M1_A10

Sofia Cantu

2024-08-30

La recta de mejor ajuste (Primera entrega)

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

| Estatura | Peso

Etatura Peso | 1.000 | 0.803 | 0.803 | 1.000

```
# Cargar los datos
M = read.csv("~/Downloads/ArchivosCodigos/Estatura-peso_HyM.csv")

# Crear el data frame
datos <- data.frame(Estatura = M$Estatura, Peso = M$Peso)

# Calcular la matriz de correlación
correlacion <- cor(datos)
print(correlacion)
```

```
##           Estatura      Peso
## Estatura 1.0000000 0.8032449
## Peso     0.8032449 1.0000000
```

```
cat("\nInterpretación: La correlación entre la estatura y el peso es positiva y fuerte (0.803), lo que indica una relación significativa. A medida que la estatura aumenta, también tiende a aumentar el peso de manera considerable.\n")
```

```
##
## Interpretación: La correlación entre la estatura y el peso es positiva y fuerte (0.803), lo que indica una relación significativa. A medida que la estatura aumenta, también tiende a aumentar el peso de manera considerable.
```

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

n=4
d=matrix(NA,ncol=7,nrow=n)
for(i in 1:n){
  d[i,]<-c(as.numeric(summary(M1[,i])),sd(M1[ ,i]))
```

```r
}
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")
```

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

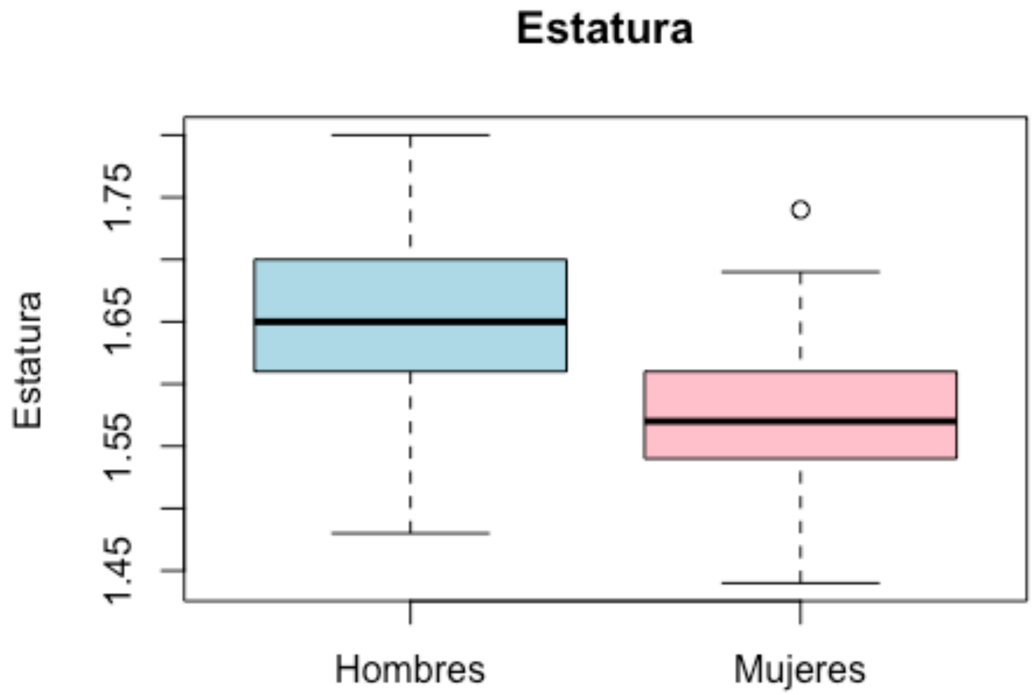
```r
# Resumen estadístico de los datos
resumen <- summary(datos)
print(resumen)
```

```
##     Estatura          Peso
##  Min.   :1.440   Min.   :37.39
##  1st Qu.:1.560   1st Qu.:54.49
##  Median :1.610   Median :64.53
##  Mean   :1.613   Mean   :63.97
##  3rd Qu.:1.660   3rd Qu.:73.22
##  Max.   :1.800   Max.   :90.49
```

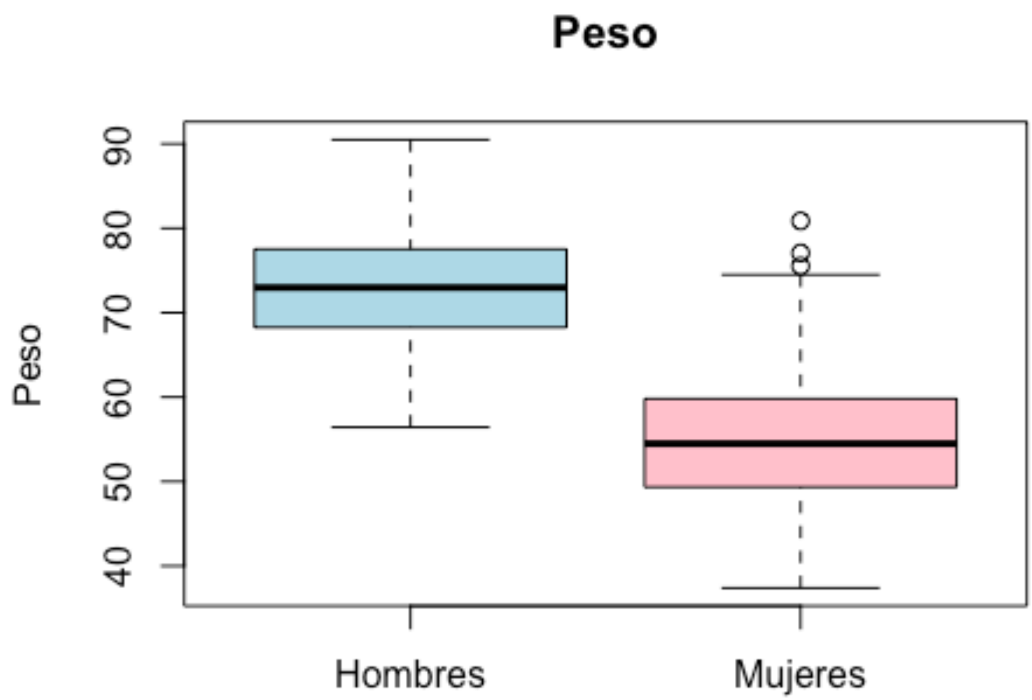
```r
# Calcular la desviación estándar
desviacion_estandar <- apply(datos, 2, sd)
print(desviacion_estandar)
```

```
##    Estatura        Peso
##  0.06929171 11.54161456
```

```r
boxplot(M$Estatura~M$Sexo, ylab="Estatura", xlab="",
col=c("lightblue","pink"), names=c("Hombres", "Mujeres"), main="Estatura")
```




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



3. Encuentra la ecuación de regresión de mejor ajuste:

3.1 Realiza la regresión entre las variables involucradas

```r
# Ajustar el modelo de regresión lineal
modelo <- lm(Peso ~ Estatura, data = datos)

# Resumen del modelo
summary(modelo)
```

```
##
## Call:
## lm(formula = Peso ~ Estatura, data = datos)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -28.8653  -3.7654   0.6706   5.0142  15.6006
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -151.883      7.655  -19.84   <2e-16 ***
## Estatura     133.793      4.741   28.22   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 6.883 on 438 degrees of freedom
## Multiple R-squared:  0.6452, Adjusted R-squared:  0.6444
## F-statistic: 796.5 on 1 and 438 DF,  p-value: < 2.2e-16
```

Hombres

Hipótesis: • $H_0: \beta_1 = 0$ • $H_1: \beta_1 \neq 0$

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

```
##
## Call:
## lm(formula = Estatura ~ Peso, data = MH)
##
## Coefficients:
## (Intercept)         Peso
##    1.101770     0.007576
```

```
summary(Modelo1H)
```

```
##
## Call:
## lm(formula = Estatura ~ Peso, data = MH)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.091473 -0.020942  0.001445  0.024020  0.082089
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1017704  0.0235832   46.72   <2e-16 ***
## Peso        0.0075758  0.0003223   23.51   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03291 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
```

Mujeres

```
Modelo1M = lm(Estatura~Peso, MM)
Modelo1M
```

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

```
## (Intercept)        Peso
##     1.38622     0.00339
```

```r
summary(Modelo1M)
```

```
##
## Call:
## lm(formula = Estatura ~ Peso, data = MM)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.11162 -0.02611 -0.00174  0.02806  0.12814
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3862212  0.0207336  66.859   <2e-16 ***
## Peso        0.0033900  0.0003727   9.096   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04298 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
```

3.2 Verifica el modelo:

3.2.1 Verifica la significancia del modelo con un alfa de 0.03. 3.2.1 Verifica la significancia de $\hat{\beta}i$ con un alfa de 0.03. 3.2.1 Verifica el porcentaje de variación explicada por el modelo

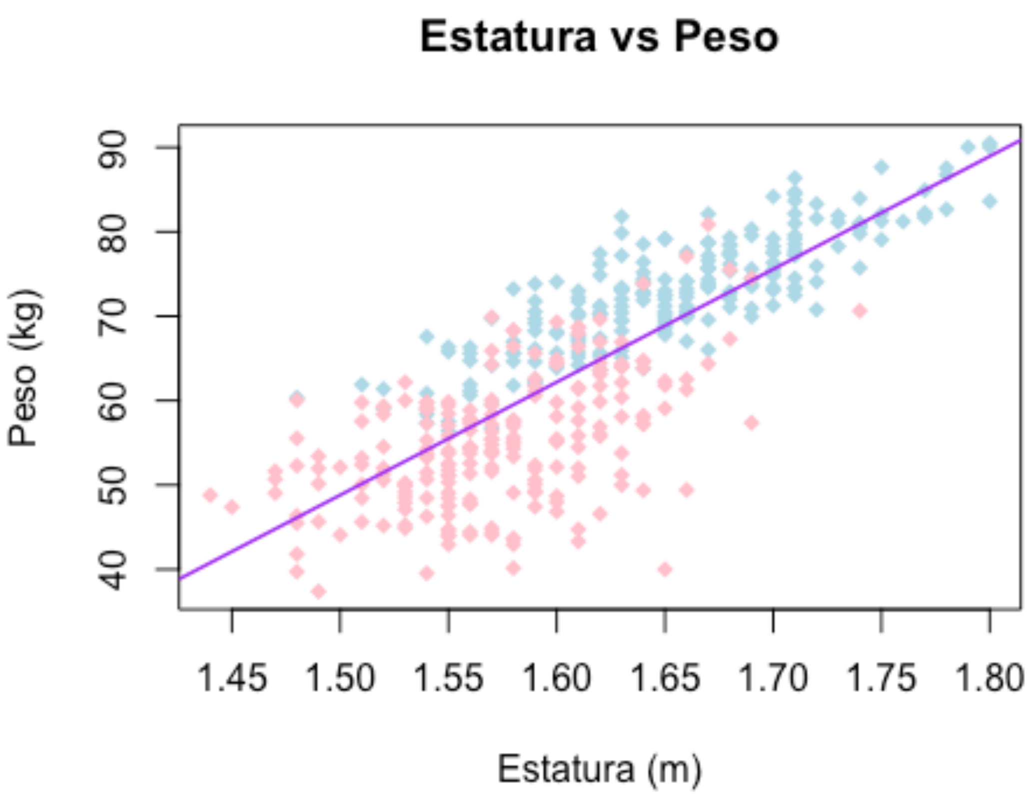
```r
# Verificar la significancia del modelo y los coeficientes
summary(modelo)
```

```
##
## Call:
## lm(formula = Peso ~ Estatura, data = datos)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -28.8653  -3.7654   0.6706   5.0142  15.6006
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -151.883      7.655  -19.84   <2e-16 ***
## Estatura     133.793      4.741   28.22   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.883 on 438 degrees of freedom
## Multiple R-squared:  0.6452, Adjusted R-squared:  0.6444
## F-statistic: 796.5 on 1 and 438 DF,  p-value: < 2.2e-16
```

```r
# Obtener el R cuadrado
r_squared <- summary(modelo)$r.squared
print(r_squared)

## [1] 0.6452023
```

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

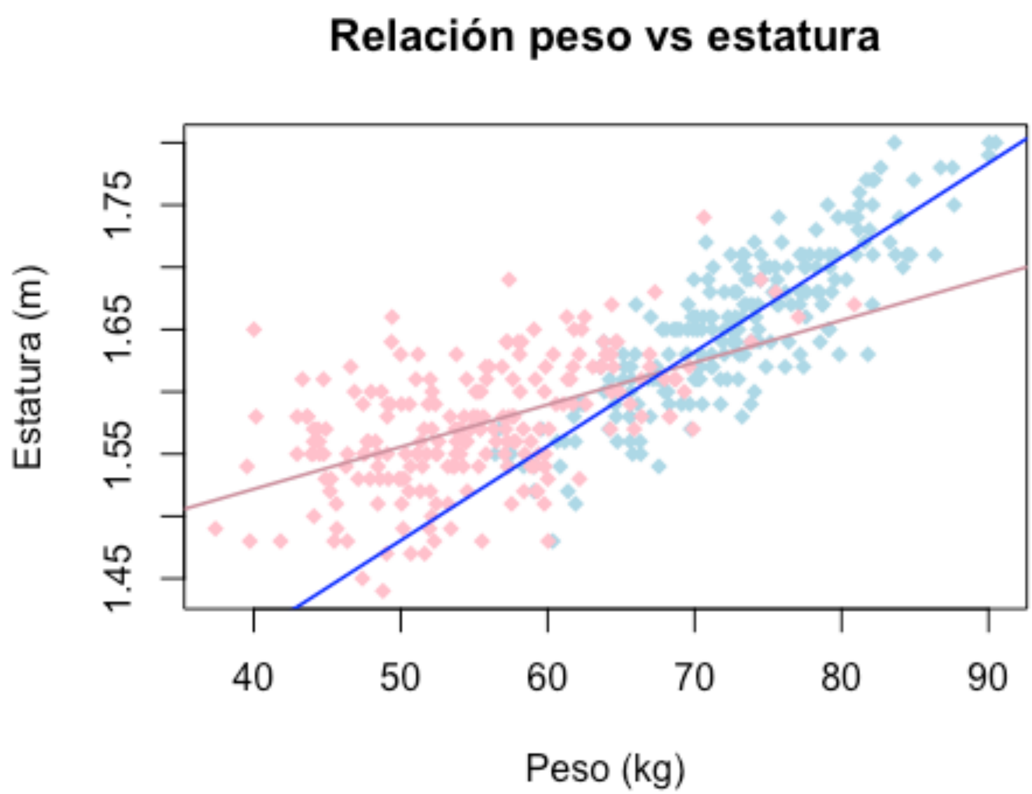
```r
colors <- ifelse(M$Sexo == "H", "lightblue", "pink")
# Diagrama de dispersión y recta de mejor ajuste
plot(datos$Estatura, datos$Peso, main="Estatura vs Peso",
    ylab="Peso (kg)", xlab="Estatura (m)", col=colors, pch=18)
abline(modelo, col="purple", lwd=1.5)
```



A 0.05 si es significativo y los modelos quedarían:

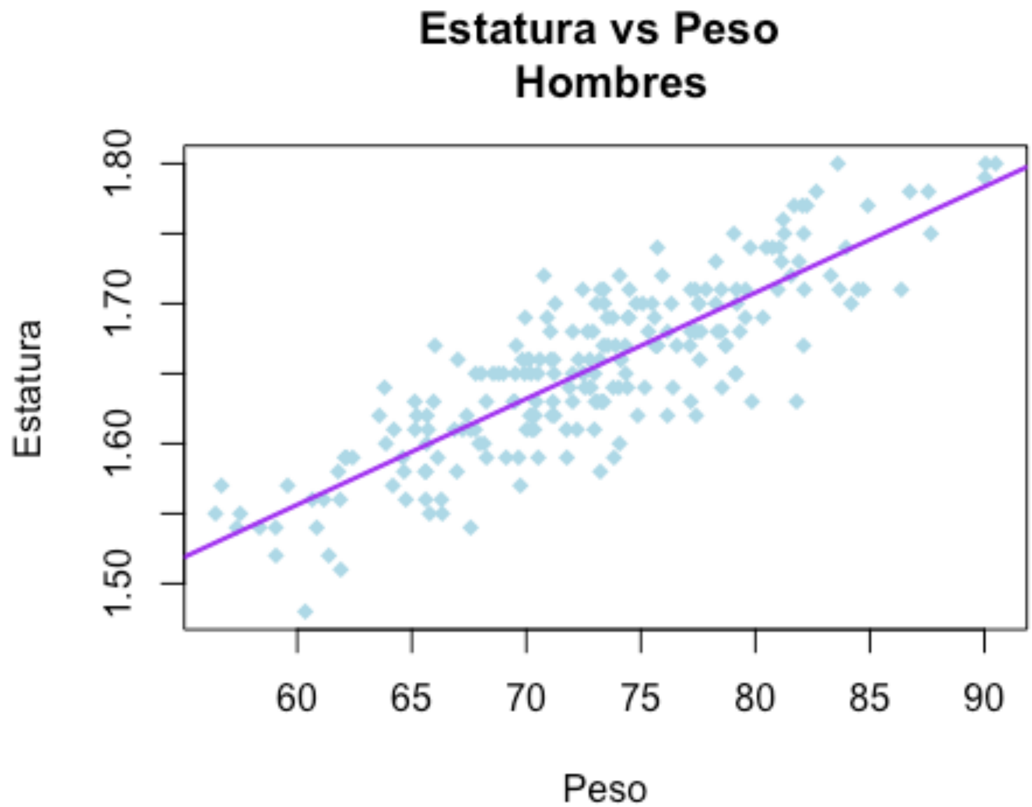
```r
# Extraer coeficientes del modelo
b0 <- modelo$coefficients[1]
b1 <- modelo$coefficients[2]
b2 <- modelo$coefficients[3]

# Definir funciones para las líneas
Ym <- function(x){b0 + b2 + b1*x}
Yh <- function(x){b0 + b1*x}
```

```r
# Definir colores
colors <- ifelse(M$Sexo == "H", "lightblue", "pink")

# Crear el gráfico de dispersión
plot(M$Peso, M$Estatura, col=colors, pch=18,
     xlab="Peso (kg)", ylab="Estatura (m)", main="Relación peso vs estatura")

# Generar secuencia para x
x <- seq(1.40, 1.80, 0.01)

# Agregar líneas
lines(x, Ym(x), col="pink3", lwd=2)
lines(x, Yh(x), col="blue", lwd=2)

abline(Modelo1M, col="pink3", lwd=1.5)
abline(Modelo1H, col="blue", lwd=1.5)
```



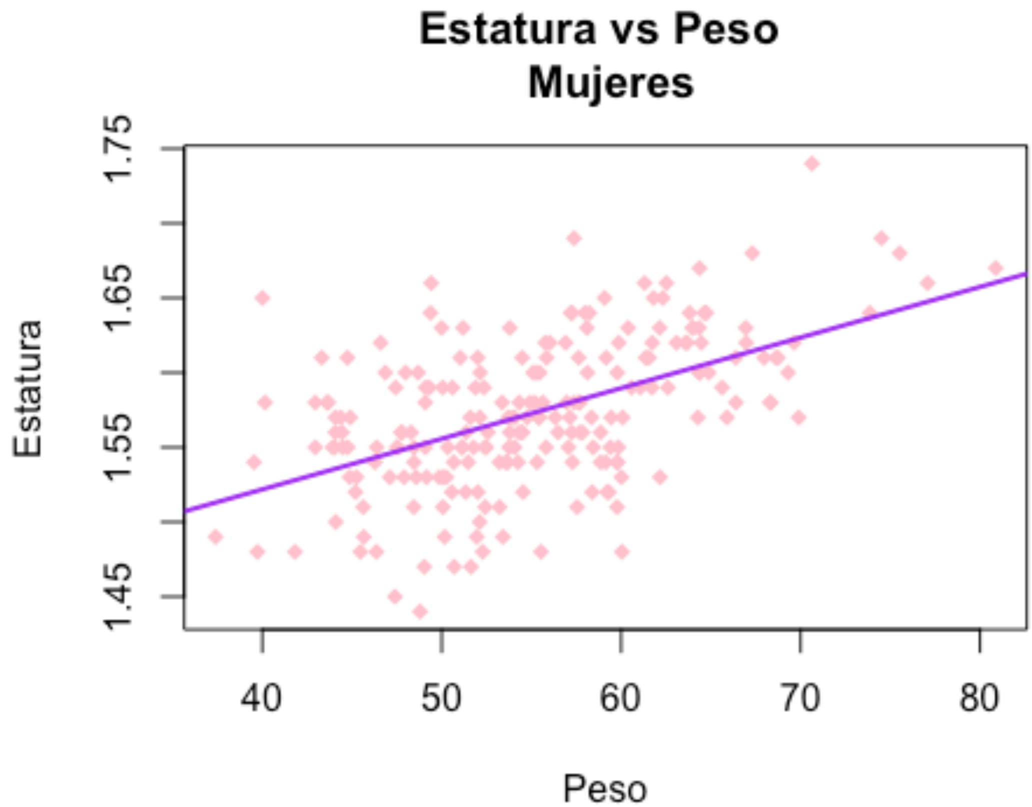
Hombres

```r
plot(MH$Peso, MH$Estatura, col="lightblue", main="Estatura vs Peso \n
Hombres", ylab="Estatura", xlab="Peso", pch=18)
abline(Modelo1H, col="purple", lwd=2)
```



Mujeres

```
plot(MM$Peso, MM$Estatura, col="pink", main="Estatura vs Peso \n Mujeres",
ylab="Estatura", xlab="Peso", pch=18)
abline(Modelo1M, col="purple", lwd=2)
```



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

- Correlación entre Estatura y Peso: La matriz de correlación muestra una relación positiva y significativa entre la estatura y el peso (r = 0.803). Esto indica que, a medida que aumenta la estatura de una persona, es probable que su peso también aumente. Esta relación es fuerte, sugiriendo que el peso es un buen indicador de la estatura en este grupo de datos.

- Distribución Descriptiva: Los resúmenes estadísticos para estatura y peso muestran que, en promedio, los hombres tienden a ser más altos y más pesados que las mujeres. Las diferencias en los valores mínimos, máximos y los cuartiles reflejan la variabilidad dentro de cada grupo. Por ejemplo, los hombres tienen una estatura media de 1.613 m y un peso medio de 63.97 kg, mientras que las mujeres tienden a ser más bajas y ligeras.

- Diagramas de Caja para Estatura y Peso por Género: Los diagramas de caja indican que la distribución de la estatura y el peso difiere entre hombres y mujeres. Los hombres tienden a tener valores más altos de estatura y peso. La presencia de valores atípicos en ambos géneros sugiere que algunas observaciones están significativamente alejadas de la mediana.

- Regresión Lineal: Los análisis de regresión muestran una relación significativa entre la estatura y el peso tanto para hombres como para mujeres. El valor de R-cuadrado ajustado en ambos casos sugiere que una parte considerable de la variabilidad en la estatura puede ser explicada por el peso, aunque esta relación es más fuerte en hombres (R^2 ajustado ≈ 0.717) que en mujeres (R^2 ajustado ≈ 0.271).

- Gráficos de Dispersión con Líneas de Regresión: Los gráficos muestran una clara tendencia ascendente, lo que refuerza la relación positiva entre estatura y peso. Las líneas de regresión para hombres y mujeres tienen pendientes distintas, lo que sugiere que el peso afecta la estatura de manera diferente según el género.

6. Interpreta en el contexto del problema:

6.1 ¿Qué información proporciona $\hat{\beta}0$ sobre la relación entre la estatura y el peso de hombres y mujeres?

Para hombres: El número que obtuvimos no tiene mucho sentido en la vida real (nadie mide 1.10 metros con peso cero), pero nos ayuda a hacer cálculos y predicciones. Para mujeres: Pasa lo mismo, el número (1.38 metros) no es realista, pero nos sirve para comparar. Nos da una pista de que, en general, las mujeres son un poco más bajitas que los hombres a igual peso.

6.2 ¿Cómo interpretas $\hat{\beta}1$ en la relación entre la estatura y el peso de hombres y mujeres?

Para hombres: Por cada kilo que sube el peso, la altura aumenta en promedio 0.0076 metros (menos de un centímetro). No parece mucho, pero indica que hay una relación bastante fuerte entre peso y altura en los hombres. Para mujeres: Aquí, por cada kilo extra, la altura aumenta en promedio 0.00339 metros. Es menos que en los hombres, lo que sugiere que en las mujeres, el peso no está tan directamente relacionado con la altura.