**SAS HOMEWORK 3 (Group Homework)**

**Note : This data has been provided for teaching purpose in this course and may not be used in any other research or distributed to anyone else without prior permission.**

I have provided a dataset of 7401 credit card customers. Of these credit card customers some are inactive (i.e., have never used the card) and the rest are active. We have the following variables.

1. The mode of acquisition (whether they were acquired through direct mail (DM), direct selling (DS), telephone sales (TS) or through internet (NET))
2. Whether they have a Reward card (i.e., a card that gives points for every dollar purchased)
3. Whether they have an affinity card and the type of affinity card they have.
4. The type of card that they were given (that is, whether they have a standard, gold, platinum or quantum card). Note: Quantum > Platinum > Gold > Standard card in terms of credit worthiness.
5. Note that profit = totfc + 1.6%\*TotalTrans (approximately)

|  |  |  |
| --- | --- | --- |
|  | **HID** | ID of the account |
|  | **Rewards** | whether the customer has a reward card (=1) or not (=0) |
|  | **Limit** | credit limit of the customer |
|  | **numcard** | number of cards that the customer has from this bank |
| **Mode of acquisition** | **DM** | whether the customer was acquired though direct mail (1=Yes, 0=No) |
| **DS** | whether the customer was acquired though direct selling (1=Yes, 0=No) |
| **TS** | whether the customer was acquired though telephone selling (1=Yes, 0=No) |
| **NET** | whether the customer was acquired though internet (1=Yes, 0=No) |
| **Type of card** | **Gold** | whether the customer has a GOLD card (1=Yes, 0=No) |
| **Platinum** | whether the customer has a PLATINUM card (1=Yes, 0=No) |
| **Quantum** | whether the customer has a QUANTUM card (1=Yes, 0=No) |
| **Standard** | whether the customer has a STANDARD card (1=Yes, 0=No) |
|  | **Totaltrans** | Total transaction amount (money spent) by the customer over a 3 year period |
|  | **Totfc** | Total finance charges paid by the customer over a 3 year period |
|  | **Age** | Age in years |
|  | **Dur** | Duration: Number of months a customer has stayed with the firm |
| **Types of Affinity cards** | **sectorA** | No affinity – card is not associated with affinity to an organization |
| **SectorB** | Affinity card affiliated with Professional organization (e.g. Am. Medical. Assoc) if a customer has an affinity card of this type value =1 else 0. |
| **SectorC** | Affinity card affiliated with Sports |
| **SectorD** | Affinity card affiliated with Financial institution |
| **SectorE** | Affinity card affiliated with University (e.g. UTD card) |
| **SectorF** | Affinity card affiliated with Commercial (e.g. Macy’s card) |

1. If profit is negative, set it to 0, since profit cannot be negative.
2. IF TOTTRANS=0 THEN CREATE A NEW VARIABLE CALLED ACTIVE THAT TAKES THE VALUE=0, ELSE IT TAKES THE VALUE =1. Only active customers have positive transactions.
3. Create a new variable climit = limit/10000
4. Create a new variable ttrans=tottrans/10000;
5. Create a new variable profit that is = totfc + 1.6%\*TotalTrans

Run the following Tobit model (Use PROC QLIM)

Model profit = age ttrans rewards climit numcard, modes of acquisition, type of card, types of affinity

1. Write a summary of the results. Focus on important effects, interpretation, model fit etc.
2. Which mode of acquisition generates the highest profit?
3. Order the modes of acquisition from high to low in terms of profit.

2. Run a selection model (Use PROC QLIM)

Model active = age, rewards, climit, numcard, modes of acquisition, type of card, types of affinity

Model totfc = age, ttrans, rewards, climit, numcard, modes of acquisition, type of card, types of affinity

1. Write a summary of the results. Focus on important effects, interpretation, model fit etc.
2. Which type of affinity card generates the most profit? Order the types from best to worst in terms of profit.

3: Survival analysis

Note that duration is censored if its value is 37 as we have only 37 months of data.

Create a new variable Censor which takes the value=1 if dur=37 (the maximum value) and value=0 otherwise. Use this as a censoring variable.

1. Run a proportional hazards model (PROC PHREG)

**Duration** = age, ttrans, rewards, climit, numcard, modes of acquition, type of card, types of affinity

Write a summary of the results. Focus on important effects, interpretation, model fit etc.

4. Run the same model as above using PROC LIFEREG with Weibull distribution.

Write a summary of the results. Focus on important effects, interpretation, model fit etc.

5. Use PROC LIFETEST to test whether survivor function of affinity groups are significantly different from that of non-affinity groups. (that is compare sectorA with other sectors)

What do you conclude?