



# NCCU DA

政大數據分析社

## 2.0 儲備幹部數據專案

台灣發票資料分析

WELCOME



# 數據專案 - 題目說明

## 台灣發票資料

---



taiwan\_invoice\_data.csv

■ **9 columns**

■ **65549 rows**

資料前處理

解決問題

# 資料輪廓 - 台灣發票資料

發票年月	縣市	縣市代號	行業別	載具類別	電子發票張數	電子發票金額	年	月
2018/1/1	臺北市	A	零售業	無載具	58357391	29622747754	2018	1
2018/1/1	臺北市	A	批發業	無載具	4915556	24494764143	2018	1
2018/1/1	臺中市	B	零售業	無載具	45143168	18490197906	2018	1
2018/1/1	新北市	F	零售業	無載具	63005459	18477975597	2018	1
2018/1/1	高雄市	E	零售業	無載具	40751017	15813468704	2018	1

## 資料日期

以年月表示

資料區間：

2018/01-2021/04

ex. 2021/09

## 縣市

含縣市名與代號

ex. 高雄市 / E

## 載具類別

含縣市名與代號

ex. 高雄市 / E

## 行業別

106 個獨特值

ex. 不動產開發業

## 張數 / 金額

對應條件之發票張數

與合計金額

# 關鍵字

## ■ 資料前處理

- 缺失值
- 重複值
- 離群值
- 資料類型
- EDA

## ■ Python

- **pandas**
  - DataFrame
- **numpy**

## ■ 編輯器

- Google Colab
- jupyter notebook
- VS Code

# 資料輪廓 - 行業別對照表

產業分類	行業別
支援服務業	建築物及綠化服務業
支援服務業	旅行及相關服務業
支援服務業	租賃業
支援服務業	住宿服務業
支援服務業	旅行及相關代訂服務業
支援服務業	業務及辦公室支援服務業
支援服務業	保全及私家偵探服務業
出版影音及資通訊業	廣播、電視節目編排及傳播業
出版影音及資通訊業	電信業
出版影音及資通訊業	影片及電視節目業；聲音錄製及音樂發行業
出版影音及資通訊業	出版業
出版影音及資通訊業	資訊服務業
出版影音及資通訊業	電腦程式設計、諮詢及相關服務業
出版影音及資通訊業	影片服務、聲音錄製及音樂出版業
出版影音及資通訊業	傳播及節目播送業
出版影音及資通訊業	電腦系統設計服務業
出版影音及資通訊業	資料處理及資訊供應服務業
用水供應及污染整治業	污染整治業
用水供應及污染整治業	廢水及污水處理業
用水供應及污染整治業	廢棄物清除、處理及資源回收處理業
用水供應及污染整治業	用水供應業
用水供應及污染整治業	廢（污）水處理業
用水供應及污染整治業	廢棄物清除、處理及資源回收業
住宿及餐飲業	餐飲業
住宿及餐飲業	住宿業

## 行 業 別

共 106 種獨立值



## 產 業 分 類

整理為 19 種類別

「串  
資  
料」  
嘗  
試  
看  
看  
！



# 資料合併後預期成果

發票年月	縣市	縣市代號	行業別	產業分類	載具類別	電子發票張數	電子發票金額	年	月
2018/1/1	臺北市	A	不動產開發業	不動產業	無載具	1638762	2553965042	2018	1
2018/1/1	臺北市	A	餐飲業	住宿及餐飲業	無載具	13071869	3384901815	2018	1
2018/1/1	臺北市	A	電信業	出版影音及資通訊業	使用載具	16307118	15424465043	2018	1
2018/1/1	臺北市	A	電信業	出版影音及資通訊業	無載具	264998	7069430238	2018	1
2018/1/1	臺北市	A	影片服務、聲音錄製及音樂出版業	出版影音及資通訊業	使用載具	1850398	3472000049	2018	1
2018/1/1	臺北市	A	資料處理及資訊供應服務業	出版影音及資通訊業	使用載具	5503415	3120768143	2018	1
2018/1/1	臺北市	A	零售業	批發及零售業	無載具	58357391	29622747754	2018	1
2018/1/1	臺北市	A	批發業	批發及零售業	無載具	4915556	24494764143	2018	1
2018/1/1	臺中市	B	零售業	批發及零售業	無載具	45143168	18490197906	2018	1
2018/1/1	新北市	F	零售業	批發及零售業	無載具	63005459	18477975597	2018	1
2018/1/1	高雄市	E	零售業	批發及零售業	無載具	40751017	15813468704	2018	1
2018/1/1	桃園市	H	零售業	批發及零售業	無載具	37804037	13457180120	2018	1
2018/1/1	臺南市	D	零售業	批發及零售業	無載具	26970811	9928428815	2018	1
2018/1/1	臺北市	A	零售業	批發及零售業	使用載具	10754807	6799819367	2018	1
2018/1/1	新竹市	O	零售業	批發及零售業	無載具	9974546	5009651090	2018	1
2018/1/1	新北市	F	批發業	批發及零售業	無載具	3610074	4458622470	2018	1
2018/1/1	彰化縣	N	零售業	批發及零售業	無載具	12110208	4457810524	2018	1
2018/1/1	臺中市	B	批發業	批發及零售業	無載具	1828745	4357813677	2018	1
2018/1/1	屏東縣	T	零售業	批發及零售業	無載具	9007431	2798522963	2018	1
2018/1/1	臺北市	A	租賃業	支援服務業	無載具	109716	15784308526	2018	1
2018/1/1	高雄市	E	電力及燃氣供應業	電力及燃氣供應業	使用載具	956242	6959295398	2018	1
2018/1/1	臺中市	B	電力及燃氣供應業	電力及燃氣供應業	使用載具	1055567	5936179445	2018	1
2018/1/1	桃園市	H	電力及燃氣供應業	電力及燃氣供應業	使用載具	777694	5747003719	2018	1
2018/1/1	臺南市	D	電力及燃氣供應業	電力及燃氣供應業	使用載具	730973	5449054676	2018	1
2018/1/1	新竹市	O	電力及燃氣供應業	電力及燃氣供應業	使用載具	336598	3654956564	2018	1
2018/1/1	臺北市	A	電力及燃氣供應業	電力及燃氣供應業	使用載具	698768	3516098861	2018	1
2018/1/1	新北市	F	電力及燃氣供應業	電力及燃氣供應業	使用載具	1100808	3169822712	2018	1

## ■ 資料合併

- **Merge**
- **Key Value**
- **Inner / Outer / Left / Right Join**

# 題目

你是一位在花蓮縣經營「DA餐酒館」老闆，由於生意興旺，你想要把品牌連鎖化，在其他縣市設立加盟店。

當你和團隊討論地點及運營體系的時候，突然有整合行銷公司打電話給你，詢問你是否最近有想要拓店？若需要電子載具、多元化支付等服務可以為你提供一條龍的服務。對方的服務層次都很完善，價格也是合理，不過首先作為老闆的你需要先跟團隊討論以下問題：



# 題目

1. 現在手上的流轉資本僅可設立兩家店面，你會選擇在哪兩個城市設店（也可以兩家設立在同個城市）？原因是什麼？
2. 承上，在該城市開設店面後，是否有需要再花一筆額外資金找尋整合行銷公司導入電子載具服務？還是電子發票就可以滿足需求了？
3. 配合國內旅遊盛行的風氣，你想要把店面拓展到離島；請問哪個離島縣市會是你擴展店面的首選呢？原因是什麼？

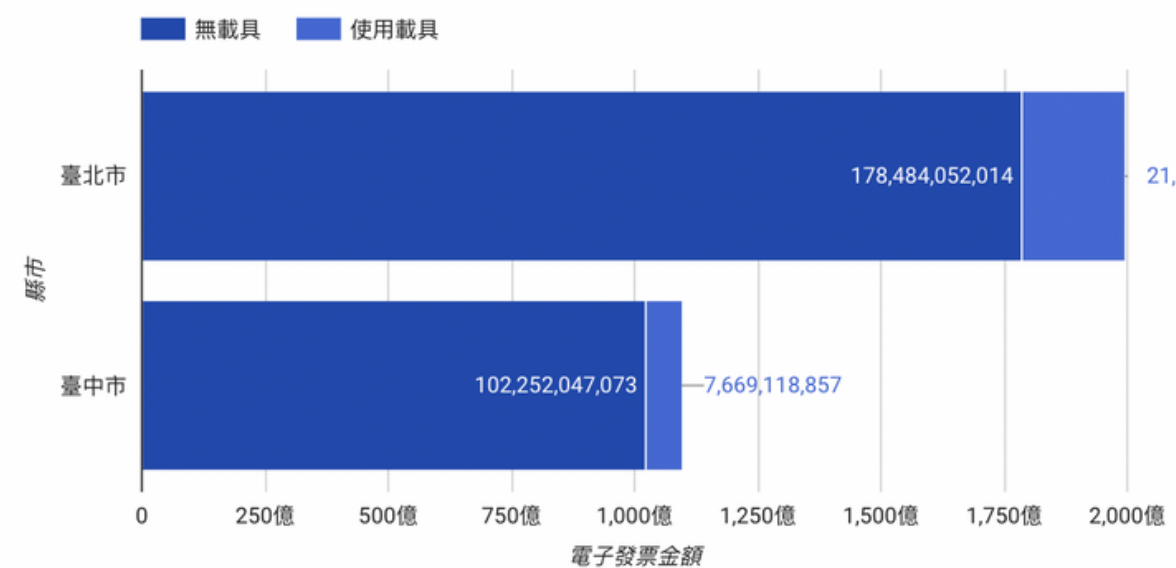
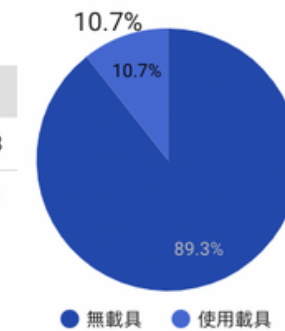
# 參考解題脈絡

Question 2 : 承上，在該城市開設店面後，是否有需要再花一筆額外資金找尋整合行銷公司導入電子載具服務？還是電子發票就可以滿足需求了？

臺北市電子載具使用比例

縣市...	行業別	載具類別	電子發票張數	電子發票金額	平均電子發票...
1. 臺北市	餐飲業	使用載具	62,784,106	21,320,705,446	399.58
2. 臺北市	餐飲業	無載具	565,530,143	178,484,052,014	315.01
3. 臺中市	餐飲業	使用載具	29,631,063	7,669,118,857	238.07
4. 臺中市	餐飲業	無載具	434,999,076	102,252,047,073	234.1

縣市...	行業別	載具類別	電子發票張...	電子發票金額	平均電子發...
1. 臺北市	餐飲業	使用載具	62,784,106	21,320,705,446	399.58
2. 臺北市	餐飲業	無載具	565,530,143	178,484,052,014	315.01



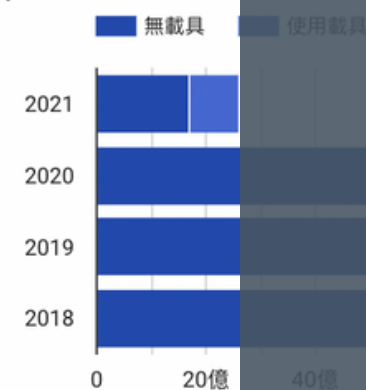
## Insights

臺北市：電子載具比例高達 10%，總金額高達兩百億，故可以考慮跟電子載具公司合作  
-> 不過仍要確認 (1) 電子載具這三年來的使用趨勢是否在成長 (2) 用電子載具的消費平均金額

臺中市：電子載具使用比例僅 7%，電子發票總額僅佔十億元，不需要再額外花錢跟整合行銷公司合作

待補資訊：(1) 六都的載具使用率 Benchmark : (2) 要不要鎖定在 2020 年的數據

(1) V



可使用：

- 視覺化工具如：Tableau, Data Studio

- Python 也有視覺化套件如：matplotlib, seaborn

- 其他分析方法與工具

【統整】

## ■ 你會拿到 ——

- 兩份資料集 (link)
  - taiwan\_invoice\_data.csv
  - industry\_category.csv
- 兩份示範專案
  - EDA (link)
  - 解題流程 (link)
- 這份簡報
- Datacamp 學習帳號

## ■ 你需要 ——

在 **6/15 (三) 23:59** 前：  
繳交含「完整解題流程與邏輯」  
的專案至專案資料夾 (link)

**Good Luck!**



# NCCU DA

政大數據分析社

歡迎加入！期待一起學習

結尾

