

Programming language: Python

Amazon.csv

A Beginner's Guide, Java: The Complete Reference, Java For Dummies, Android Programming: The Big Nerd Ranch,

A Beginner's Guide, Java: The Complete Reference, Java For Dummies,

A Beginner's Guide, Java: The Complete Reference, Java For Dummies, Android Programming: The Big Nerd Ranch, Head First Java 2nd Edition

Android Programming: The Big Nerd Ranch, Head First Java 2nd Edition, Beginning Programming with Java,,

Android Programming: The Big Nerd Ranch, Beginning Programming with Java, Java 8 Pocket Guide,

A Beginner's Guide, Android Programming: The Big Nerd Ranch, Head First Java 2nd Edition,

A Beginner's Guide, Head First Java 2nd Edition, Beginning Programming with Java,

Java: The Complete Reference, Java For Dummies, Android Programming: The Big Nerd Ranch,

Java For Dummies, Android Programming: The Big Nerd Ranch, Head First Java 2nd Edition, Beginning Programming with Java,

Beginning Programming with Java, Java 8 Pocket Guide, C++ Programming in Easy Steps,

A Beginner's Guide, Java: The Complete Reference, Java For Dummies, Android Programming: The Big Nerd Ranch,

A Beginner's Guide, Java: The Complete Reference, Java For Dummies, HTML and CSS: Design and Build Websites,

A Beginner's Guide, Java: The Complete Reference, Java For Dummies, Java 8 Pocket Guide, HTML and CSS: Design and Build Websites

Java For Dummies, Android Programming: The Big Nerd Ranch, Head First Java 2nd Edition,

Java For Dummies, Android Programming: The Big Nerd Ranch,,,

A Beginner's Guide, Java: The Complete Reference, Java For Dummies, Android Programming: The Big Nerd Ranch.

A Beginner's Guide, Java: The Complete Reference, Java For Dummies, Android Programming: The Big Nerd Ranch.

Head First Java 2nd Edition, Beginning Programming with Java, Java 8 Pocket Guide,

Android Programming: The Big Nerd Ranch, Head First Java 2nd Edition,,,

A Beginner's Guide, "Java: The Complete Reference, Java For Dummies",,,U

BESTBUY.CSV

Desk Top, Printer, Flash Drive, Microsoft Office, Speakers, Anti-Virus

Lab Top, Flash Drive, Microsoft Office, Lab Top Case, Anti-Virus, ,,,,

Lab Top, Printer, Flash Drive, Microsoft Office, Anti-Virus, Lab Top Case, External Hard-Drive,,,

Lab Top, Printer, Flash Drive, Anti-Virus, External Hard-Drive, Lab Top Case,,,,

Lab Top, Flash Drive, Lab Top Case, Anti-Virus,,,,,

Lab Top, Printer, Flash Drive, Microsoft Office,,,,,

Desk Top, Printer, Flash Drive, Microsoft Office,,,,,,

Lab Top, External Hard-Drive, Anti-Virus,,,,,,

Desk Top, Printer, Flash Drive, Microsoft Office, Lab Top Case, Anti-Virus, Speakers, External Hard-Drive,

Digital Camera, Lab Top, Desk Top, Printer, Flash Drive, Microsoft Office, Lab Top Case, Anti-Virus, External Hard-Drive, Speakers

Lab Top, Desk Top, Lab Top Case, External Hard-Drive, Speakers, Anti-Virus, ,,,

Digital Camera, Lab Top, Lab Top Case, External Hard-Drive, Anti-Virus, Speakers,,,,

Digital Camera, Speakers,,,,,,

Digital Camera, Desk Top, Printer, Flash Drive, Microsoft Office,,,,,

Printer, Flash Drive, Microsoft Office, Anti-Virus, Lab Top Case, Speakers, External Hard-Drive,,,

Digital Camera, Flash Drive, Microsoft Office, Anti-Virus, Lab Top Case, External Hard-Drive, Speakers,,,

Digital Camera, Lab Top, Lab Top Case,,,,,,

Digital Camera, Lab Top Case, Speakers,,,,,,

Digital Camera, Lab Top, Printer, Flash Drive, Microsoft Office, Speakers, Lab Top Case, Anti-Virus,

Digital Camera, Lab Top, Speakers, Anti-Virus, Lab Top Case, ...

KMART.CSV

Decorative Pillows, "Quilts SEP!

",Embroidered Bedspread,,,,

Embroidered Bedspread, "Shams SEP,", "Kids Bedding

",Bedding Collections,"Bed Skirts Bedspreads, "Sheets Bedspreads," Sheets Bedspreads,

Decorative Pillows, "Quilts SEP!

",Embroidered Bedspread, "Shams, "Kids Bedding", "Kids Bedding",

",Bedding Collections,

Kids Bedding, Bedding Collections, "Sheets [1]", Bedspreads, "Bed Skirts [1]",

Decorative Pillows,"Kids Bedding

```
",Bedding Collections, "Sheets "," Bed Skirts "," Bedspreads,
Bedding Collections, Bedspreads, "Bed Skirts[sep]", "Sheets[sep]", "Shams[sep]", "Kids Bedding
Decorative Pillows, Quilts,,,,,
"Decorative Pillows,", Quilts, Embroidered Bedspread,,,,
Bedspreads, "Bed Skirts "," Shams "," Kids Bedding
","Sheets[SEP]",,
Quilts, Embroidered Bedspread, Bedding Collections,,,,
Bedding Collections, Bedspreads, "Bed Skirts, "Kids Bedding
","Shams[sep]",Sheets,
Decorative Pillows, Quilts,,,,,
Embroidered Bedspread, "Shams[5EP]",,,,,
Sheets, "Shams "," Bed Skirts EE ", "Kids Bedding
Decorative Pillows, Quilts,,,,,
Decorative Pillows,"Kids Bedding
","Bed Skirts[1]","Shams[1]",,,
Decorative Pillows, "Shams sep", "Bed Skirts sep", ",,,,
Quilts, Sheets, "Kids Bedding
Shams, "Bed Skirts "," Kids Bedding
",Sheets,,,
Decorative Pillows, Bedspreads, "Shams's ", Sheets, "Bed Skirts ", "Kids Bedding
```

NIKE.CSV

```
Running Shoe, Socks, Sweatshirts, Modern Pants,,,,,,

Running Shoe, Socks, Sweatshirts, Modern Pants,,,,,,

Running Shoe, Sweatshirts, Modern Pants,,,,,,

Running Shoe, Socks, Sweatshirts, Modern Pants, Soccer Shoe,,,,,

Running Shoe, Socks, Sweatshirts,,,,,,,,
```

Running Shoe, Socks, Sweatshirts, Modern Pants, Tech Pants, Rash Guard, Hoodies,,,

Swimming Shirt, Socks, Sweatshirts,,,,,,

Swimming Shirt, Rash Guard, Dry Fit V-Nick, Hoodies, Tech Pants,,,,,

Swimming Shirt, Rash Guard, Dry Fit V-Nick,,,,,,

Swimming Shirt, Rash Guard, Dry Fit V-Nick,,,,,,

Running Shoe,Swimming Shirt,Socks,Sweatshirts,Modern Pants,Soccer Shoe,Rash Guard,Hoodies,Tech Pants,Dry Fit V-Nick

Running Shoe, Swimming Shirt, Socks, Sweatshirts, Modern Pants, Soccer Shoe, Rash Guard, Tech Pants, Dry Fit V-Nick, Hoodies

Running Shoe, Swimming Shirt, Rash Guard, Tech Pants, Hoodies, Dry Fit V-Nick,,,,

Running Shoe, Swimming Shirt, Socks, Sweatshirts, Modern Pants, Dry Fit V-Nick, Rash Guard, Tech Pants,

Swimming Shirt, Soccer Shoe, Hoodies, Dry Fit V-Nick, Tech Pants, Rash Guard,,,,

Running Shoe, Socks,,,,,,,

"Socks,",Sweatshirts,Modern Pants,Soccer Shoe,Hoodies,Rash Guard,Tech Pants,Dry Fit V-Nick,,

Running Shoe, Swimming Shirt, Rash Guard,,,,,,

Running Shoe, Swimming Shirt, Socks, Sweatshirts, Modern Pants, Soccer Shoe, Hoodies, Tech Pants, Rash Guard, Dry Fit V-Nick

How to run the file: python project.py datasetname.csv minimum_support(in decimal) minimum_confidence(in decimal)

```
1 from csv import reader
2 import pandas as pd
 4 Edef load_data_set():
         Load a sample data set
          A data set: A list of transactions. Each transaction contains several items.
10
11
12
13
         print("Hello, Please enter which data set you need \n 1) Press 1 for Amazon \n 2) Press 2 for BestBuy \n 3) Press 3 for Nike \n 4) Press 4 for KMart")
         while True:
             choice of data = input()
             if (choice of data == '1'):
14
                 data = 'amazon.csv'
                 print('User chose amazon dataset')
 16
                  break
             elif(choice_of_data == '2'):
18
19
20
                data = 'bestbuy.csv'
                 print('User chose bestbuy dataset')
                  break
21
             elif (choice of data == '3'):
               data = 'Nike.csv'
                 print('User chose Nike dataset')
24
                 break
              elif (choice_of_data == '4'):
26
27
                data = 'kmart.csv'
                 print('User chose KMart dataset')
                 break
29 E
              else:
                print("Please enter the right one")
 31
        with open(data, 'r') as read_obj:
             # pass the file object to reader() to get the reader object
33
34
             csv_reader = reader(read_obj)
35
36
37
38
39
             # Pass reader object to list() to get a list of lists
              data set = list(csv reader)
              for items in data set:
                 for j in range(0, len(items)):
                      for items1 in items:
40
41 E
                         if items1 == "":
                          items.remove("")
42
43
44
         return data set
```

```
47 pdef create_C1(data_set):
49
         Create frequent candidate 1-itemset C1 by scaning data set.
           data set: A list of transactions. Each transaction contains several items.
52
         C1: A set which contains all frequent candidate 1-itemsets """
 54
         C1 = set()
56 E
         for t in data set:
          for item in t:
 58
              item_set = frozenset([item])
 59
                C1.add(item set)
         return C1
62
63 | def is_apriori(Ck_item, Lksub1):
 65
         Judge whether a frequent candidate k-itemset satisfy Apriori property.
 66
         Args:
 67
            Ck_item: a frequent candidate k-itemset in Ck which contains all frequent
 68
                     candidate k-itemsets.
 69
           Lksub1: Lk-1, a set which contains all frequent candidate (k-1)-itemsets.
         Returns:
           True: satisfying Apriori property.
         False: Not satisfying Apriori property.
 74 🖨
         for item in Ck_item:
            sub Ck = Ck item - frozenset([item])
 76
            if sub Ck not in Lksub1:
              return False
         return True
 79
81 pdef create_Ck(Lksub1, k):
82 🛱
         Create Ck, a set which contains all all frequent candidate k-itemsets
84
         by Lk-1's own connection operation.
         Args:
          Lksubl: Lk-1, a set which contains all frequent candidate (k-1)-itemsets.
86
            k: the item number of a frequent itemset.
         Return:
89
         Ck: a set which contains all all frequent candidate k-itemsets.
```

```
Ck = set()
            len Lksub1 = len(Lksub1)
 92
 93
            list_Lksub1 = list(Lksub1)
 94 🛱
            for i in range(len_Lksub1):
 95
                for j in range(1, len_Lksub1):
 96
                      11 = list(list_Lksub1[i])
 97
                     12 = list(list_Lksub1[j])
 98
                     11.sort()
 99
                     12.sort()
                     if 11[0:k-2] == 12[0:k-2]:
                          Ck_item = list_Lksub1[i] | list_Lksub1[j]
                          # pruning
                          if is apriori(Ck item, Lksub1):
104
105
106
107
                              Ck.add(Ck_item)
            return Ck
108 | def generate Lk by Ck(data set, Ck, min support, support data):
109 | """
110 | Generate Lk by executing a delete policy from Ck.
111 | Args:
                data set: A list of transactions. Each transaction contains several items. Ck: \overline{A} set which contains all all frequent candidate k-itemsets.
                min support: The minimum support.
support_data: A dictionary. The key is frequent itemset and the value is support.
114
116
            Returns:
            Lk: A set which contains all all frequent k-itemsets.
            Lk = set()
 119
            item_count = {}
121 日
122 日
123 日
124 日
            for t in data_set:
                for item in Ck:
                    if item.issubset(t):
                         if item not in item_count:
125
126 B
                              item_count[item] = 1
                              item count[item] += 1
            t num = float(len(data set))
           for item in item count:
             if (item_count[item] / t_num) >= min_support:
                     Lk.add(item)
                     support data[item] = item count[item] / t num
            return Lk
```

```
137 E
138
139
           Generate all frequent itemsets.
           Args:
140
141
               data_set: A list of transactions. Each transaction contains several items. 
 k: Maximum number of items for all frequent itemsets.
               min_support: The minimum support.
 142
 143
           Returns:
           support_data: A dictionary. The key is frequent itemset and the value is support.
 145
 146
 147
           support_data = {}
 148
           C1 = create_C1(data_set)
 149
150
           L1 = generate_Lk_by_Ck(data_set, C1, min_support, support_data)
           Lksub1 = L1.copy()
           L = [1]
           L.append (Lksub1)
153 B
154
           for i in range(2, k+1):
             Ci = create_Ck(Lksub1, i)
               Li = generate_Lk_by_Ck(data_set, Ci, min_support, support_data)
 156
               Lksub1 = Li.copy()
               L.append(Lksub1)
           return L, support_data
160
161 | Hedef generate_big_rules(L, support_data, min_conf):
162 | """
163 | Generate_big_rules from frequent itemsets.
 164
           Args:
 166
               support_data: A dictionary. The key is frequent itemset and the value is support.
              min_conf: Minimal confidence.
           Returns:
            big_rule_list: A list which contains all big rules. Each big rule is represented
 169
                   as a 3-tuple.
           big_rule_list = []
           sub_set_list = []
for i in range(0, len(L)):
 174
175 自
176 自
               for freq set in L[i]:
                   for sub set in sub set list:
                       if sub_set.issubset(freq_set):
                            conf = support_data[freq_set] / support_data[freq_set - sub_set]
179
                            big_rule = (freq_set - sub_set, sub_set, conf)
```

```
if conf >= min_conf and big_rule not in big_rule_list:|
    # print freq_set-sub_set, " => ", sub_set, "conf: ", conf
big_rule_list.append(big_rule)
                    sub_set_list.append(freq_set)
184
185
            return big_rule_list
data set = load data set()
            L, support_data = generate_L(data_set, k=3, min_support = float(input("Enter the minimum support: ")))
193
194
195
196
197
198
199
-
200
            big_rules_list = generate_big_rules(L, support_data, min_conf= float(input("Enter the minimum confidence: ")))
            for Lk in L:
                print("="*50)
                print("Fought" + str(len(list(Lk)[0])) + "-itemsets\t\tsupport")
print("="*50)
for freq_set in Lk:
                    print(freq_set, support_data[freq_set])
            print("Big Rules")
            for item in big_rules_list:
               print(item[0], "=>", item[1], "conf: ", item[2])
```

INPUT

```
C:\Users\admin\Desktop\Mid_term_DATA_MINING>python apriori.py
Hello, Please enter which data set you need

1) Press 1 for Amazon

2) Press 2 for BestBuy

3) Press 3 for Nike

4) Press 4 for KMart

3
User chose Nike dataset
Enter the minimum support: 0.2
Enter the minimum confidence: 0.4
```

OUTPUT

```
citer the skinian confidence 0.2

dirich the skinian confidence 0.4

dirich the skinian confidence 0.4

forcement([Schming Shirt']) 0.55

forcement([Chy fit V Hick']) 0.5

forc
```

```
Treatment States Support
Frequent States Support
Frequent (States) Summing State*, "Neder Parks*) 0.2
frequent(State Good*, "Suming State*, "Neder Parks*)) 0.2
frequent(State Good*, "Suming State*, "Neder Parks*)) 0.3
frequent(State Good*, "Suming State*, "Neder Parks*)) 0.2
frequent(State Good*, "Suming State*, "Neder*)) 0.2
frequent(State Good*, "Runing State*, "Neder*) 0.2
frequent(State Good*, "Runing State*, "Neder*) 0.2
frequent(State Good*, "Suming State*, "Neder*) 0.2
frequent(State Good*, "Neder*, "Neder*, "Neder*) 0.2
frequent(State Good*, "Neder*, "
```