**QUESTIONS**

* Are there any interesting patterns or item relationships within Electronidex's transactions?
* Would Blackwell benefit from selling any of Electronidex's items?
* In your opinion, should Blackwell acquire Electronidex?
* If Blackwell does acquire Electronidex, do you have any recommendations for Blackwell? (Ex: cross-selling items, sale promotions, should they remove items, etc.)

**CODING**

**UPLOAD TRANSACTIONS**

> Trans <- read.transactions("ElectronidexTransactions2017.csv", format = "basket",sep = ",", rm.duplicates = TRUE)

> Trans

transactions in sparse format with

9835 transactions (rows) and

125 items (columns)

> summary(Trans)

transactions as itemMatrix in sparse format with

9835 rows (elements/itemsets/transactions) and

125 columns (items) and a density of 0.03506172

most frequent items:

iMac HP Laptop CYBERPOWER Gamer Desktop

2519 1909 1809

Apple Earpods Apple MacBook Air (Other)

1715 1530 33622

element (itemset/transaction) length distribution:

sizes

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

2 2163 1647 1294 1021 856 646 540 439 353 247 171 119 77 72 56

16 17 18 19 20 21 22 23 25 26 27 29 30

41 26 20 10 10 10 5 3 1 1 3 1 1

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.000 2.000 3.000 4.383 6.000 30.000

includes extended item information - examples:

labels

1 1TB Portable External Hard Drive

2 2TB Portable External Hard Drive

3 3-Button Mouse

**PLOTTING, VISUALIZE TRANSACTIONS**

Consider thinking about these questions when plotting: Is there a way to only plot a certain number of transactions? Is there a way to plot using certain metrics? Which plots might provide the most insight?

Solution, use image() Function:

> image(sample(Trans, 500))

**APRIORI**

There are three parameters controlling the number of rules to be generated *viz.* **SUPPORT** and **CONFIDENCE**. Another parameter **LIFT** is generated using Support and Confidence and is one of the major parameters to filter the generated rules.

* ***Support***is an indication of how frequently the itemset appears in the dataset. Consider only the two transactions from the above output. The support of the item *citrus fruit* is 1/2 as it appears in only 1 out of the two transactions.
* ***Confidence***is an indication of how often the rule has been found to be true. We will discuss more about confidence after generating the rules.

> Rules<- apriori (Trans, parameter = list(supp = 0.1, conf = 0.8))

Apriori

Parameter specification:

confidence minval smax arem aval originalSupport maxtime support minlen maxlen

0.8 0.1 1 none FALSE TRUE 5 0.1 1 10

target ext

rules FALSE

Algorithmic control:

filter tree heap memopt load sort verbose

0.1 TRUE TRUE FALSE TRUE 2 TRUE

Absolute minimum support count: 983

set item appearances ...[0 item(s)] done [0.00s].

set transactions ...[125 item(s), 9835 transaction(s)] done [0.00s].

sorting and recoding items ... [10 item(s)] done [0.00s].

creating transaction tree ... done [0.00s].

checking subsets of size 1 2 done [0.00s].

writing ... [**0 rule(s)**] done [0.00s].

creating S4 object ... done [0.00s].

> summary(Rules)

set of 0 rules

* Receiving 0 rules means that you will need to experiment with the **Support** and **Confidence** values.

> Rules<- apriori (Trans, parameter = list(supp = 0.01, conf = 0.1, minlen = 1))

Apriori

Parameter specification:

confidence minval smax arem aval originalSupport maxtime support minlen maxlen

0.1 0.1 1 none FALSE TRUE 5 0.01 1 10

target ext

rules FALSE

Algorithmic control:

filter tree heap memopt load sort verbose

0.1 TRUE TRUE FALSE TRUE 2 TRUE

Absolute minimum support count: 98

set item appearances ...[0 item(s)] done [0.00s].

set transactions ...[125 item(s), 9835 transaction(s)] done [0.01s].

sorting and recoding items ... [82 item(s)] done [0.00s].

creating transaction tree ... done [0.00s].

checking subsets of size 1 2 3 4 done [0.00s].

writing ... [527 rule(s)] done [0.00s].

creating S4 object ... done [0.00s].

*Minlen Parameter:*

1 = 527 rules => 10

2 = 517 rules => 385

3 = 132 rules

4 = 0 rules

> summary(Rules)

set of 527 rules

rule length distribution (lhs + rhs):sizes

1 2 3

10 385 132

Min. 1st Qu. Median Mean 3rd Qu. Max.

1.000 2.000 2.000 2.231 2.500 3.000

summary of quality measures:

support confidence lift count

Min. :0.01007 Min. :0.1000 Min. :0.6931 Min. : 99.0

1st Qu.:0.01159 1st Qu.:0.1469 1st Qu.:1.3269 1st Qu.: 114.0

Median :0.01505 Median :0.2171 Median :1.6085 Median : 148.0

Mean :0.02066 Mean :0.2457 Mean :1.6703 Mean : 203.2

3rd Qu.:0.02222 3rd Qu.:0.3236 3rd Qu.:1.9478 3rd Qu.: 218.5

Max. :0.25613 Max. :0.6023 Max. :3.3596 Max. :2519.0

mining info:

data ntransactions support confidence

Trans 9835 0.01 0.1

> inspect(head(sort(Rules, by = "confidence"), 10))

**LHS RHS SUPPORT CONFIDENCE LIFT COUNT**

{Acer Aspire,ViewSonic Monitor} => {HP Laptop} 0.01077783 0.6022727 3.102856 106

{ASUS 2 Monitor,Lenovo Desktop Computer => {iMac} 0.01087951 0.5911602 2.308083 107

{Apple Magic Keyboard, Dell Desktop} => {iMac} 0.01016777 0.5847953 2.283232 100

{ASUS Monitor,HP Laptop} => {iMac} 0.01179461 0.5829146 2.275889 116

{ASUS 2 Monitor,HP Laptop} => {iMac} 0.01108287 0.5828877 2.275784 109

{Dell Desktop,ViewSonic Monitor} => {HP Laptop} 0.01525165 0.5747126 2.960869 150

{Dell Desktop, ViewSonic Monitor} => {iMac} 0.01474326 0.5555556 2.169071 145

{Lenovo Desktop Computer,ViewSonic Monitor}=> {iMac} 0.01576004 0.5555556 2.169071 155

{HP Laptop,Microsoft Office Home and Student 2016} =>{iMac}0.01291307 0.5521739 2.155868 127

{Acer Desktop,ViewSonic Monitor} => {iMac} 0.01006609 0.5439560 2.123782 99

Also by Items

> ACER.Rules <- apriori(Trans, parameter = list(supp=0.01, conf=0.1),appearance = list(default="lhs",rhs="ACER"))

To Be Tested

> plot(Rules)

A screenshot of a cell phone

Description automatically generated

> plot(TopRules, method = "graph", engine = "htmlwidget")

A screenshot of a cell phone

Description automatically generated

*Plot according to number of Items Per Rules*

A close up of a logo

Description automatically generated