1 (1) $\frac{1}{18}(4+5+6+5+9+4+7+6+4+6+5+4+0+5+6+7+10+6)$

$=\frac{99}{18}=5.5$ (点)

(2) データを小さい順に並べると

0, 4, 4, 4, 4, 5, 5, 5, 5, 6, 6, 6, 6, 7, 7, 9, 10

よって，このデータの中央値は $\frac{5+6}{2}=5.5$ (点)

(3) このデータの最頻値は 6 点

2 (1) ヒストグラムから，このデータの最頻値は 2 (人)

組数は 30 組であるから，小さい方から 15 番目と 16 番目の平均値が中央値となる。

小さい方から 15 番目の人数は 2 人，

16 番目の人数は 3 人

よって，中央値は $\frac{2+3}{2}=2.5$ (人)

(2) 平均値は $\frac{1}{30}(1\times6+2\times9+3\times6+4\times7+5\times1+6\times1)=\frac{81}{30}=2.7$ (人)

3 度数が最も大きい階級の階級値は，27.5 個である。

よって，このデータの最頻値は 27.5 個

4 データの平均値が 30 であるとき $\frac{1}{7}(21+33+10+49+47+12+a)=30$

これを解いて $a=38$

5 各組の点数の合計は，(平均値)×(人数) で与えられる。

よって，20 人全員の点数の合計は $6.4\times5+7.5\times8+6.0\times7=134$

したがって，全員の点数の平均値は $\frac{134}{20}=6.7$ (点)

6 (1) 得点の平均値が最小となるのは，すべての得点が階級内の最小の値となるときであるから $\frac{1}{20}(40\times2+50\times6+60\times8+70\times4)=\frac{1140}{20}=57$ (点)

(2) 得点の平均値が最大となるのは，すべての得点が階級内の最大の値となるときであるから $\frac{1}{20}(49\times2+59\times6+69\times8+79\times4)=\frac{1320}{20}=66$ (点)

7 データを大きさの順に並べると

23, 24, 26, 29, 30

(1) 中央値は 26 人

平均値は $\frac{1}{5}(23+24+26+29+30)=\frac{132}{5}=26.4$ (人)

(2) 実際の平均値は (1) で求めたものより 0.4 小さいので，どれかが 2 人多い。

1 つ選んで 2 人減らした結果，中央値が 24 になるものは，26 という誤ったデータのみである。

よって，誤っている数値は 26 人，正しい数値は 24 人

8 (1) 中央値が 350 人を超えているものであるから，C 店である。

(2) 第 1 四分位数が 250 人未満であるから，A 店である。

(3) B 店の最小値は 200 人未満であり，第 1 四分位数は 200 人を超えている。

よって，少なくとも 1 日は 200 人未満であるが，それ以外の日は 200 人を超えた可能性がある。

したがって，最大で 30日。

9 (1) それぞれのデータを大きさの順に並べると

A市 2, 5, 5, 7, 10, 11, 13, 13, 13, 14, 14, 18

B市 4, 5, 7, 7, 8, 9, 10, 12, 13, 13, 14, 15

よって、それぞれのデータの範囲は

A市 $18-2=16$　　B市 $15-4=11$

(2), (3), (4)

A市のデータについて、最小値，第 1 四分位数，中央値，第 3 四分位数，最大値は，それぞれ

A 市 2, $Q_1=\frac{5+7}{2}=6$, $\frac{11+13}{2}=12$, $Q_3=\frac{13+14}{2}=13.5$, 18

である。従って、

四分位範囲は $Q_3-Q_1=13.5-6=7.5$ (日) 四分位偏差は $\frac{Q_3-Q_1}{2}=\frac{7.5}{2}=3.75$ (日)

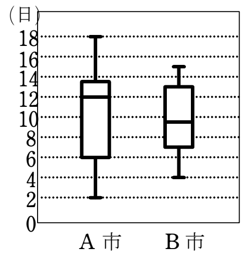
またB市のデータについて、最小値，第 1 四分位数，中央値，第 3 四分位数，最大値は，

それぞれ B 市 4, $Q_1=\frac{7+7}{2}=7$, $\frac{9+10}{2}=9.5$, $Q_3=\frac{13+13}{2}=13$, 15

である。従って、

四分位範囲は $Q_3-Q_1=13-7=6$ (日)　　四分位偏差は $\frac{Q_3-Q_1}{2}=\frac{6}{2}=3$ (日)

したがって，箱ひげ図は [図] のようになる。



(5) (4) より，A 市の箱ひげ図の方が B 市の箱ひげ図より長く，分布の範囲が広いため，A 市の方がデータの散らばりの度合いが大きいと考えられる。

10 平均値が $a+1$ であるから $\frac{1}{5}(6+11+15+17+a)=a+1$

よって $a+49=5(a+1)$

これを解いて $a=11$

したがって $\overline{x}=a+1=12$

分散 s^2 は $s^2=\frac{1}{5}\times72=14.4$

x	6	11	15	17	11	計 60
$(x-\overline{x})^2$	36	1	9	25	1	計 72

11 (1) 15 個のデータの和は $9\times10+6\times5=120$

よって，このデータ全体の平均値は $\frac{120}{15}=8$

(2) 10 個の値の 2 乗の平均値を a とすると，分散と平均値の関係式により $3=a-9^2$

よって $a=84$

同様に，残り 5 個の値の 2 乗の平均値を b とすると $9=b-6^2$

よって $b=45$

ゆえに，15 個の値の 2 乗の和は $a\times10+b\times5=84\times10+45\times5=1065$

したがって，このデータ全体の分散は，分散と平均値の関係式により

$\frac{1065}{15}-8^2=71-64=7$

12 (1) 平均値 $\overline{x}$ は $\overline{x}=\frac{1}{6}(5+1+8+10+4+8)=\frac{1}{6}\times36=6$

偏差の 2 乗は右の表のようになるから，分散 s^2 は

x	5	1	8	10	4	8	計 36
$(x-\overline{x})^2$	1	25	4	16	4	4	計 54

$s^2=\frac{1}{6}(1+25+4+16+4+4)$

$=\frac{1}{6}\times54=9$

(2) $\overline{x^2}=\frac{1}{6}(5^2+1^2+8^2+10^2+4^2+8^2)=\frac{1}{6}\times270=45$

分散 s^2 は $s^2=\overline{x^2}-(\overline{x})^2=45-6^2=9$

(3) 標準偏差 s は $s=\sqrt{s^2}=\sqrt{9}=3$

13 (1) 変量 x のデータの各値と仮平均 $x_0=630$ との差を表にすると，次のようになる。

x	672	693	644	665	630	644	計
$x-x_0$	42	63	14	35	0	14	y

$x-x_0$ の合計 y の値は

$y=42+63+14+35+0+14=168$

よって $\overline{x}=630+\frac{168}{6}=630+28=658$

(2) $u=\frac{x-630}{7}$ とおくと，次のようになる。

x	672	693	644	665	630	644	計
u	6	9	2	5	0	2	24
u^2	36	81	4	25	0	4	150

$\overline{u^2}-(\overline{u})^2=\frac{150}{6}-\left(\frac{24}{6}\right)^2=25-16=9$

よって，u のデータの標準偏差は $\sqrt{9}=3$

ゆえに，x のデータの標準偏差は $3\times7=21$

したがって，x のデータの分散は $21^2=441$

14 x, y のデータの平均値は $\overline{x}=\frac{60}{10}=6$, $\overline{y}=\frac{70}{10}=7$

番号	x	y	$x-\overline{x}$	$y-\overline{y}$	$(x-\overline{x})(y-\overline{y})$	$(x-\overline{x})^2$	$(y-\overline{y})^2$
1	7	6	1	−1	−1	1	1
2	5	7	−1	0	0	1	0
3	9	7	3	0	0	9	0
4	6	10	0	3	0	0	9
5	3	3	−3	−4	12	9	16
6	5	9	−1	2	−2	1	4
7	8	7	2	0	0	4	0
8	7	7	1	0	0	1	0
9	4	7	−2	0	0	4	0
10	6	7	0	0	0	0	0
計	60	70			9	30	30

表から，相関係数 r は $r=\frac{9}{\sqrt{30\times30}}=\frac{9}{30}=0.3$