

## Sample Code & 作業內容

作業 1：

請上 **Kaggle**, 在 **Competitions** 或 **Dataset** 中找一組競賽或資料並寫下：

1. 你選的這組資料為何重要
2. 資料從何而來 (tips: 譬如提供者是誰、以什麼方式蒐集)
3. 蒐集而來的資料型態為何
4. 這組資料想解決的問題如何評估

作業 2：

想像你經營一個自由載客車隊，你希望能透過數據分析以提升業績，請你思考並描述你如何規劃整體的分析/解決方案：

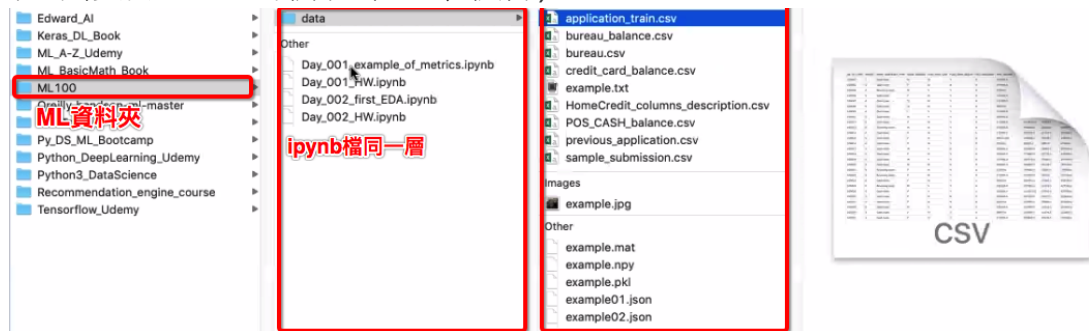
1. 核心問題為何 (tips: 如何定義「提升業績 & 你的假設」)
2. 資料從何而來 (tips: 哪些資料可能會對你想問的問題產生影響 & 資料如何蒐集)
3. 蒐集而來的資料型態為何
4. 你要回答的問題，其如何評估 (tips: 你的假設如何驗證)

作業 3：

請點選下方檢視範例依照 **Day\_001\_example\_of\_metrics.ipynb** 完成 **Mean Squared Error** 的函式

資料夾結構建立規則提醒：ML100 > data 資料夾 & homework 資料夾 (ipynb 檔)

(請注意 data 的存放位置，建議放在\*.ipynb 同一個目錄下，這樣才能在不修改 code 的情況下正常執行)



提交作業請上傳 **Day\_001\_HW.ipynb** 檔案到您的 **github** 帳號之後，再提供連結，(作業 1.2 申論題目可一併寫在 **HW** 檔案裡)

提醒：同步 Github 本機資料夾時，不需將 Data 一起同步(可 ignore)，只需要將作業檔同步即可。