Homework 07 Product Recommendation

BADS7105 – CRM Analytics and Intelligence Yupalapas Panomvongkasem

01 Data preparation

Importing libraries

```
import networkx as nx
import pandas as pd
import numpy as np
import seaborn as sns
import sklearn
from sklearn.decomposition import TruncatedSVD
from scipy.spatial.distance import cosine
```

[1] !pip install -U --pre matplotlib

```
Requirement already up-to-date: matplotlib in /usr/local/lib/python3.7/dist-packages (3.4.2)

Requirement already satisfied, skipping upgrade: cycler>=0.10 in /usr/local/lib/python3.7/dist-packages (from matplotlib) (0.10.0)

Requirement already satisfied, skipping upgrade: kiwisolver>=1.0.1 in /usr/local/lib/python3.7/dist-packages (from matplotlib) (1.3.1)

Requirement already satisfied, skipping upgrade: pyparsing>=2.2.1 in /usr/local/lib/python3.7/dist-packages (from matplotlib) (2.4.7)

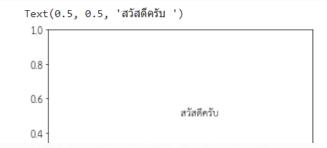
Requirement already satisfied, skipping upgrade: python-dateutil>=2.7 in /usr/local/lib/python3.7/dist-packages (from matplotlib) (2.8.1)

Requirement already satisfied, skipping upgrade: pillow>=6.2.0 in /usr/local/lib/python3.7/dist-packages (from matplotlib) (7.1.2)

Requirement already satisfied, skipping upgrade: numpy>=1.16 in /usr/local/lib/python3.7/dist-packages (from matplotlib) (1.19.5)

Requirement already satisfied, skipping upgrade: six in /usr/local/lib/python3.7/dist-packages (from cycler>=0.10->matplotlib) (1.15.0)
```

- #Adding TH font for visualisation import matplotlib as mpl mpl.__version__ # ต้อง 3.2 ขึ้น ถึงจะเรียก addfont() ได้
- [→ '3.4.2'
- [3] !wget -q https://github.com/Phonbopit/sarabun-webfont/raw/master/fonts/thsarabunnew-webfont.ttf
- [4] mpl.font_manager.fontManager.addfont('thsarabunnew-webfont.ttf')
- [5] mpl.rc('font', family='TH Sarabun New', size=12)
- [6] plt = mpl.pyplot
- [7] #Test Font plt.text(0.5, 0.5, 'สวัสดีครับ ', fontsize=12)



Data importing and pre-processing

```
[25] df = pd.read_excel("Customer Survey.xlsx")
    df.rename(columns = {'Unnamed: 0':'date'}, inplace=True)
    df.head()
```

	Transaction	จักรยาน เสือ ภูเขา	คอ ล ลา เจน	Brownie อวกาศ	แม่ โขง	Botox	กัญชา	กัญ ชง	เครื่อง วัด ความ ดัน	หลอด ไฟ เปลี่ยน ส็ได้	คอร์ส bikini wax	เก้าอี้ นวด ไฟฟ้า	ต้นไม้ ประดับ ห้อง	หมอนข้าง	ตาชั่ง	คอร์ส โบท็ อก กราม	แสง โสม	กะหรี่(ปั๊บ)	Lego	แฟน ที่แปล ว่า พัดลม	ยาด อง
0	1	ไม่เคย	ไม่ เคย	ไม่เคย	ไม่ เคย	ไม่เคย	ไม่เคย	ไม่ เคย	ไม่เคย	ไม่เคย	ไม่เคย	ไม่เคย	ไม่เคย	เคย	ไม่ เคย	ไม่ เคย	ไม่ เคย	เคย	เคย	เคย	ไม่ เคย
1	2	ไม่เคย	เคย	ไม่เคย			ไม่เคย		เคย	ไม่เคย	ไม่เคย	ไม่เคย	ไม่เคย	เคย	เคย	ไม่ เคย	เคย	เคย	ไม่ เคย	เคย	ไม่ เคย
2	3	เคย	ไม่ เคย	ไม่เคย	ไม่ เคย	ไม่เคย	ไม่เคย	ไม่ เคย	เคย	เคย	ไม่เคย	ไม่เคย	เคย	เคย	เคย	ไม่ เคย	ไม่ เคย	เคย	เคย	เคย	ไม่ เคย
3	4	ไม่เคย	ไม่ เคย	ไม่เคย	เคย	ไม่เคย	เคย	ไม่ เคย	ไม่เคย	ไม่เคย	ไม่เคย	เคย	เคย	เคย	เคย	ไม่ เคย	เคย	เคย	เคย	เคย	เคย
4	5	ไม่เคย	เคย	ไม่เคย	ไม่ เคย	เคย	ไม่เคย	ไม่ เคย	ไม่เคย	ไม่เคย	ไม่เคย	ไม่เคย	เคย	ไม่เคย	ไม่ เคย	เคย	ไม่ เคย	เคย	ไม่ เคย	เคย	ไม่ • เคย

<class 'pandas.core.frame.DataFrame'>

```
RangeIndex: 44 entries, 0 to 43
Data columns (total 62 columns):
     Column
                                         Non-Null Count Dtype
     Transaction
                                         44 non-null
                                                          int64
     จักรยานเสือภเขา
                                     44 non-null
                                                       object
     คอลลาเจน
                                         44 non-null
                                                          object
     Brownie อวกาศ
                                         44 non-null
                                                          object
     แม่โขง
                                        44 non-null
                                                         object
                                                          object
 5
     Botox
                                         44 non-null
     กัญชา
 6
                                         44 non-null
                                                          object
     กัญชง
                                         44 non-null
                                                          object
     เครื่องวัดความดัน
                                    44 non-null
                                                      object
     หลอดไฟเปลี่ยนสีได้
                                      44 non-null
                                                       object
     คอร์ส bikini wax
                                                         object
                                        44 non-null
     เก้าอื้นวดไฟฟ้า
11
                                     44 non-null
                                                      object
     ต้นไม้ประดับห้อง
                                     44 non-null
                                                       object
     หมอนข้าง
                                        44 non-null
                                                          object
     ตาชั้ง
                                       44 non-null
                                                        object
14
     คอร์สโบท็อกกราม
                                       44 non-null
                                                         object
     แสงโสม
16
                                         44 non-null
                                                           object
     กะหรี(ปั้บ)
17
                                     44 non-null
                                                      object
                                                          object
18
     Lego
                                         44 non-null
     แฟน ที่แปลว่าพัดลม
 19
                                      44 non-null
                                                        object
                                                          object
 20
     ยาดอง
                                         44 non-null
     ชดนอนไม่ได้นอน
                                       44 non-null
                                                        object
     ที่ยกน้ำหนัก (Dumbell)
                                      44 non-null
                                                       object
     เครื่องดูดฝุ่นอัตโนมัติ (robot)
                                 44 non-null
                                                  object
     เนื้อวากิล
                                     44 non-null
                                                       object
     หมแดดเดียว
                                       43 non-null
                                                         object
     drone
                                         44 non-null
                                                          object
 26
```

44	กรรไกรตัดขนจมูก	44 non-null	object
45	PS5	44 non-null	object
46	เครื่องตรวจน้ำตาลในเลือด	43 non-null	object
47	โรตี	44 non-null	object
48	กะปิดอง	44 non-null	object
49	น้องหมา	44 non-null	object
50	แพ็คเกจไทยเที่ยวไทย	43 non-null	object
51	เกมส์ FIFA Online	44 non-null	object
52	มะม่วงเบา	44 non-null	object
53	ห้องน้ำแมวอัตโนมัติ	44 non-null	object
54	กระเป๋า chanel	43 non-null	object
55	ข้าวขาหมูคากิคาจักออนไลน์	44 non-null	object
56	ถั่งเช่า	43 non-null	object
57	หมอนสุขภาพสำหุรับคนนอนไม่หลับ	44 non-null	object
58	อูาหารแมวคลีนเพื่อสุขภาพลดหุ่น	44 non-null	object
59	น้ำส้มดีโด้	44 non-null	object
60	Software ถูกลิขสิทธิ์	44 non-null	object
61	ครีมกันแดด	44 non-null	object
dtyp	es: int64(1), object(61)		
memo	ry usage: 21.4+ KB		
None			

```
[27] #Consider portion of data pd.melt(df, id_vars=['Transaction'])['value'].value_counts()
```

ไม่เคย 1642 เคย 1036 Name: value, dtype: int64

DT.	df.isnull().head(5)
y -	az()caa(5)

	Transaction	จักรยาน เสือ ภูเขา	คอล ลา เจน	Brownie อวกาศ	แม่ โขง	Botox	กัญชา	กัญ ชง	เครื่อง วัด ความ ดัน	หถอด ไฟ เปถี่ยน สีได้	คอร์ส bikini wax	เก้าอี้ นวด ไฟฟ้า	ต้นไม้ ประดับ ห้อง	หมอนข้าง	ตาชั่ง	คอร์ส โบท็ อก กราม	แสง โสม	กะหรี่(ปั๊บ)
0	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False

4

[37] #Remove Null data and drop column for item-item collaborating
 df1 = df.dropna().reset_index(drop=True)
 df2 = df1.drop('Transaction',1)
 df2.head(5)

	จักรยาน เสือ ภูเขา	คอ ถ ถา เจน	Brownie อวกาศ		Botox	กัญชา	กัญ ชง	เครื่อง วัด ความ ดัน	หถอด "ไฟ เปลี่ยน สีได้		เก้าอี้ นวด ไฟฟ้า	ประดับ	หมอนข้าง	ตาชั่ง	คอร์ส โบท็ อก กราม	แสง โสม	กะหรี่(ปั๊บ)	Lego	แฟน ที่แปล ว่า พัดลม	ยาด อง
0	ไม่เคย	ไม่ เคย	ไม่เคย	ไม่ เคย	ไม่เคย	ไม่เคย	ไม่ เคย	ไม่เคย	ไม่เคย	ไม่เคย	ไม่เคย	ไม่เคย	เคย	ไม่ เคย	ไม่ เคย	ไม่ เคย	เคย	เคย	เคย	ไม่ เคย
1	ไม่เคย	เคย	ไม่เคย	เคย	ไม่เคย	ไม่เคย	ไม่ เคย	เคย	ไม่เคย	ไม่เคย	ไม่เคย	ไม่เคย	เคย	เคย	ไม่ เคย	เคย	เคย	ไม่ เคย	เคย	ไม่ เคย
2	เคย	ไม่ เคย	ไม่เคย	ไม่ เคย	ไม่เคย	ไม่เคย	ไม่ เคย	เคย	เคย	ไม่เคย	ไม่เคย	เคย	เคย	เคย	ไม่ เคย	ไม่ เคย	เคย	เคย	เคย	ไม่ เคย

0

0

0

```
[38] df2.shape
     (39, 61)
[39] #Convert String to Value
     for i in df2.columns:
       df2[i] = df2[i].replace(['ไม่เคย'],0)
       df2[i] = df2[i].replace(['เคย'],1)
     df2.head(5)
                                                                                                                                                          ชุด
                                                        เครื่อง
                                                                                                               คอร์ส
                                                                หลอด
                                                                                                                                             แฟน
                  คอ
                                                                        คอร์ส
                                                                                เก้าอื้
                                                                                       ต้นไม้
         จักรยาน
                                                                                                                                                        นอน
                                                                                                                โบท็
                       Brownie แม่
                                                                                                                     แสง
                                                                                                                                            ทีแปล ยาด
            เสือ
                                                                       bikini
                                                                                      ประดับ หมอนข้าง ตาชั่ง
                                                                                                                           กะหรี่(ป็บ) Lego
                                                                                นวด
                                                                                                                                                          ไม่
                                     Botox กัญชา
                                                                เปลี่ยน
                                                                                                                     โสม
                        อวกาศ โขง
                                                         ความ
                                                    ชง
                                                                                                                 อก
                                                                                                                                               ว่า
                                                                                                                                                    อง
                                                                          wax ไฟฟ้า
                                                                                                                                                          ได้
                                                                                        ห้อง
           ภูเขา
                                                                  สีได้
                                                           ดัน
                                                                                                                                            พัดลม
                  เจน
                                                                                                               กราม
                                                                                                                                                        นอน
      0
               0
                    0
                                         0
                                                0
                                                     0
                                                                    0
                                                                            0
                                                                                   0
                                                                                           0
                                         0
                                                0
                                                     0
                                                                    0
                                                                            0
                                                                                   0
                                                                                           0
```

0

0

0

2

3

0

0

0

0

0

0

0

0

0

0

0

0

02 Data Exploratory Analysis

Exploratory Data Analysis

[40] #Explore statistic of data
 df2.describe()

	จักรยาน เสือภูเขา	คอลลาเจน	Brownie อวกาศ	แม่โขง	Botox	กัญชา	กัญชง	เครื่องวัด ความดัน	หลอดไฟ เปลี่ยนสี ได้	คอร์ส bikini wax	เก้าอึ้นวด ไฟฟ้า	ตันไม้ ประดับ ห้อง	หมอนข้าง	ตาชั่ง
count	39.000000	39.000000	39.000000	39.000000	39.000000	39.000000	39.000000	39.000000	39.000000	39.000000	39.000000	39.000000	39.000000	39.000000
mean	0.358974	0.538462	0.102564	0.307692	0.179487	0.153846	0.076923	0.358974	0.256410	0.076923	0.128205	0.666667	0.871795	0.666667
std	0.485971	0.505035	0.307355	0.467572	0.388776	0.365518	0.269953	0.485971	0.442359	0.269953	0.338688	0.477567	0.338688	0.477567
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	0.000000
50%	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.000000	1.000000	1.000000
75%	1.000000	1.000000	0.000000	1.000000	0.000000	0.000000	0.000000	1.000000	0.500000	0.000000	0.000000	1.000000	1.000000	1.000000
max	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

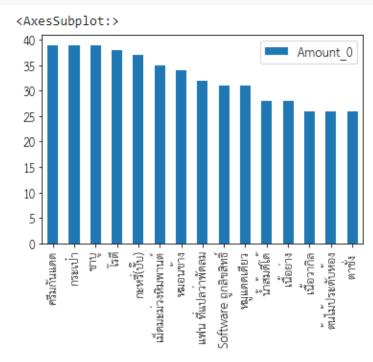
```
[52] #List all products and Sorting
     df3 = df2.sum().sort_values(ascending=False).to_frame()
     df3 = df3.add_prefix('Amount_')
     from google.colab import data_table
     data_table.DataTable(df3)
```

1 to 25 of 61 entries Filter



index	Amount_0
ครีมกันแดด	39
กระเป๋า	39
ชาบู	39
โรตี	38
กะหรี่(ปั๊บ)	37
เม็ดมะม่วงหิมพานต์	35
หมอนข้าง	34
แฟน ที่แปลว่าพัดลม	32
Software ถูกลิขสิทธิ์	31
หมูแดดเดียว	31
น้ำส้มดีโด้	28
เนื้อย่าง	28
เนื้อวากิล	26
ต้นไม้ประดับห้อง	26
ตาชั่ง	26
Board game	26
Lego	25
ที่ยกน้ำหนัก (Dumbell)	24

```
[42] #Visualise top 15
    df3.head(15).plot(kind = "bar")
```



03 Collaborative filtering

item-based collaborative filtering

```
[53] #create item-item matrix
   item_item_matrix = pd.DataFrame(index=df2.columns,columns=df2.columns)
   item_item_matrix.head(5)
```

	จักรยาน เสือ ภูเขา	ลา	Brownie อวกาศ	แม่ โขง	Botox	กัญชา	กัญ ชง	เครื่อง วัด ความ ดัน	หถอด ไฟ เปลี่ยน สีได้	คอร์ส bikini wax	เก้าอี้ นวด ไฟฟ้า	ต้นไม้ ประดับ ห้อง	หมอนข้าง	ตาชั่ง	คอร์ส โบท็ อก กราม	แสง โสม	กะหรี่(ปั๊บ)	Lego	
จักรย เสือภูเ		NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
คอลเ เจน	Nan	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	/
Brow	Nan	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
แม่โร	ขง NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	1
Boto	x NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	/

	จักรยาน เสือภูเขา	คอลลา เจน	Brownie อวกาศ	แม่โขง	Botox	กัญชา	กัญชง	เครื่องวัด ความดัน	หลอดไฟ เปลี่ยนสี ได้	คอร์ส bikini wax	เก้าอึ้นวด ไฟฟ้า	ต้นไม้ ประดับ ห้อง	หมอนข้าง	ตาชั่ง	คอร์สโบ ท็อก กราม
จักรยาน เสือภูเขา	1	0.583212	0.267261	0.385758	0.404061	0.218218	0.154303	0.5	0.507093	0.154303	0.119523	0.524142	0.595854	0.576557	0.267261
คอลลา เจน	0.583212	1	0.327327	0.377964	0.57735	0.178174	0.251976	0.408248	0.345033	0.377964	0.09759	0.641941	0.673633	0.641941	0.436436
Brownie อวกาศ	0.267261	0.327327	1	0.433013	0.377964	0.612372	0.57735	0.133631	0.158114	0.288675	0.447214	0.392232	0.257248	0.294174	0.25
แม่โขง	0.385758	0.377964	0.433013	1	0.218218	0.471405	0.333333	0.385758	0.273861	0.166667	0.516398	0.566139	0.495074	0.396297	0.144338
Botox	0.404061	0.57735	0.377964	0.218218	1	0.154303	0.218218	0.303046	0.358569	0.218218	0.169031	0.2965	0.388922	0.2965	0.755929

links = item_item_matrix.stack().reset_index() links.columns = ['item1', 'item2','cosine_value'] links

	item1	item2	cosine_value
0	จักรยานเสือภูเขา	จักรยานเสื้อภูเขา	1
1	จักรยานเสือภูเขา	คอลลาเจน	0.583212
2	จักรยานเสือภูเขา	Brownie อวกาศ	0.267261
3	จักรยานเสือภูเขา	แม่โขง	0.385758
4	จักรยานเสือภูเขา	Botox	0.404061
3716	ครีมกันแดด	หมอนสุขภาพสำหรับคนนอนไม่หลับ	0.620174
3717	ครีมกันแดด	อาหารแมวคลืนเพื่อสุขภาพลดหุ่น	0.27735
3718	ครีมกันแดด	น้าส้มดีโด้	0.847319
3719	ครีมกันแดด	Software ถูกลิขสิทธิ์	0.891556
3720	ครีมกันแดด	ครีมกันแดด	1
3721 rd	ws × 3 columns		

Product Recommendation

```
#Filter cosine to address weak link
links_filter = links.loc[ (links['cosine_value'] > 0.6) & (links['item1'] != links['item2']) ]
links_filter.sort_values(by='cosine_value',ascending=True)
```

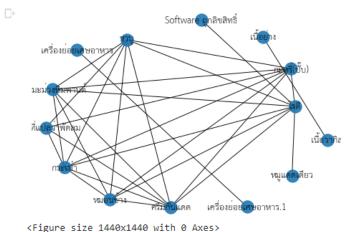
₽		item1	item2	cosine_value
	2300	Omakase	กรรไกรตัดขนจมูก	0.602464
	2660	กรรไกรตัดขนจมูก	Omakase	0.602464
	1507	หมูแดดเดียว	กรรไกรตัดขนจมูก	0.602861
	2647	กรรไกรตัดขนจมูก	หมูแดดเดียว	0.602861
	46	จักรยานเสื้อภูเขา	โรตี	0.606977
	2161	เครื่องย่อยเศษอาหาร.1	เครื่องย่อยเศษอาหาร	1
	2012	ชาบู	ครีมกันแดด	1
	1621	เครื่องย่อยเศษอาหาร	เครื่องย่อยเศษอาหาร.1	1
	1951	กระเป๋า	ครีมกันแดด	1
	3692	ครีมกันแดด	ชาบู	1

504 rows × 3 columns

```
[103] # Build Network graph
import networkx as nx

G=nx.from_pandas_edgelist(links_filter, 'item1', 'item2','cosine_value')
labels_params = {'font_family': 'TH Sarabun New', 'alpha':.9, 'font_size':12}
pos = nx.spring_layout(G, k=5, weight='cosine_value', iterations=80, seed=90, scale=2.5)
#pos = nx.spring_layout(G)
nx.draw(G, pos,**labels_params)
nx.draw_networkx_labels(G, pos,**labels_params)
edge_labels = nx.get_edge_attributes(G,'cosine_value')
plt.figure(figsize=(20,20))

plt.show()
```



04 Interpretation

หลังจาก run model ด้วยการหา Cosine Similarity item-item และ filter item similarity ช่วง > 0.8 และ > 0.85 พบว่า มีหลาย กรณีที่ของทั้งสองอย่างไม่น่าจะมีความสัมพันธ์ หรือมีทิศทางเดียวกัน แต่กลับมีค่า cosine similarity ที่ค่อนข้างสูง เช่น

- คนชอบกระเป๋า chanel มีแนว์โน้มชอบ Brownie อวกาศ 86%, คนซื้อ software ถูกลิขสิทธิ์มีแนว์โน้มชอบกินกะหรี่ปั้ป 85%
- คนชอบกระหรี่ปั๊ป จะชอบเล่น Board game และชอบต้นไม้ประดับห้อง สูงถึง 80%
- 🗠 ส่วนที่ไม่แปลกก็คือ แสงโสม Vs แม่โขง ที่มี cosine 88% เนื่องจากเป็นสิ่งที่มีลักษณะใกล้เคียงกัน

	item1	item2	cosine_value
1069	Lego	ชาบู	0.800641
3677	ครีมกันแดด	Lego	0.800641
1969	ชาบู	Lego	0.800641
1068	Lego	กระเป๋า	0.800641
1908	กระเป๋า	Lego	0.800641
1097	Lego	ครีมกันแดด	0.800641
1127	แฟน ที่แปลว่าพัดลม	เนื้อย่าง	0.801784
1787	เนื้อย่าง	แฟน ที่แปลว่าพัดลม	0.801784
1293	ที่ยกน้ำหนัก (Dumbell)	หมอนข้าง	0.805161
753	หมอนข้าง	ที่ยกน้ำหนัก (Dumbell)	0.805161
2090	Board game	กะหรี่(ปั๊บ)	0.806032
1419	เนื้อวากิล	กะหรี่(ปั๊บ)	0.806032
987	กะหรี่(ปั๊บ)	ต้นไม้ประดับห้อง	0.806032
687	ต้นไม้ประดับห้อง	กะหรี่(ปั๊บ)	0.806032
1010	กะหรี่(ปั๊บ)	Board game	0.806032

	item1	item2	cosine_value
1035	กะหรี่(บั๊บ)	Software ถูกลิขสิทธิ์	0.856281
3615	Software ถูกลิขสิทธิ์	กะหรี่(ปั๊บ)	0.856281
1815	เนื้อย่าง	โรตี	0.858395
2835	โรดี	เนื้อย่าง	0.858395
2347	เม็ดมะม่วงหิมพานต์	เนื้อย่าง	0.862483
1807	เนื้อย่าง	เม็ดมะม่วงหิมพานต์	0.862483
3235	กระเป๋า chanel	Brownie อวกาศ	0.866025
175	Brownie อวกาศ	กระเป๋า chanel	0.866025
3554	น้าส้มดีโด้	กะหรี่(บึ๊บ)	0.869918
1034	กะหรี่(บั๊บ)	น้าส้มดีโด้ 0.86	
1110	แฟน ที่แปลว่าพัดลม	หมอนข้าง	0.879192
750	หมอนข้าง	แฟน ที่แปลว่าพัดลม	0.879192
1502	หมูแดดเดียว	เม็ดมะม่วงหิมพานต์	0.880406
2342	เม็ดมะม่วงหิมพานต์	หมูแดดเดียว	0.880406
918	แสงโสม	แม่โขง	0.880705

04 Interpretation

สำหรับ Cosine Similarity item-item และ filter item similarity ช่วง > 0.9 และ > 0.95 พบว่า มีหลายกรณีที่ของทั้งสอง อย่างไม่น่าจะมีความส้มพันธ์ หรือมีทิศทางเดียวกัน แต่กลับมีค่า cosine similarity ที่ค่อนข้างสูง เช่น

- คนชอบกินโรตี กระหรี่ปั๊ปมีแนวโน้มชอบทาครีมกันแดด และชอบกระเป่า โดยมีค่า cosine สู่งถึง 97-98%
- คนที่ชอบพัดลม ก็มีแนวโน้มที่จะชอบ ชาบู ครีมกันแดด กระเป๋า และกระหรี่ปั๊ป โดยมีค่า Cosine 90%

```
#Filter cosine to address weak link
links_filter = links.loc[ (links['cosine_value'] & 0.90) & (links['item1'] != links['item2']) 1
links_filter.sort_values(by='cosine_value',ascending=True).head(15)
#Filter cosine to address weak link
links_filter = links.loc[ (links['cosine_value'] & 0.95) & (links['item1'] != links['item2']) ]
links_filter.sort_values(by='cosine_value',ascending=True).head(15)
```

	item1	item2	cosine_value
994	กะหรี่(บึ๊บ)	แฟน ที่แปลว่าพัดลม	0.900919
1114	แฟน ที่แปลว่าพัดลม	กะหรี่(ปั๊บ)	0.900919
748	หมอนข้าง	กะหรี่(ปั๊บ)	0.902214
988	กะหรี่(บั๊บ)	หมอนข้าง	0.902214
1510	หมูแดดเดียว	โรตี	0.903211
3645	Software ถูกลิขสิทธิ์	โรตี	0.903211
2865	โรตี	Software ถูกลิขสิทธิ์	0.903211
2830	โรตี	หมูแดดเดียว	0.903211
1158	แฟน ที่แปลว่าพัดลม	ครีมกันแดด	0.905822
1130	แฟน ที่แปลว่าพัดลม	ชาบู	0.905822
1129	แฟน ที่แปลว่าพัดลม	กระเป๋า	0.905822
1909	กระเป๋า	แฟน ที่แปลว่าพัดลม	0.905822
3678	ครีมกันแดด	แฟน ที่แปลว่าพัดลม	0.905822
1970	ชาบู	แฟน ที่แปลว่าพัดลม	0.905822
778	หมอนข้าง	โรตี	0.918085

	item1	item2	cosine_value
1022	กะหรี่(บั๊บ)	โรตี	0.960085
2822	โรตี	กะหรี่(ปั๊บ)	0.960085
1007	กะหรี่(ปั๊บ)	กระเป๋า	0.974022
1008	กะหรี่(ปั๊บ)	ชาบู	0.974022
1036	กะหรี่(บั๊บ)	ครีมกันแดด	0.974022
3676	ครีมกันแดด	กะหรี่(ปั๊บ)	0.974022
1907	กระเป๋า	กะหรี่(ปั๊บ)	0.974022
1968	ชาบู	กะหรี่(ปั๊บ)	0.974022
2866	โรตี	ครีมกันแดด	0.987096
2838	โรตี	ชาบู	0.987096
2837	โรตี	กระเป๋า	0.987096
3706	ครีมกันแดด	โรตี	0.987096
1937	กระเป๋า	โรตี	0.987096
1998	ชาบู	โรดี	0.987096

