



# Introduction to Spatial Data Analysis

Yuting Fan

Oct 7, 2020

# 下載資料

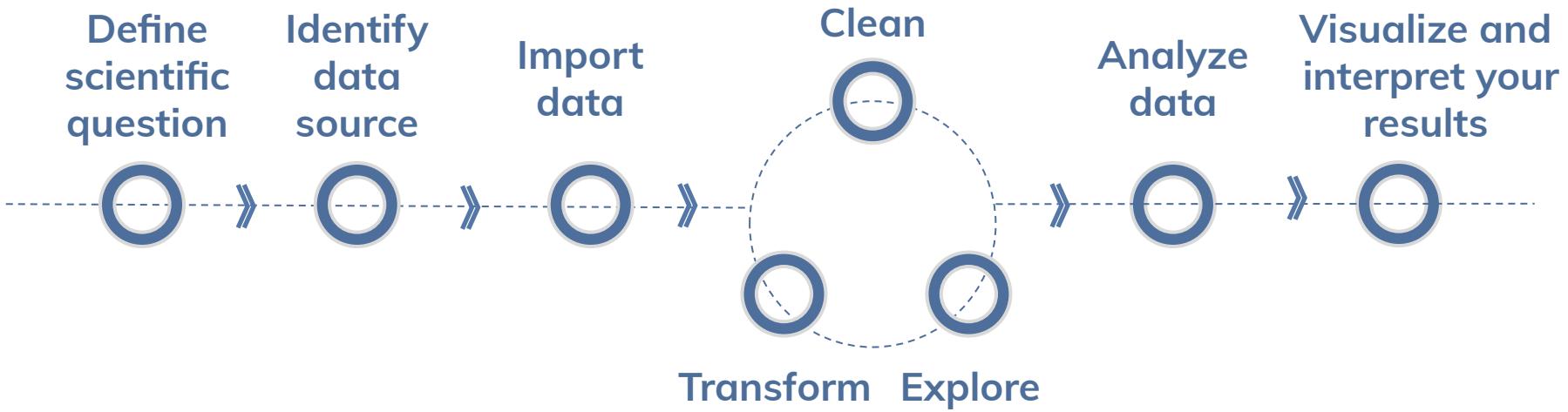
請從 <https://reurl.cc/avdgL9> 下載今天需要用到的資料

- Data
  - 台灣縣市 shapefile (folder)
  - 台灣鄉鎮 shapefile (folder)
  - 104 台南登革熱.csv
  - 108TaiwanPopulation.csv
  - pm.RData
- PDF
  - Spatial data introduction.pdf
- Script
  - #checkLibrary.R
  - 0\_spPackage
  - 1\_demo\_Taiwan.R
  - 2\_demo\_pm25.R
  - 3\_exercise\_denguefever\_1.R
  - 4\_demo\_population.R
  - 5\_exercise\_denguefever\_2.R

開啟檔案時請注意，本次提供文件均為 UTF-8 編碼，請以此編碼打開文件！

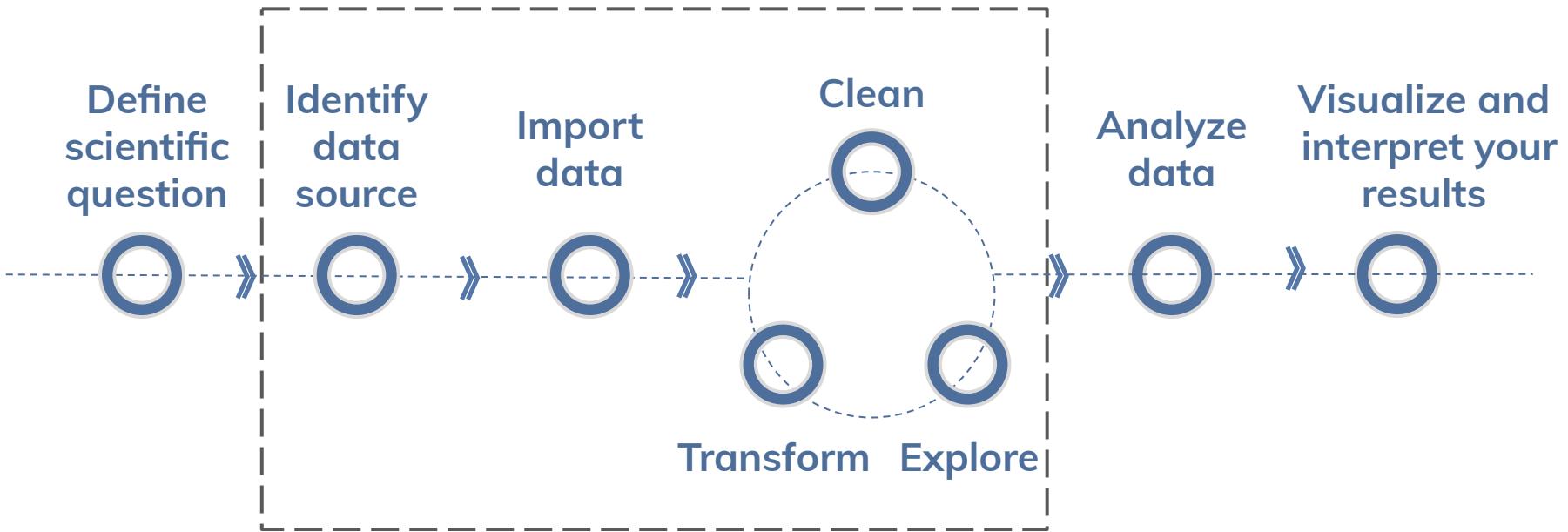
# 資料分析流程

根據不同的需求，資料分析流程會有所不同

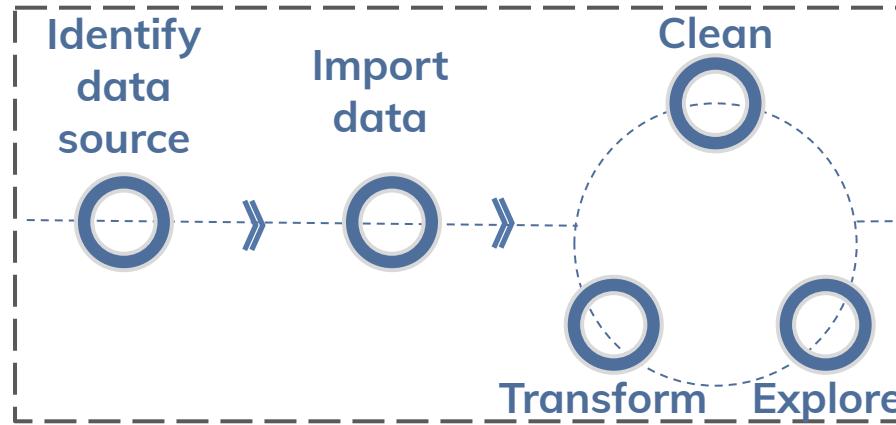


# 資料分析流程

本次分享主要會介紹空間資料分析部分流程中的知識



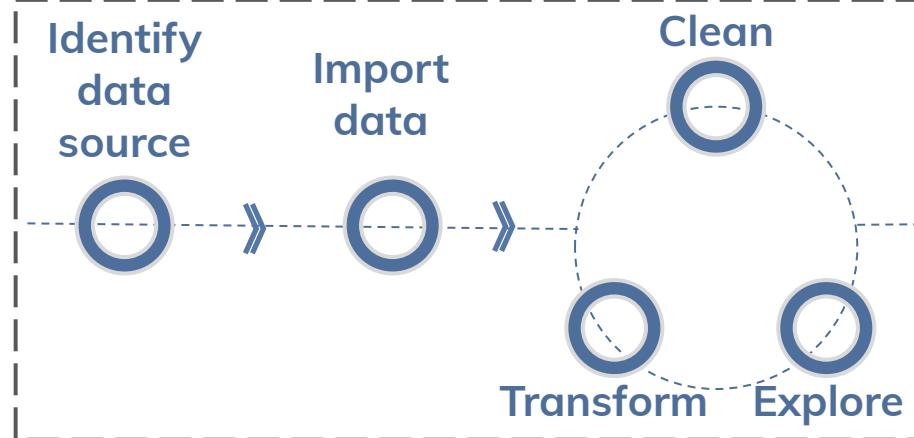
# 流程



## 第一部分：空間資料 第二部分：視覺化探索

- 空間資料
- 資料格式
- 資料來源
- 簡介工具
- R中空間資料物件類型
- 如何利用R探索空間資料
  - 台灣邊界
  - 空污 Pm 2.5 分佈
  - 登革熱分佈
  - 台灣人口數分佈
  - 登革熱和臺南人口密度分佈

# 第一部分： 空間資料



## 第一部分：空間資料

- 空間資料
- 資料格式
- 資料來源

## 第二部分：視覺化探索

- 簡介工具
- R中空間資料物件類型
- 如何利用R探索空間資料
  - 台灣邊界
  - 空污 Pm 2.5 分佈
  - 登革熱分佈
  - 台灣人口數分佈
  - 登革熱和臺南人口密度分佈

# 空間資料記錄方式: 手繪

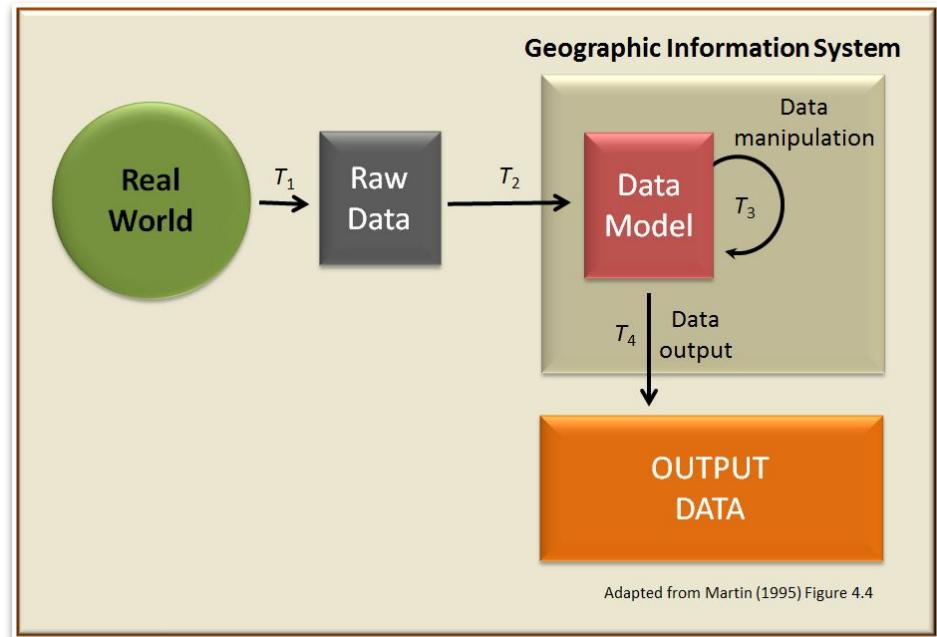
地圖是提供空間資訊的方式，沒有電腦的年代，利用手繪紀錄空間資訊



# 空間資料記錄方式: GIS

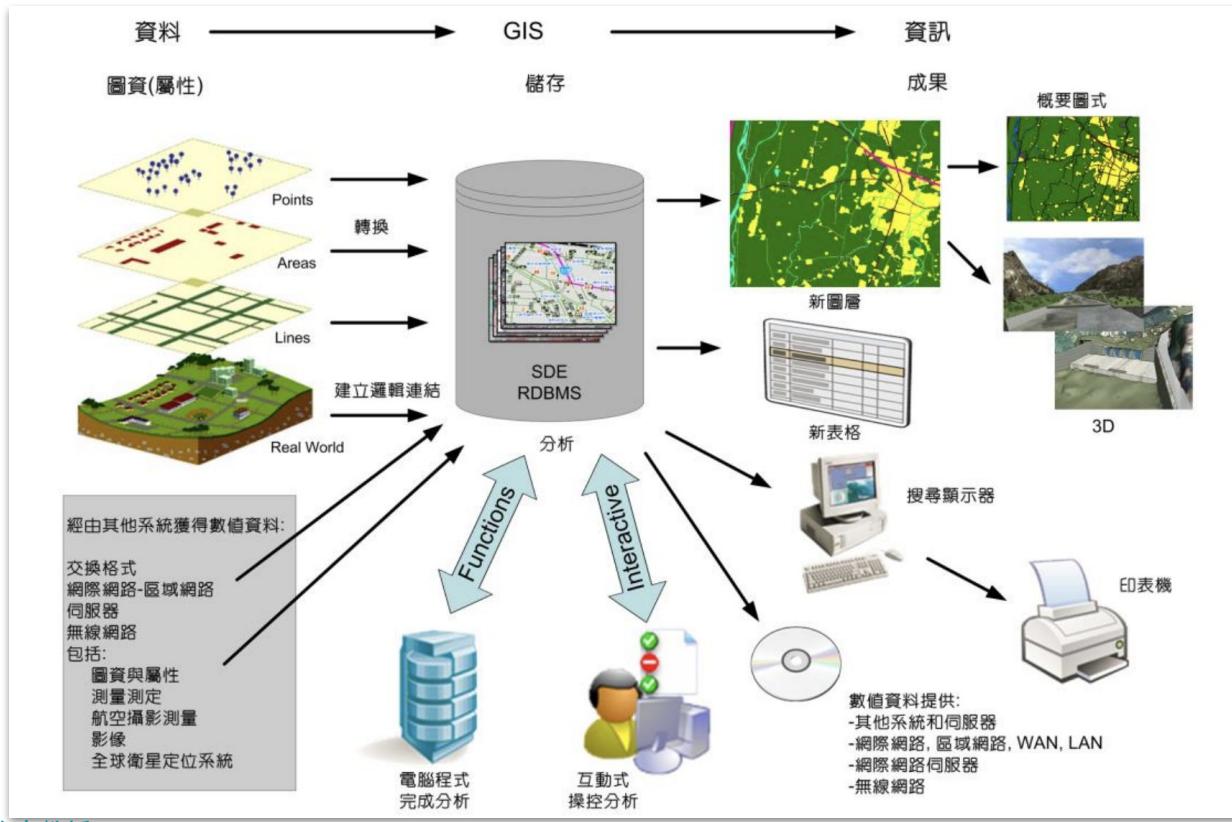
## 何謂GIS?

A geographic information system (GIS) is a conceptualized framework that provides the ability to capture and analyze spatial and geographic data.



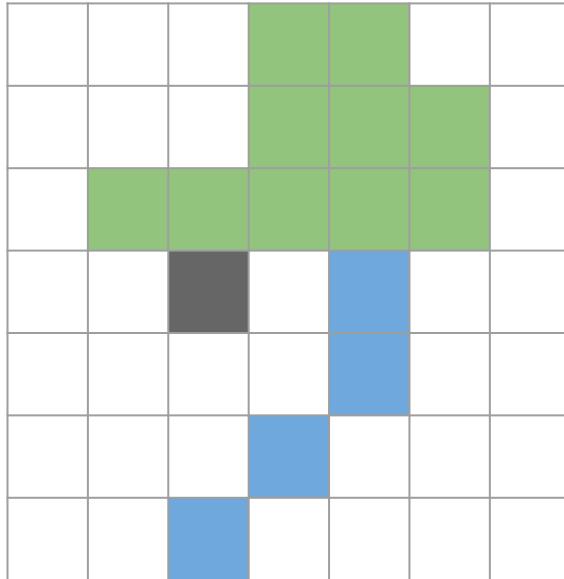
# 空間資料記錄方式: 有了電腦以後

何謂GIS?

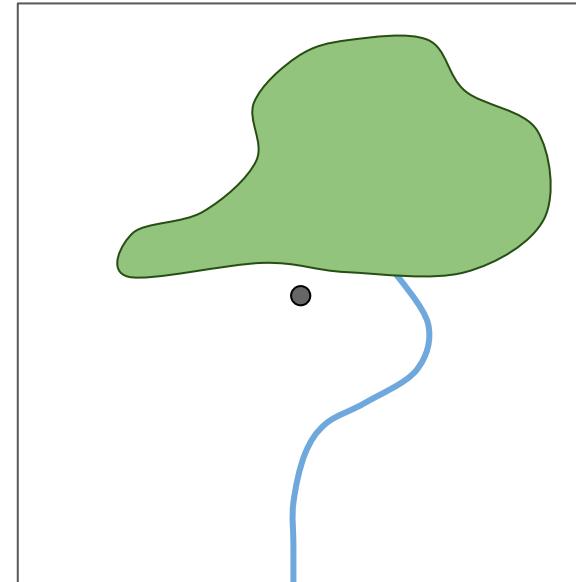
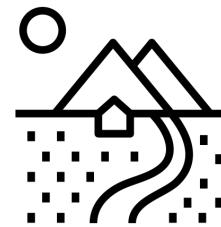


# 空間資訊記錄方式: Spatial data models in GIS

有哪些格式? → vector data or raster data



Raster data

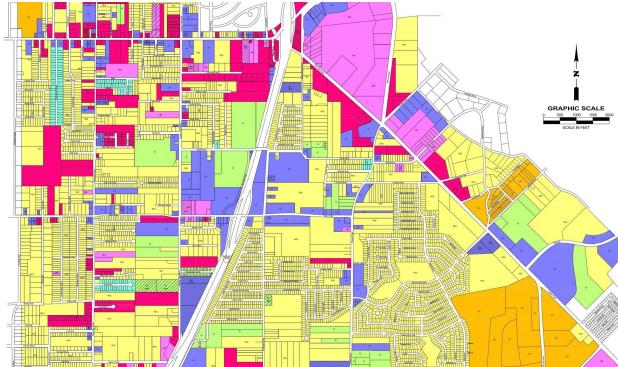


Vector data

# 空間資訊記錄方式: Spatial data models in GIS

有哪些格式? → vector data or raster data

衛星影像



都市土地利用分類圖

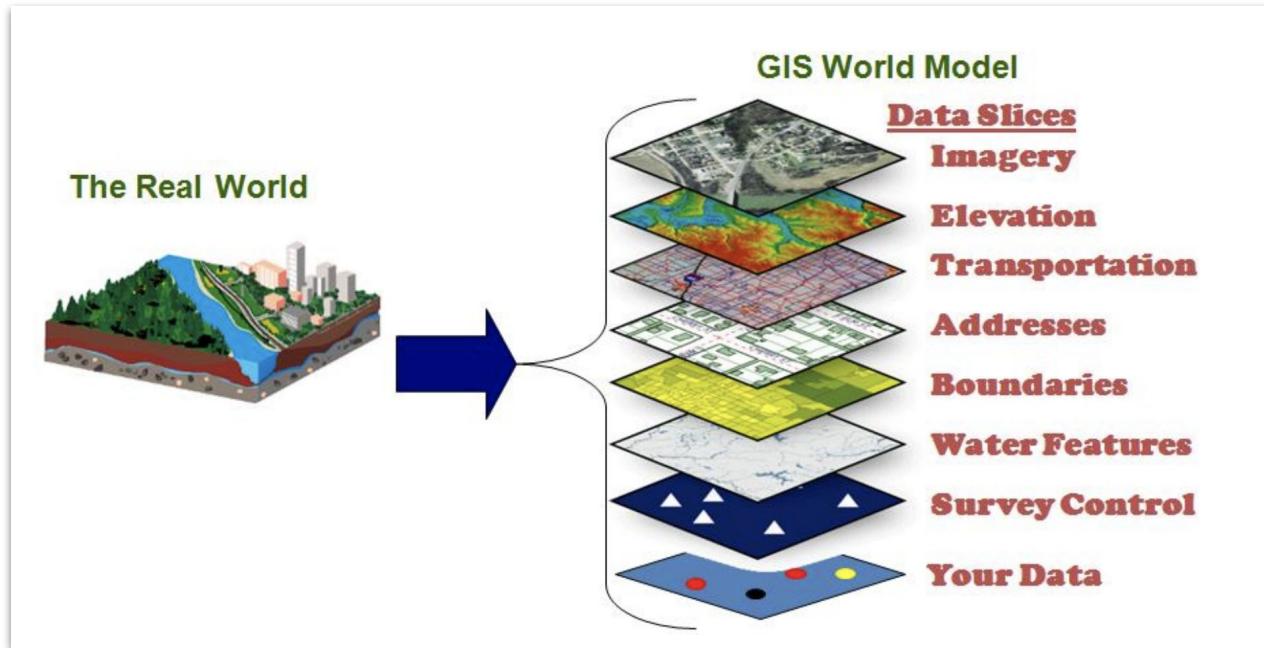
航空照片



Google map

# 空間資訊記錄方式

## GIS 模式的圖層



# 地理座標系統

地理坐標系統(Geographic coordinate system)又可稱大地坐標系統

地球



大地基準

橢球



投影與否

全球定位系統: WGS84

區域座標系: TWD67, TWD97

大地基準	WGS84	TWD67	TWD97
參考橢球體	WGS84 (全球性大地基準)	International 1967 (地區性大地基準)	GRS80 (地區性大地基準)
長半徑 (a)	6378137.000	6378160.000	6378137.000
短半徑 (b)	6356752.3142	6356774.7192	6356752.3141
扁率 (f)	1/298.257223563	1/298.25	1/298.257222101
大地基準點	地球質量中心	南投埔里虎子山	內政部 8 個 GPS 衛星追蹤站坐標值
地圖投影	無投影	TM 二度分帶	TM 二度分帶
單位	經緯度	公尺	公尺
備註	<ul style="list-style-type: none"><li>● GPS 使用</li><li>● TWD97 與 WGS84 近似</li></ul>		TWD97 與 WGS84 近似

# 空間資料: 空間 + 屬性

There are two components to GIS data: attribute data and spatial information (coordinate and projection information for spatial features).



# 資料檔案格式



PDF

1. 台北市
  - a. 氣溫: 23 度
  - b. 濕度: 76.6 %
2. 高雄市
  - a. 氣溫: 25.1度
  - b. 濕度: 75.9 %
3. 台中市
  - a. 氣溫: 23.3 度
  - b. 濕度: 75.6 %



CSV

縣市	氣溫	濕度
台北市	23	76.6
高雄市	25.1	75.9
台中市	23.3	75.6



JSON

```
{  
    "縣市": "台北市",  
    "氣溫": 23,  
    "濕度": 76.6  
}  
  
{  
    "縣市": "高雄市",  
    "氣溫": 25.1,  
    "濕度": 75.9  
}  
  
{  
    "縣市": "台中市",  
    "氣溫": 23.3,  
    "濕度": 75.6  
}
```

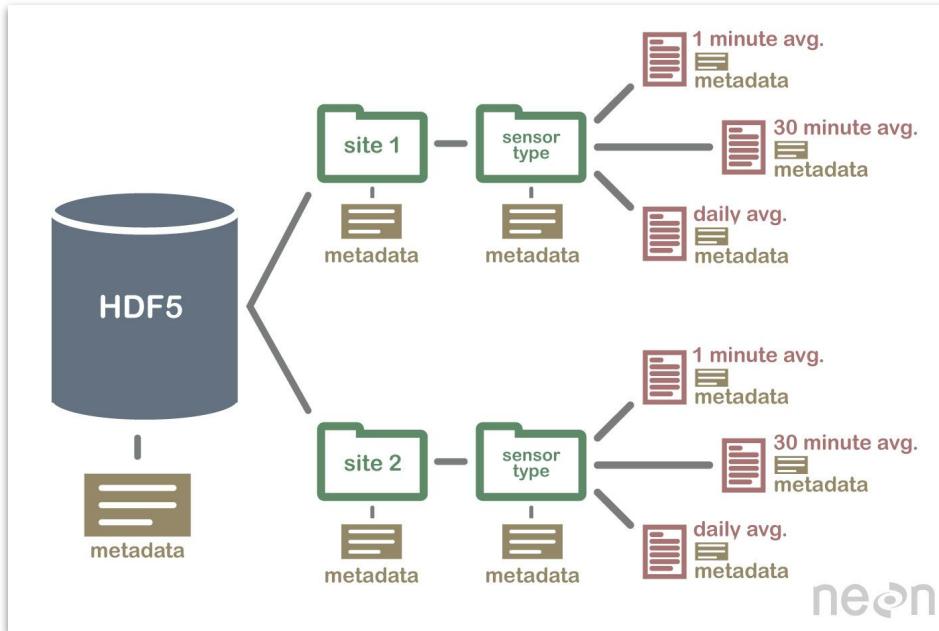


XML

```
<object>  
<縣市>台北市</縣市>  
<氣溫>23</氣溫>  
<濕度>76.6</濕度>  
</object>  
<object>  
<縣市>高雄市</縣市>  
<氣溫>25.1</氣溫>  
<濕度>75.9</濕度>  
</object>  
<object>  
<縣市>台中市</縣市>  
<氣溫>23.3</氣溫>  
<濕度>75.6</濕度>  
</object>
```

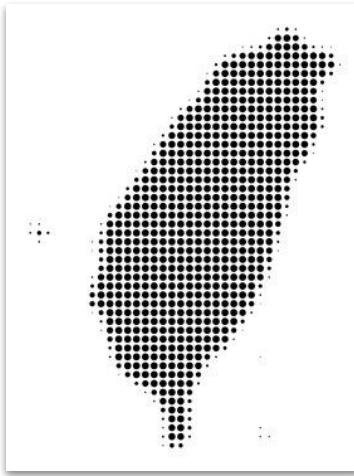
# 資料檔案格式:HDF5

存取巨大的檔案的格式之一



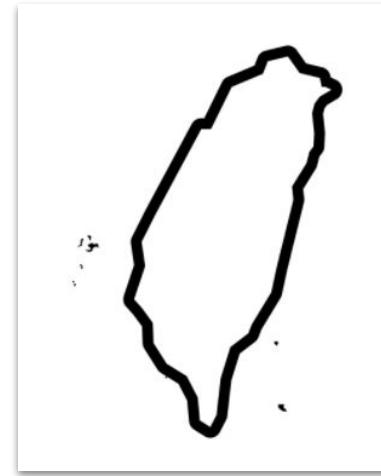
	group	name	otype	dclass	dim
0	/	site1	H5I_GROUP		
1	/site1	sensor a	H5I_GROUP		
2	/site1/sensor a	1-min	H5I_DATASET	INTEGER	5 x 2
3	/site1/sensor a	30-mins	H5I_DATASET	FLOAT	2
4	/site1/sensor a	daily avg.	H5I_DATASET	COMPOUND	5
5	/	site2	H5I_GROUP		
6	/site2	sensor a	H5I_GROUP		
7	/site2/sensor a	1-min	H5I_DATASET	INTEGER	5 x 2
8	/site2/sensor a	30-mins	H5I_DATASET	FLOAT	2
9	/site2/sensor a	daily avg.	H5I_DATASET	COMPOUND	5

# 空間資料檔案格式



Raster data

1. JPEG
2. PNG
3. TIFF (Tagged Image File Format)



Vector data

1. SHP (Shapefile)
2. GeoJson
3. KML (Keyhole Markup Language )

# 空間資料檔案格式: KML

KML 是 OGC 規範格式，廣泛用於 Google 平台上，且包含了其在 Google Earth 上的樣式資料，基本上都使用 ESPG4326 的經緯度坐標

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2">
<Document>
  <Placemark>
    <name>台北市</name>
    <ExtendedData>
      <Data name="氣溫">
        <value>23</value>
      </Data>
      <Data name="濕度">
        <value>76.6</value>
      </Data>
    </ExtendedData>
    <Point>
      <coordinates>121.5, 25.03</coordinates>
    </Point>
  </Placemark>
</Document>
</kml>
```

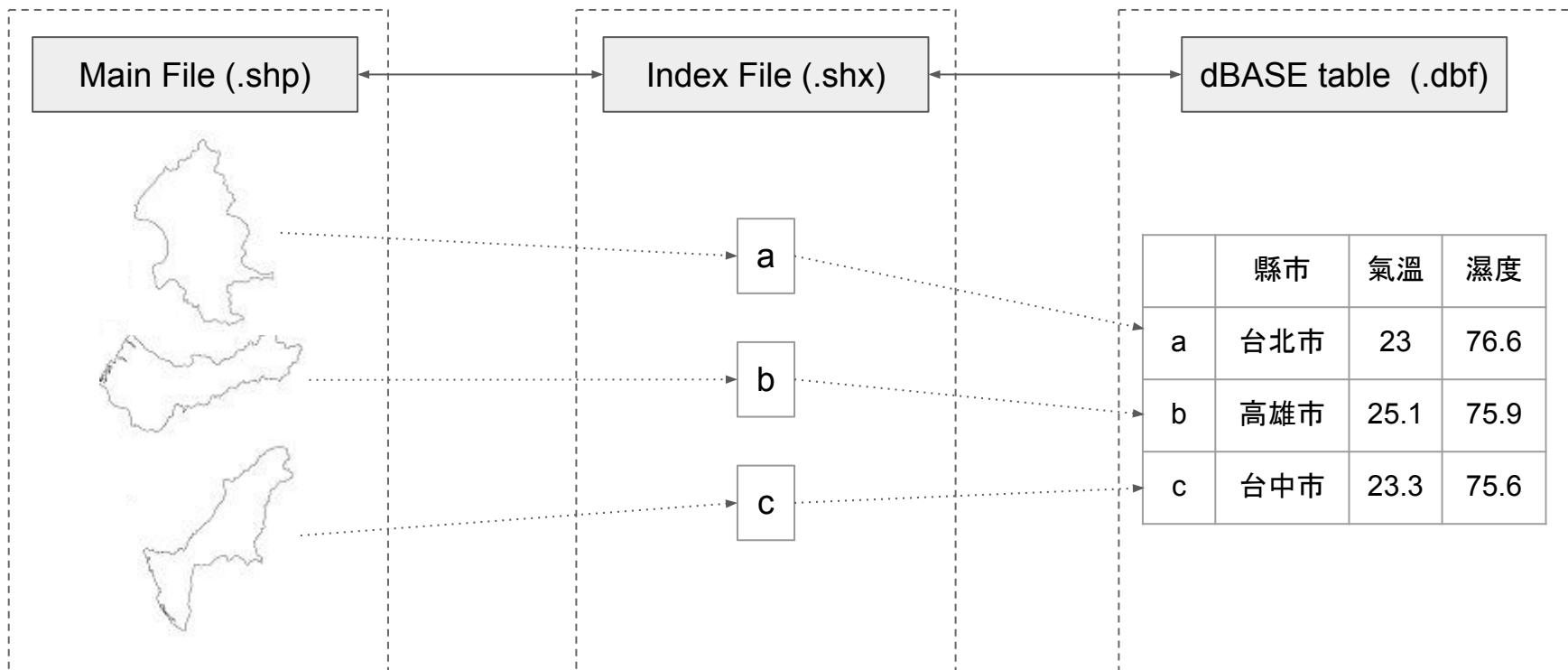
# 空間資料檔案格式: GEOJSON

Geojson 非 OGC (Open Geospatial Consortium) 標準，是由 IETF (Internet Engineering Task Force) 組織訂定的，因為是 json 物件的方式表現，使用起來簡單好用，於 D3.js 與 WebGIS 廣為流通

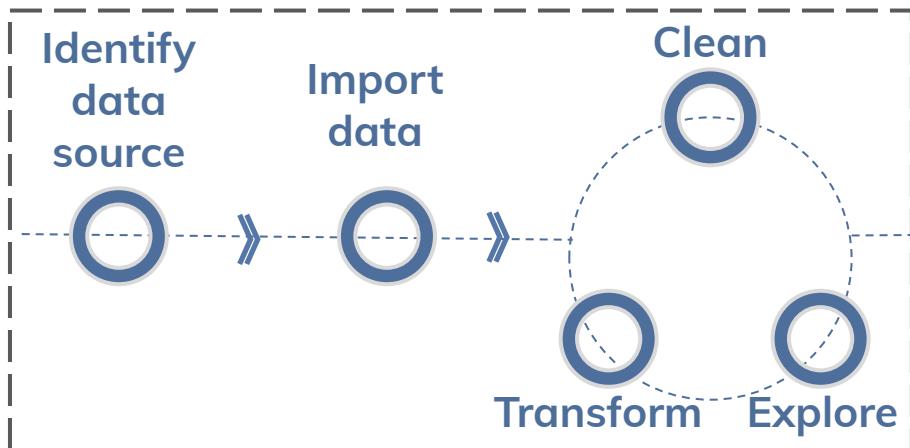
自動產生 geojson 的網站: [geojson.io](https://geojson.io)

```
{  
  "type": "Feature",  
  "geometry": {  
    "type": "Point",  
    "coordinates": [121.5, 25.03]  
  },  
  "properties": {  
    "縣市": "台北市",  
    "氣溫": 23,  
    "濕度": 76.6  
  }  
}
```

# 空間資料檔案格式: Shapefile



# 小結一下



## 第一部分：空間資料

- 空間資料
  - 空間資料：屬性資料 + 空間位置
  - GIS記錄的方式 (Spatial data models): raster data and vector data
- 資料格式
  - Raster data: JPEG, PNG, TIFF
  - Vector data: SHP, GeoJson, KML
- 資料來源

# 臺灣政府開放資料平台

<https://data.gov.tw/>

政府資料開放平臺  
DATA.GOV.TW

全部資料集 網站導覽 互動專區 資料故事館 NEW 最新消息 諮詢小組 關於平臺 EN 資料開放應用

輸入您想要搜尋的關鍵字 (資料集)

進階搜尋器 | 進階搜尋請參考這裡

資料集服務分類

- 生育保健(508)
- 出生及收養(89)
- 求學及進修(762)
- 服兵役(276)
- 求職及就業(575)
- 飲食及營養(111)
- 婚姻(35)
- 投資理財(1716)
- 休閒旅遊(947)
- 交通及通訊(1955)
- 就醫(1145)
- 選舉及投票(110)
- 生活安全及品質(2889)
- 退休(21)
- 老年安養(236)
- 生命禮儀(115)

智能客服 | 線上  
政府資料開放平臺

全部資料集 網站導覽 互動專區 資料故事館 NEW 最新消息

檔案格式

- CSV 31624
- XML 11630
- JSON 11409
- ZIP 3626
- PDF 1569
- ODS 931
- XLSX 924
- XLS 852
- TZ 716
- KML 534
- WEBSERVICES 522
- API 406
- 壓縮檔 372
- RAR 264
- SHP 238
- DOC 207
- 其他 126
- ODT 84
- DOCX 65
- TXT 53
- GEOJSON 31
- RSS 26
- RDF 20
- GML 15
- WMS 14
- KMZ 11
- IGC 10

108年度臺中市總決算附屬單位決算及綜計表-營業基金-現金流量綜計表(依主要欄位說明:項目(單位:新臺幣元)、合計、豐原農產品股份有限公司、臺中市政府主計處 / 註釋資料更新時間: 2020/09/26 16:41)  
47 0 0 CSV JSON XML

108年度臺中市總決算附屬單位決算及綜計表-營業流量項目分列(依主要欄位說明:項目(單位:新臺幣元)、預算數、決算核定數、比較增減金額、臺中市政府主計處 / 註釋資料更新時間: 2020/09/26 16:40)  
46 0 0 CSV JSON XML

108年度臺中市總決算附屬單位決算及綜計表-營業別分列(依主要欄位說明:項目(單位:新臺幣元)、合計金額、合計百分比、豐原農產品股份有限公司、捷運股份有限公司金額、臺中捷運股份有限公司百分比、臺中市政府主計處 / 註釋資料更新時間: 2020/09/26 16:38)  
50 0 0 CSV JSON XML

# US data.gov

<https://www.data.gov/>



DATA TOPICS ▾ RESOURCES STRATEGY DEVELOPERS CONTACT

## The home of the U.S. Government's open data

Here you will find data, tools, and resources to conduct research, develop web and mobile applications, design data visualizations, and [more](#).

For information regarding the Coronavirus/COVID-19, please visit [Coronavirus.gov](#).

### GET STARTED

SEARCH OVER 226,531 DATASETS



*Manufacturing & Trade Inventories & Sales*



# CSIRO 聯邦科學與工業研究組織

<https://data.csiro.au/collections/>

The screenshot shows the homepage of the CSIRO Data Access Portal. At the top, there is a dark header bar with the text "CSIRO Data Access Portal" on the left, and "Contact", "Help", and "API" links on the right. To the right of the header is the CSIRO logo. Below the header is a large banner with a dark background featuring a blurred image of a circuit board and glowing blue and green dots. The banner contains the text "Access research data, software and other digital assets across a range of disciplines." In the center of the page, below the banner, are three search options: "Search by keyword" (with a magnifying glass icon), "Search by location" (with a globe icon), and "Search by domain" (with a flask icon). A horizontal line separates these from a search bar at the bottom. The search bar has a placeholder text "Search CSIRO collections" and a magnifying glass icon on its right. Below the search bar, a note states: "Each word is searched with results returned for one of the words, any combination or all of the words." The footer of the page is a light gray color.

# Google dataset search

<https://datasetsearch.research.google.com/>

## Dataset Search

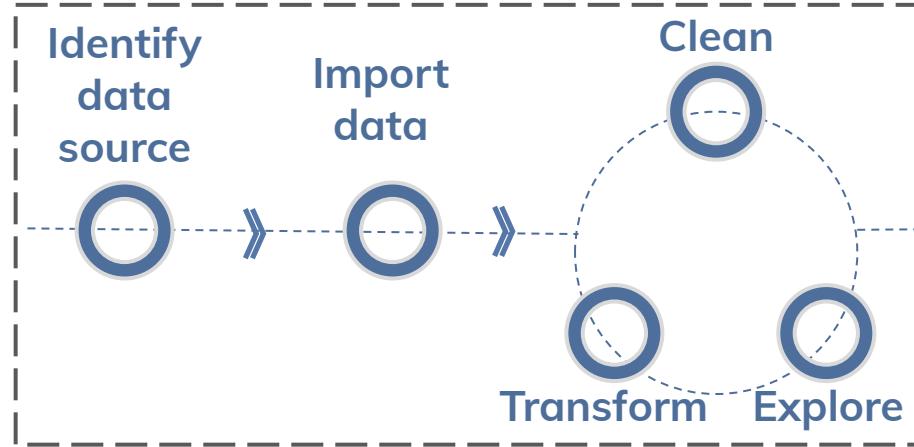
Search for Datasets



Try [coronavirus covid-19](#) or [global temperatures](#).

[Learn more](#) about including your datasets in Dataset Search.

# 第二部分： 視覺化探索



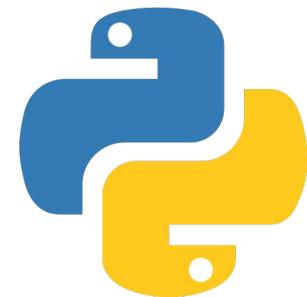
## 第一部分：空間資料

- 空間資料
- 資料格式
- 資料來源

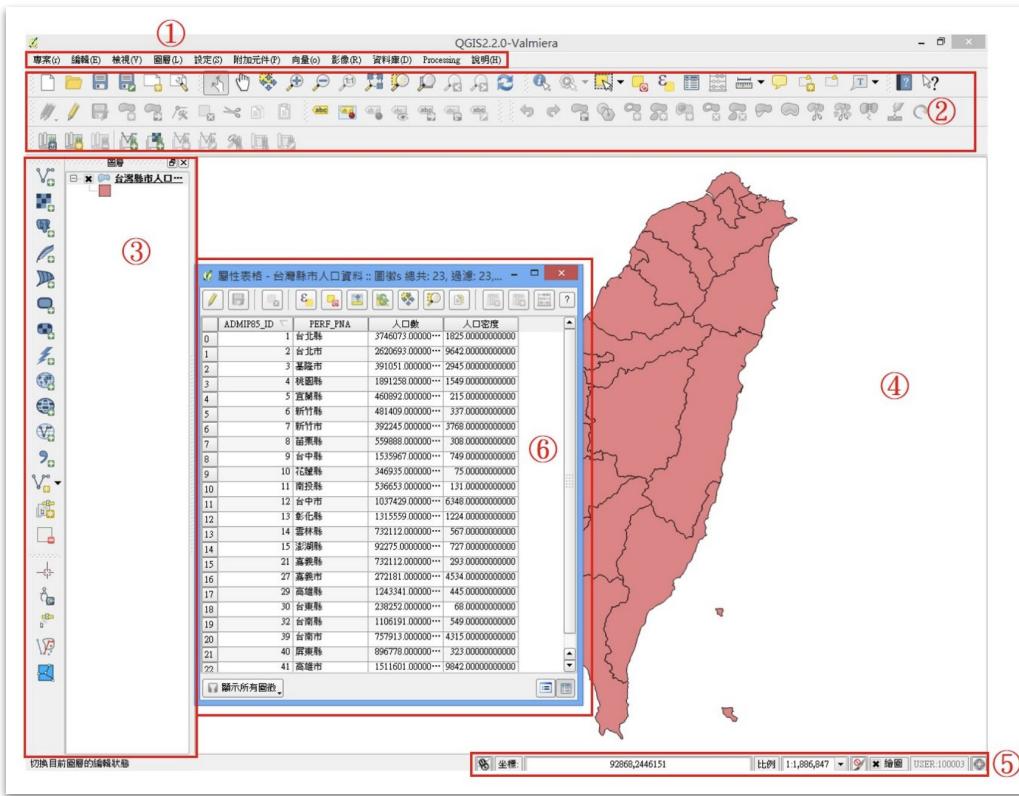
## 第二部分：視覺化探索

- 簡介工具
- R中空間資料物件類型
- 如何利用R探索空間資料
  - 台灣邊界
  - 空污 Pm 2.5 分佈
  - 登革熱分佈
  - 台灣人口數分佈
  - 登革熱和臺南人口密度分佈

# 空間資料分析工具



# 空間資料分析工具: QGIS



1. 選單列: 提供下拉式選單
2. 工具列: 選單功能以圖形化的介面表示
3. 圖層管理: 顯示加入 QGIS 的各類圖層, 包含向量與網格圖層
4. 圖層展示視窗: 透過平移、放大、縮小等功能檢視圖層內容
5. 狀態列: 顯示坐標值、比例尺、繪圖功能的標示、快速啟動坐標設定
6. 屬性表格: 檢視圖層屬性

# 空間資料分析工具: QGIS的學習資源

<http://gis.rchss.sinica.edu.tw/qgis/>

# 空間資料分析工具



ArcGIS



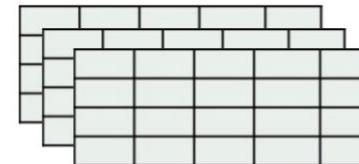
# Data types in R

Variables	Example
integer	100
numeric	0.05
character	“hello”
logical	TRUE
factor	“Green”

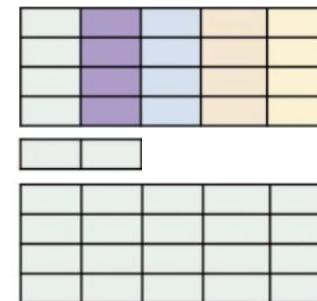
Vector



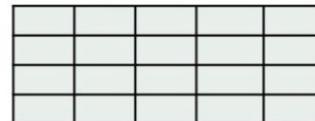
Array



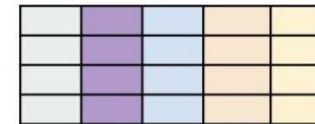
List



Matrix



Data frame



# 空間資料物件類型 types in sp package

The sp package provides classes and methods for dealing with spatial data in R.

[https://cran.r-project.org/web/packages/sp/vignettes/intro\\_sp.pdf](https://cran.r-project.org/web/packages/sp/vignettes/intro_sp.pdf)

## 1. Points

## 2. Lines

## 3. Polygons

- w/o attributes

- w/ attributes

data type	class	attributes	contains
points	SpatialPoints	No	Spatial
points	SpatialPointsDataFrame	data.frame	SpatialPoints
multipoints	SpatialMultiPoints	No	Spatial
multipoints	SpatialMultiPointsDataFrame	data.frame	SpatialMultiPoints
pixels	SpatialPixels	No	SpatialPoints
pixels	SpatialPixelsDataFrame	data.frame	SpatialPixels
			SpatialPointsDataFrame
full grid	SpatialGrid	No	SpatialPixels
full grid	SpatialGridDataFrame	data.frame	SpatialGrid
line	Line	No	
lines	Lines	No	Line list
lines	SpatialLines	No	Spatial, Lines list
lines	SpatialLinesDataFrame	data.frame	SpatialLines
polygons	Polygon	No	Line
polygons	Polygons	No	Polygon list
polygons	SpatialPolygons	No	Spatial, Polygons list
polygons	SpatialPolygonsDataFrame	data.frame	SpatialPolygons

# (0)Point

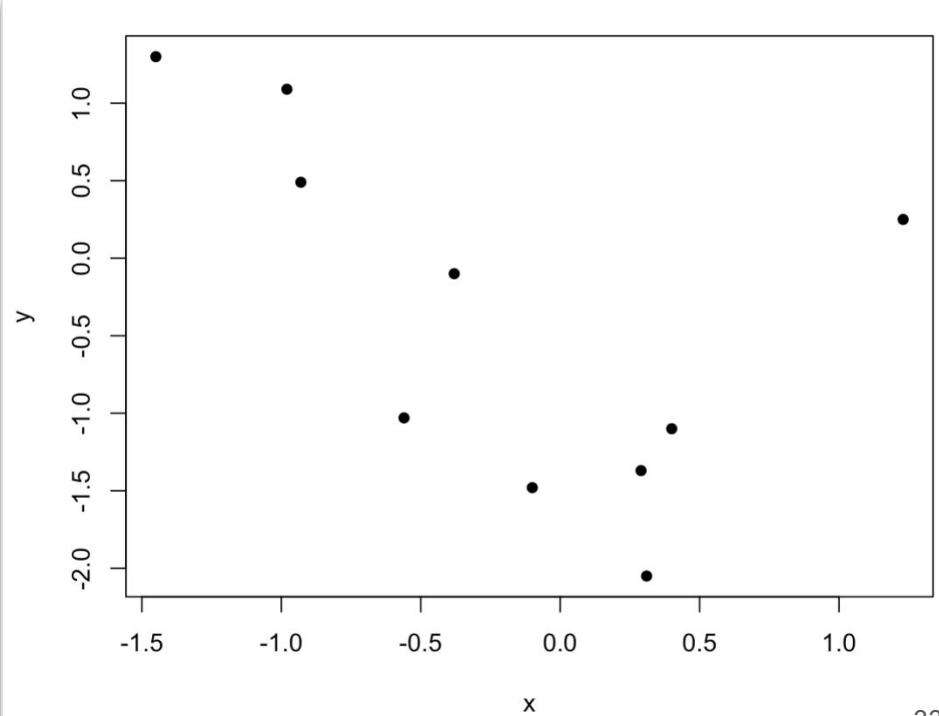
```
# coordinates
  x = round(rnorm(10), 2)
  y = round(rnorm(10), 2)
  xy = cbind(x, y)

# attributes
  df = data.frame(z1 = 1:10, z2 = 11:20)

# w/o attributes
  xy.sp = SpatialPoints(xy)
  summary(xy.sp)
  coordinates(xy.sp)
  plot(coordinates(xy.sp), pch = 16)

# w/ attributes - 1
  xy.df1 = SpatialPointsDataFrame(xy, df)

# w/ attributes - 2
  xy.df2 = data.frame(xy, df)
  coordinates(xy.df2) = c("x", "y")
```



# (0)Line

# coordinates

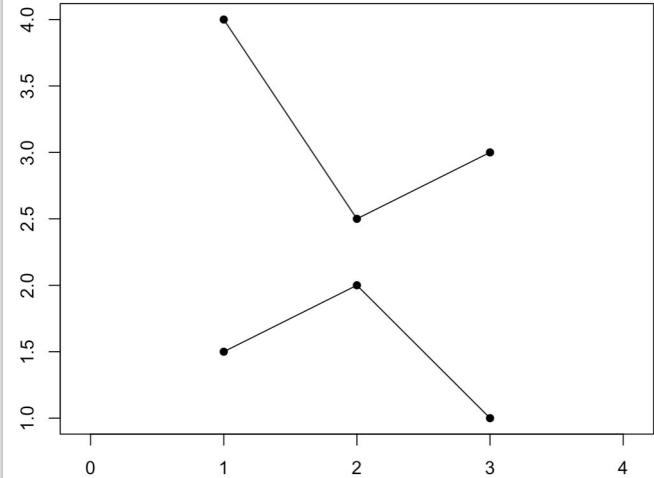
```
pt1 <- cbind(c(1,2,3),c(4,2.5,3))
pt2 <- cbind(c(1,2,3),c(1.5,2,1))
l1 = Line(pt1); l2 = Line(pt2)
Sl1 = Lines(list(l1), ID="a")
Sl2 = Lines(list(l2), ID="b")
```

# w/o attributes

```
Sline = SpatialLines(list(Sl1, Sl2))
plot(Sline, axes = T)
points(pt1, pch = 16); points(pt2, pch = 16)
```

# w/ attributes

```
df = data.frame(z = c(1,2),
                 row.names=apply(slot(Sline, "lines"), function(x) slot(x, "ID")))
Sl df = SpatialLinesDataFrame(Sline,
                             data = df)
Sl df@data
```



# (0) Polygon

# coordinates

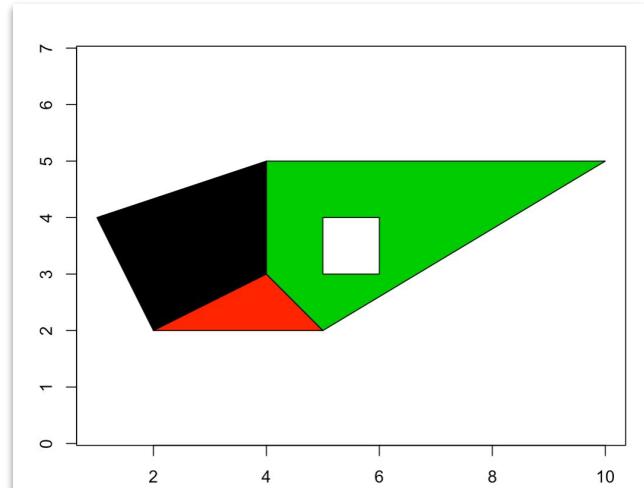
```
Sr1 = Polygon(cbind(c(2,4,4,1,2),c(2,3,5,4,2)))  
Sr2 = Polygon(cbind(c(5,4,2,5),c(2,3,2,2)))  
Sr3 = Polygon(cbind(c(4,4,5,10,4),c(5,3,2,5,5)))  
Sr4 = Polygon(cbind(c(5,6,6,5,5),c(4,4,3,3,4))), hole = TRUE)  
Srs1 = Polygons(list(Sr1), "s1")  
Srs2 = Polygons(list(Sr2), "s2")  
Srs3 = Polygons(list(Sr3, Sr4), "s3/4")
```

# w/o attributes

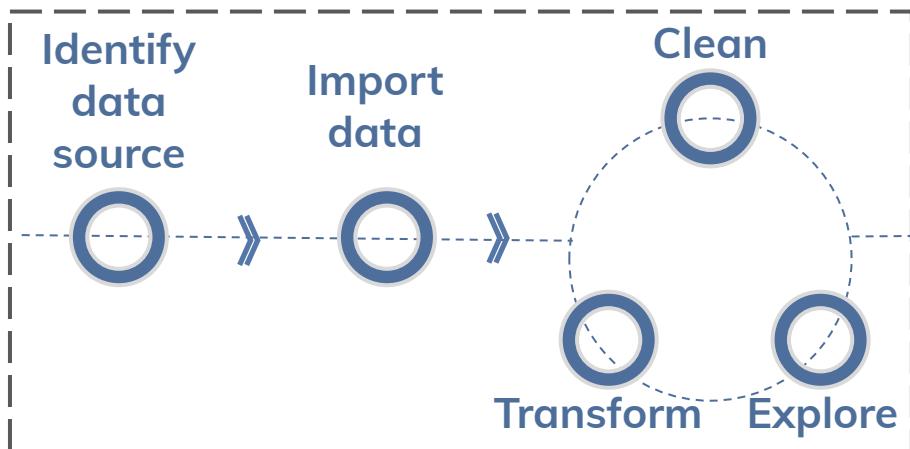
```
SpP = SpatialPolygons(list(Srs1,Srs2,Srs3), 1:3)  
plot(SpP, col = 1:3, axes = T)
```

# w/ attributes

```
attr = data.frame(a=1:3, b=3:1, row.names=c("s3/4", "s2", "s1"))  
SrDf = SpatialPolygonsDataFrame(SpP, attr)  
SrDf@data
```



# 小結一下



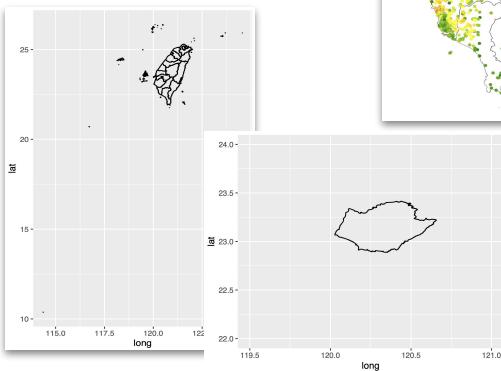
**第一部分：空間資料** → Focus on SHP format of vector data models

**第二部分：視覺化探索**

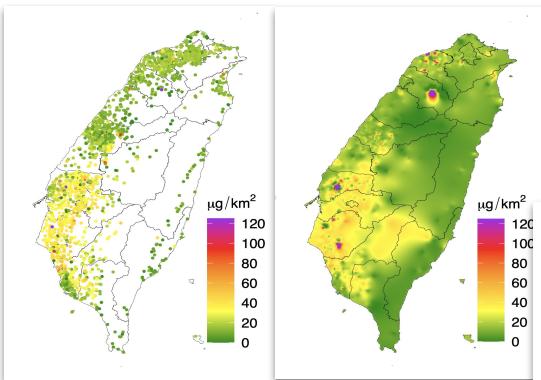
- 簡介工具 → Focus on R
- R中空間資料物件類型 → sp package
  - SpatialPointsDataFrame
  - SpatialLinesDataFrame
  - SpatialPolygonsDataFrame
- 如何利用R探索空間資料
  - 台灣邊界
  - 空污 Pm 2.5 分佈
  - 登革熱分佈
  - 台灣人口數分佈
  - 登革熱和臺南人口分佈

# 如何利用R探索空間資料: Overview (1/2)

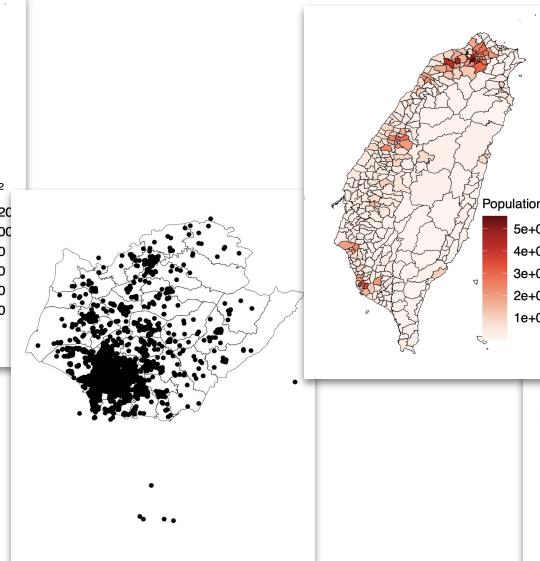
1 台灣邊界 demo



2 空污Pm2.5分佈 demo

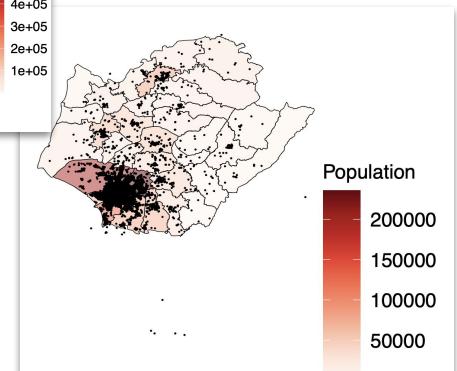


4 台灣人口數分佈 demo



3 登革熱分佈 exercise

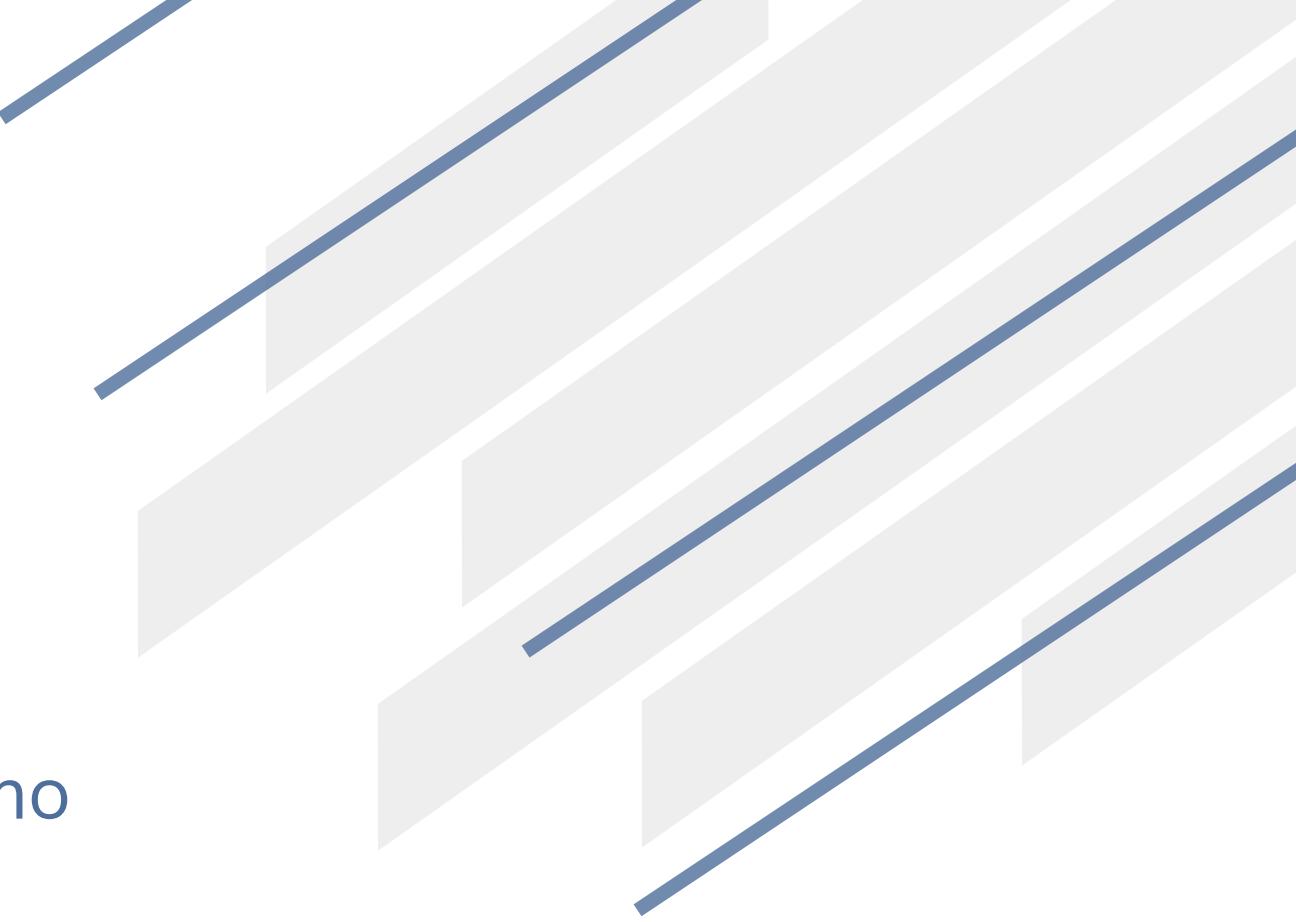
5 登革熱和臺南人口密度分佈 exercise



# 如何利用R探索空間資料: Overview (2/2)

Case name	R script	Required data files	Goal
台灣邊界	1_demo_Taiwan.R	<ul style="list-style-type: none"><li>COUNTY_MOI_1090820.shp</li></ul>	<ul style="list-style-type: none"><li>讀 .shp 檔案</li><li>利用geom_path畫邊界圖</li><li>例用subset選擇部分地區</li></ul>
空污 Pm 2.5 分佈	2_demo_pm25.R	<ul style="list-style-type: none"><li>COUNTY_MOI_1090820.shp</li><li>pm.RData</li></ul>	<ul style="list-style-type: none"><li>讀.RData檔案</li><li>利用geom_point畫散佈圖</li><li>利用geom_raster畫網格圖</li></ul>
登革熱分佈	3_exercise_denguefever_1.R	<ul style="list-style-type: none"><li>TOWN_MOI_1081121.shp</li><li>104台南登革熱.csv</li></ul>	
台灣人口數分佈	4_demo_population.R	<ul style="list-style-type: none"><li>TOWN_MOI_1081121.shp</li><li>108TaiwanPopulation.csv</li></ul>	<ul style="list-style-type: none"><li>利用 fortify 轉換 points object 成為 data frame</li><li>結合csv和shp資料</li><li>利用 geom_polygon 畫區域分佈圖</li></ul>
登革熱和臺南人口密度分佈	5_exercise_denguefever_2.R	<ul style="list-style-type: none"><li>TOWN_MOI_1081121.shp</li><li>108TaiwanPopulation.csv</li><li>104台南登革熱.csv</li></ul>	

## (1) 台灣邊界 demo



# 下載臺灣縣市邊界圖

資料下載來源：<https://data.gov.tw/dataset/7442>

政府資料開放平臺  
DATA.GOV.TW

全部資料集 網站導覽 互動專區 資料故事館 NEW 最新消息 諮詢小組 關於平臺

首頁 » 資料集 » 直轄市、縣市界線(TWD97經緯度)

## 直轄市、縣市界線(TWD97經緯度)

**相關資料集**

**資料集評分:** ★★★★☆  
平均 3.5 (53 人次投票)

**資料集描述:** 我國各直轄市以及縣(市)行政區域界線圖資

**主要欄位說明:** COUNTYID、COUNTYCODE、COUNTYNAME、COUNTYENG

**資料下載網址:** [其他](#) [檢視資料](#) 直轄市、縣市界線(TWD97經緯度)GM.....  
[SHP](#) [檢視資料](#) 直轄市、縣市界線(TWD97經緯度)SH.....

**提供機關:** 內政部國土測繪中心

**提供機關聯絡人姓名:** 吳嘉隆先生 (04-22522966#335)

**更新頻率:** 每週

# (1) 讀資料:台灣縣市邊界圖

```
> taiwan <- rgdal::readOGR("COUNTY_MOI_1090820.shp")
> class(taiwan)
[1] "SpatialPolygonsDataFrame"
attr(,"package")
[1] "sp"
```

`readOGR` is a function for reading vector-based spatial data from the `rgdal` package.

**SpatialPolygonsDataFrame** with default ID matching checks the data frame row names against the Polygons ID slots.

# (1) 資料結構

```
> slotNames(taiwan)
[1] "data" "polygons" "plotOrder" "bbox" "proj4string"
> taiwan@data
COUNTYID COUNTYCODE COUNTYNAME COUNTYENG
0      Z     09007    連江縣      Lienchiang County
1      G     10002    宜蘭縣      Yilan County
> names(taiwan)
[1] "COUNTYID" "COUNTYCODE" "COUNTYNAME" "COUNTYENG"
> taiwan@data$COUNTYNAME
[1] 連江縣 宜蘭縣 ...
22 Levels: 南投縣 嘉義市 ...
> head(fortify(taiwan))
Regions defined for each Polygons
  long   lat order hole piece id group
1 119.9170 26.17518  1 FALSE   1 0  0.1
2 119.9171 26.17517  2 FALSE   1 0  0.1
```

## Slot

- What is slot? → a component of the object
- Where are they? → Use `slotNames` to find the name of slots and the operator "@" to extract.

## Fortify

- `fortify` convert curves and points object to a data frame.

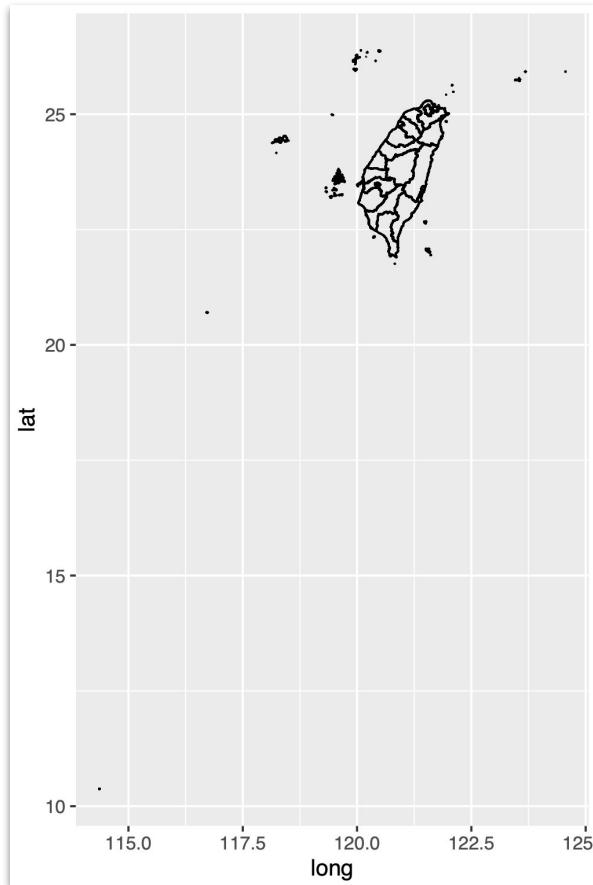
# (1)畫邊界圖

畫臺灣縣市邊界圖

```
ggplot() +  
  geom_path(data = taiwan,  
            aes(x = long, y = lat, group = group))
```

中華民國領土主權實際控制四極點為：

- 東：基隆市中正區棉花嶼，東經  $122^{\circ}06'25''$
- 南：高雄市旗津區南沙群島太平島，北緯  $10^{\circ}22'38''$
- 西：高雄市旗津區南沙群島太平島，東經  $114^{\circ}21'59''$
- 北：連江縣東引鄉北固礁，北緯  $26^{\circ}23'58''$



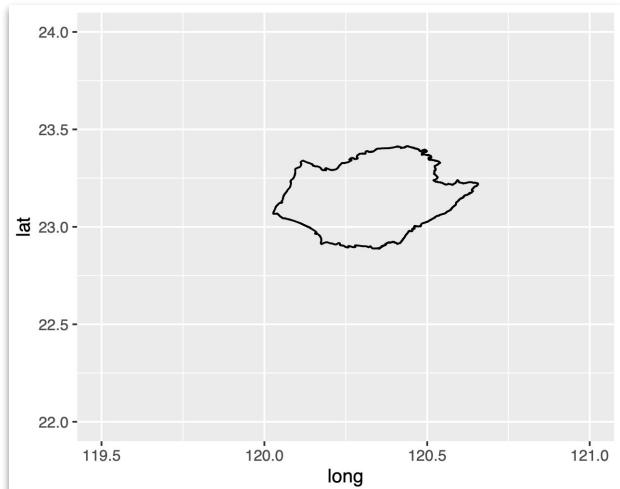
資料來源：<https://zh.wikipedia.org/wiki/%E4%B8%AD%E8%8F%AF%E6%B0%91%E5%9C%8B%E7%96%86%E5%9F%9F>

# (1)部分區域

畫臺南邊界圖

```
map <- subset(taiwan, taiwan$COUNTYNAME %in% c("臺南市"))

ggplot() +
  geom_path(data = map,
            aes(x = long, y = lat, group = group)) +
  xlim(119.5, 122) +
  ylim(21, 24.5)
```



## (2) 空汚 Pm 2.5 分佈 demo

## (2) 讀資料:台灣縣市邊界圖+PM2.5濃度

```
taiwan <- rgdal::readOGR("COUNTY_MOI_1090820.shp")
load(file = "pm.RData")
```

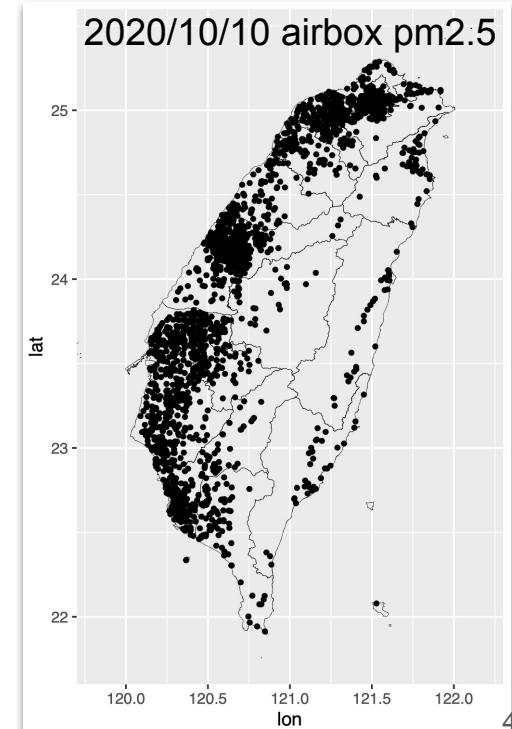
Data	
data_points	2976 obs. of 3 variables
data_raster	32326 obs. of 3 variables
taiwan	Large SpatialPolygonsDataFrame (22 elements, 6.1 Mb)

### data\_points

- 資料來源: 2020/10/10 airbox pm2.5
- 變數: 經度、緯度、pm2.5 濃度
- 資料點: 2976 個

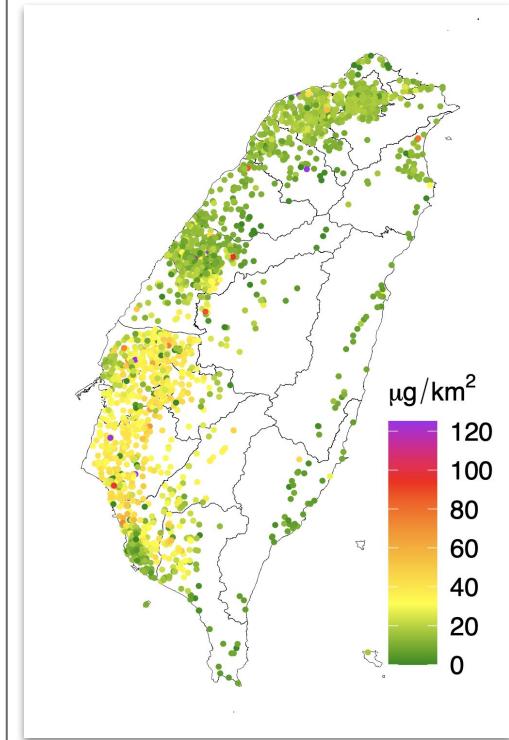
### data\_raster

- 變數: 經度、緯度、pm2.5 濃度
- 資料點: 32326 個 (台灣邊界內間隔為0.01之網格點)
- 利用 kriging 作為未觀測點之pm2.5濃度之估計值



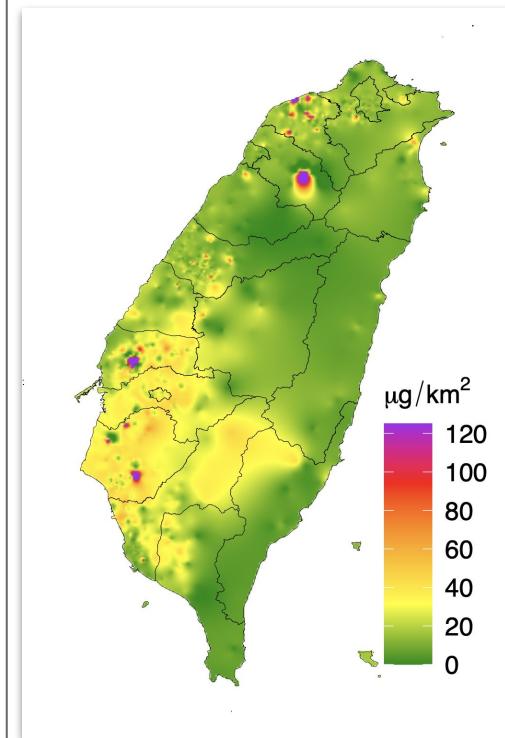
## (2)Point: 畫PM2.5濃度分佈

```
ggplot(data_points, aes(x=lon, y=lat, color=pm25)) +  
  geom_point(size=1) +  
  geom_path(data = taiwan,  
            aes(x = long, y = lat, group = group),  
            color = 'black', size = .1) +  
  scale_colour_gradientn(colours = c("green4","yellow","orange","red","purple"),  
                        values = seq(0,1, length.out = 10),  
                        limits = c(0,125),  
                        guide = "colourbar",  
                        name = bquote(mu * g / km^2),  
                        breaks = seq(0,120,by=20),  
                        labels = seq(0,120,by=20)) +  
  coord_cartesian(xlim = c(119.7, 122.3), ylim = c(21.6,25.6), expand = FALSE) +  
  theme_void() +  
  theme(legend.position=c(0.85,0.3),  
        legend.text = element_text(size = 15),  
        legend.key.size = unit(1, "cm"),  
        legend.title = element_text(size = 15))
```



## (2)Raster: 畫估計的PM2.5濃度分佈

```
ggplot(data_raster, aes(x=lon, y=lat, color=pm25)) +  
  geom_raster(aes(fill = pm25)) +  
  geom_path(data = taiwan,  
            aes(x = long, y = lat, group = group),  
            color = 'black', size = .1) +  
  scale_fill_gradientn(colours = c("green4","yellow","orange","red","purple"),  
                       values = seq(0,1, length.out = 10),  
                       limits = c(0,125),  
                       guide = "colourbar",  
                       name = bquote(mu * g / km^2),  
                       breaks = seq(0,120,by=20),  
                       labels = seq(0,120,by=20)) +  
  coord_cartesian(xlim = c(119.7, 122.3), ylim = c(21.6,25.6), expand = FALSE) +  
  theme_void() +  
  theme(legend.position=c(0.85,0.3),  
        legend.text = element_text(size = 15),  
        legend.key.size = unit(1, "cm"),  
        legend.title = element_text(size = 15))
```

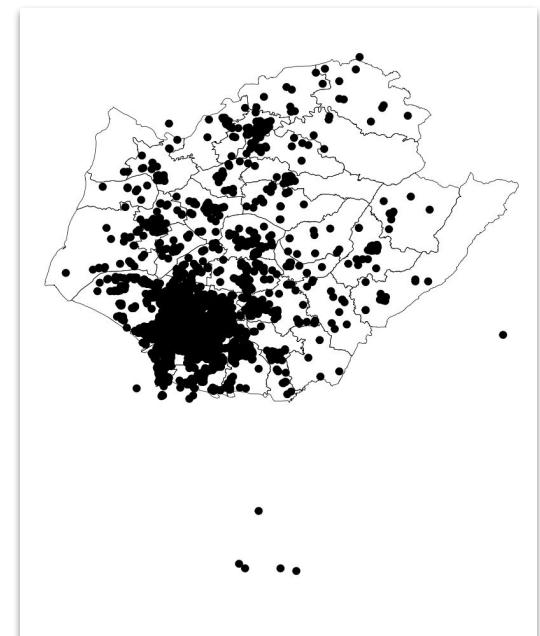


### (3) 登革熱分佈 exercise

## (3) 畫104年臺南登革熱分佈 exercise

Step 1: 讀資料

Step 2: 畫圖



## (4) 台灣人口數分佈 demo

taiwan  
(sp)

taiwan\_map  
(df)

## (4) 讀資料1：台灣縣市鄉鎮邊界圖

```
taiwan <- rgdal::readOGR("COUNTY_MOI_1090820.shp")
taiwan_map <- fortify(taiwan)
taiwan_map$id <- as.numeric(as.character(taiwan_map$id))
```

`fortify` convert curves and points object to a data frame. (1135666 x 7)

▲	long	lat	order	hole	piece	id	group
1	121.4098	23.21369	1	FALSE	1	0	0.1
2	121.4098	23.21366	2	FALSE	1	0	0.1
3	121.4099	23.21362	3	FALSE	1	0	0.1
4	121.4099	23.21361	4	FALSE	1	0	0.1
5	121.4100	23.21351	5	FALSE	1	0	0.1
6	121.4098	23.21346	6	FALSE	1	0	0.1

## (4) 讀資料2: 108年台灣人口統計

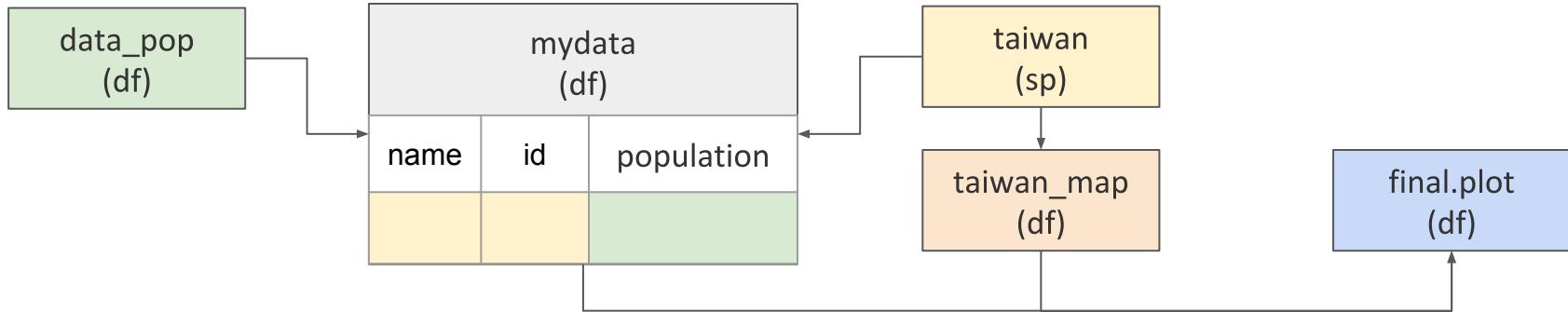
data\_pop  
(df)

```
data_pop <- read.csv(file = "108TaiwanPopulation.csv", header = T)
data_pop <- data_pop[-c(1,370:375),]
data_pop$people_total <- as.numeric(as.character(data_pop$people_total))
data_pop$site_id <- as.character(data_pop$site_id)
```

	statistic_yyy	site_id	people_total	area	population_density
1	統計年	區域別	年底人口數	土地面積	人口密度
2	108	新北市板橋區	556897	23.1373	24069
3	108	statistic_yyy	site_id	people_total	area
4	108	370	108	東沙群島	...
	371	108	南沙群島	...	0.5045
	372				
	373		說明：1.人口密度係...		
	374		2.96年12月起，我...		
	375		，由高雄市代管；原...		

資料來源：[政府開放資料](#) 13

## (4)結合資料



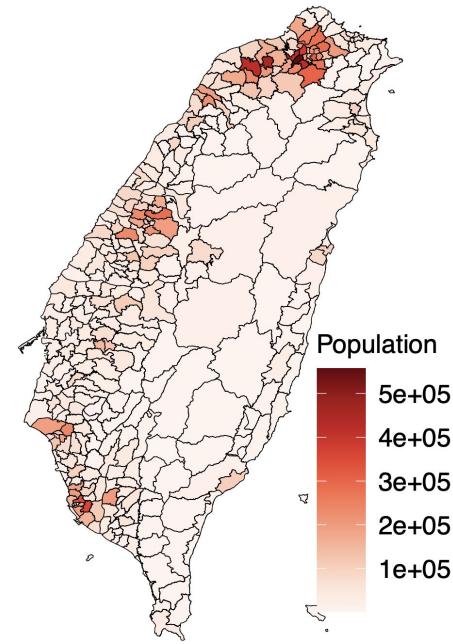
```
mydata <- data.frame(name = paste0(taiwan@data$COUNTYNAME, taiwan@data$TOWNNAME, sep = ""),
                      id      = row.names(taiwan@data),
                      population = 0)

for(i in 1:nrow(mydata)){
  mydata$population[i] <- data_pop$people_total[which(data_pop$site_id == mydata$name[i])]
}

final.plot <- merge(taiwan_map, mydata, by="id", all.x=T)
```

## (4) polygon: 畫人口數分佈圖

```
ggplot() +  
  geom_polygon(data = final.plot,  
               aes(x = long, y = lat, group = group, fill = population),  
               color = "black", size = 0.2) +  
  scale_fill_gradientn(colours = brewer.pal(9,"Reds"),  
                       name = "Population") +  
  coord_cartesian(xlim = c(119.7, 122.3),  
                  ylim = c(21.6,25.6),  
                  expand = FALSE) +  
  theme_void() +  
  theme(legend.position=c(0.85,0.3),  
        legend.text = element_text(size = 15),  
        legend.key.size = unit(1, "cm"),  
        legend.title = element_text(size = 15))
```





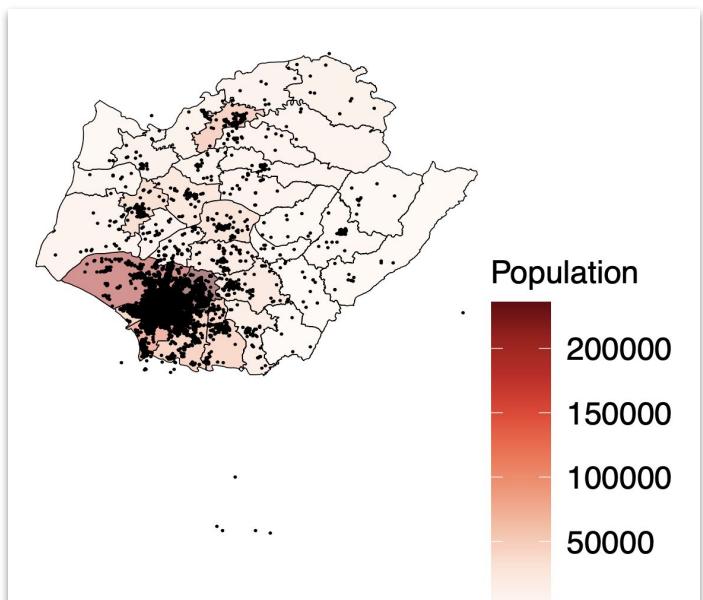
## (5) 登革熱和臺南鄉鎮人口密度分佈 exercise

# (5) 畫登革熱和臺南鄉鎮人口密度分佈 exercise

Step 1: 讀資料

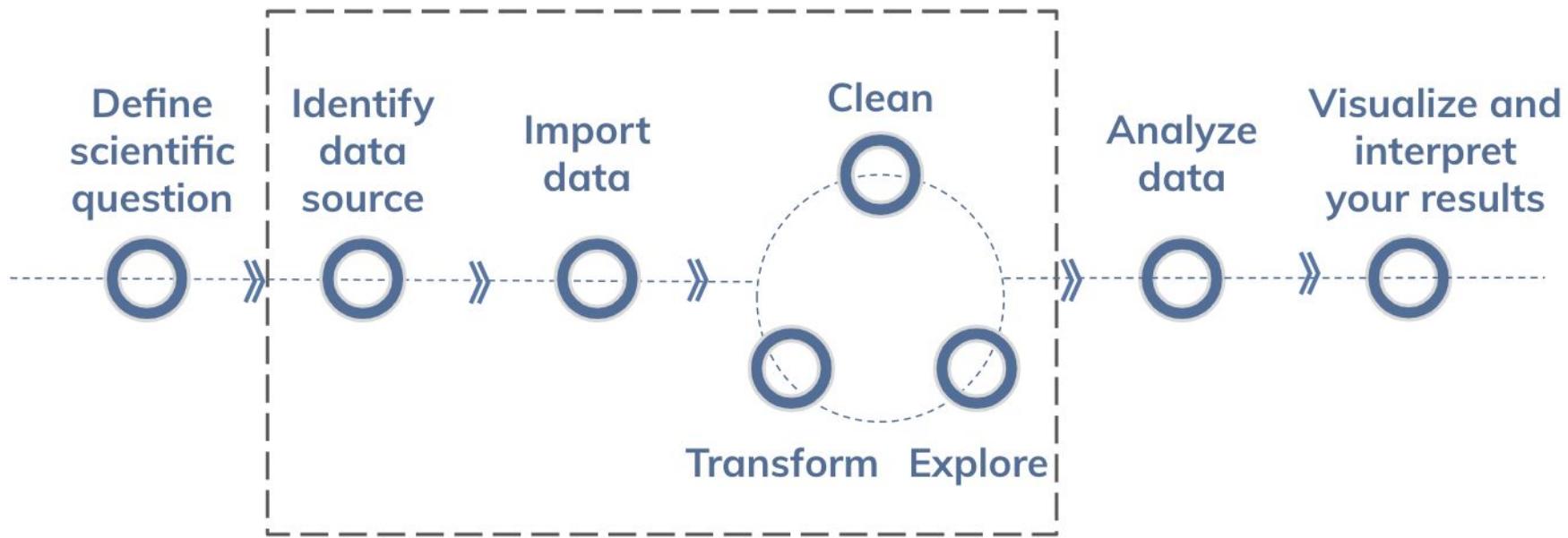
Step 2: 結合資料

Step 3: 畫圖



# Take home message

## Take home message (1/3)



# Take home message (2/3)

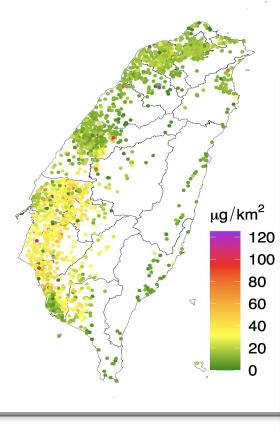
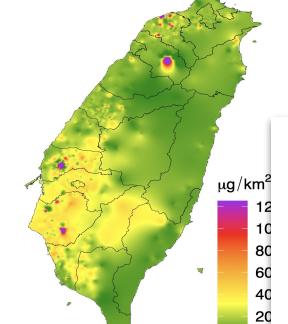
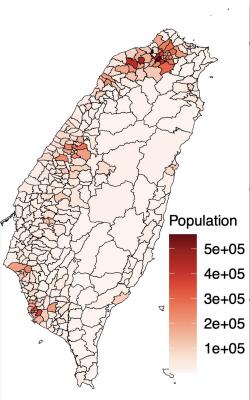
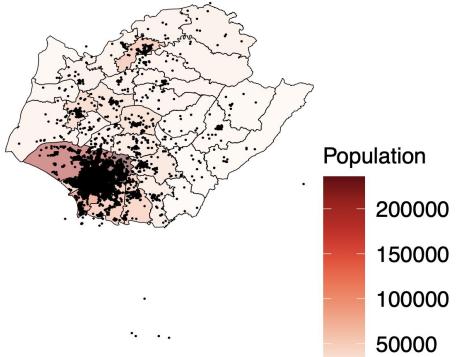
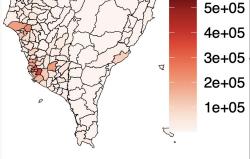
## 第一部分：空間資料

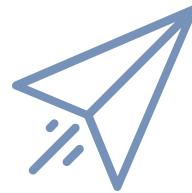
- 空間資料
  - 空間資料：屬性資料 + 空間位置
  - GIS記錄的方式 (Spatial data models): raster data and vector data
- 資料格式
  - Raster data: JPEG, PNG, TIFF
  - Vector data: SHP, GeoJson, KML
- 資料來源
  - 各國政府開放資料
  - 研究機構開放資料
  - Dataset Search

## 第二部分：視覺化探索

- 簡介工具
- R中空間資料物件類型 → R sp package
  - SpatialPointsDataFrame
  - SpatialLinesDataFrame
  - SpatialPolygonsDataFrame
- 如何利用R探索空間資料
  - 台灣邊界
  - 空污 Pm 2.5 分佈
  - 登革熱分佈
  - 台灣人口數分佈
  - 登革熱和臺南人口分佈

# Take home message (3/3)

Task name		Goal
台灣邊界		<ul style="list-style-type: none"> <li>● 讀 .shp 檔案</li> <li>● 利用 geom_path 畫邊界圖</li> <li>● 例用 subset 選擇部分地區</li> </ul>
空污 Pm 2.5 分佈		<ul style="list-style-type: none"> <li>● 讀.RData 檔案</li> <li>● 利用 geom_point 畫散佈圖</li> <li>● 利用 geom_raster 畫網格圖</li> </ul>
登革熱分佈		
台灣人口數分佈		<ul style="list-style-type: none"> <li>● 利用 fortify 轉換 points object 成為 data frame</li> <li>● 結合 csv 和 shp 資料</li> <li>● 利用 geom_polygon 畫區域分佈圖</li> </ul>
登革熱和臺南人口密度分佈		



# Thank you!

Feel free to contact me if you have any question.

Email: [ytfan.st07g@nctu.edu.tw](mailto:ytfan.st07g@nctu.edu.tw)