**Lab #6: Computing Support, Confidence, and Lift**

Here are the baskets from eight shoppers:

|  |  |
| --- | --- |
| **Basket** | **Items** |
| 1 | Coke, Pop-Tarts, Donuts |
| 2 | Cheerios, Coke, Donuts, Napkins |
| 3 | Waffles, Cheerios, Coke, Napkins |
| 4 | Bread, Milk, Coke, Napkins |
| 5 | Coffee, Bread, Waffles |
| 6 | Coke, Bread, Pop-Tarts |
| 7 | Milk, Waffles, Pop-Tarts |
| 8 | Coke, Pop-Tarts, Donuts, Napkins |

Compute the support, confidence, and lift for the following rules:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Rule** | | **Support** | **Confidence** | **Lift** |
| 1 | {Coke, Pop-Tarts} 🡪{Donuts} | **2/8=0.25** | **0.25/0.375=0.67** | **0.25/ (0.375\*0.375) =1.78** |
| 2 | {Coke} 🡪 {Pop-Tarts, Donuts} | **2/8=0.25** | **0.25/0.75=0.33** | **0.25/ (0.75\*0.25) =1.33** |
| 3 | {Coffee} 🡪 {Bread, Waffles} | **1/8=0.125** | **0.125/0.125=1** | **0.125/ (0.125\*0.125) =8** |
| 4 | {Coke} 🡪 {Donuts} | **3/8=0.375** | **0.375/0.75=0.50** | **0.375/ (0.75\*0.375) =1.33** |

1. Which rule has the strongest association? How do you know?

**Rule 3 {Coffee} 🡪 {Bread, Waffles}**

**Highest confidence 1.00**

1. Consider a customer who is walking through the store with only a bottle of coke in their shopping cart. You then see them put pop-tarts in their cart. Do you become more or less sure than you were before that they will buy donuts? Explain.

**Rule 1: A: Coke B: Donut lift= (Coke, donut)=1.33**

**Rule 2 A: Coke and pop-tarts B: donuts lift=(Coke and pop-tarts, donuts)**

**=1.78**

**More sure**

***Computing lift based on aggregate purchase numbers***

1. Consider two products, the Squishee and the Peanut Butter Bowl. Here’s a profile of 18,500 customers:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Squishee | | |  |
| Peanut Butter Bowl |  | No | Yes |  | |
| No | 10000 | 2000 |  | |
| Yes | 1500 | 5000 |  | |
|  |  |  |  | Total: 18500 | |

What is the lift for the rule {Peanut Butter Bowl} 🡪 {Squishee}?  
(Are people who bought a Peanut Butter Bowl more likely than chance to buy a Squishee too?)

**Support (Peanut Butter Bowl, Squishee) = 5000/18500 = 0.270**

**Support (Peanut Butter Bowl) = 6500/18500 = 0.351**

**Support (Squishee) = 7000/18500 = 0.378**

**Lift= 0.27/(0.351\*0.378) = 2.03>1**

**So these products appear in the same basket more often than what you'd expect by chance (i.e., Lift > 1). If a customer buys a Peanut Butter Bowl, they are more likely than chance to also buy a Squishee.**

1. Consider two products, Potato Chips and Krusty-O’s. Here’s a profile of 10,500 customers:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Krusty-O’s | | |  |
| Potato Chips |  | No | Yes |  | |
| No | 5000 | 1000 |  | |
| Yes | 4000 | 500 |  | |
|  |  |  |  | Total: 10500 | |

What is the lift for the rule {Potato Chips} 🡪 {Krusty-O’s}?  
(Are people who bought Potato Chips more likely than chance to buy Krusty-O’s too?)

**Support(Potato Chips, Krusty-O's) = 500/10500 = 0.048**

**Support(Potato Chips) = 4500/10500 = 0.429**

**Support(Krusty-O's) = 1500/10500 = 0.143**

**Lift= 0.048/(0.429\*0.143) = 0.782<1**

**So these products appear in the same basket less often than what you'd expect by chance (i.e., Lift < 1). If a customer buys Potato Chips, they are less likely than chance to also buy Krusty-O's**