Definitions

The data is available in the unit4.csv document in the folder. Below you will find the description of the features used in the data.

| **Column** | **Description** |
| --- | --- |
| STATE | State abbreviation (a nominal/symbolic field) |
| PVASTATE | EPVA State or PVA State - Indicates whether the donor lives in a state served by the organization's EPVA chapter (P = PVA State, E = EPVA State (Northeastern US)) |
| DOB | Date of birth (YYMM, Year/Month format.) |
| RECP3 | P3 File Flag (\_ = Not a P3 Record, X = Donor has given to PVA's P3 program |
| MDMAUD | The Major Donor Matrix code <sup>\*</sup> |
| GENDER | Gender of the donor |
| DOMAIN | Domain/Cluster code. A nominal or symbolic field. <sup>\*\*</sup> |
| HOMEOWNR | Home Owner Flag (H = Homeowner; U = Unknown) |
| INCOME | Household Income |
| HV1 | Median Home Value in hundreds |
| HV2 | Average Home Value in hundreds |
| HV3 | Median Contract Rent in hundreds |
| HV4 | Average Contract Rent in hundreds 'IC1' |
| IC1 | Median Household Income in hundreds |
| IC2 | Median Family Income in hundreds |
| IC3 | Average Household Income in hundreds |
| IC4 | Average Family Income in hundreds |
| IC5 | Per Capita Income |
| VETERANS | Veterans (Y/N) |
| RFA\_2 | Donor's RFA status as of 97NK promotion date |
| CARDPROM | Lifetime number of card promotions received to date. Card promotions are promotion type FS, GK, TK, SK, NK, XK, UF, UU. |
| MAXADATE | Date of the most recent promotion received (in YYMM, Year/Month format) |
| NUMPROM | Lifetime number of promotions received to date |
| CARDPM12 | Number of card promotions received in the last 12 months (in terms of calendar months translates into 9603-9702) |
| NUMPRM12 | Number of promotions received in the last 12 months (in terms of calendar months translates into 9603-9702) |
| NGIFTALL | Number of lifetime gifts to date |
| TIMELAG | Number of months between first and second gift neighborhood demographics |
| AGE901 | Median Age of Population |
| AGE902 | Median Age of Adults 18 or Older |
| AGE903 | Median Age of Adults 25 or Older |
| AVGGIFT | (often the thing we want to predict) the size of the average gift given in dollars – tip, HINT : format to 0 dp |

Complex strings and definitions

**RFA –**

the codes describe frequency and amount of giving for donors who have given a $100+ gift at any time in their giving history. An RFA (recency/frequency/monetary) field. The (current) concatenated version is a nominal or symbolic field. The individual bytes could separately be used as fields and refer to the following:

* First byte: Recency of Giving C=Current Donor L=Lapsed Donor I=Inactive Donor D=Dormant Donor
* 2nd byte: Frequency of Giving 1=One gift in the period of recency 2=Two-Four gifts in the period of recency 5=Five+ gifts in the period of recency
* 3rd byte: Amount of Giving L=Less than $100(Low Dollar) C=$100-499(Core) M=$500-999(Major) T=$1,000+(Top)
* 4th byte: Blank/meaningless/filler 'X' indicates that the donor is not a major donor.

**Domain -**

* 1st byte = Urbanicity level of the donor's neighbourhood U=Urban C=City S=Suburban T=Town R=Rural
* 2nd byte = Socio-Economic status of the neighbourhood 1 = Highest SES 2 = Average SES 3 = Lowest SES (except for Urban communities, where 1 = Highest SES, 2 = Above average SES, 3 = Below average SES, 4 = Lowest SES.)