

# PVA\_RAW\_DATA Description

A national veterans' organization seeks to better target its solicitations for donation. By soliciting only the most likely donors, less money is spent on solicitation efforts and more money is available for charitable concerns. Solicitations involve sending a small gift to an individual with a request for donation. Gifts include mailing labels and greeting cards.

The organization has more than 3.5 million individuals in its mailing database. These individuals were classified by their response behaviors to previous solicitation efforts. Of particular interest is the class of individuals who were identified as lapsing donors. These individuals made their most recent donations between 12 and 24 months ago. The organization found that by predicting the response behavior of this group, they could use the model to rank all 3.5 million individuals in their database. With this ranking, a decision can be made to either solicit or ignore an individual in the current solicitation campaign. The current campaign refers to a greeting card mailing sent in June of 1997. It is identified in the raw data as the 97NK campaign.

The raw analysis data were reduced for this course. A subset of slightly more than 19,000 records was selected for modeling. This subset was not chosen arbitrarily. In addition, the 481 fields were reduced to 47. Some of the fields were eliminated when their potential association with the analysis objective was considered. (For example, it is doubtful that CD player ownership is strongly correlated with donation potential.) Other fields were combined to form summaries of a specific customer behavior. In general, if the variable has **PROM** in its name, then the variable is related to the number of items that the organization sent to the customer. However, if the variable has **GIFT** in its name, then the variable is related to the amount of money that the customer sent to the organization.

The data are in a data set called **pva\_raw\_data**. The following table details the variables and descriptions:

Variable	Description
CARD_PROM_12	Count of card promotions in the past 12 months
CLUSTER_CODE	Socio-Economic Cluster Code
CONTROL_NUMBER	ID
DONOR_AGE	Donor Age
DONOR_GENDER	Donor Gender
FREQUENCY_STATUS_97NK	Count of Donations between June 1995 and June 1996 (capped at 4)
HOME_OWNER	Home Owner flag
INCOME_GROUP	Income Bracket, from 1 to 7
IN_HOUSE	Flag for <i>In-House</i> donor program
LAST_GIFT_AMT	Amount of most recent donation
LIFETIME_AVG_GIFT_AMT	Average donation amount, ever
LIFETIME_CARD_PROM	Number of card promotions, ever

Variable	Description
LIFETIME_GIFT_AMOUNT	Total donation amount, ever
LIFETIME_GIFT_COUNT	Total number of donations, ever
LIFETIME_GIFT_RANGE	Maximum gift amount less minimum gift amount
LIFETIME_MAX_GIFT_AMT	Maximum gift amount, ever
LIFETIME_MIN_GIFT_AMT	Minimum gift amount, ever
LIFETIME_PROM	Count of solicitations ever sent
MEDIAN_HOME_VALUE	Census data
MEDIAN_HOUSEHOLD_INCOME	Census data
MONTHS_SINCE_FIRST_GIFT	Months since first donation
MONTHS_SINCE_LAST_GIFT	Months since most recent donation
MONTHS_SINCE_ORIGIN	Months since entry onto the file
MOR_HIT_RATE	Data recorded by a third-party Mail-Order Response rate
NUMBER_PROM_12	Count of promotions in the past 12 months
OVERLAY_SOURCE	Source of Demographic overlay
PCT_MALE_MILITARY	Census data
PCT_MALE_VETERANS	Census data
PCT_OWNER_OCCUPIED	Census data
PCT_VIETNAM_VETERANS	Census data
PCT_WWII_VETERANS	Census data
PEP_STAR	Flag to identify consecutive donors
PER_CAPITA_INCOME	Census data
PUBLISHED_PHONE	Flag
REGENCY_STATUS_96NK	Categorization of donation patterns
RECENT_AVG_CARD_GIFT_AMT	Average donation amount to card promotions since June 1994
RECENT_AVG_GIFT_AMT	Average donation amount to promotions since June 1994
RECENT_CARD_RESPONSE_COUNT	Count of responses to card promotions since June 1994

Variable	Description
RECENT_CARD_RESPONSE_PROP	Proportion of responses to card promotions since June 1994
RECENT_RESPONSE_COUNT	Count of responses to promotions since June 1994
RECENT_RESPONSE_PROP	Proportion of responses to promotions since June 1994
RECENT_STAR_STATUS	STAR status flag, since June 1994
SES	A clustering of the levels of <b>CLUSTER_CODE</b>
TARGET_B	B=Binary, flag for response to 97NK—Target Variable
TARGET_D	Dollar amount of response to 97NK
URBANICITY	Categorization of residency
WEALTH_RATING	Measures wealth relative to others within state