Exercício 2

Gustavo Beretta Gonçalves

#lendo a base de dados  
base <- read.csv("car\_base.csv", sep=";")  
  
#a) retirando as observações "drivewheel=4wd"  
base <- subset(base, drivewheel != "4wd")  
  
#b) retirando a amostra  
set.seed(09092003)  
base = base[sample(nrow(base), 120),]  
  
#c) ajuste do modelo  
modelo <- lm(price ~ carwidth \* drivewheel, data = base)  
  
#d) análise do modelo  
summary(modelo)

##   
## Call:  
## lm(formula = price ~ carwidth \* drivewheel, data = base)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -10.9051 -1.8043 -0.3891 1.0053 12.2902   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -106.3569 16.0691 -6.619 1.18e-09 \*\*\*  
## carwidth 1.7741 0.2462 7.205 6.31e-11 \*\*\*  
## drivewheelrwd -68.5288 23.6652 -2.896 0.00452 \*\*   
## carwidth:drivewheelrwd 1.1161 0.3561 3.134 0.00218 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.765 on 116 degrees of freedom  
## Multiple R-squared: 0.7776, Adjusted R-squared: 0.7718   
## F-statistic: 135.2 on 3 and 116 DF, p-value: < 2.2e-16

Neste caso, o p-valor para a interação é significativo (0.00218), o que sugere que a interação melhora o ajuste do modelo aos dados.

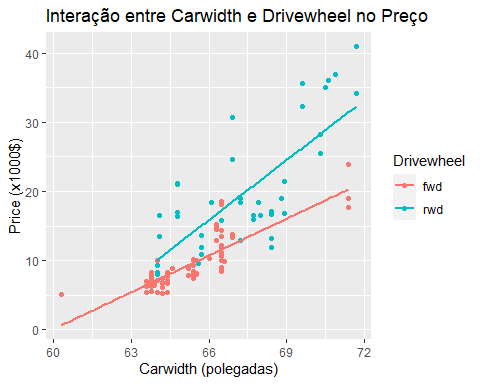
**Intercepto (-106.3569)**: Este é o valor esperado de price quando todas as outras variáveis são 0. No entanto, como drivewheel é uma variável categórica, essa interpretação não é muito útil.

**carwidth (1.7741)**: Este é o aumento esperado no price para cada aumento de uma unidade em carwidth, assumindo que drivewheel é igual a fwd (a categoria de referência).

**drivewheelrwd (-68.5288)**: Este é a diferença esperada no price entre um carro com drivewheel igual a rwd e um carro com drivewheel igual a fwd, assumindo que carwidth é 0. Novamente, essa interpretação não é muito útil porque carwidth nunca é 0.

**carwidth:drivewheelrwd (1.1161)**: Este é o efeito da interação. Ele representa a diferença na inclinação da relação entre price e carwidth para carros com drivewheel igual a rwd em comparação com carros com drivewheel igual a fwd. Em outras palavras, para carros com drivewheel igual a rwd, o aumento esperado no price para cada aumento de uma unidade em carwidth é (1.7741 + 1.1161) = 2.8902.

#e) gráfico  
library(ggplot2)  
ggplot(base, aes(x = carwidth, y = price, color = drivewheel)) +  
 geom\_point() +  
 stat\_smooth(method = "lm", se = FALSE) +  
 labs(title = "Interação entre Carwidth e Drivewheel no Preço",  
 x = "Carwidth (polegadas)",  
 y = "Price (x1000$)",  
 color = "Drivewheel")



#f) intervalos  
dados <- data.frame(carwidth = 70, drivewheel = "fwd")  
predict(modelo, dados, interval = "confidence")

## fit lwr upr  
## 1 17.83286 15.35951 20.30621

predict(modelo, dados, interval = "prediction")

## fit lwr upr  
## 1 17.83286 9.977202 25.68851

**Intervalo de Confiança (15.36, 20.31)**: Este intervalo fornece uma estimativa do intervalo dentro do qual você pode esperar que a média do preço dos carros esteja, com 95% de confiança.

**Intervalo de Predição (9.98, 25.69)**: Este intervalo fornece uma estimativa do intervalo dentro do qual você pode esperar que o preço de um único carro esteja, com 95% de confiança. Este é

mais amplo porque está tentando capturar a variabilidade inerente em previsões individuais.