



Class period 13

บทที่ 6 การแสดงและการกระจายของข้อมูล (ต่อ)
Visualize_Data_Distribution part2

1



quiz3 (5 นาที)

- โหลดตาราง คค57 กับ ตาราง พย57
- จากไฟล์ จัดชื่อ คค57-คค58
- จาก https://data.go.th/dataset/item_fadc318e-7743-4bc2-84d1-a19210a9bbd6

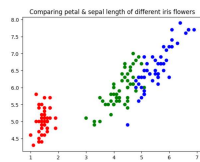
1

2



การเพิ่มชื่อกราฟ

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



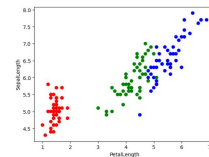
2

3



การเพิ่มชื่อแกนในกราฟ

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



3

4



การแยกข้อมูลเพื่อกำหนดลักษณะ marker แต่ละกลุ่ม

- โดยการแยกข้อมูลเพื่อกำหนดลักษณะ marker แต่ละกลุ่ม จำเป็นที่จะต้องรู้ลำดับและจำนวน record ของแต่ละกลุ่มก่อน เช่น
- ในข้อมูลดอกไม้ Iris มีอยู่ 3 สายพันธุ์ คือ
- Iris-setosa, Iris-versicolor และ Iris-verginiga
- เมื่อรู้ข้อมูลลำดับและจำนวน record แต่ละสายพันธุ์ สามารถกำหนด record ที่ต้องการแยกได้
- Iris-setosa คือ [:50] ข้อมูล record แรกถึง record ที่ 49
- Iris-versicolor คือ [50:100] ข้อมูล record ที่ 50 ถึง record ที่ 99
- Iris-verginiga คือ [100:] ข้อมูล record ที่ 99 ถึง record สุดท้าย

4

5



ตัวอย่างการใช้การแยกข้อมูลเพื่อ 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:])`

5

6

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

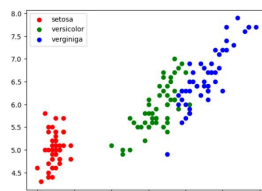
- สามารถกำหนดชื่อ marker แต่ละกลุ่มได้ โดยการใส่ Parameter: label เพิ่มในข้อมูลแยกตามลำดับและจำนวน record ของแต่ละกลุ่ม
- และใช้คำสั่ง plt.legend() เพื่อแสดงชื่อแบบกราฟ เช่น

```
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='versicolor')
plt.scatter(df2['PetalLength'][100:], df2['SepalLength'][100:], c=df2['Name'][100:], label='virginica')
plt.legend()
```

6

7

ผลลัพธ์ตัวอย่าง กำหนดชื่อและแสดงชื่อ marker แต่ละกลุ่ม



7

8

Parameter: s ใช้ขนาดของ marker แสดงค่าของ feature

สามารถกำหนดขนาดของ marker โดยการใส่และกำหนด Parameter: s= nomination ที่ 3 สำหรับดูการกระจายของข้อมูลระหว่าง 3 nomination (คอลัมน์) โดย nomination ที่ 3 จะถูกเปรียบเทียบในรูปแบบขนาด ค่ายก marker จะขนาดใหญ่ ค่า น้อย marker จะขนาดเล็ก เช่น

- ตัวอย่างการใช้งาน Parameter: s ต้องการเพิ่มคอลัมน์ PetalWidth เป็น nomination ที่ 3


```
plt.scatter(df2['PetalLength'][:50], df2['SepalLength'][:50], c=df2['Name'][:50], label='setosa', s=df2['PetalWidth'][:50])
```

8

9

ผลลัพธ์ตัวอย่างการใช้งาน Parameter: s

```
plt.scatter(df2['PetalLength'][:50], df2['SepalLength'][:50], c=df2['Name'][:50], label='setosa', s=df2['PetalWidth'][:50])
plt.scatter(df2['PetalLength'][50:100], df2['SepalLength'][50:100], c=df2['Name'][50:100], label='versicolor', s=df2['PetalWidth'][50:100])
plt.scatter(df2['PetalLength'][100:], df2['SepalLength'][100:], c=df2['Name'][100:], label='virginica', s=df2['PetalWidth'][100:])
plt.legend()
```



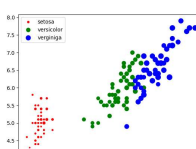
9

10

การเพิ่มขนาด marker ที่เล็กเกินไป

- สามารถใช้ตัวเลขคูณเข้าไปในข้อมูลผ่านค่านับ nomination ที่ 3 ใน Parameter: s เพื่อเพิ่มขนาดได้โดย เช่น

```
s=40*df2['PetalWidth'][:50:100]
```



10

11

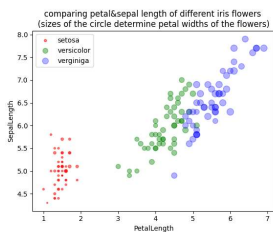
ตัวอย่างการใช้งาน plt.scatter() และ parameter

```
plt.scatter(df2['PetalLength'][:50], df2['SepalLength'][:50], alpha=0.5, c=df2['Name'][:50], label='setosa', s=40*df2['PetalWidth'][:50])
plt.scatter(df2['PetalLength'][50:100], df2['SepalLength'][50:100], alpha=0.4, c=df2['Name'][50:100], label='versicolor', s=40*df2['PetalWidth'][50:100])
plt.scatter(df2['PetalLength'][100:], df2['SepalLength'][100:], alpha=0.3, c=df2['Name'][100:], label='virginica', s=40*df2['PetalWidth'][100:])
plt.legend()
plt.xlabel('PetalLength')
plt.ylabel('SepalLength')
plt.title('comparing petal&sepal length of different iris flowers \n(sizes of the circle determine petal widths of the flowers)');
```

11

12

ผลลัพธ์ตัวอย่างการใช้งาน plt.scatter() และ parameter



12

13

Parameter: marker เปลี่ยนหน้าตาของ Marker

- สามารถเปลี่ยนหน้าตา marker ได้ตามที่ต้องการเพื่อความสวยงาม โดยการใส่และกำหนด Parameter: marker
- สัญลักษณ์หน้าตา marker ต่างๆ สามารถทำได้จาก
- https://matplotlib.org/stable/api/markers_api.html#module-matplotlib.markers
- ตัวอย่าง

+	+	plus
x	x	x
•	•	x (filled)
◆	◆	diamond
◆	◆	thin_diamond
		vline
—	—	hline

13

14

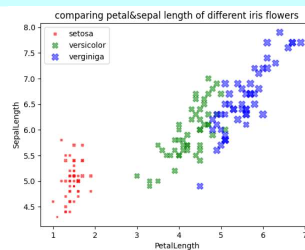
การใช้งาน Parameter: marker

- สามารถใช้สัญลักษณ์ที่ต้องการใน Parameter: marker ได้เลย เช่น marker='x'
 - ตัวอย่าง
- ```
plt.scatter(df2['Petal.Length'][:50], df2['Sepal.Length'][:50], alpha=0.5,
c=df2['Name'][:50], label='setosa', s=40*df2['Petal.Width'][:50], marker='x')
plt.scatter(df2['Petal.Length'][50:100], df2['Sepal.Length'][50:100], alpha=0.5,
c=df2['Name'][50:100], label='versicolor', s=40*df2['Petal.Width'][50:100],
marker='x')
plt.scatter(df2['Petal.Length'][100:], df2['Sepal.Length'][100:], alpha=0.5,
c=df2['Name'][100:], label='virginica', s=40*df2['Petal.Width'][100:], marker='x')
plt.legend()
plt.xlabel('Petal.Length')
plt.ylabel('Sepal.Length')
plt.title('comparing petal&sepal length of different iris flowers');
```

14

15

## ผลลัพธ์ตัวอย่างการใช้งาน Parameter: marker



15

16

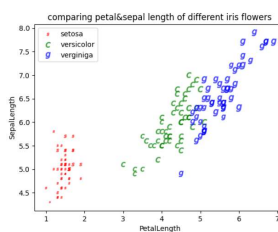
## สามารถกำหนดหน้าตาของ Marker ตามกลุ่มที่ต้องการได้

- ใช้สัญลักษณ์ที่ต้องการใน Parameter: marker แต่ละกลุ่มได้เลย เช่น
- ```
plt.scatter(df2['Petal.Length'][:50], df2['Sepal.Length'][:50], alpha=0.5, c=df2['Name'][:50], label='setosa', s=40*df2['Petal.Width'][:50], marker='gs')
plt.scatter(df2['Petal.Length'][50:100], df2['Sepal.Length'][50:100], alpha=0.5, c=df2['Name'][50:100], label='versicolor', s=40*df2['Petal.Width'][50:100], marker='gs')
plt.scatter(df2['Petal.Length'][100:], df2['Sepal.Length'][100:], alpha=0.5, c=df2['Name'][100:], label='virginica', s=40*df2['Petal.Width'][100:], marker='gs')
plt.legend()
plt.xlabel('Petal.Length')
plt.ylabel('Sepal.Length')
plt.title('comparing petal&sepal length of different iris flowers');
```

16

17

ผลลัพธ์ตัวอย่างกำหนด Marker ตามกลุ่ม



17

18