

Projet analyse des facteurs influençant le risque de maladies cardiaques

Vous êtes consultant.e dans une entreprise de santé et vous avez été chargé.e de mener une analyse des facteurs influençant le risque de maladies cardiaques chez les patients. Pour cela, vous utilisez le Heart Disease Dataset (Cleveland), qui contient plusieurs variables importantes liées aux caractéristiques cliniques des patients. Le tableau ci-dessous décrit les variables présentes dans ce dataset.

Variable	Description
age	Âge du patient.
sex	Sexe du patient (1 = homme, 0 = femme).
cp	Type de douleur thoracique (4 catégories).
trestbps	Pression artérielle au repos (en mm Hg).
chol	Taux de cholestérol sérique (mg/dl).
fbs	Glycémie à jeun > 120 mg/dl (1 = vrai; 0 = faux).
restecg	Résultats de l'électrocardiogramme au repos (0, 1, 2).
thalach	Fréquence cardiaque maximale atteinte.
exang	Angine induite par l'exercice (1 = oui; 0 = non).
oldpeak	Dépression ST induite par l'exercice par rapport au repos.
slope	Pente du segment ST lors de l'exercice (3 valeurs).
ca	Nombre de vaisseaux principaux colorés par fluoroscopie (0-3).

thal	Résultat de l'examen sanguin (3 = normal, 6 = fixe, 7 = réversible).
target	Variable cible indiquant la présence (1) ou l'absence (0) de maladie cardiaque.