

單元四 資料的分布量

課文A： 四分位數與盒狀圖

在前面的單元當中，我們學了一些統計量：「眾數」、「平均數」、「中位數」，它們都可以拿來代表整組數據。

如果我們想知道一組資料的分布情況（集中或分散），該如何進行呢？
以下分別為班上的兩次數學段考成績（分）：

第一次 (12個)	59、60、64、65、67、70、72、74、75、76、78、80
第二次 (12個)	58、61、67、68、69、71、71、72、72、73、77、81

初步我們可以計算這兩組數據，最大值及最小值相差多少：

$$\text{第一次段考 } 80 - 59 = 21, \text{ 第二次段考 } 81 - 58 = 23。$$

而這個最大數值與最小數值的差我們就稱為「全距」。



我們知道「第一次段考的全距是 21、第二次段考的全距是 23」，
所以應該第二次段考的資料比較分散。

小安

如果只單純從資料的最大值和最小值去做推論，好像不是那麼的準確，
所以我們會希望想要再多一點對資料的描述。

現在我們再將資料多做分割，根據資料的個數等分成 4 等分，例如第一次段考，想想看要如何將這組資料分成 4 等分？

59、60、64、65、67、70、72、74、75、76、78、80

如上所示，要將這 12 個資料由小排到大分成 4 等分，需要 3 個等分點，這三個等分點的值分別會是多少呢？

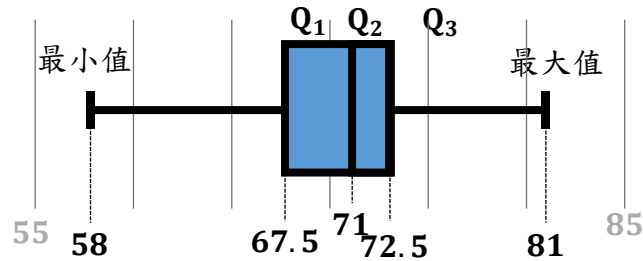
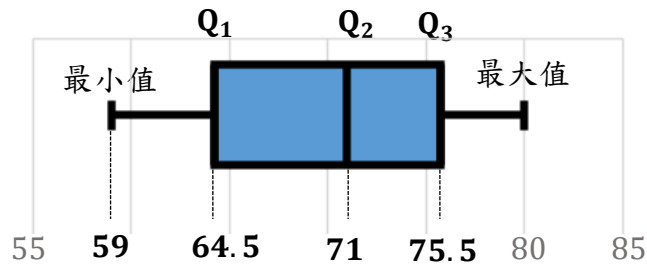
第一個等分點的值會在 64 與 65 中間，我們取 64 與 65 的平均值 64.5 當作等分點，稱為第 1 四分位數(可以記做 Q_1)。

第二個等分點稱為第 2 四分位數(記做 Q_2)，也就是中位數 71。

第三個等分點的值會在 75 與 76 中間，我們取 75 和 76 的平均值 75.5 當作等分點，稱為第 3 四分位數(記做 Q_3)。

59、60、64、65、67、70、72、74、75、76、78、80
64.5 71 75.5

我們利用直線標示出最小值、 Q_1 、 Q_2 、 Q_3 、最大值，再將 Q_1 、 Q_2 、 Q_3 以長方形連接起來，畫出第一次段考的統計圖形。用同樣的方式，畫出第二次段考的統計圖形，將兩組資料的圖放在一起比較看看：



這個圖形我們稱它為「盒狀圖」。從盒狀圖可以知道最小值、 Q_1 、 Q_2 、 Q_3 、最大值，也可以看出大致的分布情形。

第一段考分布的比較均勻，而第二段考的中間 $\frac{1}{2}$ 的分布比較集中、前 $\frac{1}{4}$ 及後 $\frac{1}{4}$ 分布的比較分散。

另外，我們可以計算第 1 四分位數與第 3 四分位數的差，像是

第一段考的 $Q_3 - Q_1 = 19 - 7 = 12$ ，表示中間 $\frac{1}{2}$ 的分布範圍大小有 12，

第二段考的 $Q_3 - Q_1 = 15.5 - 10.5 = 5$ ，中間 $\frac{1}{2}$ 的分布範圍大小有 5，

因此就可以說第二段考的中間 $\frac{1}{2}$ 分布比第一段考的中間 $\frac{1}{2}$ 分布還要密集。

第 1 四分位數與第 3 四分位數的差值我們稱為「四分位距」。現在試著練習看看一題練習題！

例題一：老師將班上分成兩組，兩組的人數分別是 10 人、11 人，下表為各組每人投籃 15 次中的進球數，請求各組進球數的第 1 四分位數、中位數與第 3 四分位數，畫出盒狀圖，並比較各組的分散趨勢。

第一組	0、3、7、11、12、13、14、14、15、15
第二組	0、1、2、3、4、5、6、7、11、13、15

※解題思維：資料分成一半（正中間）是中位數，一半的數再取正中間就是第 1 四分位數與第 3 四分位數。

解：先來看第一組的進球數：

0、3、7、11、12、13、14、14、15、15

我們先將整組資料分成一半，

$10 \div 2 = 5$ ，代表中位數是由小到大的 5 和 6 個資料的平均：

0、3、7、11、12、13、14、14、15、15
|
12.5

再將前半部分的 5 個資料分成一半，選擇由小到大第 3 個資料當第 1 四分位數：

0、3、7、11、12、13、14、14、15、15
| |
7 12.5

再將後半部分的 5 個資料分成一半，選擇由小到大第 8 個資料當第 3 四分位數：

0、3、7、11、12、13、14、14、15、15
| | |
7 12.5 14

這樣就找出 $Q_1 = 7$ 、 $Q_2 = 12.5$ 、 $Q_3 = 14$ ，再搭配最小值 = 0 和最大值 = 15 就可以畫出盒狀圖了！

再來看第二組的進球數：

0、1、2、3、4、5、6、7、11、13、15

我們一樣先找出中位數，將整組分成一半：

0、1、2、3、4、5、6、7、11、13、15

再將前半部分的 5 個資料分成一半當第 1 四分位數：

0、1、2、3、4、5、6、7、11、13、15

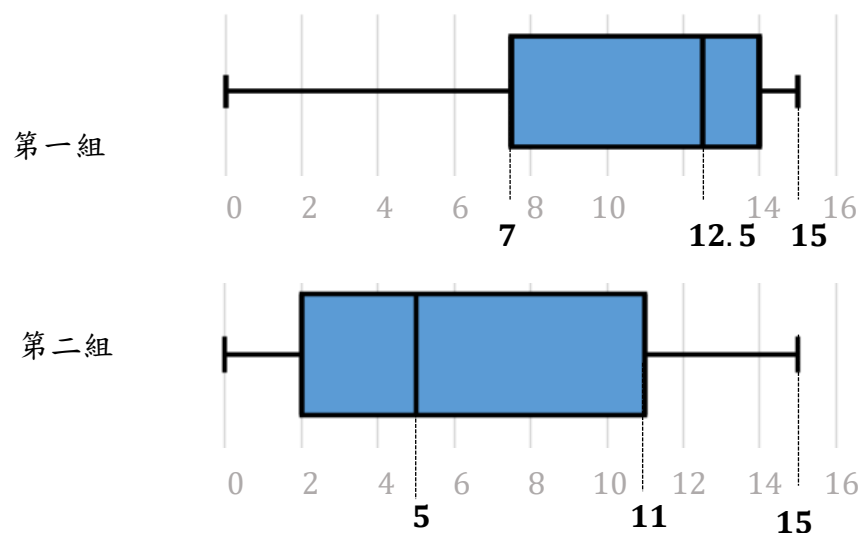
後半部分的 5 個資料也分成一半當第 3 四分位數：

0、1、2、3、4、5、6、7、11、13、15

這樣就找出 $Q_1 = 2$ 、 $Q_2 = 5$ 、 $Q_3 = 11$ ，再搭配最小值 = 0 和最大值 = 15 就可以畫出盒狀圖了！

根據前面所求出來的第 1 四分位數、中位數、第 3 四分位數，再搭配最小值及最大值，就可以畫出各組的盒狀圖了。

若要同時分析討論兩組以上性質相近的資料時，可以將其畫在一起，這樣可以更清楚的看出這些資料分布的情形：



試著說說看，從這兩個盒狀圖可以看出來這兩組的分布有什麼差異嗎？

(1) 這兩組的數據中，第一組進球數的中位數高於第二組。

(2) 這兩組的數據中，第一組進球數的四分位距小於第二組；可得知第一組進球數中間 50% 的人密集、第二組進球數中間 50% 的人分散。

(3) 從盒狀圖可以看出，將這兩組的數據由小排到大，

第一組資料的後 $\frac{1}{2}$ 分布比前 $\frac{1}{2}$ 還要密集；

第二組資料的前 $\frac{1}{2}$ 分布比後 $\frac{1}{2}$ 還要密集。

★省思：

其實四分位數，也就是中位數的延伸，之前在單元三找中位數時：

如果總共有 11 個數，那麼 $11 \times \frac{1}{2} = 5.5$ ，無條件進位 = 6，中位數就是從小排到大的第 6 個數。

如果總共有 10 個數，那麼 $10 \times \frac{1}{2} = 5$ ， $5 + 1 = 6$ ，中位數就是從小排到大的(第 5 個數 + 第 6 個數) ÷ 2。

我們可以看出來如果：總數 $\times \frac{1}{2}$ 有小數，中位數就是無條件進位。

總數 $\times \frac{1}{2}$ 是整數，中位數就是兩個相加除以 2。

所以例題一第一組的四分位數也可以這樣算：

0、3、7、11、12、13、14、14、15、15
7 12.5 14

第 1 四分位數 $10 \times \frac{1}{4} = 2.5$ ，有小數就無條件進位 = 3，

也就是第 3 個資料 = 7，當作第 1 四分位數。

第 2 四分位數 $10 \times \frac{1}{2} = 5$ ，是整數就兩個相加除以 2，

也就是第 5 和 6 個資料相加除以 2=12.5，當作第 2 四分位數。

第 3 四分位數 $10 \times \frac{3}{4} = 7.5$ ，有小數就無條件進位= 8，

也就是第 8 個資料=14，當作第 3 四分位數。

而例題一的第二組的四分位數就可以這樣算：

0、1、2、3、4、5、6、7、11、13、15

第 1 四分位數 $11 \times \frac{1}{4} = 2.75$ ，有小數就無條件進位= 3，

也就是第 3 個資料=2，當作第 1 四分位數。

第 2 四分位數 $11 \times \frac{1}{2} = 5.5$ ，有小數就無條件進位= 6，

也就是第 6 個資料=5，當作第 2 四分位數。

第 3 四分位數 $11 \times \frac{3}{4} = 8.25$ ，有小數就無條件進位= 9，

也就是第 9 個資料=11，當作第 3 四分位數。

有時候當資料數據非常多時，只會呈現整理過的圖表，這時候我們就要藉由圖表的資訊看出次數，並找出 $\frac{1}{4}$ 位置的值！

例題二：下表是班上每位同學家庭人口數的次數分配表，請求出班上同學家庭人口數的三個四分位數。

家庭人口數(口)	3	4	5	6	7	8	9
學生人數(人)	7	6	3	1	1	2	1

※解題思維：先計算班上總共有多少人，再乘以 $\frac{1}{4}$ 、 $\frac{1}{2}$ 、 $\frac{3}{4}$ 找出三個四分位數。

解：班上總共 $7 + 6 + 3 + 1 + 1 + 2 + 1 = 21$ 位同學。

$21 \times \frac{1}{4} = 5.25$ ，有小數就無條件進位=6，所以由小到大的第 6 個資料當

第 1 四分位數：

想想看，從小排到大的第 6 個資料會是哪個數字呢？

家庭人口數(口)	3	4	5	6	7	8	9
學生人數(人)	7	6	3	1	1	2	1

由小到大的第 6 個資料為 3。

$21 \times \frac{1}{2} = 10.5$ ，有小數就無條件進位=11，所以由小到大的第 11 個資料

當中位數：可以利用累積次數來看！

家庭人口數(口)	3	4	5	6	7	8	9
學生人數(人)	7	6	3	1	1	2	1
累積人數(人)	7	13	16	17	18	20	21

從累積次數來看，由小到大的第 11 個資料為 4。

$21 \times \frac{3}{4} = 15.75$ ，有小數就無條件進位=16，所以由小到大的第 16 個資料

當第 3 四分位數：從累積次數來看，由小到大的第 16 個資料為 5。

重點提問

1.根據上面的課文，請用自己的話解釋「四分位數」的意思，它又有什麼意義？

2.如何找到一組資料的四分位數？有哪些步驟？

3.想要畫出一組資料的盒狀圖需要資料中的哪些數據點？

• 隨堂練習：

1. 以下為班上女生身高（公分）的資料：

148、150、151、153、157、159、159、160、162、164、165、165、167、168、170、172

(1) 求班上女生身高的第 1 四分位數 Q_1 、第 2 四分位數 Q_2 和第 3 四分位數 Q_3 分別為何？

(2) 請畫出班上女生身高的盒狀圖。

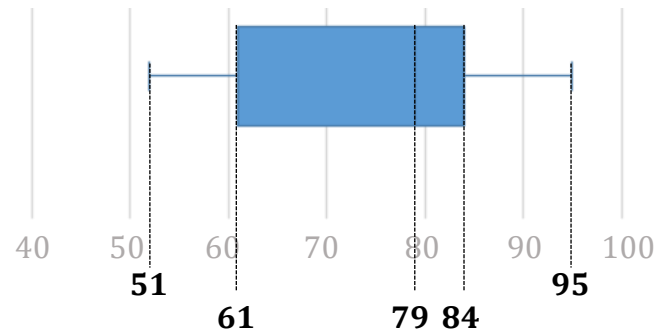
2. 下表為班上同學的投籃進球數：

進球數	0	1	2	3	4	5
次數(人)	1	5	6	7	4	3

(1) 班上投籃進球數的全距、第 1 四分位數 Q_1 、第 2 四分位數 Q_2 、第 3 四分位數 Q_3 和四分位距分別為何？

(2) 請畫出班上投籃進球數的盒狀圖。

3.下圖是班上成績的盒狀圖：



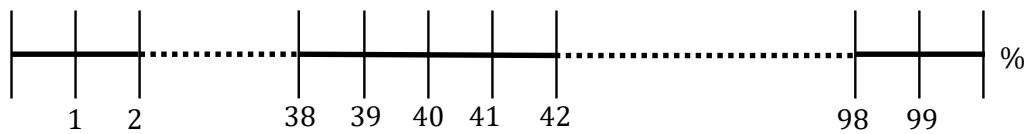
回答下列問題：

- (1)這次班上成績的全距是多少分？
- (2)這次班上成績的四分位距為多少分？
- (3)成績在 50 ~ 79 分大約占全班人數的百分比是多少？

課文B： 百分位數

單元四的課文 A 有介紹了四分位數，並利用這些數字就描述出整組資料的分布。當中可以知道資料的 前 $\frac{1}{4}$ 、前 $\frac{1}{4} \sim \frac{2}{4}$ 、前 $\frac{2}{4} \sim \frac{3}{4}$ 、後 $\frac{1}{4}$ 大概的範圍，但是當數據非常多的時候，我們可能會想要知道更細的情形，這時候就要將資料切得更細了。

通常我們會將全部的資料個數切分成 100 等份，這時候就會需要 99 個等分點了，如下圖：



如同四分位數一樣，我們將這些分點就是「百分位數」，而第 m 個等分點就稱為「第 m 百分位數」，以 P_m 表示，例如第 40 個等分點就稱為「第 40 百分位數」，以 P_{40} 表示。而計算也和四分位數一樣：

總數 \times 比例有小數，就無條件進位
總數 \times 比例是整數，就兩個相加除以 2

現在我們就來練習找出一些百分位數！

例題一：數成國中三年級有 200 位學生，數學段考成績由低排到高如下表：

10	14	16	18	18	21	23	23	24	24
25	25	27	27	28	30	31	31	32	32
35	35	35	35	35	37	38	38	38	38
40	40	40	41	42	42	43	44	44	44
45	45	45	45	47	48	48	49	49	49
49	50	50	50	50	50	52	52	53	53
55	55	55	55	57	57	57	58	58	58
60	60	60	60	60	61	61	61	62	62
62	62	62	65	65	65	65	65	65	66
66	66	66	67	68	68	68	68	68	69
70	70	70	72	72	72	72	73	73	73
73	73	73	73	75	75	75	75	75	75
75	78	78	78	78	78	78	79	79	79
79	79	79	79	80	80	82	82	82	84
84	84	84	84	85	85	85	85	85	85
85	85	85	85	85	85	85	86	86	86
86	86	86	88	88	88	88	89	89	90
90	90	90	92	92	92	93	94	94	94
94	95	95	95	95	95	97	97	97	98
98	98	98	100	100	100	100	100	100	100

(1)請問三年級學生數學成績的 P_{25} (第 25 百分位數)是多少分？

(2)請問三年級學生數學成績的 P_{60} (第 60 百分位數)是多少分？

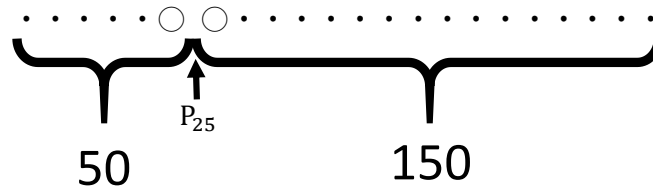
※解題思維：

第一四分位數是 $\frac{1}{4} = \frac{25}{100} = 25\%$ ，也就是 P_{25} 。把人數乘以 25% 與 60%

解(1)： $200 \times 25\% = 50$ ，是整數就兩個相加除以 2，所以第 25 百分位數是由小到大的第 50 與第 51 個資料的平均。

第 50 資料為 35 分、第 51 個資料為 37 分

$$P_{25} = \frac{35+37}{2} = 36 \text{ (分)}$$

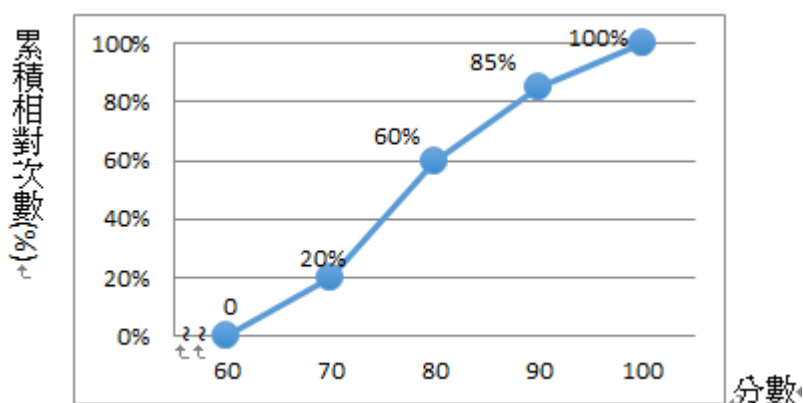


解(2)： $200 \times 60\% = 120$ ，是整數就兩個相加除以 2，所以第 60 百分位數是由小到大的第 120 與 121 個資料的平均。

第 120 資料為 75 分、第 121 個資料也是 75 分， $P_{60} = \frac{75+75}{2} = 75 \text{ (分)}$ 。

如果我們得到的資料是經過整理過的資料而不是個別數值資料，此時可以將其繪製成累積相對次數分配折線圖，那怎麼透過累積相對次數分配折線圖找出各個百分位數呢？

例題二：下圖是班上數學科段考成績累積相對次數折線圖：



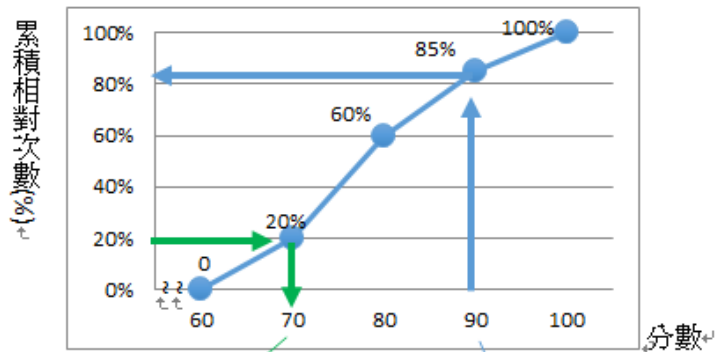
根據累積相對次數分配折線圖，

(1) 數學分數小於 90 分的人數占全部的多少比例？

(2) 全班 20% 的同學數學分數低於幾分？

※解題思維：從累積相對次數圖中就能看到百分比。

解：



從累積相對次數圖找到20%，對應回去的分數是70分，
表示全班20%的同學低於70分，
我們也稱第20百分位數 $P_{20} = 70$ (分)。

90分的累積相對人數為85%，
代表大約有85%的同學少於90分，
也可以說第85百分位數 $P_{85} = 90$ (分)。

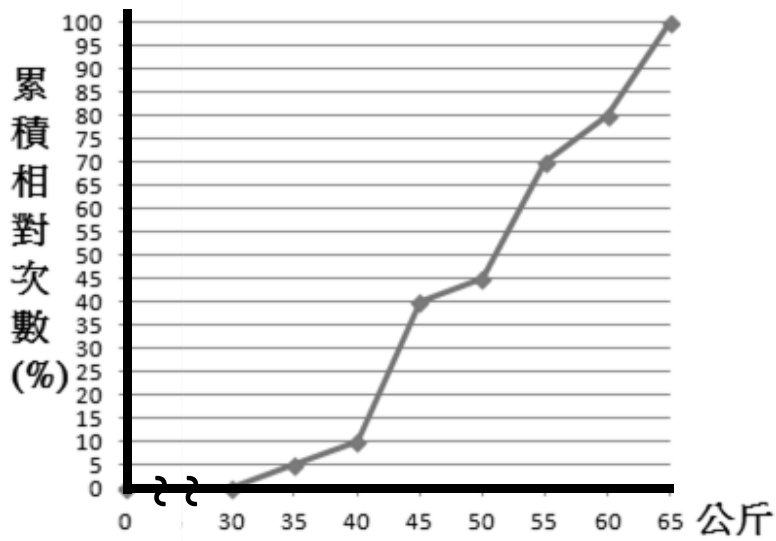
重點提問

1. 根據上面的課文，請用自己的話解釋「百分位數」的意思，它又有什麼意義？

2. 根據上面的課文，如何找到一組資料的百分位數？請舉一個例子作說明。

• 隨堂練習：

1. 下圖為動物園中袋鼠體重的累積相對次數分配折線圖：



(1) 袋鼠體重的 P_{70} 是多少公斤？

(2) 袋鼠體重的 P_{30} 在哪一組？

2. 下表為全校每人投籃 10 次的進球數：

進球數(球)	0	1	2	3	4	5	6	7	8	9	10
次數(人)	10	10	14	13	16	19	25	21	23	20	19

求學生投籃進球數的 P_{25} 與 P_{80} 。