Sec. 3-1 (p.130)

8.

Top-Paid CEOs

The data shown are the total compensation (報酬) (in millions of dollars) for the 50 top-paid CEOs for a recent year. Compare the averages, and state which one you think is the best measure.

17.5	18	36.8	31.7	31.7
17.3	24.3	47.7	38.5	17
23.7	16.5	25.1	17.4	18
37.6	19.7	21.4	28.6	21.6
19.3	20	16.9	25.2	19.8
25	17.2	20.4	20.1	29.1
19.1	25.2	23.2	25.9	24
41.7	24	16.8	26.8	31.4
16.9	17.2	24.1	35.2	19.1
22.9	18.2	25.4	35.4	25.5

- 排序後的資料

```
     [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
[1,] 16.5 16.8 16.9 16.9 17.0 17.2 17.2 17.3 17.4  17.5
[2,] 18.0 18.0 18.2 19.1 19.1 19.3 19.7 19.8 20.0  20.1
[3,] 20.4 21.4 21.6 22.9 23.2 23.7 24.0 24.0 24.1  24.3
[4,] 25.0 25.1 25.2 25.2 25.4 25.5 25.9 26.8 28.6  29.1
[5,] 31.4 31.7 31.7 35.2 35.4 36.8 37.6 38.5 41.7  47.7
```

$$mean = \frac{\sum X}{n} = 17.5 + 18 + \cdots + 25.5 = 24.4$$

$$median = \frac{X_{25} + X_{26}}{2} = \frac{23.2 + 23.7}{2} = 23.45$$

$$mode = 16.9, 17.2, 18, 19.1, 24, 25.2, 31.7$$

$$midrange = \frac{X_1 + X_{50}}{2} = \frac{16.5 + 47.7}{2} = 32.1$$

It appears that mean and median are good measures of average.

10.

Foreign Workers

The number of foreign workers' certificates for the New England states and the northwestern states is shown. Find the mean, median, and mode for both areas and compare the results.

New England states	Northwestern states
6768	1870
3196	622
1112	620
819	23
1019	172
1795	112

- 排序 (小 → 大) ⇒ $n = 6$ → 偶數 ⇒ $median = (X_{\frac{n}{2}} + X_{\frac{n}{2}+1}) / 2$

New England states	819	1019	1112	1795	3196	6768
Northwestern states	23	112	172	620	622	1870

	New England states	Northwestern states
Mean	= (819 + 1019 + … + 6768) / 6 = 2451.5	= (23 + 112 + … + 1870) / 6 = 569.8
Median	= (1112 + 1795) / 2 = 1453.5	= (172 + 620) / 2 = 396
Mode	沒有眾數 (∵每個值都只有一次)	沒有眾數 (∵每個值都只有一次)

The measure of central tendency are much larger for New England states.

16.

Percentage of Foreign-Born People

The percentage of foreign-born population for each of the 50 states is represented here. Find the mean and modal class for the data. Do you think the mean is the best average for this set of data? Explain.

Percentage	Frequency
0.8 - 4.4	26
4.5 - 8.1	11
8.2 - 11.8	4
11.9 - 15.5	5
15.6 - 19.2	2
19.3 - 22.9	1
23.0 - 26.6	1

Percentage	Frequency f	Midpoint X_m	$f \cdot X_m$
0.8 - 4.4	26	2.6	67.6
4.5 - 8.1	11	6.3	69.3
8.2 - 11.8	4	10	40
11.9 - 15.5	5	13.7	68.5
15.6 - 19.2	2	17.4	34.8
19.3 - 22.9	1	21.1	21.1
23.0 - 26.6	1	24.8	24.8
	n = 50		Total = 326.1

$$mean = \frac{326.1}{50} = 6.52$$

$$modal\ class \Rightarrow\ Class\ of\ 0.8 - 4.4$$

The mean is probably not the best measure of central rendency for this data because the data is “bottom heavy”.

20.

Commissions Earned

This frequency distribution represents the commission (傭金) earned (in dollars) by 100 salespeople employed at several branches of a large chain store. Find the mean and modal class for the data.

Class limits	Frequency
150 -158	5
159 - 167	16
168 - 176	20
177 - 185	21
186 - 194	20
195 - 203	15
204 - 212	3

Class limits	Frequency	Midpoint X_m	$f \cdot X_m$
150 -158	5	154	770
159 - 167	16	163	2608
168 - 176	20	172	3440
177 - 185	21	181	3801
186 - 194	20	190	3800
195 - 203	15	199	2985
204 - 212	3	208	624
	n = 100		Total = 18028

$$mean = \frac{18028}{100} = 180.28$$

$$modal\ class \Rightarrow\ Class\ of\ 177 - 185$$

28.

Final Grade

Another instructor gives four 1-hour exams and one final exam, which counts as two 1-hour exams. Find a student's grade if she received 62, 83, 97, and 90 on the 1-hour exams and 82 on the final exam.

$$\bar{X} = \frac{\sum wX}{\sum w} = \frac{1 \cdot 62 + 1 \cdot 83 + 1 \cdot 97 + 1 \cdot 90 + 2 \cdot 82}{1 + 1 + 1 + 1 + 2} = 82.7$$

30.

Describe which measure of central tendency — mean, median, or mode — was probably used in each situation.

a. One-half of the factory workers make more than $5.37 per hour, and one-half make less than $5.37 per hour.
(一半的工廠工人每小時賺超過 5.37 美元，而另一半每小時賺不到 5.37 美元)

b. The average number of children per family in the Plaza Heights Complex is 1.8.
(Plaza Heights 複合區每户家庭的平均子女數是 1.8)

c. Most people prefer red convertibles over any other color.

(大多數人更喜歡紅色敞篷車而不是其他任何顏色)

d. The average person cuts the lawn once a week.
(一般人每周修剪草坪一次)

e. The most common fear today is fear of speaking in public.
(如今最普遍的恐懼是對公開演講的恐懼)

f. The average age of college professors is 42.3 years.
(大學教授的平均年齡是 42.3 歲)

a. Median
b. Mean
c. Mode
d. Mode (一般人屬於大眾)
e. Mode
f. Mean

Sec. 3-2 (p.151)

8. **Cigarette Taxes**

The increases (in cents) in cigarette taxes for 17 states in a 6-month period are

60, 20, 40, 40, 45, 12, 34, 51, 30, 70, 42, 31, 69, 32, 8, 18, 50

Find the range, variance, and standard deviation for the data. Use the range rule of thumb to estimate the standard deviation. Compare the estimate to the actual standard deviation.

X	X^2
60	3600
20	400
40	1600
40	1600
45	2025
12	144
34	1156
51	2601
30	900
70	4900
42	1764
31	961
69	4761
32	1024
8	64
18	324
50	2500
$\sum X = 652$ $n = 17$	$\sum X^2 = 30324$

1. **Range:**

$$R = max - min$$
$$= 70 - 8 = 62$$

2. **Variance:**

$$s^2 = \frac{n(\sum X^2) - (\sum X)^2}{n(n-1)}$$
$$= \frac{17 \cdot 30324 - 652^2}{17 \cdot 16}$$
$$= \frac{90404}{272} \approx 332.4$$

3. **Standard deviation:**

$$s = \sqrt{s^2} = \sqrt{332.4} \approx 18.2$$

4. **Range rule of thumb of s:**

$$s = \frac{R}{4} = \frac{62}{4} = 15.5$$

5. **Conclusion:**

This is close to the actual standard deviation of 18.2.

18.

Use the data from Exercises 7, 15, and 17 (spending, laws, precipitation days) and compare the standard deviation with that obtained by the range rule of thumb(R/4). Comment on the results.

7. Traveler Spending

20.1	33.5	21.7	58.4	23.2	110.8	30.9	24.0	74.8	60.0

15. Laws Passed

283	394	383	580	498	460	377	482

17. Annual Precipitation Days (年降雨天数)

135	128	136	78	116	77	111	79	44	97
116	123	88	102	26	82	156	133	107	35
112	98	45	122	125					

Data	**Traveler Spending**	**Laws Passed**	**Annual Precipitation Days**
Standard deviation $s = \sqrt{\frac{n(\sum X^2) - (\sum X)^2}{n(n-1)}}$	$n = 10, \sum X = 457.4$ $\sum X^2 = 28948.44$ $s = 29.864 \approx 29.9$	$n = 8, \sum X = 3457$ $\sum X^2 = 1552471$ $s = 91.507 \approx 91.5$	$n = 25, \sum X = 2471$ $\sum X^2 = 271995$ $s = 34.011 \approx 34.0$
Range rule of thumb $s = \frac{R}{4}$	$s = \frac{110.8 - 20.1}{4}$ $= \frac{90.7}{4} = 22.675$ ≈ 22.7	$s = \frac{580 - 283}{4}$ $= \frac{297}{4} = 74.25$ ≈ 74.3	$s = \frac{156 - 26}{4}$ $= \frac{130}{4} = 32.5$
difference	7.2	17.2	1.5

The closest estimate is for precipitation. The estimate for spending is also close.

20.

Automotive Fuel Efficiency

Thirty automobiles (汽車) were tested for fuel efficiency (燃油效率測試) (in miles per gallon). This frequency distribution was obtained. Find the variance and standard deviation for the data.

Class boundaries	**Frequency**
7.5-12.5	3
12.5-17.5	5
17.5-22.5	15
22.5-27.5	5
27.5-32.5	2

Class boundaries	Frequency	X_m	$f \cdot X_m$	$f \cdot X_m^2$
7.5-12.5	3	10	30	300
12.5-17.5	5	15	75	1125
17.5-22.5	15	20	300	6000
22.5-27.5	5	25	125	3125
27.5-32.5	2	30	60	1800
Total	30		590	12350

1. Sample variance

$$s^2 = \frac{n(\sum f \cdot X_m^2) - (\sum f \cdot X_m)^2}{n(n-1)}$$

$$= \frac{30 \cdot 12350 - 590^2}{30 \cdot 29} = 25.747 \approx 25.75$$

2. Sample standard deviation

$$s = \sqrt{s^2} = \sqrt{25.75} = 5.074 \approx 5.07$$

28. **Suspension Bridges**

The lengths (in feet) of the main span (跨距) of the longest suspension bridges (吊橋) in the United States and the rest of the world are shown below. Which set of data is more variable?

United States	4205	4200	3800	3500	3478	2800	2800	2310
World	6570	5538	5328	4888	4626	4544	4518	3970

比較 → 使用變異係數: $CVar = \frac{s}{\bar{X}} \cdot 100$

US:

$$\bar{X}_{US} = \frac{\sum x}{n} = \frac{4205 + \cdots + 2310}{8} = 3386.63 \approx 3386.6$$

$$s = \sqrt{\frac{n\sum x^2 - (\sum x)^2}{n(n-1)}} = \sqrt{\frac{8 \cdot 95124609 - 27093^2}{8 \cdot 7}} = 693.93 \approx 693.9$$

$$CVar = \frac{s}{\bar{X}} \cdot 100 = \frac{693.9}{3386.6} \cdot 100 = 20.49\%$$

World:

$$\bar{X}_W = \frac{\sum x}{n} = \frac{6570 + \cdots + 3970}{8} = 4997.75 \approx 4997.8$$

$$s = \sqrt{\frac{n\sum x^2 - (\sum x)^2}{n(n-1)}} = \sqrt{\frac{8 \cdot 204335508 - 39982^2}{8 \cdot 7}} = 803.16 \approx 803.2$$

$$CVar = \frac{s}{\bar{X}} \cdot 100 = \frac{803.2}{4997.8} \cdot 100 = 16.07\%$$

Conclusion: The data of US is more variable.

30. **Ages of Accountants**

The average age of the accountants (會計師) at Three Rivers Corp. (corporation, 公司) is 26 years, with a standard deviation of 6 years; the average salary of the accountants is $31,000, with a standard deviation of $4000. Compare the variations of age and income.

	Age	Income
$CVar = \frac{s}{\bar{X}} \cdot 100$	$CVar_{age} = \frac{6}{26} \cdot 100 = 23.1\%$	$CVar_{income} = \frac{4{,}000}{31{,}000} \cdot 100 = 12.9\%$

Age is more variable.

33. In a distribution of 160 values with a mean of 72, at least 120 fall within the interval 67-77. Approximately what percentage of values should fall in the interval 62-82? Use Chebyshev's theorem.

Chebyshev's Thm.: 落在 k 倍標準差內的比例至少有 $1-\frac{1}{k^2}$

1. $\frac{120}{160}\cdot 100=75\%$ 的資料落在 $67-77$ 區間內$\Rightarrow 1-\frac{1}{k^2}=0.75\Rightarrow k=2$

$(72-2\cdot s, 72+2\cdot s)=(67,77)\Rightarrow s=2.5$

2. $(62,82)=(72-k\cdot 2.5, 72+k\cdot 2.5)\Rightarrow k=4$

$$1-\frac{1}{4^2}=0.9375$$

At least 93.75 percentage of values should fall in the interval 62-82.

34. **Calories in Bagels**

The average number of calories in a regular-size bagel is 240. If the standard deviation is 38 calories, find the range in which at least 75% of the data will lie. Use Chebyshev's theorem.

$1-\frac{1}{k^2}=0.75\Rightarrow k=2$

Range: $(240-2\cdot 38, 240+2\cdot 38)=(164,316)$

At least 75% of the data will fall in the interval 164-316.

38. **Trials to Learn a Maze**

The average of the number of trials it took a sample of mice to learn to traverse (通過) a maze (迷宮) was 12. The standard deviation was 3. Using Chebyshev's theorem, find the minimum percentage of data values that will fall in the range of 4-20 trials.

$(4,20)=12\pm k\cdot 3\Rightarrow k=\frac{8}{3}$

$1-\frac{1}{k^2}=1-\frac{1}{\left(\frac{8}{3}\right)^2}=\frac{55}{64}=0.859$

At least 85.9% of the data will fall in the interval 4-20.

42. **Work Hours for College Faculty**

The average fulltime faculty member (全職教員) in a postsecondary degree-granting institution (授予高等教育學位的機構) works an average of 53 hours per week.

a. If we assume the standard deviation is 2.8 hours, what percentage of faculty members work more than 58.6 hours a week?

b. If we assume a bell-shaped distribution, what percentage of faculty members work more than 58.6 hours a week?

a. $58.6=53+k\cdot 2.8\Rightarrow k=2$

$1-\frac{1}{2^2}=0.75\Rightarrow$ By Chebyshev's Thm. 有 75%的工作時間落在 2 倍標準差內.

因為不知道資料分布是否為對稱的，我們可以說每周工作超過 58.6 小時的全職教員不超過

25%。

b. 假設資料為對稱分布，根據經驗法則(Empirical rule)，有 95%的資料落在 2 倍標準差內

則有 $\frac{5\%}{2} = 2.5\%$ 的員工工作超過 58.6 小時。

46. **Bonuses (獎金)**

The mean and standard deviation of the bonuses that the employees of a company received 10 years ago were, respectively, $2,000 and $325. Today the amount of the bonuses is 5 times what it was 10 years ago. Find the mean and standard deviation of the new bonuses.

$$X_{today} = 5 \cdot X_{ago}$$
$$\bar{X}_{today} = 5\bar{X}_{ago} = 5 \cdot \$2{,}000 = \$10{,}000$$
$$s_{today} = 5s_{ago} = 5 \cdot \$325 = \$1{,}625$$

Sec. 3-3 (p.167)

10. **Age of Senators**

The average age of Senators (參議員) in the 114th congress (國會) was 61.7 years. If the standard deviation was 10.6, find the z scores of a senator who is 48 years old and one who is 66 years old.

48 years old	$z = \frac{X - \mu}{\sigma}$	$z = \frac{48 - 61.7}{10.6} = -1.292 \approx -1.29$
66 years old		$z = \frac{66 - 61.7}{10.6} = 0.405 \approx 0.41$

12. **Teacher's Salary**

The average teacher's salary in a particular state is $54,166. If the standard deviation is $10,200, find the salaries corresponding to the following z scores.

a. 2 b. -1 c. 0 d. 2.5 e. -1.6

$$z = \frac{X - \bar{X}}{s} \Rightarrow X = s \cdot z + \bar{X}$$

a.	b.	c.	d.	e.
X	X	X	X	X
$= 10{,}200 \cdot 2$	$= 10{,}200 \cdot (-1)$	$= 10{,}200 \cdot 0$	$= 10{,}200 \cdot 2.5$	$= 10{,}200 \cdot (-1.6)$
$+ 54{,}166$	$+ 54{,}166$	$+ 54{,}166$	$+ 54{,}166$	$+ 54{,}166$
$= 74566$	$= 43966$	$= 54166$	$= 79666$	$= 37846$

15. **Annual Miles Driven**

The average miles driven (行駛里程) annually per licensed driver in the United States is approximately 14,090 miles. If we assume a fairly mound-shaped distribution with a standard deviation of approximately 3500 miles, find the following:

a. z score for 16,000 miles

b. z score for 10,000 miles

c. Number of miles corresponding to z scores of 1.6, -0.5, and 0.

$z = \frac{X - \bar{X}}{s}$		$X = s \cdot z + \bar{X}$		
a.	b.	c.		
$z = \frac{16000 - 14090}{3500} = -0.545 \approx -0.55$	$z = \frac{10000 - 14090}{3500} = -1.168 \approx -1.17$	$X = 3500 \cdot 1.6 + 14090 = 19690$	$X = 3500 \cdot (-0.5) + 14090 = 12340$	$X = 3500 \cdot 0 + 14090 = 14090$

17.

The data show the population (in thousands) for a recent year of a sample of cities in South Carolina (南卡羅萊納州).

29	26	15	13	17	58	14	25	37
19	40	67	23	10	97	12	129	27
20	18	120	35	66	21	11	43	22

Find the data value that corresponds to each percentile.

a. 40th percentile b. 75th percentile c. 90th percentile d. 30th percentile

Using the same data, find the percentile corresponding to the given data value.

e. 27 f. 40 g. 58 h. 67

Given percentile: $c = \frac{n \cdot p}{100}$, $n = 27$

a.	b.	c.	d.
$c = \frac{27 \cdot 40}{100} = 10.8$ $\Rightarrow c = 11$ $11^{th}: 21$	$c = \frac{27 \cdot 75}{100} = 20.25$ $\Rightarrow c = 21$ $21^{st}: 43$	$c = \frac{27 \cdot 90}{100} = 24.3$ $\Rightarrow c = 25$ $25^{th}: 97$	$c = \frac{27 \cdot 30}{100} = 8.1$ $\Rightarrow c = 9$ $9^{th}: 19$

Given data value: $percentile = \frac{\text{比 X 小的個數}+0.5}{\text{總個數}} \cdot 100$

e.	f.	g.	h.
$p = \frac{15+0.5}{27} \cdot 100$	$p = \frac{19+0.5}{27} \cdot 100$	$p = \frac{21+0.5}{27} \cdot 100$	$p = \frac{23+0.5}{27} \cdot 100$
$= 57.4 \approx 57$	$= 72.2 \approx 72$	$= 79.6 \approx 80$	$= 87.0 \approx 87$
27 高於 57%的資料	40 高於 72%的資料	58 高於 80%的資料	67 高於 87%的資料

20.

Airplane Speeds

The airborne (飛行中的) speeds in miles per hour of 21 planes are shown.

Class	Frequency
366-386	4
387-407	2
408-428	3
429-449	2
450-470	1
471-491	2
492-512	3
513-533	4
Total	21

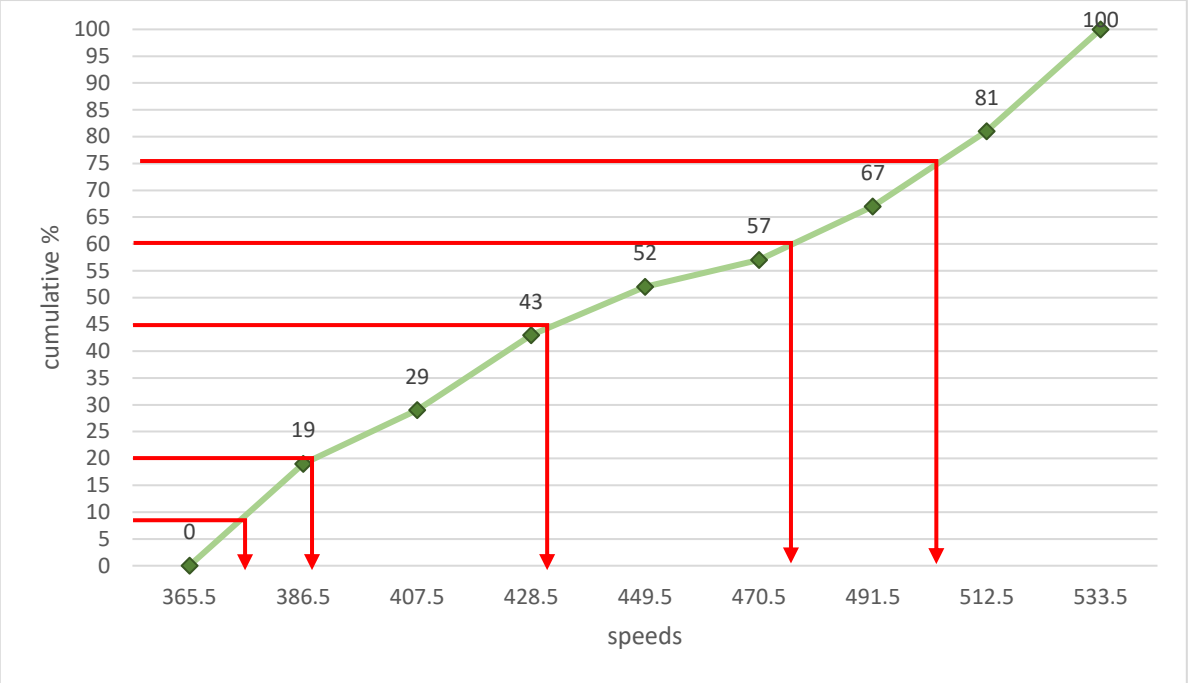
Find the approximate values that correspond to the given percentiles by constructing a percentile graph.

a. 9th b. 20th c. 45th d. 60th e. 75th

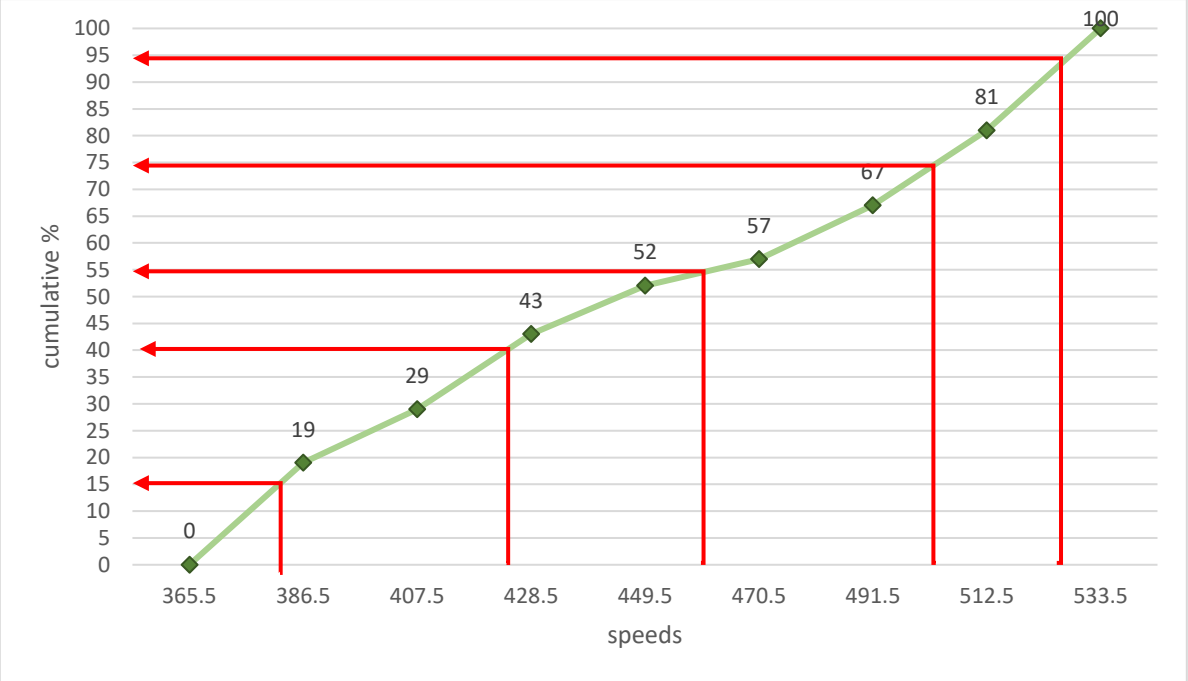
Using the same data, find the approximate percentile ranks of the following speeds in miles per hour (mph).

f. 380 mph g. 425 mph h. 455 mph i. 505 mph j. 525 mph

Class	Boundaries		Frequency	Cumulative freq.	Cumulative percent
366-386	365.5	386.5	4	4	19
387-407	386.5	407.5	2	6	29
408-428	407.5	428.5	3	9	43
429-449	428.5	449.5	2	11	52
450-470	449.5	470.5	1	12	57
471-491	470.5	491.5	2	14	67
492-512	491.5	512.5	3	17	81
513-533	512.5	533.5	4	21	100
Total			21		



	a. 9th	b. 20th	c. 45th	d. 60th	e. 75th
Approximate value	376	387	429	477	502
解答的答案	375	389	433	477	504



	f. 380 mph	g. 425 mph	h. 455 mph	i. 505 mph	j. 525 mph
Approximate percentile	15	40	55	75	95
解答的答案	13^{th}	40^{th}	54^{th}	76^{th}	92^{nd}

25. **Taxes**

The data for a recent year show the taxes (in millions of dollars) received from a random sample of 10 states. Find the first and third quartiles and the IQR.

13	15	32	36	11	24	6	25	11	71

排序	6	11	11	13	15	24	25	32	36	71

Method 1	Method 2
$n = 10 \rightarrow median = \frac{X_5 + X_6}{2} = \frac{15 + 24}{2} = 19.5$ $Q_1 = 11, Q_3 = 32$	$Q_1 = P_{25} \rightarrow c = \frac{n \cdot p}{100} = \frac{10 \cdot 25}{100} = 2.5 \rightarrow c = 3$ $Q_3 = P_{75} \rightarrow c = \frac{10 \cdot 75}{100} = 7.5 \rightarrow c = 8$
$IQR = Q_3 - Q_1 = 32 - 11 = 21$	

28. **Police Calls in Schools**

The number of incidents (事件) in which police were needed for a sample of 9 schools in Allegheny County (阿利根尼縣，美國賓夕法尼亞州西部) is 7, 37, 3, 8, 48, 11, 6, 0, 10. Find the first and third quartiles for the data.

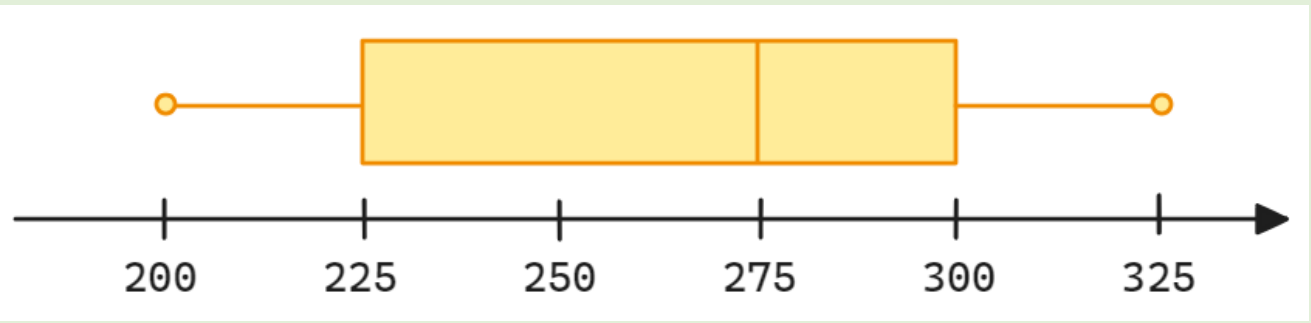
排序	0	3	6	7	8	10	11	37	48

Method 1	Method 2
$n = 9 \rightarrow median = X_5 = 8$ $Q_1 = \frac{3 + 6}{2} = 4.5, Q_3 = \frac{11 + 37}{2} = 24$	$Q_1 = P_{25} \rightarrow c = \frac{n \cdot p}{100} = \frac{9 \cdot 25}{100} = 2.25 \rightarrow c = 3$ $Q_3 = P_{75} \rightarrow c = \frac{9 \cdot 75}{100} = 6.75 \rightarrow c = 7$ $c = 2.25$ 取 3，小樣本誤差會大 內插法: $Q_1 = \frac{1 \cdot 6 + 3 \cdot 3}{4} = 3.75$ $Q_3 = \frac{3 \cdot 11 + 1 \cdot 10}{4} = 10.75$
$IQR = Q_3 - Q_1 = 32 - 11 = 21$	$IQR = 7$

法 1 比較簡單、(✿˚▽˚)ノ

Sec. 3-4 (p.180)

8. Use each boxplot to identify the maximum value, minimum value, median, first quartile, third quartile and interquartile range.



Max	Min	Median = Q2	Q1	Q3	IQR = Q3 – Q1
325	200	275	225	300	300 – 225 = 75

12. **Innings Pitched**

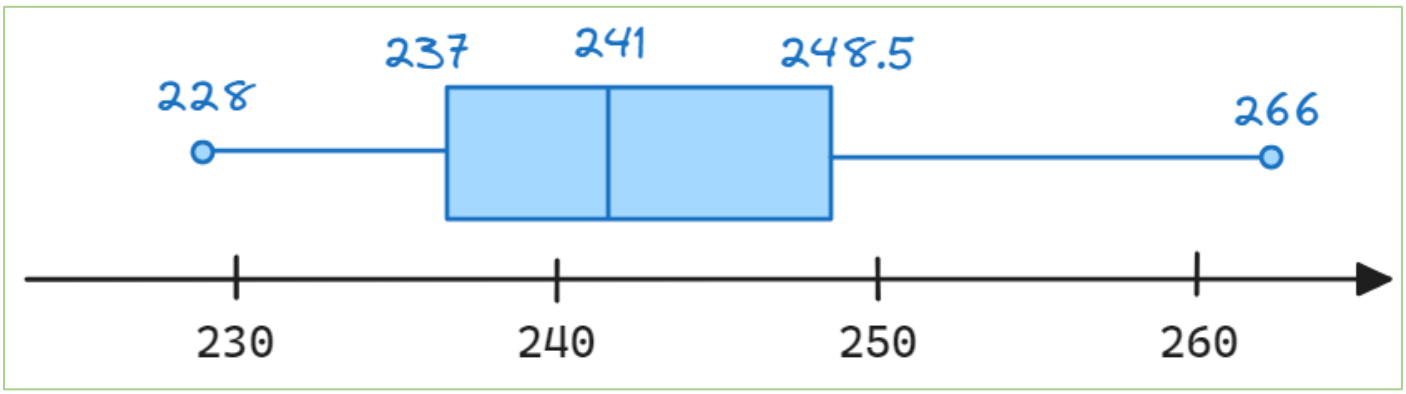
Construct a boxplot for the following data which represent the number of innings pitched (IP, 投球局數) by the ERA (earned run average, 投手防禦率) leaders for the past few years. Comment on the shape of the distribution.

239 266 245 236 241 246 240
249 251 238 228 248 232

排序	228	232	236	238	239	240	241	245	246	248	249	251	266

$$n = 13 \rightarrow median = X_7 = 241$$

$$Q_1 = \frac{236 + 238}{2} = 237, Q_3 = \frac{248 + 249}{2} = 248.5$$



The distribution is slightly right-skewed. (∵右邊的線段較長且中位數偏箱子左邊)

($mean = \bar{X} = 243 > median = 241$)

16.

Size of Dams

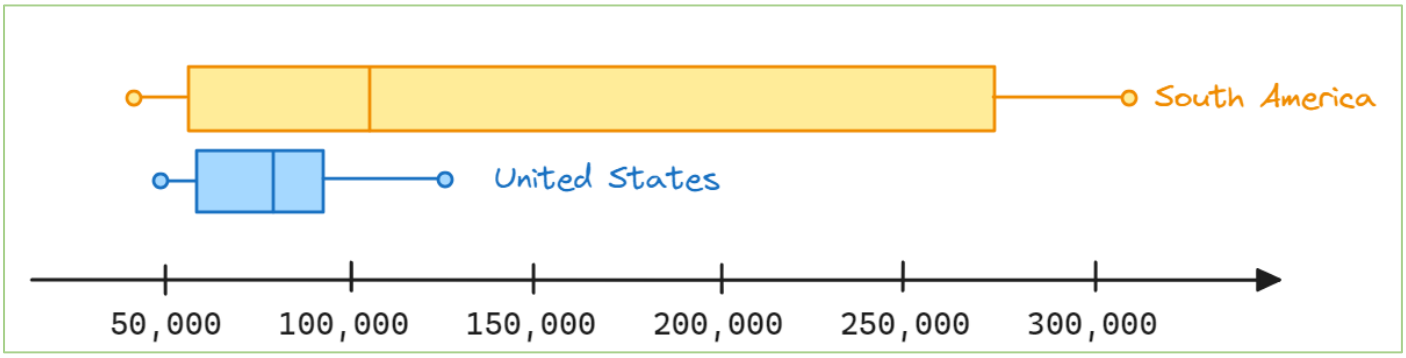
These data represent the volumes (體積) in cubic yards (立方碼) of the largest dams (水壩) in the United States and in South America. Construct a boxplot of the data for each region and compare the distributions.

United States	South America
125,628	311,539
92,000	274,026
78,008	105,944
77,700	102,014
66,500	56,242
62,850	46,563
52,435	
50,000	

排序

United States	50,000	52,435	62,850	66,500	77,700	78,008	92,000	125,628
South America	46,563	56,242	102,014	105,944	274,026	311,539		

	min	max	Q1	Q2	Q3	mean
United States	50,000	125,628	57,642.5	72,100	85,004	75,640
South America	46,563	311,539	56,242	103,979	274,026	149,388



The range and variation of the capacity of the dams in South America is considerably larger than those of the United States. Both United States and South America are right-skewed.

18. **Number of Tornadoes** (龍捲風)

(hurricane, 颶風，和龍捲風相比影響範圍較大) (typhoon, 颱風，和颶風一樣同屬熱帶氣旋)

A four-month record for the number of tornadoes in 2013-2015 is given here.

	2013	2014	2015
April	80	130	170
May	227	133	382
June	126	280	184
July	69	90	116

a. Which month had the highest mean number of tornadoes for this 3-year period?

b. Which year has the highest mean number of tornadoes for this 4-month period?

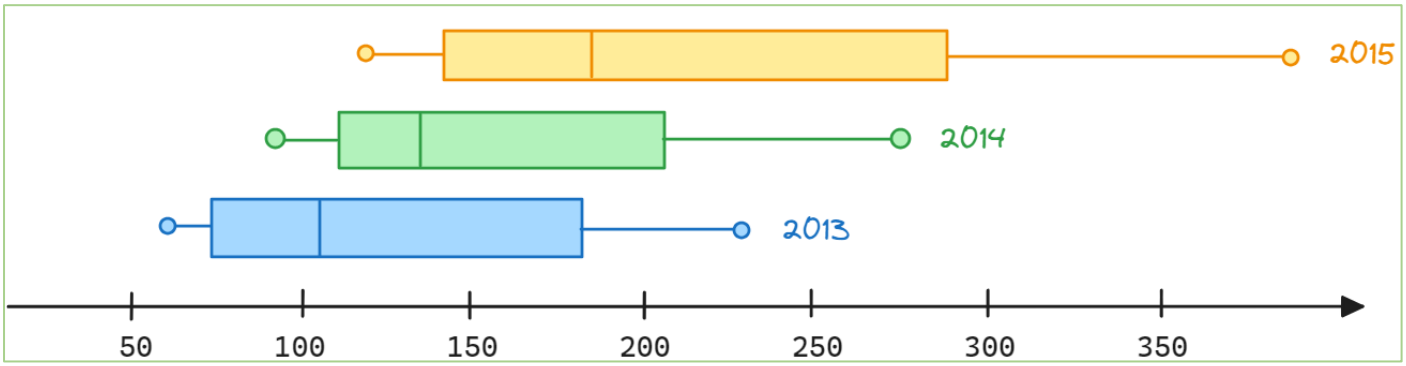
c. Construct three boxplots and compare the distributions.

	2013	2014	2015	mean
April	80	130	170	126.67
May	227	133	382	247.33
June	126	280	184	196. 67
July	69	90	116	91. 67
mean	125.5	158.25	213	

a. The month with the highest mean number of tornadoes is May.

b. The year with the highest mean number of tornadoes is 2015.

	min	Q1	Q2	Q3	max	mean
2013	69	(69+80)/2=74.5	(126+80)/2=103	(126+227)/2=176.5	227	125.5
2014	90	(90+130)/2=110	(130+133)/2=131.5	(133+280)/2=206.5	280	158.25
2015	116	(116+170)/2=143	(170+184)/2=177	(184+382)/2=283	382	213



The distribution for 2013, 2014, and 2015 are positively skewed. The data for 2013 appears to be the least variable