**Name:** Shukhrat Khuseynov **ID:** 0070495 **Date:** 01.11.2019

**Course:** DASC 501 **Assignment 2B - Affinity Analysis (multiple LHS)**

Discussion (Interpretation)

**Part 1.**

The output for Part 1:

The rules:

1. cheese -> apples: Support = 0.22 Confidence = 0.56 Lift = 1.31

acceptable & interesting

2. bread -> cheese: Support = 0.05 Confidence = 0.18 Lift = 0.46

3. milk -> bread: Support = 0.13 Confidence = 0.25 Lift = 0.89

4. bananas -> milk: Support = 0.27 Confidence = 0.47 Lift = 0.91

5. apples -> bananas: Support = 0.27 Confidence = 0.63 Lift = 1.10

acceptable & interesting

(acceptable: confidence > 50% & interesting: positively associated)

The definition for the terms “acceptable” and “interesting” was obtained from the slides of assignment description.

These 5 rules were just obtained randomly by logic. Numpy arrays were employed in the initial step. Calculating support, confidence and lift was not difficult for 5 rules.

It is difficult to choose one rule among two both acceptable and interesting ones, but I recommend the rule #1 because its lift is proportionally much greater than the confidence of rule #**5**, if to compare both.

**Part 2.**

The output for Part 2 is stored in Rules(B).txt after running the code and choosing part 2 as an input. The chosen ones among them (only acceptable and interesting rules):

4. bread -> bananas:

Support = 0.16 Confidence = 0.57 Lift = 1.00 acceptable & interesting

11. cheese -> apples:

Support = 0.22 Confidence = 0.56 Lift = 1.31 acceptable & interesting

15. apples -> cheese:

Support = 0.22 Confidence = 0.51 Lift = 1.31 acceptable & interesting

16. apples -> bananas:

Support = 0.27 Confidence = 0.63 Lift = 1.10 acceptable & interesting

24. bread, cheese -> milk:

Support = 0.03 Confidence = 0.60 Lift = 1.15 acceptable & interesting

27. bread, apples -> milk:

Support = 0.06 Confidence = 0.67 Lift = 1.28 acceptable & interesting

44. cheese, apples -> bananas:

Support = 0.14 Confidence = 0.64 Lift = 1.12 acceptable & interesting

47. cheese, bananas -> apples:

Support = 0.14 Confidence = 0.70 Lift = 1.63 acceptable & interesting

50. apples, bananas -> cheese:

Support = 0.14 Confidence = 0.52 Lift = 1.33 acceptable & interesting

51. bread, milk, cheese -> apples:

Support = 0.02 Confidence = 0.67 Lift = 1.55 acceptable & interesting

57. bread, cheese, apples -> milk:

Support = 0.02 Confidence = 1.00 Lift = 1.92 acceptable & interesting

61. bread, apples, bananas -> milk:

Support = 0.03 Confidence = 0.60 Lift = 1.15 acceptable & interesting

66. milk, cheese, bananas -> apples:

Support = 0.02 Confidence = 0.50 Lift = 1.16 acceptable & interesting

(acceptable: confidence > 50% & interesting: positively associated)

Note: if lift is written as 1.00 and it shows “interesting”, it means that lift is greater than 1 exhibiting positive association, but after rounding it became 1.00, nothing is wrong there.

The 75 rules with multiple LHS were obtained as a list of tuples in multiple loops creating combination of LHS and RHS elements, with a separate step for each amount of LHS. Numpy arrays were employed in the initial step. The support, confidence and lift were computed in multiple loops, requiring the frequencies/supports of single elements calculated beforehand and of multiple elements calculated in loops throughout the algorithm and saved as a dictionary. Then, the invalid rules were determined, not allowing to have any Zero Division Error in calculations.

As seen in the output, there are only 13 acceptable and interesting rules, where only 5 have to be recommended. I analyzed them mainly through the lift and confidence parameters, checking the support as a last step. I could recommend the rules #57, #51 or even #27 which seem very good, but since their support is quite low, more evidence is needed to prove them, so gathering more transaction samples might help. Sorting the remaining rules and picking among them, I recommend rule #47, rule #11 (inverse of #15 but better), rule #50, rule #44 and rule #16 from highest to lowest degree of my recommendation respectively.