

Actividad10

Facundo Colasurdo Caldironi

2024-08-30

Analiza la base de datos de estatura y peso Download datos de estatura y peso de los hombres y mujeres en México y obten el mejor modelo de regresión para esos datos ## La recta de mejor ajuste

analisis descriptivo

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

```
M=read.csv("file:///Users/facundocolasurdocaldironi/Downloads/Estatura-peso_HyM.csv") #leer la base de datos
head(M)
```

```
##   Estatura  Peso Sexo
## 1    1.61  72.21   H
## 2    1.61  65.71   H
## 3    1.70  75.08   H
## 4    1.65  68.55   H
## 5    1.72  70.77   H
## 6    1.63  77.18   H
```

```
MM = subset(M,M$Sexo=="M")
MH = subset(M,M$Sexo=="H")
M1=data.frame(MH$Estatura,MH$Peso,MM$Estatura,MM$Peso)
```

correlacion

```
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
```

Se observa que la hay una fuerte relacion entre el peso y la altura, pero solo dentro del mismo género, pero casi ninguna cuando se cruza entre generos.

En resumen, se observa una fuerte relación entre la estatura y el peso dentro del mismo género, pero poca o ninguna correlación cruzada entre géneros.

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

```

```

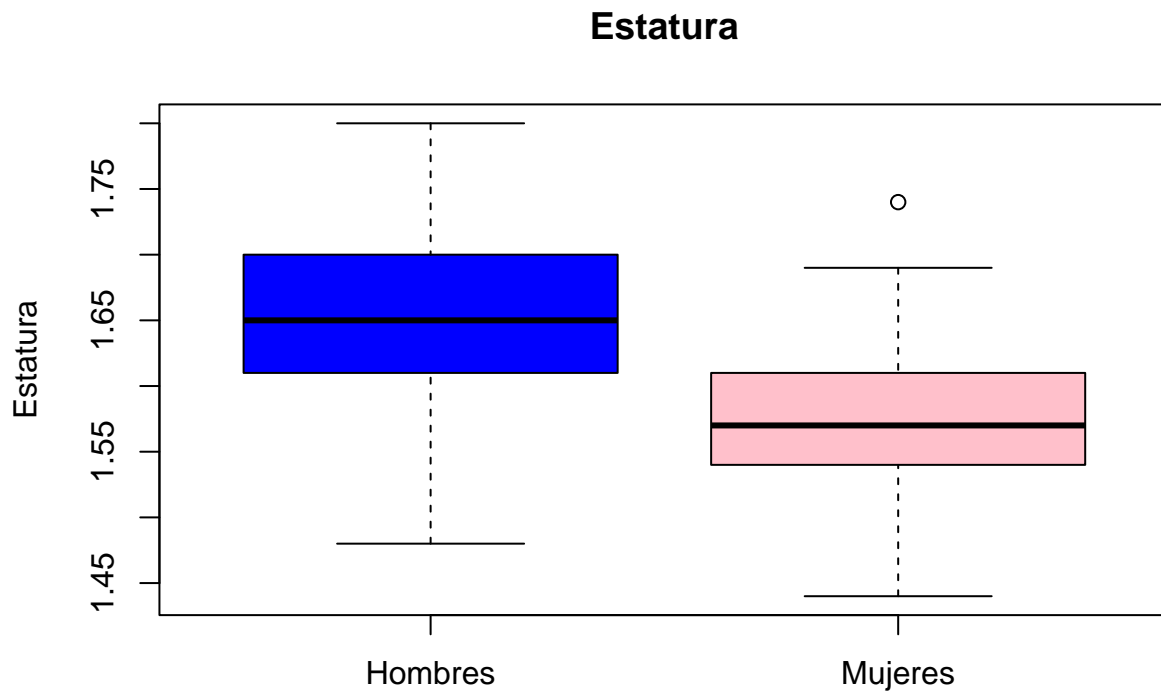
##           Minimo      Q1 Mediana      Media      Q3 Máximo      Desv Est
## H-Estatura   1.48  1.6100   1.650  1.653727  1.7000   1.80  0.06173088
## H-Peso       56.43 68.2575   72.975 72.857682 77.5225   90.49  6.90035408
## M-Estatura   1.44  1.5400   1.570  1.572955  1.6100   1.74  0.05036758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

```

```

boxplot(M$Estatura~M$Sexo, ylab="Estatura", xlab="", col=c("blue","pink"), names=c("Hombres", "Mujeres"))

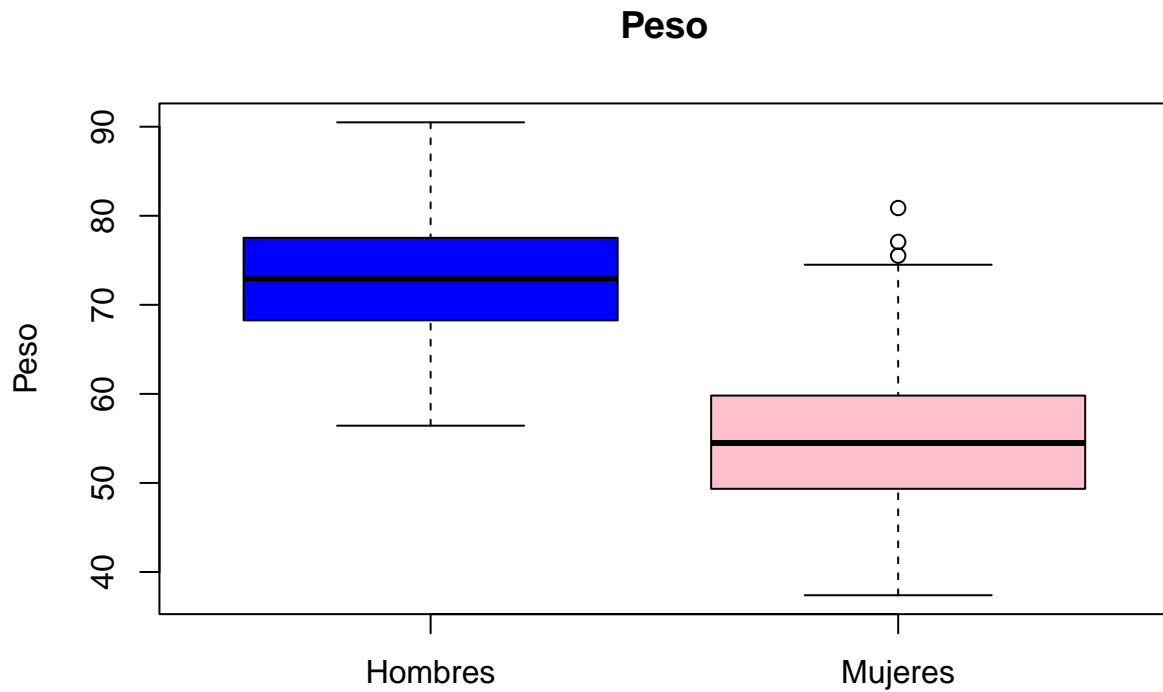
```



```

boxplot(M$Peso~M$Sexo, ylab="Peso",xlab="", names=c("Hombres", "Mujeres"), col=c("blue","pink"), main="")

```



la recta de mejor ajuste Encuentra la ecuación de regresión de mejor ajuste:

```
modelo1H = lm(Peso ~ Estatura, data = MH)
modelo1H
```

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

```
modelo1M = lm(Peso ~ Estatura, data = MM)
modelo1M
```

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

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

##Hombres

```
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
```

El 71 % de la variabilidad de y esta explicada por el modelo, mientras que el otro porcentaje esta en los errores en los datos de los hombres.

##Mujeres

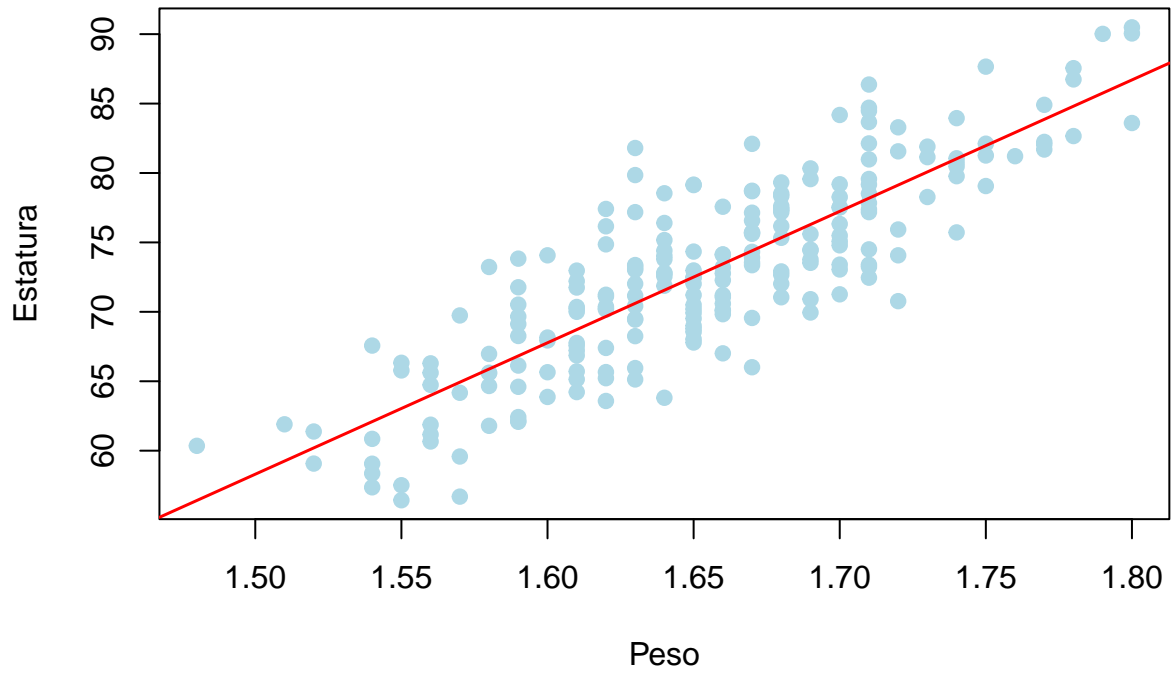
```
summary(modelo1M)
```

```
##
## 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
```

El 27.51 % de la variabilidad de y esta explicada por el modelo, mientras que el otro porcentaje esta en los errores en los datos de los mujeres

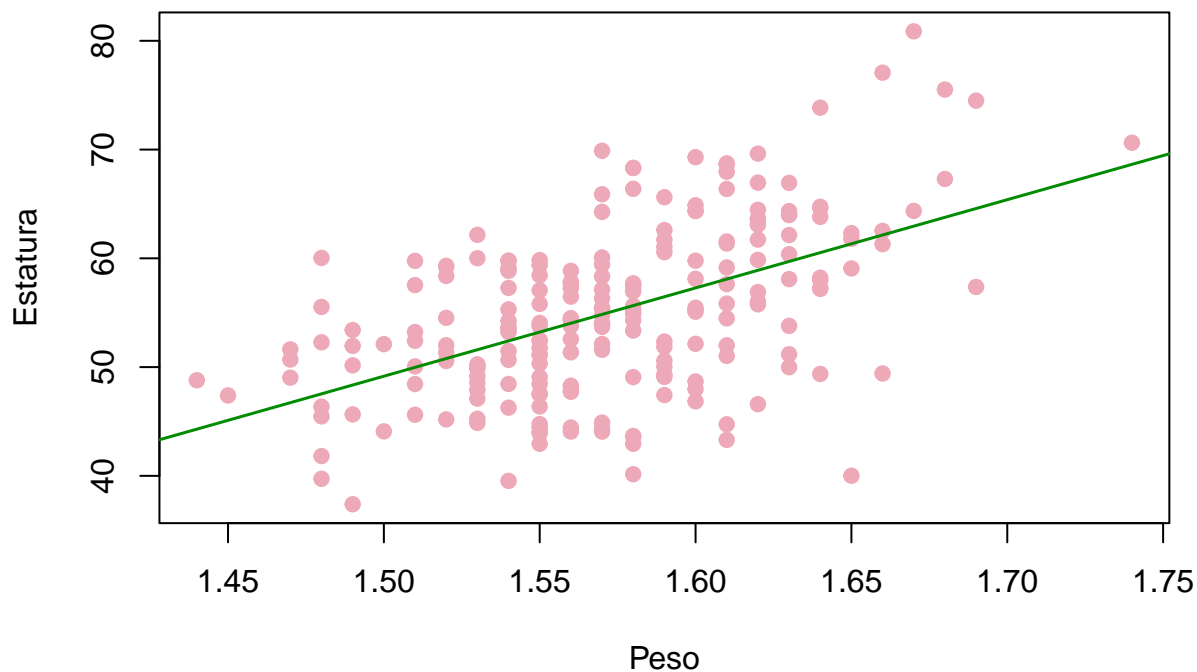
```
plot(MH$Estatura, MH$Peso, col="lightblue", main="Estatura vs Peso \n Hombres", ylab="Estatura", xlab =
abline(modelo1H, col="red", lwd=1.5)
```

Estatura vs Peso Hombres



```
plot(MM$Estatura, MM$Peso, col="pink2", main="Estatura vs Peso \n Mujeres", ylab="Estatura",xlab = "Peso",  
abline(modelo1M, col="green4", lwd=1.5))
```

Estatura vs Peso Mujeres



Realiza la regresión entre las variables involucradas

###Un modelo

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

```
##
## Call:
## lm(formula = Peso ~ Estatura + Sexo, data = M)
##
## Coefficients:
## (Intercept)      Estatura      SexoM
##      -74.75         89.26        -10.56
```

```
summary(Modelo2)
```

```
##
## Call:
## lm(formula = Peso ~ Estatura + Sexo, data = M)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.9505  -3.2491   0.0489   3.2880  17.1243
##
## Coefficients:
```

```
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) -74.7546      7.5555  -9.894  <2e-16 ***
## Estatura    89.2604      4.5635  19.560  <2e-16 ***
## SexoM       -10.5645      0.6317 -16.724  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.381 on 437 degrees of freedom
## Multiple R-squared:  0.7837, Adjusted R-squared:  0.7827
## F-statistic: 791.5 on 2 and 437 DF,  p-value: < 2.2e-16
```

A 0.05 si es significativo y los modelos quedarian:

SexoM = 1 -> Mujer SexoM = 0 -> Hombre

Mujeres: $\text{Estatura} = 1.38622 + 0.00339P$ ($P = \text{Peso}$) $E = 1.2727097 + 0.0052296P + 0.0121799\text{SexoM}$ $E = 1.2848 + 0.0052296P$

$\text{Peso} = -72.560 + 81.149E$ ($E = \text{Estatura}$)

Hombres= $\text{Estatura} = 1.101770 + 0.007576P$ ($P = \text{Peso}$) $E = 1.2727097 + 0.0052296P$

$\text{Peso} = -83.68 + 94.66E$ ($E = \text{Estatura}$)

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

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

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

```
colores =c("blue", "pink3")
```

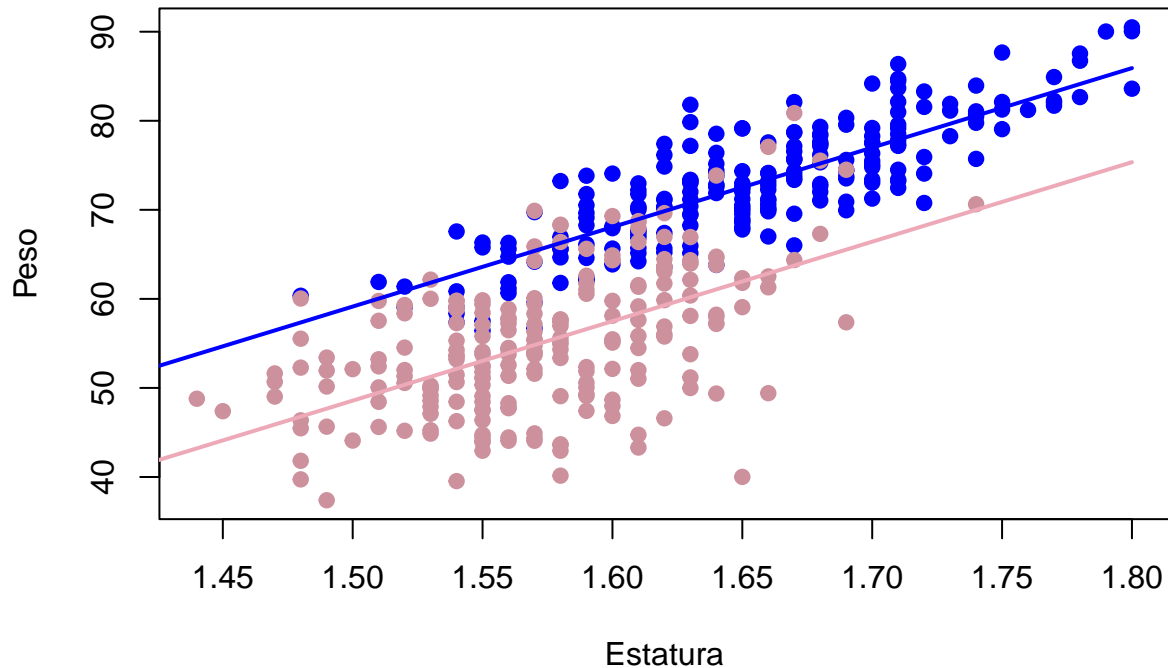
```
plot(M$Estatura, M$Peso, , col= colores[factor(M$Sexo)], pch=19, ylab="Peso", xlab="Estatura", main="Re")
```

```
x= seq(1.40, 1.80, 0.01)
```

```
lines(x,Ym(x),col="pink2", lwd = 2)
```

```
lines(x,Yh(x), col="blue", lwd = 2)
```

Relacion Peso vs Estatura



Esta grafica nos deja en claro que los hombres estarán una estatura mayor que las mujeres, al mismo tiempo que el peso depende de la estatura de la persona y que en general, podemos decir que las mujeres son de menor estatura y de menor peso, mientras que los hombres son los que tienen una mayor tamaño y peso

interpreta el contexto del problema

¿Qué información proporciona b_0 sobre la relación entre la estatura y el peso de hombres y mujeres? Proporciona el peso esperado para una persona con estatura cero, que es teórico y solo sirve como punto de ajuste de tanto hombres como mujeres

¿Cómo interpretas b_1 en la relación entre la estatura y el peso de hombres y mujeres? Muestra cómo cambia el peso con la estatura, independientemente del sexo de la persona, al mismo tiempo, demuestra que hay una relación entre ambas, en donde mayor el peso, mayor la estatura, las graficas demuestran que los hombres tienen una pendiente mayor, lo cual nos indica que los hombres son más pesados de manera general, todo lo anterior nos demuestra que tienen una relación significativa.