Full tree summary

1. Cross validation results

> printcp(fit)

Regression tree:

rpart(formula = GNI.per.capita..constant.2005.US..\_next ~ .,

data = training\_set, method = "anova")

Variables actually used in tree construction:

[1] GNI.per.capita..constant.2005.US..

Root node error: 2.7751e+11/2280 = 121716693

n= 2280

CP nsplit rel error xerror xstd

1 0.849520 0 1.000000 1.000936 0.0276907

2 0.075504 1 0.150480 0.151086 0.0066297

3 0.025513 2 0.074977 0.075797 0.0032171

4 0.014011 3 0.049463 0.050374 0.0024830

5 0.011918 4 0.035452 0.036466 0.0014315

6 0.010000 5 0.023534 0.024377 0.0011325

2. Summary

> summary(fit) # detailed summary of splits

Call:

rpart(formula = GNI.per.capita..constant.2005.US..\_next ~ .,

data = training\_set, method = "anova")

n= 2280

CP nsplit rel error xerror xstd

1 0.84951970 0 1.00000000 1.00093640 0.027690695

2 0.07550373 1 0.15048030 0.15108625 0.006629717

3 0.02551317 2 0.07497657 0.07579695 0.003217050

4 0.01401112 3 0.04946340 0.05037436 0.002483007

5 0.01191811 4 0.03545228 0.03646629 0.001431486

6 0.01000000 5 0.02353417 0.02437688 0.001132489

Variable importance

GNI.per.capita..constant.2005.US..

19

GDP.per.capita..constant.2005.US..

18

CountryName

16

Birth.rate..crude..per.1.000.people.

13

Population.ages.65.and.above....of.total.

13

Age.dependency.ratio..old....of.working.age.population.

13

Adjusted.net.national.income.per.capita..constant.2005.US..

2

Household.final.consumption.expenditure.per.capita..constant.2005.US..

2

GNI.per.capita..Atlas.method..current.US..

2

GDP.per.capita..current.US..

1

Node number 1: 2280 observations, complexity param=0.8495197

mean=9388.39, MSE=1.217167e+08

left son=2 (1560 obs) right son=3 (720 obs)

Primary splits:

GNI.per.capita..constant.2005.US.. < 13143.51 to the left, improve=0.8495197, (0 missing)

GDP.per.capita..constant.2005.US.. < 12222.06 to the left, improve=0.8454019, (0 missing)

CountryName splits as LLRRRLLLLLRLLLLLLRLLLLLRRRRRLLRRLRRLLLRRLLLLLLLLLLLRLRRLLLRLLLLLRLRLLLLRRLLL, improve=0.8397542, (0 missing)

Adjusted.net.national.income.per.capita..constant.2005.US.. < 10648.21 to the left, improve=0.8157672, (142 missing)

Household.final.consumption.expenditure.per.capita..constant.2005.US.. < 7279.751 to the left, improve=0.8157559, (125 missing)

Surrogate splits:

GDP.per.capita..constant.2005.US.. < 13096.31 to the left, agree=0.992, adj=0.974, (0 split)

CountryName splits as LLRRRLLLLLRLLLLLLRLLLLLRLRRRLLRRLRRLLLRRLLLLLLLLLLLRLRRLLLRLLLLLRLRLLLLRRLLL, agree=0.984, adj=0.950, (0 split)

Birth.rate..crude..per.1.000.people. < 16.30539 to the right, agree=0.943, adj=0.818, (0 split)

Population.ages.65.and.above....of.total. < 9.705437 to the left, agree=0.941, adj=0.814, (0 split)

Age.dependency.ratio..old....of.working.age.population. < 14.167 to the left, agree=0.937, adj=0.800, (0 split)

Node number 2: 1560 observations, complexity param=0.02551317

mean=2480.176, MSE=6799668

left son=4 (1254 obs) right son=5 (306 obs)

Primary splits:

GNI.per.capita..constant.2005.US.. < 4043.95 to the left, improve=0.6674781, (0 missing)

GDP.per.capita..constant.2005.US.. < 4302.384 to the left, improve=0.6668615, (0 missing)

CountryName splits as LR---LLRLL-RLLLLL-LLLLL-R---RL-RL--LLL--LRRRLLLLRRL-L--LLR-LLLRLRL-LLRL--RRL, improve=0.6323191, (0 missing)

Household.final.consumption.expenditure.per.capita..constant.2005.US.. < 1870.33 to the left, improve=0.5586147, (124 missing)

Adjusted.net.national.income.per.capita..constant.2005.US.. < 2415.803 to the left, improve=0.5281439, (80 missing)

Surrogate splits:

GDP.per.capita..constant.2005.US.. < 4185.147 to the left, agree=0.997, adj=0.984, (0 split)

CountryName splits as LR---LLLLL-LLLLLL-LLLLL-R---RL-RL--LLL--LRLLLLLLRRL-L--LLR-LLLRLRL-LLRL--LRL, agree=0.956, adj=0.778, (0 split)

Adjusted.net.national.income.per.capita..constant.2005.US.. < 3574.087 to the left, agree=0.917, adj=0.575, (0 split)

GDP.per.capita..current.US.. < 2889.287 to the left, agree=0.917, adj=0.575, (0 split)

GNI.per.capita..Atlas.method..current.US.. < 2790 to the left, agree=0.913, adj=0.559, (0 split)

Node number 3: 720 observations, complexity param=0.07550373

mean=24356.19, MSE=4.326794e+07

left son=6 (398 obs) right son=7 (322 obs)

Primary splits:

GNI.per.capita..constant.2005.US.. < 24367.23 to the left, improve=0.6725966, (0 missing)

GDP.per.capita..constant.2005.US.. < 24792.82 to the left, improve=0.6696300, (0 missing)

Household.final.consumption.expenditure.per.capita..constant.2005.US.. < 13516.11 to the left, improve=0.6520190, (1 missing)

Adjusted.net.national.income.per.capita..constant.2005.US.. < 20438.31 to the left, improve=0.5816414, (62 missing)

GDP.per.capita..current.US.. < 16450.51 to the left, improve=0.5626488, (0 missing)

Surrogate splits:

GDP.per.capita..constant.2005.US.. < 24336.85 to the left, agree=0.986, adj=0.969, (0 split)

Household.final.consumption.expenditure.per.capita..constant.2005.US.. < 13599.71 to the left, agree=0.943, adj=0.873, (0 split)

Adjusted.net.national.income.per.capita..constant.2005.US.. < 20438.31 to the left, agree=0.893, adj=0.761, (0 split)

GDP.per.capita..current.US.. < 13471.56 to the left, agree=0.868, adj=0.705, (0 split)

GNI.per.capita..Atlas.method..current.US.. < 13355 to the left, agree=0.865, adj=0.699, (0 split)

Node number 4: 1254 observations

mean=1427.792, MSE=1260614

Node number 5: 306 observations

mean=6792.888, MSE=6360818

Node number 6: 398 observations, complexity param=0.01191811

mean=19503.89, MSE=1.099668e+07

left son=12 (193 obs) right son=13 (205 obs)

Primary splits:

GNI.per.capita..constant.2005.US.. < 18967.26 to the left, improve=0.7556970, (0 missing)

GDP.per.capita..constant.2005.US.. < 19154.49 to the left, improve=0.7529048, (0 missing)

Household.final.consumption.expenditure.per.capita..constant.2005.US.. < 10823.64 to the left, improve=0.7056150, (1 missing)

Adjusted.net.national.income.per.capita..constant.2005.US.. < 16216.79 to the left, improve=0.5908564, (47 missing)

CountryName splits as --RRR-----R------------RLRRR--RL-RR---RR-L------L--R-RR--LL-----L-R----RR---, improve=0.4527160, (0 missing)

Surrogate splits:

GDP.per.capita..constant.2005.US.. < 19149.22 to the left, agree=0.985, adj=0.969, (0 split)

Household.final.consumption.expenditure.per.capita..constant.2005.US.. < 10865.26 to the left, agree=0.935, adj=0.865, (0 split)

Adjusted.net.national.income.per.capita..constant.2005.US.. < 16216.79 to the left, agree=0.842, adj=0.674, (0 split)

CountryName splits as --RRR-----R------------RLLRR--RL-LR---RR-L------L--R-RR--LL-----L-R----RR---, agree=0.771, adj=0.528, (0 split)

GNI.per.capita..Atlas.method..current.US.. < 7130 to the left, agree=0.739, adj=0.461, (0 split)

Node number 7: 322 observations, complexity param=0.01401112

mean=30353.74, MSE=1.808351e+07

left son=14 (215 obs) right son=15 (107 obs)

Primary splits:

GNI.per.capita..constant.2005.US.. < 30957.83 to the left, improve=0.6677584, (0 missing)

GDP.per.capita..constant.2005.US.. < 31256.14 to the left, improve=0.6568353, (0 missing)

Adjusted.net.national.income.per.capita..constant.2005.US.. < 25671 to the left, improve=0.6287778, (15 missing)

Household.final.consumption.expenditure.per.capita..constant.2005.US.. < 16706.67 to the left, improve=0.5218500, (0 missing)

GNI.per.capita..Atlas.method..current.US.. < 25120 to the left, improve=0.4934209, (0 missing)

Surrogate splits:

GDP.per.capita..constant.2005.US.. < 31106.29 to the left, agree=0.981, adj=0.944, (0 split)

Adjusted.net.national.income.per.capita..constant.2005.US.. < 25892.28 to the left, agree=0.910, adj=0.729, (0 split)

Household.final.consumption.expenditure.per.capita..constant.2005.US.. < 17401.27 to the left, agree=0.873, adj=0.617, (0 split)

GNI.per.capita..Atlas.method..current.US.. < 25120 to the left, agree=0.851, adj=0.551, (0 split)

GDP.per.capita..current.US.. < 24027.73 to the left, agree=0.848, adj=0.542, (0 split)

Node number 12: 193 observations

mean=16532.89, MSE=2759925

Node number 13: 205 observations

mean=22300.98, MSE=2617417

Node number 14: 215 observations

mean=27902.28, MSE=3661250

Node number 15: 107 observations

mean=35279.56, MSE=1.072371e+07