William Stults - Association Rules and Lift Analysis (D212 Task 3)

February 12, 2023

---

1 Part I: Research Question

1.1 Research Question

My data set for this data mining exercise includes data on a telecommunications company's sales history, dating back to 2 years prior, with a focus on technology related merchandise. Data analysis performed on the dataset will be aimed with this research question in mind: what are the top 3 association rules we can determine based on the raw sales data? The telecommunications company's data is only loosely organized and will require some cleaning and restructuring.

1.2 Objectives and Goals

The goal of my data analysis will be to determine 3 rules best suited to illustrate the relationships between items frequently purchased together, and offer advice on how those rules and other insights might be actionable by the telecommunications company.

---

2 Part II: Market Basket Justification

2.1 Market Basket Analysis

Market Basket Analysis is a technique used by retailers to determine associations between items. The algorithm discovers associations between different items and products that may be purchased together. This helps retailers to make business and marketing decision, such as the right product placement or promotions likely to succeed. The algorithm presents this information as association rules, which can be thought of as "if, then" type rules. The two components of these rules are the antecedent (the "if" component) and the consequent (the "then" component) (Deb, 2019).

The quality of the association rules mined by the algorithm is determined by three metrics:

- Support - the fraction of transactions which contain item "A" and "B". Support reveals the frequently bought items or combinations of items.
- Confidence - how often the items "A" and "B" are purchased together, based on the number times "A" is purchased.
- Lift - the strength of a rule over random instances of "A" and "B". Lift is commonly used as the authoritative indicator of how strong a rule is.

The Apriori algorithm, which I'll be using for this market basket analysis, begins by identifying frequently purchased individual items in a data set of transactions. Each item is assigned a "support" measure, which again is determined by how frequently the item is purchased. It then proceeds to take items that meet a minimum support threshold and looks for frequent item combinations, grouping them into item sets. This process continues until the algorithm can no longer find larger item sets that meet the minimum support threshold. Association rules can then be created using minimum threshold values for the other metrics, "confidence" and "lift" (Deb, 2019).

The expected outcome of this exercise will be a group of association rules that exhibit the strongest values for support, confidence and lift. The insights provided by these rules can be used to drive business related decisions.

2.2 Transactions

Transactions within the transformed data set will be more clearly shown further on in this document, but one example of a transaction that would appear in the data is shown here, where a customer purchased three items: "Apple Lightning to Digital AV Adapter", "Apple Pencil", and "TP-Link AC1750 Smart WiFi Router", denoted by the "TRUE" values in the columns for those items.

---

1 FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE TRUE FALSE FALSE TRUE FALSE TRUE FALSE

2.3 Assumption

Market basket analysis assumes that if an item set is frequent, then the subsets of that item set (whether that consists of an individual item or multiple items) must also be frequent (Ranjan, 2020).

---

3 Part III: Data Preparation and Analysis

My first steps will be to import the Python libraries needed for my data analysis and then import the complete data set and execute functions that will give me information on its size and the data types of its variables.

[1]:
```python
# Imports and housekeeping
import pandas as pd
```

```python
import numpy as np
from mlxtend.preprocessing import TransactionEncoder
from mlxtend.frequent_patterns import apriori
from mlxtend.frequent_patterns import association_rules
```

[2]:
```python
# Import the main dataset
df = pd.read_csv('teleco_market_basket.csv')
```

[3]:
```python
# Column names, non-null counts and dtypes
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 15002 entries, 0 to 15001
Data columns (total 20 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Item01  7501 non-null   object
 1   Item02  5747 non-null   object
 2   Item03  4389 non-null   object
 3   Item04  3345 non-null   object
 4   Item05  2529 non-null   object
 5   Item06  1864 non-null   object
 6   Item07  1369 non-null   object
 7   Item08  981 non-null    object
 8   Item09  654 non-null    object
 9   Item10  395 non-null    object
 10  Item11  256 non-null    object
 11  Item12  154 non-null    object
 12  Item13  87 non-null     object
 13  Item14  47 non-null     object
 14  Item15  25 non-null     object
 15  Item16  8 non-null      object
 16  Item17  4 non-null      object
 17  Item18  4 non-null      object
 18  Item19  3 non-null      object
 19  Item20  1 non-null      object
dtypes: object(20)
memory usage: 2.3+ MB
```

[4]:
```python
# Preview top 5 rows
df.head()
```

[4]:
```
                                Item01                           Item02  \
0                                  NaN                              NaN
1           Logitech M510 Wireless mouse                       HP 63 Ink
2                                  NaN                              NaN
3  Apple Lightning to Digital AV Adapter  TP-Link AC1750 Smart WiFi Router
4                                  NaN                              NaN
```

```
       Item03                      Item04                   Item05  \
0         NaN                         NaN                      NaN
1    HP 65 ink  nonda USB C to USB Adapter  10ft iPHone Charger Cable
2         NaN                         NaN                      NaN
3 Apple Pencil                        NaN                      NaN
4         NaN                         NaN                      NaN

       Item06                        Item07  \
0         NaN                           NaN
1 HP 902XL ink  Creative Pebble 2.0 Speakers
2         NaN                           NaN
3         NaN                           NaN
4         NaN                           NaN

                              Item08                        Item09  \
0                                NaN                           NaN
1 Cleaning Gel Universal Dust Cleaner  Micro Center 32GB Memory card
2                                NaN                           NaN
3                                NaN                           NaN
4                                NaN                           NaN

                                  Item10                       Item11  \
0                                    NaN                          NaN
1 YUNSONG 3pack 6ft Nylon Lightning Cable  TopMate C5 Laptop Cooler pad
2                                    NaN                          NaN
3                                    NaN                          NaN
4                                    NaN                          NaN

                   Item12                        Item13  \
0                     NaN                           NaN
1 Apple USB-C Charger cable  HyperX Cloud Stinger Headset
2                     NaN                           NaN
3                     NaN                           NaN
4                     NaN                           NaN

                     Item14                          Item15  \
0                       NaN                             NaN
1 TONOR USB Gaming Microphone  Dust-Off Compressed Gas 2 pack
2                       NaN                             NaN
3                       NaN                             NaN
4                       NaN                             NaN

                        Item16                 Item17  \
0                          NaN                    NaN
1 3A USB Type C Cable 3 pack 6FT  HOVAMP iPhone charger
2                          NaN                    NaN
```

```
3                             NaN                     NaN
4                             NaN                     NaN

                Item18                               Item19  \
0                  NaN                                  NaN
1  SanDisk Ultra 128GB card  FEEL2NICE 5 pack 10ft Lighning cable
2                  NaN                                  NaN
3                  NaN                                  NaN
4                  NaN                                  NaN

                               Item20
0                                 NaN
1  FEIYOLD Blue light Blocking Glasses
2                                 NaN
3                                 NaN
4                                 NaN
```

[5]:
```
# Show column data types
print(df.dtypes)
```

```
Item01    object
Item02    object
Item03    object
Item04    object
Item05    object
Item06    object
Item07    object
Item08    object
Item09    object
Item10    object
Item11    object
Item12    object
Item13    object
Item14    object
Item15    object
Item16    object
Item17    object
Item18    object
Item19    object
Item20    object
dtype: object
```

[6]:
```
# Dimensions of data set
df.shape
```

[6]: (15002, 20)

Once this is done, I determine whether null data points exist in the data set, and if so, I remove them.

[7]:
```
# Check the data frame for null values
print(df.isnull().sum())
```

```
Item01     7501
Item02     9255
Item03    10613
Item04    11657
Item05    12473
Item06    13138
Item07    13633
Item08    14021
Item09    14348
Item10    14607
Item11    14746
Item12    14848
Item13    14915
Item14    14955
Item15    14977
Item16    14994
Item17    14998
Item18    14998
Item19    14999
Item20    15001
dtype: int64
```

[8]:
```
# Drop null values from the data frame
df = df.dropna(how = 'all')
```

Reviewing the changes made to the data set by removing the null data points, I see that the data set size has been reduced from 15001 rows to 7501.

[9]:
```
# Dimensions of data set with no nulls
df.shape
```

[9]:
```
(7501, 20)
```

With my null values removed, I can proceed with transactionalizing the data set. This is done by creating an array of data points from the data set, then fitting and transforming the array of data points using mlxtend's TransactionEncoder function. I will create a new data frame from the transactionalized data, named "prep_df".

[10]:
```
# Initialize "trans" array and populate with data points
trans = []
```

```
for i in range (0, 7501):
    trans.append([str(df.values[i, j]) for j in range (0,20)])
```

[11]:
```
# Transactionalize the data in the "trans" array
te = TransactionEncoder()
array = te.fit(trans).transform(trans)
```

[12]:
```
# Create a data frame from the transactionalized data
prep_df = pd.DataFrame(array, columns = te.columns_)
prep_df
```

[12]:
```
      10ft iPHone Charger Cable  10ft iPHone Charger Cable 2 Pack  \
0                          True                             False
1                         False                             False
2                         False                             False
3                         False                             False
4                         False                             False
…                           …                               …
7496                      False                             False
7497                      False                             False
7498                      False                             False
7499                      False                             False
7500                      False                             False

      3 pack Nylon Braided Lightning Cable  3A USB Type C Cable 3 pack 6FT  \
0                                    False                            True
1                                    False                           False
2                                    False                           False
3                                    False                           False
4                                    False                           False
…                                    …                              …
7496                                 False                           False
7497                                 False                           False
7498                                 False                           False
7499                                 False                           False
7500                                 False                           False

      5pack Nylon Braided USB C cables  ARRIS SURFboard SB8200 Cable Modem  \
0                                False                               False
1                                False                               False
2                                False                               False
3                                False                               False
4                                False                               False
…                                …                                  …
7496                             False                               False
7497                             False                                True
7498                             False                               False
```

```
7499                            False                               False
7500                            False                               False

      Anker 2-in-1 USB Card Reader  Anker 4-port USB hub  \
0                            False                 False
1                            False                 False
2                            False                 False
3                            False                 False
4                            False                 False
…                              …                   …
7496                         False                 False
7497                         False                 False
7498                         False                 False
7499                         False                 False
7500                         False                 False

      Anker USB C to HDMI Adapter  Apple Lightning to Digital AV Adapter  …  \
0                           False                                  False  …
1                           False                                   True  …
2                           False                                  False  …
3                           False                                  False  …
4                           False                                  False  …
…                             …                                  …  …
7496                        False                                  False  …
7497                        False                                   True  …
7498                        False                                  False  …
7499                        False                                  False  …
7500                        False                                  False  …

      iFixit Pro Tech Toolkit  iPhone 11 case  iPhone 12 Charger cable  \
0                       False           False                    False
1                       False           False                    False
2                       False           False                    False
3                       False           False                    False
4                       False           False                    False
…                        …              …                      …
7496                    False           False                    False
7497                    False           False                    False
7498                    False           False                    False
7499                    False           False                    False
7500                    False           False                    False

      iPhone 12 Pro case  iPhone 12 case  iPhone Charger Cable Anker 6ft  \
0                  False           False                          False
1                  False           False                          False
2                  False           False                          False
3                  False           False                          False
```

```
4                    False           False                               False
…                    …            …                              …
7496                 False           False                               False
7497                 False           False                               False
7498                 False           False                               False
7499                 False           False                               False
7500                 False           False                               False

      iPhone SE case    nan  nonda USB C to USB Adapter  seenda Wireless mouse
0              False  False                        True                  False
1              False   True                       False                  False
2              False   True                       False                  False
3              False   True                       False                  False
4              False   True                       False                  False
…              …    …                        …                    …
7496           False   True                       False                  False
7497           False   True                       False                  False
7498           False   True                       False                  False
7499           False   True                       False                  False
7500           False   True                       False                  False

[7501 rows x 120 columns]
```

With my data transformed, I will check the columns of the new dataframe to see if any null (nan) columns are present.

[13]:
```python
# List columns in data frame
for col in prep_df.columns:
    print(col)
```

```
10ft iPHone Charger Cable
10ft iPHone Charger Cable 2 Pack
3 pack Nylon Braided Lightning Cable
3A USB Type C Cable 3 pack 6FT
5pack Nylon Braided USB C cables
ARRIS SURFboard SB8200 Cable Modem
Anker 2-in-1 USB Card Reader
Anker 4-port USB hub
Anker USB C to HDMI Adapter
Apple Lightning to Digital AV Adapter
Apple Lightning to USB cable
Apple Magic Mouse 2
Apple Pencil
Apple Pencil 2nd Gen
Apple Power Adapter Extension Cable
Apple USB-C Charger cable
AutoFocus 1080p Webcam
```

BENGOO G90000 headset
Blue Light Blocking Glasses
Blue Light Blocking Glasses 2pack
Brother Genuine High Yield Toner Cartridge
Cat 6 Ethernet Cable 50ft
Cat8 Ethernet Cable
CicTsing MM057 2.4G Wireless Mouse
Cleaning Gel Universal Dust Cleaner
Creative Pebble 2.0 Speakers
DisplayPort ot HDMI adapter
Dust-Off Compressed Gas
Dust-Off Compressed Gas 2 pack
FEEL2NICE 5 pack 10ft Lighning cable
FEIYOLD Blue light Blocking Glasses
Falcon Dust Off Compressed Gas
HOVAMP Mfi 6pack Lightning Cable
HOVAMP iPhone charger
HP 61 2 pack ink
HP 61 Tri-color ink
HP 61 ink
HP 62XL Tri-Color ink
HP 62XL ink
HP 63 Ink
HP 63 Tri-color ink
HP 63XL Ink
HP 63XL Tri-color ink
HP 64 Tri-Color ink
HP 64 ink
HP 65 ink
HP 902XL ink
HP 952 ink
HP ENVY 5055 printer
HP952XL ink
HooToo USB C Hub
HyperX Cloud Stinger Headset
Jelly Comb 2.4G Slim Wireless mouse
Leader Desk Pad Protector
Logitech M510 Wireless mouse
Logitech MK270 Wireless Keyboard/Mouse
Logitech MK345 Wireless combo
Logitech USB H390 headset
M.2 Screw kit
Mfi-Certified Lightning to USB A Cable
Micro Center 32GB Memory card
Microsot Surface Dock 2
Moread HDMI to VGA Adapter
Mpow HC6 USB Headset
NETGEAR CM500 Cable Modem

NETGEAR Nighthawk WiFi Router
NETGEAR Orbi Home Mesh WiFi System
Nylon Braided Lightning to USB cable
PS4 Headset
Premium Nylon USB Cable
RUNMUS Gaming Headset
SAMSUNG 128GB card
SAMSUNG 256 GB card
SAMSUNG EVO 32GB card
SAMSUNG EVO 64GB card
Sabrent 4-port USB 3.0 hub
SanDisk 128GB Ultra microSDXC card
SanDisk 128GB card
SanDisk 128GB microSDXC card
SanDisk 32GB Ultra SDHC card
SanDisk 32GB card
SanDisk Extreme 128GB card
SanDisk Extreme 256GB card
SanDisk Extreme 32GB 2pack card
SanDisk Extreme Pro 128GB card
SanDisk Extreme Pro 64GB card
SanDisk Ultra 128GB card
SanDisk Ultra 256GB card
SanDisk Ultra 400GB card
SanDisk Ultra 64GB card
Screen Mom Screen Cleaner kit
Stylus Pen for iPad
Syntech USB C to USB Adapter
TONOR USB Gaming Microphone
TP-Link AC1750 Smart WiFi Router
TP-Link AC4000 WiFi router
TopMate C5 Laptop Cooler pad
UNEN Mfi Certified 5-pack Lightning Cable
USB 2.0 Printer cable
USB C to USB Male Adapter
USB Type C Cable
USB Type C to USB-A Charger cable
VIVO Dual LCD Monitor Desk mount
VicTsing Mouse Pad
VicTsing Wireless mouse
Vsco 70 pack stickers
Webcam with Microphone
XPOWER A-2 Air Pump blower
YUNSONG 3pack 6ft Nylon Lightning Cable
hP 65 Tri-color ink
iFixit Pro Tech Toolkit
iPhone 11 case
iPhone 12 Charger cable

```
iPhone 12 Pro case
iPhone 12 case
iPhone Charger Cable Anker 6ft
iPhone SE case
nan
nonda USB C to USB Adapter
seenda Wireless mouse
```

There is one null column (third from the bottom), so I will remove it and check my columns once more to ensure no nulls exist.

[14]:
```python
# Remove null columns from data frame
prep_df = prep_df.drop(['nan'], axis = 1)
```

[15]:
```python
# List columns in data frame
for col in prep_df.columns:
    print(col)
```

```
10ft iPHone Charger Cable
10ft iPHone Charger Cable 2 Pack
3 pack Nylon Braided Lightning Cable
3A USB Type C Cable 3 pack 6FT
5pack Nylon Braided USB C cables
ARRIS SURFboard SB8200 Cable Modem
Anker 2-in-1 USB Card Reader
Anker 4-port USB hub
Anker USB C to HDMI Adapter
Apple Lightning to Digital AV Adapter
Apple Lightning to USB cable
Apple Magic Mouse 2
Apple Pencil
Apple Pencil 2nd Gen
Apple Power Adapter Extension Cable
Apple USB-C Charger cable
AutoFocus 1080p Webcam
BENGOO G90000 headset
Blue Light Blocking Glasses
Blue Light Blocking Glasses 2pack
Brother Genuine High Yield Toner Cartridge
Cat 6 Ethernet Cable 50ft
Cat8 Ethernet Cable
CicTsing MM057 2.4G Wireless Mouse
Cleaning Gel Universal Dust Cleaner
Creative Pebble 2.0 Speakers
DisplayPort ot HDMI adapter
Dust-Off Compressed Gas
Dust-Off Compressed Gas 2 pack
```

FEEL2NICE 5 pack 10ft Lighning cable
FEIYOLD Blue light Blocking Glasses
Falcon Dust Off Compressed Gas
HOVAMP Mfi 6pack Lightning Cable
HOVAMP iPhone charger
HP 61 2 pack ink
HP 61 Tri-color ink
HP 61 ink
HP 62XL Tri-Color ink
HP 62XL ink
HP 63 Ink
HP 63 Tri-color ink
HP 63XL Ink
HP 63XL Tri-color ink
HP 64 Tri-Color ink
HP 64 ink
HP 65 ink
HP 902XL ink
HP 952 ink
HP ENVY 5055 printer
HP952XL ink
HooToo USB C Hub
HyperX Cloud Stinger Headset
Jelly Comb 2.4G Slim Wireless mouse
Leader Desk Pad Protector
Logitech M510 Wireless mouse
Logitech MK270 Wireless Keyboard/Mouse
Logitech MK345 Wireless combo
Logitech USB H390 headset
M.2 Screw kit
Mfi-Certified Lightning to USB A Cable
Micro Center 32GB Memory card
Microsot Surface Dock 2
Moread HDMI to VGA Adapter
Mpow HC6 USB Headset
NETGEAR CM500 Cable Modem
NETGEAR Nighthawk WiFi Router
NETGEAR Orbi Home Mesh WiFi System
Nylon Braided Lightning to USB cable
PS4 Headset
Premium Nylon USB Cable
RUNMUS Gaming Headset
SAMSUNG 128GB card
SAMSUNG 256 GB card
SAMSUNG EVO 32GB card
SAMSUNG EVO 64GB card
Sabrent 4-port USB 3.0 hub
SanDisk 128GB Ultra microSDXC card

SanDisk 128GB card
SanDisk 128GB microSDXC card
SanDisk 32GB Ultra SDHC card
SanDisk 32GB card
SanDisk Extreme 128GB card
SanDisk Extreme 256GB card
SanDisk Extreme 32GB 2pack card
SanDisk Extreme Pro 128GB card
SanDisk Extreme Pro 64GB card
SanDisk Ultra 128GB card
SanDisk Ultra 256GB card
SanDisk Ultra 400GB card
SanDisk Ultra 64GB card
Screen Mom Screen Cleaner kit
Stylus Pen for iPad
Syntech USB C to USB Adapter
TONOR USB Gaming Microphone
TP-Link AC1750 Smart WiFi Router
TP-Link AC4000 WiFi router
TopMate C5 Laptop Cooler pad
UNEN Mfi Certified 5-pack Lightning Cable
USB 2.0 Printer cable
USB C to USB Male Adapter
USB Type C Cable
USB Type C to USB-A Charger cable
VIVO Dual LCD Monitor Desk mount
VicTsing Mouse Pad
VicTsing Wireless mouse
Vsco 70 pack stickers
Webcam with Microphone
XPOWER A-2 Air Pump blower
YUNSONG 3pack 6ft Nylon Lightning Cable
hP 65 Tri-color ink
iFixit Pro Tech Toolkit
iPhone 11 case
iPhone 12 Charger cable
iPhone 12 Pro case
iPhone 12 case
iPhone Charger Cable Anker 6ft
iPhone SE case
nonda USB C to USB Adapter
seenda Wireless mouse

3.1 Copy of Prepared Data Set

With my data set cleaned and prepared I will export the data frame. Below is the code used to export the prepared data set to CSV format.

[16]:
```
# Export prepared dataframe to csv
prep_df.to_csv(r'C:\Users\wstul\d212\transactions_cleaned.csv')
```

I can now begin data mining using the Apriori algorithm. The first step will be to determine which items within the transactionalized data meet a minimum "support" threshold, in this case 0.05, meaning the items are included in no fewer than 5% of purchases.

[17]:
```
# Narrow the data set using a support value of 0.05 as the cutoff
fi = apriori(prep_df, min_support = 0.05, use_colnames = True)
fi
```

[17]:

	support	itemsets
0	0.050527	(10ft iPHone Charger Cable 2 Pack)
1	0.068391	(Anker USB C to HDMI Adapter)
2	0.087188	(Apple Lightning to Digital AV Adapter)
3	0.179709	(Apple Pencil)
4	0.132116	(Apple USB-C Charger cable)
5	0.062525	(Cat8 Ethernet Cable)
6	0.238368	(Dust-Off Compressed Gas 2 pack)
7	0.065858	(FEIYOLD Blue light Blocking Glasses)
8	0.059992	(Falcon Dust Off Compressed Gas)
9	0.163845	(HP 61 ink)
10	0.058526	(HP 62XL Tri-Color ink)
11	0.079323	(HP 63XL Ink)
12	0.071457	(Logitech M510 Wireless mouse)
13	0.095321	(Nylon Braided Lightning to USB cable)
14	0.051060	(Premium Nylon USB Cable)
15	0.052393	(SAMSUNG EVO 32GB card)
16	0.063325	(SanDisk Ultra 128GB card)
17	0.098254	(SanDisk Ultra 64GB card)
18	0.129583	(Screen Mom Screen Cleaner kit)
19	0.095054	(Stylus Pen for iPad)
20	0.081056	(Syntech USB C to USB Adapter)
21	0.076523	(TopMate C5 Laptop Cooler pad)
22	0.170911	(USB 2.0 Printer cable)
23	0.080389	(USB Type C to USB-A Charger cable)
24	0.174110	(VIVO Dual LCD Monitor Desk mount)
25	0.050927	(Apple Pencil, Dust-Off Compressed Gas 2 pack)
26	0.052660	(Dust-Off Compressed Gas 2 pack, HP 61 ink)
27	0.059725	(Dust-Off Compressed Gas 2 pack, VIVO Dual LCD…

With this new set of "frequent items", I can use Apriori to data mine association rules. I want a view of the strongest rules only, so I set a minimum "lift" threshold of 1.

[18]:
```
# Mine association rules using a lift value of 1 as the cutoff
rules = association_rules(fi, metric = 'lift', min_threshold = 1)
rules
```

[18]:
```
                      antecedents                        consequents  \
0                  (Apple Pencil)   (Dust-Off Compressed Gas 2 pack)
1  (Dust-Off Compressed Gas 2 pack)                   (Apple Pencil)
2  (Dust-Off Compressed Gas 2 pack)                     (HP 61 ink)
3                     (HP 61 ink)   (Dust-Off Compressed Gas 2 pack)
4  (Dust-Off Compressed Gas 2 pack)  (VIVO Dual LCD Monitor Desk mount)
5 (VIVO Dual LCD Monitor Desk mount)  (Dust-Off Compressed Gas 2 pack)

   antecedent support  consequent support   support  confidence      lift  \
0            0.179709            0.238368  0.050927    0.283383  1.188845
1            0.238368            0.179709  0.050927    0.213647  1.188845
2            0.238368            0.163845  0.052660    0.220917  1.348332
3            0.163845            0.238368  0.052660    0.321400  1.348332
4            0.238368            0.174110  0.059725    0.250559  1.439085
5            0.174110            0.238368  0.059725    0.343032  1.439085

   leverage  conviction
0  0.008090    1.062815
1  0.008090    1.043158
2  0.013604    1.073256
3  0.013604    1.122357
4  0.018223    1.102008
5  0.018223    1.159314
```

With a small set of association rules to work with, I can determine the strongest rules by eliminating those rules that have a "lift" value less than 1.15 and a "confidence" value less than 0.26, then list the top 3 rules resulting from the data mining.

[19]:
```
# List the top 3 rules using a lift threshold of 1.15 and a confidence␣
↪threshold of 0.26
rules[(rules['lift'] >= 1.15) &
      (rules['confidence'] >= 0.26)].nlargest(n = 3, columns = 'lift')
```

[19]:
```
                      antecedents                       consequents  \
5 (VIVO Dual LCD Monitor Desk mount)  (Dust-Off Compressed Gas 2 pack)
3                     (HP 61 ink)  (Dust-Off Compressed Gas 2 pack)
0                  (Apple Pencil)  (Dust-Off Compressed Gas 2 pack)

   antecedent support  consequent support   support  confidence      lift  \
5            0.174110            0.238368  0.059725    0.343032  1.439085
```

```
3           0.163845            0.238368  0.050927    0.321400  1.348332
0           0.179709            0.238368  0.050927    0.283383  1.188845

   leverage  conviction
5  0.018223    1.159314
3  0.013604    1.122357
0  0.008090    1.062815
```

The rules can be summarized as follows (values rounded to 2 decimals):

1. IF **"VIVO Dual LCD Monitor Desk mount"** is purchased THEN **"Dust-Off Compressed Gas 2 pack"** is also purchased

   lift = 1.44, confidence = 0.34, support = 0.06

2. IF **"HP 61 ink"** is purchased THEN **"Dust-Off Compressed Gas 2 pack"** is also purchased

   lift = 1.35, confidence = 0.32, support = 0.05

3. IF **"Apple Pencil"** is purchased THEN **"Dust-Off Compressed Gas 2 pack"** is also purchased

   lift = 1.19, confidence = 0.28, support = 0.05

---

4 Part IV: Data Summary and Implications

4.1 Support, Lift, and Confidence

To recap from my earlier explanation of Apriori association rule mining:

- Support - the fraction of transactions which contain item "A" and "B". Support reveals the frequently bought items or combinations of items.
- Confidence - how often the items "A" and "B" are purchased together, based on the number times "A" is purchased.
- Lift - the strength of a rule over random instances of "A" and "B". Lift is commonly used as the authoritative indicator of how strong a rule is.

Based on my results, in order of "lift" measure descending:

1. Customers purchase a VIVO Dual LCD Monitor Desk mount together with a Dust-Off Compressed Gas 2 pack 6% of the time, and if a customer purchases a VIVO Dual LCD Monitor Desk mount, there is a 34% likelihood they will also purchase a Dust-Off Compressed Gas 2 pack.
2. Customers purchase an HP 61 ink together with a Dust-Off Compressed Gas 2 pack 5% of the time, and if a customer purchases an HP 61 ink, there is a 32% likelihood they will also purchase a Dust-Off Compressed Gas 2 pack.

3. Customers purchase an Apple Pencil together with a Dust-Off Compressed Gas 2 pack 5% of the time, and if a customer purchases an Apple Pencil, there is a 28% likelihood they will also purchase a Dust-Off Compressed Gas 2 pack.

“Dust-Off Compressed Gas 2 pack” appears to be not only a frequently purchased item (28% of purchases, regardless of associations), but is frequently purchased with other products. Due to its position as the “consequent” in each of these rules, we might conclude that customers purchase these as an “add-on” while they are already shopping for other items.

Based upon these insights it might prove advantageous to position the “Dust-Off Compressed Gas 2 pack” in multiple locations, such as on endcaps throughout the store. Additionally the retailer could position items they would like to sell more of near the “Dust-Off Compressed Gas 2 pack” in its department of the store.

---

5 Part V: Demonstration

Panopto Video Recording

A link for the Panopto video has been provided separately. The demonstration includes the following:

- Demonstration of the functionality of the code used for the analysis
- Identification of the version of the programming environment

---

6 Web Sources

https://www.section.io/engineering-education/apriori-algorithm-in-python/

https://medium.com/edureka/apriori-algorithm-d7cc648d4f1e

---

7 References

Deb, S. (2019, June 20). *Apriori Algorithm — Know How to Find Frequent Itemsets.* Medium. https://medium.com/edureka/apriori-algorithm-d7cc648d4f1e

Ranjan, A. (2020, December 3). *Apriori Algorithm in Association Rule Learning.* Medium. https://medium.com/analytics-vidhya/apriori-algorithm-in-association-rule-learning-9287fe17e944