

Class period 13

บทที่ 6 การแสดงผลการกระจายข้อมูล (ต่อ)
Visualize_data_distribution part2

1

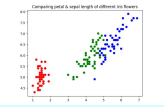
quiz3 (5 นาที)

- โปรดทราบ คัด 57 กับ สาราน ๓๘57
- จากไฟล์ quiz3 คัด ๓๘57
- จาก https://data.gov.th/dataset/chem_fac313e-7763-4bc2-8461-a19210u8iud6

2

การเพิ่มชื่อกราฟ

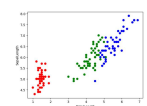
- สามารถใส่ชื่อกราฟตามที่อยากได้โดยการใส่ชื่อ plot() เช่น
- `plt.scatter(df2['PetalLength'], df2['SepalLength'], c=df2['Name'])`
- `plt.title('Comparing petal & sepal length of different iris flowers')`



3

การเพิ่มชื่อแกนในกราฟ

- สามารถใส่ชื่อแกนในกราฟได้ตามต้องการโดยการใส่ชื่อ `plt.xlabel()` และ `plt.ylabel()` เช่น
- `plt.scatter(df2['PetalLength'], df2['SepalLength'], c=df2['Name'])`
- `plt.xlabel('PetalLength')`
- `plt.ylabel('SepalLength')`



4

การแยกข้อมูลเพื่อกำหนดลักษณะ marker แต่ละกลุ่ม

- โดยการแยกข้อมูลเป็นกลุ่มตามค่าของ marker แต่ละกลุ่ม จึงเป็นที่มาของโค้ดด้านล่างและจำนวน record ของแต่ละกลุ่มตาม เช่น
- iris.setosa มี record 50 ชื่อ setosa
- iris.vernecula มี record 50 ชื่อ vernecula
- เมื่อรู้ข้อมูลค่าและจำนวน record แต่ละชื่อแล้ว สามารถกำหนด record ที่ต้องการมาได้
- iris.setosa คือ [50] มีชื่อ record 50 ชื่อ record ที่ ๕๑
- iris.vernecula คือ [50:100] มีชื่อ record ที่ 50 ถึง record ที่ ๙๙
- iris.vernecula คือ [100] มีชื่อ record ที่ ๙๙ ถึง record สุดท้าย

5

ตัวอย่างการแยกข้อมูลเพื่อ plot กราฟซ้อนกัน

- `plt.scatter(df2['PetalLength'][:50], df2['SepalLength'][:50], c=df2['Name'][:50])`
- `plt.scatter(df2['PetalLength'][50:100], df2['SepalLength'][50:100], c=df2['Name'][50:100])`
- `plt.scatter(df2['PetalLength'][100:], df2['SepalLength'][100:], c=df2['Name'][100:])`

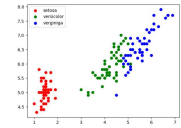
6

Parameter: label และคำสั่ง plt.legend() ใช้กำหนดชื่อและแสดงชื่อ marker แต่ละกลุ่ม

- สามารถกำหนดชื่อ marker แต่ละกลุ่มได้โดยการใส่ Parameter: label เป็นชื่อที่ต้องการมาใส่และจำนวน record ของแต่ละกลุ่ม
- และใช้คำสั่ง `plt.legend()` เพื่อแสดงชื่อกราฟตาม เช่น
- `plt.scatter(df2['PetalLength'][:50], df2['SepalLength'][:50], c=df2['Name'][:50], label='setosa')`
- `plt.scatter(df2['PetalLength'][50:100], df2['SepalLength'][50:100], c=df2['Name'][50:100], label='vernalculade')`
- `plt.scatter(df2['PetalLength'][100:], df2['SepalLength'][100:], c=df2['Name'][100:], label='verginliga')`
- `plt.legend()`

7

ผลลัพธ์ตัวอย่าง กำหนดชื่อและแสดงชื่อ marker แต่ละกลุ่ม

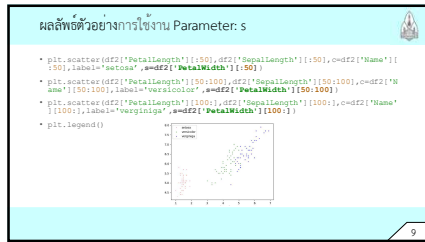


8

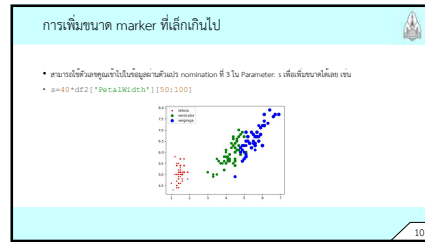
Parameter: s ใช้ขนาดของ marker แสดงค่าของ feature

- สามารถกำหนดขนาดของ marker โดยการใส่ค่าของ Parameter: s = notation ที่ 3 สำหรับค่าการกระจายข้อมูล
- ขนาด 3 notation ที่เขียนไว้ คือ notation ที่ 3 เลขที่เขียนไว้เป็นตัวเลขขนาด สามารถ marker ขนาดใหญ่ เล็ก
- และชื่อ marker ขยายตามได้ เช่น
- ตัวอย่างการใช้ Parameter: s เป็นการเพิ่มขนาด marker เป็น notation ที่ 3
- `plt.scatter(df2['PetalLength'][:50], df2['SepalLength'][:50], c=df2['Name'][:50], label='setosa', s=df2['PetalLength'][:50])`

9



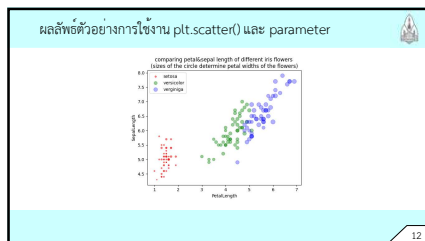
10



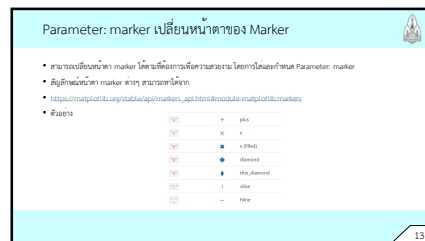
11



12



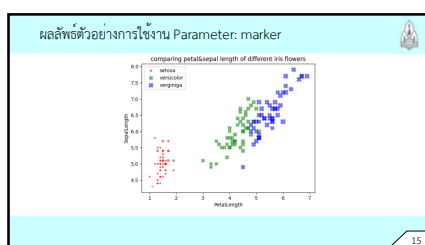
13



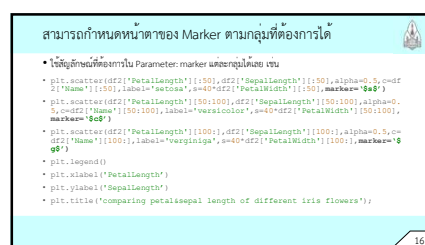
14



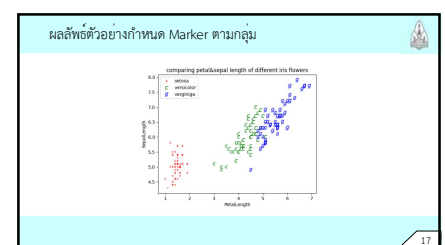
15



16



17



18