

100道題目 > D4 : EDA: 欄位的資料類型介紹及處理

## D4 : EDA: 欄位的資料類型介紹及處理

全螢幕瀏覽

主 資 專

頁 訊 區

↓ PDF 下載

### Sample Code & 作業內容

A、在初步 EDA 的過程，我們無可避免會想問的問題

- 不同資料類型各有多少個欄位？
- 類別型欄位 (pandas 中的 object) 的類別數量？
- 模型怎麼處理類別型的資料？有什麼表示方法？



Label encoding: 把每個類別 mapping 到一個整數，通常會增加一個欄位

- One Hot encoding: 為每個類別新增一個欄位，用 0/1 表示是否

Label Encoding

Food Name	Categorical #	Calories
Apple	1	95
Chicken	2	231
Broccoli	3	50

One Hot Encoding

Apple	Chicken	Broccoli	Calories
1	0	0	95
0	1	0	231
0	0	1	50

專  
區

C、作業1：請詳細閱讀[label-encoder-vs-one-hot-encoder](#)

D、作業2：請點檢視範例，作業請提交Day\_004\_HW.ipynb檔

↓ 檢視範例

## 參考資料

課後補充:

**Pandas 數據類型**

提交作業

