

1. **age**

- **Tipo:** Numérica (continua)
- **Descripción:** Indica la edad del paciente en años. Esta variable suele ser importante ya que el riesgo de infarto varía con la edad.

2. **sex**

- **Tipo:** Categórica (convertida a factor)
- **Descripción:** Representa el género del paciente. En muchos conjuntos de datos, se codifica como 0 y 1 o “Male”/“Female”. La influencia del sexo puede estar relacionada con diferencias en la incidencia y presentación clínica del infarto.

3. **cp (tipo de dolor en el pecho)**

- **Tipo:** Categórica (o numérica con pocos niveles, convertida a factor)
- **Descripción:** Indica el tipo de dolor en el pecho. Por ejemplo, los niveles pueden corresponder a: dolor típico de angina, atípico, no anginoso y asintomático. Este factor ayuda a diferenciar la probabilidad de enfermedad coronaria.

4. **trestbps (presión arterial en reposo)**

- **Tipo:** Numérica (continua)
- **Descripción:** Mide la presión arterial en reposo (generalmente en mm Hg). Es un indicador de la salud cardiovascular y puede influir en el riesgo de sufrir un infarto.

5. **chol (colesterol sérico)**

- **Tipo:** Numérica (continua)
- **Descripción:** Nivel de colesterol en la sangre, normalmente medido en mg/dl. Altos niveles pueden ser un factor de riesgo importante en enfermedades cardiovasculares.

6. **fbs (azúcar en ayunas)**

- **Tipo:** Categórica (o numérica con pocos niveles, convertida a factor)
- **Descripción:** Indica si el nivel de azúcar en sangre en ayunas es superior a un cierto umbral (por ejemplo, 120 mg/dl). Generalmente se

codifica como 0 = no y 1 = sí, representando la presencia de hiperglucemia.

7. restecg (resultados del electrocardiograma en reposo)

- **Tipo:** Categórica
- **Descripción:** Resume los hallazgos del electrocardiograma en reposo (por ejemplo, normal, con anomalías en la repolarización, hipertrofia ventricular izquierda). Ayuda a evaluar la función eléctrica del corazón.

8. thalach (frecuencia cardíaca máxima alcanzada)

- **Tipo:** Numérica (continua)
- **Descripción:** Registra el número máximo de latidos por minuto alcanzado durante una prueba de esfuerzo. Es un indicador del rendimiento cardíaco y la reserva funcional.

9. exang (angina inducida por el ejercicio)

- **Tipo:** Categórica (o numérica con pocos niveles, convertida a factor)
- **Descripción:** Indica la presencia o ausencia de angina inducida por el ejercicio (por ejemplo, 0 = no, 1 = sí). Su aparición durante una prueba de esfuerzo es un indicador de enfermedad coronaria.

10. oldpeak (depresión del ST)

- **Tipo:** Numérica (continua)
- **Descripción:** Representa la magnitud de la depresión del segmento ST durante el ejercicio, comparada con el reposo. Valores mayores pueden indicar isquemia miocárdica.

11. slope (pendiente del segmento ST)

- **Tipo:** Categórica (o numérica discreta, convertida a factor)
- **Descripción:** Refleja la inclinación de la pendiente del segmento ST en el electrocardiograma durante el ejercicio. Variaciones en esta medida pueden ayudar a diferenciar entre diferentes grados de isquemia.

12. ca (número de vasos principales coloreados por fluoroscopia)

- **Tipo:** Numérica o categórica (según la codificación)

- **Descripción:** Indica el número de vasos sanguíneos principales (usualmente entre 0 y 3 o 4) afectados o visualizados mediante técnicas de imagen. Es un marcador de la extensión de la enfermedad arterial coronaria.

13. thal (resultado de la prueba de talasemia o estado del miocardio)

- **Tipo:** Categórica
- **Descripción:** Suele codificar la condición del tejido cardíaco, por ejemplo, “normal”, “defecto fijo” o “defecto reversible”. Esta variable aporta información sobre la perfusión miocárdica.

14. heart_attack

- **Tipo:** Categórica (variable objetivo)
- **Descripción:** Indica si el paciente sufrió un infarto al miocardio. Se codifica con etiquetas “No” y “Yes” (o 0 y 1 en origen). Esta es la variable que se intenta predecir en el modelo.

Variables no numéricas

. ID o Identificador del Paciente

- **Tipo:** Categórica (puede tratarse de un identificador único en formato de texto)
- **Descripción:** Generalmente se incluye un identificador único para cada paciente. Esto permite rastrear o distinguir registros de manera individual, sin que este dato influya en el análisis predictivo.

2. Género (sex)

- **Tipo:** Categórica (factor)
- **Descripción:** Indica el género del paciente, normalmente codificado como “Male”/“Female” o “M”/“F”. El género es un factor relevante en el análisis de riesgo, puesto que existen diferencias en la incidencia y presentación de la enfermedad cardiovascular entre hombres y mujeres.

3. Tipo de Dolor Torácico (cp)

- **Tipo:** Categórica (factor)
- **Descripción:** Representa la descripción o clasificación del dolor en el pecho. Por ejemplo, los valores pueden indicar dolor típico de angina, dolor atípico, o

ausencia de síntomas. Esta variable es un indicador clave en la evaluación de la probabilidad de enfermedad coronaria.

4. Prueba de Azúcar en Ayunas (fbs)

- **Tipo:** Categórica (factor)
- **Descripción:** Indica si el nivel de azúcar en sangre en ayunas supera un umbral determinado (por ejemplo, 120 mg/dl). Se suele codificar como “Yes” o “No”, lo cual sugiere la presencia o ausencia de hiperglucemia, un factor de riesgo cardiovascular.

5. Resultados del Electrocardiograma en Reposo (restecg)

- **Tipo:** Categórica (factor)
- **Descripción:** Resume los hallazgos del electrocardiograma en reposo. Los valores pueden representar resultados como “normal”, “anormalidad en la repolarización” o “hipertrofia ventricular”. Esta información es útil para evaluar el estado eléctrico del corazón.

6. Angina Inducida por el Ejercicio (exang)

- **Tipo:** Categórica (factor)
- **Descripción:** Indica si durante una prueba de esfuerzo se presenta angina (dolor en el pecho inducido por el ejercicio). Generalmente se codifica en dos categorías (“Yes” y “No”), lo que permite identificar a pacientes con signos de isquemia inducida por el esfuerzo.

7. Pendiente del Segmento ST (slope)

- **Tipo:** Categórica (factor)
- **Descripción:** Se refiere a la inclinación de la pendiente del segmento ST en el electrocardiograma durante la prueba de esfuerzo. La clasificación (por ejemplo, “upsloping”, “flat”, “downsloping”) ayuda a diferenciar entre distintos patrones de respuesta cardíaca ante el ejercicio y puede indicar la presencia o gravedad de isquemia.

8. Resultados de la Prueba de Talasemia (thal)

- **Tipo:** Categórica (factor)
- **Descripción:** Mide el estado del miocardio o la perfusión mediante pruebas especializadas; los resultados suelen agruparse en categorías tales como

“normal”, “defecto fijo” o “defecto reversible”. Esta variable aporta información sobre la condición del tejido cardíaco.

9. Estado de Fumador (smoker)

- **Tipo:** Categórica (factor)
- **Descripción:** Indica el estatus respecto al consumo de tabaco, con categorías como “Nunca”, “Exfumador” o “Fumador actual”. Dado que el tabaquismo es un factor de riesgo cardiovascular, esta variable resulta relevante para la predicción del infarto.

10. Infarto (heart_attack)

- **Tipo:** Categórica (factor, variable objetivo)
- **Descripción:** Variable de salida que indica si el paciente sufrió un infarto al miocardio. Se codifica comúnmente con etiquetas “No” y “Yes”. Es la variable que se intenta predecir en el modelo, y su correcta identificación es crucial para entrenar y evaluar los algoritmos predictivos.