Apriori_Algorithm

Tree Directory

data: Has all the csv files for the stores

Milestones: Has a pdf of the previous milestone documentation

Screenshots: Contains screenshots/images of outputs from running the py and ipynb file and is used in the README.md file

src: A source folder of a python script to generate a csv file of random transactions given store and items

Aprior_Algo.ipynb: The Jupyter Notebook

Documentation.ipynb: Jupyter Notebook of the creation of the README.md file

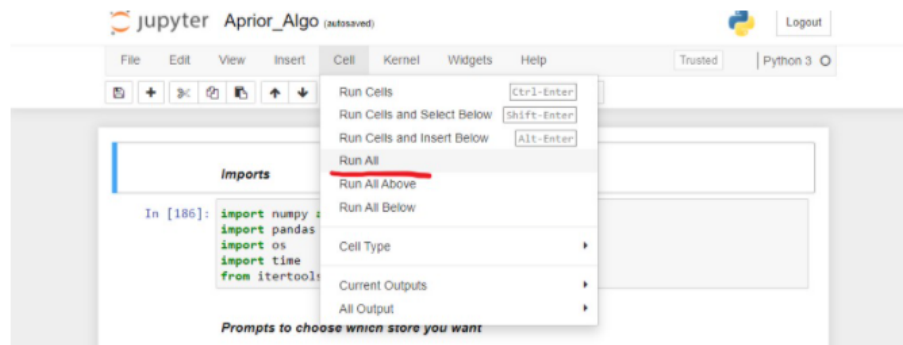**Aprior_Algo.py** : py script version of the algorithm

LICENSE: MIT LICENSE

Implementation
After learning the subject of the Apriori Algorithm in Data Mining, it was in my best interest to use the concept of sets, counters, dataframes, and lambdas, loops, and well defined functions to carry out specific tasks to achieve the Apriori Algorithm.

Running Aprior_Algo.ipynb (Jupyter Notebook)

Make sure to run the ENTIRE workbook

Jupyter Aprior_Algo (autosaved)   Logout

| File | Edit | View | Insert | Cell | Kernel | Widgets | Help | | Trusted | Python 3 ○ |

Cell menu:
- Run Cells .... Ctrl-Enter
- Run Cells and Select Below .... Shift-Enter
- Run Cells and Insert Below .... Alt-Enter
- Run All
- Run All Above
- Run All Below
- Cell Type ▸
- Current Outputs ▸
- All Output ▸

Imports

```
In [186]: import numpy a
          import pandas
          import os
          import time
          from itertool
```

Prompts to choose which store you want

Answer the following prompts

Prompts to choose which store you want

```
In [*]: print("Welcome to Apriori 2.0!")
        store_num = input("Please select your store \n 1. Amazon \n 2. Nike \n 3. Best B
        print(store_num)
        support_percent = input("Please enter the percentage of Support you want?\n")
        print(support_percent)
        confidence_percent = input("Please enter the percentage of Confidence you want?\
        print(confidence_percent)
```

```
Welcome to Apriori 2.0!
Please select your store
 1. Amazon
 2. Nike
 3. Best Buy
 4. K-Mart
 5. Walmart
1
1
Please enter the percentage of Support you want?
50
50
Please enter the percentage of Confidence you want?
50
```

Scroll to the end of the notebook and Viola!! You have your associations!!

Final Associations

```
In [143]: print("Store Name: "+ str(ns(int(store_num))))
          print("\nFinal Associations that meet the user standards....")
          print("Support: " + str(support_percent) + "%" + "\t" + "Confidence: " + str(con
          df_final
```

```
Store Name: Amazon

Final Associations that meet the user standards....
Support: 50%    Confidence: 50%
```

Out[143]:

	Association	Support	Confidence
0	(Java: The Complete Reference, Java For Dummies)	50.0	100.0
1	(Java For Dummies, Java: The Complete Reference)	50.0	77.0

Example (Jupyter Notebook)

This will be with a different store with Best Buy, Support = 40% and Confidence = 70%

Screenshot 1

Prompts to choose which store you want

```
[145]: print("Welcome to Apriori 2.0!")
       store_num = input("Please select your store \n 1. Amazon \n 2. Nike \n 3. Best B
       print(store_num)
       support_percent = input("Please enter the percentage of Support you want?\n")
       print(support_percent)
       confidence_percent = input("Please enter the percentage of Confidence you want?\
       print(confidence_percent)
```

```
Welcome to Apriori 2.0!
Please select your store
 1. Amazon
 2. Nike
 3. Best Buy
 4. K-Mart
 5. Walmart
3
3
Please enter the percentage of Support you want?
40
40
Please enter the percentage of Confidence you want?
70
70
```

Screenshot 2

Store Name: Best Buy

Final Associations that meet the user standards....
Support: 40%   Confidence: 70%

Out[162]:

	Association	Support	Confidence
0	(Lab Top Case, Lab Top)	50.0	71.0
1	(Lab Top, Lab Top Case)	50.0	83.0
2	(Anti-Virus, Lab Top)	50.0	71.0
3	(Lab Top, Anti-Virus)	50.0	83.0
4	(Printer, Flash Drive)	50.0	100.0
5	(Flash Drive, Printer)	50.0	77.0
6	(Printer, Microsoft Office)	45.0	90.0
7	(Microsoft Office, Printer)	45.0	82.0
8	(Flash Drive, Microsoft Office)	55.0	85.0
9	(Microsoft Office, Flash Drive)	55.0	100.0
10	(Flash Drive, Anti-Virus)	50.0	77.0
11	(Anti-Virus, Flash Drive)	50.0	71.0
12	(Microsoft Office, Anti-Virus)	40.0	73.0
13	(Speakers, Lab Top Case)	45.0	82.0
14	(Speakers, Anti-Virus)	45.0	82.0
15	(Lab Top Case, Anti-Virus)	60.0	86.0
16	(Anti-Virus, Lab Top Case)	60.0	86.0
17	(Lab Top Case, Anti-Virus, Lab Top)	45.0	75.0
18	(Lab Top, Lab Top Case, Anti-Virus)	45.0	90.0
19	(Lab Top, Anti-Virus, Lab Top Case)	45.0	90.0
20	(Printer, Flash Drive, Microsoft Office)	45.0	90.0
21	(Printer, Microsoft Office, Flash Drive)	45.0	100.0
22	(Flash Drive, Microsoft Office, Printer)	45.0	82.0
23	(Flash Drive, Anti-Virus, Microsoft Office)	40.0	80.0
24	(Flash Drive, Microsoft Office, Anti-Virus)	40.0	73.0
25	(Microsoft Office, Anti-Virus, Flash Drive)	40.0	100.0
26	(Lab Top Case, Anti-Virus, Flash Drive)	45.0	75.0
27	(Flash Drive, Lab Top Case, Anti-Virus)	45.0	100.0
28	(Flash Drive, Anti-Virus, Lab Top Case)	45.0	90.0
29	(Speakers, Lab Top Case, Anti-Virus)	40.0	89.0
30	(Speakers, Anti-Virus, Lab Top Case)	40.0	89.0

Running Aprior_Algo.py (Python File)

run the following command in the project directory

python Aprior_Algo.py

Follow and answer the following prompts

```
(base) C:\Users\PScout\Desktop\Github\Apriori_Algorithm>python Aprior_Algo.py
Welcome to Apriori 2.0!
Please select your store
 1. Amazon
 2. Nike
 3. Best Buy
 4. K-Mart
 5. Walmart
1
1
Please enter the percentage of Support you want?
50
50
Please enter the percentage of Confidence you want?
50
50
   transaction id                                transaction
```

Then at the end of the terminal you will get your final associations

```
Store Name: Amazon
Final Associations that meet the user standards....
Support: 50%    Confidence: 50%
                                        Association  Support  Confidence
0  (Java: The Complete Reference, Java For Dummies)     50.0       100.0
1  (Java For Dummies, Java: The Complete Reference)     50.0        77.0
```

EXAMPLE

This will be with a different store with Best Buy, Support = 40% and Confidence = 70%

Screenshot 3

```
(base) C:\Users\PScout\Desktop\Github\Apriori_Algorithm>python Aprior_Algo.py
Welcome to Apriori 2.0!
Please select your store
 1. Amazon
 2. Nike
 3. Best Buy
 4. K-Mart
 5. Walmart
3
3
Please enter the percentage of Support you want?
40
40
Please enter the percentage of Confidence you want?
70
70
```

Screenshot 4

```
Store Name: Best Buy
Final Associations that meet the user standards....
Support: 40%    Confidence: 70%
                                  Association  Support  Confidence
0                    (Lab Top Case, Lab Top)     50.0        71.0
1                    (Lab Top, Lab Top Case)     50.0        83.0
2                     (Anti-Virus, Lab Top)      50.0        71.0
3                     (Lab Top, Anti-Virus)      50.0        83.0
4                    (Printer, Flash Drive)      50.0       100.0
5                    (Flash Drive, Printer)      50.0        77.0
6                 (Printer, Microsoft Office)    45.0        90.0
7                 (Microsoft Office, Printer)    45.0        82.0
8              (Microsoft Office, Flash Drive)   55.0       100.0
9              (Flash Drive, Microsoft Office)   55.0        85.0
10                  (Anti-Virus, Flash Drive)    50.0        71.0
11                  (Flash Drive, Anti-Virus)    50.0        77.0
12              (Microsoft Office, Anti-Virus)   40.0        73.0
13                 (Speakers, Lab Top Case)      45.0        82.0
14                   (Speakers, Anti-Virus)      45.0        82.0
15                (Anti-Virus, Lab Top Case)     60.0        86.0
16                (Lab Top Case, Anti-Virus)     60.0        86.0
17        (Lab Top Case, Anti-Virus, Lab Top)    45.0        75.0
18        (Lab Top, Lab Top Case, Anti-Virus)    45.0        90.0
19        (Lab Top, Anti-Virus, Lab Top Case)    45.0        90.0
20     (Printer, Microsoft Office, Flash Drive)  45.0       100.0
21     (Printer, Flash Drive, Microsoft Office)  45.0        90.0
22     (Flash Drive, Microsoft Office, Printer)  45.0        82.0
23  (Microsoft Office, Anti-Virus, Flash Drive)  40.0       100.0
24  (Flash Drive, Anti-Virus, Microsoft Office)  40.0        80.0
25  (Flash Drive, Microsoft Office, Anti-Virus)  40.0        73.0
26       (Lab Top Case, Anti-Virus, Flash Drive) 45.0        75.0
27       (Flash Drive, Anti-Virus, Lab Top Case) 45.0        90.0
28       (Flash Drive, Lab Top Case, Anti-Virus) 45.0       100.0
29         (Speakers, Anti-Virus, Lab Top Case)  40.0        89.0
30         (Speakers, Lab Top Case, Anti-Virus)  40.0        89.0
```

Enjoy the Aprior Algorithm 2.0

Refer to the README file on github!
Git Repo Link:
https://github.com/MichaelWoo-git/Apriori_Algorithm