In [3]:
```
install.packages("car")
library("car")
```

```
also installing the dependency 'lme4'

  There are binary versions available but the source versions are later:
     binary source needs_compilation
lme4 1.1-26 1.1-31              TRUE
car  3.0-10  3.1-1             FALSE

  Binaries will be installed
package 'lme4' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Lenovo\AppData\Local\Temp\RtmpQJh8yQ\downloaded_packages
installing the source package 'car'

Loading required package: carData
```

In [69]:
```
data(Prestige)
help(Prestige)
```

(a) Create a new variable professional by recoding the variable type so that professionals are coded as 1, and blue and white collar workers are coded as 0 (Hint: ifelse).

In [70]:
```
# borramos los datos NAN
Prestige <- na.omit(Prestige)
```

In [72]:
```
# creamos una lista vacia
professional <- c()
# creamos un ciclo for para añadir elementos a la lista
# 1 profesionales
# 0 caso contrario
# creamos un ciclo para rellenar la lista
for (i in 1:length(Prestige$type))
{
    # agregamos los valores respectivos a la variable
    if (Prestige$type[i] == "prof")
    {
    professional <- c(professional,1)
    }
    else
    {
    professional <- c(professional,0)
    }
}
```

(b) Run a linear model with prestige as an outcome and income, professional, and the interaction of the two as predictors (Note: this is a continuous × dummy interaction.)

In [73]:
```
# creamos el modelo de regresion lineal multiple
model <- lm(Prestige$prestige ~ Prestige$income + professional)
```

(c) Write the prediction equation based on the result.

In [74]:
```
# imprimimos la ecuacion
```

```
print(paste("Equation: prestige = ",coef(model)[2],"*income + ",coef(model)[3],"*pro
```

```
[1] "Equation: prestige =  0.00137062547941901 *income +  22.7569998569787 *professi
onal +  30.6183338104581"
```

(d) Interpret the coefficient for income.

The income coefficient is almost 0, so we can conclude that their contribution affects the prestige variable very little or not at all. In other words, income does not affect a person's prestige much.

(e) Interpret the coefficient for professional.

For its part, the professional coefficient is almost 22, which is much higher than the previous one, with which we can conclude that a person's profession does affect their prestige.

(f) What is the effect of a 1,000 increase in income on prestige score for professional occupations? In other words, we are interested in the marginal effect of income when the variable professional takes the value of 1. Calculate the change in ˆy associated with a 1,000 increase in income based on your answer for (c).

In [80]:

```
# creamos una nueva lista para aumenta 1000 al ingreso de los profesionales
pro_100 <- c()
for (i in 1:length(Prestige$type))
{
    # en caso de ser profesional aumentamos 1000
    if (Prestige$type[i] == "prof")
    {
        pro_100 <- c(pro_100,Prestige$income[i] + 1000)
    }
    else
    {
        pro_100 <- c(pro_100,Prestige$income[i])
    }

}
```

In [87]:

```
# volvemos a realizar el modelo con la nueva lista
model2 <- lm(Prestige$prestige ~ pro_100 + professional)
# resultados
print(paste("Equation: prestige = ",coef(model2)[2],"*income + ",coef(model2)[3],"*p
print(paste("the value of the professional coefficient decreases in:",coef(model)[3]
```

```
[1] "Equation: prestige =  0.00137062547941901 *income +  21.3863743775597 *professi
onal +  30.6183338104581"
[1] "the value of the professional coefficient decreases in: 1.37062547941902"
```

(g) What is the effect of changing one’s occupations from non-professional to professional when her income is $6,000? We are interested in the marginal effect of professional jobs when the variable income takes the value of 6, 000. Calculate the change in ˆy based on your answer for (c).

In [90]:

```
# volvemos a calcular la variable profesional con los datos dados
professional2 <- c()

for (i in 1:length(Prestige$type))
{
    # agregamos los valores respectivos a la variable
    if (Prestige$type[i] == "prof" ||  Prestige$income[i]>= 6000)
```

```
    {
    professional2 <- c(professional2,1)
    }
    else
    {
    professional2 <- c(professional2,0)
    }
}
```

In [92]:

```
# creamos la ecuacion
model3 <- lm(Prestige$prestige ~ Prestige$income + professional2)
print(paste("Equation: prestige = ",coef(model3)[2],"*income + ",coef(model3)[3],"*p
```

```
[1] "Equation: prestige =  0.00173189528658058 *income +  15.1838673340949 *professi
onal +  26.943556264508"
```

In [95]:

```
print(paste("change income :",coef(model3)[2]-coef(model)[2]))
print(paste("change professional :",coef(model3)[3]-coef(model)[3]))
print(paste("change b :",coef(model3)[1]-coef(model)[1]))
```

```
[1] "change income : 0.00036126980716157"
[1] "change professional : -7.57313252288382"
[1] "change b : -3.6747775459501"
```

Question 2: Political Science

Researchers are interested in learning the effect of all of those yard signs on voting preferences.1 Working with a campaign in Fairfax County, Virginia, 131 precincts were randomly divided into a treatment and control group. In 30 precincts, signs were posted around the precinct that read, “For Sale: Terry McAuliffe. Don’t Sellout Virgina on November 5.”

Below is the result of a regression with two variables and a constant. The dependent variable is the proportion of the vote that went to McAuliff’s opponent Ken Cuccinelli. The first variable indicates whether a precinct was randomly assigned to have the sign against McAuliffe posted. The second variable indicates a precinct that was adjacent to a precinct in the treatment group (since people in those precincts might be exposed to the signs).

(a) Use the results from a linear regression to determine whether having these yard signs in a precinct affects vote share (e.g., conduct a hypothesis test with α = .05).

In [16]:

```
# guardamos los valores respectivos de los coeficientes
B1 =  0.042
seB1  = 0.016
# having these yard signs in a precinct affects vote share?

# hypothesis

# H0   B1 = 0   no afecta el procentaje de votos
# H1   B1 > 0   si afecta el porcentaje de votos

# Test statistic
t = (B1-0)/seB1

# p-value
n = 131
df = n-3
p = 2*pt(q = t , df = df, lower.tail = F)
```

```
alfa = 0.05

if (p < alfa)
{
    print(paste("B1 = 0 was not rejected!, which is why have these yard signs in a p
}else
{
    print(paste("B1 = 0 rejected!, which is why have these yard signs in a precinct
}
```

```
[1] "B1 = 0 was not rejected!, which is why have these yard signs in a precinct does
not affect the percentage of votes"
```

(b) Use the results to determine whether being next to precincts with these yard signs affects vote share (e.g., conduct a hypothesis test with α = .05)

In [17]:

```
# guardamos los valores respectivos de los coeficientes
B2 =  0.042
seB2  = 0.013
# having these yard signs in a precinct affects vote share?

# hypothesis

# H0   B1 = 0   no afecta el procentaje de votos
# H1   B1 > 0   si afecta el porcentaje de votos

# Test statistic
t2 = (B2-0)/seB2

# p-value
n = 131
df = n-3
p2 = 2*pt(q = t2 , df = df, lower.tail = F)

alfa = 0.05

if (p2 < alfa)
{
    print(paste("B2 = 0 was not rejected!, does not affect the percentage of votes"))
}else
{
    print(paste("B2 = 0 rejected!, if it affects the percentage of votes"))
}
```

```
[1] "B2 = 0 was not rejected!, does not affect the percentage of votes"
```

(c) Interpret the coefficient for the constant term substantively.

Given the results shown, we have that when there are no signs assigned to the enclosure or to the sides of the enclosure, the voting preference of the people is 0.302.

(d) Evaluate the model fit for this regression. What does this tell us about the importance of yard signs versus other factors that are not modeled?

The R2 coefficient is very small, which means that our model has very low precision. In other words, yard signs do not have much influence with respect to other variables that have not been considered in the study.

In [ ]: