Algunos otros hallazgos:

Se han generado histogramas y gráficos de densidad para las variables numéricas clave.

* Algunas observaciones importantes:

Edad: La distribución es aproximadamente normal con una concentración mayor entre los 40 y 60 años.

Cigarrillos por día: La mayoría de los participantes no fuman o fuman una cantidad baja de cigarrillos.

Colesterol total (totChol): La distribución muestra una ligera asimetría hacia la derecha.

Presión arterial sistólica (sysBP) y diastólica (diaBP): Ambas muestran distribuciones con picos en valores típicos de la población general.

Índice de masa corporal (BMI): La distribución es casi normal con un ligero sesgo hacia la derecha.

Frecuencia cardíaca (heartRate): Distribución normal con un pico alrededor de los 70-80 latidos por minuto.

Glucosa: La distribución está sesgada hacia la derecha con algunos valores muy altos.

* Visualización de Variables Categóricas

Algunas observaciones:

Género (gender): Hay una distribución más alta de mujeres en el dataset.

Fumador actual (currentSmoker): Casi la mitad de los participantes son fumadores actuales.

Educación (education): La mayoría de los participantes tienen niveles de educación básicos (niveles 1 y 2).

Accidente cerebrovascular prevalente (prevalentStroke): Muy pocos participantes han tenido un accidente cerebrovascular.

Hipertensión prevalente (prevalentHyp): Alrededor del 30% de los participantes tienen hipertensión.

Diabetes (diabetes): Una pequeña proporción de los participantes tiene diabetes.

Diez años de riesgo de enfermedad coronaria (TenYearCHD): Aproximadamente el 15% de los participantes tiene un riesgo positivo de enfermedad coronaria a diez años.

* El heatmap de la matriz de correlación muestra las relaciones entre las variables numéricas.
* Algunas observaciones clave:

Edad (age) tiene una correlación positiva notable con la presión arterial sistólica (sysBP) y diastólica (diaBP).

Presión arterial sistólica (sysBP) y diastólica (diaBP) están altamente correlacionadas entre sí.

Índice de masa corporal (BMI) tiene una correlación positiva con la presión arterial diastólica (diaBP) y la frecuencia cardíaca (heartRate).

Frecuencia cardíaca (heartRate) muestra una correlación negativa moderada con la presión arterial sistólica (sysBP).

* Análisis Bivariado

Algunas observaciones clave:

Edad (age): Las personas con TenYearCHD positivo tienden a ser mayores.

Presión arterial sistólica (sysBP) y diastólica (diaBP): Ambas presiones son más altas en el grupo con TenYearCHD positivo.

Índice de masa corporal (BMI) y frecuencia cardíaca (heartRate): No se observan diferencias notables entre los grupos.

Colesterol total (totChol): Levemente más alto en el grupo con TenYearCHD positivo.

Glucosa (glucose): Similar entre ambos grupos, aunque con una ligera tendencia hacia valores más altos en el grupo con TenYearCHD positivo.

* Resultados del Análisis Categórico Bivariado

Los resultados de las pruebas de chi-cuadrado para evaluar la independencia entre las variables categóricas y la variable objetivo (TenYearCHD) son los siguientes:

Género (gender):

Existe una relación significativa entre el género y el riesgo de enfermedad coronaria a diez años.

Fumador actual (currentSmoker):

No hay una relación significativa entre ser fumador actual y el riesgo de enfermedad coronaria a diez años.

Educación (education):

Existe una relación significativa entre el nivel educativo y el riesgo de enfermedad coronaria a diez años.

Accidente cerebrovascular prevalente (prevalentStroke):

Existe una relación significativa entre haber tenido un accidente cerebrovascular y el riesgo de enfermedad coronaria a diez años.

Hipertensión prevalente (prevalentHyp):

Existe una relación muy significativa entre tener hipertensión y el riesgo de enfermedad coronaria a diez años.

Diabetes (diabetes):

Existe una relación significativa entre tener diabetes y el riesgo de enfermedad coronaria a diez años.

Conclusiones Generales

Variables con Relaciones Significativas: Género, nivel educativo, accidente cerebrovascular prevalente, hipertensión prevalente y diabetes tienen relaciones significativas con el riesgo de enfermedad coronaria a diez años.

Variables sin Relaciones Significativas: Ser fumador actual no muestra una relación significativa con el riesgo de enfermedad coronaria a diez años, aunque podría tener otras implicaciones para la salud.