4

2

1

```r
rm(list=ls()); gc();  gc(); #<1>
if (!require("pacman")) install.packages("pacman") #<2>
pacman::p_load(tidyverse, magrittr,estimatr,car,modelsummary,ggrepel,patchwork)
```
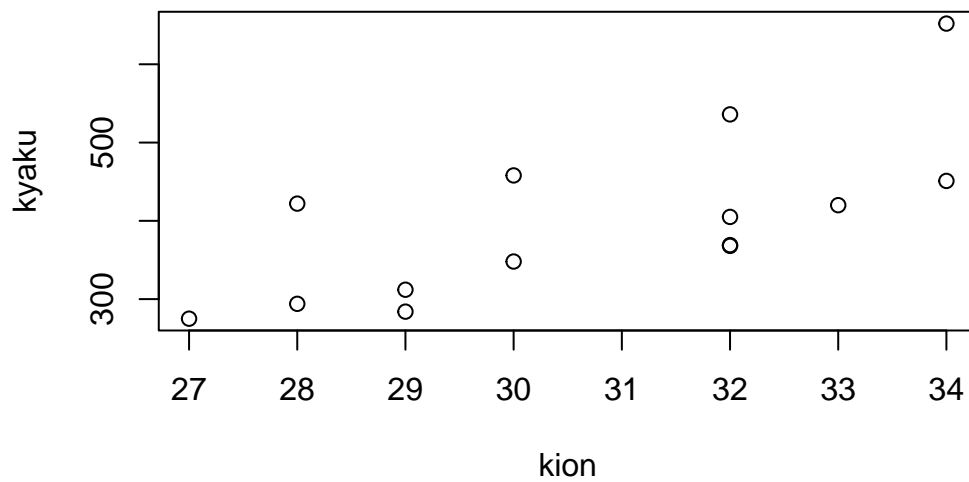
①   r
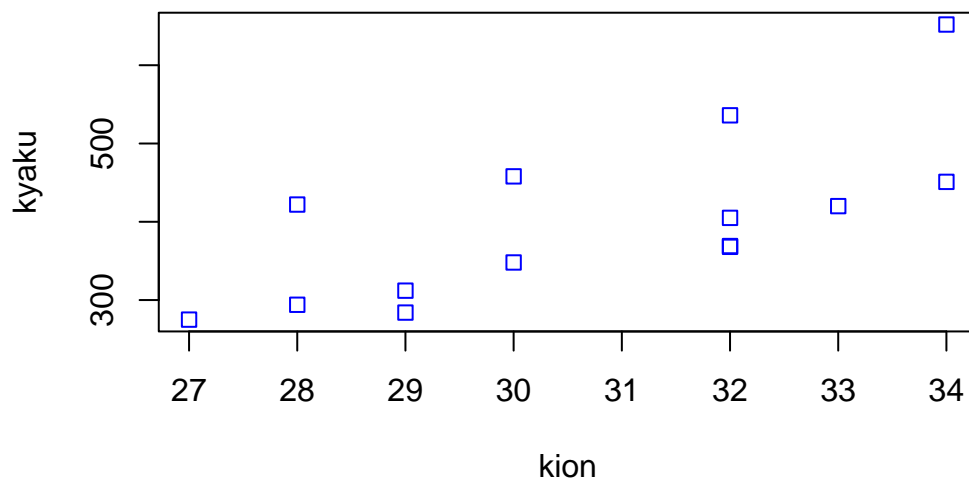②   pacman

```r
icedata <- read_csv("data/ice1_1.csv")
```

x   y

plot(x   , y   )

```r
icedata %$%
  plot(kion,kyaku)

icedata %$%
  plot(kion,kyaku)
```

```
icedata %$%
  plot(kion,kyaku,
      pch = 0, cex = 1, col = "blue")
```
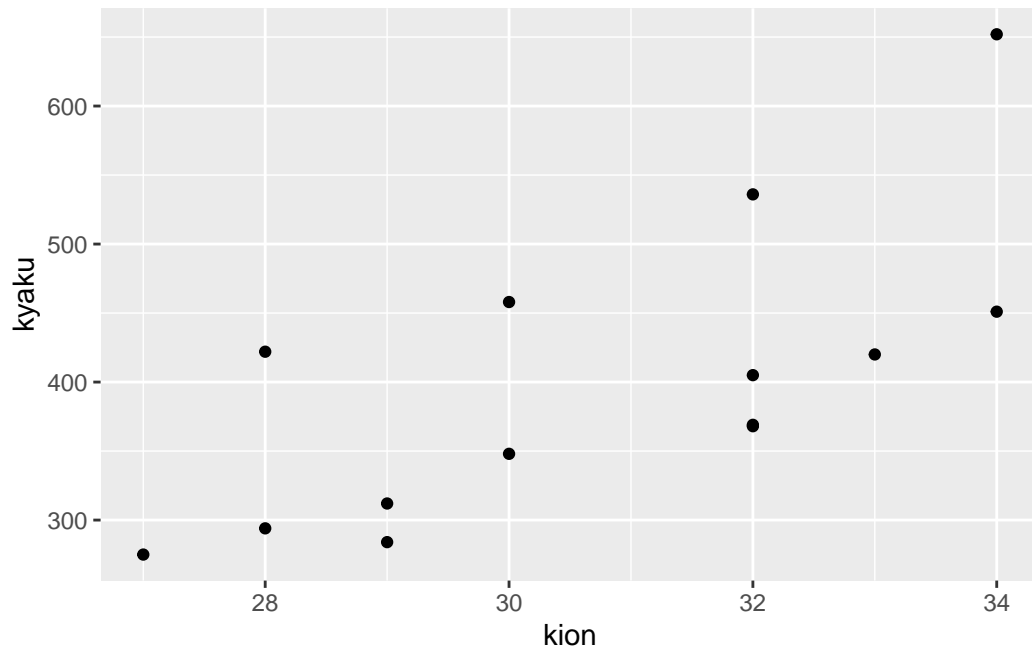

ggplot2　（tidyverse　　）

```
a <- ggplot(data = icedata, mapping = aes(kion, kyaku)) %>% #
    + geom_point() #    (point)
plot(a)
```
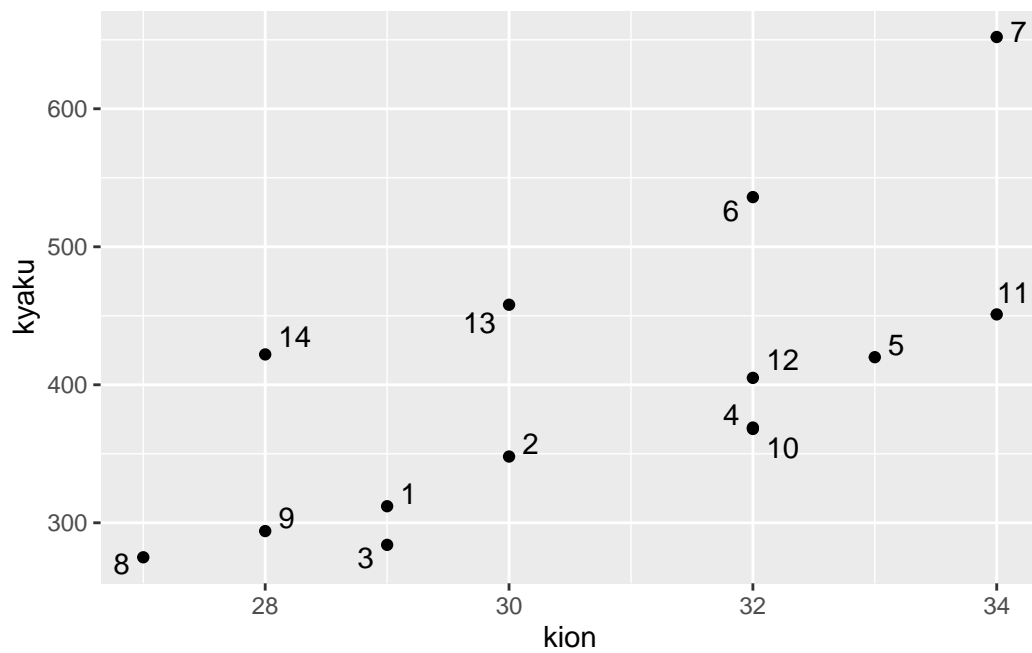
```
b <- ggplot(data = icedata, mapping = aes(kion, kyaku, label = day)) %>% #
    + geom_point() %>% #   (point)
    + geom_text_repel() #
plot(b)
```
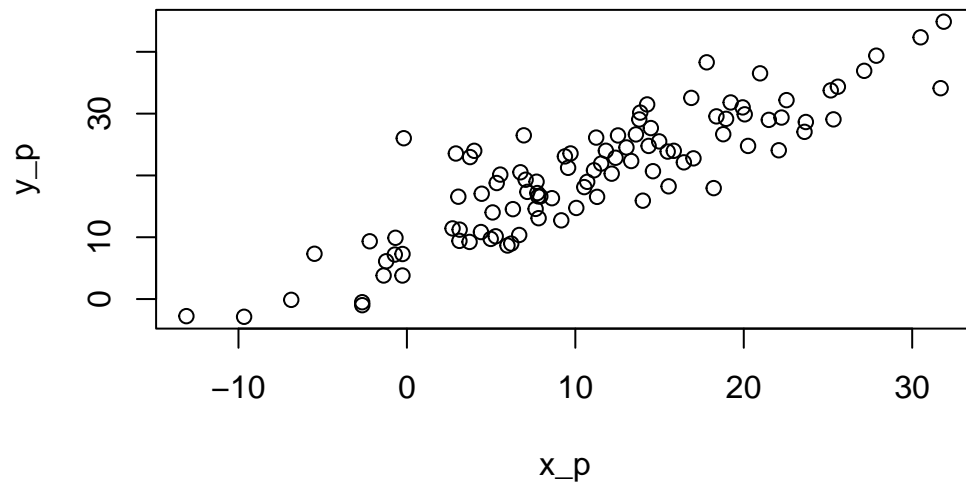

4

```
set.seed(123)
x_p <- rnorm(100,10,10)
y_p <- x_p + rnorm(100,10,5)

x <- rnorm(100,10,10)
y <- rnorm(100,10,10)

x_n <- rnorm(100,10,10)
y_n <- -x_n - rnorm(100,10,5)


plot(x_p,y_p)
```



```
plot(x,y)
```
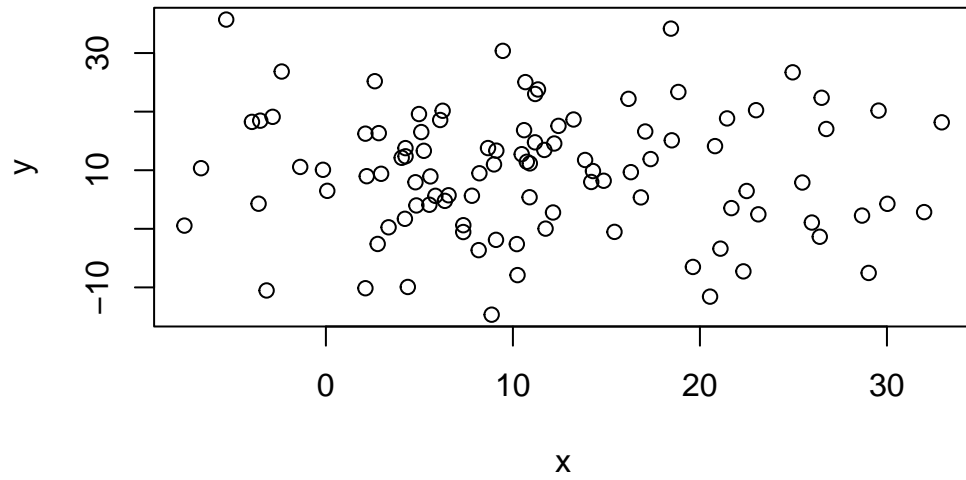
```
plot(x_n,y_n)
```
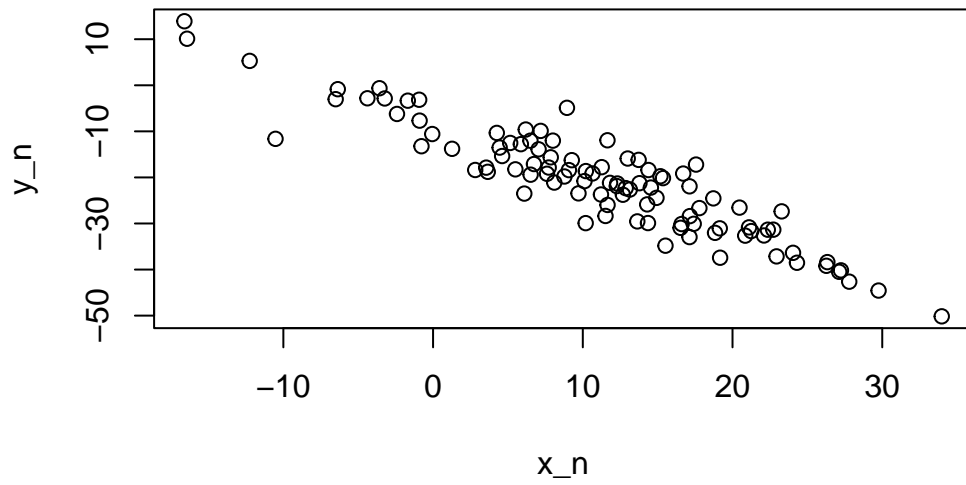

$$Cov_{xy} = \frac{1}{n}\Sigma_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)$$

1. x $\left(x_i - \bar{x}\right)$ y $\left(y_i - \bar{y}\right)$
2. $\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)$
3. 1 n $\Sigma_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)$

4.   n

R cov()

```
cov(x,y)
```

[1] -4.426973

```
cov(x_p,y_p)
```

[1] 81.13723

$$r = \frac{Cov_{xy}}{SD_x SD_y} = \frac{Cov_{xy}}{\sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(y_i - \bar{y})^2}}$$

r -1 1   -1   1

```
cov(x,y)/(sqrt(var(x))*sqrt(var(y)))
```

[1] -0.04486571

(correlation) cor()

```
cor(x_p,y_p)
```

[1] 0.8786993

```
cor(x,y)
```
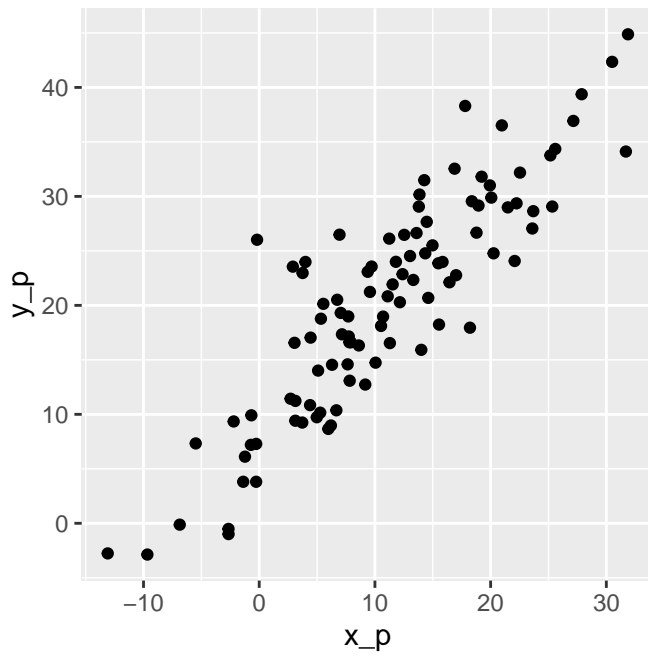
[1] -0.04486571
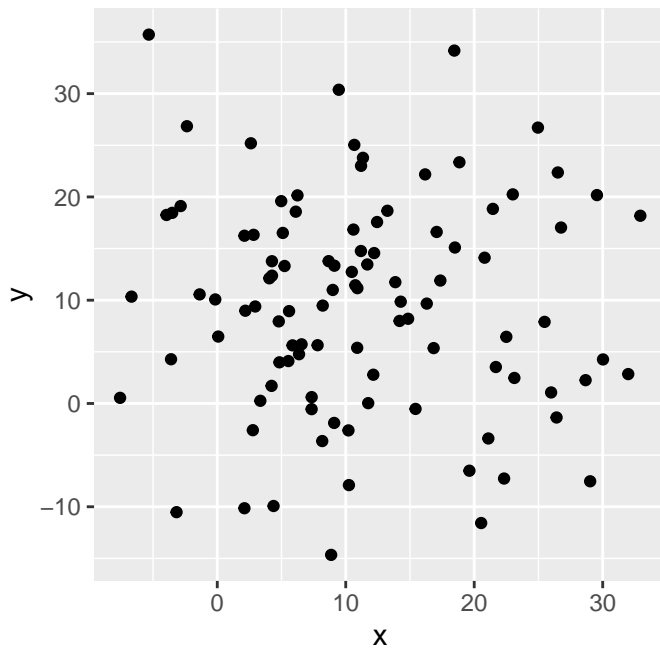
```r
cor(x_n,y_n)
```

```
[1] -0.9210776
```



```r
data.frame(x_p,y_p,x,y,x_n,y_n) |> #<1>
  datasummary_correlation() #<2>
```

①   data.frame
②   datasummary_correlation()   modelsummary

```r
anscombe # R        datasets
```

```
  x1 x2 x3 x4    y1   y2   y3   y4
1 10 10 10  8  8.04 9.14 7.46 6.58
2  8  8  8  8  6.95 8.14 6.77 5.76
```

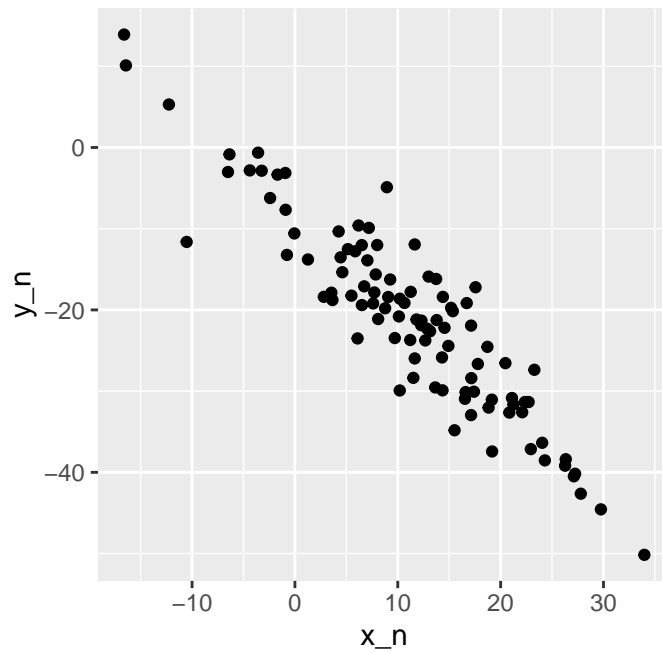|	x_p	y_p	x	y	x_n	y_n
x_p	1	.	.	.	.	.
y_p	0.88	1	.	.	.	.
x	−0.13	−0.10	1	.	.	.
y	−0.04	−0.02	−0.04	1	.	.
x_n	−0.19	−0.24	−0.02	−0.02	1	.
y_n	0.18	0.20	0.01	0.05	−0.92	1

```
3   13 13 13  8  7.58 8.74 12.74  7.71
4    9  9  9  8  8.81 8.77  7.11  8.84
5   11 11 11  8  8.33 9.26  7.81  8.47
6   14 14 14  8  9.96 8.10  8.84  7.04
7    6  6  6  8  7.24 6.13  6.08  5.25
8    4  4  4 19  4.26 3.10  5.39 12.50
9   12 12 12  8 10.84 9.13  8.15  5.56
10   7  7  7  8  4.82 7.26  6.42  7.91
11   5  5  5  8  5.68 4.74  5.73  6.89
```

```r
  anscombe_long <-
    anscombe |>
    tidyr::pivot_longer(
      tidyselect::everything(),
      names_to = c(".value", "set"),
      names_pattern = "(.)(.)")
  course_colors <- c("#364968", "#fddf97", "#e09664", "#6c4343", "#ffffff")


  anscombe_long #
```

```
# A tibble: 44 x 3
   set       x      y
   <chr> <dbl>  <dbl>
 1 1        10   8.04
 2 2        10   9.14
 3 3        10   7.46
 4 4         8   6.58
 5 1         8   6.95
 6 2         8   8.14
 7 3         8   6.77
 8 4         8   5.76
```

```
 9 1         13  7.58
10 2         13  8.74
# i 34 more rows
```

```r
  #
  # set
  # set
  #        set
  anscombe_long |>
    group_by(set) |>
    summarise(across(.cols = c(x, y), .fns = list(mean = mean, sd = sd)),
              .groups = "keep") |>
    summarise(across(.cols = contains("_"), .fns = ~ round(.x, digits = 2))) |>
    left_join(
      anscombe_long |>
        group_by(set) |>
        group_modify(~ tibble::tibble(cor = cor.test(.x$x, .x$y)$estimate)) |>
        ungroup() |>
        mutate(cor = round(cor, digits = 2)),
      by = "set")
```

```
# A tibble: 4 x 6
  set    x_mean  x_sd y_mean  y_sd    cor
  <chr>   <dbl> <dbl>  <dbl> <dbl>  <dbl>
1 1           9  3.32    7.5  2.03   0.82
2 2           9  3.32    7.5  2.03   0.82
3 3           9  3.32    7.5  2.03   0.82
4 4           9  3.32    7.5  2.03   0.82
```
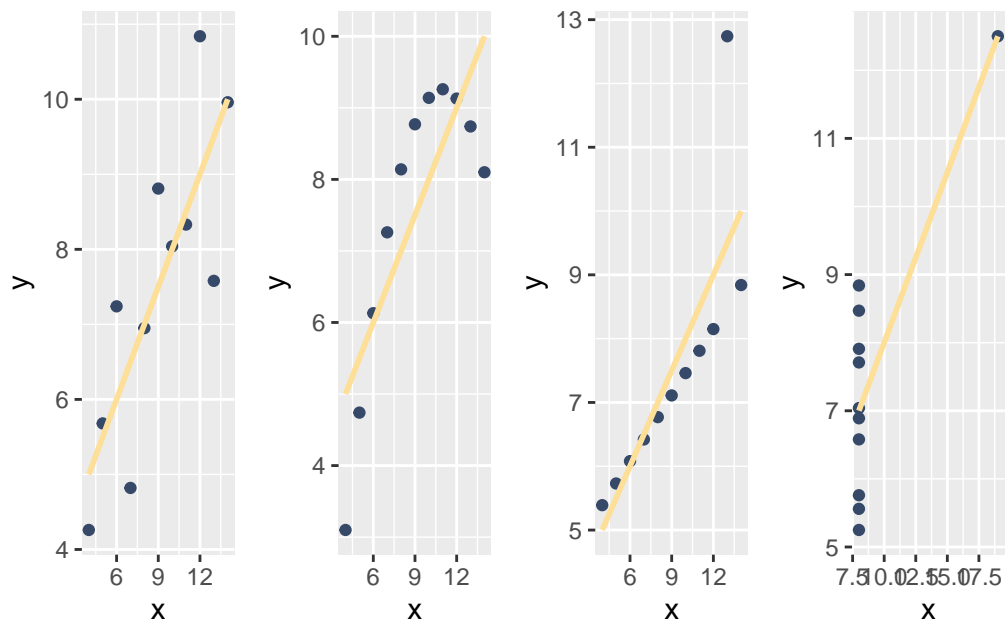
```r
  anscombe_long |>
    group_by(set) |>
    group_map(
      ~ ggplot(.x, aes(x, y)) +
        geom_point(color = course_colors[1]) +
        geom_smooth(method = lm,
                    se = FALSE,
                    color = course_colors[2])) |>
    wrap_plots(ncol = 4)
```

```
# ggsave(here("images/anscombes_quartet.png"),
#        width = 7,
#        height = 2.2)
```

```
icedata2 <- read_csv("data/ice2_5.csv")
```

（　）

plot()

```
icedata2 %$%
  plot(kyori, kyaku)
```