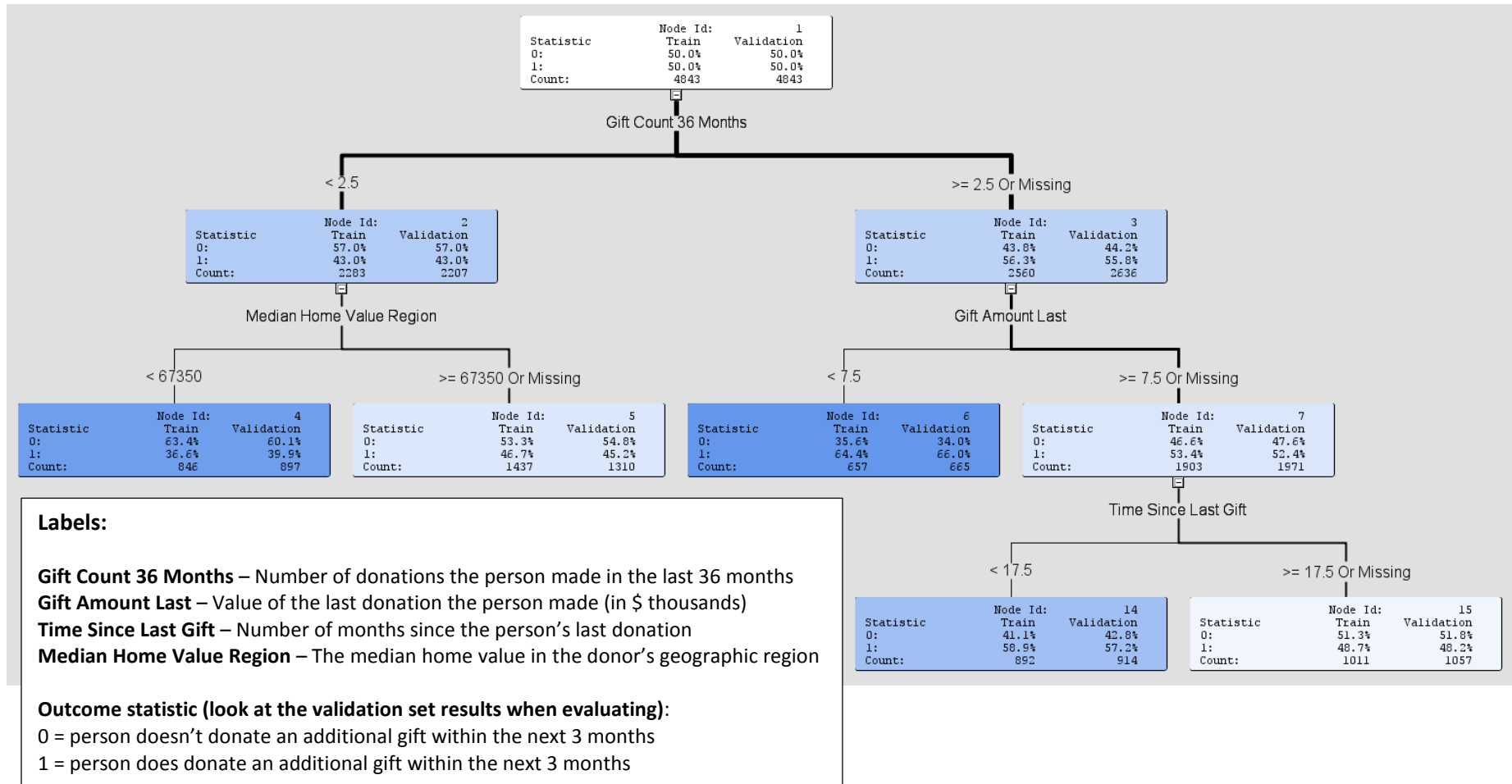


In-Class Exercise: Reading Decision Tree Output

The following is the output from a decision tree analysis in SAS Enterprise Miner. It uses the same data set (a charity donation database) you'll be using in your in-class exercise, but I've changed the model parameters so we get some different results.



Labels:

- Gift Count 36 Months** – Number of donations the person made in the last 36 months
- Gift Amount Last** – Value of the last donation the person made (in \$ thousands)
- Time Since Last Gift** – Number of months since the person's last donation
- Median Home Value Region** – The median home value in the donor's geographic region

Outcome statistic (look at the validation set results when evaluating):
 0 = person doesn't donate an additional gift within the next 3 months
 1 = person does donate an additional gift within the next 3 months

Answer the following questions about this tree on the next page:

1. What is the probability that the following people will donate an additional gift in the next 3 months?
 - a. Someone who has made two donations in the last 36 months and the median home value in their region is \$85,000.
 - b. Someone who has made five donations in the last 36 months and their last gift amount was \$3,000.
 - c. Someone who has made three donations in the last 36 months, their last gift amount was \$10,000, and it's been 12 months since their last donation.
 - d. Someone who has made three donations in the last 36 months, their last gift amount was \$10,000, and it's been 24 months since their last donation.

2. What does it mean that median home value isn't a split variable when more than three donations were made in the last 36 months?

3. Check out this screen shot from SAS Enterprise Miner. This is the list of logworth values used to choose the first split variable for the data.
 - a. According to this list, which variable is the best one to choose for that first split? Why?
 - b. How does this correspond to the final decision tree on the previous page?

Split Node 1

Target Variable: TargetB

Variable	Variable Description	-Log(p)	Branches
GiftCnt36	Gift Count 36 Months	18.4407	2
GiftAvgCard36	Gift Amount Average Card 3...	17.2404	2
GiftAvgLast	Gift Amount Last	15.0207	2
GiftAvg36	Gift Amount Average 36 Mo...	14.6342	2
GiftAvgAll	Gift Amount Average All Mon...	13.9546	2
GiftCntCard36	Gift Count Card 36 Months	12.7354	2
GiftCntCardAll	Gift Count Card All Months	11.909	2
StatusCatStarAll	Status Category Star All Mo...	10.8502	2
StatusCat96NK	Status Category 96NK	9.5243	2
GiftCntAll	Gift Count All Months	9.2782	2
PromCntCard36	Promotion Count Card 36 Mo...	8.7291	2
GiftTimeFirst	Time Since First Gift	6.6932	2
GiftTimeLast	Time Since Last Gift	5.7032	2
PromCntAll	Promotion Count All Months	5.4915	2
PromCntCardAll	Promotion Count Card All Mo...	4.5925	2
DemMedHomeValue	Median Home Value Region	4.4604	2
PromCntCard12	Promotion Count Card 12 Mo...	3.8675	2
PromCnt36	Promotion Count 36 Months	3.3498	2
PromCnt12	Promotion Count 12 Months	2.6075	2
DemPctVeterans	Percent Veterans Region	1.9857	2
DemAge	Age	1.5129	2

Buttons: Edit Rule..., OK, Cancel, Apply, Refresh