

## 2 統計圖表



**Question:**什麼是敘述統計？敘述統計的目的, 方法與要點為何？

**敘述統計：**透過統計表、統計圖與統計量，我們可以將資料有系統的整理，清楚明瞭地顯示出資料的特性

- 呈現資料的特徵、讓讀者容易閱讀其中的資訊。
- 將資料濃縮成統計表、統計圖、統計量。
- 無論是統計表、統計圖、統計量，都需要依照資料的特性，搭配適合的方法來處理與表示。

### 個案問題: 電影工業

電影工業是充滿競爭性的行業。每年有超過50家的製片公司製作300至400部的電影，每部電影在財務上的回收會有很大的差異。上映首週的票房、總票房、上映的電影院的數目、上映週數等，都是常用來衡量電影是否成功的指標。在名為2011Movies的檔案中蒐集了100部2011年出品的電影的樣本資料 (Box Mojo, March 17, 2012)。表2.19是其中的10部。

表2.19 100部電影的樣本資料

管理報告

1. 對表2.19中的四個變數的資料進行表格及圖形的彙整，並針對每個變數的彙整結果說明你對電影工業的觀察。
2. 繪製表示總票房收入及首週票房收入的關係的散布圖，並討論之。
3. 繪製表示總票房收入及上映電影院數目的關係的散布圖，並討論之。
4. 繪製表示總票房收入及上映週數的關係的散布圖，並討論之。

## 1. 統計表

### 1-1 統計表的編制

1. 在進行資料整理後，為了後續的資料分析，可先利用統計表來化繁為簡，彙整相關資訊。編制統計表的邏輯，按順序可以粗分為下面三個步驟：



**圖 2.1 編製統計表的邏輯**

- **分類**：資料經過適當的分類，以簡化龐大的資訊。
- **歸類計數**：將每個觀察個體歸屬到所屬的類別中，並且計算各類別的觀察個體數量。
- **列表**：按照特定格式，將歸類計數的結果編程簡單、有系統的表格型態。

**2. 統計表 (statistical table)**：將原始資料經過分類、歸類計數後，按特定的規則作成表格。

=> 製成統計表的功用：化繁為簡、清楚呈資料特徵、方便計算比較。

### 3. 統計表的格式

- 統計表要有標題、表身及資料來源三項。
- 標題置於統計表的上端。(時間、空間、特性)
- 表身分類項目置於統計表的左方或上方，按特定次序排列。
- 資料來源或附註說明置於統計表的下方。

Table X **← 表的編號**

*Proportion of Errors in Younger and Older Groups* **← 表的名稱**

Level of difficulty	Younger			Older		
	<i>n</i>	<i>M (SD)</i>	95% CI	<i>n</i>	<i>M (SD)</i>	95% CI
Low	12	.05 (.08)	[.02, .11]	18	.14 (.15)	[.08, .22]
Moderate	15	.05 (.07)	[.02, .10]	12	.17 (.15)	[.08, .28]
High	16	.11 (.10)	[.07, .17]	14	.26 (.21)	[.15, .39]

Note. CI = confidence interval. **← 註解**

### 4. 常用的統計表有四種：

- 次數分配表
- 相對次數分配表
- 累積次數分配表
- 累積相對次數分配表。

=> 依資料類別可分為質化（類別）資料編製步驟與量化（數字）資料編製步驟。

## 1-2 類別變數的統計表

### 1. 次數分配表

(1) 次數分配表 (frequency distribution table)

### (1)次數分配表(frequency distribution table)

- 次數分配表將資料依照類別分成若干組，然後計算各組資料的次數，以顯示資料分布狀況的表格。
- 次數分配表是檢視與描述資料常用的方法，用來了解每一個分類的頻率/次數。

### (2)類別資料編製步驟

當資料是類別變數，屬於名目尺度或順序尺度的資料時，可以編制離散型資料的次數分配表。

- Step 1:資料分類 ( 資料以類別分組 )
- Step 2: 歸類計數 ( 計算各類次數 )
- Step 3: 列表

#### 例題 2.1

捐血中心的醫務人員記錄 20 位捐血者的血型如下：

O	O	A	B	A
B	B	O	O	O
A	A	AB	A	B
O	O	A	AB	AB

如何編製次數分配表？

**解** 首先，確認當前資料的資料特性，血型是屬於名目尺度的資料，依目前的分類，共有四種可能性：O、A、B、AB，除了這四種以外，沒有其它可能。接著，確認 20 筆資料中，各分類發生的次數，也就是 O 型發生幾次，A 型發生幾次，B 型發生幾次，AB 型發生幾次，並把計數的結果記錄於表格上，就完成了離散型資料次數分配表的編製。

表 2.1 20 位捐血者血型之次數分配表

血型	人數
O	7
A	6
B	4
AB	3
合計	20



## 2. 相對次數分配表

### 1. 相對次數分配表(relative frequency distribution table)

- 相對次數分配表的編製步驟與次數分配表相似。
- 只需將次數分配表中之次數改為相對次數即可。

相對次數 = 各組次數 / 總次數

- $\text{相對次數} = \frac{\text{組次數}}{\text{總次數}}$
- 若將相對次數以百分比紀錄，則稱為相對次數百分比。

## 例題 2.4 (承例題 2.1)

試將次數分配表加入相對次數。

**解** 各組人數分別除以總人數，即得出各組的相對次數。編表如表 2.4 所示。

表 2.4 20 位捐血者血型之相對次數

血型	人數	相對次數
O	7	0.35
A	6	0.30
B	4	0.20
AB	3	0.15
合計	20	1.00



若是順序尺度資料，則可以有累積次數分配表與累積相對次數分配表。

### 3. 累積次數分配表

列出累加次數的表即稱為累積次數分配表。

- 累積次數係指將各組次數依次累加
  - 以下累積次數：由組值最小的次數累加到組值最大的次數。
  - 以上累積次數：由組值最大的次數累加到組值最小的次數。

### 4. 累積相對次數分配表

將上表中的累積次數改為累積相對次數，則可得累積相對次數分布表。

## 1-3 數量變數的統計表

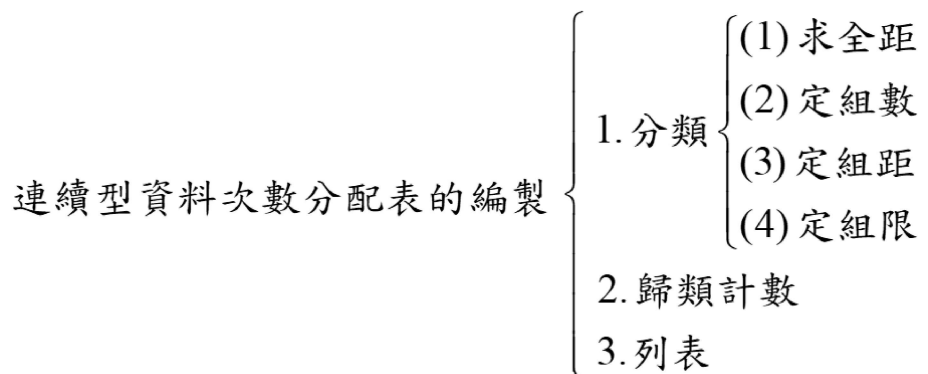
### 1. 次數分布表

#### (1) 次數分布表(frequency distribution table)

次數分布表將資料依照數量大小分成若干組，然後計算各組資料的次數，最後做成顯示資料分布狀況的表格。

#### (2) 屬量資料編製步驟

當資料是數量變數，屬於區間尺度或比例尺度的資料時，可以編制連續型資料的次數分配表。



• Step1: 資料分組

- (1)求全距(R)：將資料由小至大排序並計算出資料的全距。
  - 全距( $R$ ) = 最大值( $X_n$ ) – 最小值( $X_1$ )。
- (2)定組數(K)：依照資料多寡，組數不能太多也不能太少。
  - 依照經驗決定組數，K約為 5 ~ 20 組。
  - 滿足  $2^k \geq n$  的最小K值，n為資料總數。
  - Sturges公式： $K = 1 + 3.32 * \log_{10} n$  (課本用的分組公式)
- (3)定組距(W)：每組的大小距離
  - 組距約等於全距除以組數，即  $W = \frac{R}{K}$ 。
- (4)定組限：每組的組上界與組下界
  - 每組的最大值為組上界
  - 每組的最小值為組下界
  - 第一組下界必須小於等於最小值
  - 最後一組上界必須大於等於最大值。

• Step2: 歸類計數

- 1. 將原始資料依據組限歸入所屬各組
- 2. 計算各組劃記次數
  - 在歸類計數時，包含本組的組下限數值，但不包含本組的組上限數值。

• Step3: 列表

- 1. 最上面一列分別為變數跟次數
- 2. 最左邊一行依各分組順序排列。

### 例題 2.3

民國 108 年底佳佳經紀公司 50 名女性模特兒體重（單位：公斤）如下：

45	40	46	41	44	43	48	42	45	45
42	41	46	45	45	40	45	50	44	42
50	45	44	40	40	37	46	42	45	43
43	40	38	40	44	45	46	46	39	51
44	38	39	39	43	46	40	46	38	44

如何編製次數分配表？

**解** 在這個例子中，體重資料雖然是整數，但是並不是離散型的資料，而是連續型資料中比例尺度的資料。故編製次數分配表時，應該編製連續型資料的次數分配表。按照編製統計表的邏輯，編製程序如下：

(1) 分類

(a) 求全距 ( $R$ )： $R = \text{最大值} - \text{最小值} = 51 - 37 = 14$

(b) 定組數 ( $K$ )： $K = 1 + 3.32 \log_{10} 50 = 1 + 3.32(1.7) = 6.644$ ，取 7 組

(c) 定組距 ( $H$ )： $H = R/K = 14/7 = 2$  (公斤)

(d) 定組限：以最小值 37 為最小一組的下限，從 37 開始每隔 2 公斤為一組，直至能包括最大數 51 為止，則最後一組為 49 ~ 51 為止。

(2) 歸類計數

如下表所示。本例中之 39 分別屬於第 1 組的組上限及第 2 組的組下限，依習慣歸入第 2 組。本例中之「51」將歸類於體重 49 到 51 公斤的這一組，按照前述每組範圍來說，「51」不應該歸類於該組。但因「51」是最後一筆資料，故將其歸類於 49 到 51 公斤這一組是方便的做法；當然，若再多開一組 51 到 53 公斤，也沒有錯。這兩種做法對於次數分配的影響都很小。

**表 2.2** 108 年底佳佳經紀公司 50 名女性模特兒體重歸類計數

體重	劃記	次數
37~39	正	4
39~41	正 正	10
41~43	正 一	6
43~45	正 正	10
45~47	正 正 正 一	16
47~49	一	1
49~51	下	3
合計		50

(3) 列表：

完整之次數分配表如下：

**表 2.3** 108 年底佳佳經紀公司 50 名女性模特兒體重之次數分配表

體重	次數
37~39	4
39~41	10
41~43	6
43~45	10
45~47	16
47~49	1
49~51	3
合計	50



## 2. 相對次數分布表

- 相對次數分布表的編製步驟與次數分布表相似
- 只需將次數分布表中之次數改為相對次數即可。
  - $\text{相對次數} = \frac{\text{各組次數}}{\text{總次數}}$

## 3. 累積次數分布表

- 累積次數係指將各組次數依次累加
- 列出累加次數的表即稱為累積次數分配表。

## 4. 累積相對次數分布表

- 將累積次數改為累積相對次數，可得累積相對次數分布表。



## 例題 2.5 (承例題 2.3)

試將次數分配表加入相對次數及累積次數，並說明 45 公斤以下（不含 45 公斤）及 47 公斤以上（含 47 公斤）的女模特兒分別有多少人，所占百分比各為多少？

**解** 各組人數分別除以總人數，即得出各組的相對次數、百分比。將各組人數、百分比分別以下累加、以上累加，即得出各組的以下或以上累積次數、累積次數百分比。編表如表 2.5 所示。

表 2.5 民國 108 年底佳佳經紀公司 50 名女性模特兒體重之相對次數及累積次數

體重	次數	相對次數	百分比	以下累積 次數	以下累積 次數百分比	以上累積 次數	以上累積 次數百分比
37~39	4	0.08	8%	4	8%	50	100%
39~41	10	0.20	20%	14	28%	46	92%
41~43	6	0.12	12%	20	40%	36	72%
43~45	10	0.20	20%	30	60%	30	60%
45~47	16	0.32	32%	46	92%	20	40%
47~49	1	0.02	2%	47	94%	4	8%
49~51	3	0.06	6%	50	100%	3	6%
合計	50	1.00	100%				

（註：第一組的相對次數計算方式  $4 \div 50 = 0.08$ ，其餘類推）

由以下累積次數欄可知，45 公斤以下（該組上限）之人數共 30 名，占該經紀公司人數 60%。

由以上累積次數欄可知，47 公斤以上（該組下限）之人數為 4 人，占 8%。



=> 無論是離散型資料或連續型資料，編制次數分配表的目的是瞭解各組間的相對次數，這是機率分配 (probability distribution) 的基礎。

## 2. 統計圖

### 2-1 一張圖勝過千言萬語

#### 1. 文不如表，表不如圖

如果我們能將統計表以簡單的統計圖來表示，將更能觀察出統計資料的特徵。

- 依照資料不同的性質，選用不同的圖形，以表示資料各項的統計特徵。
  - 易於閱讀
  - 易於了解比較分析



- 易引起閱讀圖的人興趣

2. 依據資料的型態，常用的統計圖如下：

- 類別資料的統計圖
  - 長條圖(bar chart)
  - 圓餅圖(pie chart)
- 數量資料的統計圖
  - 直方圖 (histogram)
  - 多邊形圖 (polygon)
  - 肩形圖 (ogive) = 累積次數分配圖(cumulative frequency chart)

3. 先有統計表，再依統計表畫出統計圖。

## 2-2 類別橫向資料統計圖

### 1. 長條圖(bar chart)

長條圖係依資料的類別及次數（統計表），繪製成長條，而其長條的高度與次數成正比。

- 各長條間不可緊鄰，要有適當間隔。
- 長條越長，代表發生次數或相對次數更高，即該分類發生的比例更高。
- 長條圖有縱式和橫式兩種。
- 順序尺度資料的各分類排列順序不可以任意更換，名目尺度資料則可以。

### 2. 圓餅圖

圓餅圖將資料依各類別次數（比例）多寡，將一個圓分割成幾個扇面而成。

以圓形代表全部面積，各分類相對次數百分比代表圓形面積的一部分（占圓形面積的百分比）之統計圖。

## 例題 2.7 ( 承例題 2.1 )

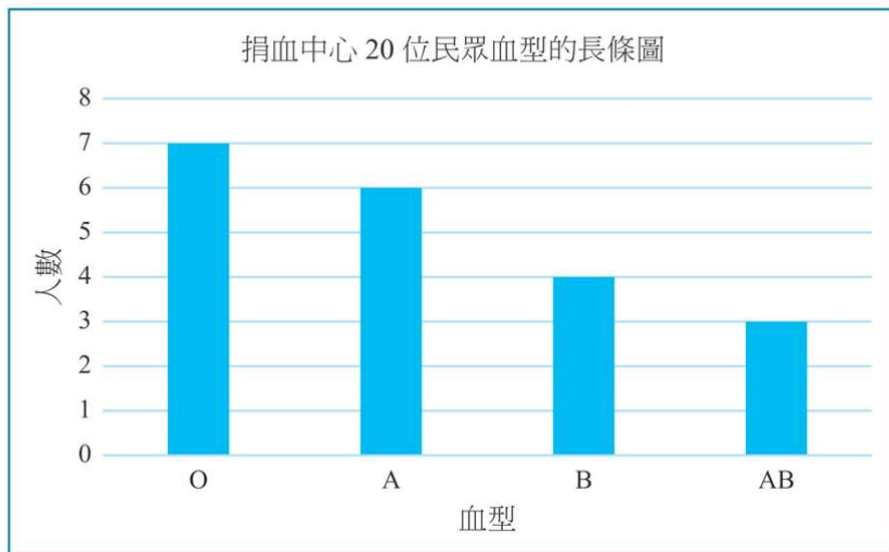
請以 Excel 完成捐血中心 20 位民眾血型之相關統計圖。

**解** 延續例題 2.2 次數分配表 Excel 操作，相關統計圖的操作如下：

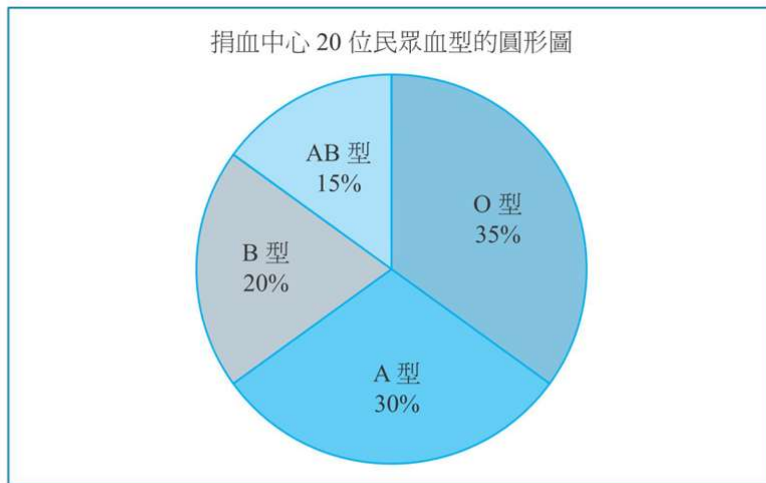
**步驟 1：**繪製統計圖，先選取 B1 到 B4 與 C1 到 C4。

**步驟 2：**滑鼠點選 <插入> 選單，在 <圖表> 區點選所需的圖樣選項。例如長條圖與圓形圖，都可以在圖樣中找到（Excel 中直條圖圖樣為本書的長條圖，圓形圖樣為本書的圓形圖），滑鼠點選圖樣後還可以對圖形做細部的調整，讓圖形更為美觀。

直條圖結果：



圓形圖結果：



需注意的是長條圖與圓形圖操作後的成品都是半成品，需加以修改、標註說明才算完整，例如長條圖的縱軸名稱、橫軸名稱、圖形標題；圓形圖的各扇形的人數、所占比例、圖形標題，使得圖形清楚易讀，然後才能使用。



## 例題 2.8

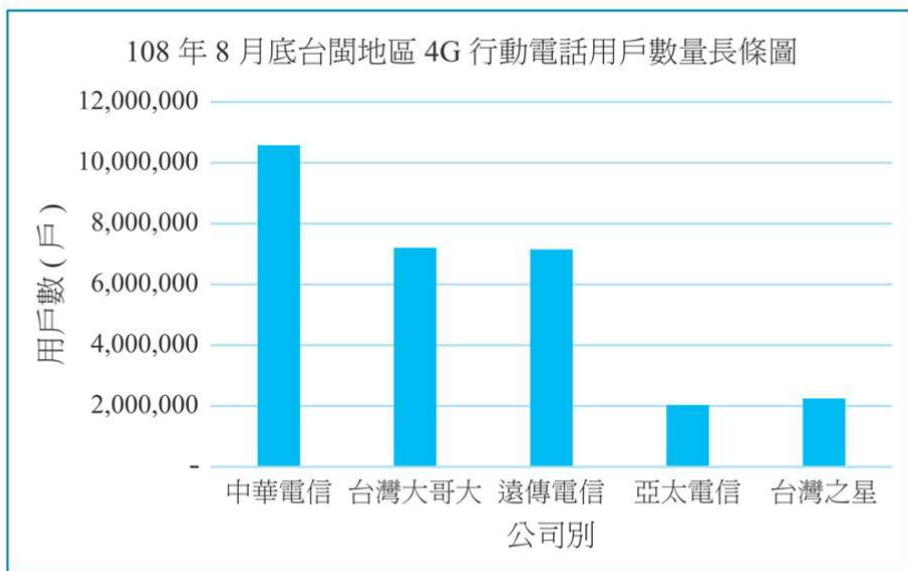
下表為 108 年 8 月底台閩地區 4G 行動電話用戶數量表，試繪製相關圖形。

公司別	用戶數	百分比
中華電信	10,613,121	36.23%
台灣大哥大	7,224,581	24.66%
遠傳電信	7,171,995	24.48%
亞太電信	2,034,746	6.95%
台灣之星	2,250,527	7.68%
合 計	29,294,970	100%

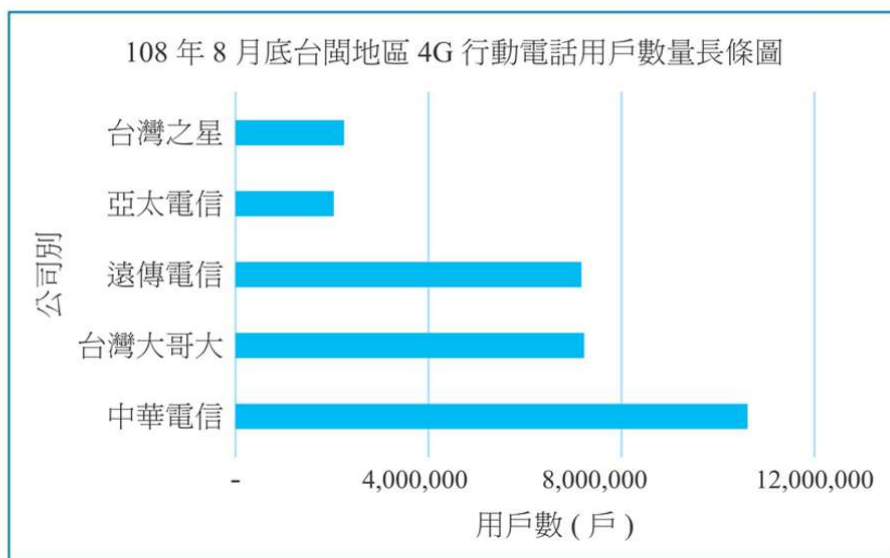
資料來源：中華民國統計資訊網

**解** 該表內之資料，是離散型資料，宜選用離散型資料所適合的圖樣，為長條圖或是圓形圖。變數是公司別，各分組有中華電信、台灣大哥大、遠傳電信、亞太電信及台灣之星。

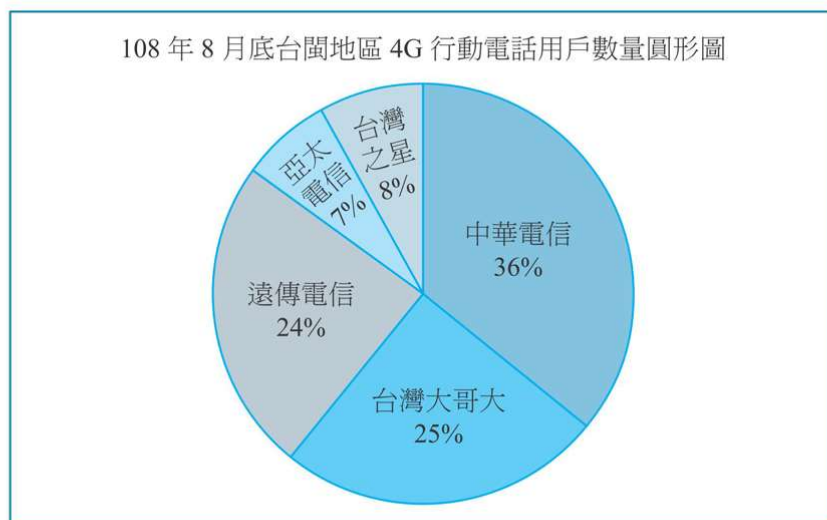
(1) 縱式長條圖：



(2) 橫式長條圖：



(3) 圓形圖：



## 2-3 數量資料的統計圖

次數分配圖乃是一種統計圖，用以表示次數分配表中次數分配的情況。因此，它只是將次數分配中的統計資料由表格形式改以圖形來陳示，故欲繪製次數分配圖之前，必須先編好次數分配表。

### 1. 直方圖 (histogram)

依照連續型資料次數分配表之分組數量，以若干長條來表示分組數量之次數的多寡所繪製之圖形，長條與長條緊密相聯。

- 適用於連續資料。
- 長條與長條間緊密相連。
- 各分組之間的順序不可任意變換。

~ 建立直方圖的步驟：

Step1: 找出最小值與最大值

Step2: 算全距

Step3: 算組數

Step4: 算組距

Step5: 計算每組的次數分配

Step6: 建立直方圖

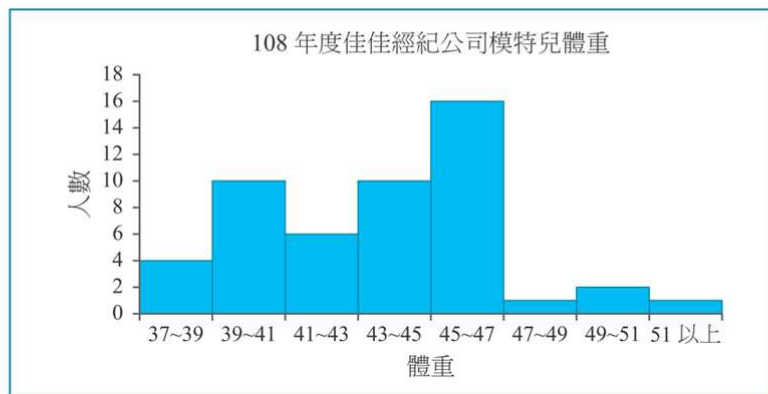


## 例題 2.9 (承例題 2.3)

請以 Excel 完成 108 年度佳佳經紀公司模特兒體重之次數分配表及相關統計圖。

**解** 此例題是模特兒體重的資料，屬連續型資料，因此編製次數分配表時必須先分類。按照例題 2.3 所提到的分類規則，共分 7 組，組距 2 公斤，第一組為 37 公斤以上，未滿 39 公斤，第二組為 39 公斤以上，未滿 41 公斤，以此類推。前文曾經提到，統計學中的規則都是組下界以上，組上界未滿，這點與 Excel 的運算方式不同，Excel 的做法是不含組下界，包含組上界，例如 41 公斤在 Excel 運算中會被分類於 39 到 41 公斤這一組，而不是我們所希望的 41 到 43 公斤這一組，為了讓這個軟體可以配合我們的目的，所以在輸入組界的時候要做些微調，請看下面的操作過程。

稍作修改後，完成圖、表如下：



108 年度佳佳經紀公司模特兒體重次數分配表

組界	人數	組界	人數
37~39	4	45~47	16
39~41	10	47~49	1
41~43	6	49~51	2
43~45	10	51 以上	1



~ 相對次數直方圖: 將次數直方圖中之次數改為相對次數即可。

~ 這題請繼續畫多邊型圖與肩型圖

## 2. 多邊形圖(polygon)

用線段連起以組中點 ( 上界+下界 ) / 2 ) 為橫(x)座標，頻率(組次數)為縱(y)座標的一種圖形。也就是說，頻率是圖形上每一點的高度。(為構成封閉曲線圖，在資料的前後須各再加組數為0的一組。)

- 可由直方圖繪出

以各組組中點為橫座標，縱座標為各組次數

- 以各組組中點為橫坐標，縱坐標不各組次數。
- 將各組中點以直線相連所繪製的圖形。

3. 累積次數分配圖 是一種表達各組別累計相對次數(累加頻率)的圖形，又稱肩形圖 (ogive)。

- 以下累積次數曲線圖
- 以上累積次數曲線圖

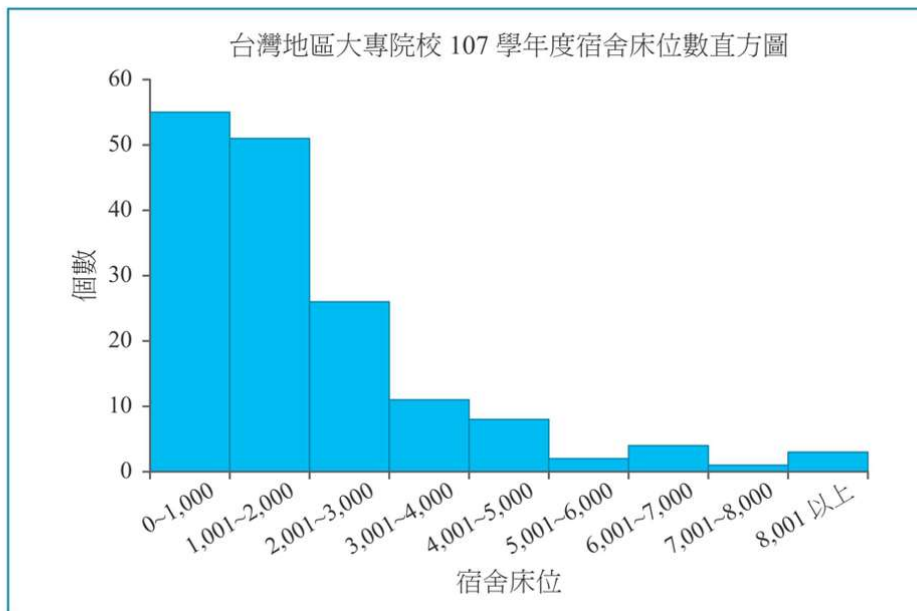
### 例題 2.10

下表為台灣地區大專院校 107 學年度宿舍床位數，試繪製直方圖與累積次數分配圖。

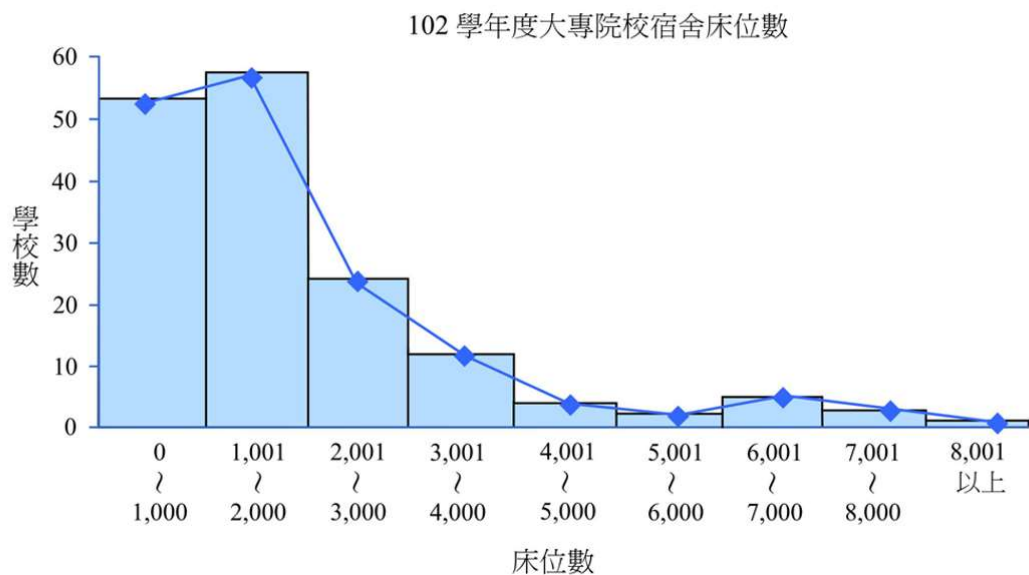
床位數	學校數	以下累積次數	以上累積次數
0~1,000	55	55	161
1,001~2,000	51	106	106
2,001~3,000	26	132	55
3,001~4,000	11	143	29
4,001~5,000	8	151	18
5,001~6,000	2	153	10
6,001~7,000	4	157	8
7,001~8,000	1	158	4
8,001 以上	3	161	3
合計	161		

資料來源：中華民國統計資訊網

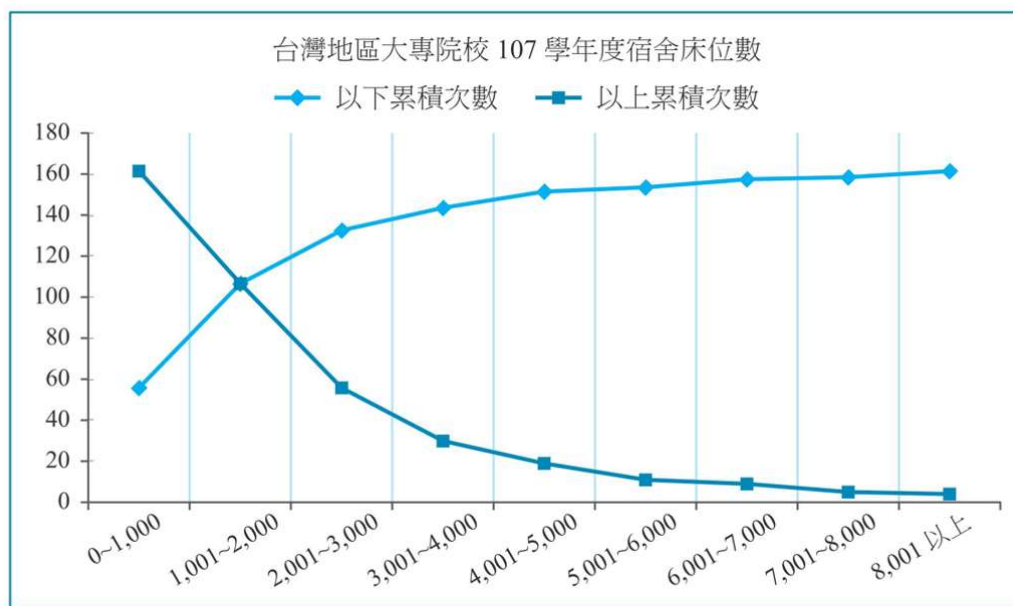
**解** (1) 直方圖：



## (2) 次數多邊圖



## (2) 累積次數分配圖：



## 2-4 其他圖形

縱向資料是按照時間的順序收集而來，宜選用時間曲線圖，俗稱（曲）線圖。線圖是以曲線之升降、傾斜等表示統計資料數量上之變化情形，又稱為歷史線圖或趨勢圖。

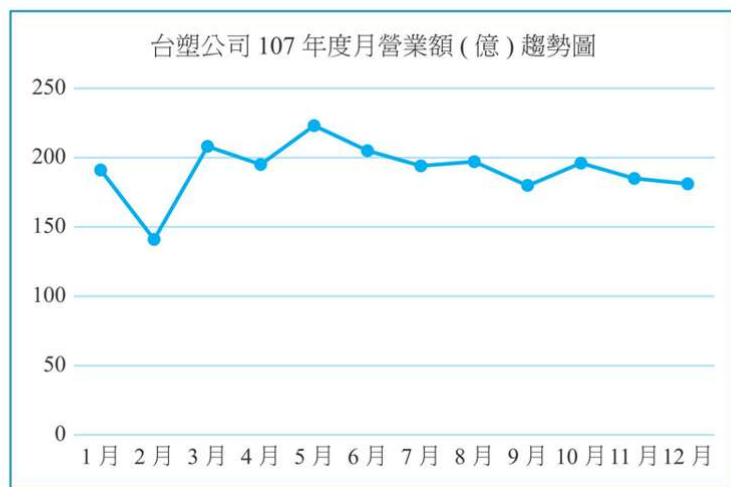
## 例題 2.6

下列為台塑公司 107 年度的月營業額（單位：億元），請以 Excel 繪製趨勢圖表示之。

月份	1 月	2 月	3 月	4 月	5 月	6 月	7 月	8 月	9 月	10 月	11 月	12 月
營業額 ( 億 )	191	141	208	195	223	205	194	197	180	196	185	181

資料來源：台塑公司營業報告。

最後將圖形稍做修飾，即可完成趨勢圖。



圖形的縱軸，是台塑公司的營業額（億），橫軸是月份，以目前這個例子來說，總共 12 個月份。將各個月份的營業額點在直角座標系上，然後將 12 個點以直線相連，就能繪製出趨勢圖。

由此趨勢圖看出在 2 月份時，營業額下降不少，其他月份的營業額是較為穩定的。



## 2-5 次數分配的型態

由於相對次數可以視為各分組發生的機率，將各組發生的機率放在一起看，則這些機率可視為此資料之機率分配。

- 單峰對稱分配：中央部分有一高峰，最高點左右兩邊次數逐漸減少，成為對稱鐘型。
- 單峰偏斜分配：高峰不在中央，呈現左右不對稱的型態。
  - 右偏分配：高峰在偏左方，長尾在右方。
  - 左偏分配：高峰在偏右方，長尾在左方。

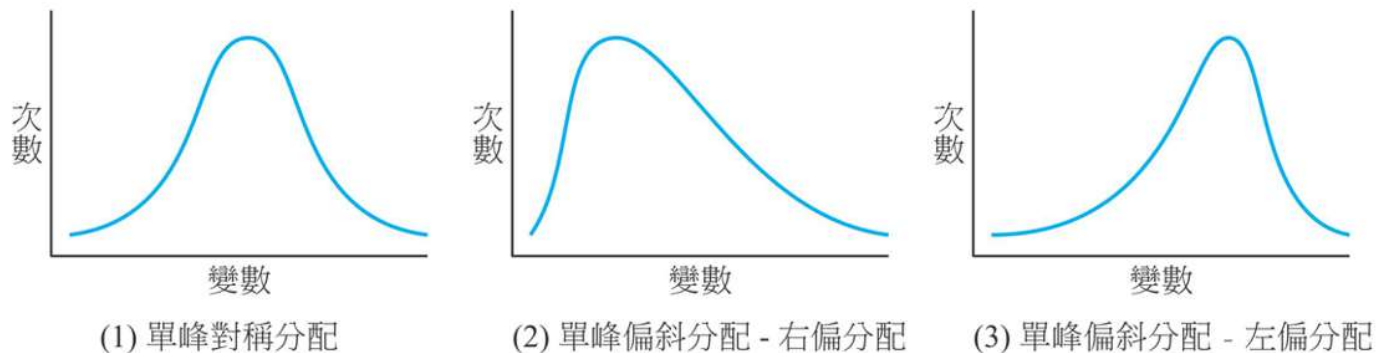


圖 2.2 機率分配型態

~ 曲線高的時候，代表發生次數較為頻繁，也就是機率比較大，曲線低的時候，發生機率較小。

## 補充說明：資料視覺化(Data Visualization)

- 商業分析(Business Analytics): 以數據分析為基礎做決策 + 用圖表進行溝通 => 說故事

實例：好的統計圖能凸顯資料分析的結果

### 單獨任務

某 T 醫院資訊組所收到的報修單，按月之統計 (2020年)

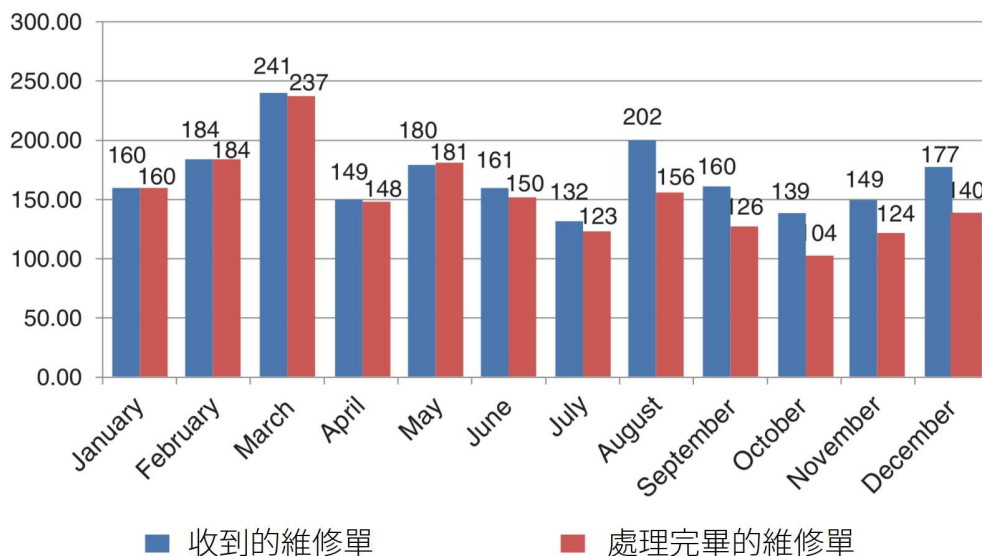
情境：您想要提醒主管，人力短缺的問題。您在院務會議中只有兩分鐘時間可以報告。

	A	B	C	
1	月份	收到數量	處理完畢	
2	Jan	160	160	
3	Feb	184	184	
4	Mar	241	237	
5	Apr	149	148	
6	May	180	181	
7	Jun	161	150	
8	Jul	132	123	
9	Aug	202	156	
10	Sep	160	126	
11	Oct	139	104	
12	Nov	149	124	
13	Dec	177	140	

單位：報修單數量 ( 單 )

您已經會做了：用軟體呈現資料圖表

本組 2020 年維修單處理數量一覽



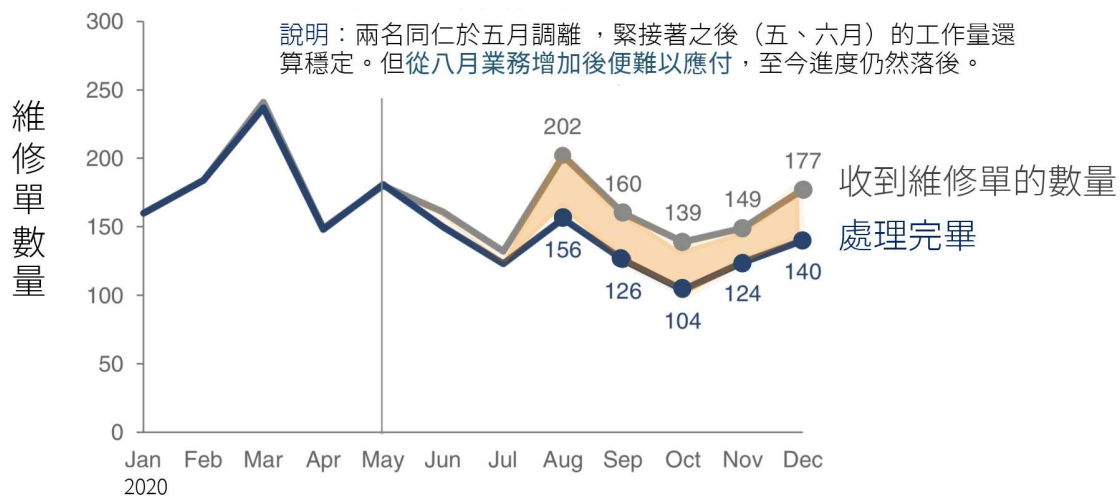
9



更好的數據溝通力

建議：請再雇用至少一位全職員工 (以填補 2020 年五月的職缺)

本組 2020 年維修單處理數量一覽



資料來源與補充說明：內部數據 (2020/12/31)，平均每人處理量與解決時間已經完成更深入的分析，若您需要可以向 ○○ 組員 (代號：9527) 索取，作為申請之參考

10