9.1 Conceptual

[Q1] In a regression tree the target variable does not have classes, therefore a regression model is fit to the target variable using each of the independent variables.

Entropy is used to calculate the amount of disorder of a specific node, meaning calculating the presence of one class of the target variable compared to the other class. The purer the data, the higher the presence of one specific class. Since, for regression tree the target variable does not have clear classes, entropy cannot be calculated and thus not used as a measure when splitting nodes. Instead other measures such as the sum of squared errors or cross validation need to be used.

[Q2]

a)

b)