

12. Regresión Lineal - Análisis de los errores

Ricardo Salinas

2024-08-30

1. Obtén la matriz de correlación de los datos que se te proporcionan. Interpreta.

```
M = read.csv("Downloads/Estatura-peso_HyM.csv")
MM = subset(M,M$Sexo=="M")
MH = subset(M,M$Sexo=="H")
M1=data.frame(MH$Estatura,MH$Peso,MM$Estatura,MM$Peso)
cor(M1)

##              MH.Estatura    MH.Peso  MM.Estatura    MM.Peso
## MH.Estatura 1.0000000000 0.846834792 0.0005540612 0.04724872
## MH.Peso      0.8468347920 1.0000000000 0.0035132246 0.02154907
## MM.Estatura 0.0005540612 0.003513225 1.0000000000 0.52449621
## MM.Peso      0.0472487231 0.021549075 0.5244962115 1.000000000

print('Se puede interpretar que los valores entre mujeres y hombres no tienen
mucha correlacion, lo cual tambien nos muestra inconsistencias en los datos
ya que valores como altura de hombre y de mujer no deberian de tener relacion
alguna, tambien podemos ver que el peso tiene una gran correlacion con la
estatura.')
```

```
## [1] "Se puede interpretar que los valores entre mujeres y hombres no
tienen mucha correlacion, lo cual tambien nos muestra inconsistencias en los
datos ya que valores como altura de hombre y de mujer no deberian de tener
relacion alguna, tambien podemos ver que el peso tiene una gran correlacion
con la estatura."
```

2. Obtén medidas (media, desviación estándar, etc) que te ayuden a analizar los datos.

```
n=4 #número de variables
d=matrix(NA,ncol=7,nrow=n)
for(i in 1:n){
  d[i,]<-c(as.numeric(summary(M1[,i])),sd(M1[,i]))
}
m=as.data.frame(d)

row.names(m)=c("H-Estatura","H-Peso","M-Estatura","M-Peso")
names(m)=c("Minimo","Q1","Mediana","Media","Q3","Máximo","Desv Est")
m

##              Minimo      Q1 Mediana      Media      Q3 Máximo      Desv Est
## H-Estatura    1.48 1.6100 1.650 1.653727 1.7000 1.80 0.06173088
## H-Peso        56.43 68.2575 72.975 72.857682 77.5225 90.49 6.90035408
## M-Estatura    1.44 1.5400 1.570 1.572955 1.6100 1.74 0.05036758
## M-Peso        37.39 49.3550 54.485 55.083409 59.7950 80.87 7.79278074
```

Hipotesis:

$$H_0: \beta_1 = 0 \quad H_1: \beta_1 \neq 0$$

3. Encuentra la ecuación de regresión de mejor ajuste:

#Realiza la regresión entre Las variables involucradas

```
Modelo1H = lm(Peso~Estatura, MH)
Modelo1H

##
## Call:
## lm(formula = Peso ~ Estatura, data = MH)
##
## Coefficients:
## (Intercept)      Estatura
##      -83.68         94.66

Modelo1M = lm(Peso~Estatura, MM)
print(Modelo1H)

##
## Call:
## lm(formula = Peso ~ Estatura, data = MH)
##
## Coefficients:
## (Intercept)      Estatura
##      -83.68         94.66

print(Modelo1M)

##
## Call:
## lm(formula = Peso ~ Estatura, data = MM)
##
## Coefficients:
## (Intercept)      Estatura
##      -72.56         81.15

Modelo2 = lm(Estatura~Peso+Sexo, M)
print(Modelo2)

##
## Call:
## lm(formula = Estatura ~ Peso + Sexo, data = M)
##
## Coefficients:
## (Intercept)      Peso      SexoM
##      1.27271      0.00523      0.01218

summary(Modelo1M)

##
## Call:
```

```
## lm(formula = Peso ~ Estatura, data = MM)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.3256  -4.1942   0.4004   4.2724  17.9114
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -72.560      14.041  -5.168 5.34e-07 ***
## Estatura      81.149       8.922   9.096 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.65 on 218 degrees of freedom
## Multiple R-squared:  0.2751, Adjusted R-squared:  0.2718
## F-statistic: 82.73 on 1 and 218 DF,  p-value: < 2.2e-16
```

summary(Modelo1H)

```
##
## Call:
## lm(formula = Peso ~ Estatura, data = MH)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.3881  -2.6073  -0.0665   2.4421  11.1883
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -83.685       6.663  -12.56 <2e-16 ***
## Estatura      94.660       4.027   23.51 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.678 on 218 degrees of freedom
## Multiple R-squared:  0.7171, Adjusted R-squared:  0.7158
## F-statistic: 552.7 on 1 and 218 DF,  p-value: < 2.2e-16
```

summary(Modelo2)

```
##
## Call:
## lm(formula = Estatura ~ Peso + Sexo, data = M)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.118876 -0.026908 -0.000819  0.027854  0.155874
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.2727097  0.0196768  64.681 <2e-16 ***
```

```
## Peso          0.0052296  0.0002674  19.560   <2e-16 ***
## SexoM         0.0121799  0.0061647   1.976   0.0488 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04118 on 437 degrees of freedom
## Multiple R-squared:  0.6483, Adjusted R-squared:  0.6467
## F-statistic: 402.8 on 2 and 437 DF,  p-value: < 2.2e-16

#A 0.05 si es significativo y Los modelos quedarían:

b0 = Modelo2$coefficients[1]
b1 = Modelo2$coefficients[2]
b2 = Modelo2$coefficients[3]

Ym = function(x){b0+b1+b2*x}
Yh = function(x){b0+b1*x}

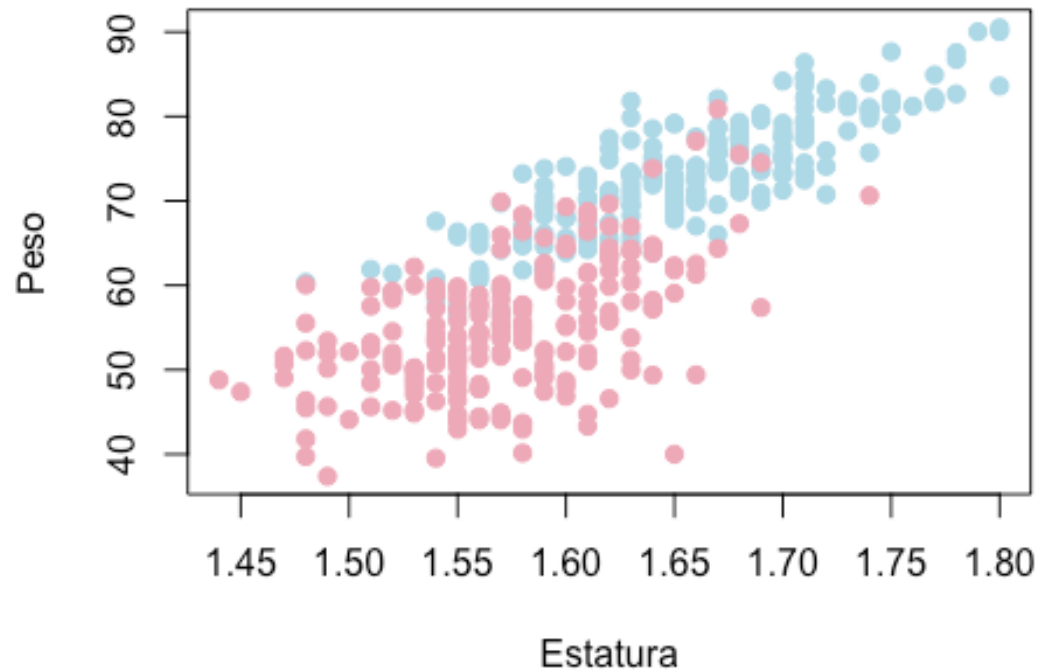
colores= c("lightblue", "pink2")
plot(M$Estatura, M$Peso, data=M, col=colores[factor(M$Sexo)], pch=19,
ylab="Peso", xlab="Estatura", main= "Relacion Peso vs Estatura")

## Warning in plot.window(...): "data" is not a graphical parameter
## Warning in plot.xy(xy, type, ...): "data" is not a graphical parameter
## Warning in axis(side = side, at = at, labels = labels, ...): "data" is not
a
## graphical parameter
## Warning in axis(side = side, at = at, labels = labels, ...): "data" is not
a
## graphical parameter

## Warning in box(...): "data" is not a graphical parameter
## Warning in title(...): "data" is not a graphical parameter

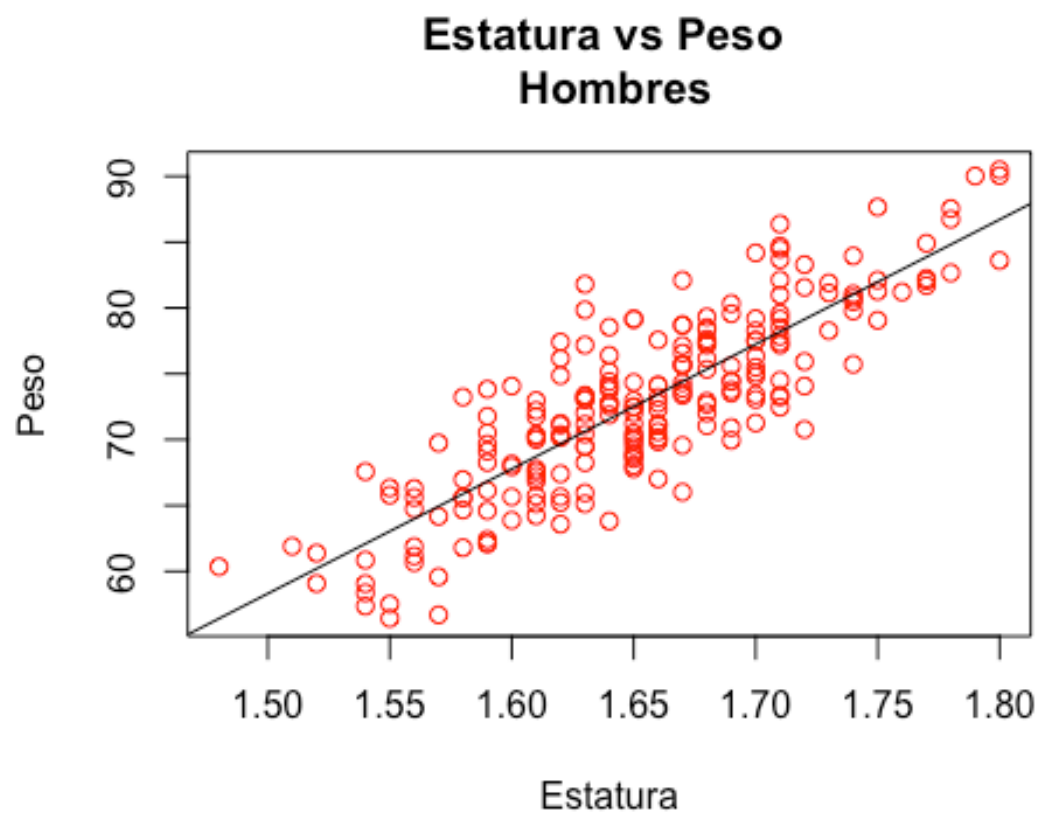
x = seq(1.40, 1.80, 0.01)
lines(x, Ym(x), col="pink2")
lines(x, Yh(x), col="lightblue")
```

Relacion Peso vs Estatura

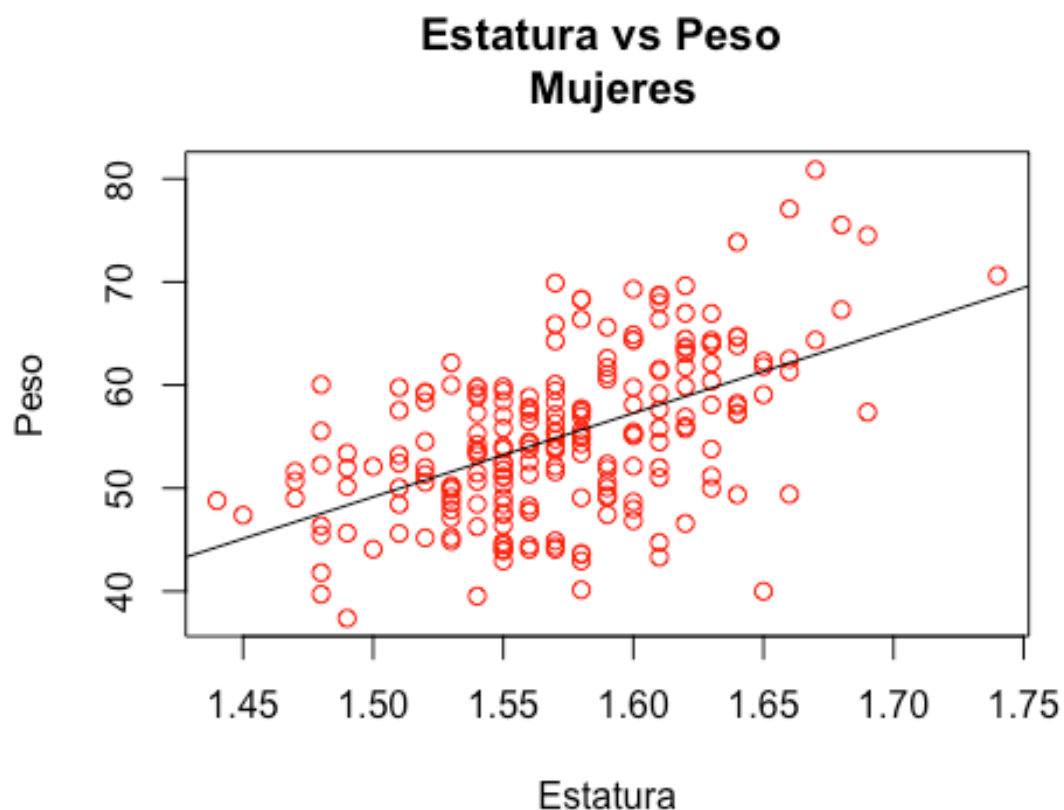


4. Dibuja el diagrama de dispersión de los datos y la recta de mejor ajuste.

```
plot(MH$Estatura,MH$Peso, col="red", main = "Estatura vs Peso \n Hombres",  
ylab = "Peso", xlab= "Estatura")  
abline(Modelo1H)
```



```
plot(MM$Estatura,MM$Peso, col="red", main = "Estatura vs Peso \n Mujeres",  
ylab = "Peso", xlab= "Estatura")  
abline(Modelo1M)
```



5. Interpreta en el contexto del problema cada uno de los análisis que hiciste.

```
print("Las graficas nos muestran relaciones entre el incremento del peso y la
estatura en ambos sexos, lo cual indica que si se tiene una clara correlacion
entre peso y estatura.")
```

```
## [1] "Las graficas nos muestran relaciones entre el incremento del peso y
la estatura en ambos sexos, lo cual indica que si se tiene una clara
correlacion entre peso y estatura."
```

6. Interpreta en el contexto del problema:

```
print("Se puede concluir que si existe una relacion entre peso y estatura, ya
que ambos tienen la tendencia de crecer mutuamente, esto se puede notar en
ambos sexos, tambien se puede reconocer que los hombres llegan a tener
valores mas altos que las mujeres.")
```

```
## [1] "Se puede concluir que si existe una relacion entre peso y estatura,
ya que ambos tienen la tendencia de crecer mutuamente, esto se puede notar en
ambos sexos, tambien se puede reconocer que los hombres llegan a tener
valores mas altos que las mujeres."
```

Nuevo Modelo

Propón un nuevo modelo. Esta vez toma en cuenta la interacción de la Estatura con el Sexo y realiza los mismos pasos que hiciste con los modelos anteriores:

```
M = read.csv("Downloads/Estatura-peso_HyM.csv")
MM1 = subset(M,M$Sexo=="M")
MH1 = subset(M,M$Sexo=="H")
M1=data.frame(MH$Estatura,MH$Sexo,MM$Estatura,MM$Sexo)
print(M1)
```

##	MH.Estatura	MH.Sexo	MM.Estatura	MM.Sexo
## 1	1.61	H	1.53	M
## 2	1.61	H	1.60	M
## 3	1.70	H	1.54	M
## 4	1.65	H	1.58	M
## 5	1.72	H	1.61	M
## 6	1.63	H	1.57	M
## 7	1.76	H	1.61	M
## 8	1.67	H	1.52	M
## 9	1.67	H	1.62	M
## 10	1.65	H	1.63	M
## 11	1.63	H	1.55	M
## 12	1.70	H	1.60	M
## 13	1.69	H	1.51	M
## 14	1.59	H	1.59	M
## 15	1.71	H	1.53	M
## 16	1.66	H	1.67	M
## 17	1.65	H	1.56	M
## 18	1.59	H	1.65	M
## 19	1.59	H	1.52	M
## 20	1.67	H	1.61	M
## 21	1.71	H	1.65	M
## 22	1.68	H	1.61	M
## 23	1.59	H	1.57	M
## 24	1.70	H	1.63	M
## 25	1.68	H	1.69	M
## 26	1.61	H	1.54	M
## 27	1.70	H	1.59	M
## 28	1.70	H	1.53	M
## 29	1.63	H	1.54	M
## 30	1.72	H	1.57	M
## 31	1.62	H	1.55	M
## 32	1.69	H	1.52	M
## 33	1.58	H	1.61	M
## 34	1.68	H	1.56	M
## 35	1.62	H	1.58	M
## 36	1.65	H	1.61	M
## 37	1.58	H	1.56	M
## 38	1.68	H	1.59	M
## 39	1.64	H	1.55	M
## 40	1.74	H	1.56	M

## 41	1.63	H	1.57	M
## 42	1.60	H	1.62	M
## 43	1.61	H	1.63	M
## 44	1.65	H	1.57	M
## 45	1.69	H	1.54	M
## 46	1.67	H	1.57	M
## 47	1.61	H	1.64	M
## 48	1.56	H	1.56	M
## 49	1.63	H	1.55	M
## 50	1.59	H	1.57	M
## 51	1.58	H	1.48	M
## 52	1.66	H	1.62	M
## 53	1.55	H	1.53	M
## 54	1.71	H	1.56	M
## 55	1.65	H	1.57	M
## 56	1.65	H	1.64	M
## 57	1.71	H	1.55	M
## 58	1.80	H	1.55	M
## 59	1.59	H	1.66	M
## 60	1.56	H	1.53	M
## 61	1.56	H	1.68	M
## 62	1.71	H	1.45	M
## 63	1.77	H	1.61	M
## 64	1.72	H	1.61	M
## 65	1.68	H	1.62	M
## 66	1.64	H	1.55	M
## 67	1.71	H	1.54	M
## 68	1.71	H	1.58	M
## 69	1.55	H	1.48	M
## 70	1.71	H	1.74	M
## 71	1.63	H	1.62	M
## 72	1.66	H	1.60	M
## 73	1.63	H	1.64	M
## 74	1.74	H	1.62	M
## 75	1.54	H	1.61	M
## 76	1.66	H	1.47	M
## 77	1.68	H	1.63	M
## 78	1.71	H	1.60	M
## 79	1.70	H	1.52	M
## 80	1.59	H	1.53	M
## 81	1.56	H	1.65	M
## 82	1.64	H	1.57	M
## 83	1.75	H	1.53	M
## 84	1.77	H	1.57	M
## 85	1.70	H	1.52	M
## 86	1.60	H	1.64	M
## 87	1.54	H	1.63	M
## 88	1.68	H	1.54	M
## 89	1.59	H	1.54	M
## 90	1.65	H	1.58	M

## 91	1.70	H	1.47	M
## 92	1.63	H	1.59	M
## 93	1.71	H	1.58	M
## 94	1.71	H	1.63	M
## 95	1.61	H	1.62	M
## 96	1.69	H	1.60	M
## 97	1.72	H	1.55	M
## 98	1.62	H	1.60	M
## 99	1.73	H	1.59	M
## 100	1.64	H	1.49	M
## 101	1.67	H	1.58	M
## 102	1.71	H	1.54	M
## 103	1.65	H	1.59	M
## 104	1.67	H	1.56	M
## 105	1.62	H	1.52	M
## 106	1.64	H	1.56	M
## 107	1.74	H	1.66	M
## 108	1.59	H	1.51	M
## 109	1.77	H	1.48	M
## 110	1.57	H	1.61	M
## 111	1.65	H	1.63	M
## 112	1.62	H	1.49	M
## 113	1.68	H	1.50	M
## 114	1.66	H	1.62	M
## 115	1.61	H	1.55	M
## 116	1.68	H	1.51	M
## 117	1.70	H	1.58	M
## 118	1.71	H	1.55	M
## 119	1.60	H	1.55	M
## 120	1.69	H	1.57	M
## 121	1.65	H	1.58	M
## 122	1.62	H	1.51	M
## 123	1.61	H	1.55	M
## 124	1.66	H	1.47	M
## 125	1.71	H	1.54	M
## 126	1.62	H	1.58	M
## 127	1.62	H	1.49	M
## 128	1.68	H	1.61	M
## 129	1.56	H	1.56	M
## 130	1.55	H	1.52	M
## 131	1.62	H	1.64	M
## 132	1.67	H	1.64	M
## 133	1.58	H	1.56	M
## 134	1.70	H	1.57	M
## 135	1.55	H	1.58	M
## 136	1.57	H	1.53	M
## 137	1.63	H	1.56	M
## 138	1.66	H	1.55	M
## 139	1.75	H	1.59	M
## 140	1.73	H	1.44	M

## 141	1.52	H	1.53	M
## 142	1.78	H	1.60	M
## 143	1.71	H	1.62	M
## 144	1.74	H	1.58	M
## 145	1.70	H	1.61	M
## 146	1.78	H	1.53	M
## 147	1.64	H	1.55	M
## 148	1.69	H	1.55	M
## 149	1.69	H	1.56	M
## 150	1.64	H	1.58	M
## 151	1.67	H	1.55	M
## 152	1.54	H	1.49	M
## 153	1.67	H	1.64	M
## 154	1.57	H	1.60	M
## 155	1.62	H	1.60	M
## 156	1.59	H	1.59	M
## 157	1.69	H	1.48	M
## 158	1.68	H	1.63	M
## 159	1.65	H	1.54	M
## 160	1.70	H	1.55	M
## 161	1.71	H	1.62	M
## 162	1.65	H	1.65	M
## 163	1.62	H	1.60	M
## 164	1.63	H	1.48	M
## 165	1.64	H	1.62	M
## 166	1.63	H	1.51	M
## 167	1.66	H	1.48	M
## 168	1.62	H	1.54	M
## 169	1.68	H	1.49	M
## 170	1.68	H	1.58	M
## 171	1.66	H	1.57	M
## 172	1.61	H	1.57	M
## 173	1.48	H	1.58	M
## 174	1.65	H	1.58	M
## 175	1.77	H	1.62	M
## 176	1.66	H	1.61	M
## 177	1.60	H	1.63	M
## 178	1.67	H	1.54	M
## 179	1.61	H	1.59	M
## 180	1.66	H	1.50	M
## 181	1.60	H	1.59	M
## 182	1.67	H	1.48	M
## 183	1.74	H	1.55	M
## 184	1.67	H	1.56	M
## 185	1.65	H	1.56	M
## 186	1.54	H	1.54	M
## 187	1.63	H	1.52	M
## 188	1.63	H	1.60	M
## 189	1.65	H	1.56	M
## 190	1.61	H	1.51	M

## 191	1.64	H	1.55	M
## 192	1.63	H	1.60	M
## 193	1.67	H	1.55	M
## 194	1.73	H	1.54	M
## 195	1.80	H	1.63	M
## 196	1.80	H	1.55	M
## 197	1.74	H	1.55	M
## 198	1.61	H	1.60	M
## 199	1.67	H	1.53	M
## 200	1.51	H	1.66	M
## 201	1.57	H	1.57	M
## 202	1.63	H	1.66	M
## 203	1.66	H	1.68	M
## 204	1.72	H	1.51	M
## 205	1.69	H	1.64	M
## 206	1.58	H	1.54	M
## 207	1.52	H	1.55	M
## 208	1.78	H	1.57	M
## 209	1.75	H	1.59	M
## 210	1.56	H	1.58	M
## 211	1.64	H	1.69	M
## 212	1.66	H	1.57	M
## 213	1.61	H	1.59	M
## 214	1.59	H	1.57	M
## 215	1.79	H	1.64	M
## 216	1.54	H	1.58	M
## 217	1.75	H	1.57	M
## 218	1.64	H	1.56	M
## 219	1.58	H	1.61	M
## 220	1.65	H	1.67	M

2. Significancia del modelo:

Hipotesis:

$$H_0: \beta_1 = 0 \quad H_1: \beta_1 \neq 0$$

```
Modelo2_ = lm(Peso~Estatura*Sexo, M)
print(Modelo2_)
```

```
##
## Call:
## lm(formula = Peso ~ Estatura * Sexo, data = M)
##
## Coefficients:
##      (Intercept)      Estatura      SexoM Estatura:SexoM
##          -83.68           94.66          11.12          -13.51
```

#Valida la significancia del modelo con un alfa de 0.03 (incluye las hipótesis que pruebas)

#Valida la significancia de β_1 con un alfa de 0.03 (incluye las hipótesis que

```
pruebas)
#Indica cuál es el porcentaje de variación explicada por el modelo.
summary(Modelo2_)
```

```
##
## Call:
## lm(formula = Peso ~ Estatura * Sexo, data = M)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.3256  -3.1107   0.0204   3.2691  17.9114
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -83.685      9.735  -8.597  <2e-16 ***
## Estatura       94.660      5.882  16.092  <2e-16 ***
## SexoM          11.124     14.950   0.744   0.457
## Estatura:SexoM -13.511      9.305  -1.452   0.147
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.374 on 436 degrees of freedom
## Multiple R-squared:  0.7847, Adjusted R-squared:  0.7832
## F-statistic: 529.7 on 3 and 436 DF,  p-value: < 2.2e-16
```

Formula: $-83.68 + 94.66(\text{Estatura}) + 11.12(\text{SexoM}) - 13.51(\text{Estatura:SexoM})$

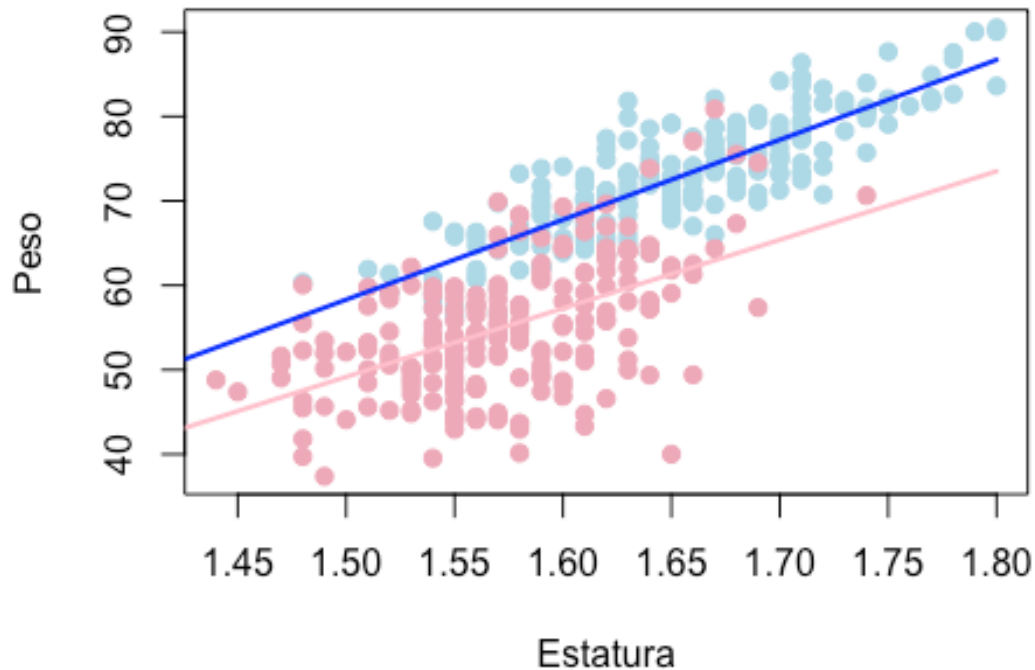
3. Dibuja el diagrama de dispersión de los datos y la recta de mejor ajuste.

```
b0 = Modelo2_$coefficients[1]
b1 = Modelo2_$coefficients[2]
b2 = Modelo2_$coefficients[3]
b3 = Modelo2_$coefficients[4]

Ym = function(x){b0+b1*x+b2+b3*x}
Yh = function(x){b0+b1*x}

colores= c("lightblue", "pink2")
plot(M$Estatura, M$Peso, col=colores[factor(M$Sexo)], pch=19, ylab="Peso",
xlab="Estatura", main= "Relacion Peso vs Estatura")
x = seq(1.40, 1.80, 0.01)
lines(x, Ym(x), col="pink", lwd=2)
lines(x, Yh(x), col="blue", lwd=2)
```

Relacion Peso vs Estatura



Interpreta en el contexto del problema cada uno de los análisis que hiciste.

```
print("Con este modelo podemos visualizar la relacion entre el sexo y la
estatura, demostrando que los hombres llegan a tener una tendencia a crecer
mucho mas que las mujeres en general, siendo que las mujeres tienen una media
mucho mas baja y una recta que indica lo mismo.")
```

```
## [1] "Con este modelo podemos visualizar la relacion entre el sexo y la
estatura, demostrando que los hombres llegan a tener una tendencia a crecer
mucho mas que las mujeres en general, siendo que las mujeres tienen una media
mucho mas baja y una recta que indica lo mismo."
```

Interpreta en el contexto del problema:

```
print("¿Qué información proporciona  $\beta\hat{0}$  sobre la relación entre la estatura y
el peso de hombres y mujeres? Interpreta y compara entre este modelo con los
3 modelos anteriores.")
```

```
## [1] "¿Qué información proporciona  $\beta\hat{0}$  sobre la relación entre la estatura y
el peso de hombres y mujeres? Interpreta y compara entre este modelo con los
3 modelos anteriores."
```

```
print("En el primer modelo donde se tiene la division entre hombres y mujeres
se tiene el peso promedio con base en la estatura, la diferencia que podmeos
```

ver con los otros modelos es que ademas de que se pueden graficar dentro del mismo scatterplot ya que no se separan por sexo, tambien podemos ver la relacion de la estatura por el sexo en vez de por el peso.")

```
## [1] "En el primer modelo donde se tiene la division entre hombres y mujeres se tiene el peso promedio con base en la estatura, la diferencia que podemos ver con los otros modelos es que ademas de que se pueden graficar dentro del mismo scatterplot ya que no se separan por sexo, tambien podemos ver la relacion de la estatura por el sexo en vez de por el peso."
```

```
print("¿Cómo interpretas  $\beta_1$  en la relación entre la estatura y el peso de hombres y mujeres? Interpreta y compara entre este modelo con los 3 modelos anteriores.")
```

```
## [1] "¿Cómo interpretas  $\beta_1$  en la relación entre la estatura y el peso de hombres y mujeres? Interpreta y compara entre este modelo con los 3 modelos anteriores."
```

```
print("Esta variable representa como aumenta el peso dependiendo de la estatura, este nos da una comparacion diferente de los otros modelos, ya que otros nos da la estatura dependiendo del sexo.")
```

```
## [1] "Esta variable representa como aumenta el peso dependiendo de la estatura, este nos da una comparacion diferente de los otros modelos, ya que otros nos da la estatura dependiendo del sexo."
```

```
print("Indica cuál(es) de los modelos probados para la relación entre peso y estatura entre hombres y mujeres consideras que es más apropiado y explica por qué.")
```

```
## [1] "Indica cuál(es) de los modelos probados para la relación entre peso y estatura entre hombres y mujeres consideras que es más apropiado y explica por qué."
```

```
print("El ultimo modelo es el mejor ya que se tiene la division entre sexos, lo cual nos deja visualizar las diferentes pendientes, las cuales representan la relacion de estatura y peso, lo cual nos da una representacion mucho mas clara de la correlacion entre las variables.")
```

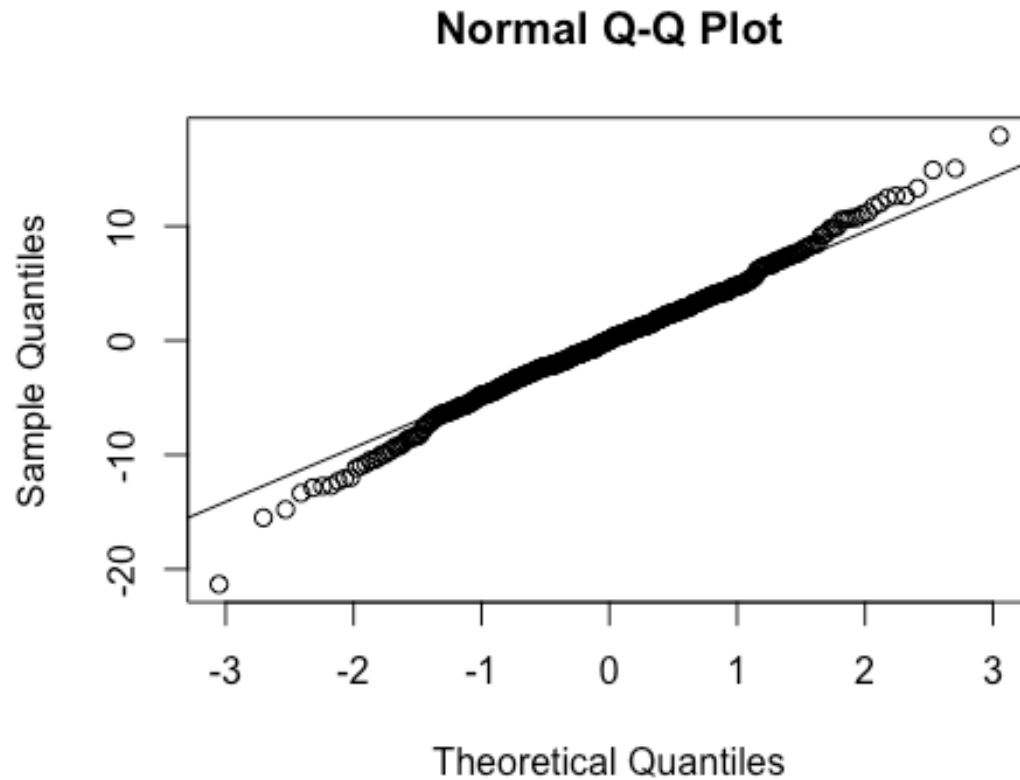
```
## [1] "El ultimo modelo es el mejor ya que se tiene la division entre sexos, lo cual nos deja visualizar las diferentes pendientes, las cuales representan la relacion de estatura y peso, lo cual nos da una representacion mucho mas clara de la correlacion entre las variables."
```

#Análisis de Residuos

H0: Los datos provienen de una población normal H1: Los datos no provienen de una población normal

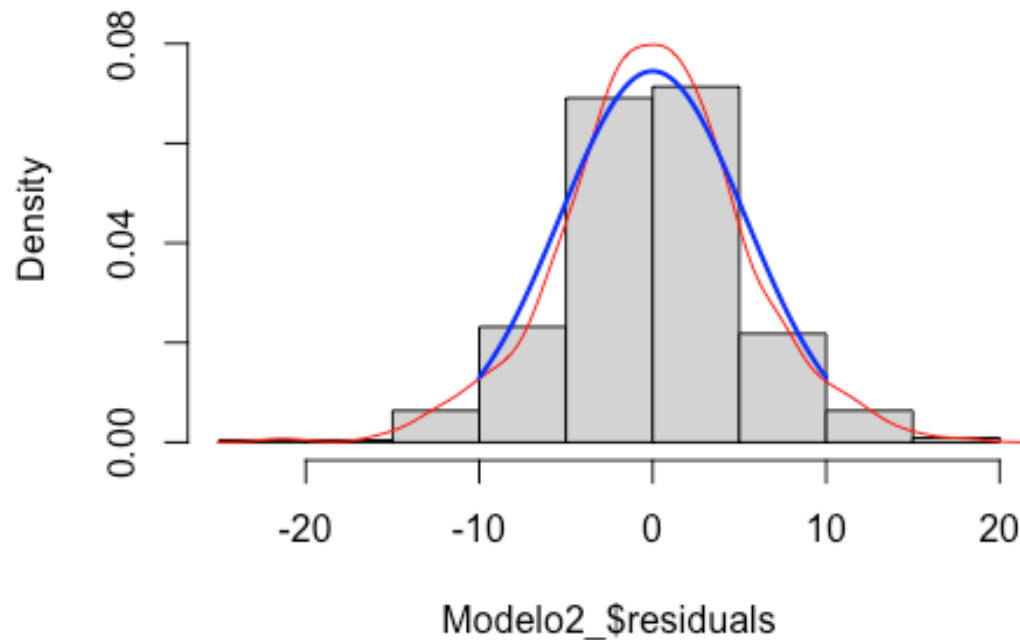
```
library(nortest)
ad.test(Modelo2$residuals)
```

```
##  
## Anderson-Darling normality test  
##  
## data: Modelo2$residuals  
## A = 0.8138, p-value = 0.03516  
  
qqnorm(Modelo2$residuals)  
qqline(Modelo2$residuals)
```



```
hist(Modelo2$residuals,freq=FALSE, ylim=c(0,0.09))  
lines(density(Modelo2$residuals),col="red")  
curve(dnorm(x,mean=mean(Modelo2$residuals),sd=sd(Modelo2$residuals)),  
from=-10, to=10, add=TRUE, col="blue",lwd=2)
```


Histogram of Modelo2_\$residuals



```
#Verificación de media cero
t.test(Modelo2_$residuals)

##
##  One Sample t-test
##
## data:  Modelo2_$residuals
## t = -3.1626e-16, df = 439, p-value = 1
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  -0.5017741  0.5017741
## sample estimates:
##      mean of x
## -8.074349e-17

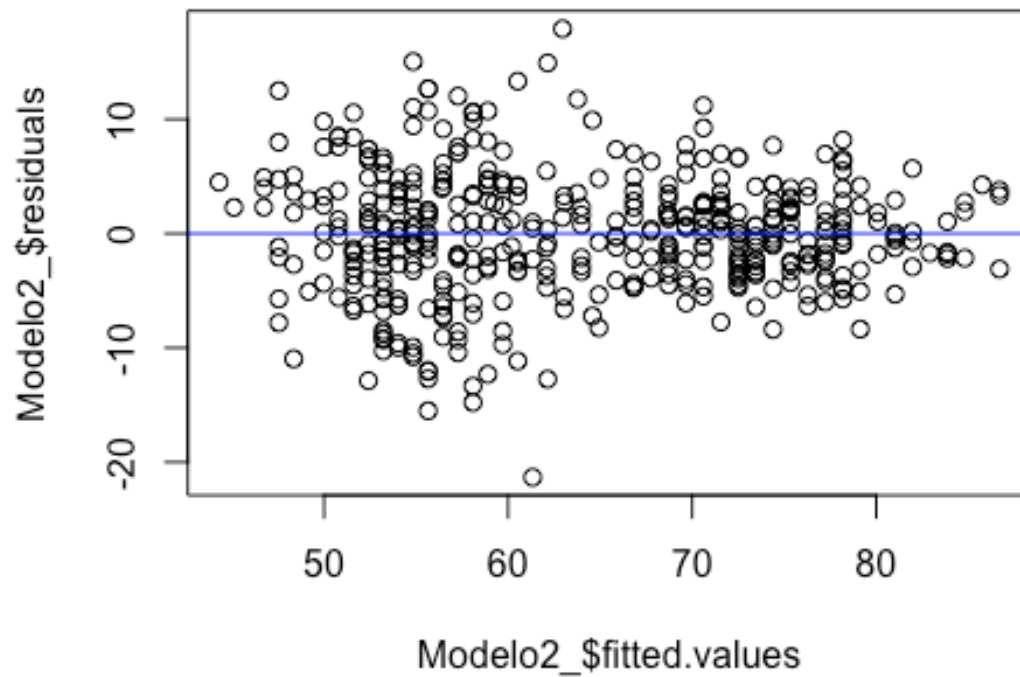
#Grafica de residuos
plot(Modelo2_$fitted.values,Modelo2_$residuals)
abline(h=0, col='blue')

#Independencia
library(lmtest)

## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##   as.Date, as.Date.numeric
```



```
dwtest(Modelo2_)

##
## Durbin-Watson test
##
## data:  Modelo2_
## DW = 1.8646, p-value = 0.07113
## alternative hypothesis: true autocorrelation is greater than 0

bgtest(Modelo2_)

##
## Breusch-Godfrey test for serial correlation of order up to 1
##
## data:  Modelo2_
## LM test = 1.3453, df = 1, p-value = 0.2461
```

```
#Homcedasticidad
```

```
library(lmtest)
```

```
bptest(Modelo2_)
```

```
##
```

```
## studentized Breusch-Pagan test
```

```
##
```

```
## data: Modelo2_
```

```
## BP = 59.211, df = 3, p-value = 8.667e-13
```

```
gqtest(Modelo2_)
```

```
##
```

```
## Goldfeld-Quandt test
```

```
##
```

```
## data: Modelo2_
```

```
## GQ = 3.2684, df1 = 216, df2 = 216, p-value < 2.2e-16
```

```
## alternative hypothesis: variance increases from segment 1 to 2
```

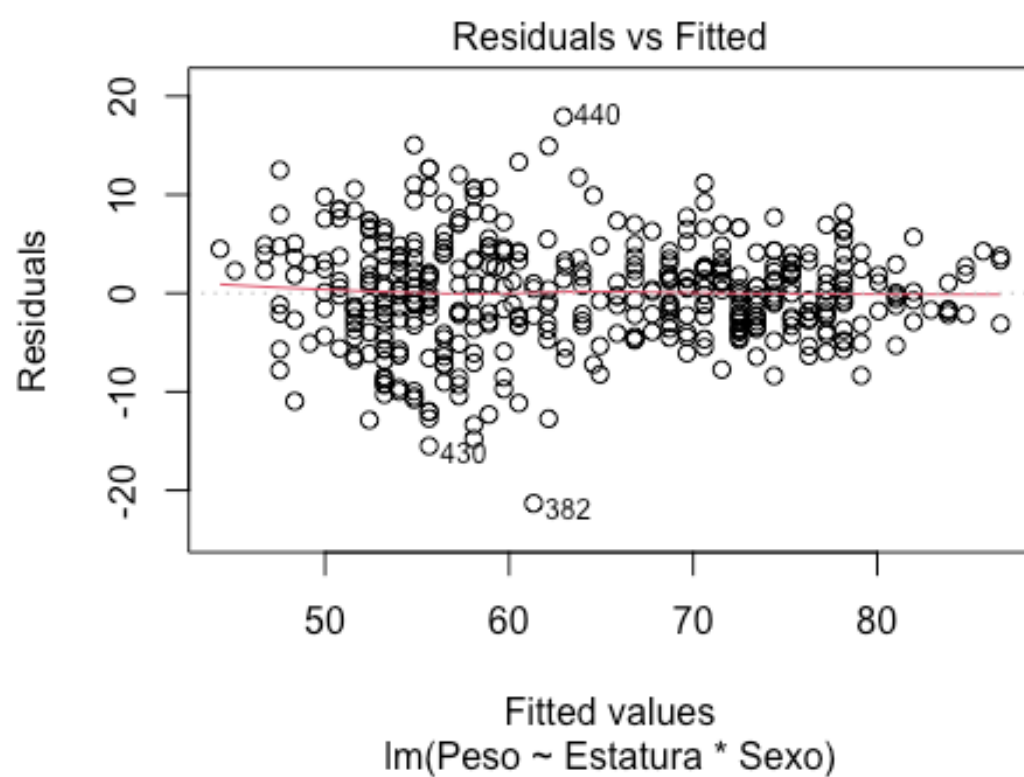
3. Interpreta en el contexto del problema cada uno de los análisis que hiciste.

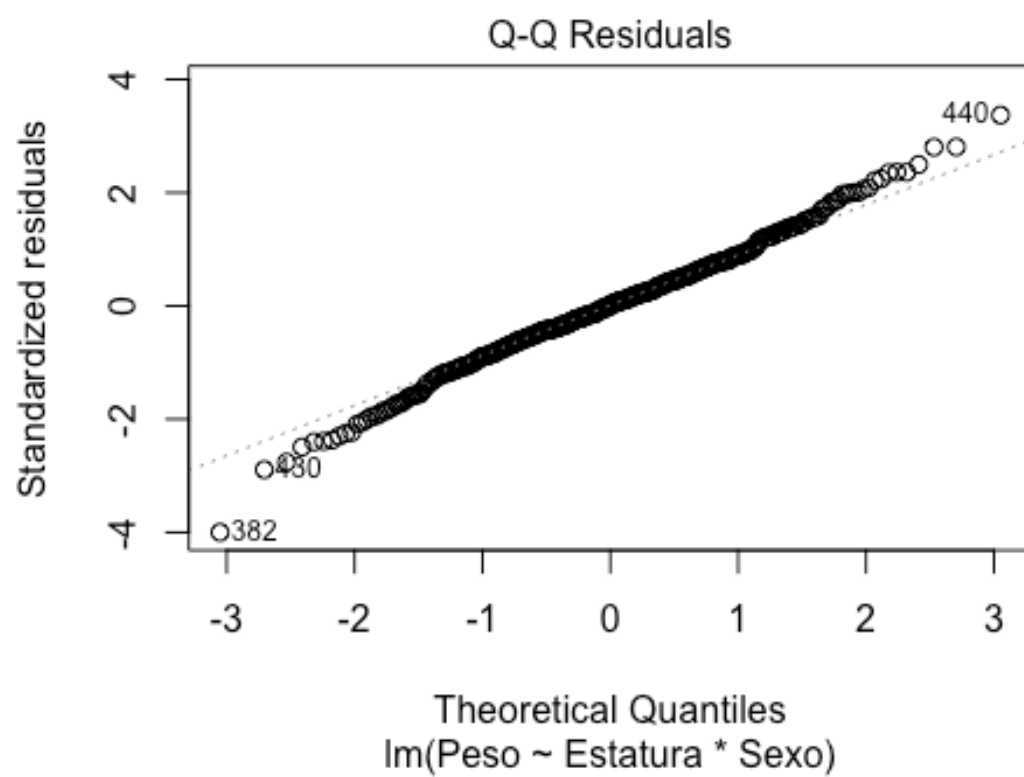
```
print("Con los analisis implementados podemos saber homocedasticidad da un  
valor muy bajo, lo cual nos indica que en algunas partes de los datos se  
tiene diferentes varianzas, tambien se llega a rechazar la hipotesis debido a  
que este nos da un valor muy bajo. Las pruebas para la independencia nos  
indica que no hay correlacion entre loos errores presentes.")
```

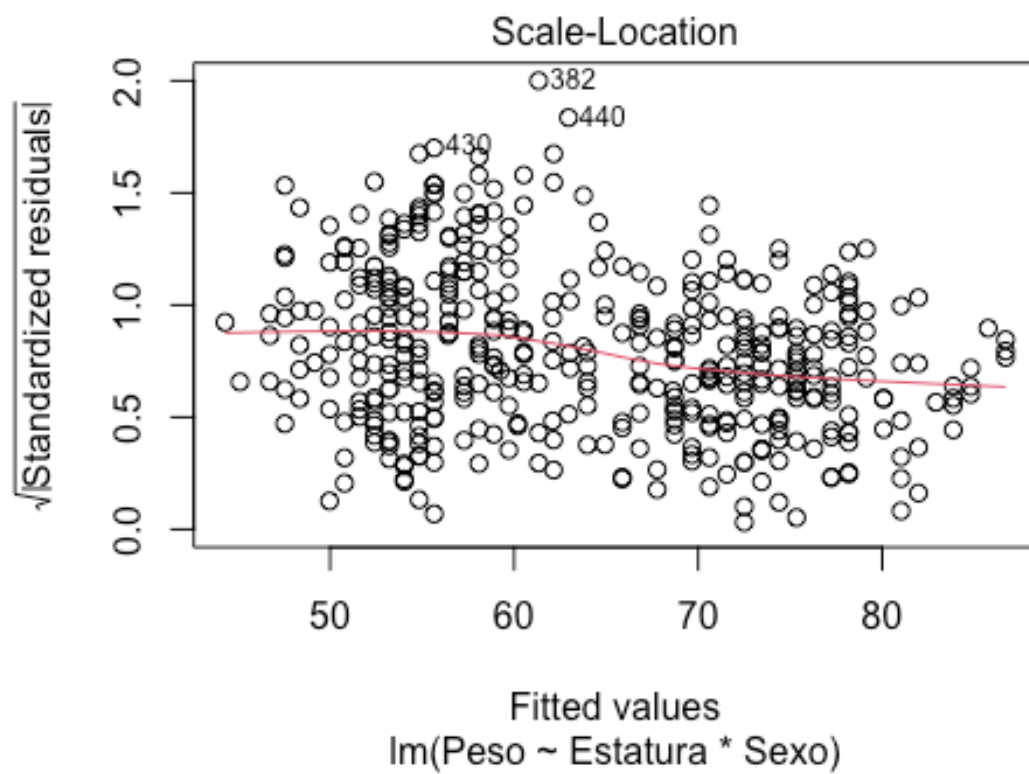
```
## [1] "Con los analisis implementados podemos saber homocedasticidad da un  
valor muy bajo, lo cual nos indica que en algunas partes de los datos se  
tiene diferentes varianzas, tambien se llega a rechazar la hipotesis debido a  
que este nos da un valor muy bajo. Las pruebas para la independencia nos  
indica que no hay correlacion entre loos errores presentes."
```

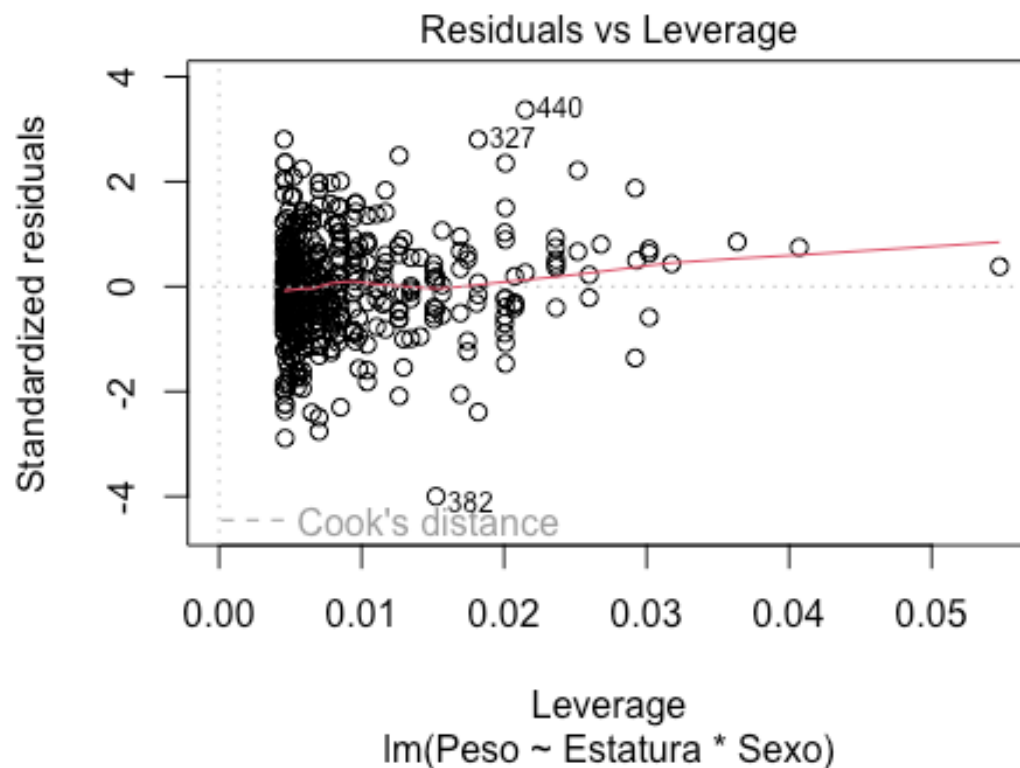
4. Utiliza el comando: plot(modelo). Observa las gráficas obtenidas y contesta:

```
plot(Modelo2_)
```









#¿Cuáles son las diferencias y similitudes de estos gráficos con respecto a los que ya habías analizado?

```
print("Los primeros dos graficos ya se habian calculado antes, en cambio el
3er y 4to grafico no se habian visto antes, la 3era grafica nos ayuda a
visualizar la homocedasticidad, siendo que se comparan los valores ajustados
y los residuos estandarizados, la 4ta grafica demuestra los residuos
estandarizados vs las influencias de las obervaciones del modelo")
```

```
## [1] "Los primeros dos graficos ya se habian calculado antes, en cambio el
3er y 4to grafico no se habian visto antes, la 3era grafica nos ayuda a
visualizar la homocedasticidad, siendo que se comparan los valores ajustados
y los residuos estandarizados, la 4ta grafica demuestra los residuos
estandarizados vs las influencias de las obervaciones del modelo"
```

#Estos gráficos, ¿cambian en algo las conclusiones que ya habías obtenido?

```
print("En realidad ya se habian calculado la mayoría de los datos que nos
brindan estos graficos, por lo cual no han cambiado las conclusiones.")
```

```
## [1] "En realidad ya se habian calculado la mayoría de los datos que nos
brindan estos graficos, por lo cual no han cambiado las conclusiones."
```

5. Emite una conclusión final sobre el mejor modelo de regresión lineal que conjunte lo que hiciste en las tres partes de esta actividad.

```
print("El ultimo modelo fue el cual se ha demostrado ser el mejor de los 4  
modelos iniciales, la interaccion nos brinda mejor visualizacion de los datos  
y mejores resultados en general.")
```

```
## [1] "El ultimo modelo fue el cual se ha demostrado ser el mejor de los 4  
modelos iniciales, la interaccion nos brinda mejor visualizacion de los datos  
y mejores resultados en general."
```

1. Con los datos de las estaturas y pesos de los hombres y las mujeres construye la gráfica de los intervalos de confianza y predicción para la estimación y predicción de Y para el mejor modelo seleccionado.

```
Ip=predict(object=Modelo2_,interval="prediction",level=0.97)
```

```
## Warning in predict.lm(object = Modelo2_, interval = "prediction", level =  
0.97): predictions on current data refer to _future_ responses
```

```
datos1=cbind(M,Ip)
```

```
MM = subset(datos1,datos1$Sexo=="M")
```

```
MH = subset(datos1,datos1$Sexo=="H")
```

```
library(ggplot2)
```

```
ggplot(MM,aes(x=Estatura,y=Peso))+
```

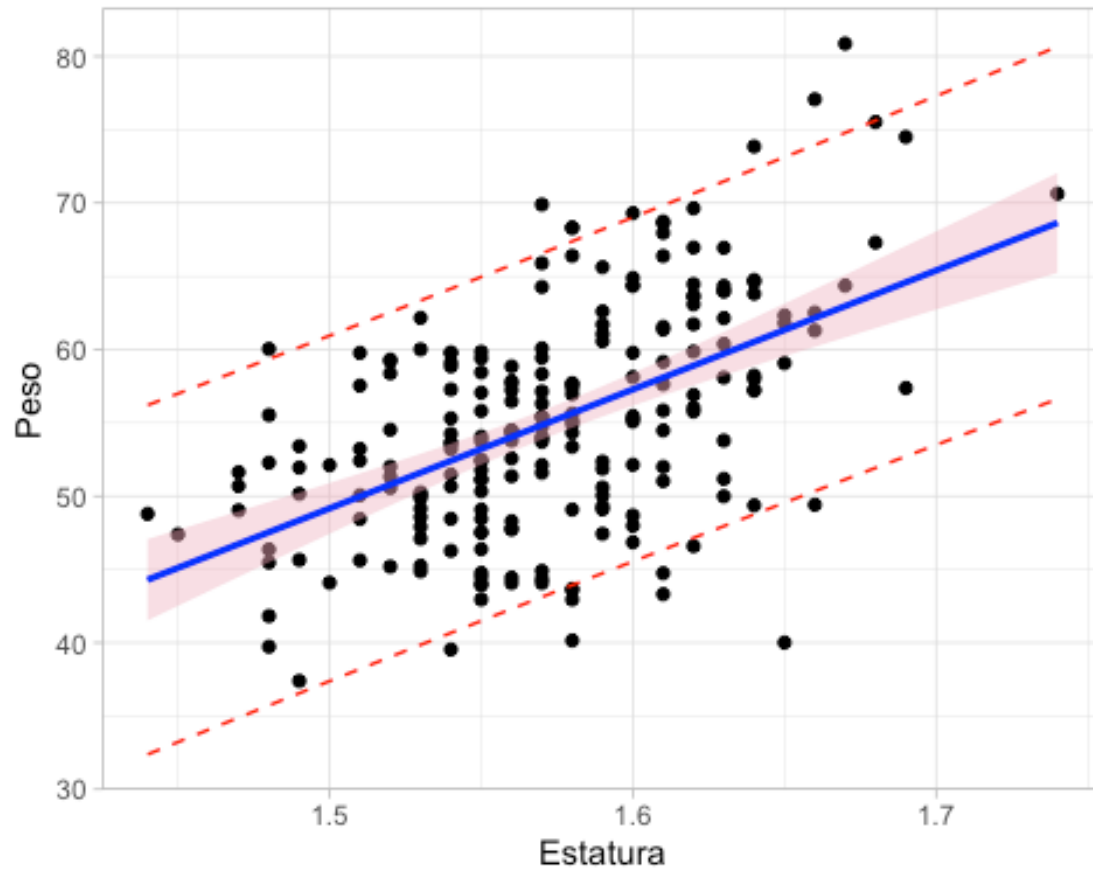
```
  geom_point()+
```

```
  geom_line(aes(y = fit), color = "blue") +
```

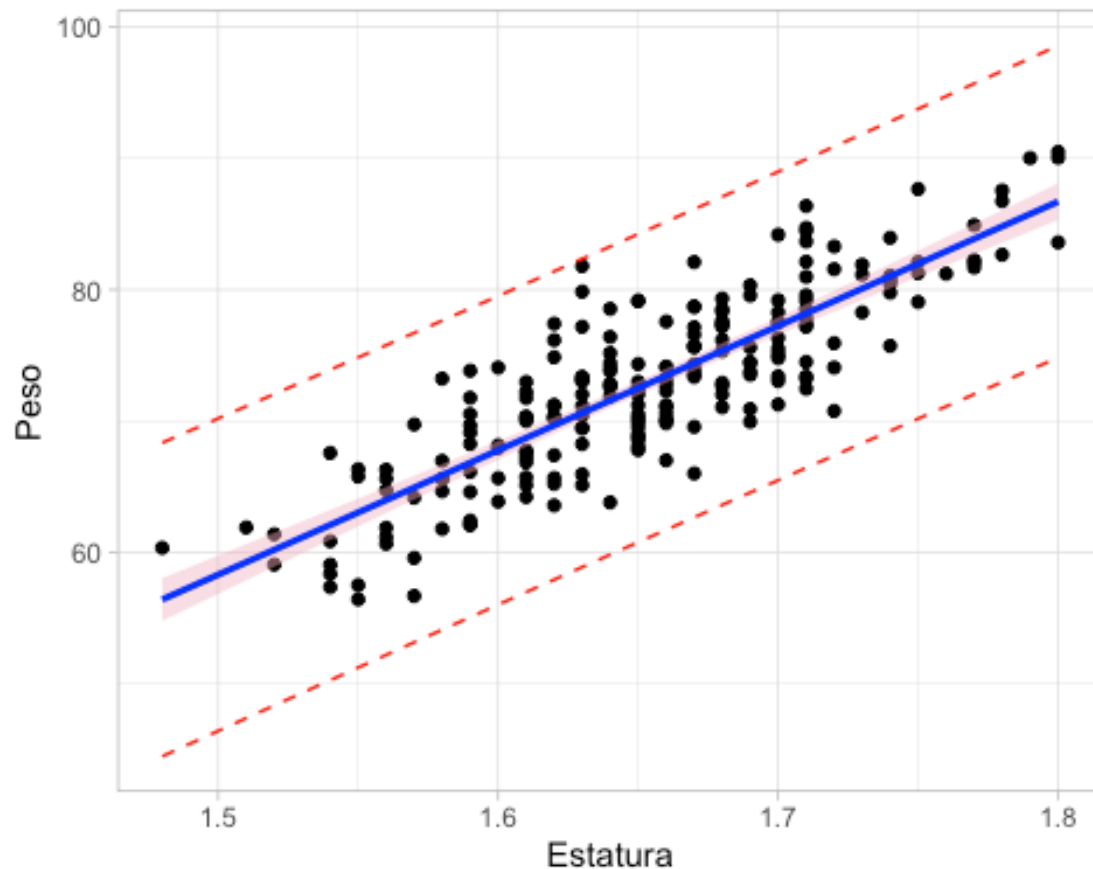
```
  geom_line(aes(y=lwr), color="red", linetype="dashed")+
```

```
  geom_line(aes(y=upr), color="red", linetype="dashed")+
```

```
  geom_smooth(method=lm, formula=y~x, se=TRUE, level=0.97, col="blue",  
fill="pink2")+  
  theme_light()
```

```
library(ggplot2)
ggplot(MH, aes(x=Estatura, y=Peso)) +
  geom_point() +
  geom_line(aes(y=lwr), color="red", linetype="dashed") +
  geom_line(aes(y=upr), color="red", linetype="dashed") +
  geom_smooth(method=lm, formula=y~x, se=TRUE, level=0.97, col="blue",
fill="pink2") +
  theme_light()
```



2. Interpreta y comenta los resultados obtenidos

```
print("En los datos de las mujeres se llegan a presentar datos mas  
irregulares que en el conjunto de los hombres, siendo que se tienen multiples  
datos los cuales salen de lo regular, ambas graficas demuestran la tendencia  
de ambos sexos en que si aumenta la estatura, igualmente aumenta el peso, y  
aunque la mayoría de los datos siguen un crecimiento regular, hay ocasiones  
donde los datos se comportan de manera diferente.")
```

```
## [1] "En los datos de las mujeres se llegan a presentar datos mas  
irregulares que en el conjunto de los hombres, siendo que se tienen multiples  
datos los cuales salen de lo regular, ambas graficas demuestran la tendencia  
de ambos sexos en que si aumenta la estatura, igualmente aumenta el peso, y  
aunque la mayoría de los datos siguen un crecimiento regular, hay ocasiones  
donde los datos se comportan de manera diferente."
```