**Vargas Carrillo Raul Paulino**

**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 m ́aximo.

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.

> HEART$AHD <- factor(HEART$AHD)

> HEART$Sex <- factor(HEART$Sex)

> HEART$ChestPain <- factor(HEART$ChestPain)

> HEART$RestECG <- factor(HEART$RestECG,

+ labels = c("Normal", "Anomalia", "Hipertrofia"))

> HEART$ExAng <- factor(HEART$ExAng,

+ labels = c("NO", "SÍ"))

> HEART$Slope <- factor(HEART$Slope,

+ labels = c("Ascendente", "Plano", "Descendente"))

> HEART$Thal <- factor(HEART$Thal)

> HEART$Ca <- factor(HEART$Ca)

> HEART$Fbs <- factor(HEART$Fbs)

> summary(HEART)

Age RestBP Chol MaxHR Oldpeak Ca

Min. :29.00 Min. : 94.0 Min. :126.0 Min. : 71.0 Min. :0.00 0 :176

1st Qu.:48.00 1st Qu.:120.0 1st Qu.:211.0 1st Qu.:133.5 1st Qu.:0.00 1 : 65

Median :56.00 Median :130.0 Median :241.0 Median :153.0 Median :0.80 2 : 38

Mean :54.44 Mean :131.7 Mean :246.7 Mean :149.6 Mean :1.04 3 : 20

3rd Qu.:61.00 3rd Qu.:140.0 3rd Qu.:275.0 3rd Qu.:166.0 3rd Qu.:1.60 NA's: 4

Max. :77.00 Max. :200.0 Max. :564.0 Max. :202.0 Max. :6.20

Sex ChestPain Fbs RestECG ExAng Slope

0: 97 asymptomatic:144 0:258 Normal :151 NO:204 Ascendente :142

1:206 nonanginal : 86 1: 45 Anomalia : 4 SÍ: 99 Plano :140

nontypical : 50 Hipertrofia:148 Descendente: 21

typical : 23

Thal AHD

fixed : 18 No :164

NA : 2 Yes:139

normal :166

reversable:117

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?

> table(HEART$Sex)

0 1

97 206

(b) ¿Cuál es el promedio de edad de los pacientes en el estudio?

> mean(HEART$Age)

[1] 54.43894

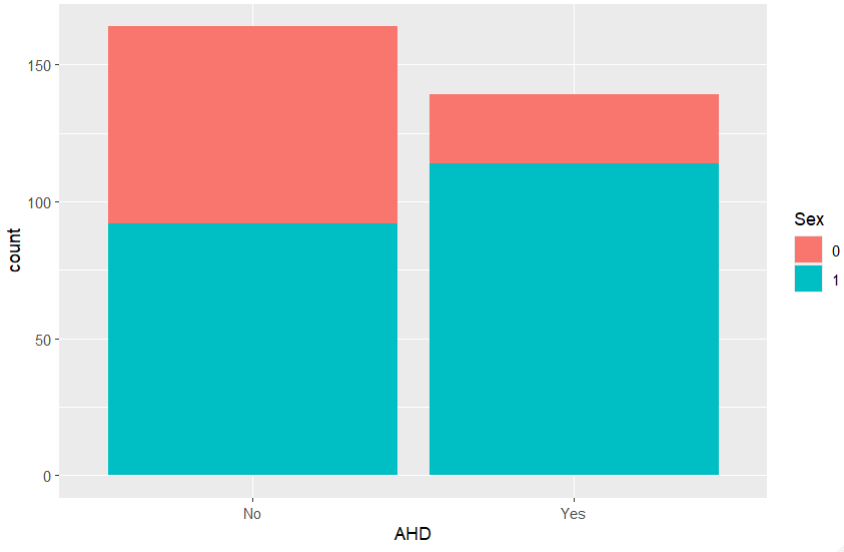
(c) ¿De los 297 pacientes en el estudio cuantos de ellos no presentaron enfermedades en el corazón y cuantos si la presentaron?

> table(HEART$AHD)

No Yes

164 139

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.

ggplot(HEART, aes(x = AHD, fill = Sex)) +  
 geom\_bar(position = "dodge")

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.

> ggplot(HEART, aes(x = ChestPain, fill = Sex)) +

+ geom\_bar(position = "dodge")

Gráfico, Gráfico de barras

Descripción generada automáticamente

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().

> HEART %>%

+ filter(AHD == "Yes") %>%

+ ggplot(aes(MaxHR, AHD)) +

+ geom\_boxplot(color = "black", fill = "red") +

+ geom\_jitter(color = "blue", alpha = 0.4)

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

6. Diseña un histograma con 7 barras para representar la distribución de edades de los pacientes.

> HEART %>%

+ ggplot(aes(Age)) +

+ geom\_histogram(color = "black", fill = "green", bins = 7)

Gráfico, Histograma

Descripción generada automáticamente

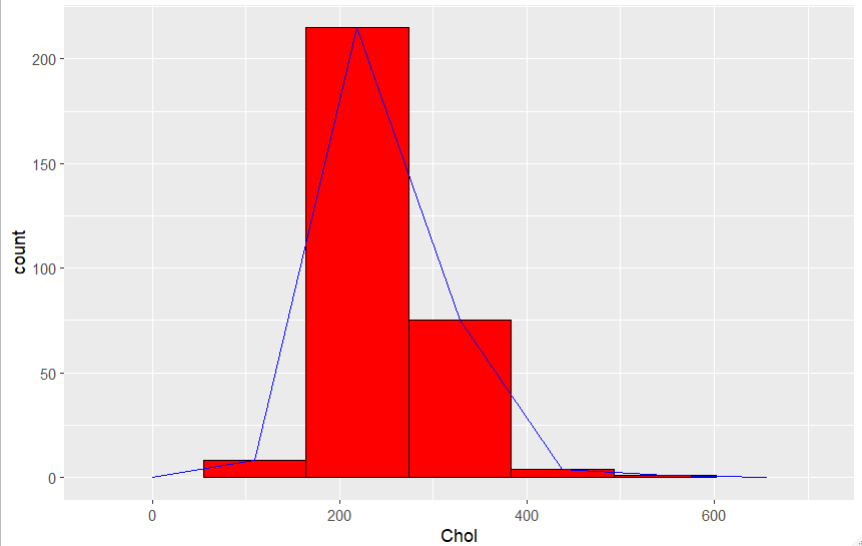
7. Crea un histograma y un polígono de frecuencias con 5 barras para representar los niveles de colesterol en sangre de los pacientes.

> HEART %>%

+ ggplot(aes(Chol)) +

+ geom\_histogram(color = "black", fill = "red", bins = 5) +

+ geom\_freqpoly(color = "blue", bins = 5)



8. Desarrolla una gráfica de dispersión que relacione la presión arterial y los niveles de colesterol en sangre de los pacientes.

> HEART %>%

+ ggplot(aes(RestBP, Chol, colour = Chol)) +

Gráfico, Gráfico de dispersión

Descripción generada automáticamente+ geom\_point()

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.

> HEART %>%

+ ggplot(aes(RestBP, Chol, colour = Sex)) +

+ geom\_point()

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

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.

> HEART %>%

+ ggplot(aes(RestBP, Chol, colour = AHD)) +

+ geom\_point() +

+ facet\_wrap(~RestECG)

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

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