

[100 道題目](#) > [D12：數值型特徵-補缺失值與標準化](#)

## D12：數值型特徵-補缺失值與標準化

**Day 12 機器學習前處理**

**數值型特徵-補缺失值與標準化**

**陳明佑**  
出題教練

[PDF 下載](#)[全螢幕](#)

### Sample Code & 作業內容

請參考範例程式碼 `Day_012_Fill_NaN_and_Scalers.ipynb`: 房價預測

執行作業範例 `Day_012_HW.ipynb`: 鐵達尼生存預測

作業1: 試著在補空值區塊, 替換並執行兩種以上填補的缺值, 看看何者比較好?

作業2: 使用不同的標準化方式 (原值 / 最小最大化 / 標準化), 搭配羅吉斯迴歸模型, 何者效果最好?

作業請提交 `Day_012_HW.ipynb`

[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

[h](#) [herohtsu](#) · 2019.09.19 · 2 回答

**day12 作業一的疑問** 算出來和答案不同請問哪裡有誤

# 空值補 -1, 做羅吉斯迴歸

```
df_m1 = df.fillna(-1) train_X = df_m1[:train_num] estimator = LogisticRegression(solver='lbfgs', multi_class='multinomial')
```

[李](#) [李衡](#) · 2019.09.12 · 3 回答

**不是很懂訓練/預測用的格式這樣撰寫的原因?**

```
# 重組資料成為訓練 / 預測用格式 train_Y = np.log1p(df_train['SalePrice']) ids = df_test['Id'] df_train = df_train.drop(['Id', 'SalePrice'], axis=1) df_test = df_test.drop(['Id'], axis=1) df = pd.concat([df_train, df_test])
```

[黃](#) [黃育仁](#) · 2019.09.10 · 3 回答

**Day012 填補空值做交叉驗證**

請問填補空值做交叉驗證為什麼會出現下圖紅色底圖的文字呢?這是正常現象嗎?我用-1,0,平均值填補都會出現

[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](#) ▾