

Analytics for Marketing Decisions - MKTG 464

Individual Assignment 4 (Total 12 Points)

Note:

- Each student is expected to work on this problem set **independently** and without any help from other students.
- You need to show the relevant R output and also **briefly explain your results in plain English**. You must also **clearly explain the managerial and/or marketing implications of the results**. Don't just report the final numbers / output as your answers.
- If after giving some thought, you have problems doing this assignment, do not hesitate to contact me (sumond@uw.edu). You can also ask clarifying questions on the Canvas Discussion board for assignments.

Questions

1. Revisiting the Conjoint Study for All-in-one Printers (3 Points)

Recall the Conjoint Analysis survey on All-in-one printers from Individual Assignment 1. Also recall, that in that survey, ratings data (on a scale of 0 - 100) for the 16 profiles were collected from 35 respondents. The relative attribute importance for the respondents, resulting from their utility functions, are provided to you in the file, **Printers_RAI.csv**. Use this data to run a Cluster Analysis (Don't standardize the variables). Find k-means segments based on individuals' relative attribute importance. Consider the three-segment solution. Describe the three segments in plain English.

2. Revisiting Consumers for the ConneCtor PDA (4 Points)

Recall the ConneCtor PDA problem from Individual Assignment 3. Suppose, based on the Needs of the respondents in the survey ($X_1 - X_{10}$), they were divided into **two broad segments**. The segment memberships of the respondents, and their demographic information ($Z_1 - Z_5$), are

provided in the file **PDA_2seg.csv**. Discuss how well you can use the demographic variables to discriminate between the two segments (estimate a suitable binary logistic regression model to predict consumers' segment membership using the Stepwise variable selection approach and report the classification accuracy). Interpret the estimated model.

Note: Split the data set into a 'Training Sample' (70%) and a 'Testing Sample' (30%).

Note: Use any five of the six occupation category dummies as potential inputs in your regression model.

3. Identifying Loyal Customers (5 Points)

Dermaglow, a skincare spa chain wants to be able to identify its loyal customers (Those who visit only *Dermaglow* for their skincare needs). They have data for 348 customers who are known to be either loyal or not loyal. The data is available in the file, **Dermaglow.csv**. It includes the following variables:

Loyal: Coded as 1 if customer is known to be loyal and 0 if customer is not loyal

Avg.spent: Avg. amount spent per visit by the customer at a *Dermaglow* spa (in \$)

Intervisit.time: Average time interval (in weeks) between visits to a *Dermaglow* spa

Mincome: Monthly income of customer (in \$1000s)

Rating: Customer's satisfaction rating for *Dermaglow* (on a scale of 1-100)

Note: Split the data set into a 'Training Sample' (70%) and a 'Testing Sample' (30%).

- a. Estimate a binary logistic regression model to predict *Dermaglow's* loyal customers using all input variables in the model. Report the classification accuracy in the Training and Testing samples.
- b. Now, use the Stepwise Model Selection approach. Report the classification accuracy in the Training and Testing samples for the selected model. Report and interpret the selected

model. Using this model, calculate the predicted probability that a customer that spends an average of \$50 per visit, visits once in six weeks, has a monthly income of \$6,000, and has a 70% satisfaction rating, will be a Loyal customer.

- c. Plot the ROC curves, and report the AUC numbers for the models in 3(a) and 3(b)?