1) In the regression framework (i.e., when we try to predict a numerical value Y instead of a class), propose an other homogeneity measure. Justify your choice.

A homogeneity measure that would be appropriate in the regression framework is the least square function. Since we are predicting numerical values, we could utilize the following function to compute the impurity of a node m:

$$Q_m(T) = \frac{1}{N_m} \sum_{i:x_i \in R_m} (y_i - \hat{c}_m)^2$$

where Nm is the number of observations in node m, and c is the average value computed on y present in node m.