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

-------------------------------ANSWER-------------------------------

Competiton Name : Prudential Life Insurance Assessment

1. 保戶風險自動分類預測模型，加速承保流程進行，以降低保戶於申請過程中放棄的機率。
2. Provided by Prudential.(英國保誠人壽)
3. 結構性資料
4. quadratic weighted kappa

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)

-------------------------------ANSWER-------------------------------

1. 規劃出一天之內各時段單位時間內車資最高的載客地點提升業績量
   1. 定義業績 = 單位時間內的車資
   2. 假設在不同地區及時間點的單位時間內車資具有顯著差異。

2. 不同時段該區域附近店家性質(有無載客服務需求)、消費人數(有無潛在市場)

3. 結構性資料

4. 使用數據分析前後業績是否具有顯著成長。

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：2nd-ML100Days > 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)