

# ปฏิบัติการที่ 5

## การใช้เครื่องมือการวิเคราะห์ตะกร้าตลาด

### วัตถุประสงค์

1. เพื่อให้สามารถใช้เครื่องมือวิเคราะห์ตะกร้าตลาด (Market Basket Analysis) ในการหา frequent item set ได้
2. เพื่อให้สามารถใช้เครื่องมือวิเคราะห์ตะกร้าตลาด (Market Basket Analysis) ในการหากฎความสัมพันธ์ (Association Rule) ได้

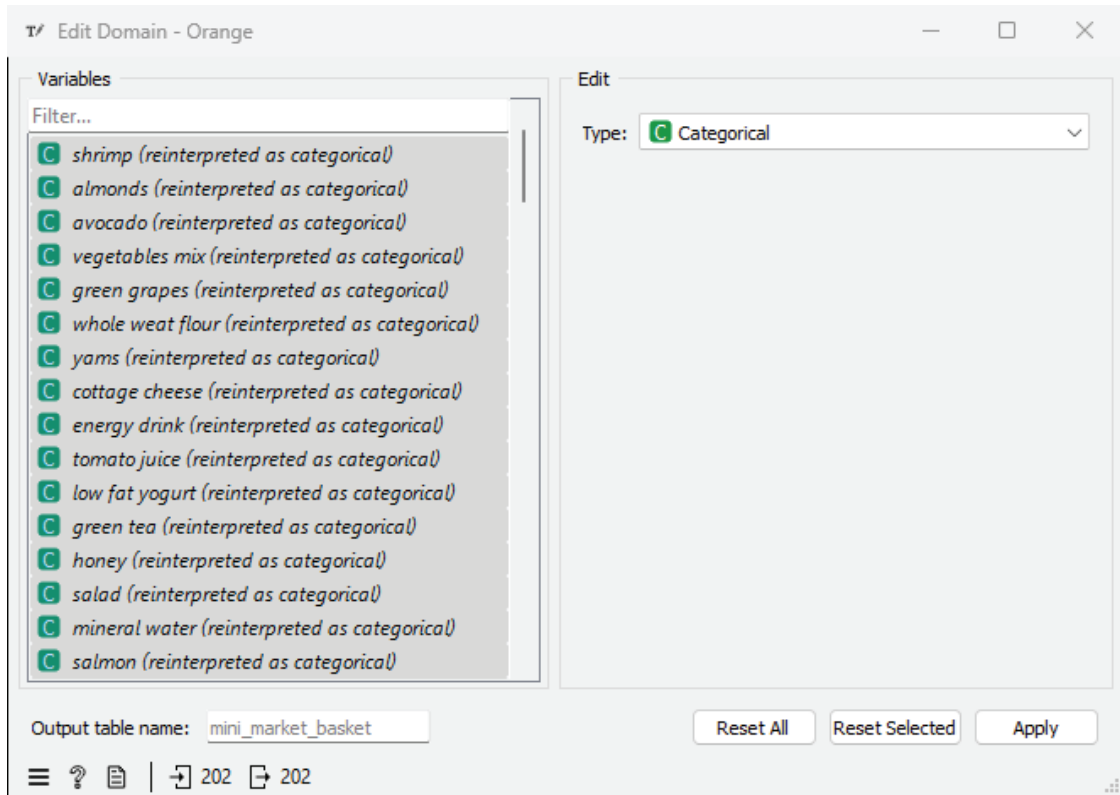
### ชุดข้อมูลปฏิบัติการ

- ชุดข้อมูล Mini Market Basket (สำหรับการสาธิต)
- ชุดข้อมูล Simulated Online Shopping Carts (สำหรับการฝึกปฏิบัติการ)

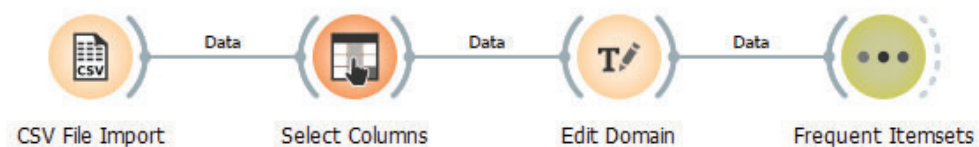
### ขั้นตอนปฏิบัติการ

ขั้นตอนปฏิบัติการ มีดังนี้

1. เปิดโปรแกรม Orange
2. ทำการบันทึก workspace โดยไปที่เมนู File เลือก Save จากนั้นทำการตั้งชื่อไฟล์ในรูปแบบ Practice\_05\_id.ows โดยแทน id ด้วยรหัสนักศึกษา แล้วกดปุ่ม Save
3. นำชุดข้อมูลจากแฟ้มข้อมูล mini\_market\_basket.csv เข้าสู่โปรแกรม Orange โดยใช้โมดูล CSV File Import
4. เลือกใช้ข้อมูลตัวแปรทุกตัวแปร ยกเว้น ตัวแปร id ในการวิเคราะห์ตะกร้าตลาด โดยใช้โมดูล Select Columns (ดูรายละเอียดในปฏิบัติการที่ 4)
5. ตรวจสอบและเปลี่ยนชนิดข้อมูลของตัวแปรที่เลือกใช้ในการวิเคราะห์ โดยใช้โมดูล Edit Domain (ดูปฏิบัติการที่ 2) กำหนดให้แต่ละตัวแปรต้องมีชนิดเป็น categorical ดังรูป



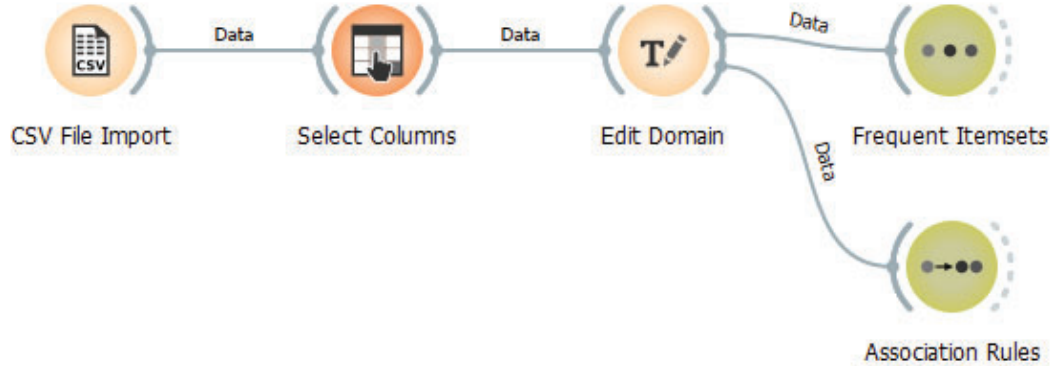
6. หา Frequent Item Sets ในชุดข้อมูล mini market basket โดยใช้โมดูล Frequent Itemsets คลิกเลือกโมดูล Frequent Itemsets จะปรากฏโมดูลใน workspace จากนั้นคลิกเชื่อมโมดูล Edit Domain (นำผลลัพธ์จากโมดูล Edit Domain ไปใช้ต่อ) จากด้าน output เข้าสู่โมดูล Frequent Itemsets ด้าน input ดังรูป



7. ดับเบิลคลิกที่โมดูล Frequent Itemsets จากปรากฏหน้าต่างตั้งค่าพารามิเตอร์ของโมดูล กำหนดค่าสนับสนุนน้อยที่สุด (Minimal support) และจำนวนไอเทมเซตมากที่สุด (Max. number of itemsets) ในที่นี้กำหนดค่าสนับสนุนน้อยที่สุด เท่ากับ 10% และจำนวนไอเทมเซตมากที่สุด เท่ากับ 10000 ไอเทมเซต จากนั้นคลิกปุ่ม Find Itemsets จะปรากฏผลลัพธ์ ดังรูป

Itemsets	Support	%
mineral water=1.0	80	39.6
spaghetti=1.0	29	14.36
eggs=1.0	24	11.88
chocolate=1.0	25	12.38
spaghetti=1.0	67	33.17
eggs=1.0	52	25.74
chocolate=1.0	50	24.75
french fries=1.0	48	23.76
green tea=1.0	45	22.28
frozen vegetables=1.0	42	20.79
milk=1.0	39	19.31
ground beef=1.0	38	18.81
burgers=1.0	29	14.36
olive oil=1.0	27	13.37
soup=1.0	25	12.38
frozen smoothie=1.0	24	11.88
pancakes=1.0	24	11.88
shrimp=1.0	22	10.89
turkey=1.0	22	10.89
cake=1.0	21	10.4

8. ผลลัพธ์จะปรากฏทางด้านขวาของหน้าต่าง คือ ไอเทมเซตที่เกิดขึ้นบ่อย (Frequent Itemsets) พร้อมกับจำนวนครั้งในการเกิดขึ้นของไอเทมเซตในชุดข้อมูลและค่าสนับสนุน
9. ทั้งนี้สามารถเลือกกรอกไอเทมเซตที่สนใจได้ โดยการกำหนดเงื่อนไขในส่วน Filter itemsets
10. หา Frequent Item Sets ในชุดข้อมูล mini market basket โดยใช้โมดูล Frequent Itemsets คลิกเลือกโมดูล Frequent Itemsets จะปรากฏโมดูลใน workspace จากนั้นคลิกเชื่อมโมดูล Edit Domain (นำผลลัพธ์จากโมดูล Edit Domain ไปใช้ต่อ) จากด้าน output เข้าสู่โมดูล Frequent Itemsets ด้าน input ดังรูป
11. หา Association Rules ในชุดข้อมูล mini market basket โดยใช้โมดูล Association Rules คลิกเลือกโมดูล Association Rules จะปรากฏโมดูลใน workspace จากนั้นคลิกเชื่อมโมดูล Edit Domain (นำผลลัพธ์จากโมดูล Edit Domain ไปใช้ต่อ) จากด้าน output เข้าสู่โมดูล Association Rules ด้าน input ดังรูป



12. ดับเบิลคลิกที่โมดูล Association Rules จากปรากฏหน้าจอต่างตั้งค่าพารามิเตอร์ของโมดูล กำหนดค่าสนับสนุนน้อยที่สุด (Min. support) ค่าความเชื่อมั่นน้อยที่สุด (Min. conf) และจำนวนกฎมากที่สุด (Max. rules) ในที่นี้กำหนดกำหนดค่าสนับสนุนน้อยที่สุดและค่าความเชื่อมั่นน้อยที่สุด เท่ากับ 10% และจำนวนไอเทมเซตมากที่สุด เท่ากับ 10000 ไอเทมเซต จากนั้นคลิกปุ่ม Find Rules จะปรากฏผลลัพธ์ ดังรูป

\*\*\* Association Rules - Orange

Info  
Rules: 6 (shown 6)

Find association rules

Min. supp.: 10 %

Min. conf.: 10 %

Max. rules: 10k

Induce only classification rules

Restrict search by below filters

Find Rules

Filter by Antecedent

Contains:

Items, min: 1 max: 999

Filter by Consequent

Contains:

Items, min: 1 max: 999

Send selection

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent	Consequent
0.144	0.362	0.396	0.838	1.093	0.012	mineral water=1.0	→ spaghetti=1.0
0.144	0.433	0.332	1.194	1.093	0.012	spaghetti=1.0	→ mineral water=1.0
0.119	0.300	0.396	0.650	1.165	0.017	mineral water=1.0	→ eggs=1.0
0.119	0.462	0.257	1.538	1.165	0.017	eggs=1.0	→ mineral water=1.0
0.124	0.312	0.396	0.625	1.262	0.026	mineral water=1.0	→ chocolate=1.0
0.124	0.500	0.248	1.600	1.262	0.026	chocolate=1.0	→ mineral water=1.0

202 | - | 6

13. ผลลัพธ์จะปรากฏทางด้านขวาของหน้าต่าง คือ กฎความสัมพันธ์ที่เกิดขึ้นบ่อยและน่าเชื่อถือ (Frequent and confident association rules) พร้อมกับค่าสนับสนุน ค่าความเชื่อมั่น ค่าลิฟท์ และอื่น ๆ
14. ทั้งนี้สามารถเลือกกฎความสัมพันธ์ที่สนใจได้ โดยการกำหนดเงื่อนไขในส่วน Filter

## แบบฝึกปฏิบัติการ

ให้นักศึกษาทำแบบฝึกปฏิบัติการ ตามลำดับขั้นตอนต่อไปนี้

1. ให้นักศึกษานำชุดข้อมูล Simulated Online Shopping Carts จากแฟ้มข้อมูล simulated\_online\_shopping\_carts.csv เข้าสู่โปรแกรม Orange
2. เลือกใช้ข้อมูลตัวแปรทุกตัวแปร ยกเว้นตัวแปร id โดยใช้โมดูล Select Columns
3. หา Frequent Itemsets ในชุดข้อมูล simulated online shopping carts โดยกำหนดค่าพารามิเตอร์ Minimal Support เท่ากับ 15%
4. หากฎความสัมพันธ์ (Association Rules) ในชุดข้อมูล simulated online shopping carts โดยกำหนดค่าพารามิเตอร์ ดังนี้
  - Minimal Support: 15%
  - Minimal Confidence: 20%
5. ทดลองเปลี่ยนค่าพารามิเตอร์ต่างๆ ของการหา Frequent Itemsets และ Association Rules
6. **สิ่งที่ต้องส่งเป็นการบ้าน** ทำการบันทึกไฟล์ workspace ของนักศึกษา โดยตั้งชื่อไฟล์ในรูปแบบ Lab\_05\_id.ows โดยแทน id ด้วยรหัสนักศึกษา ส่งผ่านเว็บไซต์ <http://hw.cs.science.cmu.ac.th>