

591租屋爬蟲分析與應用-台北市整層住家



張晴晴

整體架構



1

資料蒐集

2

資料處理

3

建立迴歸預測模型

4

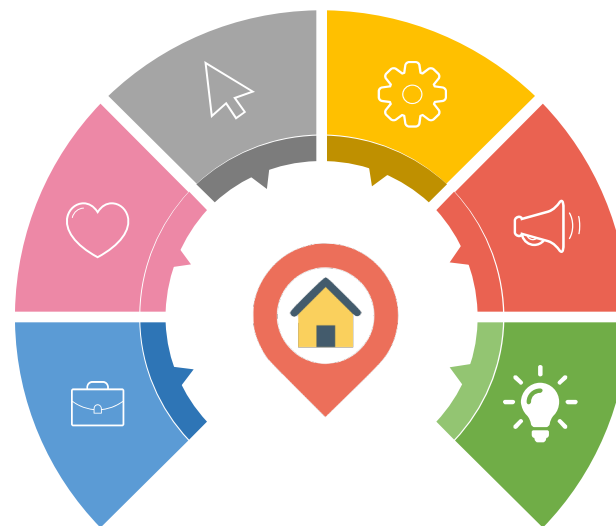
資料視覺化

房屋出租類型

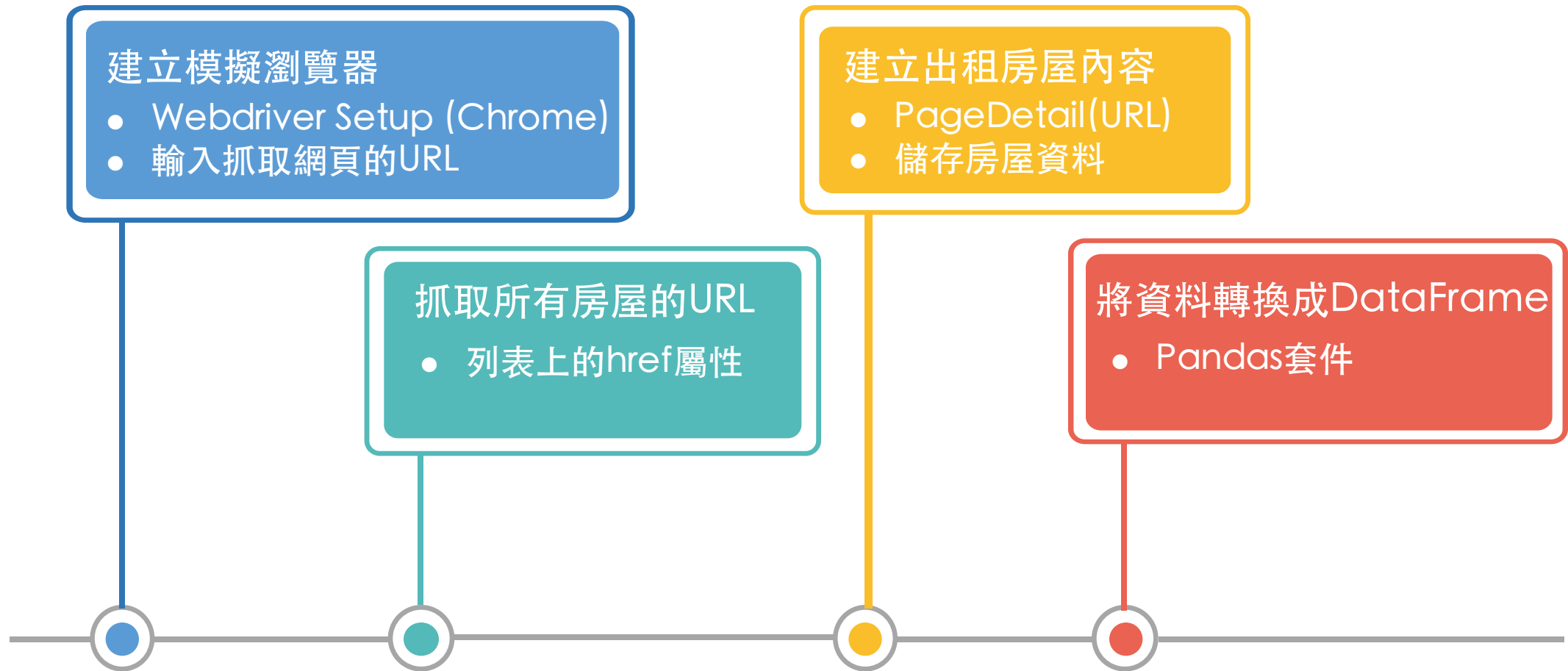
5

銀行應用情境

- 1) 增加轉貸客源
- 2) 房貸機器人
- 3) 房屋鑑價模型



資料蒐集 – Selenium動態爬蟲



資料處理



- 填補缺失值
- 過多欄位為空值即捨棄

- Dummy Variable
- 建立樓層差的變數



- 使用normalizePrice
轉換金額、坪數
9,000 元/月 -> 9000
6坪 -> 6

- 只選租屋類型為整層住家
- 建築類型排除倉庫、店面

資料筆數： 11,126 TOTAL ➡ 5,661 TOTAL

◆ 資料檔呈現

num	address	community	floor	housetype	identity	latitude	longitude	name		phone	pin	price	room	status
0	台北市 中山區 樂群二 路	大直美堤	8F/10F	電梯大樓	(屋主 聲明： 仲介勿 擾)	25.080000	121.552177	劉小 姐	//statics.591.com.tw/tools/showPhone.php?info_...		6.00	9000	5房0 廳4 衛1 陽台	雅房
1	台北市 中山區 吉林路 393巷29 號	NaN	5F/5F	公寓	(屋主 聲明： 仲介勿 擾)	25.065672	121.531527	鄭小 姐	//statics.591.com.tw/tools/showPhone.php?info_...		10.00	15500	NaN	獨立套 房
2	台北市 中山區 八德路 二段	台北科技大 學	頂樓加 蓋/5F	公寓	(屋主 聲明： 仲介勿 擾)	25.046456	121.539001	王先 生	//statics.591.com.tw/tools/showPhone.php?info_...		6.00	9800	NaN	獨立套 房

圖1.資料處理前

	latitude	longitude	price_num	pin_num	bedroom	livingroom	bathroom	balcony	Dfloor	Community	Apartment	Villa	Huaxia	Thorough	Elevator building
4	25.079166	121.511221	26800.0	35.0	2	2	1	1	0.0	0	1	0	0	0	0
5	25.108349	121.525257	125000.0	59.5	4	2	2	1	4.0	0	0	0	0	0	1
13	25.093676	121.517661	106000.0	65.0	1	2	1	1	0.0	0	0	0	0	0	1
21	25.088913	121.510113	30000.0	34.0	2	2	2	2	-1.0	0	0	0	0	0	1
23	25.084082	121.525360	27999.0	17.0	1	1	1	1	6.0	0	0	0	0	0	1

圖2.資料處理後

建置模型 – 7種迴歸預測模型比較



◆ 變數假設

依變數：租屋金額 (Y)

自變數：經度 (X_1)

緯度 (X_2)

坪數 (X_3)

房間數目 (X_4)

客廳數目 (X_5)

浴室數目 (X_6)

陽台數目 (X_7)

樓層差 (X_8)

住宅大樓 (X_9)

公寓 (X_{10})

別墅 (X_{11})

華廈 (X_{12})

透天厝 (X_{13})

電梯大樓 (X_{14})

Data
preparation

- 標準化
- 特徵選擇(Filter、Embedded)

Model
Comparison

- LinearRegression
- SGDRegressor
- Ridge
- LassoCV
- SVR
- ExtraTreesRegressor
- XGBRegressor

Model
Verification

- 參數調整
- 評估

◆ 相關係數圖

與租屋金相關性較高的前五名：

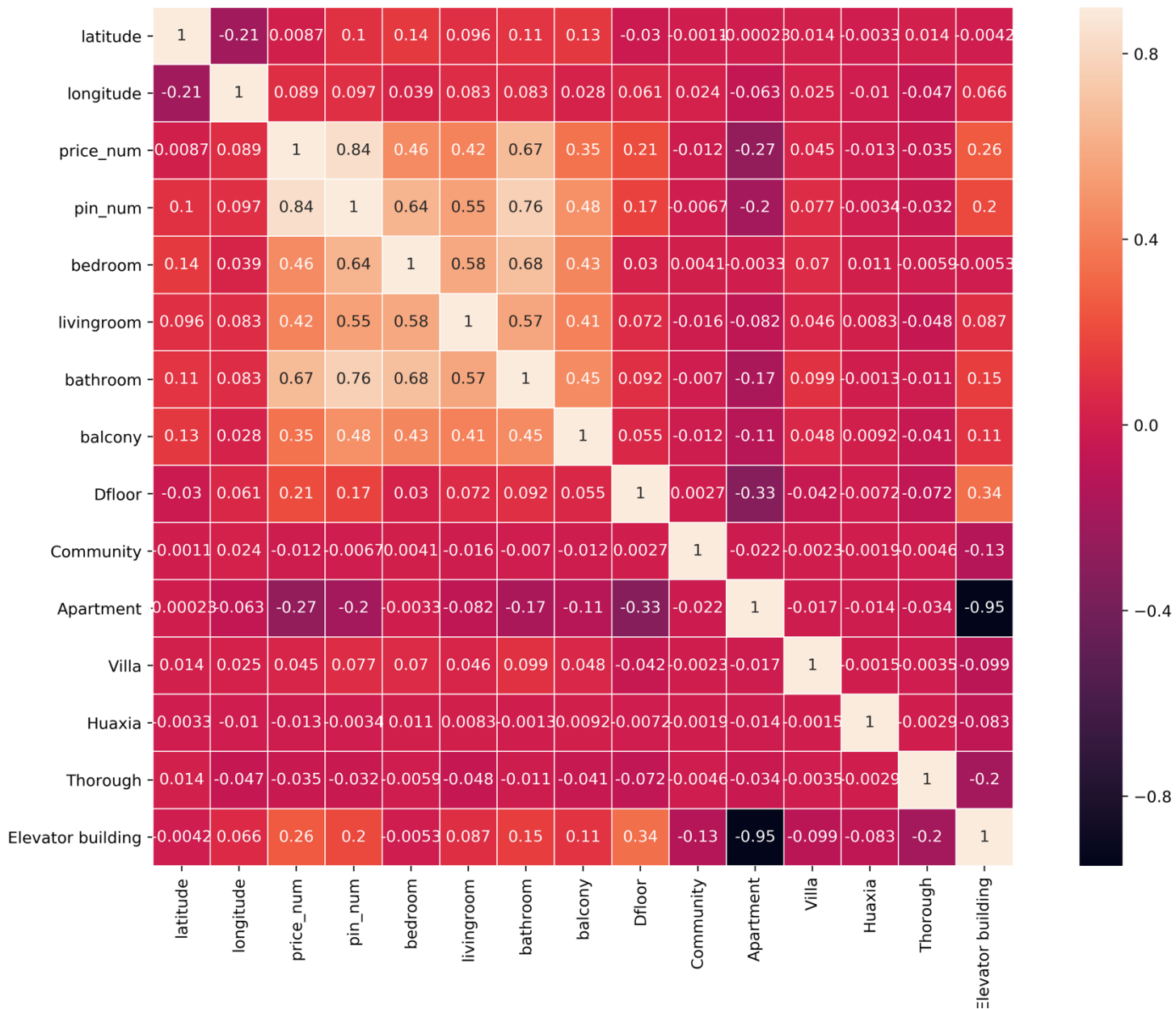
坪數 0.84

浴室數目 0.67

房間數目 0.46

客廳數目 0.42

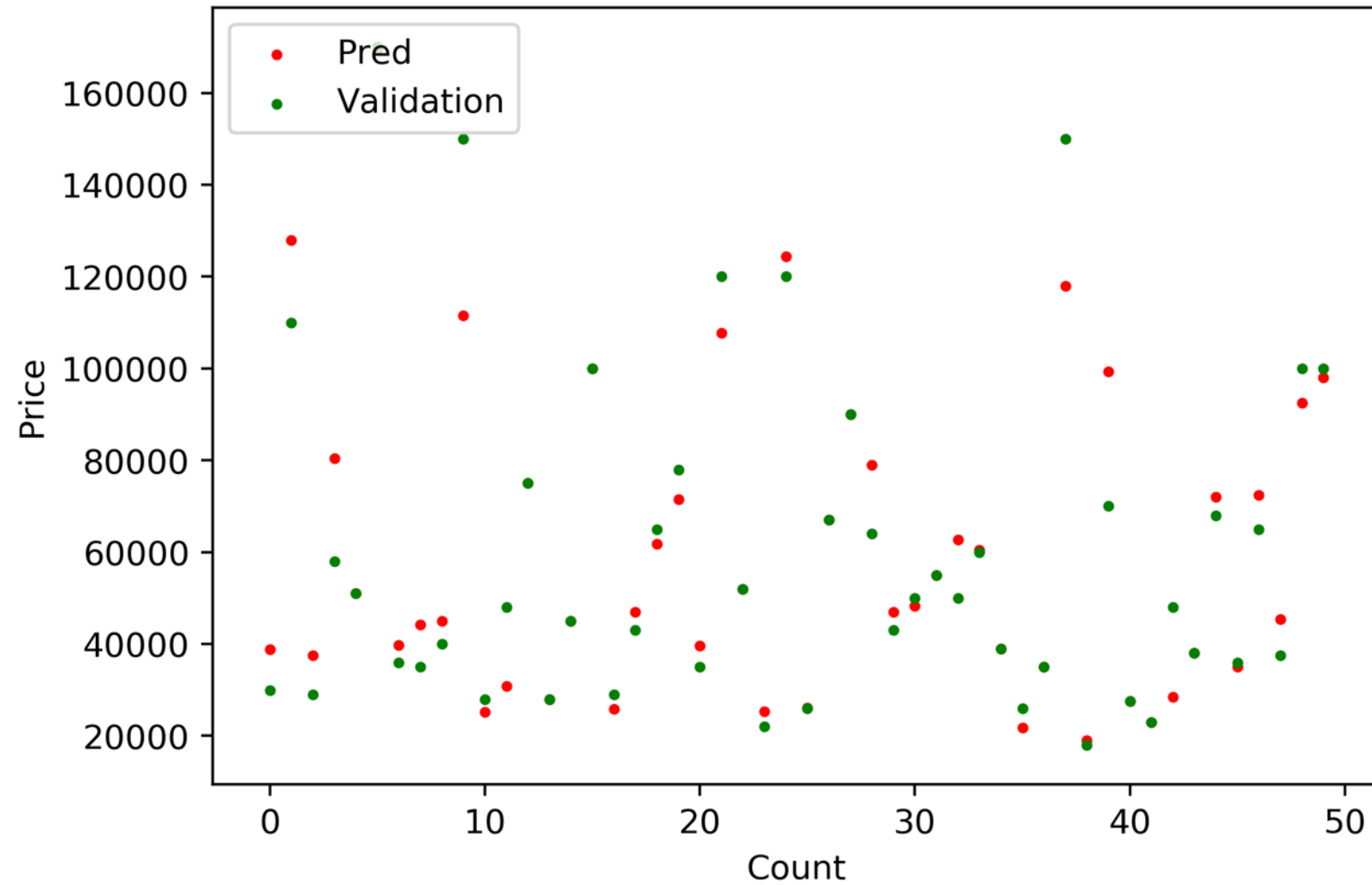
陽台數目 0.35



◆ 比較迴歸預測模型

預測模型	精確度	RMSE
LinearRegression	73%	0.58
SGDRegressor	73%	0.58
Ridge	73%	0.58
LassoCV	60%	0.72
SVR	70%	0.62
ExtraTreesRegressor	88%	0.39
XGBRegressor	82%	0.47

◆ 預測值、測試值的散佈圖(取前50筆資料)

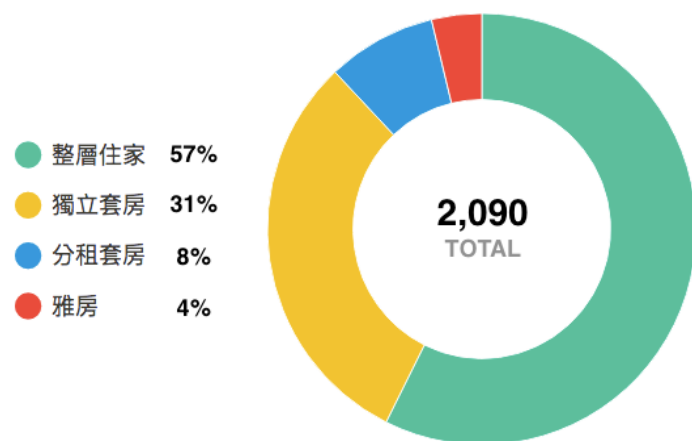


資料視覺化 — 台北市中山區



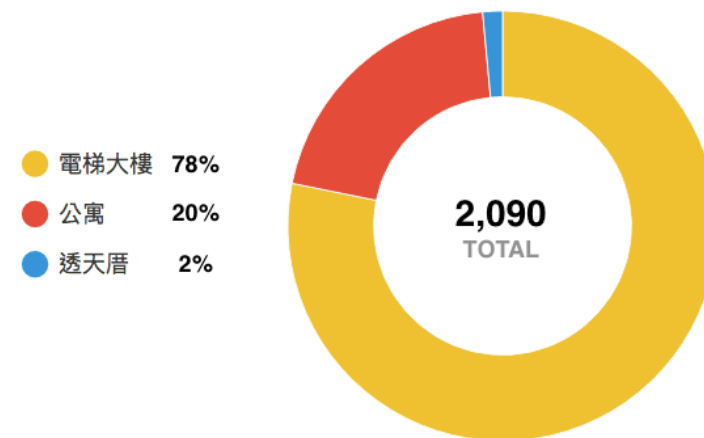
房屋出租總覽

◆ 房屋出租類型比例？



近六成的房屋出租類型為整層住家，約三成為獨立套房，近一成為分租套房，而出租雅房佔最少數。

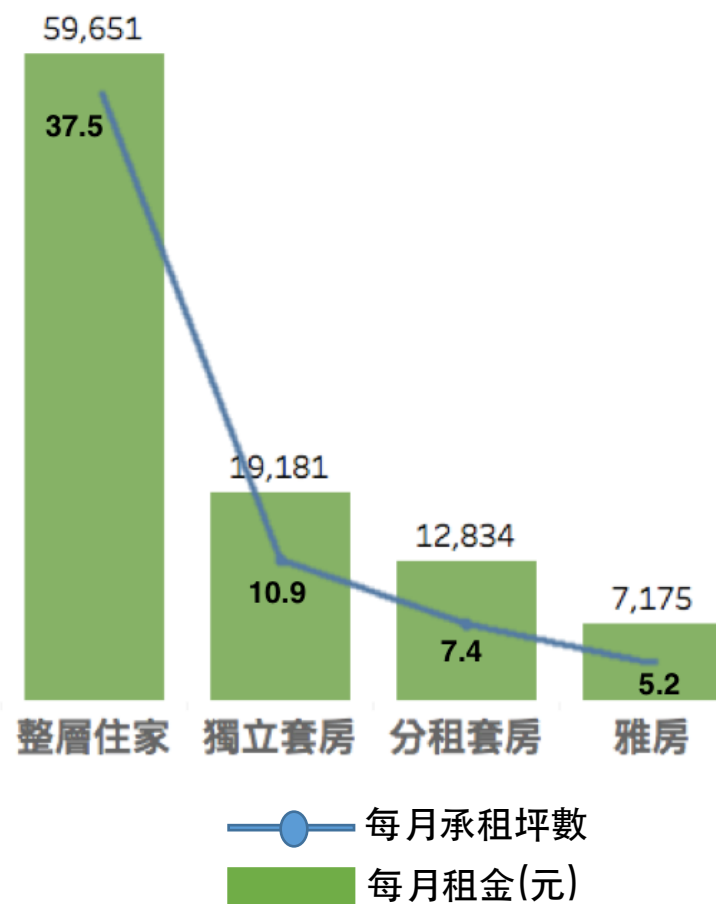
◆ 房屋出租的建築型態比例？



建築型態有近八成為電梯大樓，兩成為公寓，原因為台北市中山區的建築密集度高，建築的樓層數大多有5樓以上，因此以建電梯大樓居多。

◆ 房屋出租類型之平均承租坪數及租金比較

平均承租坪數及租金比較圖



- 根據左圖顯示，若要在中山區租一間整層住家，平均每月需負擔59,651元的房租；若要租套房，平均每月需負擔12,834元以上。

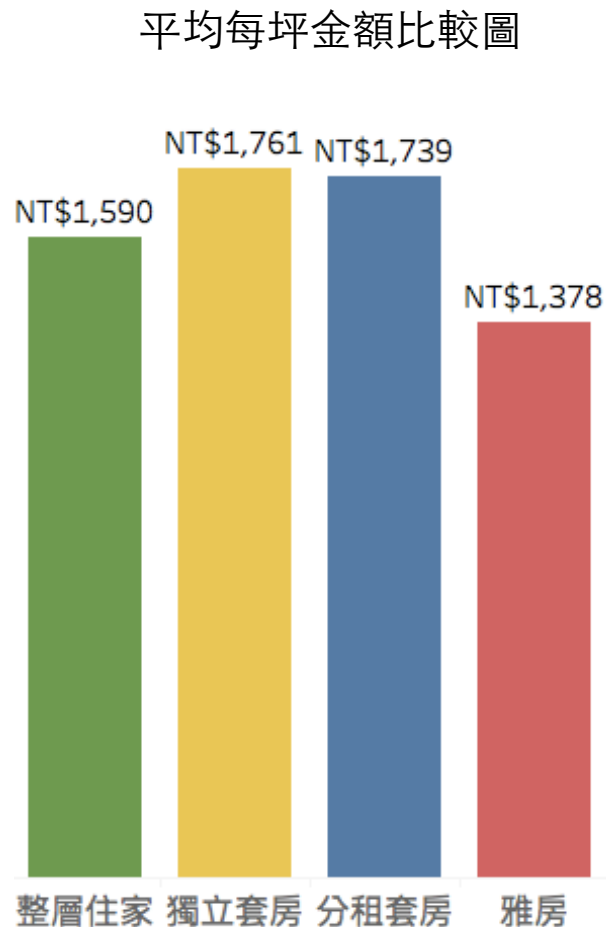
- 整層住家的租金報酬率 $\left(\frac{\text{每坪租金} \times 12 \text{個月}}{\text{房屋總價}} \right)$

以中山區出租最多的整層住家為例，其平均租金為59,651元，依據台灣地區平均住宅租金投資報酬率大約在2%~3%之間。

$$\begin{aligned} \text{房屋總價} &\rightarrow (59,651 \text{元} \times 12 \text{個月}) \div 3\% \sim (59,651 \text{元} \times 12 \text{個月}) \div 2\% \\ &\rightarrow \underline{2386 \text{萬元} \sim 3579 \text{萬元}} \end{aligned}$$

若房屋總價超過3579萬元，表示該房屋售價投資報酬率相對低，可能得重新檢視該房屋的條件。

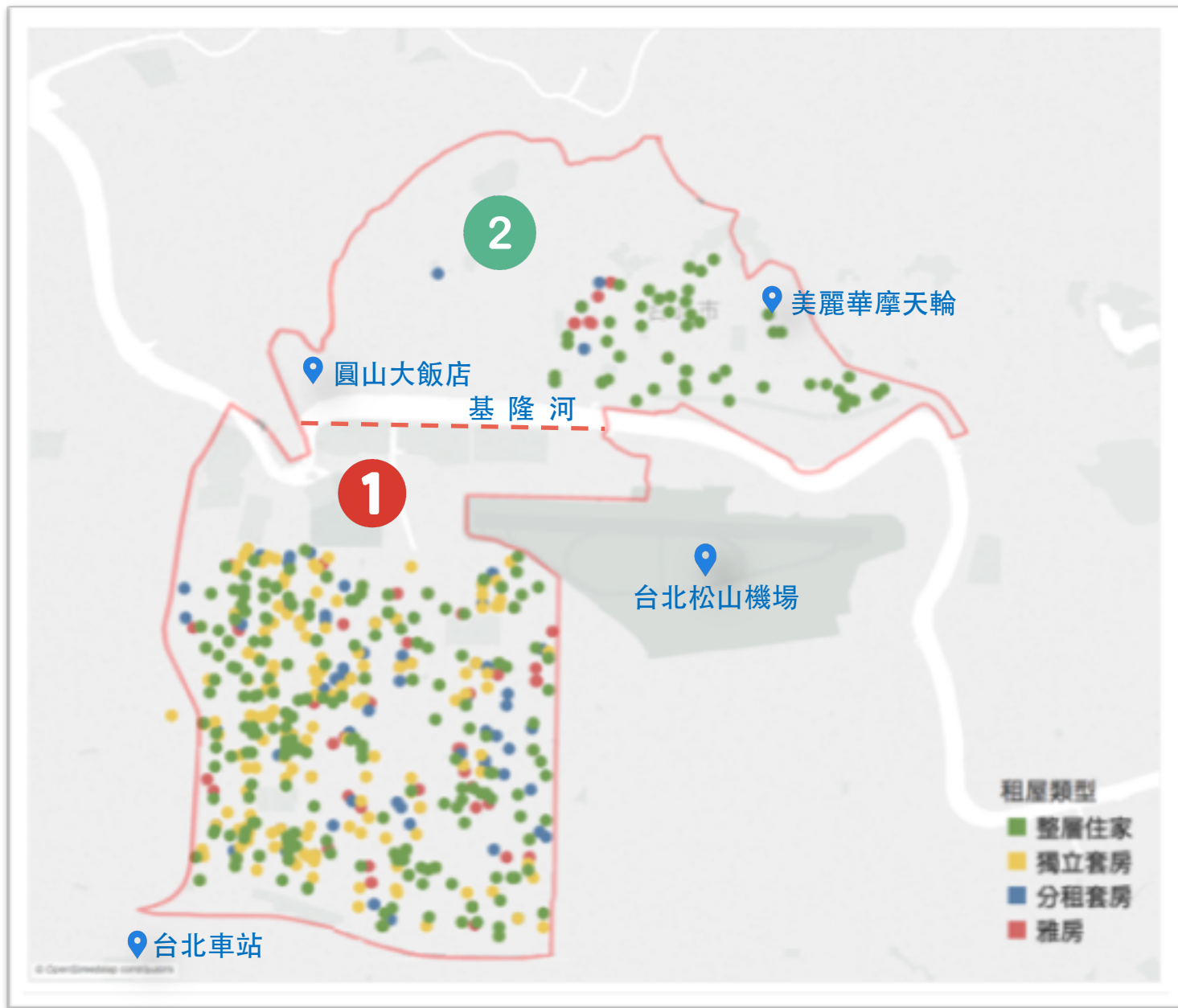
◆ 房屋出租類型之平均每坪金額



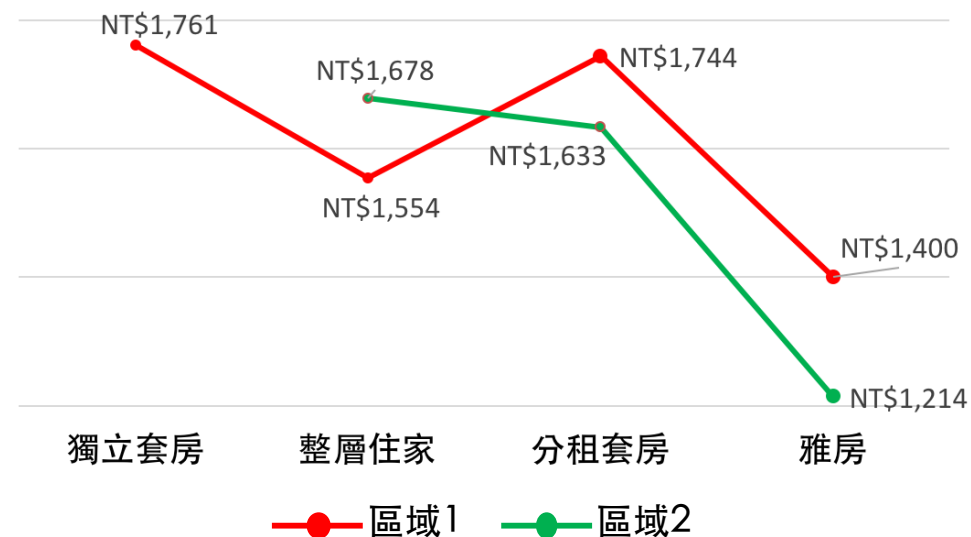
- 中山區的獨立套房和分租套房的每坪金額差異不大，兩者的平均每坪金額皆在1,700以上，而雅房的平均每坪金額最低。
- 整層住家房屋租售比 $\left(\frac{\text{每坪房價}}{\text{每坪租金}} \right)$
同樣以中山區的整層住家為例，每坪租金為1,590元
根據591房屋出售交易行情，台北市中山區整層住家的每坪房價66.46萬/坪
房屋租售比 $\rightarrow 664,600 \div 1,590 = \underline{418}$ 倍
一般認為租售比為300倍內尚算合理，即代表若打算買一間中山區的整層住家，目前仍不適合，選擇租屋會比買屋划算。



中山區房屋出租類型分佈 - 依交通方便程度將中山區劃分為以下兩區域



區域1和2的每坪金額



- 區域2以出租整層住家為主，且區域2整層住家的每坪金額高於區域1，進一步觀察，發現區域2大多數的整層住家包含車位或泳池等公設，因此每坪金額較高。

銀行應用情境



1) 增加轉貸客源

假設每位屋主出租的房屋皆有貸款，且每月的房貸金額與每月的平均出租金額收入相差不大，並結合房屋市場價格、目前房貸餘額、貸款年限…等，以此計算出利率較低的轉貸優惠專案，吸引屋主轉貸。

屋主的平均出租金額

單位：元

建築型態	租屋類型	整層住家	獨立套房	分租套房	雅房	整體
電梯大樓		50,609	17,115	12,012	7,300	25,277
公寓		24,953	12,159	11,263	6,953	12,724
透天厝		28,499	11,857	12,000	5,200	12,343
整體		39,757	14,878	11,464	6,886	18,175



EX:

以屋主整體的平均出租金額18,175元做為某位屋主的每月房貸金額，並假設目前貸款金額為500萬、30年期、房貸利率1.9%。

若該名屋主每月同樣還款18,175元，藉由房貸試算(本息平均攤還法)可以推出25年期、房貸利率為1.5%的轉貸優惠方案。

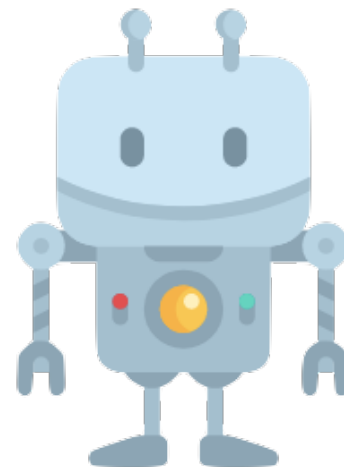
2) 房貸機器人

每月固定從591租屋網抓取全台灣各地區的出租資料(包含房屋出租類型、出租金額…等)，以計算每個地區每月的租金報酬率或房屋租售比，並根據地區或房屋類型等顯著的變數，藉由迴歸分析方法建立一個能夠預測房價趨勢的模型，進一步推估出適合買房的時機點，以及目前最適合買房的地區在哪裡，並藉由訊息推播以增加客源。

台北市中山區房屋價格下滑XX%，適合買房時機，貸款利率最低享有XX%起。

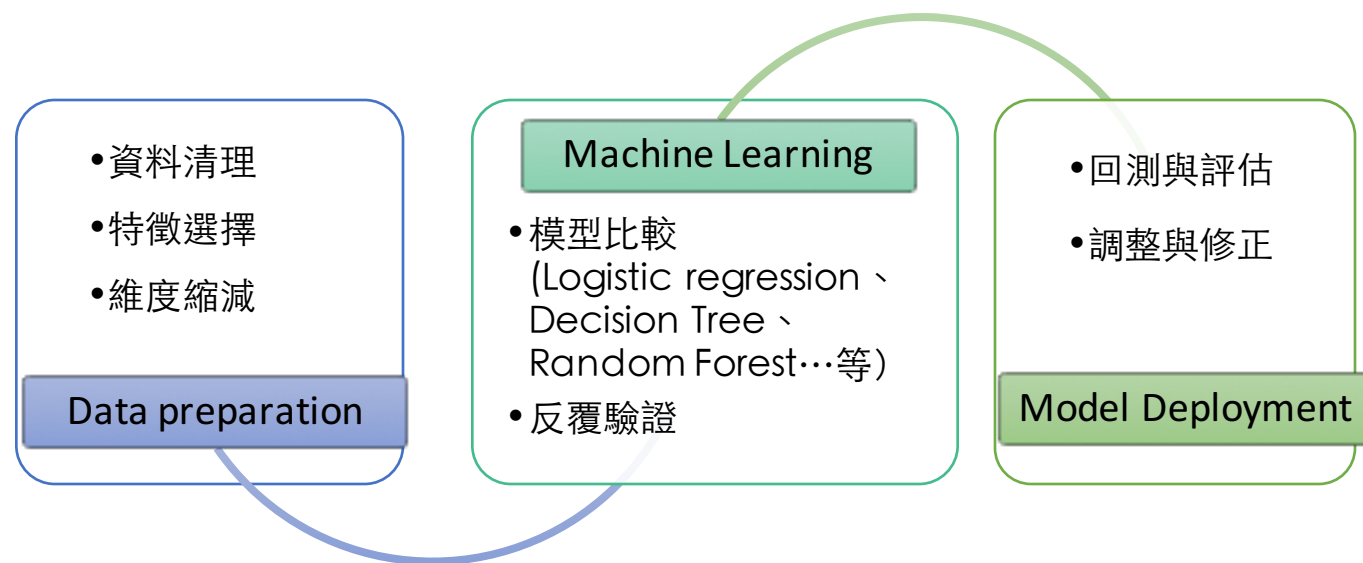


+



3) 房屋鑑價模型

蒐集更多房屋總價的open data，如：591、信義房屋、永慶房屋的出售交易行情，並考慮建築型態、屋齡、街道、生活機能等物件型態，**建立房屋鑑價模型，以分群分類的方式提高精準鑑價，減少客戶對銀行鑑價或貸款成數不如預期的情形。**



THANKS⁺