

D15: EDA from Correlation



Day 15 探索式數據分析

相關係數實作

</>

Coding 練習日

游為翔 / 杜靖愷
出題教練

Eupoy

[PDF 下載](#)[全螢幕](#)

Sample Code & 作業內容

作業目標：

請同學試著使用 `pandas.corr()` 這個函數來顯示相關係數並加以觀察結果

思考1：使用 `pandas` 有沒有什麼寫法，可以顯示欄位中最大的幾筆，以及最小幾筆呢？(Hint: 排序後列出前幾筆/後幾筆)

思考2：試著使用散佈圖，顯示相關度最大/最小的特徵與目標值的關係，如果圖形不明顯，是否有調整的方法？

參考範例程式碼 `Day_015-supplementary_correlation_and_plot_with_different_range.ipynb`

作業請提交 `Day_015_HW`

[檢視範例](#)

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

[嚴 嚴詩豪](#) · 2019.09.25 · 2 回答

[Day 15 計算數值最大及最小各15個?](#)

[依照題目](#)，我需要計算這些欄位的最大最小值各15個但我想了很久依然沒有頭緒該如何計算我的想法是將 `app_train corr['TARGET']` 儲存成一個 DataFrame 然後將這個 DataFrame 賦予兩個欄位 'Variable', 'Values' 之...

[黃 黃世安](#) · 2019.09.18 · 1 回答

[Day_015_HW](#)

[請問經過LabelEncoder後物甚麼CODE_GENDER沒有變呢](#)

[林 林敬堯](#) · 2019.09.11 · 2 回答

[correlation相關問題](#)

1. 資料裡的兩個 features, 如果高度相關的時候, 我們可能會無法解釋哪一個 feature 對 regression model 比較重要, 所以需要拿掉一個 feature, 但如果我們只在意預測的準確性, 是否可以不用在意 data 裡有沒有高度相...

[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](#) ▾

