

行人安全 X 檔案： 解碼台灣行人事故現象

解碼小組：蔡尚峰、溫庭宇、江翊瑄、葉沛瑩



目錄 TABLE OF CONTENT

- | | | |
|-----------------|--------------------|-----------------|
| 01. 專案背景 | 04. 資料清理、合併 | 07. 未來展望 |
| 02. 專案目標 | 05. 事故總覽 | |
| 03. 使用套件 | 06. 資料分析 | |

01. 專案背景

台灣最美的風景是人，但對於外國旅客來說，交通卻是地獄。

2022年12月，美國CNN報導台灣交通險象環生的情形，澳洲、加拿大、日本與美國等多個國家都明確點出台灣道路狀況問題。身為關心台灣大事的我們，進而促成了本次期末專案的製作動機。

02. 專案目標

透過分析近三年（109年-111年）政府資料開放平臺台灣六都歷史交通事故資料，探討人車事故的成因，包含事故本身條件，如當事人（車）當下的行為等，及外在因素，如事故當下的路面、天氣等。綜合內外條件，提供行人道路安全指南，讓行人得以藉由行人安全地圖，安心上路，平安回家！

03. 使用套件



- 資料合併、清理



- 資料清理
- 匯入檔案



- 分析資料、製作統計圖表



- 結合地圖繪製熱點圖

Folium



04. 資料清理、合併

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	發生年度	發生月份	發生日期	發生時間	發生時間區	事故類別名	處理單位名	發生地點	行政區	天氣名稱	事故位置子	路面狀況-路	路面狀況-路	肇因研判子	死亡人數	受傷人數	當事者區分	當事者區分	當事者屬-性	當事者事故	當事者行動	經度	緯度
2	2020	1	3	23800	02:00-03:00	A1	新北市政府	新北市中和	新北市中和	陰	交叉路口內	柏油	乾燥	搶越行人穿	1	0	小貨車(含客	自用	男	71	車的狀態	121.505006	24.993558
3	2020	1	3	85600	08:00-09:00	A1	臺北市政府	臺北市南港	臺北市南港	晴	交叉路口內	柏油	乾燥	搶越行人穿	1	0	小客車	自用	男	23	車的狀態	121.615942	25.04445
4	2020	1	6	112500		A1	臺北市政府	臺北市中正	臺北市中正	晴	交叉路口內	柏油	乾燥	未注意車前	1	0	小貨車(含客	自用	男	24	車的狀態	121.509371	25.04522
5	2020	1	6	191800		A1	新北市政府	新北市新店	新北市新店	陰	交叉路口內	柏油	乾燥	未注意車前	1	0	小貨車(含客	自用	男	48	車的狀態	121.540362	24.964583
6	2020	1	8	73000		A1	臺中市政府	臺中市市區	臺中市市區	晴	交叉路口內	柏油	乾燥	搶越行人穿	1	0	小客車	自用	女	71	車的狀態	120.6829	24.140782
7	2020	1	9	195546		A1	臺南市政府	臺南市東區	臺南市東區	晴	機車優先道	柏油	乾燥	未注意車前	1	0	機車	普通重型	男	32	車的狀態	120.220595	22.968235
8	2020	1	16	141512		A1	臺南市政府	臺南市永康	臺南市永康	晴	一般車道(未	柏油	乾燥	未注意車前	1	0	大貨車	營業用	男	40	車的狀態	120.258123	23.025031
9	2020	1	16	192100		A1	臺中市政府	臺中市西屯	臺中市西屯	晴	穿越道附近	柏油	乾燥	搶越行人穿	1	0	小客車	計程車	男	52	車的狀態	120.644295	24.166462
10	2020	1	18	80807		A1	臺南市政府	臺南市學甲	臺南市學甲	晴	快車道	柏油	乾燥	起步未注意	1	0	大貨車	營業用	男	43	車的狀態	120.182708	23.234414
11	2020	1	23	183251		A1	臺南市政府	臺南市官田	臺南市官田	晴	快車道	柏油	乾燥	未注意車前	1	0	小客車	自用	男	45	車的狀態	120.327994	23.213424
12	2020	1	24	42300		A1	高雄市政府	高雄市苓雅	高雄市苓雅	晴	交叉口附近	柏油	乾燥	搶越行人穿	1	0	小貨車(含客	自用	女	70	車的狀態	120.300289	22.612522
13	2020	1	28	24700		A1	臺中市政府	臺中市西屯	臺中市西屯	晴	交叉路口內	柏油	乾燥	搶越行人穿	1	1	機車	普通重型	男	19	車的狀態	120.616904	24.188609
14	2020	1	30	211600		A1	高雄市政府	高雄市大樹	高雄市大樹	晴	快車道	柏油	乾燥	未注意車前	1	0	小客車	自用	男	19	車的狀態	120.449403	22.771831
15	2020	1	31	92820		A1	桃園市政府	桃園市平鎮	桃園市平鎮	晴	一般車道(未	柏油	乾燥	未注意車前	1	0	小貨車(含客	自用	男	47	車的狀態	121.284624	24.934476
16	2020	1	31	181100		A1	桃園市政府	桃園市龜山	桃園市龜山	陰	行人穿越道	柏油	乾燥	未注意車前	1	0	機車	普通重型	男	38	車的狀態	121.374971	25.029664
17	2020	2	5	183000		A1	高雄市政府	高雄市鼓山	高雄市鼓山	晴	交叉路口內	柏油	乾燥	搶越行人穿	1	0	小客車	自用	男	39	車的狀態	120.2791	22.64994
18	2020	2	6	203900		A1	臺中市政府	臺中市大雅	臺中市大雅	晴	機車優先道	柏油	乾燥	未注意車前	1	2	機車	普通重型	男	17	車的狀態	120.65412	24.221181
19	2020	2	9	130500		A1	高雄市政府	高雄市前鎮	高雄市前鎮	晴	交叉路口內	柏油	乾燥	搶越行人穿	1	1	機車	普通重型	男	26	車的狀態	120.306121	22.602985
20	2020	2	10	155200		A1	新北市政府	新北市土城	新北市土城	陰	交叉口附近	柏油	乾燥	搶越行人穿	1	0	小貨車(含客	自用	男	55	車的狀態	121.430514	24.964815
21	2020	2	11	41300		A1	臺中市政府	臺中市大肚	臺中市大肚	晴	其他	柏油	乾燥	未注意車前	1	0	小貨車(含客	自用	男	53	車的狀態	120.568376	24.121843
22	2020	2	19	202000		A1	高雄市政府	高雄市岡山	高雄市岡山	晴	一般車道(未	柏油	乾燥	未注意車前	1	1	小客車	自用	男	32	車的狀態	120.309769	22.825539
23	2020	2	27	53400		A1	臺中市政府	臺中市大雅	臺中市大雅	晴	交叉口附近	柏油	乾燥	未注意車前	1	0	小客車	自用	女	46	車的狀態	120.663693	24.218526
24	2020	3	10	85000		A1	臺北市政府	臺北市士林	臺北市士林	雨	一般車道(未	柏油	濕潤	未注意車前	1	0	小客車	自用	男	61	車的狀態	121.519729	25.091098
25	2020	3	10	221200		A1	臺中市政府	臺中市北屯	臺中市北屯	晴	行人穿越道	柏油	乾燥	搶越行人穿	1	0	小客車	自用	男	29	車的狀態	120.722295	24.175928
26	2020	3	13	155900		A1	臺中市政府	臺中市豐原	臺中市豐原	晴	交叉口附近	柏油	乾燥	搶越行人穿	1	0	小貨車(含客	自用	男	48	車的狀態	120.727966	24.244573
27	2020	3	17	122300		A1	臺中市政府	臺中市后里	臺中市后里	晴	快車道	柏油	乾燥	未注意車前	1	0	小客車	計程車	男	45	車的狀態	120.674552	24.316902
28	2020	3	18	155500		A1	新北市政府	新北市五股	新北市五股	陰	交叉路口內	柏油	乾燥	未注意車前	1	0	小貨車(含客	自用	男	60	車的狀態	121.43861	25.078531
29	2020	3	21	123700		A1	新北市政府	新北市汐止	新北市汐止	晴	交叉路口內	柏油	乾燥	搶越行人穿	1	0	大客車	民營客運	男	53	車的狀態	121.635617	25.056589
30	2020	3	26	54000		A1	臺中市政府	臺中市清水	臺中市清水	晴	交叉口附近	柏油	乾燥	酒醉(後)駕駛	1	1	機車	普通重型	男	18	車的狀態	120.629006	24.292875
31	2020	3	30	120300		A1	臺中市政府	臺中市西屯	臺中市西屯	晴	行人穿越道	柏油	乾燥	搶越行人穿	1	0	小客車	自用	女	60	車的狀態	120.655328	24.158488
32	2020	4	1	160000		A1	臺南市政府	臺南市安平	臺南市安平	晴	交叉路口內	柏油	乾燥	搶越行人穿	1	1	機車	普通重型	男	21	車的狀態	120.169957	23.000949
33	2020	4	1	211700		A1	新北市政府	新北市汐止	新北市汐止	晴	其他	水泥	乾燥	未注意車前	1	1	機車	普通重型	男	28	車的狀態	121.662039	25.069739
34	2020	4	11	43233		A1	臺南市政府	臺南市中西	臺南市中西	晴	交叉口附近	柏油	乾燥	酒醉(後)駕駛	1	1	小客車	自用	女	22	車的狀態	120.195774	22.991675
35	2020	4	11	233200		A1	臺北市政府	臺北市大安	臺北市大安	雨	公車專用道	柏油	濕潤	搶越行人穿	1	0	大客車	民營公車	男	44	車的狀態	121.531626	25.017466
36	2020	4	17	202400		A1	桃園市政府	桃園市大園	桃園市大園	陰	交叉路口內	柏油	乾燥	未注意車前	1	0	小客車	自用	男	26	車的狀態	121.21206	25.031997
37	2020	4	23	140724		A1	桃園市政府	桃園市中壢	桃園市中壢	陰	交叉路口內	柏油	乾燥	未注意車前	1	0	小貨車(含客	自用	女	43	車的狀態	121.224489	24.96504
38	2020	5	2	195100		A1	臺南市政府	臺南市麻豆	臺南市麻豆	晴	一般車道(未	柏油	乾燥	未注意車前	1	0	小客車	自用	男	39	車的狀態	120.213221	23.197424
39	2020	5	3	85620		A1	桃園市政府	桃園市中壢	桃園市中壢	晴	穿越道附近	柏油	乾燥	未注意車前	1	0	小貨車(含客	自用	男	29	車的狀態	121.217684	24.955267
40	2020	5	4	94500		A1	臺北市政府	臺北市中山	臺北市中山	晴	交叉路口內	柏油	乾燥	搶越行人穿	1	1	小客車	自用	女	34	車的狀態	121.525705	25.066719
41	2020	5	11	165400		A1	臺南市政府	臺南市安南	臺南市安南	晴	交叉路口內	柏油	乾燥	未注意車前	1	0	半聯結車	營業用	男	45	車的狀態	120.196619	23.067881

原始檔案為csv檔，合併所有年份事故資料，並篩選六都、人車事故、主要肇因為駕駛人，並刪除用不到的欄位，以利匯入python進一步處理。


```

df['發生時間'] = df['發生時間'].astype(str).str.zfill(6)

df['發生時間'] = df['發生時間'].apply(convert_time_format)

df['time_category'] = df['發生時間'].apply(classify_time)

df.insert(loc = df.columns.get_loc('發生時間') + 1, column='時間分類', value=df['time_category'])

df = df.drop('time_category', axis = 1)

df

```

Python

	發生年度	發生月份	發生日期	發生時間	時間分類	經度	緯度
0	2020	1	20200103	02:38:00	半夜至清晨	121.505006	24.993558
1	2020	1	20200103	08:56:00	早上	121.615942	25.044450
2	2020	1	20200106	11:25:00	早上	121.509371	25.045220
3	2020	1	20200106	19:18:00	傍晚至晚上	121.540362	24.964583
4	2020	1	20200108	07:30:00	早上	120.682900	24.140782
...
22615	2022	12	20221231	21:50:00	傍晚至晚上	121.279148	24.960691
22616	2022	12	20221231	21:55:00	傍晚至晚上	121.488443	25.059041
22617	2022	12	20221231	22:00:00	傍晚至晚上	121.542235	25.044046
22618	2022	12	20221231	23:15:00	傍晚至晚上	121.552433	25.033015
22619	2022	12	20221231	23:19:00	傍晚至晚上	121.444859	25.043724

22620 rows x 7 columns

```

import pandas as pd
import matplotlib.pyplot as plt
from datetime import datetime

#時間格式轉換
def convert_time_format(time_str):
    time_obj = datetime.strptime(time_str, '%H%M%S')
    return time_obj.strftime('%H:%M:%S')

#時段的轉換
def classify_time(time_str):
    time_obj = datetime.strptime(time_str, '%H:%M:%S')
    hour = time_obj.hour

    if hour >= 0 and hour < 6:
        return '半夜至清晨'
    elif hour >= 6 and hour < 12:
        return '早上'
    elif hour >= 12 and hour < 18:
        return '下午'
    else:
        return '傍晚至晚上'

#讀取資料

df = pd.read_excel('109-111年 A1&A2 人車事故合併.xlsx',
                  parse_dates = ['發生時間'],
                  sheet_name = '人車(六都、駕駛人、順位1)',
                  usecols = 'A, B, C, D, AX, AY',
                  )

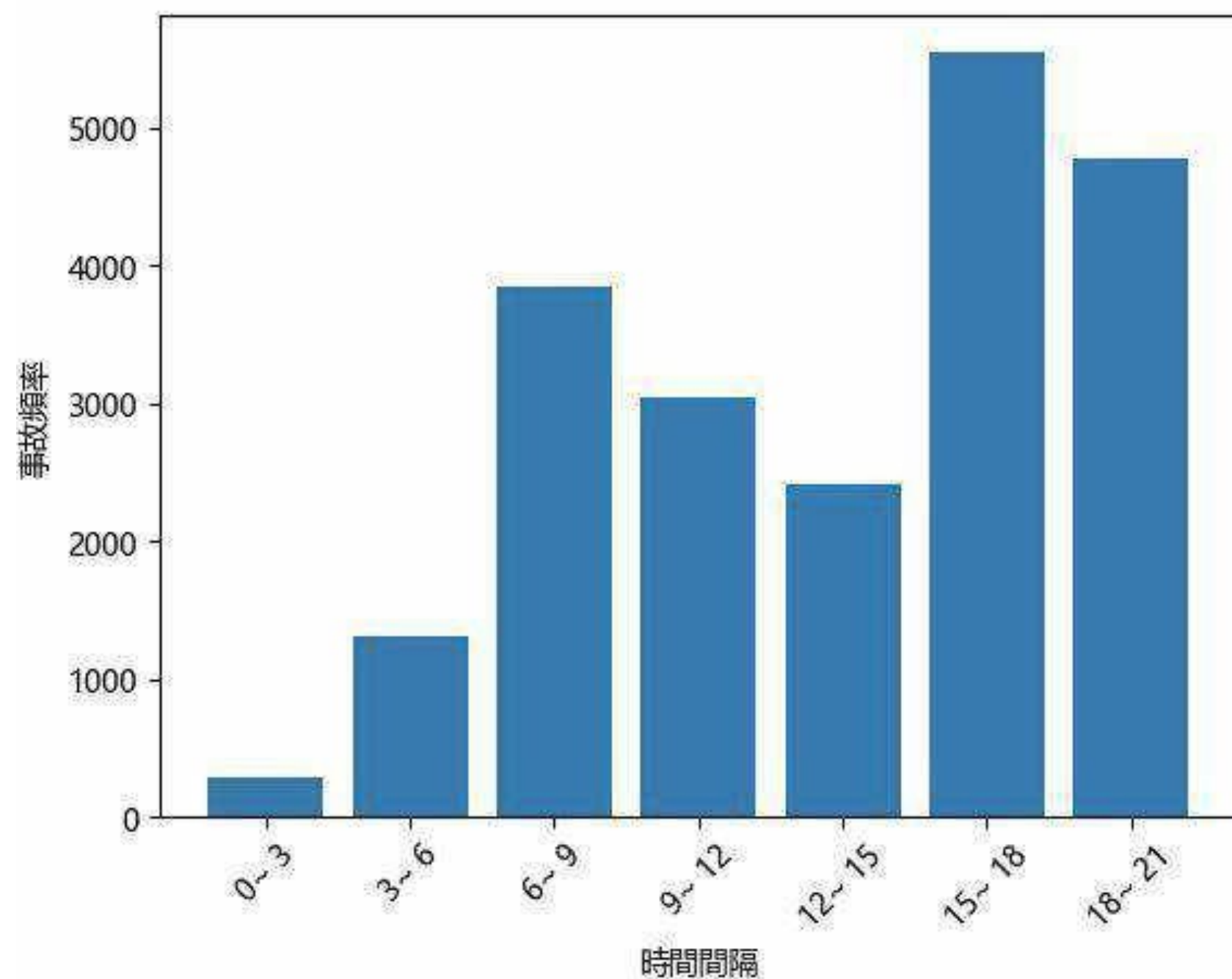
```

05. 事故總覽

擷取109-111年資料，篩選台灣六都之人車事故且主要肇事者為駕駛人之數據

- 近三年人車事故數量：22620 筆
- 台北市：3761 筆、新北市：5226 筆、桃園市：4781 筆、台中市：3642 筆、台南市：2365 筆、高雄市：2845 筆
- 共造成 22407 人受傷， 213 人死亡

06. 分析一：事故發生時間段



```
import pandas as pd
import matplotlib.pyplot as plt

# 設定中文字型為 Microsoft YaHei
plt.rcParams['font.family'] = 'Microsoft YaHei'

# 將發生時間轉換為 pandas 的時間戳格式
data['發生時間'] = pd.to_datetime(data['發生時間'])

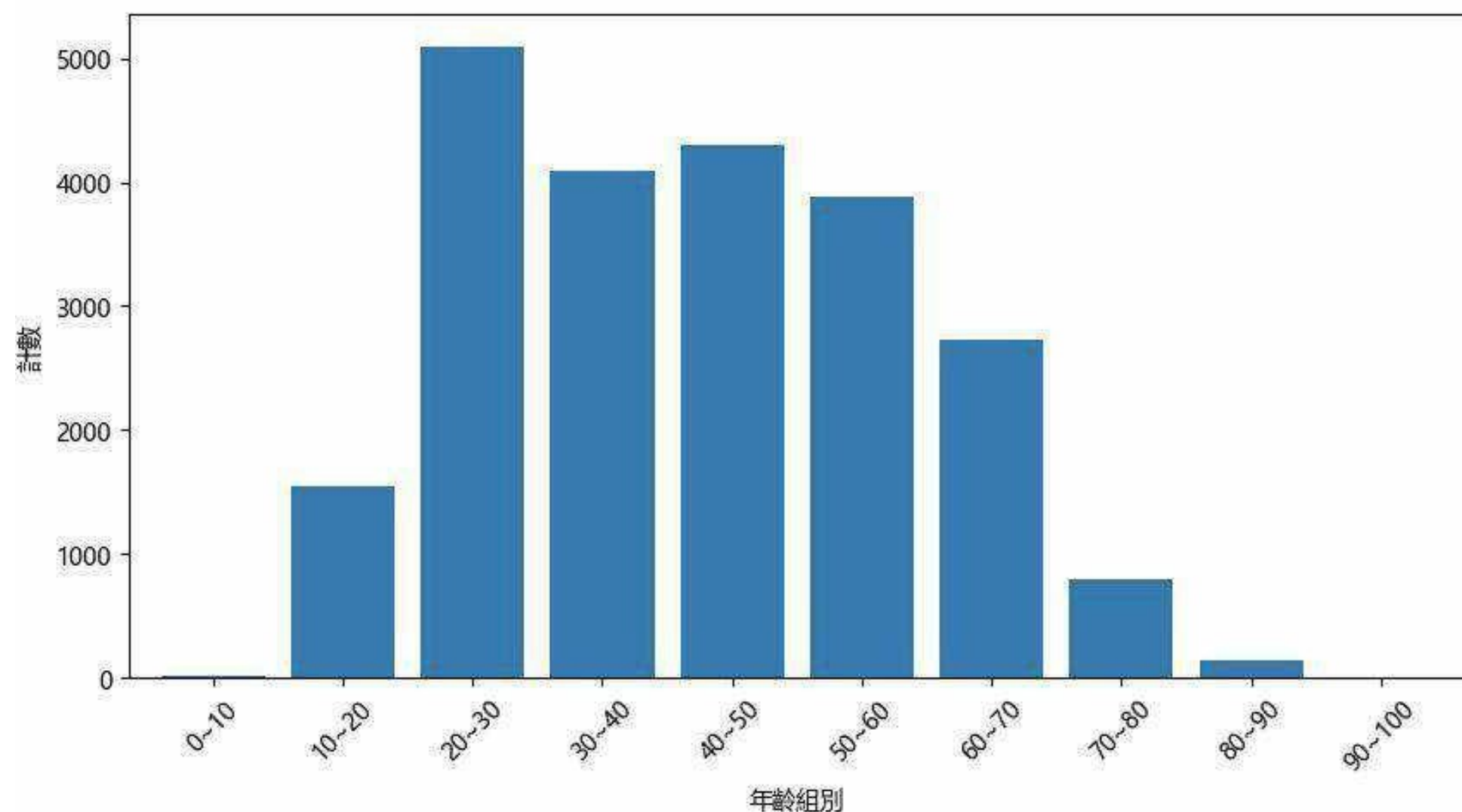
# 根據三小時間隔進行時間分組
data['時間間隔'] = pd.cut(data['發生時間'].dt.hour, bins=range(0, 24, 3))

# 統計各時間間隔的事故頻率
accident_freq = data['時間間隔'].value_counts().sort_index()

# 繪製長條圖
plt.bar(accident_freq.index.astype(str), accident_freq.values)
plt.xlabel('時間間隔')
plt.ylabel('事故頻率')
plt.title('每三小時事故頻率')
plt.xticks(rotation=45)
plt.show()
```

應加強下午三點過後、
下班時間之交通巡邏！

06. 分析二：年齡



```
import matplotlib.pyplot as plt
import pandas as pd

data["AgeGroup"] = pd.cut(data["當事者事故發生時年齡"], bins=range(0, 101, 10), right=False)
age_counts = data["AgeGroup"].value_counts().sort_index()
plt.figure(figsize=(10, 5))
age_counts.plot(kind="bar", width=0.8)
plt.xlabel("年齡組別")
plt.ylabel("計數")
plt.title("當事者事故發生時年齡分佈")
plt.show()
```

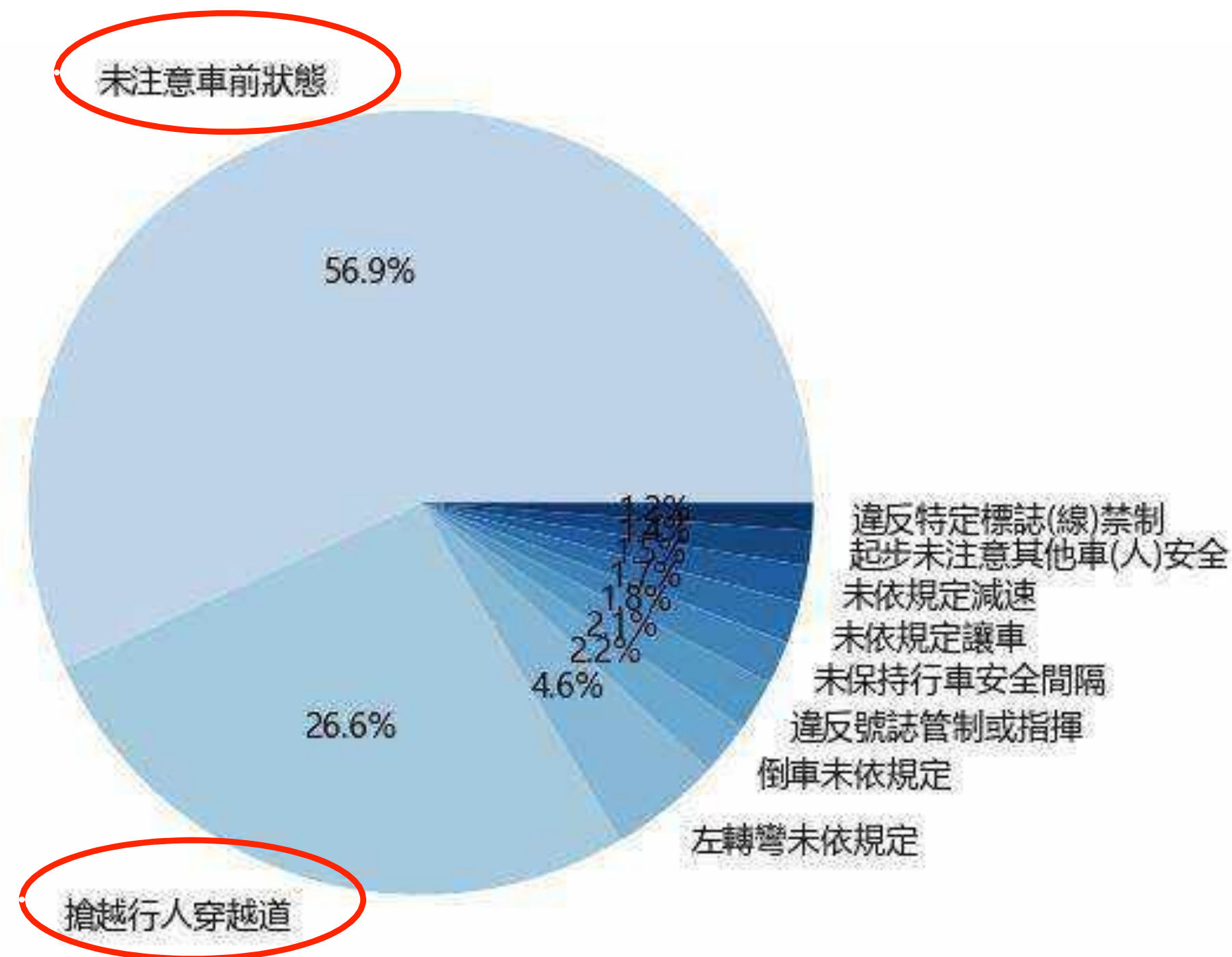
**20 到 30 歲族群
用路禮儀、觀念最需加強！**

06. 分析三：肇因

```
import matplotlib.pyplot as plt

plt.rcParams['font.family'] = 'Microsoft YaHei'
grouped = data.groupby('肇因研判子類別名稱-主要').size()
grouped = grouped.sort_values(ascending=False).head(10)
grouped.plot(kind='pie', autopct='%0.1f%%')
plt.title('肇因研判')
plt.show()
```

未注意車前狀況、搶越行人穿越道
佔事故肇因總數八成以上



06. 分析三：肇因—未注意車前狀況

```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_excel('109-111_new.xlsx', usecols = 'I, N')

df_1 = df[df['肇因研判子類別名稱-主要'] == '未注意車前狀態']
df_2 = df[df['肇因研判子類別名稱-主要'] == '搶越行人穿越道']

df_1['行政區'].value_counts()
# df_2['行政區'].value_counts()
```

✓ 13.5s

桃園市中壢區	642
桃園市桃園區	570
新北市三重區	393
新北市板橋區	375
桃園市八德區	336

...

駕駛最不專心的行政區：

桃園市中壢區，近三年共發生642件人車交通事故，
發生原因為駕駛 **未注意車前狀態**。

06. 分析三：肇因—搶越行人穿越道

```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_excel('109-111_new.xlsx', usecols = 'I, N')

df_1 = df[df['肇因研判子類別名稱-主要'] == '未注意車前狀態']
df_2 = df[df['肇因研判子類別名稱-主要'] == '搶越行人穿越道']

# df_1['行政區'].value_counts()
df_2['行政區'].value_counts()
```

✓ 24.1s

臺北市中山區	306
臺北市大安區	229
臺北市中正區	208
臺中市北屯區	202
臺北市內湖區	201

...

過馬路最危險的行政區：

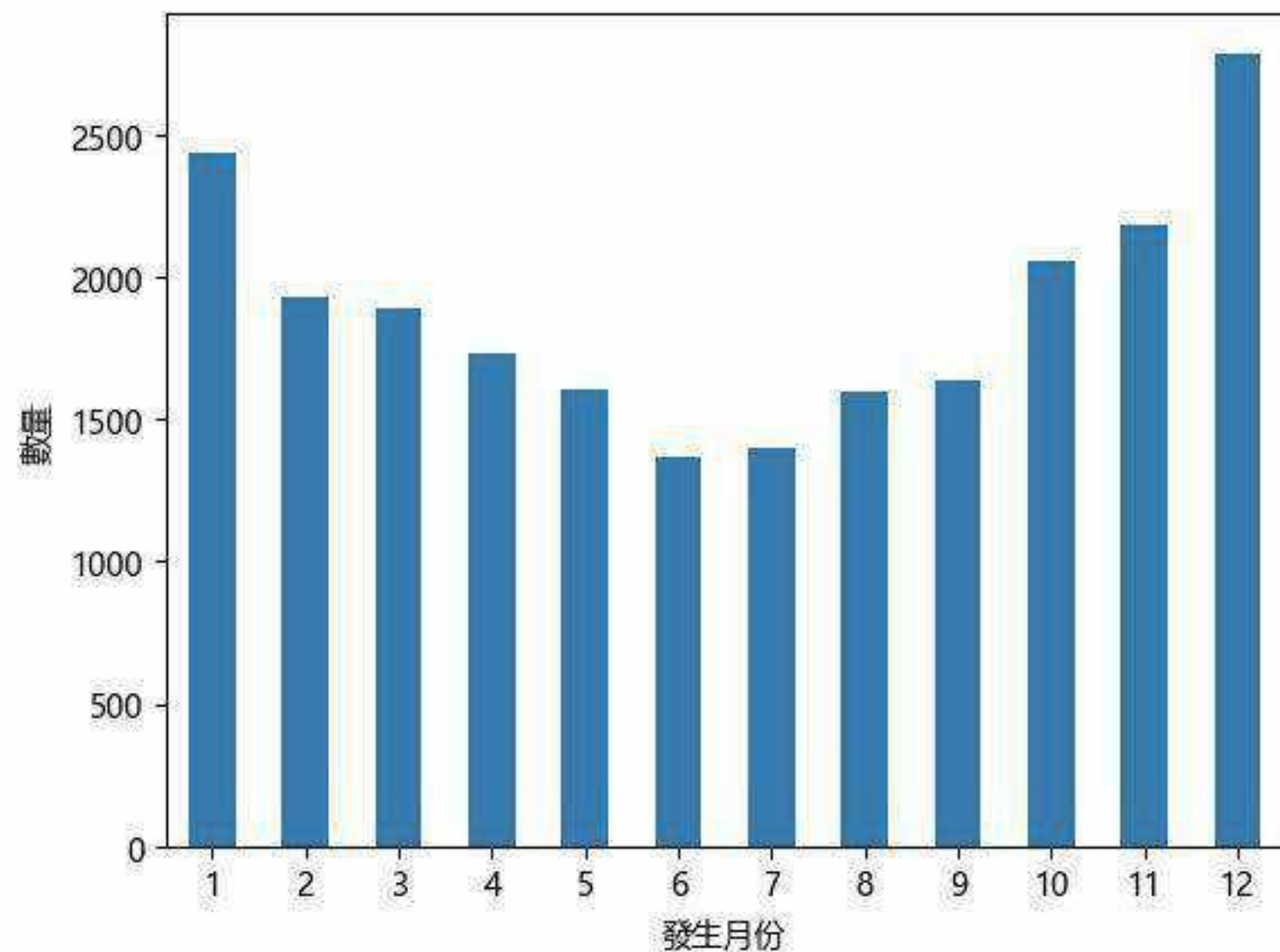
台北市中山區，近三年共發生306 件人車交通事故，
發生原因為駕駛 **搶越行人穿越道**。

06. 分析四：月份

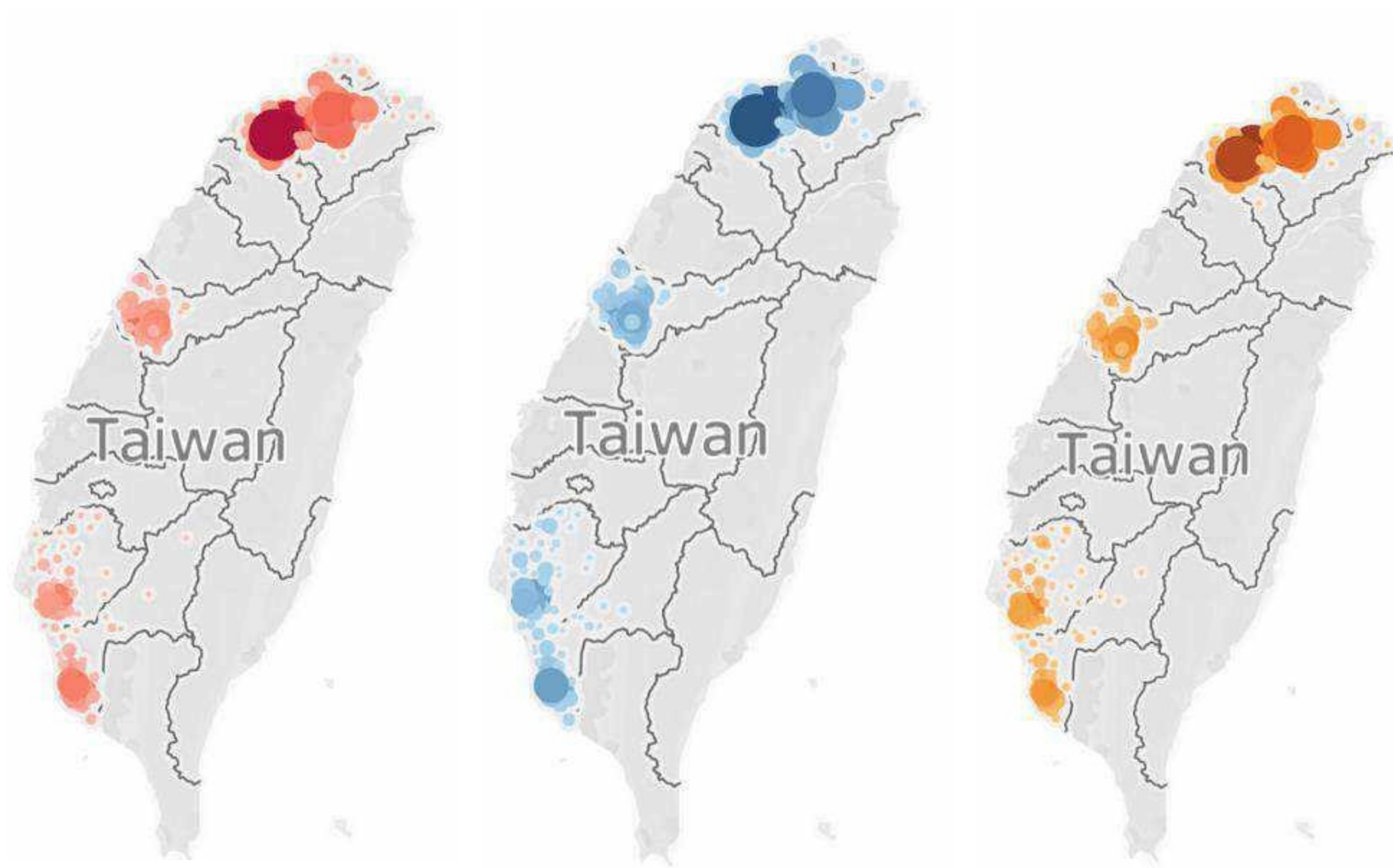
```
import matplotlib.pyplot as plt

grouped = data.groupby('發生月份').size()
grouped.plot(kind='bar')
plt.xlabel('發生月份')
plt.ylabel('數量')
plt.xticks(rotation=0)
plt.show()
```

第四季及一月份發生交通事故之比例較高，
出門應小心！



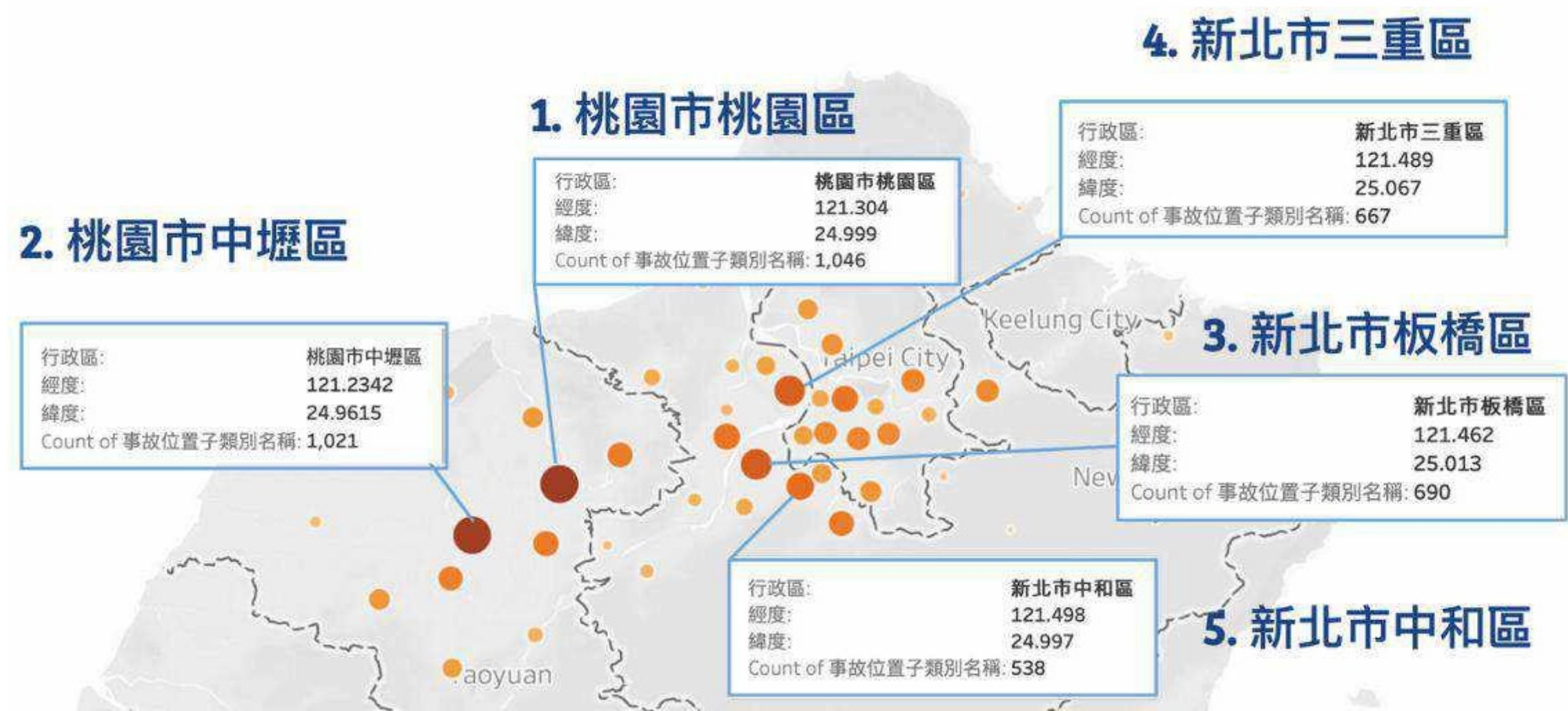
06. 分析五：六都行政區人車事故分佈（109 - 111年）



- 交通事故分佈集中在**人口密集地**
- 三年的**事故集中地**並無大變化

2020
2021
2022

06. 分析五：六都行政區人車事故排行榜（109 - 111年）

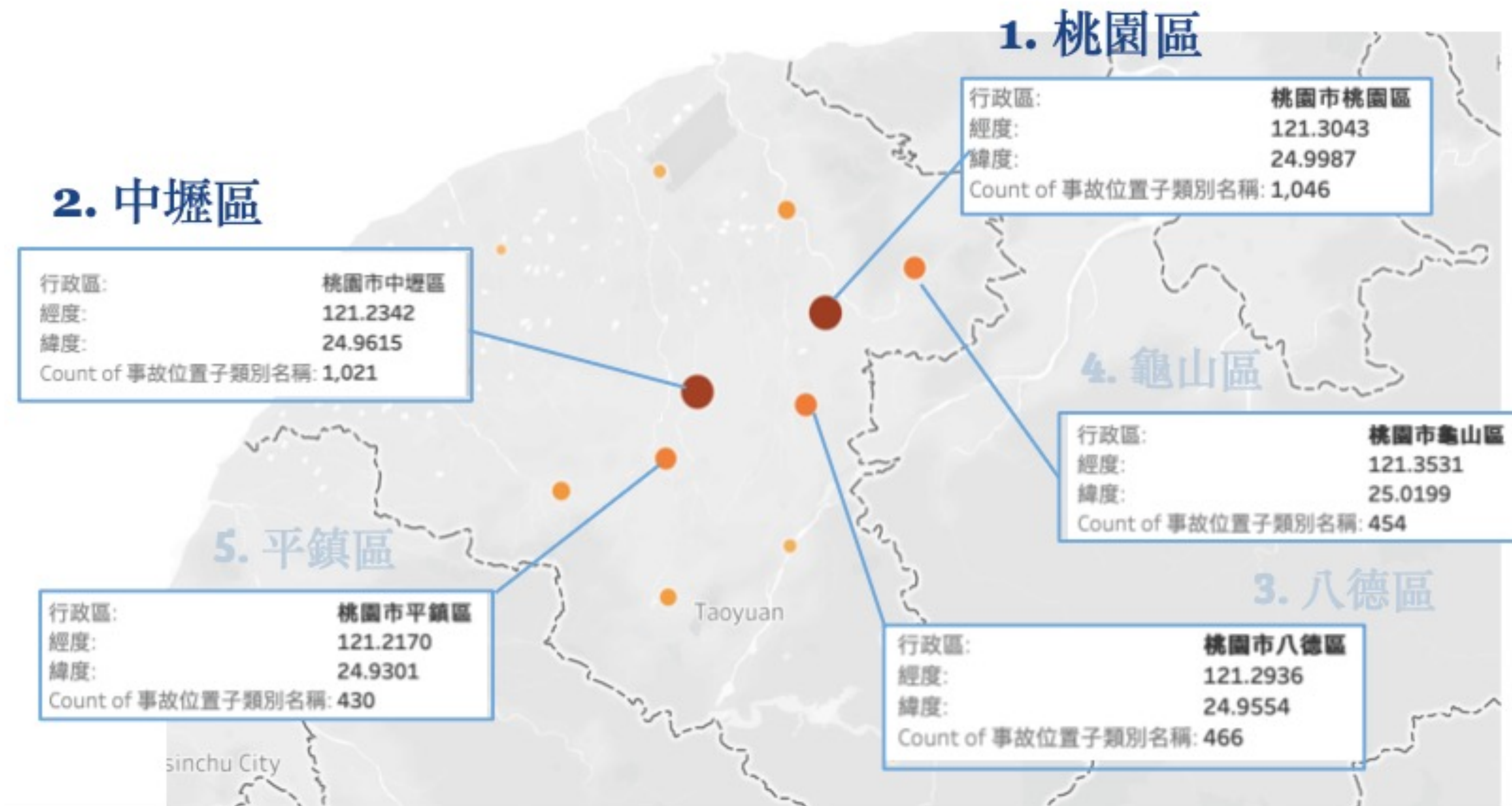


- 第一名：桃園市桃園區 1046 件
- 第二名：桃園市中壢區 1021 件
- 第三名：新北市板橋區 690 件
- 第四名：新北市三重區 667 件
- 第五名：新北市中和區 538 件

桃園的交通，亟需改善！

06. 分析五 — 縣市

桃園市行政區地獄排行



06. 分析五 — 縣市

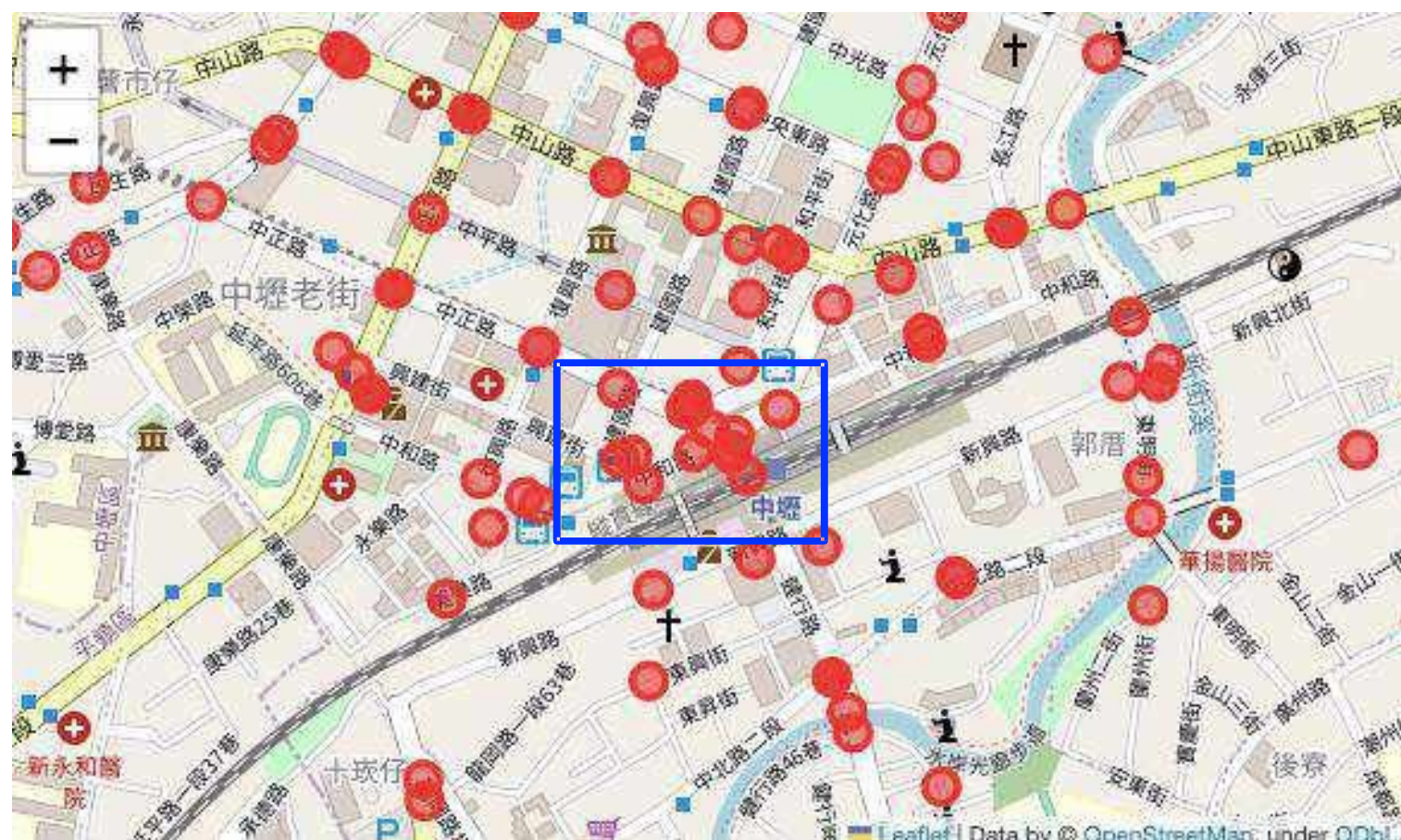
桃園市桃園區

中山路與三民路三段交界處與火車站附近皆為人車事故好發地區，行人通過以上路段須格外小心！



06. 分析五 — 縣市

桃園市中壢區



中壢火車站與中原國小周圍
為中壢區交通事故好發的兩大區域，
中壢區的朋友們通過要小心！

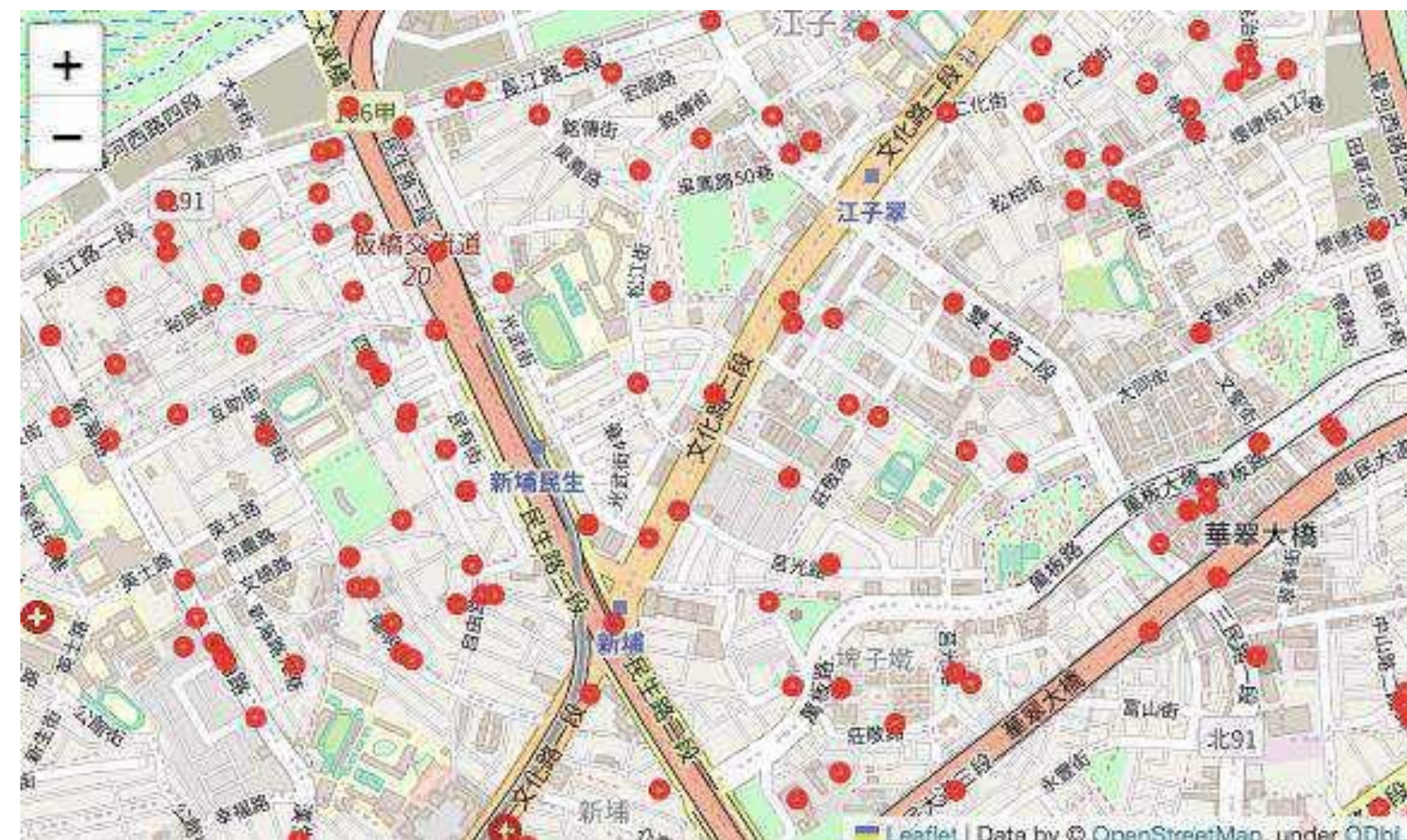
06. 分析五 — 縣市

新北市行政區地獄排行



06. 分析五 — 縣市

新北市板橋區



作為新北市人口數最多的行政區，板橋區行人交通事故分佈較分散，府中站周圍事故稍多。
以往交通較擁塞**文化民生路口**及**文化民權路口**未發現有較多人車事故。

06. 分析五 — 縣市

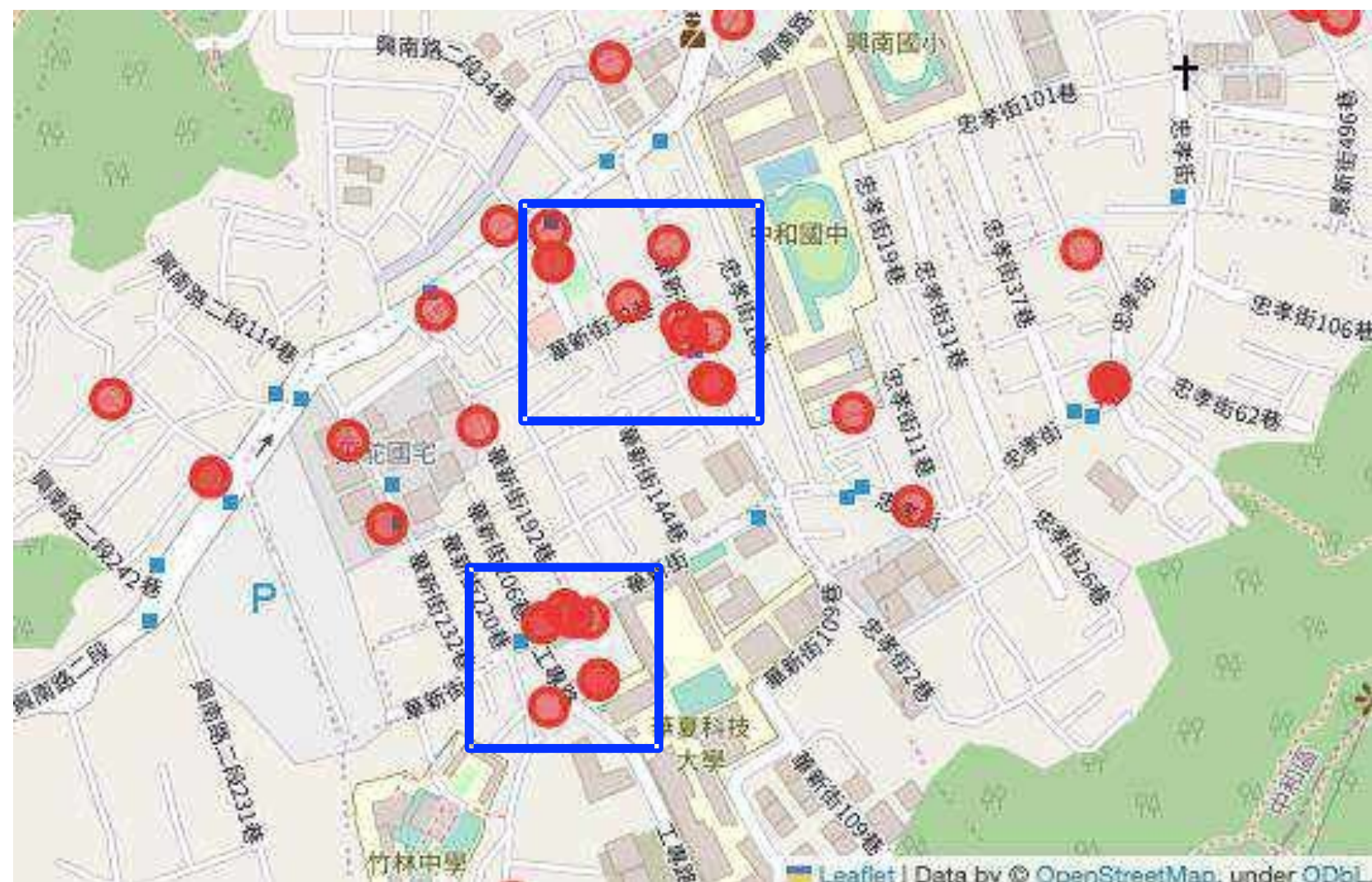
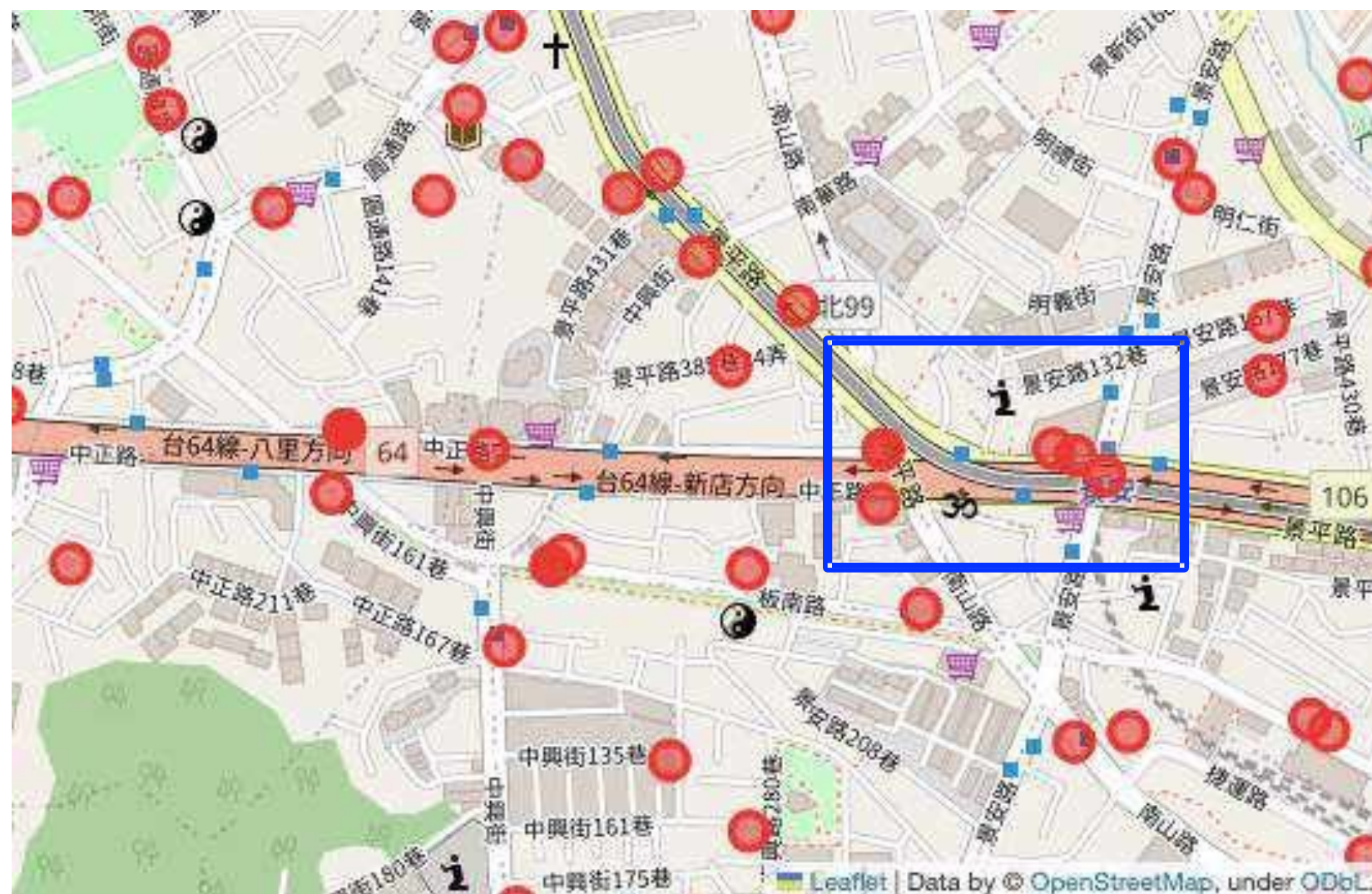
新北市三重區



**文化北路台北橋下至信義公園段，為人車事故好發地點；
仁愛街、五華街、仁義街一帶，為三重之行人天堂。**

06. 分析五 — 縣市

新北市中和區



中和區主要幹道**中正板南路口**、**中正景安路口**，事故未想像中多。
但本次分析意外發現，**華新路上**學校旁的人車事故有兩處較密集區域，原因值得深入探討。

06. 分析五 — 縣市

生活在台北市的我們...

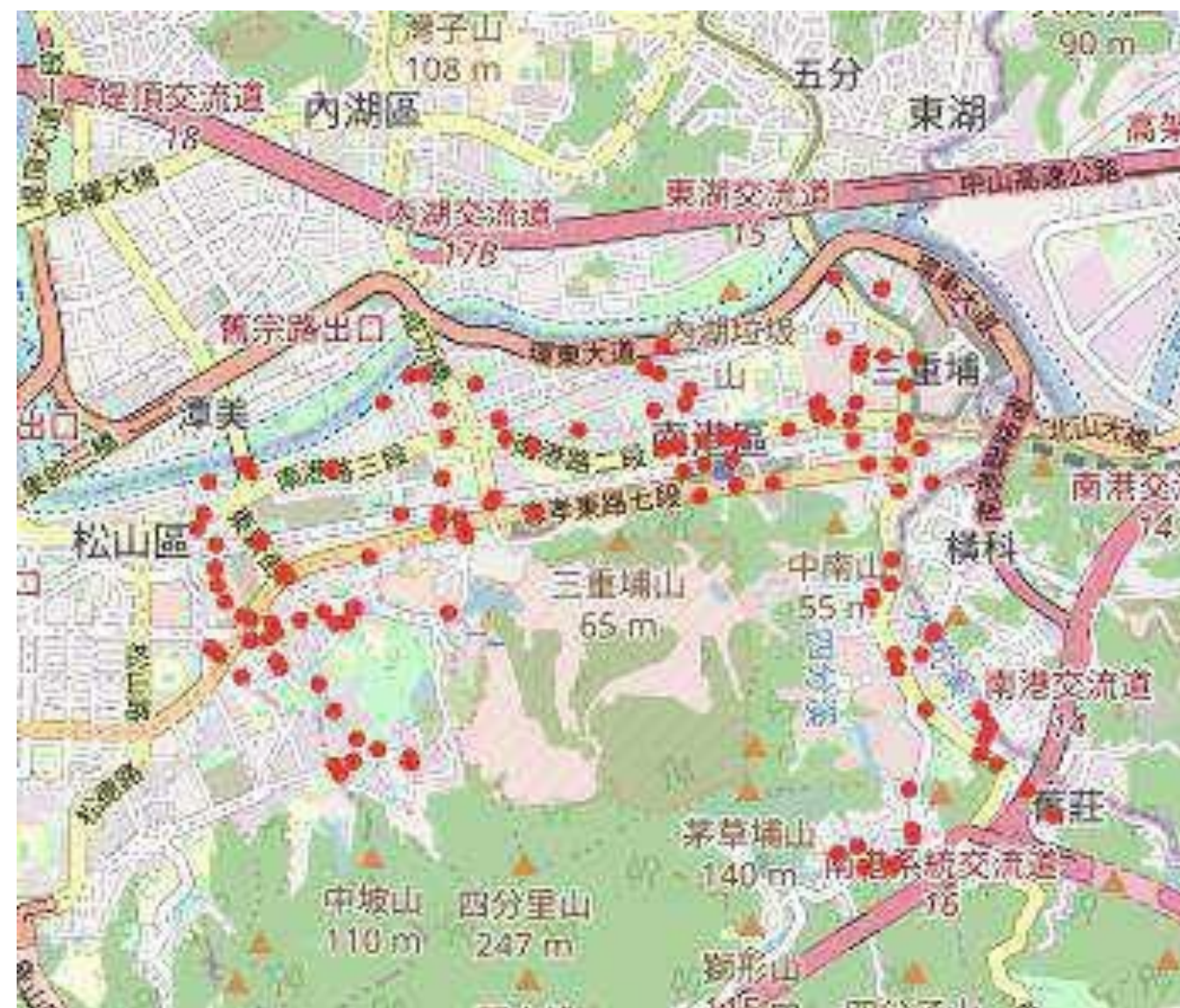


06. 分析五 — 縣市

生活在台北市的我們...



台北市中山區近三年共發生人車事故500件，
是台北市行政區行人地獄冠軍



台北市南港區近三年共發生人車事故153件，
是台北市行政區行人天堂冠軍

07.未來展望

- 分析範圍不只侷限六都，而是可以擴展到全台灣。
- 對於造成各行政區域可做更嚴謹之統計分析，來了解造成該區域好發人車事故的原因。
- 交通事故法規頻頻更新，科技執法上路，可針對科技執法的路段分析前後交通事故是否有顯著下降。
- 根據本次的分析結合 linebot，使用者可輸入前往路段，linebot 機器人可通知該路段 or 區域為交通安全 / 危險區域。
- 本次分析僅抓取人車事故資料，只是全部資料的冰山一角，還並未考慮車對車，看來在地狹人稠的台灣，交通的改善還有很長一段的路要走。

08. 參考資料

- Learning material: <https://medium.com/seaniap/pandas基礎介紹-進入資料科學的領域-be9894b3548>
- Data resource: <https://data.gov.tw/dataset/12197>
- 歷屆專案參考:
 - eBird鳥類資料庫:<https://www.youtube.com/watch?v=lq6qnUquzbA>
 - 稻作產量地圖與視覺化:<https://www.youtube.com/watch?v=D19nZTUQnxs>
 - 過敏兒小幫手:<https://www.youtube.com/watch?v=k-nCK7hNVRs>
 - 診所孵化器:<https://www.youtube.com/watch?v=zbfpI3ADgFk>