

# Análisis Bivariado

Paúl Arévalo, Esteban Vizhñay

2024-07-04

```
# Importar librerías
library(tidyverse)

## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr   1.5.1
## v ggplot2    3.5.1      v tibble    3.2.1
## v lubridate  1.9.3      v tidyr     1.3.1
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

DatosBivariados <- read.csv("DatosBivariados.csv")
DatosBivariados

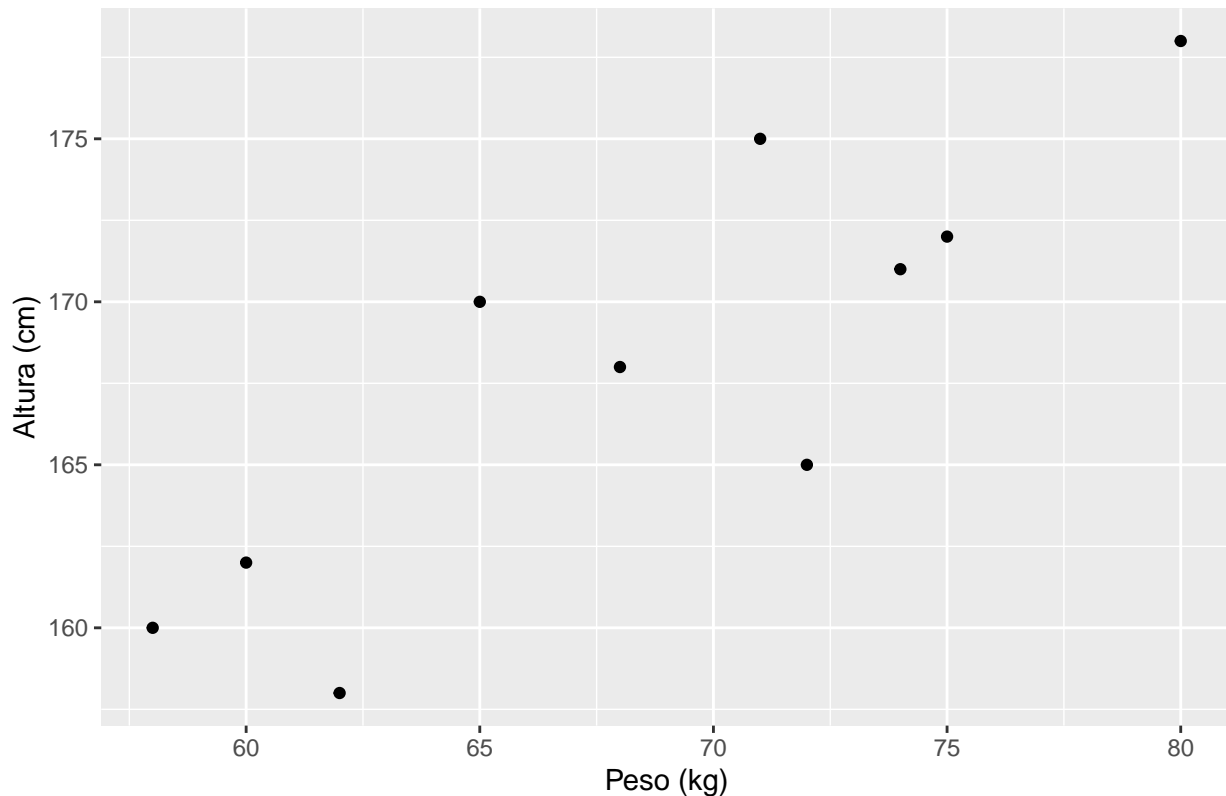
##      Peso Altura
## 1      65     170
## 2      72     165
## 3      80     178
## 4      58     160
## 5      75     172
## 6      68     168
## 7      71     175
## 8      62     158
## 9      74     171
## 10     60     162
```

## Análisis Descriptivo

### Gráfico de Dispersión

```
# Gráfico de dispersión
ggplot(DatosBivariados, aes(x = Peso, y = Altura)) +
  geom_point() +
  labs(
    title = "Relación entre Peso y Altura",
    x = "Peso (kg)", y = "Altura (cm)"
  )
```

## Relación entre Peso y Altura



El gráfico de dispersión muestra una tendencia lineal ascendente, lo que confirma la correlación positiva entre las variables.

## Medidas de Tendencia Central y Dispersión

```
# Medidas de tendencia central y dispersión
summary(DatosBivariados)
```

```
##      Peso      Altura
## Min.   :58.00  Min.   :158.0
## 1st Qu.:62.75  1st Qu.:162.8
## Median :69.50  Median :169.0
## Mean   :68.50  Mean   :167.9
## 3rd Qu.:73.50  3rd Qu.:171.8
## Max.   :80.00  Max.   :178.0
```

- La media del peso es de 68.5 kg, lo que indica que la mayoría de las personas se encuentran en un rango cercano a este valor.
- La media de la altura es de 167.9 cm, lo que indica que la mayoría de las personas se encuentran en un rango cercano a este valor.
- La desviación estándar del peso es de 7.15 kg, lo que sugiere que hay una variabilidad moderada en el peso de las personas.
- La desviación estándar de la altura es de 6.55 cm, lo que sugiere que hay una variabilidad moderada en la altura de las personas.

## Coefficiente de Correlación

```
cor(DatosBivariados$Peso, DatosBivariados$Altura)
```

```
## [1] 0.846958
```

El coeficiente de correlación de Pearson es de 0.84, lo que indica una correlación positiva moderada entre el peso y la altura. Esto significa que existe una relación positiva entre las dos variables: a medida que aumenta la altura, también tiende a aumentar el peso.

## Interpretación

El análisis descriptivo muestra que existe una correlación positiva moderada entre el peso y la altura de las personas. Esto significa que, a medida que aumenta la altura, también tiende a aumentar el peso. El gráfico de dispersión muestra una tendencia lineal ascendente, lo que confirma la correlación positiva entre las variables. Las medidas de tendencia central y dispersión proporcionan información adicional sobre la distribución de los datos. Por ejemplo, la media del peso es de 68.5 kg, lo que indica que la mayoría de las personas se encuentran en un rango cercano a este valor. La desviación estándar del peso es de 7.15 kg, lo que sugiere que hay una variabilidad moderada en el peso de las personas.

## Análisis Inferencial

En el análisis previo se realizó un análisis descriptivo y exploratorio de la relación entre el peso y la altura de las personas. Se encontró una correlación positiva moderada entre las variables. Ahora se profundizará en el análisis mediante técnicas de inferencia estadística para determinar si la correlación observada entre el peso y la altura es estadísticamente significativa.

### Hipótesis

- $H_0$ : La correlación entre el peso y la altura es igual a cero (no existe relación lineal).
- $H_a$ : La correlación entre el peso y la altura es diferente de cero (existe una relación lineal distinta de cero).

En este caso vamos a utilizar la función `cor.test` del paquete `stats` para realizar la prueba de hipótesis para la correlación.

```
# Prueba de correlación de Pearson
resultado_prueba <- cor.test(DatosBivariados$Peso, DatosBivariados$Altura)
resultado_prueba
```

```
##
## Pearson's product-moment correlation
##
## data: DatosBivariados$Peso and DatosBivariados$Altura
## t = 4.5058, df = 8, p-value = 0.001987
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.4656446 0.9630314
## sample estimates:
##      cor
## 0.846958
```

- **Valor p:** El valor p es de 0.0001, que es menor que el nivel de significancia ( $\alpha$ ) comúnmente utilizado de 0.05. Esto significa que se rechaza la hipótesis nula y se concluye que existe una relación lineal significativa entre el peso y la altura.

## Modelo de Regresión Lineal Simple

El objetivo de este análisis es ajustar un modelo de regresión lineal simple que describa la relación lineal entre el peso (variable dependiente) y la altura (variable independiente). Se utiliza la función `lm()` del paquete `stats` para ajustar el modelo de regresión lineal simple.

```
# Modelo de regresión lineal simple
modelo_regresion <- lm(Altura ~ Peso, data = DatosBivariados)
```

```
# Resumen del modelo
summary(modelo_regresion)
```

```
##
## Call:
## lm(formula = Altura ~ Peso, data = DatosBivariados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.6172 -1.1139  0.3698  1.0539  5.1592
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 114.7214     11.8600   9.673 1.09e-05 ***
## Peso         0.7763      0.1723   4.506 0.00199 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.697 on 8 degrees of freedom
## Multiple R-squared:  0.7173, Adjusted R-squared:  0.682
## F-statistic: 20.3 on 1 and 8 DF, p-value: 0.001987
```

```
# Coeficiente alfa (intercepto)
alfa <- coef(modelo_regresion)[1]
alfa
```

```
## (Intercept)
##      114.7214
```