emboth AI共學社群 我的 AI共學社群 > Python資料科學 > D20 使用 Seaborn 進行資料視覺化

問題討論

學習心得(完成)

## D20 使用 Seaborn 進行資料視覺化

範例與作業

簡報閱讀

知識點回顧

延伸閱讀



可視化數據集的分佈

重要知識點 • 了解資料視覺化應用 • 完成今日課程後你應該可以了解 # 如何安裝並使用seaborn # 結合 Matplotlib & seaborn

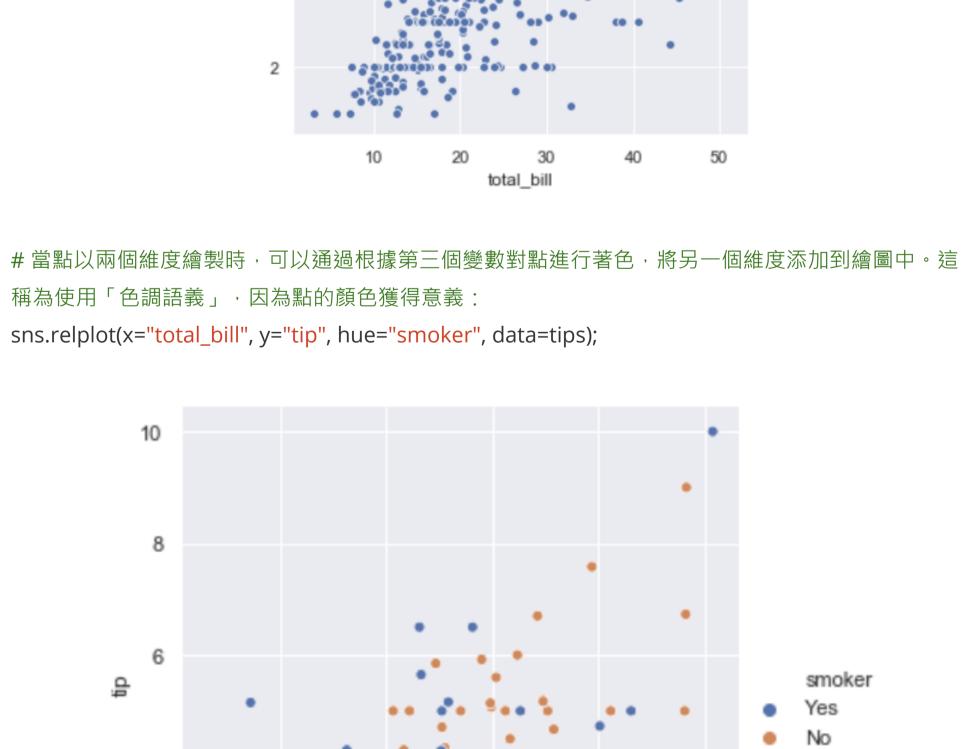
- seaborn:統計數據可視化
  - conda install seaborn pip install git+https://github.com/mwaskom/seaborn.git

pip install seaborn

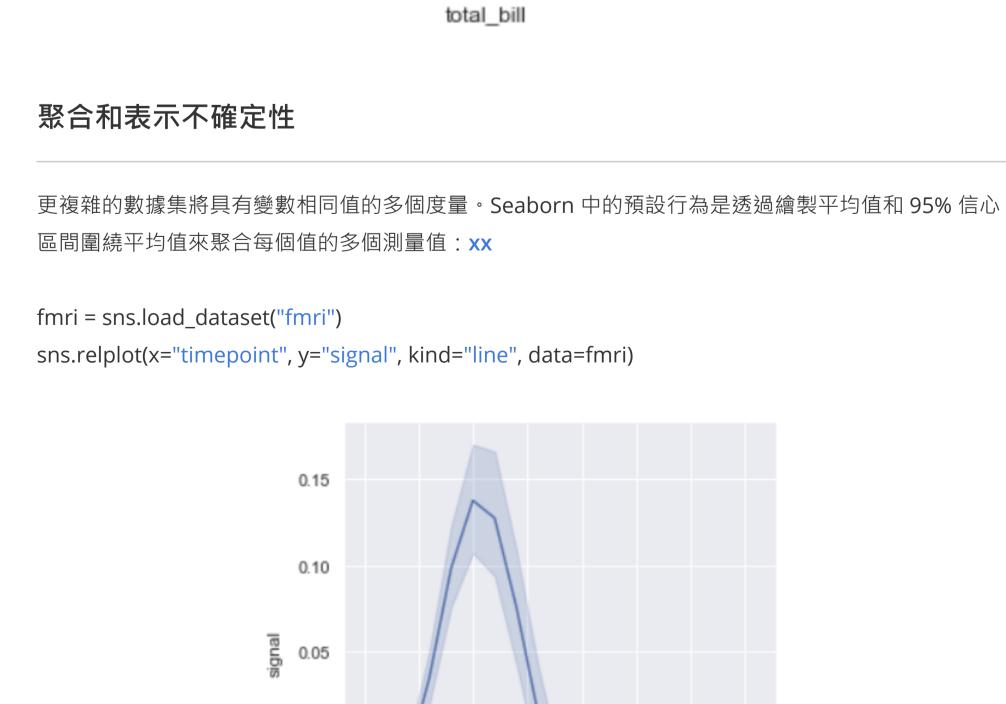
- 功能: • 可視化統計關係 • 構建結構化多繪圖網格 • 繪圖美學
- Seaborn 圖樣式 有五種預設seaborn主題:darkgrid,whitegrid,dark,white,和ticks。它們各自適合不同的應用和
- 個人喜好。默認主題是darkgrid sns.set\_style("whitegrid") # 設定圖形樣式 #設定圖示的高與寬 data = np.random.normal(size=(20, 6))
- #列印圖示 sns.boxplot(data=data);

+ np.arange(6) / 2

- -2
  - 可視化統計關係 • 統計分析是一個了解數據集中變數之間如何相互關聯以及這些關係如何依賴於其他變數的過 程。可視化可能是此過程的核心元件,因為當數據正確可視化時,人類可視化系統可以看到指示關 係的趨勢和模式。 ● 將變數與散點圖關聯 散點圖是統計可視化的支柱。它使用點雲描繪了兩個變數的聯合分佈,其中 每個點表示數據集中的觀測值。這種描述允許眼睛推斷出大量有關它們之間是否有任何有意義 的關係的資訊。 # 散點圖關聯變量 sns.set(style="darkgrid") sns.relplot(x="total\_bill", y="tip", data=tips);



ŧр



20

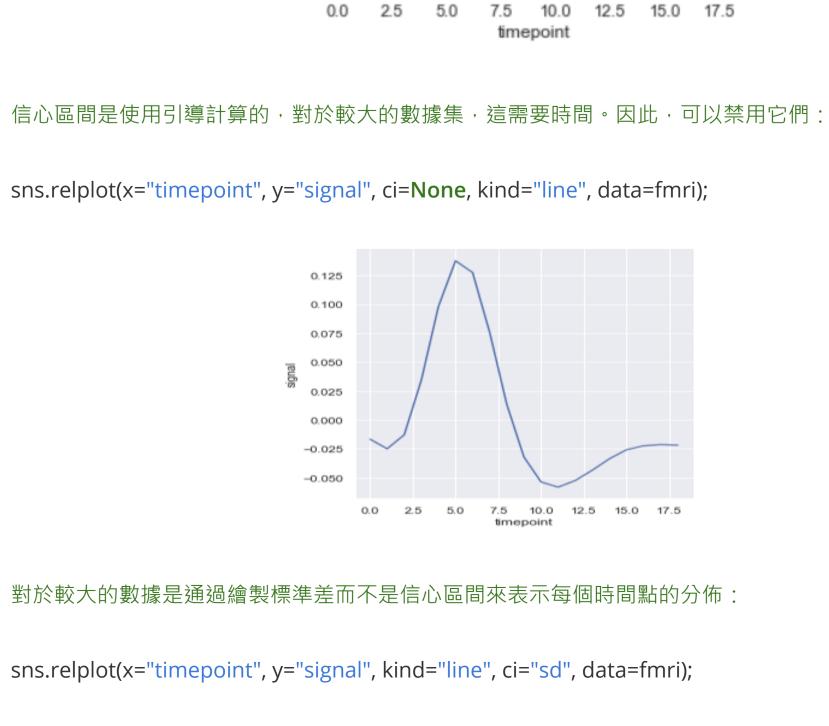
10

0.00

-0.05

40

50



-0.10 0.0 2.5 5.0 7.5 10.0 12.5 15.0 17.5 timepoint 使用語義映射繪製數據子集

sns.relplot(x="timepoint", y="signal", hue="event", kind="line", data=fmri);

0.25

0.2

0.1

0.3

0.2

kind="line", data=fmri);

現。

重要

區間:xyy~x

度。

sns.set(color\_codes=**True**)

sns.distplot(x);

知識點回顧

• 分類估計圖

延伸閱讀

x = np.random.normal(size=100)

sns.distplot(x, bins=20, kde=**False**, rug=**True**);

import seaborn as sns

sns.set(color\_codes=**True**)

tips = sns.load\_dataset("tips")

sns.regplot(x="total\_bill", y="tip", data=tips);

import matplotlib.pyplot as plt

Lineplot() 具有與 Scatterplot() 相同的靈活性:它可以通過修改繪圖元素的色調、大小和樣式來顯示最

0.25

0.20

0.15

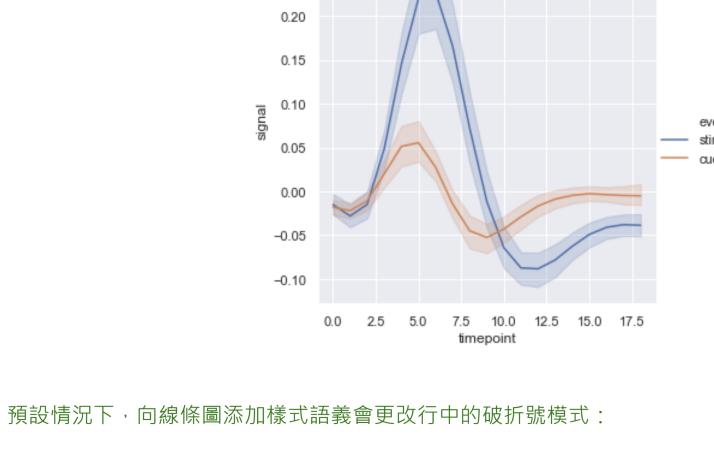
0.10

0.05

0.00

-0.05

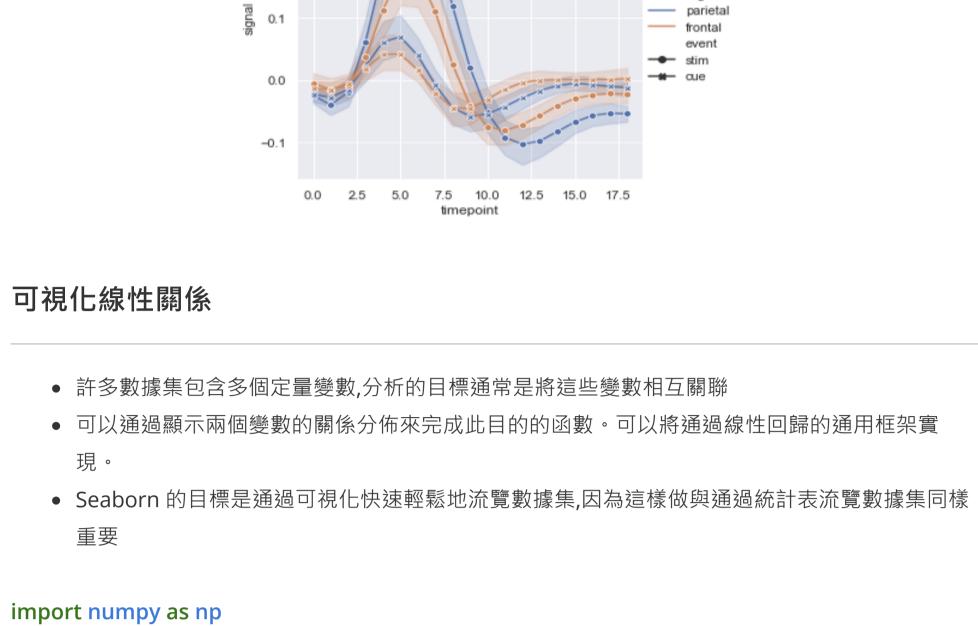
多三個附加變數

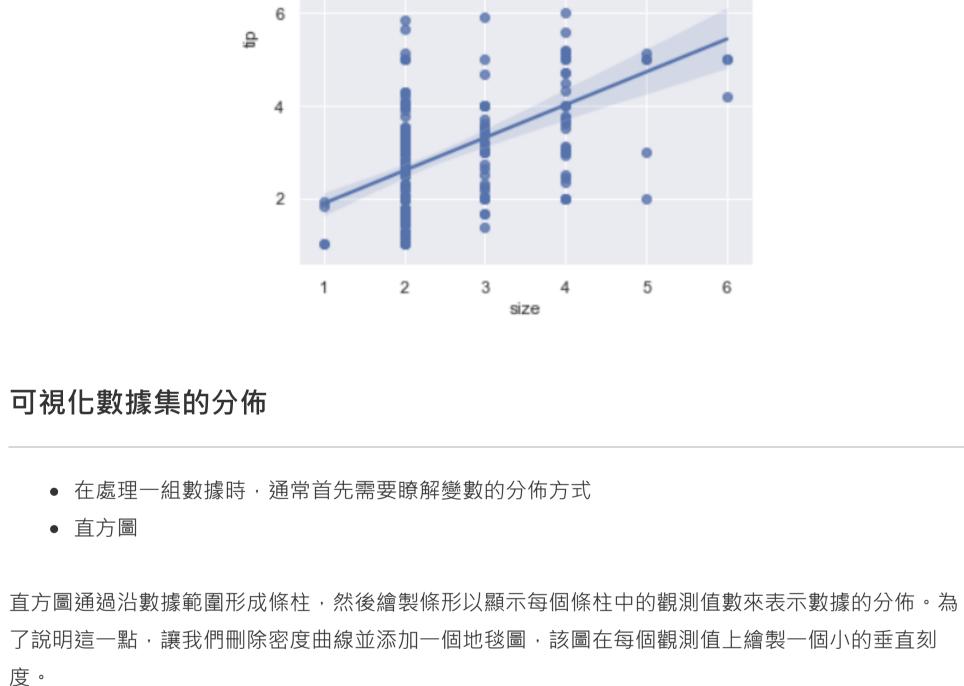


sns.relplot(x="timepoint", y="signal", hue="region", style="event", kind="line", data=fmri);

0.0 2.5 5.0 7.5 10.0 12.5 15.0 17.5

timepoint 通過每個觀測中使用的標記來標識子集,或者與破折號一起識別子集,或者代替它們: sns.relplot(x="timepoint", y="signal", hue="region", style="event", dashes=**False**, markers=**True**,





調用中,兩個函數繪製兩個變數的散點圖,然後擬合回歸模型並繪製生成的回歸線和該回歸的 95% 置信

sns.lmplot(x="size", y="tip", data=tips);

## import numpy as np import pandas as pd import seaborn as sns import matplotlib.pyplot as plt from scipy import stats

況下正確的數位,但嘗試更多或更少的 bin 可能會顯示資料中的其他特徵:

繪製直方圖時,主要選擇是要使用的條柱數和放置位置。<u>distplot()</u>使用簡單的規則來正確猜測預設情

0.3 0.2 0.1

- 在關係圖教程中,我們瞭解如何使用不同的視覺表示來顯示數據集中的多個變數之間的關係。在示例 中,我們重點介紹了兩個數值變數之間的主要關係。如果其中一個主要變數是"分類"(分為離散組),則 使用更專業化的可視化方法可能會有所説明。 在 Seaborn 中,有幾種不同的方法來可視化涉及分類數據的關係。有許多軸級函數用於以不同的方式 繪製分類數據,以及一個圖形級介面 $\underline{\mathsf{catplot}}$ ,該函數提供對它們的統一更高級別的訪問。 將不同的分類情節類型視為屬於三個不同的類別會很有幫助。它們是: • 分類散點圖 • 分類分布圖
- 使用Seaborn Python包進行可視化 網站: <u>Visualization Using Package Seaborn Python</u> ● Seaborn是一個統計繪圖庫
- 它也被設計為與Pandas數據框對像一起很好地工作。 • 使用散點圖關聯變量 • 用線圖強調連續性 • 使用語義映射繪製數據子集
- Putri Choirunisa 4 Followers · About Follow

• 它具有漂亮的默認樣式

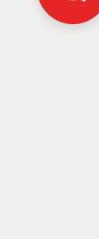
- **Visualization Using Package Seaborn Python**
- Putri Choirunisa Jan 15 · 6 min read Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative

• It has beautiful default styles

- statistical graphics **Keys Features** • Seaborn is a statistical plotting library
  - It also is designed to work very well with Pandas dataframe objects. LETS TRY!!!! In this case using the tips data in the Seaborn library. first step install and getting started. To install the latest release of seaborn, you can use pip:

 $\Box$ 

It's also possible to install the released version using conda: conda install seaborn



pip install seaborn