Boqian_Mao_Biostat_620_Lab1

Boqian Mao

```
# install.packages("datasauRus")
library(datasauRus)
```

Question 1

```
nrow(datasaurus_dozen)
```

```
[1] 1846
```

```
ncol(datasaurus_dozen)
```

```
[1] 3
```

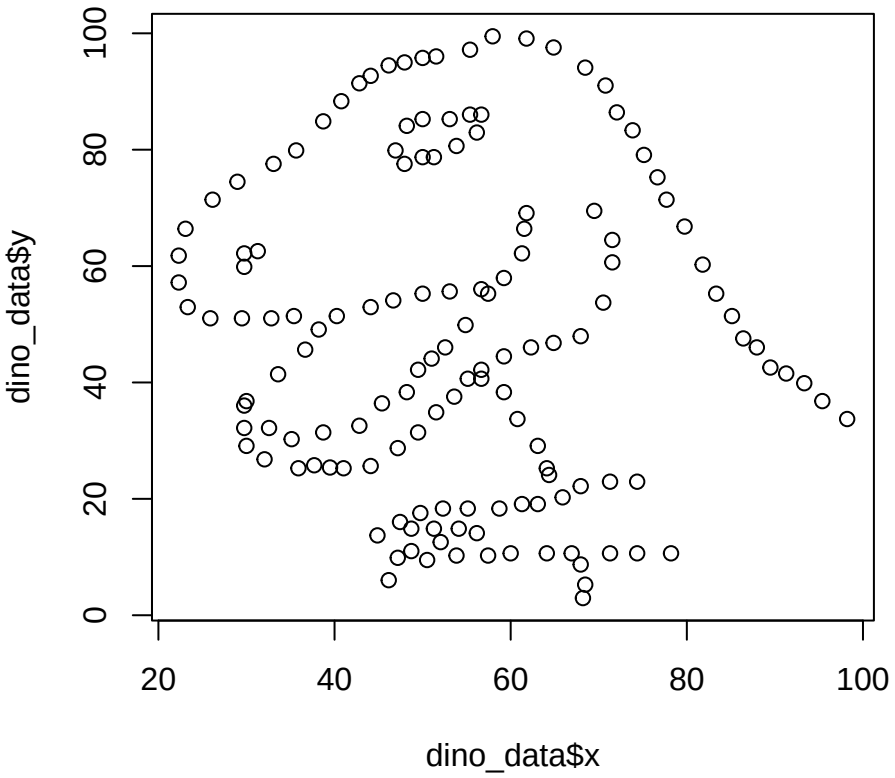
```
summary(datasaurus_dozen)
```

```
   dataset                x               y          
 Length:1846        Min.   :15.56   Min.   : 0.01512  
 Class :character   1st Qu.:41.07   1st Qu.:22.56107  
 Mode  :character   Median :52.59   Median :47.59445  
                    Mean   :54.27   Mean   :47.83510  
                    3rd Qu.:67.28   3rd Qu.:71.81078  
                    Max.   :98.29   Max.   :99.69468  
```

There are 1846 rows and 3 column in the dataset. Dataset, X, and Y as variables included in the dataset.

Question 2

```
dino_data <- datasaurus_dozen[datasaurus_dozen$dataset == 'dino', ]
plot(dino_data$x, dino_data$y)
```

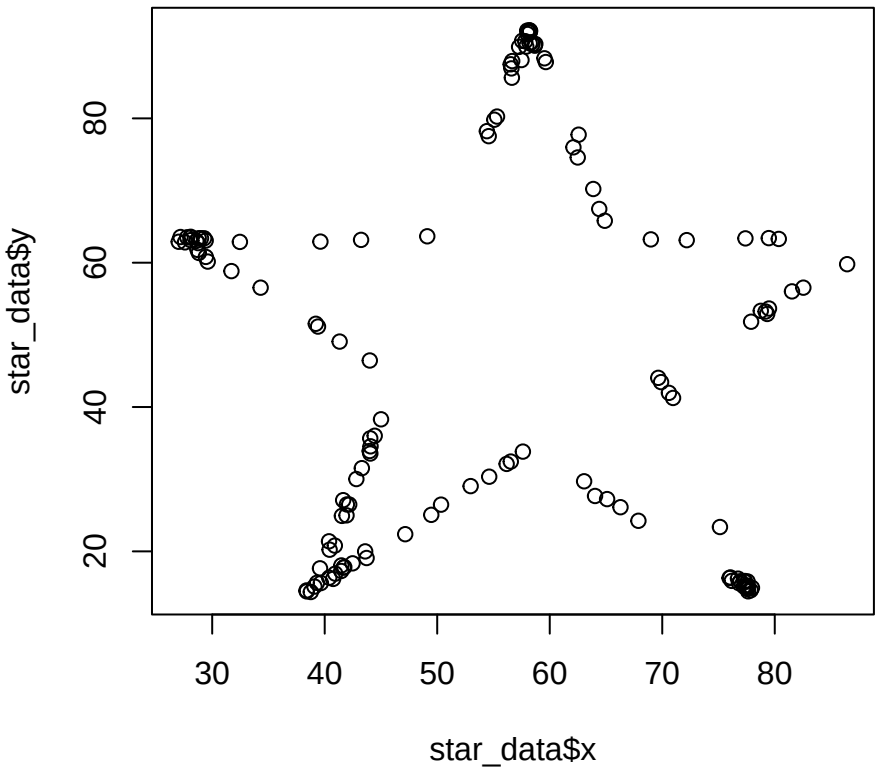



```
cor(dino_data$x, dino_data$y)
```

```
[1] -0.06447185
```

Question 3

```
star_data <- datasaurus_dozen[datasaurus_dozen$dataset == 'star', ]
plot(star_data$x, star_data$y)
```



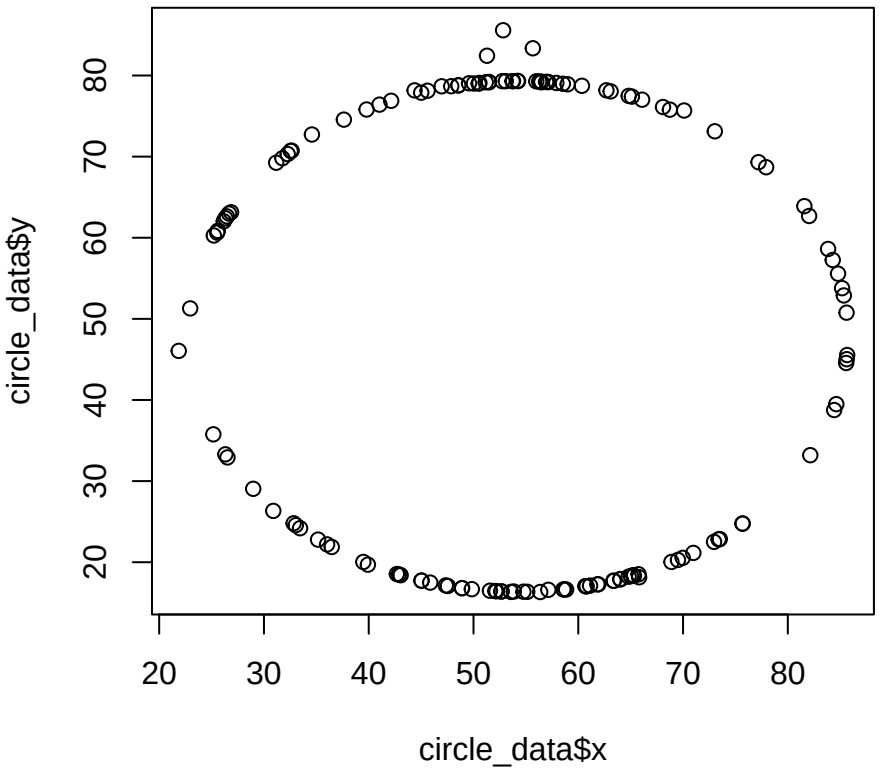

```
cor(star_data$x, star_data$y)
```

```
[1] -0.0629611
```

The correlation coefficient between x and y for dino and star dataset is very close.

Question 4

```
circle_data <- datasaurus_dozen[datasaurus_dozen$dataset == 'circle', ]
plot(circle_data$x, circle_data$y)
```



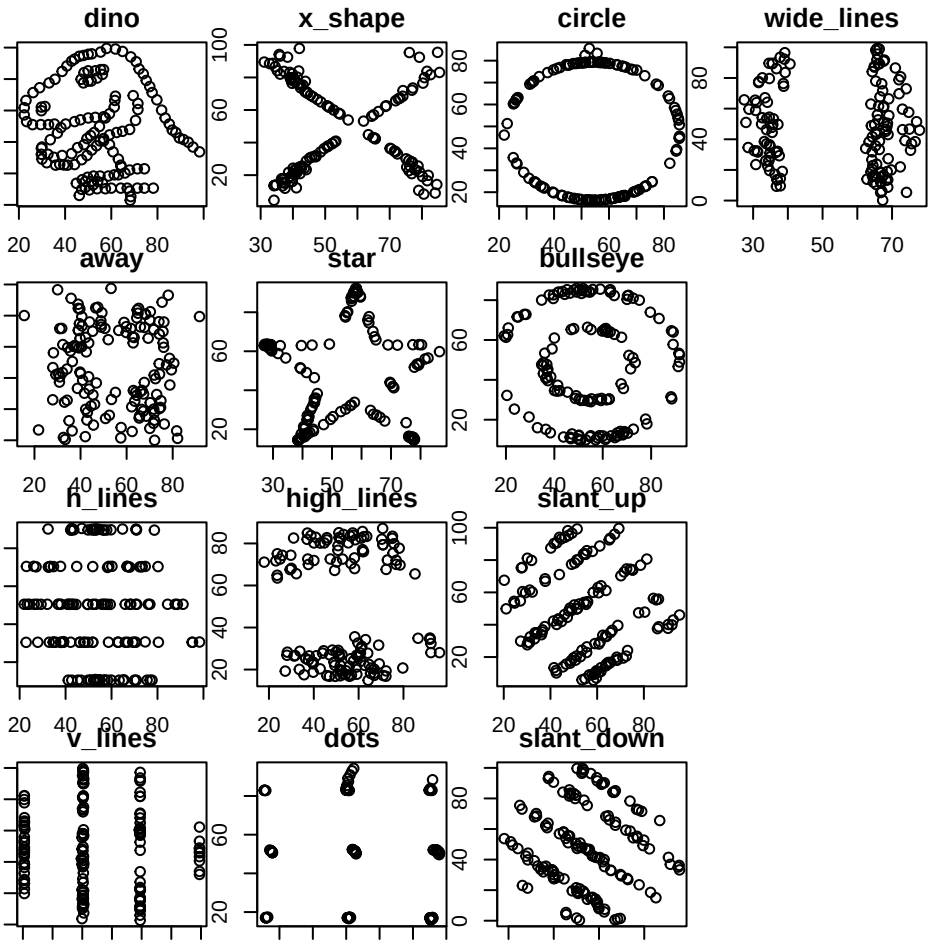

```
cor(circle_data$x, circle_data$y)
```

```
[1] -0.06834336
```

The correlation coefficient between x and y for dino and circle dataset is very close.

Question 5

```
layout(matrix(1:16, nrow=4, ncol=4))
par(mar = c(1, 1, 2, 1))
for(name in unique(datasaurus_dozen$dataset)){
  subset <- datasaurus_dozen[datasaurus_dozen$dataset == name, ]
  plot(subset$x, subset$y, main = name)
}
layout(1)
```



Question 6

```
sapply(unique(datasaurus_dozen$dataset), function(name){
    subset <- datasaurus_dozen[datasaurus_dozen$dataset == name, ]
    return(cor(subset$x, subset$y))
})
```

```
       dino         away      h_lines      v_lines      x_shape         star
-0.06447185 -0.06412835 -0.06171484 -0.06944557 -0.06558334 -0.06296110
 high_lines         dots       circle     bullseye     slant_up   slant_down
-0.06850422 -0.06034144 -0.06834336 -0.06858639 -0.06860921 -0.06897974
 wide_lines
-0.06657523
```