

## Quiz 7 หา frequent itemsets และ association rules ที่น่าสนใจ

| Transaction ID | Items Purchased          |
|----------------|--------------------------|
| T1             | Milk, Bread, Butter      |
| T2             | Milk, Eggs               |
| T3             | Bread, Butter, Jam       |
| T4             | Milk, Eggs, Butter       |
| T5             | Bread, Butter, Jam, Milk |

### หา Frequent Itemsets

- ขั้นตอนที่ 1: หาความถี่ (Frequency) ของแต่ละ Item:
  - Milk: 4 (ปรากฏใน T1, T2, T4, T5)
  - Bread: 3 (ปรากฏใน T1, T3, T5)
  - Butter: 4 (ปรากฏใน T1, T3, T4, T5)
  - Eggs: 2 (ปรากฏใน T2, T4)
  - Jam: 2 (ปรากฏใน T3, T5)
- ขั้นตอนที่ 2: หาความถี่ของ Itemsets ขนาด 2 (Pairs of items):
  - {Milk, Bread}: 2 (T1, T5)
  - {Milk, Butter}: 3 (T1, T4, T5)
  - {Milk, Eggs}: 2 (T2, T4)
  - {Milk, Jam}: 1 (T5)
  - {Bread, Butter}: 3 (T1, T3, T5)
  - {Bread, Eggs}: 1 (ไม่มีในชุด)
  - {Bread, Jam}: 2 (T3, T5)
  - {Butter, Eggs}: 1 (ไม่มีในชุด)
  - {Butter, Jam}: 2 (T3, T5)
  - {Eggs, Jam}: 1 (ไม่มีในชุด)
- ขั้นตอนที่ 3: คำนวณ Support สำหรับแต่ละ itemset และเลือกเฉพาะ itemsets ที่มี Support สูง (เช่น Support > 0.4 หรือ 40% จากจำนวน transaction ทั้งหมด)

สมมติว่า threshold ของ Support คือ 40% หรือ 2 ใน 5:

- {Milk, Bread}:  $2/5 = 0.4$  (frequent itemset)
- {Milk, Butter}:  $3/5 = 0.6$  (frequent itemset)
- {Milk, Eggs}:  $2/5 = 0.4$  (frequent itemset)
- {Bread, Butter}:  $3/5 = 0.6$  (frequent itemset)
- {Bread, Jam}:  $2/5 = 0.4$  (frequent itemset)
- {Butter, Jam}:  $2/5 = 0.4$  (frequent itemset)

itemsets ที่มี Support สูงกว่าหรือเท่ากับ 0.4 ได้แก่:

- {Milk, Bread}
- {Milk, Butter}
- {Milk, Eggs}
- {Bread, Butter}
- {Bread, Jam}
- {Butter, Jam}

## หา Association Rules

หลังจากได้ frequent itemsets แล้ว เราสามารถหากฎการเชื่อมโยง (Association Rules) โดยการคำนวณ **confidence** และ **lift** ซึ่งจะช่วยบ่งบอกถึงความน่าสนใจของกฎ

- **Rule 1:** {Milk}  $\rightarrow$  {Butter}
  - Support:  $3/5 = 0.6$
  - Confidence:  $3/4 = 0.75$  (ถ้ามี Milk จะมี Butter 75% ของกรณี)
  - Lift:  $0.75 / (\text{Support of Butter}) = 0.75 / 0.8 = 0.9375$

## Association Rules ที่น่าสนใจ:

**{Bread}  $\rightarrow$  {Butter}** (Confidence = 1.0, Lift = 1.25)



