

## D4：EDA/讀取資料與分析流程

[PDF 下載](#)[全螢幕](#)

## Sample Code & 作業內容

作業：

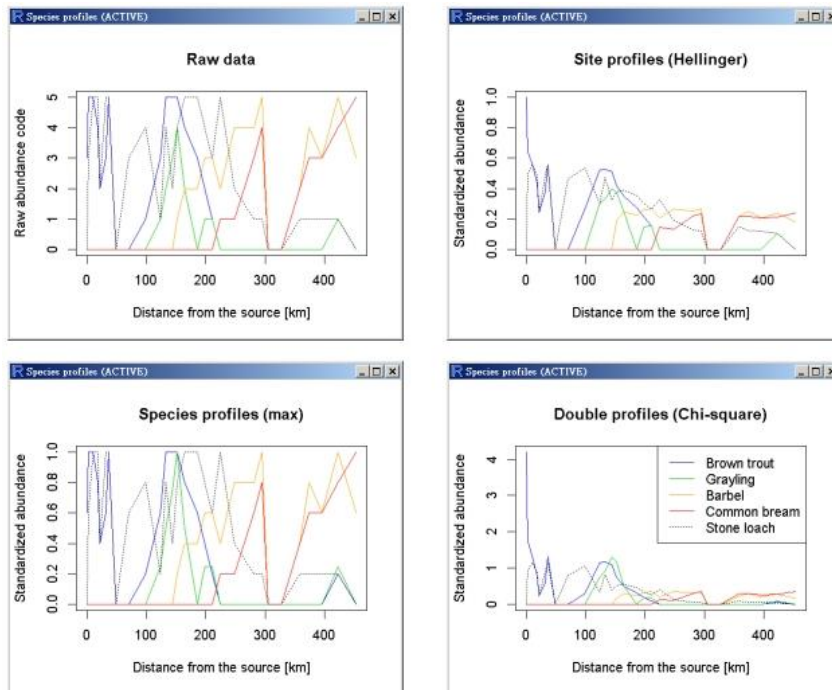
- 請下載本次馬拉松建議的 [Kaggle 資料](#) 或準備好自己的資料
- 如採用 Kaggle 資料，請通過 HomeCredit\_columns\_description.csv，了解各個欄位的意義 (csv檔可從檢視範例內的 data資料夾下載)
- 請點下方檢視範例參考 Day\_004\_first\_EDA.ipynb 檔案，了解如何讀取資料並可自行嘗試對資料進行操作，作業檔請提交Day\_004\_HW.ipynb

[檢視範例](#)

吳漢銘老師 [網頁連結](#)

這是吳老師講解探索式資料分析(EDA)的內容，比較側重於理論部分，同學可以在這裡看到許多豐富的資料，對照我們後續的課程內容，可以讓您更了解EDA的全貌。

建議同學可以閱讀自己有興趣的部分即可，有必要了解的細節，我們會在後續課程中提到。



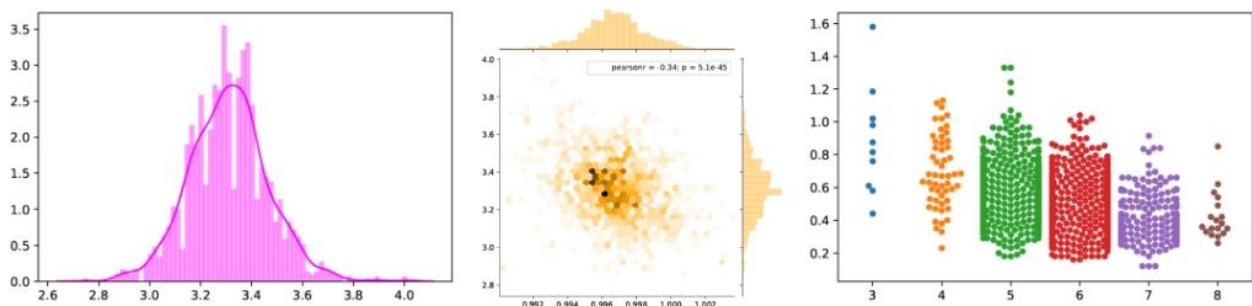
## What is Exploratory Data Analysis?

towards data science Prasad Patil

[網頁連結](#)

這是另一份 EDA 的教學，比較側重於圖形與對應的程式碼，比較接近我們後續的課程形式

從這些範例中，我們可以看到有分布圖(左)，蜂窩聯合圖(中)，分類散佈圖(右)，這些圖形都可以輕鬆藉由 **seaborn** 套件繪製，所以只要我們在後續課程中學會這些，就可以輕鬆完成資料視覺化。



## What is Exploratory Data Analysis?



Prasad Patil [Follow](#)

Mar 24, 2018 · 6 min read



## 提交作業

請將你的作業上傳至 Github，並貼上該網網址，完成作業提交

確定提交

[如何提交](#) ▾

## 熱門問答

[A](#) [Aibo](#) · 2019.09.10 · 1 回答

### [如何善用Github的討論區](#)

[請問對於Python資料分析&新聞檢索文字探勘在Github有沒有比較推薦可以參考的討論區？感謝~](#)

[朱武勳](#) · 2019.09.07 · 3 回答

### [d4 FileNotFoundError](#)

[f\\_app = os.path.join\(dir\\_data, 'application\\_train.csv'\)print\('Path of read in data: %s' % \(f\\_app\)\)app\\_train = pd.read\\_csv\(f\\_app\)在jupyter上運行時顯示FileNotFoundError請問該怎麼辦？](#)

[J](#) [Jen](#) · 2019.09.04 · 2 回答

### [請問總數是否要加1?\(因為起示為0\)](#)

[總行與列是否須要加1?](#)

[查看更多](#)

## 到 Cupoy 問答社區提問，讓教練群回答你的疑難雜症

[向專家提問](#)

[如何提問](#) ▾

