

## ASSIGNMENT 12: (Spr 2020)

1. For the posted FURNITURE data, run Extree to obtain both an additive tree and an extended tree solution. [NOTE: if you can't get EXTREE to run on your machine, I have posted the EXTREE output for the Sports data, you can just use that output to do the rest of the assignment.]
2. Using an F test, test if the extended tree solution offers improvement in fit (adjusted for # parameters) over the additive tree (see test below). Report your conclusion.
3. Discuss the interpretation of the tree and marked features in the EXTREE solution.

-----  
Alternative formula for  $F_{obs}$  to compare the fit of two nested regression models, where Model A has one or more predictors in addition to the predictors used in Model C:

$$F_{obs} = F^* = \frac{PRE / (PA - PC)}{(1 - PRE) / (N - PA)} = \frac{\Delta R^2 / (PA - PC)}{(1 - R_A^2) / (N - PA)}$$

where  $\Delta R^2 = R_A^2 - R_C^2$  = difference in fit (RSQ) between the two regression models

PA is the number of parameters (#predictors + 1) for Model A (the more complex or “augmented” model)

PC is the number of parameters (#predictors + 1) for Model C (the “compact” model)

$F_{obs}$  is tested against  $F_{crit}$  with (PA-PC) numerator d.f. and (N-PA) denominator d.f.

N is the number of data points (proximities).