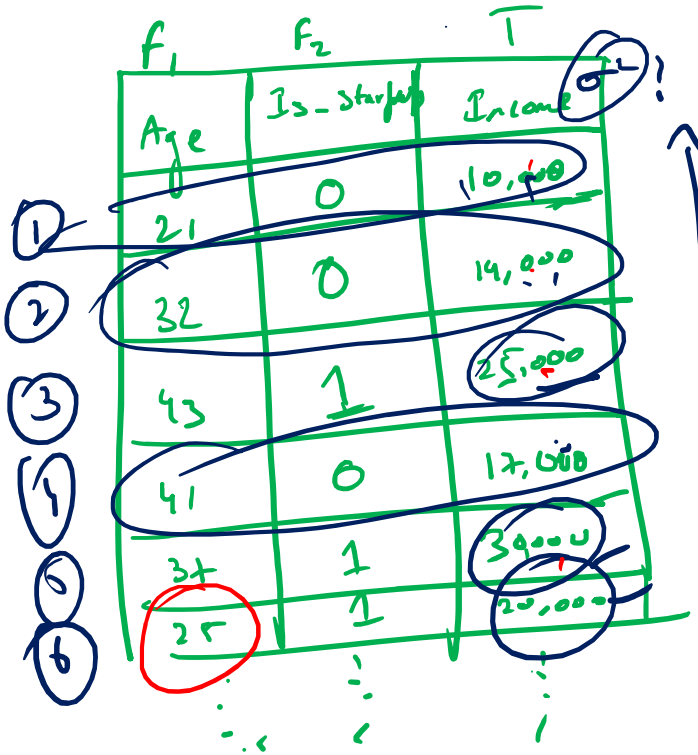
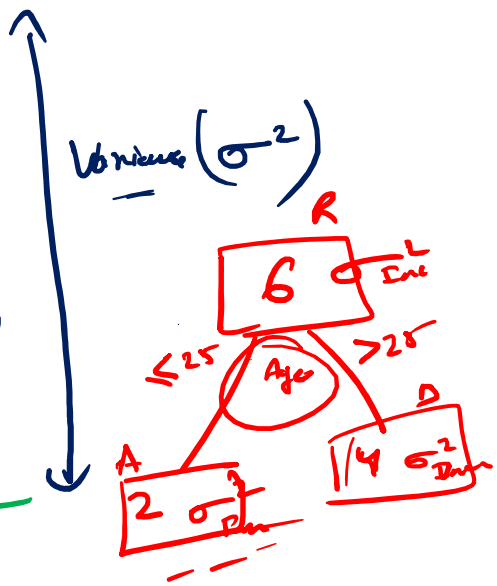


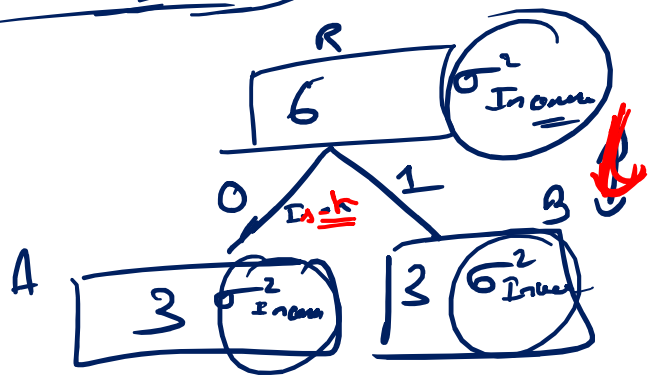


$$\sigma^2 = \frac{SS}{N} \rightarrow \frac{\sum (x - \bar{x})^2}{N}$$

Regression using D.T.

Splitting Criteria

↳ Variance Reduction



w_A ⇓ 3/6

w_B ⇓ 3/6

$$V.Red = \sigma^2_{Parent} - \left(w_A \sigma^2_A + w_B \sigma^2_B \right)$$

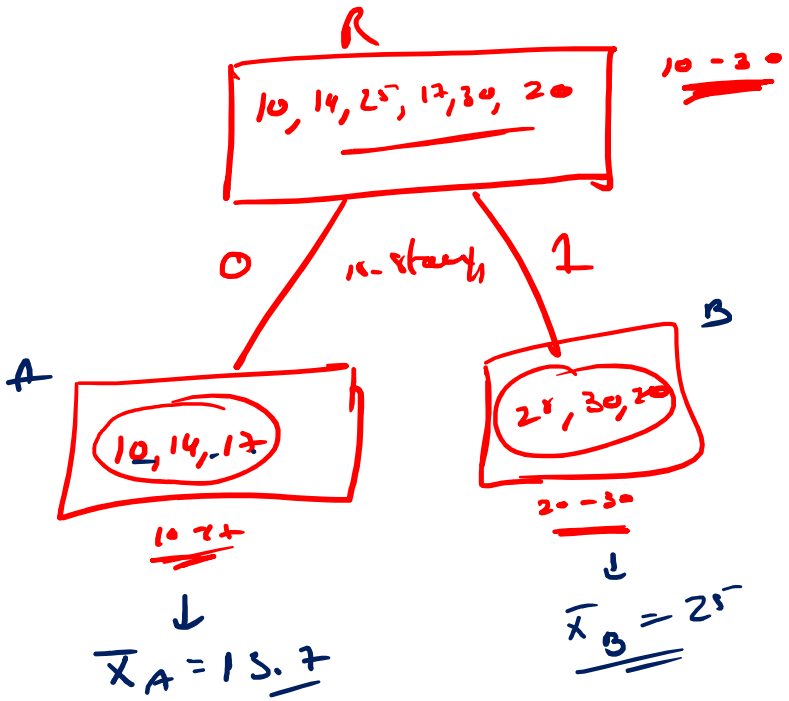
$V.Red_{Age}$, $V.Red_{Diff}$ → 1000

$V.Red_{Age}$ ⇓ 100

which one is more will be the best split

↑ σ^2 → ↑ Impurity

↓ σ^2 → ↓ Impurity



Take the data + Prob Statement

↓

Apply Linear Regression

↓

TV, Sales, Radio

↓

Decision Tree Regressor

↓

Regression Models

↳ R^2, adj R^2

↳ RMSE, MAE, MAPE