



# Decision Trees Intro

# Models we learned so far

Parametric vs. Non-parametric

Linear Regression

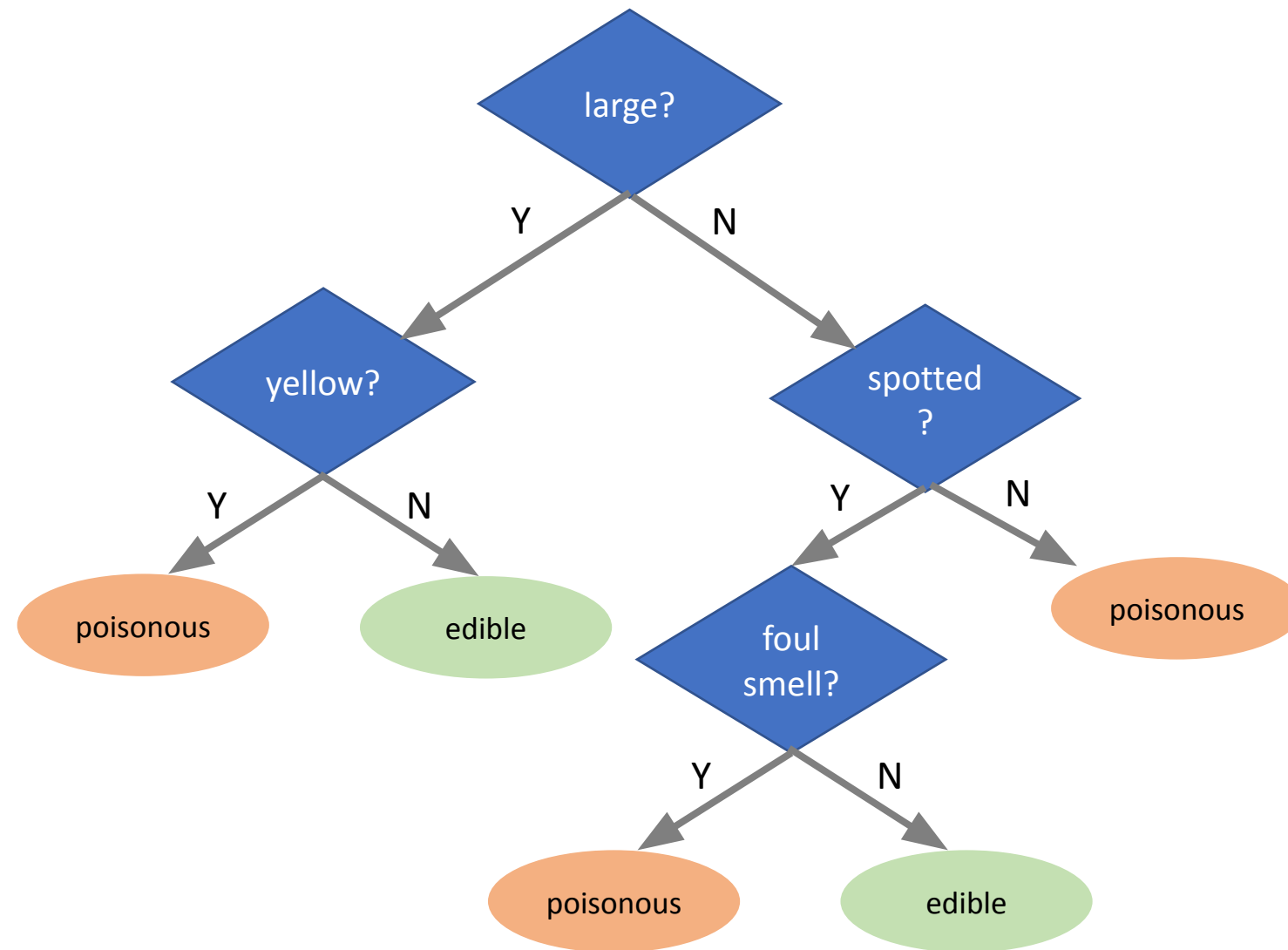
Logistic Regression

kNN

Parameters vs. Hyperparameters

Decision Tree

# What is Decision Tree?

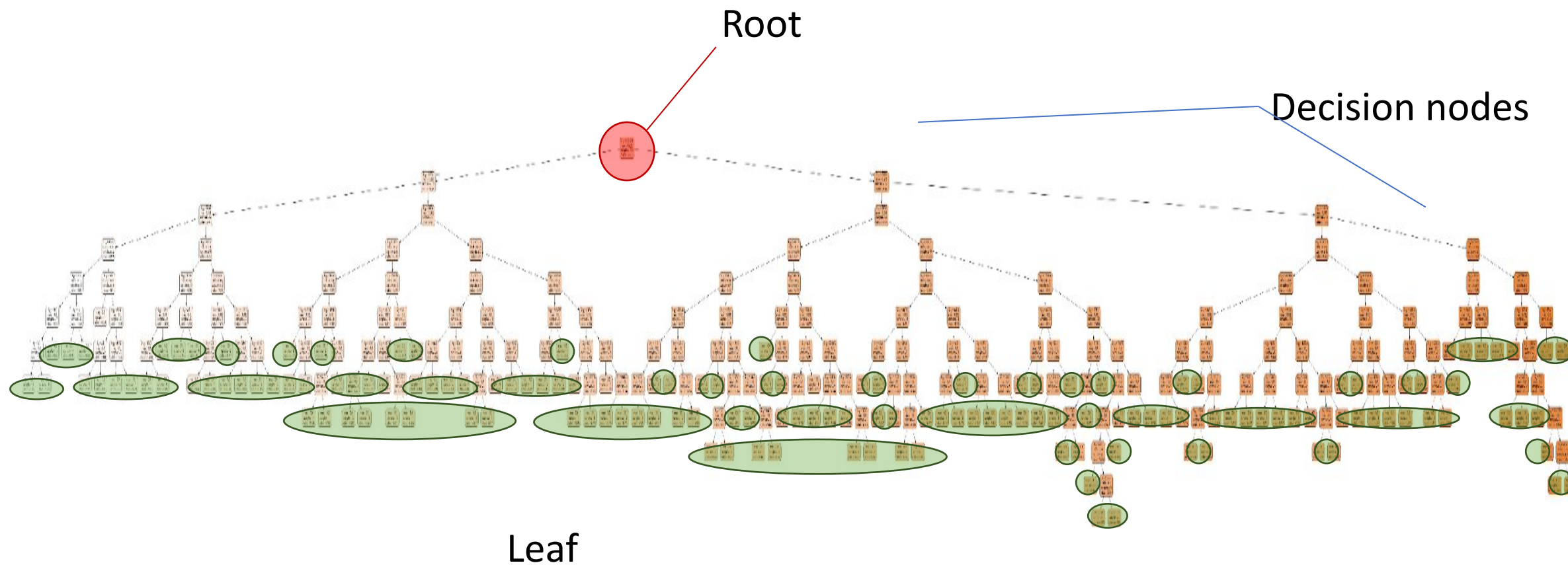


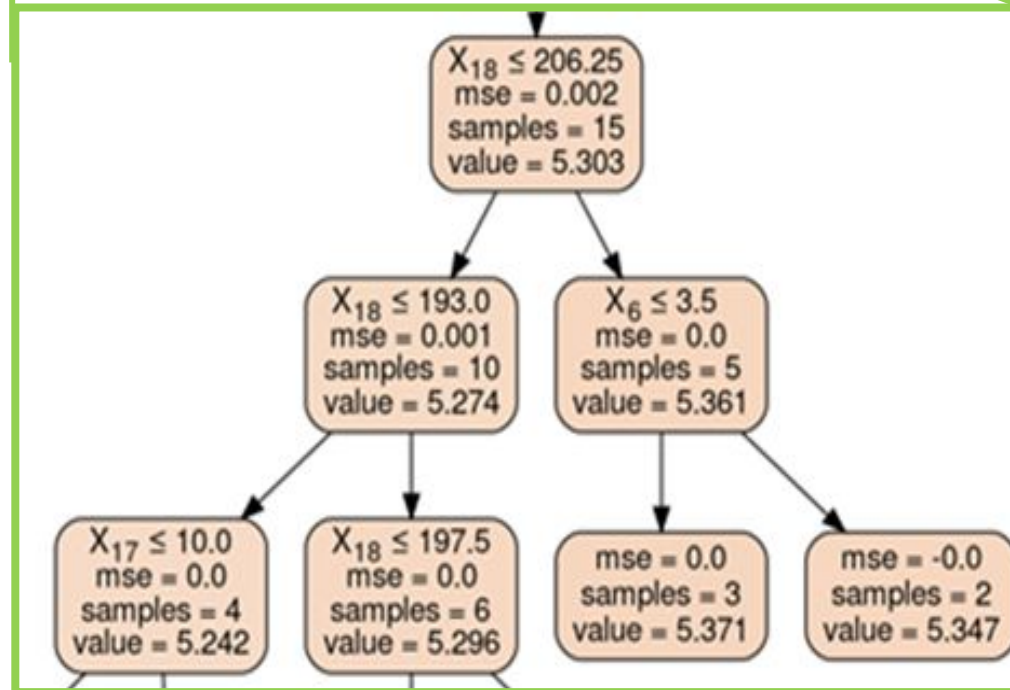
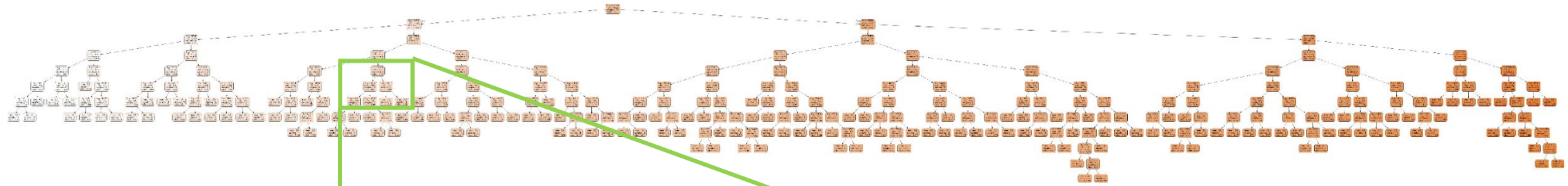
Caesar's mushroom



Death Cap

# Decision Tree Nodes





# How models “Learn” to make a Decision

Linear Regression

Minimize MSE

Logistic Regression

Minimize Cross Entropy

kNN

No optimization, but uses distance

Decision Tree

Split to minimize MSE for Regression task  
and minimize  
Cross Entropy or Gini for Classification task

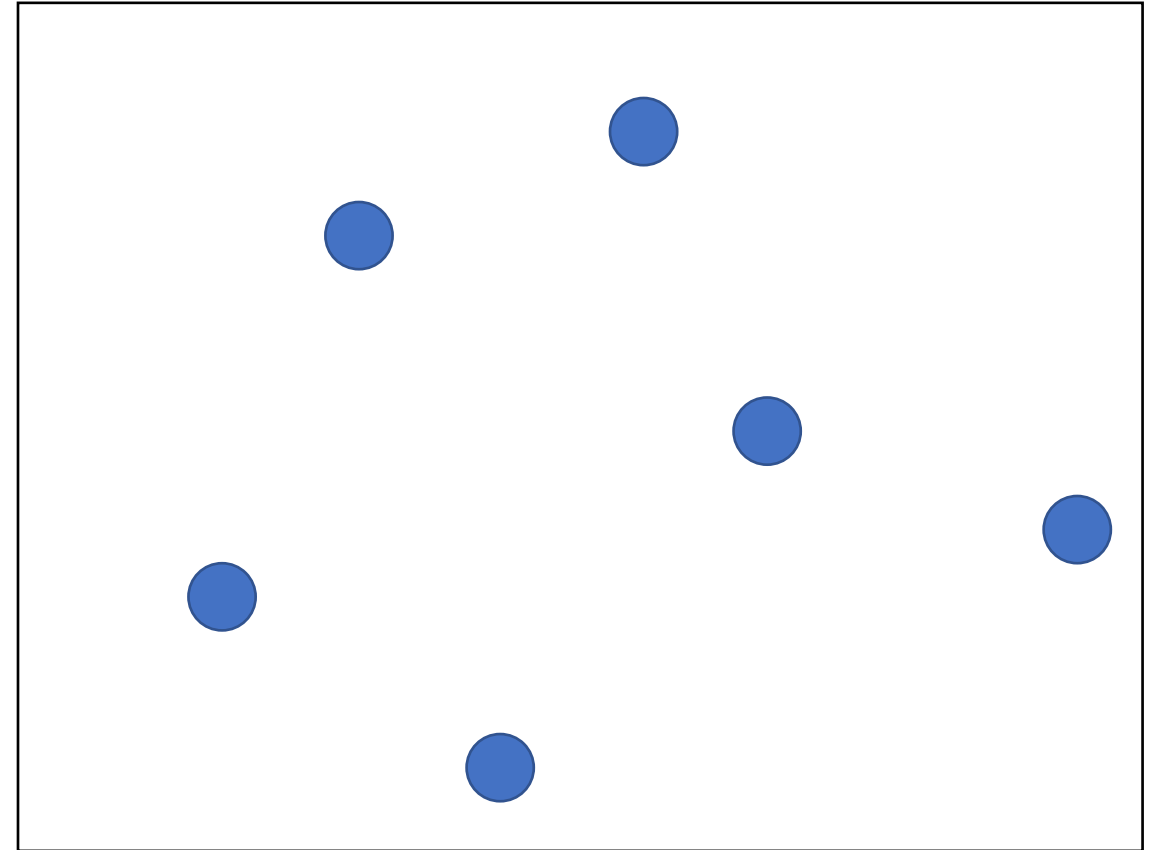
# Decision Tree Regressor

The goal is to find boxes  $R_1 \sim R_J$  such that

$$\sum_{j=1}^J \sum_{i \in R_j} (y_i - \hat{y}_{R_j})^2 \text{ is minimized.}$$

the mean of the data in  
the box

x2



x1

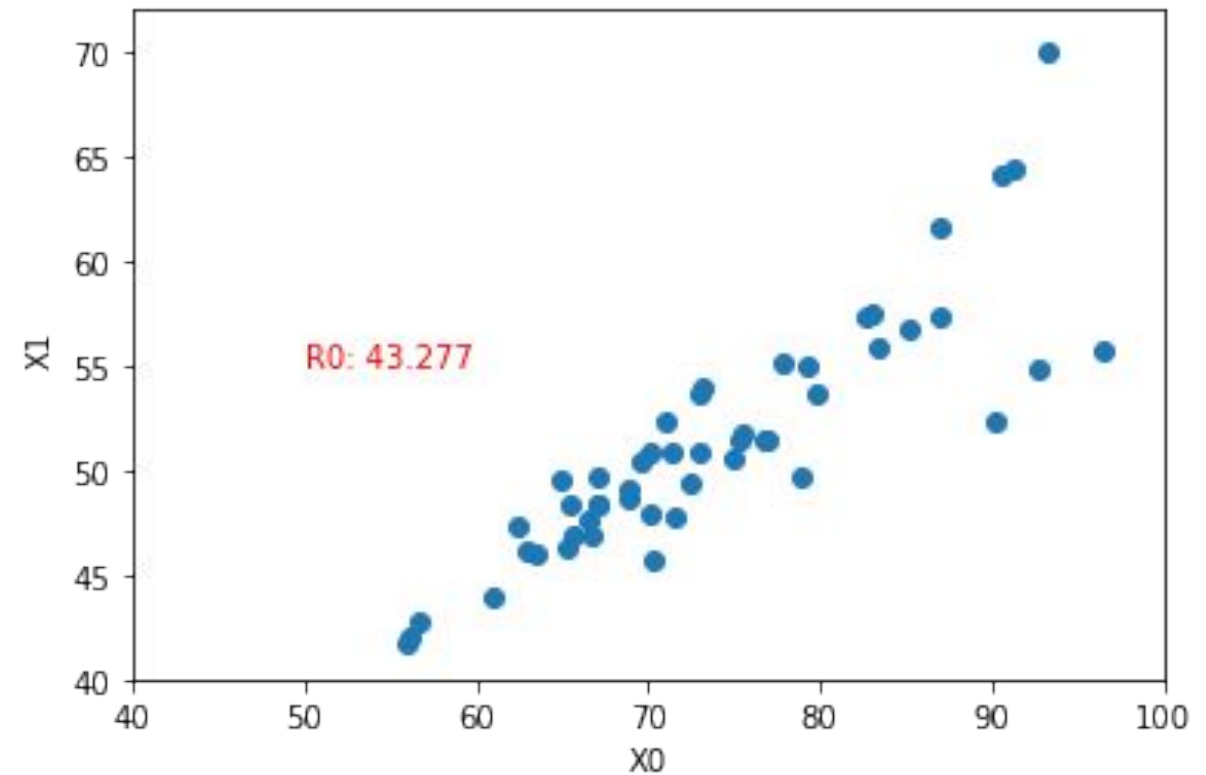
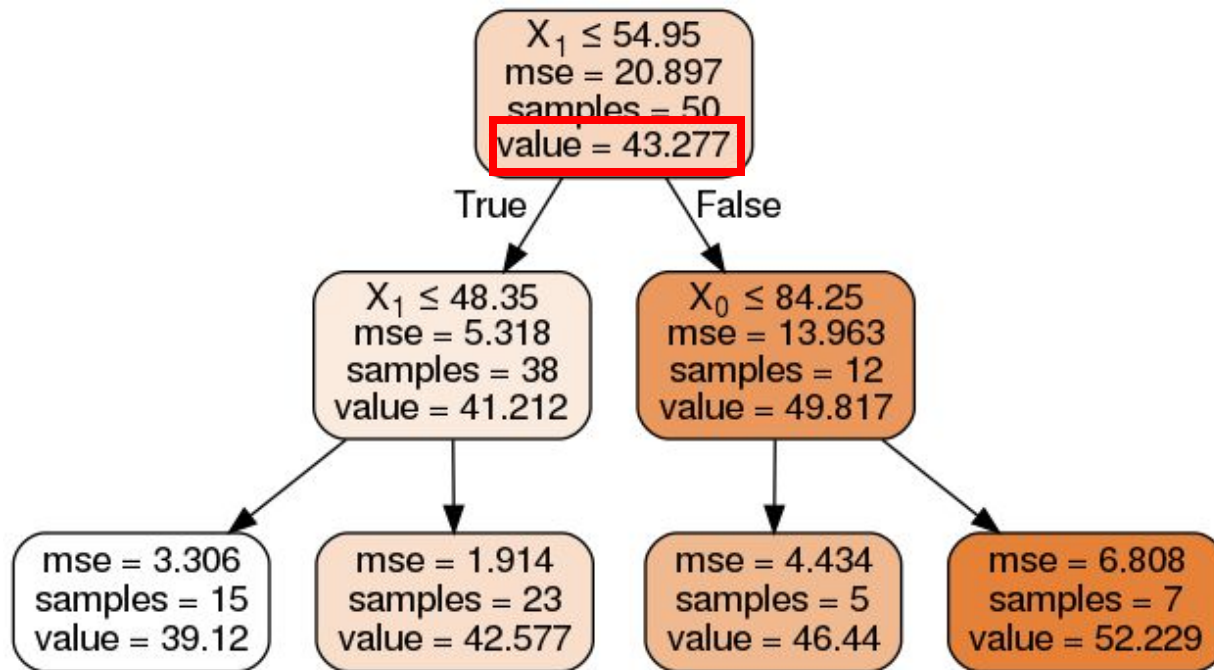


# Decision Tree Regressor

## Faculty Salary Dataset

4 features

50 samples



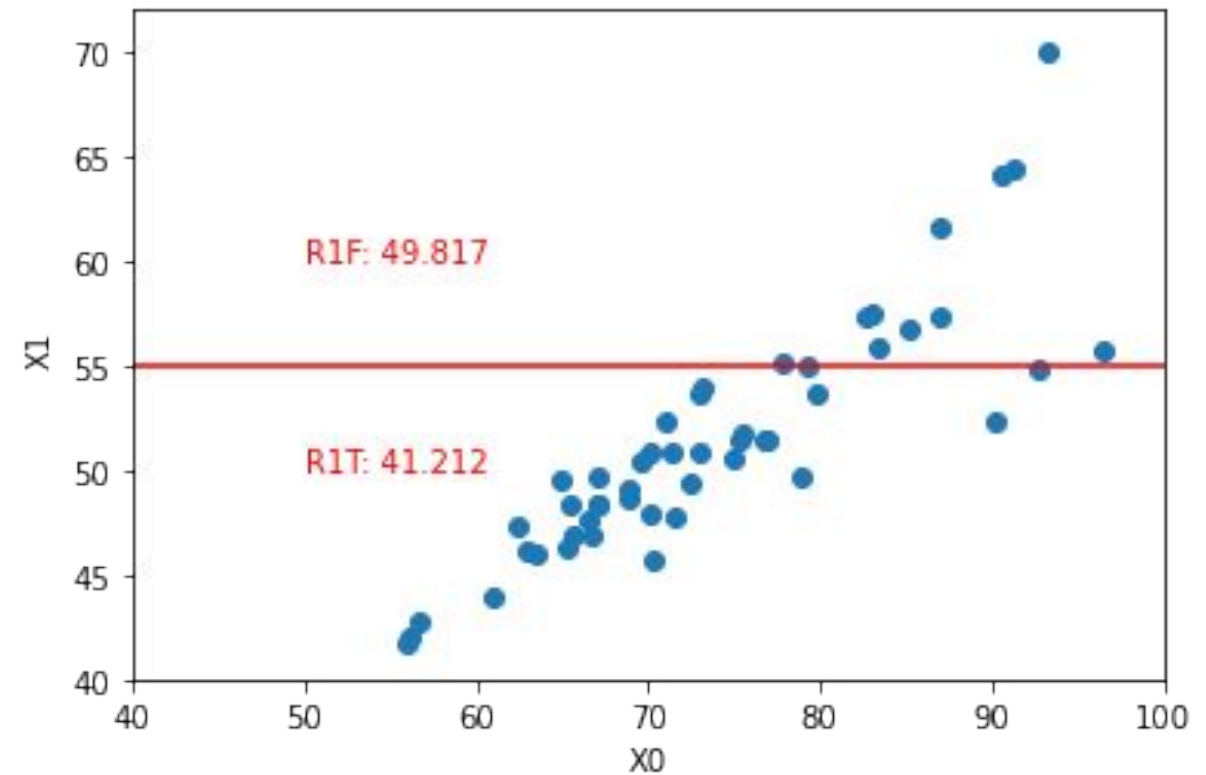
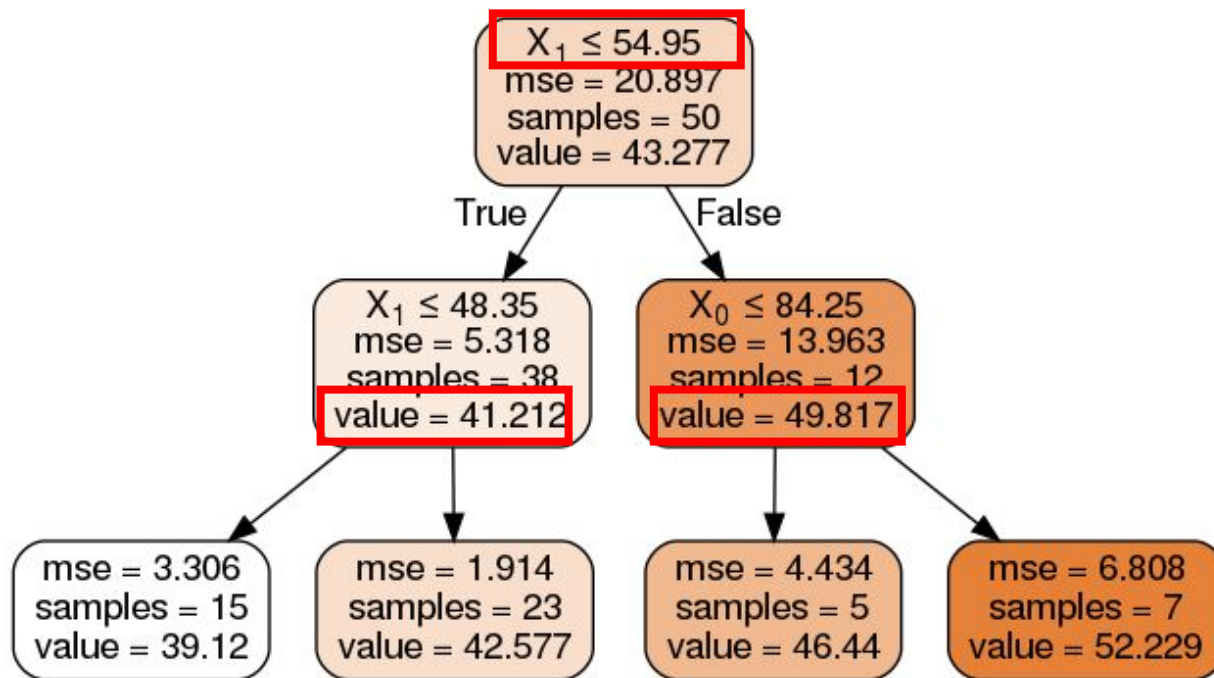


# Decision Tree Regressor

## Faculty Salary Dataset

4 features

50 samples

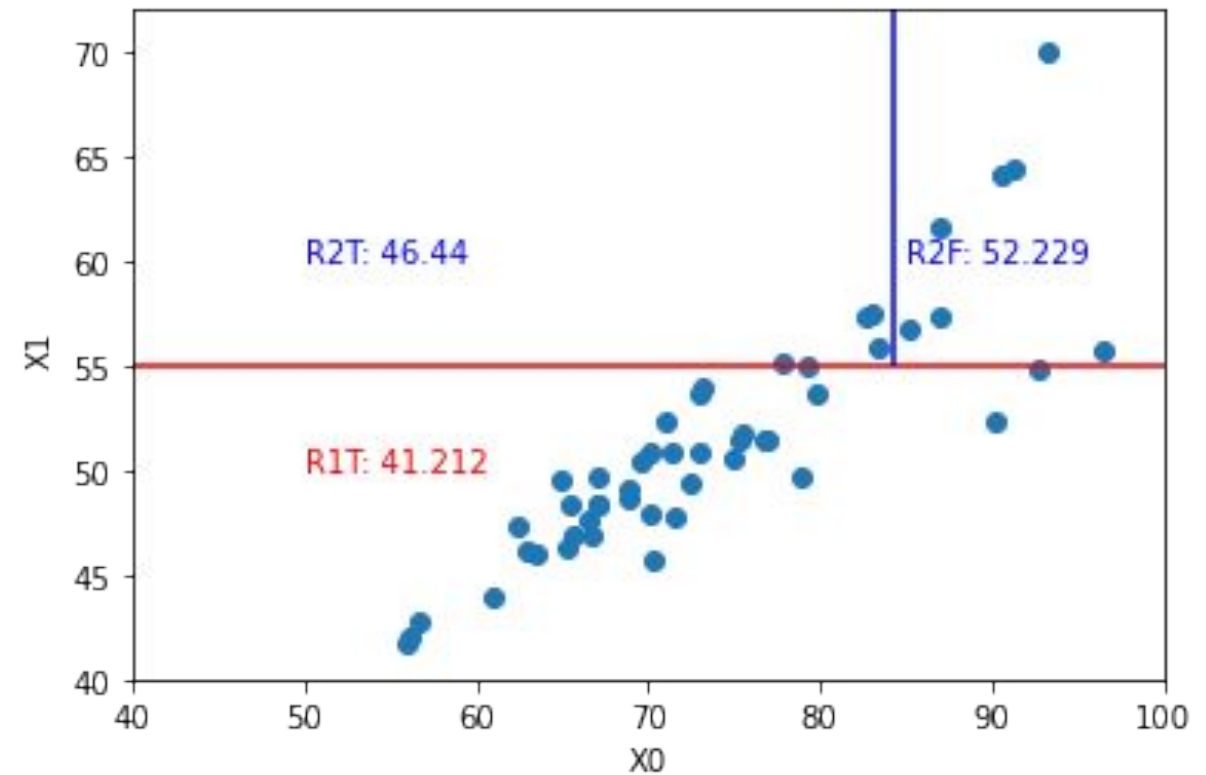
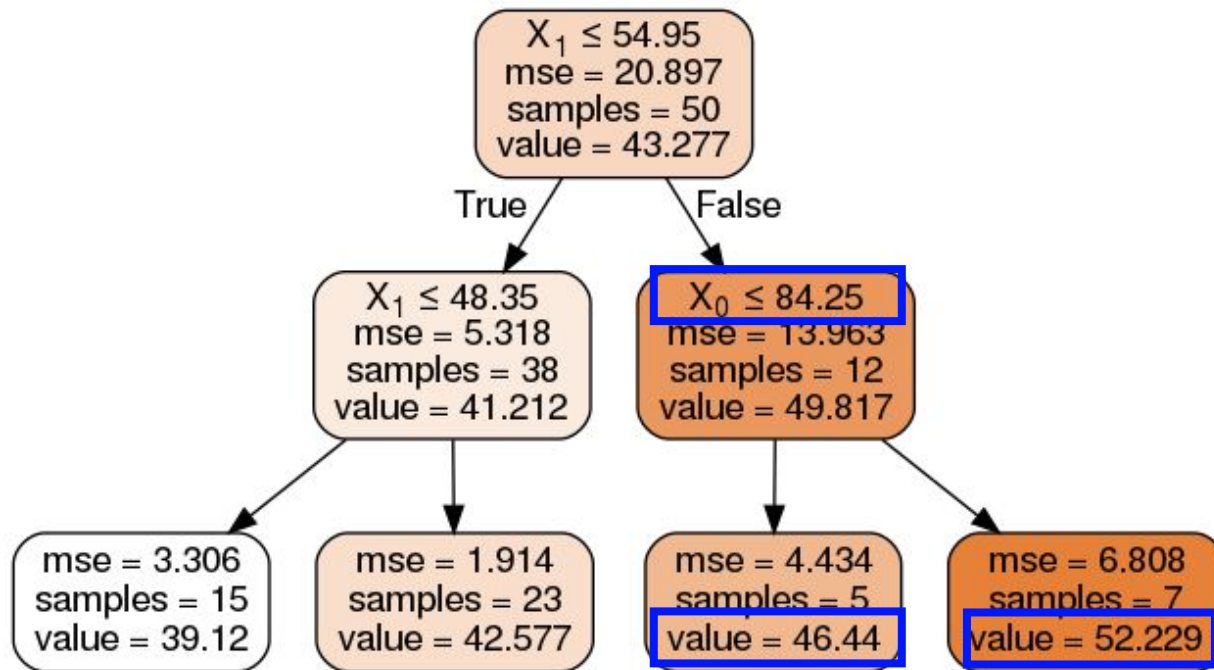


# Decision Tree Regressor

## Faculty Salary Dataset

4 features

50 samples



# Decision Tree Regressor

## Faculty Salary Dataset

4 features

50 samples

