CS 634 Data Mining

Midterm Project: Apriori Algorithm

Alfred Zane Rajan Velladurai
UCID: ar2257

Source Code: (In Python)

```
import pandas as pd

support = 50   #float(sys.argv[1])
confidence = 70     #float(sys.argv[2])
file = "db1.txt"    #(sys.argv[3])
df = pd.read_csv( file, skipinitialspace = True, names = ['0', '1','2','3','4','5','6','7','8','9','10'])
df = df.drop('0', axis=1)
df = df.T
for col in df:
   df[col] = df[col].str.replace(" ", "")
db = list(list(df[i]) for i in df)
db = [[j for j in i if not pd.isna(j)] for i in db]
#for trans in db:
#item = ["diapers", "sweaters", "tissues", "belts", "water", "noodles", "cereals", "books", "pen",
"batteries"]
items = []
for trans in db:
      for item in trans:
         if not item in items:
            items.append(item)
support = support*len(db)/100
confidence = confidence/100

count = {}
for item in items:
   count[item] = 0
for item in items:
   for trans in db:
      if item in trans:
         count[item]+=1
for item in items:
    if count[item] < support:
        items.remove(item)

itemsets = [{i} for i in items]
while itemsets:
    tempsets = []
    for iset in itemsets:
        for item in items:
            if not {item}.issubset(iset):
                match = 0
                total = 0
                for trans in db:
                    if iset.issubset(trans):
                        total+=1
                        if item in trans:
                            match+=1
                if match>=support:
                    temp = {i for i in iset}
                    temp.add(item)
                    if temp not in tempsets:
                        tempsets.append(temp)
                    if (match/total)>=confidence:
                        print(iset," -> ",item)
    itemsets = tempsets
    for item in items:
        flag = 0
        for iset in itemsets:
            if {item}.issubset(iset):
                flag = 1
        if flag == 0:
            items.remove(item)
```

P.S.: I had run the program on Spyder 3.2.8 on Windows on my PC. So I have hard coded the support, confidence and input path into the code. But I have also given the code to read from the command line (Commented).

For all Databases:
Support = 50%
Confidence = 70%

Data Base 1:

Input:

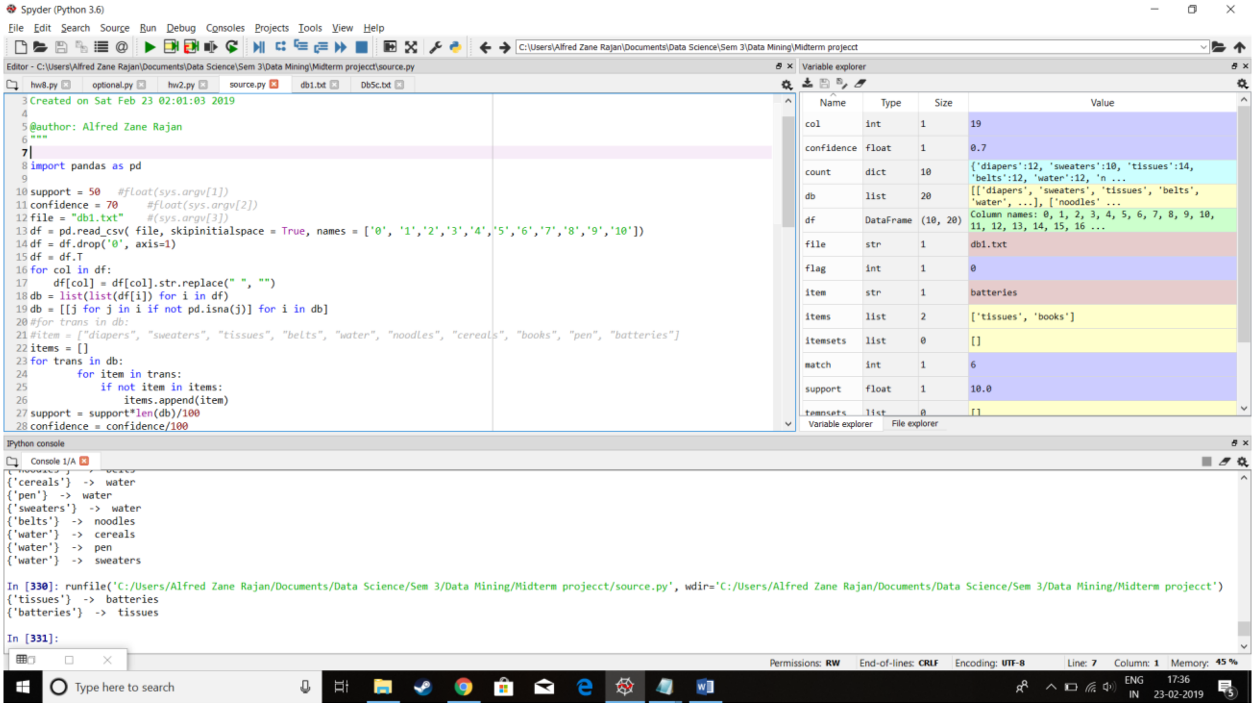
1001, diapers, sweaters, tissues, belts, water, noodles, cereals, books, pen, batteries
1002, noodles, pen, books, sweaters, water
1003, sweaters, tissues, noodles, cereals, belts, books
1004, pen, tissues, batteries, diapers, water
1005, water, diapers, belts, books
1006, sweaters, batteries, cereals, tissues, diapers, books
1007, batteries, water, sweaters, belts, noodles, books, diapers, cereals, tissues, pen
1008, water, belts, pen, books, tissues
1009, books, tissues, batteries, diapers, noodles
1010, pen, noodles, books, sweaters, batteries, belts, tissues
1011, water, noodles, cereals, sweaters, tissues, diapers, belts
1012, belts, tissues, batteries
1013, belts, water, tissues, batteries
1014, batteries, sweaters, belts, diapers, pen, books, water, cereals, tissues
1015, noodles, batteries, cereals, tissues, pen
1016, noodles, cereals, sweaters, diapers, water, batteries
1017, belts, pen
1018, diapers, sweaters, water, noodles, books, batteries
1019, diapers, tissues, water, noodles, pen, batteries
1020, diapers, belts, noodles,

Output:

runfile('C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct/source.py', wdir='C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct')

{'tissues'} -> batteries

{'batteries'} -> tissues



Editor - C:\Users\Alfred Zane Rajan\Documents\Data Science\Sem 3\Data Mining\Midterm projecct\source.py

```python
Created on Sat Feb 23 02:01:03 2019

@author: Alfred Zane Rajan
"""

import pandas as pd

support = 50    #float(sys.argv[1])
confidence = 70      #float(sys.argv[2])
file = "db1.txt"     #(sys.argv[3])
df = pd.read_csv( file, skipinitialspace = True, names = ['0', '1','2','3','4','5','6','7','8','9','10'])
df = df.drop('0', axis=1)
df = df.T
for col in df:
    df[col] = df[col].str.replace(" ", "")
db = list(list(df[i]) for i in df)
db = [[j for j in i if not pd.isna(j)] for i in db]
#for trans in db:
#item = ["diapers", "sweaters", "tissues", "belts", "water", "noodles", "cereals", "books", "pen", "batteries"]
items = []
for trans in db:
        for item in trans:
            if not item in items:
                items.append(item)
support = support*len(db)/100
confidence = confidence/100
```

Variable explorer

Name	Type	Size	Value
col	int	1	19
confidence	float	1	0.7
count	dict	10	{'diapers':12, 'sweaters':10, 'tissues':14, 'belts':12, 'water':12, 'n ...
db	list	20	[['diapers', 'sweaters', 'tissues', 'belts', 'water', ...], ['noodles' ...
df	DataFrame	(10, 20)	Column names: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 ...
file	str	1	db1.txt
flag	int	1	0
item	str	1	batteries
items	list	2	['tissues', 'books']
itemsets	list	0	[]
match	int	1	6
support	float	1	10.0
temnsets	list	0	[]

IPython console

```
{'cereals'}  ->  water
{'pen'}  ->  water
{'sweaters'}  ->  water
{'belts'}  ->  noodles
{'water'}  ->  cereals
{'water'}  ->  pen
{'water'}  ->  sweaters

In [330]: runfile('C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct/source.py', wdir='C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct')
{'tissues'}  ->  batteries
{'batteries'}  ->  tissues

In [331]:
```

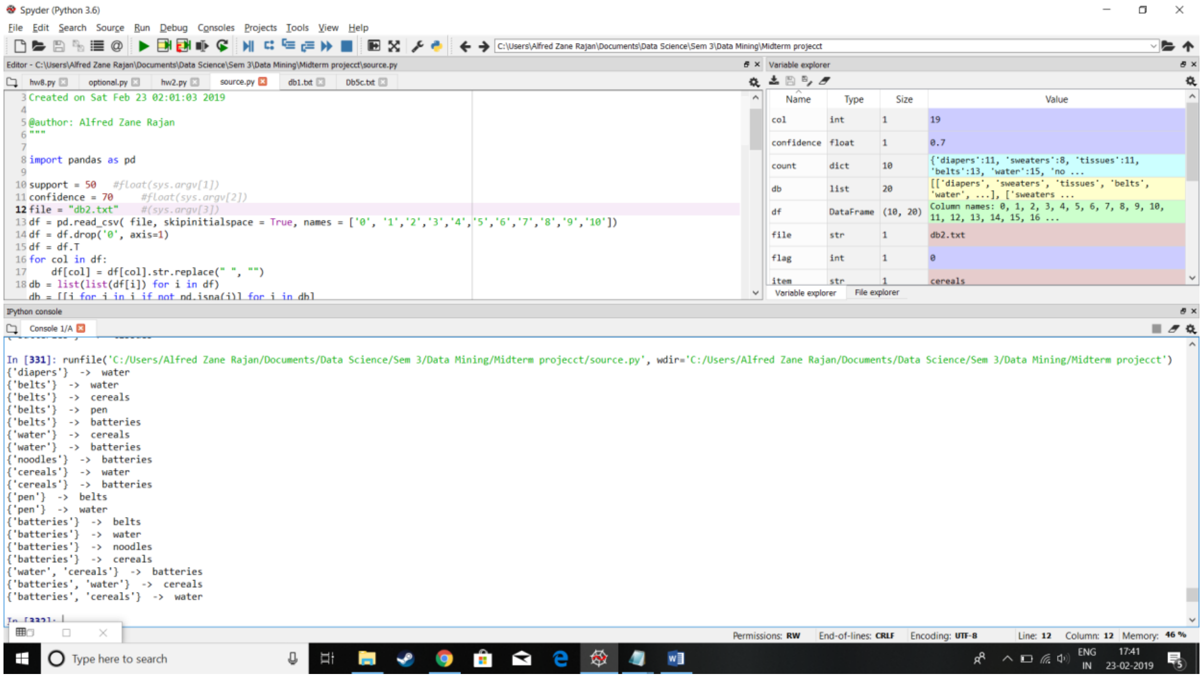
Data Base 2:

Input:

1001, diapers, sweaters, tissues, belts, water, noodles, cereals, books, pen, batteries
1002, sweaters, tissues, noodles, cereals, books, batteries
1003, sweaters, tissues, belts, noodles, cereals, pen, batteries
1004, tissues, belts, water, noodles, cereals, books, pen, batteries
1005, tissues, belts, water, cereals, pen, batteries
1006, diapers, tissues, belts, water, noodles, cereals, books, batteries
1007, sweaters, cereals, books
1008, diapers, belts, water, pen
1009, diapers, sweaters, tissues, water, pen
1010, diapers, belts, water, noodles, books, pen, batteries
1011, diapers, belts, water, noodles, cereals, pen, batteries
1012, diapers, tissues, belts, water, cereals, batteries
1013, sweaters, water, noodles, cereals, books, batteries
1014, belts, pen
1015, sweaters, tissues, belts, water, noodles, cereals, batteries
1016, diapers, tissues, water, noodles, pen
1017, diapers, water, cereals, batteries
1018, diapers, belts, cereals, pen
1019, sweaters, belts, water, noodles, cereals, pen, batteries
1020, diapers, tissues, water, cereals, books, pen

Output:

```
runfile('C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct/source.py', wdir='C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct')
{'diapers'}  ->  water
{'belts'}  ->  water
{'belts'}  ->  cereals
{'belts'}  ->  pen
{'belts'}  ->  batteries
{'water'}  ->  cereals
{'water'}  ->  batteries
{'noodles'}  ->  batteries
{'cereals'}  ->  water
{'cereals'}  ->  batteries
{'pen'}  ->  belts
{'pen'}  ->  water
{'batteries'}  ->  belts
{'batteries'}  ->  water
{'batteries'}  ->  noodles
{'batteries'}  ->  cereals
{'water', 'cereals'}  ->  batteries
{'batteries', 'water'}  ->  cereals
{'batteries', 'cereals'}  ->  water
```



Spyder (Python 3.6)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\Alfred Zane Rajan\Documents\Data Science\Sem 3\Data Mining\Midterm projecct
Editor - C:\Users\Alfred Zane Rajan\Documents\Data Science\Sem 3\Data Mining\Midterm projecct\source.py
hw8.py
optional.py
hw2.py
source.py
db1.txt
Db5c.txt
Variable explorer
File explorer
IPython console
Console 1/A
Permissions: RW
End-of-lines: CRLF
Encoding: UTF-8
Line: 12
Column: 12
Memory: 46 %
Type here to search
ENG
IN
17:41
23-02-2019

```python
Created on Sat Feb 23 02:01:03 2019

@author: Alfred Zane Rajan
"""

import pandas as pd

support = 50    #float(sys.argv[1])
confidence = 70     #float(sys.argv[2])
file = "db2.txt"    #(sys.argv[3])
df = pd.read_csv( file, skipinitialspace = True, names = ['0', '1','2','3','4','5','6','7','8','9','10'])
df = df.drop('0', axis=1)
df = df.T
for col in df:
    df[col] = df[col].str.replace(" ", "")
db = list(list(df[i]) for i in df)
db = [[i for i in i if not pd.isna(i)] for i in db]
```

Name	Type	Size	Value
col	int	1	19
confidence	float	1	0.7
count	dict	10	{'diapers':11, 'sweaters':8, 'tissues':11, 'belts':13, 'water':15, 'no ...
db	list	20	[['diapers', 'sweaters', 'tissues', 'belts', 'water', ...], ['sweaters ...
df	DataFrame	(10, 20)	Column names: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 ...
file	str	1	db2.txt
flag	int	1	0
item	str	1	cereals

```
In [331]: runfile('C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct/source.py', wdir='C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct')
{'diapers'}  ->  water
{'belts'}  ->  water
{'belts'}  ->  cereals
{'belts'}  ->  pen
{'belts'}  ->  batteries
{'water'}  ->  cereals
{'water'}  ->  batteries
{'noodles'}  ->  batteries
{'cereals'}  ->  water
{'cereals'}  ->  batteries
{'pen'}  ->  belts
{'pen'}  ->  water
{'batteries'}  ->  belts
{'batteries'}  ->  water
{'batteries'}  ->  noodles
{'batteries'}  ->  cereals
{'water', 'cereals'}  ->  batteries
{'batteries', 'water'}  ->  cereals
{'batteries', 'cereals'}  ->  water

In [332]:
```

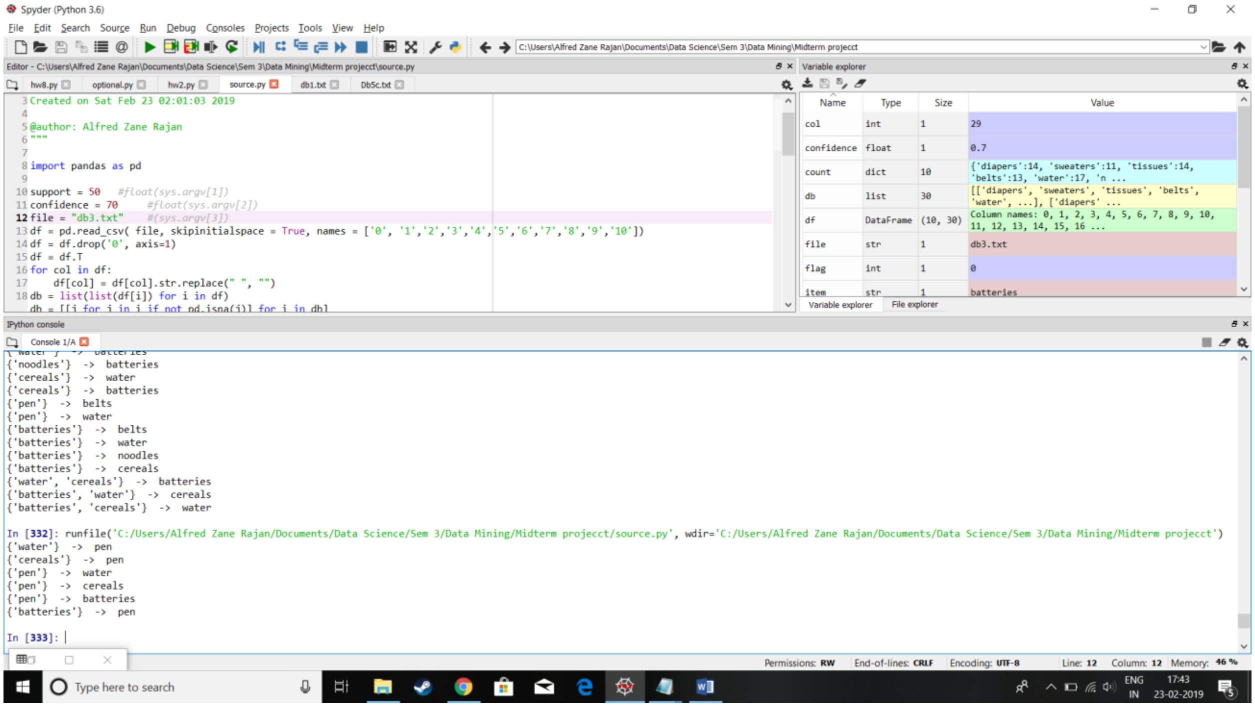
Data Base 3:

Input:

```
1001, diapers, sweaters, tissues, belts, water, noodles, cereals, books, pen, batteries
1002, diapers, tissues, belts, water, noodles, cereals, books, pen, batteries
1003, diapers, belts, water, noodles, cereals, pen, batteries
1004, diapers, tissues, belts, water, noodles, books, pen, batteries
1005, sweaters, belts, noodles, cereals, books, pen, batteries
1006, diapers, tissues, water, cereals, pen
1007, diapers, sweaters, belts, water, cereals, pen, batteries
1008, water, cereals, books, pen, batteries
1009, diapers, sweaters, tissues, belts, cereals, books, pen, batteries
1010, sweaters, tissues, belts, water, cereals, books, pen
1011, diapers, sweaters, tissues, belts, water, noodles, cereals, pen, batteries
1012, sweaters, tissues, belts, water, noodles, cereals, books, pen, batteries
1013, diapers, sweaters, tissues, noodles, cereals, pen, batteries
1014, tissues, water, noodles, cereals, pen, batteries
1015, diapers, sweaters, tissues, belts, water, noodles, pen
1016, sweaters, belts, water, noodles, pen, batteries
1017, diapers, sweaters, belts, water, cereals, books, pen, batteries
1018, diapers, tissues, water, noodles, pen
1019, diapers, tissues, water, noodles, cereals, books, pen, batteries
1010, diapers, tissues, water, noodles, cereals, books
```

Output:

```
runfile('C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct/source.py', wdir='C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct')
{'water'}  ->  pen
{'cereals'}  ->  pen
{'pen'}  ->  water
{'pen'}  ->  cereals
{'pen'}  ->  batteries
{'batteries'}  ->  pen
```



Spyder (Python 3.6)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\Alfred Zane Rajan\Documents\Data Science\Sem 3\Data Mining\Midterm projecct
Editor - C:\Users\Alfred Zane Rajan\Documents\Data Science\Sem 3\Data Mining\Midterm projecct\source.py
hw8.py
optional.py
hw2.py
source.py
db1.txt
Db5c.txt
Variable explorer

```python
Created on Sat Feb 23 02:01:03 2019

@author: Alfred Zane Rajan
"""

import pandas as pd

support = 50   #float(sys.argv[1])
confidence = 70     #float(sys.argv[2])
file = "db3.txt"     #(sys.argv[3])
df = pd.read_csv( file, skipinitialspace = True, names = ['0', '1','2','3','4','5','6','7','8','9','10'])
df = df.drop('0', axis=1)
df = df.T
for col in df:
    df[col] = df[col].str.replace(" ", "")
db = list(list(df[i]) for i in df)
db = [[i for i in i if not pd.isna(i)] for i in db]
```

Name	Type	Size	Value
col	int	1	29
confidence	float	1	0.7
count	dict	10	{'diapers':14, 'sweaters':11, 'tissues':14, 'belts':13, 'water':17, 'n ...
db	list	30	[['diapers', 'sweaters', 'tissues', 'belts', 'water', ...], ['diapers' ...
df	DataFrame	(10, 30)	Column names: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 ...
file	str	1	db3.txt
flag	int	1	0
item	str	1	batteries

```
{'water'}  ->  batteries
{'noodles'}  ->  batteries
{'cereals'}  ->  water
{'cereals'}  ->  batteries
{'pen'}  ->  belts
{'pen'}  ->  water
{'batteries'}  ->  belts
{'batteries'}  ->  water
{'batteries'}  ->  noodles
{'batteries'}  ->  cereals
{'water', 'cereals'}  ->  batteries
{'batteries', 'water'}  ->  cereals
{'batteries', 'cereals'}  ->  water

In [332]: runfile('C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct/source.py', wdir='C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct')
{'water'}  ->  pen
{'cereals'}  ->  pen
{'pen'}  ->  water
{'pen'}  ->  cereals
{'pen'}  ->  batteries
{'batteries'}  ->  pen

In [333]: 
```

Data Base 4:

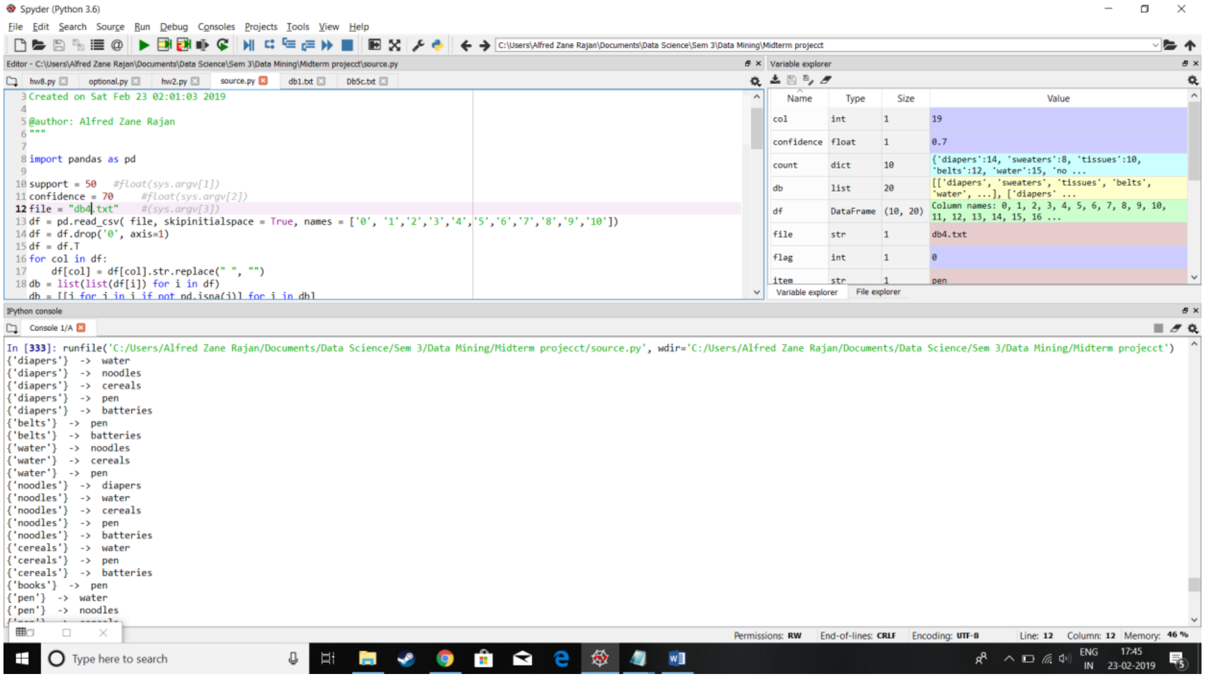
Input:

```
1001, diapers, sweaters, tissues, belts, water, noodles, cereals, books, pen, batteries
1002, diapers, tissues, belts, water, noodles, cereals, books, pen, batteries
1003, diapers, belts, water, noodles, cereals, pen, batteries
1004, diapers, tissues, belts, noodles, books, pen, batteries
1005, belts, noodles, cereals, books, pen, batteries
1006, diapers, tissues, water, cereals, pen
1007, diapers, sweaters, water, cereals, batteries
1008, water, cereals, books, pen, batteries
1009, diapers, sweaters, tissues, belts, books, pen, batteries
1010, sweaters, belts, water, cereals, books, pen
1011, diapers, belts, water, noodles, cereals, pen, batteries
1012, sweaters, belts, water, noodles, cereals, books, pen, batteries
1013, diapers, tissues, noodles, cereals, pen, batteries
1014, tissues, water, noodles, cereals, pen, batteries
1015, diapers, sweaters, tissues, belts, water, noodles, pen
1016, sweaters, belts, water, noodles, pen, batteries
1017, diapers, sweaters, belts, cereals, books, pen, batteries
1018, diapers, tissues, water, noodles, pen
1019, diapers, water, noodles, cereals, books, pen, batteries
1010, diapers, tissues, water, noodles, cereals, books
```

Output:

```
{'diapers'} -> water
{'diapers'} -> noodles
{'diapers'} -> cereals
{'diapers'} -> pen
{'diapers'} -> batteries
{'belts'} -> pen
{'belts'} -> batteries
{'water'} -> noodles
{'water'} -> cereals
{'water'} -> pen
{'noodles'} -> diapers
{'noodles'} -> water
{'noodles'} -> cereals
{'noodles'} -> pen
{'noodles'} -> batteries
{'cereals'} -> water
```

```
{'cereals'} -> pen
{'cereals'} -> batteries
{'books'} -> pen
{'pen'} -> water
{'pen'} -> noodles
{'pen'} -> cereals
{'pen'} -> batteries
{'batteries'} -> noodles
{'batteries'} -> cereals
{'batteries'} -> pen
{'belts', 'pen'} -> batteries
{'belts', 'batteries'} -> pen
{'noodles', 'water'} -> pen
{'water', 'cereals'} -> pen
{'pen', 'water'} -> noodles
{'pen', 'water'} -> cereals
{'pen', 'noodles'} -> water
{'pen', 'noodles'} -> batteries
{'batteries', 'noodles'} -> pen
{'pen', 'cereals'} -> water
{'pen', 'cereals'} -> batteries
{'batteries', 'cereals'} -> pen
{'pen', 'batteries'} -> belts
{'pen', 'batteries'} -> noodles
{'pen', 'batteries'} -> cereals
```


Spyder (Python 3.6)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\Alfred Zane Rajan\Documents\Data Science\Sem 3\Data Mining\Midterm projecct
Editor - C:\Users\Alfred Zane Rajan\Documents\Data Science\Sem 3\Data Mining\Midterm projecct\source.py
hw8.py
optional.py
hw2.py
source.py
db1.txt
Db5c.txt
Variable explorer
File explorer
IPython console
Console 1/A
Permissions: RW
End-of-lines: CRLF
Encoding: UTF-8
Line: 12
Column: 12
Memory: 46 %
Type here to search
ENG
IN
17:45
23-02-2019

```python
Created on Sat Feb 23 02:01:03 2019

@author: Alfred Zane Rajan
"""

import pandas as pd

support = 50    #float(sys.argv[1])
confidence = 70      #float(sys.argv[2])
file = "db4.txt"     #(sys.argv[3])
df = pd.read_csv( file, skipinitialspace = True, names = ['0', '1','2','3','4','5','6','7','8','9','10'])
df = df.drop('0', axis=1)
df = df.T
for col in df:
    df[col] = df[col].str.replace(" ", "")
db = list(list(df[i]) for i in df)
db = [[i for i in i if not pd.isna(i)] for i in db]
```

Name	Type	Size	Value
col	int	1	19
confidence	float	1	0.7
count	dict	10	{'diapers':14, 'sweaters':8, 'tissues':10, 'belts':12, 'water':15, 'no ...
db	list	20	[['diapers', 'sweaters', 'tissues', 'belts', 'water', ...], ['diapers' ...
df	DataFrame	(10, 20)	Column names: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 ...
file	str	1	db4.txt
flag	int	1	0
item	str	1	pen

```
In [333]: runfile('C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct/source.py', wdir='C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct')
{'diapers'}  ->  water
{'diapers'}  ->  noodles
{'diapers'}  ->  cereals
{'diapers'}  ->  pen
{'diapers'}  ->  batteries
{'belts'}  ->  pen
{'belts'}  ->  batteries
{'water'}  ->  noodles
{'water'}  ->  cereals
{'water'}  ->  pen
{'noodles'}  ->  diapers
{'noodles'}  ->  water
{'noodles'}  ->  cereals
{'noodles'}  ->  pen
{'noodles'}  ->  batteries
{'cereals'}  ->  water
{'cereals'}  ->  pen
{'cereals'}  ->  batteries
{'books'}  ->  pen
{'pen'}  ->  water
{'pen'}  ->  noodles
{'pen'}  ->  cereals
```

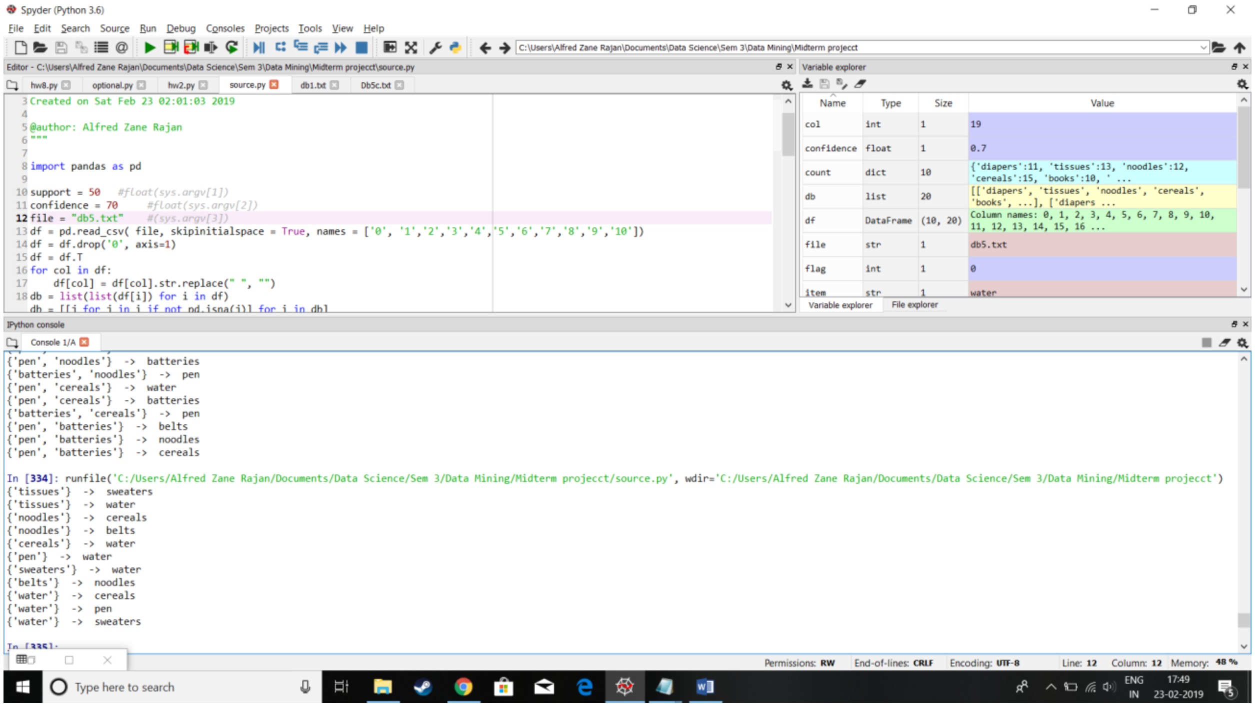
Data Base 5:

Input:

1001, diapers, tissues, noodles, cereals, books, pen, batteries
1002, diapers, sweaters, belts, water, noodles, cereals, books, pen, batteries
1003, sweaters, tissues, water, noodles, cereals, batteries
1004, diapers, sweaters, tissues, cereals
1005, sweaters, tissues, belts, water, noodles, pen
1006, belts, water, noodles, cereals, books, pen, batteries
1007, diapers, sweaters, tissues, water, cereals, pen,
1008, sweaters, tissues, belts, water, noodles, cereals, books, batteries
1009, diapers, tissues, water, cereals, books, pen, batteries
1010, diapers, tissues, belts, water, noodles, cereals, books, pen, batteries
1011, diapers, sweaters, tissues, water, books, pen, batteries
1012, sweaters, tissues, belts, noodles, books, batteries
1013, diapers, sweaters, books, pen, batteries
1014, diapers, belts, noodles, cereals
1015, diapers, sweaters, belts, water, noodles, cereals, books, pen
1016, sweaters, tissues, belts, water, noodles, cereals
1017, diapers, sweaters, tissues, belts, water, pen
1018, sweaters, tissues, water, cereals, pen, batteries
1019, sweaters, belts, water, noodles, cereals, pen, batteries
1010, sweaters, water, cereals

Output:

```
{'tissues'} -> sweaters
{'tissues'} -> water
{'noodles'} -> cereals
{'noodles'} -> belts
{'cereals'} -> water
{'pen'} -> water
{'sweaters'} -> water
{'belts'} -> noodles
{'water'} -> cereals
{'water'} -> pen
{'water'} -> sweaters
```


Spyder (Python 3.6)
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\Alfred Zane Rajan\Documents\Data Science\Sem 3\Data Mining\Midterm projecct
Editor - C:\Users\Alfred Zane Rajan\Documents\Data Science\Sem 3\Data Mining\Midterm projecct\source.py
hw8.py
optional.py
hw2.py
source.py
db1.txt
Db5c.txt
Variable explorer
File explorer
IPython console
Console 1/A
Permissions: RW
End-of-lines: CRLF
Encoding: UTF-8
Line: 12
Column: 12
Memory: 48 %
Type here to search
ENG
IN
17:49
23-02-2019

```
Created on Sat Feb 23 02:01:03 2019

@author: Alfred Zane Rajan
"""

import pandas as pd

support = 50    #float(sys.argv[1])
confidence = 70     #float(sys.argv[2])
file = "db5.txt"     #(sys.argv[3])
df = pd.read_csv( file, skipinitialspace = True, names = ['0', '1','2','3','4','5','6','7','8','9','10'])
df = df.drop('0', axis=1)
df = df.T
for col in df:
    df[col] = df[col].str.replace(" ", "")
db = list(list(df[i]) for i in df)
db = [[i for i in i if not pd.isna(i)] for i in db]
```

Name	Type	Size	Value
col	int	1	19
confidence	float	1	0.7
count	dict	10	{'diapers':11, 'tissues':13, 'noodles':12, 'cereals':15, 'books':10, ' ...
db	list	20	[['diapers', 'tissues', 'noodles', 'cereals', 'books', ...], ['diapers ...
df	DataFrame	(10, 20)	Column names: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 ...
file	str	1	db5.txt
flag	int	1	0
item	str	1	water

```
{'pen', 'noodles'}  ->  batteries
{'batteries', 'noodles'}  ->  pen
{'pen', 'cereals'}  ->  water
{'pen', 'cereals'}  ->  batteries
{'batteries', 'cereals'}  ->  pen
{'pen', 'batteries'}  ->  belts
{'pen', 'batteries'}  ->  noodles
{'pen', 'batteries'}  ->  cereals

In [334]: runfile('C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct/source.py', wdir='C:/Users/Alfred Zane Rajan/Documents/Data Science/Sem 3/Data Mining/Midterm projecct')
{'tissues'}  ->  sweaters
{'tissues'}  ->  water
{'noodles'}  ->  cereals
{'noodles'}  ->  belts
{'cereals'}  ->  water
{'pen'}  ->  water
{'sweaters'}  ->  water
{'belts'}  ->  noodles
{'water'}  ->  cereals
{'water'}  ->  pen
{'water'}  ->  sweaters

In [335]:
```