THEIL-SEN REGRESSION

This code uses **Theil-Sen Regression** to predict the PRODUCT_LENGTH of a product based on its PRODUCT_TYPE_ID. Here's a breakdown of how the code works:

- The necessary libraries are imported, including pandas for data manipulation, numpy for numerical operations, statsmodels for statistical modeling, and the mean_absolute_percentage_error function from scikit-learn for evaluating the model's performance.

- The training and test data are loaded into pandas dataframes. The NaN values in the data are replaced with 0 using the fillna method.

- The features (i.e., the PRODUCT_TYPE_ID column) and target variable (i.e., the PRODUCT_LENGTH column) are selected from the training and test data.

- A Theil-Sen regression model is trained on the training data using the RLM class from the statsmodels library. This model fits a linear regression model to subsets of the data and calculates the median of the resulting coefficients to estimate the regression line.

- The trained model is used to make predictions on the test data by calling the predict method on the result object returned by the model's fit method.

- The mean absolute percentage error between the true target values and predicted target values is calculated using the mean_absolute_percentage_error function.

- The score is printed to the console.A submission file is created using a pandas dataframe containing the predicted PRODUCT_LENGTH values and the corresponding PRODUCT_ID values from the test data. The index of this dataframe is set to the PRODUCT_ID column, and the dataframe is saved as a CSV file.