樞紐分析與綜合應用

松山高中 大學合作微課程 #6 2019/12/27 Fri. 廖皓宇 (anniehaoyu@gmail.com)

助教:吳杰恩、盧淨婕、蕭羽軒

本週內容

1. 樞紐分析

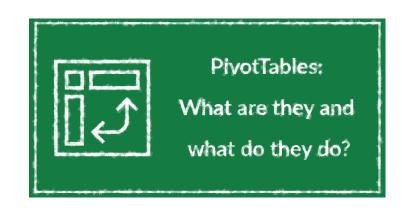
▶ 資料處理:建立樞紐分析表

▶ 資料處理:介紹分析工具

▶ 資料視覺化:從樞紐分析表格到統計圖表

2. 本日實作

3. 專題報告主題



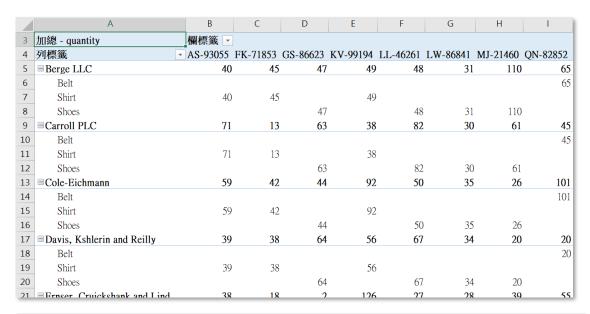
每週課程簡報及資料:

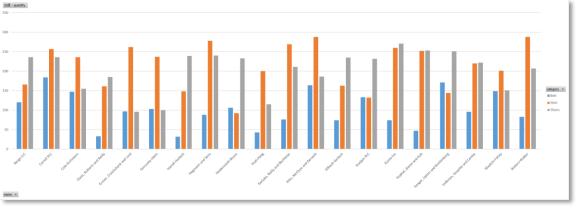
https://reurl.cc/k5kj8x

什麼是樞紐分析?

	А	В	С	D	E	F	G	Н
1	account number	name	sku	category	quantity	unit price	ext price	date
2	296809	Carroll PLC	QN-82852	Belt	13	44.48	578.24	2014/9/27 07:13
3	98022	Heidenreich-Bosco	MJ-21460	Shoes	19	53.62	1018.78	2014/7/29 02:10
4	563905	Kerluke, Reilly and Bechtelar	AS-93055	Shirt	12	24.16	289.92	2014/3/1 10:51
5	93356	Waters-Walker	AS-93055	Shirt	5	82.68	413.4	2013/11/17 20:41
6	659366	Waelchi-Fahey	AS-93055	Shirt	18	99.64	1793.52	2014/1/3 08:14
7	563905	Kerluke, Reilly and Bechtelar	AS-93055	Shirt	17	52.82	897.94	2013/12/4 02:07
8	995267	Cole-Eichmann	GS-86623	Shoes	18	15.28	275.04	2014/4/9 16:15
9	524021	Hegmann and Sons	LL-46261	Shoes	7	78.78	551.46	2014/6/18 19:25
10	929400	Senger, Upton and Breitenberg	LW-86841	Shoes	17	38.19	649.23	2014/2/10 05:55
11	563905	Kerluke, Reilly and Bechtelar	KV-99194	Shirt	12	26.98	323.76	2014/5/20 00:21
12	995267	Cole-Eichmann	KV-99194	Shirt	19	60.22	1144.18	2014/3/10 06:23
13	524021	Hegmann and Sons	QN-82852	Belt	6	13.12	78.72	2013/11/3 18:38
14	758133	Kihn, McClure and Denesik	LL-46261	Shoes	4	59.69	238.76	2014/1/11 21:48
15	555594	Ernser, Cruickshank and Lind	FK-71853	Shirt	12	97.25	1167	2014/9/19 13:20
16	201259	Koelpin PLC	GS-86623	Shoes	9	81.44	732.96	2014/8/12 08:05
17	93356	Waters-Walker	LL-46261	Shoes	18	53.33	959.94	2014/7/15 23:21
18	563905	Kerluke, Reilly and Bechtelar	KV-99194	Shirt	4	35.62	142.48	2014/10/5 23:38
19	201259	Koelnin PLC	KA-00104	Shirt	17	08 23	1660 01	2014/1/26 01:52

多筆、多欄位、缺乏彙整的資料





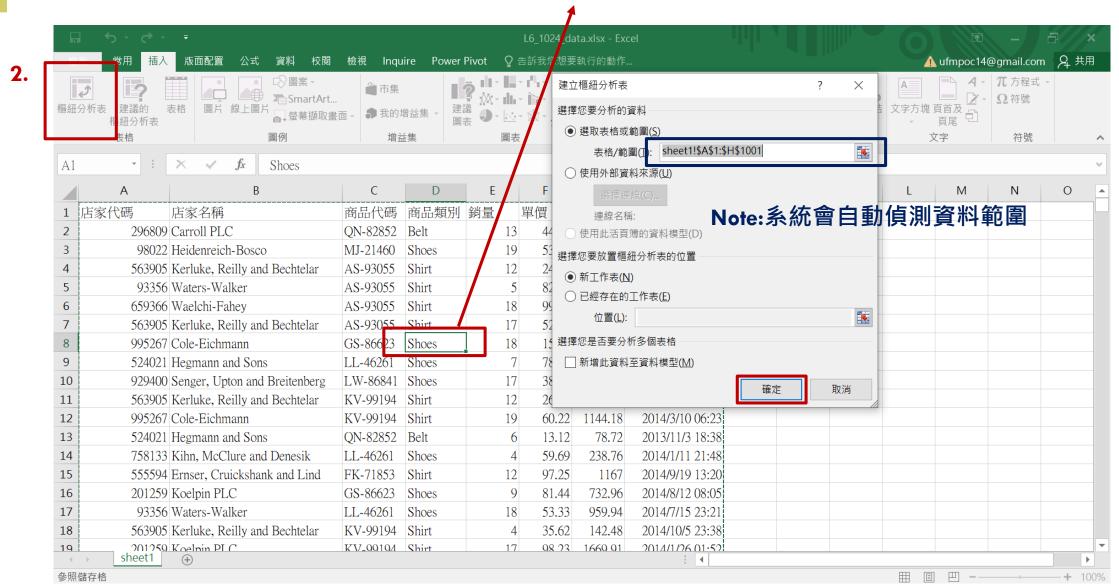
將欄位拆解成<mark>列資料</mark>與欄資料交叉比對、 彙整,產出表格或圖表

Excel中的樞紐分析

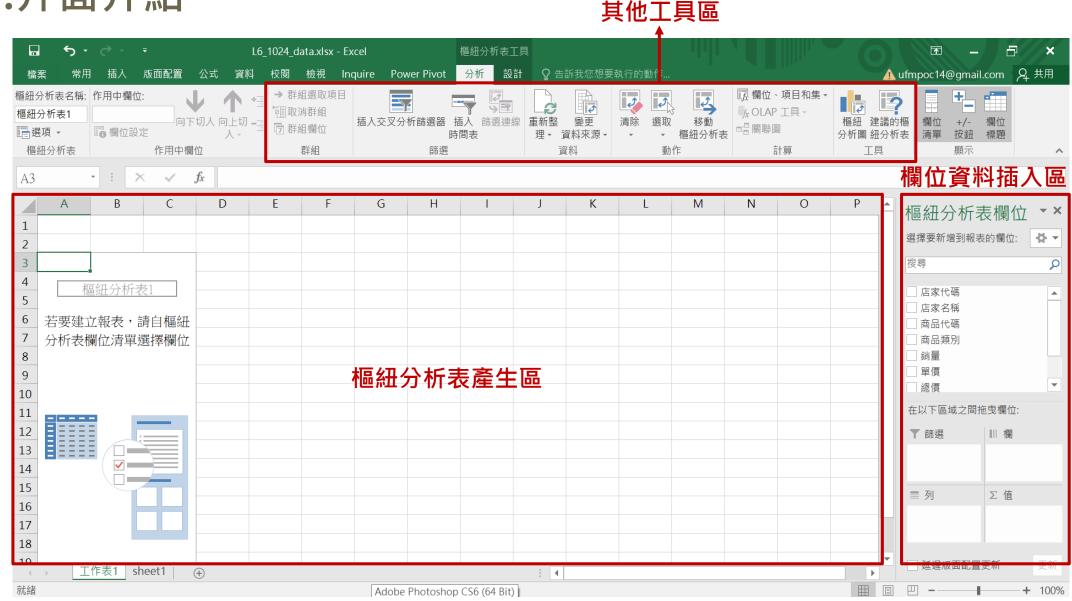
- 認識資料與樞紐分析介面
- 選取列、欄位、數值資料的彙整方式
 - ▶ 加總、平均、最大值、最小值...
- 篩選
- 排序
- 插入時間軸儀表板
- 樞紐分析之後:建立統計圖表

0.匯入資料

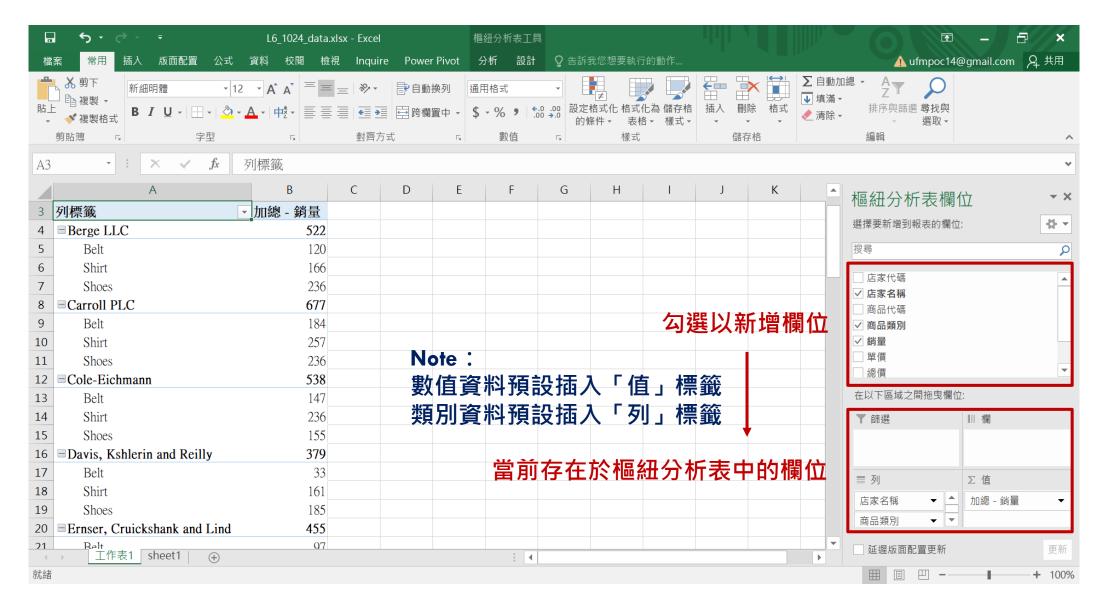
1.點選範圍內任意儲存格



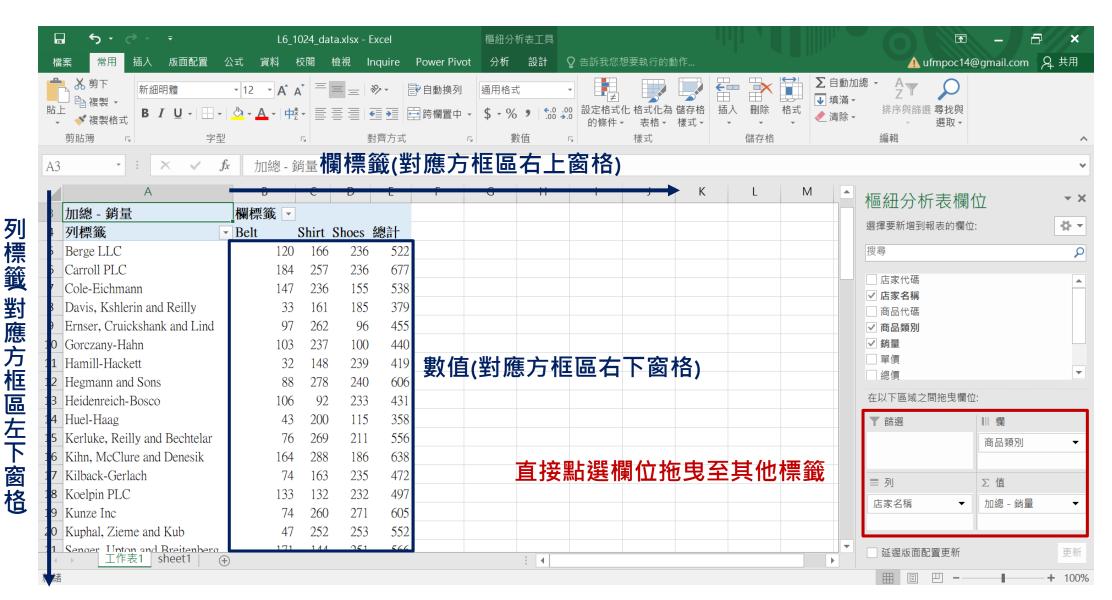
1.介面介紹



2.插入資料欄位

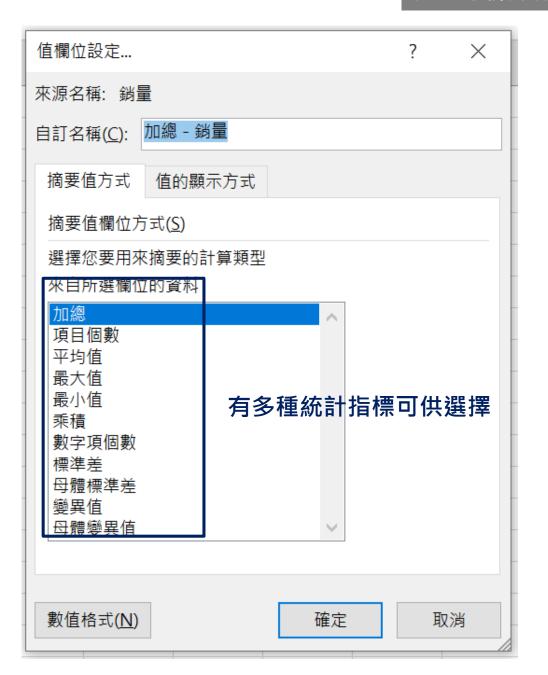


3.調整表格型態

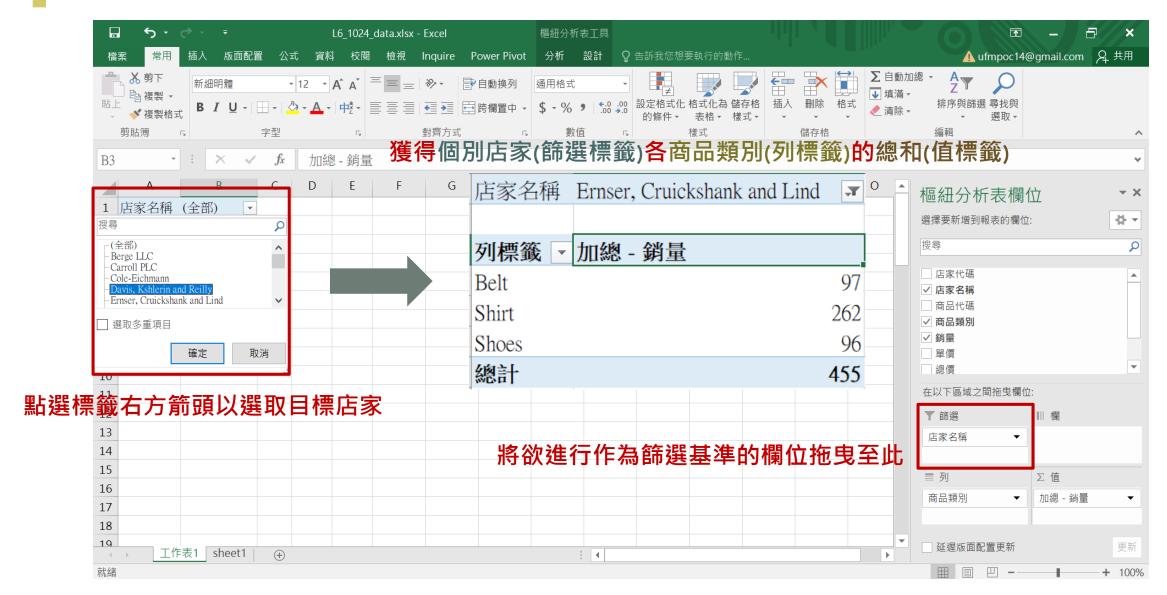


4.數值資料設定





5-1.篩選by個別欄位



677

538

556

638

605

566

538

578

5854

236

240

186

2232

欄標籤 ▼

184

257

∓ Belt

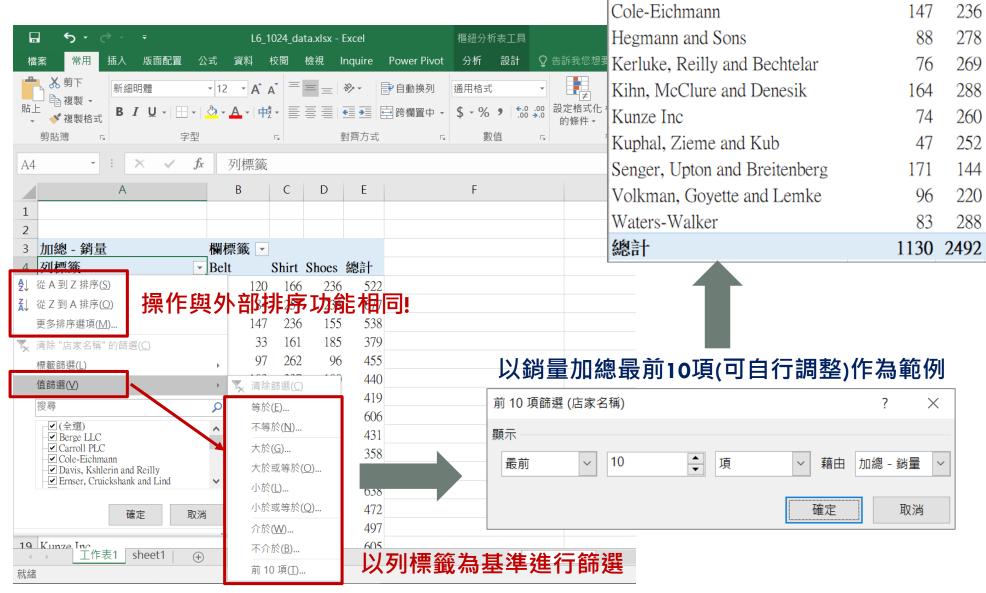
加總 - 銷量

Carroll PLC

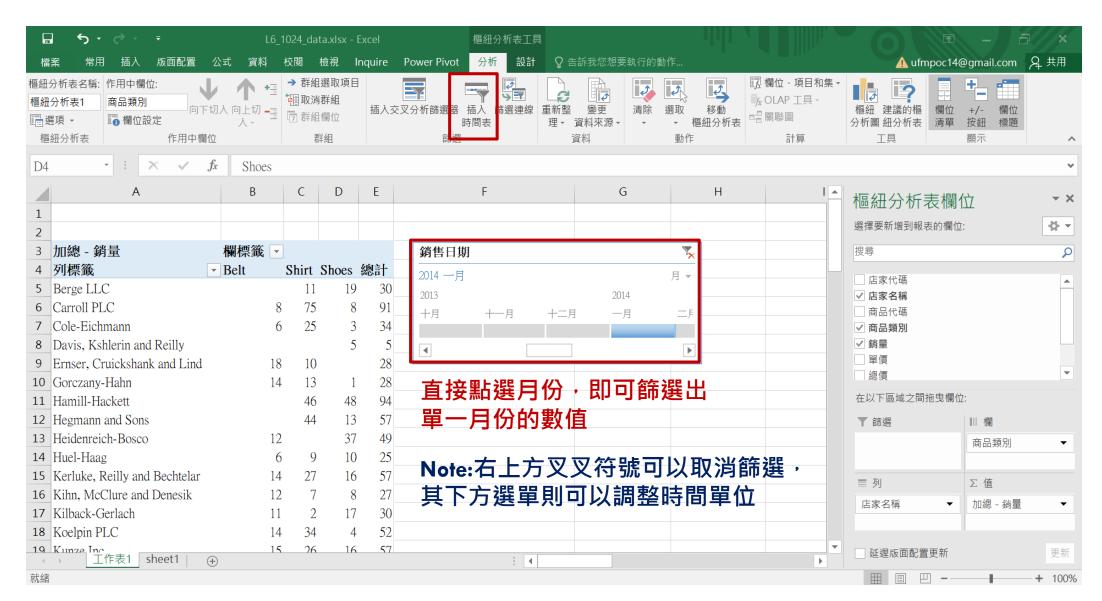
列標籤



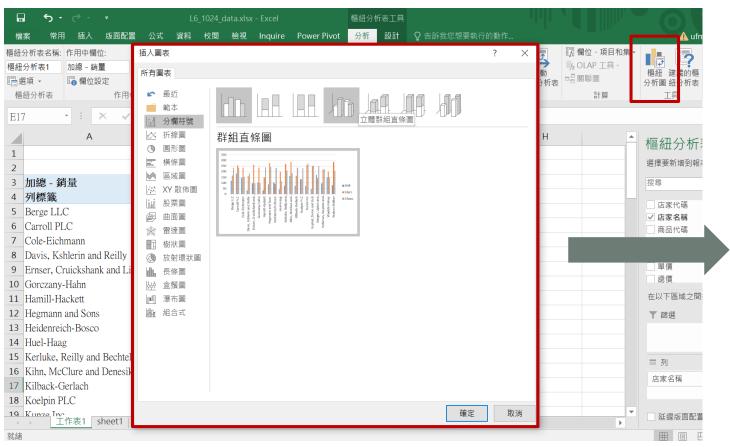
5-2.直接在標籤中篩選與排序

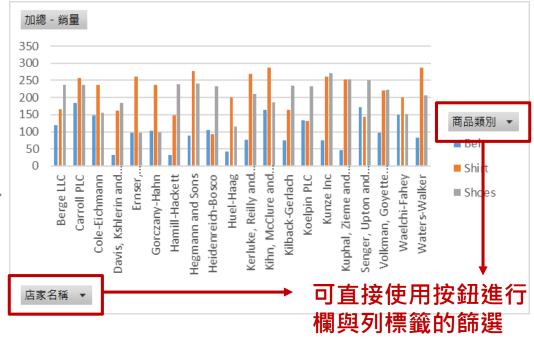


6.插入時間表



7.插入樞紐分析圖





選擇適合的圖表類型

樞紐分析: 各大成衣店銷售量評估

- 範例資料: clothing
 - ▶ 本資料為各大成衣店2013及2014年的銷售資料,包含上衣、皮帶及鞋子三種類型商品的銷量及單價。
- 1) 請找出所有商品以及各類商品總營業額最高的前五個店家。
- 2) 請找出前20%於10月總銷售量成長率最高的店家。
 - ▶ 以2013及2014年10月份的資料作為抽樣計算。
- 3) 請找出**總銷售量成長率前20%**,或**總營業額達27,000元<u>且</u>衣服類營業額達12,000** 元以上的店家。
- 4) 請以統計圖表呈現不同商品種類中,各型號於四季的熱門程度。
 - ▶ 商品種類:衣服、皮帶、鞋子類
 - ▶ 以「總銷售量」作為熱門程度指標

觀察資料

	А	В	С	D	Е	F	G
1	店家代碼	店家名稱	商品代碼	商品類別	銷量	單價	銷售日期
2	296809	Carroll PLC	QN-82852	Belt	13	44.48	2014/9/27 07:13
3	98022	Heidenreich-Bosco	MJ-21460	Shoes	19	53.62	2014/7/29 02:10
4	563905	Kerluke, Reilly and Bechtelar	AS-93055	Shirt	12	24.16	2014/3/1 10:51
5	93356	Waters-Walker	AS-93055	Shirt	5	82.68	2013/11/17 20:41
6	659366	Waelchi-Fahey	AS-93055	Shirt	18	99.64	2014/1/3 08:14
7	563905	Kerluke, Reilly and Bechtelar	AS-93055	Shirt	17	52.82	2013/12/4 02:07
8	995267	Cole-Eichmann	GS-86623	Shoes	18	15.28	2014/4/9 16:15
9	524021	Hegmann and Sons	LL-46261	Shoes	7	78.78	2014/6/18 19:25
10	929400	Senger, Upton and Breitenberg	LW-86841	Shoes	17	38.19	2014/2/10 05:55
11	563905	Kerluke, Reilly and Bechtelar	KV-99194	Shirt	12	26.98	2014/5/20 00:21
12	995267	Cole-Eichmann	KV-99194	Shirt	19	60.22	2014/3/10 06:23
13	524021	Hegmann and Sons	QN-82852	Belt	6	13.12	2013/11/3 18:38
14	758133	Kihn, McClure and Denesik	LL-46261	Shoes	4	59.69	2014/1/11 21:48
15	555594	Ernser, Cruickshank and Lind	FK-71853	Shirt	12	97.25	2014/9/19 13:20
16	201259	Koelpin PLC	GS-86623	Shoes	9	81.44	2014/8/12 08:05
17	93356	Waters-Walker	LL-46261	Shoes	18	53.33	2014/7/15 23:21
18	563905	Kerluke, Reilly and Bechtelar	KV-99194	Shirt	4	35.62	2014/10/5 23:38
19	201259	Koelpin PLC	KV-99194	Shirt	17	98.23	2014/1/26 01:52
20	758133	Kihn, McClure and Denesik	WJ-02096	Belt	15	69.52	2013/11/13 21:38
0.4							

1) 請找出總營業額最高的前五個店家。

● 提示:

- ▶ Excel計算欄位 (總營業額 = 銷量×單價)
- ▶ 選取欄資料、列資料、彙整的數值資料
- ▶ 篩選數值
- 預期成果:

列標籤	加總 -	總營業額
Kihn, McClure and Denesik		38,935.29
Carroll PLC		37,399.11
Waters-Walker		36,778.96
Hegmann and Sons		35,213.72
Kunze Inc		34,406.54
總計		182,733.62

加總 - 總營業額 列標籤	欄標籤 ☑ Belt	細總	計
Kihn, McClure and Denesik	10,11	6.90	10,116.90
Senger, Upton and Breitenberg	950	9.88	9509.88
Carroll PLC	9,35	9.26	9,359.26
Waelchi-Fahey	828	35.72	8285.72
Cole-Eichmann	81	112.7	8112.7
總計	45,38	4.46	45,384.46

加總 - 總營業額 列標籤	欄標籤 ☑ Shirt	總計	
Kihn, McClure and Denesik	18,95	6.35 1	8,956.35
Waters-Walker	18,63	3.71 1	8,633.71
Ernser, Cruickshank and Lind	l 1694	14.19	16944.19
Hegmann and Sons	16,77	4.47 1	6,774.47
Kunze Inc	15,63	8.87 1	5,638.87
總計	86,94	7.59 8	6,947.59

加總 - 總營業額 列標籤	欄標籤 ☑ Shoes	紅絲	· · · · · · · · · · · · · · · · · · ·
Volkman, Goyette and Lemke	1478	36.59	14786.59
Kunze Inc	14,50	3.08	14,503.08
Berge LLC	143	361.1	14361.1
Carroll PLC	14,32	2.24	14,322.24
Kilback-Gerlach	1421	18.95	14218.95
總計	72,19	1.96	72,191.96

2) 請找出前20%於10月的總銷售量成長率最高的店家。

提示:

- ▶ 篩選時間
- ▶ 樞紐分析 + Excel欄位運算 (總銷售量成長率 = (2014 10 月總銷售量-2013 10 月總銷售量) 2013 10 月總銷售量
- ▶ Excel中的複製、「貼上」
- ▶ Excel中的資料排序及篩選

● 預期成果:

店家名稱	2013年10月銷售量	2014年10月銷售量	銷售量增加率 📧
Kerluke, Reilly and Bechtelar	1	23	22
Waelchi-Fahey	2	13	5.5
Kilback-Gerlach	5	30	5
Ernser, Cruickshank and Lind	13	67	4.153846154

3) 請找出總銷售量成長率前20%,或總營業額達27,000元且衣服類營業額達12,000元以上的店家。

• 提示:

- ▶ 樞紐分析 + Excel欄位運算
- ▶ 多個欄位的篩選(AND)

● 預期成果:

店家名稱	2013年10月銷售量	2014年10月銷售量	銷售量增加率	Belt總營業額	Shirt總營業額	Shoes總營業額	所有商品總營業額
Carroll PLC	20	39	0.95	9359.26	13717.61	14322.24	37399.11
Cole-Eichmann	12	22	0.833333333	8112.7	14528.01	8214.71	30855.42
Ernser, Cruickshank and Lind	13	67	4.153846154	5894.38	16944.19	5250.45	28089.02
Hegmann and Sons	20	58	1.9	4909.38	16774.47	13529.87	35213.72
Kerluke, Reilly and Bechtelar	1	23	22	4102.11	12958.23	10329.09	27389.43
Kihn, McClure and Denesik	18	61	2.388888889	10116.9	18956.35	9862.04	38935.29
Kilback-Gerlach	5	30	5	2863.4	9904.85	14218.95	26987.2
Kunze Inc	29	25	-0.137931034	4264.59	15638.87	14503.08	34406.54
Kuphal, Zieme and Kub	11	18	0.636363636	2009.69	12101.14	12921.03	27031.86
Volkman, Goyette and Lemke	17	53	2.117647059	4429.01	12791.27	14786.59	32006.87
Waelchi-Fahey	2	13	5.5	8285.72	11689.05	8993.91	28968.68
Waters-Walker	15	6	-0.6	5957.24	18633.71	12188.01	36778.96

4) 請以統計圖表比較不同商品種類中,各型號

於四季的熱門程度。

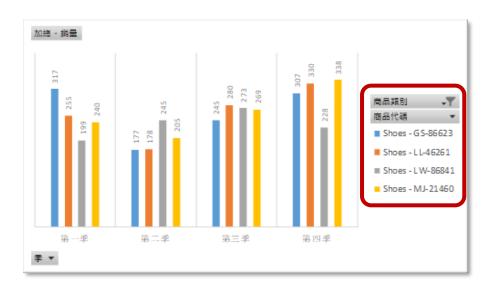
提示:

- ▶ 以「總銷售量」代表熱門程度
- ▶ 巢狀式的樞紐分析欄位
- ▶ 繪製統計圖表
- ▶ 圖表中的資料篩選

● 預期成果:







綜合練習:台北市房屋交易買賣資料分析

- 過去的房地產交易時常因為交易雙方對於價格認知的出入,而衍生許多糾紛。2012年開始,實價登錄交易平台正式上線,至今已經邁入第七年,在各方努力之下,買賣雙方之間的資訊更加的對等、透明,進而降低因為價格所產生的房地產糾紛。同時,實價登錄也成為房仲業者與相關研究人員針對市場結構、交易行為的分析素材...
- 實作資料: housing (臺北市實價登錄交易資料)

綜合練習:台北市房屋交易買賣資料分析(cont.)

- Q1. 你是一名房地產仲介新鮮人,主管要求你透過樞紐分析實價登錄的資料獲得臺 北市2018下半年至2019年初「各行政區」「不同建物型態」<u>單筆交易平均</u>的 「建物移轉總面積」以及「單價元」,產出表格並整理成書面報告交給他,你 該怎麼做?
- Q2. 你交出書面報告之後,主管覺得效果不錯,希望你可以製成圖表及統計地圖, 在隔天的會議上進行簡報,分別是:
 - ① 「各行政區」「不同建物型態」單筆交易平均的「建物移轉總面積」統計圖表
 - ② 「各行政區」「不同建物型態」<u>單筆交易平均</u>的「<u>前六高的</u>單價元」統計圖表
- **Q3.** 結合上述兩張統計圖表,以<mark>單張地圖</mark>展現:「各行政區」「不同建物型態」<u>單</u> 筆交易平均的「單價元」與「各行政區」<u>單筆交易平均</u>的「建物移轉總面積」 的空間分布
 - ※ 提示:以樞紐分析彙整數據,再整理成繪製地圖的資料格式
 - ※ 提示:新增紀錄「台北市」的欄位(以利3D地圖辨別位置)、圖層套疊

今日實作成果查核

共8小題:

- 各大成衣店銷售量評估
 - **(**1), (2), (3), (4)
- 台北市房屋交易買賣資料分析
 - ▶ Q1
 - **Q**2: (1), (2)
 - **Q**3
 - 以「組」為單位查核
 - 舉手請老師查核實作成果
 - 下課前**最快完成的組別**:該組組員各獲得**台大地理系小禮物一份**

松山高中 大學合作微課程

專題報告主題

專題報告進行方式

1/3 (#7)

- 分組專題報告討論及Q&A
- 各組進度報告
 - ▶ 每組5~8分鐘的口頭報告
 - ▶ 以powerpoint簡報呈現
 - ▶ 報告內容:題目構想、實作規劃、執行進度、(可能)遇到的困難

1/10 (#8) 各組正式口頭報告

- ▶ 每組10~15分鐘的口頭報告
- ▶ 以powerpoint簡報呈現
- ▶ 老師評分(50%)、助教評分(50%)、**同儕互評**
- ▶ 優等獎、最佳人氣獎

- → 大學合作課程結業修課證明
- → 台大地理系紀念禮物

期末主題:總統大選競選策略

你是臺北市某競選辦公室的總幹事,面對即將到來的總統大選,你想要利用統計資料來分析選民、並提出適合的競選策略。

● Q1. 分析選區情勢:

▶ 分析選區情勢可從過去選舉的得票率分布,或是一地區的地理及社會特徵著手。其中,一個地區的社會特徵包含了許多不同面向,如,人口結構、教育程度、經濟狀況、社會福利、醫療資源等。請挑選你們感興趣的主題,並使用合適的資料進行分析比較與視覺化。

○ Q2. 擬定競選策略

▶ 請以**台北市**或**挑選一(幾)個**你們所感興趣的行政區,綜合Q1的分析結果,針對這個行政區提出競選策略。

Q1. 分析選區情勢

▶ 提示:以下提供幾個方向,協助你們分析台北市各區的選區情勢:如,選舉結果、人口結構、 教育程度、經濟狀況、社會福利、醫療資源等。



2016總統大選各行政區的不同政黨的得票率分布,各政黨的勢力分布為何?



各行政區的人口結構(性別、年齡)有什麼特徵?可以用**高齡 化、老化指數、扶養比、性別比**等指標來比較



以大專以上/高中職/國中以下三個層級區分,各行政區的教育程度為何?是否有教育程度明顯較高or較低的區域?



綜合各鄰里的所得收入資料,各行政區的經濟狀況為何?是否有平均都很富有/平均都很貧窮/貧富落差大的區域?

Q2. 擬定競選策略

- 請以<u>台北市</u>或<u>挑選一(幾)個</u>你們所感興趣的行政區,綜合Q1的分析結果,針對這個行政區提出競選策略。
- 可能的呈現方式:
 - ▶ 統計圖表
 - ▶ 靜態、動態地圖
 - ▶ 圖表與地圖搭配文字

可使用的資料

- 如:
 - ▶ 臺北市各行政區總統大選投票統計
 - ▶ 臺北市各行政區人口結構
 - ▶ 臺北市各行政區教育程度
 - ▶ 臺北市各行政區所得收入
 - **...**
- 線上開放資料資源:
 - ▶ 台北市資料大平台 https://data.taipei/#/
 - ▶ 政府資料開放平台 https://data.gov.tw/