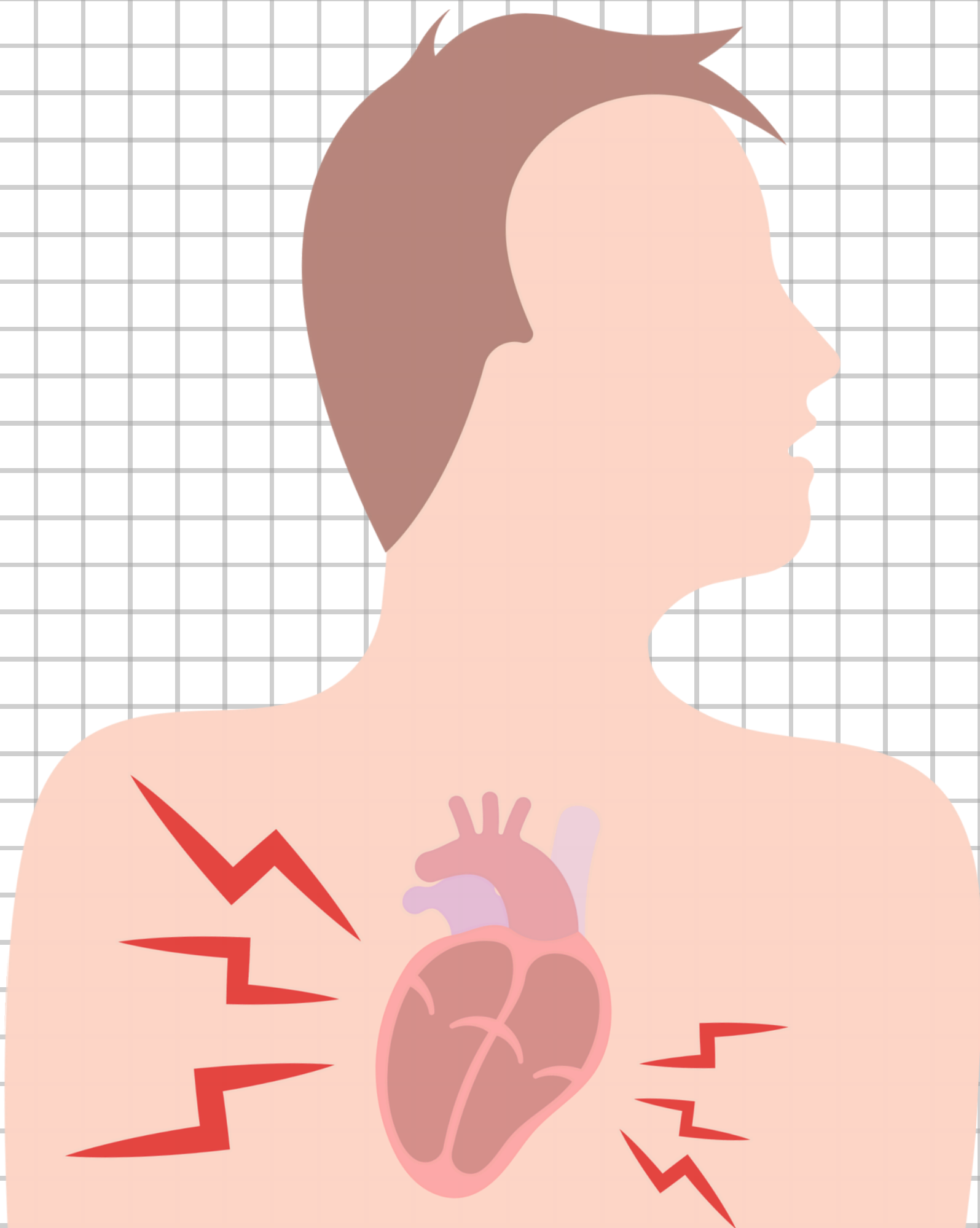


**COLESTEROL Y PRESIÓN
ARTERIAL EN PACIENTES
CON ENFERMEDADES
CARDIOVASCULARES A
PARTIR DE LOS 50 AÑOS:**

**Análisis Diferenciado por
Sexo**

Yatmelis Freites



ENFERMEDAD CARDIOVASCULAR

Enfermedad cardiovascular es un término general para una variedad de afecciones que afectan el sistema cardiovascular. Estas incluyen enfermedades de los vasos sanguíneos, enfermedades coronarias, arritmias cardíacas, defectos cardíacos congénitos y enfermedades de las válvulas cardíacas. Constituyen una de las principales causas de morbilidad y mortalidad a nivel mundial.

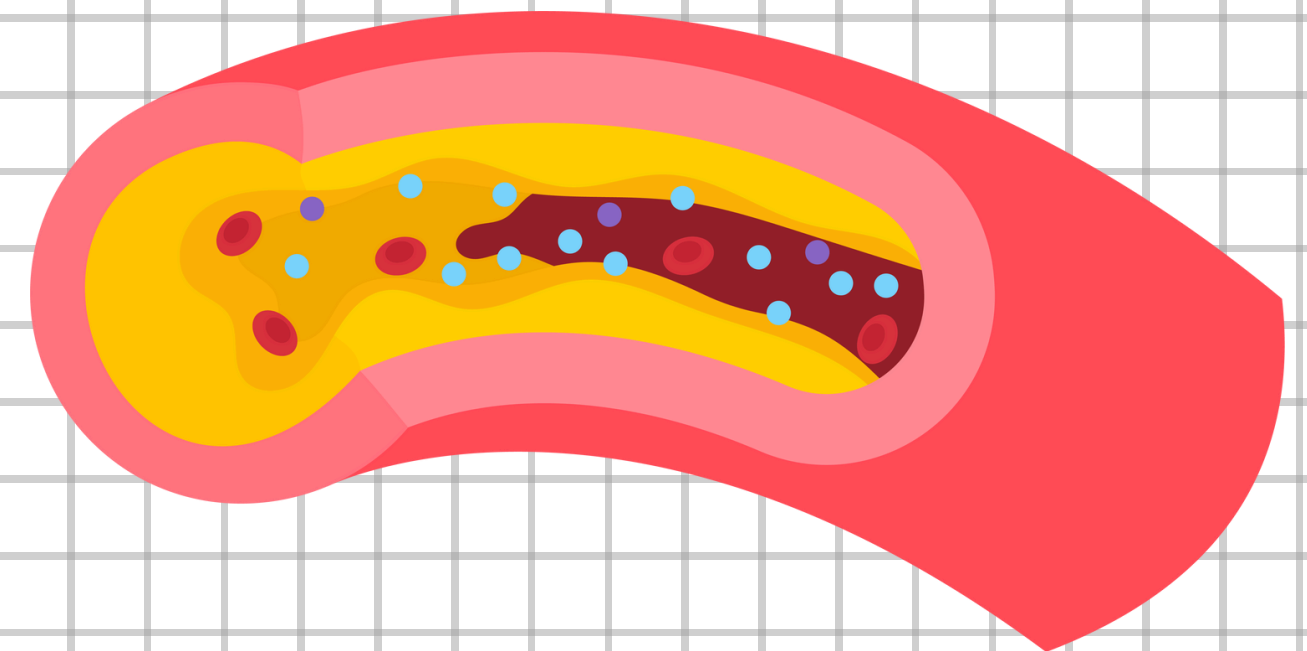
HIPERCOLESTEROLEMIA

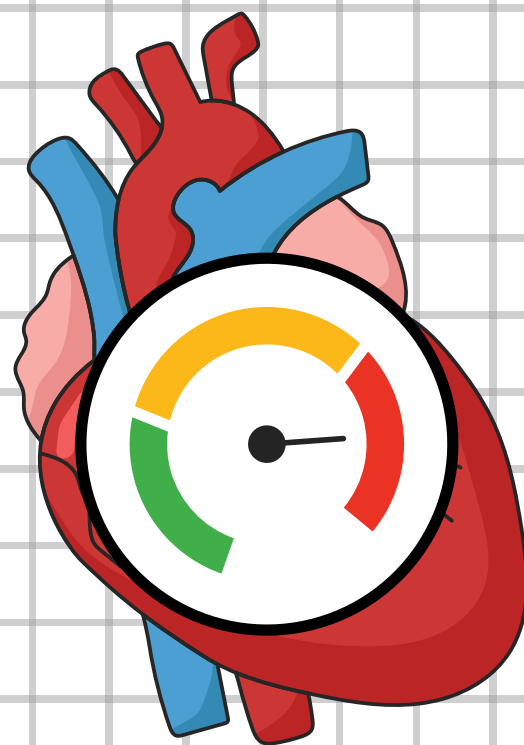
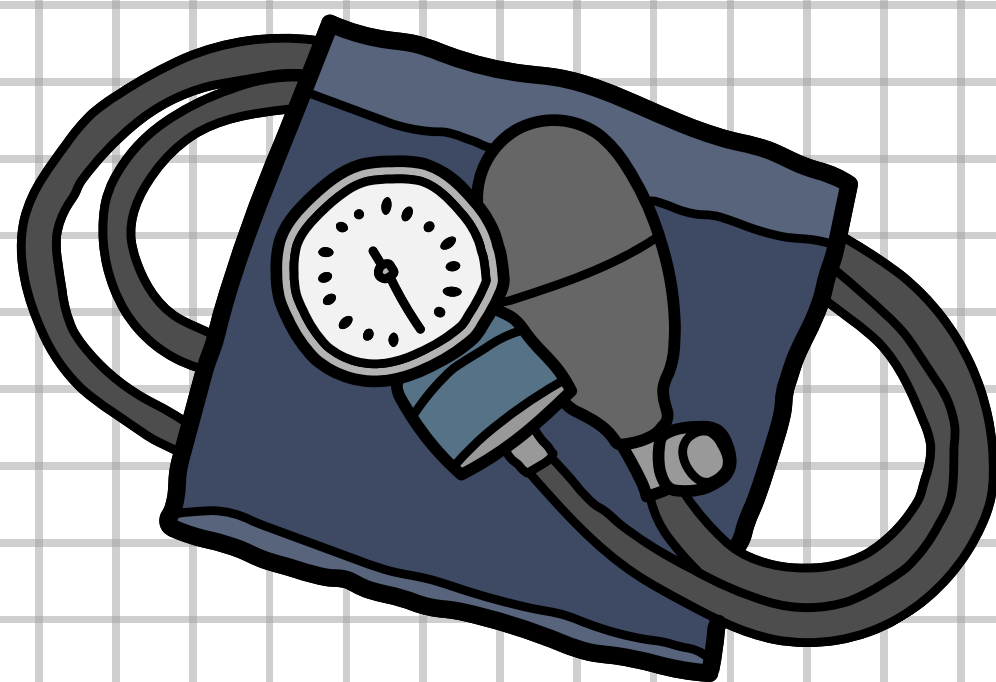
Consecuencia

Acumulación de placa de ateroma (placa compuesta de grasas y colesterol en las arterias), provoca el endurecimiento y estrechamiento de los vasos sanguíneos.

Definición

Colesterol alto en sangre, superior a 200 mg/dl





HIPERTENSIÓN ARTERIAL

Definición

Niveles en el rango entre 120 y 129 mm Hg en el caso de la sistólica y por debajo de 80 mm Hg en el de la diastólica

PA Sistólica

Fuerza que ejerce el flujo de sangre contra las paredes arteriales cuando la sangre se bombea fuera del corazón.

CONTEXTO DEL ESTUDIO

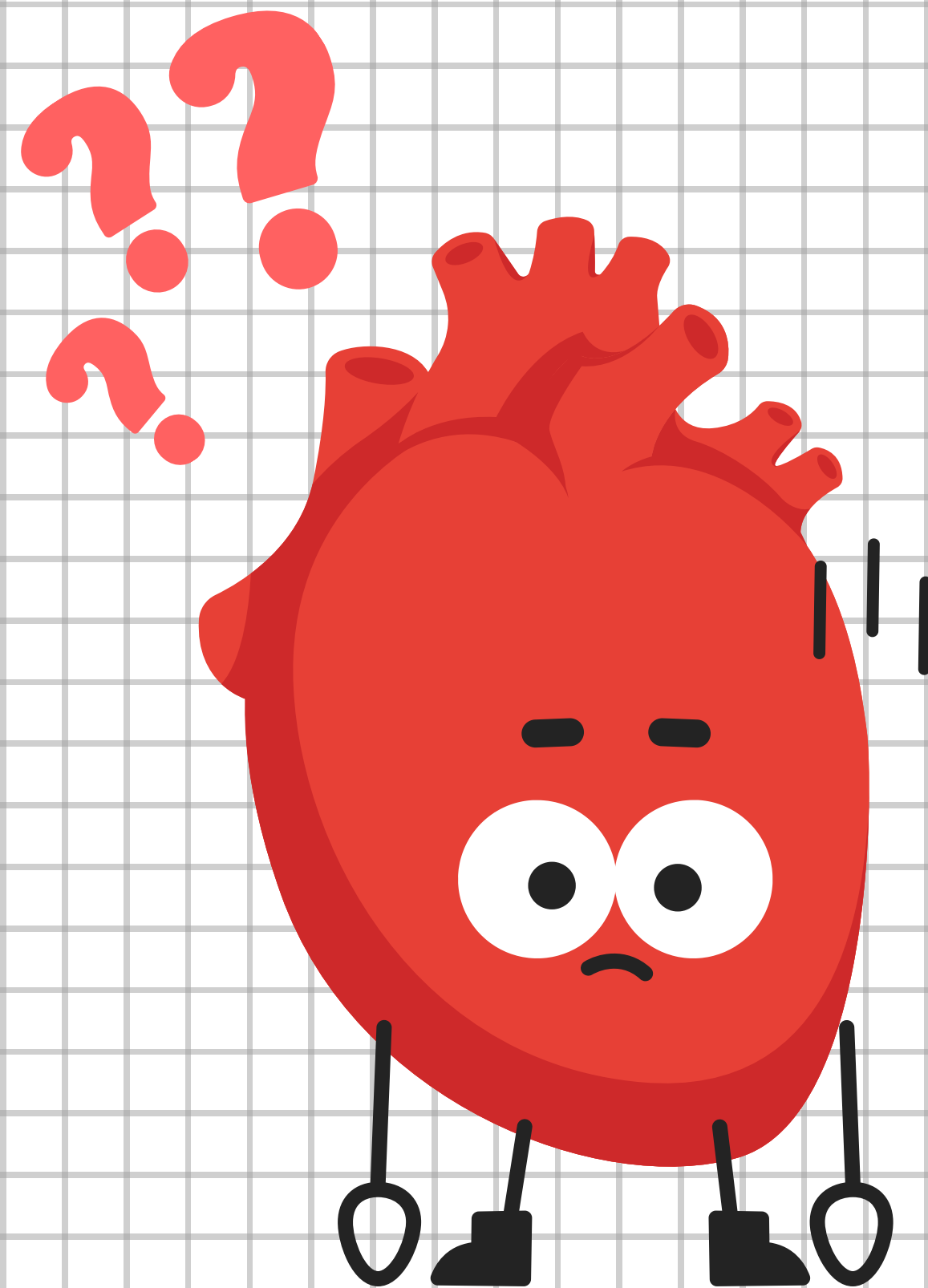
Las enfermedades cardiovasculares constituyen una de las principales causas de morbilidad y mortalidad a nivel mundial, especialmente en personas mayores de 50 años.

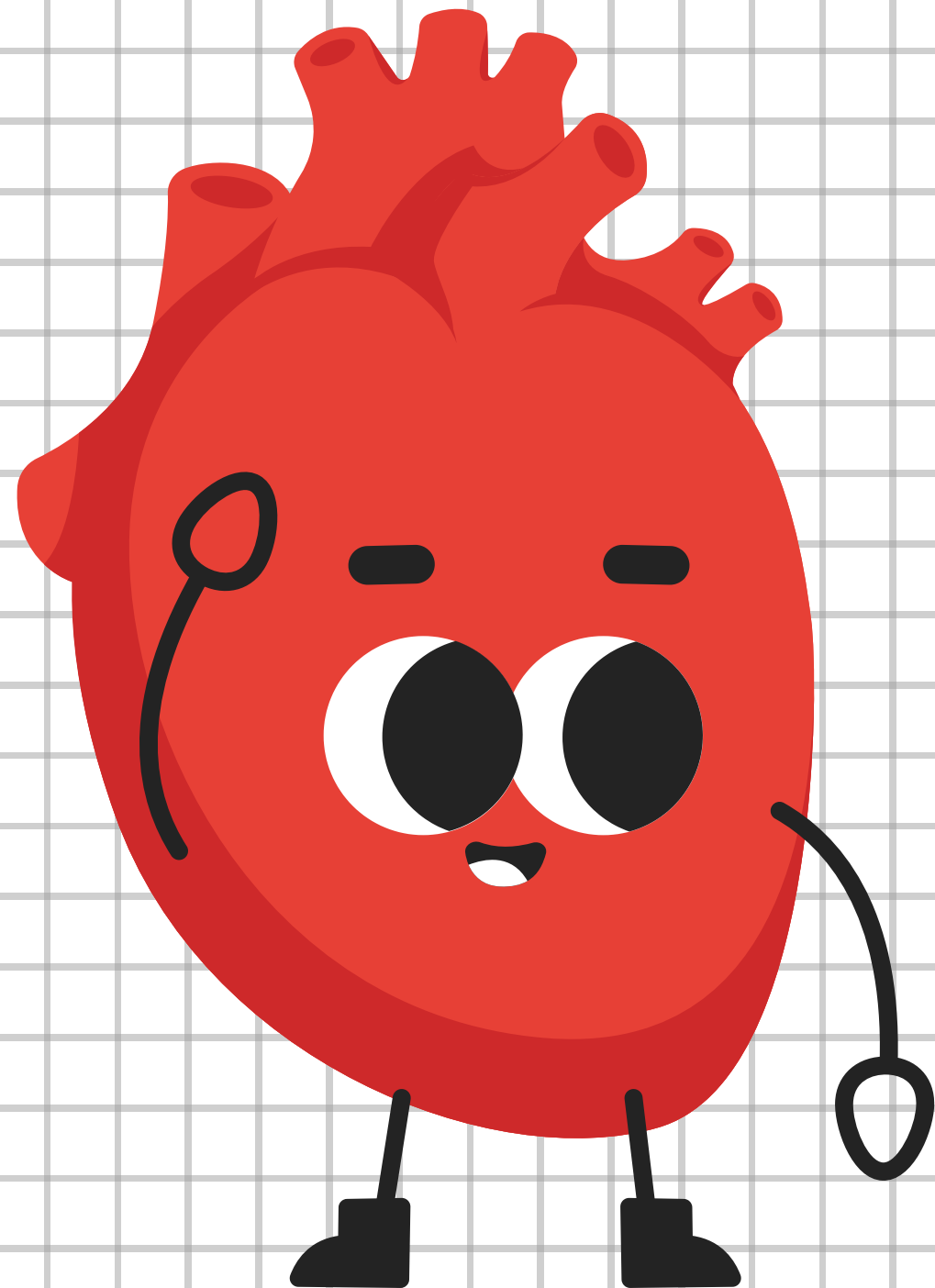
Dos factores de riesgo cruciales en el desarrollo de estas enfermedades son la hipercolesterolemia y la hipertensión arterial.

Tanto el riesgo cardiovascular como la complicación de estas enfermedades varían significativamente entre hombres y mujeres.

OBJETIVO

Analizar los valores de colesterol y la presión arterial en pacientes con enfermedades cardiovasculares mayores de 50 años, diferenciando por sexo, considerando las diferencias de riesgo y manifestación clínica entre ambos sexos.

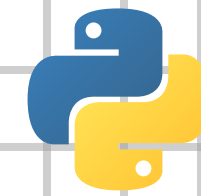




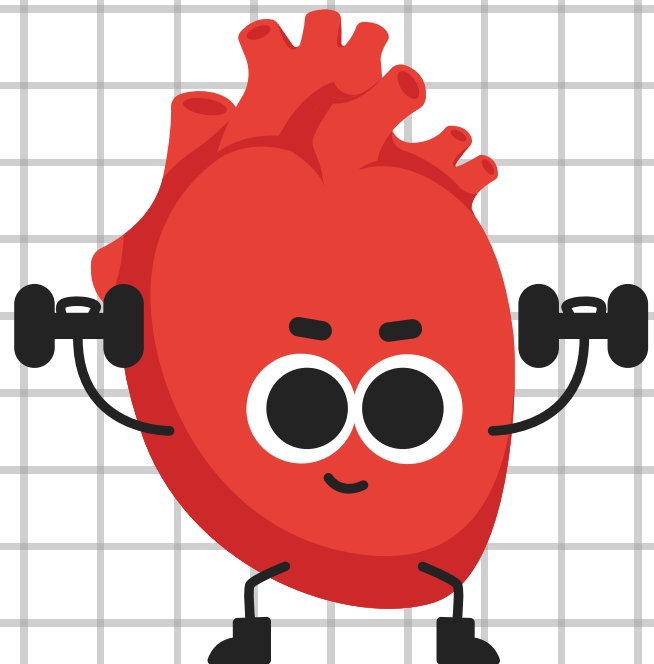
FUENTE DE DATOS

- Manu Siddhartha. (2020).
- Heart Disease Dataset (Comprehensive).
- IEEE Dataport.
- <https://dx.doi.org/10.21227/dz4t-cm36>

LIMPIEZA DE DATOS



- Cambio datos categóricos representados con números por palabras
- Renombro columnas para facilitar el análisis
- Elimino duplicados



```
heartdisease.drop_duplicates(inplace=True)
heartdisease.shape
```

```
(918, 12)
```

Data columns (total 12 columns):

#	Column	Non-Null Count	Dtype
0	age	1190 non-null	int64
1	sex	1190 non-null	int64
2	chest pain type	1190 non-null	int64
3	resting bp s	1190 non-null	int64
4	cholesterol	1190 non-null	int64
5	fasting blood sugar	1190 non-null	int64
6	resting ecg	1190 non-null	int64
7	max heart rate	1190 non-null	int64
8	exercise angina	1190 non-null	int64
9	oldpeak	1190 non-null	float64
10	ST slope	1190 non-null	int64
11	target	1190 non-null	int64

dtypes: float64(1), int64(11)
memory usage: 111.7 KB

Data columns (total 12 columns):

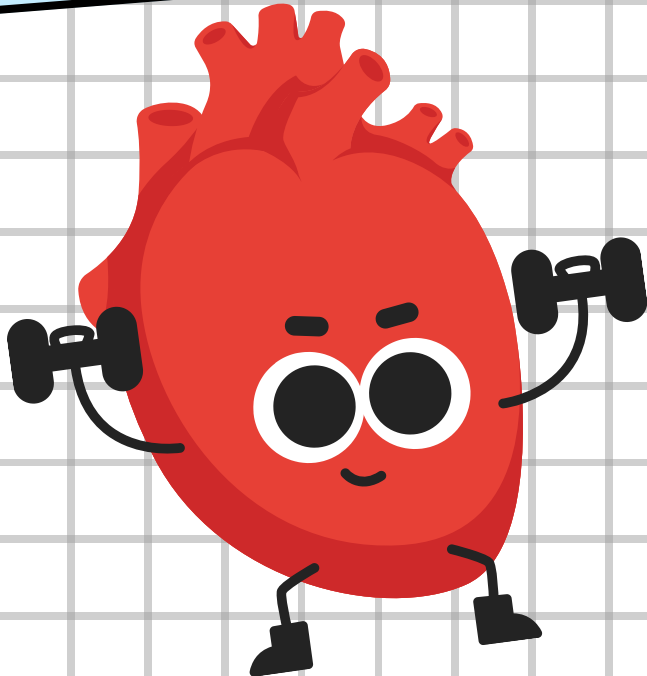
#	Column	Non-Null Count	Dtype
0	Edad	918 non-null	int64
1	Sexo	918 non-null	object
2	Tipo_Dolor_Pecho	918 non-null	object
3	Presion_Arterial_Reposo	918 non-null	int64
4	Colesterol	918 non-null	int64
5	Glicemia_Basal_Mayor_120	918 non-null	object
6	ECG_Reposo	918 non-null	object
7	Frecuencia_Cardiaca_Max	918 non-null	int64
8	Angina_Ejercicio	918 non-null	object
9	Depresion_ST	918 non-null	float64
10	Pendiente_ST	918 non-null	object
11	Estado	918 non-null	object

dtypes: float64(1), int64(4), object(7)
memory usage: 93.2+ KB

LIMPIEZA Y ANÁLISIS EXPLORATORIO DE DATOS



- Realizo filtro para delimitar población de estudio según el objetivo
- Realizo un describe para conocer estadísticas de datos numéricos
- Hago un filtro para solo incluir valores de colesterol mayores a 120 mg/DL

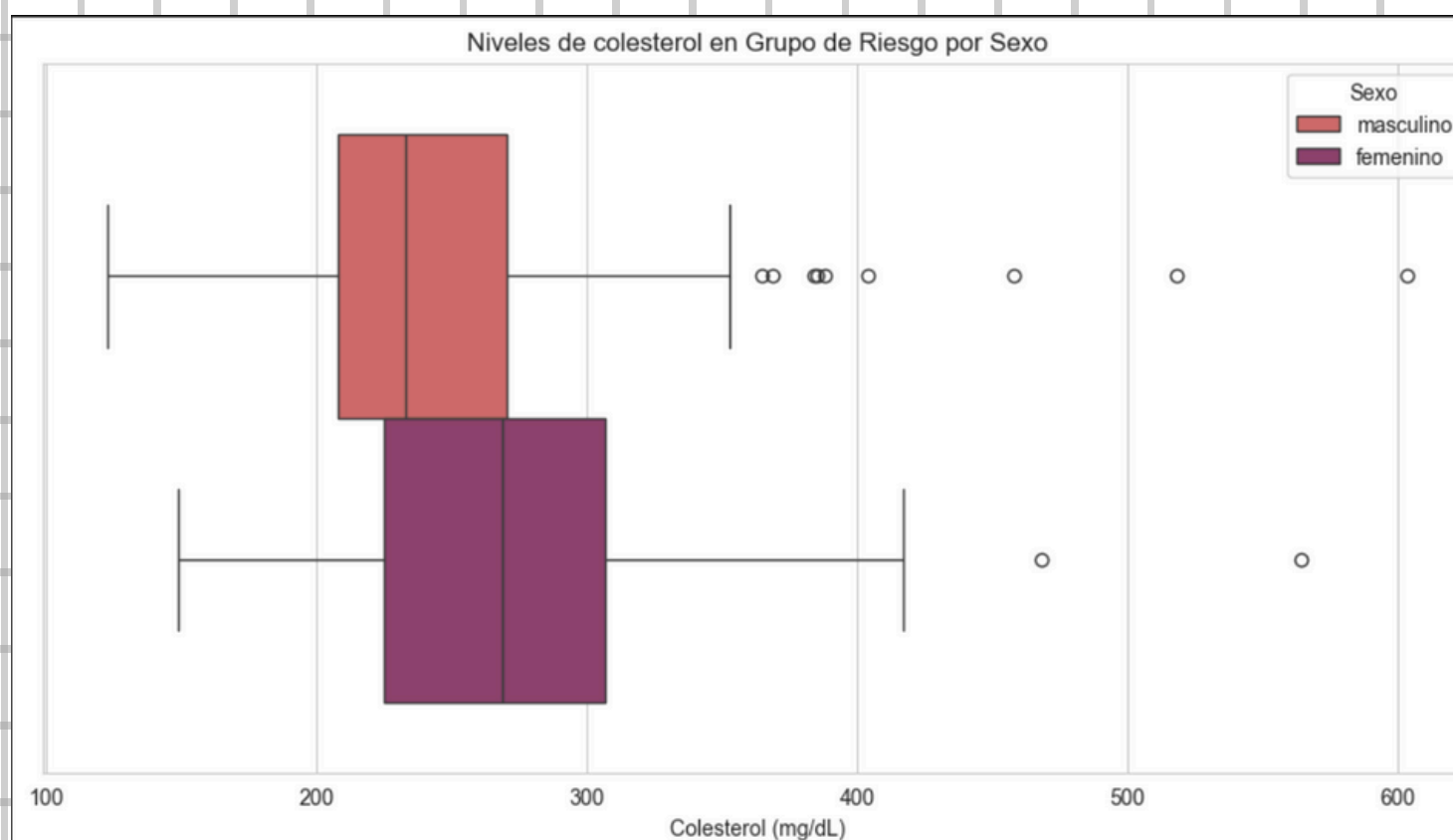


```
grupo_de_riesgo_df = heartdisease[((heartdisease['Sexo'] == 'femenino') & (heartdisease['Edad'] >= 50)) |  
                                   ((heartdisease['Sexo'] == 'masculino') & (heartdisease['Edad'] >= 50))]  
  
grupo_de_riesgo_df.head()
```

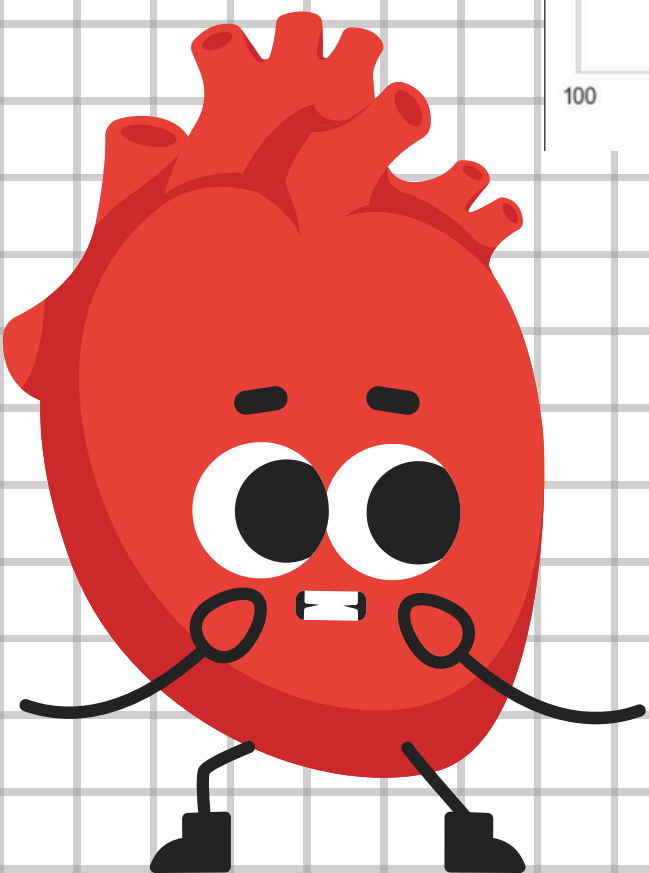
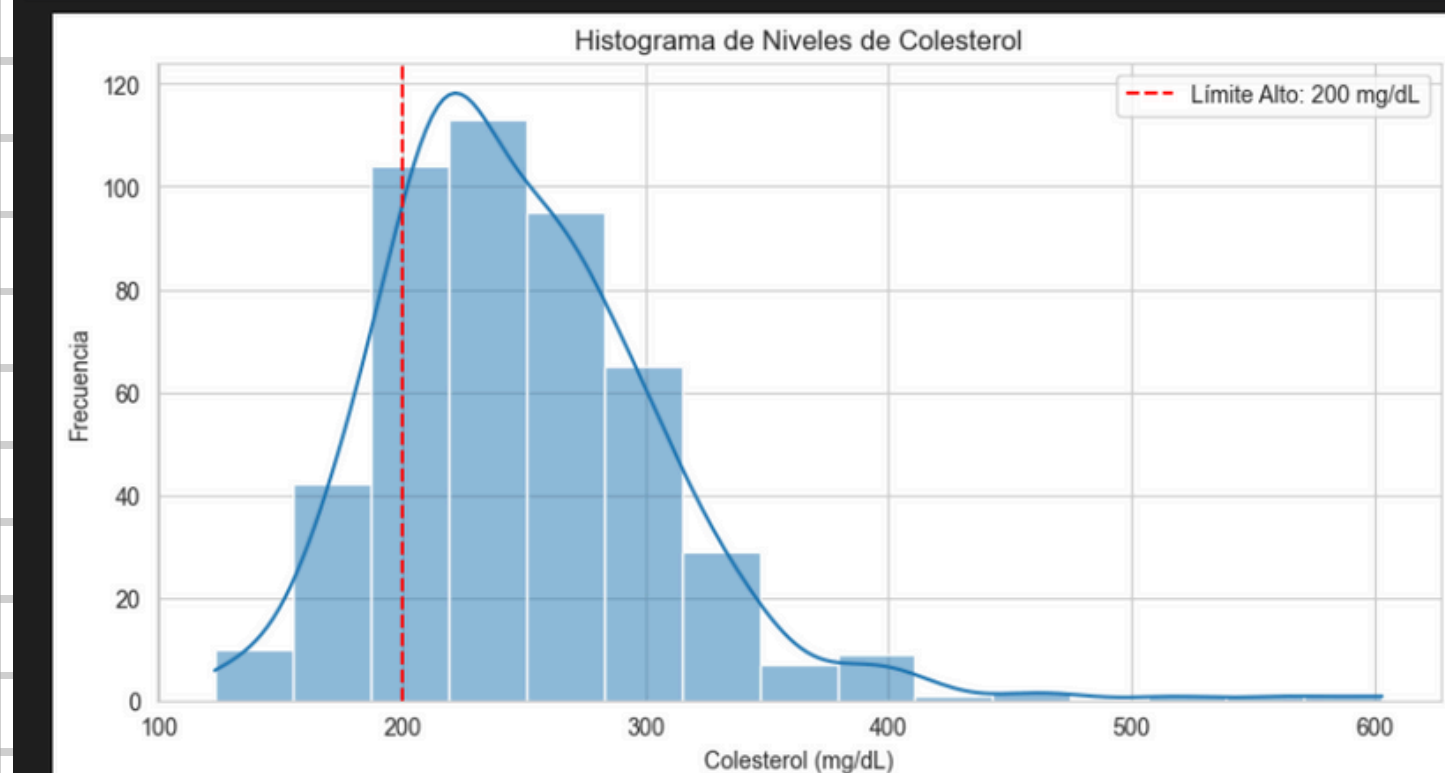
```
grupo_de_riesgo_df = grupo_de_riesgo_df[grupo_de_riesgo_df['Colesterol'] > 120]  
grupo_de_riesgo_df.describe()
```

	Edad	Presion_Arterial_Reposo	Colesterol	Frecuencia_Cardiaca_Max	Depresion_ST
count	480.000000	480.000000	480.000000	480.000000	480.000000
mean	58.479167	136.179167	249.050000	134.825000	1.093542
std	5.885439	17.375900	59.268625	23.018523	1.116782
min	50.000000	94.000000	123.000000	69.000000	0.000000
25%	54.000000	125.000000	211.000000	118.000000	0.000000
50%	58.000000	135.000000	240.000000	136.000000	1.000000
75%	62.000000	145.000000	282.000000	152.000000	1.800000
max	77.000000	200.000000	603.000000	195.000000	6.200000

OUTLIERS



```
plt.figure(figsize=(10, 5))
sns.histplot(grupo_de_riesgo_df['Colesterol'], bins=15, kde=True)
plt.title('Histograma de Niveles de Colesterol')
plt.xlabel('Colesterol (mg/dL)')
plt.ylabel('Frecuencia')
plt.axvline(x=200, color='r', linestyle='--', label='Límite Alto: 200 mg/dL')
plt.legend()
plt.show()
```



OUTLIERS



- Calculo la asimetría con Pandas. Ya que es mayor que 1, se considera fuertemente sesgada:
- Un valor positivo indica sesgo a la derecha (colas largas en el extremo superior), tal y como se observa en el histograma.

```
asimetria = grupo_de_riesgo_df['Colesterol']  
  
# Calcular asimetría con pandas  
skewness_pd = asimetria.skew()  
print("Asimetría:", skewness_pd)
```

Asimetría: 1.392750523951329

- Realizo la Prueba de Shapiro-Wilk para comprobar hipótesis nula
- El resultado es mucho menor que el nivel de significancia común de 0.05. Esto significa que es muy poco probable que estos datos sigan una distribución normal.

```
# Prueba de Shapiro-Wilk  
stat, p_value = shapiro(asimetria)  
print("Shapiro-Wilk Test:")  
print("Statistic:", stat, "P-value:", p_value)
```

```
[23]  
... Shapiro-Wilk Test:  
Statistic: 0.9250930953630504 P-value: 9.728515075658714e-15
```



MANEJO DE OUTLIERS

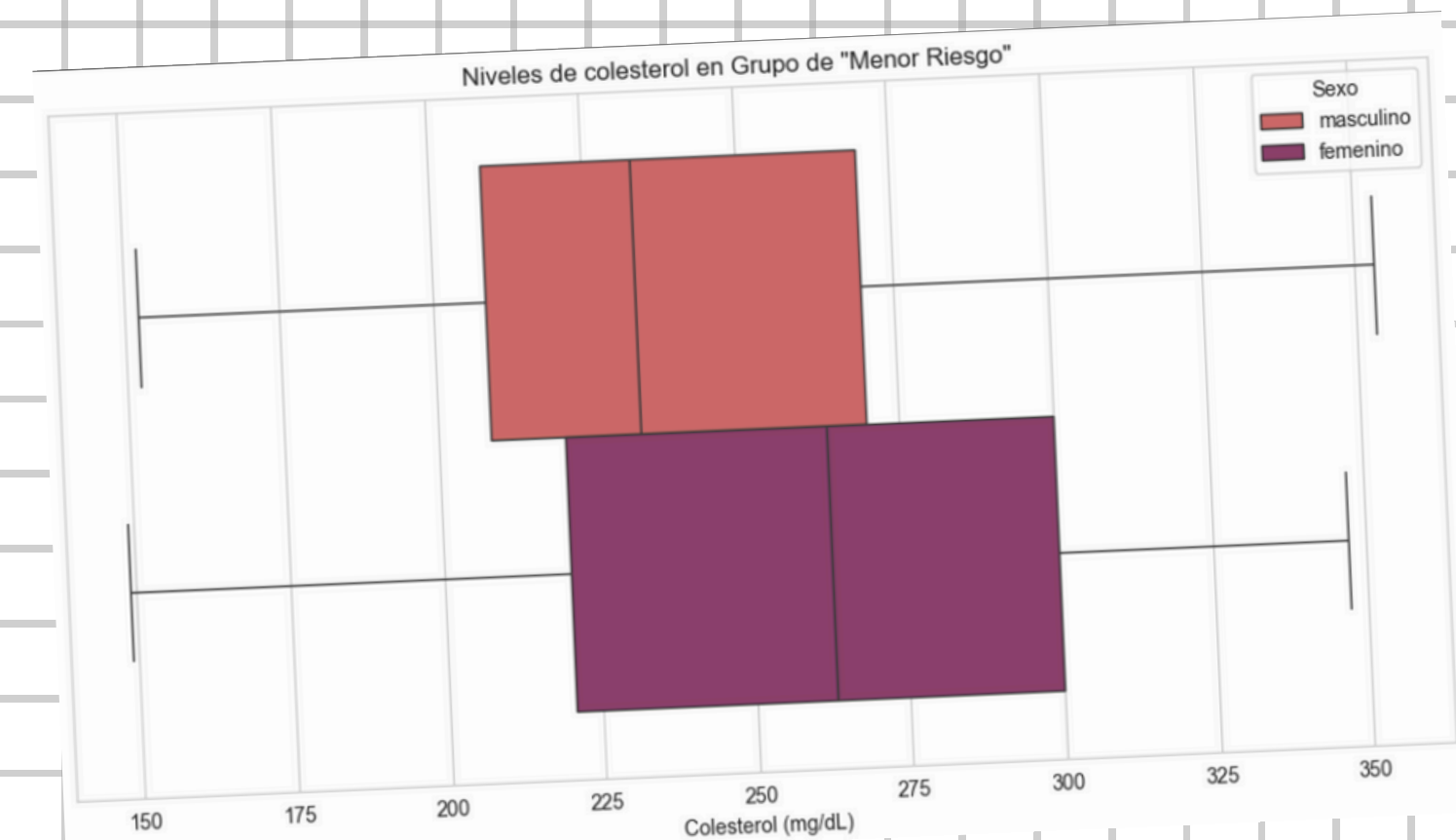
Se ajusta el IQR a 1.0 en lugar de 1.5 porque se trabaja con datos altamente variables y porque en este dataset existen valores extremos que no debo eliminar del estudio ya que siguen teniendo sentido e importancia. Este ajuste permite filtrar datos más específicos y reducir el número de outliers considerados, que podría ser apropiado para conjuntos de datos no normalmente distribuidos o con distribuciones altamente sesgadas, como es el caso.



```
[24] Q1 = grupo_de_riesgo_df['Colesterol'].quantile(0.25)
      Q3 = grupo_de_riesgo_df['Colesterol'].quantile(0.75)
      IQR = Q3 - Q1
      limite_inferior = Q1 - 1.0 * IQR
      limite_superior = Q3 + 1.0 * IQR

[25] grupo_sin_outliers = grupo_de_riesgo_df[(grupo_de_riesgo_df['Colesterol'] >= limite_inferior) &
      (grupo_de_riesgo_df['Colesterol'] <= limite_superior)]
      grupo_sin_outliers['Colesterol'].describe()
```

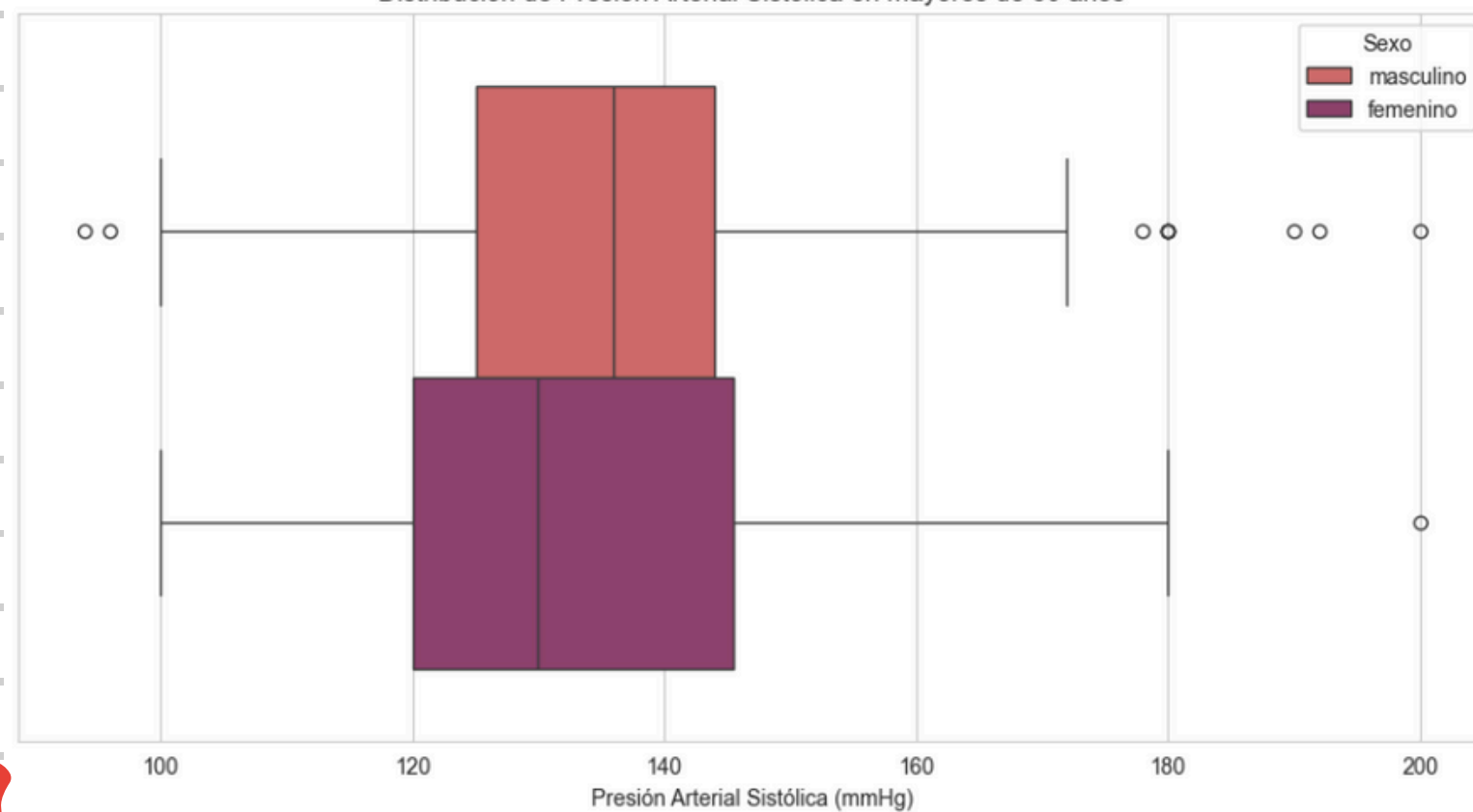
...	count	454.000000
	mean	243.605727
	std	44.804270
	min	149.000000
	25%	211.000000
	50%	238.500000
	75%	274.750000
	max	353.000000
	Name: Colesterol, dtype: float64	



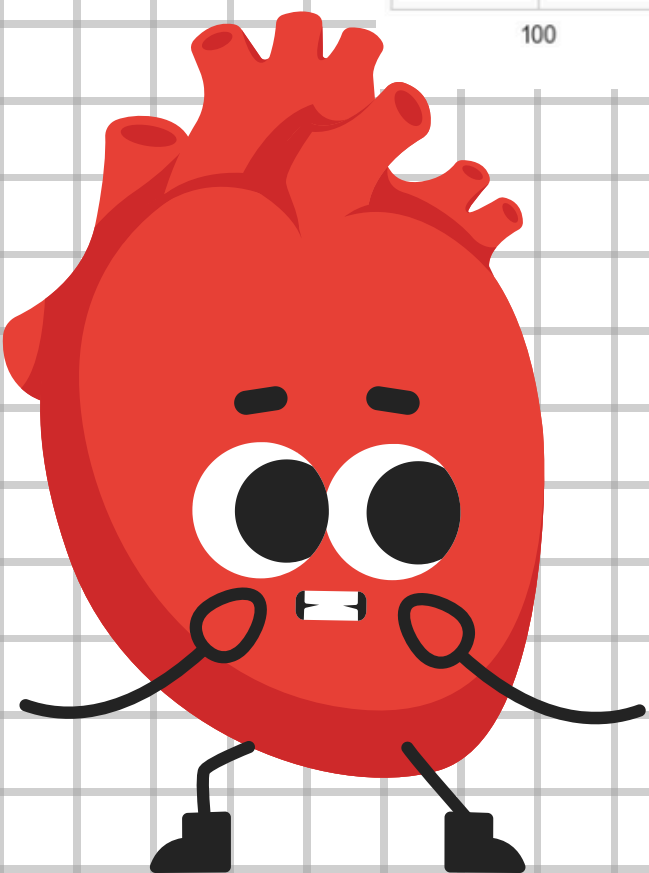
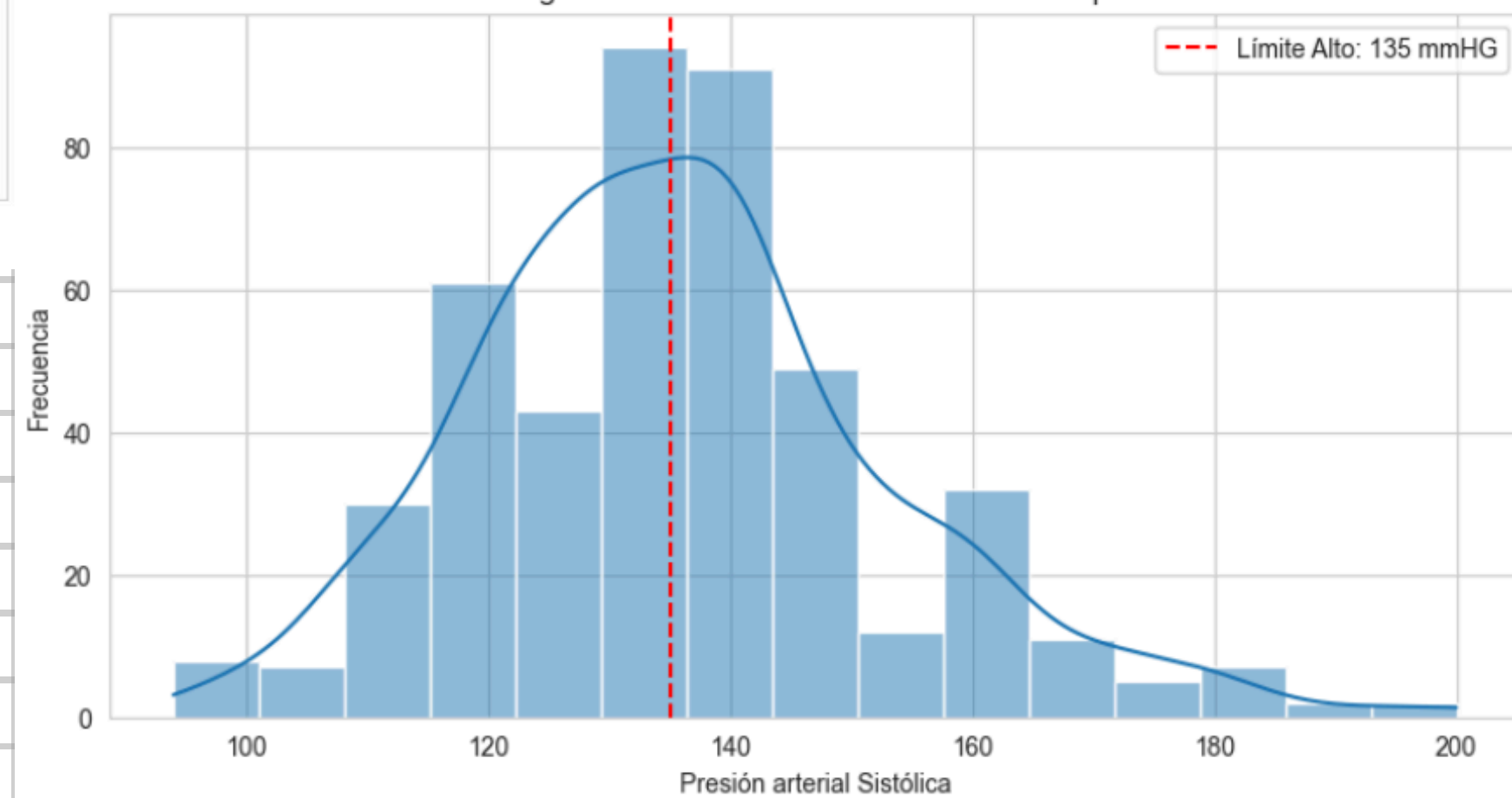
OUTLIERS



Distribución de Presión Arterial Sistólica en mayores de 50 años



Histograma de Presión Arterial Sistólica en Repoco



OUTLIERS

- Calculo la asimetría con Pandas.
- La asimetría no es extremadamente pronunciada

```
Q1 = grupo_de_riesgo_df['Presion_Arterial_Reposo'].quantile(0.25)
Q3 = grupo_de_riesgo_df['Presion_Arterial_Reposo'].quantile(0.75)
IQR = Q3 - Q1
limite_inferior = Q1 - 1.4 * IQR
limite_superior = Q3 + 1.4 * IQR

grupo_sin_outliers = grupo_de_riesgo_df[(grupo_de_riesgo_df['Presion_Arterial_Reposo'] >= limite_inferior) &
                                         (grupo_de_riesgo_df['Presion_Arterial_Reposo'] <= limite_superior)]
grupo_sin_outliers['Presion_Arterial_Reposo'].describe()
```

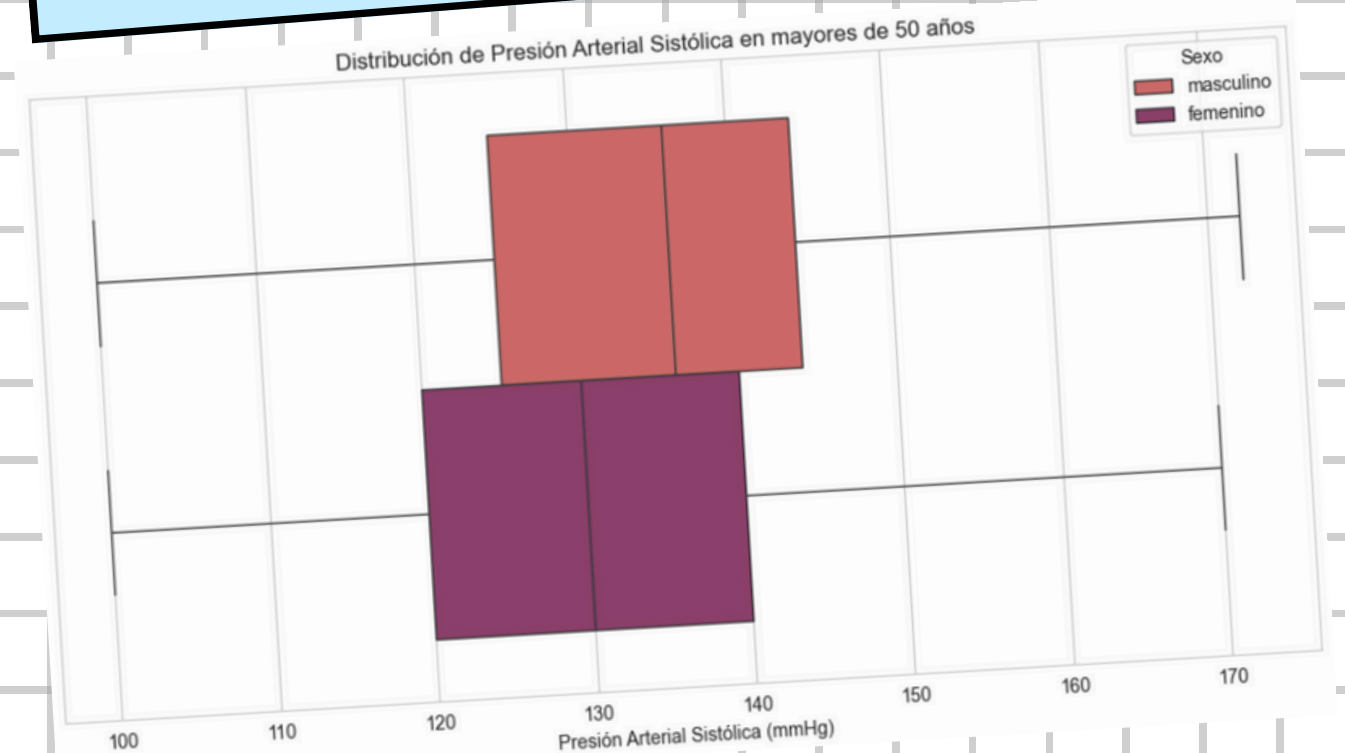
```
count    463.000000
mean     134.825054
std       15.142100
min      100.000000
25%      124.000000
50%      135.000000
75%      143.000000
max      172.000000
Name: Presion_Arterial_Reposo, dtype: float64
```

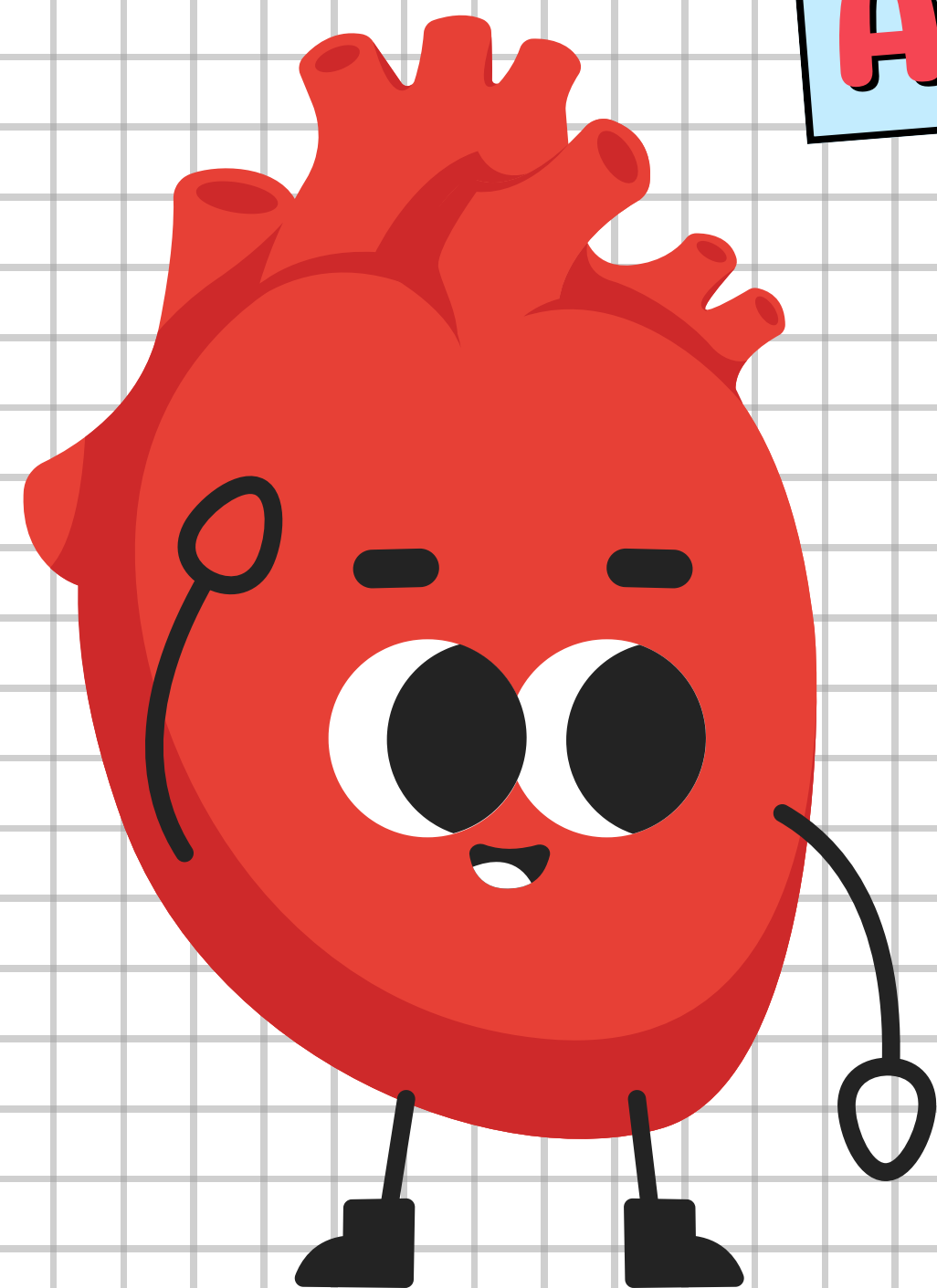
```
# análisis de asimetría de Presión Arterial Sistólica en Reposo
data = grupo_sin_outliers["Presion_Arterial_Reposo"]

skewness_pd = data.skew()
print("Asimetría:", skewness_pd)
```

[31]
... Asimetría: 0.5965969557724164

- Se ajusta el IQR a 1.4 en lugar de 1.5





ANÁLISIS DE CORRELACIÓN

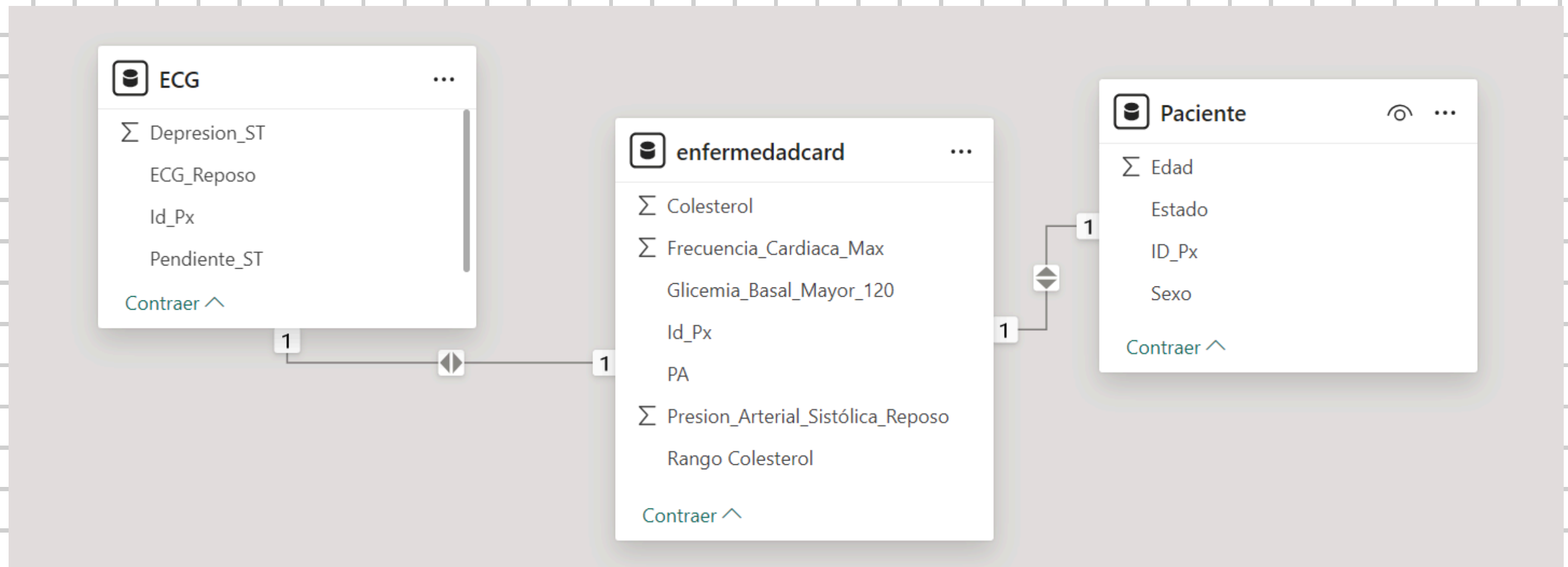
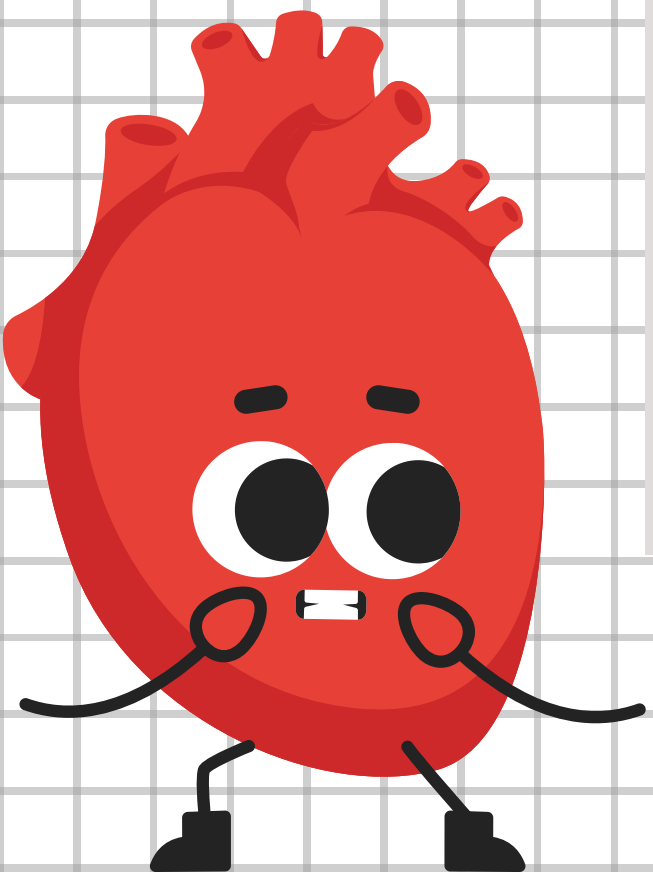
```
# Calcular la correlación y el valor p entre presión arterial y colesterol
correlation, p_value = pearsonr(grupo_sin_outliers['Presion_Arterial_Reposo'], grupo_sin_outliers['Colesterol'])

print("Correlación entre Presión Arterial y Colesterol:", correlation)
print("Valor p:", p_value)
```

Correlación entre Presión Arterial y Colesterol: 0.005436228987350045
Valor p: 0.9071307824686339

- Se observa una relación débil entre el nivel de colesterol y la presión Arterial sistólica en reposo.
- El valor p indica que la correlación no es estadísticamente significativa.

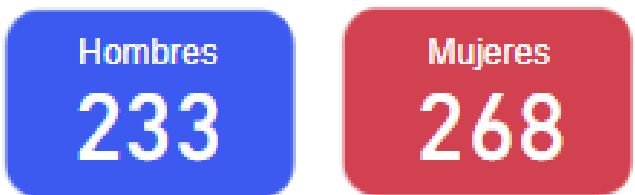
MODELADO DE DATOS



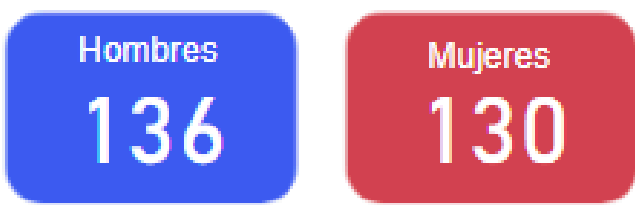
DASHBOARD

Colesterol y Presión Arterial Sistólica en Pacientes con Enfermedades Cardiovasculares a Partir de los 50 Años: Un Análisis Diferenciado por Sexo

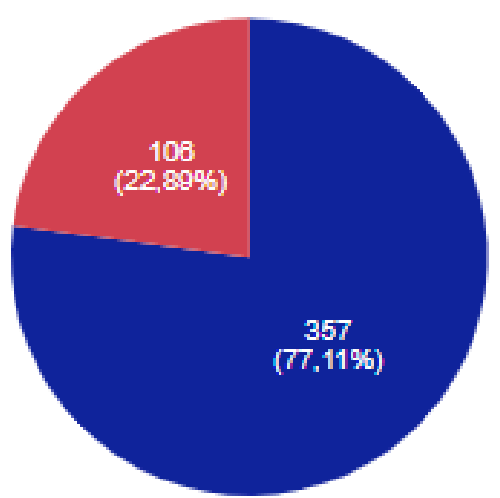
Mediana de Colesterol (mg/dl)



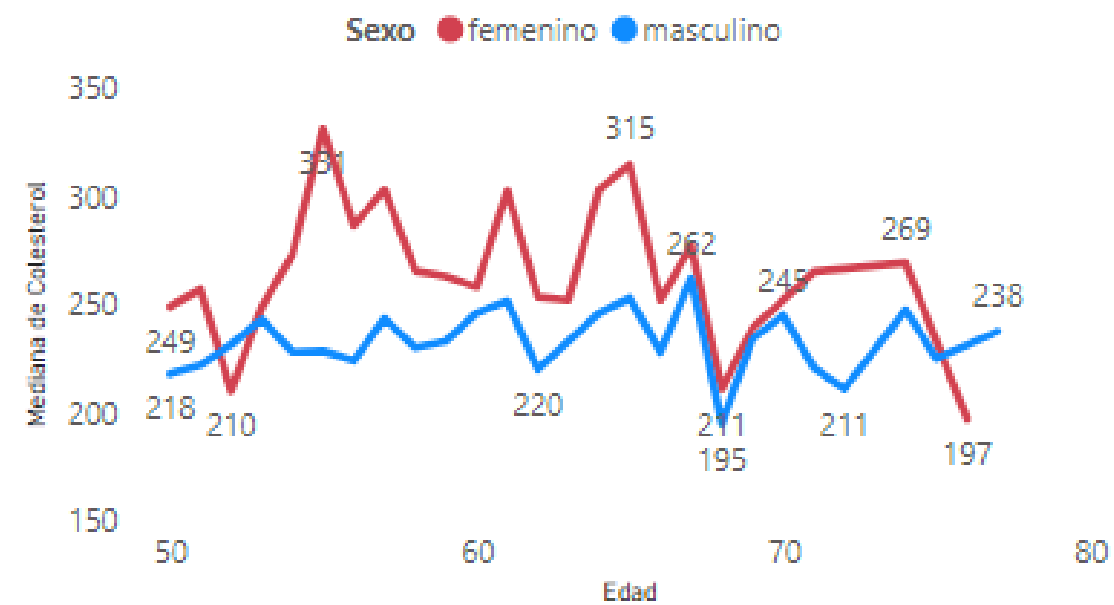
Mediana de PA Sistólica (mmHg)



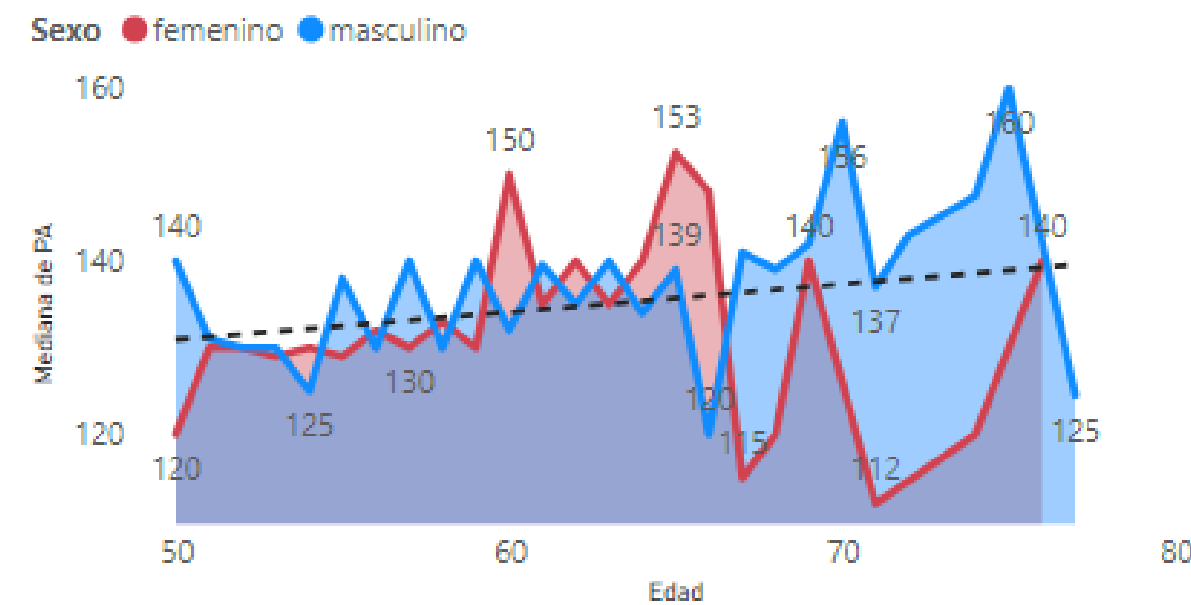
Recuento de pacientes por Sexo



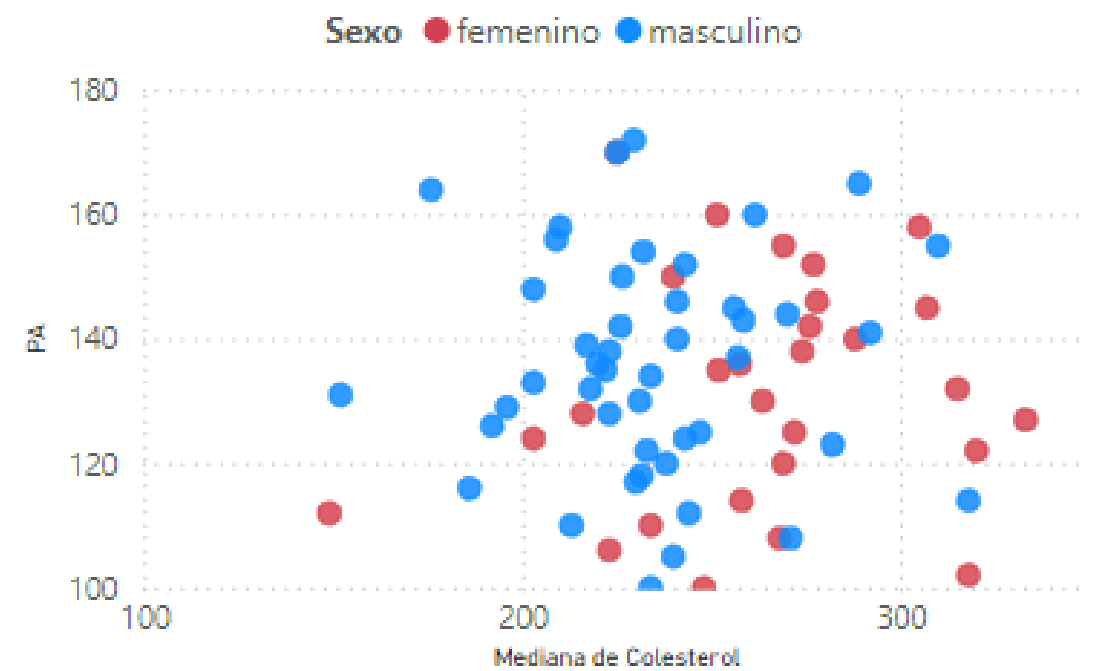
Mediana de Colesterol por edad y Sexo



Mediana de PA Sistólica por edad y Sexo



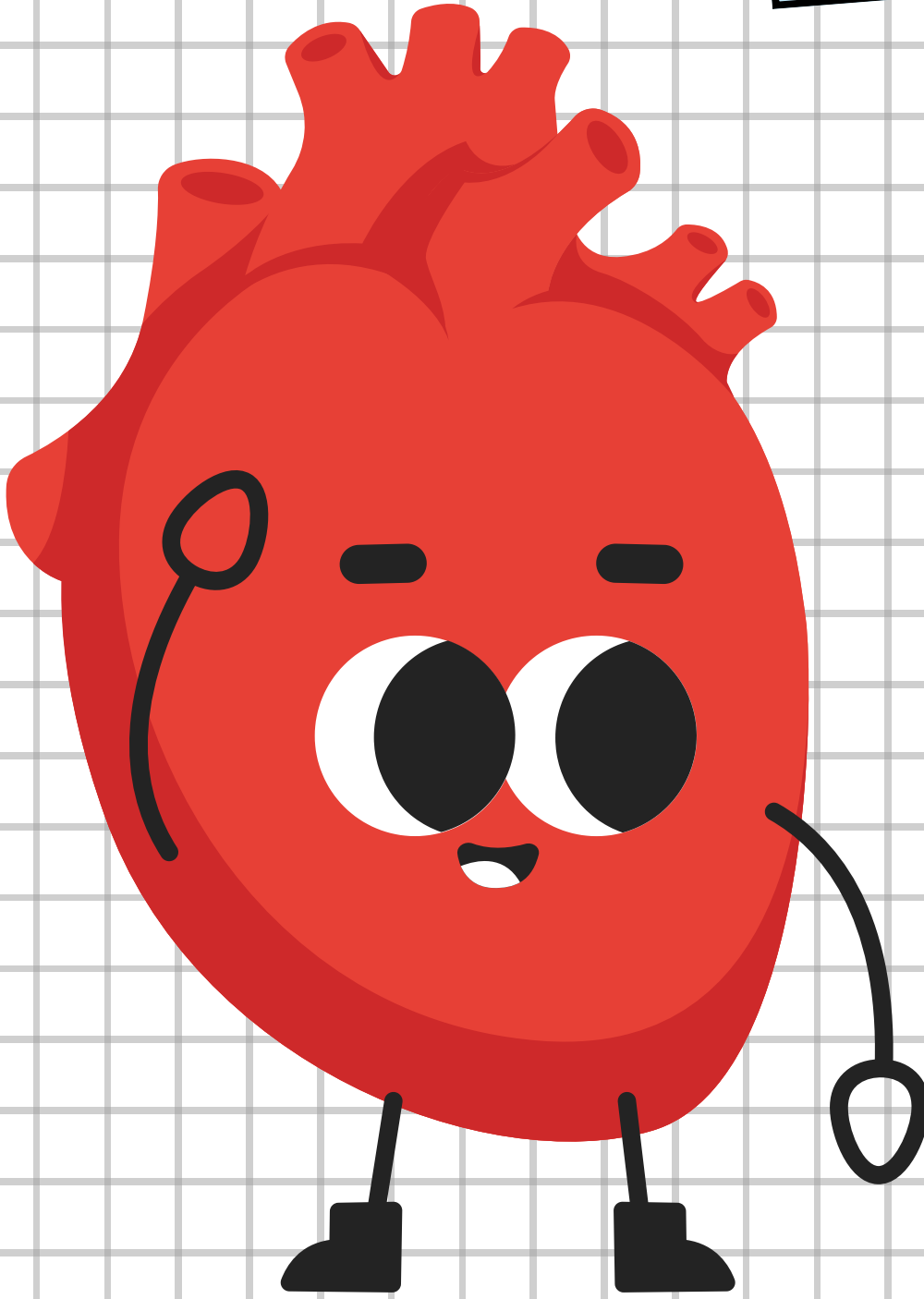
Mediana de Colesterol por Sexo y PA



Indicador de Riesgo por paciente

Edad	Estado	Sexo	Rango Colesterol	PA
69	diagnosticado	masculino	Peligroso	HIPERTENSION NIV
69	sano	masculino	Peligroso	HIPERTENSION NIV
70	diagnosticado	masculino	Peligroso	HIPERTENSION NIV
70	diagnosticado	masculino	Peligroso	HIPERTENSION NIV
70	sano	masculino	Peligroso	HIPERTENSION NIV
71	sano	femenino	Peligroso	HIPERTENSION NIV
71	sano	femenino	Peligroso	Normal
74	diagnosticado	masculino	Peligroso	HIPERTENSION NIV
74	sano	femenino	Peligroso	Elevada
75	sano	masculino	Peligroso	HIPERTENSION NIV
77	diagnosticado	masculino	Peligroso	Elevada





ANÁLISIS DE DASHBOARD

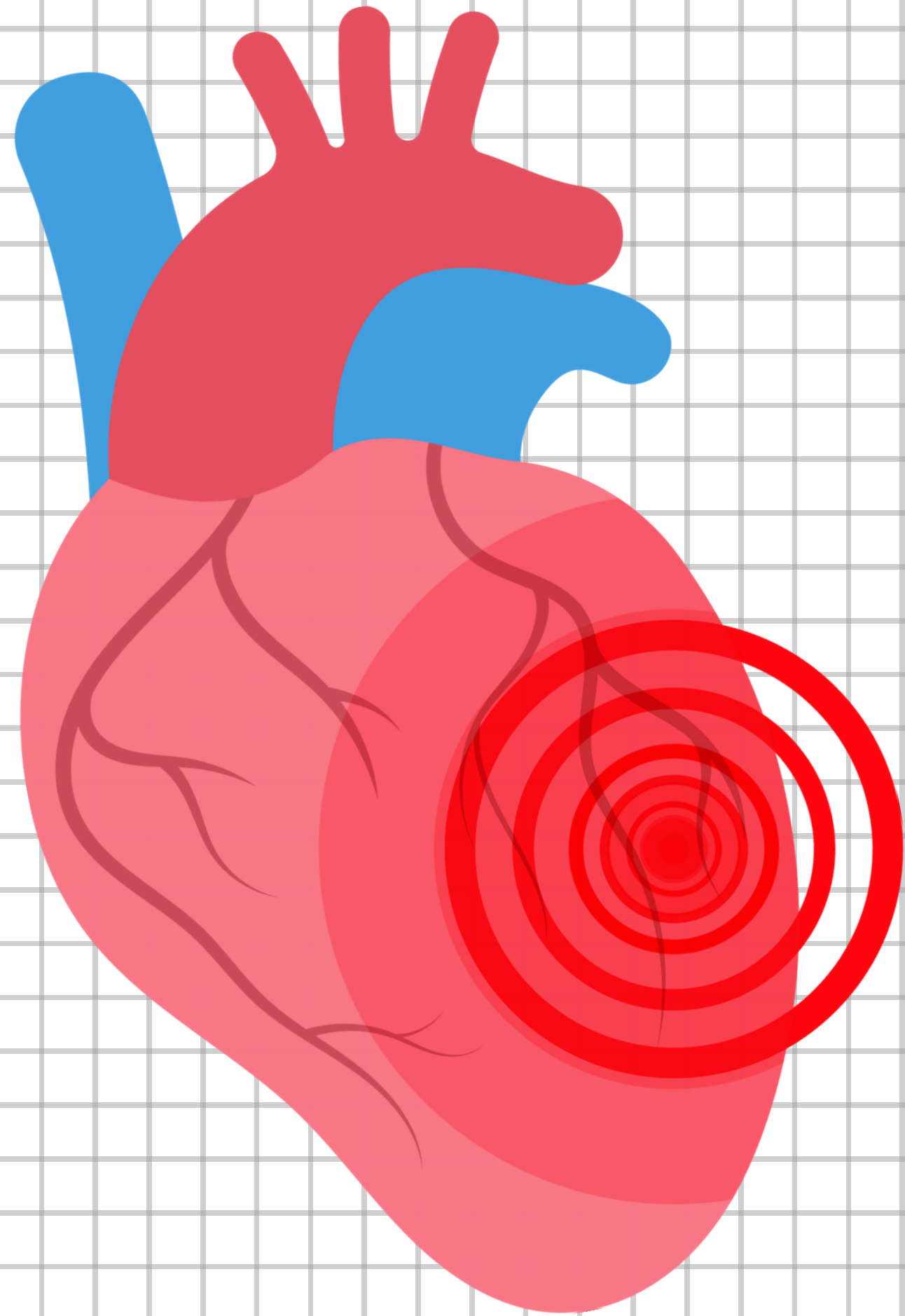
- Diferencias en los niveles de colesterol: Las mujeres presentan niveles de colesterol total más altos que los hombres, especialmente después de los 53 años.
- Diferencias en la presión arterial: Los hombres muestran una presión arterial ligeramente mayor que las mujeres, con un aumento progresivo después de los 63 años.
- Relación entre colesterol y presión arterial: Aunque no se observa una correlación fuerte, existe una ligera tendencia a que a mayor nivel de colesterol, mayor sea la presión arterial, tanto en hombres como en mujeres.
- Importancia del análisis diferenciado por sexo: Las diferencias observadas entre hombres y mujeres resaltan la importancia de realizar análisis diferenciados por sexo en la investigación cardiovascular.

CONCLUSIONES

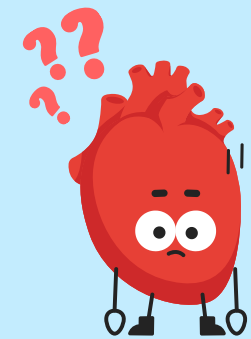
Las mujeres posmenopáusicas podrían tener un mayor riesgo cardiovascular debido a los niveles más altos de colesterol. Es fundamental que los profesionales de la salud presten especial atención a este grupo, promoviendo un estilo de vida saludable y un control regular de los niveles de colesterol.

Los hombres mayores de 63 años también requieren un seguimiento cercano de la presión arterial, ya que el aumento progresivo de la presión arterial en este grupo puede aumentar su riesgo de eventos cardiovasculares.

El análisis diferenciado por sexo es crucial para comprender las complejas interacciones entre los factores de riesgo cardiovascular. Esto permite desarrollar estrategias de prevención y tratamiento más efectivas y personalizadas, con el objetivo de mejorar la salud cardiovascular en hombres y mujeres mayores de 50 años.



**MUCHAS
GRACIAS**



Preguntas