Classification

Charles Book Club (CBC) is a leading book store in London. A new title, "The Art History of Florence", is ready for release and CBC is looking to get it to the right customers in the market. CBC always sends a random mailers to some of its customers randomly to gauge the reaction of readers.

For "The Art History of Florence", CBC has a mailing budget of 250,000 GBP. Cost of creative agency to get a copy of letter is 10,000GBP. Print and postage cost of one mailer is 1GBP. Other administrative cost for mailing is 20,000 GBP. CBC has a total customer base of 500,000 and CBC would want to prioritize its mailing to reach out to those customers who are most likely to buy this book.

This time CBC has sent a test mailing to a random sample of 1500 customers from its customer base.The customer responses have been collated with past purchase data.Each row (or case) corresponds to one market test customer. Each column is a variable with the header row giving the name of the variable. The variable names and descriptions are given in below:

|  |  |
| --- | --- |
| **Seq#** | Sequence number in the partition |
| **ID#** | Identification number in the full (unpartitioned) market test data set |
| **Gender** | O=Male 1=Female |
| **M** | Monetary- Total money spent on books |
| **R** | Recency- Months since last purchase |
| **F** | Frequency - Total number of purchases |
| **FirstPurch** | Months since first purchase |
| **ChildBks** | Number of purchases from the category: Child books |
| **YouthBks** | Number of purchases from the category: Youth books |
| **CookBks** | Number of purchases from the category: Cookbooks |
| **DoItYBks** | Number of purchases from the category Do It Yourself books |
| **RefBks** | Number of purchases from the category: Reference books (Atlases, Encyclopedias, Dictionaries) |
| **ArtBks** | Number of purchases from the category: Art books |
| **GeoBks** | Number of purchases from the category: Geography books |
| **ItalCook** | Number of purchases of book title: "Secrets of Italian Cooking." |
| **ItalAtlas** | Number of purchases of book title: "Historical Atlas of Italy." |
| **ItalArt** | Number of purchases of book title: "Italian Art." |
| **Florence** | =1  "The Art History of Florence." was bought, =0  if not |
| **Related purchase** | Number of related books purchased |

Fit a logistic regression on the given data to

* Assign probability of purchase to each record
* Identify variables which have influence on purchase likelihood

If the fitted model is used to score the whole population, how many customers can CBC reach out to? What would be the expected number of people who would respond to CBC offer?

Evaluation

**Maximum Score: 40**

Methodology &Step by Step Process used: 15 Marks

Quality of final model and variables used: 10 Marks

Explanation of Model outcome, steps and its implementation on whole population: 15 Marks

**LOGIT REGRESSION REPORT**