

人口流動與多項社經因素對六都房價之影響探討

[Group7] Final Project

王苓卉 (統計四 110304029)

李崇聖 (資訊碩一 113753209)

陳星豪 (資訊碩一 113753206)

陳凱輝 (資訊碩一 113753120)

郭祐丞 (統計四 110304058)

潘韋筑 (統計四 110304003)

目錄

1 簡介

2 全域分析

3 區域分析

4 迴歸分析

5 成果總結

專案目標

- 探討人口流動、利率、CPI、失業率等因素與六都房價之關聯
- 建立多元與進階迴歸模型預測房價變動
- 判斷各城市影響房價的主要變因是否一致
- 尋找房價與變數間的矛盾或異常模式

資料概述

- 資料期間：2013 年 1 月 – 2023 年 12 月

- $$\text{平均房價} = \frac{\text{公寓、住宅大樓、透天厝每坪單價 (元)}}{\text{交易量 (棟)}}$$

來源：[政大商學院信義不動產研究發展中心](#)

- $$\text{人口淨流入} = \text{遷入人數} - \text{遷出人數}$$

來源：[中華民國統計資訊網-歷史內政統計月報](#)

- 消費者物價指數 (CPI)

來源：[中華民國統計資訊網-重要經社指標](#)

- 五大銀行貸款利率

來源：[中華民國中央銀行全球資訊網](#)

- 每戶可支配年所得 (元)

來源：[中華民國統計資訊網-主計總處統計專區-家庭收支調查](#)

- 失業率

來源：[勞動部勞動統計查詢網](#)

研究動機

- 近年六都房價高漲，引發關注與政策討論。
- 傳統認為「人口向特定地區集中」是房價上升的主要因素。
- 台南與台北出現「人口負成長但房價上漲」的脫鉤現象。
- 單一因素無法充分解釋房價變動，需納入多元社經變數進行分析。
- 期望找出最能準確描述房價變動的統計模型。
- 為都市發展與政策規劃提供實證基礎。

研究方法

研究對象

台灣六都（台北、新北、桃園、台中、台南、高雄）住宅房價

研究方法

- 資料清理與欄位標準化
- 房價與各變數關係之視覺化探索
- 時間序列預測（ARIMA）
- 多元線性與固定效果模型建構
- 共線性與異常偵測分析
- 城市間房價影響力比較（斜率檢定）

專案架構

1 探索性資料分析 (EDA)

- 清理原始資料
- 處理遺失值、離群值 (z-score & IQR)

2 全域分析

- 人口淨流入、失業率 vs 房價
- 房價與利率變化率的即時影響分析

3 區域分析

- ARIMA 預測未來 5 年各直轄市房價走勢

4 迴歸分析

- 房價 = $\beta_0 + \beta_1$ 人口淨流入 + β_2 CPI + β_3 貸款利率 + β_4 年所得 + β_5 失業率 + ε
- 多重共線性
- 殘差分析
- 模型改進
- 模型比較

目錄

1 簡介

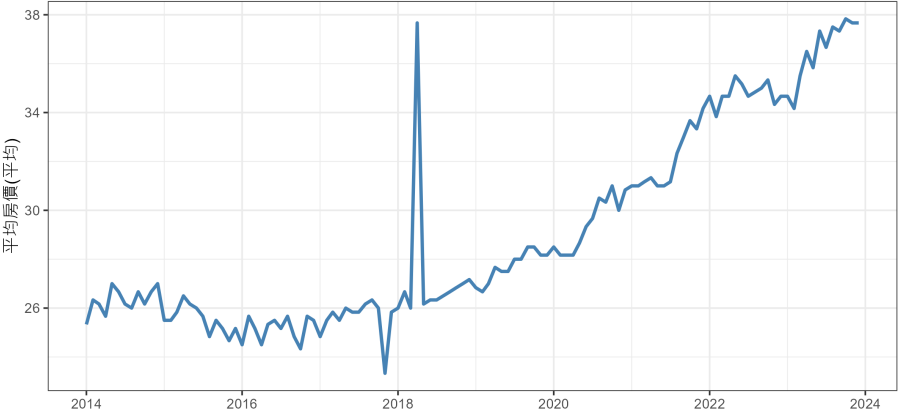
2 全域分析

3 區域分析

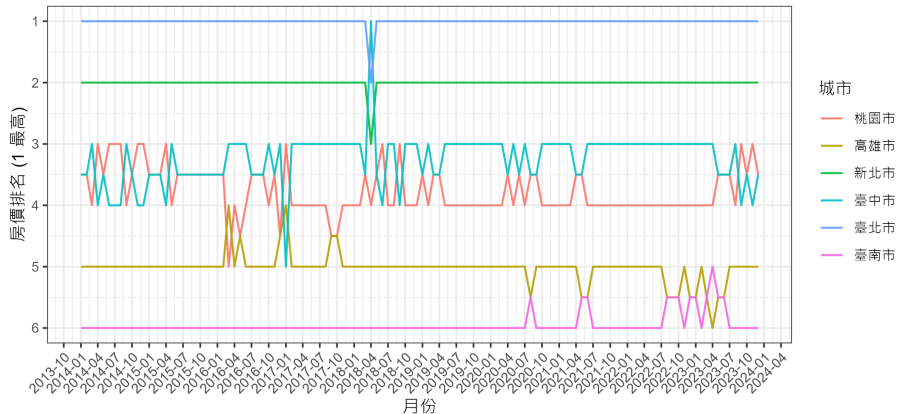
4 迴歸分析

5 成果總結

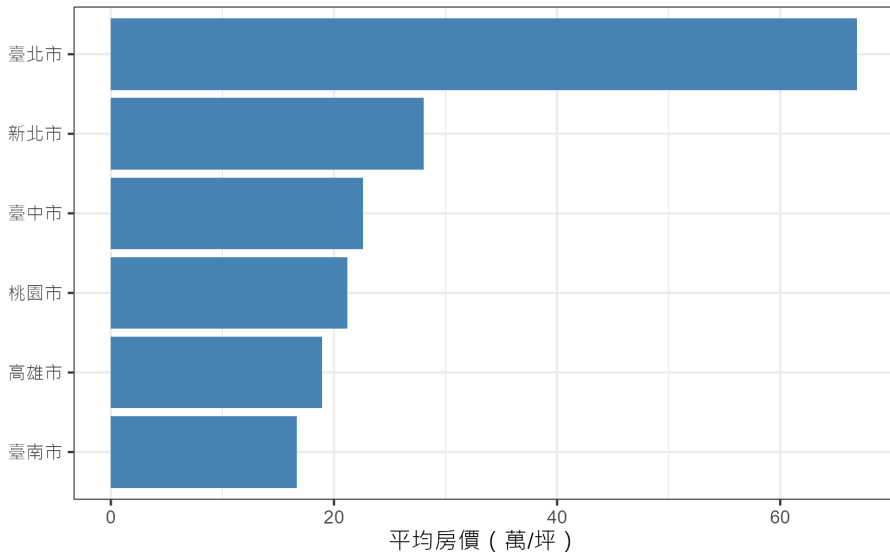
六都房價平均時序趨勢



六都房價排名變化（每坪平均價）

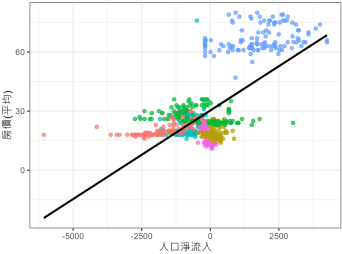


六都整體平均房價排名

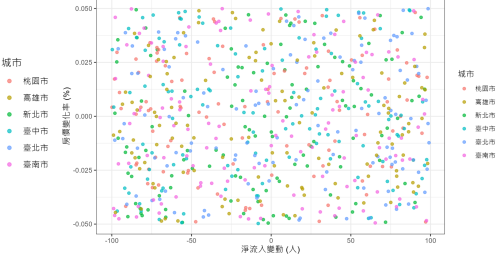


人口淨流入 vs 房價 (六都)

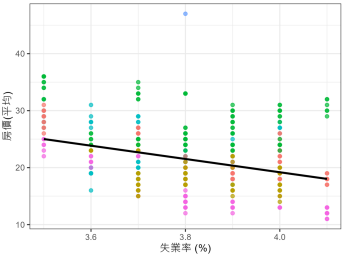
⑤ 人口淨流入 vs 房價 (六都)



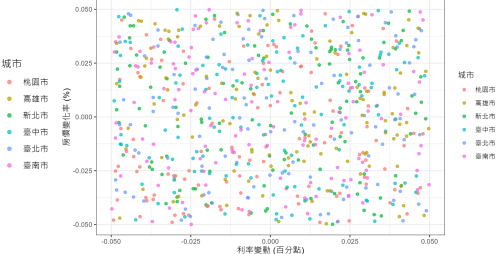
⑨ 房價變化率 vs 淨流入變動 (六都)



⑥ 失業率 vs 房價 (六都)



⑦ 房價變化率 vs 利率變動 (六都)



目錄

1 簡介

2 全域分析

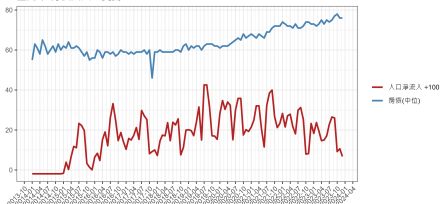
3 區域分析

4 迴歸分析

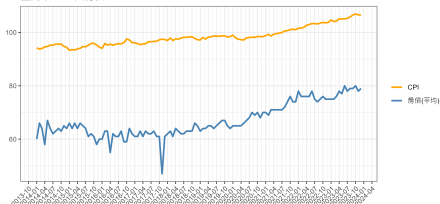
5 成果總結

房價、人口淨流入、CPI 走勢

臺北市 房價與人口變動



臺北市 CPI 與房價



分別對六個直轄市繪製房價、人口淨流入時間序列圖

- 房價大幅上升：台南、高雄、台中
- 人口流出：新北、台中、台南
- 人口流入：桃園、台北、高雄

時間序列分析

預測六個直轄市未來 5 年的房價走勢

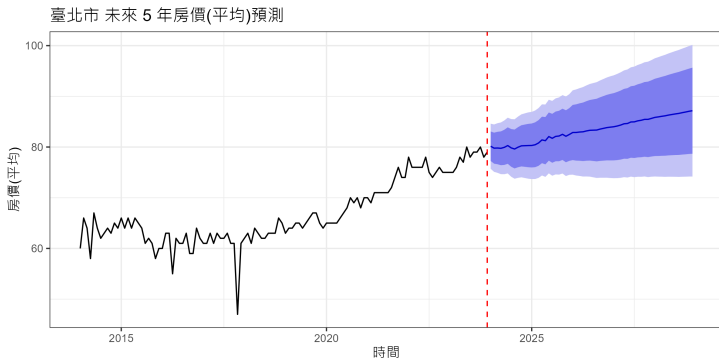
■ ARIMA 模型

- Autoregressive: 過去 p 期預測本期
- Integrated: $\nabla X_t = X_t - X_{t-1}$
- Moving Average: 過去預測誤差來修正預測
- 純時間序列預測方法
- 僅依據歷史房價進行建模

■ 模型預測能力評估與異常檢測

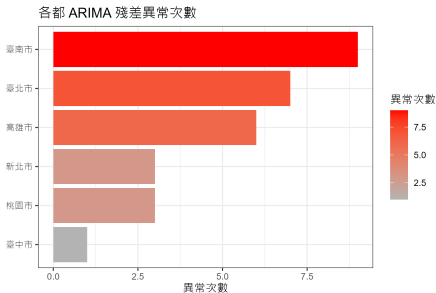
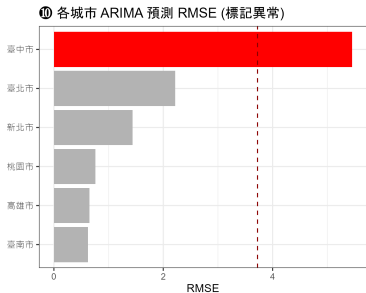
- Root Mean Square Error
- $\sum \mathbf{1}(|e_i| > 2 \cdot s_e)$

平均房價預測



- 大幅上漲：桃園、台南、高雄
- 些微上漲：台北
- 持平：新北、台中

異常檢測



- 台中為 RMSE 唯一被標記為異常的案例
- 台南、台北、高雄的殘差異常次數明顯偏高

目錄

1 簡介

2 全域分析

3 區域分析

4 迴歸分析

5 成果總結

初始模型

Model

$$\text{房價} = \beta_0 + \beta_1 \text{淨流入} + \beta_2 \text{CPI} + \beta_3 \text{利率} + \beta_4 \text{年所得} + \beta_5 \text{失業率} + \varepsilon$$

Response variable:

- 平均房價

Explanatory variable:

- 人口淨流入
- 消費者物價指數 (CPI)
- 貸款利率
- 年所得
- 失業率

初始模型

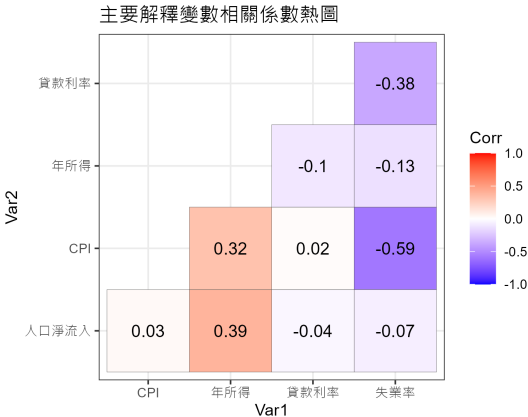
Model

$$\text{房價} = \beta_0 + \beta_1 \text{淨流入} + \beta_2 \text{CPI} + \beta_3 \text{利率} + \beta_4 \text{年所得} + \beta_5 \text{失業率} + \varepsilon$$

Results of model fitting

項目	估計值	標準誤	t 值	p-value
(截距)	-76.2313	14.3441	-5.3145	<0.01
人口淨流入	0.0048	1.86×10^{-4}	25.6642	<0.01
CPI	-0.2069	0.0851	-2.4322	0.0153
年所得	7.19×10^{-5}	1.27×10^{-6}	56.8116	<0.01
貸款利率	9.2466	1.0587	8.7326	<0.01
失業率	4.1926	1.8243	2.2982	0.0218

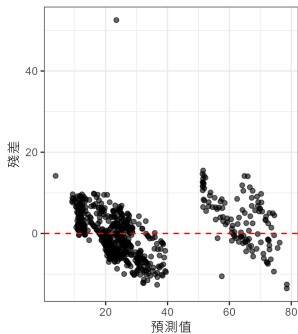
共線性診斷



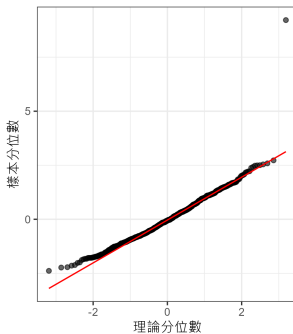
變數	VIF	等級
人口淨流入	1.2074	低
CPI	1.8327	低
年所得	1.3434	低
貸款利率	1.2865	低
失業率	1.9822	低

殘差診斷

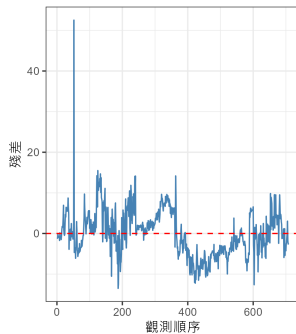
(a) 殘差 vs. 預測值



(b) 殘差常態 Q-Q 圖



(c) 殘差 vs. 順序 (檢查自相關)



- (a) 非隨機分布：異質變異、缺適性
- (b) 極端處偏離：少數離群值或偏態
- (c) 週期性波動 & 高頻震盪：自我相關

模型改進：異質變異

- Robust 標準誤
- 修正殘差異質性所造成的估計偏誤

Results of model fitting

項目	估計值	標準誤	t 值	p-value
(截距)	-76.2313	16.5120	-4.6166	<0.01
人口淨流入	0.0048	2.0×10^{-4}	23.2980	<0.01
CPI	-0.2069	0.0948	-2.1819	0.0295
年所得	1.46×10^{-6}	1.27×10^{-6}	49.3405	<0.01
貸款利率	9.2466	1.1907	7.7657	<0.01
失業率	4.1926	2.0987	1.9977	0.0461

模型改進：自我相關

- 廣義最小平方法 (GLS)
- AR(1)
- 針對每個城市的時間序列做殘差修正

Results of model fitting

項目	估計值	標準誤	t 值	p-value
(截距)	-101.0467	26.8699	-3.7606	<0.01
人口淨流入	5.3×10^{-4}	2.6×10^{-4}	2.0563	0.0401
CPI	0.4035	0.2355	1.7132	0.0871
年所得	5×10^{-4}	8×10^{-6}	6.7157	<0.01
貸款利率	1.2739	3.5421	0.3596	0.7192
失業率	4.3339	2.2865	1.8954	0.0584

模型改進：固定效果模型

- 資料兼具時間與橫截面維度
- Fixed effects model 能更準確控制不隨時間改變的特性

城市固定效果

城市	FE 係數
台北市	59.4255
新北市	20.0447
桃園市	15.3976
台中市	14.7730
台南市	7.8096
高雄市	10.4778

月份固定效果

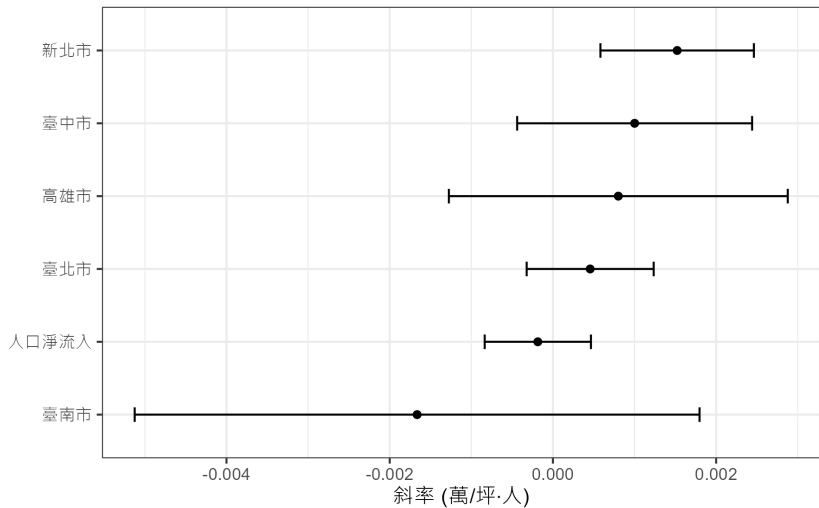
時間	FE 係數
2014-01-01	15.0187
⋮	
2020-01-01	20.3819
⋮	
2023-12-01	16.3361

模型改進：固定效果模型

- Panel Two-way Fixed Effects 模型假設各城市僅在固定效果（截距）上存在異質性
- 進一步引入「人口淨流入 × 城市」執行聯合線性假設檢定
- H_0 ：斜率一致 vs H_a ：斜率不一致
- $F=2.72$ ($p\text{-value}=0.0192$)
- 結論：不同城市中，人口淨流入對房價的影響力存在結構性差異。因此，在 panel data 結構下考量斜率異質性是有必要的。

模型改進：固定效果模型

各城市『人口淨流入 → 房價』斜率比較



模型選擇

模型比較

Model	R^2	AIC
初始模型	0.8992	4487.2972
固定效果模型	0.9740	3542.4489
GLS with AR(1)	NA	3787.2197

模型預測能力交叉驗證

Model	RMSE	MAE	R^2
初始模型	5.7180	4.4361	0.9003
lm_opt	7.5235	5.8960	0.8259
固定效果模型	2.7154	1.5869	0.9744

目錄

1 簡介

2 全域分析

3 區域分析

4 迴歸分析

5 成果總結

研究發現

綜合分析

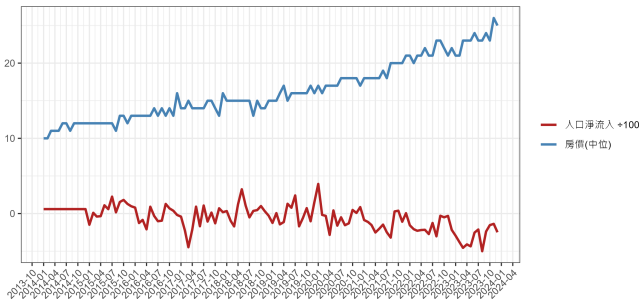
- 時間序列 (ARIMA)、複迴歸、Panel 固定效果模型
- 分析六都房價與多項社經變數 (人口、CPI、所得、利率) 的關係

主要發現

- 房價與人口流動呈正相關，但城市間存在顯著差異。
- 固定效果模型證實「城市差異」與「月份固定效果」皆顯著存在。

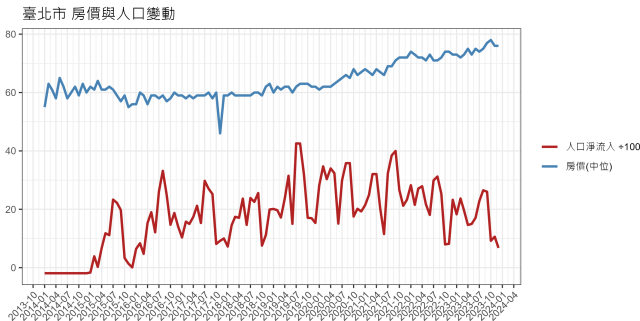
價格逆勢現象：台南市

臺南市 房價與人口變動



- 人口淨流入長期處於負值，房價卻穩定上升。
- ARIMA 預測異常最多的城市，嚴重預測失準。
- 人口對房價的迴歸斜率不顯著。
- 可能原因：區段開發吸引投資性購屋、價格受到炒作

結構性脫鉤：台北市



- 人口淨流入並未隨著房價趨勢上升。
- ARIMA 殘差異常次數高，房價難以預測。
- 人口流入迴歸係數不顯著。
- 可能原因：資金流動、土地有限、投資性購屋、豪宅市場運作

潛在政策建議

1 關注房價與需求脫鉤城市

- 強化實價登錄透明度
- 對短期持有課徵更高稅負

2 建立分區式房價預警系統

- 建立以 RMSE、殘差次數為基準的 ARIMA 異常警示系統
- 提供政府預警房市過熱的依據

3 導入多元解釋變數於都市發展評估

- 資金流動
- 基礎建設預期
- 公共運輸建設密度

Q & A