practical0

Reona

2023-05-12

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v dplyr     1.1.1     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.1     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.1
## -- Conflicts ------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```
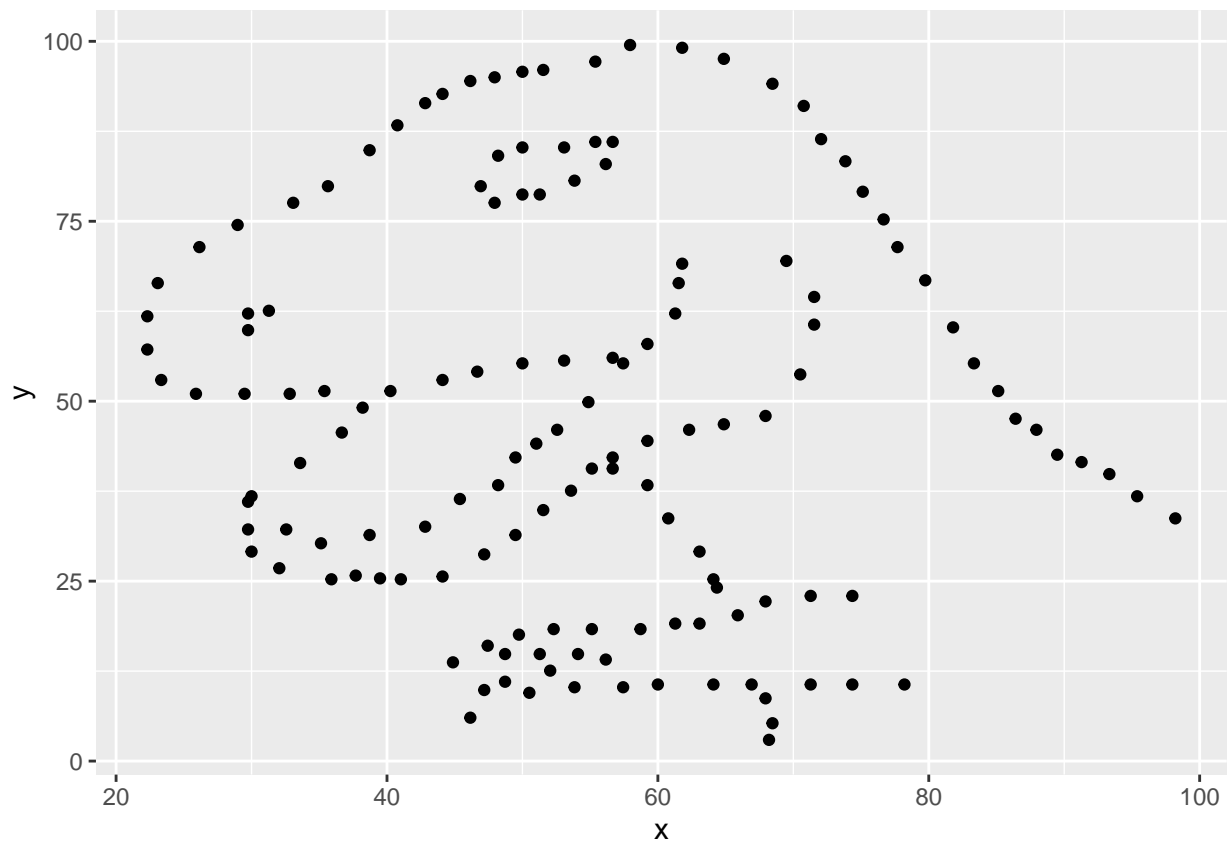
```
library(datasauRus)
```

```
datasaurus_dozen %>%
  count(dataset)
```

```
## # A tibble: 13 x 2
##    dataset        n
##    <chr>      <int>
##  1 away         142
##  2 bullseye     142
##  3 circle       142
##  4 dino         142
##  5 dots         142
##  6 h_lines      142
##  7 high_lines   142
##  8 slant_down   142
##  9 slant_up     142
## 10 star         142
## 11 v_lines      142
## 12 wide_lines   142
## 13 x_shape      142
```

```
dino_data <- datasaurus_dozen %>%
  filter(dataset == "dino")
```

```
ggplot(data = dino_data, mapping = aes(x = x, y = y)) +
  geom_point()
```
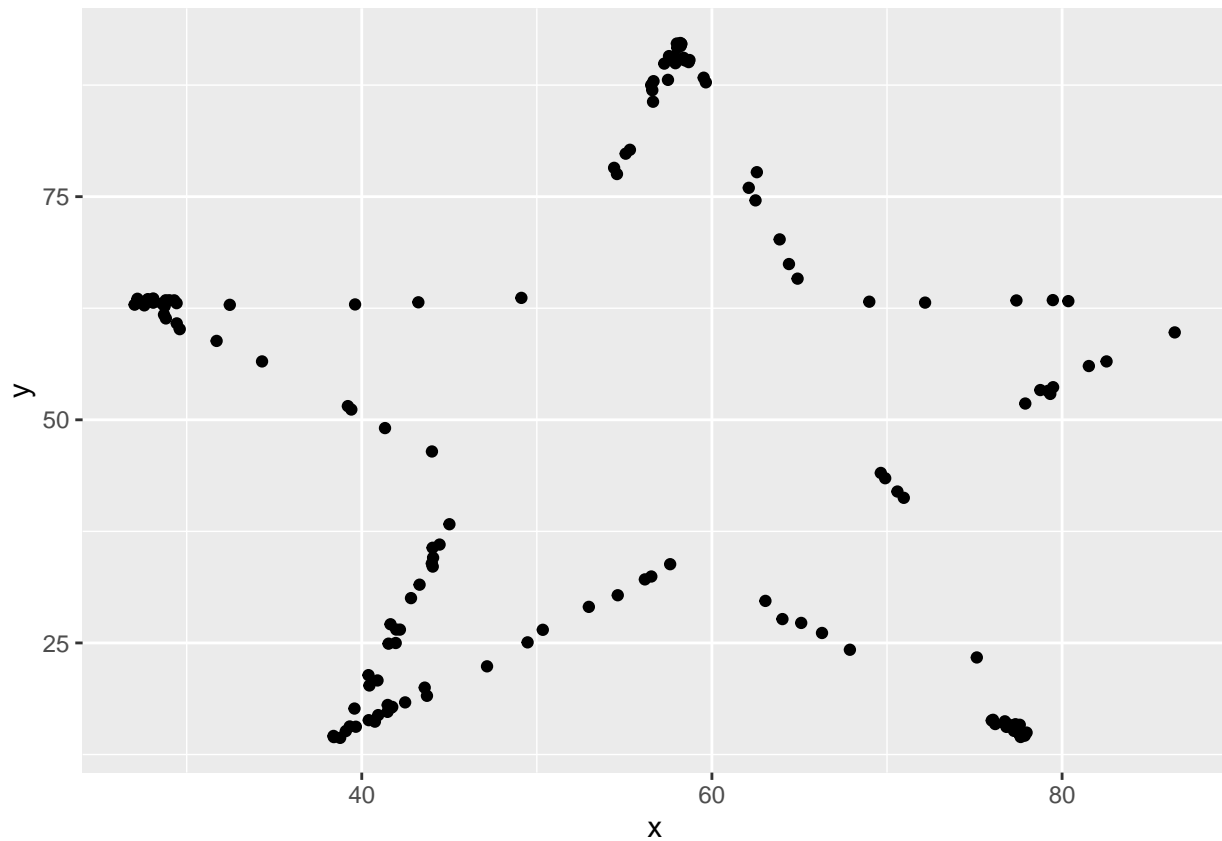
```r
dino_data %>%
  summarize(r = cor(x, y))
```

```
## # A tibble: 1 x 1
##         r
##     <dbl>
## 1 -0.0645
```

```r
star_data <- datasaurus_dozen %>%
  filter(dataset == "star")
```

```r
ggplot(data = star_data, mapping = aes(x = x, y = y)) +
  geom_point()
```
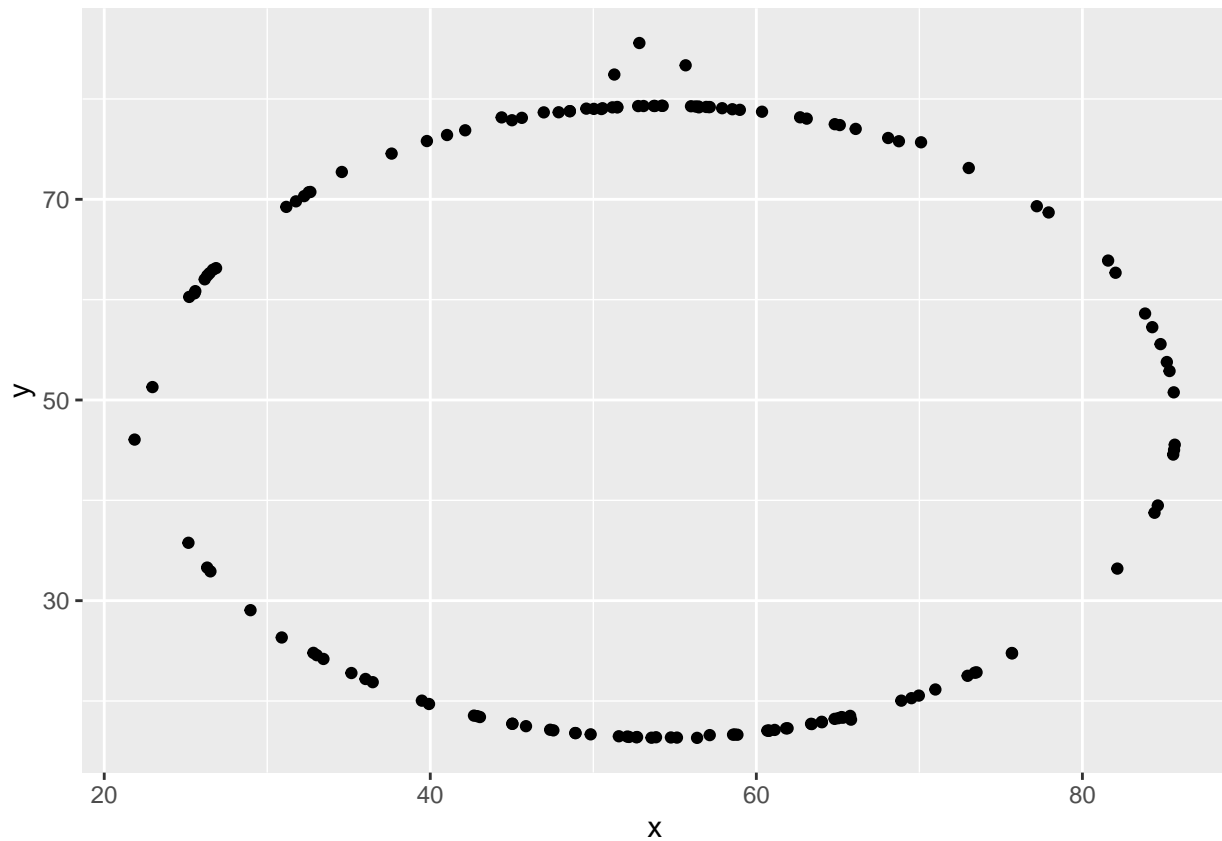
```
star_data %>%
  summarize(r = cor(x, y))
```

```
## # A tibble: 1 x 1
##        r
##    <dbl>
## 1 -0.0630
```

There is no great difference between correlations of dino plot and star plot.

```
circle_data <- datasaurus_dozen %>%
  filter(dataset == "circle")
```

```
ggplot(data = circle_data, mapping = aes(x = x, y = y)) +
  geom_point()
```
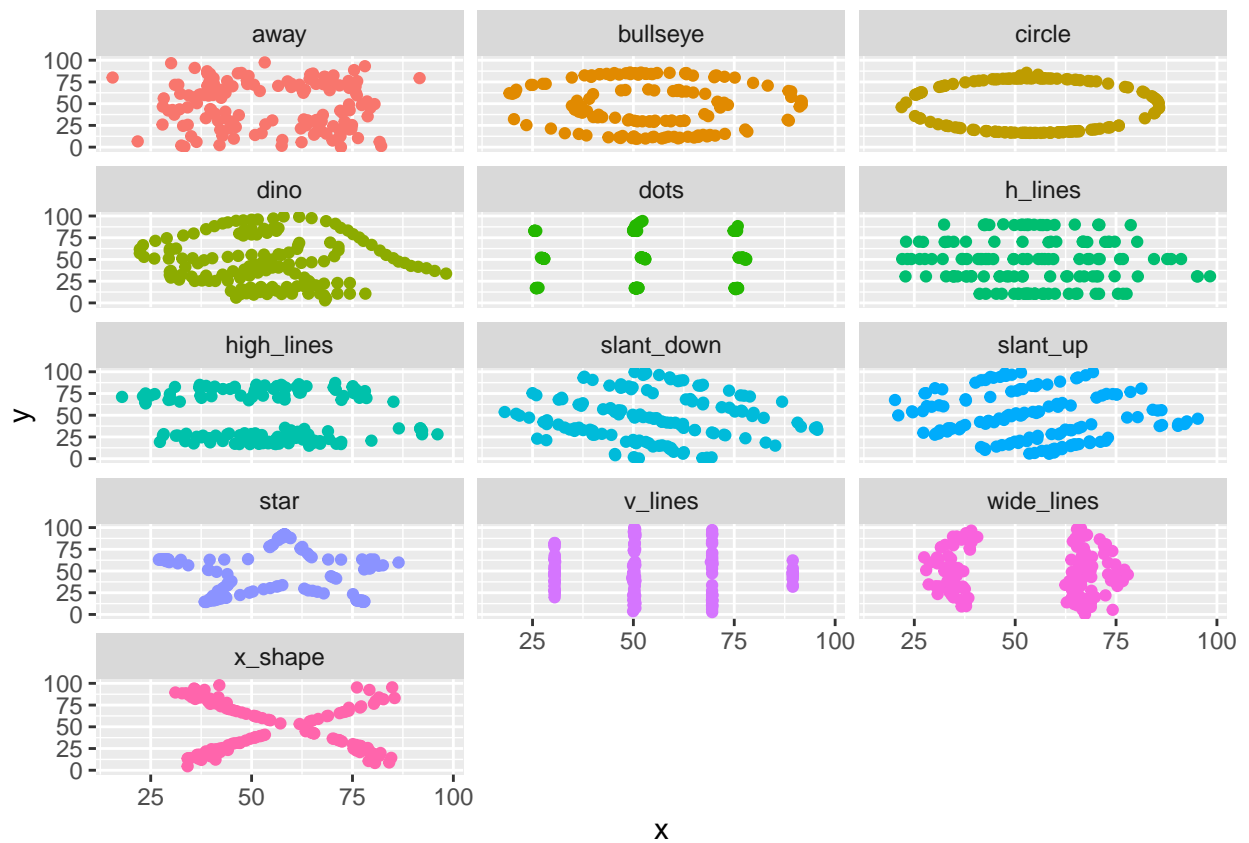
```
circle_data %>%
  summarize(r = cor(x, y))
```

```
## # A tibble: 1 x 1
##          r
##      <dbl>
## 1 -0.0683
```

There is no great difference between correlations of dino plot and circle plot.

```
ggplot(datasaurus_dozen, aes(x = x, y = y, color = dataset))+
  geom_point()+
  facet_wrap(~ dataset, ncol = 3) +
  theme(legend.position = "none")
```

```
datasaurus_dozen %>%
  group_by(dataset) %>%
  summarize(r = cor(x, y))
```

```
## # A tibble: 13 x 2
##    dataset           r
##    <chr>         <dbl>
##  1 away        -0.0641
##  2 bullseye    -0.0686
##  3 circle      -0.0683
##  4 dino        -0.0645
##  5 dots        -0.0603
##  6 h_lines     -0.0617
##  7 high_lines  -0.0685
##  8 slant_down  -0.0690
##  9 slant_up    -0.0686
## 10 star        -0.0630
## 11 v_lines     -0.0694
## 12 wide_lines  -0.0666
## 13 x_shape     -0.0656
```