

我們的小組專案旨在探討台灣六都的人口淨流入與房價變動之間的關係，同時也將多項社會經濟統計資料納入考慮，包含消費者物價指數（CPI）、地區平均年所得、房貸利率與失業率等，我們希望找出可能影響房價的關鍵因素。本專案將嘗試套用多種統計模型，評估其解釋力以找出最能準確描述房價變動的最佳模型，進而為都市發展與政策規劃提供實證基礎。

資料預處理與初步分析

剔除離群值

使用兩種常見方法以剔除資料中的離群值，以提升後續模型分析的穩定性與準確度：

方法 1：IQR（四分位距法）
若某值不在範圍 $[Q1 - 1.5 * IQR, Q3 + 1.5 * IQR]$ 內，則視為離群值並移除

方法 2：Z-score（標準差法）
若某值的 Z-score 超過 ± 3 ，即落在 $\text{mean} \pm 3 * \text{sd}$ 之外，則移除（可自行決定 coef）

初步分析

使用相關係數熱圖檢查模型中解釋變數是否存在多重共線性問題，發現模型中的各項解釋變數間沒有明顯的共線性問題，適合作為多元迴歸解釋基礎

異常分析

區域異常觀察：房價上升 vs. 人口負成長

在多數城市中，人口成長往往被視為房價上升的主要驅動因素。然而，本研究發現臺南市與臺北市，呈現出房價長期上升但人口未同步成長甚至流失的「脫鉤現象」

臺南市：價格逆勢上升的典型

- 近十年房價翻倍，從每坪 10 萬成長至 25 萬以上
- 人口淨流入多為負值，顯示人口持續外移
- 回歸模型中，人口變動對房價影響不顯著
- ARIMA 模型中為異常預測最多的城市
- 可能原因：
 - 地段開發（如沙崙科學城）、投資性購屋增加
 - 價格炒作或資金集中導致結構性拉抬

臺北市：高房價與人口流動脫鉤

- 房價持續上漲，2022 年達 75 萬元/坪
- 人口移動劇烈波動、整體成長停滯甚至下滑
- 多元回歸分析顯示，人口對房價影響微弱
- 模型預測誤差高、難以由實質需求解釋房價
- 可能原因：
 - 豪宅市場、土地稀缺、資金避險性流入
 - 結構性高房價已脫離人口基本面

模型調整與優化

基礎模型：多元線性回歸（OLS）

房價 = $\beta_0 + \beta_1 \text{人口淨流入} + \beta_2 \text{CPI} + \beta_3 \text{貸款利率} + \beta_4 \text{年所得} + \beta_5 \text{失業率} + \epsilon$

- Adjusted R^2 為 0.899，AIC = 4487
- VIF（變異膨脹因子）檢定顯示無嚴重共線性
- 殘差診斷圖顯示有異質性與非線性情況，及部分時間序列關係未被捕捉

模型優化：log 轉換與交互項調整

針對非線性與異質性問題，對變數進行 log-transform 與加入交互作用（如 CPI \times 貸款利率）。

- 儘管模型調整後更具彈性，但 AIC 增加，Adjusted R^2 反而下降，整體效果不如原始模型

固定效果模型（Fixed Effects, FE）

考量時間與地區異質性，導入 Panel Two-way Fixed Effects：

- 控制城市固定效果（跨區差異）與月份固定效果（時序波動）
- 模型解釋力大幅提升至 Adjusted $R^2 = 0.974$ ，AIC 降至 3542.45
- 斜率差異 Wald 檢定顯示城市間對人口變動反應有顯著異質性

時間結構修正：GLS with AR(1)

Durbin-Watson 檢定顯示殘差具有強烈自相關（統計量 = 0.54），因此進一步使用 Generalized Least Squares (GLS) 模型，設定自相關結構 (AR(1))

- 修正後 AIC 降至 3787，CPI 等變數由不顯著轉為顯著，顯示模型穩健度提升

透過逐步模型優化與殘差診斷，並比較各模型的 Adjusted R^2 與 AIC 分數，最終選擇 Panel Fixed Effects 模型作為主要分析依據，其能同時控制區域差異與時間因素，提供穩定、可信的估計結果，為房價政策與市場觀測提供有力依據

成果總結與政策建議

我們的實證結果顯示，人口淨流入、CPI 與失業率皆對房價具有顯著正向影響，且不同城市之間存在明顯的房價結構差異。此結果支持政府應針對人口與房價脫鉤的城市（如臺南）加強投資性購屋管理，並建立基於預測誤差的房價預警機制；同時，在房價評估中應納入非人口因素，並依各地特性推動差異化的住房政策。