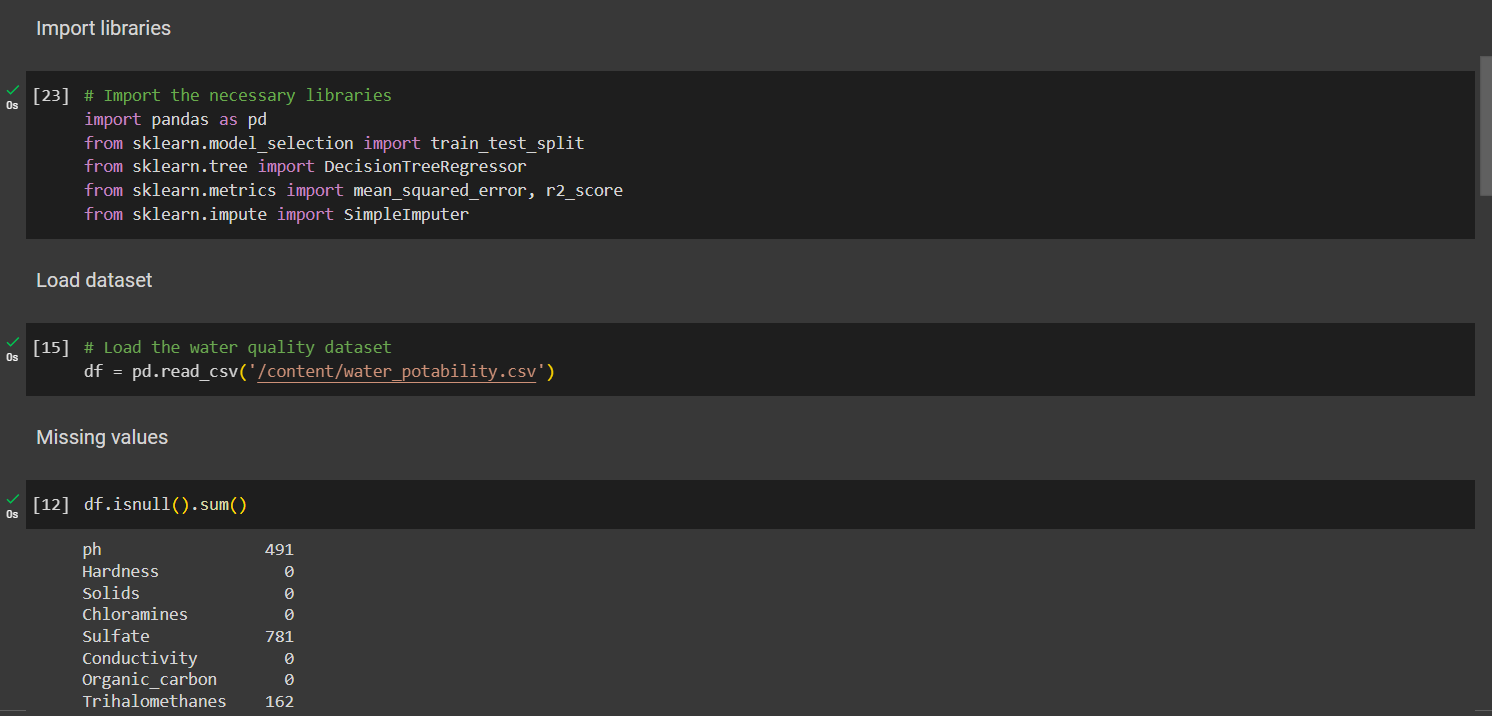
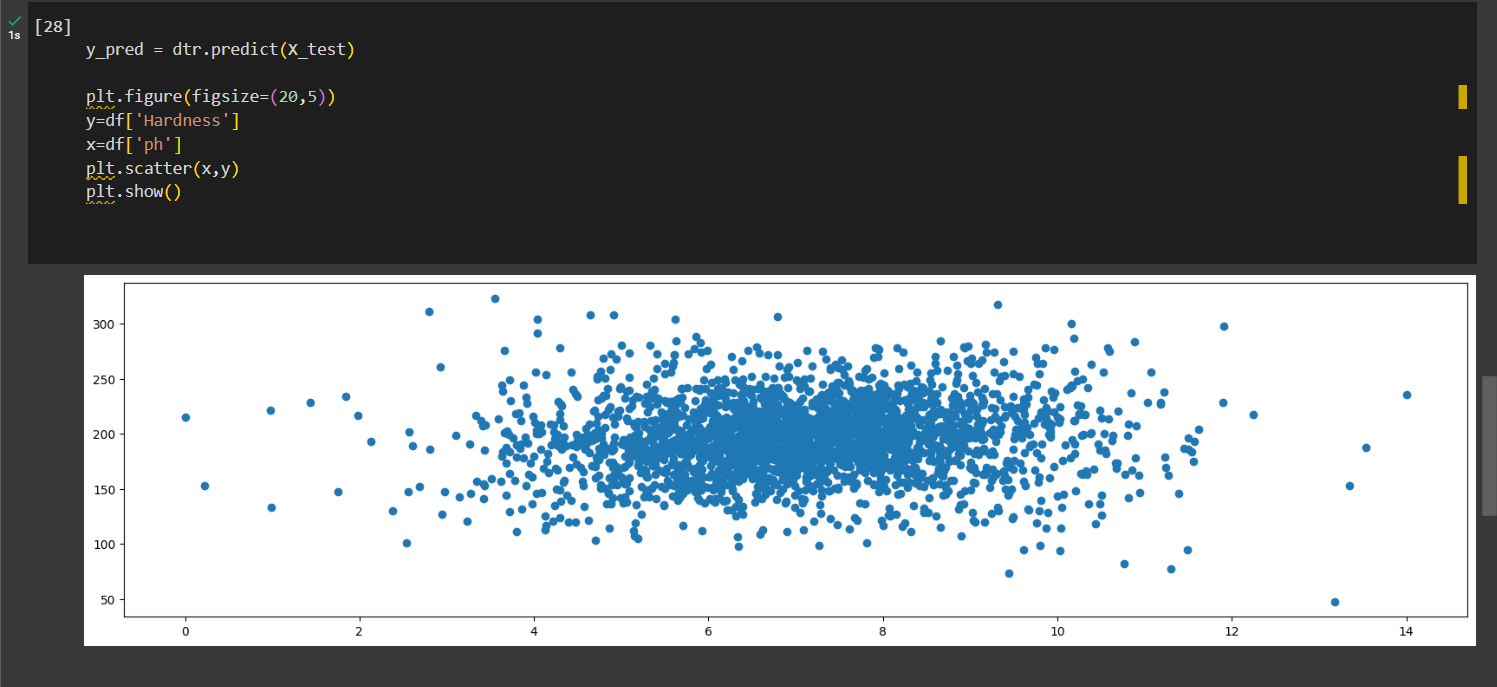
TITLE: WATER QUALITY ANALYSIS USING m5p

ALGORITHM

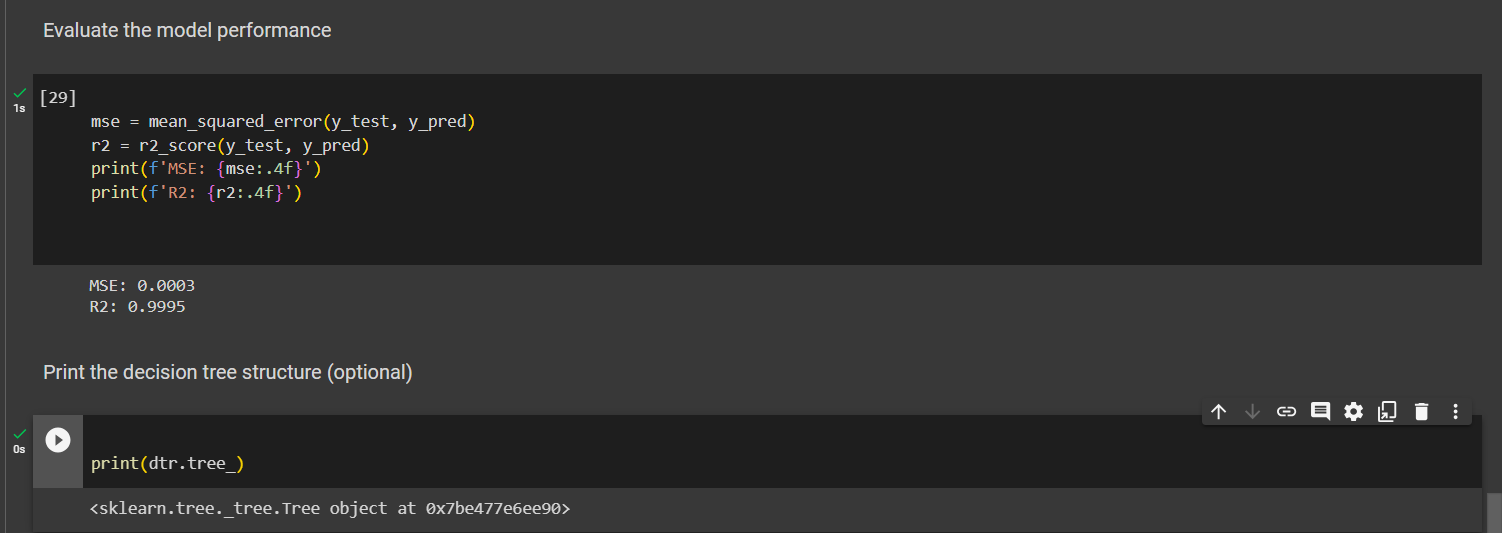
* HERE, I USED TO START BUILDING THE WATER QUALITY ANALYSIS BY PRE-PROCESSING THE DATA AND PERFORMING USING EDA WITH m5p ALGORITHM.
* WITH THAT I TO OBTAIN THE WATER QUALITY DATASET AND PRE-PROCESS IT BY HANDLING MISSING VALUES AND OUTLIERS.
* TO CONDUCT EDA TO VISUALIZE PARAMETER DISTRIBUTIONS,CORRELATIONS AND POTENTIAL DEVIATIONS FROM STANDARDS.
* FIRST I IMPORT THE NECESSARY LIBRARIES
* AFTER THAT I UPLOAD THE WATER POTABILITY CSV FILE.
* AND THEN CALLING THE MISSING VALUE FUNCTIONS.
* SPLIT THE DATASET TO TRAIN AND TEST.



* HANDLE MISSING VALUES IN X USING SIMPLEIMPUTER.
* SPLIT THE DATASET TO TRAIN AND TEST SETS.
* CREATE AND FIT THE DECISION TREE REGRESSOR MODEL.



* TO PREDICT USING THE DTR(DECISION TREE REGRESSOR AND PLOT THE FIGURE(SCATTER PLOT).



* EVALUATE THE MODEL PERFORMANCE TO FIND THE MSE(MEAN SQUARED ERROR) AND r2 SCORE.
* FINALLY I PRINT THE DECISION TREE STRUCTURE WITH THE PYTHON CODE OF print(dtr.tree\_)