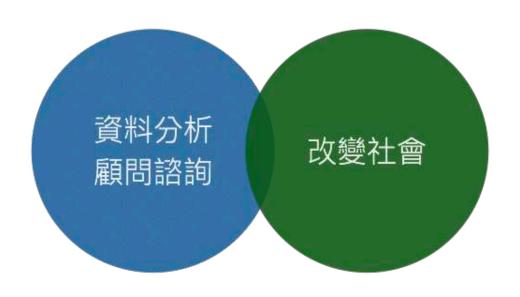


火災、水汙與家暴 資料英雄如何解決社會問題

謝宗震 (johnson@dsp.im) DSP 智庫驅動 共同創辦人 D4SG資料英雄 計畫主持人







運用數據,解決大問題



The Big Problems



D4SG資料英雄計劃



https://d4sg.org

發起單位:智庫驅動公司、開拓文教基金會、ITSA社群運算與巨量資料跨校中心





農地重金屬 污染防治



建物火災風險 預警



親密關係暴力 防治



緊急醫療流程 優化



社工排班問題 建議



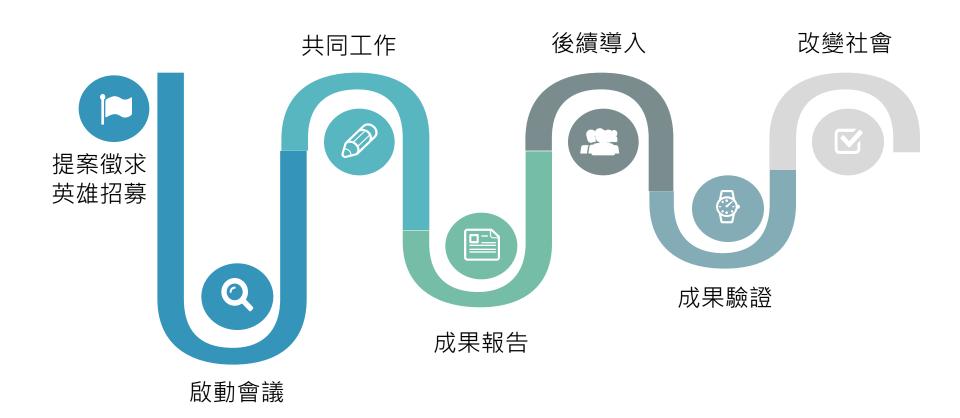
簡報大綱

- 1. D4SG 資料英雄計畫簡介
- 2. D4SG 家暴防治案例介紹
- 3. 資料科學專案的挑戰
- 4. Lessons Learned



D4SG資料英雄計畫 是怎麼進行的

D4SG Road Map



想解決什麼問題?

- 找出建築物火災熱點
- 找出建築物火災潛在危險因子
- 找出火災高風險族群
- · 找出救災(護)尖峰、離峰時段,優 化人力配置



D4SG資料英雄計畫啟動會議 (火災風險地圖)

高文宗科長 / 高雄市消防局火災預防科



共同工作會議 (誠致教育基金會 / 均一教育平台 – 如何幫助孩子學得會)



成果發表會 (新竹縣環保局 - 廢水防治看板)





D4SG 案例分享

Domestic Violence (Risk Management)

北市傳恐怖家暴 女子遭丈夫潑酸二度灼傷送醫







2017-07-01 22:26





8+

[記者陳恩惠 / 台北報導] 台北市士林區今天晚間驚傳一起夫妻吵架的家庭暴力案件! 年輕夫 妻發生口角後,女子疑遭丈夫潑灑硫酸傷害,造成上半身及頭部,百分之27的二度灼傷,所幸 女子送醫時意識仍清醒、至於事發的原因、韓區士林營分局文林派出所正在找尋女子的丈夫釐 清中。



基河路一棟7層樓電梯大樓業傳家暴案件。 (記者陳恩惠攝)

台北市消防局119勤指中心、今天晚間8時42分接獲 報案、指稱位於北市士林區基河路某大樓6樓的住 戶, 陸續傳出爭吵聲響, 隨即有女子尖叫「你竟然潑 我!」、「叫救護車!」與咒罵聲,男子隨即離家, 管理員見狀、趕緊打119表示住戶發生暴力事件急需 救護、消防局立刻派出1輛救護車3名救護人員前 往。

救護員初部檢傷發現,年約30歲的女子,頭部及上 半身遭硫酸潑灑,導致上半身有百分之27的二度灼 傷、由於女子的丈夫潑硫酸之後、隨即匆匆難家、雙 方爭吵及發生家暴的原因仍待調查,據悉,這戶從未 發生過家暴或傷害事件、是否因夫妻口角意外,仍要 釐清。

拔指甲、灌毒// 虐死2歲童 主嫌判死刑





2012-05-28



〔記者張文川/台北報導〕才兩歲五個月大的男童王昊,還來不及長大,就遭四個狠心的毒蟲 凌虐致死,生前慘遭拔指甲、烙腳底、注射海洛因、強灌安非他命,虐行之殘,令人髮指;台 北地院昨宣判主謀劉金龍死刑,另三名共犯分別判無期徒刑、十四年、十三年徒刑。



LINE



兩應五個月大的男童王昊,生前慘遭四個毒蟲 拔指甲、烙腳底、注射海洛因、強漢安非他命 等暴行凌虐致死,台北地院昨判主謀劉金龍死 刑。(資料照)

杜絕慘劇 男童姑姑籲修兒少法

男童的姑姑王薇君昨表示:「謝謝法官對吴昊處境的 感同身受,總算有法官站在被害人家屬的角度,還給 吴昊一個公道」,「希望吴昊的犧牲能促成兒少法第 五十四條之一的修法,把毒品人口列入高風險家庭, 才不會有下一個吴昊。」

審判長孫正華昨宣讀判決主文後,罕見地當庭宣讀八百字「感觸」,認為在「保障被告人權」的大旗下,訴訟制度已產生「重被告,輕被害人」的失衡現象,如此「物極必反」,對廣大善良民眾絕非善事。

審判長: 重被告人權 絕非善事

因「傷害致死」最重只能判無期徒刑,合議庭以最重可判死刑的「以強暴方式使未成年人施用毒品罪」,及槍砲、贩毒、轉讓毒品、傷害、略誘等罪,將主謀劉金龍判一個死刑、三個無期徒刑及八個月徒刑,褫奪公權終身,應執行死刑;共犯周健輝判無期徒刑及經濟八援終息。如母母判刑十二年,即呼應判刑十四



跟蹤、監視或惡性打擾(T7) 懷疑第三者介入(T13)、感情、外遇 婚姻狀態

關係暴力經驗: 武器恐嚇 (T4)

威脅殺害(T5)

威脅一起死 (T6)

求救威脅、暴力(T12)

暴力型態、武器、

持續時間、求助時間差、



關係

依附

忌妒嫉控制慾

關係結束

衝突

威脅恐嚇

嚴格的性別角色

年齡、職業、教育 經濟壓力 (T11) 精神疾病、照顧、 經濟壓力 身心障礙

> 酗酒 (T9) 子女暴力史(T2) 非家人暴力史(T10) 賭博、毒品 自殺意念及行為

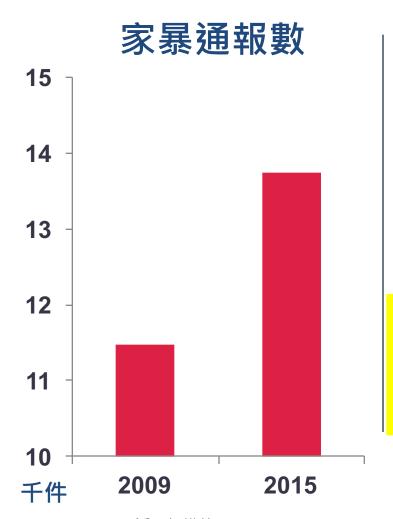
> > 居住區

₩ 社區

集中弱勢
居住穩定性
集體效能
社會聯繫
文化規範標
學校

T: TIPVDA 親密關係暴力危險量表

台北市家暴問題現況



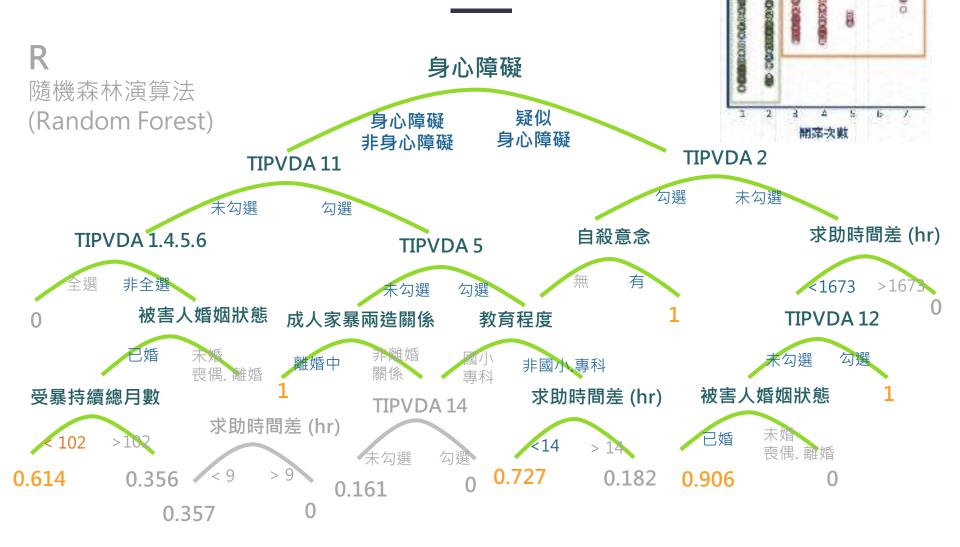
一年超過1萬3000人需要協助

平均

1位社工需要處理 116 件案子

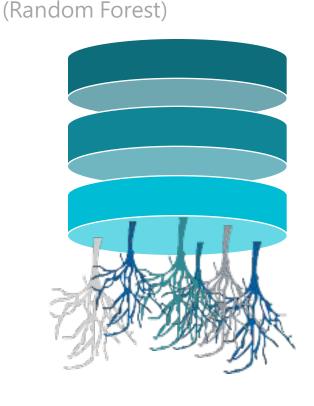
其中 56% 屬親密關係暴力 33% 為 再犯案例

再次受暴預測模型



再次受暴預測模型

R 隨機森林演算法



平均準確率 96.3% Average Accuracy

召回率 Recall Rate 88.9%

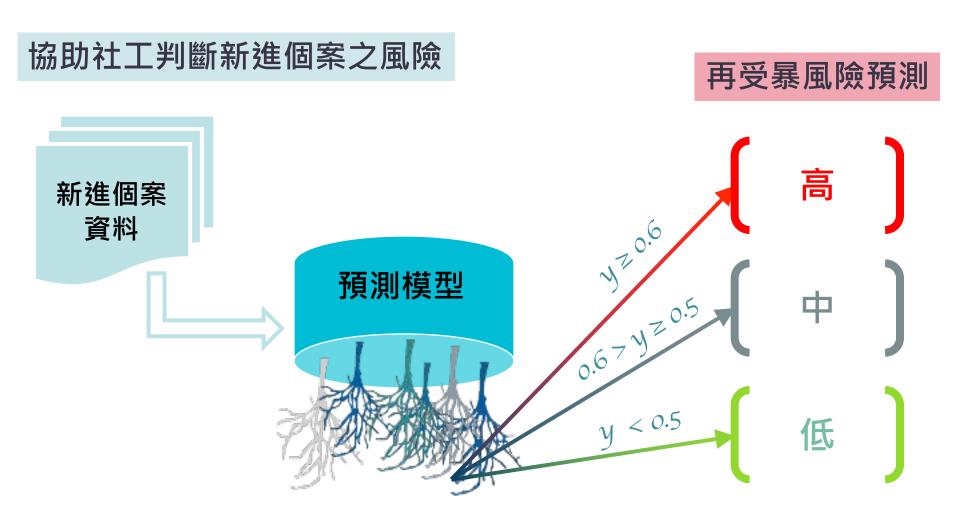
精確率 48.7% Precision

F值 F1-Measure [because of high false positives]

完整技術文章

Chen, Ya-Yun, et al. "Measuring the unmeasurable-a project of domestic violence risk prediction and management." arXiv preprint arXiv:1710.06842 (2017).

再次受暴預測模型



後續影響



增加社工人員及高階公務員的 數據素養



北市里長依據案件分布狀態制定 不同的預防政策



資料科學專案的挑戰

.....

資料專案所面臨的挑戰

最難的是把問題定義清楚

次難是成果要能落實

再來就是根本沒有資料

或者是資料真的很髒

最容易的其實資料分析



目標

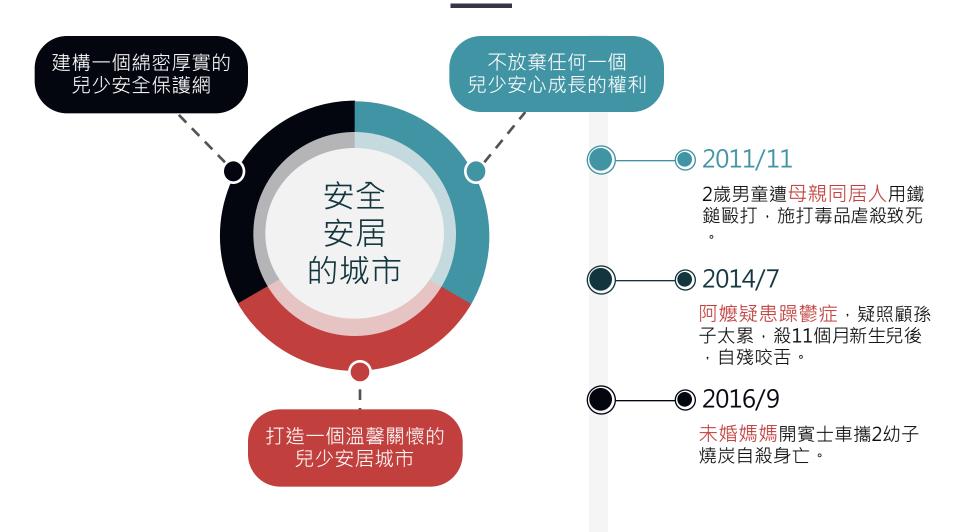
定義目標

- 大問題
- 工作流程理解
- 對齊動機
- 問題收斂



共同工作會議 (新北市高風險中心 - 危機警示模組)

大問題



工作流程理解 & 資料盤點

新北市高風險家庭服務管理中心業務流程



對齊動機

唐鳳

要公部門主動開啟對話,通常要對齊動機。

能「創造亮點」的案子->找政務官 能「降低風險」的案子->找事務官 能「簡化行政」的案子->找承辦人





收斂問題 & 預期行動

- 可以用資料科學解決的問題 (input to output)
 - Prediction, Inference, Clustering, Optimization, ...
- 可以有具體行動的問題(output to impact)
 - Prevention, Intervention, ...

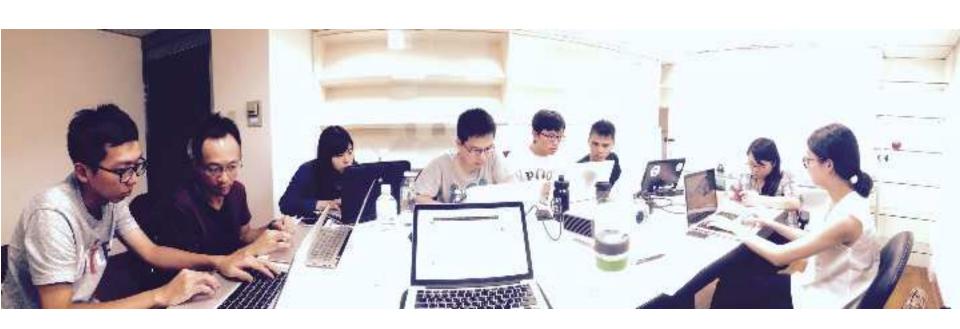


資料

資料

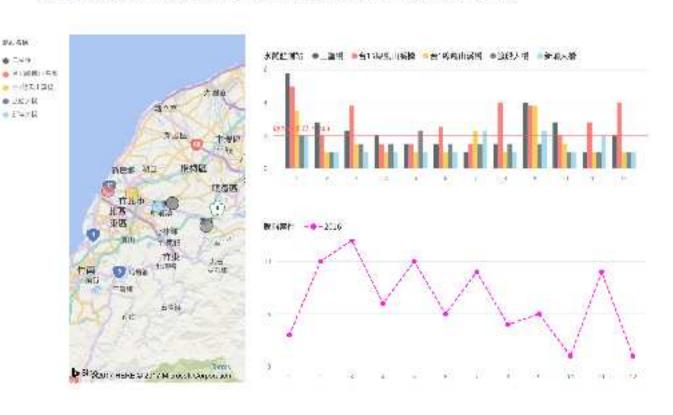
- 資料理解
- 品質檢驗

- 資料處理與合併
- 大資料表設計



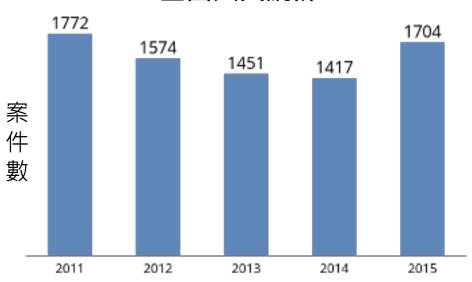
共同工作會議 (天下雜誌/國家太空中心 - 雲林農地種電大盤點)

新竹縣環保局 水汙染 裁罰案件與水質分析-河川汙染指數



工廠自主通報 的水質檢測數 據,永遠符合 管制標準

全國火災統計



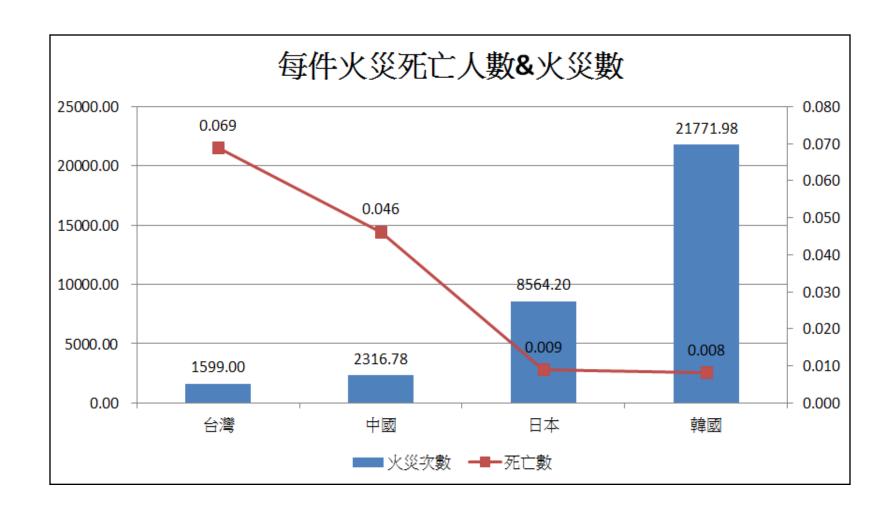
每年約1600件火災 全國消防員額約14000人 過勞???

勤務繁重過勞? 桃園消防員心肌梗塞倒下

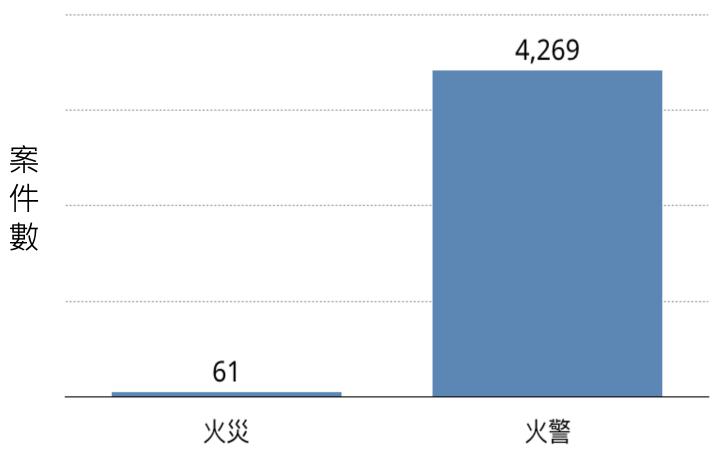
⊕2015/05/01 16:43

記者 李佳槿 / 摄影 摩宗慶 桃園 報導





2015年高雄市火災火警統計



註:2016年火災火警定義已改為 A1, A2, A3

資料 - 品質檢驗 (髒)

資訊系統的設計,可能會強調使用者經驗,

但卻不會在意資料分析者經驗。

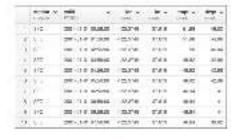
- 同一案件姓名、性別、身份證號不一致
- 未來人 (生日時間有誤)
- 只有人才看得懂的地址 (XX公園的土地公廟後面)
- 全形、半形字、行政地區劃分改變

探索性資料分析,

是理解資料、檢驗品質的最佳解法

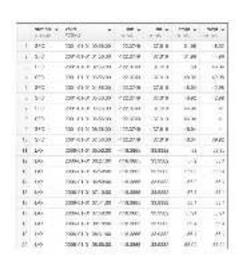
資料 - 合併與大表設計

- Data frame 是資料分析的基礎
- 資料合併才能產生價值
 - 使用情境 (火災風險熱區:村里級、建築物、地點)
 - 計算指標
 - 主鍵 (key)
 - ID、時間、空間





	mile o	1118 W	b- 9	10.0	4.0	*
Ξ	146	SHOULD STATE	115,790	23,7100	60,01	57/0
1	1.45		1779/550	moun	11.0	19.99
8	166	PROFESSION AND PERSONS.	TEMP	071787	100	. ***
	100	200 11 5 2 (200)	193046	30.000	166	F-3
4	128	STORE OF THE REAL PROPERTY.	170,000	TARRE	ist	57.00
¥	198	Design Collect	108,586	71,1180	10	335
¥	1-1	MICHAEL MYSER	direction.	disari	14	5100
4	178	Jane 1co, seriosis	19,000	MINN	80	W.O.
*	124	2009 11 9 9 145-00	13396	23,000	14	38.81
13	UK.	2005/197 (0/20)	White.	20,7790	1.3	204





黃色:重金屬超標樣區 (農試所)

紅色:列管農地(環保署)

重金屬超標與列管農地關係



分析

資料分析

- 模型建置的挑戰
- 評量指標的實務意涵

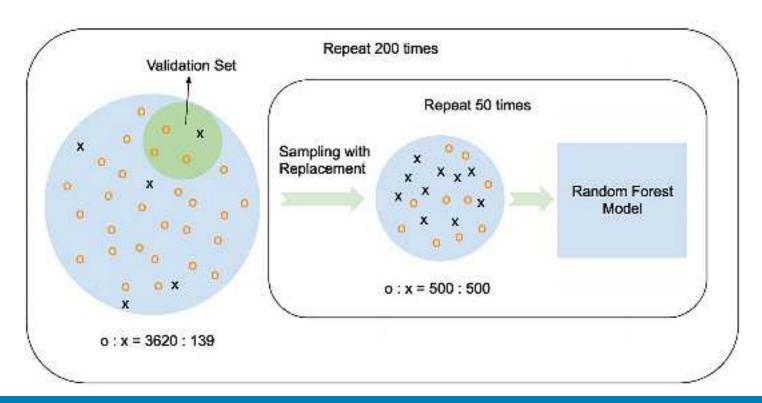


D4SG in Data science for social good conference 2017

非平衡資料

2015年台北市親密關係暴力通報數據

x:一年內受暴者三次以上者,共139筆 o:其餘親密關係暴力案件,共3620筆



評量指標的實務意涵

真實結果

		1	0	50 . 13.25
預 測	1	True positive	False positive, Type I error	$ ext{Precision} = rac{tp}{tp+fp}$
結 果	0	False negative, Type II error	True negative	

$$ext{Recall} = rac{tp}{tp+fn}$$

$$ext{Accuracy} = rac{tp+tn}{tp+tn+fp+fn}$$

Recall rate 特別在意 Type II error, 不能容忍預測不到家暴回頭客的情形。

Precision 在意 Type I error, 不喜歡狼來了的概念,尤其在資源窘迫情況下。

Lessons Learned

Key Challenges

- Connect and empower change agents across domains
- Circle of trust
- Inside-out vs. outside-in
- Data science toolbox
- From proof-of-concept (POC) to standard operation procedure (SOP)

Success ingredients

- Be the enabler, the facilitator and the translator
- "Let's call it an experiment" for on-boarding
- From quick wins to lasting impacts
- Change management



報名網址:https://d4sg.org/fellowship/

報名時間:即日起至2017/12/11

專案時間:2018年 1月至 4月



Why you should join D4SG

- 累積跨領域的實戰經驗
- 幫公共服務性組織開竅
- 可以站上資料科學年會講台XD



非均衡學習建物火災風險模型-以資料英雄計畫「高雄火災風險地圖 2.0」為例

→超速料料學本書 tray 2 11/12 (日) 16:46-1/:30 Pt2

計表語 Toby Liu 是由智序驅動陪託之第三期多料美雄計畫卷與者之一,中人經濟系畢業後,目前是漢大資應用所稱士生,研究主題包含資料探測、語意分析、情經

分析等,研究之論,他還是 書時政府年齡、台灣摩茲 且希望這個世界能藉由資料



跨界的資料英雄專案-「到院前心肺休止」(OHCA)風險地圖

台灣鐵海科學年費 Univ 2 11/12 (日) 12-15-16:00 HO

張維元(維元)#EARINE EXPREZ / 資料工程的 / DASG / EMOC

略畫一點網站前後端開發技術,並且學屬資料采物與機器學習的皮毛。目前是一個在想像與實踐之間追逐菩夢想的來寫工程師。 「透過資料力,應急診改緩運到正向的循環」、一群對資料科學與緊急需學充滿關注的營餘因為被集結,相質包含:力中、培源、宣導、婁平、中類、家場、雅婷、與提案方高雄市衛生周進行合作。 有人是醫學來學生,有人是消防隊緊急救援員志工、透有蓋院急慘仰的關節、從 D4SG 作為一個問始,讓查料科學的種子在學個緊急關實生的前芽。

問問題問問題問問題

還有什麼有趣的發現?

管別人的家務事好嗎?

參與計畫期間最有趣的地方是哪裡?

還能以什麼形式參與社會公益?

誰是危險情人?

113

只會 Excel 可以嗎?

智庫驅動是什麼單位?

我在國外也能參加D4SG嗎?我可以參加D4SG計畫嗎?

我不會分析但我能提案嗎? 我也想學資料分析...

要怎麼報名D4SG計畫?

為什麼不離開?

有國際競爭力嗎? 宅宅不善溝通怎辦...

用R還是Python?

需要具備什麼能力才能參加D4SG?

D4SG的資料科學家是學生還是上班族?

心理系畢業可以讀心嗎?

哪裡可以獲得最新資訊?

你還對什麼議題有興趣?

1999

除了家暴防治之外還有別的題目嗎?

可以Deep Learning 一下嗎?

RRRRRR

什麼是資料科學家?