**Questions 1, 2, and 3 should be solved manually. Show all steps of your work.**

1. **Assume we have 4 items in our set of items you can call them 1, 2 3, and 4**
2. **How many possibilities are shown on the itemset lattice. Create the itemset lattice.**
3. **Show the impact of the following assumption on the itemset lattice. Assume Itemset {2, 3} is infrequent**
4. **Solution:**

Set of Items= {1,2,3,4}

Itemset with 1 item = {(1),(2),(3),(4)}

Itemset with 2 items= {(1,2),(1,3),(1,4),(2,3),(2,4),(3,4)}

Itemset with 3 items = {(1,2,3),(1,2,4),(1,3,4),(2,3,4)}

Itemset with 4 items ={(1,2,3,4)}

Itemset lattice={(1),(2),(3),(4), (1,2),(1,3),(1,4),(2,3),(2,4),(3,4), (1,2,3),(1,2,4),(1,3,4),(2,3,4),(1,2,3,4)}

Total Number of possibilities =15

1. **Solution:** Assuming that the itemset {2, 3} is infrequent affects the association rule mining process by leading to the removal or pruning of this specific combination of items and any related associations. This assumption simplifies the analysis by excluding infrequent itemsets, potentially enhancing efficiency and allowing a focus on more meaningful and significant associations involving more frequent items. While it streamlines the process and highlights stronger relationships, it's essential to acknowledge that context and domain expertise remain crucial for interpreting the impact accurately.

Itemset lattice= {(1),(2),(3),(4), (1,2),(1,3),(1,4),(2,4),(3,4),(1,2,4),(1,3,4)}

1. **Consider the following 10 transaction set.**

|  |  |
| --- | --- |
| **TID** | **List of purchased items** |
| **T1** | **A, B, E** |
| **T2** | **B, D** |
| **T3** | **B, C** |
| **T4** | **A, B, D** |
| **T5** | **A, C** |
| **T6** | **B, C** |
| **T7** | **A, C** |
| **T8** | **A, B, C, E** |
| **T9** | **A, B, C** |
| **T10** | **C, D, E** |

* 1. **What is the set of items?**
  2. **Suppose minimum count required is 3, what are set of item set with the equal of greater than 3 count? You should apply Apriori algorithm**

**Solution:**

1. Set of items ={A, B, C, D, E}
2. Given Minimum count =3,

|  |  |
| --- | --- |
| Items | Frequency |
| A | 6 |
| B | 7 |
| C | 7 |
| D | 3 |
| E | 3 |

Check each item with the min count=2

|  |  |
| --- | --- |
| Rule | Support count |
| {A} | 6 |
| {B} | 7 |
| {C} | 7 |
| {D} | 3 |
| {E} | 3 |

Now lets do two frequent itemset

|  |  |
| --- | --- |
| Rule | Support count |
| {A,B} | 4 |
| {A,C} | 4 |
| {A,D} | 1 |
| {A,E} | 2 |
| {B,C} | 4 |
| {B,D} | 2 |
| {B,E} | 2 |
| {C,D} | 1 |
| {C,E} | 2 |
| {D,E} | 1 |

Now check two frequent itemset with min count

|  |  |
| --- | --- |
| Rule | Support count |
| {A,B} | 4 |
| {A,C} | 4 |
| {B,C} | 4 |

Three frequent itemset

|  |  |
| --- | --- |
| Rule | Frequency |
| {A,B,C} | 2 |

Now as you see there are no more rules to generate! Thus, algorithm output is the 1-itemset, 2-itemset, and 3-itemset with frequency greater than 3

{A},{B},{C},{D},{E},{A,B},{A,C},{B,C}

|  |  |
| --- | --- |
| Rule | Support count |
| {A} | 6 |
| {B} | 7 |
| {C} | 7 |
| {D} | 3 |
| {E} | 3 |
| {A,B} | 4 |
| {A,C} | 4 |
| {B,C} | 4 |

1. **Use the transaction list and compute the Support, Confidence, and the lift for the following rule:**

**{A, B} => {E}. evaluate the result.**

|  |  |
| --- | --- |
| **TID** | **List of purchased items** |
| T1 | A, B, E |
| T2 | B, D |
| T3 | B, C |
| T4 | A, B, D, E |
| T5 | A, C |
| T6 | B, C |
| T7 | A, C |
| T8 | A, B, C, E |
| T9 | A, B, C |
| T10 | C, D, E |

**Solution:** Total number of transactions=10

|  |  |
| --- | --- |
| Rule | Frequency |
| A | 4 |
| B | 7 |
| E | 4 |
| A, B | 4 |
| A, B, E | 3 |

**Formulas for support and confidence:**

and

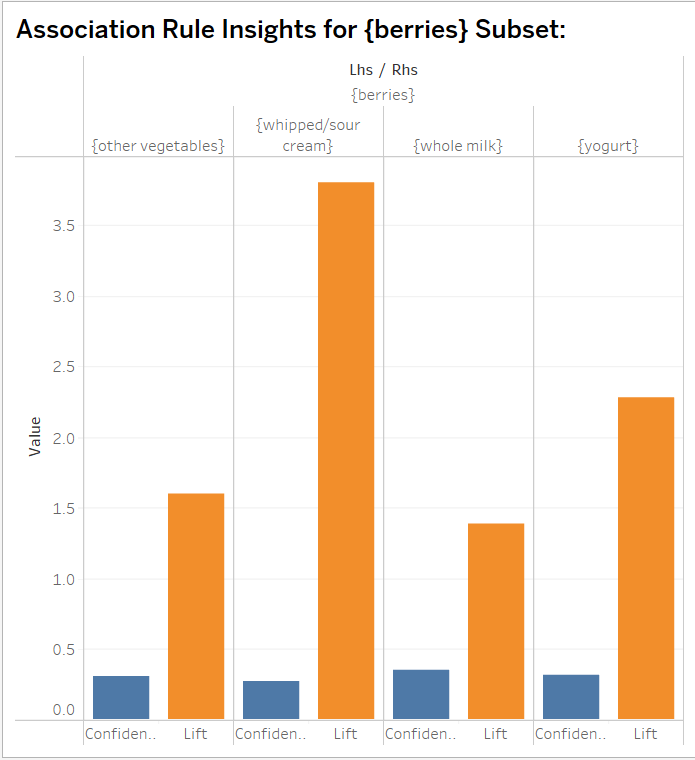
Given Rule {A,B} =>{E}

**Support:** Supp({{A,B} =>{E})= = =0.3

**Confidence: Conf(**{{A,B} =>{E})= = 3/4= 0.75

**Lift:** Lift({A,B} =>{E}) = = = 0.3/0.4\*0.4= 1.875

1. **Use basketrules.csv file to create a visualization in Tableau. Interpret the visualization.**

****