# การหากฎความสัมพันธ์
# (ASSOCIATION RULE)

# จุดประสงค์การเรียนรู้

- ผู้เรียนสามารถอธิบายส่วนต่าง ๆ ของ RapidMiner Studio ได้
- ผู้เรียนสามารถอธิบายวิธีการหาความสัมพันธ์ของข้อมูล(Association rules) ใน RapidMiner Studio ได้

# เนื้อหา

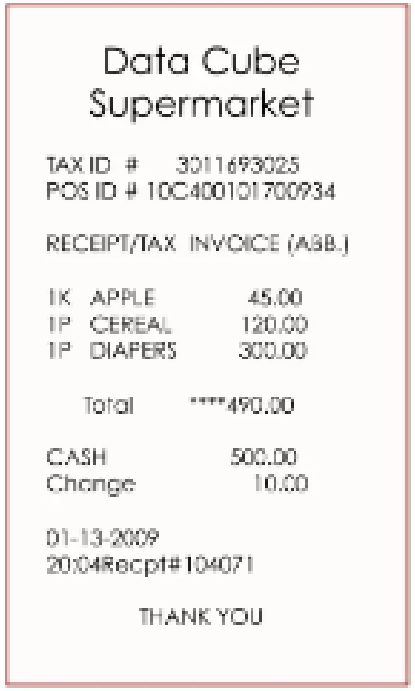
- การหากฎความสัมพันธ์ (Association Rules)
- Apriori
- FP-Growth

# กฎความสัมพันธ์ (Association Rule)

- กฎความสัมพันธ์ (Association Rule)
  - เป็นกระบวนการหนึ่งในการทำเหมืองข้อมูล ที่ได้รับความนิยมนำมาใช้ในการวิเคราะห์หาความสัมพันธ์ของสองชุด หรือมากกว่าสองชุดขึ้นไป
  - การหากฎความสัมพันธ์ที่เป็นที่รู้จักและใช้อย่างแพร่หลาย ได้แก่
    - Apriori
    - FP-Growth
  - ตัวอย่าง Association Rule เช่น Market Basket Analysis ที่ใช้วิเคราะห์ข้อมูลเพื่อดูว่าสินค้าอะไรบ้างที่ลูกค้าซื้อร่วมกันบ่อย ๆ หรือหาความสัมพันธ์ของสินค้าที่ลูกค้ามักจะซื้อพร้อมกัน เพื่อใช้ในการจัดรายการส่งเสริมการขายต่อไป

# กฎความสัมพันธ์ (Association Rule)

- Market Basket Analysis
- ข้อมูลที่ได้จากการซื้อสินค้าแต่ละครั้ง


Data Cube
Supermarket

TAX ID # 3011693025
POS ID # 10C400101700934

RECEIPT/TAX INVOICE (ABB.)

1K APPLE 45.00
1P CEREAL 120.00
1P DIAPERS 300.00

Total ****490.00

CASH 500.00
Change 10.00

01-13-2009
20:04Recpt#104071

THANK YOU

| TID | Transaction time | Product |
|---|---|---|
| 1 | 01-13-2014 20:04 | Apple |
| 1 | 01-13-2014 20:04 | Cereal |
| 1 | 01-13-2014 20:04 | Diapers |

# กฎความสัมพันธ์ (Association Rule)

- Market Basket Analysis
- ข้อมูลที่ได้จากการซื้อสินค้าแต่ละครั้ง

| TID | Transaction time | Product |
|---|---|---|
| 1 | 01-13-2014 20:04 | Apple |
| 1 | 01-13-2014 20:05 | Cereal |
| 1 | 01-13-2014 20:06 | Diapers |
| 2 | 01-14-2014 11:30 | Bear |
| 2 | 01-14-2014 11:31 | Cereal |
| 2 | 01-14-2014 11:32 | Eggs |

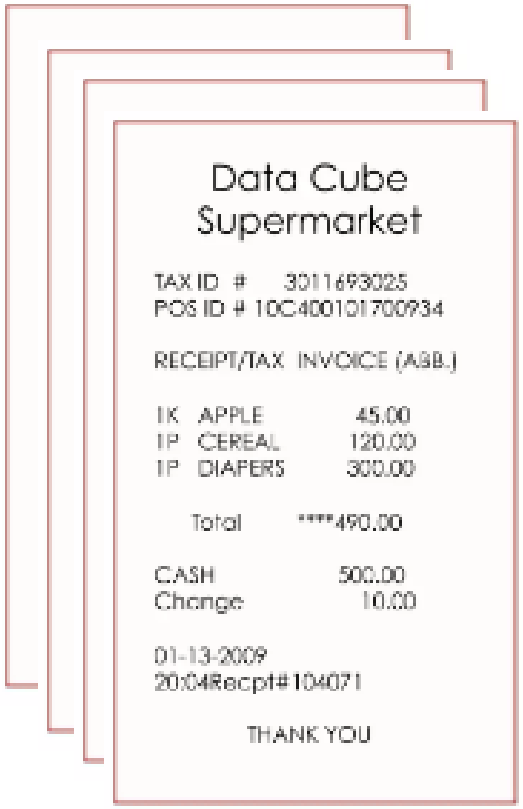
# กฎความสัมพันธ์ (Association Rule)

- Market Basket Analysis
- ข้อมูลที่ได้จากการซื้อสินค้าแต่ละครั้ง

| TID | Transaction time | Product |
|---|---|---|
| 1 | 01-13-2014 20:04 | Apple |
| 1 | 01-13-2014 20:05 | Cereal |
| 1 | 01-13-2014 20:06 | Diapers |
| 2 | 01-14-2014 11:30 | Bear |
| 2 | 01-14-2014 11:31 | Cereal |
| 2 | 01-14-2014 11:32 | Eggs |
| 3 | 01-15-2014 14:30 | Apple |
| 3 | 01-15-2014 14:31 | Bear |
| 3 | 01-15-2014 14:32 | Cereal |
| 3 | 01-15-2014 14:33 | Eggs |

# กฎความสัมพันธ์ (Association Rule)

- Market Basket Analysis
- ข้อมูลที่ได้จากการซื้อสินค้าแต่ละครั้ง




| TID | Transaction time | Product |
|---|---|---|
| 1 | 01-13-2014 20:04 | Apple |
| 1 | 01-13-2014 20:05 | Cereal |
| 1 | 01-13-2014 20:06 | Diapers |
| 2 | 01-14-2014 11:30 | Bear |
| 2 | 01-14-2014 11:31 | Cereal |
| 2 | 01-14-2014 11:32 | Eggs |
| 3 | 01-15-2014 14:30 | Apple |
| 3 | 01-15-2014 14:31 | Bear |
| 3 | 01-15-2014 14:32 | Cereal |
| 3 | 01-15-2014 14:33 | Eggs |
| 4 | 01-162014 14:15 | Bear |
| 4 | 01-162014 14:16 | Eggs |

# กฎความสัมพันธ์ (Association Rule)

- แปลงข้อมูลจาก POS database เป็น transaction database โดย group by ตามเวลาที่ซื้อสินค้า

| TID | Transaction time | Product |
|---|---|---|
| 1 | 01-13-2014 20:04 | Apple |
| 1 | 01-13-2014 20:05 | Cereal |
| 1 | 01-13-2014 20:06 | Diapers |
| 2 | 01-14-2014 11:30 | Beer |
| 2 | 01-14-2014 11:31 | Cereal |
| 2 | 01-14-2014 11:32 | Eggs |
| 3 | 01-15-2014 14:30 | Apple |
| 3 | 01-15-2014 14:31 | Beer |
| 3 | 01-15-2014 14:32 | Cereal |
| 3 | 01-15-2014 14:33 | Eggs |
| 4 | 01-162014 14:15 | Beer |
| 4 | 01-162014 14:16 | Eggs |

| Transaction ID | Items |
|---|---|
| 1 | Apple, Cereal, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Apple, Beer, Cereal, Eggs |
| 4 | Beer, Eggs |

# กฎความสัมพันธ์ (Association Rule)

- Support
  - นับจำนวนครั้งการซื้อสินค้าแต่ละชนิดคิดเป็น %
  - Apple พบว่ามีการซื้อในครั้งที่ 1 และ 3 support(Apple) คือ 2/4 = 50%
  - Beer พบว่ามีการซื้อในครั้งที่ 2,3 และ 4 support(Beer) คือ ¾ = 75%

| Transaction ID | Items |
|---|---|
| 1 | Apple, Cereal, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Apple, Beer, Cereal, Eggs |
| 4 | Beer, Eggs |

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Apple | 1 | 0 | 1 | 0 | 2/4 = 50% |
| Beer | 0 | 1 | 1 | 1 | 3/4 =75% |
| Cereal | 1 | 1 | 1 | 0 | 3/4 = 75% |
| Diapers | 1 | 0 | 0 | 0 | 1/4 = 25% |
| Eggs | 0 | 1 | 1 | 1 | 3/4 = 75% |

# กฎความสัมพันธ์ (Association Rule)

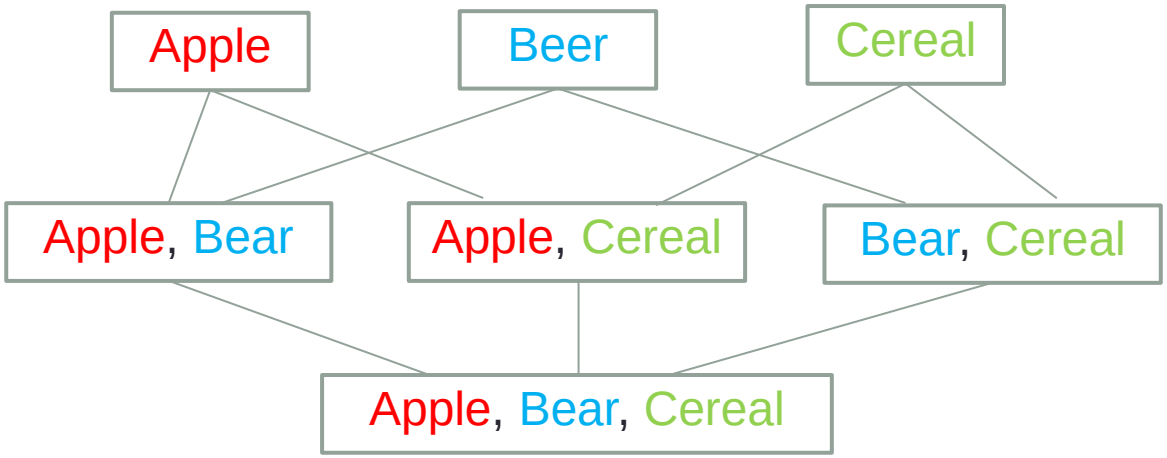
- Frequent itemset
  - รูปแบบของการซื้อสินค้า (itemset) ที่มีค่า support >= mininum support
  - กำหนดค่า mininum support = 50% ตัวอย่างของ Frequent itemset คือ {Apple},{Beer},{Cereal},{Eggs}
- Infrequent itemset
  - รูปแบบของการซื้อสินค้า (itemset) ที่มีค่า support < mininum support

| Transaction ID | Items |
|---|---|
| 1 | Apple, Cereal, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Apple, Beer, Cereal, Eggs |
| 4 | Beer, Eggs |

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Apple | 1 | 0 | 1 | 0 | 2/4 = 50% |
| Beer | 0 | 1 | 1 | 1 | 3/4 =75% |
| Cereal | 1 | 1 | 1 | 0 | 3/4 = 75% |
| Diapers | 1 | 0 | 0 | 0 | 1/4 = 25% |
| Eggs | 0 | 1 | 1 | 1 | 3/4 = 75% |

# กฎความสัมพันธ์ (Association Rule)

- สินค้าใน supermarket มีจำนวน 3 ชนิด
  - Apple, Beer, Cereal
  - มีโอกาสที่เป็นไปได้ทั้งหมด (หรือ subset ทั้งหมด๗ คือ $2^3$ -1 = 7 รูปแบบ หรือ คำนวณได้จาก $2^N$ - 1 โดยที่ N คือ จำนวนของสินค้า (item)



# กฎความสัมพันธ์ (Association Rule)

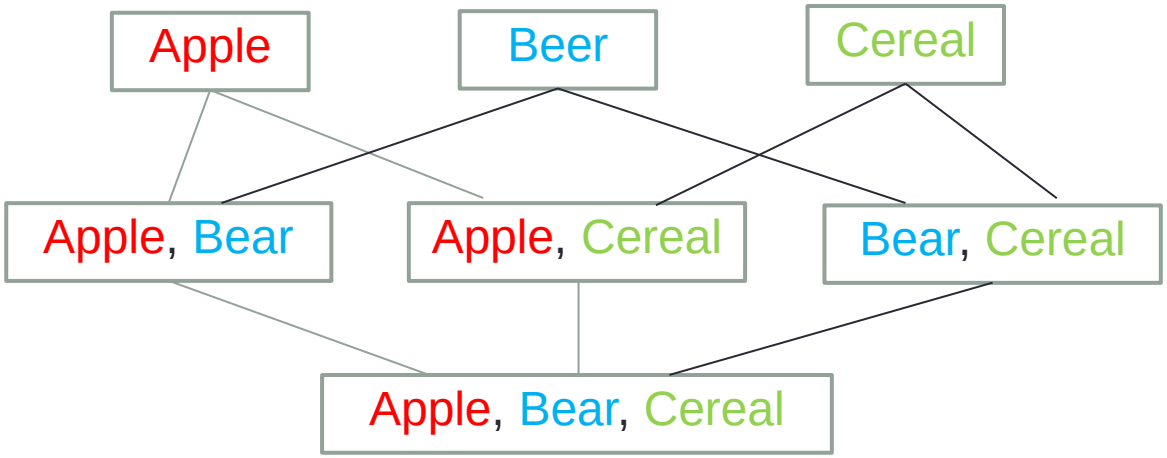
- การหากฎความสัมพันธ์ (association rule) มี 2 ขั้นตอนใหญ่ ๆ
  - ขั้นตอนที่ 1 หา frequent itemset ซึ่งใช้เวลานานกว่าขั้นตอนที่ 2
    - มี 2 เทคนิคที่นิยมใช้คือ
      - เทคนิค Apriori
      - เทคนิค FP-Growth
  - ขั้นตอนที่ 2 สร้างกฎความสัมพันธ์จาก frequent itemset ที่หาได้จากขั้นตอนที่ 1

# กฎความสัมพันธ์ (Association Rule)

- เทคนิคการหากฎความสัมพันธ์ (Association Rules)
  - Apriori (Agrawal and Srikant,1994)
    - สร้างรูปแบบของสินค้า (itemset) ที่มีจำนวนเพิ่มขึ้นที่ 1
    - คำนวณค่า support จากในฐานข้อมูล
    - ข้อเสียคือ ต้องดึงข้อมูลจากฐานข้อมูลหลายรอบทำให้ทำงานช้า
  - FP-Growth (Han, Pei and Yin,2000)
    - อ่านข้อมูลในฐานข้อมูลและสร้าง FP-Tree
    - คำนวณค่า support จาก FP-Tree
    - ทำงานได้เร็วกว่าวิธี Apriori

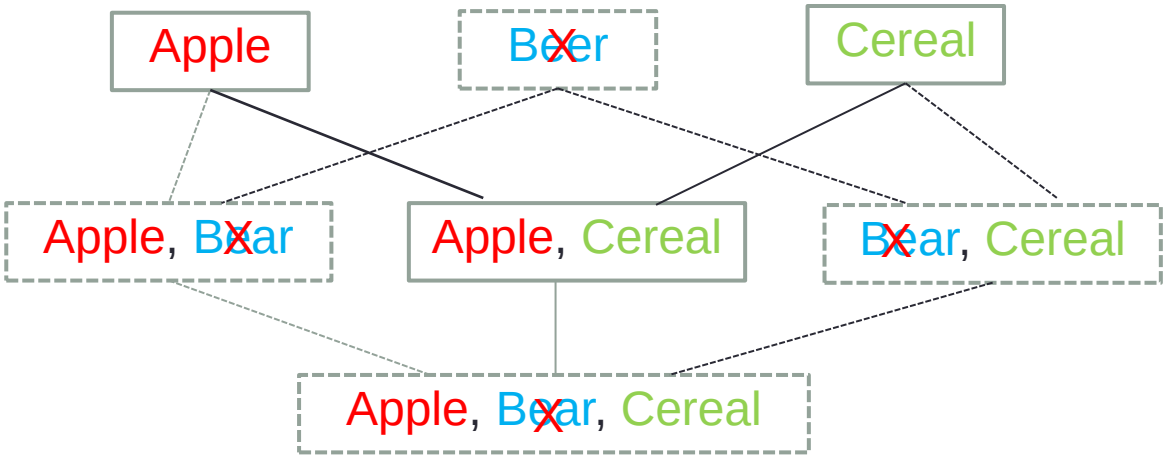
# กฎความสัมพันธ์ (Association Rule)

- Apriori
  - หลักการของ Apriori
    - Superset ของรูปแบบที่เป็น frequent itemset มีโอกาสจะเป็น frequent itemset ที่มีความยาวมากขึ้น



# กฎความสัมพันธ์ (Association Rule)

- Apriori
  - หลักการของ Apriori
    - Superset ของรูปแบบที่เป็น frequent itemset มีโอกาสจะเป็น frequent itemset ที่มีความยาวมากขึ้น
    - Superset ของรูปแบบที่เป็น infrequent itemset จะมีค่า support < minimum support สามารถตัดรูปแบบที่เป็น superset ที่มีความยาวมากขึ้นทิ้งได้



# กฎความสัมพันธ์ (Association Rule)

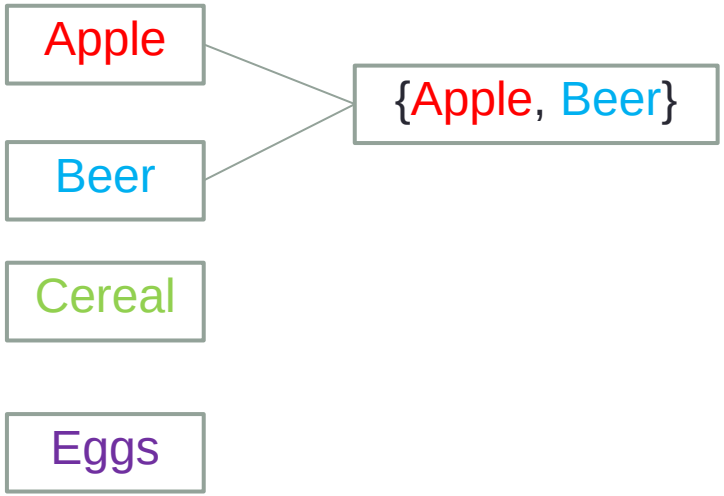
- Apriori
  - วิเคราะห์เพื่อหาสินค้าที่มีการซื้อมากกว่าหรือเท่ากับ 50%
  - Support
    - นับจำนวนครั้งการซื้อสินค้าแต่ละชนิดคิดเป็น % ของการซื้อสินค้า

| Transaction ID | Items |
|---|---|
| 1 | Apple, Cereal, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Apple, Beer, Cereal, Eggs |
| 4 | Beer, Eggs |

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Apple | 1 | 0 | 1 | 0 | 2/4 = 50% |
| Beer | 0 | 1 | 1 | 1 | 3/4 =75% |
| Cereal | 1 | 1 | 1 | 0 | 3/4 = 75% |
| ~~Diapers~~ | ~~1~~ | ~~0~~ | ~~0~~ | ~~0~~ | ~~1/4 = 25%~~ |
| Eggs | 0 | 1 | 1 | 1 | 3/4 = 75% |

# กฎความสัมพันธ์ (Association Rule)

- รูปแบบ (Pattrn) ที่มีสินค้า 2 ชนิด
  - รูปแบบที่มี Apple อยู่

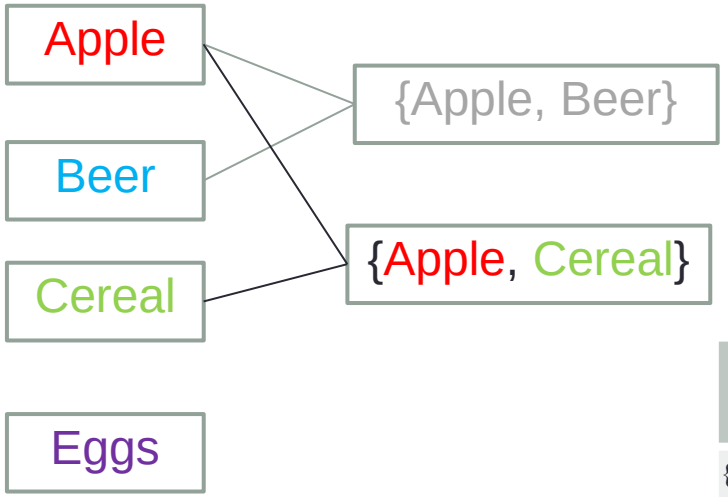



| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Apple | 1 | 0 | 1 | 0 | 2/4 = 50% |
| Beer | 0 | 1 | 1 | 1 | 3/4 =75% |

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| {Apple,Beer} | 0 | 0 | 1 | 0 | 1/4 = 25% |

# กฎความสัมพันธ์ (Association Rule)

- รูปแบบ (Pattrn) ที่มีสินค้า 2 ชนิด
  - รูปแบบที่มี Apple อยู่

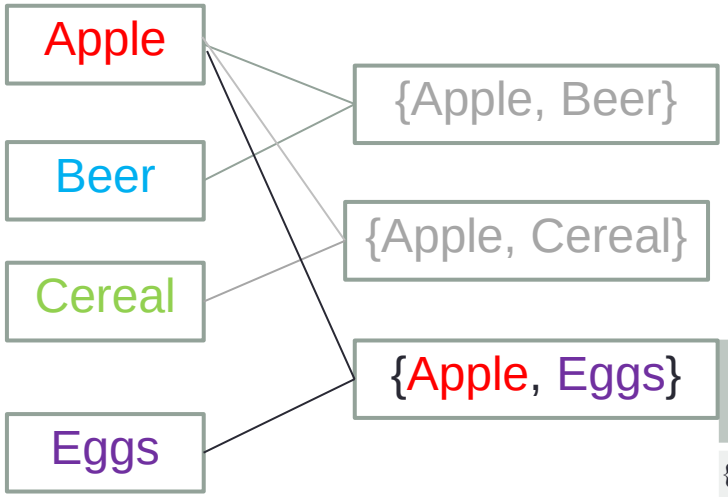



| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Apple | 1 | 0 | 1 | 0 | 2/4 = 50% |
| Cereal | 1 | 1 | 1 | 0 | 3/4 = 75% |

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| {Apple,Cereal} | 1 | 0 | 1 | 0 | 2/4 = 50% |

# กฎความสัมพันธ์ (Association Rule)

- รูปแบบ (Pattrn) ที่มีสินค้า 2 ชนิด
  - รูปแบบที่มี Apple อยู่




| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Apple | 1 | 0 | 1 | 0 | 2/4 = 50% |
| Eggs | 0 | 1 | 1 | 1 | 3/4 = 75% |

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| {Apple,Eggs} | 0 | 0 | 1 | 0 | 1/4 = 25% |

# กฎความสัมพันธ์ (Association Rule)

- รูปแบบ (Pattrn) ที่มีสินค้า 2 ชนิด
  - รูปแบบที่มี Beer อยู่

Apple

Beer

Cereal

Eggs

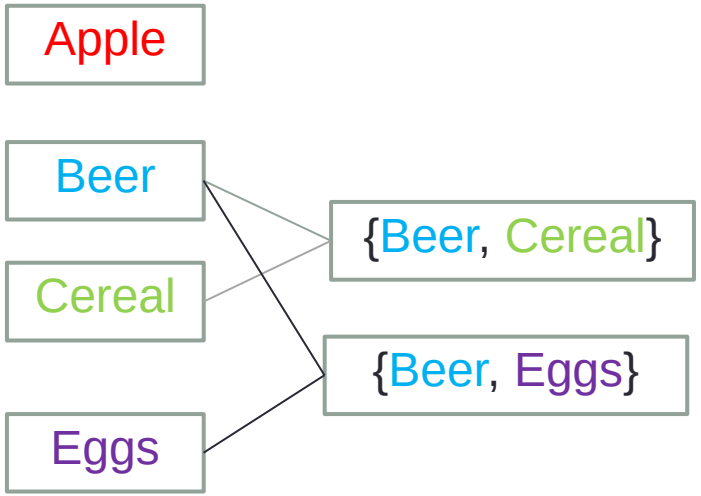
{Beer, Cereal}

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Beer | 0 | 1 | 1 | 1 | 3/4 =75% |
| Cereal | 1 | 1 | 1 | 0 | 3/4 = 75% |

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| {Beer,Cereal} | 0 | 1 | 1 | 0 | 2/4 = 50% |

# กฎความสัมพันธ์ (Association Rule)

- รูปแบบ (Pattrn) ที่มีสินค้า 2 ชนิด
  - รูปแบบที่มี Beer อยู่




| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Beer | 0 | 1 | 1 | 1 | 3/4 =75% |
| Eggs | 0 | 1 | 1 | 1 | 3/4 = 75% |

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| {Beer,Eggs} | 0 | 1 | 1 | 1 | 3/4 = 75% |

# กฎความสัมพันธ์ (Association Rule)

- รูปแบบ (Pattrn) ที่มีสินค้า 2 ชนิด
  - รูปแบบที่มี Cereal อยู่

Apple

Beer

Cereal

Eggs

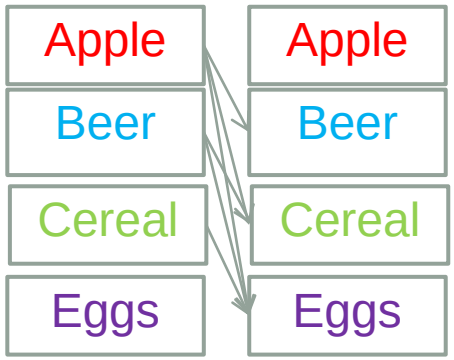
{Cereal, Eggs}

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Cereal | 1 | 1 | 1 | 0 | 3/4 = 75% |
| Eggs | 0 | 1 | 1 | 1 | 3/4 = 75% |

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| {Cereal,Eggs} | 0 | 1 | 1 | 0 | 2/4 = 50% |

# กฎความสัมพันธ์ (Association Rule)

- นับจำนวนครั้งการซื้อสินค้า 2 ชนิด




- {Apple, Beer} {Apple, Cereal},{Apple,Eggs}
- {Beer,Cereal} {Beer,Eggs}
- {Cereal, Eggs}

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| {Apple,Beer} | 0 | 0 | 1 | 0 | 1/4 = 25% |
| {Apple,Cereal} | 1 | 0 | 1 | 0 | 2/4 =50% |
| {Apple,Eggs} | 0 | 0 | 1 | 0 | 1/4 = 25% |
| {Beer,Cereal} | 0 | 1 | 1 | 0 | 2/4 = 50% |
| {Beer,Eggs} | 0 | 1 | 1 | 1 | 3/4 = 75% |
| {Cereal,Eggs} | 0 | 1 | 1 | 0 | 2/4 = 50% |

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| ~~{Apple,Beer}~~ | ~~0~~ | ~~0~~ | ~~1~~ | ~~0~~ | ~~1/4 = 25%~~ |
| {Apple,Cereal} | 1 | 0 | 1 | 0 | 2/4 =50% |
| ~~{Apple,Eggs}~~ | ~~0~~ | ~~0~~ | ~~1~~ | ~~0~~ | ~~1/4 = 25%~~ |
| {Beer,Cereal} | 0 | 1 | 1 | 0 | 2/4 = 50% |
| {Beer,Eggs} | 0 | 1 | 1 | 1 | 3/4 = 75% |
| {Cereal,Eggs} | 0 | 1 | 1 | 0 | 2/4 = 50% |

# กฎความสัมพันธ์ (Association Rule)

- นับจำนวนครั้งการซื้อสินค้า 3 ชนิด

| | |
|---|---|
| Apple,Cereal | Apple,Cereal |
| Beer,Cereal | Beer,Cereal |
| Beer, Eggs | Beer, Eggs |
| Cereal,Eggs | Cereal,Eggs |

- {Beer, Cereal, Eggs}

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Beer | 0 | 1 | 1 | 1 | 3/4 =75% |
| Cereal | 1 | 1 | 1 | 0 | 3/4 = 75% |
| Eggs | 0 | 1 | 1 | 1 | 3/4 = 75% |

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| {Beer, Cereal ∩Eggs} | 0 | 1 | 1 | 0 | 2/4 = 50% |

# กฎความสัมพันธ์ (Association Rule)

- สินค้าที่มีการซื้อมากกว่า หรือเท่ากับ 50% เรียกว่า frequent itemset

| Items | Support | Size |
|---|---|---|
| {Apple} | 2/4 = 50% | 1 |
| {Beer} | 3/4 =75% | 1 |
| {Cereal} | 3/4 =75% | 1 |
| {Eggs} | 3/4 =75% | 1 |
| {Apple, Cereal} | 2/4 = 50% | 2 |
| {Beer, Cereal} | 2/4 = 50% | 2 |
| {Beer, Eggs} | 3/4 =75% | 2 |
| {Cereal, Eggs} | 2/4 = 50% | 2 |
| {Beer, Cereal, Eggs} | 2/4 = 50% | 3 |

# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - หลักการทำงานของเทคนิค FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree
    - การสร้าง FP-Tree อ่านข้อมูลจากฐานข้อมูลเพียงแค่ 2 ครั้งเท่านั้น
    - หาค่า support ของรูปแบบการซื้อสินค้า (item) ที่มีความยาว 1
    - ในแต่ละ transaction เรียงลำดับ (sort) ตามค่า support
    - สร้าง FP-Tree จากการอ่านข้อมูลในฐานข้อมูลเพียงหนึ่งรอบ
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent itemset) จาก FP-Tree

# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree
    - การสร้าง FP-Tree อ่านข้อมูลจากฐานข้อมูลเพียงแค่ 2 ครั้งเท่านั้น
    - หาค่า support ของรูปแบบการซื้อสินค้า (item) ที่มีความยาว 1

| Transaction ID | Items |
|---|---|
| 1 | Apple, Cereal, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Apple, Beer, Cereal, Eggs |
| 4 | Beer, Eggs |

| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Apple | 1 | 0 | 1 | 0 | 2/4 = 50% |
| Beer | 0 | 1 | 1 | 1 | 3/4 =75% |
| Cereal | 1 | 1 | 1 | 0 | 3/4 = 75% |
| Diapers | 1 | 0 | 0 | 0 | 1/4 = 25% |
| Eggs | 0 | 1 | 1 | 1 | 3/4 = 75% |

# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree
    - การสร้าง FP-Tree อ่านข้อมูลจากฐานข้อมูลเพียงแค่ 2 ครั้งเท่านั้น
    - หาค่า support ของรูปแบบการซื้อสินค้า (item) ที่มีความยาว 1
    - เรียงลำดับตามค่า support ที่คำนวณได้

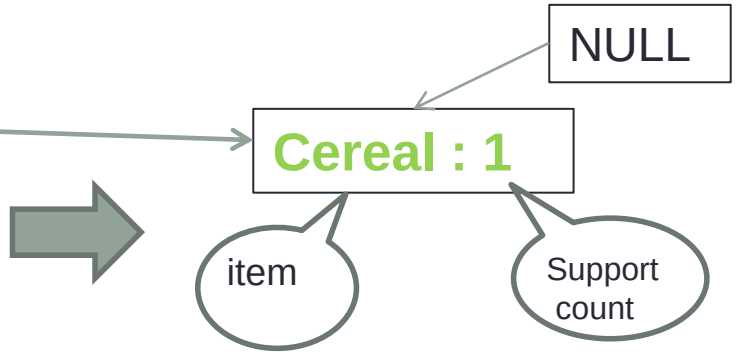
| Items | Transaction Id | | | | Support |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | |
| Apple | 1 | 0 | 1 | 0 | 2/4 = 50% |
| Beer | 0 | 1 | 1 | 1 | 3/4 =75% |
| Cereal | 1 | 1 | 1 | 0 | 3/4 = 75% |
| Diapers | 1 | 0 | 0 | 0 | 1/4 = 25% |
| Eggs | 0 | 1 | 1 | 1 | 3/4 = 75% |

| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |

# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree
    - การสร้าง FP-Tree อ่านข้อมูลจากฐานข้อมูลเพียงแค่ 2 ครั้งเท่านั้น
    - หาค่า support ของรูปแบบการซื้อสินค้า (item) ที่มีความยาว 1
    - เรียงลำดับตามค่า support ที่คำนวณได้
    - สร้าง FP-Tree จากการอ่านข้อมูลในฐานข้อมูลเพียงหนึ่งรอบ

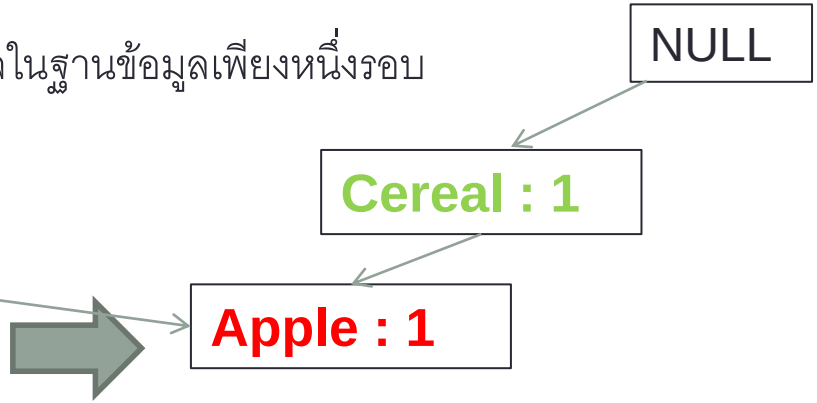
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree
    - การสร้าง FP-Tree อ่านข้อมูลจากฐานข้อมูลเพียงแค่ 2 ครั้งเท่านั้น
    - หาค่า support ของรูปแบบการซื้อสินค้า (item) ที่มีความยาว 1
    - เรียงลำดับตามค่า support
    - สร้าง FP-Tree จากการอ่านข้อมูลในฐานข้อมูลเพียงหนึ่งรอบ

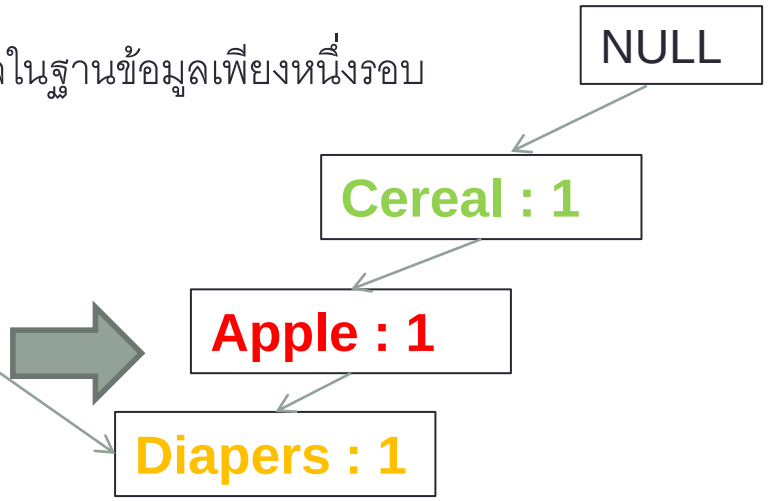
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree
    - การสร้าง FP-Tree อ่านข้อมูลจากฐานข้อมูลเพียงแค่ 2 ครั้งเท่านั้น
    - หาค่า support ของรูปแบบการซื้อสินค้า (item) ที่มีความยาว 1
    - เรียงลำดับตามค่า support
    - สร้าง FP-Tree จากการอ่านข้อมูลในฐานข้อมูลเพียงหนึ่งรอบ

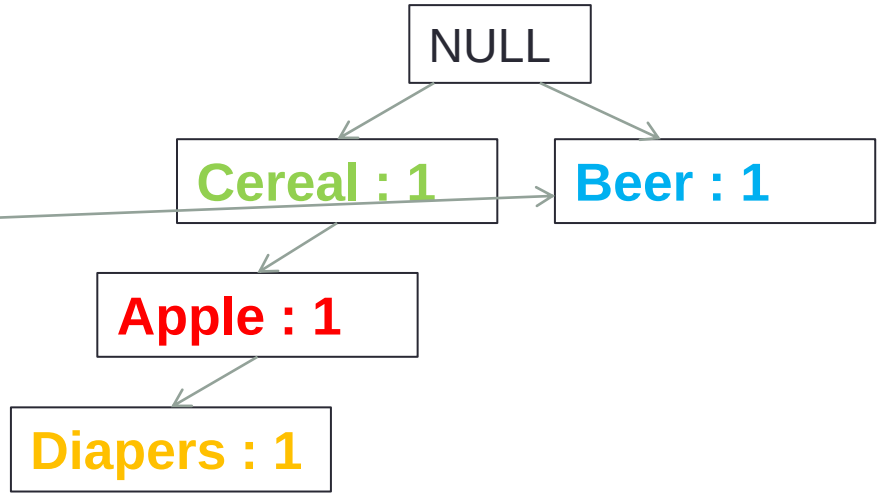
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree

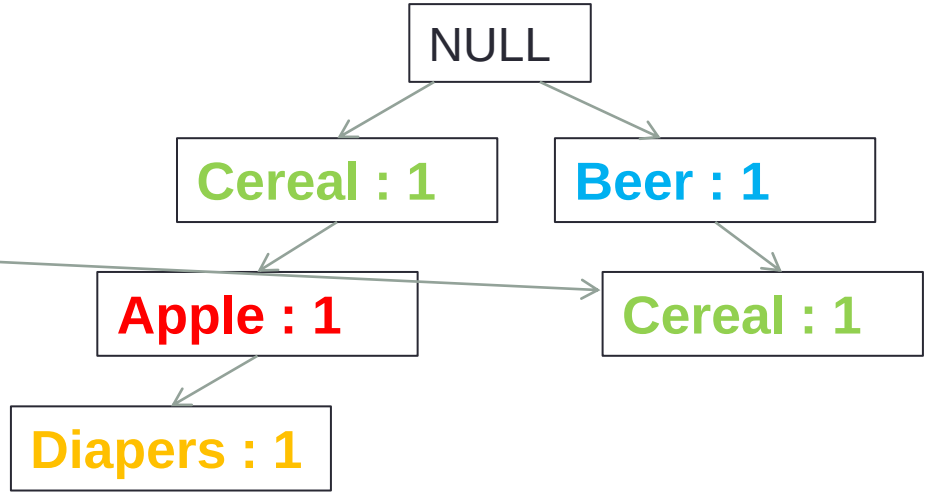
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree

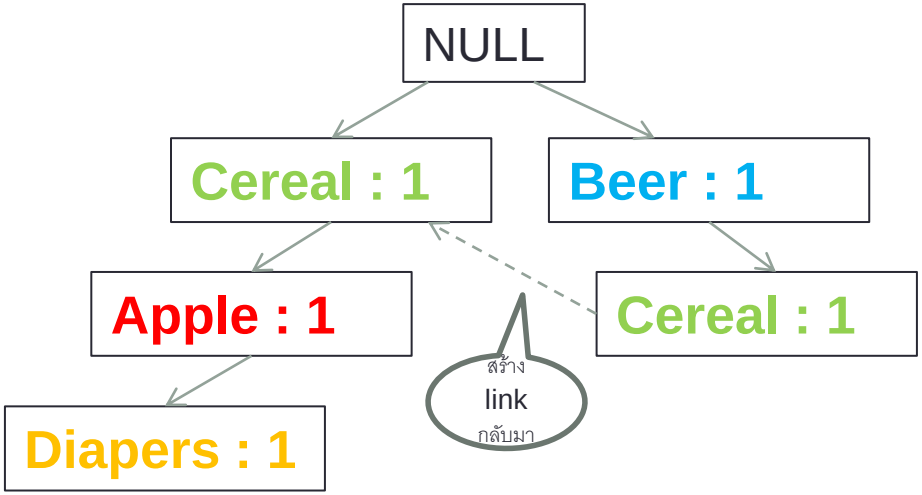
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree

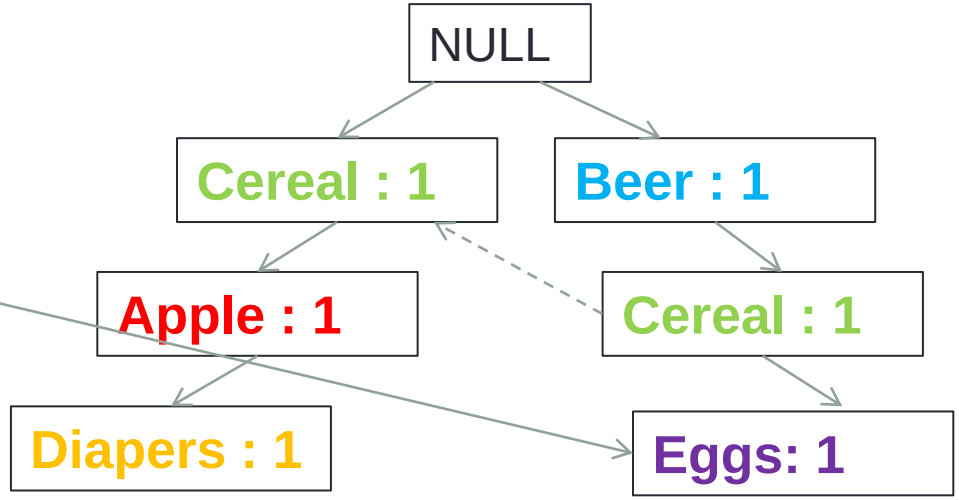
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree

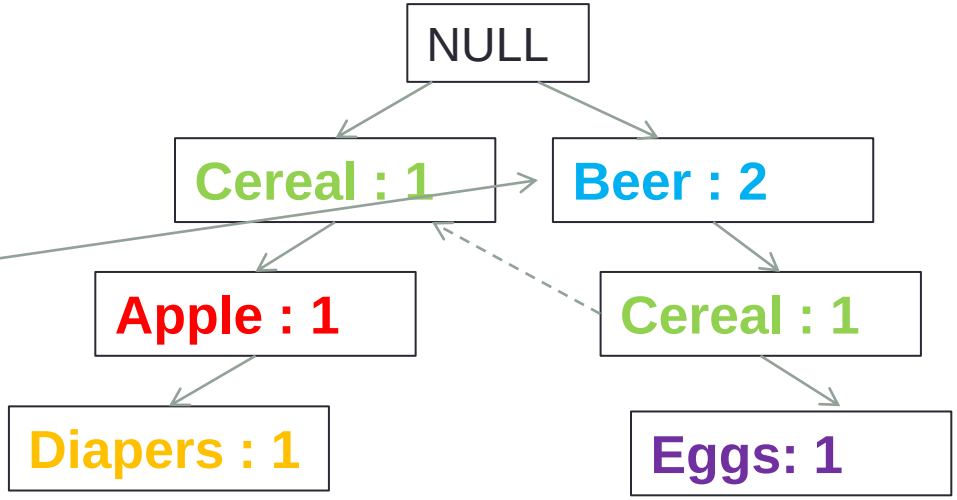
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree

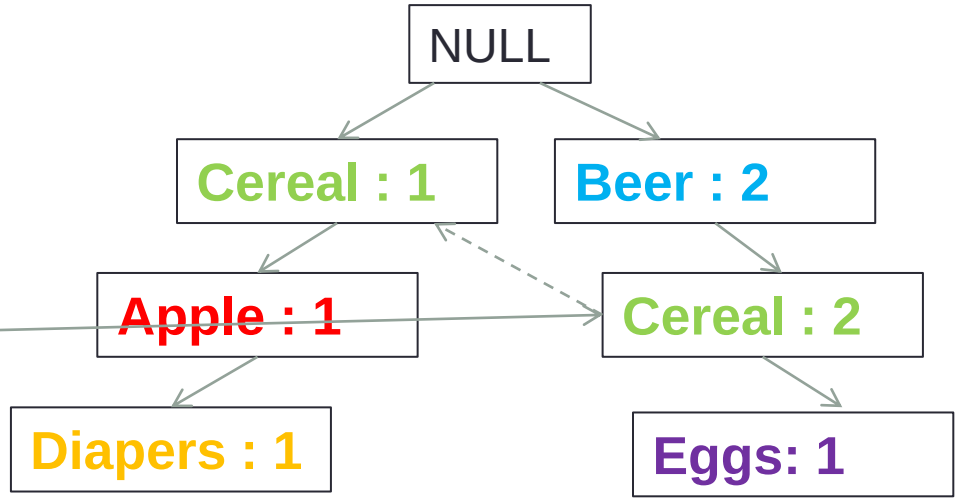
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree

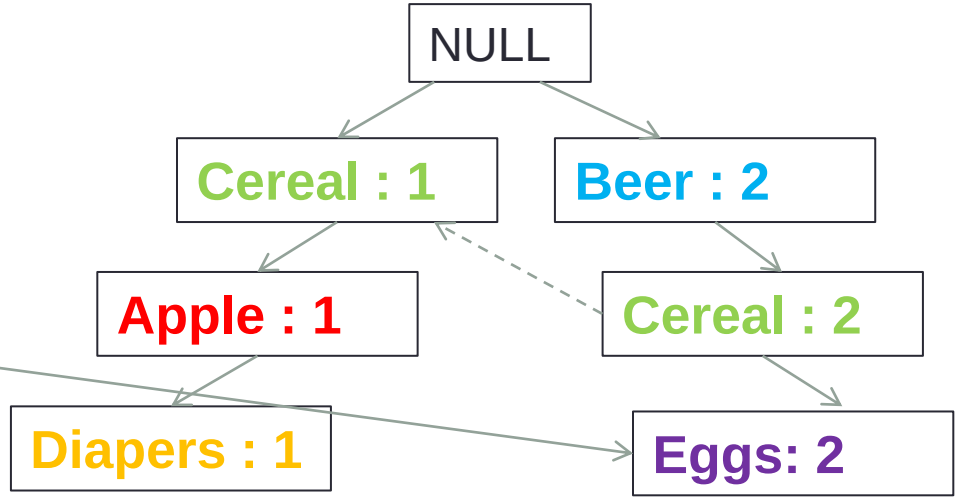
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree

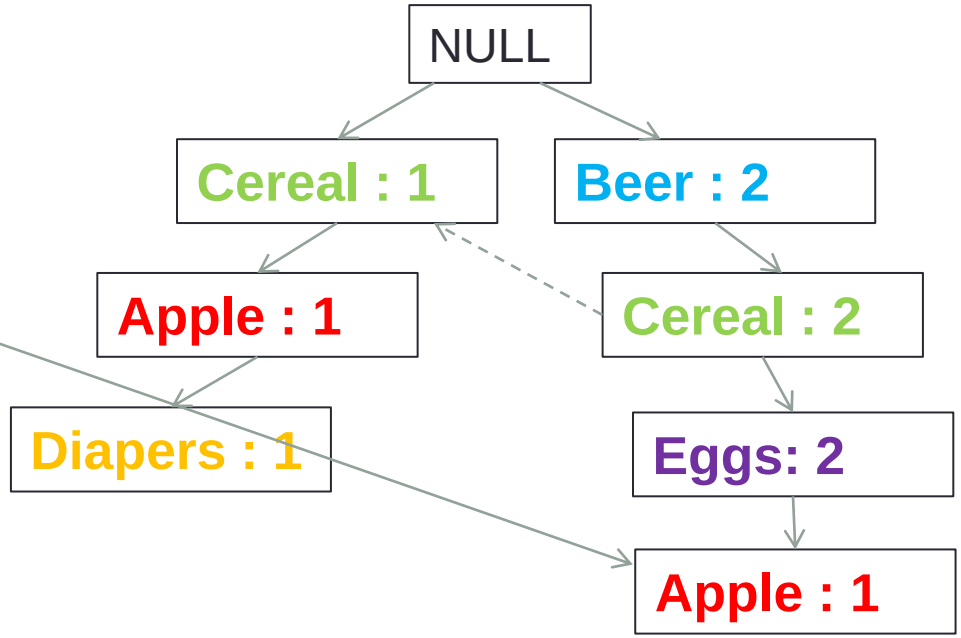
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree

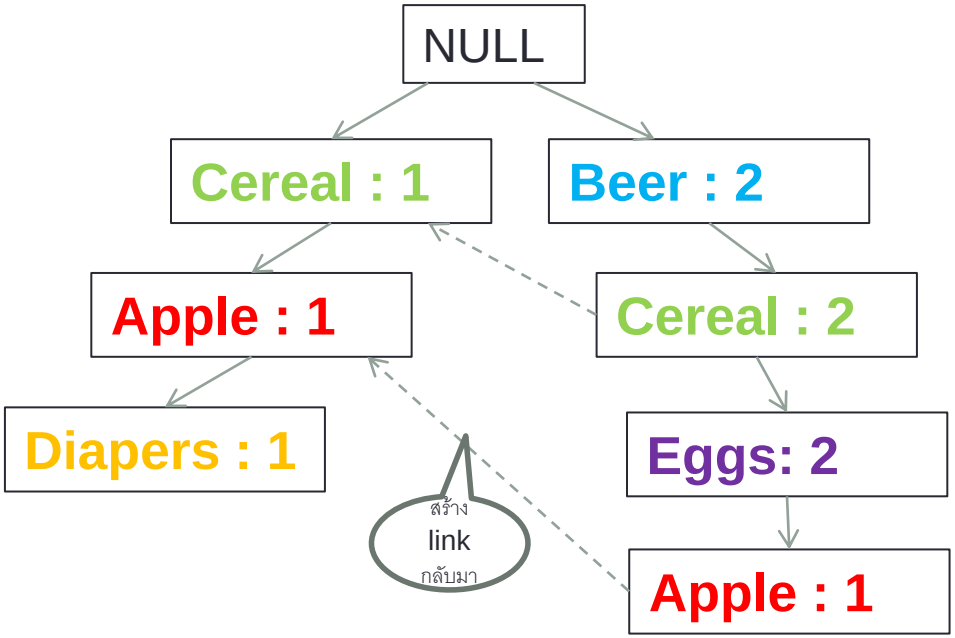
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree

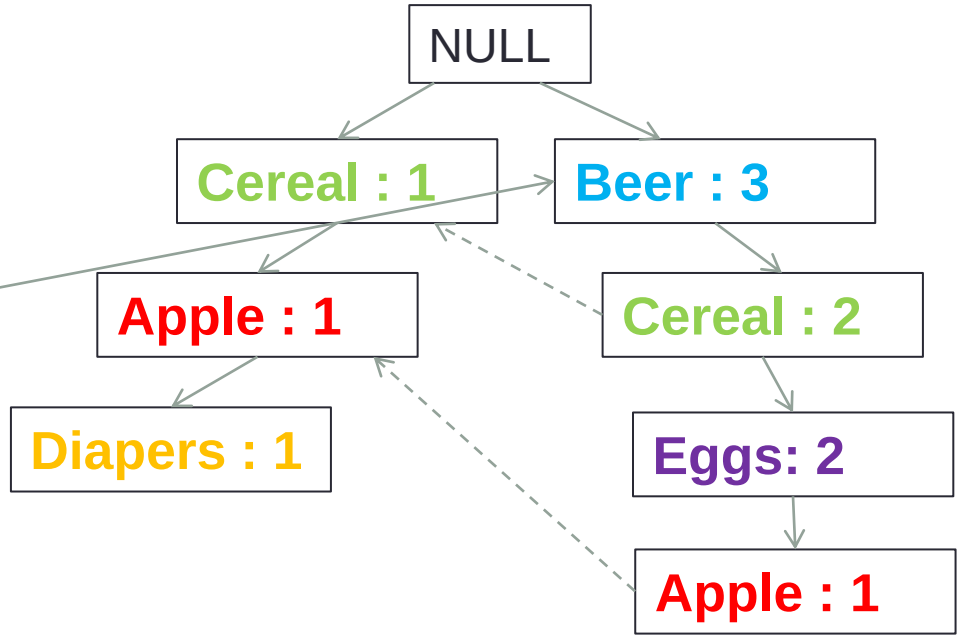
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree

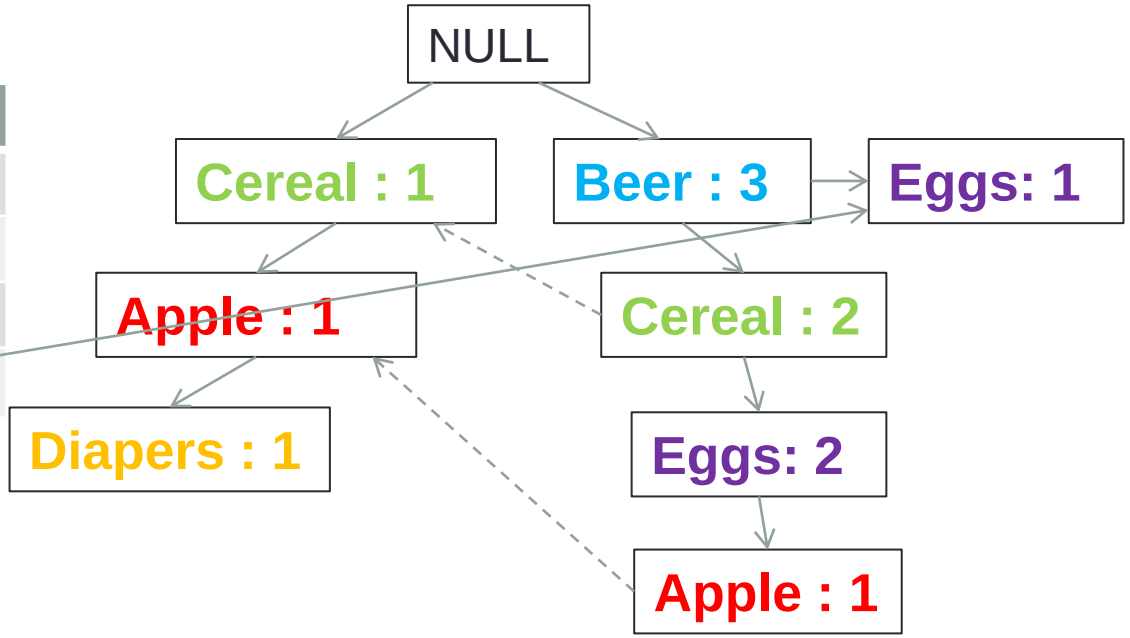
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree

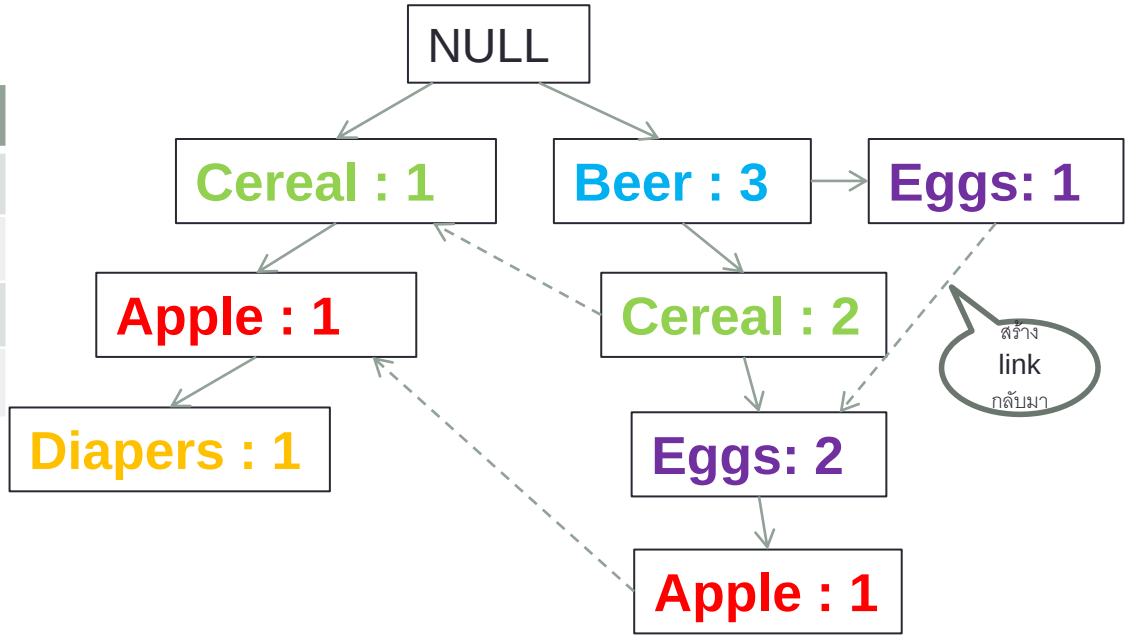
| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



# กฎความสัมพันธ์ (Association Rule)

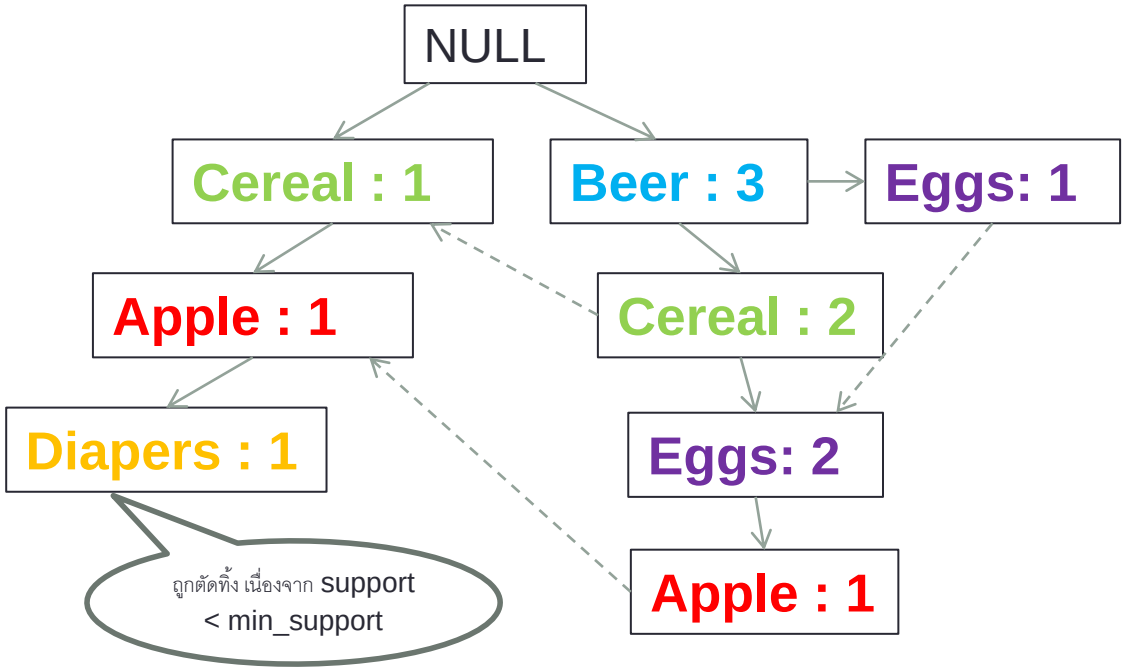
- FP-Growth
  - ขั้นตอนที่ 1 สร้าง compact data structure ที่เรียกว่า FP-Tree

| Transaction ID | Items |
|---|---|
| 1 | Cereal , Apple, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Beer, Cereal,, Eggs ,Apple |
| 4 | Beer, Eggs |



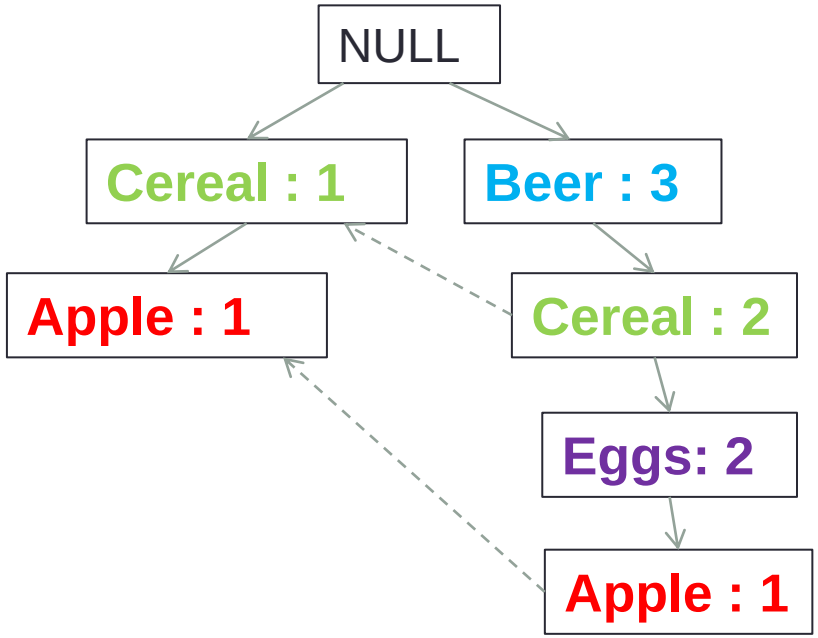
# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent items) จาก FP-Tree
    - กำหนด min_support = 50%



# กฎความสัมพันธ์ (Association Rule)

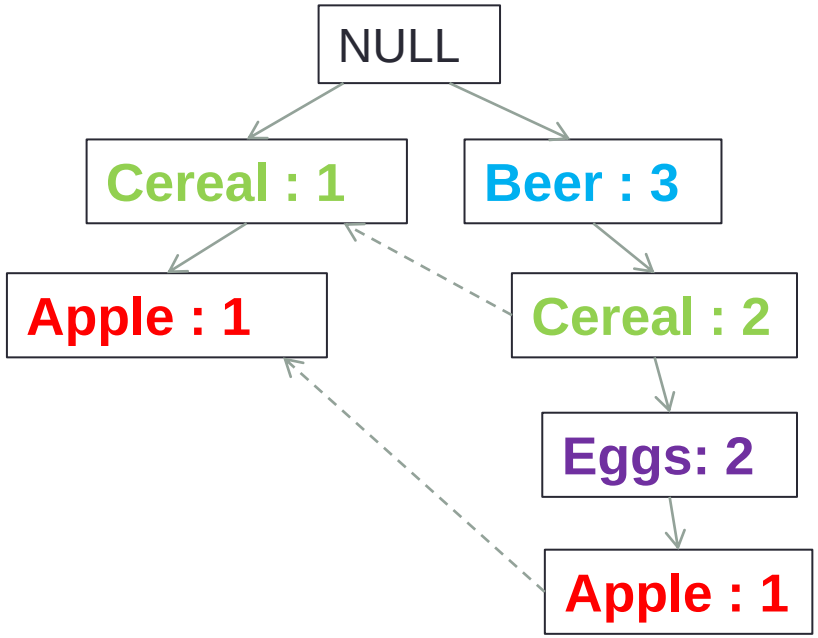
- FP-Growth
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent items) จาก FP-Tree
    - กำหนด min_support = 50%




| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |

# กฎความสัมพันธ์ (Association Rule)

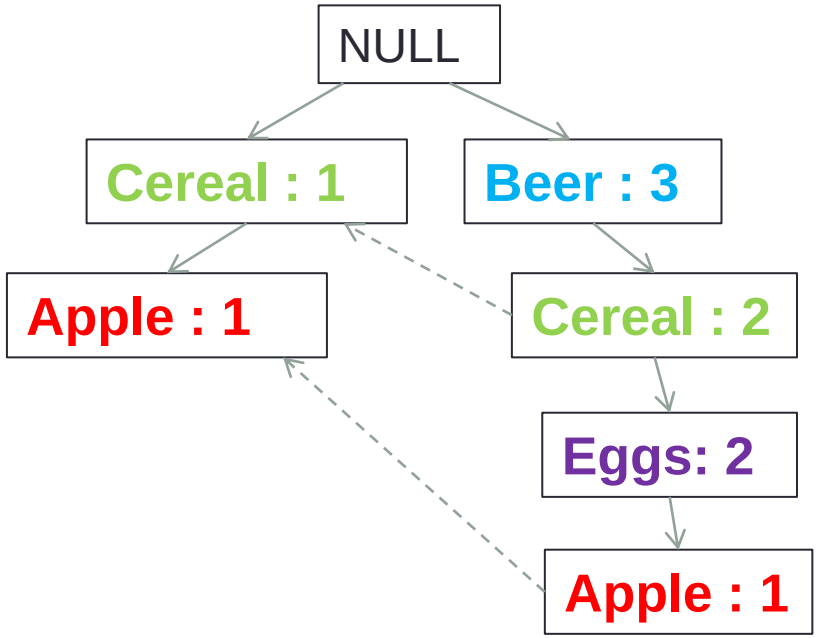
- FP-Growth
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent items) จาก FP-Tree
    - กำหนด min_support = 50%




| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |
| ~~{Apple,Eggs}~~ | ~~1/4 = 25%~~ |

# กฎความสัมพันธ์ (Association Rule)

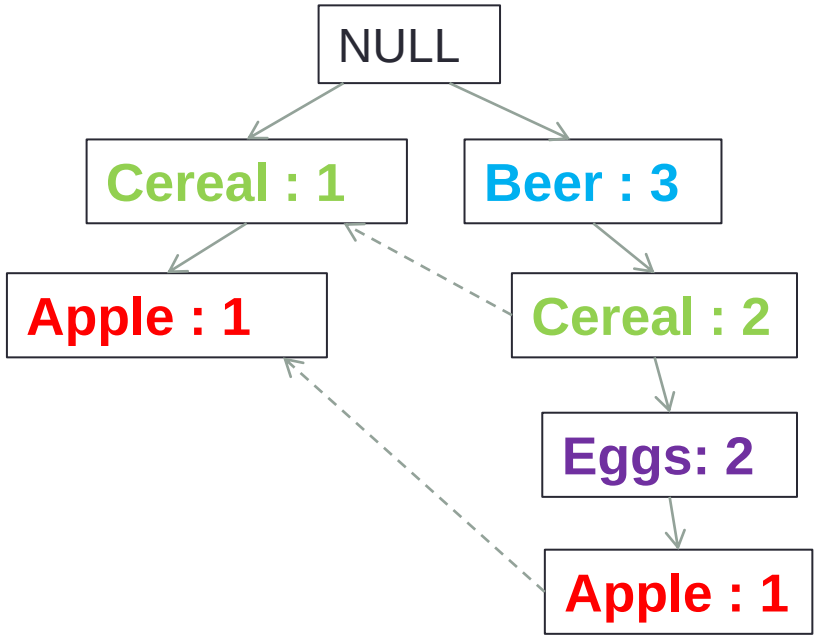
- FP-Growth
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent items) จาก FP-Tree
    - กำหนด min_support = 50%




| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |
| {Apple,Cereal} | 2/4 = 50% |

# กฎความสัมพันธ์ (Association Rule)

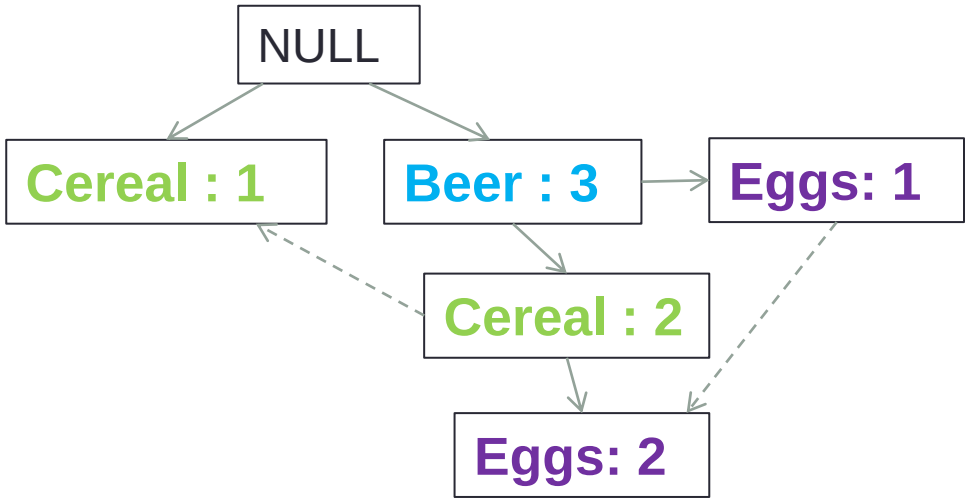
- FP-Growth
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent items) จาก FP-Tree
    - กำหนด min_support = 50%




| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |
| {Apple,Cereal} | 2/4 = 50% |
| ~~{Apple,Beer}~~ | ~~1/4 = 25%~~ |

# กฎความสัมพันธ์ (Association Rule)

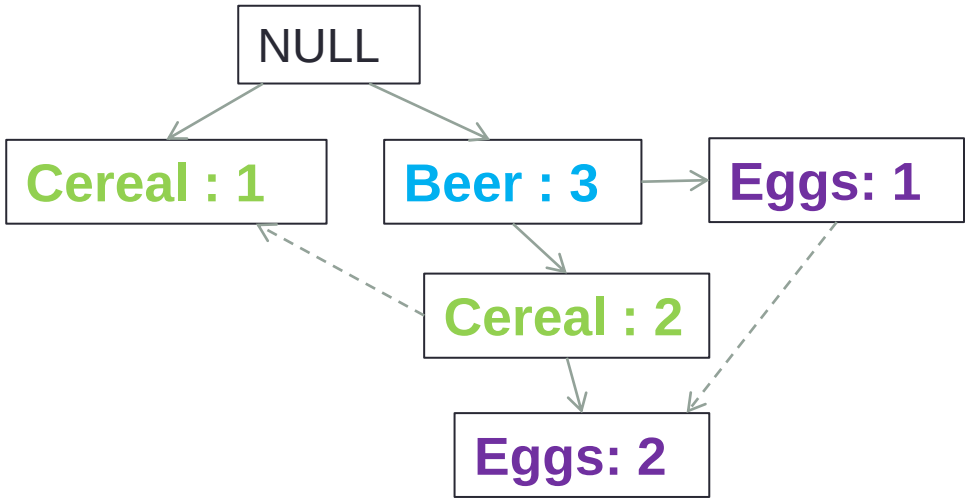
- FP-Growth
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent items) จาก FP-Tree
    - กำหนด min_support = 50%




| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |
| {Apple,Cereal} | 2/4 = 50% |
| {Eggs} | 3/4 = 75% |

# กฎความสัมพันธ์ (Association Rule)

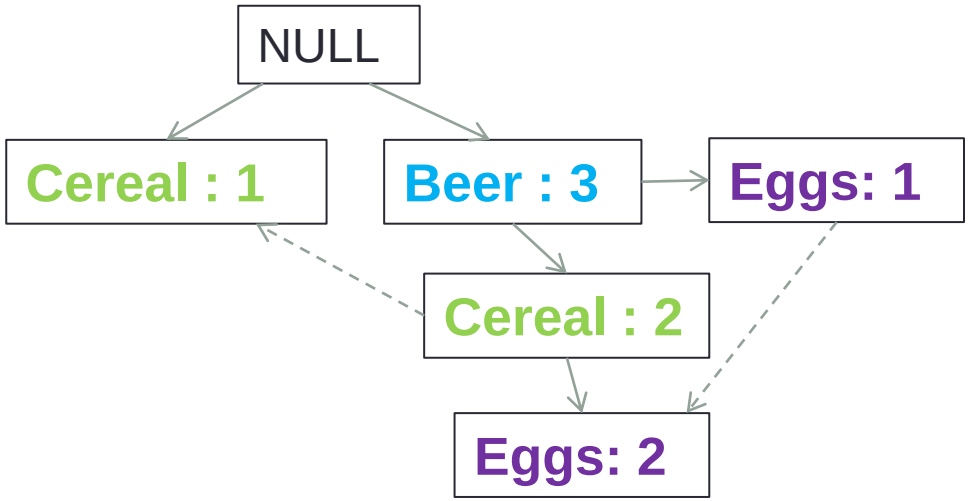
- FP-Growth
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent items) จาก FP-Tree
    - กำหนด min_support = 50%




| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |
| {Apple,Cereal} | 2/4 = 50% |
| {Eggs} | 3/4 = 75% |
| {Cereal, Eggs} | 2/4 = 50% |

# กฎความสัมพันธ์ (Association Rule)

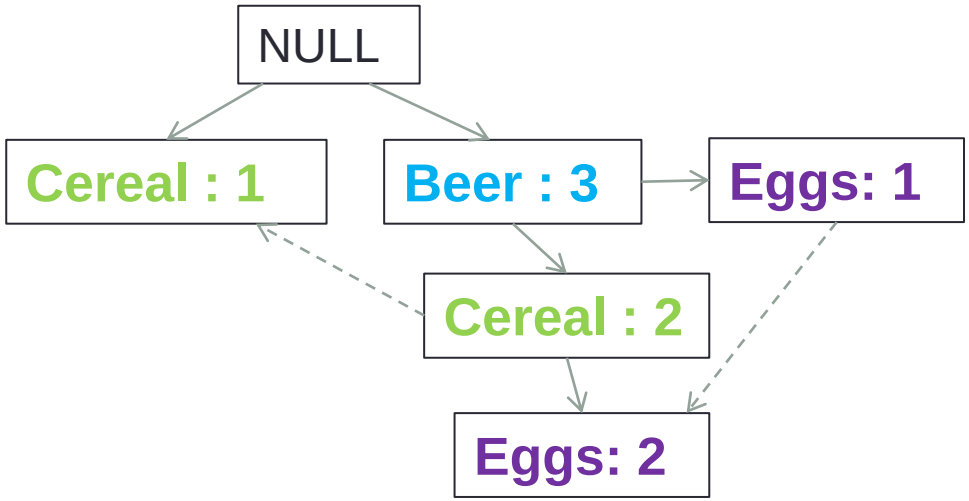
- FP-Growth
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent items) จาก FP-Tree
    - กำหนด min_support = 50%




| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |
| {Apple,Cereal} | 2/4 = 50% |
| {Eggs} | 3/4 = 75% |
| {Cereal, Eggs} | 2/4 = 50% |
| {Beer,Eggs} | 3/4 = 75% |

# กฎความสัมพันธ์ (Association Rule)

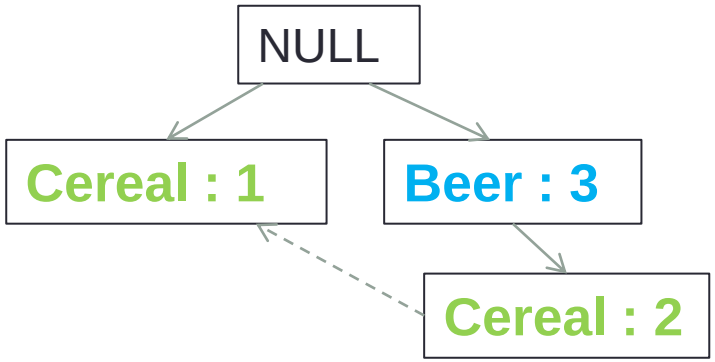
- FP-Growth
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent items) จาก FP-Tree
    - กำหนด min_support = 50%




| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |
| {Apple,Cereal} | 2/4 = 50% |
| {Eggs} | 3/4 = 75% |
| {Cereal, Eggs} | 2/4 = 50% |
| {Beer,Eggs} | 3/4 = 75% |
| {Beer, Cereal, Eggs} | 2/4 = 50% |

# กฎความสัมพันธ์ (Association Rule)

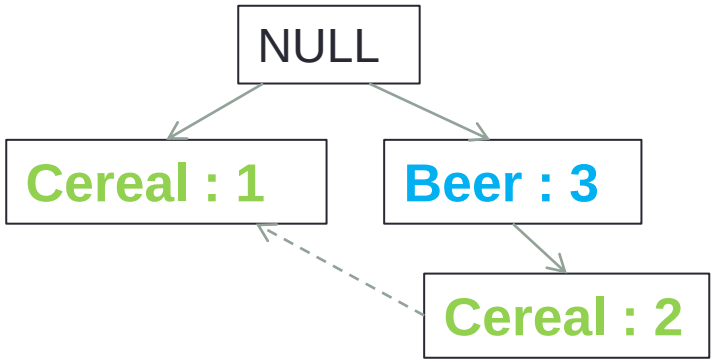
- FP-Growth
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent items) จาก FP-Tree
    - กำหนด min_support = 50%




| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |
| {Apple,Cereal} | 2/4 = 50% |
| {Eggs} | 3/4 = 75% |
| {Cereal, Eggs} | 2/4 = 50% |
| {Beer,Eggs} | 3/4 = 75% |
| {Beer, Cereal, Eggs} | 2/4 = 50% |
| {Cereal} | 3/4 = 75% |

# กฎความสัมพันธ์ (Association Rule)

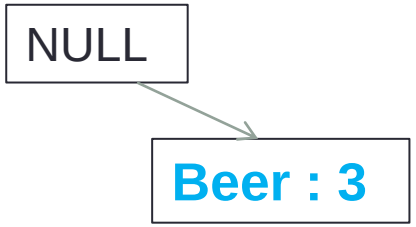
- FP-Growth
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent items) จาก FP-Tree
    - กำหนด min_support = 50%




| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |
| {Apple,Cereal} | 2/4 = 50% |
| {Eggs} | 3/4 = 75% |
| {Cereal, Eggs} | 2/4 = 50% |
| {Beer,Eggs} | 3/4 = 75% |
| {Beer, Cereal, Eggs} | 2/4 = 50% |
| {Cereal} | 3/4 = 75% |
| {Beer,Cereal} | 2/4 = 50% |

# กฎความสัมพันธ์ (Association Rule)

- FP-Growth
  - ขั้นตอนที่ 2 สร้างรูปแบบของสินค้าที่ซื้อบ่อย ๆ (frequent items) จาก FP-Tree
    - กำหนด min_support = 50%




| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |
| {Apple,Cereal} | 2/4 = 50% |
| {Eggs} | 3/4 = 75% |
| {Cereal, Eggs} | 2/4 = 50% |
| {Beer,Eggs} | 3/4 = 75% |
| {Beer, Cereal, Eggs} | 2/4 = 50% |
| {Cereal} | 3/4 = 75% |
| {Beer,Cereal} | 2/4 = 50% |
| {Beer} | 3/4 = 75% |

# กฎความสัมพันธ์ (Association Rule)

- Apriori & FP-Growth

| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |
| {Beer} | 3/4 = 75% |
| {Cereal} | 3/4 = 75% |
| {Eggs} | 3/4 = 75% |
| {Apple,Cereal} | 2/4 = 50% |
| {Beer,Cereal} | 2/4 = 50% |
| {Beer,Eggs} | 3/4 = 75% |
| {Cereal, Eggs} | 2/4 = 50% |
| {Beer, Cereal, Eggs} | 2/4 = 50% |

ผลลัพท์จาก **Apriori**

| Frequent Itemset | Support |
|---|---|
| {Apple} | 2/4 = 50% |
| {Apple,Cereal} | 2/4 = 50% |
| {Eggs} | 3/4 = 75% |
| {Cereal, Eggs} | 2/4 = 50% |
| {Beer,Eggs} | 3/4 = 75% |
| {Beer, Cereal, Eggs} | 2/4 = 50% |
| {Cereal} | 3/4 = 75% |
| {Beer,Cereal} | 2/4 = 50% |
| {Beer} | 3/4 = 75% |

ผลลัพท์จาก **FP-Growth**

# กฎความสัมพันธ์ (Association Rule)

- กฎความสัมพันธ์ (Association Rule)
  - สร้างจากสินค้าที่ลูกค้าซื้อบ่อย ๆ
  - รูปแบบของกฎความสัมพันธ์ คือ LHS => RHS
    - LHS คือ Left Hand Side สินค้าที่ซื้อพร้อมกันบ่อย ๆ ด้านซ้ายของกฎ
    - RHS คือ Right Hand Side สินค้าที่ซื้อพร้อมกันบ่อย ๆ ด้านขวาของกฎ

| Items | Support | Size |
|---|---|---|
| {Apple, Cereal} | 2/4 = 50% | 2 |
| {Beer, Cereal} | 2/4 = 50% | 2 |
| {Beer, Eggs} | 3/4 =75% | 2 |
| {Cereal, Eggs} | 2/4 = 50% | 2 |
| {Beer, Cereal, Eggs} | 2/4 = 50% | 3 |

Apple => Cereal
Beer => Cereal
Beer => Eggs
Cereal => Apple
Cereal => Eggs
Cereal,Eggs => Eeer
Eggs => Beer

# กฎความสัมพันธ์ (Association Rule)

- ตัววัดประสิทธิภาพของกฎความสัมพันธ์
- Confidence
  - แสดงความเชื่อมั่นของกฎความสัมพันธ์ที่เมื่อรูปแบบ LHS เกิดขึ้นแล้วรูปแบบ RHS จะเกิดขึ้นด้วยเป็นจำนวนกี่เปอร์เซ็นต

$$\text{Confidence}(LHS \Rightarrow RHS) = \frac{support(LHS, RHS)}{support(LHS)}$$

$$\text{Confidence}(Apple \Rightarrow Cereal) = \frac{support(Apple, Cereal)}{support(Apple)}$$

$$= \frac{2/4}{2/4}$$

$$= 100\%$$

# กฎความสัมพันธ์ (Association Rule)

- ตัววัดประสิทธิภาพของกฎความสัมพันธ์
- Lift
  - ค่าที่บ่งบอกว่าการเกิดรูปแบบ LHS และ RHS มีความสัมพันธ์กันแค่ไหน โดยถ้าค่า Lift ใกล้ค่า 1 แสดงว่ารูปแบบ LHS และ RHS ไม่ขึ้นต่อกัน (independent)

$$\text{Lift}(LHS \Rightarrow RHS) = \frac{support(LHS, RHS)}{support(LHS) x\ support(RHS)}$$

$$\text{Lift}(Apple \Rightarrow Cereal) = \frac{support(Apple, Cereal)}{support(Apple) x\ support(Cereal)}$$

$$= \frac{2/4}{2/4 x\ 3/4}$$

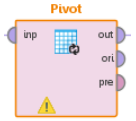
$$= 1.33$$

# กฎความสัมพันธ์ (Association Rule)

- Applications
  - ใช้ในการเพิ่มยอดขายโดยการแนะนำสินค้าที่ลูกค้ามักจะซื้อพร้อมกันบ่อย ๆ (Cross-Selling)
  - ใช้ในการจัดสินค้าในร้าน
    - วางสินค้าที่ลูกค้ามักจะซื้อพร้อมกันไว้ใกล้ ๆกัน
    - วางสินค้าที่ลูกค้ามักจะซื้อพร้อมกันไว้ไกล ๆ กัน
  - การจัดวางสินค้าใน Catalog ต่าง ๆ

# Example

- ขั้นตอนการหากฎความสัมพันธ์ (Association Rule)
  - โอเปอเรเตอร์ที่เกี่ยวข้อง

| โอเปอเรเตอร์ | | คำอธิบาย |
|---|---|---|
| | Read CSV | ใช้สำหรับนำเข้าข้อมูล CSV |
|  | Pivot | ใช้สำหรับแปลงข้อมูลให้เป็นตาราง Transaction |
| | Numerical to Binominal | ใช้สำหรับแปลงข้อมูลให้เป็นตาราง Transaction จากค่าตัวเลขให้เป็น Binominal |
| | FP-Growth | ใช้สำหรับหา frequent itemset เป็นรูปแบบของข้อมูลที่เกิดขึ้นร่วมกันบ่อย ๆ |
| | Create Association Rules | ใช้สำหรับการหากฎความสัมพันธ์ |

# Example

- ขั้นตอนที่ 1 นำเข้าข้อมูล
  - นำเข้า Data set ชื่อ POS-Association Rule.csv

- ขั้นตอนที่ 2 กำหนดหน้าที่ให้กับ attribute ID เป็นประเภท ID

| | id<br>*integer*<br>*id* | name<br>*polynominal* | value<br>*integer* |
|---|---|---|---|
| 1 | 1 | Apple | 1 |
| 2 | 1 | Beer | 0 |
| 3 | 1 | Cereal | 1 |
| 4 | 1 | Diapers | 1 |
| 5 | 1 | Eggs | 0 |
| 6 | 2 | Beer | 1 |
| 7 | 2 | Cereal | 1 |
| 8 | 2 | Eggs | 1 |
| 9 | 3 | Apple | 1 |
| 10 | 3 | Beer | 1 |
| 11 | 3 | Cereal | 1 |
| 12 | 3 | Eggs | 1 |

PREVIE...

# Example

- เมื่อนำเข้า Data set สำเร็จ ผลลัพธ์ของข้อมูลที่นำเข้าจะพบว่า ข้อมูลที่ได้จาก pos (point of sale) ซึ่งมักจะเก็บไว้ในตาราง order detail ซึ่งประกอบด้วย
  - Id แสดงหมายเลขสมาชิกของลูกค้า
  - name แสดงชื่อสินค้า
  - value แสดงจำนวนจำนวนที่ซื้อ

| Row No. | id | name | value |
|---|---|---|---|
| 1 | 1 | Apple | 1 |
| 2 | 1 | Beer | 0 |
| 3 | 1 | Cereal | 1 |
| 4 | 1 | Diapers | 1 |
| 5 | 1 | Eggs | 0 |
| 6 | 2 | Beer | 1 |
| 7 | 2 | Cereal | 1 |
| 8 | 2 | Eggs | 1 |
| 9 | 3 | Apple | 1 |
| 10 | 3 | Beer | 1 |
| 11 | 3 | Cereal | 1 |
| 12 | 3 | Eggs | 1 |
| 13 | 4 | Beer | 1 |
| 14 | 4 | Eggs | 1 |

ตาราง Exampleset

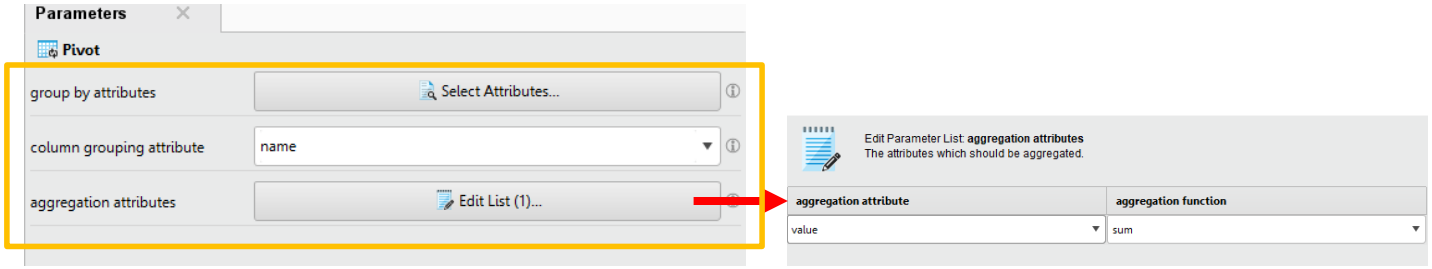
# Example

- จากผลลัพธ์ ข้อมูลที่เก็บอยู่ในลักษณะนี้ ไม่เหมาะสมที่จะนำไปหาความสัมพันธ์ของการซื้อขายต่อ ดังนั้นจึงต้องทำการแปลงข้อมูลในตาราง order detail ให้เป็นตาราง transaction เสียก่อน

| ID | Name |
|---|---|
| 1 | Apple, Cereal, Diapers |
| 2 | Beer, Cereal, Eggs |
| 3 | Apple, Beer, Cereal, Eggs |
| 4 | Beer, Eggs |

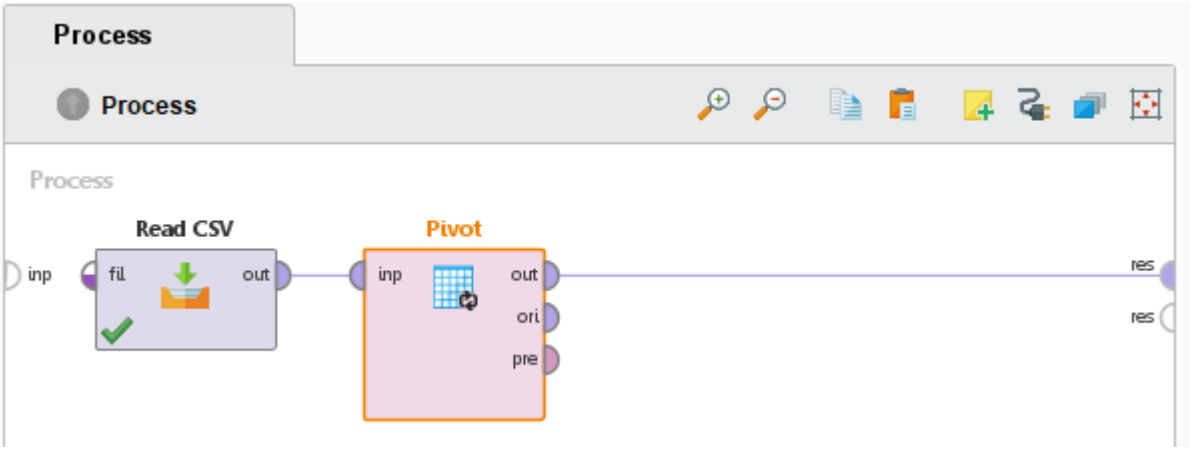
# Example

- ขั้นตอนที่ 3 แปลงข้อมูลจากตาราง order detail ให้เป็นตาราง transaction เพื่อพร้อมสำหรับสร้างกฎความสัมพันธ์
  - 1. แปลงข้อมูลจากตาราง order detail ให้เป็นตาราง Transaction โดยเรียกใช้ Pivot Operator
  - 2. ตั้งค่า parameter ให้กับ Pivot Operator โดยการกำหนด
    - กำหนด group by attributes : id (เลือก attribute ที่ต้องการรวบรวมข้อมูล)
    - กำหนด column grouping attributes : name (เลือก attribute ที่ต้องการแสดงข้อมูลหลังจากการรวมกลุ่ม)
    - กำหนด aggregation attributes : value และ aggregation function : sum



# Example

- 3. ลากเส้นเชื่อมระหว่าง Read CSV Operator กับ Pivot Operator จากนั้กดปุ่ม RUN



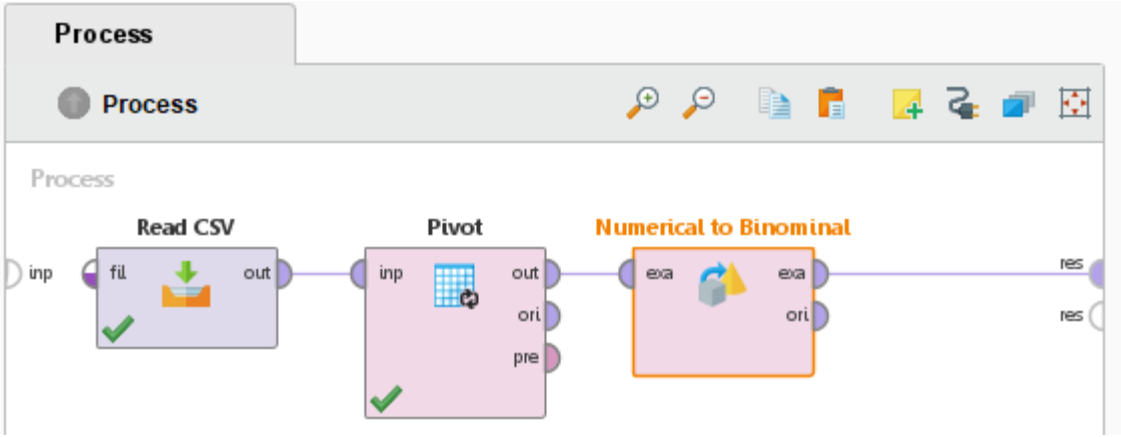
# Example

- 4. ผลลัพธ์หลังจากการแปลงข้อมูล

| Row No. | id | sum(value)_... | sum(value)_... | sum(value)_... | sum(value)_... | sum(value)_... |
|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 1 | 1 | 0 |
| 2 | 2 | ? | 1 | 1 | ? | 1 |
| 3 | 3 | 1 | 1 | 1 | ? | 1 |
| 4 | 4 | ? | 1 | ? | ? | 1 |

# Example

- ขั้นตอนที่ 4 แปลงข้อมูลตาราง Transaction จากค่าตัวเลขให้เป็น Binominal
  - 1. แปลงข้อมูล transaction จากค่าตัวเลขให้เป็นค่า Binominal โดยเรียกใช้ Numerical to Binominal Operator
  - 2. ลากเส้นเชื่อมระหว่าง Pivot Operator กับ Numerical to Binominal Operator จากนั้นกดปุ่ม Run



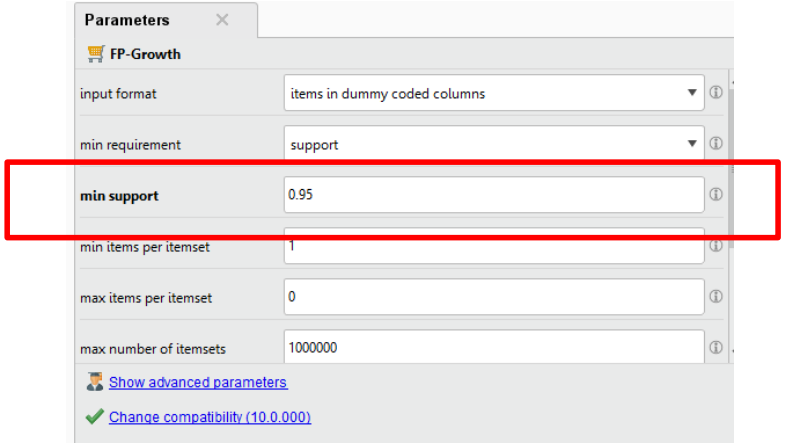
# Example

- 3. ผลลัพธ์หลังจากการแปลงข้อมูลในตาราง Transaction

| Row No. | id | sum(value)_... | sum(value)_... | sum(value)_... | sum(value)_... | sum(value)_... |
|---|---|---|---|---|---|---|
| 1 | 1 | true | false | true | true | false |
| 2 | 2 | ? | true | true | ? | true |
| 3 | 3 | true | true | true | ? | true |
| 4 | 4 | ? | true | ? | ? | true |

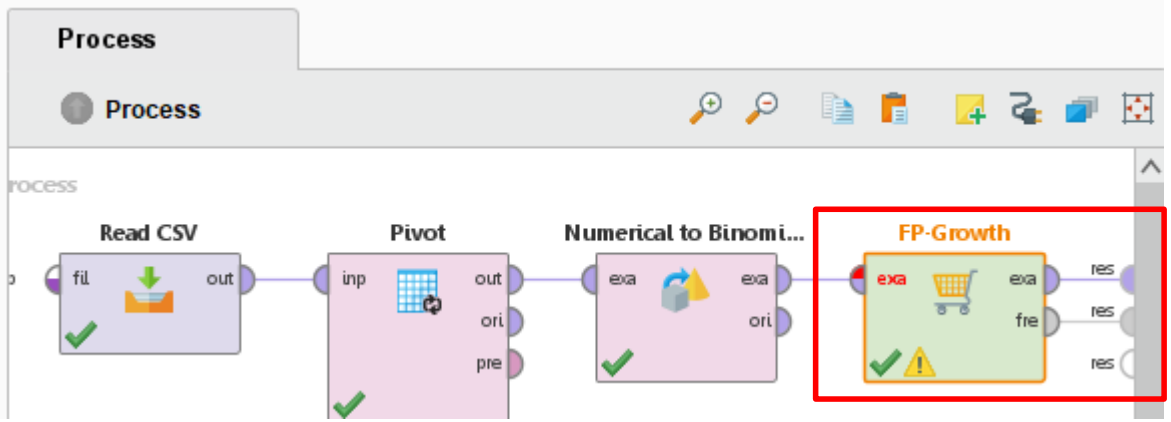
# Example

- ขั้นตอนที่ 5 หา frequent itemset
  - 1. การหา frequent itemset เป็นรูปแบบของข้อมูลที่เกิดขึ้นร่วมกันบ่อย ๆ ในฐานข้อมูล หรือมากกว่า minimum support ที่ผู้ใช้กำหนด ด้วยขั้นตอนวิธี FP-Growth โดยเรียกว่า FP-Growth Operators
  - 2. ตั้งค่า parameter ให้กับ FP-Growth Operators โดยกำหนด
    - กำหนดค่า min support : 0.5 (ค่า support ที่น้อยที่สุดที่ต้องการ ในที่นี้กำหนด 0.5 คือความถี่ในการซื้อสินค้าต่ำสุดที่เราสนใจที่มีค่า support >= 50%)



# Example

- 3. ลากเส้นเชื่อมระหว่าง Numerical to Binominal Operator กับ FP-Growth Operator และ Pot Output ทั้ง 2 ของ FP-Growth Operator ต้องเชื่อมต่อกับ result port ทั้งสองด้วยเช่นเดียวกัน จากนั้นกดปุ่ม run



# Example

- 4. ผลลัพธ์ของการหา frequent itemset ด้วยขั้นตอนวิธี FP-Growth
  - สินค้าที่ลูกค้าซื้อทั้งหมด 5 ประเภท คือ Apple, Beer, Cereal, Diapers และ Eggs
  - การหา frequent itemset โดยกำหนดค่า minimum support ไว้ที่ 50% จาก output ถ้าค่า support ต่ำกว่าค่า minimum support ที่กำหนดไว้ จะถูกตัดทิ้ง

No. of Sets: 19

Total Max. Size: 4

Min. Size: 1

Max. Size: 4

Contains Item:

Update View

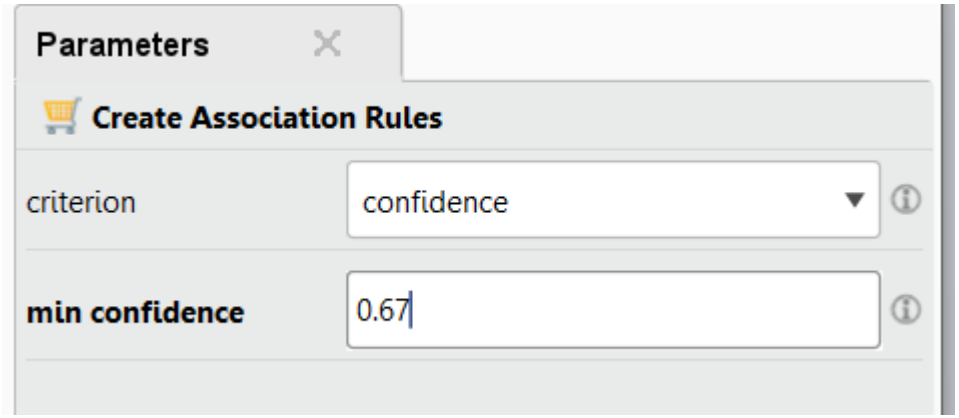
| Size | Support | Item 1 | Item 2 | Item 3 | I |
|---|---|---|---|---|---|
| 1 | 0.750 | sum(value)_Beer | | | |
| 1 | 0.750 | sum(value)_Cereal | | | |
| 1 | 0.750 | sum(value)_Eggs | | | |
| 2 | 0.750 | sum(value)_Beer | sum(value)_Eggs | | |
| 1 | 0.500 | sum(value)_Apple | | | |
| 2 | 0.500 | sum(value)_Beer | sum(value)_Cereal | | |
| 2 | 0.500 | sum(value)_Cereal | sum(value)_Eggs | | |
| 2 | 0.500 | sum(value)_Cereal | sum(value)_Apple | | |
| 3 | 0.500 | sum(value)_Beer | sum(value)_Cereal | sum(value)_Eggs | |
| 1 | 0.250 | sum(value)_Diapers | | | |
| 2 | 0.250 | sum(value)_Beer | sum(value)_Apple | | |
| 2 | 0.250 | sum(value)_Cereal | sum(value)_Diapers | | |

# Example

- ขั้นตอนที่ 6 กฎความสัมพันธ์
  - 1. สร้างกฎความสัมพันธ์ โดยเรียกใช้ Create Association Rule Operator
  - 2. ตั้งค่า parameter ให้กับ Create Association Rule Operator โดยการกำหนด
  - กำหนดค่า criterion : confidence (เกณฑ์ที่ใช้ในการสร้างกฎความสัมพันธ์)
  - เกณฑ์ที่นิยมได้แก่
    - ค่า Confidence ซึ่งถ้า Confidence กฎนั้นสูงน่าสนใจกว่า
    - ค่า Lift ซึ่งถ้า เกิน 1 แสดงว่าโอกาสจะเกิดขึ้นพร้อมกันมากกว่าแยก ถ้าน้อยกว่า 1 แสดงว่า ตัวล่างมากกว่าตัวบน แสดงว่าต้องแยกมากกว่า

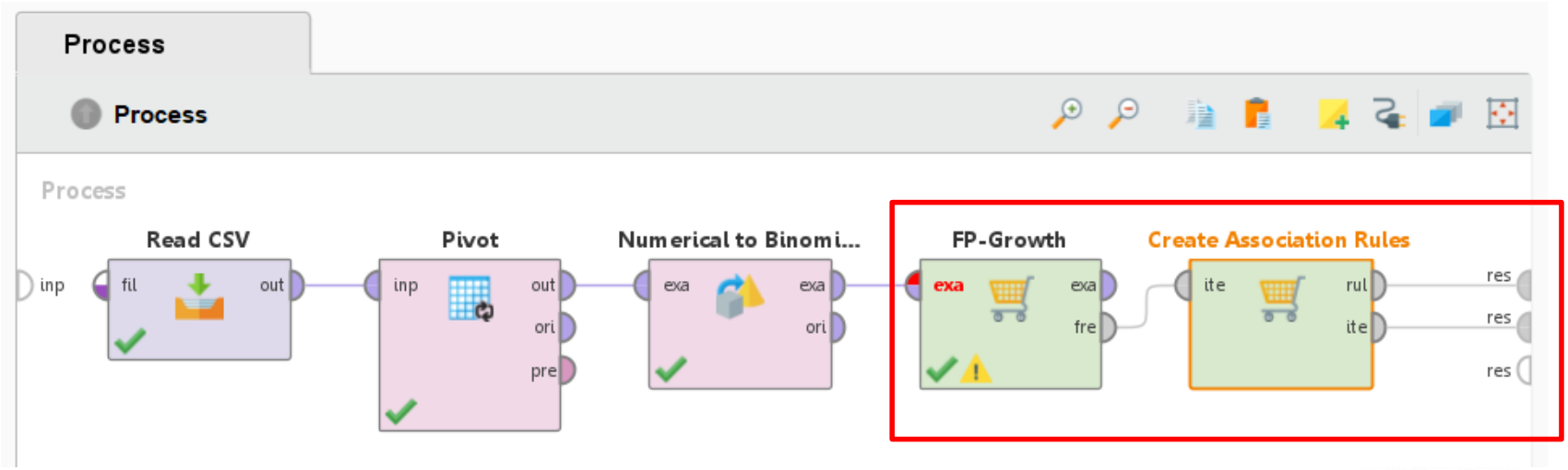
# Example

- กำหนดค่า min confidence 0.67 (ค่า confidence น้อยที่สุดที่ต้องการ คือ 0.67)



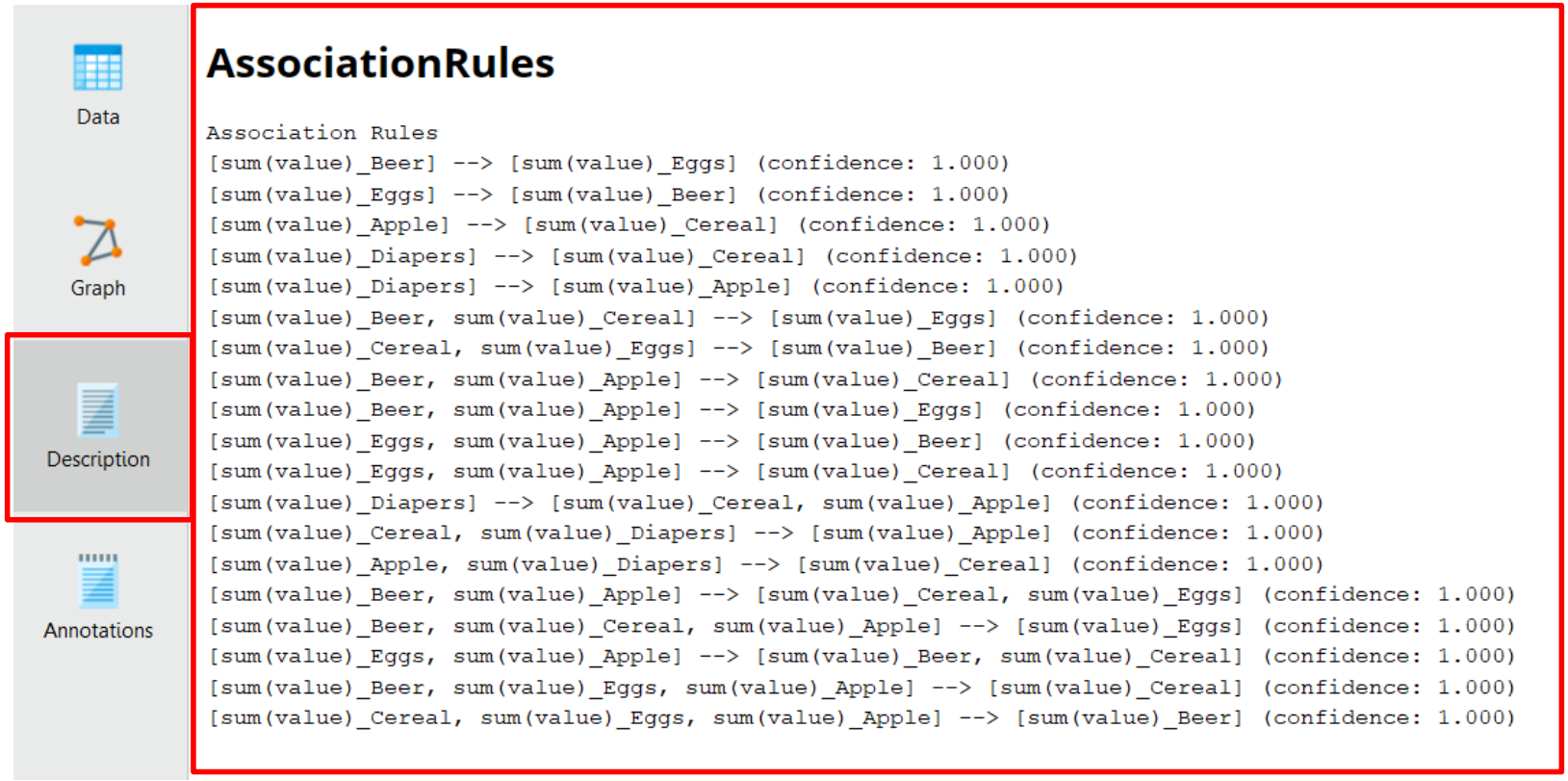
# Example

- 3. ลากเส้นเชื่อมระหว่า port output (frequent itemset) ของ FP-Growth Operator กับ port input ของ create association rule operator และ port output ทั้ง 2 ของ create association rule operator ต้องเชื่อมต่อกับ result port ทั้งสองข้างด้วยเช่นเดียวกัน จากนั้นกดปุ่ม run



# Example

- .ผลลัพธ์ของการสร้างกฎความสัมพันธ์



# Example

- 5. การอ่านผลกฎความสัมพันธ์ทั้งหมดที่สร้างได้พร้อมทั้งค่า confidence และค่า lift

| Rule NO. | Association Rule | Confidence |
|---|---|---|
| 1 | Eggs = => Beer | 100% |
| 2 | Beer = => Eggs | 100% |
| 3 | Apple = => Cereal | 100% |
| 4 | Eggs, Cereal = => Beer | 100% |
| 5 | Cereal, Beer = => Eggs | 100% |
| .. | | |
| 19 | Cereal, Beer, Apple = => Eggs | 100 % |

# Example

- การอธิบายผลลัพธ์ของกฎความสัมพันธ์
  - กฎความสัมพันธ์ที่ 1 สามารถอธิบายได้ว่า เมื่อซื้อ ไข่ จะต้องซื้อเบียร์ พร้อมกันด้วย ที่ความเชื่อมัน 100%
  - กฎความสัมพันธ์ที่ 2 สามารถอธิบายได้ว่า เมื่อซื้อ เบียร์ จะต้องการซื้อ ไข่ พร้อมกันด้วย ที่ความเชื่อมั่น 100%
  - กฎความสัมพันธ์ที่ 3 สามารถอธิบายได้ว่า เมื่อซื้อ แอบเปิ้ล จะต้องการซื้อ ซีเรียล พร้อมกันด้วย ที่ความเชื่อมั่น 100%
  - ...
  - กฎความสัมพันธ์ที่ 19 สามารถอธิบายได้ว่า เมื่อซื้อ ซีเรียล, เบียร์ และแอบเปิ้ล จะต้องการซื้อ ไข่ พร้อมกับการซื้อสินค้า 3 ชนิดดังกล่าวเสมอ ที่ความเชื่อมั่น 100%

# Example - Supermarket

- Business Understanding
  - Supermarket แห่งหนึ่ง อยากออกโปรโมชั่น เพื่อเพิ่มยอดขาย โดยโปรโมชั่นที่วางแผนไว้จะเป็นการนำสินค้าที่ลูกค้าซื้อบ่อย ๆ มาลดราคาร่วมกัน

- Data Understanding
  - supermarket แห่งนี้มีข้อมูลการซื้อสินค้าทั้งหมด 108,131 transaction
  - ในแต่ละ transaction จะเป็นสินค้าแต่ละชนิดที่ลูกค้าซื้อ
  - ข้อมูลอยู่ในไฟล์ supermarket.csv

- Data Perparation
  - เนื่องจากใน supermarket แห่งนี้มีสินค้าที่แตกต่างกันเป็นจำนวนมากทำให้อาจจะไม่พบรูปแบบที่น่าสนใจ
  - แปลงข้อมูลสินค้าแต่ละชนิดด้วยประเภทของสินค้า เช่น เนื้อหมู เนื้อไก่ จะแทนด้วย meat

# Example - Supermarket

- แอตทริบิวส์ต่าง ๆ ในชุดข้อมูล

| แอตทริบิวส์ | ประเภท | ความหมาย |
|---|---|---|
| Reciept_id | integer | หมายเลขใบเสร็จรับเงิน |
| Dessserts | Integer | สินค้าประเภทขนมหวาน |
| Meats | Integer | สินค้าประเภทเนื้อ |
| Juices | Integer | สินค้าประเภทน้ำผลไม้ |
| Paper_goods | Integer | สินค้าประเภทกระดาษ |
| Frozen_foods | integer | สินค้าประเภทแช่แข็ง |
| Snack_foods | Integer | สินค้าประเภทขนมคบเคี้ยว |
| Canned_foods | Integer | สินค้าประเภทเครื่องกระป๋อง |
| Beer_wine_spirits | Integer | สินค้าประเภทเครื่องดื่มที่มีแอลกอฮอล์ |
| Dairy | Integer | สินค้าประเภทนมสด |
| Bread | Integer | สินค้าประเภทขนมปัง |
| produce | integer | สินค้าประเภททำสำเร็จ |

# Example - Supermarket

- หมายเหตุ
  - กำหนดให้ min support ของ FP-Growth เป็น 0.1
  - กำหนดให้ min confidence เป็น 0.2

- สรุปข้อมูล ::
- (1) สินค้าประเภทแช่แข็ง มีโอกาสเกิดขึ้นพร้อมกันกับสินค้าใดบ้าง และได้ค่าความเชื่อมั่นเป็นเท่าไร
- (2) สินค้าประเภทกระดาษ มีโอกาสเกิดขึ้นพร้อมกันกับสินค้าใดบ้าง และได้ค่าความเชื่อมั่นเป็นเท่าไร

# แบบฝึกหัดท้ายบท

- 1. จากฐานข้อมูล transactions ของร้านขายคอมพิวเตอร์ ดังต่อไปนี้

| Trans_id | Itemlist |
|---|---|
| T1 | {Mouse, PC, Flashdrive, CD} |
| T2 | { Flashdrive, PC, Webcam, Ram, CD} |
| T3 | {Webcam, PC, CD, Ram} |
| T4 | {CD, PC, Flashdrive} |

- กฎที่ดีที่สุดของร้านขายคอมพิวเตอร์คือ ?
- กฎนี้มีค่า confidence เป็นเท่าไหร่ ?
- อธิบายกฎความสัมพันธ์ ได้อย่างไร ?

- หมายเหตุ กำหนดให้ min support =0.75 และ min confidence = 0.9

# แบบฝึกหัดท้ายบท

- 2. จากฐานข้อมูล transactions ของรายการผักต่าง ๆ ดังต่อไปนี้

| รายการที่ | สินค้าที่ซื้อ |
|---|---|
| 1 | บรอคโคลี่, ข้าวโพด, พริกไทยเขียว |
| 2 | แอสพาราจัส, ข้าวโพด, สควอช |
| 3 | ถั่ว, ข้าวโพด, สควอช, มะเขือเทศ |
| 4 | ถั่ว, ข้าวโพด, พริกไทยเขียว, มะเขือเทศ |
| 5 | แอสพาราจัส, ถั่ว, บรอคโคลี่ |
| 6 | แอสพาราจัส, ถั่ว, สควอช, มะเขือเทศ |
| 7 | ข้าวโพด, มะเขือเทศ |
| 8 | บรอคโคลี่, พริกไทยเขียว, มะเขือเทศ |
| 9 | แอสพาราจัส, ถั่ว, สควอช |
| 10 | ถั่ว, ข้าวโพด |
| 11 | ถั่ว,บรอคโคลี่, พริกไทยเขียว, สควอช |
| 12 | แอสพาราจัส, ถั่ว, สควอช |
| 13 | แอสพาราจัส, ถั่ว, ข้าวโพด, สควอช |
| 14 | ถั่ว, บรอคโคลี่, ข้าวโพด, พริกไทยเขียว, มะเขือเทศ |

| |
|---|
| แอสพาราจัส (asparagus) |
| ถั่ว (beans) |
| บรอคโคลี่ (broccoli) |
| ข้าวโพด (corn) |
| พริกไทยเขียว (Green peppers) |
| สควอช (squash) |
| มะเขือเทศ (tomatoes) |

- กฎที่ดีที่สุดของรายการผักต่าง ๆ คือ ?
- กฎนี้มีค่า confidence เป็นเท่าไหร่ ?
- อธิบายกฎความสัมพันธ์ ได้อย่างไร ?

- หมายเหตุ กำหนดให้ min support =0.5 และ min confidence = 0.7

- https://www.slideshare.net/sitake/practical-data-mining-with-rapidminer-studio-6-a-basic-and-intermediate
- https://my.rapidminer.com/nexus/account/index.html#downloads
-