AUTO-LAB.R

saiup

2022-09-24

library(ggplot2)  
library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──  
## ✔ tibble 3.1.8 ✔ dplyr 1.0.10  
## ✔ tidyr 1.2.1 ✔ stringr 1.4.1   
## ✔ readr 2.1.2 ✔ forcats 0.5.2   
## ✔ purrr 0.3.4   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

library(ISLR2)  
  
?ISLR2::Auto

## starting httpd help server ... done

ourData<-Auto %>% select(mpg,horsepower,weight,acceleration)  
summary(ourData)

## mpg horsepower weight acceleration   
## Min. : 9.00 Min. : 46.0 Min. :1613 Min. : 8.00   
## 1st Qu.:17.00 1st Qu.: 75.0 1st Qu.:2225 1st Qu.:13.78   
## Median :22.75 Median : 93.5 Median :2804 Median :15.50   
## Mean :23.45 Mean :104.5 Mean :2978 Mean :15.54   
## 3rd Qu.:29.00 3rd Qu.:126.0 3rd Qu.:3615 3rd Qu.:17.02   
## Max. :46.60 Max. :230.0 Max. :5140 Max. :24.80

cor(ourData$mpg,ourData$horsepower)

## [1] -0.7784268

cor(ourData$mpg,ourData$weight)

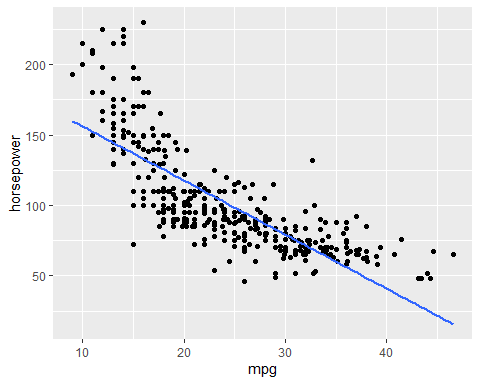
## [1] -0.8322442

cor(ourData$mpg,ourData$acceleration)

## [1] 0.4233285

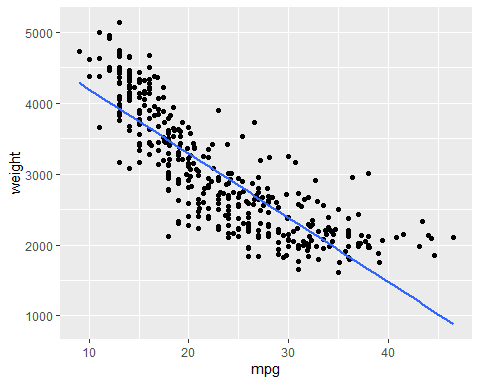
ourData %>% ggplot(aes(x=mpg, y=horsepower)) + geom\_point() + geom\_smooth(method="lm", se=FALSE)

## `geom\_smooth()` using formula 'y ~ x'



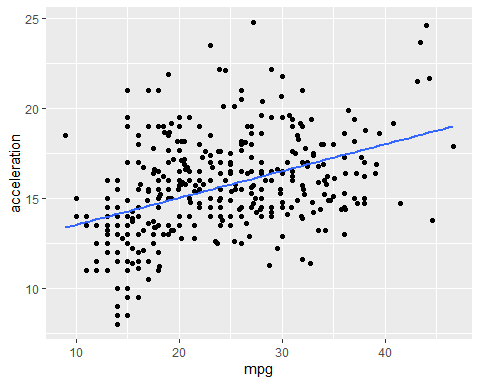
ourData %>% ggplot(aes(x=mpg, y=weight)) + geom\_point() + geom\_smooth(method="lm", se=FALSE)

## `geom\_smooth()` using formula 'y ~ x'



ourData %>% ggplot(aes(x=mpg, y=acceleration)) + geom\_point() + geom\_smooth(method="lm", se=FALSE)

## `geom\_smooth()` using formula 'y ~ x'



#Regression  
  
mpg\_regression <- lm(mpg ~ horsepower, data = ourData)  
summary(mpg\_regression)

##   
## Call:  
## lm(formula = mpg ~ horsepower, data = ourData)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -13.5710 -3.2592 -0.3435 2.7630 16.9240   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 39.935861 0.717499 55.66 <2e-16 \*\*\*  
## horsepower -0.157845 0.006446 -24.49 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.906 on 390 degrees of freedom  
## Multiple R-squared: 0.6059, Adjusted R-squared: 0.6049   
## F-statistic: 599.7 on 1 and 390 DF, p-value: < 2.2e-16

weight\_regression <- lm(mpg ~ weight, data = ourData)  
summary(weight\_regression)

##   
## Call:  
## lm(formula = mpg ~ weight, data = ourData)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.9736 -2.7556 -0.3358 2.1379 16.5194   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 46.216524 0.798673 57.87 <2e-16 \*\*\*  
## weight -0.007647 0.000258 -29.64 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.333 on 390 degrees of freedom  
## Multiple R-squared: 0.6926, Adjusted R-squared: 0.6918   
## F-statistic: 878.8 on 1 and 390 DF, p-value: < 2.2e-16

coef(mpg\_regression)

## (Intercept) horsepower   
## 39.9358610 -0.1578447

coef(weight\_regression)

## (Intercept) weight   
## 46.216524549 -0.007647343

#Prediction  
  
mpg\_prediction <- data.frame(horsepower=c(190))  
predict(mpg\_regression, mpg\_prediction)

## 1   
## 9.945362

mpg\_prediction2 <- data.frame(weight=c(4000))  
predict(weight\_regression, mpg\_prediction2)

## 1   
## 15.62715