analise.R

jenny

2023-10-05

library(data.table)  
  
base <- fread(input = paste0("car\_base.csv"), header = T, na.strings = "NA", data.table = FALSE, dec=".")  
base <- base[base$drivewheel != "4wd", ]  
  
#organizando  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:data.table':  
##   
## between, first, last

## The following objects are masked from 'package:stats':  
##   
## filter, lag

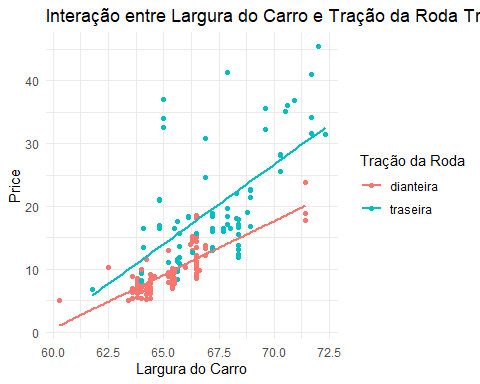
## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

base <- base %>%  
 rename(tração\_da\_roda = drivewheel,  
 largura\_do\_carro = carwidth,  
 preço = price)  
  
base <- base %>%  
 mutate(tração\_da\_roda = case\_when(  
 tração\_da\_roda == "rwd" ~ "traseira",  
 tração\_da\_roda == "fwd" ~ "dianteira",  
 TRUE ~ tração\_da\_roda  
 ))  
  
coeficiente\_correlacao <- cor(base$largura\_do\_carro, base$preço)  
  
  
# Ajuste do modelo de regressão linear com interação  
modelo <- lm(preço ~ largura\_do\_carro \* tração\_da\_roda, data = base)  
# Resumo do modelo  
summary(modelo)

##   
## Call:  
## lm(formula = preço ~ largura\_do\_carro \* tração\_da\_roda, 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  
## largura\_do\_carro 1.7103 0.2732 6.261 2.45e-09  
## tração\_da\_rodatraseira -48.4992 24.0548 -2.016 0.0452  
## largura\_do\_carro:tração\_da\_rodatraseira 0.8210 0.3640 2.256 0.0252  
##   
## (Intercept) \*\*\*  
## largura\_do\_carro \*\*\*  
## tração\_da\_rodatraseira \*   
## largura\_do\_carro:tração\_da\_rodatraseira \*   
## ---  
## 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

#grafico de interação no modelo  
library(ggplot2)  
ggplot(base, aes(x = largura\_do\_carro, y = preço, color = tração\_da\_roda)) +  
 geom\_point() +  
 geom\_smooth(method = "lm", se = FALSE) +  
 labs(title = "Interação entre Largura do Carro e Tração da Roda Traseira",  
 x = "Largura do Carro",  
 y = "Price",  
 color = "Tração da Roda") +  
 theme\_minimal()

## `geom\_smooth()` using formula = 'y ~ x'



# Intervalo de confiança e de predição para os dados da base  
tração\_da\_roda <- base$tração\_da\_roda  
largura\_do\_carro <- base$largura\_do\_carro  
preço <- base$preço  
  
IC1 <- predict(modelo, interval="confidence",level = 0.95)  
IC2 <- predict(modelo, interval="predict",level = 0.95)

## Warning in predict.lm(modelo, interval = "predict", level = 0.95): predictions on current data refer to \_future\_ responses

new <- data.frame(tração\_da\_roda,largura\_do\_carro,preço,IC1,IC2)  
  
  
# Predição para traça dianteira e tem largura de 70 polegadas   
novo = data.frame(tração\_da\_roda="dianteira", largura\_do\_carro=70)  
predict(modelo, novo, interval="confidence")

## fit lwr upr  
## 1 17.69258 14.89839 20.48677

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

## fit lwr upr  
## 1 17.69258 8.013444 27.37172