

動手「毒」理科 - STEM抗毒教育

數據科學



香港青年協會創新科學中心
HKFYG CENTRE FOR CREATIVE
SCIENCE AND TECHNOLOGY





Pre-test



https://bit.ly/DATA_03



熱身遊戲：甚麼是數據？

你認為以下有幾多項是數據的例子呢？

在聊天室輸入數字！



街訪的紙本問卷回應



訪問內容



唱片的尺寸



隱蔽吸毒個案統計數字



影視平台瀏覽紀錄



飲食平台的評論

什麼是數據

- 數據是一組通過觀測得到的特徵或資訊。 *OECD Glossary of Statistical Terms. OECD. 2008. p. 119. [ISBN 978-92-64-025561](#).*

Numerical data

Pizza的重量：
0-???(克)

辣肉腸的數量：
0-???(塊)



Categorical data

Pizza的成份：
麵包、芝士、辣肉腸、香料

味道：
鹹、辣、甜、酸、苦

食客對食物滿意：
非常同意、同意、不同意、
非常不同意

收集數據(data)就是數據科學嗎？

什麼是數據科學

- 數據科學的目標是通過從資料中提取出有價值的部分。



| | |
|----|----------|
| 地區 | 荃灣 |
| 價錢 | \$ \$ \$ |
| 座位 | 5 |



| | |
|----|-------|
| 地區 | 西貢 |
| 價錢 | \$ \$ |
| 座位 | 9 0 |



| | |
|----|-----|
| 地區 | 南區 |
| 價錢 | \$ |
| 座位 | 5 0 |

歸納後的結果可協助:

- ✓ 判斷
- ✓ 分類
- ✓ 預測

數據的重要



4.5 out of 5 stars



4 reviews

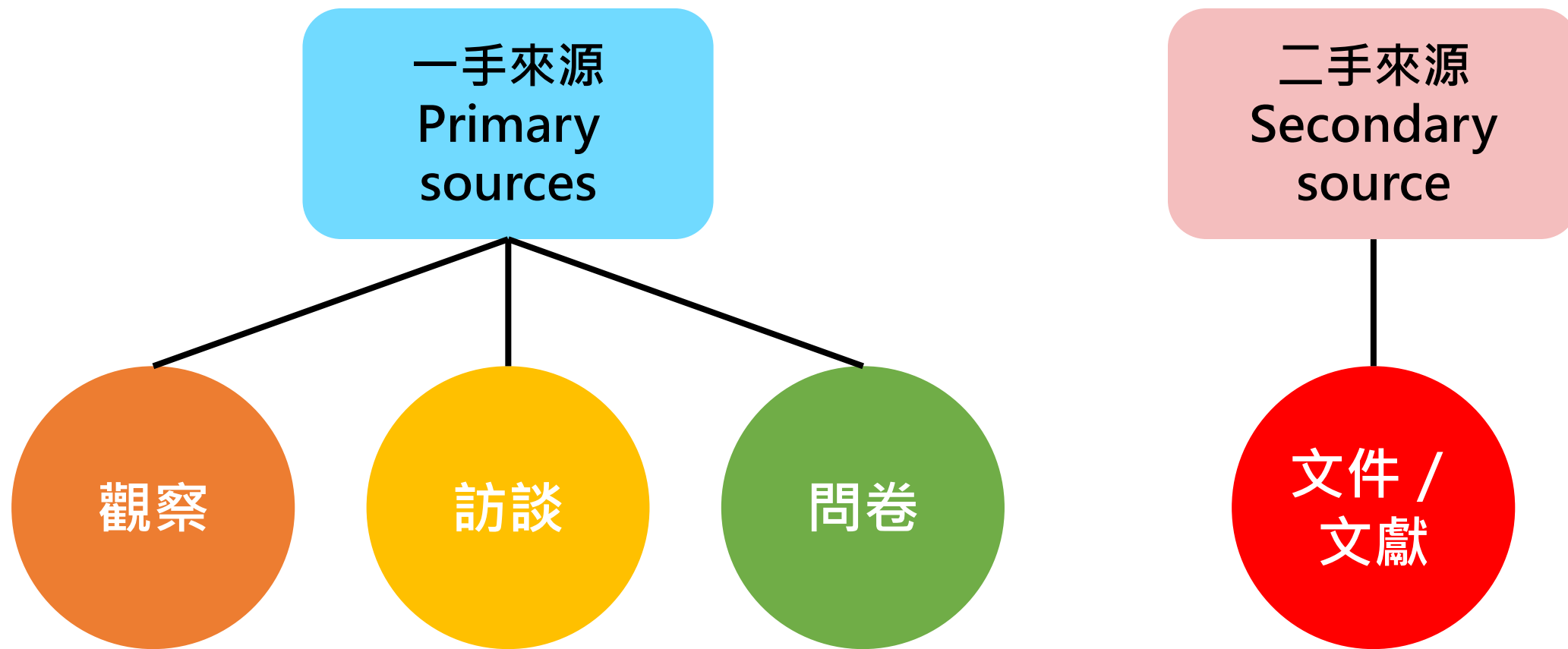


4.2 out of 5 stars



2647 reviews

第1步: 收集數據



快速問卷調查

1. 你的身高是？
2. 你的體重是？
3. 過去一周做過運動的次數？
4. 香港男團MIRROR中你最喜歡哪一個成員？
5. 你覺得現時正在吸毒的人仕（21歲以下）選擇繼續吸毒的主要原因是甚麼？
7. 你覺得毒品大麻是全天然的植物嗎？
8. 你認為自己多大程度了解吸食大麻對大腦的影響？



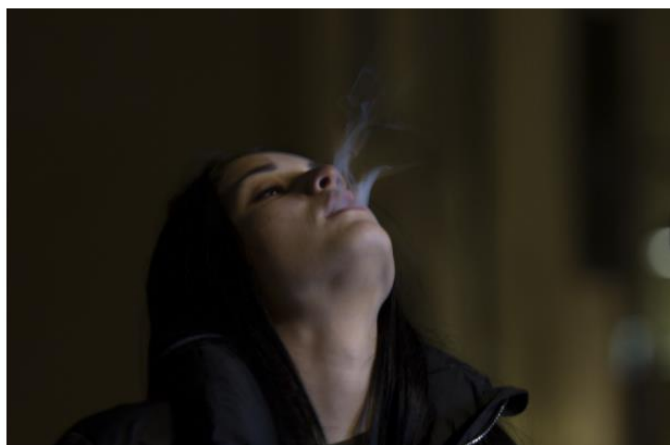
https://bit.ly/BDF_survey

網上平台對吸毒數字的影響

上述分析是偵測本港由2016年至今年6月30日，逾30萬篇毒品相關的帖文及留言。研究發現，近年社交媒體涉及毒品相關的帖文數目及瀏覽量大幅上升。帖文數目由2016年的1,001篇，上升至今年中的5,237篇，上升比率為423%；估算瀏覽量亦由2016年的440萬次，上升至今年中的1,734萬次，上升比率為298%。最多人討論及持正面態度的毒品首四位是大麻、啟靈藥（LSD）、迷姦水（GHB）和白瓜子。

Artificial Intelligence Can Now Use Instagram to Predict Substance Abuse Risk

A paper published recently developed a method of predicting someone's risk of substance abuse based on their Instagram accounts ([link](#)). Let's take a look at how they did it.



Using Artificial Intelligence to Change the Future of Drug Addiction Treatment & Narcotics Detection

Mangalore Today News Network

By Dr. G. Shreekumar Menon

社交平台上有什麼數據？

貼文內容
(文字及emojis)

圖片上的文字

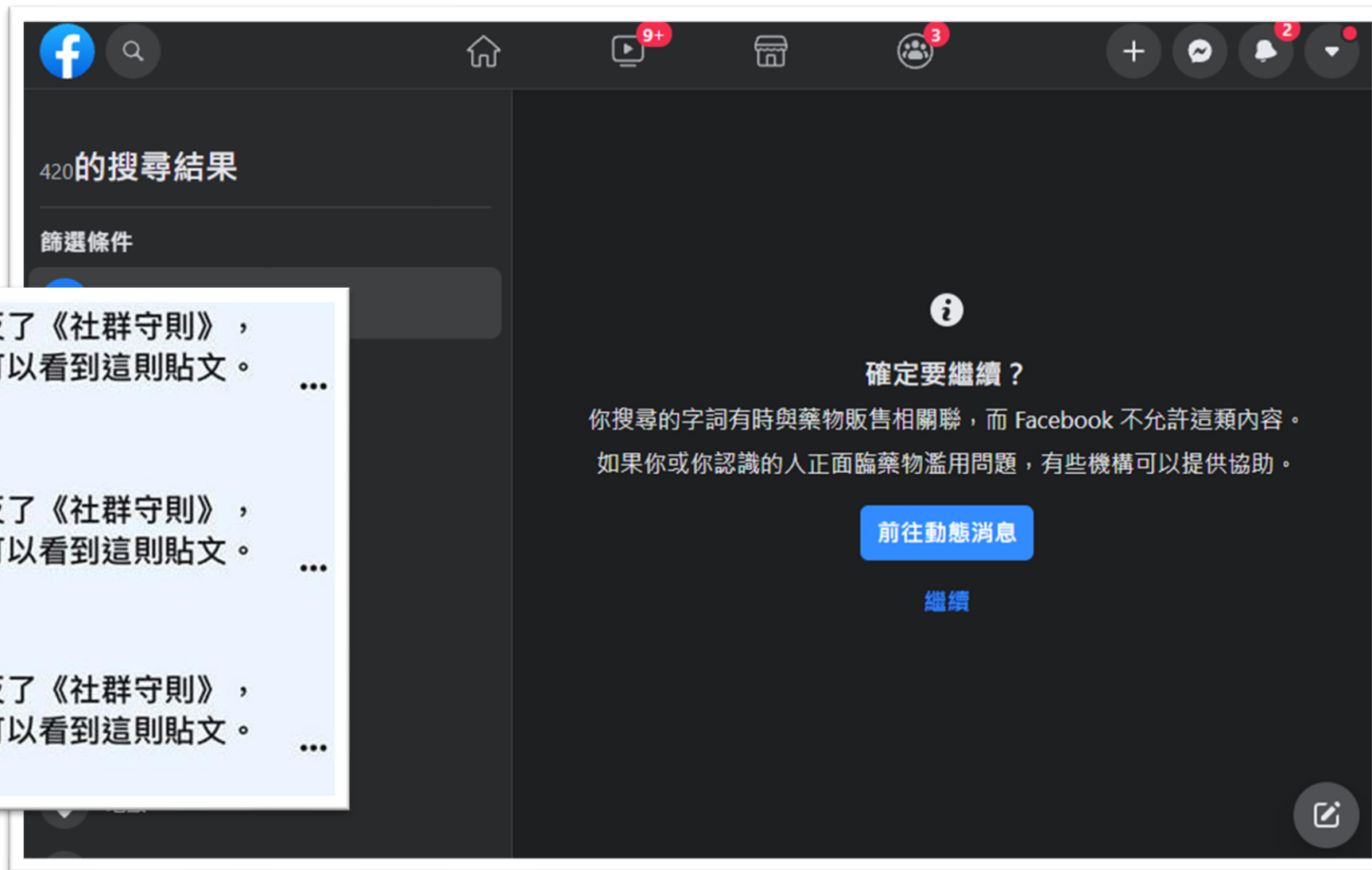
反應 Reaction
(如讚好、嬲嬲)



發貼/留言/反應用戶

圖片

留言



在社交媒體上有遇到這些情況嗎？
系統如何分析貼文是否與毒品有關？

圖片的文字



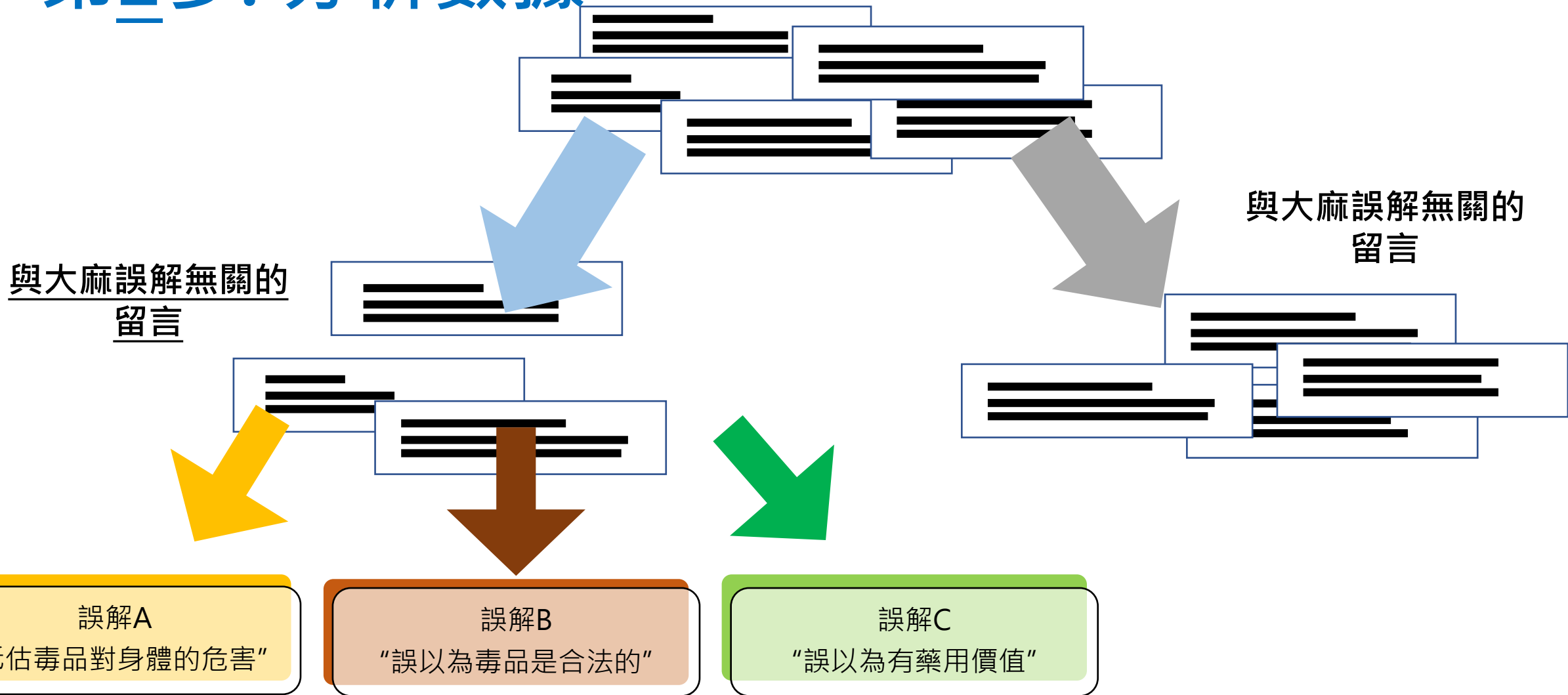
好東西需要的懂的加賴

貼文



圖片內容

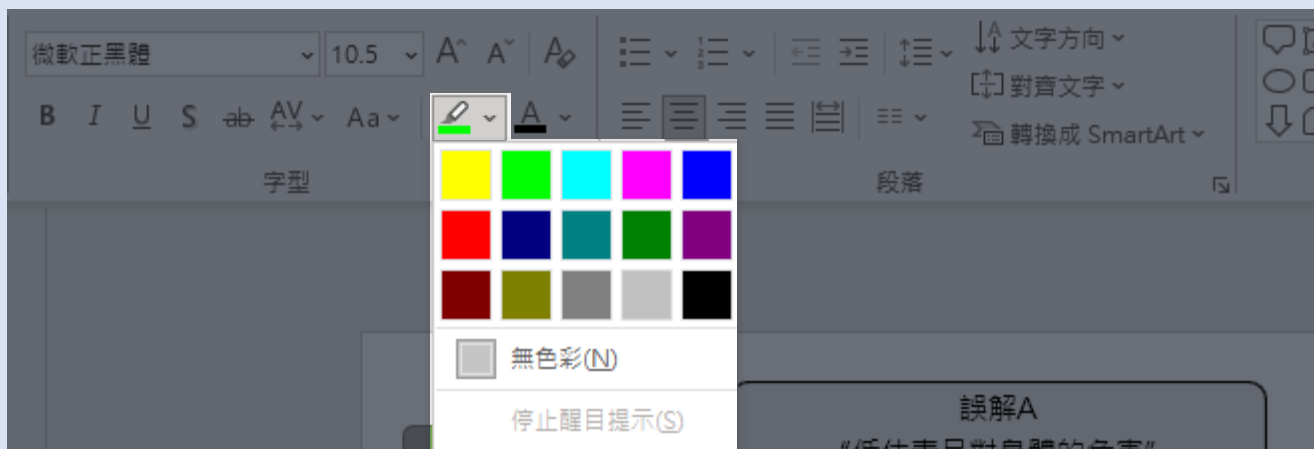
第2步: 分析數據



目標：分辨留言是否與大麻誤解有關

目標：分辨留言是否與大麻誤解有關

(1) 使用文字醒目提示色彩，將大麻誤解(A, B, C)標記到第 2 頁的留言上



誤解A
“低估毒品對身體的危害”



誤解B
“誤以為毒品是合法的”



誤解C
“誤以為有藥用價值”

(2) 將已標記的留言分類到第 3 頁上

| | |
|---|---|
| <p>誤解C “誤以為有藥用價值”</p> <p>大麻 YYYYYYYY · XXXXXXXXXXXX ·</p> | <p>誤解B “誤以為毒品是合法的”</p> <p>大麻XXXX · XXXX XXXXXXXX XXXXXXXX</p> |
|---|---|

例子

誤解A
“低估毒品對身體的危害”

誤解B
“誤以為毒品是合法的”

誤解C
“誤以為有藥用價值”

大麻有很多品種，不可一概而論。以色列甚至培育出抗癌的大麻品種，只是我們拿不到而已。

成日話smoke會搞出人命，明明草要15分鐘內隊680 kg 先會死人

XXX啲歌大部分都係stone住作

日日發夢 醒返會覺得個人好廢

發夢出海釣魚stone，四面還海，隨著海浪一高一底

唔好講啲似是而非嘅理由，大麻一定係毒品！

大麻是一種藥草、農產品、安全、天然、沒問題

大麻都唔會上癮，我返工放break都會smoke，啲人控制唔到就話毒品，冇冇了解過架

隊草係搵靈感佳品

發夢夢見stone完成日係度諗個宇宙點黎宇宙以外到底有咩空間

大麻對抑制痙攣也是非常非常重要喔

食唔係問題，但唔好上癮/倚賴/日日食囉

草係毒品咩？草係植物，草成份天然，有益身心。

咳藥水、大麻以至以前流行嘅忽得，都有藥用嘅價值，問題係有無被濫用啫

我樓上有外國人都成日隊草，物管去到就話聞唔到 唔斷正好難入罪

咳藥水、大麻以至以前流行嘅忽得，都有藥用嘅價值，問題係有無被濫用啫

今天是420國際大麻日Smoke weed every day!

相比其它同被定性為毒品嘅substances，大麻嘅毒性相對較低而且對人體造成嘅長遠健康影響相對少.....

吸食唔犯法 藏有先犯法

論毒食煙飲酒比吸大麻更傷更易上癮，我聽過人食煙飲酒食cocaine而死，但係就好少聽過人話食大麻而死

大麻是沒有副作用且十分安全，幾乎不會為人體造成負面的影響，即便發生、也是以十分溫和的方式發生。

去旅行吸過一次，其實都冇咩問題，過一排就好返

前排keep左兩個月，晚晚smoke，冇啲人講話咁多問題囉

Smoke完可以令我第二日精力充沛

有d人比保安見到 都係叫佢掉左佢算 未見過比人拉

根本冇人理你，周街都有得賣啦，話犯法個啲有冇出過街

CBD是一種營養價值極高、無精神活性的植物營養素，有紓緩和治療鎮痛、減輕壓力，緩解焦慮和改善睡眠等等的作用。

草ok的，無煙咁傷身

與大麻誤解有關的留言

與大麻誤解無關的留言

誤解A

“低估毒品對身體的危害”

誤解B

“誤以為毒品是合法的”

誤解C

“誤以為有藥用價值”

與誤解有關但未能分類

誤解A

“低估毒品對身體的危害”

誤解B

“誤以為毒品是合法的”

誤解C

“誤以為有藥用價值”

與濫毒有關的留言中，
有什麼常見的關鍵詞？

誤解D

誤解E

有其他新發現嗎？

展示數據的方法: 文字雲 (Word Cloud)

- <https://wordcloud.timdream.org/>

The image shows a screenshot of the 'HTML5 文字雲' (HTML5 Word Cloud) web application. The interface includes a title bar with 'HTML5 文字雲' and a '關於' (About) button. On the left, there are links for '本範例' (This Example), '複製/貼上文字' (Copy/Paste Text), '檔案' (Files), and '維基百科' (Wikipedia). The main area is a large text input field labeled '文字' (Text). Below the input field, a note states '您貼上的文字會直接處理，不會被上傳。' (The text you paste will be processed directly and will not be uploaded). At the bottom, there is a language dropdown menu set to '正體中文' (Traditional Chinese) and a '開始' (Start) button. Two orange callout boxes provide instructions: one on the left says '1. 貼上文字' (1. Paste text) with an arrow pointing to the text input field; another on the right says '2. 滑鼠按右鍵選擇貼上 或鍵盤按Ctrl+V' (2. Right-click mouse to choose paste or press Ctrl+V on keyboard) with arrows pointing to the text input field and the '開始' button.

1. 貼上文字

HTML5 文字雲

關於

本範例

複製/貼上文字

檔案

維基百科

文字

您貼上的文字會直接處理，不會被上傳。

正體中文

開始

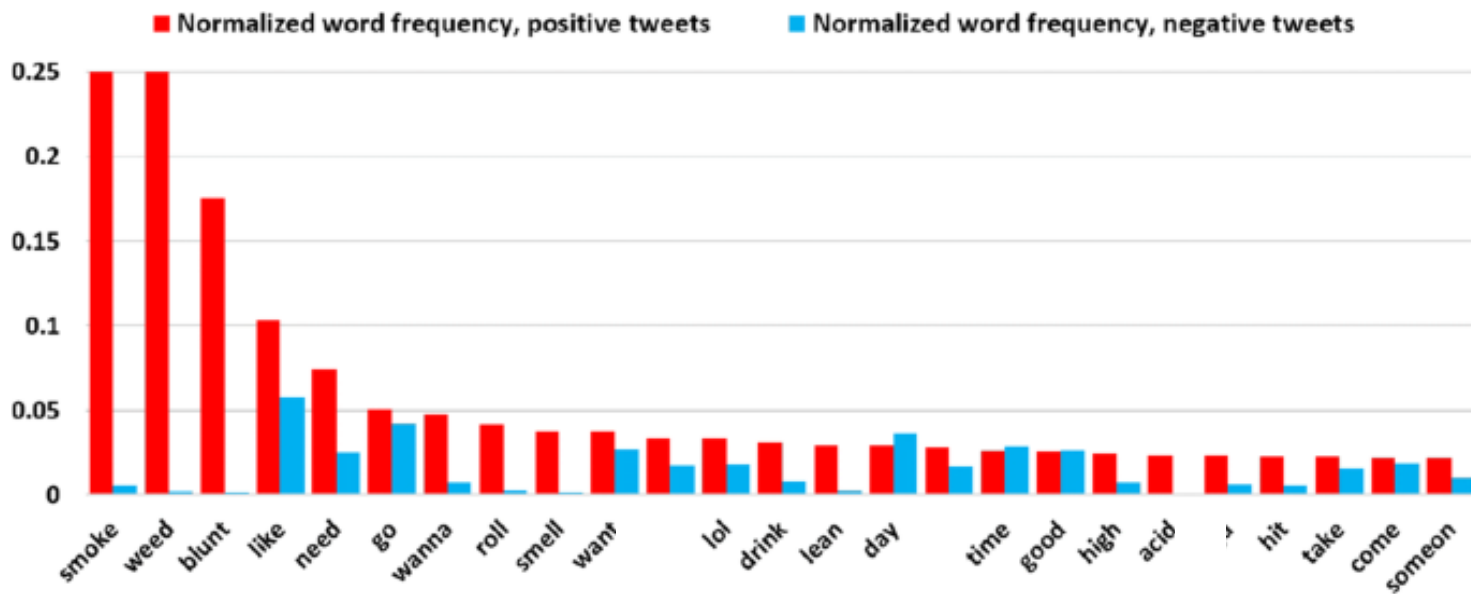
2. 滑鼠按右鍵選擇貼上
或鍵盤按Ctrl+V



a Positive tweets

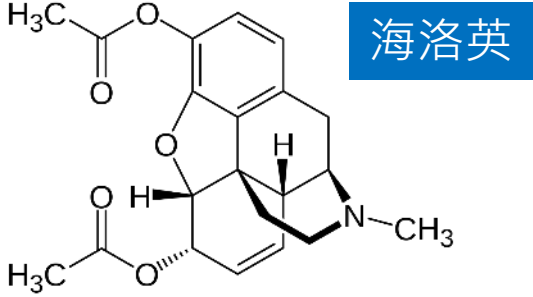
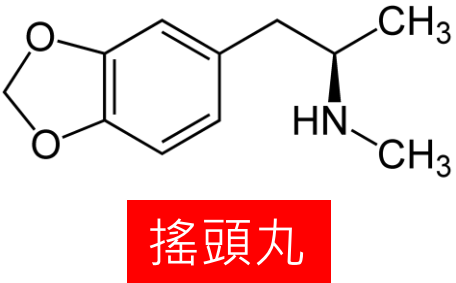
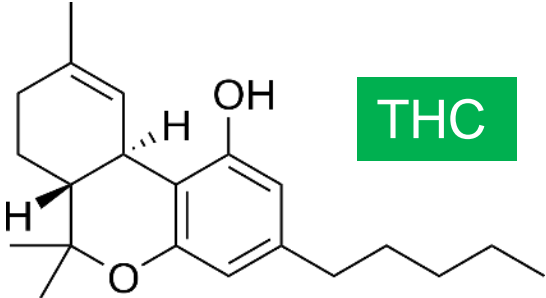


b Negative tweets



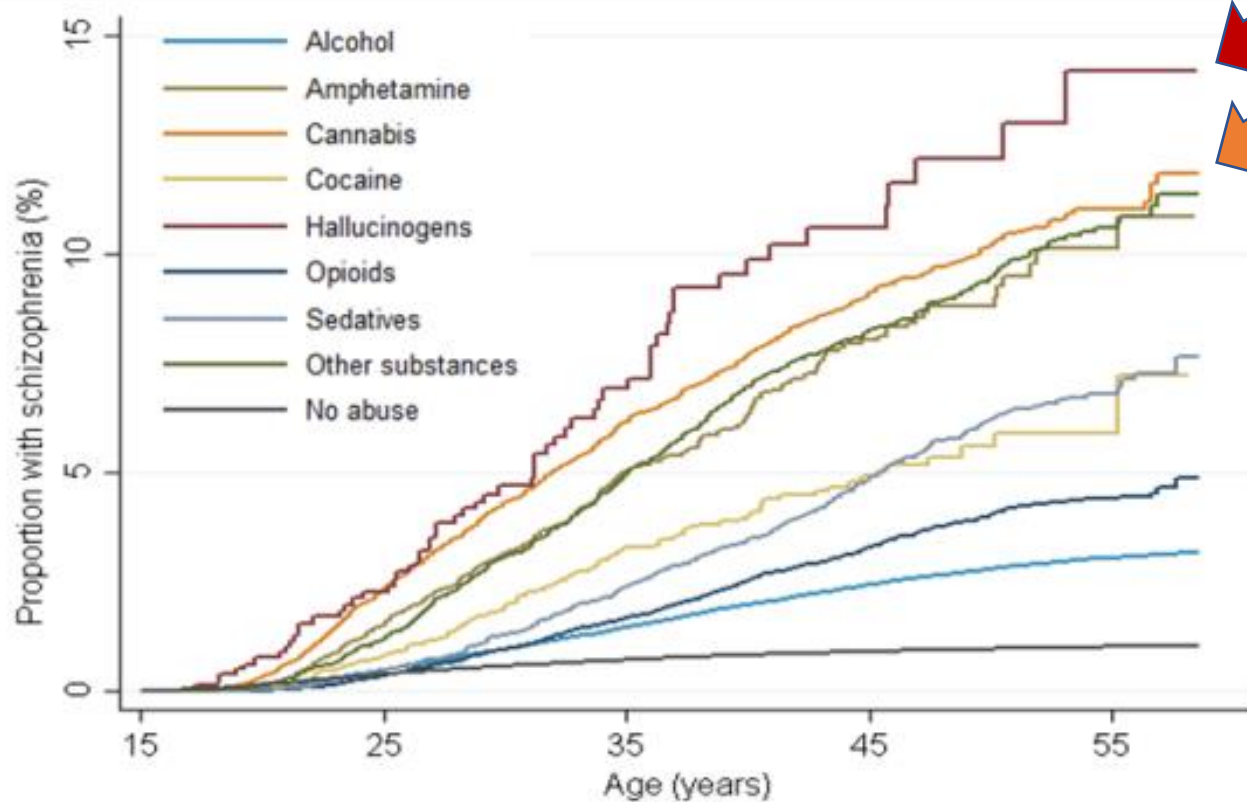
An insight analysis and detection of drug-abuse risk behavior on Twitter with self-taught deep learning

誤解A：三大類毒品對身體的影響

| 鎮抑 / 靜劑 | 興奮劑 | 迷幻劑 |
|--|--|--|
|  <p>海洛英</p> |  <p>搖頭丸</p> |  <p>THC</p> |
| 反應遲緩 | 失眠 | 幻覺 |
| 呼吸困難 | 抽筋 | 抑鬱或暴躁 |
| 失去知覺 | 心臟衰竭 | 智力 / 判斷力下降 |

誤解A：吸食大麻會對腦部造成的影響

Figure 1: Risk of Schizophrenia for All Substance Abuse Disorders



誤解B：“誤以為毒品是合法的”



誤解C “誤以為有藥用價值”



1. 在香港擁有THC會屬刑事罪行
2. 有迷幻的感覺
3. 影響行為和認知能力
4. 長期使用能導致幻聽或精神分裂



1. 在香港擁有CBD不違法，但..
2. 聲稱能有助舒緩痛症及情緒，但暫時未有完善的科學實證
3. 美國食品藥品監督管理局只批准過一種含有CBD嘅藥物，並只使用於癲癇症的病人身上
4. 提煉時仍有機會有THC殘留，並會受THC的影響或觸犯法例

警化驗大麻二酚產品 3款含受規管四 氫大麻酚 拘8人 (20:13)

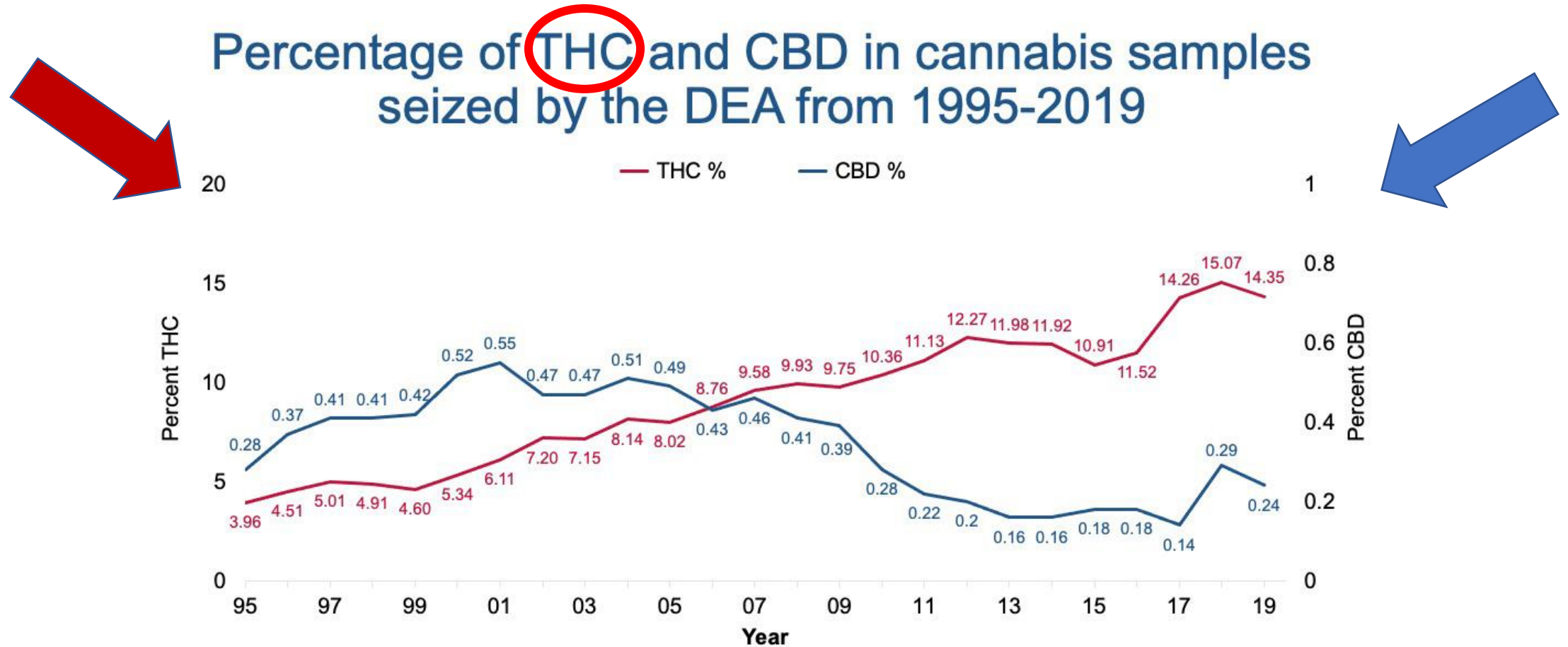
    Like 1

A+ A-    



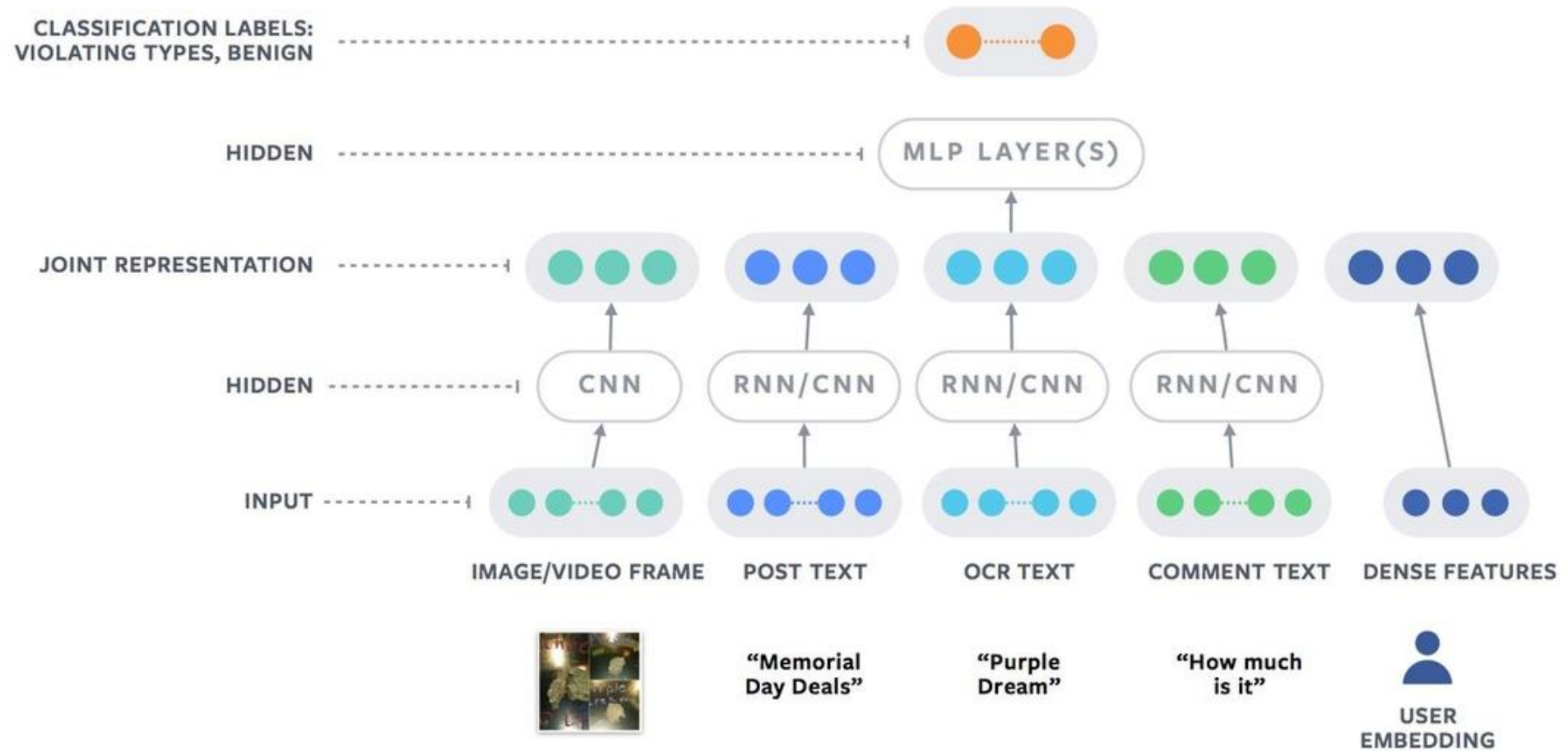
圖3之1-警方展示檢獲證物。(伍浦鋒攝)

誤解C：“誤以為有藥用價值” 你覺得毒品大麻是全天然的植物嗎？



根據美國緝毒局(DEA) 按年分析非法大麻中的THC及CBD的含量
→ 人為因素令THC 含量顯著上升

MULTIMODAL & MULTICHANNEL NETWORK





問卷調查

數據的分類

| | 數值數據 / 分類數據 | 連續 / 離散 定類 / 定序 |
|---|----------------|--------------------|
| 1. 你的身高是？ | | |
| 2. 你的體重是？ | | |
| 3. 過去一周做過運動的次數？ | | |
| 4. 香港男團MIRROR中你最喜歡哪一個成員？ | | |
| 5. 你覺得現時正在吸毒的人仕（21歲以下） 選擇繼續吸毒的主要原因是甚麼？ | | |
| 7. 你覺得毒品大麻是全天然的植物嗎？ | | |
| 8. 你認為自己多大程度了解吸食大麻對大腦 的影響？ | | |

數據的分類

| | 數值數據 / 分類數據 | 連續 / 離散 定類 / 定序 |
|---|----------------|--------------------|
| 1. 你的身高是？ | 數值 | |
| 2. 你的體重是？ | 數值 | |
| 3. 過去一周做過運動的 次數 ？ | 數值 | |
| 4. 香港男團MIRROR中你最喜歡哪一個成員？ | 分類 | |
| 5. 你覺得現時正在吸毒的人仕（21歲以下） 選擇繼續吸毒的主要原因是甚麼？ | 分類 | |
| 6. 你覺得毒品大麻是全天然的植物嗎？ | 分類 | |
| 7. 你認為自己 多大程度 了解吸食大麻對大腦 的影響？ | 分類 | |

數據的分類

| | 數值數據 / 分類數據 | 連續 / 離散 定類 / 定序 |
|---|----------------|--------------------|
| 1. 你的身高是？ | 數值 | 連續 Continuous |
| 2. 你的體重是？ | 數值 | 連續 Continuous |
| 3. 過去一周做過運動的 次數 ？ | 數值 | 離散 Discrete |
| 4. 香港男團MIRROR中你最喜歡哪一個成員？ | 分類 | 定類 Nominal |
| 5. 你覺得現時正在吸毒的人仕（21歲以下） 選擇繼續吸毒的主要原因是甚麼？ | 分類 | 定類 Nominal |
| 6. 你覺得毒品大麻是全天然的植物嗎？ | 分類 | 定類 Nominal |
| 7. 你認為自己 多大程度 了解吸食大麻對大腦 的影響？ | 分類 | 定序 Ordinal |

第3步: 分析及展示問卷調查結果 (第1及2題)

數值數據 (Numerical data)

身高

- Mean (平均數) : =AVERAGE
- Median (中位數) : =MEDIAN
- Mode (眾數) : =MODE

體重

- Mean (平均數) : =AVERAGE
- Median (中位數) : =MEDIAN
- Mode (眾數) : =MODE

https://bit.ly/BDF_datasets



檔案 編輯 查看 插入 格式 資料 工具 擴充功能 說明

1 新文件
Ctrl+O
開啟
匯入
建立副本
共用
電子郵件
下載
重新命名
移至垃圾桶
版本記錄
詳細資料
設定
列印
Ctrl+P

| | C | D |
|----|----------|-------------|
| 1 | 體重 (kg)? | 過去一周做運動的次數? |
| 2 | 73 | 6 |
| 3 | 53 | 2 |
| 4 | 57 | 0 |
| 5 | | |
| 6 | | |
| 7 | | |
| 8 | | |
| 9 | | |
| 10 | | |
| 11 | | |
| 12 | | |
| 13 | | |
| 14 | 70 | 4 |
| 15 | 70 | 10 |
| 16 | 80 | 1 |

檔案-下載-Excel

Mean

- 反映整個數據集，但容易受極端數影響
- 操場上有6歲、7歲、8歲、9歲、70歲的人共五人。其年齡平均數為**20歲**

Median

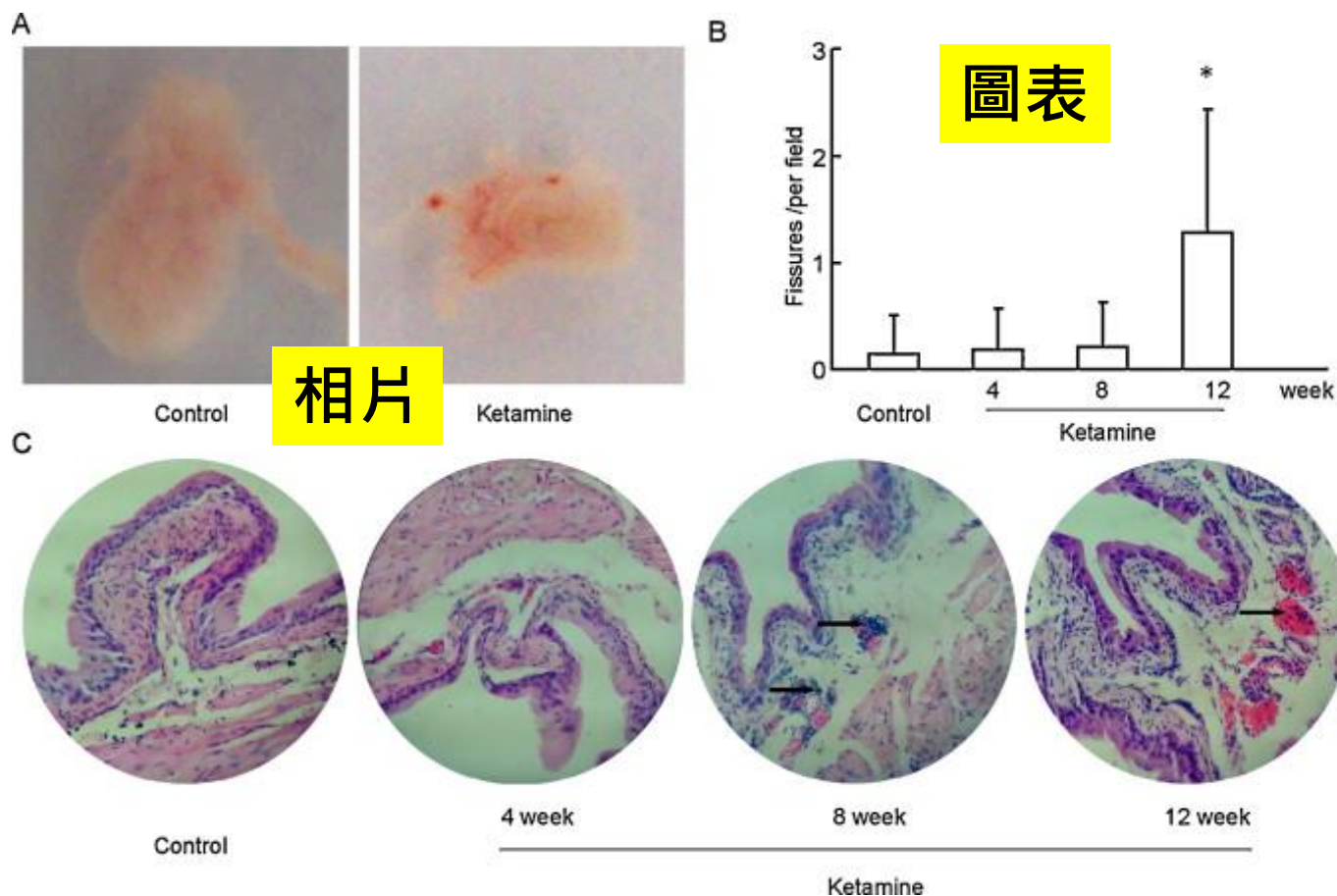
- 極端數對中位數影響較少，但**未能反映所有數據**
- 操場上有6歲、7歲、8歲、9歲、70歲的人共五人。其年齡中位數為**8歲**

Mode

- 眾數指一組數據中**出現次數最多**的變數值，適合於數據量較多時使用。
- 例如分析超級市場賣出最多的水果：
橙10次、蘋果5次、梨8次，眾數是**橙**

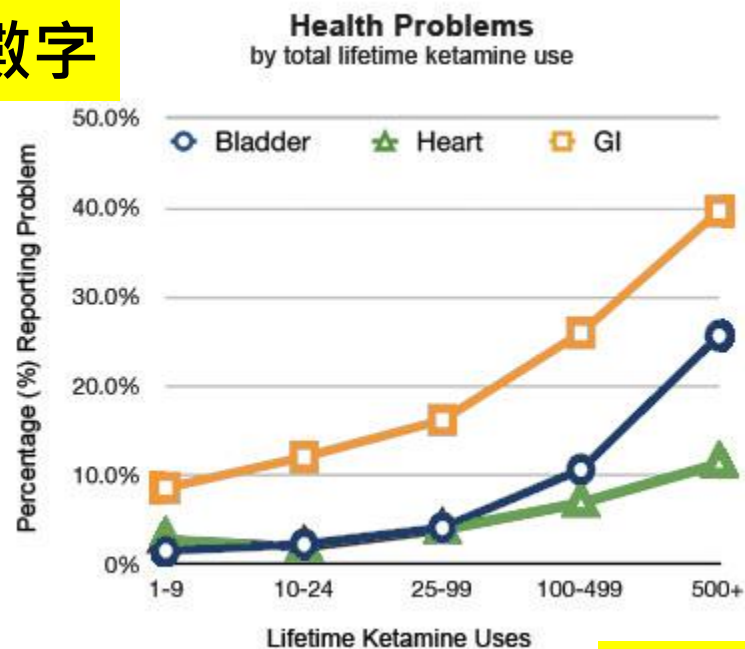
第3步: 數據形象化 (實驗 / 醫學數據)

實驗：氯胺酮對膀胱的影響



醫學：吸食氯胺酮次數與對不同身體部份的影響

數字



圖表

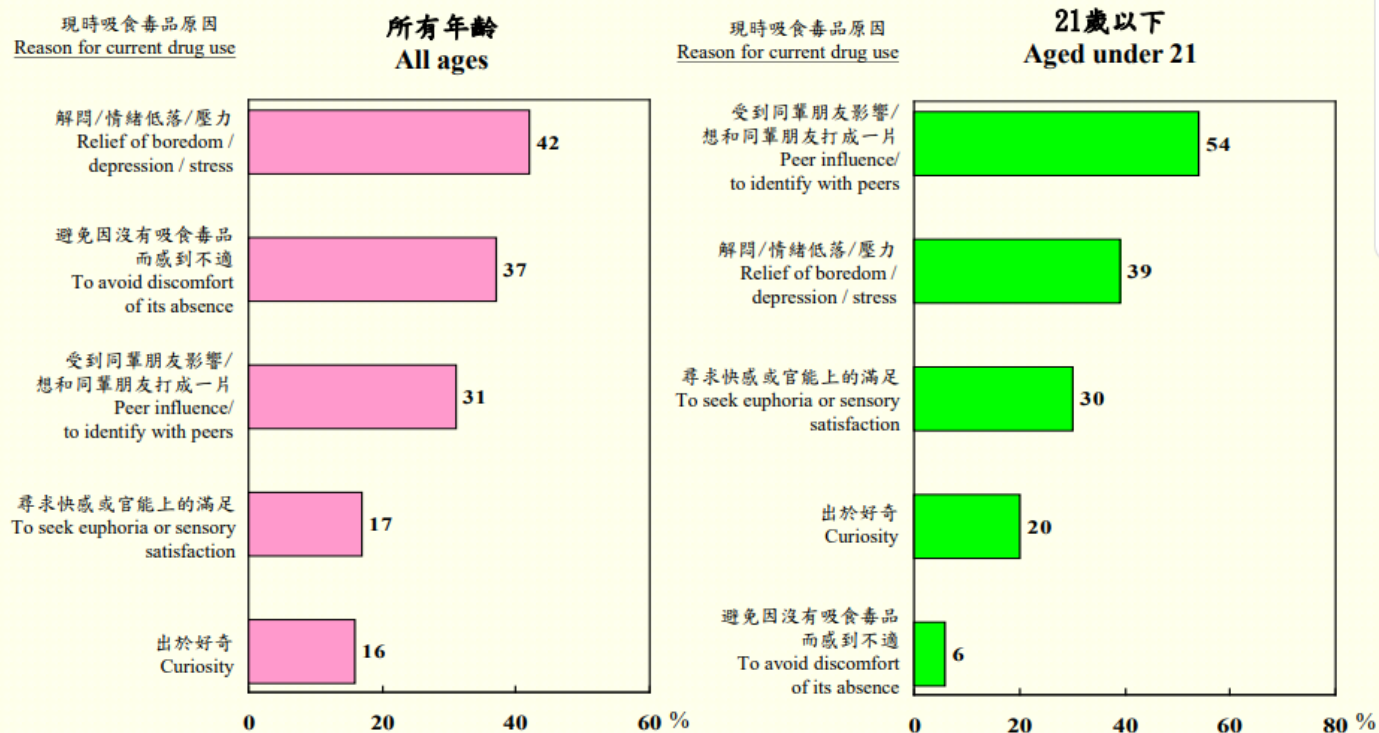
6. 你覺得一個人為何會選擇吸毒？

Excel

- =countif(range,criteria)
- 加上 \$
- 由細到大
- 插入圖表 (橫置棒形圖)

6. 你覺得一個人為何會選擇吸毒？

按現時吸食毒品原因劃分的2021年上半年被呈報吸食毒品人士
Reported drug abusers by age group by reason
for current drug use in 1st half 2021



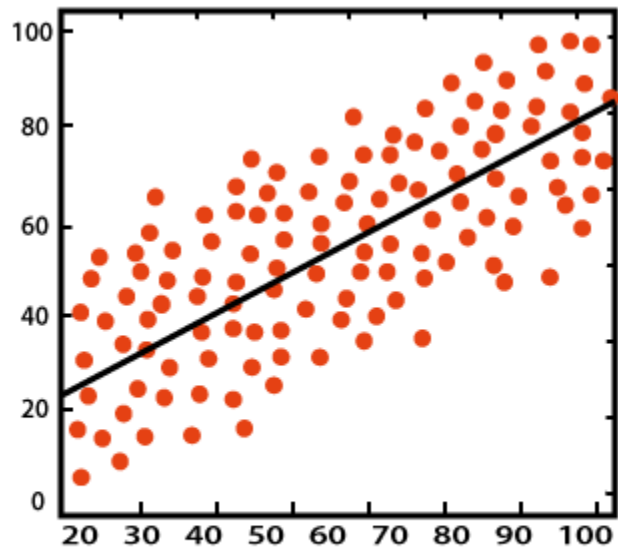
註釋：同一被呈報吸食毒品者可被呈報多於一個現時吸食毒品原因。

Note : More than one reason for current drug use may be reported for each individual reported drug abuser.

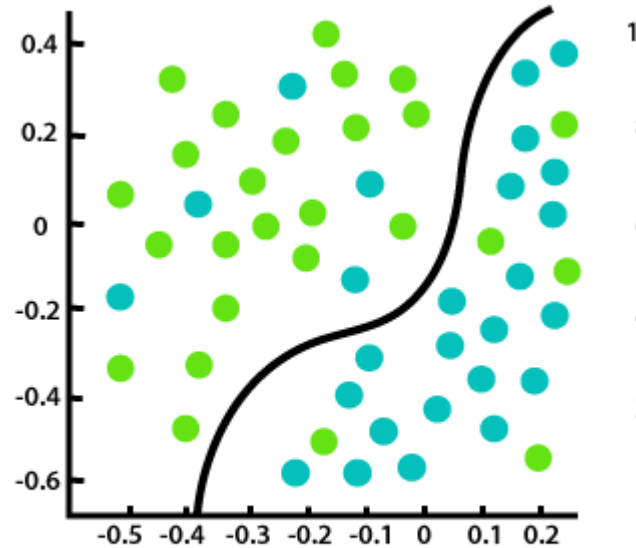
資料來源：藥物濫用資料中央檔案室
Source : Central Registry of Drug Abuse
資料更新於 16.9.2021
Updated on 16.9.2021

第3步:資料視覺化及運用 (機械學習)

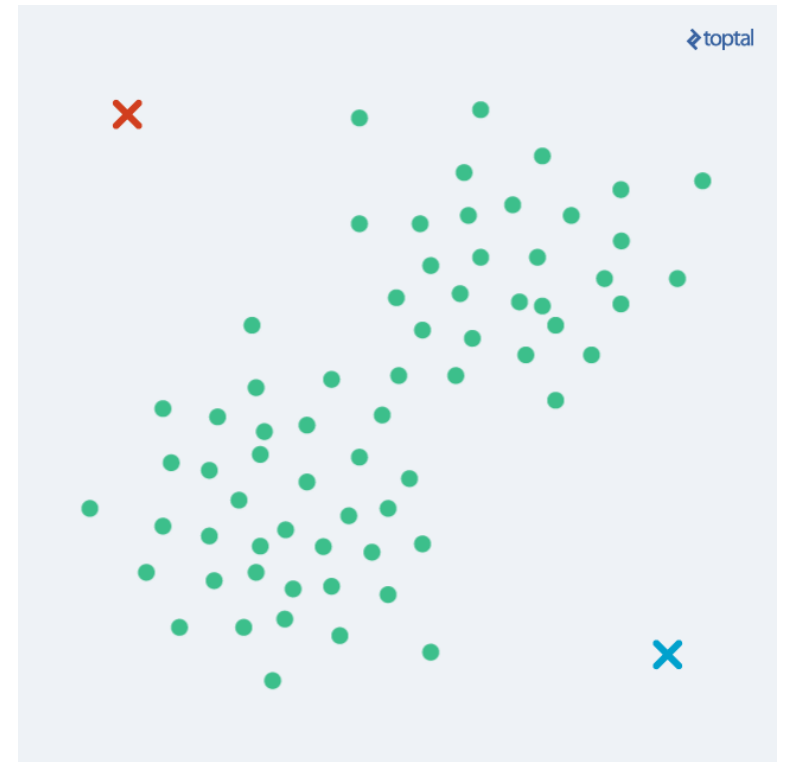
- 學習由數據點製成的模型model
- 再利用這個模型去預測下一個新的數據



Regression



Classification

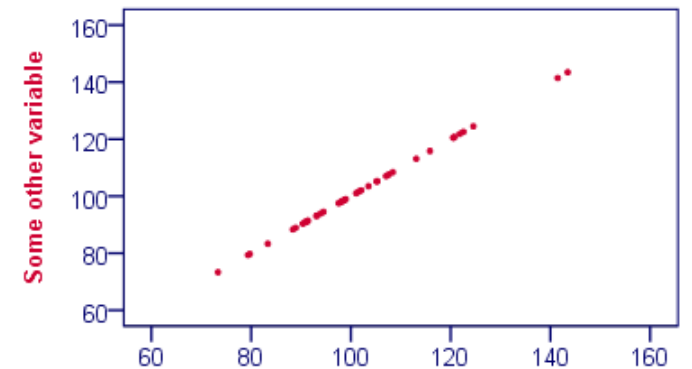


透過分析數據展示兩者的關係

- 利用散佈圖去展示每個個別的數據
- 在散佈圖上畫上線性回歸線理解兩者的關係
- 例子：問卷調查結果（第1及2題）
 - 插入圖表（散佈圖）
 - 加上趨勢線
 - 顯示R平方值



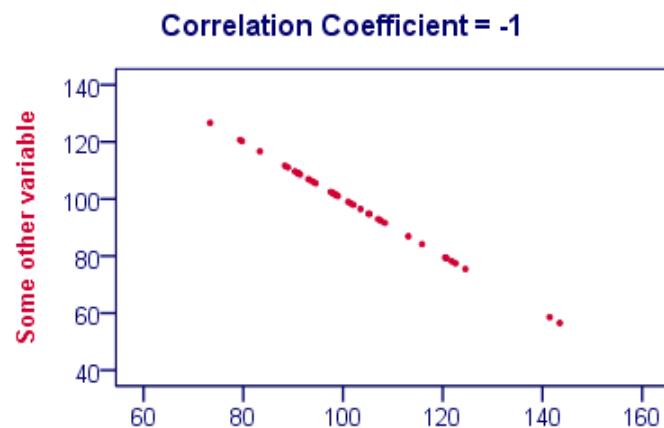
總播放時間



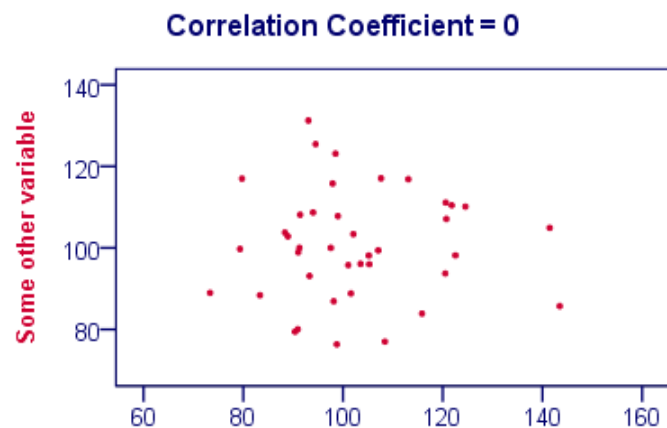
唱片大小

利用線性回歸理解兩者的關係

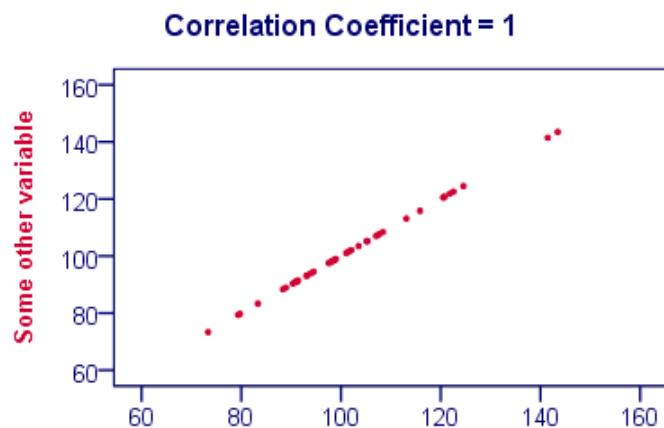
- 判定系數 (R^2) 由-1至1



強烈負相關的關係

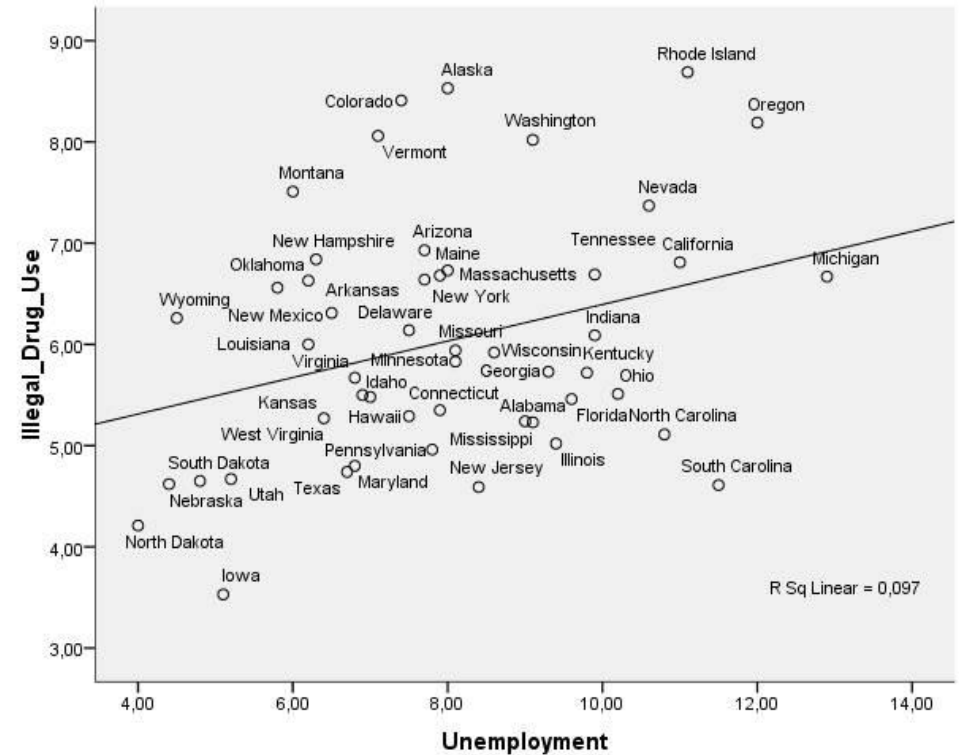
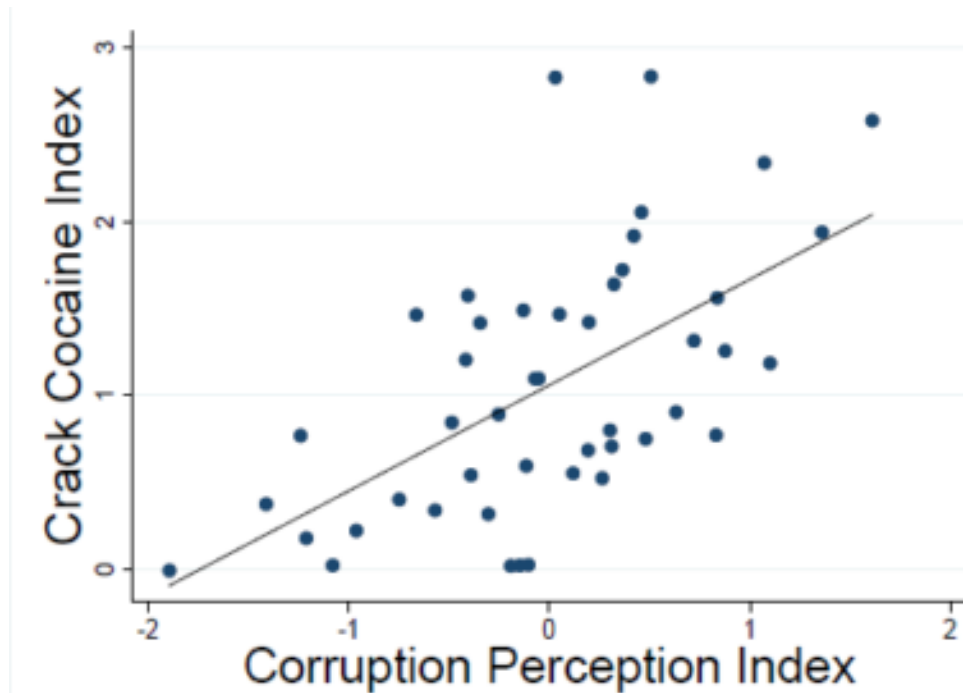


沒有明顯的關係



強烈正相關的關係

第4步: 運用數據模型 利用趨勢線作為預計下一個新數據的參考



第4步: 運用數據模型

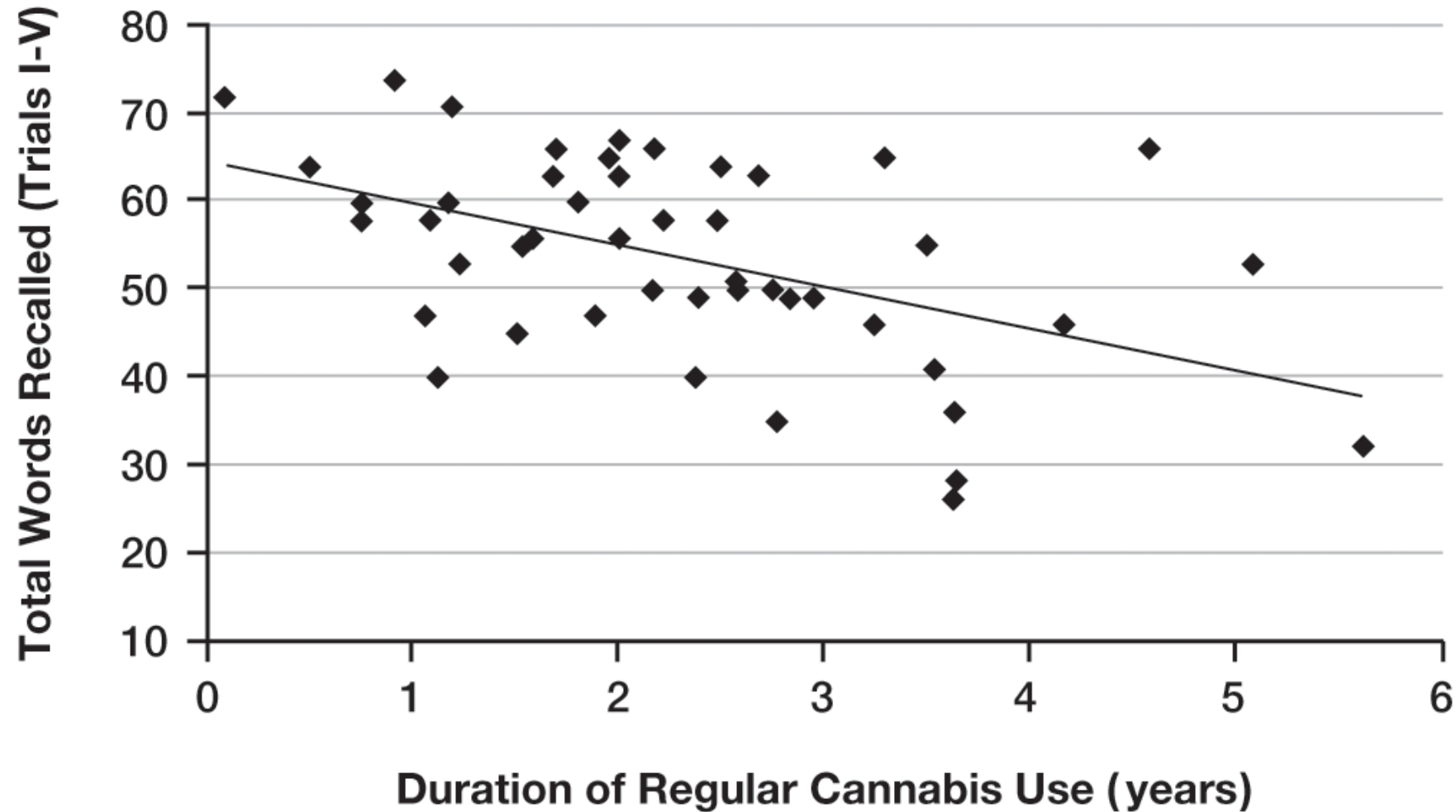
你覺得吸食大麻會對腦部造成影響嗎？(2)

- 利用線性回歸分析吸食大麻的年資及記憶力的關係
- 47個曾經吸食大麻的青少年
- 吸食大麻的年資及能夠記起文字的數目之關係

| No. | Codename | Duration of Regular Cannabis (Months) | Total Words Recalled (Trials I-V) |
|-----|----------|---------------------------------------|-----------------------------------|
| 1 | AA | 1 | 72 |
| 2 | AB | 7 | 64 |
| 3 | AC | 10 | 60 |
| 4 | AD | 10 | 57 |
| 5 | AE | 12 | 74 |
| 6 | AF | 15 | 71 |
| 7 | AG | 13 | 59 |
| 8 | AH | 14 | 61 |
| 9 | AI | 15 | 53 |
| 10 | AJ | 13 | 47 |
| 11 | AK | 14 | 40 |
| 12 | AL | 19 | 55 |
| 13 | AM | 20 | 56 |
| 14 | AN | 19 | 46 |
| 15 | AO | 22 | 65 |
| 16 | AP | 22 | 63 |
| 17 | AQ | 23 | 59 |
| 18 | AR | 23 | 57 |
| 19 | AS | 24 | 65 |
| 20 | AT | 25 | 67 |
| 21 | AU | 25 | 63 |
| 22 | AV | 25 | 57 |

| | | | |
|----|----|----|----|
| 25 | AY | 26 | 50 |
| 26 | AZ | 29 | 49 |
| 27 | BA | 28 | 40 |
| 28 | BB | 30 | 63 |
| 29 | BC | 30 | 58 |
| 30 | BD | 31 | 50 |
| 31 | BE | 31 | 51 |
| 32 | BF | 33 | 63 |
| 33 | BG | 34 | 36 |
| 34 | BH | 34 | 51 |
| 35 | BI | 35 | 50 |
| 36 | BJ | 35 | 50 |
| 37 | BK | 40 | 65 |
| 38 | BL | 43 | 55 |
| 39 | BM | 39 | 48 |
| 40 | BN | 43 | 42 |
| 41 | BO | 44 | 36 |
| 42 | BP | 44 | 28 |
| 43 | BQ | 43 | 26 |
| 44 | BR | 50 | 46 |
| 45 | BS | 56 | 66 |
| 46 | BT | 61 | 51 |
| 47 | BU | 69 | 31 |

利用線性回歸分析吸食大麻的年資及記憶力的關係



有關毒品的邀請

- 在哪裡？
- 有誰會邀請你？
- 他們會說甚麼說話去邀請？

食啦~係咪friend先？

唔係咁冇膽呀？

大麻係天然架嘛，好似食菜咁，
邊會咁易有事嘞？

食少少，最多咪瞓覺怕咩啫？
我食咁耐都冇事。

拒絕的方法

1. 堅持拒絕法

2. 告知理由法

3. 友誼勸服法

6. 遠離現場法

5. 轉移話題法

4. 自我解嘲法

拒絕的方法及技巧

1. 堅持拒絕法

2. 告知理由法

3. 友誼勸服法

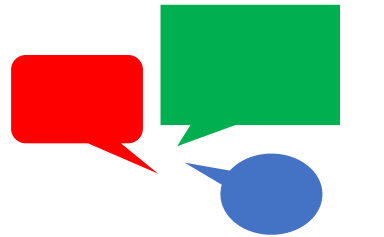
6. 遠離現場法

5. 轉移話題法

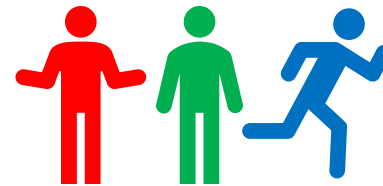
4. 自我解嘲法



眼神、表情

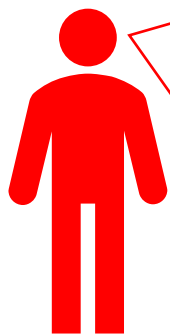


聲線、語氣



身體語言、
距離

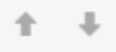
數據的重要性



• 剛剛

顯示 #1

食少少，最多咪瞓覺怕咩啫？我食咁耐都冇事。

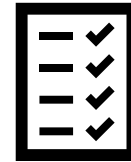


數據真確性及代表性





Post-test



https://bit.ly/DATA_03

保安局禁毒處

求助／諮詢



186 186



98 186 186



唔take嘢
narcotics.divisionhk



唔take嘢
narcotics.divisionhk



香港青年協會 創新科學中心
HKFYG CENTRE FOR CREATIVE
SCIENCE AND TECHNOLOGY

