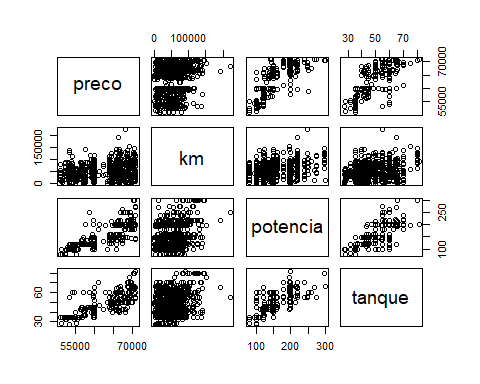
Atividade\_1.R

adans

2025-03-31

# Amostra  
set.seed(25032004) # COLOQUE O SEU DIA MÊS E ANO DE NASCIMENTO NA SET.SEED  
base = read.csv2("Atividade1.csv")  
base1 = base[sample(nrow(base), 800),]  
  
# Questão 2 - Gráfico de Dispersão  
  
base1 = base1[,-2]  
base1 = base1[,-5]  
plot(base1)



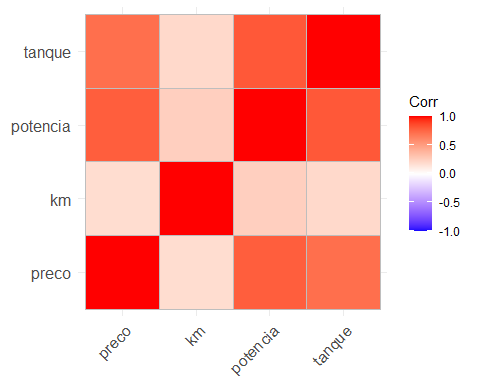
# Questão 3 - Coeficiente de Correlação  
  
# Coeficiente de correlacao  
cor(base1)

## preco km potencia tanque  
## preco 1.0000000 0.1784641 0.7901731 0.7218294  
## km 0.1784641 1.0000000 0.2510449 0.2037052  
## potencia 0.7901731 0.2510449 1.0000000 0.8106699  
## tanque 0.7218294 0.2037052 0.8106699 1.0000000

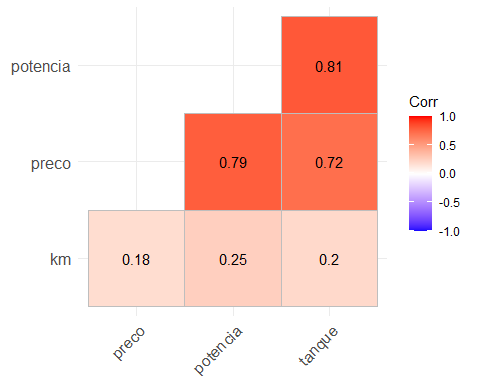
correlacao <- cor(base1)  
  
# Grafico da Matriz do correlacao  
library(ggcorrplot)

## Carregando pacotes exigidos: ggplot2

ggcorrplot(correlacao)



ggcorrplot(correlacao, hc.order = TRUE, type = "lower",  
 lab = TRUE)



library(corrplot)

## corrplot 0.95 loaded

corrplot(cor(base1), method = 'number')  
  
# Questão 4 - Intervalo de Confiança  
  
# Intervalo de Confianca  
cor.test(base1$preco, base1$potencia, conf.level=0.95)$conf.int

## [1] 0.7626283 0.8148574  
## attr(,"conf.level")  
## [1] 0.95

# Teste de Hipoteses  
cor.test(base1$preco, base1$potencia, alternative='two.sided')

##   
## Pearson's product-moment correlation  
##   
## data: base1$preco and base1$potencia  
## t = 36.42, df = 798, p-value < 2.2e-16  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.7626283 0.8148574  
## sample estimates:  
## cor   
## 0.7901731

# Questão 5 - Modelo de Regressão  
  
# Coeficientes do modelo  
modelo = lm(preco ~ potencia, data=base1)  
coef(modelo)

## (Intercept) potencia   
## 48135.04958 93.77769

# Coeficiente de determinacao  
summary(modelo)$r.squared

## [1] 0.6243736

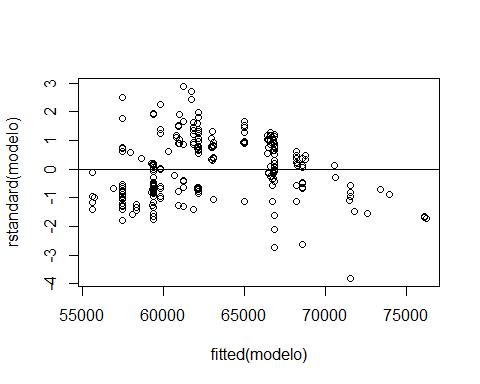
# Reta de regressao no modelo  
abline(modelo)



novo = data.frame("potencia"=700)  
predict(modelo,novo)

## 1   
## 113779.4

# Análise de Resíduos  
  
plot(fitted(modelo),rstandard(modelo))  
abline(0,0)



# Teste de Significância do Modelo  
  
summary(modelo)

##   
## Call:  
## lm(formula = preco ~ potencia, data = base1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12760.7 -2318.4 180.4 2618.6 9673.0   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 48135.050 409.353 117.59 <2e-16 \*\*\*  
## potencia 93.778 2.575 36.42 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3356 on 798 degrees of freedom  
## Multiple R-squared: 0.6244, Adjusted R-squared: 0.6239   
## F-statistic: 1326 on 1 and 798 DF, p-value: < 2.2e-16

# Questão 6 - Gráfico de Dispersão  
  
base1 = base[sample(nrow(base), 800),]  
base1 = base1[,-5]  
base1 = base1[,-3]  
base1 = base1[,-2]  
  
summary(modelo)

##   
## Call:  
## lm(formula = preco ~ potencia, data = base1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12760.7 -2318.4 180.4 2618.6 9673.0   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 48135.050 409.353 117.59 <2e-16 \*\*\*  
## potencia 93.778 2.575 36.42 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3356 on 798 degrees of freedom  
## Multiple R-squared: 0.6244, Adjusted R-squared: 0.6239   
## F-statistic: 1326 on 1 and 798 DF, p-value: < 2.2e-16

ggplot(base1, aes(x = preco, y = potencia, color = transm)) +  
 geom\_point(size = 4) + # Tamanho dos pontos  
 labs(title = "Preço X Potência: Destaca-se Transmissão",  
 x = "Preço",  
 y = "Potência",  
 color = "Transmissão",   
 shape = "Transmissão") +  
 theme\_minimal() + # Estilo minimalista  
 theme(legend.position = "top") # Posição da legenda

