

不論是透過問卷、訪談或觀察而得來的資料，若要用統計方法處理，首先便要決定會否使用電腦協助分析，考慮的原則如下：

- ◆ 是否在研究開始前已決定使用電腦進行統計分析？
- ◆ 事前已學懂使用合適的電腦軟件嗎？
- ◆ 若有大量資料需要輸入電腦，有足夠人手幫助嗎？
- ◆ 是否只需要簡單的統計分析？若是，會否用計算機計算更為快捷方便？
- ◆ 明白電腦軟件所提供的統計技術的意義和用途嗎？切勿人云亦云，盲目追求未有能力掌握的分析技術。

這些原則說明電腦只是工具，不是研究結果的保證，更不能倚靠電腦為你解決在分析數據時遇到的疑難。一些個人電腦專業統計軟件，例如 SPSS, SAS 等，需要花費以萬元計的金錢來購買，若只是要求簡單的統計，便不值得購買了，可以普通的計算機，或應用一些常見的軟件（例如是 Excel）來計算。

#### 4.1 不同的量度層次

應用統計方法來分析資料，其實是以數字為一些事物或意見作總結，並以一些統計圖表輔助，然後用文字解釋這些數字的意義。不論統計方法是簡單或複雜，都要求先把事物以不同的數字代表，這其實是一個編碼或量度的過程，以數學特性來區分，有四個量度層次：

1. 分類 (nominal level)：為事物依照一些準則來分類，就算是用數字來代表不同的類別，也不能作加、減、乘、除的運算，更不能比較優劣。常見的例子是性別（男、女）、國籍（如中國、英國、美國、日本）、宗教（如佛教、基督教、天主教、回教）等。數學上的特性便是「=」或「≠」的分別。
2. 分級 (ordinal level)：除了具備分類的特性外，更可互相比較高低等級，排列出彼此間的大小次序。常見的例子是會考科目成績的等級（A1, A2, B3, B4, C5, C6 .....）、一百公尺短跑的名次、教育程度（小學、中學、大學）等。數學上的特性是「=」或「≠」的分別、「>」或「<」的比較。
3. 等距 (interval level)：除了具備分類和分級的特性外，以等距層次量度出來相同數量的差異，其相差的程度也是一樣。常見的例子是量度溫度多少（攝氏或華氏），一百度跟五十度的相差，與五十度跟零度的相差，其差異的幅度也是一樣。數學上的特性是「=」或「≠」的分別、「>」或「<」的比較、「+」或「-」的計算。

4. 比率 (ratio level)：除了具備上述三個層次的特性外，更可計算出一個數值是另一個的多少倍。常見的例子是距離長短、重量多少，等等。必須小心的是，一些看似是比率層次的量度，其實只是等距量度，兩者的最重要不同，是比率層次要求一個「真正零點」，即要容許「一無所有」。因此，以攝氏或華氏量度的溫度多少，都只可以計算相差多少單位，卻不能說一百度較五十度熱多一倍，這是由於攝氏或華氏的零度，其實都是有溫度的，絕非毫無溫度。數學上的特性是「=」或「≠」的分別、「>」或「<」的比較、「+、-、×、÷」的計算，並要具備「真正零點」。

了解不同層次的量度特性，除了有助選擇適當的統計方法外，更可在設計問卷時預先考慮將會應用甚麼統計方法，務求充分利用所得資料並作正確的分析。

## 4.2 選擇適當的統計方法

統計方法的複雜程度，視同一時間處理多少項資料而定，最簡單的是單項，複雜一點的是兩項，最複雜的是多項分析，即同時處理三項或以上的資料。單項分析的例子是：受訪者中男女各佔多少、平均年齡多大、最多人信奉的宗教是甚麼。兩項分析的例子是：受訪者中男女的平均年齡有沒有差異、男女的教育程度有甚麼分別。多項分析的例子是：不同教育程度的男女，每月的平均工資有沒有差異。

由於多項分析要求具備豐富的基礎統計知識，這裏不會詳細介紹。統計方法分「描述統計」和「推論統計」兩大類，前者是描述樣本的情況，後者是以一些統計測試來判斷樣本結果能否推論到整體的研究對象。推論統計要求較複雜的統計技術和知識，故此不會在這裏介紹。

## 4.2.1 單項分析的常用統計方法

統計的作用是以具特別意義的數值來簡化描述一些集體的性質或特徵。在分析單項資料時，常用三類統計數字表達：頻數（frequency）或百分率（percentage）、集中程度（central tendency）、離散程度（dispersion）。不同量度層次的資料，適用的統計數值各有不同。從下表得知，量度層次愈高，適用的統計數值愈多。

統計數值	量度層次		
	分類	分級	等距 或 比率
頻數或百分率	✓	✓	✓
集中程度			
眾數（mode）	✓	✓	✓
中位數（median）		✓	✓
平均數（mean）			✓
離散程度			
最大值（maximum）		✓	✓
最小值（minimum）		✓	✓
標準差 （standard deviation）			✓

## 4.2.1 兩項分析的常用統計方法

若其中一項資料的量度層次是等距或比率，另一項是分類或分級，可在不同的類別或級別中，各自計算等距或比率資料平均值，然後比較差距，但要留意無論差距多少都只是針對樣本而論。

若兩項資料都是分類或分級，可用交互分類的方式表列出來，以頻率或百分率顯示不同交互分類的分佈情況。以百分率表示時，要先選擇其中一項資料做根據，按當中各類的小總數作計算的基礎。選擇那項資料做計算的基礎，視乎你的分析需要，一般是選兩者中的「原因」或發生時間較早的一項。以百分率表示的好處，是可直接指出基礎項資料的各類別中，另一項資料的分佈是否有不同。無論這些分佈的差異有多少，都只是針對樣本而論。

### 4.3 統計圖表

進行統計分析後得出的結果，利用適當的統計圖表去表達，不但清楚易明，更可省卻長篇大論。統計表是以清楚而簡單的格式表達數字，統計圖是把統計數字圖形化，以突出結果。常用的統計圖有條形圖 (bar graph)、圓餅圖 (pie graph)、矩形圖 (histogram) 頻數圖 (frequency polygon) 累積頻數圖 (cumulative frequency polygon)，等等。

#### 4.3.1 統計表

單項或兩項分析中，若資料的類別數目有限，即是分類或分級層次，適合以表列形式展示結果，例子如下：

##### 單項分析

表 1. 教育程度分佈

教育程度	(%)
小學或以下	16
中學	52
大學或以上	32
(N)	(150)

為求簡潔，統計表只有橫線而沒有直線，基本的項目包括：

1. 表號及標題 (表 1 教育程度分佈)
2. 所分析的資料是甚麼 (教育程度)
3. 出現百分比的符號 (%) 是表明所列數字的性質，若沒有表示，便會把數字視為頻數。
4. 樣本總數 (N)，用括號是為了清楚區分百分比和實數。
5. 若數據是從一些刊物或研究中取來使用，應在表下註明資料來源。

## 兩項分析

表 2. 男女的教育程度分佈 (%)

教育程度	性別	
	男	女
小學或以下	16	43
中學	52	36
大學或以上	32	21
合計	100	100
(N)	(85)	(65)

兩項分析的統計表跟單項分析的形式相同，但要注意：

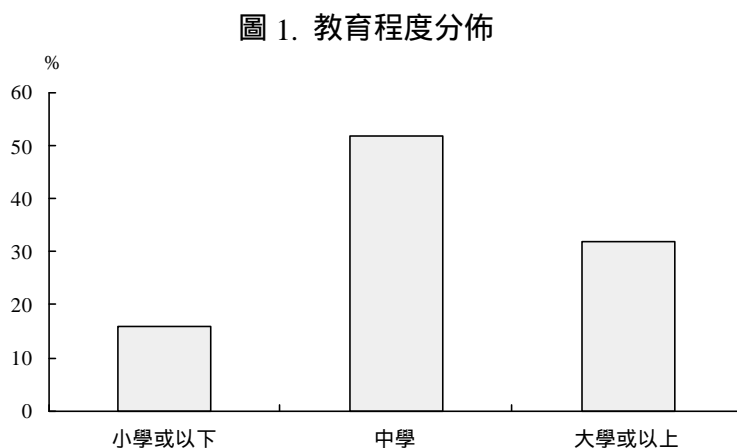
1. 把選做基礎分類的一項資料按欄 (column) 排列 (男、女)
2. 表內所計算的百分比，必須是垂直相加為 100%，即以每類性別的小總數 (男=85、女=65) 做計算基礎，並以「合計」註明。

## 4.3.2 統計圖

使用那類統計圖，視乎數據的量度層次而定。分類和分級的數據，適用的統計圖有條形圖、圓餅圖、頻數圖和累積頻數圖。等距或比率的數據，應按相等距離分組，然後以矩形圖顯示每組的數量。

繪製統計圖的原則跟編列統計表差不多，應以簡潔易明為主，若以彩色列印，慎防太多色彩使人看得眼花撩亂。統計圖內必備的項目就如以下的直條圖般包括：

1. 圖號及標題（圖1 教育程度分佈）
2. 數量的單位（%）
3. 註明每組圖列所代表的項目或種類（小學或以下、中學、大學或以上）



加插適量的統計圖可突出一些重要的數據，並增加報告的可讀性。其實不管是圖或表，都不能完全代替文字說明，切勿為求美觀而胡亂插入圖表，卻又不加說明。此外，不要把同一組數字既以圖又以表來重複顯示，以免令人看得生厭。有些人喜歡把一切不重要或不曾引述的數據全都繪製成圖表，當作附頁收錄在報告末。這樣做只會浪費紙張及時間，並無實際幫助或效用。