⑮ 四分位数と箱ひげ図

▶練習問題 ⇨ P123

学習の基本 81 四分位数

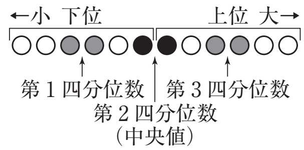
① データを小さい順に並べたときに，4 等分する位置にくる 3 つの値のことを**四分位数**といい，小さい方から順に，**第 1 四分位数，第 2 四分位数，第 3 四分位数**という。第 2 四分位数は中央値のことであり，データを 2 等分したときの下位データの中央値が第 1 四分位数，上位データの中央値が第 3 四分位数である。

② 第 3 四分位数から第 1 四分位数をひいた差を**四分位範囲**といい，四分位範囲を 2 でわった値を**四分位偏差**という。

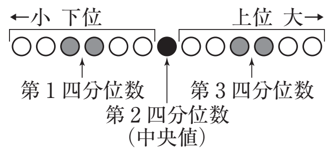
(四分位範囲)＝(第 3 四分位数)－(第 1 四分位数)

(四分位偏差)＝(四分位範囲)÷2

〈データの個数が偶数のとき〉



〈データの個数が奇数のとき〉



問題 **右のデータは生徒10人の垂直跳びの記録である。このデータの四分位数，四分位範囲，四分位偏差をそれぞれ求めよ。**

36	28	30	45	29
39	42	30	40	26

(単位：cm)

解 データを小さい順に並べると次のようになる。

←小　下位　　　　上位　大→

26　28　29　30　30　36　39　40　42　45

第 1 四分位数は，下位データの中央値で，29 cm

第 2 四分位数は，データ全体の中央値で，(30＋36)÷2＝33(cm)

第 3 四分位数は，上位データの中央値で，40 cm

四分位範囲は，40－29＝11(cm)，四分位偏差は，11÷2＝5.5(cm)

答 第 1 四分位数…29 cm，第 2 四分位数…33 cm，第 3 四分位数…40 cm，四分位範囲…11 cm，四分位偏差…5.5 cm

応 323 次のデータの四分位数，四分位範囲，四分位偏差をそれぞれ求めよ。

□(1) 3　16　9　10　5　11　9　17　11　15　(単位：分)

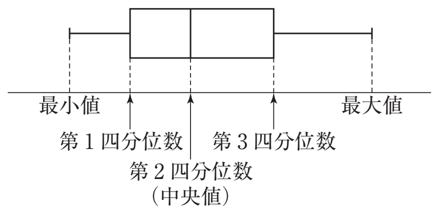
□(2) 30　47　29　58　40　39　47　39　42　56　35　(単位：kg)

□(3) 15　14　17　11　22　14　12　20　17　19　23　22　(単位：m)

□(4) 81　74　64　65　73　68　85　64　75　65　86　79　67　(単位：kg)

学習の基本 82 箱ひげ図

最小値，最大値，四分位数を用いてデータの分布のようすを示した右のような図を**箱ひげ図**という。



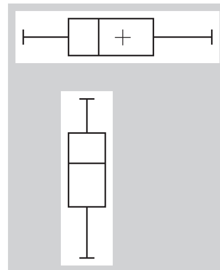
<箱ひげ図のかき方>

① 第１四分位数を左端，第３四分位数を右端とする長方形(箱)をかく。

② ①の長方形(箱)の中に第２四分位数(中央値)を示す縦線をひく。

③ 最小値，最大値を示す縦線をひき，箱と線分(ひげ)でつなぐ。

右の図のように，箱ひげ図は縦にかくことがあり，また，箱ひげ図に平均値を記入する場合は，「＋」で示す。



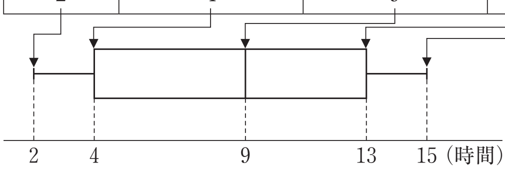
問題 下のデータは，生徒12人の先月１か月間の読書時間を短い順に並べたものである。これをもとに箱ひげ図をかけ。

2　2　3　5　6　8　10　11　12　14　14　15　（単位：時間）

解

最小値	第１四分位数	第２四分位数	第３四分位数	最大値
2	4	9	13	15

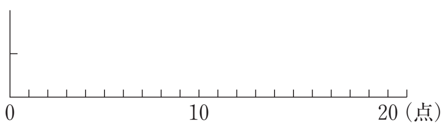
（単位：時間）



ひげの端から端までの長さが範囲(全体の散らばり)を表し，箱の長さが四分位範囲(中央値付近の散らばり)を表す。
箱の区間には，中央値付近の約50％(前後約25％)のデータの値が含まれる。

答 上の図

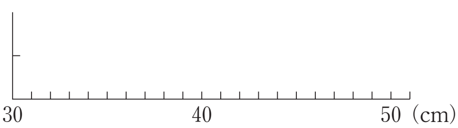
324 あるデータの最小値は５点，第１四分位数は７点，第２四分位数は10点，第３四分位数は15点，最大値は19点であった。これをもとに箱ひげ図をかけ。



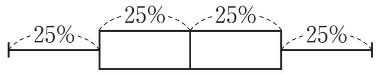
応 **325** 次のデータについて，箱ひげ図をかき，平均値の位置を「＋」で示せ。

31　36　38　39　39　43　46　48　49

（単位：cm）



学習の基本 83 箱ひげ図を読みとる



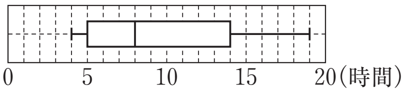
箱ひげ図の各部分にふくまれるデータの個数のおよその割合は右の図のようになっている。

ただし，箱ひげ図の読みとりでは，次のことに注意する。

① 上記のデータの個数の割合は目安であり，一般に，正確な割合は読みとれるとは限らない。

② ある値のデータが実際に存在するかどうかは読みとれない場合が多い。

問題 **右の箱ひげ図は，生徒40人の1か月間の読書時間を表したものである。次のア～エのうち，この図から読みとれる内容として正しいものをすべて選べ。**



ア　第1四分位数は1時間である。

イ　四分位範囲は9時間である。

ウ　平均値は8時間である。

エ　半数以上の人が7時間以上読書をしている。

解 **ア**　第1四分位数は5時間だから，正しくない。

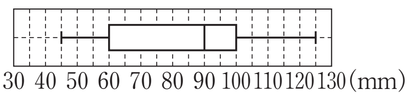
イ　四分位範囲は14－5＝9(時間) だから，正しい。

ウ　平均値は箱ひげ図から読みとれないから，このデータからはわからない。

エ　中央値が8時間だから，正しい。

答 **イ，エ**

326 □ 右の箱ひげ図は，ある都市の1年間の月ごとの降水量のデータを表したものである。次の**ア**～**エ**のうち，この図から読みとれる内容として正しいものをすべて選べ。



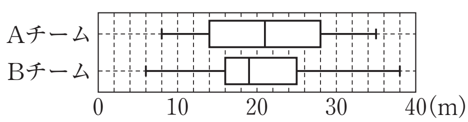
ア　中央値は90mmである。

イ　範囲は40mmである。

ウ　平均値は90mmである。

エ　降水量が60mm以上の月が6か月以上ある。

応 **327** □ 右の箱ひげ図は，Aチーム15人とBチーム15人のハンドボール投げの記録を表したものである。次の**ア**～**エ**のうち，この図から読みとれる内容として正しいものを1つ選べ。



ア　平均値はAよりBの方が大きい。

イ　範囲も四分位範囲も，AよりBの方が大きい。

ウ　記録が20m以上の人は，AよりBの方が多い。

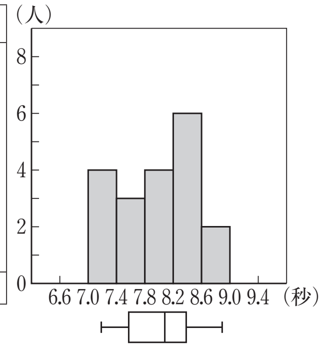
エ　記録が15m以下の人は，BよりAの方が多い。

学習の基本 84　ヒストグラムと箱ひげ図の関係

次の表，図は，1組男子と2組男子の50m走の記録の度数分布表，ヒストグラム，箱ひげ図である。

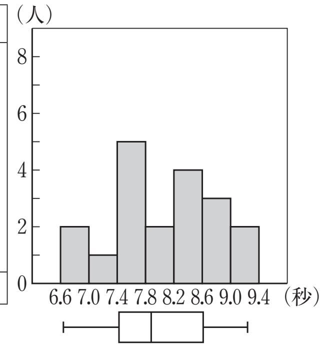
1組男子

階級(秒)	度数(人)
以上　未満	
6.6〜7.0	0
7.0〜7.4	4
7.4〜7.8	3
7.8〜8.2	4
8.2〜8.6	6
8.6〜9.0	2
9.0〜9.4	0
計	19



2組男子

階級(秒)	度数(人)
以上　未満	
6.6〜7.0	2
7.0〜7.4	1
7.4〜7.8	5
7.8〜8.2	2
8.2〜8.6	4
8.6〜9.0	3
9.0〜9.4	2
計	19

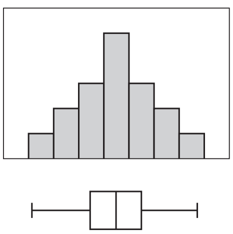


この図のように，ヒストグラムはデータの分布のようすや最頻値はわかりやすいが，中央値はわかりにくい。それに比べて，箱ひげ図は，データの散らばりぐあいや中央値がわかりやすいという特徴がある。

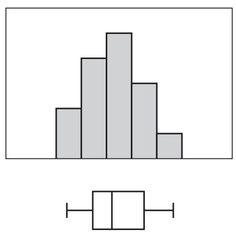
1組は，データが比較的中央に集まっている傾向がある。よって，箱の長さは短くなり，中央値も真ん中付近にある。一方，2組はデータが広く散らばっているので，箱の長さは長くなる。

ヒストグラムと箱ひげ図の関係

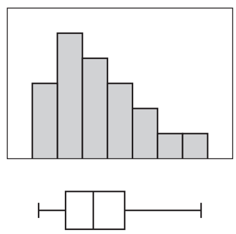
左右対称な分布



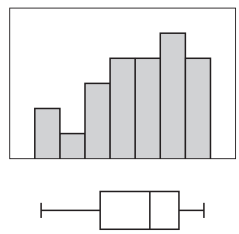
散らばりの小さい分布



左にかたよった分布

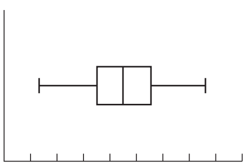


右にかたよった分布

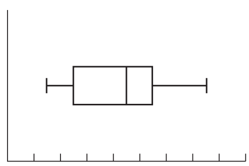


応 328 次の(1)，(2)，(3)の箱ひげ図に対応するヒストグラムを，ア，イ，ウの記号で答えよ。

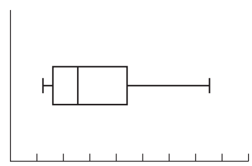
(1)



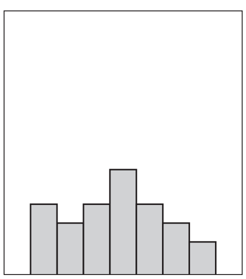
(2)



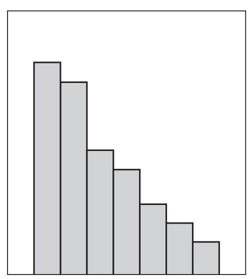
(3)



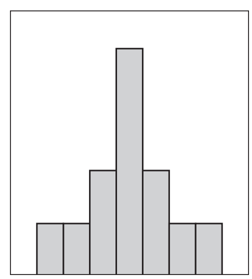
ア



イ



ウ



練習問題　15 四分位数と箱ひげ図

329 次のデータの最小値，最大値，四分位数を求めよ。

□(1) 164　170　184　157　172　180　165　172　177　164　170　（単位：cm）

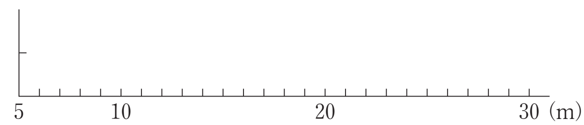
□(2) 3　4　2　7　8　10　9　5　4　5　9　5　7　3　10
（単位：冊）

応 **330** 次のデータの最小値，最大値，四分位数，四分位範囲，範囲を求めよ。

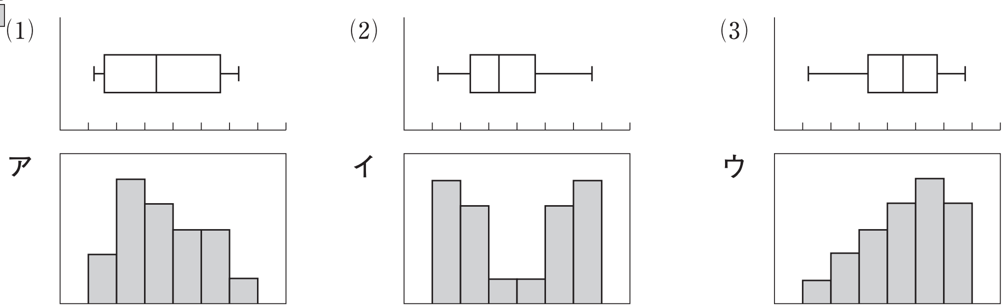
□(1) 1　0　5　2　2　4　0　1　2　1　4　3　（単位：点）

□(2) 13　32　8　22　29　13　48　29　30　25　17　21　7
（単位：分）

331 □ あるデータの最小値は7m，第１四分位数は17m，第２四分位数は25m，第３四分位数は27m，最大値は30m，平均値は23mであった。これをもとに箱ひげ図をかき，平均値の位置を「＋」で示せ。

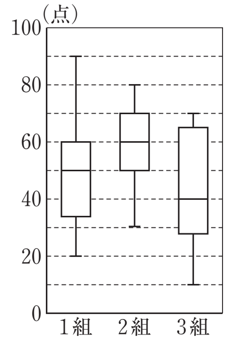


応 **332** 次の(1)，(2)，(3)の箱ひげ図に対応するヒストグラムを，**ア**，**イ**，**ウ**の記号で答えよ。

□


応 **333** □ 右の図は，ある学校の２年生１組，２組，３組それぞれ27人が受けた100点満点のテストの得点の箱ひげ図である。このとき，次の**ア**～**エ**のうち，右の図から読みとれる内容として正しいものをすべて選べ。

ア　得点の範囲がもっとも大きいのは１組である。
イ　得点が40点以下の生徒はどの組にも７人以上いる。
ウ　得点が50点以上の生徒は，２組がもっとも多い。
エ　得点が60点の生徒は，どのクラスでも上位10番以内に入る。



5章の確認問題

334 代表値　右の表は，あるクラスの生徒の体重を測ってまとめたものである。次の問いに答えよ。

階級(kg)	階級値(kg)	度数(人)	(階級値)×(度数)
以上　未満 35～40	37.5	3	112.5
40～45	42.5	11	467.5
45～50	47.5	14	665.0
50～55	52.5	8	420.0
55～60	57.5	4	230.0
計		40	1895.0

□(1)　平均値を，四捨五入によって小数第1位まで求めよ。

□(2)　中央値はどの階級に属するか。

□(3)　この度数分布表での最頻値を求めよ。

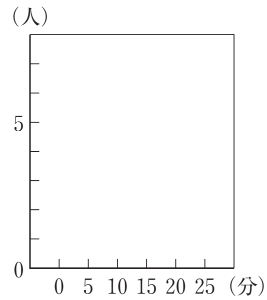
335 データの整理　右のデータは，1組15人の通学時間を調べたものである。次の問いに答えよ。

24	7	11	15	24	5	14	5
11	18	6	12	20	10	9	(単位：分)

□(1)　右の度数分布表に整理せよ。

階級(分)	度数(人)
以上　未満 0～ 5	
5～10	
10～15	
15～20	
20～25	
計	15

□(2)　右の図にヒストグラムをかき入れよ。



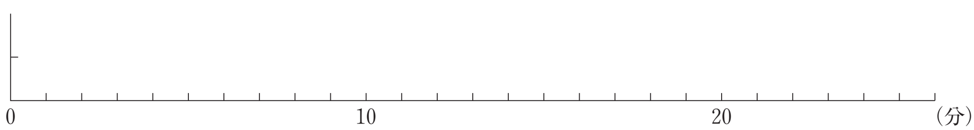
□(3)　最小値，最大値，四分位数，四分位範囲，範囲を求め，下の表にまとめよ。

最小値	第1四分位数	第2四分位数	第3四分位数	最大値

四分位範囲	範囲

(単位：分)

□(4)　箱ひげ図を下の図にかき入れよ。



5章の応用問題

336 右の表は，ある中学校の１年生全員に行った漢字テストの結果を男女別にまとめたものであり，女子の得点の平均値は5.56点，男女合わせた全員の得点の平均値は5.48点であった。

得点(点)	男子(人)	女子(人)
2	2	0
3	16	7
4	12	12
5	25	y
6	18	16
7	13	9
8	9	6
9	4	4
10	1	0
計	100	x

□(1)　男子の得点の平均値を求めよ。

□(2)　男子について，中央値と最頻値を求めよ。

□(3)　男子，女子それぞれの得点の範囲を求めよ。

□(4)　表の x，y の値を求めよ。

★★ **337** 右の表１は，ある水田の稲の穂を20本取り，その穂についているもみ粒の個数を調べた結果であり，表２はこれを度数分布表にまとめたものである。

また，表の中の x，y はもみ粒の個数を，m，n は穂の本数を表したものである。

表１

82	77	x
79	80	73
76	86	78
y	81	79
81	84	75
80	79	82
77	74	

表２

階級(もみ粒の個数)	度数(本数)
以上　以下	
72〜74	m
75〜77	n
78〜80	7
81〜83	5
84〜86	2
計	20

□(1)　個数が78以上80以下の階級の相対度数を求めよ。

□(2)　m，n の値を求めよ。

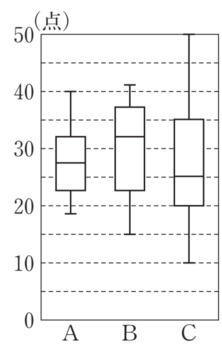
□(3)　$x-y=5$ であるとき，x，y の値を求めよ。

★★ **338** □ ある中学校のクイズ研究会で，全国中学生クイズ大会に出場する選手を，３人の会員A，B，Cの中から１人選ぶことになった。

右の箱ひげ図は，過去の大会で出題された問題を，地理や天文，ことわざなど12の分野に分けた各50点満点のクイズに，A，B，Cの３人が挑戦したときの得点を表したものである。

あなたなら，大会に出場する選手として，A，B，Cのどの会員を選ぶか。

記号で答え，その会員を選んだ理由を説明せよ。ただし，どの会員を選んで説明してもよい。



●放課後数学クラブ●　数学力を身につけよう　5章　データの活用

先　生：先日の校内実力テストの結果はどうだったかな？

か　い：後半の関数と空間図形が難しくて時間内に解ききれず，数学は77点でした。国語と英語も60点でいまいちな結果に…。

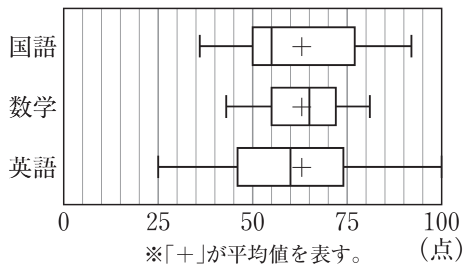
まどか：私は空間図形の問題は解けたけど，前半の基本問題でつまらない計算ミスをしちゃって，数学が80点。苦手な国語は52点，英語が76点でした。

先　生：実は，最後の空間図形の問題が解けていたのは，クラスでまどかさんだけだったんですよ。数学の最高点は81点でしたから，前半のミスさえなければまどかさんがクラス１位だったかもしれないですね。

まどか：うーん，悔しい！　でも，計算ミスも実力がまだまだ足りない証拠ですね。気を引き締めなくちゃ。

か　い：ところで先生，今回のテストは偶然にも，３教科ともクラスの平均値が63点だったんですが，ぼくは国語も英語も得点が平均値より少し$\boxed{\text{ア}}$だったのに，国語の順位はクラスの真ん中より少し$\boxed{\text{イ}}$で，英語の順位はクラスのちょうど真ん中あたりだったのが不思議でした。

先　生：君達のクラス40人の得点の分布の様子を箱ひげ図で表すと，右の図のようになっているんですよ。国語の$\boxed{\text{ウ}}$値は60点より低く，英語の$\boxed{\text{ウ}}$値は60点ちょうどなので，かいさんの順位がそのような結果になるのも当然です。



※「＋」が平均値を表す。

あなた：得点が平均値より上だからといって，必ずしも上位$\boxed{\text{エ}}$%のグループに含まれるとは限らないし，逆に，平均値より下だからといって，必ずしも下位$\boxed{\text{エ}}$%のグループに含まれるとは限らないということですね。

先　生：そうです。順位が全体の真ん中より上か下かが知りたければ，平均値ではなく$\boxed{\text{ウ}}$値を基準にすべきです。ところで，この箱ひげ図について，次のような問題を考えてみてください。

問題　次の㋐～㋓のうち，右上の箱ひげ図から判断できることとして，正しいものには○を，正しくないものには×をつけよ。

㋐　どの教科でも平均値の63点を下回ると，クラスでの順位は真ん中より下になる。

㋑　どの教科でも75点以上を取れば，クラスの上位25%に入っているといえる。

㋒　英語は得点の範囲が最も広いので，高得点から低得点まで比較的均等に得点が散らばっているといえる。

㋓　３教科の中では，最低点が最も高かった数学がいちばん問題がやさしかったといえる。

か　い：さっきの話の通り，クラスの真ん中より順位が下になるかどうかは，平均値ではなく$\boxed{\text{ウ}}$値から判断するんだよね。そうすると，「平均値を下回ったときにクラスの下位$\boxed{\text{エ}}$％に入っている」と確実にいえるのは，3教科の中では$\boxed{\text{オ}}$だけだから，㋐は$\boxed{\text{a}}$だよね。

まどか：㋑は$\boxed{\text{b}}$よね。「クラスの上位25％に入っている」かどうかは，箱の右端の値である第$\boxed{\text{カ}}$数を見て判断するけど，3教科の中では$\boxed{\text{キ}}$だけが75点以上を取っても上位25％に確実に入っているとはいえないわ。

あなた：㋒は，文章前半の「英語は得点の範囲が最も広い」という部分は正しいけれど，後半の「均等に得点が散らばっている」の部分は正しいとはいえないよ。
例えば，英語の上位25％の10人について，その散らばりを考えると，

説明の続きを書こう！❶

下位10人についても同様に考えられるので，箱ひげ図からだけでは，㋒が正しいとはいいきれないよ。したがって，㋒は$\boxed{\text{c}}$だと思うね。

先　生：そうですね。**箱ひげ図は全体の中で自分がどのあたりのグループにいるかを把握するのには役立ちます**が，分布の様子はわかりにくいです。外れ値，つまり，極端に高い得点や低い得点を取る人が1人いるだけでも分布の範囲は広がってしまいますし，分布のかたよりも見えてきません。**分布の様子を知りたいなら度数分布表やヒストグラムの方がいい**ですね。線分上に得点を「•」などを使って表すドットプロットなどもわかりやすいですよ。

まどか：㋓は正しいんじゃないかしら…。「最低点が最も高い」ということは基本的な問題が多かったということでしょ。実際に数学の前半の問題はやさしかったわ。私は計算ミスをしちゃったけれど。

か　い：そうかなぁ…。確かに前半は簡単だったけれど後半の関数や図形は難しかったよ。みんなもほとんどできてなかったんでしょ。それに「最低点が最も高い教科のテストはやさしい」と判断することと「数学のテストがやさしい」と結論付けることは，なんだか※矛盾していないかい？

※矛盾…つじつまの合わないこと。

まどか：どうして矛盾だといえるの？

あなた：「最低点が最も高い教科のテストはやさしい」と考えるのなら，逆に，

説明の続きを書こう！❷

よって，その考えは矛盾しているといえるんだ。つまり，㊁は d ということになるね。それに，数学のテストがやさしいか難しいかは，誰か１人が取った最低点や最高点だけでは判断できないと思うよ。

先　生：何をもってそのテストの難易度を判断するのかは，とても難しいことです。分布の様子が大きく異なっても平均値は一致することもあるので，平均値だけではテストの難しさは計れませんし，中央値もまたしかりです。

か　い：最高点・最低点からわかる分布の範囲だけでも判断できないですよね。例えば，英語は100点の人がいるからやさしかったかというと，たまたま得意な人がいただけかもしれないし，逆に25点の人がいるから難しかったかというと，問題自体が難しかったのか，極端に苦手な人がいたのか，テスト中に体調が悪かった人がいたのか，解答欄を間違えたのか，とかいろいろ考えられます。

先　生：結局のところ，そのテストが難しいか簡単かというのは主観的なものでもありますし，異なる教科のテストの難易度を一律に比較しても仕方ないですね。**与えられたデータを見ていろいろと考えることは大切ですが，データからだけでは判断できないことを結論として決め付けるのもよくありませんよ。**

考えてみよう！　会話を読んで，次の問いに答えよう。

① ア ～ キ にあてはまる言葉や数を答えよう。

② a ～ d には，「○」「×」のどちらがあてはまるのか，答えよう。

③ ❶，❷に続きを書いて，説明を完成させよう。

<たすく先生の数学力向上ポイント>

- **箱ひげ図，度数分布表，ヒストグラムなどはデータの何を知りたいかによって，それぞれ向き・不向きがある。度数分布表やヒストグラムの方が分布の様子を見るには適していることもある。**
- **与えられたデータからいろいろと推測することはよいが，データだけでは判断できないことを「結論」としてはいけない。**

●放課後数学クラブ●　**生活の中の数学を知ろう**　5章　データの活用

◆見た目のデータに騙されないようにしよう！

2600	1160	820	540	
400	350	320	270	200

（単位：万円）

くじなどでかたよりなく選ばれた９世帯の年収が右のようになっていたとします。この９世帯の年収の和は6660万円ですから，年収の平均値は，6660÷9＝［ア］（万円）です。９世帯の年収の中央値は［イ］万円なので，平均値が中央値よりだいぶ高く感じられますが，これは「年収が高い2600万円の世帯が平均値を大きく押し上げている」と考えることもできます。

話を単純にするために，上の例では９世帯という少ないデータで考えましたが，実際の世界でも**ごく少数の飛びぬけた値が平均値に大きく影響する**ことが起こります。年度にもよりますが，日本の年収の平均値は440万円程度，中央値は370万円程度になっていて，平均値を上回ると，年収が高い方から上位35％に入ってしまう年もあります。平均値だから真ん中ぐらいという考えは誤りで，**平均値より上か下かだけを見て判断すると，データの特徴を見誤ってしまう**こともあるのです。

次に，年収500万円以上を「高所得」，500万円未満を「低所得」というグループにして，それぞれのグループの平均値を求めると，

高所得グループ…（2600＋1160＋820＋540）÷4＝［ウ］（万円）

低所得グループ…（400＋350＋320＋270＋200）÷5＝［エ］（万円）

となりますが，ここで不景気が起こり，翌年の各世帯の年収が１割ずつ減ったとします。すると，年収540万円だった世帯の年収は，540×（1−0.1）＝［オ］（万円）になり，低所得グループに移ってしまうので，それぞれのグループの平均値は，

高所得グループ…（2340＋1044＋738）÷3＝［カ］（万円）

低所得グループ…（［オ］＋360＋315＋288＋243＋180）÷6＝［キ］（万円）

となります。つまり，不景気で各世帯の年収は下がったはずなのに，それぞれのグループの年収の平均値は上がってしまうのです。９世帯全体の平均値は，［ア］×（1−0.1）＝［ク］（万円）となるので，全体の平均値からは確かに不景気の影響が感じられますが，～～のようなデータの見方をすると，**「不景気なのに各世帯の年収が増えている」と錯覚する**結果になってしまいます。

ニュース番組や新聞などで「景気回復」「失業率改善」などといわれることもありますが，大切なのは，**どんなデータをどのように見たか**ということです。**データを見るときには，その扱われ方・見せ方がどのようになっているのかにも注意しましょう。**

考えてみよう！　上の説明を読んで，次の問いに答えよう。

①　［ア］～［ク］にあてはまる数を答えよう。

②　高校生と中学生がある検定試験を受験し，試験は２つの会場Ａ，Ｂに分けて行いました。得点の平均値を会場ごとにまとめたら，右の表のようになりました。会場Ｂの方が高校生も中学生も平均値が高いのに，全体では会場Ａの方が平均値が高くなっています。この理由として考えられることを，説明しよう。

	高校生	中学生	全体
A	90	60	84
B	94	70	82

（単位：点）

※高校生も中学生も同じ検定試験を受けていて，どちらの会場にも高校生，中学生が含まれるものとする。