

第二單元

量化資料分析

統計資料整理與信效度分析

從資料品質、重新編碼到可靠且有效的量表證據

授課教師：邱郁仁 博士

把原始資料變成可相信的決策證據

01

資料檢核與基礎統計

先確認資料範圍、缺漏與分布

02

資料轉換與重新編碼

把規則寫清楚，再產生分析變項

03

信度與效度分析

用一致性與因素結構支持解釋

先確保資料正確，才談方法進階；先留下可重做規則，才談結果可信。

AI 加速流程；SPSS 驗證結果；人類負責判斷

1

AI 助理

列檢查清單、整理輸出與解釋草稿

2

SPSS 驗證

執行運算、保留設定、產生可重現的表格與圖形

3

專業判斷

確認研究問題、代碼規則、理論意義與決策責任

原則：先去識別化；AI 的數值、門檻與結論都要回到原始資料與 SPSS 輸出核對。

資料品質決定分析的上限

統計不會自動修正錯誤代碼、漏填資料或不合理範圍。

輸入錯誤 + 處理規則不明 + 錯誤的摘要 = 錯誤決策

管理者的第一責任：確認資料能否被信任，而不是急著選模型。

AI 初步檢驗資料



AB品牌服飾銷售交...
XLSX



Data Analytics 閱讀並檢視上傳檔案的資料

資料尺度決定可以使用的統計語言

尺度	能表達什麼	商業例子	建議摘要
名目 Nominal	分類	部門、品牌、是否購買	次數、百分比、眾數
順序 Ordinal	分類 + 高低順序	滿意度等級、偏好排名	中位數、百分位數、次數
等距 Interval	順序 + 等距	標準化分數、部分態度量表	平均數、標準差、相關
比率 Ratio	順序 + 等距 + 真零	營收、年齡、件數、時間	平均數、標準差、比率

類別資料看構成；連續資料看位置與離散。先判斷尺度，再選統計量。

次數分配是最有效的第一道資料檢核

教育程度	次數	百分比	有效百分比	累計百分比
高中職	18	18.0%	18.6%	18.6%
大專	46	46.0%	47.4%	66.0%
研究所	33	33.0%	34.0%	100.0%
缺漏	3	3.0%	—	—
總計	100	100.0%	100.0%	—

檢查四件事

- 範圍：是否出現不可能代碼
- 缺漏：缺漏率是否可接受
- 分布：組別是否過度不均
- 分母：百分比與有效百分比不同

有效百分比排除缺漏值；累計百分比只對有順序的類別有意義。

描述統計要同時看位置、離散與形狀

一組數字，至少回答三個問題

位置 | 平均數、中位數、眾數

離散 | 全距、四分位距、變異數、標準差

形狀 | 偏態、峰度、離群值與分布圖

相對位置 | 百分位數、Z 分數

位置

M / Md

離散

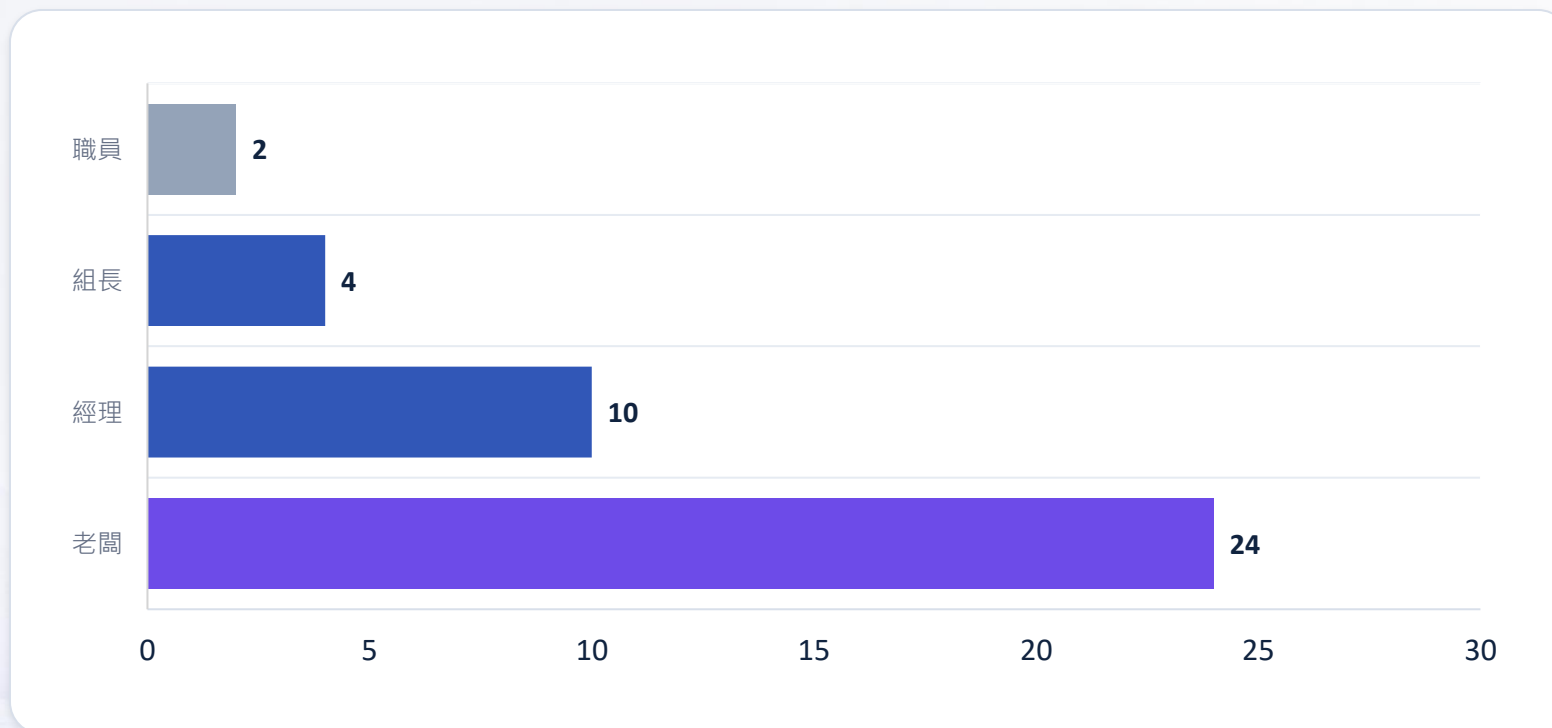
SD / IQR

形狀

Skew / Plot

只報平均數，等於只看資料的一個角度。

平均數可能掩蓋真正的薪資分布



平均月薪

6 萬

中位數：4 萬
多數職員：2 萬

結論：報告平均數時，同時揭露中位數、分布圖與樣本結構。

SPSS 基礎統計：先選問題，再選輸出

1

Frequencies

類別分布、缺漏、眾數、百分位數、圖表

2

Descriptives

平均數、標準差、最小值、最大值、Z 分數

3

Explore

常態性、離群值、箱形圖與分組比較

共同起點：Analyze → Descriptive Statistics；輸出越多不代表分析越好。

AI 輔助資料檢核與描述統計：先提問，再核對

1

找異常

給代碼表與去識別摘要；列範圍、缺漏、重複 ID、離群檢查

2

解讀分布

貼 SPSS 表格；比較平均 / 中位、SD / IQR、偏態與箱形圖

3

SPSS 複核

用 Frequencies、Descriptives、Explore 重跑；禁止補造數值

提示詞：每個結論標註證據欄位；資料不足時只列待確認事項，不要猜測。

錯誤值與缺漏值：能查證就修正，不能查證就標記

可回查原始資料

- ✓ 修正為正確代碼
- ✓ 記錄修改前後值
- ✓ 保留資料來源與日期

無法確認正確答案

- △ 設為缺漏值
- △ 評估缺漏比例與型態
- △ 說明排除或估補規則

禁止：為了讓結果『好看』而任意改值、刪值或補值。

資料轉換不是修飾，而是建立分析規則

每一次重新編碼，都應回答：為什麼轉、怎麼轉、轉完如何驗證？

WHY

商業問題需要什麼變項

RULE

保留完整對照與公式

CHECK

用次數、範圍與樣本數核對

原始變項保留不動；分析變項另存新名稱，是最安全的預設。

重新編碼常見三種任務：轉向、合併與計數

A

轉向

反向題 $1 \leftrightarrow 5$ 、 $2 \leftrightarrow 4$

讓所有題目方向一致

B

合併

年齡區間、部門群組

讓報表或決策規則更清楚

C

計數

計算一系列中某代碼出現次數

形成事件數或選擇數

SPSS : Transform → Recode / Automatic Recode / Count Values within Cases

反向題必須先轉向，再形成量表分數

五點量表公式

$$\text{新分數} = 6 - \text{原分數}$$

$$1 \leftrightarrow 5$$

$$2 \leftrightarrow 4$$

$$3 \rightarrow 3$$

原分數	轉換後	檢查
1	5	最小→最大
2	4	方向相反
3	3	中點不變
4	2	方向相反
5	1	最大→最小

轉完立即跑次數分配：範圍仍應是 1–5，樣本數與缺漏數不應無故改變。

合成分數前先定義計分與缺漏規則

加總 SUM

適合題數固定、想保留總分尺度

例：Q1 + Q2 + Q3 + Q4

平均 MEAN

保留原量尺，跨構面比較較直觀

例：MEAN(Q1,Q2,Q3,Q4)

缺漏值規則必須事先決定

至少回答幾題才計分？

缺一題時採平均或直接缺漏？

不同構面是否使用相同規則？

計分公式是否寫入分析紀錄？

先反向處理 → 再檢查信度 / 因素結構 → 最後計算構面分數。

分組會損失資訊：只有在有決策理由時使用

原始分數

1 2 3 4 5 6 7 8 9 10



合理分組

法規門檻、客戶等級、產品方案
具有明確可行動的切點

不建議分組

只為方便、任意二分、事後調門檻
會降低差異與關係的資訊

分析時優先保留連續變項；報告分組時同時揭露分組規則與各組樣本數。

可靠不等於有效：信度與效度是兩種證據

信度問『結果是否一致』；效度問『是否真的測到想測的構念』。

信度 Reliability

一致性 | 穩定性 | 測量誤差

效度 Validity

內容 | 建構 | 效標關聯

高信度是必要條件，但不是高效度的保證。

Cronbach's alpha 要搭配題項診斷

α

內部一致性

不是效度
也不是越高越好

判讀四個層次

方向：反向題是否已正確轉向

整體：alpha 是否符合研究階段與用途

題項：修正後題總相關是否合理

刪題：alpha if item deleted 是否提示異常題

教材例：數學焦慮 $\alpha=.905$ ；數學態度 $\alpha=.706$ 。解讀需搭配題數、構念與使用情境。

探索性因素分析把多題濃縮成少數構面



EFA 用於『不知道構面究竟如何形成』的探索；CFA 用於驗證既定理論結構。

EFA 判讀順序：適配性、因素數、負荷量、命名

KMO

0.839

> 0.70

Bartlett

$p < .001$

相關矩陣非單位矩陣

因素數

2

特徵值與理論並看

累積解釋

72.527%

兩因素合計

轉軸後因子矩陣的判讀

同一因素中的題項應有共同內容
題項跨兩個因素時，要檢查語意與理論
因素名稱由題項內容決定，不是由軟體決定

統計門檻只是起點；最後仍要回到題項內容與管理意義。

讓 AI 做第二讀者：因素、效度與信度判讀

1

因素適配

看 KMO、Bartlett、陡坡圖與因素數；門檻只作參考

2

建構效度

比較轉軸負荷、跨負荷與題項語意；命名回到理論

3

信度診斷

整理 alpha、題總相關、刪題後 alpha；不憑單一數字刪題

提示詞：結論標註 SPSS 證據欄位；不足就列待確認，不得虛構表格或自行命名因素。

建立可重做、可解釋、可驗證的分析流程

- 1 檢核 ——— 範圍、缺漏、分布、離群
- 2 轉換 ——— 保留原始欄位與完整規則
- 3 計分 ——— 反向題、加總 / 平均、缺漏門檻
- 4 驗證 ——— EFA 結構 + alpha 題項診斷
- 5 報告 ——— 程序、輸出、判讀與限制

真正的專業，不只是跑出數字，而是能說明每一個數字如何產生。