Regression Assignment

Background:

A bank wants to understand how customer banking habits contribute to revenues and profitability. The bank has customer age and bank account information (see table below).

The goal is to build a model that allows the bank to predict profitability for a given customer. A surrogate for customer profitability available in our data set is the Total Revenue a customer generates through their accounts and transactions. The resulting model will be used to forecast bank revenues and guide the bank in future marketing campaigns.

The Data:

|  |  |
| --- | --- |
| Rev\_Total | Total revenue generated by the customer over a 6 month period |
| Bal\_Total | Total of all account balances, across all counts held by the customer |
| Offer | An indicator of whether the customer has received a special promotional offer in the previous one month period (1 if offer was received, 0 if it was not) |
| AGE | Customer’s age |
| CHQ | Debit card activity: 0 = low or no activity, 1 = higher activity |
| CARD | Credit card activity: 0 = low or no activity, 1 = higher activity |
| SAV1 | Primary savings account activity: 0 = low or no activity, 1 = higher activity |
| LOAN | Personal loan account activity: 0 = low or no activity, 1 = higher activity |
| MORT | Mortgage account tier: 0 = lower tier, less important to bank portfolio  1 = higher tier, important to bank portfolio |
| INSUR | Insurance account activity: 0 = low or no activity, 1 = higher activity |
| PENS | Pension (retirement) account tier: 0 = lower balance / less important  1 = higher tier / more important to bank |
| Check | Checking account activity: 0 = low or zero activity, 1 = higher activity |
| AccountAge | Number of years as a customer of the bank |

Your task:

Build a multiple linear regression model to predict total revenue.

* Start by fitting the full model to the data
* Examine the residuals and make any necessary transformations
* If you transform the data, check to see whether the numeric predictor variables (Bal\_Total, AGE, or AccountAge) need to be transformed in order to maintain a linear relationship with the transformed response
* Reduce the model using forward, backward or stepwise regression
* Make any necessary modifications to arrive at the final model
* Write a short paragraph explaining what the bank can learn from this model. In particular, try to answer the question: What can the bank do to increase customer profitability (e.g., total revenue)?

Your deliverable:

You may either submit a knitted html file that is well-documented and easy to follow – OR you may copy and paste the relevant information into a Word document for submission.