Homework 7

The mtcars dataset is a classic and widely used dataset in R that contains specifications and performance data for 32 different car models from the 1974 Motor Trend magazine. Each row represents a unique vehicle, and each column records a specific attribute related to engine performance, design, or efficiency. Some key variables include mpg (miles per gallon), hp (gross horsepower), wt (weight in 1000 lbs), drat (rear axle ratio), and qsec (quarter-mile time). Additionally, the dataset includes categorical variables encoded as numeric values, such as cyl (number of cylinders), am (transmission type), and gear (number of forward gears).

This dataset is frequently used in regression modeling and statistical learning due to its compact size, real-world relevance, and mixture of quantitative and categorical variables. Analysts often model fuel efficiency (mpg) as a function of other variables to understand how engine power, vehicle weight, or gear ratios impact gas mileage. With its balance of complexity and interpretability, mtcars serves as a great playground for developing skills in exploratory data analysis, model selection, variable interpretation, and diagnostics in both teaching and applied settings.

library(dplyr) ## Data Transformation  
library(tidyr) ## Data Transformation  
library(stringr) ## Data Transformation  
library(ggplot2) ## Data Visualization

mod\_1\_mtcars <- mtcars %>%   
 mutate(vs = case\_when(  
 vs == 1 ~ "straight",  
 TRUE ~ "v-shaped"  
 ))

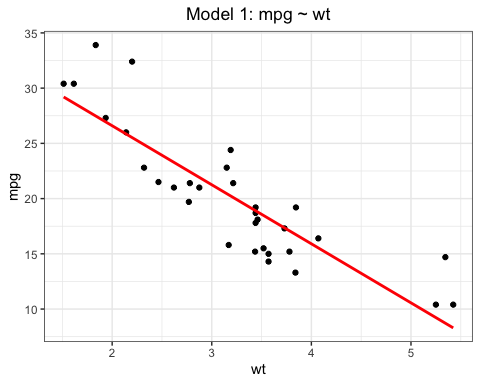
## Model 1: mpg ~ wt

model\_1 <- lm(mpg ~ wt, data = mod\_1\_mtcars)  
summary(model\_1)

Call:  
lm(formula = mpg ~ wt, data = mod\_1\_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

# Plot Model 1  
ggplot(mod\_1\_mtcars, aes(x = wt, y = mpg)) +  
 geom\_point() +  
 geom\_smooth(method = "lm", se = FALSE, color = "red") +  
 ggtitle("Model 1: mpg ~ wt") +   
 theme\_bw() +  
 theme(plot.title = element\_text(hjust = 0.5))

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



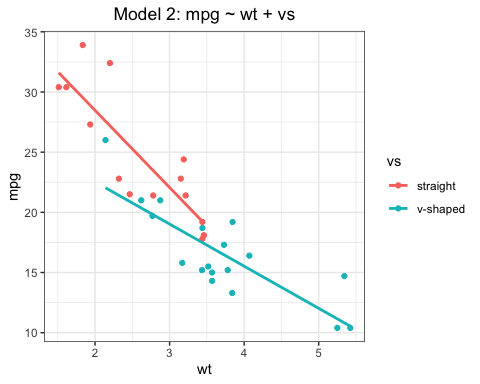
## Model 2: mpg ~ wt + vs

model\_2 <- lm(mpg ~ wt + vs, data = mod\_1\_mtcars)  
summary(model\_2)

Call:  
lm(formula = mpg ~ wt + vs, data = mod\_1\_mtcars)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-3.7071 -2.4415 -0.3129 1.4319 6.0156   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 36.1586 1.7656 20.480 < 2e-16 \*\*\*  
wt -4.4428 0.6134 -7.243 5.63e-08 \*\*\*  
vsv-shaped -3.1544 1.1907 -2.649 0.0129 \*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 2.78 on 29 degrees of freedom  
Multiple R-squared: 0.801, Adjusted R-squared: 0.7873   
F-statistic: 58.36 on 2 and 29 DF, p-value: 6.818e-11

ggplot(mod\_1\_mtcars, aes(x = wt, y = mpg, color = vs)) +  
 geom\_point() +  
 geom\_smooth(method = "lm", se = FALSE) +  
 ggtitle("Model 2: mpg ~ wt + vs") +   
 theme\_bw() +  
 theme(plot.title = element\_text(hjust = 0.5))

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



## Model 3: mpg ~ wt \* vs

model\_3 <- lm(mpg ~ wt \* vs, data = mod\_1\_mtcars)  
summary(model\_3)

Call:  
lm(formula = mpg ~ wt \* vs, data = mod\_1\_mtcars)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-3.9950 -1.7881 -0.3423 1.2935 5.2061   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 41.2981 2.7002 15.294 4.01e-15 \*\*\*  
wt -6.4110 0.9998 -6.412 6.08e-07 \*\*\*  
vsv-shaped -11.7667 3.7638 -3.126 0.0041 \*\*   
wt:vsv-shaped 2.9097 1.2157 2.393 0.0236 \*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 2.578 on 28 degrees of freedom  
Multiple R-squared: 0.8348, Adjusted R-squared: 0.8171   
F-statistic: 47.16 on 3 and 28 DF, p-value: 4.497e-11

# Plot Model 3  
ggplot(mod\_1\_mtcars, aes(x = wt, y = mpg, color = vs)) +  
 geom\_point() +  
 geom\_smooth(method = "lm", se = FALSE, formula = y ~ x \* vs) +  
 ggtitle("Model 3: mpg ~ wt \* vs") +   
 theme\_bw() +  
 theme(plot.title = element\_text(hjust = 0.5))

Warning: Failed to fit group 1.  
Caused by error:  
! object 'vs' not found

Warning: Failed to fit group 2.  
Caused by error:  
! object 'vs' not found

