

人工智慧導論-第1章 課程介紹 (iClass)

大數據分析

- R/Python/Julia/SQL 程式設計與應用
(R/Python/Julia/SQL Programming and Application)
- 資料視覺化 (Data Visualization)
- 機器學習 (Machine Learning)
- 統計品管 (Statistical Quality Control)
- 最佳化 (Optimization)



李明昌博士

<https://www.youtube.com/@alan9956>

<http://rwepa.blogspot.com/>

alan9956@gmail.com

主題

第1章 課程介紹 (iClass)

第2章 AI與程式語言簡介

第3章 Anaconda簡介與安裝

第4章 資料匯入探索

第5章 資料處理探索

第6章 資料視覺化探索

第7章 資料整合應用探索

大綱：

1.1 教師簡介

1.2 課程介紹與成績評量

1.3 課程必備技能

1.4 資料科學的心法

1.1 教師簡介

教師簡介 <http://rwepa.blogspot.com/>

- 姓名：李明昌 (ALAN LEE)
- 現職：中華R軟體學會 常務理事
臺灣資料科學與商業應用協會 常務理事
- 學歷：中原大學 工業與系統工程所 博士
- 經歷：
 - 淡江大學 兼任教師
 - 佛光大學 兼任教師
 - 國立台北商業大學 兼任教師
 - 東吳大學 兼任教師
 - 育達科技大學 資訊管理系(所) 專任助理教授
 - 崇友實業 行銷企劃專員
 - 國航船務代理股份有限公司 海運市場運籌管理員
 - 大專院校、資策會、工業技術研究院、國家發展委員會、中央氣象局、公平交易委員會、各縣市政府與日本名古屋產業大學等公民營單位 演講達360餘場，3414小時以上。
- 連絡資訊：alan9956@gmail.com



• iPAS AI應用規劃師 證照推廣
• iPAS 營運智慧分析師 證照推廣

1.2 課程介紹與成績評量

課程介紹與成績評量

- iClass demo

- 成績評量：

- 出席率： 15 %
- 平時評量： 15 %
- 期中評量： 30 %
- 期末評量： 40 %



1.3 課程必備技能

課程必備技能

- 準時上課/下課/練習教材檔案
- 下載與安裝 Anaconda
- 開啟/儲存 Jupyter-notebook 檔案
- 使用/儲存 Spyder 開啟 *.py 檔案
- 理解資料匯入 pandas 模組的使用

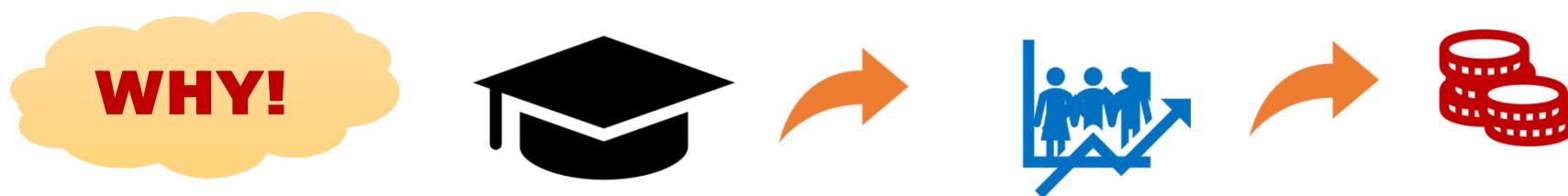
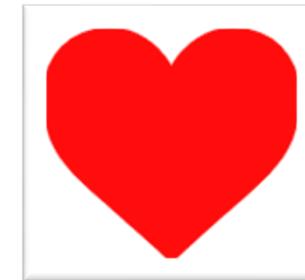
100+



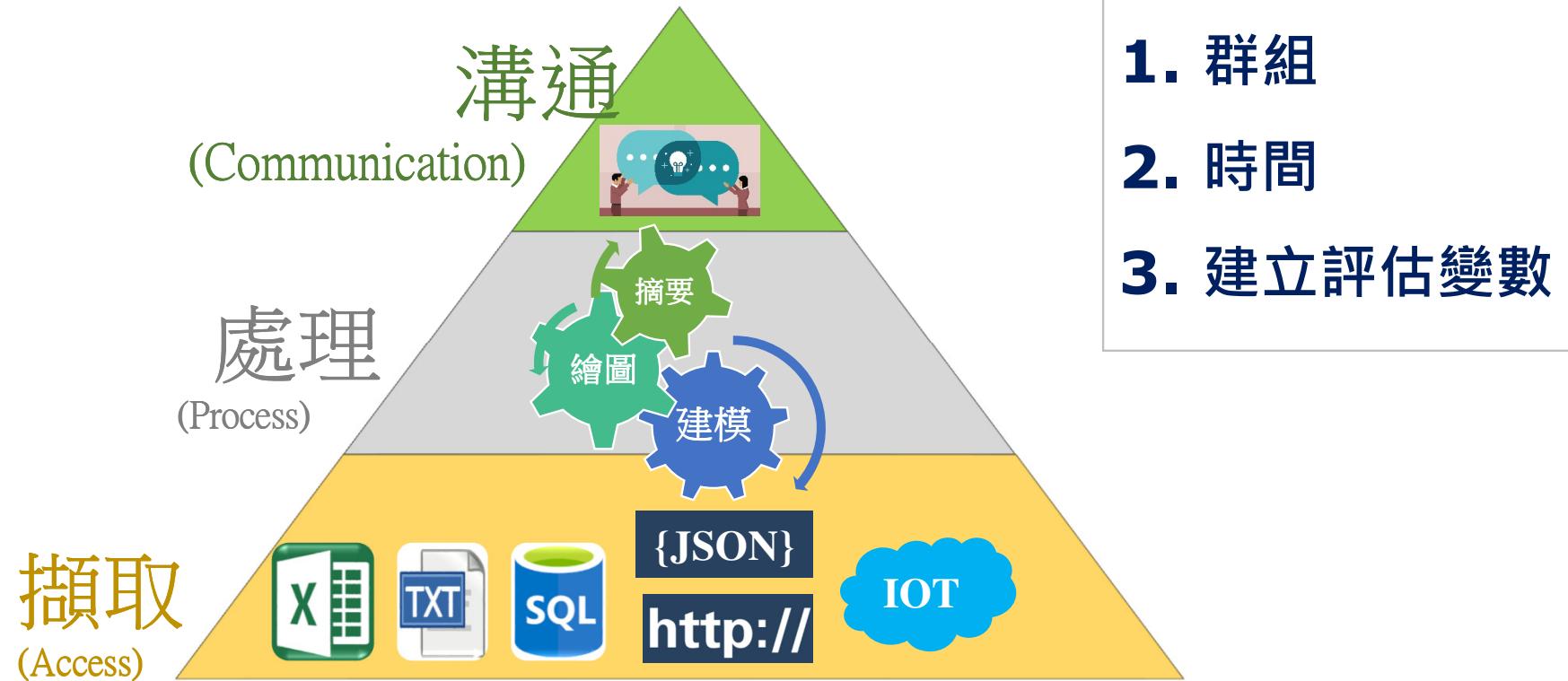
1.4 資料科學的心法

資料科學的心法

1. 熟悉教材內容
2. 掌握 APC方法
3. 掌握 摘要, 繪圖, 建模
4. 遇到問題時, 想辦法尋找答案
5. 將教材的資料集改為工作資料集
6. 參考網路應用文章 (進階) & 學術論文



★★★資料分析架構→APC方法



Python 程式設計-李明昌 免費電子書

- <http://rwepa.blogspot.com/2020/02/pythonprogramminglee.html>

主題: Python 程式設計-李明昌 - ipynb

檔名: Python_Programming_Lee_ipynb.zip

包括 Python 程式設計-李明昌電子書的原始 ipynb 檔案, 圖檔, 部分資料集

下載: https://github.com/rwepa/DataDemo/blob/master/Python_Programming_Lee_ipynb.zip



Python_Programming_Lee_ipynb.zip > python.book.lee >	
名稱	類型
.ipynb_checkpoints	檔案資料夾
data	檔案資料夾
img	檔案資料夾
Python程式設計-李明昌.ipynb	IPYNB 檔案

R 入門資料分析與視覺化應用(7小時28分鐘)

- <https://mastertalks.tw/products/r?ref=MCLEE>

課程提供教學範例的原始程式檔案與資料集



- **主題**
 1. R, RStudio簡介與套件使用
 2. 認識資料物件
 3. 資料處理與分析
 4. 資料視覺化應用
- **特色**
 1. 資料分析的**關鍵八步**
 2. 提供必備**ggplot2**套件的應用知識與使用情境
 3. 提供日期時間**zoo, xts**套件的整合應用操作
 4. 提供**人力資源**資料與**銷售資料**，強化**實務資料**操作能力

R 商業預測應用(8小時53分鐘)

- <https://mastertalks.tw/products/r-2?ref=MCLEE>



課程提供教學範例的原始程式檔案與資料集

- **主題**
 1. R , RStudio 工具操作
 2. 非監督式學習商業預測
 3. 監督式學習商業預測
 4. 財金資料預測應用
- **特色**
 1. 採用最有效率方式學習大數據R語言，並應用於職場資料分析與商業預測應用
 2. 提供多元線性迴歸的必備知識
 3. 提供財金資料商業預測應用的基礎與進階必學技能
 4. 提供學員人力資源資料與台指期 tick 資料預測演練

Excel 應用

Excel 下拉式選單與條件式格式設定教學

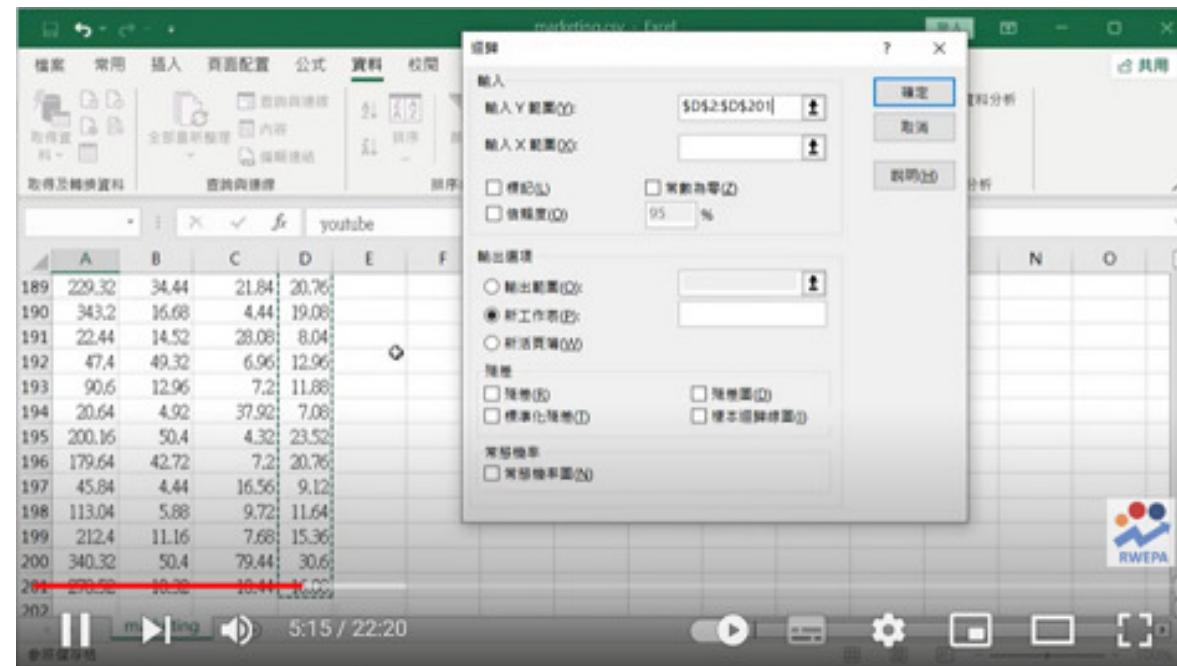
- ✿ YouTube (包括中文字幕) : <https://youtu.be/OVA4dvkrsBM>
- LINK: <https://rwepa.blogspot.com/2023/10/excel-drop-down-list-and-conditional-formatting.html>

	A	B	C	D	E
1	date	state	name		state
2	2023/10/1	簽約	AN		簽約
3	2023/10/2	訂金	LEE		訂金
4	2023/10/3	簽約	RWEPA		安裝
5	2023/10/4	安裝	alan9956@gmail.com		完工
6	2023/10/5	完工	RWEPA		
7	2023/10/6	簽約	PYTHON		

Excel 限制?

Regression Analysis in Excel (Excel 迴歸分析)

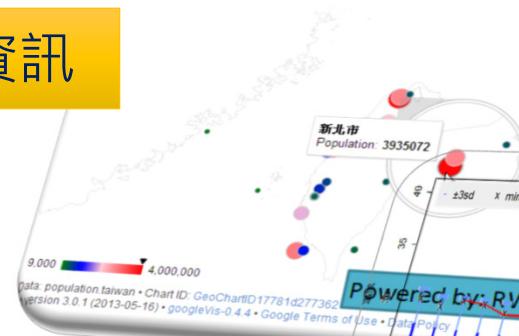
- YouTube : https://youtu.be/i5_urp8XzEs
- LINK: <https://rwepa.blogspot.com/2022/05/httpsrwepa.blogspot.com202205regression-analysis-in-excel.html.html>



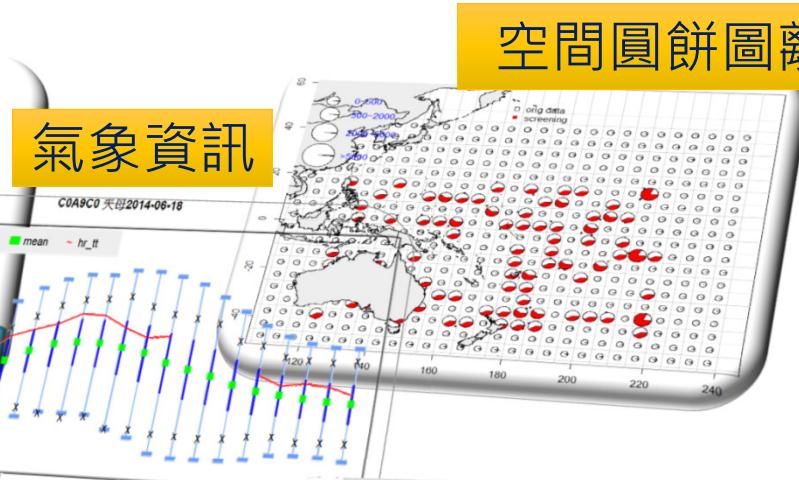
R+Shiny, Python+Streamlit 應用

R+Shiny, Python+Streamlit 互動式平台

地理資訊



氣象資訊



空間圓餅圖離群值分析

保險預測



登山路線視覺化



顧客連結資訊



中央氣象局 1,600萬筆資料(14,328檔案)

網頁呈現

The screenshot displays the CMAQ data visualization and analysis platform interface. Key features shown include:

- Top Navigation:** 氣象資料視覺化暨互動式分析平台_v0.16.5.1, 檔案上傳, 觀測站檢視, QC圖, Reg_QC圖/表 (highlighted with a red box), Reg_QC全部.
- Left Panel (客製化選單):** 選取測站 (466900 淡水), 選取月份(1-12) (滑鼠移動到 6), 選取小時(1-24) (滑鼠移動到 1), 選取年之範圍 (1998-2014), 選取QC閾值 (-10 to 10).
- Middle Panel (異常偵測):** QC_UKref/hr_tt_06mm_01hr_466900.txt scatter plot showing Actual Value - TT vs Predicted Value - UK. A red 'X' marks an outlier at approximately (20, 24).
- Right Panel (動態繪圖):** 學生化殘差直方圖 (Histogram of Studentized Residuals) showing frequency distribution.
- Bottom Panel (R統計運算):** Statistics table for entry 466900 on 2003-06-01 at 1:00. Values include: 觀測站 466900, 日期 20030601, 時 1, 觀測值TT 23.8, 推估值UK 20.6, 學生化殘差 5.163, CMT 23.751, SD 2.159, Outliers 0.0227.

客製化選單

R統計運算

保險預測模型

The screenshot shows the iInsurance interactive analysis platform version v.16.3.24. A yellow callout box on the left is labeled "機率模型閾值調整" (Probability Model Threshold Adjustment) and points to a slider control labeled "機率模型閾值" (Probability Model Threshold) with a value of 0.1. Another yellow callout box on the right is labeled "預測結果" (Prediction Results) and points to a table on the right side of the screen.

The table displays 10 entries of predicted results. The columns include: 編號 (Entry ID), 性別 (Gender), 女性 (Female), 車輛種類 (Vehicle Type), 私家車 (Private Car), 曝露風險 (Exposure Risk), 曝露風險對數 (Exposure Risk Log), 無索償折扣 (No Claim Discount), 被保險人年齡 (Insured Person Age), 私家車一車齡 (Private Car Age 1), 私家車二車齡 (Private Car Age 2), 私家車三車齡 (Private Car Age 3), 私家車車齡組合 (Private Car Age Combination), 車齡 (Age), 預測機率 (Prediction Probability), and 理賠 (Claim). The last two columns are highlighted with red boxes.

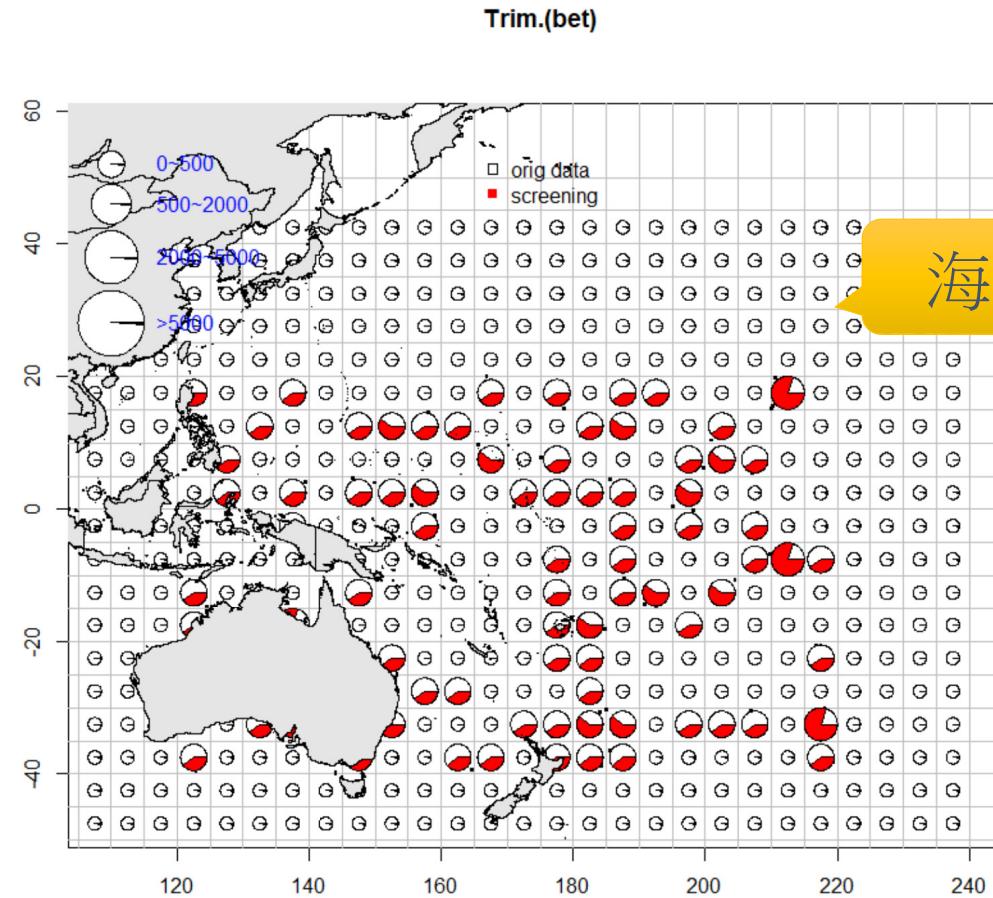
編號	性別	女性	車輛種類	私家車	曝露風險	曝露風險對數	無索償折扣	被保險人年齡	私家車一車齡	私家車二車齡	私家車三車齡	私家車車齡組合	車齡	預測機率	理賠	
1	M	0	A	1	0.9144422	-0.08944106	50	4	1	0	0	1	0	2	0.1069	有
2	M	0	A	1	0.8158795	-0.20348856	20	4	0	0	1	1	2	2	0.1441	有
3	M	0	A	1	0.8377823	-0.17699695	50	3	0	0	1	1	2	2	0.1866	有
4	M	0	A	1	0.4325804	-0.83798702	50	6	0	1	0	1	1	2	0.0944	無
5	M	0	A	1	0.7173169	-0.33223755	50	4	0	0	1	1	2	2	0.1218	有
6	M	0	A	1	0.8377823	-0.17699695	50	4	0	0	1	1	2	2	0.1495	有
7	M	0	A	1	0.8487337	-0.16400975	50	5	0	0	1	1	2	2	0.1422	有
8	F	1	A	1	0.8268309	-0.19015503	10	3	0	0	1	1	2	2	0.1733	有
9	M	0	A	1	0.7145791	-0.33606164	0	5	1	0	0	1	0	2	0.0694	無
10	M	0	A	1	0.3340178	-1.09656101	0	3	0	0	1	1	2	2	0.0783	無

Showing 1 to 10 of 12 entries

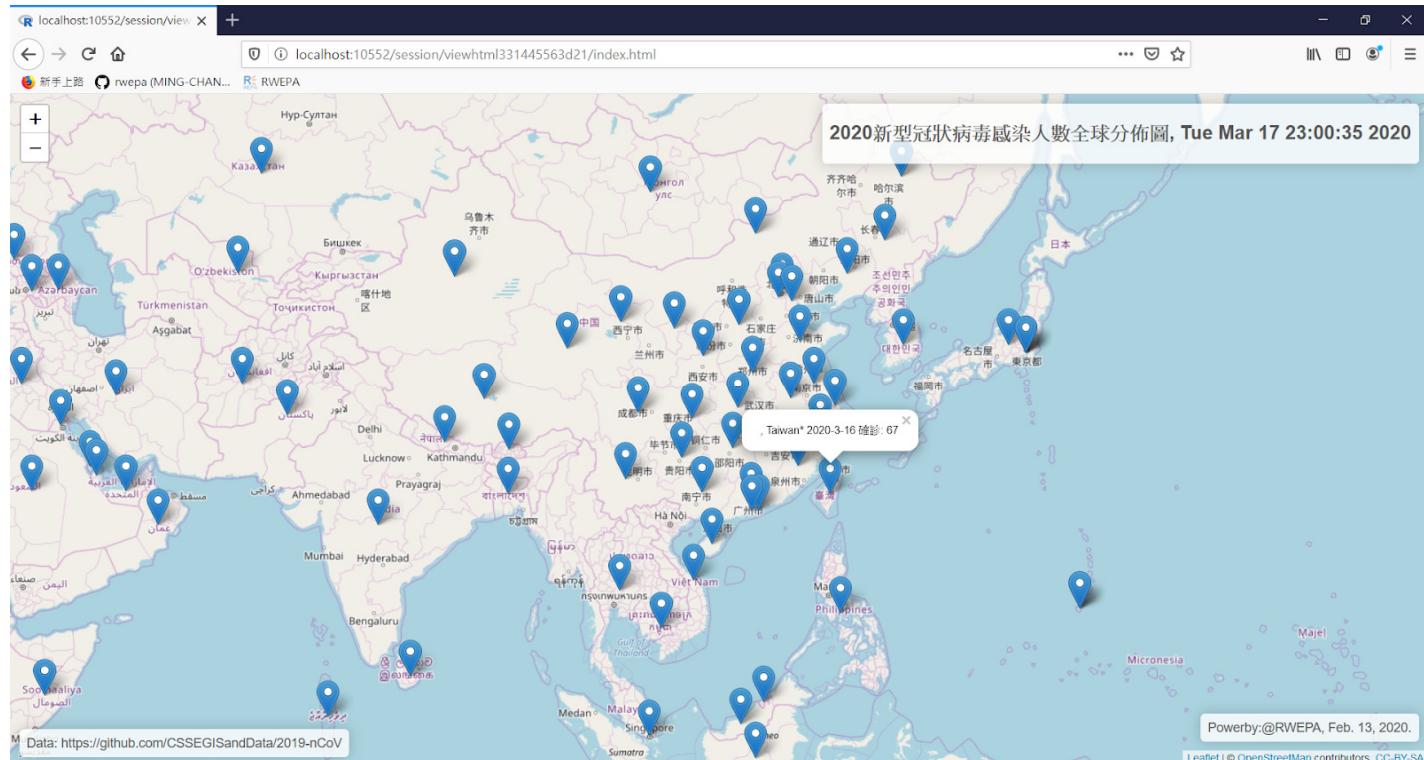
Previous 1 2 Next

127.0.0.1:6177/#tab-9487-2

空間圓餅圖離群值分析



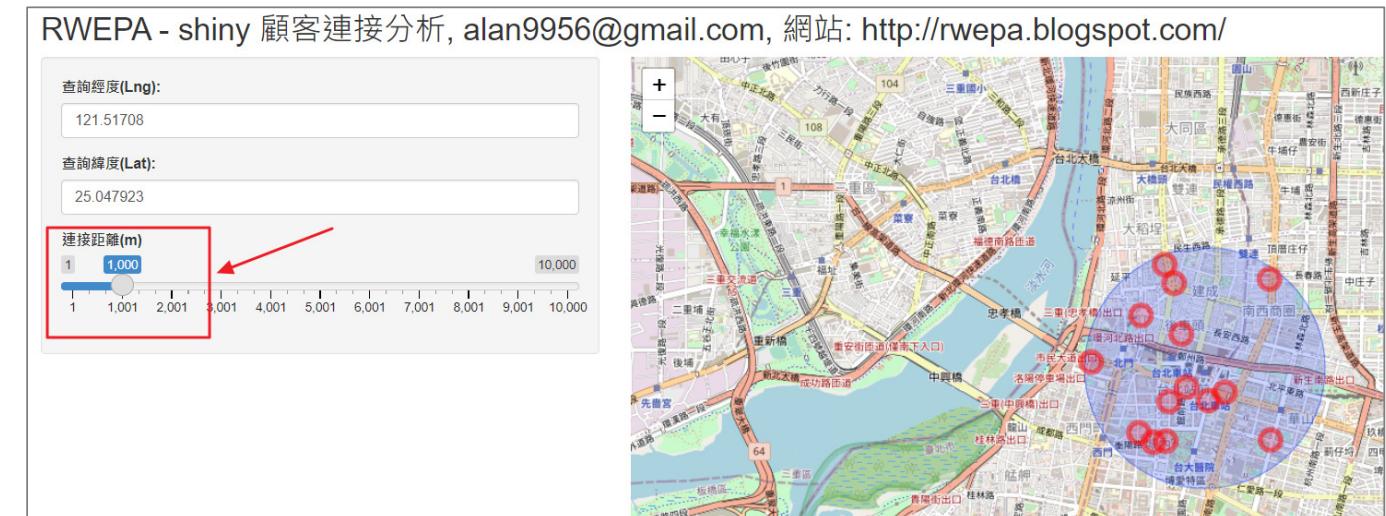
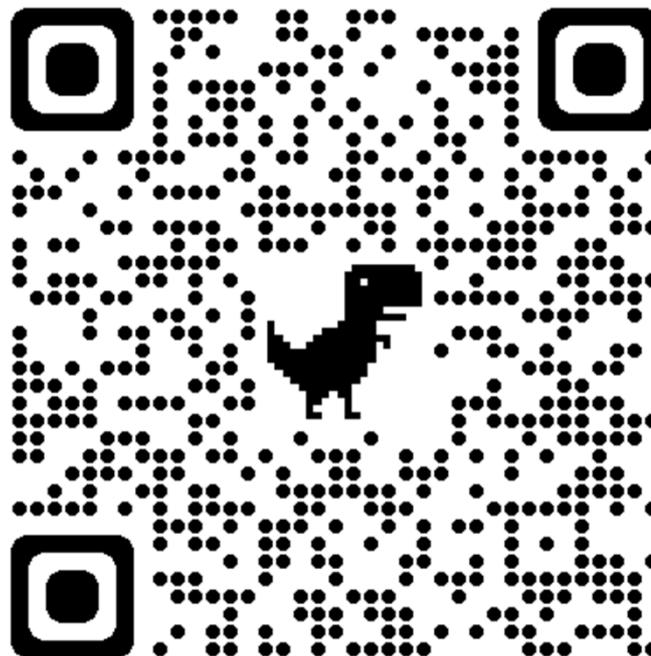
2020新型冠狀病毒視覺化



<http://rwepa.blogspot.com/2020/02/2019nCoV.html>

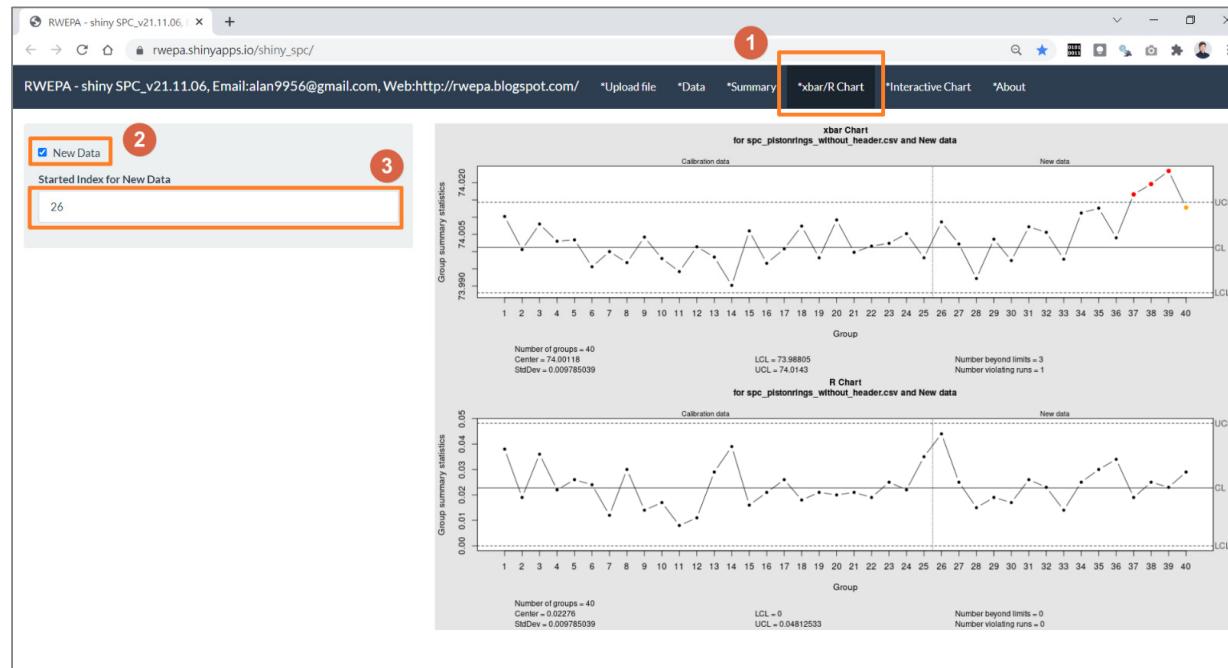
shiny 顧客連接分析

- <https://rwepa.shinyapps.io/shinyCustomerConnect/>



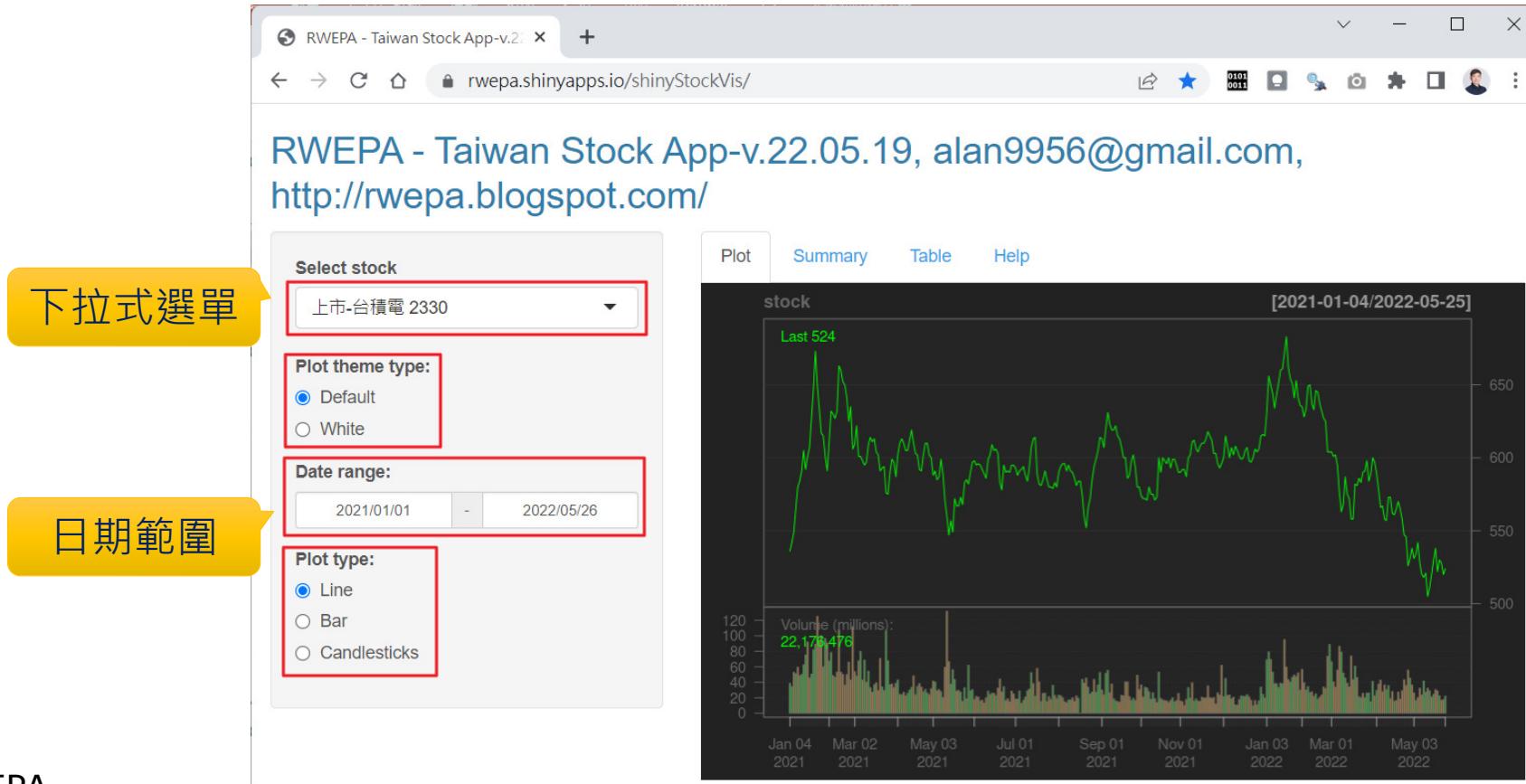
品質管制圖(quality control chart)應用

- 說明: <http://rwepa.blogspot.com/2021/10/r-shiny-quality-control-chart.html>
- 資料1: https://github.com/rwepa/shiny_spc/blob/main/data/spc_wafer_with_header.csv
- 資料2: https://github.com/rwepa/shiny_spc/blob/main/data/spc_pistonrings_without_header.csv
- 線上示範: https://rwepa.shinyapps.io/shiny_spc/

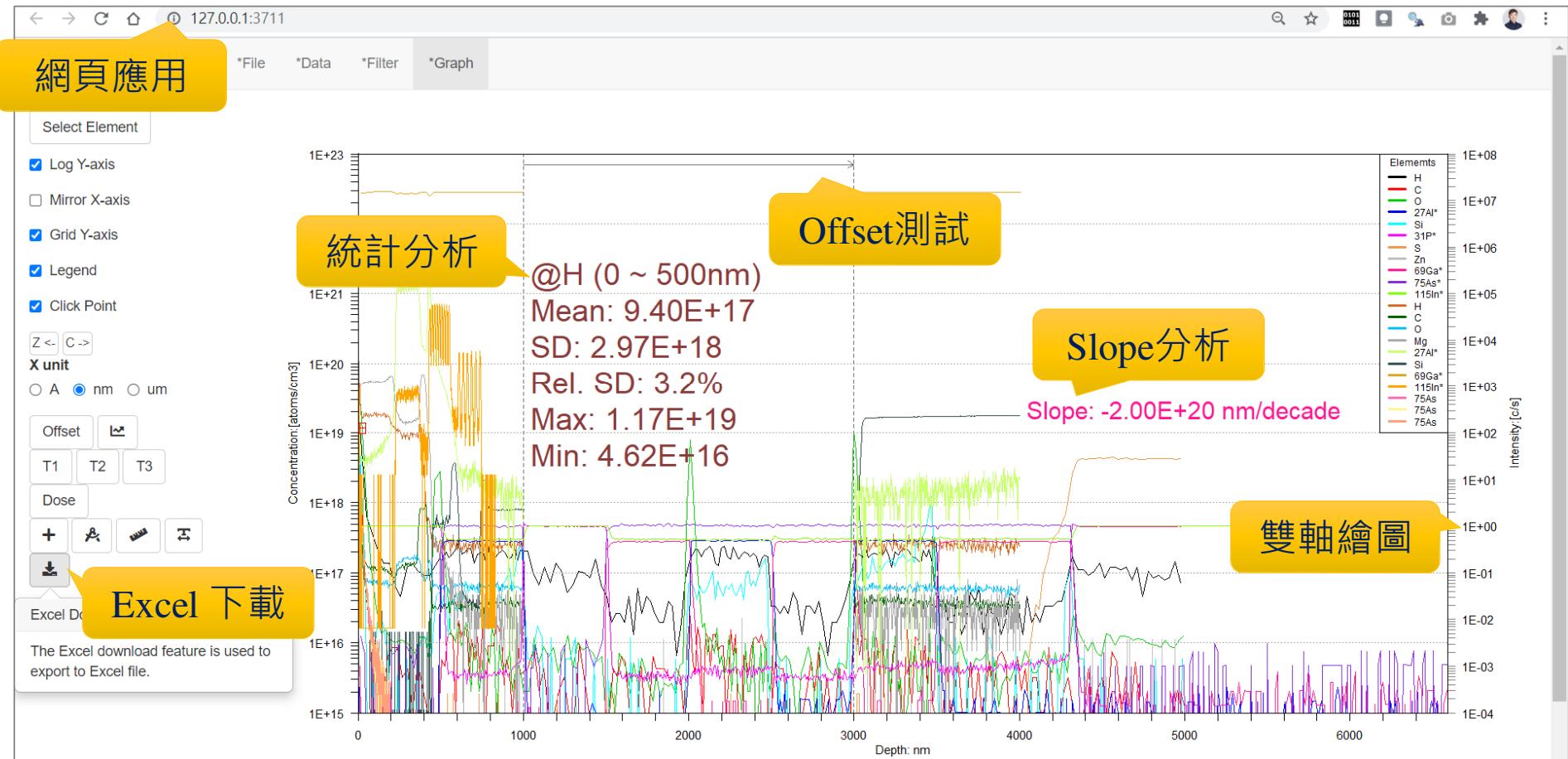


Taiwan Stock App

- <https://rwepa.shinyapps.io/shinyStockVis/>

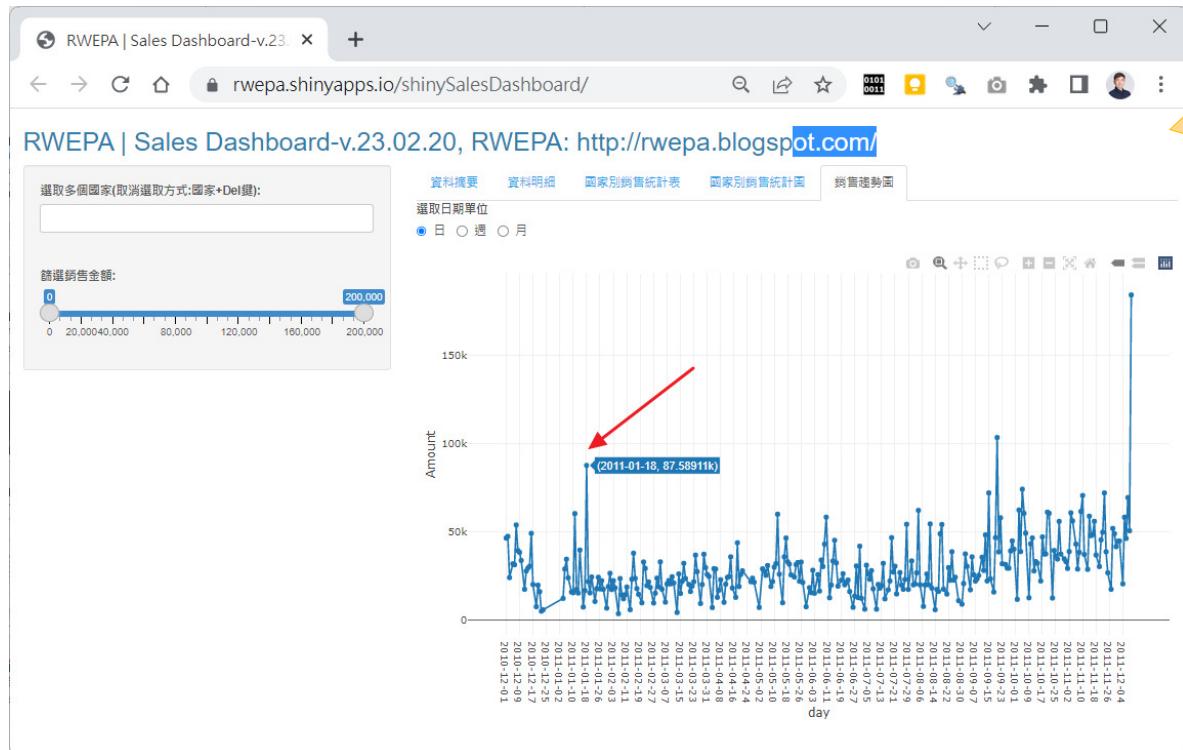


離子資料分析與視覺化應用



RWEPA | shiny企業實務應用 第4集-shiny銷售儀表板

- Shiny: <https://rwepa.shinyapps.io/shinySalesDashboard/>
- YouTube: <https://youtu.be/4GgZlf8heQk>



謝謝 ^_ ^

訂閱 + 讚 + 開啟小鈴鐺

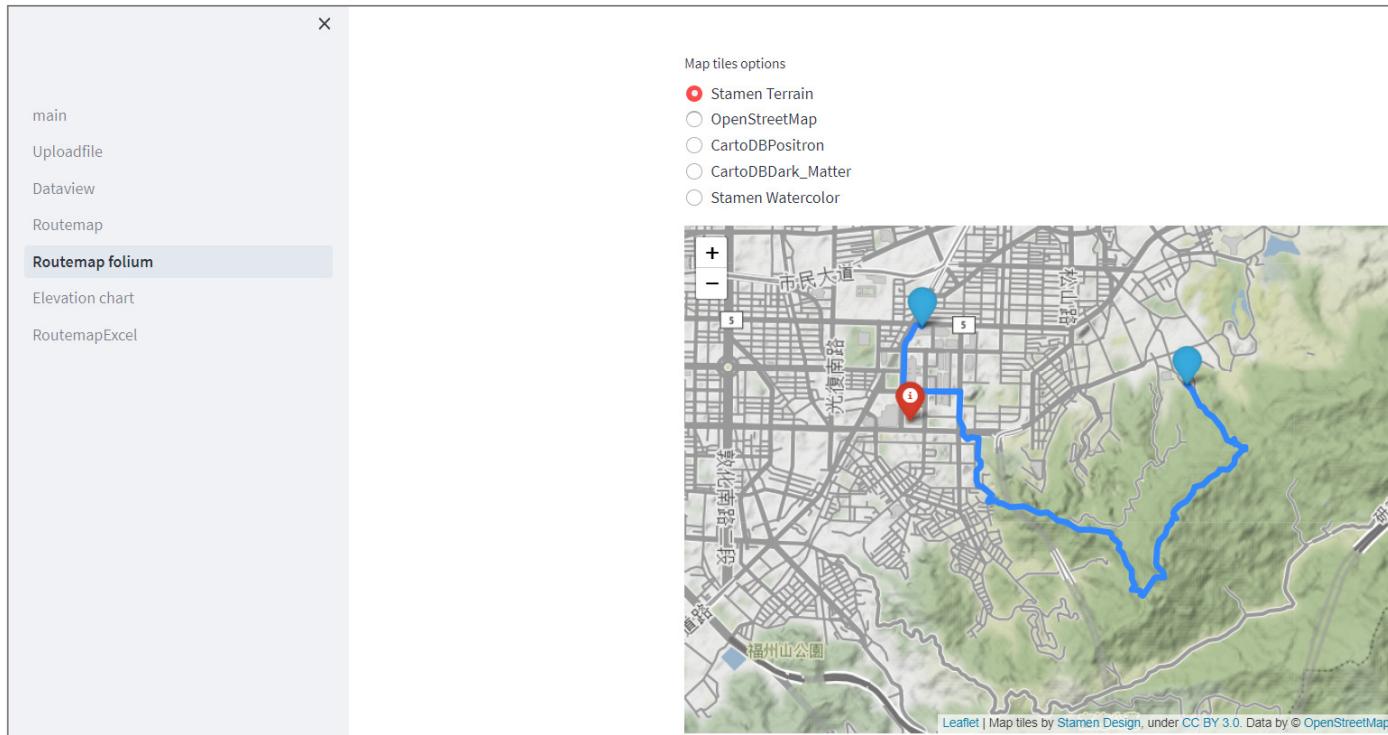
shiny企業實務應用 第6集-小明算命師(下) - 第1季完結篇

- Ubuntu Shiny Server: <https://shiny.rwepa.net/shiny-hr-teller/>
- YouTube: <https://youtu.be/rrD6KV3eV-w>



登山路線視覺化分析平台 (Python + Streamlit)

- YouTube : https://youtu.be/-_zghs2qrlg
- 系統展示 <https://rwepa-climb.streamlit.app/>



銷售儀表板2025 (Python + Streamlit)

- YouTube : <https://youtu.be/QmvIYHspvns>
- 系統展示: <https://rwepa-sales-dashboard.streamlit.app/>



Power BI 應用

Power BI - 客戶輪廓分析

The screenshot displays a Power BI Desktop interface with the following components:

- Top Bar:** Includes tabs for 檔案 (File), 常用 (Home), 插入 (Insert), 模型化 (Modeling), 檢視 (View), 最佳化 (Optimize), and 說明 (Help). It also shows the current file name: PowerBI_customer_profile_analysis - Power BI Desktop.
- Left Sidebar:** Contains sections for 剪貼簿 (Clipboard) and 資料 (Data), which includes options for 取得資料 (Get Data) from various sources like Excel, 資料中樞 (Data Catalog), and SQL Server.
- Right Sidebar:** Shows the 視覺效果 (Visualizations) pane with a list of available visual types and the 資料 (Data) pane with a search bar and a list of data tables: DAX量值管理表, 分行資料表, 日期對照表, 刷卡交易資料表, 刷卡國別對照表, 刷卡類別對照表, 客戶資料表, 理財產品代碼表, 理財產品交易資料表, 業務員資料表, and Dynamic_Calendar.
- Dashboard Content:**
 - 客戶輪廓Overview:** A main card showing the total number of customers (666), average age (34.8), and gender distribution (FEMALE: 378, MALE: 288).
 - 年齡組距 (Age Groups):** A bar chart showing the percentage of customers in different age groups: 01. 25歲以下 (11%), 02. 25-35歲 (48%), 03. 35-50歲 (32%), 04. 50-60歲 (5%), 05. 超過60歲 (3%).
 - 職業 (Occupation):** A horizontal bar chart showing the percentage of customers in various professional categories.
 - 縣市別 (County/Municipality):** A horizontal bar chart showing the percentage of customers across different counties and municipalities.
 - 區域別 (Region):** A horizontal bar chart showing the percentage of customers across different regions.

Power BI – RFM分析

- YouTube : <https://youtu.be/Lkr9HmzLTtg>
- LINK: <https://rwepa.blogspot.com/2023/07/rwepa-rfm-analysis-using-power-bi.html>

Customer Segmentation Using RFM Analysis, 2023



Calendar



最近消費 (recency) :
顧客上次消費時間愈近，用戶價值愈大。

消費頻率 (frequency) :
顧客在一段時間中，總購買次數，購買頻率愈高，用戶價值愈大。

消費金額 (monetary) :
顧客總消費金額，消費金額愈高，用戶價值愈大。

Author : Ming-Chang Lee
YouTube : <https://www.youtube.com/@alan9956>
RWEPA : <http://rwepa.blogspot.tw/>
GitHub : <https://github.com/rwepa>
Email : alan9956@gmail.com

RFM分析 X RFM標準化分析 RECENTY FREQUENCY Monetary +

Tableau 應用

Tableau - Superstore

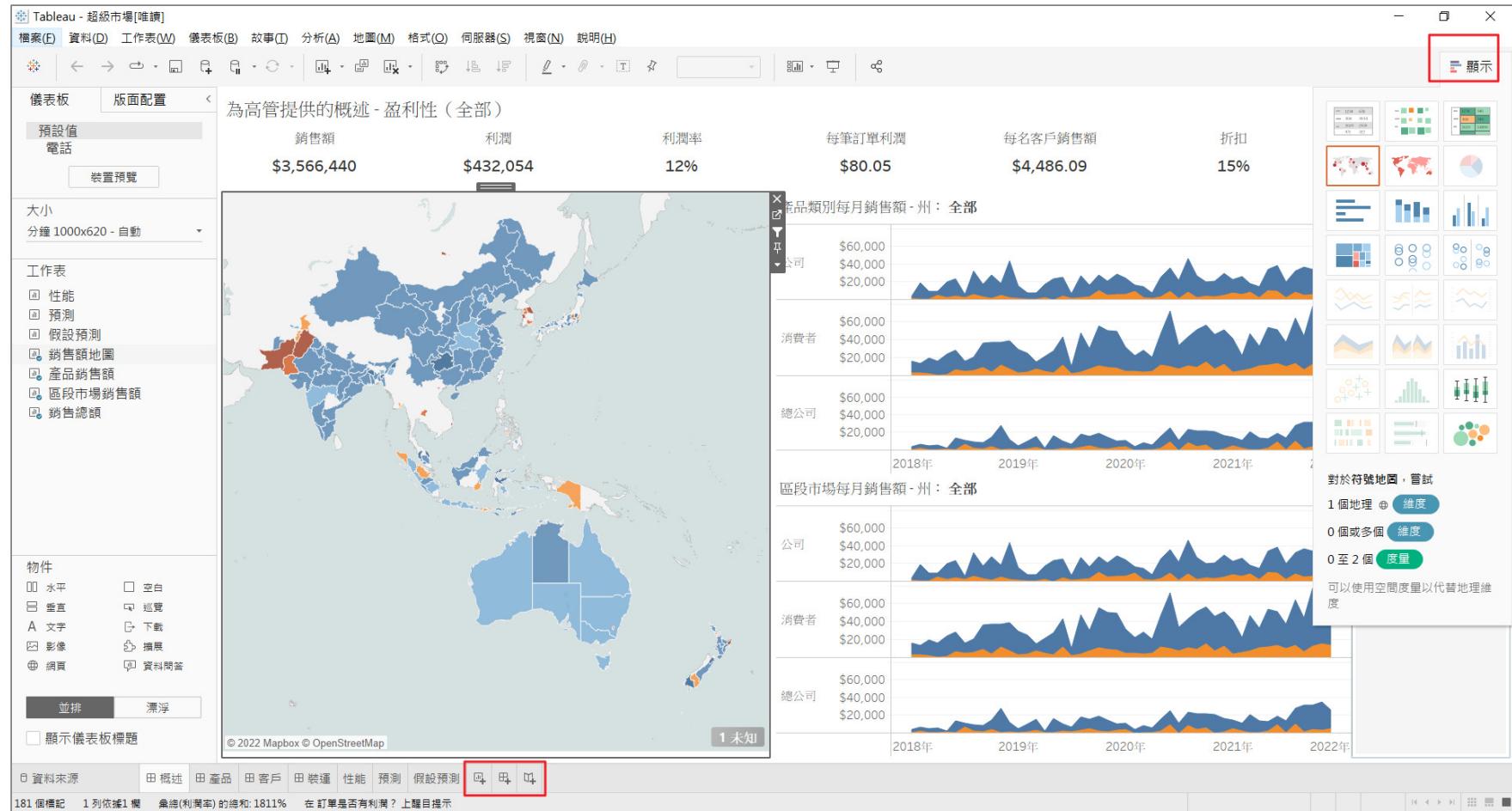
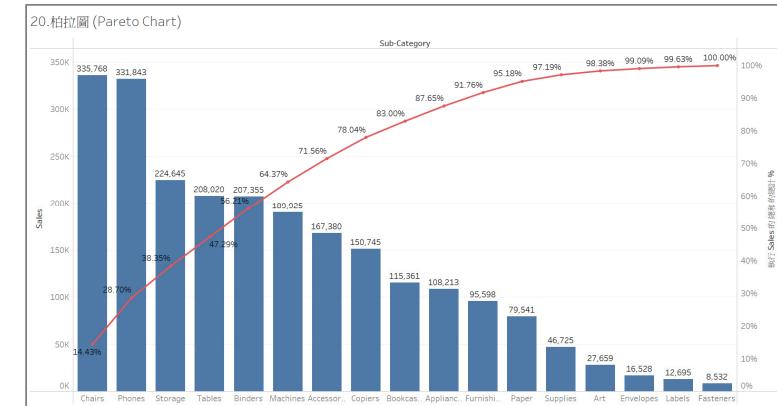
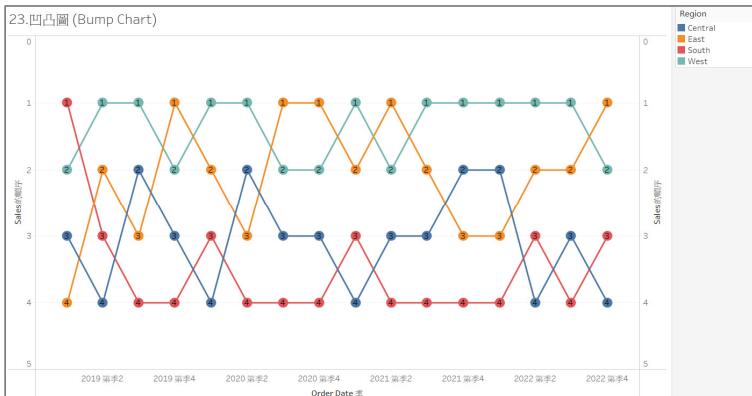
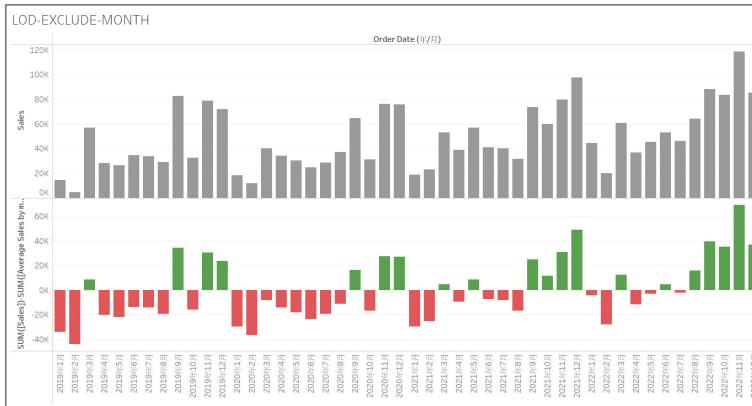


Tableau - 智慧製造應用

- <https://github.com/rwepa/Talks>
- <https://public.tableau.com/app/profile/ming.chang.lee/vizzes>

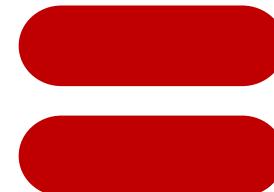
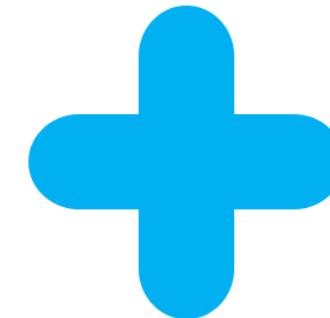


25.購物籃分析 (Market Basket Analysis)

Sub-Category

Sub-Cat.	Access.	Applianc.	Art	Binders	Bookcas.	Chairs	Copiers	Enviro.	Fasten.	Furnishi.	Label	Machin.	Paper	Phones	Storage	Supplies	Tables	
Accessories	718	60	89	161	29	64	7	6	32	27	115	47	22	153	118	103	25	44
Appliances	60	459	63	135	13	49	9	21	23	82	21	9	110	78	56	24	32	
Art	89	63	756	159	34	87	7	29	30	105	47	12	152	124	100	29	41	
Binders	161	135	159	1,339	56	126	19	54	62	200	82	30	276	199	201	43	74	
Bookcas.	29	13	34	56	228	21	2	11	7	30	21	6	48	42	37	10	9	
Chairs	64	49	87	126	21	591	10	29	28	107	36	19	133	91	85	19	36	
Copiers	6	9	7	19	2	10	70	4	5	12	6	2	20	11	14	6	4	
Enviro.	32	21	29	54	11	29	4	251	11	32	11	6	59	41	38	5	12	
Fasteners	27	23	30	62	7	28	5	11	226	48	15	8	59	39	39	13	14	
Furnishi.	115	82	105	200	30	107	12	32	48	919	57	29	181	154	140	31	46	
Labels	47	21	47	82	21	36	6	11	15	57	348	8	80	59	61	18	17	
Machin.	22	9	12	30	6	19	2	6	8	29	8	114	28	22	21	1	6	
Paper	153	110	152	276	48	133	20	59	59	181	80	28	1,205	179	178	45	54	
Phones	116	78	124	199	42	91	11	41	39	154	58	22	179	826	117	22	54	
Storage	103	56	100	201	37	95	14	38	39	140	61	21	178	117	797	33	50	
Supplies	25	24	29	43	10	19	6	5	13	31	18	1	45	22	33	189	13	
Tables	44	32	41	74	9	36	4	12	14	46	17	6	54	54	50	13	314	

學習目標



Python 模組

模組	功能	
Numpy	大型多維度陣列,矩陣	
Scipy	最佳化,線性代數,積分,內插法,FFT,信號和圖像處理	
Pandas	使用 DataFrame 進行資料分析,繪圖	
Matplotlib	靜態, 動畫圖形視覺化	
Statsmodels	統計模型	
Scikit-learn	機器學習	
Tensorflow	深度學習 (例: PyTorch)	
Biopython	生物計算	
Scanpy	單細胞分析	

參考資料

- RWEPA
 - <http://rwepa.blogspot.com/>
- Python 程式設計-李明昌 <免費電子書>
 - <http://rwepa.blogspot.com/2020/02/pythonprogramminglee.html>
- R入門資料分析與視覺化應用教學(付費)
 - <https://mastertalks.tw/products/r?ref=MCLEE>
- R商業預測與應用(付費)
 - <https://mastertalks.tw/products/r-2?ref=MCLEE>

謝謝您的聆聽

Q & A



李明昌

EMAIL: alan9956@gmail.com

<https://www.youtube.com/@alan9956>

WEB: <http://rwepa.blogspot.com/>