

## DICCIONARIO DE VARIABLES - BD: Heart\_Disease\_Prediction.xlsx

La base de datos utilizada para este proyecto fue la "Heart\_Disease\_Prediction", extraída del sitio web Kaggle ([visualizar acá](#)), compuesta por un total de 14 variables y 270 observaciones. A continuación una breve descripción de cada una:

- *pat\_id* → identificador numérico del paciente. Variable cuantitativa discreta
- *age* → edad del paciente. Variable cuantitativa discreta
- *sex* → sexo del paciente (1 = masculino, 0 = femenino). Variable cualitativa nominal
- *chest\_pain\_type* → tipo de dolor de pecho o angina de pecho (1 = Angina típica, 2 = Angina atípica, 3 = Sin angina, 4 = Asintomático). Variable cualitativa nominal
- *bp* → presión sanguínea (mm Hg). Variable cuantitativa discreta
- *cholesterol* → nivel de colesterol (mg/dl). Variable cuantitativa discreta
- *lbs\_over\_120* → nivel de azúcar en sangre por encima de los 120 mg/dl (1 = si, 0 = no). Variable cualitativa nominal
- *ekg\_results* → resultados del electrocardiograma (ECG) en reposo (0 = normal, 1 = Anomalía en el segmento ST, 2 = Hipertrofia ventricular). Variable cualitativa nominal
- *max\_hr* → frecuencia cardíaca máxima (ppm). Variable cuantitativa discreta
- *exercise\_angina* → angina de pecho inducida por el esfuerzo/ejercicio (0 = falso, 1 = verdadero). Variable cualitativa nominal
- *st\_depression* → la pendiente del segmento ST máximo inducida por el ejercicio. Variable cuantitativa discreta
- *slope\_of\_st* → orientación de la pendiente del segmento ST (1 = Pend. positiva, 2 = Plana, 3 = Pend. negativa). Variable cualitativa nominal
- *number\_vessels* → número de grandes vasos del corazón por fluroscopia (de 0 a 3). Variable cualitativa ordinal
- *thallium* → resultado de la prueba de flujo sanguíneo con talio (3 = Normal, 6 = Defecto no reversible, 7 = Defecto reversible). Variable cualitativa nominal
- *heart\_disease* → probabilidad de tener una enfermedad cardíaca (EC) (*Absence* = <50% de acortamiento del diámetro, *Presence* = >50% de acortamiento del diámetro). Variable cualitativa nominal
- *hd\_dummy* → la probabilidad de tener una enfermedad cardíaca llevada a términos de 0 y 1 (booleana) (EC) (0 = <50% de acortamiento del diámetro, 1 = >50% de acortamiento del diámetro). Variable cualitativa nominal

## PRIMERAS CONCLUSIONES DERIVADAS DEL AED:

- 1) La media de la variable edad es de 54 años, teniendo como paciente más joven a uno de 29 años y como más viejo a uno de 77 años
- 2) La muestra es predominantemente masculina, con un 68% aproximadamente versus un 32% de mujeres
- 3) La frecuencia máxima promedio de los pacientes ronda las 160 ppm
- 4) La distribución de los valores de colesterol en los pacientes presenta una asimetría positiva, es decir, la media (aprox 250 mg/dl) supera a la mediana (245 mg/dl)
- 5) Se tienen 129 pacientes asintomáticos en el dolor de pecho, lo cual representa un mayor riesgo de padecer una enfermedad cardíaca