GGPLOT與資料處理

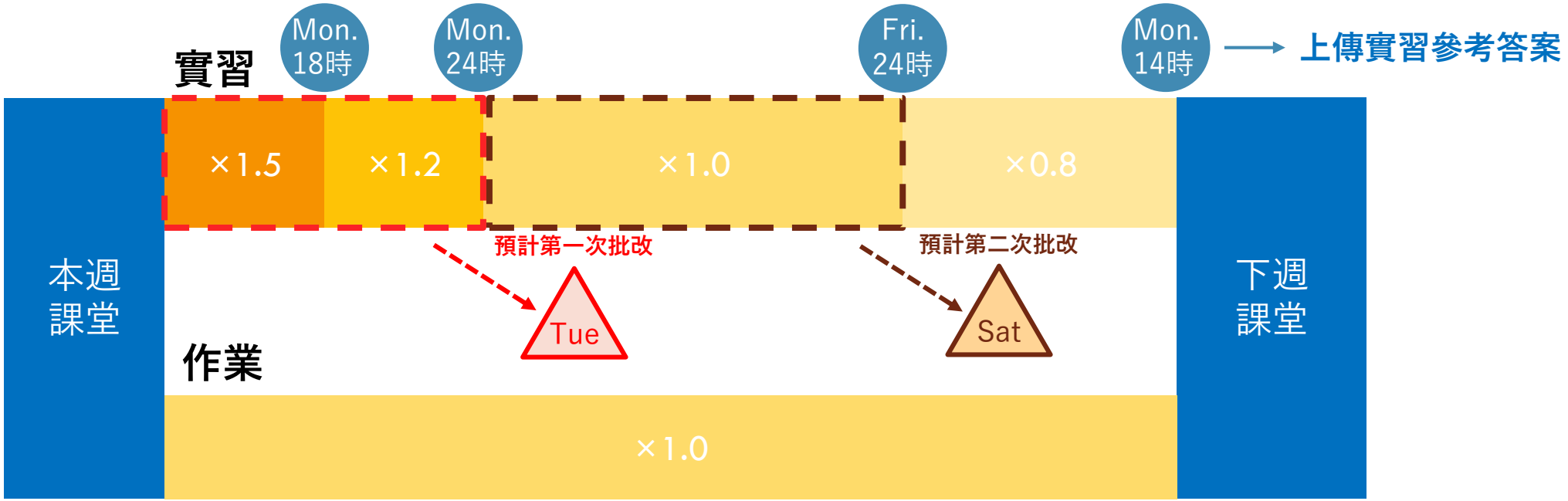
空間分析 2019.02.25 杜承軒

助教

杜承軒　r07228005@ntu.edu.tw　　地理系館501室

劉　恒　b03101049@ntu.edu.tw　　Office Hour : 寄信約時間

實習與作業繳交

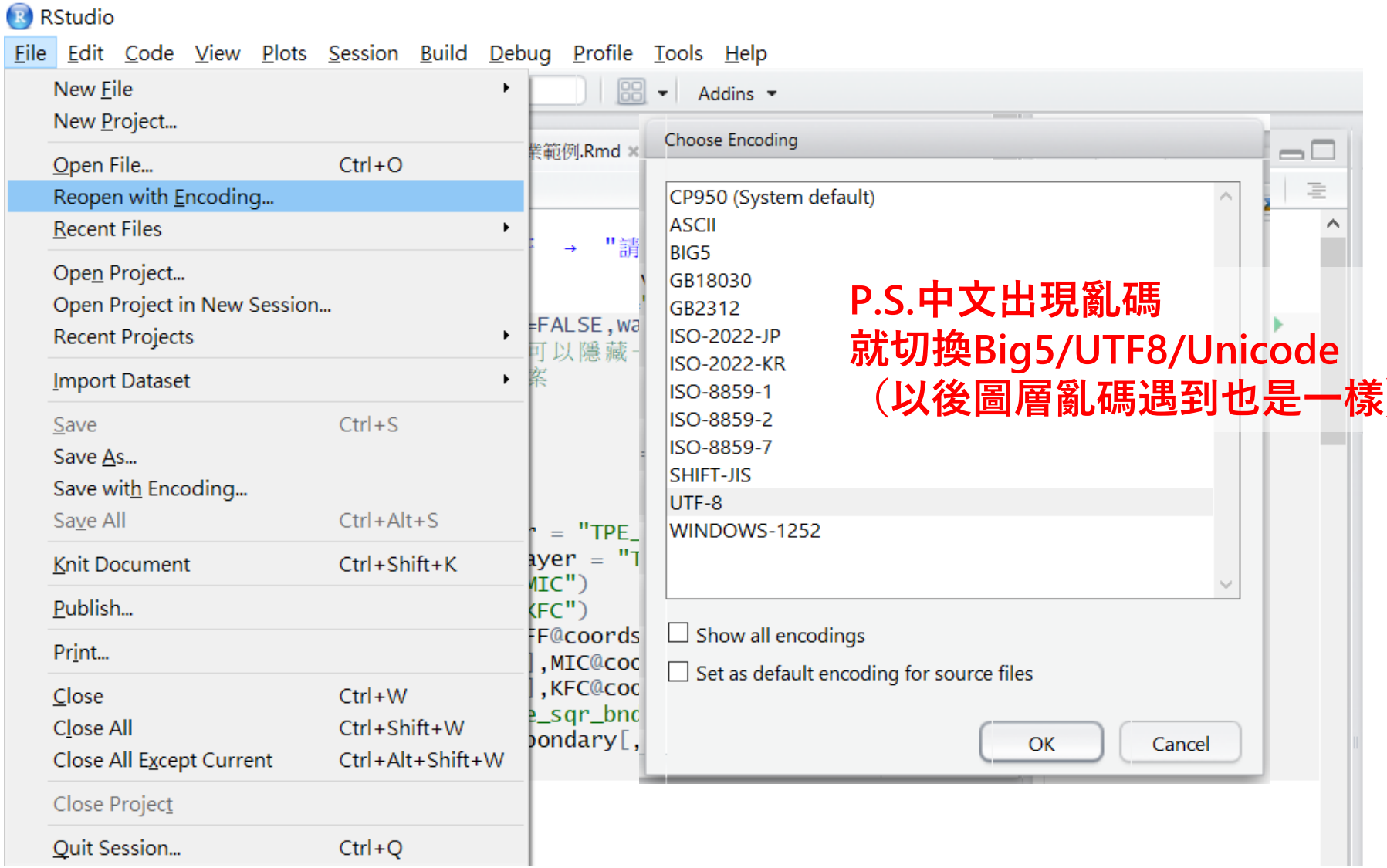


1. **繳交時程以講義和上述規定為主**
   (ceiba上的時間可能因為批改因素調動)
2. **請依規定格式繳交，不符規定予以扣分**
3. **一學期採前十高成績計算，遲交不計分**

RMarkdown 程式碼繳交格式

請搭配RMarkdown作業範例.zip

P.S. 範例中 .Rmd檔 如果亂碼， 請依以下步驟操作



RMarkdown
程式碼繳交格式

第一個R段落會先呼叫library和讀檔案
為了隱藏不需要的訊息可以使用 results/message/warning

```
```{r results='hide', message=FALSE, warning=FALSE}
library(GISTools);library(rgdal);library(splancs)
setwd("D:/1072SA/Data")
tpe=readOGR(dsn = ".", layer = "Vill",
            encoding="utf8",verbose = F)
```
```

路徑請盡量
避免使用中文
（善用捷徑）

讀完shapefile不會再回傳訊息

☺

```
#上面的results/message/warning可以隱藏一些
#可以在這個段落跑library和讀檔案
library(GISTools)
library(rgdal)
library(splancs)
setwd("D:/1072SA/Data") #路徑請盡量不使用
#讀取資料與格式轉換
tpe=readOGR(dsn = ".", layer = "Taipei_V
FF=readOGR(dsn = ".", layer = "Tpe_Fastf
```

```
$STORE=="MIC")
$STORE=="KFC")
ords[,1],FF@coords[
coords[,1],MIC@coor
coords[,1],KFC@coor
```

☹

```
library(splancs)
```

```
## Warning: package 'splancs' was built under R version 3.4.4
```

```
##
## Spatial Point Pattern Analysis Code in S-Plus
##
##  Version 2 - Spatial and Space-Time analysis
```

```
setwd("D:/1072SA/Data") #路徑請盡量不使用中文
#讀取資料與格式轉換
tpe=readOGR(dsn = ".", layer = "Taipei_Vill", encoding="utf8")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: ".", layer: "Taipei_Vill"
## with 456 features
## It has 9 fields
## Integer64 fields read as strings:  CENSUS FASTFOOD
```

P.S.
個人習慣把程式在R檔寫好測試後
才放到Rmd來做排版
再注意有沒有順序上的錯誤等等

RMarkdown 程式碼繳交格式

```
> Q1.
> 題號請用">"當前綴來標示

換行最後請空2~4格(讓它辨識你要換行)或打<br/>(支援html語法)

***

> Q2.題目通常都是畫圖和解釋

```{r}    #回答的時候請保留程式碼和結果以利批改
plot(tpe, col="#DDDDDD")
pointmap(MIC.pt, pch=21, bg="yellow", add=T)
pointmap(KFC.pt, pch=21, bg="red", add=T)
```

####文字解釋的部分可以直接撰寫在RMarkDown中

####或在
```{r}
return("R裡面用return函示寫出來，會在")  #註記程式碼
```
```

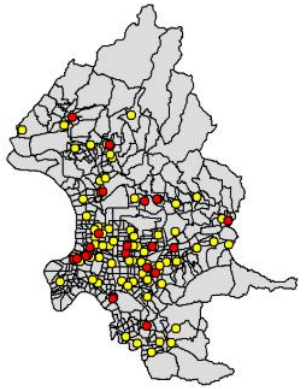
※不需要留不必要的資訊(i.e.列出data的內容)

> Q1.以下有些RMarkDown的教學題號請用">"當前綴來標示

一題結束可以用分隔線隔開

> Q2.題目通常都是畫圖和解釋

```
#回答的時候請保留程式碼以利批改
plot(tpe, col="#DDDDDD")
pointmap(MIC.pt, pch=21, bg="yellow", add=T)
pointmap(KFC.pt, pch=21, bg="red", add=T)
```



文字解釋的部分可以直接撰寫在RMarkDown中

或在

```
return("R裡面用return函示寫出來，會在") #井號是在註記程式碼的
```

```
## [1] "R裡面用return函示寫出來，會在"
```

隨堂小考

14:25→14:55　上傳繳交 R script (*.R)

建立執行平方和的**函數**：ssum(arg1, arg2, arg3)
注意：函數名稱請使用ssum

引數 arg1 & arg2，表示起始值與最終值。(arg1, arg2須為整數)
引數 arg3，表示數字間隔。(arg3須為正整數)

若輸入錯誤格式的arguments，則中止程式，並顯示 "wrong inputs" 的訊息。

引數→	arg1	arg2	arg3	
格式→	整數	整數	正整數	→錯誤格式：顯示"wrong inputs"後return
	2	5	1	$2^2+3^2+4^2+5^2=54$
	2	5	2	$2^2+4^2=20$
	6	3	1	$6^2+5^2+4^2+3^2=86$
	6	3	2	$6^2+4^2=52$

分數	人數
20	2
18	1
15	4
10	10
5	11
3	9
0	5

隨堂小考檢討

```
ssum=function(x,y,z){

  判斷是否為整數  x%%1==0

  if(x%%1!=0||y%%1!=0||z%%1!=0||z<0) {
    print("wrong inputs");
    return();
  }

  累進

  seq

  if(x>y) z=-z
  sum=sum(seq(x,y,z)^2)
  return(sum)


}
```

P.S. is.integer()是檢查型別（格式）的

any(c(x,y,z)%%1!=0)

※ 不能交換from-to的順序

while

```
sum=0; t=x;
if(x<=y) {
  while(t<=y){
      sum=sum+t^2
      t=t+z }
}else{
  while(t>=y){
      sum=sum+t^2
      t=t-z }
}
return(sum)
```

R資料處理

- **factor → number**

  問題：將數字辨識成factor格式

  (x)直接轉成factor

  (o)先轉成character再轉成factor

```
> fee
[1] 1245 2567 432  135  742
Levels: 135 432 742 1245 2567
> as.numeric(fee)
[1] 4 5 2 1 3
> as.numeric(as.character(fee))
[1] 1245 2567  432  135  742
```

- 選取：
  1. 邏輯判斷
  2. which
  3. 直接使用欄位對位選取
  4. subset

- **subset(data,判斷式)**

常用表格或 shapefile資料

逐列偵測

%in%：判斷是否在其中

```
> fast
  mic kfc
1   1   4
2   3   7
3   5   3
4   8   7
5   9   2
6   4   1
7   6   5
```

```
mic                  1 3 5 8 9 4 6
kfc                  4 7 3 7 2 1 5
mic>5                F F F T T F T
x=which(mic>5)           4 5   7
kfc[x]                   7 2   5
kfc[mic>5]           →   7 2   5
subset(kfc,mic>5)    →   7 2   5
```

- **xtabs()：樞紐分析表**

  xtabs(sum~group)

  i.e. xtabs(~TOWN+STORE)

  沒有參數→代表count

```
fast[mic>5,]          →
subset(fast,mic>5)
```

```
  mic kfc
4   8   7
5   9   2
7   6   5
```

- **aggregate(欄位,by=group, FUN=mean)**

R 處理 GIS資料

```
setwd("D:/1072SA/Data") #設定路徑
TPE=readOGR(dsn = ".", layer = "Vill", encoding="utf8",verbose=F)
```

"./" 當前資料夾
"../" 當前資料夾的上層

```
setwd("D:/1072SA")
TPE=readOGR(dsn = "Data", layer = "Vill)
```

- TPE@data
  屬性工作表（格式data.frame）
  可用$呼叫欄位：**TPE@data$ID**（直接 **TPE$ID** 也可以）

- TPE@proj4string or proj4string(TPE)

```
CRS arguments:
 +proj=tmerc +lat_0=0 +lon_0=121 +k=0.9999 +x_0=250000 +y_0=0
 +ellps=GRS80 +units=m +no_defs
```

- FastFood@cords
  點資料的x,y座標

- poly.areas(TPE)
  面資料的面積

ggplot

請搭配ggplot2-cheatsheet.zip

data

```
ggplot(mpg, aes(hwy, cty)) +
 geom_point(aes(color = cyl)) +
 geom_smooth(method ="lm") +
 coord_cartesian() +
 scale_color_gradient() +
 theme_bw()
```

add layers, elements with +

layer = geom + default stat + layer specific mappings

additional elements

```
ggplot(data = <DATA>) +
  <GEOM_FUNCTION>(
    mapping = aes(<MAPPINGS>),
    stat = <STAT>,
    position = <POSITION>
  ) +
  <COORDINATE_FUNCTION> +
  <FACET_FUNCTION> +
  <SCALE_FUNCTION> +
  <THEME_FUNCTION>
```

Required

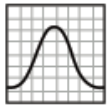
Not required, sensible defaults supplied

One Variable

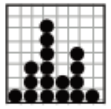
Continuous

a <- ggplot(mpg, aes(hwy))

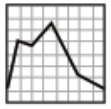
a + **geom_area(stat = "bin")**
x, y, alpha, color, fill, linetype, size
b + geom_area(aes(y = ..density..), stat = "bin")



a + **geom_density(**kernel = "gaussian"**)**
x, y, alpha, color, fill, linetype, size, weight
b + geom_density(aes(y = ..county..))



a + **geom_dotplot()**
x, y, alpha, color, fill

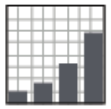


a + **geom_freqpoly()**
x, y, alpha, color, linetype, size
b + geom_freqpoly(aes(y = ..density..))

a + **geom_histogram(**binwidth = 5**)**
x, y, alpha, color, fill, linetype, size, weight
b + geom_histogram(aes(y = ..density..))

Discrete

b <- ggplot(mpg, aes(fl))

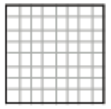


b + **geom_bar()**
x, alpha, color, fill, linetype, size, weight

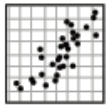
Two

Continuous X, Continuous Y

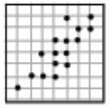
f <- ggplot(mpg, aes(cty, hwy))



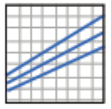
f + **geom_blank()**



f + **geom_jitter()**
x, y, alpha, color, fill, shape, size



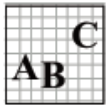
f + **geom_point()**
x, y, alpha, color, fill, shape, size



f + **geom_quantile()**
x, y, alpha, color, linetype, size, weight

f + **geom_rug(**sides = "bl"**)**
alpha, color, linetype, size

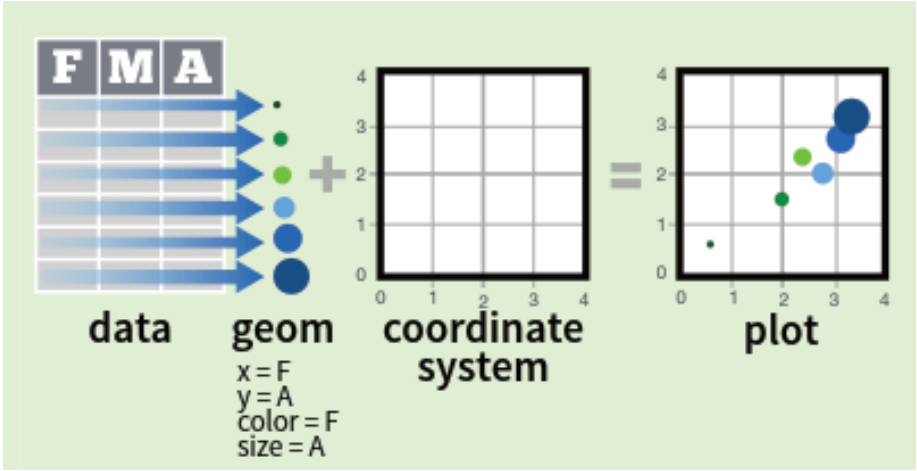
f + **geom_smooth(**model = lm**)**
x, y, alpha, color, fill, linetype, size, weight



f + **geom_text(**aes(label = cty)**)**
x, y, label, alpha, angle, color, family, fontfa
hjust, lineheight, size, vjust

ggplot

```
ggplot(data,aes(x=…,y=…))+ #放在也geom中可以
  geom_xxx(  ) +
  scale_xxx_xxx(  ) +
  labs()+ #座標軸
  theme() #主題
```



用ggplot畫GIS多邊形

- 以TOWN這欄來合併畫圖

```
TOWN.f = fortify(Vill, region="TOWN")
TOWN.f = merge(TOWN.f, Vill@data,
                    by.x = "id", by.y = "TOWN") #只是把資料對回去

ggplot()+
  geom_polygon(data = TOWN.f,
       aes(x=long, y = lat, group = group),
       fill="khaki1", color="black")+
  coord_fixed(1.0)
```

ggplot:
scale

Scales

Scales map data values to the visual values of an aesthetic. To change a mapping, add a new scale.

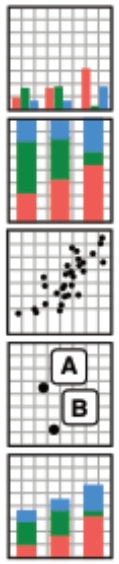
(n <- d + geom_bar(aes(fill = fl)))

scale_ | aesthetic to adjust | prepackaged scale to use | scale specific arguments

n + **scale_fill_manual(**
values = c("skyblue", "royalblue", "blue", "navy"),
limits = c("d", "e", "p", "r"), **breaks** =c("d", "e", "p", "r"),
name = "fuel", **labels** = c("D", "E", "P", "R"))

range of values to include in mapping | title to use in legend/axis | labels to use in legend/axis | breaks to use in legend/axis

```
scale_color_manual("速食店",
  values=c("red","blue"),
  labels = c("肯德基","麥當勞"))
```



s <- ggplot(mpg, aes(fl, fill = drv))

s + geom_bar(**position = "dodge"**)
Arrange elements side by side

s + geom_bar(**position = "fill"**)
Stack elements on top of one another, normalize height

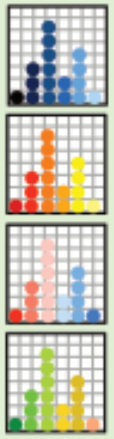
e + geom_point(**position = "jitter"**)
Add random noise to X and Y position of each element to avoid overplotting

e + geom_label(**position = "nudge"**)
Nudge labels away from points

s + geom_bar(**position = "stack"**)
Stack elements on top of one another

Color and fill scales (Continuous)

o <- c + geom_dotplot(aes(fill = ..x..))



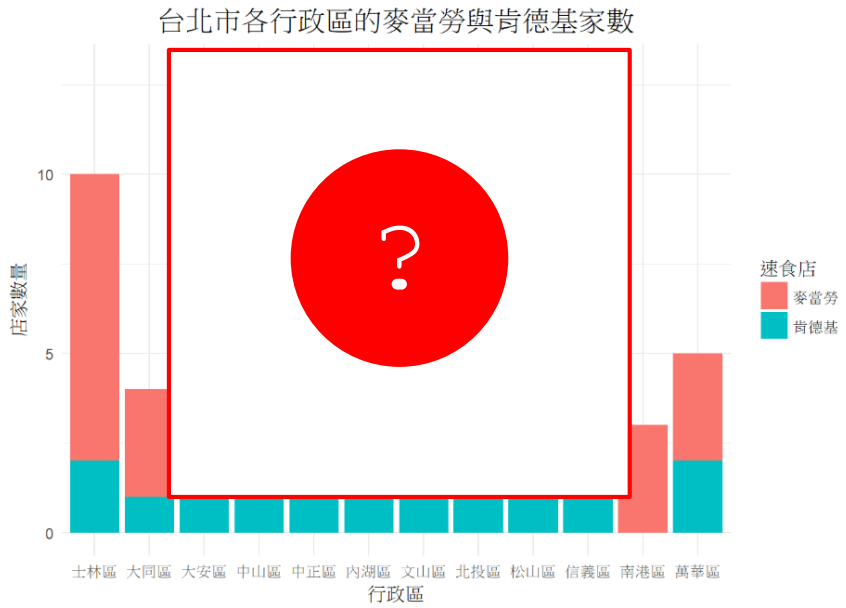
o + **scale_fill_distiller(**palette = "Blues"**)**

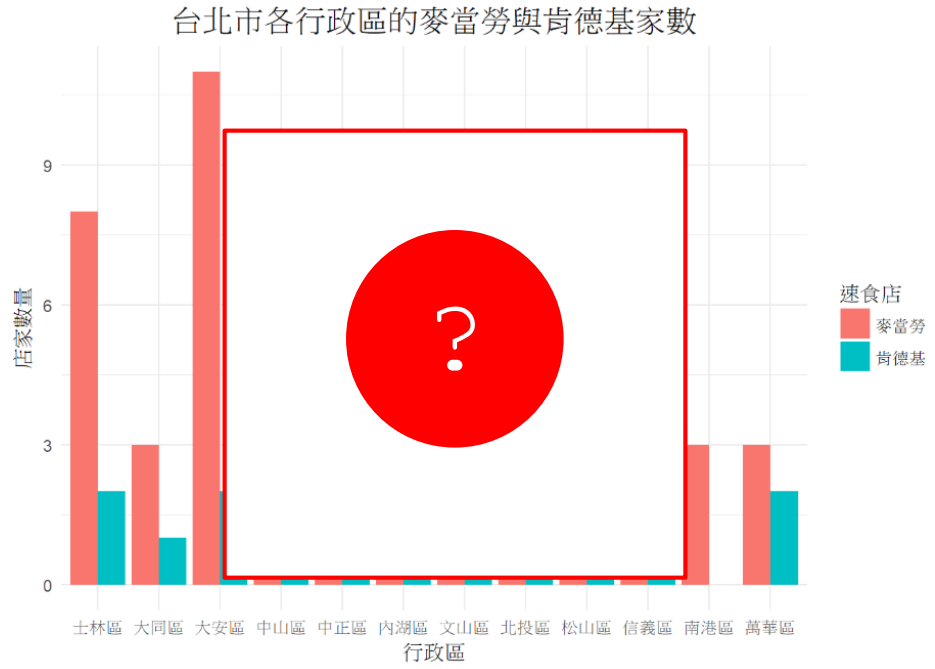
o + **scale_fill_gradient(**low="red", high="yellow"**)**

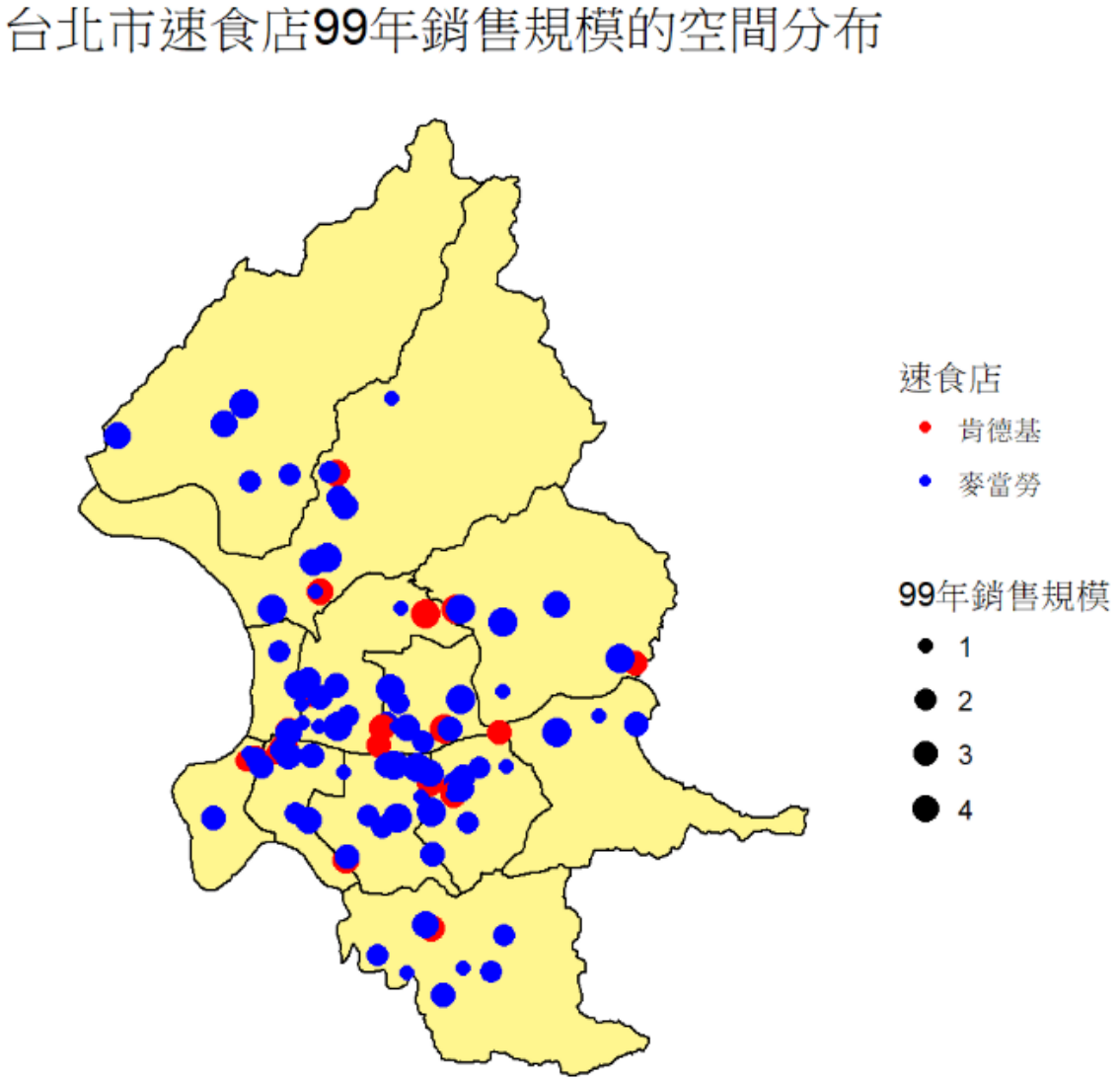
o + **scale_fill_gradient2(**low="red", high="blue", mid = "white", midpoint = 25**)**

o + **scale_fill_gradientn(**colours=topo.colors(6)**)**
Also: rainbow(), heat.colors(), terrain.colors(), cm.colors(), RColorBrewer::brewer.pal()

Lab 1 參考圖

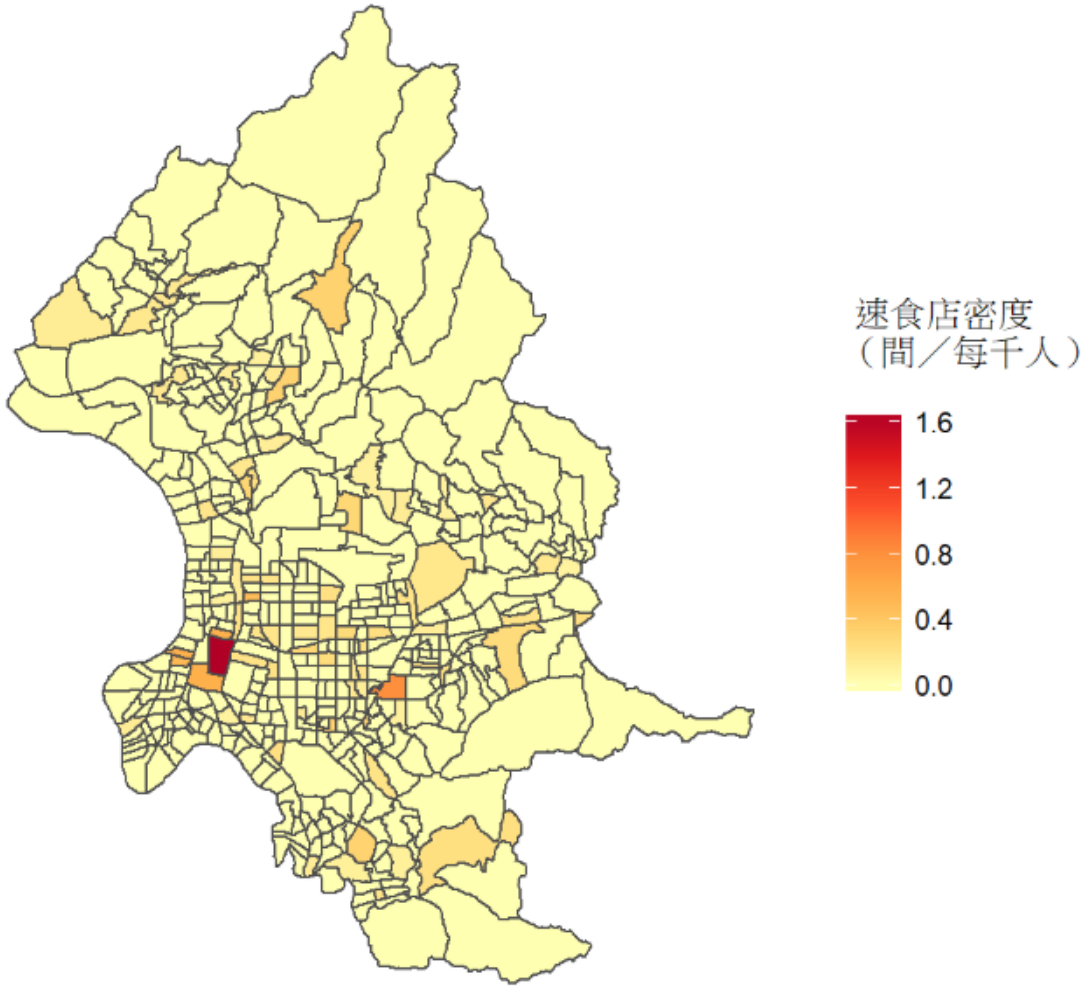

台北市各行政區的麥當勞與肯德基家數
?
店家數量
10
5
0
士林區 大同區 大安區 中山區 中正區 內湖區 文山區 北投區 松山區 信義區 南港區 萬華區
行政區
速食店
麥當勞
肯德基


台北市各行政區的麥當勞與肯德基家數
?
店家數量
9
6
3
0
士林區 大同區 大安區 中山區 中正區 內湖區 文山區 北投區 松山區 信義區 南港區 萬華區
行政區
速食店
麥當勞
肯德基

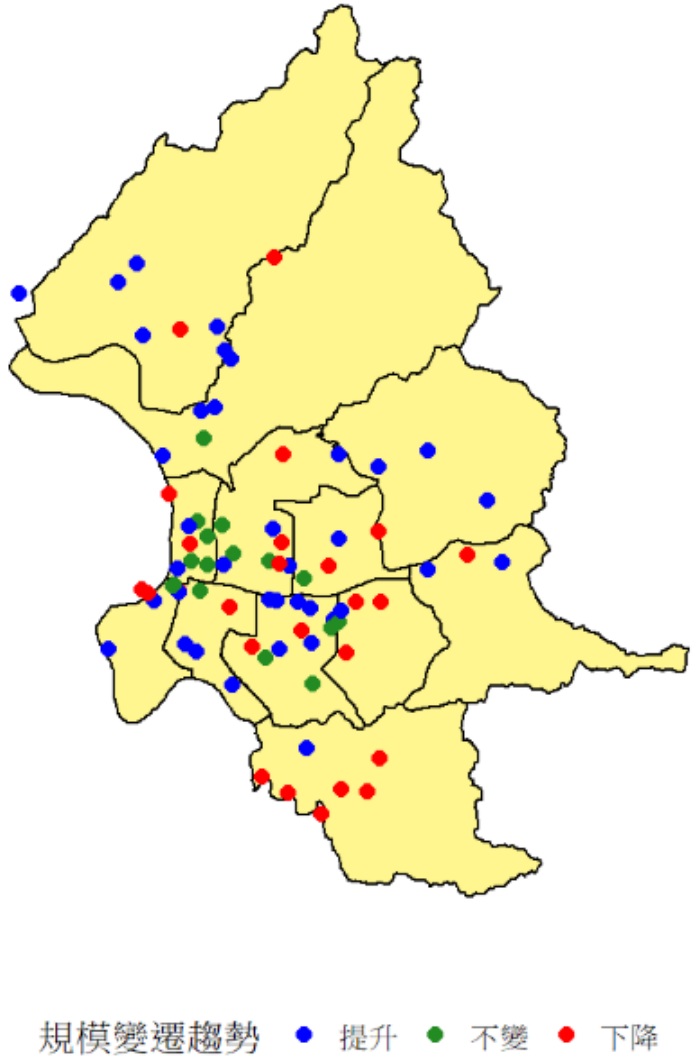

台北市速食店99年銷售規模的空間分布
速食店
肯德基
麥當勞
99年銷售規模
1
2
3
4

HW2 參考圖

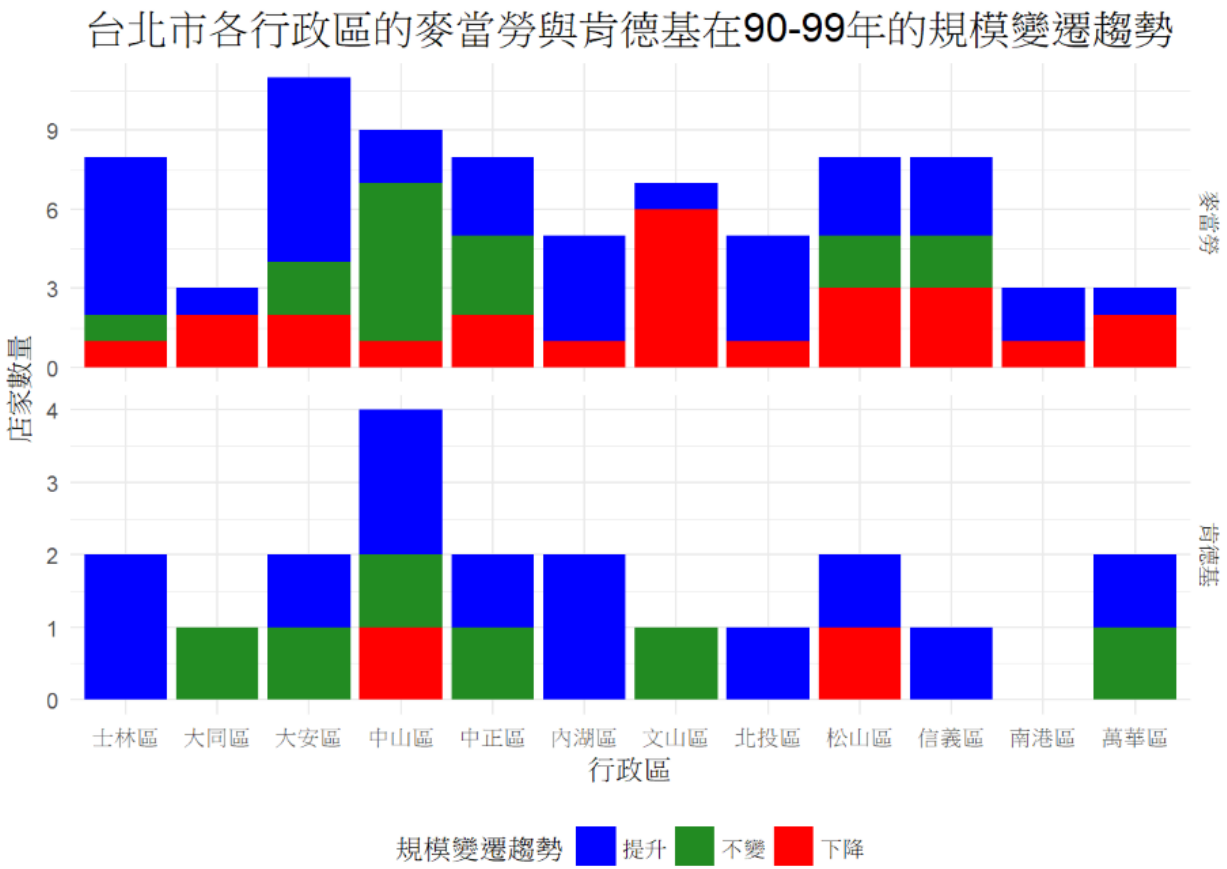
台北市各里速食店密度

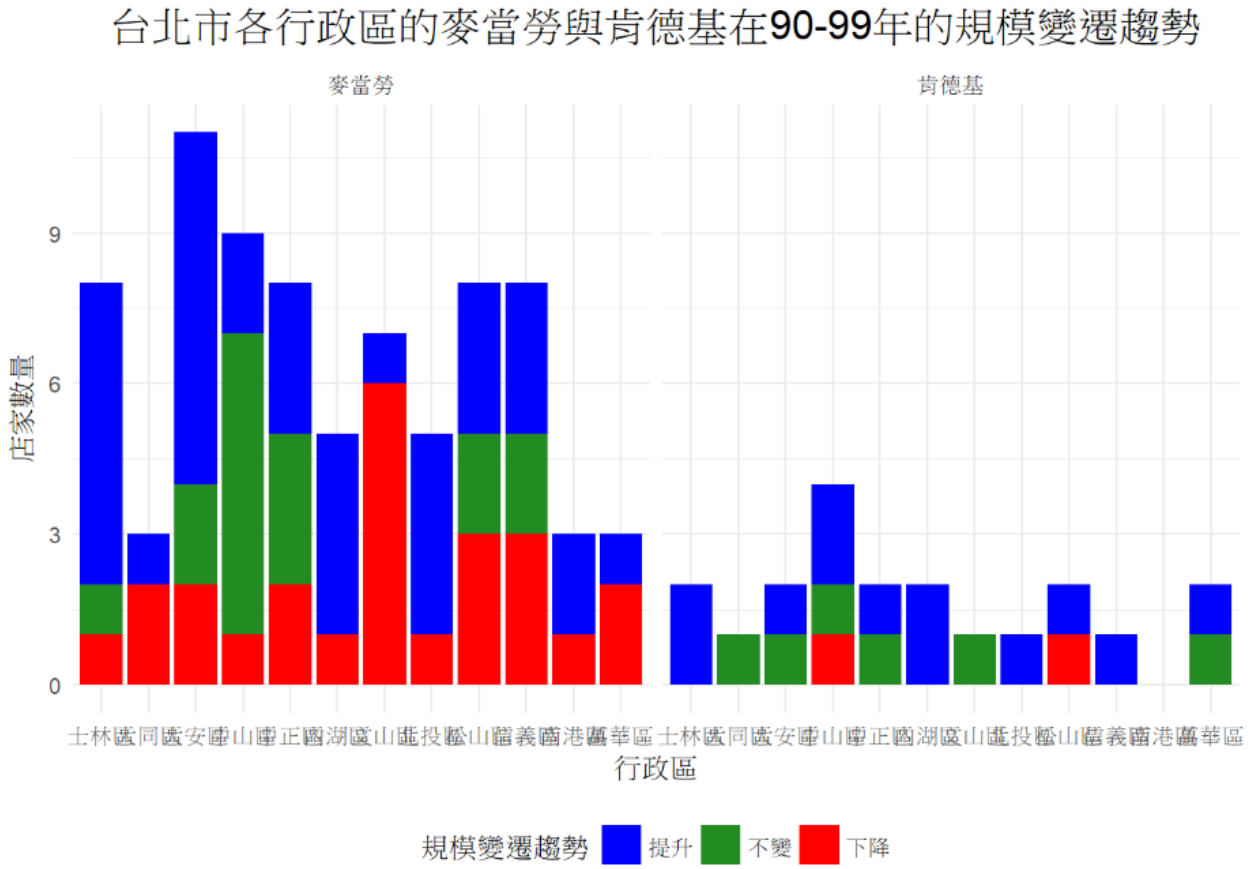


台北市麥當勞在90-99年的店家規模變遷趨勢



HW2 參考圖


台北市各行政區的麥當勞與肯德基在90-99年的規模變遷趨勢
店家數量
麥當勞
肯德基
士林區
大同區
大安區
中山區
中正區
內湖區
文山區
北投區
松山區
信義區
南港區
萬華區
行政區
規模變遷趨勢
提升
不變
下降


台北市各行政區的麥當勞與肯德基在90-99年的規模變遷趨勢
麥當勞
肯德基
店家數量
行政區
規模變遷趨勢
提升
不變
下降