drive', 'Printer'}) 0.3
frozenset({'Lab Top Case', 'Flash Drive', 'Lab Top'}) 0.3
frozenset({'Microsoft Office', 'Lab Top', 'Flash Drive'}) 0.25
```

```
Command Prompt
*****
Association Rules with Confidence
*****
frozenset({'Flash Drive'}) => frozenset({'Speakers'}) confidence: 0.4615384615384615
frozenset({'Speakers'}) => frozenset({'Flash Drive'}) confidence: 0.5454545454545454
frozenset({'Printer'}) => frozenset({'Speakers'}) confidence: 0.5
frozenset({'Speakers'}) => frozenset({'Printer'}) confidence: 0.4545454545454545
frozenset({'Anti-Virus'}) => frozenset({'External Hard-Drive'}) confidence: 0.6428571428571429
frozenset({'External Hard-Drive'}) => frozenset({'Anti-Virus'}) confidence: 1.0
frozenset({'Lab Top'}) => frozenset({'Speakers'}) confidence: 0.4166666666666667
frozenset({'Speakers'}) => frozenset({'Lab Top'}) confidence: 0.4545454545454545
frozenset({'Desk Top'}) => frozenset({'Microsoft Office'}) confidence: 0.8333333333333334
frozenset({'Microsoft Office'}) => frozenset({'Desk Top'}) confidence: 0.4545454545454545
frozenset({'Desk Top'}) => frozenset({'Anti-Virus'}) confidence: 0.6666666666666667
frozenset({'Anti-Virus'}) => frozenset({'Desk Top'}) confidence: 0.28571428571428575
frozenset({'Desk Top'}) => frozenset({'External Hard-Drive'}) confidence: 0.5
frozenset({'External Hard-Drive'}) => frozenset({'Desk Top'}) confidence: 0.3333333333333333
frozenset({'Digital Camera'}) => frozenset({'External Hard-Drive'}) confidence: 0.3333333333333333
frozenset({'External Hard-Drive'}) => frozenset({'Digital Camera'}) confidence: 0.3333333333333333
frozenset({'Flash Drive'}) => frozenset({'Lab Top Case'}) confidence: 0.6923076923076923
frozenset({'Lab Top Case'}) => frozenset({'Flash Drive'}) confidence: 0.6428571428571429
frozenset({'Flash Drive'}) => frozenset({'Microsoft Office'}) confidence: 0.8461538461538461
frozenset({'Microsoft Office'}) => frozenset({'Flash Drive'}) confidence: 1.0
frozenset({'Printer'}) => frozenset({'Microsoft Office'}) confidence: 0.9
frozenset({'Microsoft Office'}) => frozenset({'Printer'}) confidence: 0.8181818181818181
frozenset({'Printer'}) => frozenset({'Lab Top Case'}) confidence: 0.6
frozenset({'Lab Top Case'}) => frozenset({'Printer'}) confidence: 0.4285714285714286
frozenset({'Lab Top'}) => frozenset({'Microsoft Office'}) confidence: 0.4166666666666667
frozenset({'Microsoft Office'}) => frozenset({'Lab Top'}) confidence: 0.4545454545454545
frozenset({'Desk Top'}) => frozenset({'Lab Top Case'}) confidence: 0.5
frozenset({'Lab Top Case'}) => frozenset({'Desk Top'}) confidence: 0.2142857142857143
frozenset({'Lab Top'}) => frozenset({'Anti-Virus'}) confidence: 0.8333333333333334
frozenset({'Anti-Virus'}) => frozenset({'Lab Top'}) confidence: 0.7142857142857143
frozenset({'Speakers'}) => frozenset({'External Hard-Drive'}) confidence: 0.5454545454545454
frozenset({'External Hard-Drive'}) => frozenset({'Speakers'}) confidence: 0.6666666666666667
frozenset({'Flash Drive'}) => frozenset({'Digital Camera'}) confidence: 0.3076923076923077
frozenset({'Digital Camera'}) => frozenset({'Flash Drive'}) confidence: 0.4444444444444444
frozenset({'Printer'}) => frozenset({'Digital Camera'}) confidence: 0.3
frozenset({'Digital Camera'}) => frozenset({'Printer'}) confidence: 0.3333333333333333
frozenset({'Speakers'}) => frozenset({'Anti-Virus'}) confidence: 0.8181818181818181
```


3.KMART

```
Select Command Prompt
Hello, welcome to association rule mining
Please choose the dataset for which you would like to get the Association rules
Please type
1 for Amazon
2 for BestBuy
3 for KMART
4 for Nike
5 for Custom
3
You selected KMART Dataset
Enter the minimum support: 0.3
Enter the minimum Confidence: 0.4
*****
frequent 1-itemsets      support
*****
frozenset({'Kids Bedding \n'}) 0.55
frozenset({'Bed Skirtsâ€'}) 0.55
frozenset({'Shamsâ€'}) 0.5
frozenset({'Bedding Collections'}) 0.35
frozenset({'Embroidered Bedspread'}) 0.3
frozenset({'Bedspreads'}) 0.35
frozenset({'Quilts'}) 0.3
frozenset({'Decorative Pillows'}) 0.45
*****
frequent 2-itemsets      support
*****
frozenset({'Bedspreads', 'Bed Skirtsâ€'}) 0.35
frozenset({'Shamsâ€', 'Bed Skirtsâ€'}) 0.4
frozenset({'Kids Bedding \n', 'Shamsâ€'}) 0.4
frozenset({'Bedspreads', 'Kids Bedding \n'}) 0.3
```

```
Select Command Prompt
*****
frequent 3-itemsets      support
*****
frozenset({'Socks', 'Sweatshirts', 'Modern Pants'}) 0.4
frozenset({'Hoodies', 'Tech Pants', 'Dry Fit V-Nick'}) 0.35
frozenset({'Swimming Shirt', 'Tech Pants', 'Dry Fit V-Nick'}) 0.35
frozenset({'Swimming Shirt', 'Sweatshirts', 'Modern Pants'}) 0.35
frozenset({'Swimming Shirt', 'Rash Guard', 'Dry Fit V-Nick'}) 0.45
frozenset({'Tech Pants', 'Rash Guard', 'Dry Fit V-Nick'}) 0.4
frozenset({'Running Shoe', 'Socks', 'Modern Pants'}) 0.4
frozenset({'Running Shoe', 'Socks', 'Sweatshirts'}) 0.45
frozenset({'Running Shoe', 'Socks', 'Sweatshirts'}) 0.5
frozenset({'Hoodies', 'Rash Guard', 'Tech Pants'}) 0.4
frozenset({'Hoodies', 'Rash Guard', 'Dry Fit V-Nick'}) 0.35
*****
Association Rules with Confidence
*****
frozenset({'Dry Fit V-Nick'}) => frozenset({'Hoodies'}) confidence: 0.7
frozenset({'Hoodies'}) => frozenset({'Dry Fit V-Nick'}) confidence: 0.8749999999999999
frozenset({'Swimming Shirt'}) => frozenset({'Swimming Shirt'}) confidence: 0.9
frozenset({'Swimming Shirt'}) => frozenset({'Dry Fit V-Nick'}) confidence: 0.8181818181818181
frozenset({'Rash Guard'}) => frozenset({'Running Shoe'}) confidence: 0.5833333333333333
frozenset({'Running Shoe'}) => frozenset({'Rash Guard'}) confidence: 0.5
frozenset({'Sweatshirts'}) => frozenset({'Modern Pants'}) confidence: 0.7692307692307692
frozenset({'Modern Pants'}) => frozenset({'Sweatshirts'}) confidence: 1.0
frozenset({'Sweatshirts'}) => frozenset({'Socks'}) confidence: 0.8461538461538461
frozenset({'Socks'}) => frozenset({'Sweatshirts'}) confidence: 0.9166666666666667
frozenset({'Dry Fit V-Nick'}) => frozenset({'Rash Guard'}) confidence: 1.0
frozenset({'Rash Guard'}) => frozenset({'Dry Fit V-Nick'}) confidence: 0.8333333333333333
frozenset({'Rash Guard'}) => frozenset({'Tech Pants'}) confidence: 0.75
frozenset({'Tech Pants'}) => frozenset({'Rash Guard'}) confidence: 1.0
frozenset({'Rash Guard'}) => frozenset({'Hoodies'}) confidence: 0.6666666666666667
frozenset({'Hoodies'}) => frozenset({'Rash Guard'}) confidence: 1.0
frozenset({'Dry Fit V-Nick'}) => frozenset({'Tech Pants'}) confidence: 0.8
frozenset({'Tech Pants'}) => frozenset({'Dry Fit V-Nick'}) confidence: 0.8888888888888889
frozenset({'Rash Guard'}) => frozenset({'Swimming Shirt'}) confidence: 0.8333333333333333
frozenset({'Swimming Shirt'}) => frozenset({'Rash Guard'}) confidence: 0.9090909090909091
frozenset({'Socks'}) => frozenset({'Modern Pants'}) confidence: 0.6666666666666667
frozenset({'Modern Pants'}) => frozenset({'Socks'}) confidence: 0.8
```

```
Select Command Prompt
frozenset({'Swimming Shirt', 'Dry Fit V-Nick'}) => frozenset({'Rash Guard'}) confidence: 1.0
frozenset({'Swimming Shirt', 'Rash Guard'}) => frozenset({'Dry Fit V-Nick'}) confidence: 0.9
frozenset({'Rash Guard'}) => frozenset({'Swimming Shirt', 'Dry Fit V-Nick'}) confidence: 0.75
frozenset({'Swimming Shirt'}) => frozenset({'Rash Guard', 'Dry Fit V-Nick'}) confidence: 0.8181818181818181
frozenset({'Dry Fit V-Nick'}) => frozenset({'Swimming Shirt', 'Rash Guard'}) confidence: 0.9
frozenset({'Rash Guard', 'Dry Fit V-Nick'}) => frozenset({'Tech Pants'}) confidence: 0.8
frozenset({'Tech Pants', 'Dry Fit V-Nick'}) => frozenset({'Rash Guard'}) confidence: 1.0
frozenset({'Tech Pants', 'Rash Guard'}) => frozenset({'Dry Fit V-Nick'}) confidence: 0.8888888888888889
frozenset({'Tech Pants'}) => frozenset({'Rash Guard', 'Dry Fit V-Nick'}) confidence: 0.8888888888888889
frozenset({'Dry Fit V-Nick'}) => frozenset({'Tech Pants', 'Rash Guard'}) confidence: 0.8
frozenset({'Rash Guard'}) => frozenset({'Tech Pants', 'Dry Fit V-Nick'}) confidence: 0.6666666666666667
frozenset({'Running Shoe', 'Socks'}) => frozenset({'Modern Pants'}) confidence: 0.7272727272727273
frozenset({'Running Shoe', 'Modern Pants'}) => frozenset({'Socks'}) confidence: 0.8888888888888889
frozenset({'Socks', 'Modern Pants'}) => frozenset({'Running Shoe'}) confidence: 1.0
frozenset({'Running Shoe'}) => frozenset({'Socks', 'Modern Pants'}) confidence: 0.5714285714285715
frozenset({'Modern Pants'}) => frozenset({'Running Shoe', 'Socks'}) confidence: 0.8
frozenset({'Socks'}) => frozenset({'Running Shoe', 'Modern Pants'}) confidence: 0.6666666666666667
frozenset({'Running Shoe', 'Sweatshirts'}) => frozenset({'Modern Pants'}) confidence: 0.8181818181818181
frozenset({'Modern Pants', 'Sweatshirts'}) => frozenset({'Running Shoe'}) confidence: 0.9
frozenset({'Running Shoe', 'Modern Pants'}) => frozenset({'Sweatshirts'}) confidence: 1.0
frozenset({'Running Shoe'}) => frozenset({'Sweatshirts', 'Modern Pants'}) confidence: 0.6428571428571429
frozenset({'Modern Pants'}) => frozenset({'Running Shoe', 'Sweatshirts'}) confidence: 0.9
frozenset({'Sweatshirts'}) => frozenset({'Running Shoe', 'Modern Pants'}) confidence: 0.6923076923076923
frozenset({'Running Shoe', 'Sweatshirts'}) => frozenset({'Socks'}) confidence: 0.9090909090909091
frozenset({'Socks', 'Sweatshirts'}) => frozenset({'Running Shoe'}) confidence: 0.9090909090909091
frozenset({'Running Shoe', 'Socks'}) => frozenset({'Sweatshirts'}) confidence: 0.9090909090909091
frozenset({'Running Shoe'}) => frozenset({'Socks', 'Sweatshirts'}) confidence: 0.7142857142857143
frozenset({'Socks'}) => frozenset({'Running Shoe', 'Sweatshirts'}) confidence: 0.8333333333333333
frozenset({'Sweatshirts'}) => frozenset({'Running Shoe', 'Socks'}) confidence: 0.7692307692307692
frozenset({'Hoodies', 'Rash Guard'}) => frozenset({'Tech Pants'}) confidence: 1.0
frozenset({'Tech Pants', 'Rash Guard'}) => frozenset({'Hoodies'}) confidence: 0.8888888888888889
frozenset({'Hoodies', 'Tech Pants'}) => frozenset({'Rash Guard'}) confidence: 1.0
frozenset({'Hoodies'}) => frozenset({'Tech Pants', 'Rash Guard'}) confidence: 1.0
frozenset({'Tech Pants'}) => frozenset({'Hoodies', 'Rash Guard'}) confidence: 0.8888888888888889
frozenset({'Rash Guard'}) => frozenset({'Tech Pants', 'Hoodies'}) confidence: 0.6666666666666667
frozenset({'Rash Guard', 'Dry Fit V-Nick'}) => frozenset({'Hoodies'}) confidence: 0.7
frozenset({'Hoodies', 'Dry Fit V-Nick'}) => frozenset({'Rash Guard'}) confidence: 1.0
frozenset({'Hoodies', 'Rash Guard'}) => frozenset({'Dry Fit V-Nick'}) confidence: 0.8749999999999999
frozenset({'Rash Guard'}) => frozenset({'Hoodies', 'Dry Fit V-Nick'}) confidence: 0.8333333333333333
frozenset({'Hoodies'}) => frozenset({'Rash Guard', 'Dry Fit V-Nick'}) confidence: 0.8749999999999999
frozenset({'Dry Fit V-Nick'}) => frozenset({'Hoodies', 'Rash Guard'}) confidence: 0.7
```


5. Custom

```
C:\Users\VINAY>midterm datamining_vinay\python_midterm_apriori.py
Hello, welcome to association rule mining
Please choose the dataset for which you would like to get the Association rules
Please type
1 for Amazon
2 for BestBuy
3 for Kmart
4 for Nike
5 for Custom
5
You selected Custom Dataset
Enter the minimum support: 0.1
Enter the minimum Confidence: 0.1
.....
Frequent 1-itemsets      support
.....
Frozenset({'BOURNVITA'}) 0.2
Frozenset({'MAGGI'})    0.25
Frozenset({'JAM'})       0.1
Frozenset({'BISCUIT'})   0.35
Frozenset({'TEA'})        0.35
Frozenset({'MILK'})       0.25
Frozenset({'CORNFLOAKES'}) 0.3
Frozenset({'BREAD'})     0.65
Frozenset({'COCK'})       0.15
Frozenset({'COFFEE'})     0.4
Frozenset({'SUGER'})      0.3
.....
Frequent 2-itemsets      support
.....
Frozenset({'TEA', 'BREAD'}) 0.2
Frozenset({'TEA', 'CORNFLOAKES'}) 0.1
Frozenset({'COFFEE', 'BISCUIT'}) 0.1
Frozenset({'BISCUIT', 'MAGGI'}) 0.1
Frozenset({'BREAD', 'BOURNVITA'}) 0.15
Frozenset({'MILK', 'BISCUIT'}) 0.1
Frozenset({'TEA', 'MAGGI'}) 0.2
Frozenset({'CORNFLOAKES', 'COCK'}) 0.1
Frozenset({'CORNFLOAKES', 'MILK'}) 0.1
Frozenset({'BREAD', 'MILK'}) 0.2
Frozenset({'SUGER', 'BOURNVITA'}) 0.1
.....
Command Prompt
.....
Frozenset({'BREAD', 'BOURNVITA'}) 0.2
Frozenset({'COCK', 'BISCUIT'}) 0.1
Frozenset({'JAM', 'MAGGI'}) 0.1
Frozenset({'JAM', 'BREAD'}) 0.1
.....
Frequent 3-itemsets      support
.....
Frozenset({'BREAD', 'JAM', 'MAGGI'}) 0.1
Frozenset({'TEA', 'BISCUIT', 'MAGGI'}) 0.1
Frozenset({'CORNFLOAKES', 'COCK', 'BISCUIT'}) 0.1
Frozenset({'BREAD', 'MILK', 'BISCUIT'}) 0.1
Frozenset({'TEA', 'BREAD', 'BOURNVITA'}) 0.1
Frozenset({'COFFEE', 'CORNFLOAKES', 'COCK'}) 0.1
Frozenset({'SUGER', 'BREAD', 'COFFEE'}) 0.1
Frozenset({'COFFEE', 'CORNFLOAKES', 'BISCUIT'}) 0.1
Frozenset({'COFFEE', 'COCK', 'BISCUIT'}) 0.1
Frozenset({'TEA', 'BREAD', 'MAGGI'}) 0.1
.....
Association Rules with Confidence
.....
Frozenset({'BREAD'}) => frozenset({'TEA'}) confidence: 0.3076923076923077
Frozenset({'TEA'}) => frozenset({'BREAD'}) confidence: 0.5714285714285715
Frozenset({'CORNFLOAKES'}) => frozenset({'TEA'}) confidence: 0.3333333333333333
Frozenset({'TEA'}) => frozenset({'CORNFLOAKES'}) confidence: 0.28571428571428575
Frozenset({'COFFEE'}) => frozenset({'BISCUIT'}) confidence: 0.25
Frozenset({'BISCUIT'}) => frozenset({'COFFEE'}) confidence: 0.28571428571428575
Frozenset({'MAGGI'}) => frozenset({'MAGGI'}) confidence: 0.28571428571428575
Frozenset({'MAGGI'}) => frozenset({'BISCUIT'}) confidence: 0.4
Frozenset({'BREAD'}) => frozenset({'BOURNVITA'}) confidence: 0.23076923076923075
Frozenset({'BOURNVITA'}) => frozenset({'BREAD'}) confidence: 0.7499999999999999
Frozenset({'MILK'}) => frozenset({'BISCUIT'}) confidence: 0.4
Frozenset({'BISCUIT'}) => frozenset({'MILK'}) confidence: 0.28571428571428575
Frozenset({'TEA'}) => frozenset({'MAGGI'}) confidence: 0.5714285714285715
Frozenset({'MAGGI'}) => frozenset({'TEA'}) confidence: 0.8
Frozenset({'COCK'}) => frozenset({'CORNFLOAKES'}) confidence: 0.6666666666666667
Frozenset({'CORNFLOAKES'}) => frozenset({'COCK'}) confidence: 0.3333333333333333
Frozenset({'CORNFLOAKES'}) => frozenset({'MILK'}) confidence: 0.3333333333333333
Frozenset({'MILK'}) => frozenset({'CORNFLOAKES'}) confidence: 0.4
Frozenset({'BREAD'}) => frozenset({'MILK'}) confidence: 0.3076923076923077
```

Command Prompt

```
frozenset({'COFFEE'}) => frozenset({'BREAD'}) confidence: 0.37499999999999994
frozenset({'BREAD'}) => frozenset({'COFFEE'}) confidence: 0.23076923076923075
frozenset({'COFFEE'}) => frozenset({'COCK'}) confidence: 0.37499999999999994
frozenset({'COCK'}) => frozenset({'COFFEE'}) confidence: 1.0
frozenset({'TEA'}) => frozenset({'BOURNVITA'}) confidence: 0.28571428571428575
frozenset({'BOURNVITA'}) => frozenset({'TEA'}) confidence: 0.5
frozenset({'SUGER'}) => frozenset({'BREAD'}) confidence: 0.6666666666666667
frozenset({'BREAD'}) => frozenset({'SUGER'}) confidence: 0.3076923076923077
frozenset({'CORNFLAKES'}) => frozenset({'BISCUIT'}) confidence: 0.5
frozenset({'BISCUIT'}) => frozenset({'CORNFLAKES'}) confidence: 0.4285714285714286
frozenset({'BREAD'}) => frozenset({'BISCUIT'}) confidence: 0.3076923076923077
frozenset({'BISCUIT'}) => frozenset({'BREAD'}) confidence: 0.5714285714285715
frozenset({'COCK'}) => frozenset({'BISCUIT'}) confidence: 0.6666666666666667
frozenset({'BISCUIT'}) => frozenset({'COCK'}) confidence: 0.28571428571428575
frozenset({'JAM'}) => frozenset({'MAGGI'}) confidence: 1.0
frozenset({'MAGGI'}) => frozenset({'JAM'}) confidence: 0.4
frozenset({'BREAD'}) => frozenset({'JAM'}) confidence: 0.15384615384615385
frozenset({'JAM'}) => frozenset({'BREAD'}) confidence: 1.0
frozenset({'BREAD', 'JAM'}) => frozenset({'MAGGI'}) confidence: 1.0
frozenset({'BREAD', 'MAGGI'}) => frozenset({'JAM'}) confidence: 0.6666666666666667
frozenset({'JAM', 'MAGGI'}) => frozenset({'BREAD'}) confidence: 1.0
frozenset({'JAM'}) => frozenset({'BREAD', 'MAGGI'}) confidence: 1.0
frozenset({'BREAD'}) => frozenset({'JAM', 'MAGGI'}) confidence: 0.15384615384615385
frozenset({'MAGGI'}) => frozenset({'JAM', 'BREAD'}) confidence: 0.4
frozenset({'TEA', 'BISCUIT'}) => frozenset({'MAGGI'}) confidence: 1.0
frozenset({'TEA', 'MAGGI'}) => frozenset({'BISCUIT'}) confidence: 0.5
frozenset({'MAGGI', 'BISCUIT'}) => frozenset({'TEA'}) confidence: 1.0
frozenset({'TEA'}) => frozenset({'BISCUIT', 'MAGGI'}) confidence: 0.28571428571428575
frozenset({'BISCUIT'}) => frozenset({'TEA', 'MAGGI'}) confidence: 0.28571428571428575
frozenset({'MAGGI'}) => frozenset({'TEA', 'BISCUIT'}) confidence: 0.4
frozenset({'CORNFLAKES', 'COCK'}) => frozenset({'BISCUIT'}) confidence: 1.0
frozenset({'COCK', 'BISCUIT'}) => frozenset({'CORNFLAKES'}) confidence: 1.0
frozenset({'CORNFLAKES', 'BISCUIT'}) => frozenset({'COCK'}) confidence: 0.6666666666666667
frozenset({'BISCUIT'}) => frozenset({'CORNFLAKES', 'COCK'}) confidence: 0.28571428571428575
frozenset({'COCK'}) => frozenset({'CORNFLAKES', 'BISCUIT'}) confidence: 0.6666666666666667
frozenset({'CORNFLAKES'}) => frozenset({'COCK', 'BISCUIT'}) confidence: 0.3333333333333333
frozenset({'BREAD', 'MILK'}) => frozenset({'BISCUIT'}) confidence: 0.5
frozenset({'BREAD', 'BISCUIT'}) => frozenset({'MILK'}) confidence: 0.5
frozenset({'MILK', 'BISCUIT'}) => frozenset({'BREAD'}) confidence: 1.0
frozenset({'BREAD'}) => frozenset({'MILK', 'BISCUIT'}) confidence: 0.15384615384615385
frozenset({'BISCUIT'}) => frozenset({'BREAD', 'MILK'}) confidence: 0.28571428571428575
```

Command Prompt

```
frozenset({'BISCUIT'}) => frozenset({'BREAD', 'MILK'}) confidence: 0.28571428571428575
frozenset({'MILK'}) => frozenset({'BREAD', 'BISCUIT'}) confidence: 0.4
frozenset({'TEA', 'BREAD'}) => frozenset({'BOURNVITA'}) confidence: 0.5
frozenset({'BREAD', 'BOURNVITA'}) => frozenset({'TEA'}) confidence: 0.6666666666666667
frozenset({'TEA', 'BOURNVITA'}) => frozenset({'BREAD'}) confidence: 1.0
frozenset({'BOURNVITA'}) => frozenset({'TEA', 'BREAD'}) confidence: 0.5
frozenset({'TEA'}) => frozenset({'BREAD', 'BOURNVITA'}) confidence: 0.28571428571428575
frozenset({'BREAD'}) => frozenset({'TEA', 'BOURNVITA'}) confidence: 0.15384615384615385
frozenset({'COFFEE', 'COCK'}) => frozenset({'CORNFLAKES'}) confidence: 0.6666666666666667
frozenset({'COFFEE', 'CORNFLAKES'}) => frozenset({'COCK'}) confidence: 0.5
frozenset({'CORNFLAKES', 'COCK'}) => frozenset({'COFFEE'}) confidence: 1.0
frozenset({'COFFEE'}) => frozenset({'CORNFLAKES', 'COCK'}) confidence: 0.25
frozenset({'COCK'}) => frozenset({'COFFEE', 'CORNFLAKES'}) confidence: 0.6666666666666667
frozenset({'CORNFLAKES'}) => frozenset({'COFFEE', 'COCK'}) confidence: 0.3333333333333333
frozenset({'SUGER', 'COFFEE'}) => frozenset({'BREAD'}) confidence: 0.5
frozenset({'SUGER', 'BREAD'}) => frozenset({'COFFEE'}) confidence: 0.5
frozenset({'BREAD', 'COFFEE'}) => frozenset({'SUGER'}) confidence: 0.6666666666666667
frozenset({'BREAD'}) => frozenset({'SUGER', 'COFFEE'}) confidence: 0.15384615384615385
frozenset({'SUGER'}) => frozenset({'BREAD', 'COFFEE'}) confidence: 0.3333333333333333
frozenset({'COFFEE'}) => frozenset({'SUGER', 'BREAD'}) confidence: 0.25
frozenset({'COFFEE', 'CORNFLAKES'}) => frozenset({'BISCUIT'}) confidence: 0.5
frozenset({'COFFEE', 'BISCUIT'}) => frozenset({'CORNFLAKES'}) confidence: 1.0
frozenset({'CORNFLAKES', 'BISCUIT'}) => frozenset({'COFFEE'}) confidence: 0.6666666666666667
frozenset({'CORNFLAKES'}) => frozenset({'COFFEE', 'BISCUIT'}) confidence: 0.3333333333333333
frozenset({'BISCUIT'}) => frozenset({'COFFEE', 'CORNFLAKES'}) confidence: 0.28571428571428575
frozenset({'COFFEE'}) => frozenset({'CORNFLAKES', 'BISCUIT'}) confidence: 0.25
frozenset({'COFFEE', 'COCK'}) => frozenset({'BISCUIT'}) confidence: 0.6666666666666667
frozenset({'COFFEE', 'BISCUIT'}) => frozenset({'COCK'}) confidence: 1.0
frozenset({'COCK', 'BISCUIT'}) => frozenset({'COFFEE'}) confidence: 1.0
frozenset({'COCK'}) => frozenset({'COFFEE', 'BISCUIT'}) confidence: 0.6666666666666667
frozenset({'BISCUIT'}) => frozenset({'COFFEE', 'COCK'}) confidence: 0.28571428571428575
frozenset({'COFFEE'}) => frozenset({'COCK', 'BISCUIT'}) confidence: 0.25
frozenset({'TEA', 'BREAD'}) => frozenset({'MAGGI'}) confidence: 0.5
frozenset({'BREAD', 'MAGGI'}) => frozenset({'TEA'}) confidence: 0.6666666666666667
frozenset({'TEA', 'MAGGI'}) => frozenset({'BREAD'}) confidence: 0.5
frozenset({'MAGGI'}) => frozenset({'TEA', 'BREAD'}) confidence: 0.4
frozenset({'BREAD'}) => frozenset({'TEA', 'MAGGI'}) confidence: 0.15384615384615385
frozenset({'TEA'}) => frozenset({'BREAD', 'MAGGI'}) confidence: 0.28571428571428575
```

C:\Users\VINAY\midterm_datamining_vinay>