

## 10. Regresion Lineal

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 1.00000000

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}

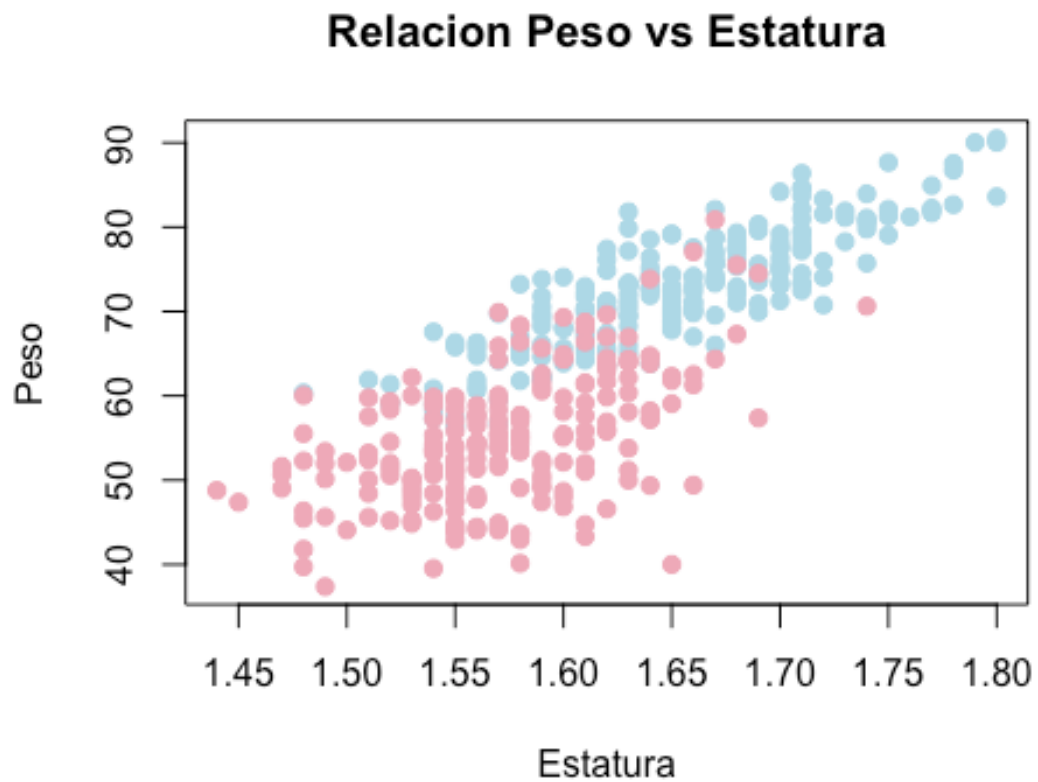
colores= c("lightblue", "pink2")
plot(M$Estatura, M$Peso, data=M, col=colores[factor(M$Sexo)], pch=19,
ylab="Peso", xlab="Estatura", main= "Relacion Peso vs Estatura")

## 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

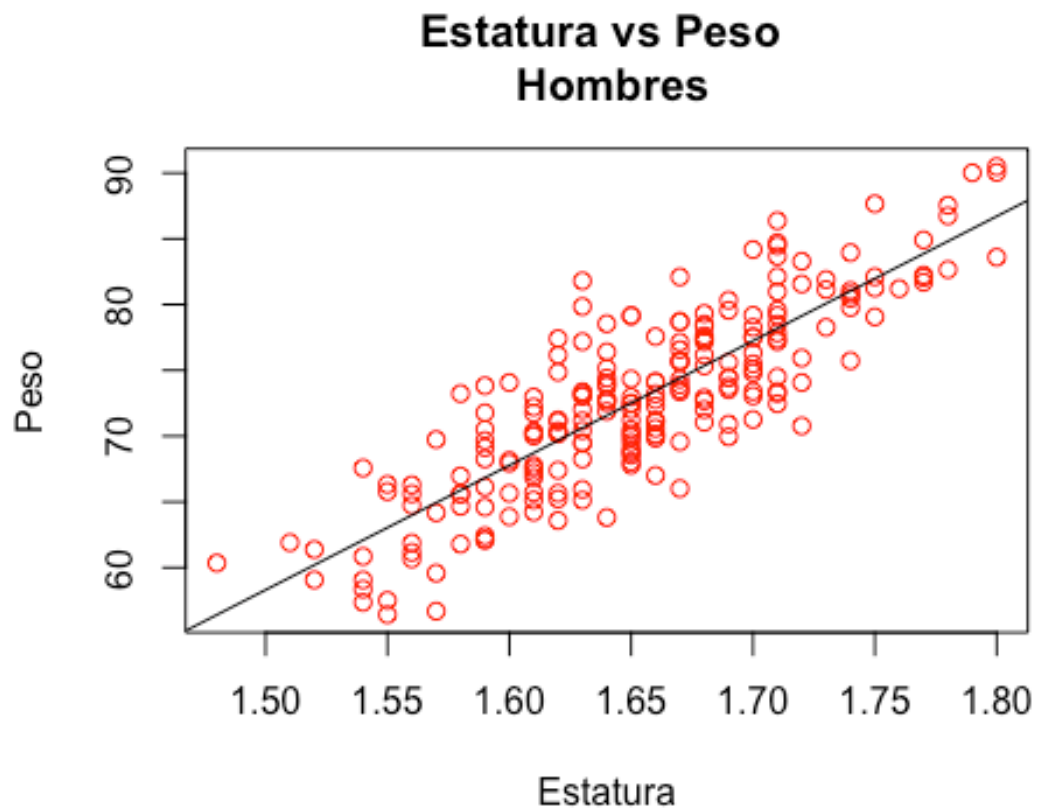
## Warning in title(...): "data" is not a graphical parameter

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

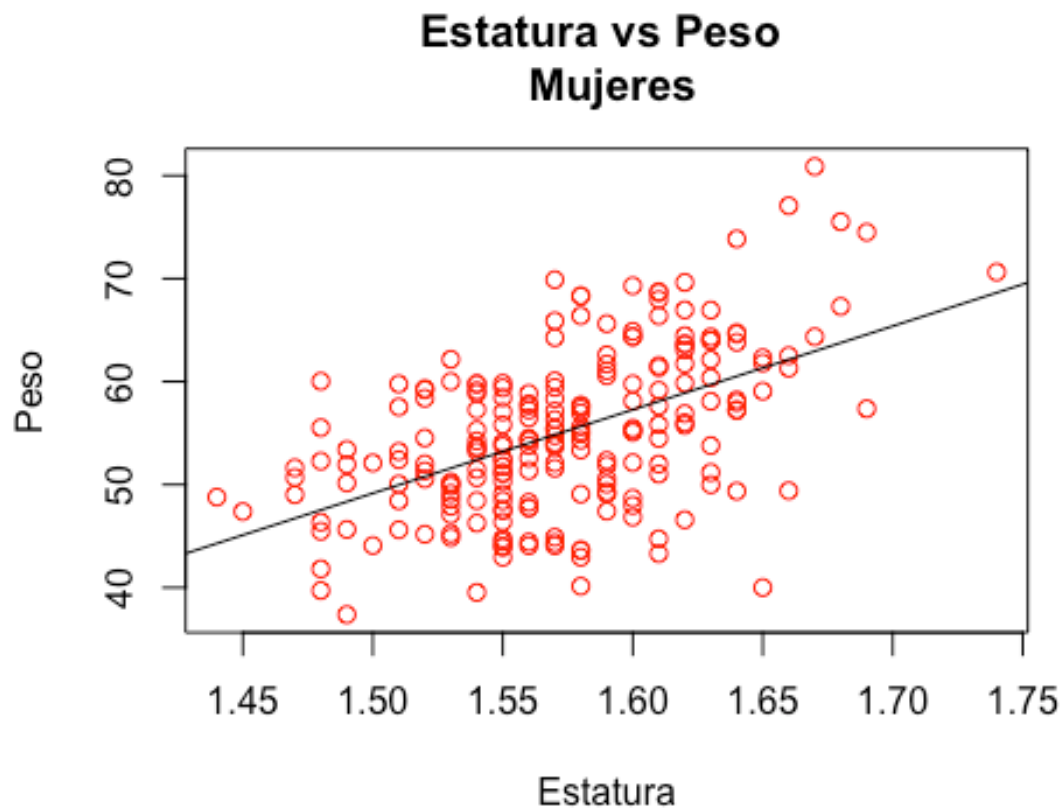


4. Dibuja el diagrama de dispersión de los datos y la recta de mejor ajuste.

```
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."
```