Objective:

1. Binary classification of 10-year risk of coronary heart disease dataset into 2 classes i.e, having risk or not using Logistic Regression.
2. Evaluate the model accuracy (evaluation metrices)

Dataset Details:
Attributes information in details:

Demographic:
Sex: male or female
Age: Age of the patient
Education: No further information was provided

Behavioral:
Current Smoker: whether or not the patient is a current smoker
Cigs Per Day: the number of cigarettes that the person smoked on average in one day.

Information on medical history:
BP Meds: whether or not the patient was on blood pressure medication
Prevalent Stroke: whether or not the patient had previously had a stroke
Prevalent Hyp: whether or not the patient was hypertensive
Diabetes: whether or not the patient had diabetes

Information on current medical condition:
Tot Chol: total cholesterol level
Sys BP: systolic blood pressure
Dia BP: diastolic blood pressure
BMI: Body Mass Index
Heart Rate: heart rate (In medical research, variables such as heart rate though in fact discrete, are considered continuous because of a large number of possible values.)
Glucose: glucose level

Target variable to predict:
10-year risk of coronary heart disease (CHD) - (binary: “1”, means “Yes”, “0” means “No”)