



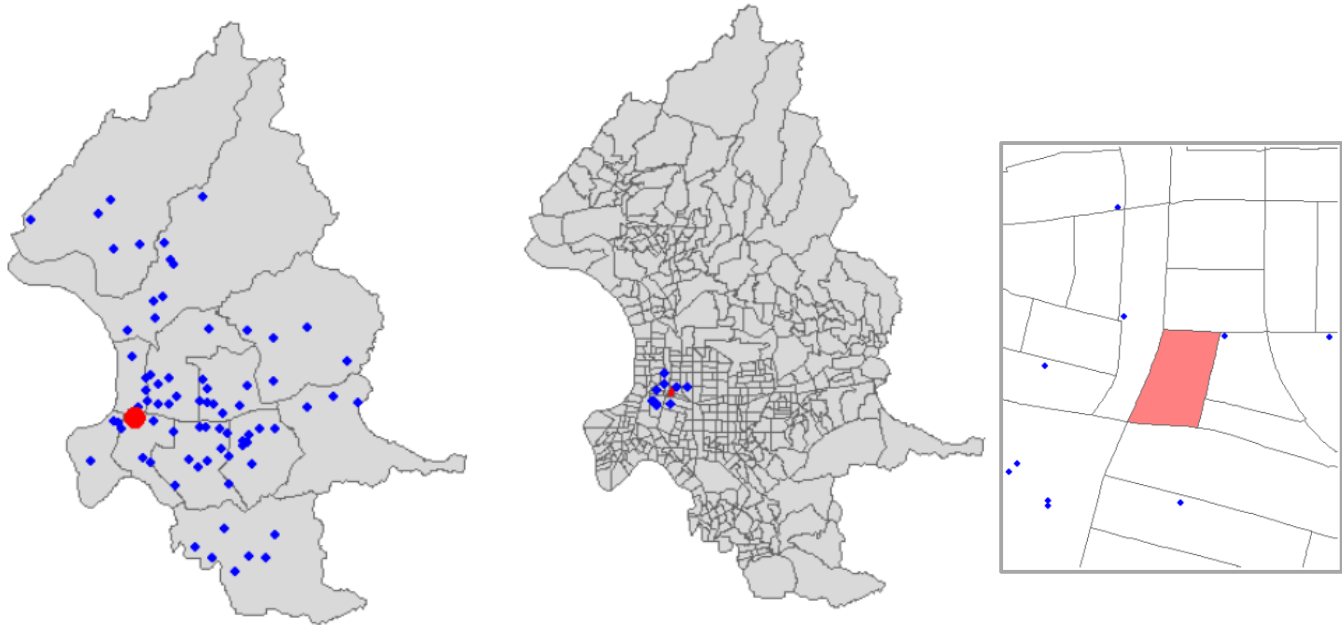
# R實作GIS

空間分析 2021.03.15  
TA 杜承軒

### 擷取麥當勞店家位置

1. 麥當勞 1 km為服務範圍內所涵蓋的麥當勞分店數，定義為該家麥當勞店家的連鎖密度。請問哪一家麥當勞的連鎖密度最高？繪製在地圖上，並標示該店家名稱。
2. 台北市各里中心點是否在涵蓋該麥當勞的服務範圍，作為判斷該麥當勞是否能服務到該里的標準。請問哪個里可被麥當勞服務的家數最多？繪製在地圖上，並標示該里的位置及可及的麥當勞店家。

(名稱以文字標示在Rmarkdown即可，位置標示在地圖上)



```

way 1. Buffer + Clip
way 2. Buffer + Intersect
way 3. Distance
way 4. WithinDistance(list)
way 5. WithinDistance(matrix)
→ 3/20 (Sat.) 公布參考答案

```

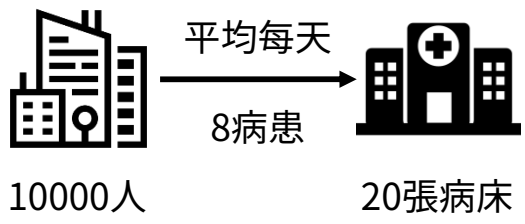
```
## [1] "麥當勞( ANS )"
```

```
## [1] " ANS "
```



# 隨堂考一

Q 1 : 7天中3天以上病床不足機率？



1. 一天病床不足的機率 → X: 一天病患人數 單位區間內的次數

$$X \sim B\left(n = 10000, p = \frac{8}{10000}\right) \rightarrow \text{Poisson}(\lambda = 8)$$

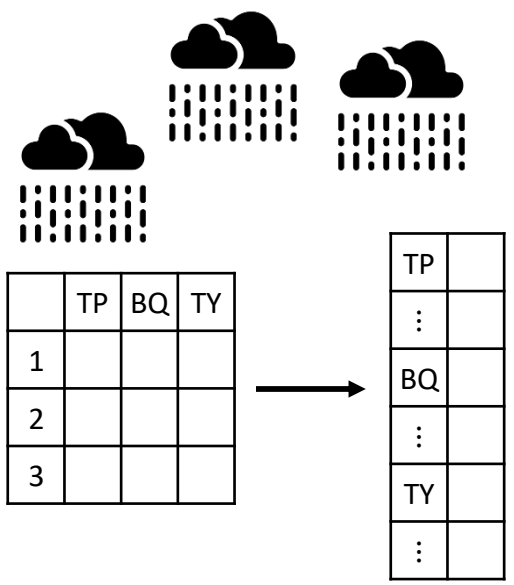
$$P_{one} = P(X > 20) = \text{pbinom}(20, n, p, F) \approx \text{ppois}(20, \lambda, F)$$

2. 七天中三天以上病床不足的機率 → W: 病床不足的天數

$$W \sim B(n = 7, p = P_{one})$$

$$\text{Pr} = P(W \geq 3) = \text{pbinom}(2, 7, p.one, F)$$

Q 2 : 三地降雨量有空間變異？



1. 假設檢定  
 $H_0$ : 台北、板橋與桃園三地平均降雨量皆相同 (無顯著空間變異)  
 $H_a$ : 台北、板橋與桃園三地平均降雨量有不同  
 顯著水準設定為0.05

2. 計算統計量、p value  
`ANOVA=aov(rain~city,Rain.data);summary(ANOVA)`

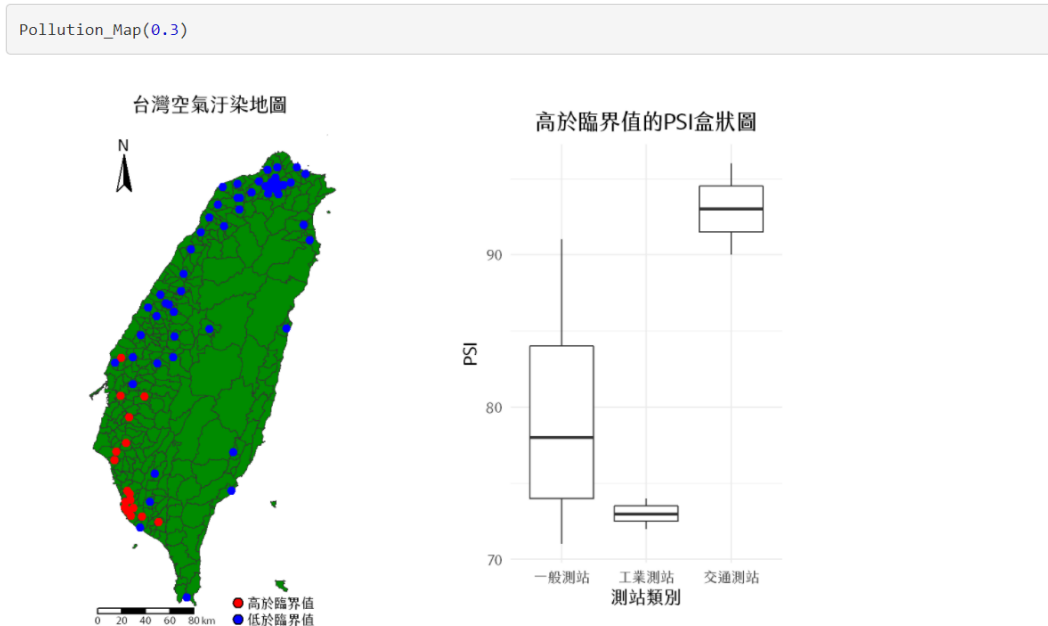
```
##           Df Sum Sq Mean Sq F value Pr(>F)
## city      2  10696    5348   0.346  0.71
## Residuals 33 509525   15440
```

(無母數ANOVA) `kruskal.test(rain~city,Rain.data)`

3. 結論  
 $p.value > 0.05$   
 因此不拒絕 $H_0$ ，表示降雨量在北台灣無顯著空間變異。

# 實習一

- 寄信、參考答案
- 上傳作業時間
- 段落(chunk)
- sf 格式
- 檔案路徑
- 避免人工資料篩選
- %in% 判斷式



## 地圖邊界

- `tm_shape(TW, xlim=c(146500,351000), ylim=c(2400000,2850000))`
- `TW.WGS84 = st_transform(TW,4326)`  
`tm_shape(TW.WGS84, xlim=c(120,122), ylim=c(21.5,25.5))`
- `tm_shape(EPA)+.....+tm_shape(TW)+.....`

## 自訂圖例

- `tm_shape()+.....+ tm_add_legend("symbol",labels=c("高","低"),col=c("red","blue"))`

## MAC中文繪圖

- `ggplot() + ..... + theme(text=element_text(family="黑體-繁 中黑"))`
- `tm_shape() + ..... +tm_layout(fontfamily = "Heiti TC Light")`

# GIS Spatial Processing

## Spatial Join

(Clip/Point-in-Polygon)

## Intersection

## Buffer

## Merge

## Centroid

## Distance

**sf1[sf2,]** : 選取sf1中，與sf2空間上有重疊 / 相交的部分

**st\_join(sf1, sf2, join = st\_intersects, ...)**

**st\_intersection(sf1, sf2)** : 截切 / 交集，並同時保留原本sf1及sf2的屬性

(1) 點×面→點：建議點放前面(sf1)

(2) 面×面→面

**st\_buffer(sf, dist)**

**st\_union(sf)**

**summarize(group\_by(sf, field))** : 依欄位合併

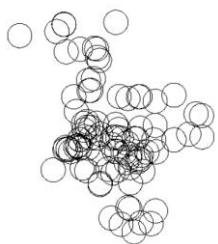
**st\_centroid(sf)**

**st\_distance(sf1, sf2)**

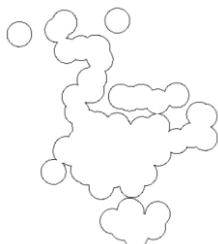
※**st\_distance(sf1, sf2) < set\_units(1,km)**

**st\_is\_within\_distance(sf1, sf2, dist, sparse = TRUE)**

T: sparse list; F: matrix



**st\_buffer(pts,1000)**



**st\_union(st\_buffer(pts,1000))**



**st\_union(TP)**



**summarize(group\_by(TP,TOWN\_ID))**

		sf2		
		0	1	2
sf1	1	128517.49	126907.54	120264.35
	2	123956.87	122449.94	116278.79
	3	29738.99	38177.79	46334.56
	4	243555.73	239797.71	224000.15
	5	47813.73	55654.62	63906.59

		0	1	2
1	FALSE	FALSE	FALSE	
2	FALSE	FALSE	FALSE	
3	TRUE	TRUE	TRUE	
4	FALSE	FALSE	FALSE	
5	TRUE	TRUE	TRUE	

**st\_distance(sf1,sf2)**

**st\_is\_within\_distance(sf1,sf2,10^5)**

# 資料處理

## droplevels

移除序列(factor)資料中沒有用到的級別(level)

## xtabs

樞紐分析表，用途同table()

**xtabs**(sum~group)

i.e. xtabs(~TOWN+STORE); xtabs(POP~COUNTY)

↑  
沒有參數→代表count

## left\_join

**left\_join**(x, y, by = c("name.x" = "name.y"))

1. 確認兩欄的格式要一樣 (事先型別轉換)
2. 配對的兩欄名稱不同 → 用by連接

```

> x
  id name
1  1  甲
2  2  乙
3  3  丙
> y
  id2 name2
1  1     A
2  2     B
3  4     D
> left_join(x,y,by=c('id'='id2'))
  id name name2
1  1  甲     A
2  2  乙     B
3  3  丙    <NA>
> full_join(x,y,by=c('id'='id2'))
  id name name2
1  1  甲     A
2  2  乙     B
3  3  丙    <NA>
4  4 <NA>     D

```

**left\_join :**

- x 被保留
- 只在x→NA值
- 只在y→消失

**full\_join :**

- x, y 都被保留
- 只在其中一者→NA值

# apply

- apply (X, MARGIN, FUN)

MARGIN:

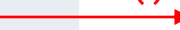
1 by row

2 by column

- lapply(LIST, FUN)

- mapply(FUN, arg1, arg2, ....)

```
> M
      [,1] [,2] [,3] [,4]
[1,]    1    3    5    7
[2,]    2    4    6    8
> apply(M,1,sum)
[1] 16 20
> apply(M,2,sum)
[1]  3  7 11 15
```

input format	function	output format
list data.frame array	apply	vector matrix
list data.frame	lapply	list  vector
parameters	mapply	vector matrix



1. 將實習所定義麥當勞的連鎖密度，建立chainstore(d)的自訂函數，可繪製服務半徑(d) vs.麥當勞的關係圖表。
2. 比較A區(文山+大安+中正)與B區(信義+南港+松山)的麥當勞連鎖密度：  
利用統計檢定方法，評估A區的平均每家麥當勞連鎖密度是否顯著高於B區。(服務半徑(d) = 1.5 km)  
(需列出虛無假設與對立假設，並說明檢定的顯著水準)。

「連鎖密度」：

當服務範圍為d公里時，服務範圍內所涵蓋的「**麥當勞分店數**」，定義為該家麥當勞店家的連鎖密度。

作答要求

Q1. 函數：chainstore(d)

輸入：d為半徑(公里)

輸出：折線圖；x軸請從0(km)開始，以0.5(km)為間隔，畫到d(km)

列出 **chainstore(3)** 的結果

Q2. 統計兩區差異

請完整寫出**假設、檢定方法、檢定標準、統計量、結論**等步驟

(P.S.可以視Fast\_Food的資料為麥當勞中的一部分抽樣，來推論兩區麥當勞的平均差異)