LINEAR REGRESSION LEARNING

KERAMA JAMES THE DATA ANALYST

2023-04-10

#Learning on how to do the hypothesis testing in R programming using the mtcars dataset  
#Load the mtcars dataset  
data("mtcars")  
#creation of the two groups of the mpg with 5 gears and with 4 gears  
mpg\_4\_gears <- mtcars$mpg[mtcars$gear == 4]  
mpg\_5\_gears <- mtcars$mpg[mtcars$gear == 5]  
#Performing the t-test using the t-tes function  
t.test(mpg\_4\_gears, mpg\_5\_gears)

##   
## Welch Two Sample t-test  
##   
## data: mpg\_4\_gears and mpg\_5\_gears  
## t = 0.94271, df = 6.2123, p-value = 0.381  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -4.964154 11.270820  
## sample estimates:  
## mean of x mean of y   
## 24.53333 21.38000

#Doing of the regression modelling of the mtcars datasets  
data(mtcars)  
#Fixing of the linear regression model  
model <- lm(mpg ~ wt, data = mtcars)  
summary(model)

##   
## Call:  
## lm(formula = mpg ~ wt, data = mtcars)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4.5432 -2.3647 -0.1252 1.4096 6.8727   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 37.2851 1.8776 19.858 < 2e-16 \*\*\*  
## wt -5.3445 0.5591 -9.559 1.29e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.046 on 30 degrees of freedom  
## Multiple R-squared: 0.7528, Adjusted R-squared: 0.7446   
## F-statistic: 91.38 on 1 and 30 DF, p-value: 1.294e-10

#Prediction of the mpg values in the mtcars dataset  
newdata <- data.frame(wt = c(2.5, 3.0, 3.5))  
predictions <- predict(model, newdata = newdata)