**Case Problem**

**Predictive Analytics**

**Module 2 – Lesson 4**

**Professor: Lee**

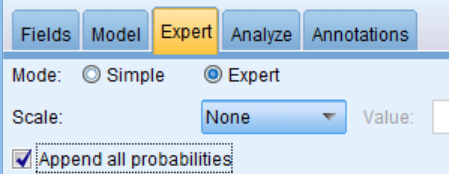
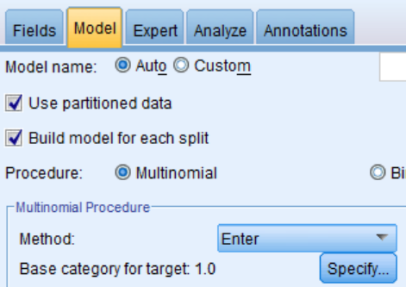
**PART ONE**

Use the FinancialCondition.xlsx which contains two common measures used to assess the financial condition of banks:

* ‘TotExp\_Assets’ is the ratio of Total Expenses to Total Assets
* “TotLnsandLses\_Assets” is the ratio of Total Loans and Leases to Total Assets.

The response is the financial condition (1=weak and 0=otherwise).

Remember to choose Append all probabilities and set the Base category to 1.0:

 ****

1. Write the estimated equation that associates the financial condition of a bank with its two predictors in three formats
   1. the logit

logit= 14.88 + (-79.964)\*TotExp\_assets + (-9.173)\*TotLnsandLses\_Assests

* 1. the odds

14.88 + (-79.964)\*TotExp\_assets + (-9.173)\*TotLnsandLses\_Assests

odds = e

* 1. probability of getting class=0

= (1/1+e^(-(14.88 + (-79.964)\*TotExp\_assets + (-9.173)\*TotLnsandLses\_Assests))

1. Consider a new bank whose total loans-&-leases-to-assets ratio = 0.6 and total-expenses-to-assets is 0.11. From your logistic regression model, estimate the following:
   1. Logit

logit= 14.88 + (-79.964)\*TotExp\_assets + (-9.173)\*TotLnsandLses\_Assests

logit = 14.88 + (-79.964)\*0.11+(-9.173)\*0.6

= 0.5842

* 1. Odds

14.88 + (-79.964)\*TotExp\_assets + (-9.173)\*TotLnsandLses\_Assests

Odds = e

= e^(0.5842) = 1.793

* 1. Probability of having class=0

= 1/(1+e^(-0.5842))

=1 /(1+.5575)

=.6420

d)Predicted classification of the bank

Predicted classification of the bank is 0 as probability is .6420, which is greater than 0.5.

The cutoff value of 0.5 is used in conjunction with the probability of being financially weak. Compute the threshold that should be used if we want to make a classification based on

* 1. the odds of being financially weak

Probability of success/ Probability of failure = 0.5 /0.5=1

* 1. the corresponding logit.

Log(1) = 0

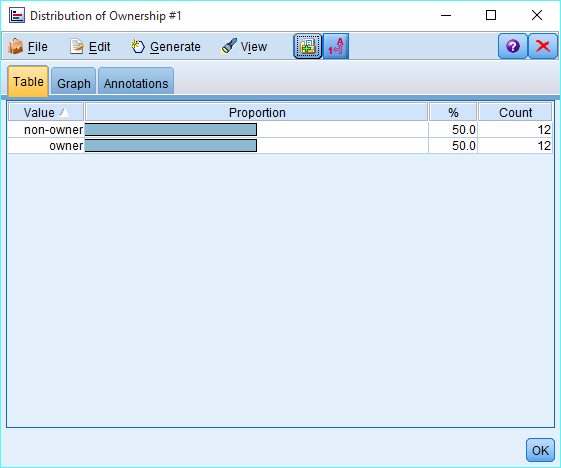
**PART TWO**

Use the RidingMowers.txt data which contains information about who the best sales prospects are for an intensive sales campaign.

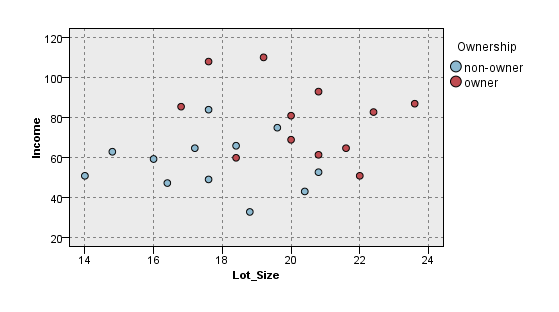
* The INCOME (in 1000s) and the LOT\_SIZE (in 1000 ft2) for a random sample of 24 homes is listed along with whether those home had a riding lawn mower.

1. What percentage of households in the study were owners of a riding mower?

**50 %**



1. Create a scatterplot of Lot\_Size by Income. Overlay Ownership by Color. Which class seems to have the higher income? Which class seems to have the larger lot size?



From scatter plot, Owner class seems to have higher income

From scatter plot, Owner class seems to have larger lot size

1. Run a logistic regression predicting ownership with INCOME and LOT\_SIZE. Append All Probabilities. Set the base category to owner. Create a confusion matrix. What is Error Rate?

|  |  |  |
| --- | --- | --- |
|  | PREDICTED | |
| ACTUAL | Owner | Non-owner |
| Owner | 10 | 2 |
| Non-owner | 2 | 10 |

Error rate = (False Positive + False Negative) / Total = 4/24 = 16.6%

1. What are the odds that a household with $50,000 income and a lot size of 20,000 sq ft is a non-owner? (Note: Keep all your decimal places to get accuracy)

25.94 + (-.1109)\*Income + (-.9638)\*Lot\_size

Odds = e

25.94 + (-.1109)\*50+ (-.9638)\*20

= e

25.94 + (-.1109)\*50 + (-.9638)\*20

= e

= e^(1.119)

=3.061

Odds of being a owner is greater than 3 which means there is chances of being an owner is 3.061 to 1 Or we can say odds of being a non owner is 1 is to 3.061 i.e .3266

1. Assuming a cutoff of .5, create a rule for classification of non-owner based on the odds. What is the classification for a household with $60,000 income and a lot size of 20,000 sq ft?

25.94 + (-.1109)\*60 + (-.9638)\*20

Odds= e

Odds=1.01

Logit=25.94 + (-.1109)\*60 + (-.9638)\*20

logit=0.01

P(0)=1/1+e^(-0.01) = .5025

Since probability is greater than 0.5 , it means it is non owner

1. What is the minimum income that a household with 16,000 sq ft lot size should have before it is classified as an owner?

Logit =ln(odds)=ln(p/(1-p)) so ln(1)=0

0=25.94 + (-.1109)\*Income + (-.9638)\*16

15.4208-25.94=-.1109\*income

-10.51=-.1109\*income

Income = 10.51/.1109 = 94.7

And sine income is in 1000 = 94770 is the minimum income to be classified as owner

**Submit your completed Word file and your stream file on Canvas. In the comments, put the names of the folks who worked on this assignment.**