**Lab 4 วิชา การเรียนรู้แบบไม่มีผู้สอนของเครื่อง (Unsupervised Machine Learning)**

**หัวข้อ**

1. Hierarchical Clustering
2. Dendrogram

**คำสั่งส่วนที่ 1**

1. *การทดลองในครั้งนี้ อนุญาตให้ใช้ Library ได้ทั้งหมด*
2. ใช้ชุดข้อมูล load\_iris()

iris = datasets.load\_iris()  
X = iris.data  
y = iris.target

1. ตรวจสอบหมายเลข Label ของข้อมูลแต่ละตัว บันทึกผลลงในช่อง target
2. ใช้ Hierarchical ในการวิเคราะห์คลัสเตอร์ของ load\_iris() โดยใช้ข้อมูลทั้ง 100% ในการเทรน โดยไม่กำหนด target ให้ model (เทรนโดยใช้แค่ data) บันทึกผลลัพธ์ของการทำคลัสเตอริ่ง อาจเลือกใช้วิธี Agglomerative หรือ Divisive Analysis ก็ได้
3. บันทึกผลการทดลองลงในตารางบันทึกผลส่วนที่ 1 ให้สมบูรณ์

**คำสั่งส่วนที่ 2**

1. ให้นักศึกษาสร้าง Dendrogram **ที่สมบูรณ์**ของชุดข้อมูล load\_iris()
2. บันทึกกราฟที่สมบูรณ์นี้ลงในการบันทึกผลการทดลองส่วนที่ 2
3. นำกราฟ Dendrogram ที่พล็อตได้ไปแก้ไข เพื่อบันทึกผลการทดลองส่วนที่ 2 ให้สมบูรณ์ (อนุญาตให้ใช้โปรแกรม edit ภาพได้ ไม่จำเป็นต้องใช้ภาษา python ในการบันทึกผลการทดลองส่วนนี้)
4. รายละเอียดการบันทึกผล
   1. บันทึกผลลงในใบงาน (หน้า 2 เป็นต้นไป) ไม่ต้องกังวลเกี่ยวกับแบบฟอร์มหรือจำนวนหน้า แม้ว่า ตารางจะตกหน้า แยกออกจากกัน หรือทำให้มีพื้นที่ว่างในบางหน้ามากเกินไป ก็ไม่มีปัญหา
   2. เซฟไฟล์นี้ ตั้งชื่อว่า Lab04\_result\_รหัสนักศึกษา.docx ส่งพร้อมโค้ด ตั้งชื่อว่า Lab04\_code\_รหัสนักศึกษา.py (แยกไฟล์กันมา)
   3. ส่งไฟล์โดยอัปโหลดไปยัง <https://silpakorn-my.sharepoint.com/:f:/g/personal/aonpong_p_su_ac_th/EmRn8pOWnXJOvxtOcoRs98IBI6ex0H6-AwN_m-BxJYG3gg?e=XSs2Zt>

\* พยายามควบคุมพารามิเตอร์ โค้ดที่ส่งจะต้องไม่มี Warning (เช่น Convergence Warning หรือ Warning อื่น ๆ)

**630710440 พีท อ่อนทอง**

**บันทึกผลการทดลอง ส่วนที่ 1**

**ผลการทดลอง Hierarchical clustering**

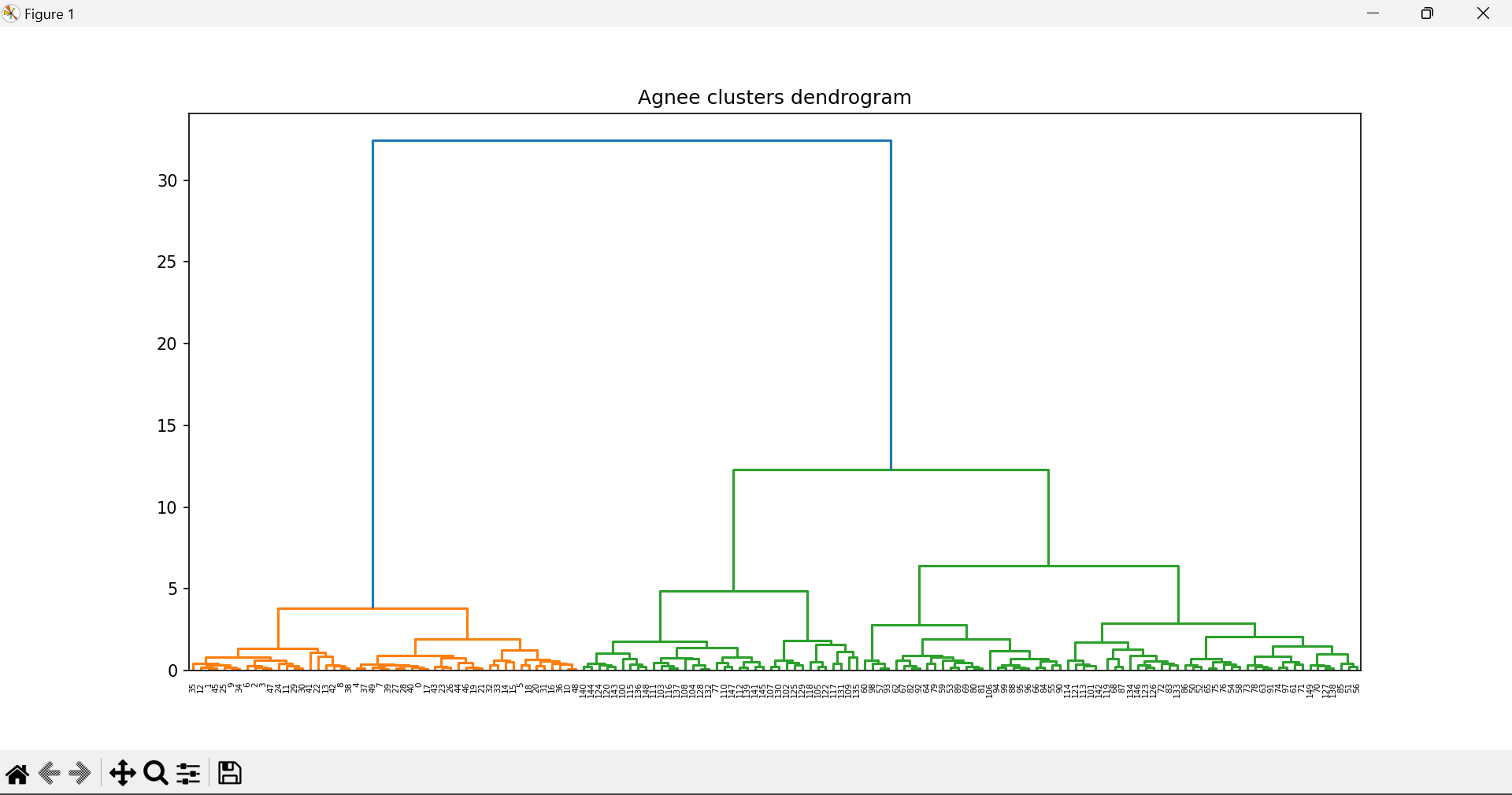
|  |  |  |
| --- | --- | --- |
| Clustering result | | |
| Hierarchical (…Agglomerative? / Divisive?...) | k-Mean (Lab03) | Target |
| 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1  1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0 0 | 2 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1  1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0  0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0 0 |
| 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1  1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2 0 2 2 2 2 0 2 2 2 2  2 2 0 0 2 2 2 2 0 2 0 2 0 2 2 0 0 2 2 2 2 2 0 0 2 2 2 0 2 2 2 0 2 2 2 0 2  2 0 | 3 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1  1 1 1 1 1 1 1 1 1 1 1 1 1 0 2 0 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2  2 2 2 0 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 0 2 0 0 0 0 2 0 0 0 0  0 0 2 2 0 0 0 0 2 0 2 0 2 0 0 2 2 0 0 0 0 0 2 0 0 0 0 2 0 0 0 2 0 0 0 2 0  0 2 |
| 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1  1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 3 2 3 2 3 2 3 3 2 3 2 3 2 3 3 2 3 2 2 2 2  2 2 2 0 2 3 3 3 3 2 3 2 2 2 3 3 3 2 3 3 3 3 3 2 3 3 0 2 0 0 0 0 3 0 0 0 0  0 0 2 2 0 0 0 0 2 0 2 0 2 0 0 2 2 0 0 0 0 0 2 2 0 0 0 2 0 0 0 2 0 0 0 2 0  0 2 | 4 | 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1  1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 2 0 2 0 2 0 2 2 2 2 0 2 0 2 2 0 2 0 2 0 0  0 0 0 0 0 2 2 2 2 0 2 0 0 0 2 2 2 0 2 2 2 2 2 0 2 2 3 0 3 0 3 3 2 3 3 3 0  0 3 0 0 3 3 3 3 0 3 0 3 0 3 3 0 0 3 3 3 3 3 0 0 3 3 0 0 3 3 3 0 3 3 3 0 0  3 0 |
| 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  0 0 0 0 0 0 0 0 0 0 0 0 0 2 2 2 1 2 1 2 1 2 1 1 2 1 2 1 2 1 1 2 1 2 2 2 2  2 2 2 4 2 1 1 1 1 2 1 2 2 2 1 1 1 2 1 1 1 1 1 2 1 1 4 2 3 4 4 3 1 3 4 3 4  4 4 2 2 4 4 3 3 2 4 2 3 2 4 3 2 2 4 3 3 3 4 2 2 3 4 4 2 4 4 4 2 4 4 4 2 4  4 2 | 5 | 4 1 1 1 4 4 1 4 1 1 4 1 1 1 4 4 4 4 4 4 4 4 1 4 1 1 4 4 4 1 1 4 4 4 1 1 4  4 1 4 4 1 1 4 4 1 4 1 4 1 0 0 0 2 0 2 0 2 0 2 2 2 2 0 2 0 2 2 0 2 0 2 0 0  0 0 0 0 0 2 2 2 2 0 2 0 0 0 2 2 2 0 2 2 2 2 2 0 2 2 3 0 3 0 3 3 2 3 3 3 0  0 3 0 0 3 3 3 3 0 3 0 3 0 3 3 0 0 3 3 3 3 3 0 0 3 3 0 0 3 3 3 0 3 3 3 0 0  3 0 |
| 2 5 5 5 2 2 5 2 5 5 2 5 5 5 2 2 2 2 2 2 2 2 5 2 5 5 2 2 2 5 5 2 2 2 5 5 2  2 5 2 2 5 5 2 2 5 2 5 2 2 0 0 0 1 0 1 0 1 0 1 1 0 1 0 1 0 1 1 0 1 0 0 0 0  0 0 0 4 0 1 1 1 1 0 1 0 0 0 1 1 1 0 1 1 1 1 1 0 1 1 4 0 3 4 4 3 1 3 4 3 4  4 4 0 0 4 4 3 3 0 4 0 3 0 4 3 0 0 4 3 3 3 4 0 0 3 4 4 0 4 4 4 0 4 4 4 0 4  4 0 | 6 | 4 1 1 1 4 4 1 4 1 1 4 1 1 1 4 4 4 4 4 4 4 4 1 4 1 1 4 4 4 1 1 4 4 4 1 1 4  4 1 4 4 1 1 4 4 1 4 1 4 1 5 5 5 2 5 2 5 2 5 2 2 5 2 5 2 5 2 2 5 2 0 2 5 5  5 5 5 5 5 2 2 2 2 0 2 5 5 5 2 2 2 5 2 2 2 2 2 5 2 2 3 0 3 0 3 3 2 3 3 3 0  0 3 0 0 0 0 3 3 0 3 0 3 0 3 3 0 0 0 3 3 3 0 0 0 3 0 0 0 3 3 0 0 3 3 0 0 0  0 0 |
| 2 5 5 5 2 2 5 2 5 5 2 5 5 5 2 2 2 2 2 2 2 2 5 2 5 5 2 2 2 5 5 2 2 2 5 5 2  2 5 2 2 5 5 2 2 5 2 5 2 2 1 1 1 0 1 0 1 0 1 0 0 1 0 1 0 1 0 0 6 0 1 1 6 1  1 1 1 4 1 0 0 0 0 6 0 1 1 6 0 0 0 1 0 0 0 0 0 1 0 0 4 6 3 4 4 3 0 3 4 3 4  4 4 6 6 4 4 3 3 6 4 6 3 6 4 3 6 1 4 3 3 3 4 6 6 3 4 4 1 4 4 4 6 4 4 4 6 4  4 1 | 7 | 4 1 1 1 4 4 1 4 1 1 4 1 1 1 4 4 4 4 4 4 4 4 1 4 1 1 4 4 4 1 1 4 4 4 1 1 4  4 1 4 4 1 1 4 4 1 4 1 4 1 5 5 5 2 5 0 5 2 5 2 2 5 2 5 2 5 0 2 5 2 0 2 0 5  5 5 5 5 5 2 2 2 2 0 0 5 5 5 2 2 2 5 2 2 2 2 2 5 2 2 6 0 3 6 6 3 2 3 6 3 6  6 6 0 0 6 6 3 3 0 6 0 3 0 6 3 0 0 6 3 3 3 6 0 0 3 6 6 0 6 6 6 0 6 6 6 0 6  6 0 |

\*เพราะคลัสเตอริ่ง เป็นวิธีการที่ไม่สนใจว่า label ของกลุ่มคืออะไร รู้เพียงแค่ว่าเป็นกลุ่มเดียวกันหรือไม่ ดังนั้น เราอาจพบเหตุการณ์ที่หมายเลข label ไม่ตรงกันได้บ่อยครั้ง “จากโมเดลที่เทรนมาด้วยกัน ถ้า label ที่ควรเป็นกลุ่มเดียวกันมีหมายเลขเดียวกัน ถือว่าใช้ได้”

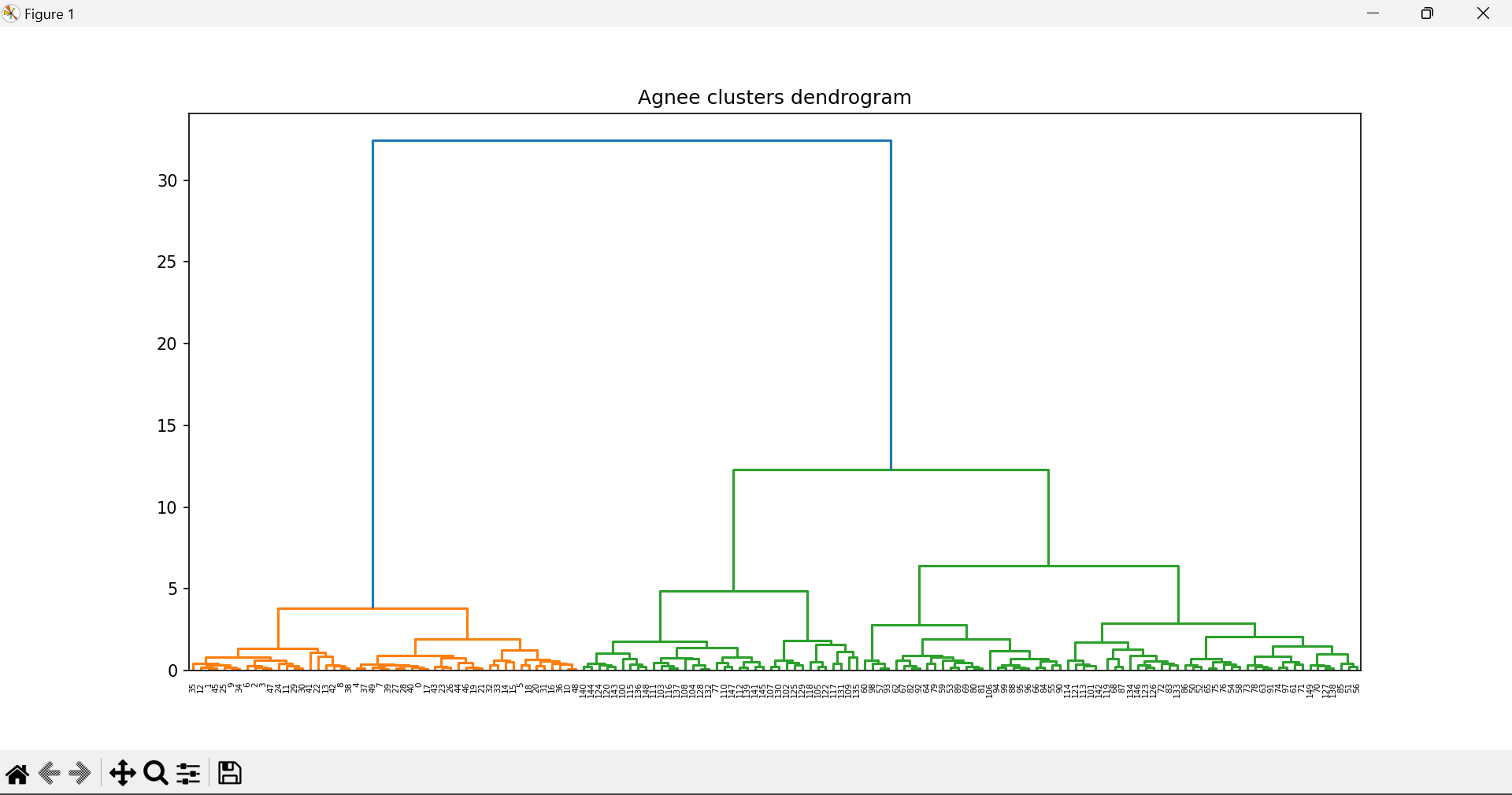
\*\*ผลการทดลองในตารางนี้ เมื่อวางข้อมูลทั้งหมดลงมาแล้วอาจยาวหลายหน้า ไม่ต้องสนใจ วางได้เลย

**บันทึกผลการทดลอง ส่วนที่ 2**

1. ภาพกราฟต้นไม้ แสดง Dendrogram **โดยสมบูรณ์**ของข้อมูล load\_iris()



1. ให้นักศึกษาขีดเส้นตัดขวางบนภาพกราฟต้นไม้ Dendrogram **เพื่อแยกข้อมูลออกเป็น 3 คลัสเตอร์**



1. ให้นักศึกษาขีดเส้นตัดขวางบนภาพกราฟต้นไม้ Dendrogram **เพื่อแยกข้อมูลออกเป็น 5 คลัสเตอร์**

