

# Gráficas con ggplot en R

Realizar los gráficos utilizando la librería ggplot de R deberá adjuntar tanto el código utilizado como los resultados obtenidos.

Los datos HEART contienen una variable binaria (HD) de 303 pacientes que se presentaron con dolor en el pecho.

Un valor de resultado de “Yes” indica la presencia de enfermedad del corazón basado en una prueba angiográfica, mientras que “No” significa que no hay enfermedades del corazón.

Descripción de las variables:

- AHD: Diagnóstico de enfermedades del corazón (Yes,No).
- Age: Edad en años.
- Sex: Sexo del paciente (0 Femenino 1 Masculino).
- ChestPain: Tipo de dolor en el pecho (asymptomatic, nonanginal,nontypical, typical).
- RestBP: Presión Arterial en Reposo.
- Chol: Colesterol en sangre en mg/dl (25 a 200 mg/dL normal).
- Fbs: Azúcar en sangre (fasting blood sugar ¿ 120 mg/dl ;1 = SI; 0 = NO) Entre 100 y 125 mg/dL se diagnostica como prediabetes. 126 mg/dL o más en dos pruebas distintas se diagnostica como diabetes.
- RestECG: Resultados Electrocardiográficos en Reposo:
  - 0: normal
  - 1: tener anomalía en la onda ST
  - 2: mostrar hipertrofia ventricular izquierda
- MaxHR: Frecuencia cardíaca máxima alcanzada.
- ExAng: Angina de pecho inducida por el ejercicio (1=SI 0=N0).
- Oldpeak: Depresión sanguínea inducida por el ejercicio en relación con el reposo.

- Slope: Pendiente del segmento ST(Segmento del electrocardiograma entre la onda S y la T) de ejercicio máximo.
- 1: ascendente
- 2: plano
- 3: pendiente descendente
- Ca: Número de vasos principales (0-3) coloreados por fluoroscopia.
- Thal: 3 = normal, 6 = defecto fijo, 7 = defecto reversible.

## 1 Reto

1. Importa el DataSet Heart a R y se transforma las variables correspondientes a factor.
2. Calcula la estadística descriptiva del conjunto de datos con la función `summary()`, con la información obtenida responde las siguientes preguntas:
  - (a) ¿Cantidad de pacientes en el estudio por género?
  - (b) ¿Cual es el promedio de edad de los pacientes en el estudio?
  - (c) ¿De los 297 pacientes en el estudio cuantos de ellos no presentaron enfermedades en el corazón y cuantos si la presentaron?
3. Crea una gráfica de barras para los pacientes que presentaron enfermedades cardíacas, destacando en cada barra la proporción de pacientes hombres y mujeres mediante el uso de colores diferenciados.
4. Elabora una gráfica de barras para los pacientes que presentaron enfermedades cardíacas, separando las barras (utilizando la opción `position = "dodge"`) según el tipo de dolor en el pecho.
5. Crea una gráfica de cajas y bigotes para la frecuencia cardiaca máxima alcanzada, comparando a los pacientes que presentaron enfermedades cardiacas. Además, incluye la dispersión de las observaciones utilizando `geom_jitter()`.
6. Diseña un histograma con 7 barras para representar la distribución de edades de los pacientes.
7. Crea un histograma y un polígono de frecuencias con 5 barras para representar los niveles de colesterol en sangre de los pacientes.
8. Desarrolla una gráfica de dispersión que relacione la presión arterial y los niveles de colesterol en sangre de los pacientes.
9. Diseña una gráfica de dispersión que relacione la presión arterial y los niveles de colesterol en sangre, diferenciando los puntos por género mediante el uso de colores.

10. Elabora un conjunto de gráficas de dispersión que relacione la presión arterial y los niveles de colesterol en sangre, coloreando los puntos según la variable AHD y utilizando `facet_wrap()` para agregar como tercera variable los resultados del electrocardiograma.

**Nota:** A cada una de las gráficas se le deberá añadir un título y etiquetas en cada uno de los ejes.