

# METHODOLOGY STEPS

## 1- SELECTING THE DATASET

An appropriate open-source dataset is chosen for analysis.

## 2- DEFINING THE LEARNING TASK

The machine learning task is identified as regression or classification.

## 4- IMPORTING THE DATA

The dataset is imported into a Jupyter Notebook environment.

## 3- SPECIFYING THE OUTPUT VARIABLE

The variable to be predicted is determined for supervised learning.

## 5- EXPLORING THE DATA STRUCTURE

The size of the dataset and sample records are examined.

## 6- UNDERSTANDING THE DATASET CHARACTERISTICS

Feature types, missing values, and basic statistics are reviewed.

