Packages_functions

Karine Almeida

2023-10-22

Install and load packages

Functions that will be used throughout this notebook

```
metricas2 <- function(p_var, y_var){
  SQE <- sum((y_var - p_var)**2)

  # Cálculo do SSE (Sum of Squares Total)
  SST <- sum((y_var - mean(y_var))**2)

  # Cálculo do R-quadrado
  R_squared <- 1 - SQE/SST

  # RMSE
  RMSE <- sqrt(mean((y_var - p_var)^2))

  # Imprimindo os resultados
  cat("SQE: ", SQE, "QME : ", SQE/length(y_var), "\n")
  cat("SST: ", SST, "QMT: ", SST/length(y_var), "\n")
  cat("RMSE:", RMSE, "\n")
  cat("R-quadrado: ", R_squared, "\n")

}
```

Graphical analysis

Function to plot the predicted value, real value and residual

```
scatterplot_color <- function(data, x_var, y_var, r_var) {
  ggplot(data) +
    geom_point(aes(x = !!sym(x_var), y = !!sym(y_var), color = !!sym(r_var))) +
    theme(legend.position="bottom") +
    ggtitle("Scatterplot") +
    scale_color_viridis_c()
}
```

Shapiro-Francia normality adherence test

```r
sf.test2 <- function (x)
{
  DNAME <- deparse(substitute(x))
  x <- sort(x[complete.cases(x)])
  n <- length(x)
  if ((n < 5 || n > 500000))
    stop("sample size must be between 5 and 500000")
  y <- qnorm(ppoints(n, a = 3/8))
  W <- cor(x, y)^2
  u <- log(n)
  v <- log(u)
  mu <- -1.2725 + 1.0521 * (v - u)
  sig <- 1.0308 - 0.26758 * (v + 2/u)
  z <- (log(1 - W) - mu)/sig
  pval <- pnorm(z, lower.tail = FALSE)
  RVAL <- list(statistic = c(W = W), p.value = pval, method = "Shapiro-Francia normality test MBA ",
    data.name = DNAME)
  class(RVAL) <- "htest"
  return(RVAL)
}
```