

MIDTERM PROJECT MILESTONE REPORT

1. Software:

I'm using python version 3.8 on Jupyter lab with some libraries in my codes:

- Numpy
- Pandas
- Itertools
- Jupyter lab

2. Data:

I used the data that the professor provided: amazon, bestbuy, kmart, nike and custom data

3. Hardware configuration:

I code this project on Macbook – my personal laptop.

4. Status:

I already done coding. Here's a brief preview of my work so far.

```
In [1]: import pandas as pd
import numpy as np
import os
import itertools
```

1. Load Data

```
In [2]: print("Welcome to the simulation of the apriori algorithms. \nPlease choose t
while True:
    choice_of_data = input()
    if (choice_of_data == '1'):
        data = 'amazon.csv'
        print('User chose amazon dataset')
        break
    elif (choice_of_data == '2'):
        data = 'bestbuy.csv'
        print('User chose bestbuy dataset')
        break
    elif (choice_of_data == '3'):
        data = 'kmart.csv'
        print('User chose kmart dataset')
        break
    elif (choice_of_data == '4'):
        data = 'nike.csv'
        print('User chose nike dataset')
        break
    elif (choice_of_data == '5'):
        data = 'custom.csv'
        print('User chose custom dataset')
        break
    else:
        print("Input is not recognized, please input the number corresponding
print("Loading and processing data ... ")
```

Welcome to the simulation of the apriori algorithms.

Please choose the data:

- amazon (1)
- bestbuy (2)
- kmart (3)
- nike (4)
- custom (5)

User chose amazon dataset

Loading and processing data ...

```
In [3]: filepath = 'data'
file = os.path.join(filepath, data)
df = pd.read_csv(file)
df.head()
```

Out[3]:

	TransactionID	Transaction
0	Trans1	A Beginner's Guide, Java: The Complete Referen...
1	Trans2	A Beginner's Guide, Java: The Complete Referen...
2	Trans3	A Beginner's Guide, Java: The Complete Referen...
3	Trans4	Android Programming: The Big Nerd Ranch, Head ...
4	Trans5	Android Programming: The Big Nerd Ranch, Begin...

```
In [4]: list_of_items = []
for tran in df[df.columns[-1]]:
    for item in str(tran).replace(u'\xa0', u' ').split(', '):
        list_of_items.append(item)
list_of_items = set(list_of_items)
items = {}
i = 1
for item in list_of_items:
    items[item] = i
    i += 1
items
```

```
Out[4]: {'HTML and CSS: Design and Build Websites': 1,
'A Beginner's Guide': 2,
'Head First Java 2nd Edition ': 3,
'Head First Java 2nd Edition': 4,
'Android Programming: The Big Nerd Ranch': 5,
'Beginning Programming with Java': 6,
'Java 8 Pocket Guide': 7,
'Android Programming: The Big Nerd Ranch, ': 8,
'Java: The Complete Reference': 9,
'Java For Dummies': 10,
'C++ Programming in Easy Steps': 11,
'Beginning Programming with Java, ': 12}
```

```
In [5]: for item in items:
    arr = np.zeros(df.shape[0]).astype(int)
    i = 0
    for tran in df[df.columns[1]]:
        if item in str(tran).replace(u'\xa0', u' ').split(', '):
            arr[i] = 1
            i += 1
    df[item] = arr
df.head()
```

Out[5]:

	TransactionID	Transaction	HTML and CSS: Design and Build Websites	A Beginner's Guide	Head First Java 2nd Edition	Head First Java 2nd Edition	Android Programming: The Big Nerd Ranch	Begin Program with
0	Trans1	A Beginner's Guide, Java: The Complete Referen...	0	1	0	0	1	
1	Trans2	A Beginner's Guide, Java: The Complete Referen...	0	1	0	0	0	
2	Trans3	A Beginner's Guide, Java: The Complete Referen...	0	1	0	1	1	
3	Trans4	Android Programming: The Big Nerd Ranch, Head ...	0	0	1	0	1	
4	Trans5	Android Programming: The Big Nerd Ranch, Begin...	0	0	0	0	1	

2. Input Support & Confidence

```
In [6]: print("Done processing the data!")
print("Please choose the support rate (>= 0 & <= 1)")
while True:
    try:
        support_rate = float(input())
        if (support_rate < 0 or support_rate > 1):
            print("input is not in range (0,1)")
            continue
        break
    except ValueError:
        print("input is not float, please input in range (0,1)")
print("User chose support rate: %f" % support_rate)
```

Done processing the data!
 Please choose the support rate (≥ 0 & ≤ 1)
 User chose support rate: 0.200000

```
In [7]: print("Please choose the confidence rate ( $\geq 0$  &  $\leq 1$ )")
while True:
    try:
        confidence_rate = float(input())
        if (confidence_rate < 0 or confidence_rate > 1):
            print("input is not in range (0,1)")
            continue
        break
    except ValueError:
        print("input is not float, please input in range (0,1)")
print("User chose confidence rate: %f" % confidence_rate)
```

Please choose the confidence rate (≥ 0 & ≤ 1)
 User chose confidence rate: 0.400000

```
In [8]: min_support = support_rate*df.shape[0]
```

3. Apriori Algorithms

```
In [9]: count_dict = {}
for item in list_of_items:
    if np.sum(df[item]) >= min_support:
        count_dict[item] = np.sum(df[item])
set(count_dict.keys())
```

```
Out[9]: {'A Beginner's Guide',
'Android Programming: The Big Nerd Ranch',
'Beginning Programming with Java',
'Head First Java 2nd Edition',
'Head First Java 2nd Edition ',
'Java 8 Pocket Guide',
'Java For Dummies',
'Java: The Complete Reference'}
```

```
In [10]: list_frequent = []
for i in range(2, len(count_dict) + 1):
    ls = list(itertools.combinations(count_dict.keys(), i))
    for per in ls:
        if (np.sum(np.sum(df[list(per)], axis = 1) == i) >= int(min_support))
            list_frequent.append(per)
    # print(len(list_frequent), list_frequent, per)
```

```
In [11]: for fred in list_frequent:
    print("For frequent list,", fred, "we have rules:")
    sp2 = np.sum(np.sum(df[list(fred)], axis = 1) == len(list(fred)))
    # print("Support: ", sp2)
    ls = list(itertools.permutations(fred, len(fred)))
    for per in ls:
        passed = False
        sp1 = np.sum(np.sum(df[list(per)[:-1]], axis = 1) == len(list(list(per)[:-1])))
        if (sp2/sp1 >= confidence_rate):
            passed = True
        if(passed):
            res = "passed"
        else:
            res = "not passed"
    print(" - ", list(per)[:-1], "=>", list(per)[-1], "with confidence:", sp2/sp1)
```

For frequent list, ('A Beginner's Guide', 'Android Programming: The Big Nerd Ranch') we have rules:

- ['A Beginner's Guide'] => Android Programming: The Big Nerd Ranch with confidence: 0.6 (passed)
- ['Android Programming: The Big Nerd Ranch'] => A Beginner's Guide with confidence: 0.5 (passed)

For frequent list, ('A Beginner's Guide', 'Java: The Complete Reference') we have rules:

- ['A Beginner's Guide'] => Java: The Complete Reference with confidence: 0.8 (passed)
- ['Java: The Complete Reference'] => A Beginner's Guide with confidence: 0.8888888888888888 (passed)

For frequent list, ('A Beginner's Guide', 'Java For Dummies') we have rules:

- ['A Beginner's Guide'] => Java For Dummies with confidence: 0.8 (passed)
- ['Java For Dummies'] => A Beginner's Guide with confidence: 0.6666666666666666 (passed)

For frequent list, ('Head First Java 2nd Edition', 'Android Programming: The Big Nerd Ranch') we have rules:

- ['Head First Java 2nd Edition'] => Android Programming: The Big Nerd Ranch with confidence: 1.0 (passed)
- ['Android Programming: The Big Nerd Ranch'] => Head First Java 2nd Edition with confidence: 0.3333333333333333 (not passed)

For frequent list, ('Android Programming: The Big Nerd Ranch', 'Java: The Complete Reference') we have rules:

- ['Android Programming: The Big Nerd Ranch'] => Java: The Complete Reference with confidence: 0.4166666666666667 (passed)
- ['Java: The Complete Reference'] => Android Programming: The Big Nerd Ranch with confidence: 0.5555555555555556 (passed)

For frequent list, ('Android Programming: The Big Nerd Ranch', 'Java For Dummies') we have rules:

- ['Android Programming: The Big Nerd Ranch'] => Java For Dummies with confidence: 0.6666666666666666 (passed)
- ['Java For Dummies'] => Android Programming: The Big Nerd Ranch with confidence: 0.6666666666666666 (passed)

For frequent list, ('Beginning Programming with Java', 'Java 8 Pocket Guide') we have rules:

- ['Beginning Programming with Java'] => Java 8 Pocket Guide with confidence: 0.75 (passed)
- ['Java 8 Pocket Guide'] => Beginning Programming with Java with confidence: 0.75 (passed)

```

: 0.75 (passed)
For frequent list, ('Java: The Complete Reference', 'Java For Dummies') we have rules:
- ['Java: The Complete Reference'] => Java For Dummies with confidence: 1.0 (passed)
- ['Java For Dummies'] => Java: The Complete Reference with confidence: 0.75 (passed)
For frequent list, ('A Beginner's Guide', 'Android Programming: The Big Nerd Ranch', 'Java: The Complete Reference') we have rules:
- ['A Beginner's Guide', 'Android Programming: The Big Nerd Ranch'] => Java: The Complete Reference with confidence: 0.8333333333333334 (passed)
- ['A Beginner's Guide', 'Java: The Complete Reference'] => Android Programming: The Big Nerd Ranch with confidence: 0.625 (passed)
- ['Android Programming: The Big Nerd Ranch', 'A Beginner's Guide'] => Java: The Complete Reference with confidence: 0.8333333333333334 (passed)
- ['Android Programming: The Big Nerd Ranch', 'Java: The Complete Reference'] => A Beginner's Guide with confidence: 1.0 (passed)
- ['Java: The Complete Reference', 'A Beginner's Guide'] => Android Programming: The Big Nerd Ranch with confidence: 0.625 (passed)
- ['Java: The Complete Reference', 'Android Programming: The Big Nerd Ranch'] => A Beginner's Guide with confidence: 1.0 (passed)
For frequent list, ('A Beginner's Guide', 'Android Programming: The Big Nerd Ranch', 'Java For Dummies') we have rules:
- ['A Beginner's Guide', 'Android Programming: The Big Nerd Ranch'] => Java For Dummies with confidence: 0.8333333333333334 (passed)
- ['A Beginner's Guide', 'Java For Dummies'] => Android Programming: The Big Nerd Ranch with confidence: 0.625 (passed)
- ['Android Programming: The Big Nerd Ranch', 'A Beginner's Guide'] => Java For Dummies with confidence: 0.8333333333333334 (passed)
- ['Android Programming: The Big Nerd Ranch', 'Java For Dummies'] => A Beginner's Guide with confidence: 0.625 (passed)
- ['Java For Dummies', 'A Beginner's Guide'] => Android Programming: The Big Nerd Ranch with confidence: 0.625 (passed)
- ['Java For Dummies', 'Android Programming: The Big Nerd Ranch'] => A Beginner's Guide with confidence: 0.625 (passed)
For frequent list, ('A Beginner's Guide', 'Java: The Complete Reference', 'Java For Dummies') we have rules:
- ['A Beginner's Guide', 'Java: The Complete Reference'] => Java For Dummies with confidence: 1.0 (passed)
- ['A Beginner's Guide', 'Java For Dummies'] => Java: The Complete Reference with confidence: 1.0 (passed)
- ['Java: The Complete Reference', 'A Beginner's Guide'] => Java For Dummies with confidence: 1.0 (passed)
- ['Java: The Complete Reference', 'Java For Dummies'] => A Beginner's Guide with confidence: 0.8888888888888888 (passed)
- ['Java For Dummies', 'A Beginner's Guide'] => Java: The Complete Reference with confidence: 1.0 (passed)
- ['Java For Dummies', 'Java: The Complete Reference'] => A Beginner's Guide with confidence: 0.8888888888888888 (passed)
For frequent list, ('Android Programming: The Big Nerd Ranch', 'Java: The Complete Reference', 'Java For Dummies') we have rules:
- ['Android Programming: The Big Nerd Ranch', 'Java: The Complete Reference'] => Java For Dummies with confidence: 1.0 (passed)
- ['Android Programming: The Big Nerd Ranch', 'Java For Dummies'] => Java: The Complete Reference with confidence: 0.625 (passed)
- ['Java: The Complete Reference', 'Android Programming: The Big Nerd Ranch'] => Java For Dummies with confidence: 1.0 (passed)

```

```

- ['Java: The Complete Reference', 'Java For Dummies'] => Android Programming: The Big Nerd Ranch with confidence: 0.5555555555555556 (passed)
- ['Java For Dummies', 'Android Programming: The Big Nerd Ranch'] => Java: The Complete Reference with confidence: 0.625 (passed)
- ['Java For Dummies', 'Java: The Complete Reference'] => Android Programming: The Big Nerd Ranch with confidence: 0.5555555555555556 (passed)
For frequent list, ('A Beginner's Guide', 'Android Programming: The Big Nerd Ranch', 'Java: The Complete Reference', 'Java For Dummies') we have rules:
- ['A Beginner's Guide', 'Android Programming: The Big Nerd Ranch', 'Java: The Complete Reference'] => Java For Dummies with confidence: 1.0 (passed)
- ['A Beginner's Guide', 'Android Programming: The Big Nerd Ranch', 'Java For Dummies'] => Java: The Complete Reference with confidence: 1.0 (passed)
- ['A Beginner's Guide', 'Java: The Complete Reference', 'Android Programming: The Big Nerd Ranch'] => Java For Dummies with confidence: 1.0 (passed)
- ['A Beginner's Guide', 'Java: The Complete Reference', 'Java For Dummies'] => Android Programming: The Big Nerd Ranch with confidence: 0.625 (passed)
- ['A Beginner's Guide', 'Java For Dummies', 'Android Programming: The Big Nerd Ranch'] => Java: The Complete Reference with confidence: 1.0 (passed)
- ['A Beginner's Guide', 'Java For Dummies', 'Java: The Complete Reference'] => Android Programming: The Big Nerd Ranch with confidence: 0.625 (passed)
- ['Android Programming: The Big Nerd Ranch', 'A Beginner's Guide', 'Java: The Complete Reference'] => Java For Dummies with confidence: 1.0 (passed)
- ['Android Programming: The Big Nerd Ranch', 'A Beginner's Guide', 'Java For Dummies'] => Java: The Complete Reference with confidence: 1.0 (passed)
- ['Android Programming: The Big Nerd Ranch', 'A Beginner's Guide', 'Java For Dummies'] => Java: The Complete Reference with confidence: 1.0 (passed)
- ['Android Programming: The Big Nerd Ranch', 'Java: The Complete Reference', 'A Beginner's Guide'] => Java For Dummies with confidence: 1.0 (passed)
- ['Android Programming: The Big Nerd Ranch', 'Java: The Complete Reference', 'Java For Dummies'] => A Beginner's Guide with confidence: 1.0 (passed)
- ['Android Programming: The Big Nerd Ranch', 'Java For Dummies', 'A Beginner's Guide'] => Java: The Complete Reference with confidence: 1.0 (passed)
- ['Android Programming: The Big Nerd Ranch', 'Java For Dummies', 'Java: The Complete Reference'] => A Beginner's Guide with confidence: 1.0 (passed)
- ['Java: The Complete Reference', 'A Beginner's Guide', 'Android Programming: The Big Nerd Ranch'] => Java For Dummies with confidence: 1.0 (passed)
- ['Java: The Complete Reference', 'A Beginner's Guide', 'Java For Dummies'] => Android Programming: The Big Nerd Ranch with confidence: 0.625 (passed)
- ['Java: The Complete Reference', 'Android Programming: The Big Nerd Ranch', 'A Beginner's Guide'] => Java For Dummies with confidence: 1.0 (passed)
- ['Java: The Complete Reference', 'Android Programming: The Big Nerd Ranch', 'Java For Dummies'] => A Beginner's Guide with confidence: 1.0 (passed)
- ['Java: The Complete Reference', 'Java For Dummies', 'A Beginner's Guide'] => Android Programming: The Big Nerd Ranch with confidence: 0.625 (passed)
- ['Java: The Complete Reference', 'Java For Dummies', 'Android Programming: The Big Nerd Ranch'] => A Beginner's Guide with confidence: 1.0 (passed)
- ['Java For Dummies', 'A Beginner's Guide', 'Android Programming: The Big Nerd Ranch'] => Java: The Complete Reference with confidence: 1.0 (passed)
- ['Java For Dummies', 'A Beginner's Guide', 'Java: The Complete Reference'] => Android Programming: The Big Nerd Ranch with confidence: 0.625 (passed)
- ['Java For Dummies', 'Android Programming: The Big Nerd Ranch', 'A Beginner's Guide'] => Java: The Complete Reference with confidence: 1.0 (passed)
- ['Java For Dummies', 'Android Programming: The Big Nerd Ranch', 'Java: The Complete Reference'] => A Beginner's Guide with confidence: 1.0 (passed)
- ['Java For Dummies', 'Java: The Complete Reference', 'A Beginner's Guide'] => Android Programming: The Big Nerd Ranch with confidence: 0.625 (passed)
- ['Java For Dummies', 'Java: The Complete Reference', 'Android Programming: The Big Nerd Ranch'] => A Beginner's Guide with confidence: 1.0 (passed)

```


In []: