



大ざっぱに言って
 この直線の周囲に
 データがありそう。
 → $y = ax + b$
 で表せそう。
 ($a > 0$)
 正の相関が
 ありそうで
 ある。

rep3ex.txt

```

> sidou<- read.csv("table2-1.csv")
> sidou
  No name sex math toukei sinri toukei1 toukei2 sidou
1   1  NA   m    n     y    13      6     10     c
2   2  NA   m    n     y    14     10     13     b
3   3  NA   m    y     y     7      6      8     b
4   4  NA   m    y     y    12     10     15     a
5   5  NA   m    n     n    10      5      8     b
6   6  NA   m    n     n     6      3      6     c
7   7  NA   m    n     n     8      5      9     a
8   8  NA   m    n     n    15      9     10     d
9   9  NA   m    n     n     4      3      7     d
10  10 NA   m    y     n    14      3      3     d
11  11 NA   f    y     y     9     11     18     a
12  12 NA   f    n     y     6      6     14     a
13  13 NA   f    y     y    10     11     18     a
14  14 NA   f    n     n    12      9     11     c
15  15 NA   f    n     y     5      7     12     b
16  16 NA   f    y     n    12      5      5     d
17  17 NA   f    n     n     8      8      7     c
18  18 NA   f    n     n     8      7     12     c
19  19 NA   f    n     n    12      7      7     b
20  20 NA   f    n     n    15      9      7     d

```

```

> cor(sidou$toukei1, sidou$toukei2)
[1] 0.749659 ← 統計テスト1と統計テスト2の相関係数
> cross <- table(sidou$math, sidou$toukei)
> cross

```

強い相関があると考えられ。
 (散布図で相関が正でどの程度
 大きくなったことがわかってい
 ることとあう。)

統計
 数字 n 10 y 4
 y 2 4
 ← 両方とも嫌いな人が多い。

```

> math10 <- ifelse(sidou$math=="y", 1, 0)
> toukei10 <- ifelse(sidou$toukei=="y", 1, 0)
> cor(math10, toukei10)

```

[1] 0.3563483 ← 数字が好きで統計が好き人のカイ係数 弱い相関があると考えられ
 (両方とも嫌いな人が多いことがクロス集計表からわかるが、
 それとあう結果になっていると考えられる。)