Day 19 探索式數據分析

Subplots

出題教練 游為翔 / 杜靖愷

知識地圖 探索式數據分析 繪圖排版 - 多圖檢視

CUPOY

機器學習概論 Introduction of Machine Learning

監督式學習
Supervised Learning

- 前處理 Processing
- 探索式數據分析 Exploratory Data Analysis
- 特徵工程 Feature Engineering
- 模型選擇 Model selection
- 參數調整 Fine-tuning
- 集成 Ensemble

非監督式學習
Unsupervised Learning

- 分群 Clustering
- 降維 Dimension Reduction

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

統計值的視覺化

- 相關係數
- 核密度函數
- 離散化
- 繪圖排版
- 常用圖形
- 模型體驗

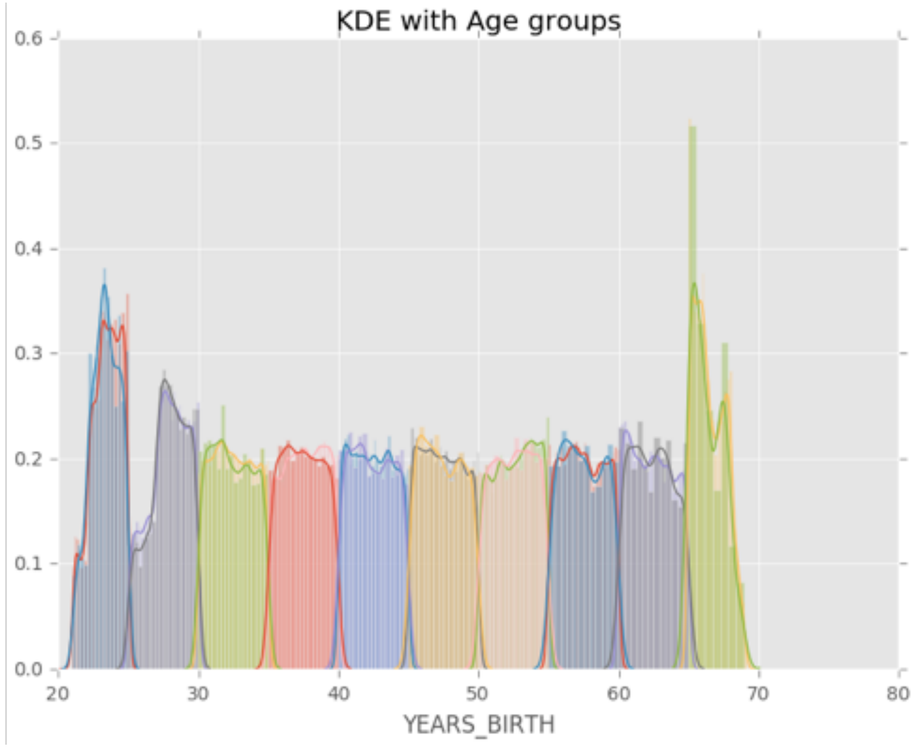
本日知識點目標

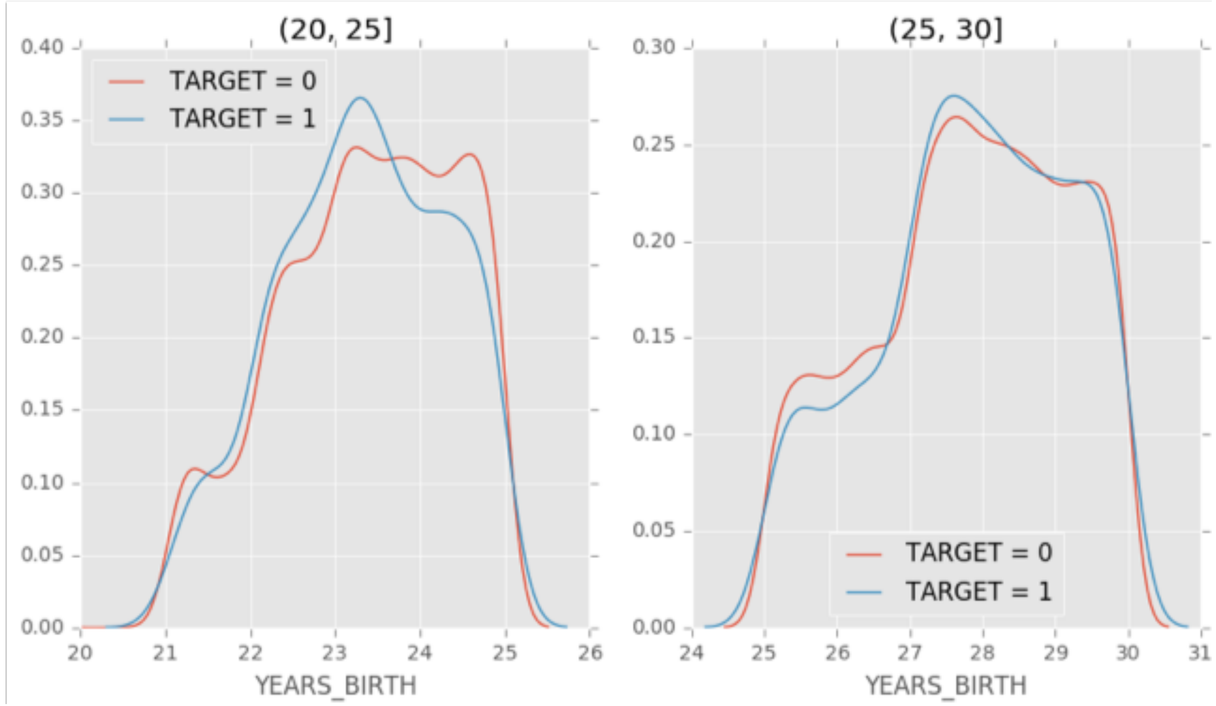
學會如何使用 subplots

Subplot - 讓圖可以適度地呈現

什麼時候該使用 subplot?

- 有很多相似的資訊要呈現時 (如不同組別的比較)
- 同一組資料，但想同時用不同的圖型呈現





Subplot

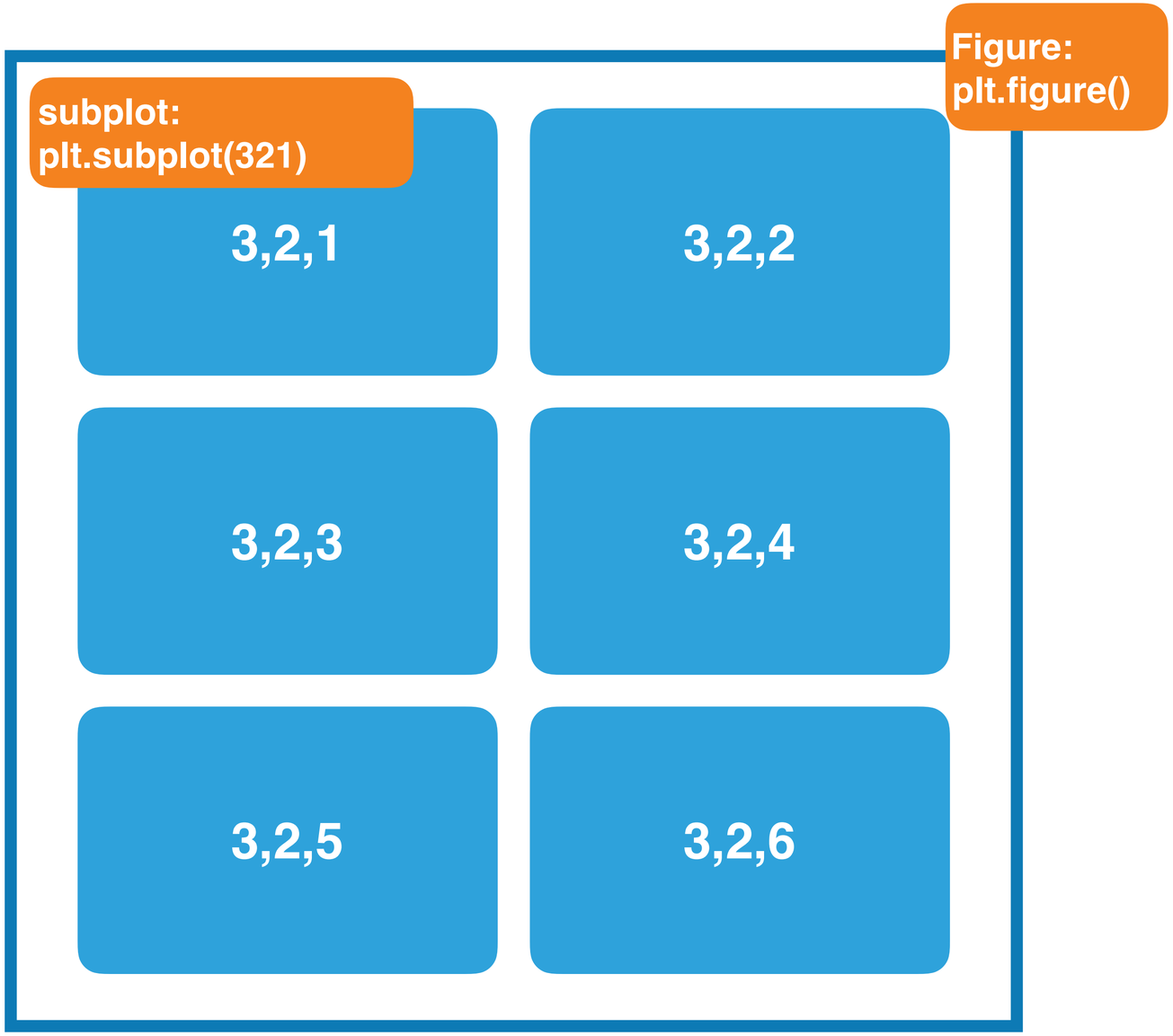
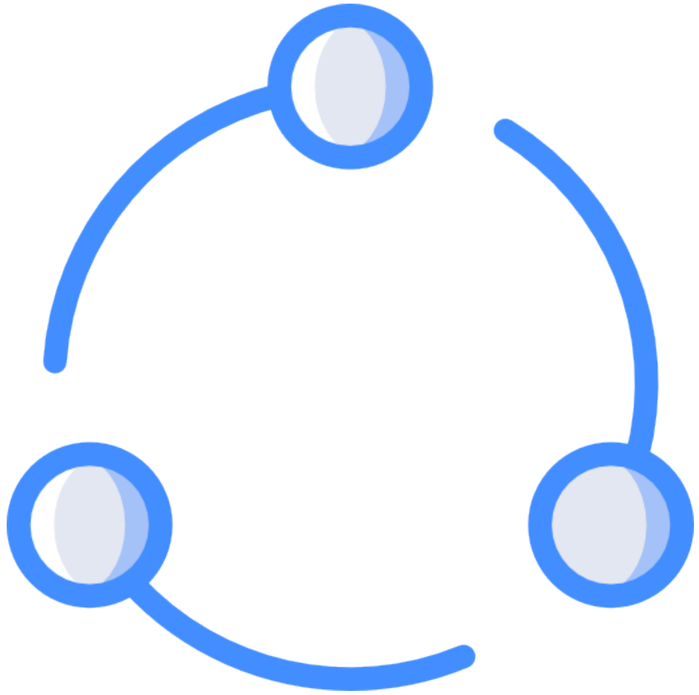


Figure:
plt.figure()
subplot:
plt.subplot(321)
3,2,1
3,2,2
3,2,3
3,2,4
3,2,5
3,2,6

重要知識點複習



- 適度地將圖片分格呈現，有助於資訊傳達
- 但過度使用 subplots 也將會使得資訊太多反而讓重點混淆
- Subplot 的坐標系 (列-欄-位置), 如
  - (321) 代表在一個 3列2欄的最左上角 (列1欄1)
  - (232) 為在一個 2列3欄的 (列1欄2) 位

延伸閱讀

除了每日知識點的基礎之外，推薦的延伸閱讀能補足學員們對該知識點的了解程度，建議您解完每日題目後，若有多餘時間，可再補充延伸閱讀文章內容。

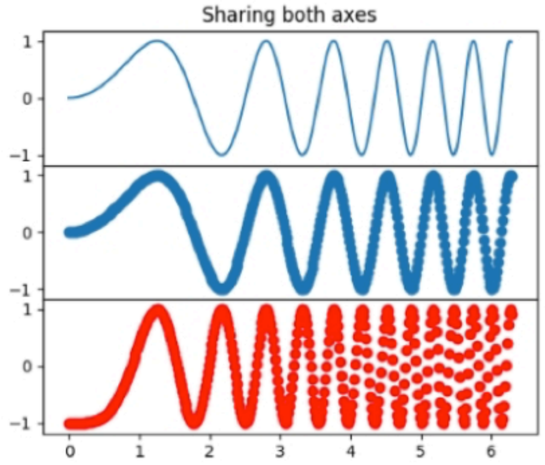
推薦延伸閱讀

Subplot

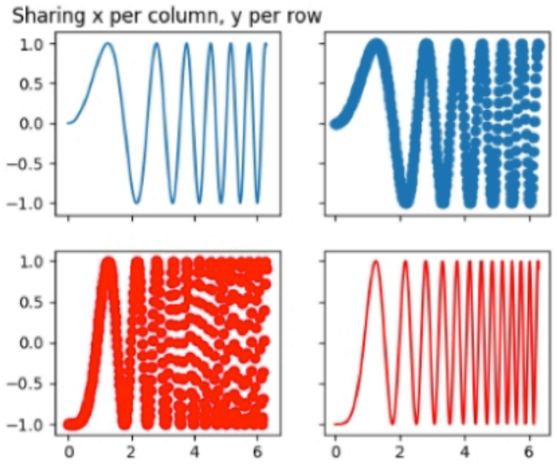
matplotlib 官方範例 網頁連結

- 這是 matplotlib 官方介紹 subplot 的排版範例，裡面有縱向並列，橫向並列等各式排版，還附上對應的語法
- 如果同學使用的方式並不複雜，可以在此直接查詢套用

(png, pdf)



(png, pdf)



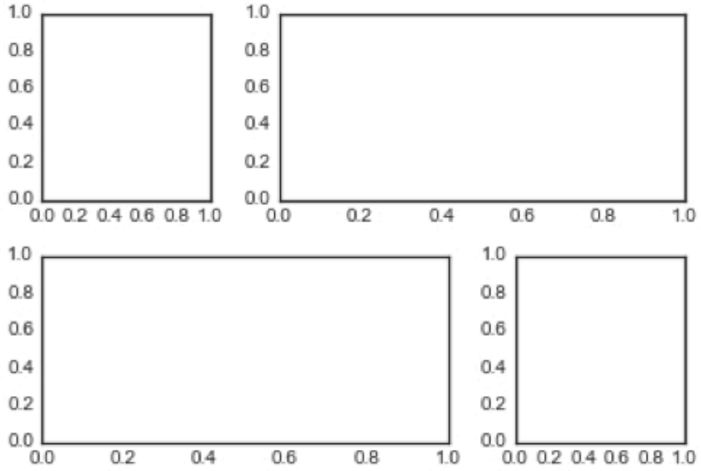
推薦延伸閱讀

複雜版 subplot 寫法

Python Data Science Handbook 網頁連結

- 如果需要複雜一點的 subplot，這個網頁有一些更詳細的寫法，也包含了下圖這種奇怪的排版方式
- 並不是要同學故意去學習這種奇怪的排版，而是當您有特殊的排版需求時，可以查詢這些內容，組合出您需要的排版語法。

```
In [9]: plt.subplot(grid[0, 0])
        plt.subplot(grid[0, 1:])
        plt.subplot(grid[1, :2])
        plt.subplot(grid[1, 2]);
```

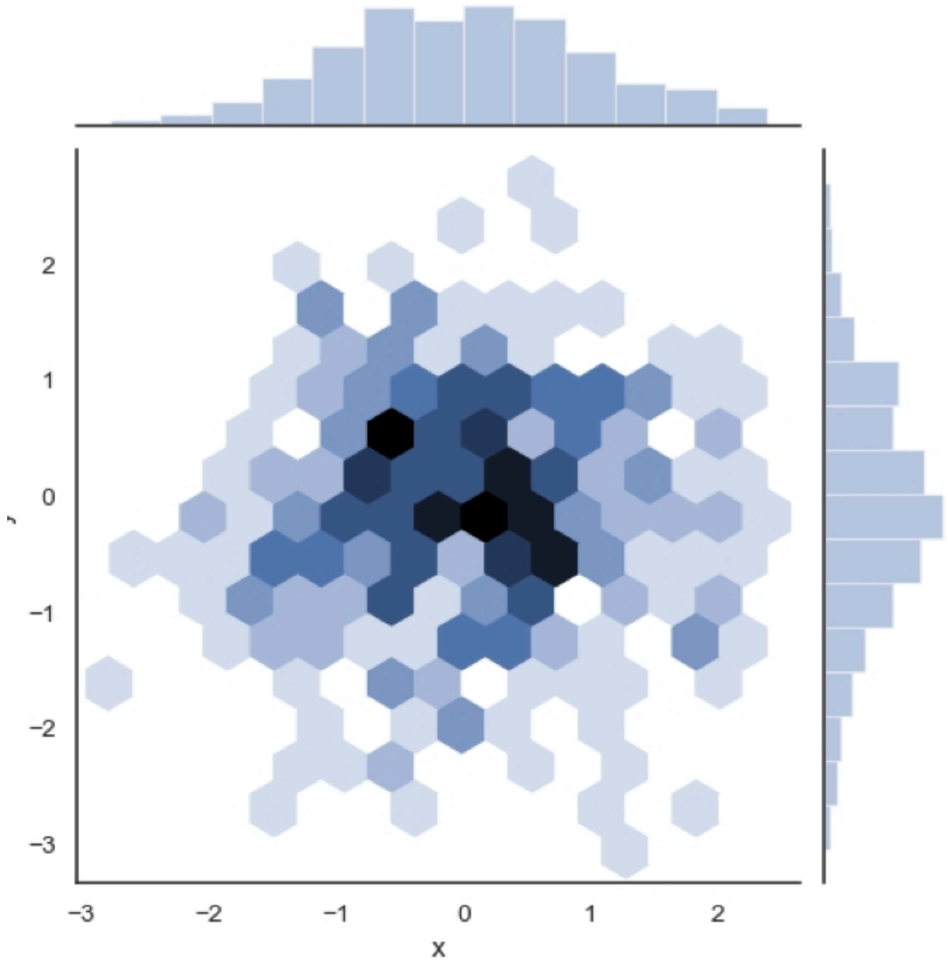


推薦延伸閱讀

另類子圖 Seaborn.jointplot

官方說明文件 網頁連結

- 除了圖本身的排版，搭配排版的圖組，有時也是不錯的選擇
- 例如最經典的 Seaborn.jointplot，除了繪製兩個變數間的散佈圖外，變數本身的分布長條圖，也會列在對應的軸上，讓人一目了然，中央的散佈圖也有點狀 / 等高線 / 蜂巢等不同選擇
- EDA 的重點是讓人看懂資料，因此圖形是否直覺是相當重要的一環



解題時間

It’s Your Turn

請跳出PDF至官網Sample Code＆作業
開始解題