**Advantages of Regression Trees:**

1. Simple

* The Results are Simplistic: Regression tree allows for the rapid classification of new observations, just evaluating just one or two logical conditions (if...than, if else…..)
* Easier to explain with true or false

1. Nonparametric & Nonlinear

* No assumptions
* No fix the model
* well suited to data mining

1. Perform Feature Selection

* the top few nodes on which the tree is split are the most important variables (first layer, second layer)

**Disadvantages of Regression Trees:**

1. overfitting

* It occurs when the tree takes into a lot of noise that exists in the data and comes up with an inaccurate result
* Noise: too many unregularized features

1. High variance

* a small variance in the data can lead to a very high variance in prediction, thereby affecting the stability of the outcome.

1. Low bias

* A regression tree that is very complicated usually has a low bias. This makes it very difficult for the model to merge any new data.

**Data explanation**:

`IncomeTwenty` is the first layer variable. If the income in 2020 is greater than \$586,500, the average death case is 270,000 people, accounting for 16% of total death cases. In contrast, if the state income in 2020 is less than \$586,500, the average number of deaths is 4700 people, accounting for 84% of total death cases. Let us continue to concentrate on the states where the average income in 2020 is less than \$586,500. The second important independent variable is `Black\_or\_African\_American\_alone`. According to the tree plot, if the state includes less than 332000 people of the black or African American race, the average number of deaths is 2000, accounting for 47 percent of total death cases. Conversely, if the state includes greater than 332000 people of them, the average number of deaths is 8200, accounting for 37 percent of total death cases.