hw6.R

Muhammad

2021-09-29

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --

## v ggplot2 3.3.5 v purrr 0.3.4  
## v tibble 3.1.4 v dplyr 1.0.7  
## v tidyr 1.1.3 v stringr 1.4.0  
## v readr 2.0.1 v forcats 0.5.1

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

setwd("C:/Users/Muhammad/ISYE/hw6")  
crime\_data<- read.table("uscrime.txt", header = TRUE)  
  
#Regression model Dependent variable is Crime and indendent variables are M,So,Ed,P01....Time.  
regression\_model<- lm(Crime~., crime\_data)  
summary(regression\_model)

##   
## Call:  
## lm(formula = Crime ~ ., data = crime\_data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -395.74 -98.09 -6.69 112.99 512.67   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -5.984e+03 1.628e+03 -3.675 0.000893 \*\*\*  
## M 8.783e+01 4.171e+01 2.106 0.043443 \*   
## So -3.803e+00 1.488e+02 -0.026 0.979765   
## Ed 1.883e+02 6.209e+01 3.033 0.004861 \*\*   
## Po1 1.928e+02 1.061e+02 1.817 0.078892 .   
## Po2 -1.094e+02 1.175e+02 -0.931 0.358830   
## LF -6.638e+02 1.470e+03 -0.452 0.654654   
## M.F 1.741e+01 2.035e+01 0.855 0.398995   
## Pop -7.330e-01 1.290e+00 -0.568 0.573845   
## NW 4.204e+00 6.481e+00 0.649 0.521279   
## U1 -5.827e+03 4.210e+03 -1.384 0.176238   
## U2 1.678e+02 8.234e+01 2.038 0.050161 .   
## Wealth 9.617e-02 1.037e-01 0.928 0.360754   
## Ineq 7.067e+01 2.272e+01 3.111 0.003983 \*\*   
## Prob -4.855e+03 2.272e+03 -2.137 0.040627 \*   
## Time -3.479e+00 7.165e+00 -0.486 0.630708   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 209.1 on 31 degrees of freedom  
## Multiple R-squared: 0.8031, Adjusted R-squared: 0.7078   
## F-statistic: 8.429 on 15 and 31 DF, p-value: 3.539e-07

M<- 14.0  
So<- 0  
Ed<-10.0  
Po1<-12.0  
Po2<- 15.5  
LF<-0.640  
M.F<-94.0  
Pop<-150  
NW<-1.1  
U1<-0.120  
U2<-3.6  
Wealth<-3200  
Ineq<-20.1  
Prob<-0.04  
Time<-39.0  
  
pred\_grid <- expand.grid(M = M, So = So, Ed=Ed, Po1=Po1, Po2=Po2,LF=LF, M.F = M.F, Pop = Pop, NW= NW, U1= U1, U2=U2, Wealth=Wealth, Ineq=Ineq,Prob=Prob, Time=Time)  
  
prediction= predict(regression\_model, newdata = pred\_grid)  
prediction

## 1   
## 155.4349

#I wouldn't rely on this model at all because data is too small and adjusted R is too small too. We could use less predictors even then 47 observations are just too low.