

---

# 檢視租屋物件與租金關係，守護租房者的居住正義

## ——透過數據，協助租屋者做出符合利益的租屋決策

**Huang, Chi-Hung**

National Taiwan University  
Taipei, 106216, Taiwan  
b06705002@ntu.edu.tw

**Chang, Yun-Chi**

National Taiwan University  
Taipei, 106216, Taiwan  
b06106046@ntu.edu.tw

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

檢視租屋物件與租金關係，守護租房者的居住正義——透過數據，協助租屋者做出符合利益的租屋決策, January 7, 2020, Taipei, Taiwan.

© 2020 Copyright is held by the owner/author(s).

### 摘要

透過檢視《開放台灣民間租屋資料》這個開源資料集中的變數，例如交通條件、生活機能、房型、建築類型等，解釋變數對租金的影響，進而協助使用者在租屋時做出符合自己利益的決策。我們透過迴歸分析探討數據，並在最後將研究限縮在可以利用最多變數的「整層住家」的房型。我們發現其他房型在現有變數上的分析限制，希望後續可以搭配質性研究，了解影響其他房型租金的因子，做出更精準且豐富的解釋模型。

### 作者關鍵字

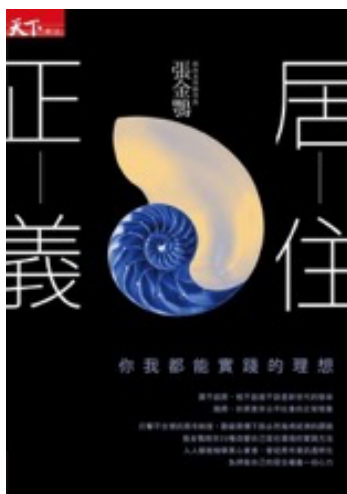
居住正義；租屋；房租；解釋型研究

### CSS 概念

• Computing methodologies~Machine learning~Learning paradigms~Supervised learning~Supervised learning by regression

### 緒論

近年來政府積極推動社會公義相關政策，是行政院의 五大施政目標之一，在其網站寫道：「公義，是國家最基本的價值，政府要讓每一位國民，都能尊嚴地生活，並在合理資源分配下，獲得安定、安全的生活保障。」在多项改革之中，有關居住正義的議題吸引了研究團隊的興趣。根據政大地政系張金鶚教授在其書籍《居住正義：你我都能實踐的理想》中提到「台灣房地產長期缺乏市場機制的獎優懲劣制度，結果買賣雙方互不信任，



圖一：張金鶚《居住正義：你我都能實踐的理想》

不但市場沒有效率，也造成許多不公平。要讓房市重回該有的機制與軌道」，而為了達到居住正義，可以從四個方向著手：

1. 提供透明的房市交易資訊，確保購屋品質，讓買房子的人住得安心。
2. 改革房地產相關稅制，增加囤房、炒房稅賦，減輕自住者的繳稅壓力，提升房屋市場的公平與效率。
3. 讓買不起房子的民眾至少也租得起房子，保障弱勢族群的居住權。
4. 改善生活環境，提升居住品質，讓每個居住在台灣的人都有居住尊嚴。

本團隊成員皆為應屆大學生，比起買房，租房對我們而言是在近期更有可能需要面對的問題。當我們離開學校的保護，沒有宿舍，離鄉背井找工作時，如何在租房市場中找到合理定價且符合需求的物件，對沒有任何租房經驗的我們來說，可能會是一大挑戰。因此，我們希望透過所學，用數據協助決策，實踐我們的居住正義。

在本次的研究中，我們希望探討北北基地區會影響租金的因子，透過檢視資料集中的數據，進而理解各變數對租金的影響，例如地點、交通便利性、房屋條件、家具配置等。這樣的做法也可以檢視是否有房東惡意哄抬價格，保護租屋者的權利。此外，除了檢驗租房市場中，物件價格的合理性以外，我們也希望透過檢視租屋者需求，進而協助其評估需求物件的價格是否合理，作為其在租屋時輔助判斷依據。

## 文獻回顧

### 一、居住正義的重要性

根據聯合國的《經濟社會文化權利國際公約》之一般性意見第四號，「住房相關的個人或家人花費應該在一定的範圍之內，而不導致其他方面的基本需求滿足被影響」。政府應採取適當的措施，保護租戶不因為受到不合理的租金以及提高租金而降低生活品質。

### 二、影響租金之可能變數

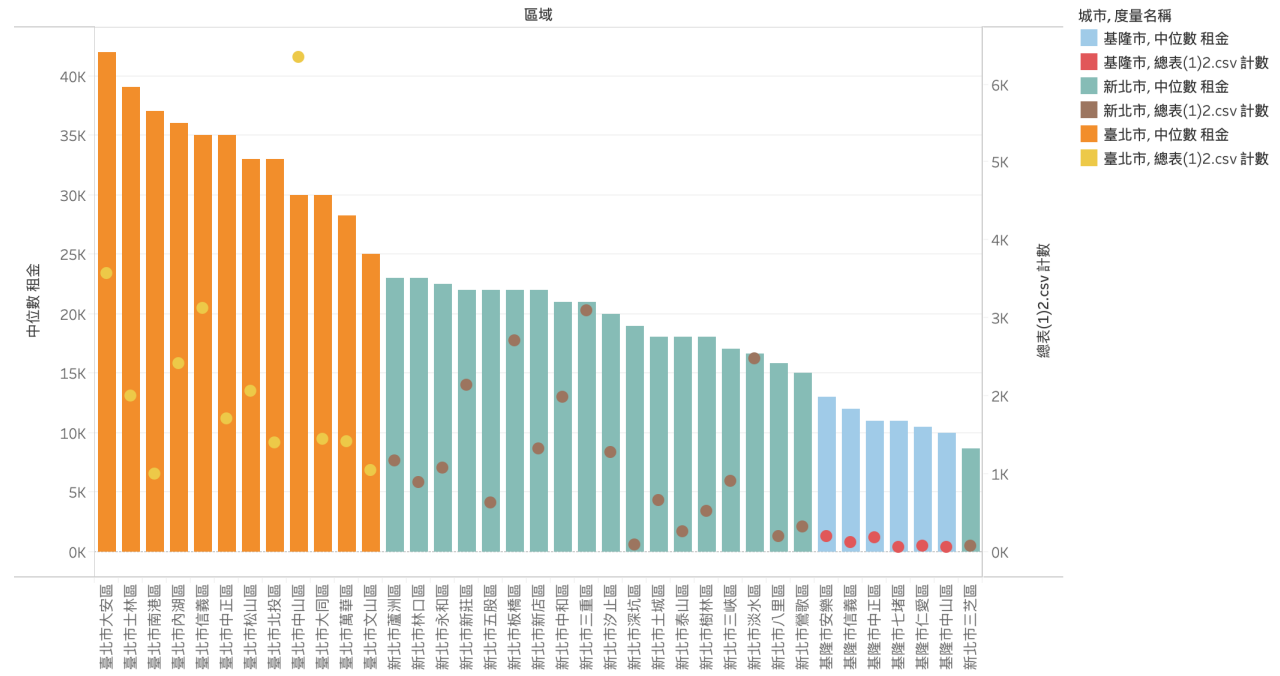
會影響租金的變數可以分成幾個類型，如周遭環境、交通狀況、建物狀況、物件狀況四類。以下分別列出相對應的變數。

1. 周遭環境：與購物商場、超市、學校、大學、開放空間、醫院的距離等。
2. 交通狀況：與地鐵、公車、火車的距離等。
3. 建物狀況：建物類型、住宅結構、屋齡、區域、是否曾經淹水、水質等。
4. 物件狀況：坪數、裝潢程度、對住宅的滿意程度、房間數、衛浴數量、大廳數量等。

### 三、租金價格探討

在評估居住品質及租金時，應避免使用單一因素歸因，可能會忽略其他有互相影響的因子，進而得到錯誤的結論，如林祖嘉（1993）在其研究中就提到，個別變數在房價與房租的關係不易區分，因此在歸納環境對租金的影響時，應用變數群組的方式進行分析。

行政區「整層住家」租金中位數



每個區域的中位數租金與總表(1)2.csv 計數。顏色顯示有關城市, 中位數租金與總表(1)2.csv 計數的詳細資訊。資料按房型進行篩選, 這會保留整層住家。檢視按區域進行篩選, 這會選擇多個成員。

圖二：北北基地區各行政區整層住家租金中位數。圖中圓點為物件數量。  
透過中位數檢視房價現況可以避免極值影響整體數據, 更中立地呈現數值。

## 資料處理

### 1. 搜集資料

本次主題的資料集來源為《[開放台灣民間租屋資料](#)》，此資料集是透過爬蟲技術取得 591 租屋網上的租賃相關資料，除了具有長期的詳細資料可以比較以外，在去除了有關著作權及隱私權的資料之後開放給公眾使用。

我們在挑選子資料集時，除了希望使用整年的資料以外，也希望避免 2020 因為疫情的相關影響，因此在本次研究中，我們選用《2019 年度資料集（消除重複住宅）》。

	月租金	縣市	物件類型	建築類型	鄉鎮市區
0	32000	臺北市	整層住家	電梯大樓	臺北市松山區
1	28000	臺北市	整層住家	電梯大樓	臺北市松山區
2	38000	臺北市	整層住家	電梯大樓	臺北市松山區
3	53000	臺北市	整層住家	電梯大樓	臺北市松山區
4	16000	臺北市	整層住家	電梯大樓	臺北市松山區
...	...	...	...	...	...
92679	8000	新北市	整層住家	透天	新北市石門區
92680	15000	新北市	整層住家	透天	新北市石門區
92681	5000	新北市	分租套房	公寓	新北市雙溪區
92682	30000	新北市	整層住家	透天	新北市石碇區
92683	7000	新北市	整層住家	透天	新北市平溪區

圖三：各區各類型案件月租金資料處理結果

## 2. 探索資料

圖一為原始資料之所有屬性。在這之中，我們認為存在著許多不會對月租金造成影響之屬性，如不具數值意義、重複資訊、與租金無關等特性。舉例來說，物件編號是該物件於原始網站的編號，並不具數值意義；格局可以透過房間數、陽台數等得知，因此編碼為重複資訊；仲介資訊為網頁上登記之公司名稱，應與屋主的個人偏好有關，而非租金。因此，我們決定將包含幾種情況的屬性排除，以便後續的研究。

一開始我們原本預計針對四種不同房型：雅房、分租套房、獨立套房及整層住家，分別進行建模，並找出對租金的影響因子。但在初步測試資料時發現僅有整層住家具有分析價值，推測是因為在資料集中，在資料搜集時便使用較適合整層住家的特性，至於選擇其他房型的租客可能考慮的因素則無法透過資料集得知。例如其他房型可能會面臨與陌生人分享公共空間的問題，或是案件本身坪數就較小，寢室可能同時也有客廳的用途。

因為上述的發現，所以我們決定在後續分析時，僅使用需要考慮最多特性的**整層住家**進行分析及建模。

## 3. 資料前處理

在本次的研究中，為了讓更多人了解目前的北北基租屋概況。在探索資料之後，我們分成不同階段對資料做了不同方式的處理。三階段分別是(1) 各區各類型案件月租金，(2) 整層住家租屋資料和(3) 各地理位置之月租金。

### (1) 各區各類型案件月租金

我們希望透過此資料集，搭配資料視覺化的方式，呈現出目前各租屋案件不同特性之間的租屋概況。因此，

此資料集需要包含月租金、縣市、鄉鎮市區、建築類型和物件類型。處理步驟如下：

1. 讀取資料後，只保留需要的欄位，刪除不需要的欄位。
2. 因為許多保留的屬性是用編碼紀錄，因此讀進編碼表後，利用 merge 將編碼轉換為文字。
3. 篩選條件，只保留北北基地區和已成交的案件。
4. 刪除不完整資料，只保留有做完整紀錄的案件方便分析。
5. 輸出為 csv 檔，以利後續視覺化之需求。

### (2) 整層住家租屋資料

我們希望透過此資料集建構針對整層住家類型之迴歸模型，利用不同案件之各欄位的屬性去解釋該案件的月租金。因此，此資料集需要包含我們認為有意義的所有欄位，例如各類型房間數量、附近交通機能等。處理步驟如下：

1. 讀取資料後，只保留需要的欄位，刪除不需要的欄位。
2. 因為許多保留的屬性是用編碼紀錄，因此讀進編碼表後，利用 merge 將編碼轉換為文字。
3. 對類別變數取虛擬變數。
4. 篩選條件，只保留整層住家類型、北北基地區和已成交的案件。

5. 刪除不完整資料，只保留有做完整紀錄的案件。
6. 調整資料，將 T 和 F 分別轉換為 1 和 0。
7. 利用平均值填補押金月數欄位之遺失值。
8. 輸出為 csv 檔，以利後續建模之需求。

### (3) 各地理位置之月租金

我們希望透過此資料集，搭配資料視覺化的方式，呈現目前每個租屋案件的地理位置與月租金。因此，此資料集需要包含月租金、經度和緯度。處理步驟如下：

1. 讀取資料後，只保留需要的欄位，刪除不需要的欄位。
2. 因為許多保留的屬性是用編碼紀錄，因此讀進編碼表後，利用 merge 將編碼轉換為文字。
3. 篩選條件，只保留北北基地區和已成交的案件。
4. 刪除不完整資料，只保留有做完整紀錄的案件。
5. 利用字串處理，約略地點範圍的經度和緯度分別獨立建欄位。
6. 輸出為 csv 檔，以利後續視覺化之需求。

	月租金	經度	緯度
0	8000	121.501625	25.042458
1	8500	121.564288	24.990190
2	9500	121.542199	25.050234
3	11000	121.511494	25.069666
4	6000	121.548142	24.986602
...	...	...	...
92679	6500	121.755338	25.093698
92680	6500	121.720940	25.122426
92681	12500	121.737106	25.129175
92682	3000	121.786385	25.136435
92683	5000	121.702330	25.091590

圖四：各地理位置之月租金資料處理結果

## 4. 模型建置

透過林祖嘉（1993）的研究，我們認為對不同類型的欄位做迴歸可以較精準地得到某個特定類型的各特性會如何影響月租金。我們會針對(1) 房屋屬性和(2) 交通機能與環境屬性建立迴歸模型。

### (1) 房屋屬性

從結果可以得知，調整後的 $R^2$ 為 0.524，代表對於模型中的預測變數個數進行了調整之後，月租金變異的 52.4%被房屋屬性的變異所解釋。而剩下的 47.6%則沒有被解釋。此迴歸模型通過 F 檢定（顯著值為 0.00），為有效。我們可以得知，房屋屬性是影響月租金的其中一個主因。房屋屬性中所有的自變數均通過 T 檢定（ $p < 0.05$ ），代表所有的自變數都是有效的。

從數據可以得知，對整層住家的需求者來說：

1. 比起公寓與透天，更加偏好於電梯大樓，且所在樓層越低越好。
2. 自報頂加和月租金呈負相關
3. 月租金對坪數和衛浴數為正相關，而房數、客廳術與陽台數則呈負相關。房數和月租金呈負相關。

雖然和直覺不合，但是在搜尋相關文獻後，我們發現有其他台灣的研究能夠支援這樣的結論，可能的原因與台灣家庭規模小，加上每坪平均價格高昂有關，房數減少可以提高室內空間使用效率，增加居住品質（李馨蘋、劉代洋，1999）

### (2) 交通機能與環境屬性

從結果可以得知，調整後的 $R^2$ 為 0.069，代表對於模型中的預測變數個數進行了調整之後，月租金變異的 6.9%被交通機能與環境屬性的變異所解釋。而剩下的 93.1%則沒有被解釋。此迴歸模型通過 F 檢定（顯著值為 0.00），為有效。我們可以得知，雖然交通機能與環境屬性通過 F 檢定，但是只解釋了月租金的一點點變異。

房屋屬性中附近有學校和附近有超商沒有通過 T 檢定 ( $p>0.05$ )，這代表著沒有足夠的證據推斷，在其餘自變數相同的情況下，這兩項自變數的不同會具有不同的月租金。

從數據可以得知，對整層住家的需求者來說：

1. 月租金對附近有公園、百貨公司、超商、醫療機構和附近的捷運站數呈正相關。
2. 月租金對附近有學校、附近有傳統市場、附近有夜市、附近的公車站數、附近的火車站數和附近的公共自行車數呈負相關。

經過推論，我們認為雖然上述屬性均代表著便利，但是會造成正相關負相關的原因在於環境。舉例來說，雖然有夜市代表著便利，然而卻也會連帶著噪音、環境問題。

### (3) 綜合比較

在最後，我們決定將所有因素混合放入模型中，查看月租金在綜合比較上的結果。在這一步驟中，我們做了下列處理：

1. 因為對公寓和電梯大樓來說，所在樓層代表的意義不一樣，因此針對這兩項屬性做相乘的動作。
2. 除了以上兩大類，新增八項屬性，分別為代表\_基隆市、代表\_新北市、代表\_臺北市、可寵、提供車位且需要停車費、提供車位但不需要停車費、押金月數及有身份限制。

3. 最後，先做一次迴歸後，刪除顯著值太低 ( $p\text{-value}>0.5$ ) 的屬性，重新進行迴歸分析，得出以下結果。

從結果可以得知，調整後的  $R^2$  為 0.680，代表對於模型中的預測變數個數進行了調整之後，月租金變異的 68% 被這些屬性的變異所解釋。而剩下的 32% 則沒有被解釋。此迴歸模型通過 F 檢定 (顯著值為 0.00)，為有效。我們可以得知，雖然交通機能與環境屬性通過 F 檢定，但是只解釋了月租金的一點點變異。而押金月數、附近有公園、附近的公共自行車數和公寓沒有通過 T 檢定 ( $p>0.05$ )，這代表著沒有足夠的證據推斷，在其餘自變數相同的情況下，這兩項自變數的不同會具有不同的月租金。

從圖中可以看出來對整層住家的需求者來說：

1. 以月租金來說，台北市 >> 新北市 > 基隆市。
2. 以公寓來說，所在樓層越高，租金越便宜；對電梯大樓來說，所在樓層越高，租金越貴。
3. 只要有提供車位，租金都會較貴。
4. 我們認為這些推論符合直覺，越遠離地面，蚊蟲和噪音皆會越少，租客可以得到較好的居住品質。
5. 不過若是沒有電梯，租客還需考慮要自己爬樓梯的問題。

## 結論

根據資料集中 2019 年一整年的資料，我們建立了對於預測位於北北基地區整層住家類型之租金模型，希望能讓未來的租屋族作為評判租金合理程度的輔助之一，也可以一併檢視房東是否有惡意哄抬價格之行為，保護自己的居住權益。

在這次的研究中，我們認為現有的研究資料不足以進行有效分析，因此無法分別針對雅房、分租套房以及獨立套房建模，可能的原因是影響這類房型租金的因子無法透過現有的資料集獲得。未來或許可以搭配對房東及租客的質化研究，以利了解箇中原因，進而建立有效的分析模型，協助租客對租房市場有更多的認識、避免房東惡意提高價格，達成租房市場的穩定，維護台灣的居住正義。

## 參考資料

- [1] 李永然（2019）。我國如何落實「居住權」的保障？—從台灣實施《租賃住宅市場發展及管理條例》談起。人權會訊，(131)，8-11。
- [2] ICESCR | 經濟社會文化權利國際公約. (n.d.). Retrieved December 26, 2020, from <https://covenantwatch.org.tw/un-core-human-rights-treaties/icescr/>
- [3] Zhang, S., Wang, L., & Lu, F. (2019). Exploring Housing Rent by Mixed Geographically Weighted Regression: A Case Study in Nanjing. *ISPRS International Journal of Geo-Information*, 8(10), 431. doi:10.3390/ijgi8100431
- [4] 林祖嘉、林素菁（1993）。台灣地區環境品質與公共設施對房價與房租影響之分析。住宅學報，(1)，21-45。doi:10.6375/JHS.199301.0021
- [5] 【租屋分析系列一】哪些因素會讓租金變高呢？- StockFeel 股感. (2020, December 07). Retrieved December 26, 2020, from <https://www.stockfeel.com.tw/%E3%80%90%E7%A7%9F%E5%B1%8B%E5%88%86%E6%9E%90%E7%B3%BB%E5%88%97%E4%B8%80%E3%80%91%E5%93%AA%E4%BA%9B%E5%9B%A0%E7%B4%A0%E6%9C%83%E8%AE%93%E7%A7%9F%E9%87%91%E8%AE%8A%E9%AB%98%E5%91%A2%EF%BC%9F/>
- [6] 【租屋分析系列二】雙北哪個行政區租金最貴？最便宜？- StockFeel 股感. (2020, December 07). Retrieved December 26, 2020, from <https://www.stockfeel.com.tw/%E3%80%90%E7%A7%9F%E5%B1%8B%E5%88%86%E6%9E%90%E7%B3%BB%E5%88%97%E4%BA%8C%E3%80%91%E9%9B%99%E5%8C%97%E5%93%AA%E5%80%8B%E8%A1%8C%E6%94%BF%E5%8D%80%E7%A7%9F%E9%87%91%E6%9C%80%E8%B2%B4%EF%BC%9F%E6%9C%80/>