Simple Linear Regression Lab Activity

Learning outcomes:

1. Simple linear regression models
2. Understanding and interpreting the output

Activity

Problem Statement

A large Toyota car dealership offers purchasers of new cars, the option to buy their used car as part of a trade-in. In particular, a new promotion promises to pay high prices for the used Toyota Corollas for purchasers of a new car. The dealers then sell the used car for a small profit. To ensure a reasonable profit, the dealer needs to be able to predict the price that the dealership will get for the used cars. For that reason, data was collected on all previous sales of used Toyota Corollas at the dealership. The goal is to predict the price of a used Toyota Corolla based on its age. Age is given in months.

Steps

1. Read the data 'Toyota_SimpleReg.csv' into R.
2. Understand the structure of the data and perform the required pre-processing steps
    a. Drop the drop the "Id", "model" attributes
    b. Check for missing values etc.
3. Check correlation between independent and dependent variable.
4. Split the data into train and test data sets.
5. Build the linear regression model and interpret the results lm(DV~IV, data = dataframe_name)
6. Review residual plots and analyze the model summary
7. Evaluate error metrics evaluation on train data and test data library(DMwR)

    regr.eval (train$DV,model$fitted.values)

    regr.eval (test$DV, predict(model, test))