Decision Tree

- you want a short tree (rather than a tall tree)
    - allows you to run through the tree faster

Gini impurity

- function:
    - $G_i = 1 - \sum[p_(i,k)^2]$
- the more spread out it is, the more impure it is

CART cost function

- $J(k, t_k) = M_l/mMG_l) + M_r/M(G_r)$
    - where l = leftside, r = rightside

Where to Stop

- Definitely stop when impurity cannot be reduced
- Hyperparameters such as max depth, min samples per leaf
- Stop if improvement is not statistically significant