

ปฏิบัติการที่ 3

การคำนวณค่าสถิติเชิงพรรณนาจากตารางข้อมูลหลัก

วัตถุประสงค์

1. เพื่อให้สามารถใช้เครื่องมือวิเคราะห์ข้อมูลในการคำนวณค่าสถิติเชิงพรรณนา เพื่ออธิบายลักษณะของข้อมูลได้
2. เพื่อให้สามารถใช้เครื่องมือวิเคราะห์ข้อมูลในการคำนวณหาค่าสหสัมพันธ์ (Correlation) ระหว่างตัวแปรได้

ชุดข้อมูลปฏิบัติการ

- ชุดข้อมูล Iris (สำหรับการสาธิต)
- ชุดข้อมูล Airfoil Self-Noise (สำหรับการฝึกปฏิบัติการ)

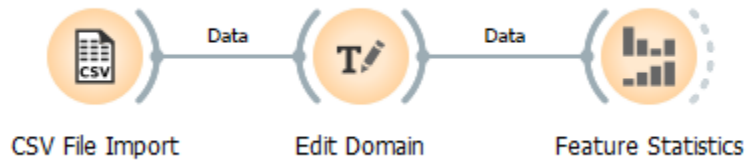
ขั้นตอนปฏิบัติการ

ขั้นตอนปฏิบัติการ มีดังนี้

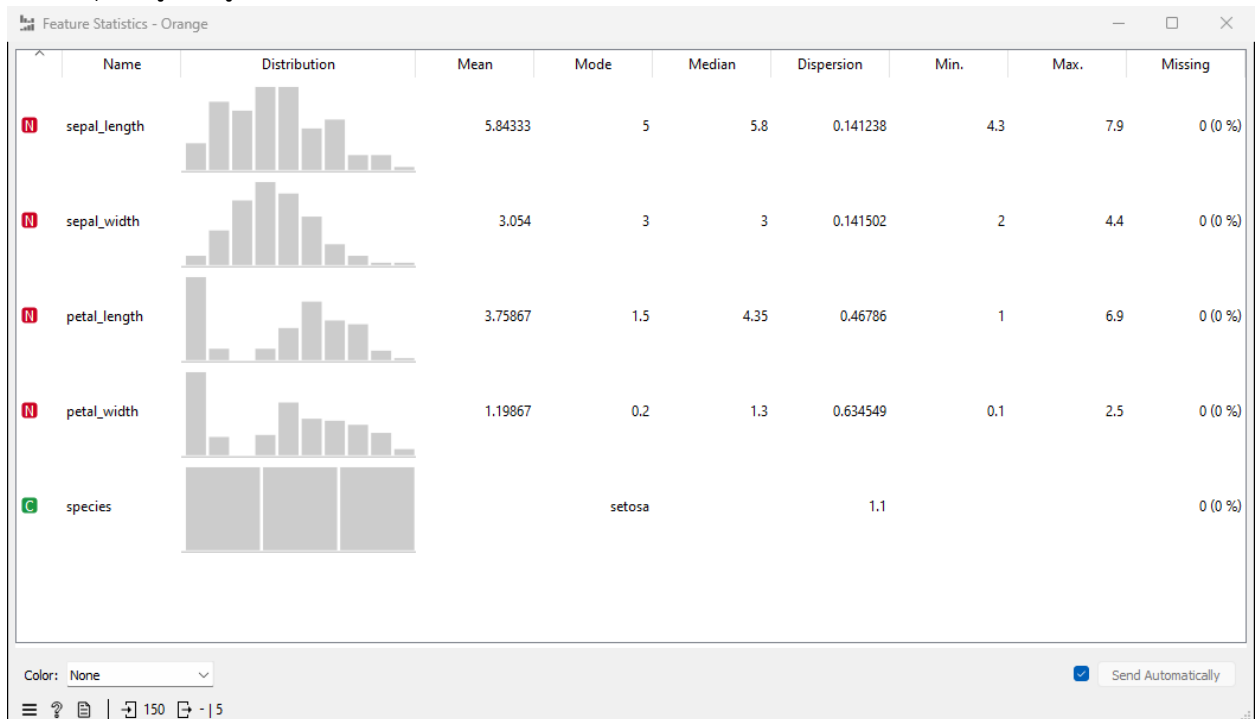
1. เปิดโปรแกรม Orange
2. ทำการบันทึก workspace โดยไปที่เมนู File เลือก Save จากนั้นทำการตั้งชื่อไฟล์ในรูปแบบ Practice_03_id.ows โดยแทน id ด้วยรหัสนักศึกษา แล้วกดปุ่ม Save
3. นำชุดข้อมูลจากแฟ้มข้อมูล iris.csv เข้าสู่โปรแกรม Orange โดยใช้โมดูล CSV File Import
4. ตรวจสอบและเปลี่ยนชนิดข้อมูลของตัวแปร โดยใช้โมดูล Edit Domain (ดูปฏิบัติการที่ 2) กำหนดให้แต่ละตัวแปรต้องมีชนิดข้อมูล ดังนี้

ตัวแปร	ชนิดข้อมูล
sepal_length	Numeric Feature
sepal_width	Numeric Feature
petal_length	Numeric Feature
petal_width	Numeric Feature
species	Categorical Feature

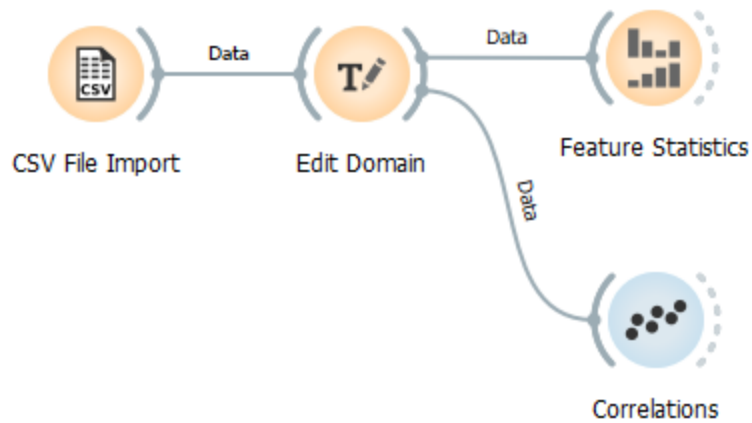
5. หากนำโมดูล Feature Statistics มาต่อเข้ากับโมดูล Edit Domain ดังรูป



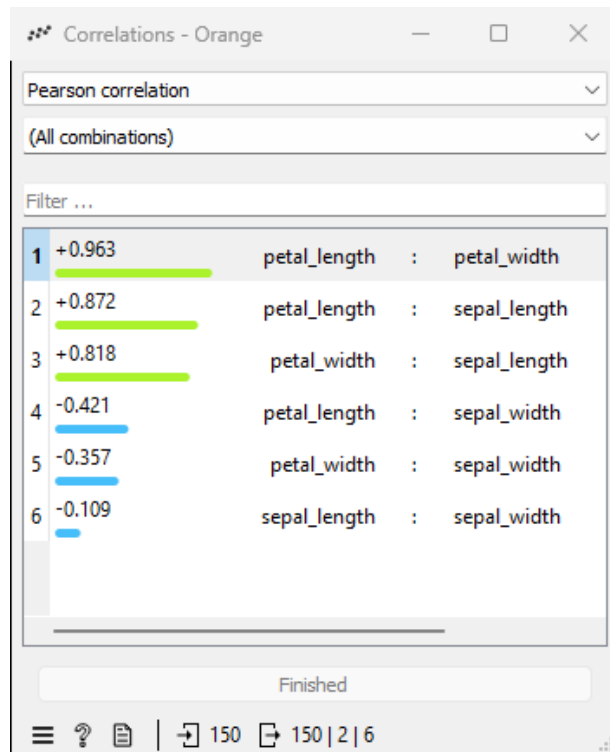
6. ดับเบิลคลิกที่โมดูล Feature Statistics จากปรากฏหน้าต่างแสดงผลค่าสถิติเชิงพรรณนาของแต่ละตัวแปรในชุดข้อมูล ดังรูป



- สังเกตลักษณะการกระจายและค่าสถิติเชิงพรรณนาของข้อมูลแต่ละตัวแปร
- คำนวณหาค่าสหสัมพันธ์ระหว่างตัวแปรเชิงตัวเลขแต่ละตัวแปร โดยใช้โมดูล Correlations
- คลิกเลือกโมดูล Correlations จะปรากฏโมดูลใน workspace จากนั้นคลิกเชื่อมโมดูล Edit Domain (นำผลลัพธ์จากโมดูล Edit Domain ไปใช้ต่อ) จากด้าน output เข้าสู่โมดูล Correlations ด้าน input ดังรูป



10. ดับเบิลคลิกที่โมดูล Correlations จากปรากฏการณ์ต่างแสดงค่าสหสัมพันธ์ระหว่างตัวแปรเชิงตัวเลขแต่ละคู่ที่เป็นไปได้ ดังรูป



11. สังเกตและอภิปรายค่าสหสัมพันธ์ระหว่างตัวแปรเชิงตัวเลขแต่ละคู่ที่เป็นไปได้

แบบฝึกปฏิบัติการ

ให้นักศึกษาทำแบบฝึกปฏิบัติการ ตามลำดับขั้นตอนต่อไปนี้

1. ให้นักศึกษานำชุดข้อมูล Airfoil Self-Noise จากแฟ้มข้อมูล airfoil_self_noise.csv เข้าสู่โปรแกรม Orange
2. คำนวณค่าสถิติเชิงพรรณนาของแต่ละตัวแปรในชุดข้อมูล airfoil self-noise โดยใช้โมดูล Feature Statistics แล้วศึกษาและอภิปรายผลลัพธ์
3. คำนวณค่าสหสัมพันธ์ของแต่ละคู่ตัวแปรในชุดข้อมูล airfoil self-noise โดยใช้โมดูล Correlations แล้วศึกษาและอภิปรายผลลัพธ์
4. **สิ่งที่ต้องส่งเป็นการบ้าน** ทำการบันทึกไฟล์ workspace ของนักศึกษา โดยตั้งชื่อไฟล์ในรูปแบบ Lab_03_id.ows โดยแทน id ด้วยรหัสนักศึกษา ส่งผ่านเว็บไซต์ <http://hw.cs.science.cmu.ac.th>