

●資料科學，打火救人—當資料科學遇上火災火警，以高雄市火災火警資料探勘為例

高雄市政府消防局火災預防科 高文宗

壹、前言

對抗火災的最佳策略是採取預防性攻擊，防患於未然，若以資料科學的語言描述，就是定義問題、資料盤點與清理、分析建模、預測、決策支援。

以高雄市為例，每年的火災案件不到一百件，但是消防隊員還是忙不完，有一個重要的原因是雖然每年「火災」數量不多，但是「火警」的案件卻是數以千計，火警和火災，差別只在一線間，家中瓦斯爐燒開水未關，鄰居報案後，消防隊來得及破門而入關掉瓦斯，就只是火警；來不及關掉，就可能演變為造成生命與財物損失的火災。

不論火災或火警，其危險因子（例如起火原因、建築物特性、人口特徵等等）可能都是共通的，若能評估火警風險，或許就等同建立了火災風險模型，這即是 D4SG（Data for Social Good Fellowship）資料英雄計畫「火災風險地圖」專案正在設法解決的公共問題。

很榮幸，高雄市政府消防局與一群「用資料力，做公益」的資料英雄合作，利用週末和晚上，共同分析過去數年的消防案件，從無到有，開始打造臺灣第一個「資料科學，打火救人」的實戰經驗。

貳、資料來源

- 一、高雄市政府消防局火災紀錄資料
- 二、高雄市政府消防局火災分析表
- 三、高雄市政府消防局火警出動人車數
- 四、高雄市政府消防局各大隊補助安裝住宅用火災警報器場所清冊
- 五、高雄市政府社會局資料（低收入戶、獨居長者、身心障礙）
- 六、高雄市政府稅捐處建築物資料

參、問題解決

- 一、優化人力資源配置的

（一）問題：如何使勤務編排更有效率？

高雄市政府消防局基層消防人員的輪班待命時間為早上 8 點到隔天 9 點，共 25 小時。如能從歷年的火災紀錄看出忙碌的時段差異，在較不忙碌的時段安排較少人力，在滿足基本戰力與鄰近分隊即時支援的前提下，不必讓每個隊員皆值勤 25 小時的時間，

能夠讓隊員有更多時間休息，避免不必要的人為疏失，並作更有效率之人力配置。

(二) 分析方法：每個分隊在什麼時段可精簡人力，什麼時段須要多加派人力？

1、從派遣次數的角度出發：

分隊的派遣次數越多代表該時段火災火警發生越頻繁，消防人員出動的次數越多，以線圖呈現分隊忙碌的時段。

2、以出動車數的角度出發：

出動車數越多，代表該時段、該分隊除必須處理很多火警火災的報案外，棘手案件也比較多，以線圖呈現分隊車輛出動情形（如圖 1）。

3、派遣車數與出動車數做比較：

經由派遣車數與出動車數一起做比較的話，可以很容易地看出哪些時間點是分隊出動頻繁且又需要較多人手的時候。

(三) 分析結果：

經統計分析 100 年 1 月至 105 年 5 月底，火災火警通報次數與出動車次顯示，上午 10 時至下午 6 時是火災火警派遣與出動車次的高峰期。

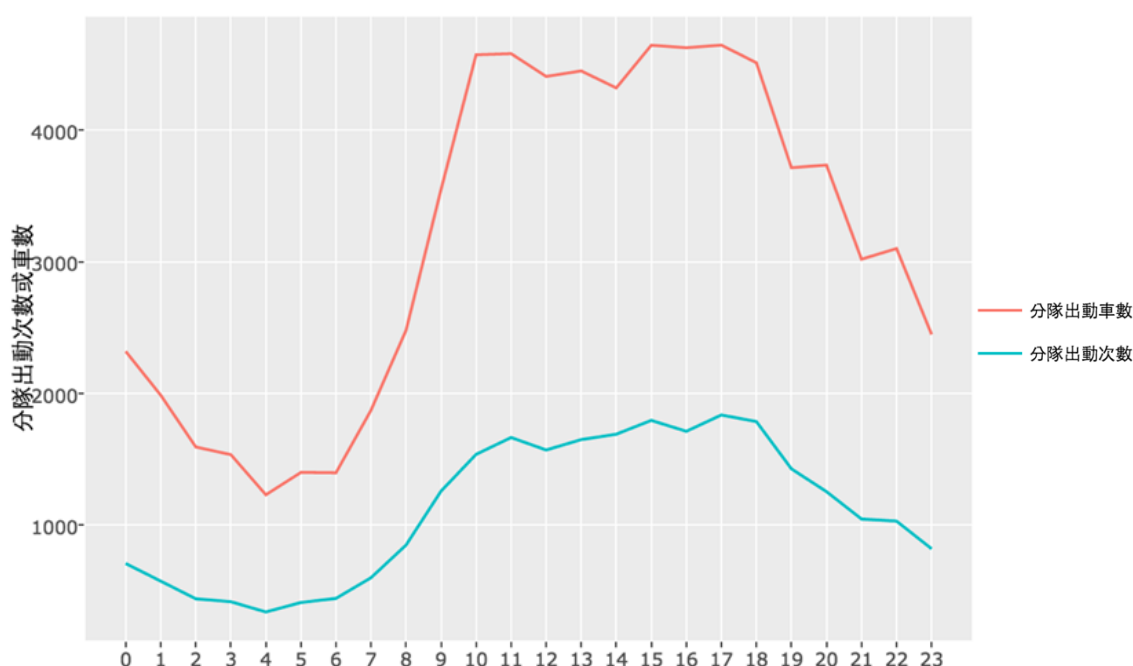


圖 1：火災火警紀錄的發生時間

二、住警器發放

(一) 問題：

「住宅安裝住警器、家人安全有保庇」，這是消防機關宣導安裝住宅用火災警報器（住警器）的標語，安裝住警器可以降低火災災害的嚴重性。消防局每年都會收到善心人士捐贈的住警器，而目前面臨到的問題是，如何發放這些愛心住警器，得以讓社會的愛心發揮最大效益。

(二) 分析方法：

觀察各區里的火災風險指標，探討往年的發放情形，在各區里的發放比例上，是否有其他的改善空間。

1、資料蒐集

- (1) 高雄市消防局住警器發放列表【103,104 年】
- (2) 高雄市火災火警資訊【消防局】
- (3) 高雄市弱勢族群（身心障礙、獨居長者）資訊【社會局】
- (4) 高雄市建築物資訊【稅捐處】

2、警報器發放指標

(1) 指標設計目的：

希望透過一個指標，去瞭解高雄市各個區里發放的狀況，希望將有限的資源，做合理的發放，達到最大效益。

(2) 指標定義：

指標的設計主要是藉由風險分數以及住警器申請數量分數，定義出住警器發放指標，定義如下：

$$\text{【住警器發放分數】}_i = \text{【住警器申請數量分數】}_i - \text{【風險分數】}_i$$

(3) 住警器發放指標說明：

每個指標的分數，都是將原始數值標準化後（重新將級距縮放為 0 到 100 之間）的結果，因此不同指標之間的分數，可以相互比較。

對於住警器發放分數的定義，可以將指標看作兩個部分，第一個部分是住警器申請數量分數，第二個部分則是風險分數：

A、第一部分－住警器申請數量分數：

統計各區里已經發放的住警器申請數量，並將原始數值標準化，分數越高，代表發放數量越多。

B、第二部分－風險分數：

在火災機率成因檢定中，藉由建立統計模型，去預測每個區里發生火災的機率，並將機率標準化，分數越高，代表該區發生火災機率越高。

透過第一部分所計算的住警器發放數量分數，減去第二部分的風險分數，便可以得到各區里現狀的住警器涵蓋分數。

(三) 分析結果

由上述發放指標分數資訊，可以得到各區里現狀的住警器涵蓋分數，呈現如圖 2，如某區里指標分數大於 0（地圖上顯示藍色），則代表住警器發放情況過剩，反之如某區里指標分數小於 0（地圖上顯示紅色），則代表住警器發放情況缺乏，將來消防局發放愛心住警器時，可利用上述資訊來調整各區里發放數量的比例，以達到效益極大化之目的。

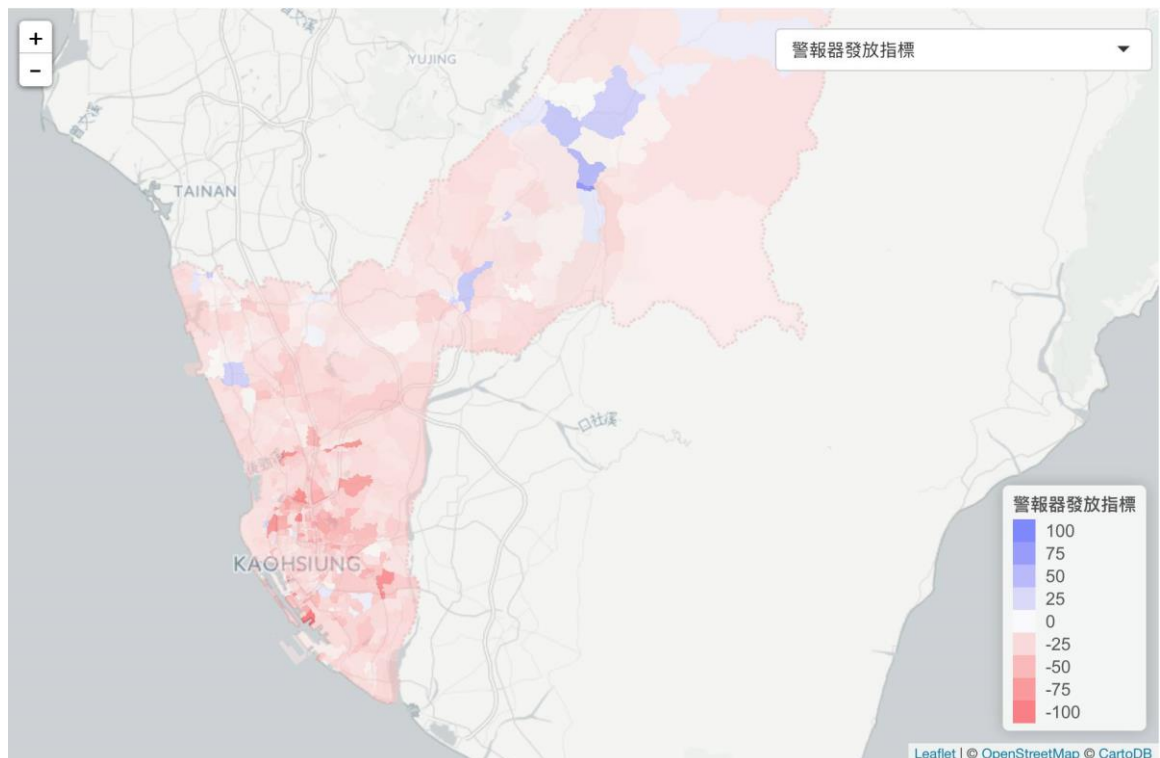


圖 2：住警器發放指標地圖

三、火災火警發生原因

(一) 問題：

瞭解高雄市地區火災火警真實發生原因為何？並藉由發掘此原因，提供市民最有效的防災與預防措施，進而讓高雄市火災火警風險降至最低，與市民共同創造一個安全無虞的環境。

(二) 分析方法：

彙集高雄市 100 年 1 月至 105 年 5 月底所有建築物火災火警通報資料，包含地籍資料、火災火警發生時間／原因／場所，以地區做整合，提供出一個初步的資料庫，並從中取得想要的資訊。

1、火災火警主要成因

利用整併後資料庫，將高雄市所有火災火警的數量對火災火警成因作圖（如圖 3），可以發現造成高雄市火災火警主要 Top3 成因為「原因不明」、「煮食不慎」、「電器因素」。（註：「原因不明」為沒有紀錄或顯示原因待查）

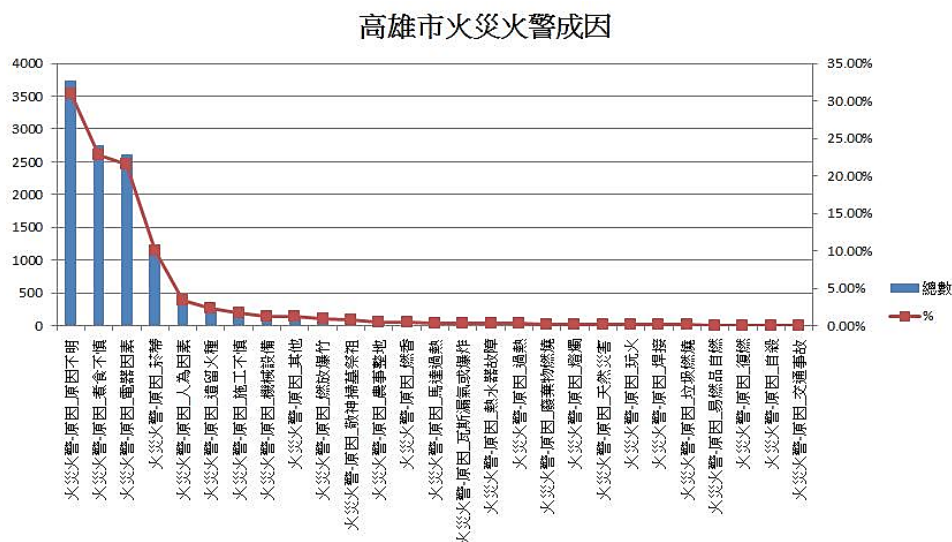


圖 3：高雄市火災火警成因

2、火災火警主要場所

利用整併後資料庫，將高雄市所有火災火警的數量對火災火警發生場所作圖（如圖 4），可以發現造成高雄市火災火警主要 Top3 場所為「一般住宅」、「其他」、「工廠及倉庫」。

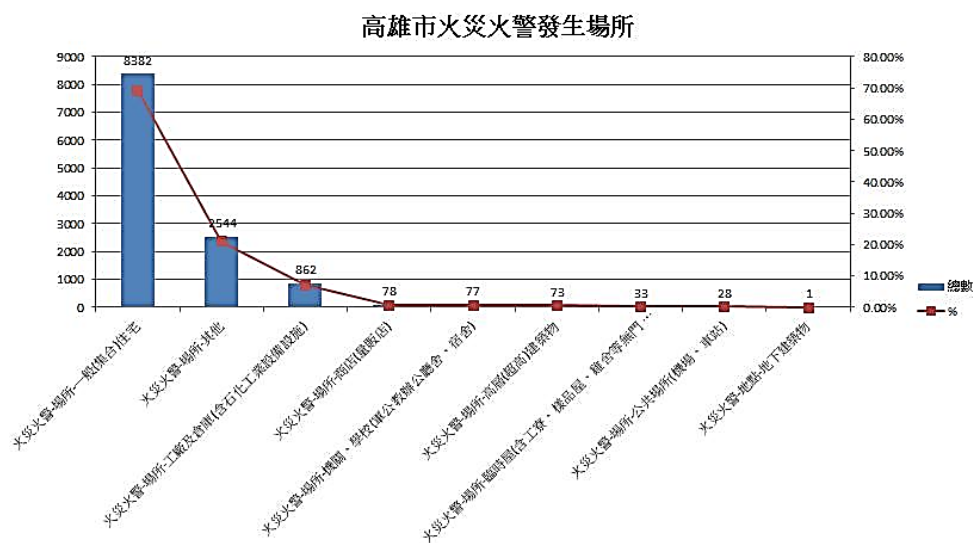


圖 4：高雄市火災火警發生場所

3、火災火警主要發生地區

將高雄市所有火災火警的數量 Top 10 的地區作圖（如圖 5），可以發現左營區佔前三名，分別是「福山里」、「新上里」及「菜公里」。

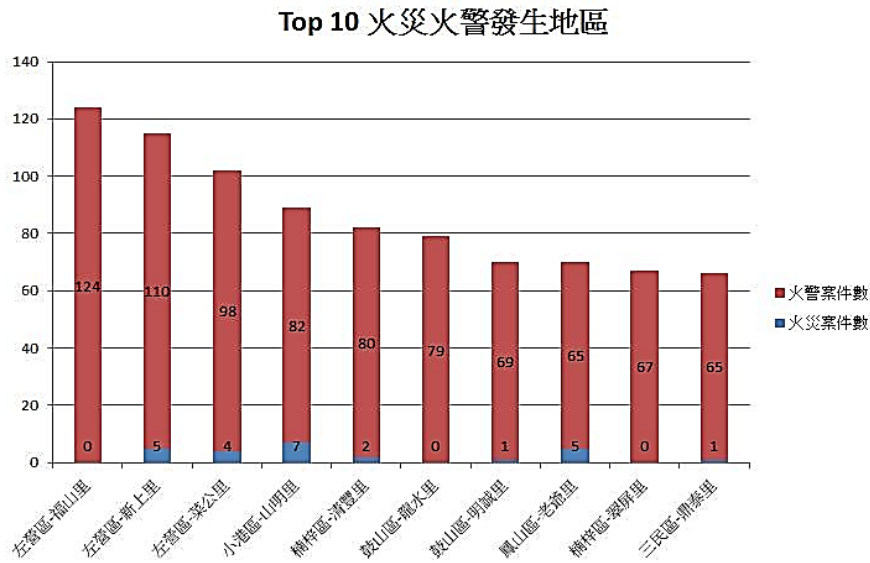


圖 5：Top 10 火災火警發生地區

（三）分析結果：

由上述分析得知，高雄市整體火災火警主要是發生在住宅中的「煮食不慎」、「電器因素」，以及在紀錄上原因較難辨別的項目。另外，發生最多火災火警的區域為左營區的三個里（福山里、新上里、菜公里），且其火災火警成因前三名也是煮食不慎、電器因素與原因不明。

藉由火災火警通報資料，加上開放資料中的區里界圖資訊，將火災火警發生次數以熱圖方式呈現（圖 6），紅色越深代表該區火災火警發生次數越多，反之綠色則是次數較少區域，並提供各區里火災火警成因前三名以及發生場所前兩名，方便使用者查詢各區域狀況，制定符合各區域最有效率的防範火災火警措施。

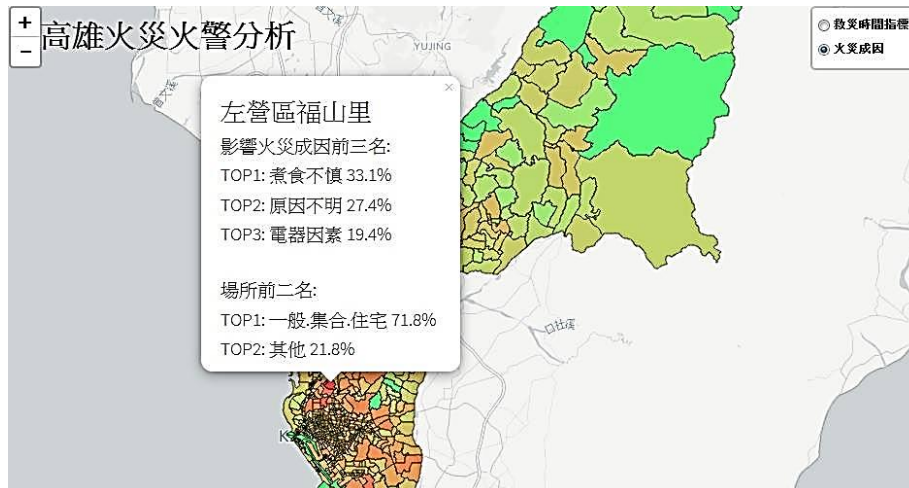


圖 6：高雄市火災火警互動式地圖分析

(四) 火災機率成因檢定

透過統計分析「身心障礙」、「獨居長者」及「房屋資訊」這三大類變數，來預測風險機率值，運用邏輯思回歸(Logistic Regression)進行分析，並搭配逐步回歸(Stepwise Regression)的方式挑選對風險指標預測有顯著效果的變數，建立模型與評價。

經過十次訓練模型的結果，至少出現過一次顯著性的變數如下：身心障礙-嚴重程度-中度、身心障礙-人數、房屋資訊-營業用面積平均、房屋資訊-透天住戶數及房屋資訊-大樓住戶數。

從顯著的變數可以看出，影響一個地區(里)火災機率的高低，取決於身心障礙的人數及房屋資訊。獨居長者相關的資料，在此樣本模型中，無顯著效果。