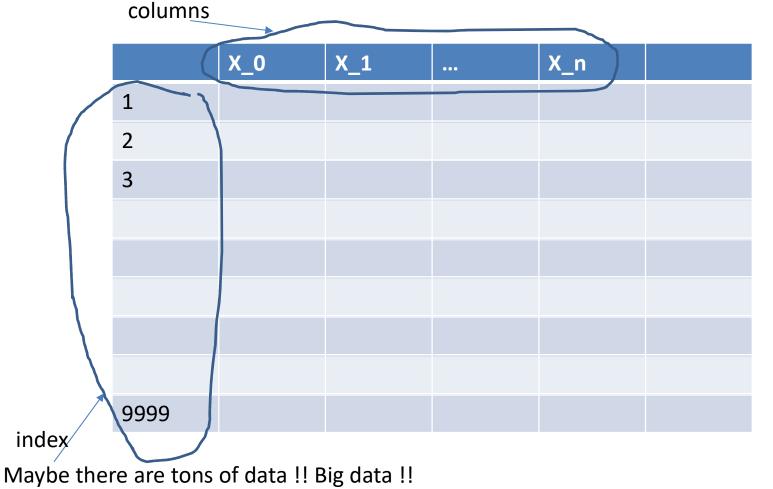
Project

- 3 種主要資料格式: html, json, csv
- 政府資料平台(json, csv): https://data.gov.tw/
 - 腸病毒資料集: FGU-Class/Project/Health_data_0.ipynb
 - 大家可以任選資料集....
- Preprocessing: 清理, 轉換
- Visualization: 折線圖, 長條圖,....
- 我們可以從資料中學習到什麼?
 - Learning: regression or classification
 - Classification: e.g. spam filter
 - Regression: e.g. 房價 股價 predictor
 - google: Hello World Machine Learning Recipes
 - google: The 7 Steps of Machine Learning

Goal: From dataset to generate the following DataFrame

There are n data features:





全部資料集

互動專區▼ 最新消息

諮詢小組

關於平臺▼ ENGL

首頁 » 資料集 » 健保門診及住院就診人次統計-陽病毒

健保門診及住院就診人次統計-腸病毒

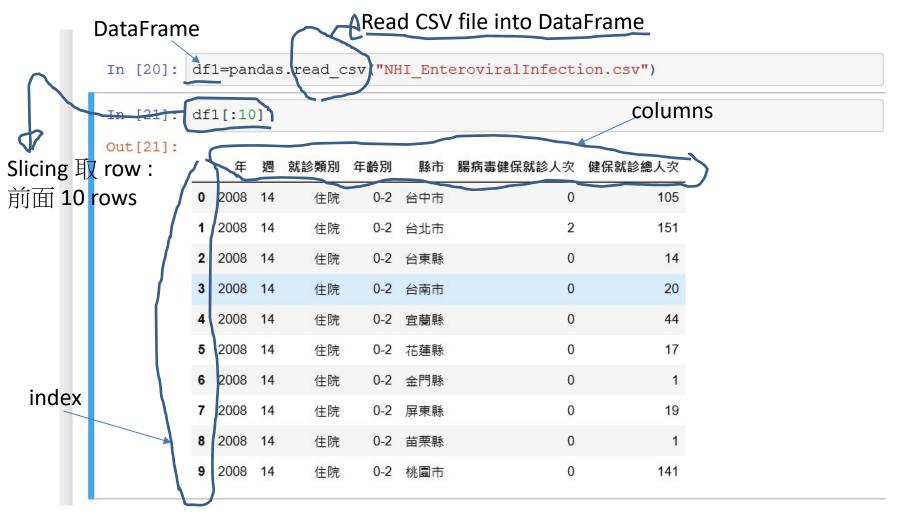
*** 資料集評分: 平均 3 (2 人次投票) 資料集描述: 各縣市、年齡別、年週之陽病毒門診及住院就診人次統計 年、週、就診類別、年齡別、縣市、**陽**病毒健保就診人次 主要欄位說明: 資料資源: ○ 檢視資料 陽病毒 CSV 格式-◯ 檢視資料 腸病毒 JSON 🖾 提供機關: 衛生福利部疾病管制署 提供機關聯絡人: 王 先生 (02-23959825#4032) 更新頻率: 每月 授權方式: 政府資料開放授權條款-第1版 計費方式: 免費 上架日期: 2015/05/01

https://data.gov.tw/dataset/14590

年,週,就診類別,年齡別,縣市,腸病毒健保就診人次,健保就診總人次2008,14,住院,0-2,台中市,0,1052008,14,住院,0-2,台北市,2,1512008,14,住院,0-2,台南市,0,202008,14,住院,0-2,官蘭縣,0,442008,14,住院,0-2,花蓮縣,0,172008,14,住院,0-2,金門縣,0,12008,14,住院,0-2,蘇東縣,0,192008,14,住院,0-2,輔票縣,0,12008,14,住院,0-2,桃園市,0,1412008,14,住院,0-2,蔣康市,0,21

PCSV 格式

逗號分隔值(Comma-Separated Values, CSV, 有時也稱為字元分隔值,因為分隔字元也可以不是逗號),其檔案以純文字形式儲存表格資料(數字和文字)。(維基百科,自由的百科全書



FGU-Class/Project/Health_data_0.ipynb

• 取欄:

- ex: df1['縣市']

```
df1.info()
In [22]:
                   總共有 111319 筆資料
        <class 'pandas.core.frame.DataFrame'>
        RangeIndex: (111319 entries, 0 to 111318
        Data columns (total 7 columns):
        年
                    111319 non-null int64
                    111319 non-null int64
        就診類別
                     111319 non-null object
        年齡別
                    111319 non-null object
        縣市 111319 non-null object
        腸病毒健保就診人次 111319 non-null int64
        健保就診總人次
                   111319 non-null int64
        dtypes: int64(4), object(3)
```

memory usage: 5.9+ MB