Sepehr Akbari

April 22nd, 2024

Regression and correlation in R

For each of the following problems,

(a) Compute the correlation between the variables.

(b) Find the equation of the regression line.

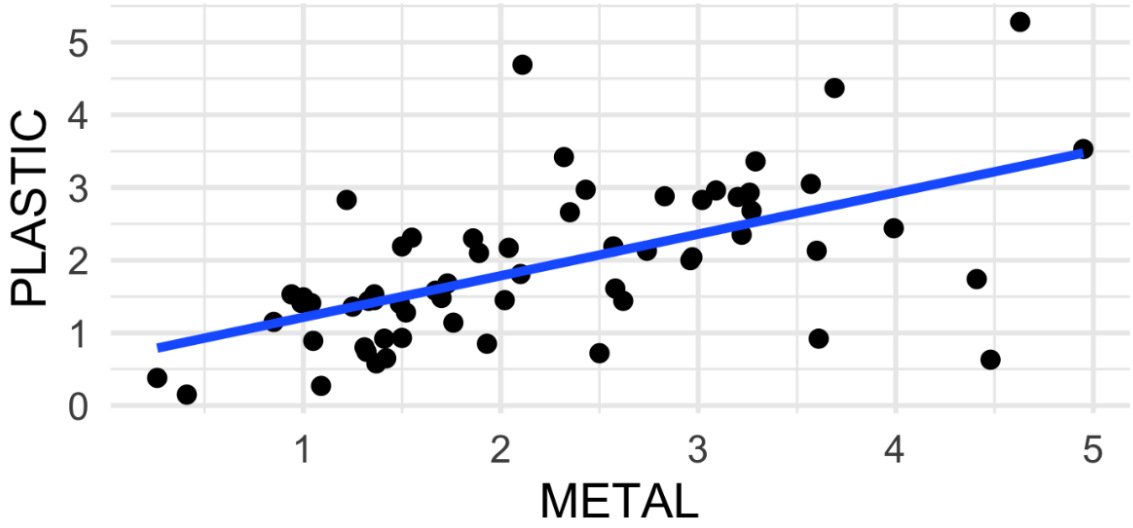
(c) Interpret the slope of the regression line in ordinary human language.

(d) Find the fitted value and residual for the specified observation, or say why doing so would be inappropriate.

Use R for all calculations. Include both code and output with your solutions.

Problem 1

Using the `garbage weight` data set, available on Moodle, model the weight of plastic waste using the weight of metal waste.



Consider an observation with 2.11 pounds of metal waste and 4.69 pounds of plastic waste.

Answer:

Loading the dataset:

```
library(readxl)
garbage_weight <- read_excel("/Users/sepehrakbari/Documents/LFC/Semester 2/MATH 150/DSs/garba
```

(a) correlation between the variables

```
cor(garbage_weight$PLASTIC, garbage_weight$METAL)
```

```
[1] 0.5862485
```

(b) equation of the regression line

```
model <- lm(PLASTIC ~ METAL, data = garbage_weight)

m <- coef(model)[["METAL"]]
b <- coef(model)[["(Intercept)"]]

cat("y =",round(m,2),"* x","+",round(b,2))
```

```
y = 0.57 * x + 0.64
```

(c) Interpretation

The analysis reveals a shallow upward trend between the amount of metal and plastic in the garbage. While the correlation is not very strong, on average, a one-unit increase in metal mass is associated with a modest increase of 0.57 units in plastic weight.

(d) fitted value and residual

```
x_metal <- 2.11
x_plastic <- 4.69

f <- ((m * x_metal) + b)
r <- x_plastic - f

cat("Fitted Value:", f,"\n")
```

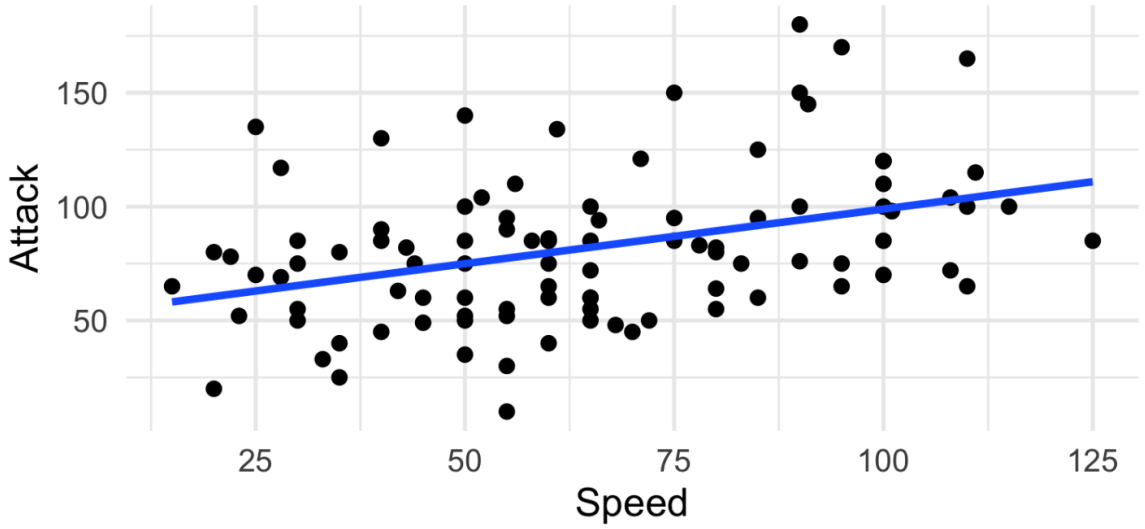
```
Fitted Value: 1.849123
```

```r
cat("Risidual:", r,"\n")
```

```
Risidual: 2.840877
```

Problem 2

Using the `pokemon sample` data set, available on Moodle, model the attack ratings of Pokemon using their speeds.



Consider a Pokemon with a speed of 150 and an attack of 100.

Answer:

Loading the dataset:

```r
library(readxl)
pokemon_sample <- read_excel("/Users/sepehrakbari/Documents/LFC/Semester 2/MATH 150/DSs/poke
```

(a) correlation between the variables

```r
cor(pokemon_sample$Speed, pokemon_sample$Attack)
```

```
[1] 0.4120036
```

(b) equation of the regression line

```
model <- lm(Attack ~ Speed, data = pokemon_sample)

m <- coef(model)[["Speed"]]
b <- coef(model)[["(Intercept)"]]

cat("y =",round(m,2),"* x","+",round(b,2))
```

```
y = 0.49 * x + 50.19
```

(c) Interpretation

A slightly positive trend can be observed in the data. On average, for every one-unit increase in speed, there's a modest increase of 0.5 units in attack.

(d) fitted value and residual

```
x_speed <- 150
x_attack <- 100

f <- ((m * x_speed) + b)
r <- x_attack - f

cat("Fitted Value:", f,"\n")
```

```
Fitted Value: 124.1666
```

```
cat("Risidual:", r,"\n")
```

```
Risidual: -24.1666
```