

exercicio2.R

caiop

2023-10-03

```
# Caio Broering Pinho
# Igor Zimmer Gonçalves
# Gustavo da Rocha Pereira

# CARREGAR O PACOTE
library(data.table)
# LEITURA DA BASE
base = fread(
  input = paste0("car_base.csv"),
  header = T,
  na.strings = "NA",
  data.table = FALSE,
  dec = ",",
)
# Convertendo para valores numéricos
base$price = as.numeric(base$price)
base$carwidth = as.numeric(base$carwidth)

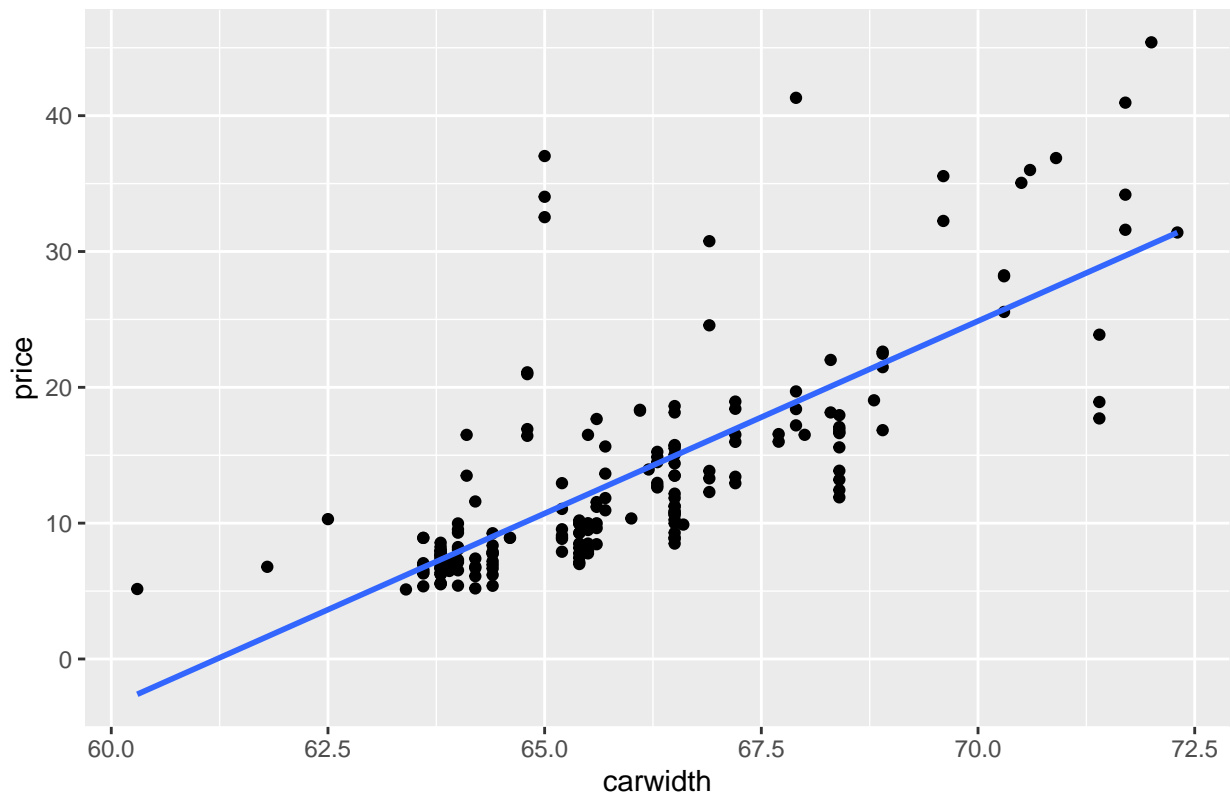
# Classificação das variáveis qualitativas
base$drivewheel = as.factor(base$drivewheel)

# Removendo os valores "4wd" da base, pois tem poucas observações
base = base[base$drivewheel != "4wd",]

# Gráfico do efeito da variável quantitativa
library(ggplot2)
ggplot(data = base, aes(x = carwidth, y = price)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +
  labs(title = "Gráfico do efeito da largura do carro",
       x = "carwidth",
       y = "price")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

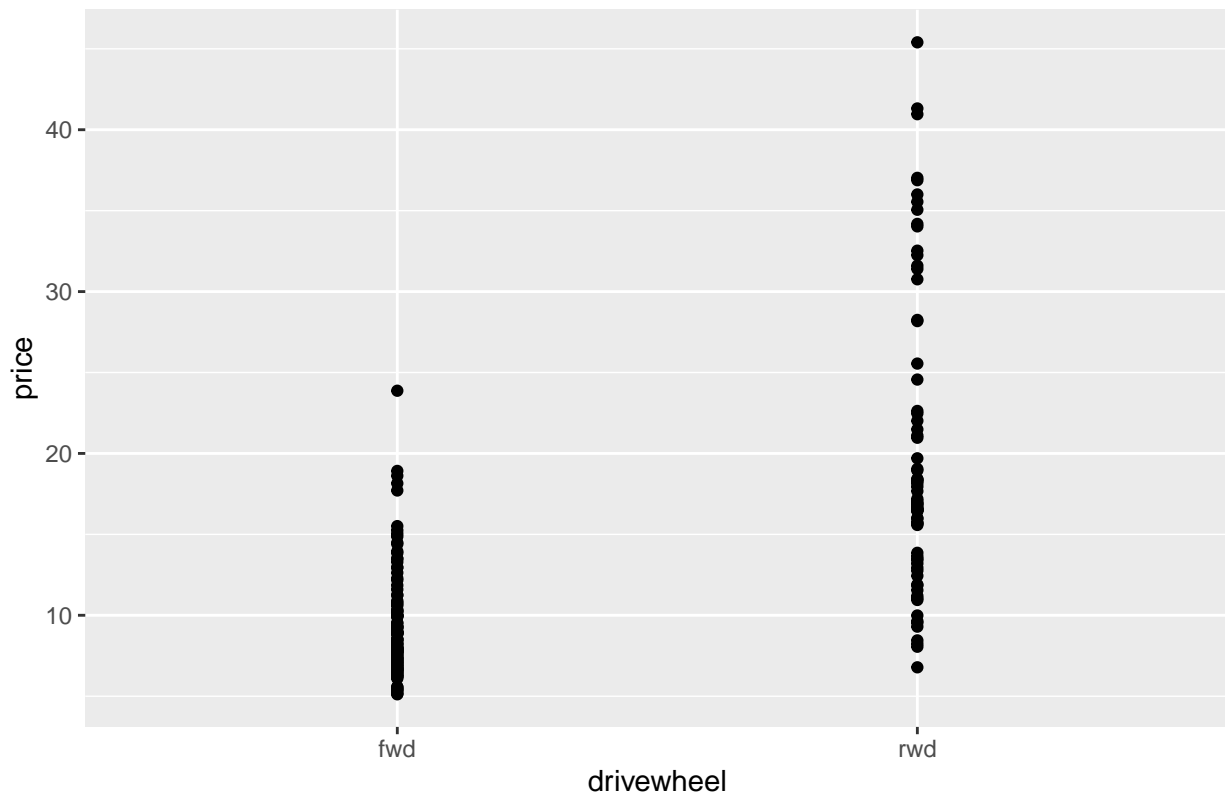
Grafico do efeito da largura do carro



```
# Gráfico do efeito da variável quantitativa
library(ggplot2)
ggplot(data = base, aes(x = drivewheel, y = price)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +
  labs(title = "Gráfico do efeito da tração",
       x = "drivewheel",
       y = "price")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

Grafico do efeito da tração



```
# Modelo
modelo = lm(price ~ carwidth * drivewheel, data = base)
summary(modelo)

##
## Call:
## lm(formula = price ~ carwidth * drivewheel, data = base)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.7120  -2.0507  -0.4600   0.9686  23.0225
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -102.0301    17.7772  -5.739 3.66e-08 ***
## carwidth         1.7103     0.2732   6.261 2.45e-09 ***
## drivewheelrwd   -48.4992    24.0548  -2.016  0.0452 *
## carwidth:drivewheelrwd  0.8210     0.3640   2.256  0.0252 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.698 on 192 degrees of freedom
## Multiple R-squared:  0.6701, Adjusted R-squared:  0.6649
## F-statistic: 130 on 3 and 192 DF, p-value: < 2.2e-16
```

```
# Com tração dianteira, para cada polegada de largura do carro é esperado que  
# o preço aumente em 1710 dólares.
```

```
# Com tração traseira, para cada polegada de largura do carro é esperado que  
# o preço aumente em 2530 dólares.
```

```
library(car)
```

```
## Carregando pacotes exigidos: carData
```

```
Anova(modelo)
```

```
## Anova Table (Type II tests)
```

```
##
```

```
## Response: price
```

```
##           Sum Sq Df F value    Pr(>F)  
## carwidth      3197.7  1 144.8559 < 2.2e-16 ***  
## drivewheel    1126.0  1  51.0080 1.854e-11 ***  
## carwidth:drivewheel 112.3  1   5.0875  0.02523 *  
## Residuals      4238.3 192
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

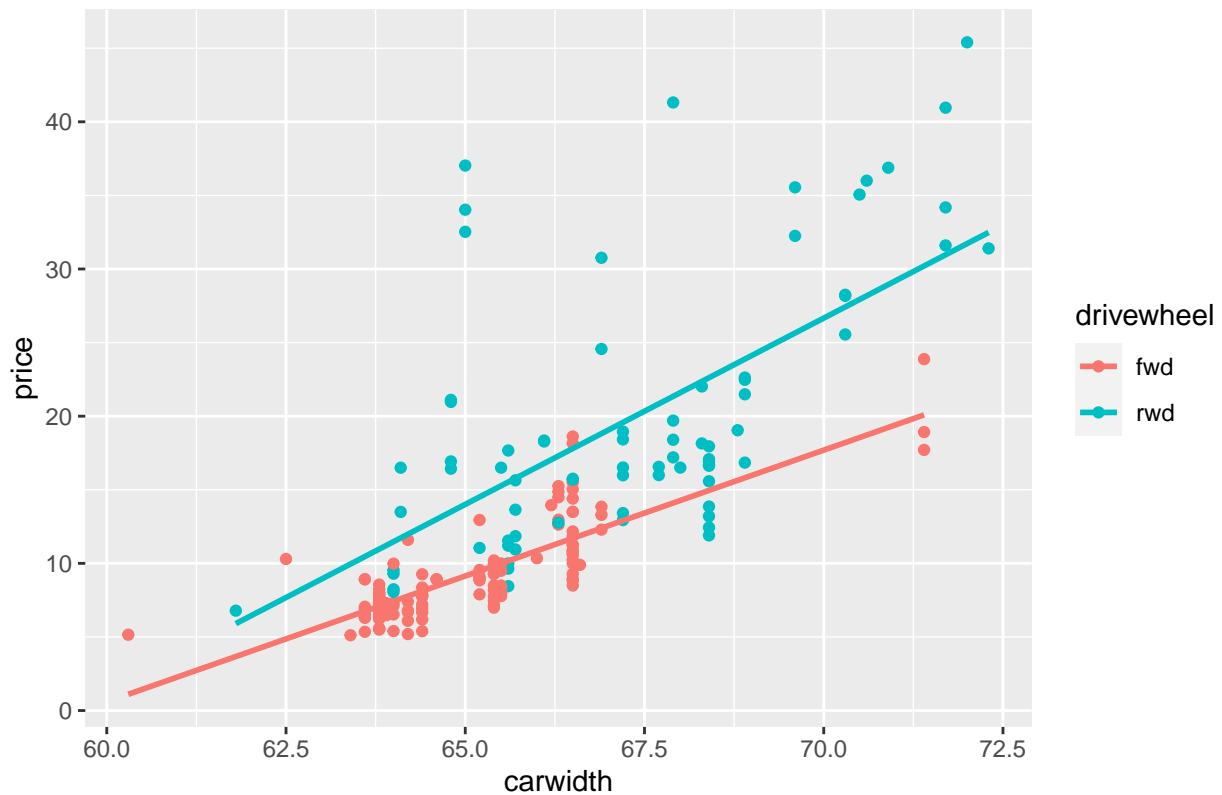
```
# Como o p-valor da interação entre largura e tração é menor que 5% ele é significativo  
# Como o coeficiente da interação entre os dois é positivo, significa que  
# quanto maior a largura maior será a diferença entre as variáveis qualitativas
```

```
# Gráfico da interação
```

```
ggplot(data = base, aes(x = carwidth, y = price, color = drivewheel)) +  
  geom_point() +  
  geom_smooth(method = "lm", se = FALSE) +  
  labs(title = "Gráfico da Interação",  
        x = "carwidth",  
        y = "price")
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```

Grafico da Interacao



```
# Predição para 70 polegadas e tração dianteira
novo = data.frame(carwidth = 70, drivewheel = "fwd")
predict(modelo, novo, interval = "confidence")
```

```
##          fit      lwr      upr
## 1 17.69258 14.89839 20.48677
```

A média do preço da população de carros com esses parametros deve estar entre 14890.39 e 20486.77 dol.

```
predict(modelo, novo, interval = "predict")
```

```
##          fit      lwr      upr
## 1 17.69258  8.013444 27.37172
```

O preço de um carros com esses parametros deve estar entre 8013.44 e 27371.72 dólares