

## 最終任務

### Final Project

**項目:** 請從以下列表挑出一項參加即可, 各項題目之任務細節請以連結方式開啟。

FX1 : 房屋越來越小了?

FX2 : 交通紅利: 捷運站對房價之影響

FX3 : 樓層對房價之影響

FX? : 其他房屋交易相關之議題

FY1 : 颱風對蔬菜交易量與價格之影響

FY2 : 估計各種蔬菜的需求價格彈性

FY? : 其他與蔬菜市場相關之議題

FZ1 : 旅館不同國籍之客群分析

FZ? : 其他旅館訂房相關之議題

F?? : 其他自行發想之議題

註: 其中 FX?, FY?, FZ?, 以及 F?? 四個項目之題目需要經過教師同意才可以進行, 請先來信 Email 約討論時間。

#### 參加及獎勵方式

- 可以用個人身份參加, 也可以與同學組隊參加 (一組最多 4 人), 但需要每個組員都已經開啟此任務。
- 若以個人參加, 獎勵則為分數百分比乘以 15000EXP, 例如: 報告分數 90 分, 則拿  $0.9 \times 15000 = 13500\text{EXP}$ 。
- 若以團體參加, 每個人之獎勵則為分數百分比乘以 10000EXP, 例如: 報告分數 90 分, 則每個人拿  $0.9 \times 10000 = 9000\text{EXP}$ 。
- 詳讀以下各項目議題以及資料內容後, 欲參加者請填以下[連結](#)之問卷, 未填問卷者不予採計。
- 在執行最終任務的過程, 歡迎來找我討論, 可以以 Email 方式約線上討論時間。

## 房屋越來越小了？

**概述：**近年來房屋市場交易的房屋似乎有越來越小趨勢，此任務想利用實價登錄資料探討房屋交易的建物面積趨勢。

**資料：**大台北地區 2012-08 至 2017-08 之房屋交易樣本，變數細節請參考連結，[房屋資料說明](#)，資料下載[連結](#)。

### 主要問題：

- 交易房屋的建物面積是不是隨著時間有下降的趨勢？

### 其他次要問題：

- 是不是每一種建物型態都有建物面積下降的趨勢？建物型態之間又有什麼差異呢？
- 除了建物面積以外，房間數量是否也有減少的趨勢？
- 新舊房屋是否都存在建物面積下降的趨勢？屋齡對此效果的影響為何？
- 此現象在不同區的差異又為何？
- 有人猜測房價高漲是建物面積縮小的主因，該如何檢驗這個假說呢？

**指引：**了解資料變數後，可以使用圖形及迴歸分析來分析建物面積 (build\_size) 的趨勢，藉此來回答主要問題，其他次要問題則可以挑一些有興趣的來回答。此外，有新聞報導猜測各種原因，連結如下：[https://hhh.com.tw/HHH\\_NEW/columns\\_detail/7783.php](https://hhh.com.tw/HHH_NEW/columns_detail/7783.php)，若有想到可以檢測的因素，也可以結合外部資料來回答問題。

請以 R markdown 製作 html 檔案完成以上內容，內容須包含 Code 以及結果，完成後只需上傳 html 檔案即可。

## 交通紅利：捷運站對房價之影響

**概述：**房屋的價值似乎與其交通有關，越便利的交通將使得房屋價值越高，此任務想利用實價登錄資料，結合捷運站的位置探討捷運站如何影響房屋的交易價格。

### 資料：

1. 大台北地區 2012-08 至 2017-08 之房屋交易樣本，變數細節請參考連結，[房屋資料說明](#)，資料下載[連結](#)。
2. 捷運站經緯度位置資料，資料下載[連結](#)。

### 主要問題：

- 交易房屋的價格是不是隨著越接近捷運站越貴？其距離的效果為何？

### 其他次要問題：

- 是不是每一種建物型態都可以觀察到此效果？建物型態之間又有什麼差異呢？
- 是不是每一條捷運站的影響效果都一樣？不同線有何差別？或是轉運站是否有更大的效果？不同捷運站周邊的房屋組成是不是也不一樣？

**指引：**了解資料變數後，可以使用座標資訊將兩個資料結合，房屋資料有 TWD97 的座標 (cor\_x, cor\_y)，捷運站位置資料也有相對應的 TWD97 格式 (TWD97\_X, TWD97\_Y)，基本上 TWD97 可以視為直角座標系統，因此距離可以直接計算，更多細節請參閱資料說明。當結合兩個資料後，便可以利用不同的迴歸模型來檢驗捷運站距離之效果，其他次要問題則可以挑一些有興趣的來回答。此外，若有捷運站相關的資訊，也可以自行創造想回答之次要問題。

請以 R markdown 製作 html 檔案完成以上內容，內容須包含 Code 以及結果，完成後只需上傳 html 檔案即可。

## 樓層對房價之影響

**概述:** 房屋的價值似乎與其樓層有關, 但越高的樓層就越好嗎? 此任務想利用實價登錄資料探討樓層對房價的影響。

**資料:** 大台北地區 2012-08 至 2017-08 之房屋交易樣本, 變數細節請參考連結, [房屋資料說明](#), 資料下載[連結](#)。

### 主要問題:

- 交易房屋的價格是不是隨著樓層越高而上升? 該建築物的高度又扮演什麼角色?

### 其他次要問題:

- 是不是每一種建物型態都可以觀察到此效果? 建物型態之間又有什麼差異呢?
- 是不是有一些特別的樓層大家會比較不喜歡或比較喜歡? 那又如何影響其價格?

**指引:** 了解資料變數後, 會發現有兩個變數跟樓高有關, 一個是該建物本身的高度 (`build_floor`), 一個是交易的樓層 (`trans_floor`), 所以除了考慮交易樓層以外, 可能還要考量其在建物的相對位置。其他次要問題則可以挑一些有興趣的來回答, 或者若有想到其他與樓高相關的次要問題也可以加入回答。

請以 R markdown 製作 html 檔案完成以上內容, 內容須包含 Code 以及結果, 完成後只需上傳 html 檔案即可。

## 其他房屋交易相關之議題

**概述:** 此交易資料為實際實價登錄上之交易資料, 範圍為大台北地區, 因此任何與房市場相關之議題應該都可以拿來檢驗與使用。加上資料有包含經緯度, 因此可以與其他地理資訊做結合。

**資料:** 大台北地區 2012-08 至 2017-08 之房屋交易樣本, 變數細節請參考連結, [房屋資料說明](#), 資料下載[連結](#)。

**指引:** 了解資料變數後, 可以思考是否有任何與房屋市場相關之政策或討論可以檢驗, 若有想到有興趣的議題, 記得要來找我確認题目的可行性, 以避免處理到太複雜或太簡單的議題。

請以 R markdown 製作 html 檔案完成以上內容, 內容須包含 Code 以及結果, 完成後只需上傳 html 檔案即可。

## 颱風對蔬菜交易量與價格之影響

**概述:** 颱風通常在夏季侵襲台灣, 每當颱風來襲, 蔬菜價格就成為新聞熱門議題, 此任務是利用蔬菜價格以及交易量資料探討颱風來襲對價格與交易量之影響。

**資料:** 台北第一果菜市場從 2000 年至 2004 年的每日各類蔬菜之交易量與平均交易價格, 變數細節請參考連結, [蔬菜市場資料說明](#), 資料下載[連結](#)。

### 主要問題:

- 颱風來襲對蔬菜的交易量與平均價格有何影響？

### 其他次要問題:

- 是不是每一種蔬菜的變化幅度都一樣？蔬菜間的差異為何？
- 颱風的大小是不是也會對此產生影響？
- 若颱風只有警報但沒有登陸, 影響會有不一樣嗎？

**指引:** 資料期間有不少颱風, 可以用全部的樣本來分析, 也可以只針對某一個特定的颱風來分析, 氣象局有記載颱風的資訊, 如以下連結: [https://rdc28.cwa.gov.tw/TDB/public/warning\\_typhoon\\_list/](https://rdc28.cwa.gov.tw/TDB/public/warning_typhoon_list/)。資料內除了記載颱風警報期間 (WR) 以外, 也包含了登陸的期間 (landfall) 以及未登陸的其他警報時間 (warning), 因此應該可以藉此了解颱風侵襲的時間點, 以及有無登陸。除了主要問題以外, 其他次要問題則可以挑一些有興趣的來回答, 或結合氣象局的資料回答一些額外自行發想與颱風和蔬菜有關之問題。

請以 R markdown 製作 html 檔案完成以上內容, 內容須包含 Code 以及結果, 完成後只需上傳 html 檔案即可。

## 估計各種蔬菜的需求價格彈性

**概述:** 在估計需求曲線的時候, 我們常會利用供給的衝擊來當工具變數, 颱風的來襲便是一個很好的外生衝擊, 此任務是利用颱風來襲來估計蔬菜之需求價格彈性。

**資料:** 台北第一果菜市場從 2000 年至 2004 年的每日各類蔬菜之交易量與平均交易價格, 變數細節請參考連結, [蔬菜市場資料說明](#), 資料下載[連結](#)。

### 主要問題:

- 蔬菜的需求價格彈性為何?

### 其他次要問題:

- 是不是每一種蔬菜的彈性都不一樣? 哪些類別的蔬菜彈性最大?
- 颱風的大小是不是也會對估計產生影響?

**指引:** 資料內有不少蔬菜, 主要還分為四大類: 葉菜類, 根莖菜類, 花果菜類, 以及菇菌類, 因此可以按類別來估計彈性。甚至每一種蔬菜的名稱 (item\_des) 也在資料內, 可以挑特定常見或重要的蔬菜來分析其彈性, 並且回憶經濟學之內容, 檢驗是否合理。工具變數也可以結合氣象局颱風的資訊, 如以下連結: [https://rdc28.cwa.gov.tw/TDB/public/warning\\_typhoon\\_list/](https://rdc28.cwa.gov.tw/TDB/public/warning_typhoon_list/)。除了主要問題以外, 其他次要問題則可以挑一些有興趣的來回答, 或結合其他農業相關知識回答一些額外自行發想與價格彈性有關之問題。

請以 R markdown 製作 html 檔案完成以上內容, 內容須包含 Code 以及結果, 完成後只需上傳 html 檔案即可。

## 其他與蔬菜市場相關之議題

**概述:** 此為真實台北第一果菜市場批發的平均價格與交易量, 並且結合當天是否有颱風, 由於是日資料, 也可以用來結合其他日層級的相關資料, 因此任何與蔬菜市場相關之議題或事件應該都可以拿來檢驗與使用。

**資料:** 台北第一果菜市場從 2000 年至 2004 年的每日各類蔬菜之交易量與平均交易價格, 變數細節請參考連結, [蔬菜市場資料說明](#), 資料下載[連結](#)。

**指引:** 了解資料變數後, 可以思考是否有任何任何事件或政府政策會影響到蔬菜市場的供給與需求, 若有想到有興趣的議題, 記得要來找我確認题目的可行性, 以避免處理到太複雜或太簡單的議題。

請以 R markdown 製作 html 檔案完成以上內容, 內容須包含 Code 以及結果, 完成後只需上傳 html 檔案即可。



## 旅館不同國籍之客群分析

**概述:** 旅客在訂房的時候通常有不同的習慣, 以旅館的角度在面對不同的旅客也可能採取不一樣的訂價策略, 此任務是利用旅館訂房資料來探討不同國籍之旅客是否有需求上的差異或是價格上之差異。

**資料:** 某旅館在 2015-2017 之間的訂房紀錄, 變數細節請參考連結, [旅館資料說明](#), 資料下載[連結](#)。

### 主要問題:

- 不同國籍旅客之訂房偏好是否有差異? 面對同樣的房型價格上是否也有差異?

### 其他次要問題:

- 不同國籍旅客之訂房時間上是否有差異? 所住宿的長度是否也有不同?
- 不同國籍旅客之訂房房型上是否有差異?
- 旅館價格之波動是否影響到不同國籍旅客所接受價格之差異?

**指引:** 此為真實旅館訂房資料, 必須詳細了解各變數的名稱, 價格資料為變數 `adr`, 另外要小心有些樣本是屬於「取消」的部分, 並非所有樣本都是交易。主要問題關於偏好的部分相對籠統, 可以自己定義一些變數, 如其他次要問題所述來檢驗。此外, 國籍 (`country`) 變數包含的國家非常多, 可以限縮在一些樣本較多的國家來分析即可。若有其他次要問題也可以加入討論或回答。

請以 R markdown 製作 html 檔案完成以上內容, 內容須包含 Code 以及結果, 完成後只需上傳 html 檔案即可。

## 其他旅館訂房相關之議題

**概述:** 此為真實旅館訂房資料, 主要來源為 kaggle [[連結](#)], 因此有許多有趣的問題可以回答。

**資料:** 某旅館在 2015-2017 之間的訂房紀錄, 變數細節請參考連結, [旅館資料說明](#), 資料下載[連結](#)。

**指引:** 在充分了解資料後, 可以運用資料去檢驗各種差別取價, 或是了解在一年的時間內旅館價格的波動, 以及波動所帶來的各種影響。若有想到有興趣的議題, 記得要來找我確認题目的可行性, 以避免處理到太複雜或太簡單的議題。

請以 R markdown 製作 html 檔案完成以上內容, 內容須包含 Code 以及結果, 完成後只需上傳 html 檔案即可。

## 其他自行發想之議題

**概述:** 此部分主要是鼓勵學生自行找尋有趣的議題及資料。

**指引:** 可以想想自己對哪些議題感興趣, 再從中去尋找資料, 或者反過來也可以, 先找到一個有趣的資料, 在嘗試問問看有什麼有趣的議題。若有想到有興趣的議題以及搭配之資料, 記得要來找我確認题目的可行性, 以避免處理到太複雜或太簡單的議題。

請以 R markdown 製作 html 檔案完成以上內容, 內容須包含 Code 以及結果, 完成後只需上傳 html 檔案即可。

## 房屋資料說明

**概述:** 此資料為大台北地區 2012-08 至 2017-08 之房屋交易樣本, 來源為內政部實價登錄, 連結為<https://lvr.land.moi.gov.tw/>。此樣本只針對有房屋買賣的交易, 且做過一些資料的清理, 因此並非完整樣本。

### 變數說明:

- id: 交易代碼 (unique)
- county: 台北市, 新北市
- town: 區
- trans\_obj: 房地 (土地 + 建物), 房地 (土地 + 建物)+ 車位
- land\_size: 土地大小 (平方公尺)
- build\_form: 建物型態, 包含住宅大樓, 華廈, 公寓, 以及套房。
- build\_size: 建物面積 (平方公尺)
- num\_bedroom/livingroom/bathroom: 臥房/客廳/浴室的數量
- guard: 是否有管理組織
- total\_price: 總價 (新台幣元)
- unit\_price: 單位價格 (新台幣元/每平方公尺)
- trans\_date\_mdy: 交易時間 YYYY/MM/DD
- finished\_date\_mdy: 完工日期 YYYY/MM/DD
- housing\_age\_y: 屋齡 (年)
- build\_floor: 建物高度 (樓)
- trans\_floor: 交易標的所在樓層
- cor\_x, cor\_y: TWD97 座標 (單位: 公尺)。例如: 若有兩個座標:  $(x_1, y_1)$  and  $(x_2, y_2)$  在 TWD97 系統中, 兩點的距離 (公尺):

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

## 蔬菜市場資料說明

**概述:** 此資料為台北第一果菜市場 2000 年至 2004 年蔬菜批發交易資料, 來源為<https://amis.afa.gov.tw/main/Main.aspx>, 蔬菜種類經過篩選, 只留下樣本較為齊全的蔬菜。

### 變數說明:

- year/month/day: 年/月/日
- avg\_price: 平均交易價格 (元/公斤)
- quantity: 交易量 (公斤)
- item\_code: 蔬菜代碼
- item\_des: 蔬菜代碼說明
- group\_code: 蔬菜所屬類別代碼
- group\_des: 蔬菜所屬類別
- WR: 是否有颱風警報
- landfall: 是否颱風登陸台灣 (註: 此登陸採用的定義是颱風的暴風半徑是否包含任何台灣的土地)
- warning: 颱風警報期間, 但颱風沒有登陸

## 旅館資料說明

**概述:** 此資料為某旅館在 2015-2017 之間的訂房紀錄, 來源為 kaggle [\[連結\]](#)。

**變數說明:** 請參閱以下論文連結說明: <https://www.sciencedirect.com/science/article/pii/S2352340918315191>