Data Collection and Preprocessing Phase

Date	15 July 2024
Team ID	740044
Project Title	**One Year Life Expectancy post on Thoracic Surgery using Machine Learning**
Maximum Marks	2 Marks

Data Quality Report:

The Data Quality Report will summarize data quality issues from the selected source, including severity levels and resolution plans. It will aid in systematically identifying and rectifying data discrepancies.

Data Quality Report:

Data Source	Data Quality Issue	Severity	Resolution Plan

	Index(['Diagnosis', 'FVC', 'FEV1', 'Performance', 'Pain', 'Hemoptysis', 'Dyspnea', 'Cough', 'Weakness', 'Tumor Size', 'Diabetes_Mellitus', 'MI_6mo', 'PAD', 'Smoking', 'Asthma', 'Age', 'Death_1yr'],	Moderate	Use mean/median imputation.
	Categorical data in the dataset	Moderate	Encoding has to be done in