## Python資料科學程式馬拉松

► Pandas 大型資料處理與效能調校



陪跑專家: Hong

## 作業



- 作業目標:
  - 了解效能優化方法
- 作業重點:
  - 優化有很多面相可以切入,目前可以先以投影片方法為主

## 作業題目



- 1. 在速度較慢的時候,可以先從哪邊開始檢查?
  - 1. 看是否是用對格式, e.g. xlst太慢換成csv, 可以用內建的函式跑比較快, 用向量的方式計算較快
- 2. 資料過大時應採取什麼方式讓記憶體占用量下降?
  - 2. 將浮點數型態 float64 改成 float32