

iPhone 的二手市場 -

以 PTT Macshop 版 為例



109-1ECON-DSSI

2021.1.11



陳柏瑜 R08323004 r08323004@ntu.edu.tw  
高翊傑 R07323031 r07323031@ntu.edu.tw  
鍾新敏 B06302157 b06302157@ntu.edu.tw



# Outline



- Core Question
- Data Description
- Descriptive Statistics
- - ×
    - 「女用機」比較值錢？
  - ✓
    - Estimating the second-hand iPhone Demand
    - Estimating the Demand Curve via IV Regression
    - The probability of successfully selling an iPhone 6s

# Core Questions





# Outline

- 「女用機」比較值錢？
- Estimating the second-hand iPhone Demand
- The probability of successfully selling an iPhone 6s
- Analysis based on the geographical distribution of second hand
- iPhone selling and buying

# Data Description





## Data Description

- 使用網路爬蟲，爬取：價格資料、是否成交、地理位置、文字描述、產品型號、交易時間
- 用IV regression估計Demand curve
- 將是否為女用機作為dummy variable加入回歸式，以係數估計是否顯著異於0，判斷刻板印象是否正確
- 使用文字探勘爬取地理位置資訊，觀察不同地區是否存在價格差異

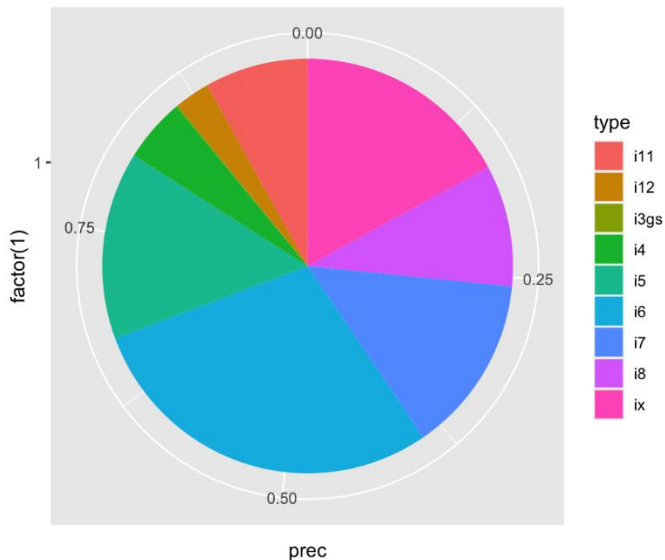
# Descriptive Statistics





## 圖一 iPhone各個型號占總貼文數的比例

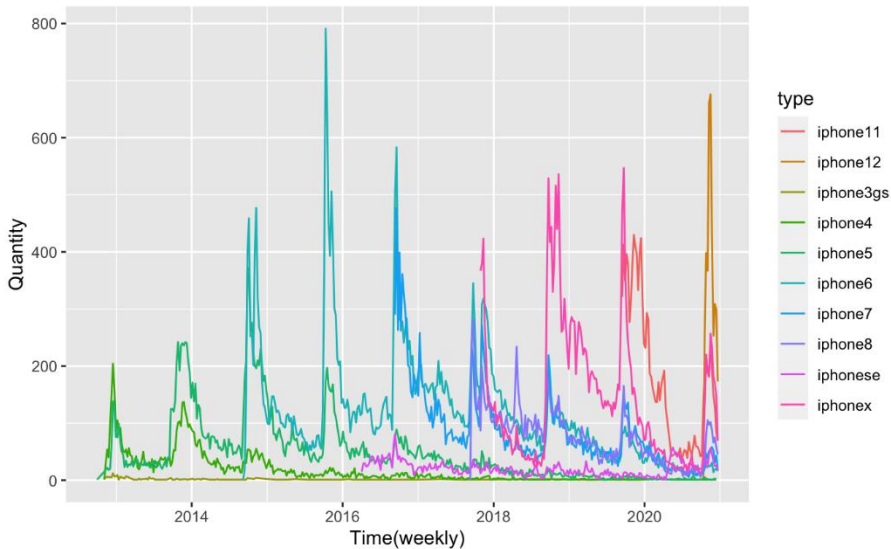
可知清理完後的資料共有144848筆，其中以iPhone6/6s的件數最多，有40386筆，約占28%。第二多的是ix/ixs/ixr，約占17%。







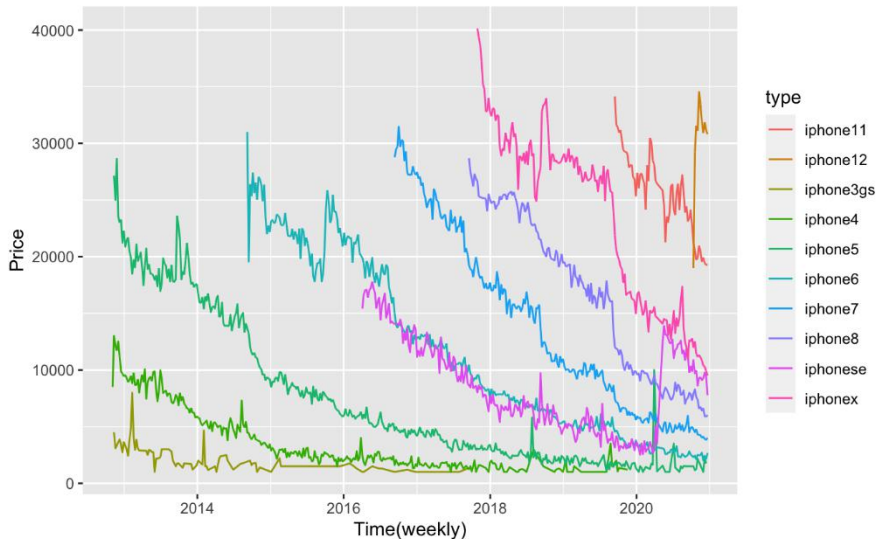
X軸為時間趨勢，Y軸為貼文數量，依顏色區分不同的型號(type)，可以看出剛推出新品時，上一代手機的貼文數會突然大增，可見舊機換新機到風氣。





### 圖三 iPhone不同型號時間與價格走勢

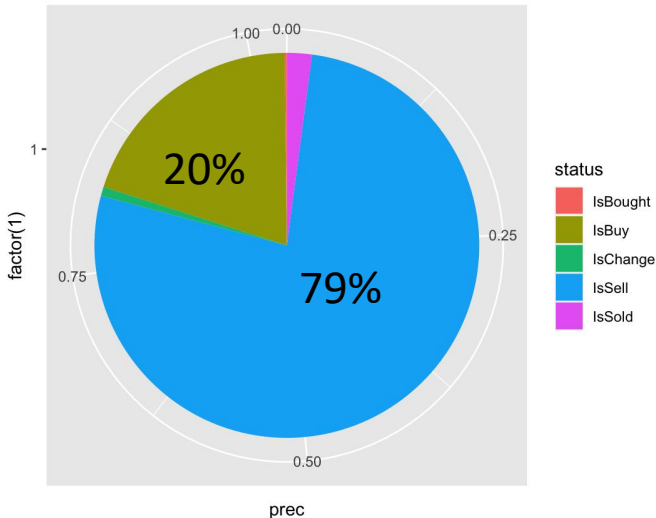
x軸為時間趨勢，y軸為價格，依顏色區分不同的型號(type)，可以看出不論哪種 iPhone 型號，推出新品後一年約折價1萬元，一年之後折價速度會減緩。



## 圖四 iPhone不同交易狀態占總貼文數的比例 (IsBuy, IsBought, IsSell, IsSold, IsChange)



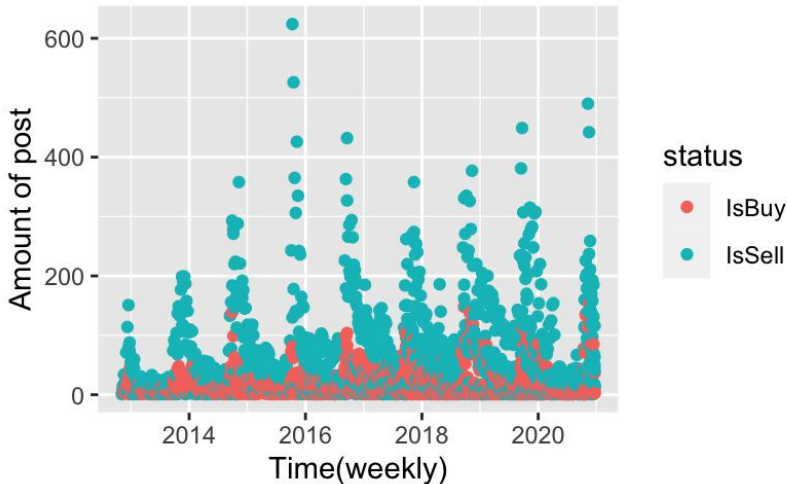
可知以買/賣的貼文數最多，但已成交(已售/已購買)的件數卻非常少。





## 圖五「出售」、「購買」數量與時間的走勢

其中，再以占大部分貼文數的「出售」以及「購買」來畫圖，可看出賣的人出價比買的人高很多。



Results



## 「女用機」比較值錢？

- 欲出售 iphone 的貼文內標注「女用機」或「女生用」
- 「女用機」真的可以賣得比較貴嗎？
- 將此 “premium ”稱為 “gender rent ”



## 「女用機」比較值錢？

在 163950 則交易 iPhone 的貼文

- 標注「男用機」：83 則；標注「女用機」：2922 則
- 需要區別「成交價格」以及「非成交價格」

考量線性迴歸模型：

$$price_i = \beta_0 + \beta_1 D_{female,i} + \beta_2 D_{male,i} + \gamma W + u$$

where  $W$  is the control variable and  $W$  contains:  
*TimeUsed, ROM, color*





# 「女用機」比較值錢？

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	6.6765e+04	3.8878e+03	17.1730	< 2.2e-16	***
IsFemale	-5.2049e+03	5.8177e+02	-8.9466	< 2.2e-16	***
IsMale	-6.4808e+03	2.8031e+03	-2.3120	0.02084	*
TimeUsed	-3.9740e-05	2.6729e-06	-14.8675	< 2.2e-16	***
ROM	8.7480e+01	2.4696e+00	35.4230	< 2.2e-16	***
color_紅	3.9561e+03	6.8944e+02	5.7382	1.037e-08	***
color_灰	2.6941e+03	4.6818e+02	5.7544	9.428e-09	***
color_金	2.5873e+03	3.6616e+02	7.0661	1.909e-12	***
color_藍	9.7082e+03	1.5156e+03	6.4054	1.696e-10	***
color_綠	9.4491e+03	1.1865e+03	7.9641	2.220e-15	***
color_玫瑰	1.0511e+03	4.8286e+02	2.1768	0.02956	*
color_銀	2.0776e+03	4.6752e+02	4.4440	9.099e-06	***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1





## 「女用機」比較值錢？

- 與預期相反，在所有型號中且是「已售出」或「已徵得」的 iPhone 中，標注有「女用機」的貼文的成交價格是比較低的。
- 若是限制在 iPhone 6s 這個機型的話，則沒有顯著地異於零，但大致的方向仍是負的，意味著標示「女用機」並沒有辦法「提升價格」。

在 PTT 的文化中，更常見的是嘲諷標注女用機的貼文者



## 「女用機」比較值錢？

考量線性迴歸模型，並將樣本限制在成交的 iPhone 6s 貼文中：

$$price_i = \beta_0 + \beta_1 D_{female,i} + \beta_2 D_{male,i} + \gamma W_i + u_i$$

where  $W$  is the control variable and  $W$  contains:

*TimeUsed, ROM, color*



## 「女用機」比較值錢？

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t )							
(Intercept)	21343.68352	1480.37045	14.4178	< 2.2e-16	***						
IsFemale	-457.14882	1492.68059	-0.3063	0.7596							
TimeUsed	-15.22156	0.55273	-27.5388	< 2.2e-16	***						
ROM	50.24563	8.26413	6.0800	3.39e-09	***						
color_灰	-572.08213	1215.11728	-0.4708	0.6381							
color_金	941.82397	1267.18374	0.7432	0.4579							
color_玫瑰	1200.32805	1205.09409	0.9960	0.3200							
color_銀	-294.11541	1313.01866	-0.2240	0.8229							
---											
Signif. codes:	0	'***'	0.001	'**'	0.01	'*'	0.05	'.'	0.1	' '	1



# Definition

- Demand function characterizes the quantity demanded  $q^d$  when given a price  $p$

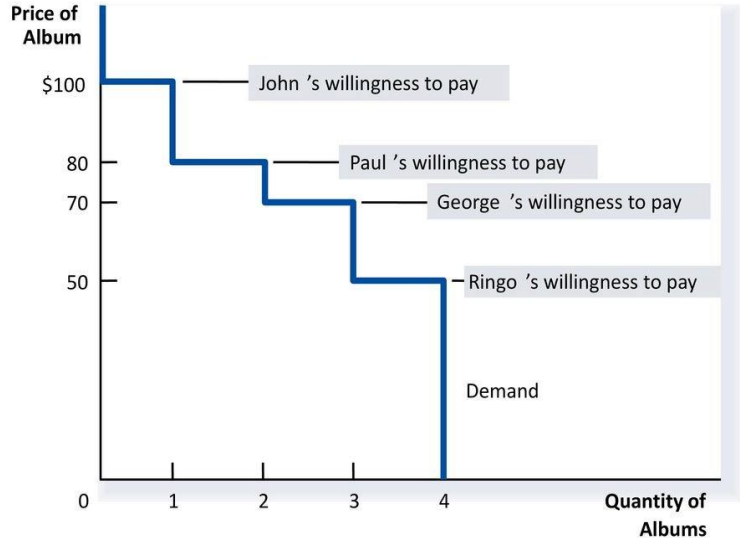
Notice that  $q^d$  is a function of  $p$

- The Inverse Demand function characterizes the price  $p$  when given a quantity demanded  $q^d$
- Economists usually use inverse demand to characterize the "Willingness to Pay"



# Definition

## The Demand Curve



## Estimating Demand?

- If we do not restrict the data on the SOLD and BOUGHT one, then we are not at the equilibrium.
- If we do not restrict the data on the SOLD and BOUGHT one, then we are not at the equilibrium.
- However, we have the label: Whether the poster is a potential buyer or seller.
- Thus, we're looking the Willingness to pay (WTP) for buyers. What about sellers? "Willingness to be paid" (WTBP)



## Estimating WTP and WTBP

若分別以 inverse supply function 以及 inverse demand function 的角度視之，我們可以分別寫下：

$$p_t = \alpha_0 + \alpha_1 q_t^s + \alpha W_t + u_t$$

$$p_t = \beta_0 + \beta_1 q_t^d + \beta W_t + v_t$$

where  $q_t^s, q_t^d$  分別代表時間  $t$  時，iPhone 6s 的供給以及需求



## Estimating WTP and WTBP

在此，由於我們有區別是在 supply side 或 demand side 的分類標籤，因此無須處理 simultaneous equation 的問題。當然，我們所估計的上面兩條迴歸式尚且不能稱之為供給及需求函數，但可以作為此二函數的近似。在此，我們尚且稱呼此二式為 Willingness To Be Paid (WTBP) 及 Willingness To Pay (WTP)





# Estimating WTBP

WTBP:  $p = 22346.26 + 20.54q^s$

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	22346.26298	543.25356	41.1341	< 2.2e-16	***
quantity	20.53982	5.76329	3.5639	0.000436	***
TimeUsed	-12.53909	0.35069	-35.7552	< 2.2e-16	***
ROM	3.27603	4.90388	0.6680	0.504706	
Is_6s_plus	160.45261	289.86773	0.5535	0.580381	
---					
Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' ' 1



# Estimating WTP

WTP:

$$p = 23691.78 - 8.54q^d$$

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	23691.7842	954.1712	24.8297	< 2e-16	***
quantity	-8.5358	55.3531	-0.1542	0.87762	
TimeUsed	-12.4746	0.7538	-16.5489	< 2e-16	***
ROM	-18.4653	8.3369	-2.2149	0.02799	*
Is_6s_plus	1134.3397	552.4897	2.0531	0.04147	*

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1



# Definition

Recall the Introduction to Econometrics class:

- Quantity and Price are determined simultaneously in a system
- $q$  and  $p$  are endogenous variables in this system
- OLS fails without identification



## Definition

Take supply function estimation for example:

Supply side:

$$Q = \beta_1 P + \varepsilon$$

Demand side:

$$Q = \alpha_1 P + \alpha_2 \textit{Income} + u$$

where *Income* is an exogenous variable (determined outside the system).



# Definition

We cannot estimate  $\beta_1$  for the supply function via estimating the following regression model with OLS:

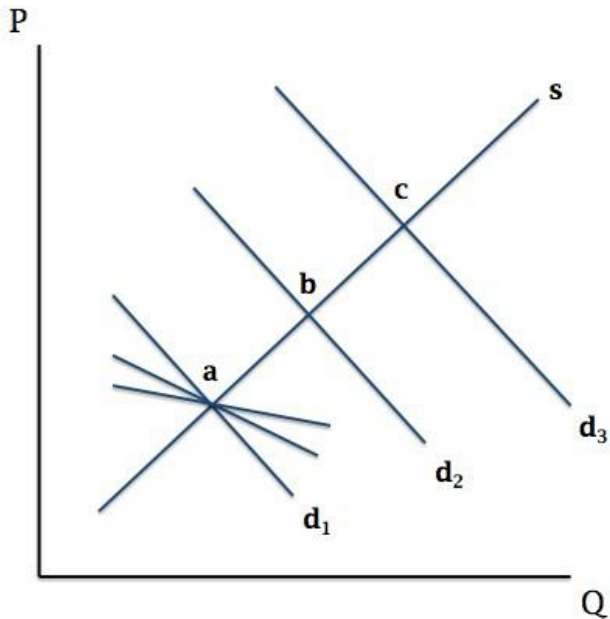
$$Q = \beta_1 P + \varepsilon$$

Why? Because of Endogeneity!

We need to find the demand shifter!  $\rightarrow$  Income



# Definition



# Definition

- A demand shifter only affects the quantity demanded, and do no effects on quantity supplied.
- Similarly, as long as we want to find the demand curve, we need to find a supply shifter.
- A supply shifter is simply an IV (Instrument Variable)



## IV: TSMC or Samsung?

- 我們 propose 一個 IV，它是「販賣或購買 iPhone 6s 的貼文內是否標注了『台積電/TSMC 晶片』或『三星/Samsung 晶片』」，因此我們的 IV 為兩個 Dummy Variable，分別以 *IsTSMC* 及 *IsSamsung* 稱之
- 由於 2015.9.25 發表 iPhone 6s 及 6s Plus 之後，隨即發生了「晶片門」事件





## IV: TSMC or Samsung?

我們真實想估計的 demand function (非 inverse demand function) 是

$$q_t^d = \beta_0 + \beta_1 p_t + \beta_3 W_t + \varepsilon_t$$

where  $W_t$  表示其他外生變數。但由於  $p_t$  存在內生性問題

$Cov(p_t, \varepsilon_t) \neq 0$ ，因此我們可以透過 *IsTSMC* 以及 *IsSamsung*

此二外生變數來 serve  $p_t$



## IV:TSMC orSamsung?

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	4284.6034	18815.6238	0.228	0.820
price	-0.1833	0.8156	-0.225	0.822
TimeUsed	-2.3286	10.1989	-0.228	0.820
ROM	0.1215	1.1386	0.107	0.915
Is_6s_plus	36.8174	167.6908	0.220	0.826

Diagnostic tests:

	df1	df2	statistic	p-value
Weak instruments	2	251	0.026	0.97427
Wu-Hausman	1	251	10.771	0.00118 **
Sargan	1	NA	0.013	0.90755

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1



## IV Regression

也就是 demand function 是：

$$q^d = 4284.60 - 0.18328p$$

移項得到 inverse demand function:

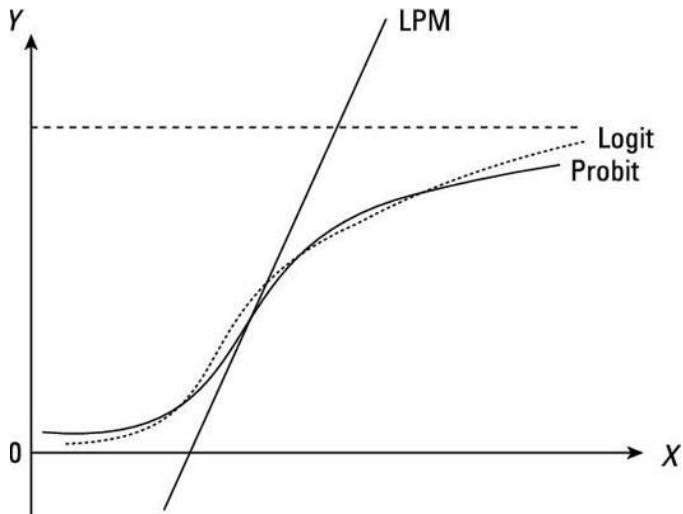
$$p = 23377.34 - 5.4561q^d$$

Recall the WTP<sub>estimated</sub>:

$$p = 23691.78 - 8.64q^d$$



# Binary Response Model



# Binary Response Model

以「是否賣出」為被解釋變數

$$IsSold = \beta_0 + \beta_1 price + \beta_2 ROM + \beta_3 i6sPlus + \beta_4 TimeUsed + \beta_5 D_{female} + \beta_6 color$$



# Binary Response Model:LPM

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	3.7299e-02	1.3766e-02	2.7095	0.006748	**
avg_price	-1.3428e-06	4.7938e-07	-2.8011	0.005101	**
ROM	5.6730e-06	4.9910e-05	0.1137	0.909505	
Is_6s_plus	1.0190e-02	3.5333e-03	2.8839	0.003935	**
TimeUsed	-1.0668e-05	9.0650e-06	-1.1768	0.239304	
IsFemale	-3.4201e-03	9.0412e-03	-0.3783	0.705230	
color_灰	1.0223e-02	7.0065e-03	1.4590	0.144582	
color_金	1.2327e-02	6.9306e-03	1.7787	0.075320	.
color_玫瑰	1.3615e-02	6.5062e-03	2.0927	0.036399	*
color_銀	9.1335e-03	7.3937e-03	1.2353	0.216739	

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1



# Binary Response Model:Probit

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	-1.7569e+00	2.4610e-01	-7.1389	9.407e-13	***
avg_price	-2.4706e-05	8.0708e-06	-3.0611	0.002205	**
ROM	1.3855e-04	8.2493e-04	0.1679	0.866625	
Is_6s_plus	1.6549e-01	5.6492e-02	2.9294	0.003396	**
TimeUsed	-2.0327e-04	1.4356e-04	-1.4159	0.156794	
IsFemale	-5.6058e-02	1.4617e-01	-0.3835	0.701349	
color_灰	2.1121e-01	1.5505e-01	1.3622	0.173143	
color_金	2.3960e-01	1.5243e-01	1.5719	0.115985	
color_玫瑰	2.6174e-01	1.4728e-01	1.7772	0.075532	.
color_銀	1.8686e-01	1.6171e-01	1.1555	0.247867	
---					

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1



# Binary Response Model:Logit

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	-3.1681e+00	5.9295e-01	-5.3430	9.144e-08	***
avg_price	-5.9276e-05	1.9055e-05	-3.1107	0.001866	**
ROM	2.2993e-04	1.9502e-03	0.1179	0.906144	
Is_6s_plus	4.0372e-01	1.3258e-01	3.0451	0.002326	**
TimeUsed	-4.8572e-04	3.3617e-04	-1.4449	0.148493	
IsFemale	-1.2858e-01	3.4595e-01	-0.3717	0.710136	
color_灰	4.8629e-01	3.8594e-01	1.2600	0.207661	
color_金	5.6200e-01	3.7943e-01	1.4812	0.138555	
color_玫瑰	6.1361e-01	3.6748e-01	1.6698	0.094962	.
color_銀	4.3925e-01	4.0170e-01	1.0935	0.274191	
---					
Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' ' 1





## Binary Response Model

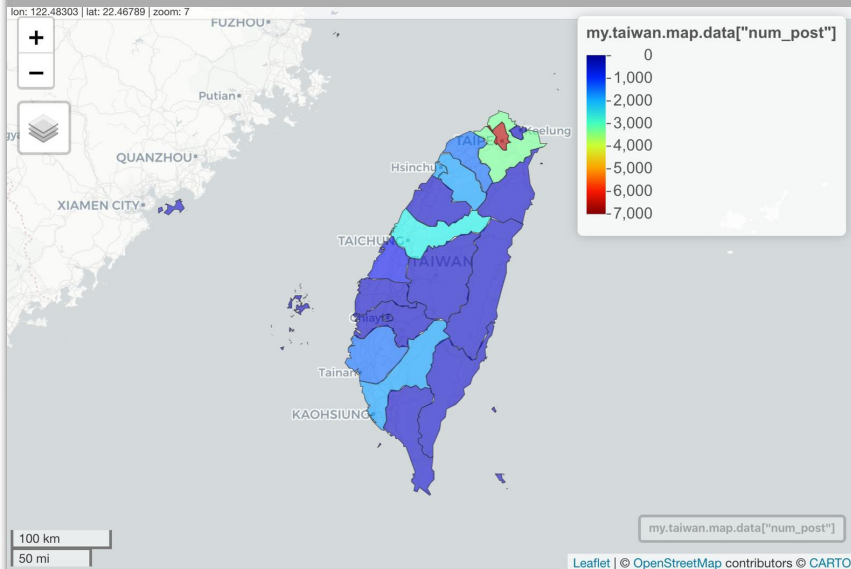
- 最重要的因素仍是價格:價格越低，成交的機率就越高
- 「玫瑰金」在 2015 年推出 iPhone 6s 時是新出的顏色



# Geographical Visualization

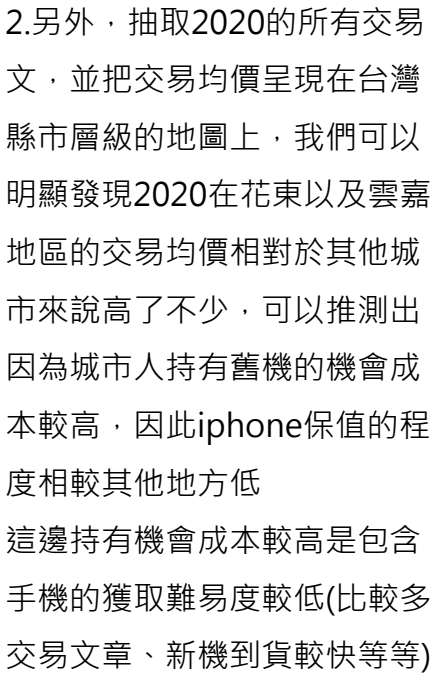


1.左邊可以看到的是台北市的面交地點相較其他地方的來得多很多，另外北北基、桃園面交地點也頗多，可以看得出來台北身為大都會的地位有擴散他新北市跟桃園市；此外其他直轄市也有許多面交地點。

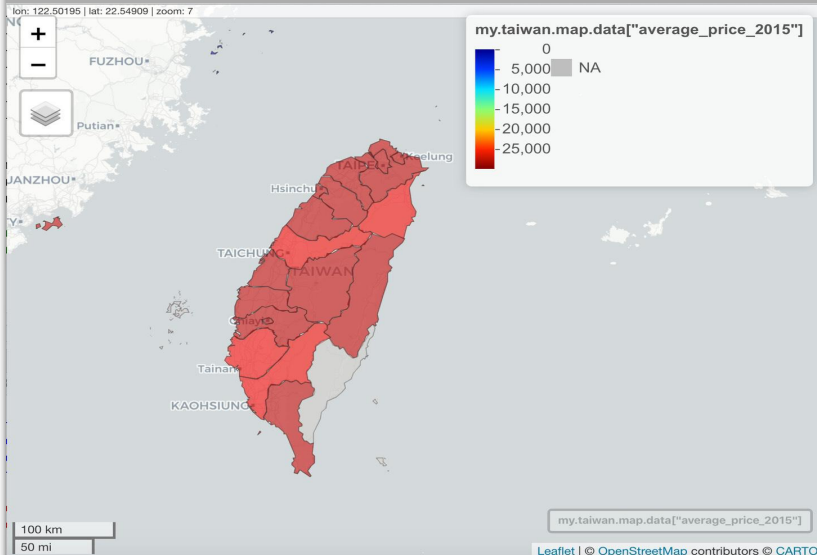


台灣各縣市iPhone 6s交易貼文數量圖





3.可以看到2015年時的二手iphone6s的交易市場各地方均價是差不多的



2015台灣各縣市iPhone 6s平均價格

