Bookbinders Book Club Case

Assignment 2: Identifying Target Customers
(Individual assignment: 12% of your course grade)

Due on 10/14 at the beginning of class

Case & Data: The case and the Excel data file for this homework is available at the Decision Pros website. You can also find them on Blackboard. There are two data files related to this homework assignment, one called "Estimation Sample" and the other called "Holdout Sample".

BBB Club's management wants to know whom to send a specially produced brochure promoting a new art book *The Art History of Florence*. Thus, it runs an experiment, as we discussed in class. You will work through the analysis.

Question 1: Estimation using the Estimation Sample
Using ME-XL Excel Add-In and the data in the Estimation Sample, estimate a logit model predicting probability of response as a function of all the available variables:

- Gender
- Amt_purchased
- Frequency
- Last_Purchase
- First_purchase
- P_Child
- P_Youth
- P_Cook
- P_DIY
- P_Art

To do this, select ME-XL→ Customer Choice (Logit). Once you've filled in your information properly, you should be able to run a logit model.

Look at your logit model results. **Report only the items that are requested below.**

(a) (5 points) **Report:** Logit model coefficients and t-statistics. Interpret the sign of each coefficient. Specifically, note which variables increase the probability that the customer makes a purchase and which decrease this probability. Highlight which factors most influenced the customers' decision to buy or not to buy. *Hint: you do not need to compute any probabilities because higher coefficients imply higher probabilities*.

Question 2: Applying the scoring model to the Holdout Sample
These steps follow the same process we used in our in-class RFM exercise. Unless a particular step says "report", you do not need to report anything in your write-up.
Hint: please refer to the Excel file "Bookbinders Book Club Student Worksheets" in Blackboard for useful formulae

(c) Compute a logit score for each person in the Holdout Sample. *Hint: the instructor already provided the logit score calculation for the first respondent for your reference.*

(d) Sort the 2300 prospects in the Holdout Sample in decreasing order of logit scores.

(e) (5 points) Compute the Decile Scores based on logit scores. Create a plot with the deciles on the horizontal axis and the percent cumulative response on the vertical axis. This is the Lift Curve.
Report: The decile scores and the lift curve plot. Comment on this plot. What conclusions can you draw about the quality of your scoring rule? Is it a good rule? How does it compare with the lift curve from the RFM exercise we did in class? *Hint: see instructor's formulae in the "RFM sorted" worksheet as guidelines*

(f) (10 points) Bookbinders is considering a similar mail campaign in the Midwest where it has data for 50,000 customers. Such mailings typically promote several books. The allocated cost of the mailing is $0.65/addressee (including postage) for the art book, and the book costs Bookbinders $15 to purchase and mail. The company allocates overhead to each book at 45 percent of cost. The selling price of the book is $31.95. Based on the logit model, how many percentiles of customers should Bookbinders target? How would Bookbinders' profit be at this percentile? How much more profit would you expect the company to generate using this model as compared to sending the mail offer to the entire list? *Hint: see instructor's formulae in the "Model Evaluator" worksheet as guidelines*

Question 3: Repeat analyses listed in Q1 and Q2 above using a regression model.

Answer the following questions:
(l) (5 points) **Report:** regression analysis results (be sure to report both coefficients and t-statistics); interpret the sign of each coefficient; and the decile scores and the lift curve plot. How many percentiles of customers should Bookbinders target? How much profit will the company make based on the regression model?

(m) (5 points) compare and contrast with results from the logit; the regression; and the RFM analysis. Summarize the advantages and limitations of each of the modeling approaches.