



2020 資料創新應用競賽

組別：內政大數據資料應用組

設組單位：內政部統計處

大綱

01 本組主題及競賽方式說明

02 內政大數據模擬資料介紹

03 內政大數據 SEGIS 介紹

04 AWS雲端服務介紹

05 競賽作品主題發想

01

本組主題及競賽方式說明

歷年設組主題

2018

- 社會經濟空間資料暨戶政應用組

2019

- 社會經濟空間資料暨地政應用組

2020

- 內政大數據資料應用組

運用資源及作品型式



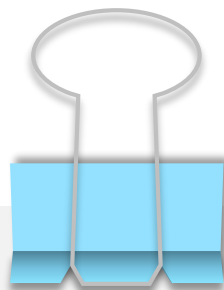
使用資料

1. 內政大數據**模擬**資料
2. 結合「國土資訊系統社會經濟資料服務平台」(SEGIS)之各項開放資料
3. 搭配其他開放資料



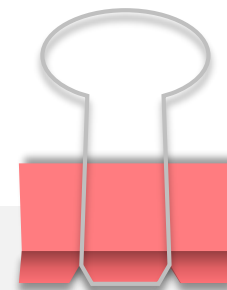
SEGIS服務

1. 統計地圖 **API**
2. 統計資料網路服務 (**Web service**)



AWS雲端服務

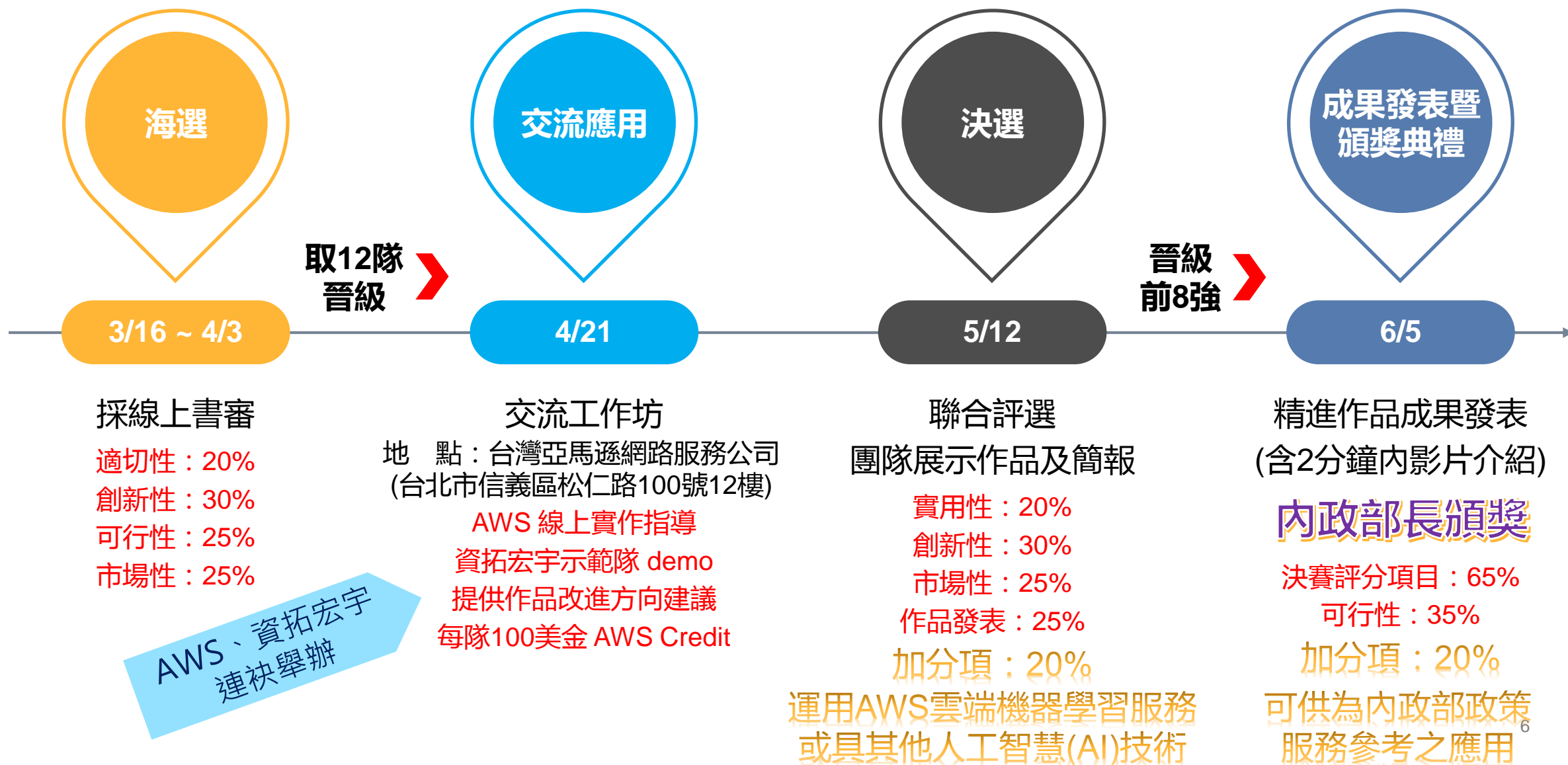
1. Amazon S3 (物件儲存服務)
 2. Amazon SageMaker (機器學習服務)
- P.S.作品亦可運用其他雲端服務或人工智慧技術



作品型式

1. 分析模型
2. **Web** 或 **APP**(iOS、Android)服務

競賽流程及評分方式



獎勵方式

決選



頒發績優團隊獎8隊

每隊各提供AWS績優點數

成果發表暨
頒獎典禮

內政部長頒獎

金獎15萬1隊



銀獎8萬1隊



銅獎5萬1隊



榮譽獎1萬5隊



成果發表暨頒獎典禮

(一) 時間：109年6月5日(星期五)

(二) 特色：*邀請各界產學研單位共襄盛舉，了解內政大數據之應用趨勢，並促進資料應用人才與產業代表媒合機會*

(三) 議程：



時間	議程
12:30~14:00	績優團隊最終評選
13:30~14:00	來賓報到&攤位導覽
14:00~14:15	短講：內政大數據 服務驅動力！ ★由內政部代表分享大數據將如何帶動政府服務的變革
14:15~14:30	短講：政府大數據雲端分析暨增值應用 ★邀請AWS分享國內外政府大數據及AI技術於雲端的運用
14:30~14:45	開場儀式 ★由內政部長與各界代表共同為內政大數據應用揭幕
14:45~15:15	內政大數據 資料應用閃電秀 ★邀請8隊績優隊伍以2分鐘內之成果影片+Demo方式，精采呈現創意構想
15:15~15:30	資料創新應用競賽 頒獎典禮 ★將邀請部長前來頒獎，讓獲獎團隊倍感榮耀
15:30~16:30	跨界交流媒合會 ★擴大邀約本組海選晉級團隊進行展示及媒合 ★邀約各界產學研單位前來參觀與交流，擴大比賽效益
16:30~16:40	尾聲

02

內政大數據模擬資料介紹

內政大數據模擬資料

為配合辦理「資料創新應用競賽」及本部109年建教合作試辦計畫，針對前2年本部以戶籍人口及地籍建物資料串接的成果，產製約5%之模擬資料，其中至少10項欄位之特徵與整體資料相符。

- 內政部統計處自產初步版，完成時間約2月底。
- 大數據案委外廠商產製版，完成時間約3月底。



屋主資料 (HOUSE_OWN_SUBSET_SUMY_RAWDATA)

資料期間:107年8月底、108年8月底

項次	資料欄位名稱	項次	資料欄位名稱
1	建物所有權人序號	10	持有面積
2	建物所在縣市	11	年齡
3	建物所在鄉鎮市區	12	性別
4	用途分類名稱	13	婚姻狀況
5	建材分類	14	教育程度註記
6	層數組別	15	第一個小孩出生日期
7	登記原因	16	初次持有房屋年份
8	屋齡	17	建物總樓層數
9	持有年數		

人+建物+地理資訊資料 (MOI_PEOPLE_HOUSE_DTL)

項次	資料欄位名稱	項次	資料欄位名稱
1	序號	17	最近遷入日期
2	戶長序號	18	最近遷出戶籍縣市鄉鎮區
3	戶號(序號)	19	第一個小孩出生日期
4	戶籍縣市鄉鎮區	20	初次持有房屋年份
5	戶籍村里	21	出生子女數
6	稱謂(與戶長關係)	22	是否與子女戶籍同縣市
7	性別	23	有無殼分類
8	出生日期	24	建物層數組別
9	出生地	25	主要用途分類名稱
10	教育程度	26	主要建材分類
11	婚姻狀況	27	建物面積(平方公尺)
12	原住民身分	28	屋齡
13	戶口組織型態_依家庭收支調查	29	戶籍街路門牌二級統計區
14	身心障礙類別	30	斷層帶註記
15	中低收入戶列冊款別	31	土壤液化註記
16	遷徙次數	32	順逆向坡註記

1: 單人戶
2: 夫婦戶
3: 單親戶
4: 核心戶
5: 祖孫戶
6: 三代戶
7: 其他戶

A: 標準有殼(本人有殼)
B: 標準無殼(本人無殼)
C1: 與屋主關係為成年子女/其他人士(住宅為本戶所有)
C2: 與屋主關係為成年子女/其他人士(住宅不為本戶所有)
D1: 與屋主關係為配偶父母未成年子女(住宅為本戶所有)
D2: 與屋主關係為配偶父母未成年子女(住宅不為本戶所有)
E: 持有房屋不在戶籍同縣市(住宅不為本戶所有)
F: 持有房屋在戶籍同縣市(住宅不為本戶所有)
NULL: 無法分類

03

SEGIS資料與服務介紹

SEGIS 資料查詢下載

SEGIS 社會經濟資料服務平台

關於本分組 分組成果 資料與服務 統計地圖

現在位置: 首頁

資料與服務查詢

最新消息

產品公告

統計資料網路服務

開放產品合併下載

統計圖比對服務

影音專區

活用SEGIS 生活有意思

財利經濟 運輸觀光

各類產品數及下載次數

產品總數: 80,497

產品累計下載次數: 527,451

統計地圖專區

統計區MAP

臺灣通用電子地圖

關閉底圖

資料時間或統計期

107年

108年06月

108年06月

107年

107年

107年

107年

107年

106年

106年

99年

SEGIS © 2020 內政部

<https://segis.moi.gov.tw>

現在位置: 首頁 > 資料與服務 > 產品與服務查詢

開放資料 一般查詢 空間性別統計

關鍵字

縣市搜尋: 全國(22縣) 請選擇類別: 人口 不限小類

產品名稱 產品欄位名稱 查詢

請選擇中類

資料搜尋條件: 人口

15中類

加入購物車即可下載或申請資料

縮小搜尋範圍

資料時間

不限

今年

近三年

~年 確定

空間統計單元

不限

縣市別

鄉鎮市區別

村里別

二級發布區

一級發布區

最小統計區

開放程度

開放資料

需申請

資料類型

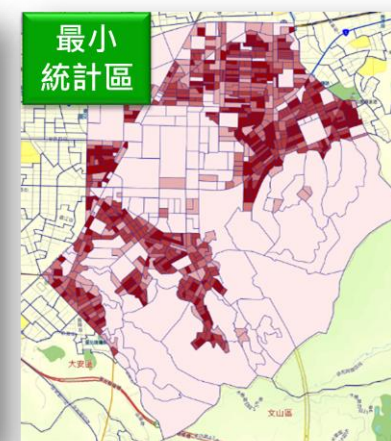
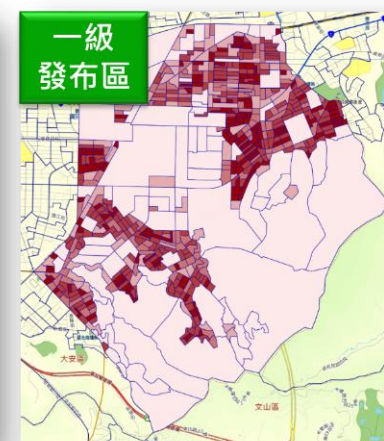
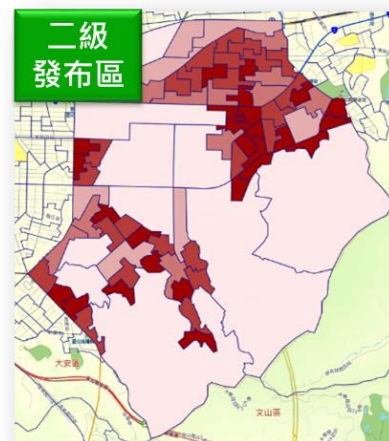
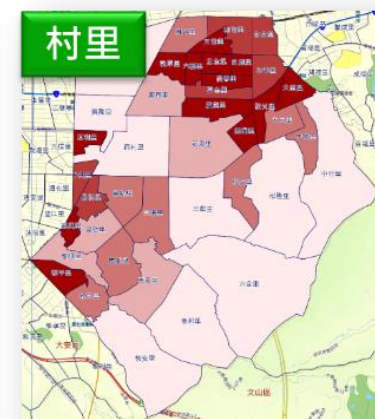
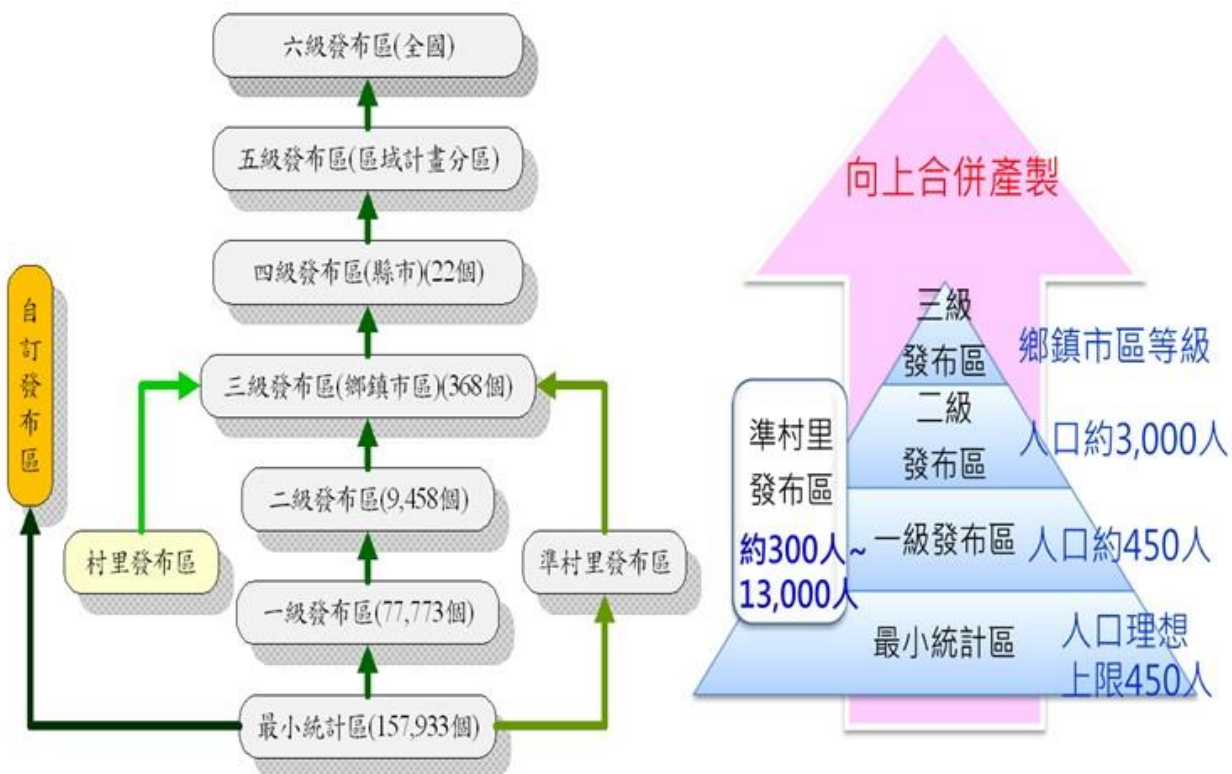
產品

網路服務

搜尋條件

產品與服務名稱	空間統計單元	共1203筆
1 108年9月...	村里	
2 108年9月...	村里	
3 108年9月...	村里	
4 108年9月...	村里	
5 108年9月...	村里	
6 108年9月...	村里	
7 108年9月...	村里	
8 108年9月...	村里	
9 108年9月...	村里	
10 108年9月...	村里	
11 108年9月...	村里	
12 108年9月...	村里	
13 108年9月...	村里	
14 108年8月...	村里	
15 108年6月...	村里	
16 108年6月...	村里	

統計區分類系統



產製競賽用資料集

公布日詳見SEGIS社會經濟資料服務平台

1. 全國各縣市最小統計區統計資料集

最小統計區代碼	欄位1	欄位2	欄位3	欄位4	欄位5	欄位6	欄位7	...	資料時間
A0204-0001-00									
A0204-0003-00									
A0204-0002-00									
A0204-0004-00									
A0204-0006-00									
A0204-0005-00									
:									

2. 全國各村里統計資料集

村里代碼	欄位1	欄位2	欄位3	欄位4	欄位5	欄位6	欄位7	...	資料時間
10002100-005									
10002100-017									
10002100-015									
10002100-010									
10002100-002									
10002100-003									
:									

3. 機構點位圖層、機構別統計

醫療衛生

1. 醫療院所分布圖
2. 護產機構分布圖

教育文化

1. 各級學校校別概覽(國小、國中等)
2. 各類宗教分布圖(寺廟、法人教會等)

社會福利

1. 社會福利機構
2. 社區照顧關懷據點

退輔會資料

所屬機構別概覽之統計(包含安養機構、農林機構、醫療機構、服務機構)

全國各縣市最小統計區統計資料集

人口統計

1. 人口統計(戶數、男女人口數等)
 2. 人口指標(性比例、人口密度等)
 3. 五歲年齡組性別人口統計
 4. 三段年齡組性別人口統計
 5. 分齡兒童及少年性別人口統計
 6. 15歲以上人口五歲年齡組與性別與教育程度人口統計
 7. 15歲以上人口五歲年齡組與性別與婚姻狀況統計
 8. 15歲以上人口十歲年齡組與性別與教育程度與婚姻狀況統計
 9. 原住民人口指標
 10. 原住民人口統計(平地、山地等)
 11. 原住民戶數統計
 12. 原住民15歲以上人口十歲年齡組與性別與教育程度人口統計
 13. 原住民15歲以上人口十歲年齡組與性別與婚姻狀況人口統計
-等

土地及營建統計

1. 國土利用調查統計
2. 不動產實價登錄建物成交資訊(交易日)—按屋齡分、按建物型態分

醫療衛生統計

1. 醫療院所統計
2. 護產機構統計

教育文化統計

單歲年齡組學齡人口統計

台北市
電腦公會

A6301-0729-00

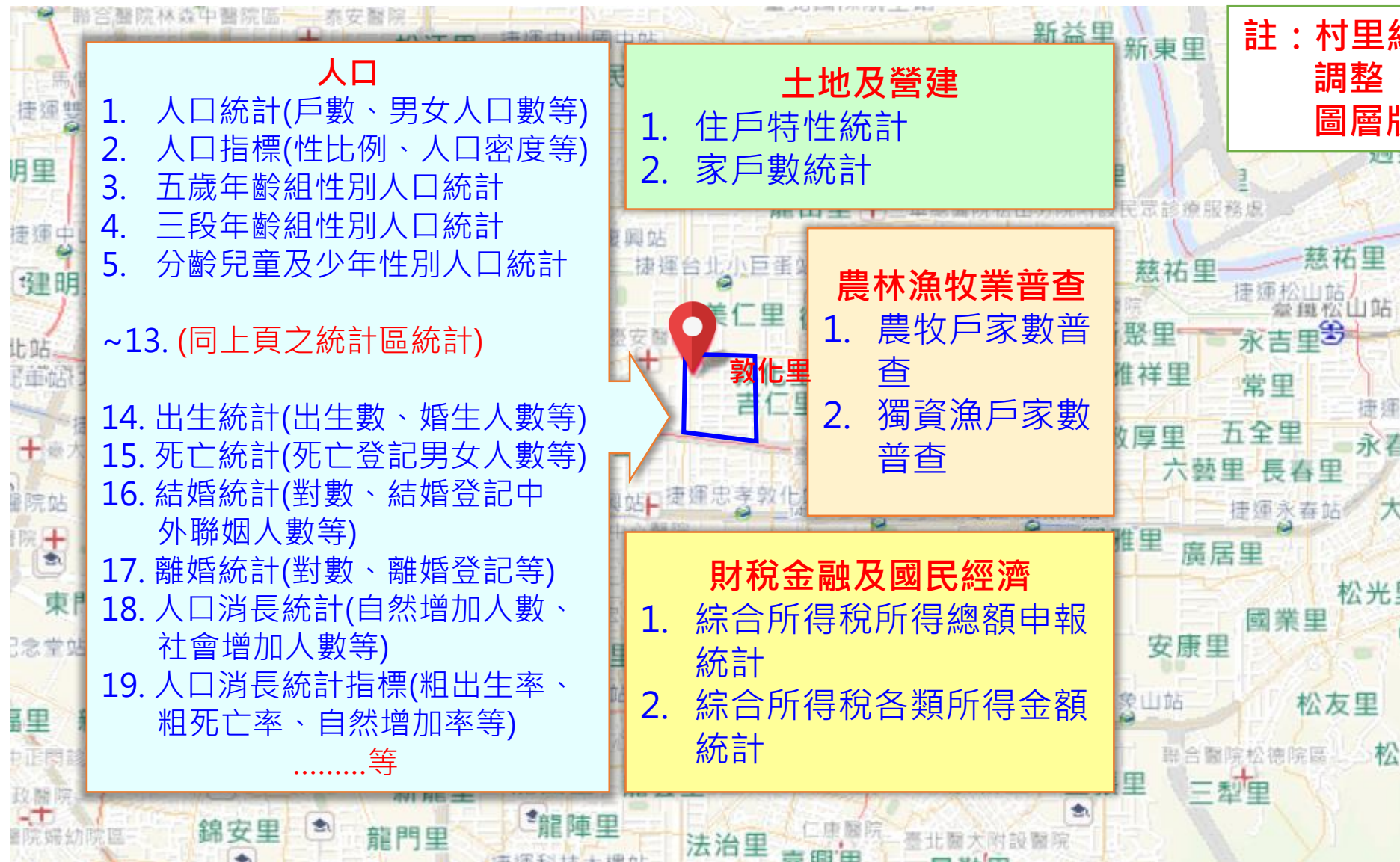
工業、商業及其他服務業統計

1. 連鎖便利商店家數統計
2. 工商家數(工商業總家數,農、林、漁、牧業,....等各行業別家數)

社會福利統計

1. 低收入戶統計指標(低收入戶占總戶數比例等)
2. 中低收入戶統計指標
3. 老人安養機構統計
4. 身心障礙者統計指標
5. 社區關懷據點統計
6. 社會福利機構統計(兒福、婦幼、老人、身障、少年福利服務機構家數等)

全國各村里統計資料集



統計地圖API雲端服務 <https://segisapi.moi.gov.tw>

統計地圖API
範例網站



Edit Code: 重置 展開程式碼 複製程式碼 檢視成果 >>

```
1 <!--程式碼範例-->
2 <!DOCTYPE html>
3 <html>
4 <head>
5   <title>行政區層級下探</title>
6   <link type="text/css" rel="stylesheet" href="/STAT_API_SAMPLE/Content/Adv/css/sample2.css" />
7   <link href="/STAT_API_SAMPLE/Content/Adv/div_mask.css" rel="stylesheet" />
8   <script src="/STAT_API_SAMPLE/Content/Adv/jscolor/jscolor.js"></script>
9   <script src="/STAT_API_SAMPLE/Content/Adv/events/my_event.js"></script>
10  <script language="javascript" type="text/javascript" src="https://segisapi.moi.gov.tw/STAT_API_SAMPLE/js/jquery"></script>
11  <script language="javascript" type="text/javascript" src="https://segisapi.moi.gov.tw/STAT_API_SAMPLE/js/jquery"></script>
12  <script language="javascript" type="text/javascript" src="https://segisapi.moi.gov.tw/StatAPI/statistics?appid=
13  <!-- 下載後請將yourID及yourkey取代之為您申請所取得的APPID及APIKEY方能正確顯示服務-->
14  <script type="text/javascript" src="https://api.tgoss.gov.tw/TGOS_API/tgoss?ver=2&appid=TGOSID&apikey=TGOSKEY" char:
15  <!-- 下載後請將TGOSID及TGOSKEY取代之為您申請之APPID及APIKEY方能正確顯示服務-->
16  <script type="text/javascript">
17    var myMap;
18    var dataDB;
19    var tgossMap;
20    var tgKML = null;
21    var opacityStart = 0.5;
22    var opacityStop = 1;
23    var draw_status = 0; //判斷是否為首次繪圖
24    var timeList = ["104Y"];
25    var regionCode = ['10002', '10004', '10005', '10007', '10008', '10009',
26    '10010', '10013', '10014', '10015', '10017', '10018', '10020',
27    '63000', '64000', '65000', '66000', '67000', '68000'];
28    var infoType = {
29      STATINFO: "STATINFO",
30      STATLEGEND: "STATLEGEND"
31    };
32    function init() {
33      myMap = new StatMap();
34      myChart = new StatChart(document.getElementById("chartView"));
35      //共用變數區
36      var dataSet = new Array(); //建立空陣列，準備納入統計資料設定集合
37      var tmp = { //建立統計資料設定
38        tableName: "AA9FE_A4C3", //指定統計表
39        fieldsId: ["STU_CNT"] //指定使用欄位
40      };
41      dataSet.push(tmp);
42      var timeOpts = { //指定統計時間
43        timeSeries: true
```

可線上修改
範例程式

行政區層級下探

資料設定

Step 1
產品選擇: 行政區國民小學統計
使用統計欄位: 國小學生數
選擇時序年度: 104年

Step 2
統計地圖參數:
地圖類型: 面量圖
分級法: 等距法
分級數: 5
配色: 最小值色彩: FFD6C1
 最大值色彩: 420000
 預設色彩: D0D0D0

Step 3
繪製統計地圖

100公里 / 100英里
巴士海峽
TGOS © 2020 內政部

統計地圖API雲端服務

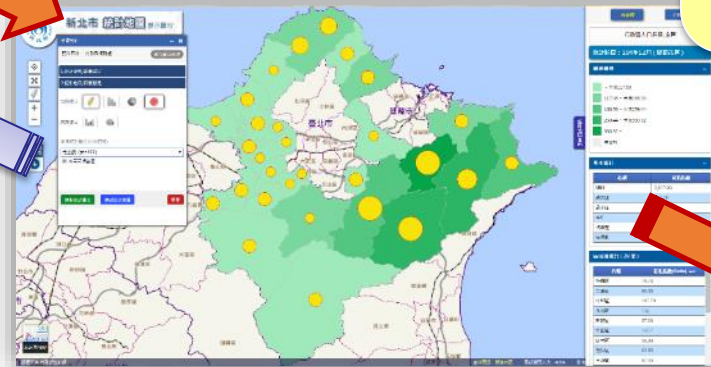


前端使用者

送出查詢需求

1. 送出應用資料庫資料欄位參數+繪製參數
2. 送出自有資料XML+繪製參數

展示統計圖以及相關資訊供前端使用者瀏覽

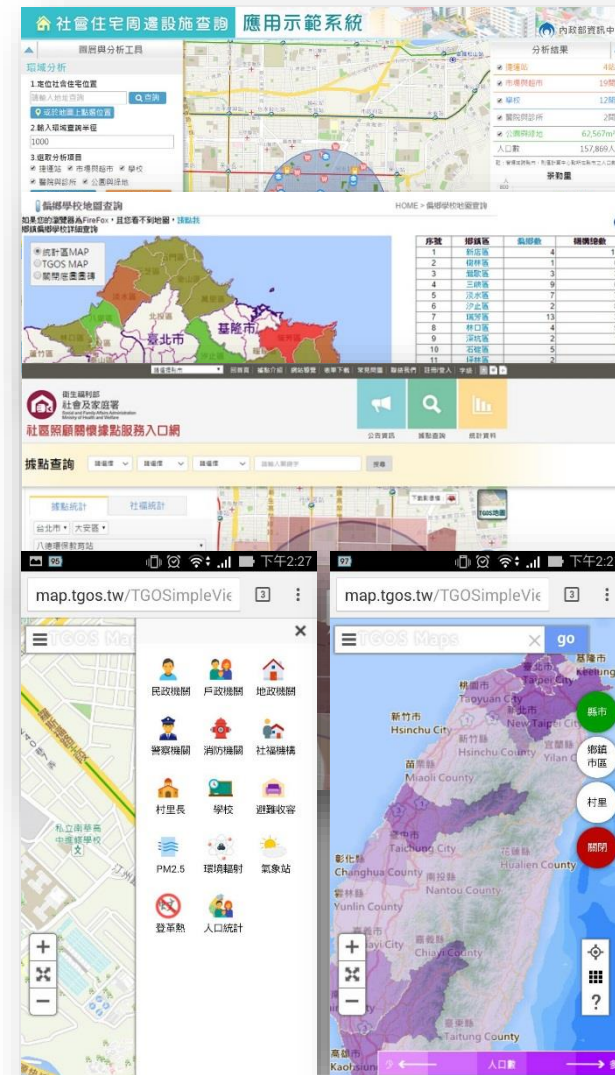


介接API的開發網站

回傳統計圖、互動式介面或相關統計資訊



API 伺服器



網路服務(Web Service)

資料時間

不限

今年

近三年

~ 年

空間統計單元

不限

縣市別

鄉鎮市區別

村里別

二級發布區

一級發布區

最小統計區

開放程度

開放資料

需申請

資料類型

產品

網路服務



20 第 1 共 3 頁

產品與服務名稱

- 1 [統計地圖API \(Javascript API\)](#)
- 2 [統計資料網路服務](#)
- 3 [統計區比對服務 V2.0](#)
- 4 [農產品交易行情](#)
- 5 [每月盛產農產品產地](#)
- 6 [臺灣地區雜糧生產概況](#)
- 7 [農產品生產成本查詢](#)
- 8 [臺灣地區蔬菜生產概況](#)
- 9 [糧價查詢](#)
- 10 [臺灣地區果品生產概況](#)
- 11 [臺灣地區特用作物生產概況](#)
- 12 [臺灣地區蔬菜生產概況-菇類](#)
- 13 [臺灣地區花卉生產概況](#)
- 14 [臺灣地區農產品生產量值](#)
- 15 [農產品產地價格\(月平均價格\)](#)

資料與服務專區

產品與服務查詢



最新消息



產品公告



統計資料網路服務



開放產品合併下載



統計區比對服務



請選擇要前往的網路服務

統計資料網路服務



單點統計查詢網路服務



查詢單點坐標所屬之各空間階層較為重要且開放的社會經濟統計資料

04

AWS雲端服務介紹

AWS雲端服務介紹

- 競賽使用服務
 - Amazon S3 (物件儲存服務)
 - Amazon SageMaker (機器學習服務)
- 線上免費資源
 - AWS 免費方案：<https://aws.amazon.com/tw/free/>
 - Amazon SageMaker 教學：
<https://tinyurl.com/uwoyq3q>
- 技術支援：
 - AWS User Group Taiwan：<https://tinyurl.com/rhnkc83>
 - AWS Developer Forum：<https://tinyurl.com/r6sh6vr>



<https://tinyurl.com/y5vt2x2w>

- 可使用免費方案
與網路教學了解
並學習



Amazon SageMaker

Amazon SageMaker

適用於每個開發人員和資料科學家的機器學習

Amazon SageMaker 是一項全受管服務，能讓所有開發人員和資料科學家快速輕鬆地建立、訓練及部署機器學習 (ML) 模型。SageMaker 可消除機器學習過程中每個步驟的繁重工作，讓開發高品質模型變得輕而易舉。

傳統的 ML 開發是一個複雜、昂貴、反覆的過程，因為沒有用於整個機器學習工作流程的整合工具，讓這一切變得更加困難。您需要將工具和 workflows 拼接在一起，這既耗時又容易出錯。

SageMaker 透過在單一工具集中，提供機器學習所用的全部元件來解決這一難題，因此，模型能以更少的工作和更低的成本更快地投入生產。

The screenshot displays the Amazon SageMaker Studio IDE interface. The central pane shows a Jupyter notebook with the following content:

- Have the predictor variable in the first column
- Not have a header row

But first, let's convert our categorical features into numeric features.

```
[ ]: model_data = pd.get_dummies(churn)
model_data = pd.concat([model_data[ 'Churn: True, '], model_data.drop(['Churn: True, '], axis=1)]

...

And now let's split the data into training, validation, and test sets. This will help prevent us from overfitting the model, and allow us to test the models accuracy on data it hasn't already seen.
```

```
[ ]: train_data, validation_data, test_data = np.split(model_data.sample(frac=1, random_state=42), [int(len(model_data) * 0.7), int(len(model_data) * 0.8)])
train_data.to_csv('train.csv', header=False, index=False)
validation_data.to_csv('validation.csv', header=False, index=False)

...

Now we'll upload these files to S3.
```

```
[ ]: boto3.Session().resource('s3').Bucket(bucket).Object(os.path.join(prefix, 'train.csv')).upload_file(train_data.to_csv('train.csv', header=False, index=False))
boto3.Session().resource('s3').Bucket(bucket).Object(os.path.join(prefix, 'validation.csv')).upload_file(validation_data.to_csv('validation.csv', header=False, index=False))

...

Trial Component Chart: A line graph showing 'trainloss_last' on the y-axis (ranging from 0.0 to 0.4) and 'period' on the x-axis (ranging from 0 to 10). Multiple lines represent different trials, showing a general downward trend in loss over time.
```

使用 IDE 進行 ML 開發。例如，對筆記本內部的模型進行更新，並使用筆記本的並排檢視和訓練實驗，來查看變更如何影響模型品質。

05

競賽作品主題發想

關心「蟻居族」

想解決什麼問題	已確定可取用資料	仍需再取得之資料(含不確定是否有此資料)或其他建議
「蟻居」是被高房價所催生出一種居住型態，經濟弱勢族群和社會新鮮人買不起房，只能租生活品質差的便宜房子，由於需求高，房東也樂得多隔幾間來租。可研究找出這些族群之特性，並給予相關政策協助。	<ol style="list-style-type: none"> 1.串連戶籍檔及建物檔，找出同一門牌設籍人數較多者，且由與戶長關係判斷是否為承租者，並估算平均可使用面積。 2.再串連中低收入戶檔、身心障礙檔，並結合年齡，與家人同住狀況，綜合判斷其弱勢程度。 	<ol style="list-style-type: none"> 1.各縣市租金補貼最高補貼金額 2.申請租金補貼資料檔 3.出租成交行情資料 4.社會住宅物件資料 5.包租代管物件資料 6.其他相關議題：無殼族居住特性(面積、區位、家戶組成)及因應之道



蘋果即時
升級壹會員
🔍

最新
焦點
熱門
政治
娛樂時尚

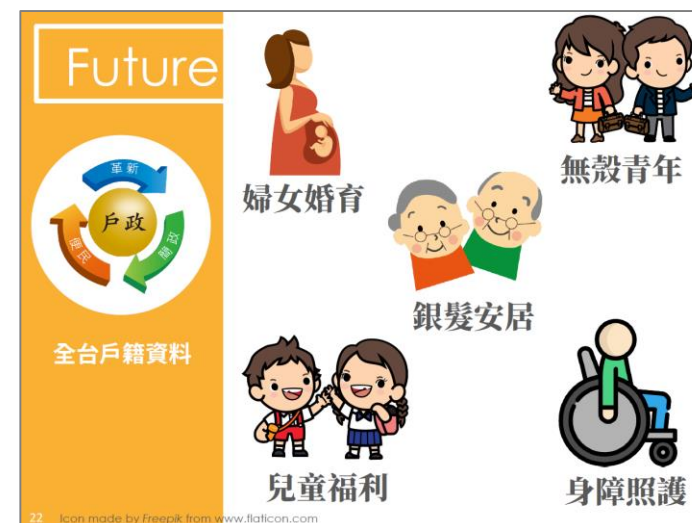
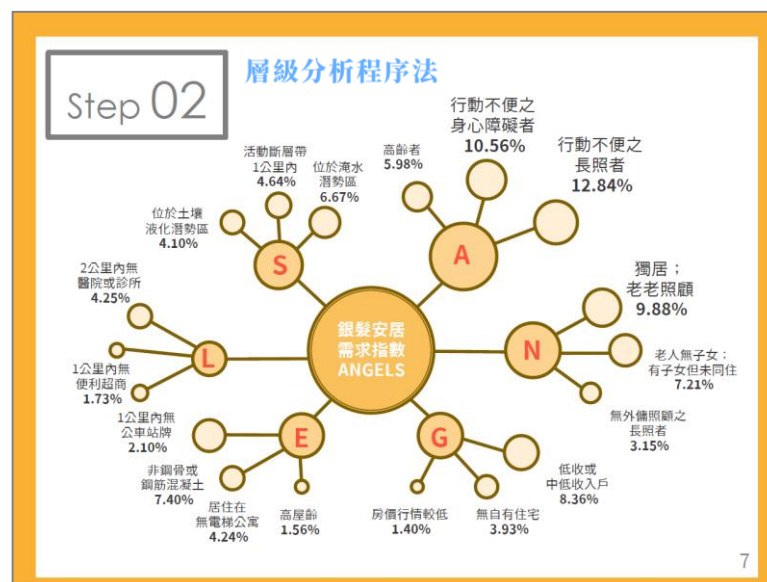
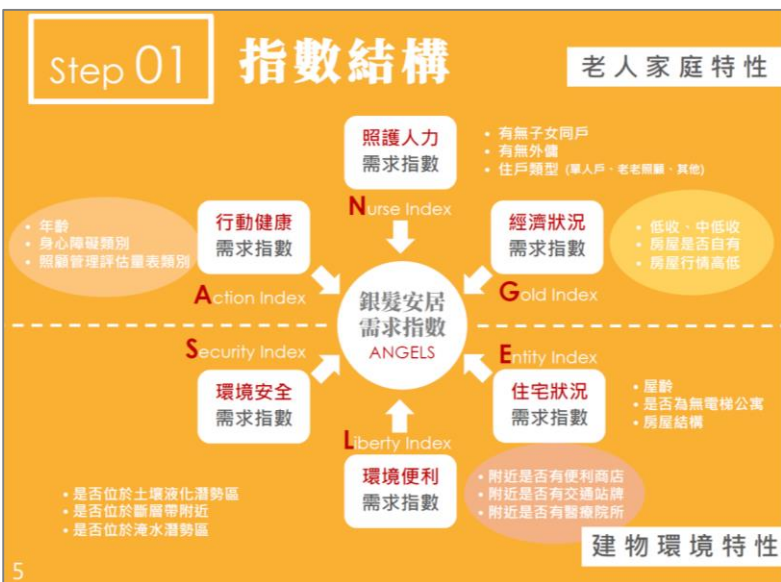
【蟻居人生1】菩薩腳下的孤老租屋 龍山寺「下流老人」2坪求生滿是感恩

建立時間：2019/04/26 21:00

尚餘0篇
f
LINE
⋮

銀髮安居開放資料

想解決什麼問題	已確定可取用資料	仍需再取得之資料(含不確定是否有此資料)或其他建議
內政部與衛生福利部合作之 銀髮安居計畫 ，以18項統計項目找出最需要幫助的老人，獲總統盃黑客松肯定。現欲在不揭露個資前提下，產製開放資料供外界應用，應選取哪些項目及分類標準，才最符外界需求？歡迎提出研究觀點。	1.年齡2.身心障礙類別3.照顧量表類別4.有無子女5.有無外傭6.住戶類型7.中低收入8.房屋權屬9.房屋行情10.屋齡11.是否無電梯公寓12.房屋結構13.附近有否便利商店(14.站牌15.醫療院所)16.位於土壤液化區(17.斷層帶18.易淹水區)	1.歡迎各行各業從產業需求角度提出建議，如醫療照顧產業(需照顧不能行動的老人在哪裡)、金融保險業(高收入男性分布在哪裡)等。 2.另可參考銀髮安居計畫，設計「青年安居」、「托兒安心」等計畫。



週邊環境智慧寫手

想解決什麼問題

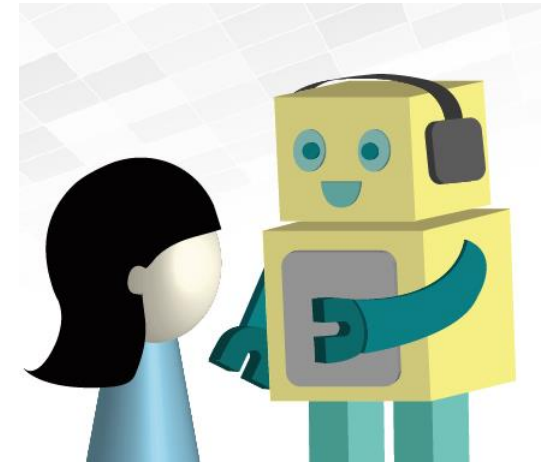
以**簡單扼要的文字**，描述某地（鄉鎮市區、村里）特性資訊，如人口、工商、財稅之統計概況與變化，並以線上互動方式，依使用者所需類別，階段下鑽式提供更細緻空間單元資料或交通、醫療、社福等服務據點資訊。

已確定可取用資料

- 1.內政部社會經濟資料服務平台(SEGIS)
- 2.內政部公務統計資料庫
- 3.國家發展委員會TESAS地方創生資料庫

仍需再取得之資料(含不確定 是否有此資料)或其他建議

- 1.各部會、地方政府發布之縣市別、鄉鎮市區別、村里別資料(例如中央氣象局鄉鎮市區天氣預報資料、農委會農業生產資料)。
- 2.若能與**聊天機器人**結合應用更佳。



SEGIS 社會經濟資料服務平台

關於本分組：分組成果、資料與服務、統計地圖

產品與服務：產品與服務、最新消息、產品公告、統計資料網的業務、開放產品合併下載、統計區比對服務

統計地圖專區：統計地圖、統計地圖API、統計地圖小常識、主題圖集、統計部應用案例

縣市重要統計指標：資料主題、資料時間或統計、指標人口率、107年、低收入戶、108年06月、老人安養機構家數、108年06月、自來水供水普及率、107年

各類產品數及下載次數：產品總數：80,497、產品累計下載次數：527,451

內政部 統計處
Ministry of the Interior Department of Statistics

關於統計：統計處簡介、最新消息、重要統計事項或公告、出版品目錄、內政概要、法令規章

統計報告：我國生命表、內政統計通報、內政統計月報、內政統計年報、內政國際指標、重要參考指標、內政統計指標縣市排名、各縣市內政統計指標、統計調查摘要、專題分析、性別統計專區

人口：你知道國人平均壽命有多長嗎？
84.54, 82.38, 82.62, 82.83, 83.36, 83.19, 83.62, 83.41, 83.70, 84.05
78.57, 79.01, 79.88, 79.15, 79.51, 80.02, 79.81, 80.00, 79.97, 80.69
75.59, 76.03, 76.13, 75.96, 76.05, 76.01, 76.72, 77.01, 76.81, 77.28, 77.55

你知道國人平均壽命有多長嗎？
107年國人平均壽命達80.7歲，其中男性77.5歲，女性84.0歲，皆創歷年新高，女性平均較男性為高，兩者差距於6.2歲至6.7歲之間變動，且國人的平均壽命長期呈增加趨勢，顯示國人已越來越長壽。

TESAS

南投縣草屯鎮

Dashboard: Population, Income, Industry, Transportation, Tourism, Land, Housing, Environment, Education, Culture, Welfare, Medical

人口數(人) 97,509
4G基地台數量 411
(南投縣之)平均每年全住宅支出(元) 123,132
歷史建築數量 7

(南投縣之)可支配所得(元) 758,388
旅遊民宿數量 18
(南投縣之)每萬人公園、綠地面積數(公頃) 2.39
低收入戶全區總戶數(%) 0.96

工商各業別之公司登記家數總計 5,505
(南投縣之)都市計畫工業區(公頃) 311.19
大學以上人數 25,906
診所家數 108

統計地圖：北屯區、新社區、西區、中區、東區、南區、大里區、太平區、霧峰區、草屯鎮



THANKS

歡迎報名參加本組競賽

諮詢窗口

賽程資訊及SEGIS : 王暉鈞 02-23566083
江文意 02-23678720#22
內政大數據模擬資料 : 黃騰暘 02-23565364