

100 道題目 > D23：數值型特徵 - 去除偏態

D23：數值型特徵 - 去除偏態

Day 23 特徵工程

數值型特徵-去除偏態

陳明佑
出題教練

[PDF 下載](#)[全螢幕](#)

Sample Code & 作業內容

請參閱作業範例Day_023_Reduce_Skewness.ipynb：鐵達尼生存預測

作業1：試著在鐵達尼的票價 (Fare) 欄位中使用對數去偏 (log1p)，結果是否更好？

作業2：最後的 boxcox 區塊直接執行會造成錯誤，起因為輸入值有負值，請問如何修正後可以使用 boxcox? (Hint：試圖修正資料)

作業請繳交：Day_023_HW.ipynb

[檢視範例](#)

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

[陳 陳美心](#) · 2019.09.22 · 2 回答

[HW_023 boxcox 寫法問題](#)

在[Day_023_Reduce_Skewness](#)中，使用的程式碼是`f_fixed['LotArea'] = stats.boxcox(df['LotArea'], lmbda=0.5)`這邊我理解boxcox的參數是0.5，對應的就是sqrt但，在[Day_023_HW](#)中，使用的程式碼是...

[汪 汪君優](#) · 2019.09.22 · 2 回答

[DAY_23 作業二 欄位中的負數處理](#)

我的程式碼如上反白處。我想不透是甚麼問題導致它跑出以下錯誤：`cannot reindex from a duplicate axis`不知這題有沒有更好的處理方法？

[廖 廖崇嶂](#) · 2019.09.21 · 1 回答

[D23 的講義關於box-cox問題？](#)

講義最後有提到使用box-cox， λ 參數要介於0 到0.5 之間，那為甚麼講義中的轉換表 λ 的值會在-2到2之間

[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](#) ▾