**北大 x 好域**

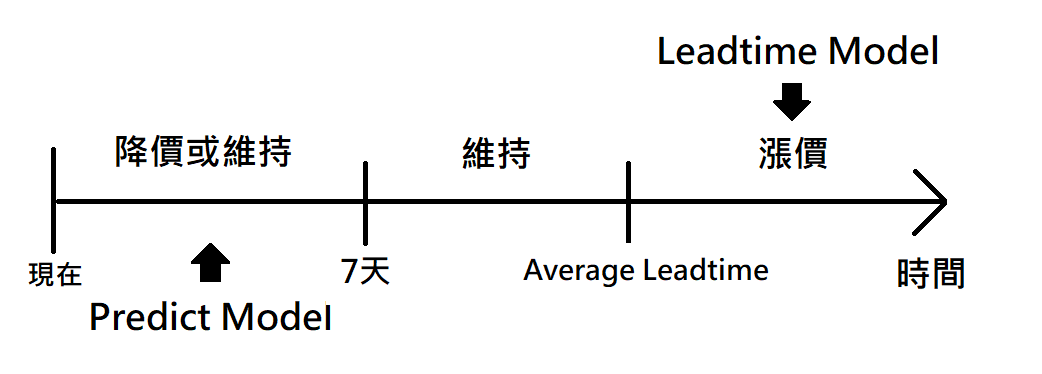
時間：2020/2/19（三） 14:00

參與人：沈易星學長、林奕銜、蕭琮峻、許程閔、黃俊友

上次[會議紀錄](https://docs.google.com/document/d/1QuZiHb5_2FIpLVcXvhKxig72p1xEcTZXgNad3nf9Q8Y/edit)，[Working Process](https://docs.google.com/document/d/1AUf4gI5kxPYSEiniz-MhMeHtMWicBmLqcMjot6EHihM/edit?usp=sharing)

**目前進度與討論：**

1. space準確率在80%以下的模型，可能是under fitting的問題。
   1. 可先將模型依照準確率高低排序，並按模型數量平均分為四群，試圖找出原因。
      1. 將資料分群後，分別依照 space\_data，order\_data，使用率去找尋系統性差異。目前找出跟使用率有一定的關係。[連結](https://drive.google.com/open?id=19or49Xbc7ajKFXNvNNFf5CjQPhVk6K8_)
   2. 就學理上的瞭解，將 under fitting 的解決方式套用在模型上，並觀察模型績效是否有提升。
      1. 原本想運用課程中圖像化的方式來找出 x 與 y 的關係以提升模型準確率，但是發現圖像化的方式只適用於數值型的 y ，而不適用於分類型的 y。
2. 動態定價：以兩種模式來決定價格
   1. Predict Model：
      1. 7天以內之價格，以預測模型預測出的出租機率作為依據。
      2. 出租機率高可維持價格；出租機率低則降低價格。
   2. Leadtime Model：
      1. 7天到 Average Leadtime 之間，維持價格。
      2. 超過 Average Leadtime 的時間，依超出標準差多寡決定價格。



**下次進度：**

1. 分析顧客的消費模式，是不是符合常態分配。
2. 程式化繼續隨目前進度持續修改及改善。[連結](https://colab.research.google.com/drive/19vzpn55u9JkY9UOkAcP9E13Uae3pSy4A)
   1. Leadtime Model 產生 Leadtime 正負一個標準差內區間。
   2. Predict Model 依據leadtime進行出租預測，並做出最終定價。
   3. 預定率小於60%不調價，小於75% -> 調漲5%，大於75% ->調漲10%。
3. 研究如何使用 Python -> MySQL 對資料庫的操作：
   1. 從資料庫撈出資料
   2. 進行資料處理、預測
   3. 產生 Space Leadtime Price Table
      1. 屬性
4. 做一個完整的專案整理文件，包含資料整理過程、內容，以及最後產出。

**上次進度：**

1. Lead\_time和「預測模型」進行交叉驗證，看兩者是否有關聯。[連結](https://drive.google.com/open?id=1PO5idTNtdLMo17SHZuB2WSjRqN0D3sYQ)  
   相關係數為 : -0.008

[lead\_time程式連結](https://drive.google.com/open?id=1sC2rraCRhLQOAPgTTynnB9beRutSzYk8)

[lead\_time資料表](https://drive.google.com/open?id=1bYBjGyOczGyViVLoHmBXvhW_VZF3y-xg)

1. space準確率在75%以下的模型，可能是under fitting的問題。
   1. 可先將模型依照準確率高低排序，並按模型數量平均分為四群，試圖找出原因。
   2. 就學理上的瞭解，將under fitting的解決方式套用在模型上，並觀察模型績效是否有提升。

**要注意的事情：**

1. 羅吉斯回歸預測模型，ROC Curve要切的機率點可以再討論和確定。
2. 思考最適價格的預測（與當前定價比較）。
3. 可思考依據那些因素（季節、使用率）觀測市場的需求變化，可再依此做出價格調整。
4. 純粹根據房間資訊來進行預測，找出同一個locate當中，不同space的價格符不符合目前的價格模型，再進一步來進行調整。
5. 自動化檢定比較新舊資料差異（可從使用率進行比較）。

會使用到的原始資料：

* spaces.csv
* spaces\_times.csv
* spaces\_devices.csv
* orders.csv
* god\_account\_view.csv

核心程式：[連結](https://colab.research.google.com/drive/19vzpn55u9JkY9UOkAcP9E13Uae3pSy4A)

* 下列為程式會產生出的檔案

1. Phase 1: Create necessary data
   1. \* [Space\_info](https://drive.google.com/open?id=1xlW-pEEos8QoUz_YUNNQdCXpMET4XZi1) (For Generate Fake Order)
   2. \* [Real Order](https://drive.google.com/open?id=1l8ByaEIDClP0PBZ_Zb_o3FnnjbxTZd13) (For Generate Fake Order)
2. Phase 2: Generate Fake Order
   1. [Fake\_Order\_in\_Empty\_Time](https://drive.google.com/open?id=1fAhTkt9oKpUKtHgtbVj8IRz2jGrtUczz) (For c. combine)
   2. [Fake\_Order\_in\_Real\_Order\_Time](https://drive.google.com/open?id=1E8s82tr2z0uR-HP-_E0CslEA42e-KGj3) (For c. combine)
   3. \* [Fake & Real Order Combine](https://drive.google.com/open?id=165OJ3XqADkfGKW4kAcX1EtQWKOilKVEW) (For predict model, calculate lead\_time)
3. Phase 3: Setting Price
   1. [Locate\_Lead\_Time](https://drive.google.com/open?id=18ansq4XL-k3hySMV-51-njxY7mPnDUqv) (For setting Price)
   2. **Final Recommanded Price Table**