

Day 15 探索式數據分析

# Heatmap & Grid-plot



出題教練

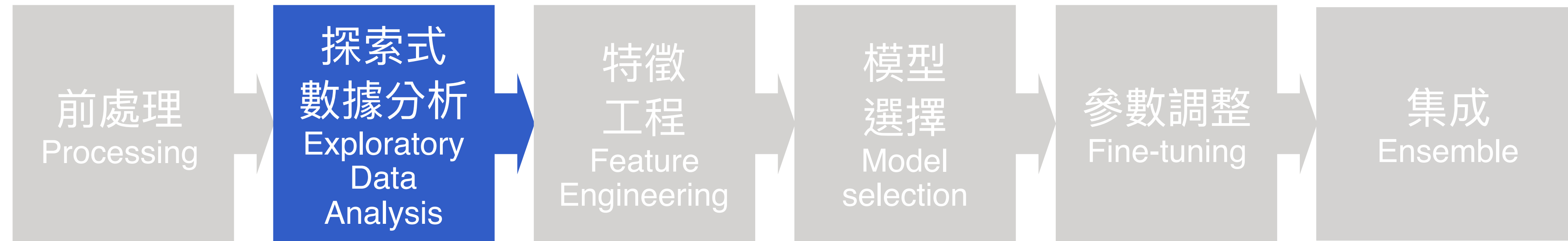
游為翔 / 杜靖愷

# 知識地圖 探索式數據分析 常用圖形 - HeatMap 與 GridPlot



## 機器學習概論 Introduction of Machine Learning

### 監督式學習 Supervised Learning

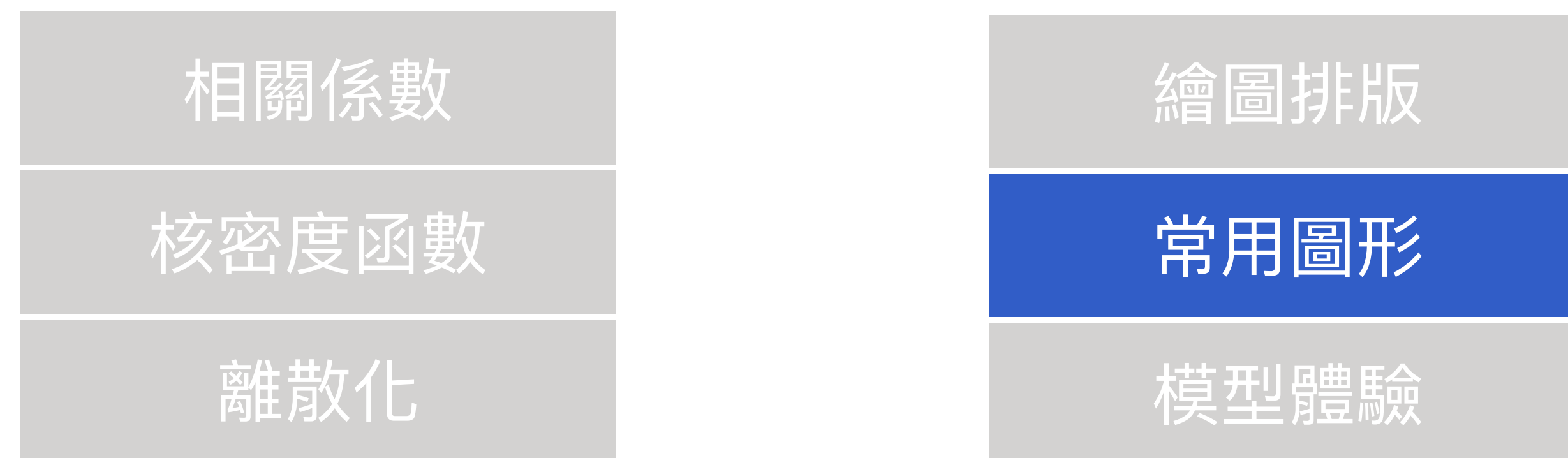


### 非監督式學習 Unsupervised Learning



## 探索式數據分析 Exploratory Data Analysis (EDA)

### 統計值的視覺化



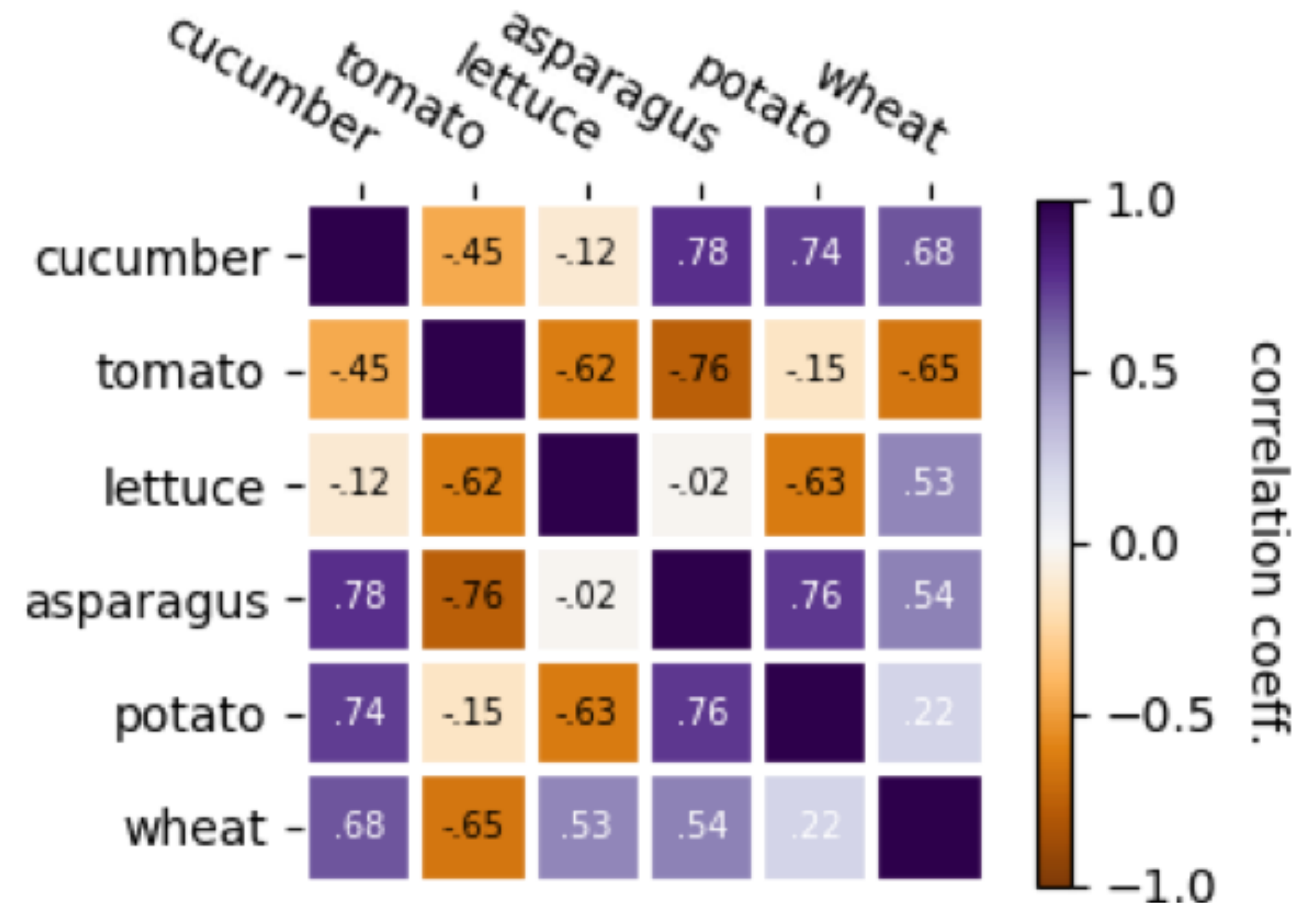


# 本日知識點目標

- 了解 heatmap 可以傳達的訊息
- 學會繪製 heatmap

# Heatmap

- 常用於呈現變數間的相關性
- 也可以用於呈現不同條件下，數量的高低關係



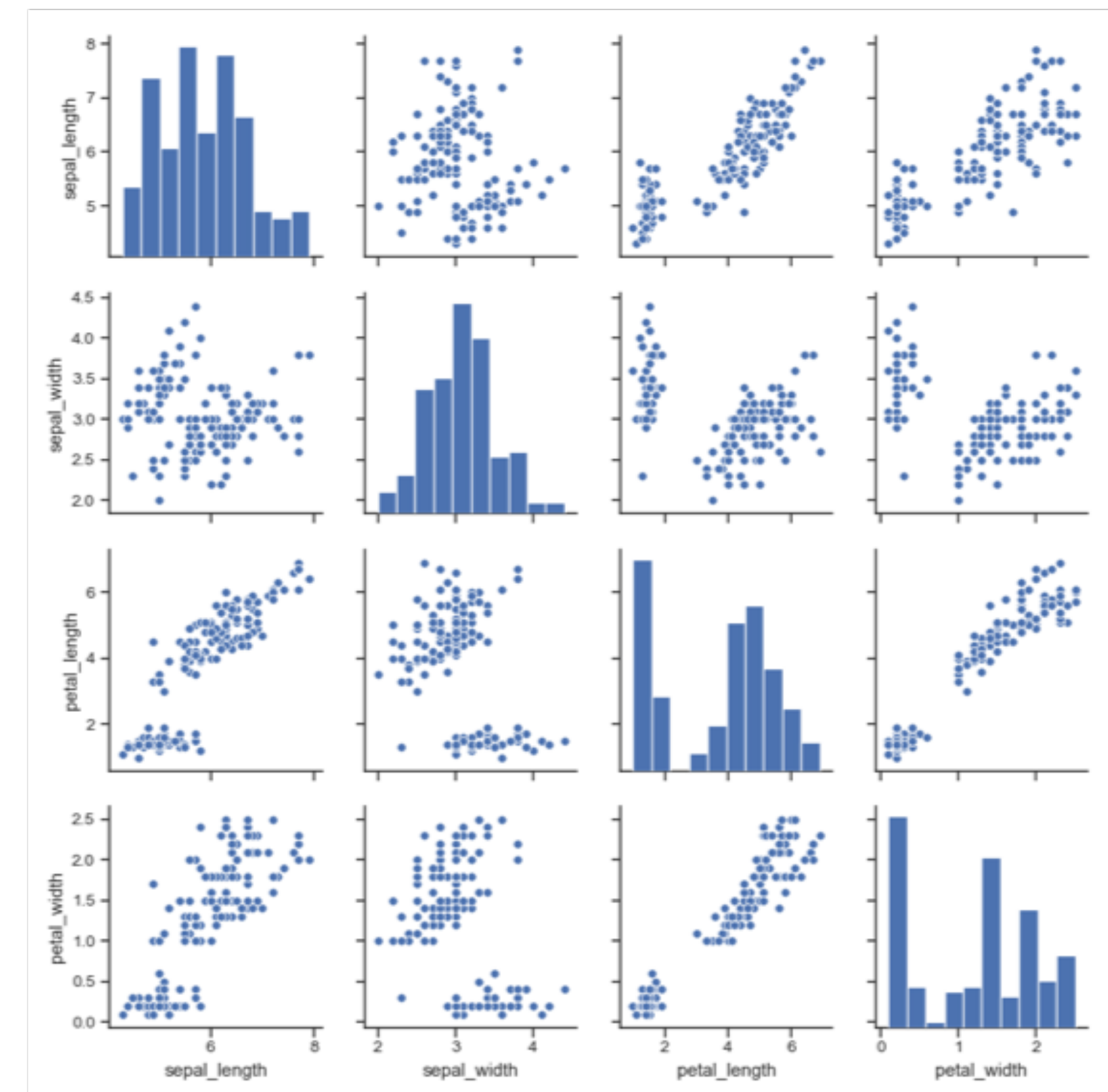
圖片來源: [matplotlib](https://matplotlib.org/)

# Gridplot

Subplot 的延伸, 但是 seaborn 幫你做的又美又好!

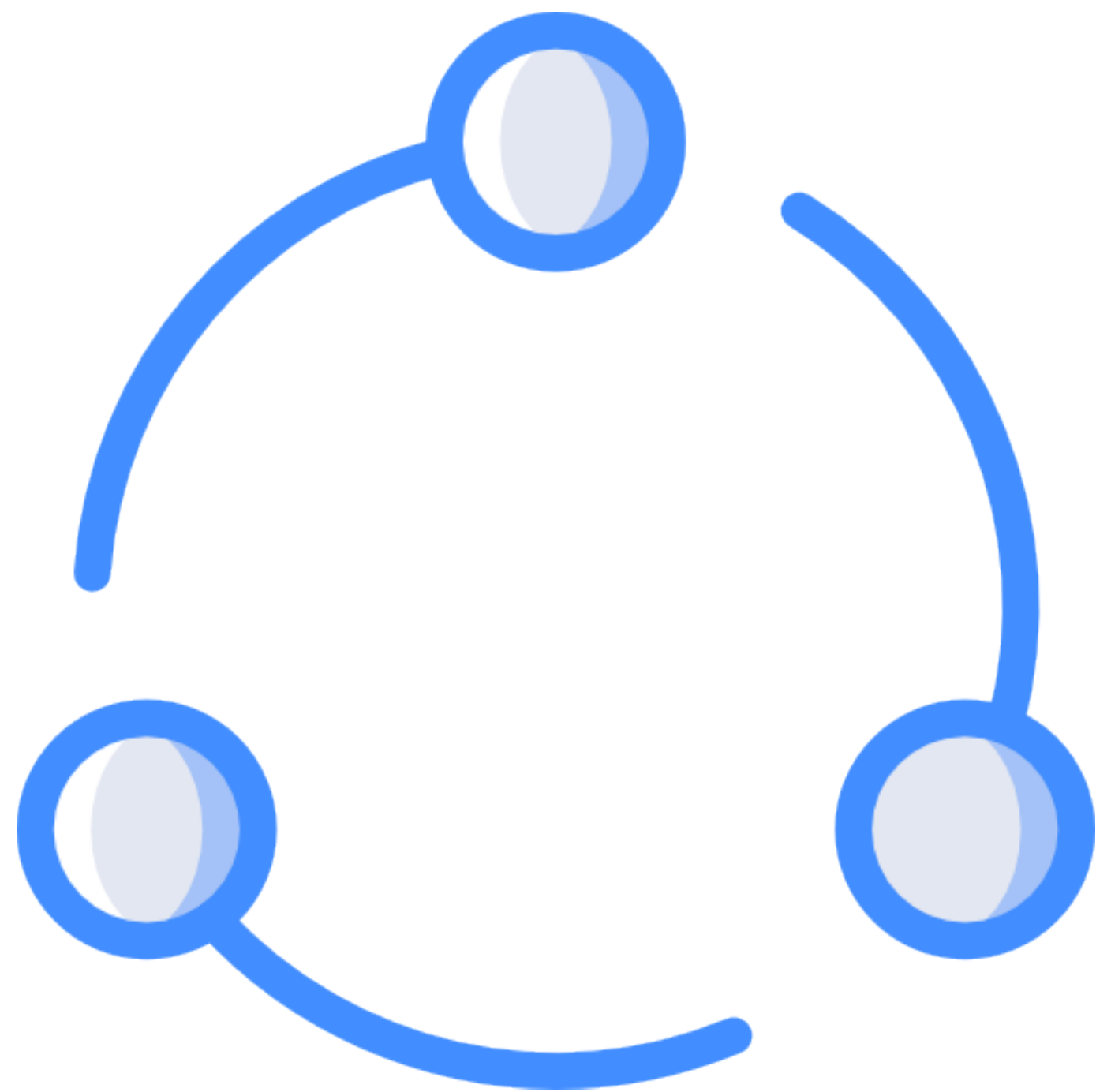
對角線：該變數的分布 (distribution)  
非對角線：兩兩變數間的散布圖

```
>>> import seaborn as sns; sns.set(style="ticks", color_codes=True)
>>> iris = sns.load_dataset("iris")
>>> g = sns.pairplot(iris)
```



圖片來源: [seaborn](https://seaborn.pydata.org/)





- Heatmap 常用於呈現訊息的強弱 (以顏色深淺呈現)，也常用於呈現混淆矩陣 (Confusion matrix, 後面的機器學習課程會再深入介紹)
- pair/gridplot 結合了 scatter plot 與 histogram 的好處來呈現變數間的相關程度

# 解題時間 It's Your Turn

請跳出PDF至官網Sample Code & 作業  
開始解題

