

Colab Link: <https://colab.research.google.com/drive/1t5ZnRBDH6e916Ge8KPRfQoHctv46FeQo?usp=sharing>

## 桃園市交通改善計畫\_109&110年事故發生位置與各年齡層酒駕之分析

(import套件 & read\_table)

```
1 !wget -O taipei_sans_tc_beta.ttf https://drive.google.com/uc?id=1eGAsTN1HBPJAkeVM57_C7ccp7hbgSz3_&export=download
2
3 import matplotlib
4 # 新增字體
5 matplotlib.font_manager.FontManager.addfont('taipei_sans_tc_beta.ttf')
6
7 # 將 font-family 設為 Taipei Sans TC Beta
8 # 設定完後，之後的圖表都可以顯示中文了
9 matplotlib.rcParams['font.family'] = 'Taipei Sans TC Beta'
10
11 from datascience import *
12 import pandas as pd
13 import numpy as np
14 import matplotlib
15 matplotlib.use('Agg')
16 %matplotlib inline
17 import matplotlib.pyplot as plt
18 plt.style.use('fivethirtyeight')
19
20 print([f for f in matplotlib.font_manager.FontManager.ttflist if 'Heiti' in f.name])
21 # matplotlib.rcParams['font.family'] = ['Heiti TC']
22
```

```
--2022-12-30 06:58:11-- https://drive.google.com/uc?id=1eGAsTN1HBPJAkeVM57_C7ccp7hbgSz3_
Resolving drive.google.com (drive.google.com)... 108.177.125.102, 108.177.125.100, 108.177.125.139, ...
Connecting to drive.google.com (drive.google.com)|108.177.125.102|:443... connected.
HTTP request sent, awaiting response... 303 See Other
Location: https://doc-0k-9o-docs.googleusercontent.com/docs/securesc/ha0ro937gcuc7l7deffksulhg5h7mbp1/7d01jp8t54tfh4n755op3b5vq2m3ro61/1672393
Warning: wildcards not supported in HTTP.
--2022-12-30 06:58:12-- https://doc-0k-9o-docs.googleusercontent.com/docs/securesc/ha0ro937gcuc7l7deffksulhg5h7mbp1/7d01jp8t54tfh4n755op3b5vq
Resolving doc-0k-9o-docs.googleusercontent.com (doc-0k-9o-docs.googleusercontent.com)... 64.233.188.132, 2404:6800:4008:c06::84
Connecting to doc-0k-9o-docs.googleusercontent.com (doc-0k-9o-docs.googleusercontent.com)|64.233.188.132|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 20659344 (20M) [application/x-font-ttf]
Saving to: 'taipei_sans_tc_beta.ttf'
```

```
taipei_sans_tc_beta 100%[=====>] 19.70M --.-KB/s in 0.09s
```

```
2022-12-30 06:58:13 (228 MB/s) - 'taipei_sans_tc_beta.ttf' saved [20659344/20659344]
```

```
[]
```

109年桃園市交通事故資料表-含第二當事人 資料表

table_109 = Table().read_table('https://data.tycg.gov.tw/opendata/datalist/datasetMeta/download?id=77c34235-6fd9-4c1f-8ebd-1f675cdc7c02&rid=3fa9285a-b5f5-4f20-91e2-b628fcc489ac')																																	
table_109																																	
發生日期	發生時間	發生年度	發生月份	發生星期	到場處理日期	到場處理時間	結束_事故排除日期	結束_事故排除時間	GPS經度	GPS緯度	事故類別名稱	地址類型名稱	發生縣市名稱	發生市區鄉鎮名稱	發生地址_村里代碼	發生地址_村里名稱	發生地址_鄰	發生地址_路街代碼	發生地址_路街	發生地址_段	發生地址_巷	發生地址_弄	發生地址_號	發生地址_前幾公尺	發生地址_側名稱	發生交叉路口_村里名稱	發生交叉路口_路街口	發生交叉路口_路街口	發生交叉路口_巷	發生交叉路口_弄	發生地址_其他	死亡人數_24小時內	死亡人數_2_30日內
20200101	56	2020	1	三	20200101	815	20200101	2315	121.208	25.0058	A2	其他	桃園市	中壢區	nan	nan	nan	nan	nan	nan	nan	0		nan	nan	nan	0	0					
高鐵南路與公園路路																																	

110年桃園市交通事故資料表-含第二當事人 資料表

table_110 = Table().read_table('https://data.tycg.gov.tw/opendata/datalist/datasetMeta/download?id=ae5d2d9b-59ad-4f15-a9d2-d0f5f3de2475&rid=8b93ee39-5f31-4807-bdfb-4da6bb71db4b')																																
table_110																																
發生日期	發生時間	發生年度	發生月份	發生星期	到場處理日期	到場處理時間	結束_事故排除日期	結束_事故排除時間	GPS經度	GPS緯度	事故類別名稱	地址類型名稱	發生縣市名稱	發生市區鄉鎮名稱	發生地址_村里代碼	發生地址_村里名稱	發生地址_鄰	發生地址_路街代碼	發生地址_路街	發生地址_段	發生地址_巷	發生地址_弄	發生地址_號	發生地址_前幾公尺	發生地址_側名稱	發生交叉路口_村里名稱	發生交叉路口_路街口	發生交叉路口_路街口	發生交叉路口_巷	發生交叉路口_其他	死亡人數_24小時內	死亡人數_2_30日內
20210101	1600	2021	1	五	20210101	2600	20210101	4100	121.248	24.9571	A2	交叉路口	桃園市	中壢區	nan	nan	nan	6.8e+15	環中東路	nan	nan	nan	0	口	新中北路	nan	nan	0	0			

# PartI 桃園市事故位置類別分析

## 桃園市事故位置類別分析 EDA

1. 將每個區域的事故位置類型與數量做統計與製圖  
(擷取整個圖表 各區域事故位置類型名稱與數量 並分區域統計)
2. 由直方圖可說明，各區應加強佔比較多之事故位置的巡邏和相關交通安全宣導  
(e.g. 109年及110年，中壢區前三大事故位置類別皆為 <1>交叉路口內 <2>一般車道(未劃分快慢車道) <3>交叉口附近，  
因此中壢區警察局應特別加強這三類位置之巡邏和相關交通安全宣導)

```
# 109年 各區域+事故位置類別 (之後統計各區各事故位置數量 要用)
location_109 = table_109.select(['發生市區鄉鎮名稱', '事故位置子類別名稱'])
location_109
```

發生市區鄉鎮名稱	事故位置子類別名稱
中壢區	交叉路口內
中壢區	交叉路口內
中壢區	交叉路口內
平鎮區	交叉路口內
平鎮區	交叉路口內
八德區	交叉路口內
八德區	交叉路口內
平鎮區	路肩、路緣
平鎮區	路肩、路緣
平鎮區	路肩、路緣

... (109586 rows omitted)

```
# 109年 各區域事故數量
location_109_total = table_109.group('發生市區鄉鎮名稱')
location_109_total
```

發生市區鄉鎮名稱	count
中壢區	25833
八德區	10314
大園區	3994
大溪區	4357
平鎮區	10526
復興區	283
新屋區	2018
桃園區	21271
楊梅區	6378
蘆竹區	7643

... (3 rows omitted)

```
# 110年 各區域+事故位置類別 (之後統計各區各事故位置數量 要用)
location_110 = table_110.select(['發生市區鄉鎮名稱','事故位置子類別名稱'])
location_110
```

發生市區鄉鎮名稱	事故位置子類別名稱
中壢區	交叉路口內
中壢區	交叉路口內
中壢區	交叉路口內
中壢區	交叉路口內
中壢區	交叉路口內
中壢區	交叉路口內
中壢區	交叉路口內
中壢區	交叉路口內
桃園區	交叉路口內
桃園區	交叉路口內

... (92240 rows omitted)

```
# 110年 各區域事故數量
location_110_total = table_110.group('發生市區鄉鎮名稱')
location_110_total
```

發生市區鄉鎮名稱	count
中壢區	20546
八德區	9257
大園區	3600
大溪區	3740
平鎮區	8604
復興區	227
新屋區	1621
桃園區	17344
楊梅區	5419
蘆竹區	6850

... (3 rows omitted)

```
# 109年 <各區域><事故位置類別> 數量統計表
loc_109 = [0 for i in range(13)]
total_109 = [0 for i in range(13)]
a_109 = [0 for i in range(13)]
a2_109 = [0 for i in range(13)]
pd_109 = [0 for i in range(13)]

# 各區域的各事故位置 的數量 之table
for i in range(13):
    loc_109[i] = location_109_total.column(0).item(i) # 各區域名稱
    total_109[i] = location_109.where('發生市區鄉鎮名稱',are.equal_to(loc_109[i])).group('事故位置子類別名稱') # 各區域 分別統計 各事故位置 數量
    pd_109[i] = ((total_109[i].to_df()).rename(columns=('count':loc_109[i]))).set_index('事故位置子類別名稱') # 轉換為dataframe 且 修正欄位名稱(count->各區域名稱)

# 將所有區域table 合併成一個table
for i in range(12):
    loc_109_total = pd_109[i].join(pd_109[i+1], how='outer')
    pd_109[i+1] = loc_109_total

loc_109_total = loc_109_total.astype('Int64') # float轉int
loc_109_total.fillna(0, inplace=True) # 將缺失值填入0
loc_109_total
```

	中壢區	八德區	大園區	大溪區	平鎮區	復興區	新屋區	桃園區	楊梅區	蘆竹區	觀音區	龍潭區	龜山區
事故位置子類別名稱													
一般車道(未劃分快慢車道)	8535	3863	1555	1830	3171	230	881	7091	2701	2870	930	1955	3728
交叉口附近	4021	2094	528	722	2298	17	207	2648	1129	751	239	589	1490
交叉路口內	11410	3611	1385	1273	3739	11	655	9733	1899	3137	933	2019	2881
交通島(含精化線)	75	9	11	15	49	0	2	26	4	10	2	10	28
人行道	80	10	5	14	9	0	2	54	4	9	2	6	14
公車專用道	0	0	0	0	0	0	0	7	0	0	0	0	2
其他	344	72	53	55	129	8	34	237	194	157	25	24	44
加速車道	5	2	0	2	2	0	0	5	0	0	0	0	0
快車道	218	120	88	69	215	2	61	239	43	113	65	85	154
慢車道	449	279	139	95	272	0	66	494	159	198	141	136	658
機車停等區	77	41	19	24	30	0	3	82	26	26	4	17	22
機車優先道	24	39	50	10	133	0	20	62	25	23	28	11	89
機車專用道	65	2	16	8	95	0	5	113	21	36	2	8	22
機車待轉區	37	10	0	4	14	0	0	20	5	10	0	6	17
減速車道	0	0	0	4	2	0	0	0	0	0	0	0	0
環道匝道	2	0	0	2	0	0	0	2	0	0	0	0	2
直線匝道	4	2	2	0	29	0	5	2	3	3	0	0	0
穿越道附近	14	6	2	0	0	0	5	8	0	4	2	6	4
行人穿越道	131	23	14	17	41	0	2	127	13	31	8	22	34
路肩、路緣	340	131	127	210	282	15	68	315	148	258	117	194	202
迴轉道	2	0	0	3	16	0	2	6	4	7	0	0	2

```
# 110年 <各區域><事故位置類別> 數量統計表
loc = [0 for i in range(13)]
total = [0 for i in range(13)]
a = [0 for i in range(13)]
a2 = [0 for i in range(13)]
pd = [0 for i in range(13)]

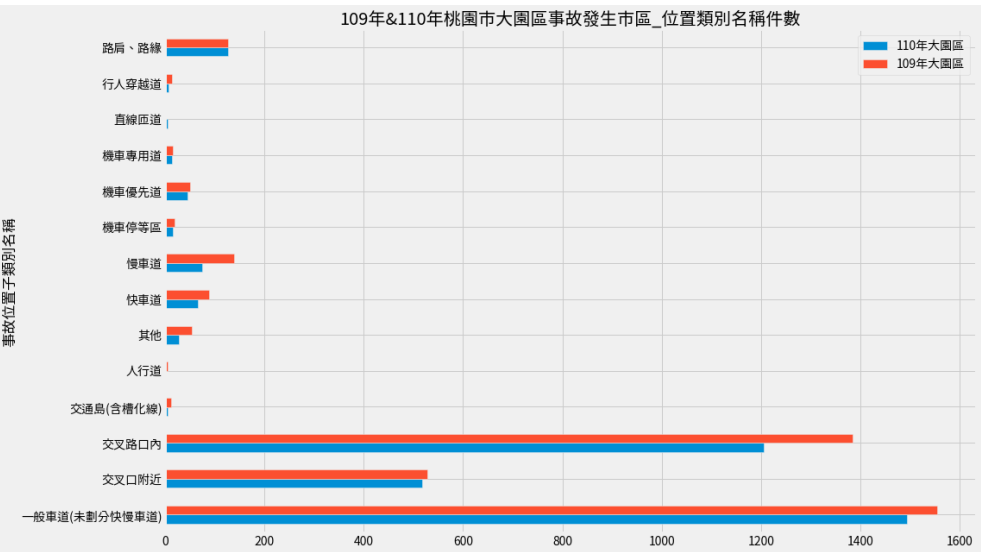
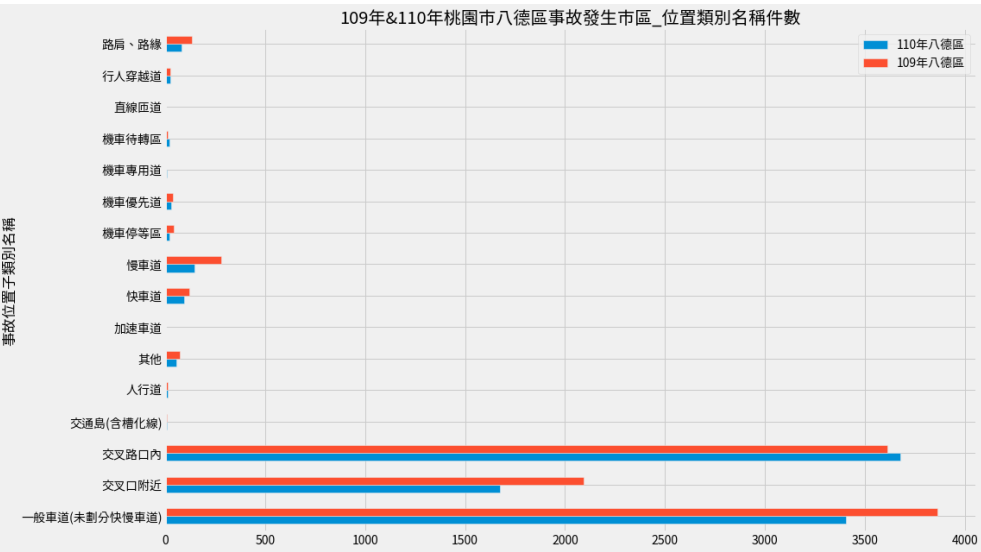
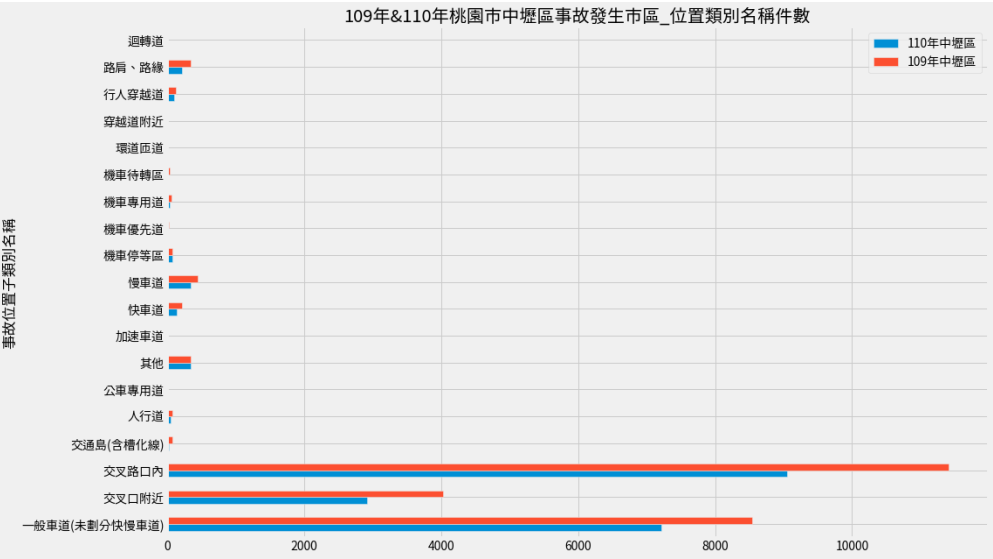
# 各區域的各事故位置 的數量 之table
for i in range(13):
    loc[i] = location_110_total.column(0).item(i) # 各區域名稱
    total[i] = location_110.where('發生市區鄉鎮名稱', are.equal_to(loc[i])).group('事故位置子類別名稱') # 各區域 分別統計 各事故位置 數量
    pd[i] = ((total[i].to_df()).rename(columns={'count': loc[i]})).set_index('事故位置子類別名稱') # 轉換為dataframe 且 修正欄位名稱(count->各區域名稱)

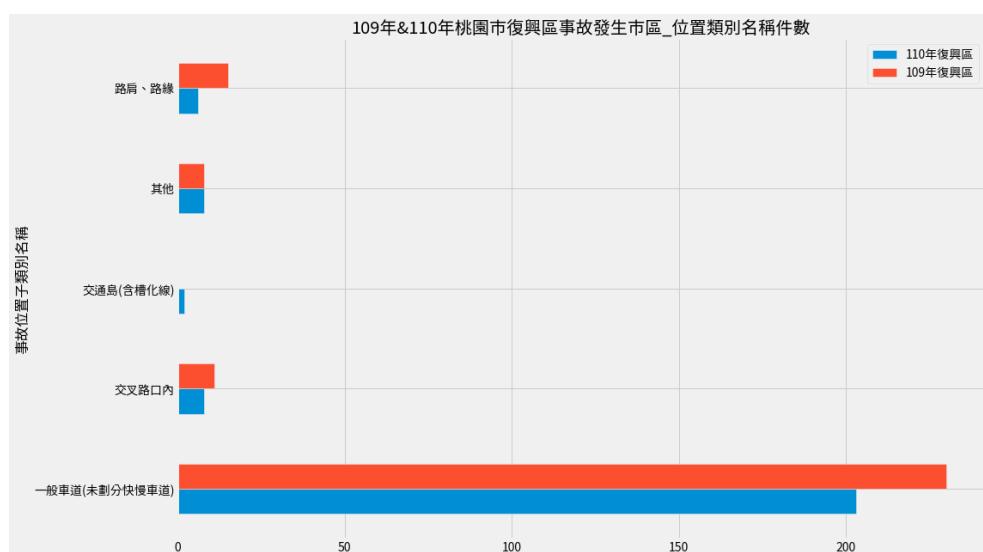
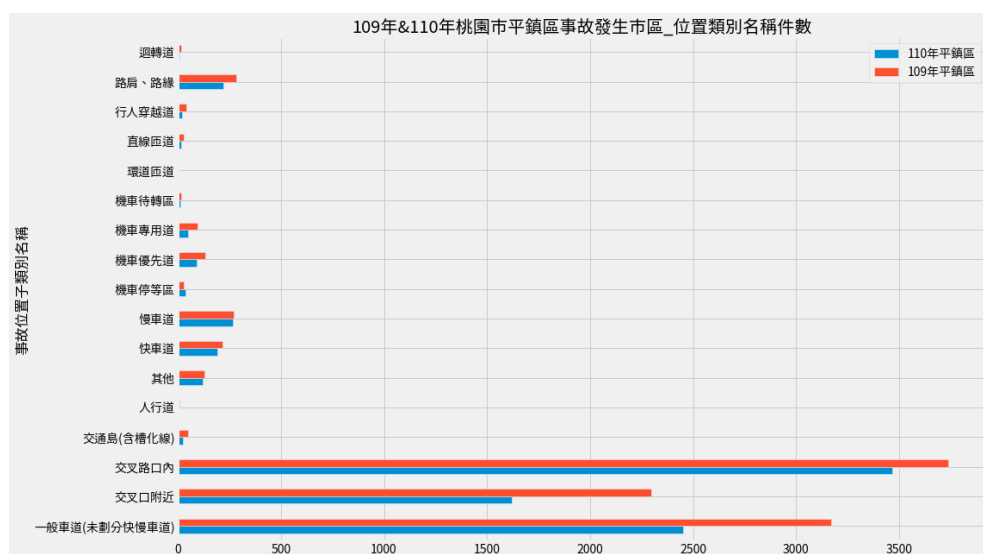
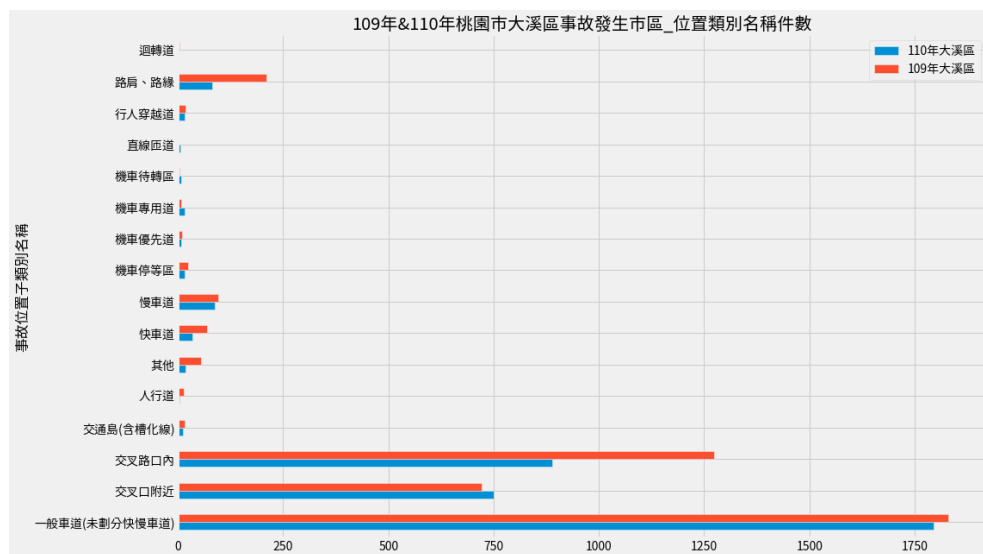
# 將所有區域table 合併成一個table
for i in range(12):
    loc_110_total = pd[i].join(pd[i+1], how='outer')
    pd[i+1] = loc_110_total

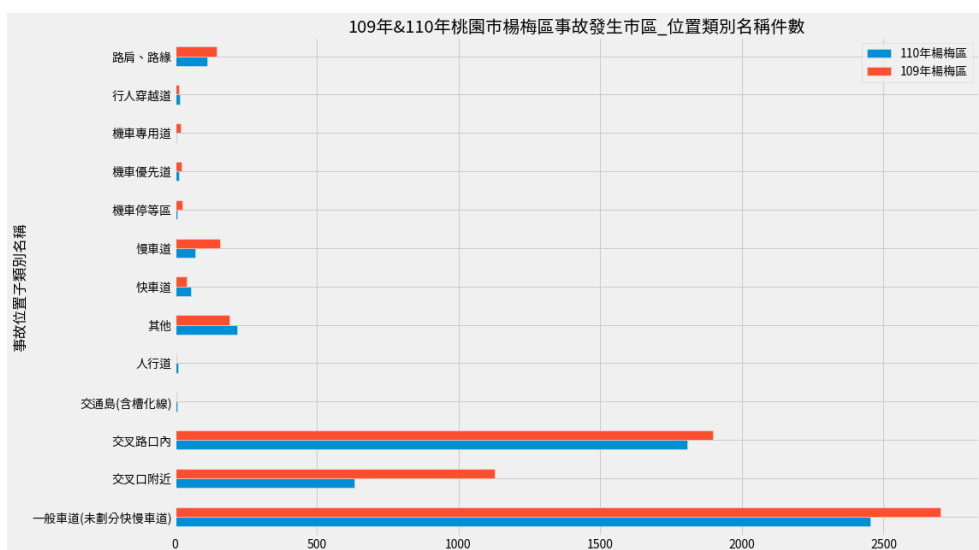
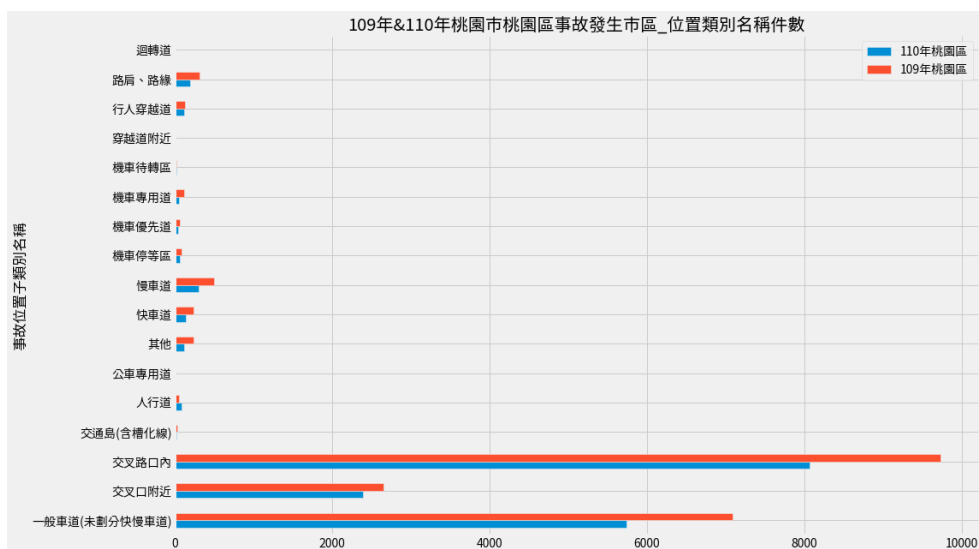
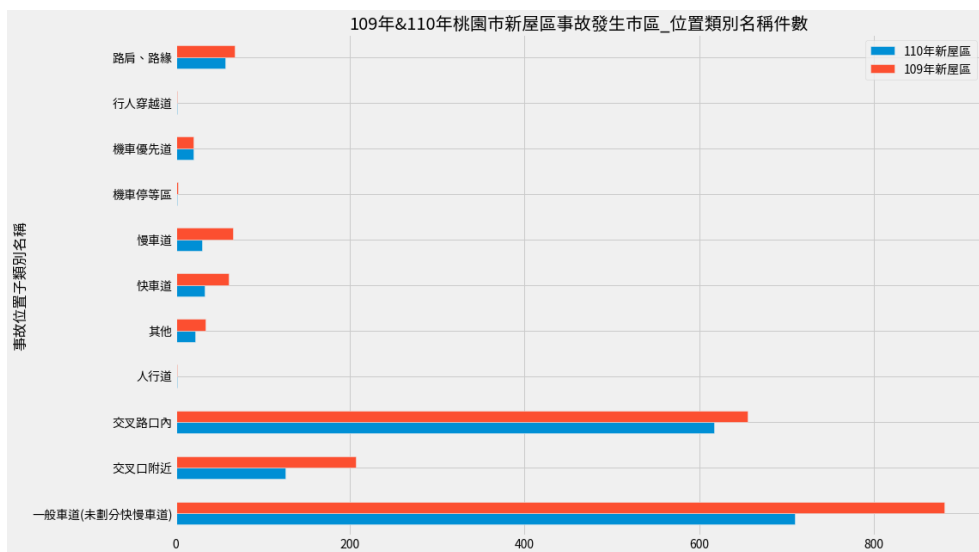
loc_110_total = loc_110_total.astype('Int64') # float轉int
loc_110_total.fillna(0, inplace=True) # 將缺失值填入0
loc_110_total
```

	中壢區	八德區	大園區	大溪區	平鎮區	復興區	新屋區	桃園區	楊梅區	蘆竹區	觀音區	龍潭區	龜山區
事故位置子類別名稱													
一般車道(未劃分快慢車道)	7211	3404	1494	1796	2450	203	710	5739	2453	2797	847	1574	3717
交叉口附近	2921	1673	517	750	1622	0	126	2389	634	605	198	434	1447
交叉路口內	9056	3675	1205	889	3466	8	617	8064	1806	2800	978	1727	2706
交通島(含精化線)	26	6	4	11	23	2	0	17	8	2	7	7	13
人行道	51	10	2	3	4	0	2	88	10	27	0	3	36
公車專用道	4	0	0	0	0	0	0	4	0	0	0	0	2
其他	345	55	27	17	122	8	22	113	220	58	6	4	38
加速車道	2	4	0	0	0	0	0	0	0	2	2	2	4
快車道	134	96	66	34	190	0	33	134	58	92	81	45	90
慢車道	339	147	74	88	266	0	30	299	71	137	88	80	384
機車停等區	79	20	15	15	38	0	2	67	9	8	4	16	28
機車優先道	16	30	45	7	91	0	20	38	14	42	12	9	39
機車專用道	39	8	14	16	51	0	0	56	6	23	2	6	4
機車待轉區	12	19	0	8	12	0	0	23	0	2	2	2	12
環道匝道	2	0	0	0	4	0	0	0	0	0	2	0	0
直線匝道	0	2	4	6	16	0	0	0	0	0	3	0	2
穿越道附近	2	0	0	0	0	0	0	5	0	0	0	0	0
行人穿越道	94	27	6	16	20	0	2	114	17	33	4	20	18
路肩、路緣	211	81	127	82	221	6	57	192	113	220	124	91	120
迴轉道	2	0	0	2	8	0	0	2	0	2	0	0	2

```
# plot 單一區域 109年&110年(mix) 橫直方圖
# total[0]: 中壢區"事故位置子類別"數量統計表
for i in range(13):
    total_bar_109 = ((total_109[i].to_df()).rename(columns={'count':'109年'+loc[i]})).set_index('事故位置子類別名稱') # 109年 轉換為dataframe 且 修正欄位名稱(count->109年各區域名稱)
    total_bar = ((total[i].to_df()).rename(columns={'count':'110年'+loc[i]})).set_index('事故位置子類別名稱') # 110年 轉換為dataframe 且 修正欄位名稱(count->110年各區域名稱)
    total_bar_mix = total_bar.join(total_bar_109) # mix 109年&110年
# plot
total_bar_mix.plot.barh(title='109年&110年桃園市'+loc[i]+'事故發生市區_位置類別名稱件數', figsize=(15,10))
plt.show()
```

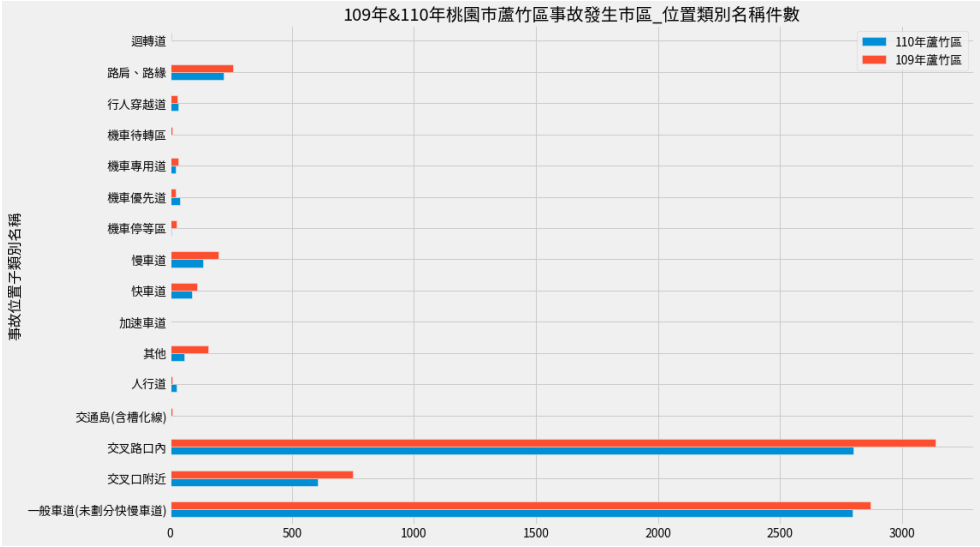




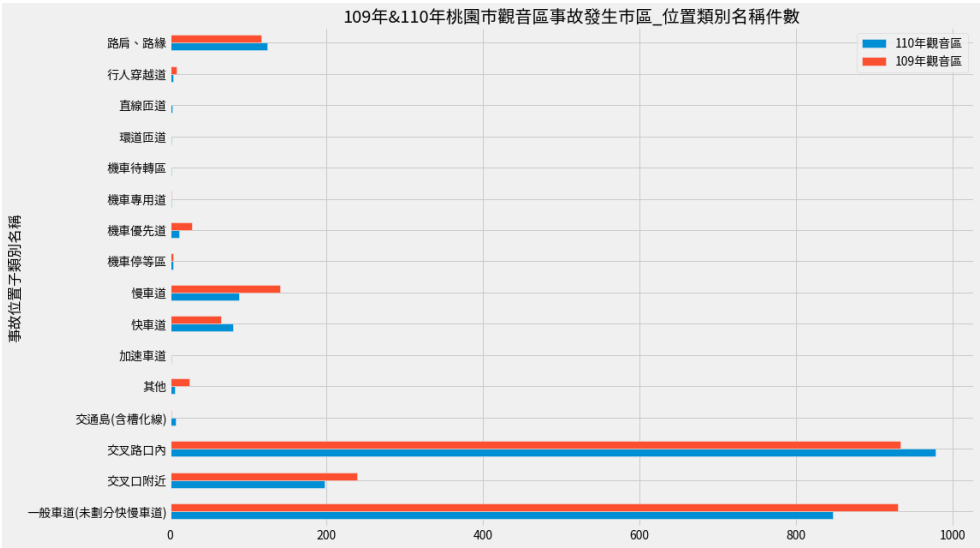




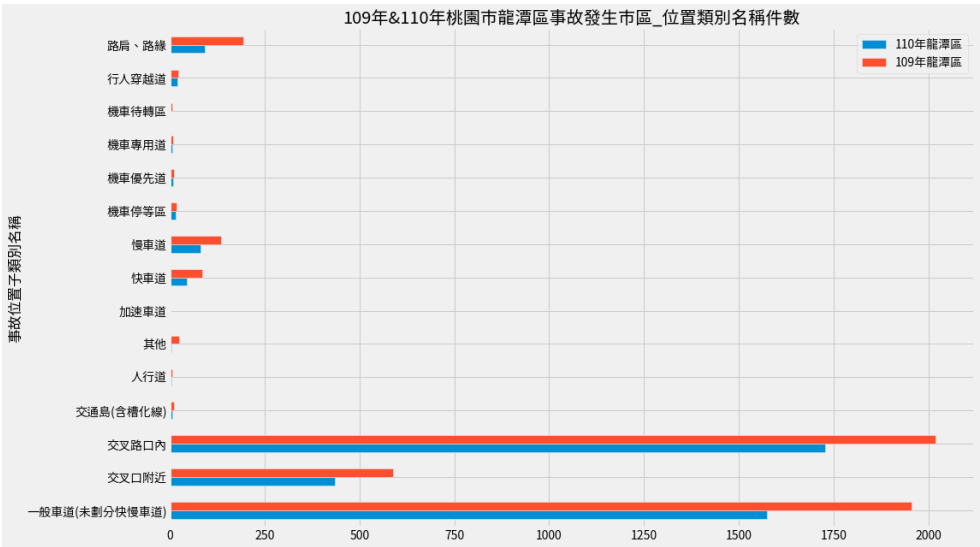
事故位置子類別名稱

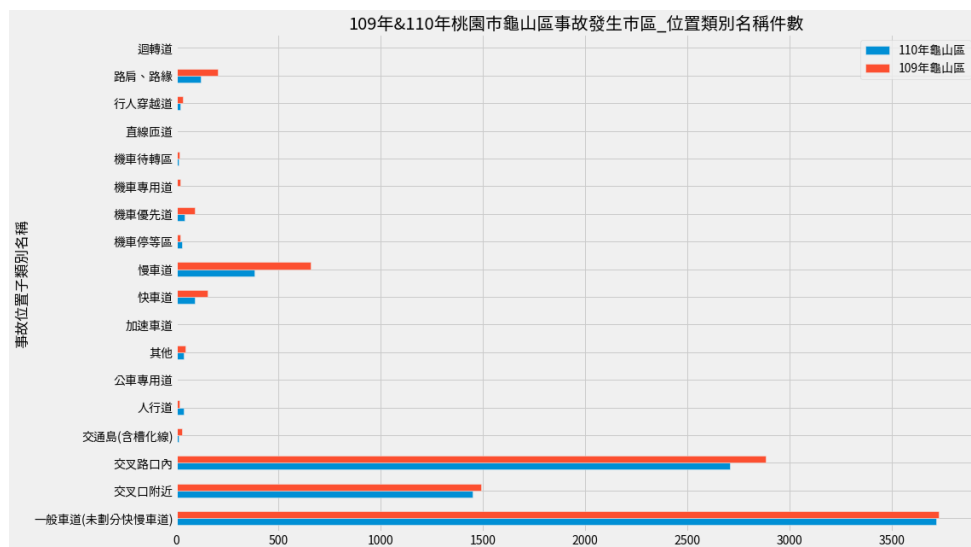


事故位置子類別名稱



事故位置子類別名稱





## 桃園市事故位置類別分析 In-depth Analysis

- 計算"加權平均" (權重: 整個區域交通事故數量/整個桃園市交通事故數量)  
(加權平均=各區域各事故位置類型之數量\*權重 之相加)  
=> 桃園市政府應加強"平均比例高的事故位置類別"之相關交通安全政令宣導

\*因各區域地理位置、大小與道路數量等環境因素皆不相同，故採取加權的方式，計算整個桃園市交通事故位置平均數量較多之類別

```
# 109年桃園市事故發生位置類別名稱_(加權)平均數量
# 計算加權平均
avg_total_109 = [0 for i in range(21)]
all = table_109.num_rows # 桃園市事故總數量
for i in range(13):
    for j in range(21):
        avg_total_109[j] += loc_109_total.iat[j,i] * (location_109_total.column(1).item(i) / all)

avg_table_109 = Table().with_columns(
    '事故位置子類別名稱', loc_109_total.index,
    '109年桃園市加權平均事故數量', avg_total_109
)
avg_table_109
```

事故位置子類別名稱	109年桃園市加權平均事故數量
一般車道(未劃分快慢車道)	4991.19
交叉口附近	2209.88
交叉路口內	6081.8
交通島(含槽化線)	33.1515
人行道	34.303
公車專用道	1.53001
其他	178.707
加速車道	2.6088
快車道	165.802
慢車道	353.134

... (11 rows omitted)

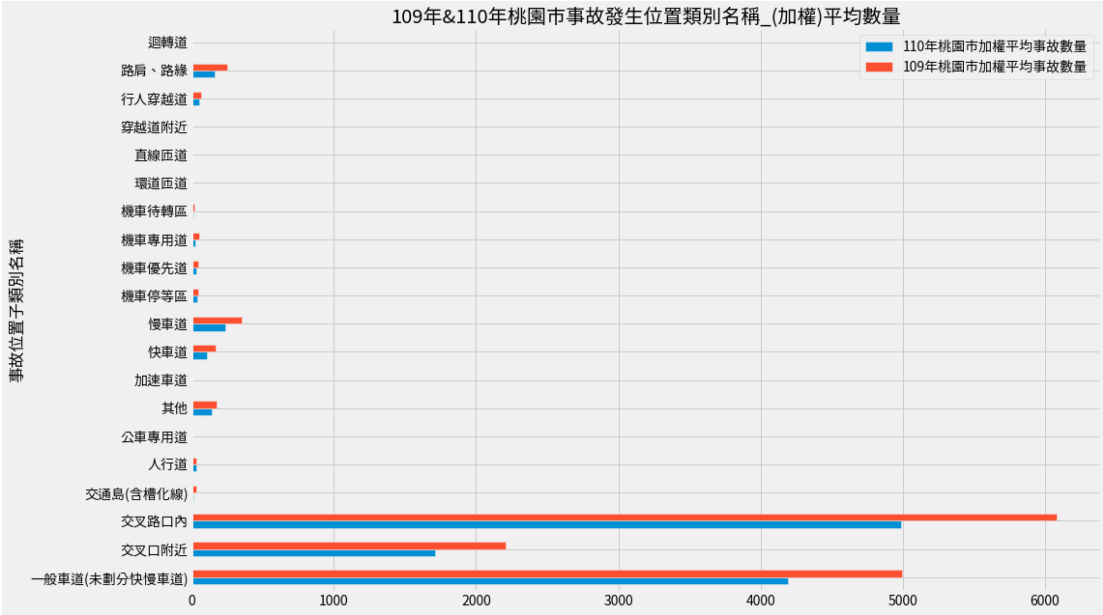
```
# 110年桃園市事故發生位置類別名稱_(加權)平均數量
# 計算加權平均
avg_total_110 = [0 for i in range(20)]
all = table_110.num_rows # 桃園市事故總數量
for i in range(13):
    for j in range(20):
        avg_total_110[j] += loc_110_total.iat[j,i] * (location_110_total.column(1).item(i) / all)

avg_table_110 = Table().with_columns(
    '事故位置子類別名稱', loc_110_total.index,
    '110年桃園市加權平均事故數量', avg_total_110
)
avg_table_110
```

事故位置子類別名稱	110年桃園市加權平均事故數量
一般車道(未劃分快慢車道)	4190.27
交叉ロ附近	1713.7
交叉路口内	4987.42
交通島(含槽化線)	14.6644
人行道	35.6184
公車專用道	1.83072
其他	138.257
加速車道	1.50925
快車道	109.649
慢車道	234.398

... (10 rows omitted)

```
# plot 109年&110年桃園市事故發生位置類別名稱_(加權)平均數量
avg_table_109 = ((avg_table_109.to_df()).rename(columns={'count':'109年桃園市'})).set_index('事故位置子類別名稱') # 109年 轉換為dataframe 且 修正欄位名稱(count->事故位置子類別名稱)
avg_table_110 = ((avg_table_110.to_df()).rename(columns={'count':'110年桃園市'})).set_index('事故位置子類別名稱') # 110年 轉換為dataframe 且 修正欄位名稱(count->事故位置子類別名稱)
avg_table_mix = avg_table_110.join(avg_table_109) # mix 109年&110年
avg_table_mix.plot.barh(title='109年&110年桃園市事故發生位置類別名稱_(加權)平均數量', figsize=(15,10))
plt.show()
```



## 桃園市事故位置類別分析 Conclusions

- (以桃園市各區域而言)由【109年&110年桃園市各區域事故發生市區\_位置類別名稱件數】之直方圖可說明，各區應加強佔比較高之事故位置的巡邏和相關交通安全宣導(依統計圖來看各區主要需加強的事故位置類別)(不同區域，事故位置類別可能不同)
- (以桃園市整體交通狀況而言)由【109年&110年桃園市事故發生位置類別名稱\_(加權)平均數量】之直方圖可說明，桃園市政府可做加強的"平均數量高的事故位置類別"之相關交通安全政令宣導/政策

# PartII 桃園市交通酒駕事故發生時的年齡分佈

## 桃園市交通酒駕事故發生時的年齡與傷亡人數分佈 EDA

### 109年酒駕發生的死亡人數與受傷人數

```
drunk_name = make_array('經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%', '經呼氣檢測0.41~0.55mg/L或血液檢測0.081%~0.11%', '經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%')
# 將有酒駕的關鍵字組合成一個變數
drunk_table109 = table_109.select('當事者事故發生時年齡', '死亡人數_24小時內', '死亡人數_2_30日內', '飲酒情形名稱').where('飲酒情形名稱', are.equal_to(drunk_name)).to_df()
# 把109年因酒駕而發生的死亡事故篩選
df = drunk_table109[(drunk_table109['死亡人數_24小時內'] > 0) | (drunk_table109['死亡人數_2_30日內'] > 0)]
df
```

當事者事故發生時年齡	死亡人數_24小時內	死亡人數_2_30日內	飲酒情形名稱
44	19	1	0 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
49	57	1	0 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
95	49	1	0 經呼氣檢測0.41~0.55mg/L或血液檢測0.081%~0.11%
171	38	1	0 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
285	54	1	0 經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%
402	31	1	0 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
472	38	1	0 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
482	19	1	0 經呼氣檢測0.41~0.55mg/L或血液檢測0.081%~0.11%
538	28	1	0 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
564	30	1	0 經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%

### 計算109年酒駕事故的各年齡段死亡總人數

```
total_18 = 0
total_26 = 0
total_30 = 0
total_40 = 0
total_50 = 0
total_60 = 0

for i in df.index:
    if 18 <= df['當事者事故發生時年齡'][i] <= 25:
        total_18 += df['死亡人數_24小時內'][i] + df['死亡人數_2_30日內'][i]
    elif 26 <= df['當事者事故發生時年齡'][i] <= 29:
        total_26 += df['死亡人數_24小時內'][i] + df['死亡人數_2_30日內'][i]
    elif 30 <= df['當事者事故發生時年齡'][i] <= 39:
        total_30 += df['死亡人數_24小時內'][i] + df['死亡人數_2_30日內'][i]
    elif 40 <= df['當事者事故發生時年齡'][i] <= 49:
        total_40 += df['死亡人數_24小時內'][i] + df['死亡人數_2_30日內'][i]
    elif 50 <= df['當事者事故發生時年齡'][i] <= 59:
        total_50 += df['死亡人數_24小時內'][i] + df['死亡人數_2_30日內'][i]
    elif df['當事者事故發生時年齡'][i] >= 60:
        total_60 += df['死亡人數_24小時內'][i] + df['死亡人數_2_30日內'][i]
```

```
df_i109 = table_109.select('當事者事故發生時年齡', '受傷人數', '飲酒情形名稱').where('飲酒情形名稱', are.equal_to(drunk_name)).to_df()
df_i109 # 因酒駕而受傷的人數
```

當事者事故發生時年齡	受傷人數	飲酒情形名稱
0	15	1 經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%
1	21	2 經呼氣檢測0.41~0.55mg/L或血液檢測0.081%~0.11%
2	39	1 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
3	55	1 經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%
4	53	2 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
...	...	...
605	57	1 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
606	61	1 經呼氣檢測0.41~0.55mg/L或血液檢測0.081%~0.11%
607	18	3 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
608	38	1 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
609	25	1 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%

計算109年酒駕事故的各年齡段受傷總人數

```
totali_18 = 0
totali_26 = 0
totali_30 = 0
totali_40 = 0
totali_50 = 0
totali_60 = 0

for i in df_i109.index:
    if 18 <= df_i109['當事者事故發生時年齡'][i] <= 25:
        totali_18 += df_i109['受傷人數'][i]
    elif 26 <= df_i109['當事者事故發生時年齡'][i] <= 29:
        totali_26 += df_i109['受傷人數'][i]
    elif 30 <= df_i109['當事者事故發生時年齡'][i] <= 39:
        totali_30 += df_i109['受傷人數'][i]
    elif 40 <= df_i109['當事者事故發生時年齡'][i] <= 49:
        totali_40 += df_i109['受傷人數'][i]
    elif 50 <= df_i109['當事者事故發生時年齡'][i] <= 59:
        totali_50 += df_i109['受傷人數'][i]
    elif df_i109['當事者事故發生時年齡'][i] >= 60:
        totali_60 += df_i109['受傷人數'][i]
```

110年酒駕發生的死亡人數與受傷人數

```
drunk_table110=table_110n.select('當事者事故發生時年齡','死亡人數_24小時內','死亡人數_2_30日內','飲酒情形名稱').where('飲酒情形名稱',are.equal_to(drunk_name)).to_df()
df1=drunk_table110[(drunk_table110['死亡人數_24小時內'] > 0) | (drunk_table110['死亡人數_2_30日內'] >0) ]
df1
```

當事者事故發生時年齡	死亡人數_24小時內	死亡人數_2_30日內	飲酒情形名稱
62	41	1	0 經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%
182	42	1	0 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
210	25	1	0 經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%
215	37	1	0 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
231	44	0	1 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
232	47	1	0 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
323	20	1	0 經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%
349	20	1	0 經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%

計算110年酒駕事故的各年齡段死亡總人數

```
total1_18=0
total1_26=0
total1_30=0
total1_40=0
total1_50=0
total1_60=0

for i in df1.index:
    if 18<= df1['當事者事故發生時年齡'][i] <=25:
        total1_18+=df1['死亡人數_24小時內'][i]+df1['死亡人數_2_30日內'][i]
    elif 26<= df1['當事者事故發生時年齡'][i] <=29 :
        total1_26+=df1['死亡人數_24小時內'][i]+df1['死亡人數_2_30日內'][i]
    elif 30<= df1['當事者事故發生時年齡'][i] <=39:
        total1_30+=df1['死亡人數_24小時內'][i]+df1['死亡人數_2_30日內'][i]
    elif 40<= df1['當事者事故發生時年齡'][i] <=49:
        total1_40+=df1['死亡人數_24小時內'][i]+df1['死亡人數_2_30日內'][i]
    elif 50<= df1['當事者事故發生時年齡'][i] <=59:
        total1_50+=df1['死亡人數_24小時內'][i]+df1['死亡人數_2_30日內'][i]
    elif df1['當事者事故發生時年齡'][i] >=60:
        total1_60+=df1['死亡人數_24小時內'][i]+df1['死亡人數_2_30日內'][i]
```

```
df_i110=table_110n.select('當事者事故發生時年齡','受傷人數','飲酒情形名稱').where('飲酒情形名稱',are.equal_to(drunk_name)).to_df()
df_i110
```

	當事者事故發生時年齡	受傷人數	飲酒情形名稱
0	18	2	經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%
1	34	1	經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%
2	22	2	經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
3	33	2	經呼氣檢測0.41~0.55mg/L或血液檢測0.081%~0.11%
4	62	1	經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%
...	...	...	...
461	53	2	經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%
462	49	1	經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%
463	39	2	經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%
464	35	1	經呼氣檢測0.41~0.55mg/L或血液檢測0.081%~0.11%
465	22	2	經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%

466 rows × 3 columns

計算110年酒駕事故的各年齡段受傷總人數

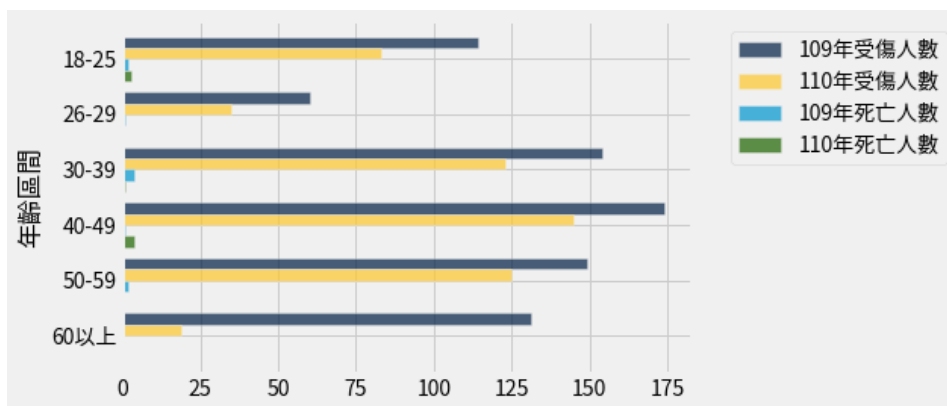
```
totali1_18=0
totali1_26=0
totali1_30=0
totali1_40=0
totali1_50=0
totali1_60=0

for i in df_i110.index:
    if 18<= df_i110['當事者事故發生時年齡'][i] <=25:
        totali1_18+=df_i110['受傷人數'][i]
    elif 26<= df_i110['當事者事故發生時年齡'][i] <=29 :
        totali1_26+=df_i110['受傷人數'][i]
    elif 30<= df_i110['當事者事故發生時年齡'][i] <=39:
        totali1_30+=df_i110['受傷人數'][i]
    elif 40<= df_i110['當事者事故發生時年齡'][i] <=49:
        totali1_40+=df_i110['受傷人數'][i]
    elif 50<= df_i110['當事者事故發生時年齡'][i] <=59:
        totali1_50+=df_i110['受傷人數'][i]
    elif df_i109['當事者事故發生時年齡'][i] >=60:
        totali1_60+=df_i110['受傷人數'][i]
```

```
die_inj_result=Table().with_columns(
    '年齡區間',make_array('18-25','26-29','30-39','40-49','50-59','60以上'),
    '109年受傷人數',make_array(totali_18,totali_26,totali_30,totali_40,totali_50,totali_60),
    '110年受傷人數',make_array(totali1_18,totali1_26,totali1_30,totali1_40,totali1_50,totali1_60),
    '109年死亡人數',make_array(total_18,total_26,total_30,total_40,total_50,total_60),
    '110年死亡人數',make_array(total1_18,total1_26,total1_30,total1_40,total1_50,total1_60)
)
die_inj_result
```

年齡區間	109年受傷人數	110年受傷人數	109年死亡人數	110年死亡人數
18-25	114	83	2	3
26-29	60	35	1	0
30-39	154	123	4	1
40-49	174	145	1	4
50-59	149	125	2	0
60以上	131	19	0	0

```
die_inj_result.barh('年齡區間')
# plot
```

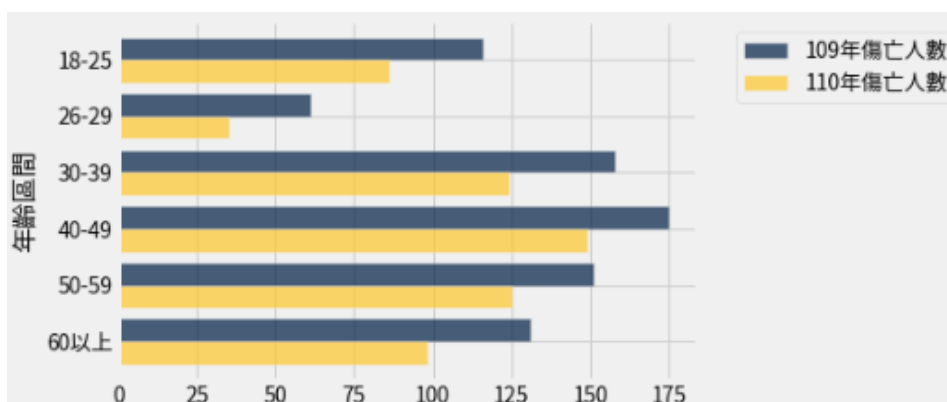


因死亡人數不多，所以將死亡與受傷人數合併計算

```
dieandinj_result = Table().with_columns(
    '年齡區間', make_array('18-25', '26-29', '30-39', '40-49', '50-59', '60以上'),
    '109年傷亡人數', die_inj_result.column('109年受傷人數') + die_inj_result.column('109年死亡人數'),
    '110年傷亡人數', die_inj_result.column('110年受傷人數') + die_inj_result.column('110年死亡人數')
)
dieandinj_result
```

年齡區間	109年傷亡人數	110年傷亡人數
18-25	116	86
26-29	61	35
30-39	158	124
40-49	175	149
50-59	151	125
60以上	131	98

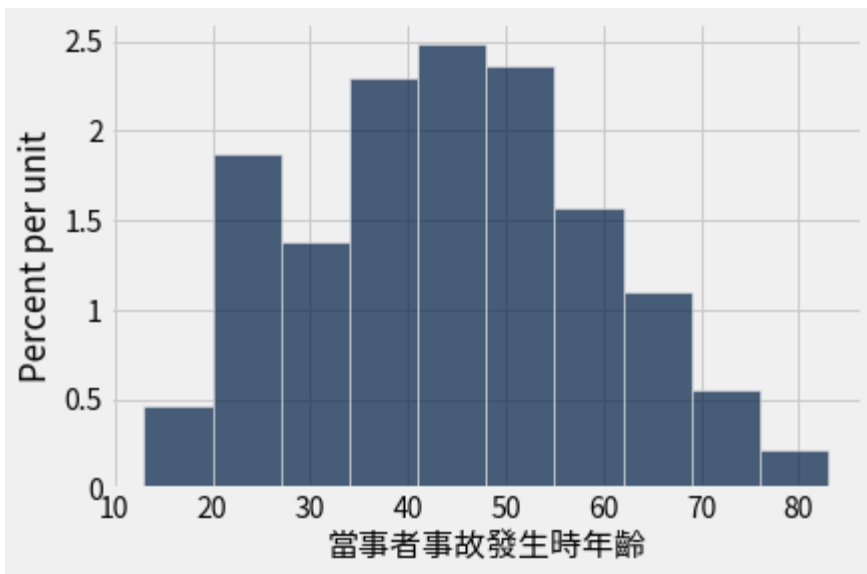
```
dieandinj_result.barh('年齡區間')
# plot
```



# 桃園市交通酒駕事故發生時的年齡與傷亡人數分佈 In-depth Analysis

## 使用bootstrap

```
drunk_name = make_array('經呼氣檢測0.26~0.40mg/L或血液檢測0.051%~0.08%', '經呼氣檢測0.41~0.55mg/L或血液檢測0.081%~0.11%', '經呼氣檢測0.56~0.80mg/L或血液檢測0.111%~0.16%')
df = table_110.select('當事者事故發生時年齡', '受傷人數', '飲酒情形名稱').where('飲酒情形名稱', are.equal_to(drunk_name))
df = df.select('當事者事故發生時年齡')
df.hist()
```



```
pop_median = percentile(50, df.column('當事者事故發生時年齡'))
```

```
pop_median
```

```
# 酒駕事故發生的中位數年齡
```

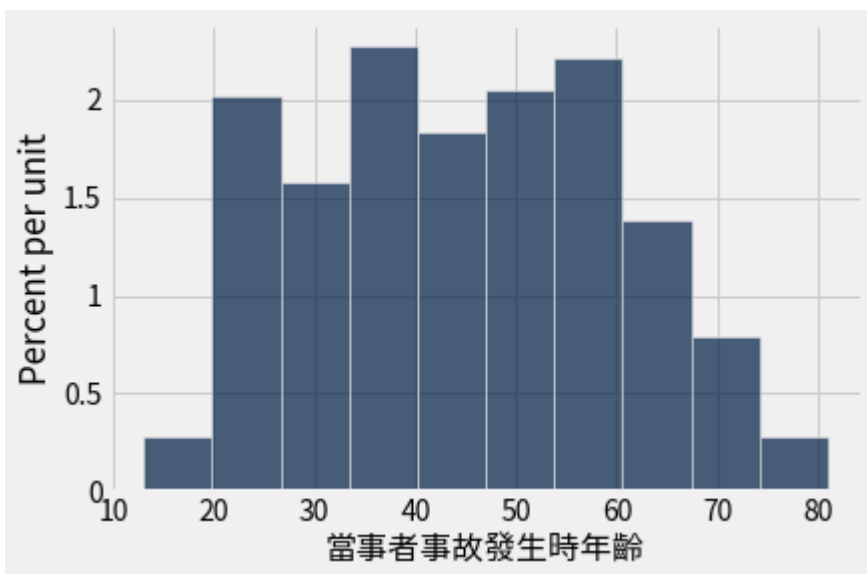
```
43
```

```
resample_1 = df.sample()
```

```
# 從table中隨機抽第一樣本
```

```
resample_1.select('當事者事故發生時年齡').hist()
```

```
# 隨機抽第一樣本的直方圖
```



```
resampled_median_1 = percentile(50, resample_1.column('當事者事故發生時年齡'))
```

```
resampled_median_1
```

```
# 隨機抽第一樣本的中位數年齡
```

```
43
```



```
resample_2 = df.sample()
resampled_median_2 = percentile(50, resample_2.column('當事者事故發生時年齡'))
resampled_median_2
# 隨機抽第二樣本的中位數年齡
```

42

```
def one_bootstrap_median():
    resampled_table = df.sample()
    bootstrapped_median = percentile(50, resampled_table.column('當事者事故發生時年齡'))
    return bootstrapped_median

one_bootstrap_median()
# bootstrap函数的中位數年齡
```

44

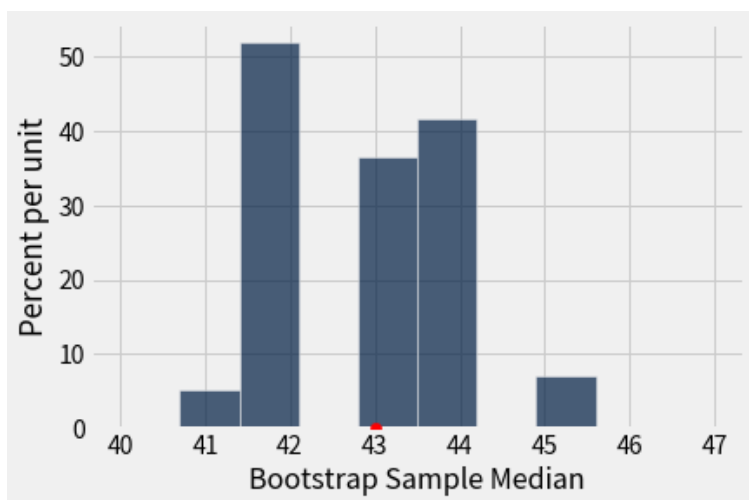
```
num_repetitions = 5000
bstrap_medians = make_array()
for i in np.arange(num_repetitions):
    bstrap_medians = np.append(bstrap_medians, one_bootstrap_median())
```

# 在每次迴圈中，我們將調用函數one\_bootstrap\_median來生成一個基於原始樣本df的bootstrapped median的值，然後我們將把bootstrapped中位數追加到集合數組bstrap\_medians中。

```
resampled_medians = Table().with_column('Bootstrap Sample Median', bstrap_medians)

resampled_medians.hist()

plt.scatter(pop_median, 0, color='red', s=40, zorder=5);
# 與原本總體的中位數年齡結果不一致
```



## 桃園市交通酒駕事故發生時的年齡與傷亡人數分佈 Conclusions

1. 由【109年&110年桃園市酒駕傷亡人數】之直方圖可說明，對比109年，110年的傷亡人數少很多。大多數傷亡人數分佈30歲到50歲之間，由此可以得知這段年齡層的職場人可能會由於應酬等會需要飲酒的場合，因此這段年齡層酒駕的人數分佈最多。有飲酒的駕駛會造成精神渙散，所以容易在車禍發生當下無法做出及時判斷，阻止事故發生。
2. 由【110年桃園市酒駕傷亡人數\_深層分析】之直方圖可說明，110年的酒駕傷亡人數明顯下降，呈現了桃園市在酒駕的法規加強起了明顯的效果，2022年全國修正《道路交通管理處罰條例》第35條及第35條之1條文，更正了酒駕罰款，提醒民眾們酒駕的嚴重性。

酒駕肇事件數於110年雖降低，但仍配合「酒駕零容忍」政策，加強酒駕執法勤務，增加罰款的金額，並廣泛地在社群媒體等宣導酒駕等嚴重性。

# 桃園市交通事故發生時的年齡 EDA

```
table_109n = table_109.select(['發生日期', '當事者屬性別名稱', '當事者事故發生時年齡', '發生市區鄉鎮名稱', '事故類型及型態大類別名稱', '肇因研判大類別名稱_個別', '受傷人數', '死亡人數_24小時內', '死亡人數_2_30日內', '飲酒情形名稱'])

table_110n = table_110.select(['發生日期', '當事者屬性別名稱', '當事者事故發生時年齡', '發生市區鄉鎮名稱', '事故類型及型態大類別名稱', '肇因研判大類別名稱_個別', '受傷人數', '死亡人數_24小時內', '死亡人數_2_30日內', '飲酒情形名稱'])
```

## 109年事故發生的年齡段

```
total_0_17 = sum(table_109n.where('當事者事故發生時年齡', are.between(0, 18)).group('當事者事故發生時年齡').column('count'))
total_18_25 = sum(table_109n.where('當事者事故發生時年齡', are.between(18, 26)).group('當事者事故發生時年齡').column('count'))
total_26_29 = sum(table_109n.where('當事者事故發生時年齡', are.between(26, 30)).group('當事者事故發生時年齡').column('count'))
total_30_39 = sum(table_109n.where('當事者事故發生時年齡', are.between(30, 40)).group('當事者事故發生時年齡').column('count'))
total_40_49 = sum(table_109n.where('當事者事故發生時年齡', are.between(40, 50)).group('當事者事故發生時年齡').column('count'))
total_50_59 = sum(table_109n.where('當事者事故發生時年齡', are.between(50, 60)).group('當事者事故發生時年齡').column('count'))
total_60_64 = sum(table_109n.where('當事者事故發生時年齡', are.between(60, 65)).group('當事者事故發生時年齡').column('count'))
total_65 = sum(table_109n.where('當事者事故發生時年齡', are.above(65)).group('當事者事故發生時年齡').column('count'))
```

## 110年事故發生的年齡段

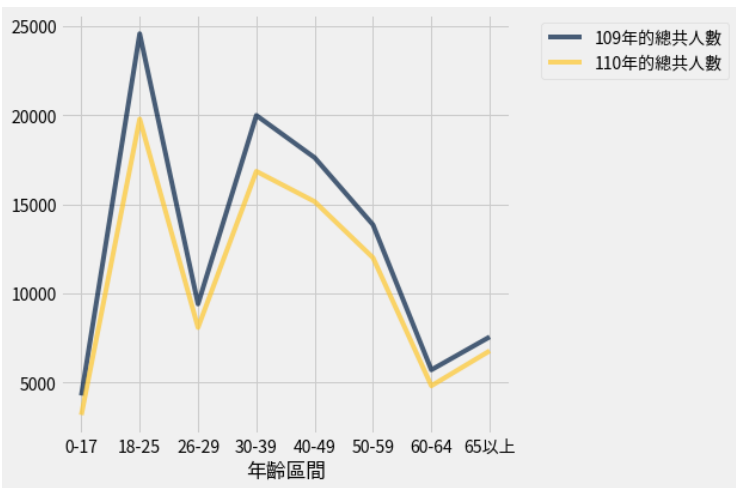
```
total1_0_17 = sum(table_110n.where('當事者事故發生時年齡', are.between(0, 18)).group('當事者事故發生時年齡').column('count'))
total1_18_25 = sum(table_110n.where('當事者事故發生時年齡', are.between(18, 26)).group('當事者事故發生時年齡').column('count'))
total1_26_29 = sum(table_110n.where('當事者事故發生時年齡', are.between(26, 30)).group('當事者事故發生時年齡').column('count'))
total1_30_39 = sum(table_110n.where('當事者事故發生時年齡', are.between(30, 40)).group('當事者事故發生時年齡').column('count'))
total1_40_49 = sum(table_110n.where('當事者事故發生時年齡', are.between(40, 50)).group('當事者事故發生時年齡').column('count'))
total1_50_59 = sum(table_110n.where('當事者事故發生時年齡', are.between(50, 60)).group('當事者事故發生時年齡').column('count'))
total1_60_64 = sum(table_110n.where('當事者事故發生時年齡', are.between(60, 65)).group('當事者事故發生時年齡').column('count'))
total1_65 = sum(table_110n.where('當事者事故發生時年齡', are.above(65)).group('當事者事故發生時年齡').column('count'))
```

## 對比這兩年每個年齡層發生的事故有沒有減少

```
happen_total = Table().with_columns(
    '年齡區間', make_array('0-17', '18-25', '26-29', '30-39', '40-49', '50-59', '60-64', '65以上'),
    '109年的總共人數', make_array(total_0_17, total_18_25, total_26_29, total_30_39, total_40_49, total_50_59, total_60_64, total_65),
    '110年的總共人數', make_array(total1_0_17, total1_18_25, total1_26_29, total1_30_39, total1_40_49, total1_50_59, total1_60_64, total1_65)
)
happen_total
```

年齡區間	109年的總共人數	110年的總共人數
0-17	4276	3183
18-25	24584	19799
26-29	9413	8093
30-39	20000	16860
40-49	17625	15158
50-59	13858	12008
60-64	5715	4820
65以上	7566	6783

```
happen_total.plot('年齡區間')
```



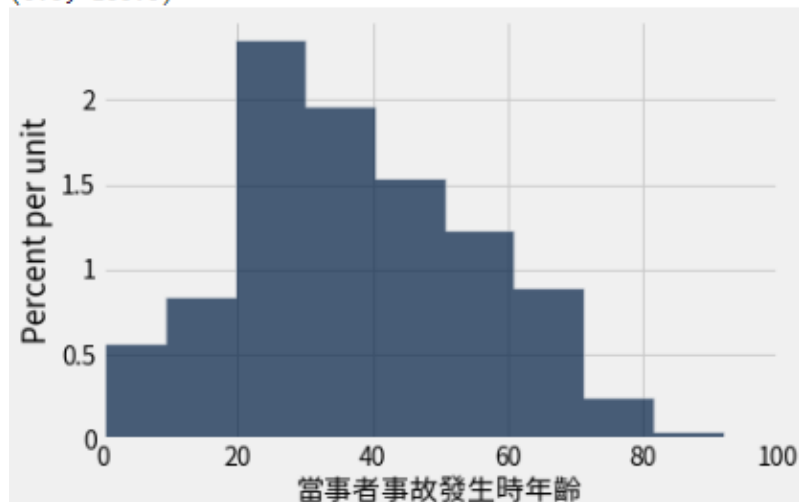
從以上的直方圖顯示藍色是109年、黃色是110年，從這邊明顯看得出18-25歲是剛考到駕照的時候，而且這個年齡層發生事故比其它年齡層都要得多。

## 桃園市交通事故發生時的年齡分佈 In-depth Analysis

使用bootstrap

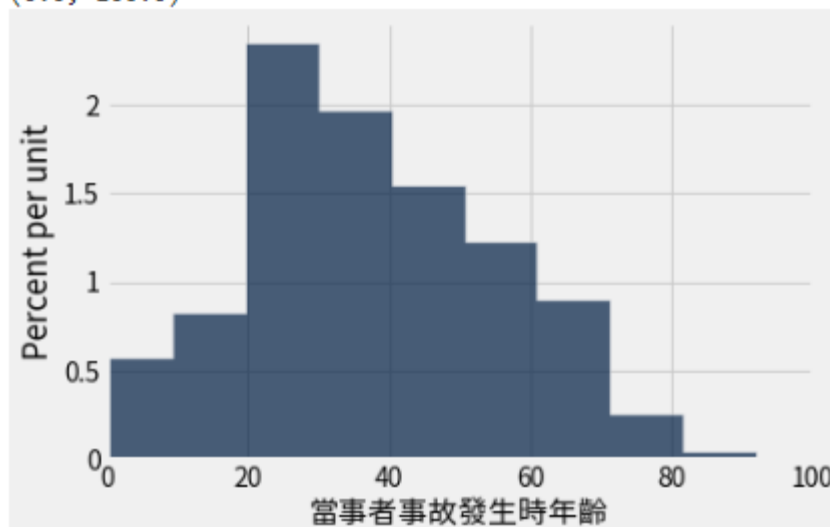
```
total = table_110.select('當事者事故發生時年齡')
pop_median1 = percentile(50, total.column('當事者事故發生時年齡'))
# 事故發生的中位數年齡
total.hist(bins=100)
plt.xlim(0, 100)
```

(0.0, 100.0)



```
resample1 = total.sample()
resample1.select('當事者事故發生時年齡').hist(bins=100)
plt.xlim(0, 100)
# 從總體抽隨機第一樣本
```

(0.0, 100.0)



```
resampled_median1 = percentile(50, resample1.column('當事者事故發生時年齡'))
resampled_median1
# 第一樣本的中位數年齡
```

36

```
resample2 = total.sample(100)
resampled_median2 = percentile(50, resample2.column('當事者事故發生時年齡'))
resampled_median2
# 從總體抽100的隨機第二樣本的中位數年齡
```

35

```
def one_bootstrap_median1():
    resampled_table = total.sample(500)
    bootstrapped_median = percentile(50, resampled_table.column('當事者事故發生時年齡'))
    return bootstrapped_median

one_bootstrap_median1()
# bootstrap函數的中位數年齡
```

36

```
num_repetitions1 = 5000
bstrap_medians1 = make_array()
for i in np.arange(num_repetitions):
    bstrap_medians1 = np.append(bstrap_medians1, one_bootstrap_median1())
```

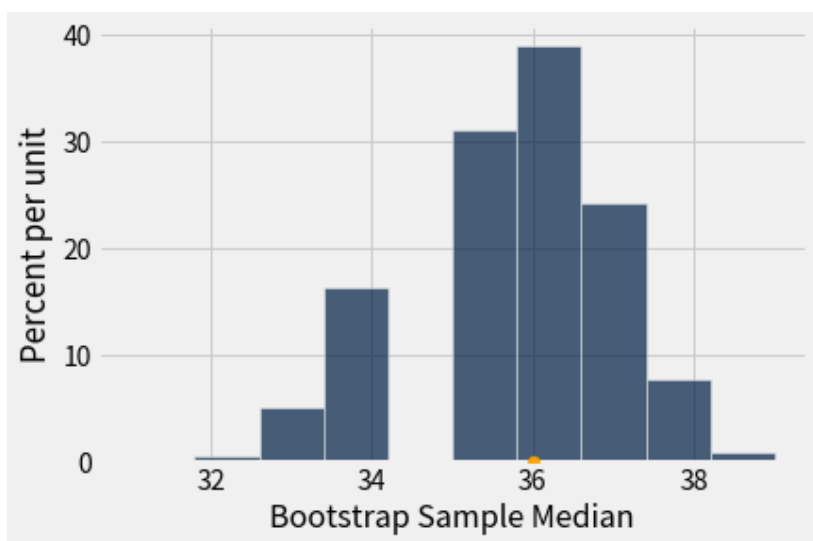
# 在每次迴圈中，我們將調用函數one\_bootstrap\_median來生成一個基於原始樣本total的bootstrapped median的值，然後我們將把bootstrapped中位數追加到集合數組bstrap\_medians中。

```
resampled_medians1 = Table().with_column('Bootstrap Sample Median1', bstrap_medians1)

resampled_medians1.hist()

plt.scatter(pop_median1, 0, color='orange', s=40, zorder=10);

# 與原本總體的中位數年齡結果相似
```



## 事故發生的年齡分佈 Conclusions

1. 由【109年&110年桃園市交通事故發生時的年齡】之直方圖說明，對比109年，110年的發生事故人數數量明顯變少，大多數分佈在18-25歲之間，這說明了大多數新手駕駛比較容易發生事故，可作為宣導及執法對象參考。
2. 由【110年桃園市事故發生人數\_深層分析】之直方圖說明了，桃園市政府可做加強“安全駕駛”之相關的交通安全的宣導及道路管制、巡邏。未來仍需持續加強交通事故防治工作。

因 18-25 歲多為大專院校學生，由前述分析可知此年齡層駕駛人對路權觀念、安全駕駛觀念及速度管理認知不足，仍應再加強機車駕駛安全教育。雖然發生事故人數數量下降了，政府依舊能改善交通環境、宣導交通安全教育、維護交通秩序等，為民眾負責任。

## Total Conclusions

由於交通問題關聯到的層面很多，因此我們取出交通事故統計表中的年齡,酒駕,事故位置等重要因素，做分布圖與深入分析，並分成兩個part，藉此了解桃園市政府在交通層面可以在做加強的部分。