HW1

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

Landmark recognition in 2019

1.為什麼這個資料重要?

在這個以消費經濟為主軸的時代，人們為了知名度高的景點而花費金錢去參訪，然而，有的時候玩得不亦樂乎卻沒有記錄到當時拍下的美景，因此導致事後整理照片時毫無頭緒。如果我們能夠發展一種應用程式，只要將圖片置入就能立即辨識該景點的詳細資訊的話，對於人們來說，這一定是非常具有吸引力的一種新技術。雖然GOOGLE早就已經有圖片搜尋的功能，但是對於該演算法來說，還是有所謂優化的空間。

2.資料重何而來?

挖掘網路圖片並化成資料型態

3.資料的型態是什麼?

圖片和資料描述

4.我們可以回答什麼問題?

圖片特徵🡪辨識景點

HW2

Q: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」)

客群預測，怎樣的人會使用服務?->預測Accuracy。

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)

用戶使用紀錄、問卷調查。

3.蒐集而來的資料型態為何?

根據使用紀錄及用戶資料整理過後的結構式資料。

4.你要回答的問題，其如何評估 (tips：你的假設如何驗證)

分群、分類問題 ->一個客戶依據使用情況分群後被分到使用或不使用的類別，並驗證準確率為何。