

BRIEF CASE



Building Models

The Paralyzed Veterans of America (PVA) is a philanthropic organization sanctioned by the U.S. government to represent the interests of those veterans who are disabled. (For more information on the PVA see the opening of Chapter 21.) To generate donations, the PVA sends out greeting cards and mailing address labels periodically with their requests for donations. To increase their efficiency, they would like to be able to model the amount of donations based on past giving and demographic variables on their donors. The dataset **PVA** contains data on 3648 donors who gave to a recent solicitation. There are 26 predictor variables and 1 response variable. The response variable (**GIFTAMNT**) is the amount of money donated by the donor to the last solicitation. Find a model from the 26 predictor variables using any model selection procedure you like to predict this amount. The variables include:

Variables based on the donor's ZIP code

- MALEVET (% Male veterans)
- VIETVETS (% Vietnam Veterans)
- WWIIVETS (% WWII veterans)
- LOCALGOV (% Employed by local government)
- STATEGOV (% Employed by state government)
- FEDGOV (% Employed by federal government)

Variables specific to the individual donor

- HOMEOWNER (N = No; Y = Yes)
- HIT (Number of times donor has responded to mail order offers other than PVAs)
- CARDPROM (Number of card promotions received lifetime)
- MAXADATE (Date of most recent promotion received in YYMM Year Month format)
- NUMPROM (Number of promotions received lifetime)
- CARDPRM12 (Number of card promotions received in last 12 months)
- NUMPRM12 (Number of promotions received in last 12 months)
- NGIFTALL (Number of gifts given lifetime to date)
- CARDGIFT (Number of gifts to card promotions given lifetime to date)

(continued)

BRIEF CASE *(continued)*

- MINRAMNT (Amount of smallest gift to date in \$)
- MINRDATE (Date associated with the smallest gift to date—YYMM format)
- MAXRAMNT (Amount of largest gift to date in \$)
- MAXRDATE (Date associated with the largest gift to date—YYMM format)
- LASTGIFT (Amount of most recent gift in \$)
- AVGGIFT (Average amount of gifts to date in \$)
- CONTROLN (Control number—unique record identifier)
- HPHONE_D (Indicator variable for presence of a published home phone number: 1 = Yes; 0 = No)
- CLUSTER2 (Marketing Cluster Code—nominal field)
- CHILDREN (Number of children living at home)

Response variable

- GIFTAMNT (Response variable—amount of last gift in \$)

Be sure to include exploratory data analysis and evaluate the relationship among these variables using graphical and correlation analysis to guide you in building your regression models. Write a report summarizing your analysis.