EE4483
DATA-MINING ASSIGNMENT
REPORT
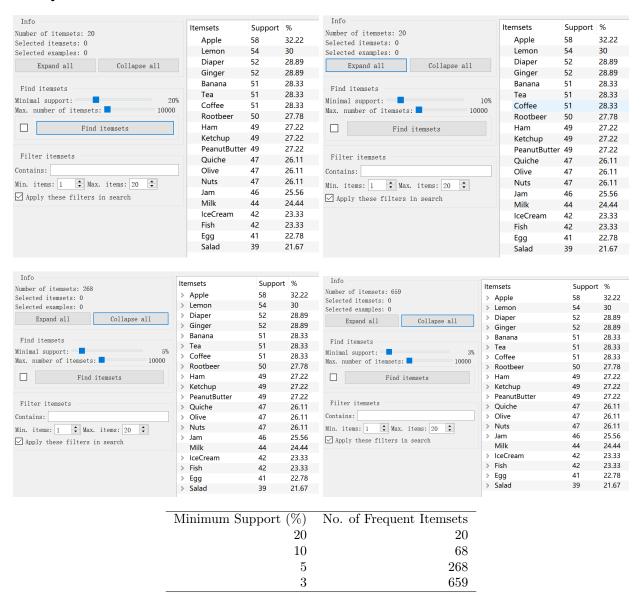
Hu Xiaoxiang
U1521319A
EEE

13 Nov, 2017
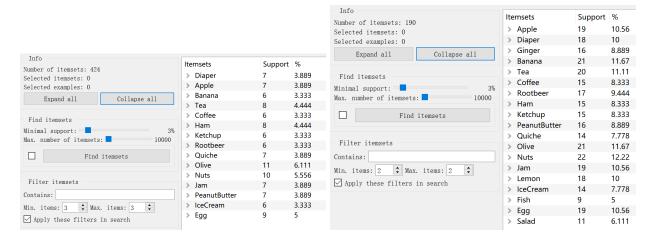
Contents

1 Question 1:

Info

Number of itemsets: 20
Selected itemsets: 0
Selected examples: 0

Expand all   Collapse all

Find itemsets

Minimal support: ▮ 20%
Max. number of itemsets: ▮ 10000

☐ Find itemsets

Filter itemsets

Contains: ____
Min. items: 1   Max. items: 20
☑ Apply these filters in search

Itemsets	Support	%
Apple	58	32.22
Lemon	54	30
Diaper	52	28.89
Ginger	52	28.89
Banana	51	28.33
Tea	51	28.33
Coffee	51	28.33
Rootbeer	50	27.78
Ham	49	27.22
Ketchup	49	27.22
PeanutButter	49	27.22
Quiche	47	26.11
Olive	47	26.11
Nuts	47	26.11
Jam	46	25.56
Milk	44	24.44
IceCream	42	23.33
Fish	42	23.33
Egg	41	22.78
Salad	39	21.67

Info

Number of itemsets: 20
Selected itemsets: 0
Selected examples: 0

Expand all   Collapse all

Find itemsets

Minimal support: ▮ 10%
Max. number of itemsets: ▮ 10000

☐ Find itemsets

Filter itemsets

Contains: ____
Min. items: 1   Max. items: 20
☑ Apply these filters in search

Itemsets	Support	%
Apple	58	32.22
Lemon	54	30
Diaper	52	28.89
Ginger	52	28.89
Banana	51	28.33
Tea	51	28.33
Coffee	51	28.33
Rootbeer	50	27.78
Ham	49	27.22
Ketchup	49	27.22
PeanutButter	49	27.22
Quiche	47	26.11
Olive	47	26.11
Nuts	47	26.11
Jam	46	25.56
Milk	44	24.44
IceCream	42	23.33
Fish	42	23.33
Egg	41	22.78
Salad	39	21.67

Info

Number of itemsets: 268
Selected itemsets: 0
Selected examples: 0

Expand all   Collapse all

Find itemsets

Minimal support: ▮ 5%
Max. number of itemsets: ▮ 10000

☐ Find itemsets

Filter itemsets

Contains: ____
Min. items: 1   Max. items: 20
☑ Apply these filters in search

Itemsets	Support	%
Apple	58	32.22
Lemon	54	30
Diaper	52	28.89
Ginger	52	28.89
Banana	51	28.33
Tea	51	28.33
Coffee	51	28.33
Rootbeer	50	27.78
Ham	49	27.22
Ketchup	49	27.22
PeanutButter	49	27.22
Quiche	47	26.11
Olive	47	26.11
Nuts	47	26.11
Jam	46	25.56
Milk	44	24.44
IceCream	42	23.33
Fish	42	23.33
Egg	41	22.78
Salad	39	21.67

Info

Number of itemsets: 659
Selected itemsets: 0
Selected examples: 0

Expand all   Collapse all

Find itemsets

Minimal support: ▮ 3%
Max. number of itemsets: ▮ 10000

☐ Find itemsets

Filter itemsets

Contains: ____
Min. items: 1   Max. items: 20
☑ Apply these filters in search

Itemsets	Support	%
Apple	58	32.22
Lemon	54	30
Diaper	52	28.89
Ginger	52	28.89
Banana	51	28.33
Tea	51	28.33
Coffee	51	28.33
Rootbeer	50	27.78
Ham	49	27.22
Ketchup	49	27.22
PeanutButter	49	27.22
Quiche	47	26.11
Olive	47	26.11
Nuts	47	26.11
Jam	46	25.56
Milk	44	24.44
IceCream	42	23.33
Fish	42	23.33
Egg	41	22.78
Salad	39	21.67

Minimum Support (%)	No. of Frequent Itemsets
20	20
10	68
5	268
3	659

2 Question 2:

	Itemsets	Support	%
Info	Diaper	7	3.889
Number of itemsets: 424	Apple	7	3.889
Selected itemsets: 0	Banana	6	3.333
Selected examples: 0	Tea	8	4.444
Expand all / Collapse all	Coffee	6	3.333
	Ham	8	4.444
Find itemsets	Ketchup	6	3.333
Minimal support: 3%	Rootbeer	6	3.333
Max. number of itemsets: 10000	Quiche	7	3.889
Find itemsets	Olive	11	6.111
	Nuts	10	5.556
Filter itemsets	Jam	7	3.889
Contains:	PeanutButter	7	3.889
Min. items: 3 Max. items: 3	IceCream	6	3.333
Apply these filters in search	Egg	9	5

	Itemsets	Support	%
Info	Apple	19	10.56
Number of itemsets: 190	Diaper	18	10
Selected itemsets: 0	Ginger	16	8.889
Selected examples: 0	Banana	21	11.67
Expand all / Collapse all	Tea	20	11.11
	Coffee	15	8.333
Find itemsets	Rootbeer	17	9.444
Minimal support: 3%	Ham	15	8.333
Max. number of itemsets: 10000	Ketchup	15	8.333
Find itemsets	PeanutButter	16	8.889
	Quiche	14	7.778
Filter itemsets	Olive	21	11.67
Contains:	Nuts	22	12.22
Min. items: 2 Max. items: 2	Jam	19	10.56
Apply these filters in search	Lemon	18	10
	IceCream	14	7.778
	Fish	9	5
	Egg	19	10.56
	Salad	11	6.111

Minimum Support (%)	No. of Frequent Itemsets	No. of Frequent 3-Itemsets	No. of Frequent 2-Itemsets
3	659	424	190

Percentage of frequent 3-itemset = 424 / 659 = 64.34%
Percentage of frequent 2-itemset = 190 / 659 = 28.83%

3 Question 3:

Info
Number of rules: 117
Filtered rules: 117
Selected rules: 0
Selected examples: 0

Find association rules
Minimal support: 5%
Minimal confidence: 50%
Max. number of rules: 10000
Induce classification (itemset → class) rules
Find rules

Info
Number of rules: 0
Filtered rules: 0
Selected rules: 0
Selected examples: 0

Find association rules
Minimal support: 10%
Minimal confidence: 50%
Max. number of rules: 10000
Induce classification (itemset → class) rules
Find rules

Minimum Support (%)	Minimum Confidence (%)	No. of Association Rules
5	50	117
10	50	0

Explanation: The smaller the minimum support is, the more the number of strong rules generate.

4 Question 4:

	Info						
Number of rules: 78							
Filtered rules: 78							
Selected rules: 0							
Selected examples: 0							

Find association rules

Minimal support: ■ 2%
Minimal confidence: ■ 100%
Max. number of rules: ■ 10000
☐ Induce classification (itemset → class) rules

☐ Find rules

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent
0.033	1.000	0.033	8.500	3.529	0.024	IceCream, Olive, Tea → Banana
0.033	1.000	0.033	9.667	3.103	0.023	Banana, Ham, Salad → Apple
0.028	1.000	0.028	8.400	4.286	0.021	Ham, Diaper, Coffee → IceCream
0.028	1.000	0.028	10.400	3.462	0.020	Tea, Ketchup, Fish → Diaper
0.028	1.000	0.028	8.200	4.390	0.021	Diaper, Coffee, Ginger → Egg
0.028	1.000	0.028	10.200	3.529	0.020	Egg, Diaper, Milk → Coffee
0.028	1.000	0.028	11.600	3.103	0.019	Ham, Egg, Nuts → Apple
0.028	1.000	0.028	9.800	3.673	0.020	Egg, Nuts, Apple → Ham
0.028	1.000	0.028	11.600	3.103	0.019	Egg, Coffee, Salad → Apple
0.022	1.000	0.022	14.500	3.103	0.015	Banana, Ketchup, Rootbeer, Ginger → Apple

Rule1: Ice Cream, Olive, Tea –> Banana (minsup=3.3%)
Rule2: Banana, Ham, Salad –> Apple (minsup=3.3%)
Rule3: Ham, Diaper, Coffee –> Ice Cream (minsup=2.8%)

5 Question 5:

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent
0.022	1.000	0.022	4.250	10.588	0.020	Coffee, Salad, Lemon → Egg, Apple
0.022	1.000	0.022	5.000	9.000	0.020	Nuts, Coffee, Salad → Apple, Ketchup
0.028	1.000	0.028	8.200	4.390	0.021	Diaper, Coffee, Ginger → Egg
0.022	1.000	0.022	10.250	4.390	0.017	Diaper, Apple, Rootbeer → Egg

Interesting Rules:
Coffee, Salad, Lemon –> Egg, Apple (Minsup=2%, Minconf=100%, Lift = 10.588)
Nuts, Coffee, Salad –> Apple, Ketchup (Minsup=2%, Minconf=100%, Lift = 9)

These are the two most interesting rules I found when minimum support is 2% and minimum confidence is 100%. The measure used to identify the interestingness of the rule is lift. Lift shows the correlaton between the two itemsets. If the lift of the rules is high, it means the probability of occurrence of the antecedent is low. The higher the value of lift is, the more positively correlated these two datasets are.