loadData.R

andrewma

2022-04-09

```r
library(haven)
library(skimr)


house <- read_dta('114542-V1/PisoFirme_AEJPol-20070024_household.dta')
indiv <- read_dta('114542-V1/PisoFirme_AEJPol-20070024_individual.dta')



test <- house[rowSums(is.na(house)) > 0,]
test2 <- indiv[rowSums(is.na(indiv)) > 0,]



tt <- skim(indiv)
```

```
## Warning: Couldn't find skimmers for class: haven_labelled, vctrs_vctr, double,
## numeric; No user-defined `sfl` provided. Falling back to `character`.
```

```r
print(tt)
```

```
## ── Data Summary ─────────────────────────
##                            Values
## Name                       indiv
## Number of rows             6693
## Number of columns          89
## _____
## Column type frequency:
##    character               1
##    numeric                 88
## _____
## Group variables            None
##
## ── Variable type: character ─────────────────────────────────────────────
##    skim_variable n_missing complete_rate    min    max empty n_unique whitespace
## 1 S_gender               0             1      1      1     0        2          0
##
## ── Variable type: numeric ───────────────────────────────────────────────
##     skim_variable         n_missing complete_rate      mean       sd         p0
##  1 dpisofirme                     0             1   4.90e-1 5.00e-1          0
```

```
##  2 idcluster              67   0.990    2.09e+8 1.37e+8 70000537
##  3 coord_x               472   0.929   -1.03e+2 5.52e-2    -104.
##  4 coord_y               472   0.929    2.56e+1 3.71e-2     25.4
##  5 idmun                   0   1        2.10e+1 1.37e+1        7
##  6 idmza                   0   1        9.84e+1 2.59e+2        1
##  7 S_age                   0   1        1.34e+1 1.43e+1        0
##  8 S_childma            2601   0.611    9.65e-1 1.83e-1        0
##  9 S_childmaage         2441   0.635    2.94e+1 9.43e+0       14
## 10 S_childmaeduc        2444   0.635    6.69e+0 2.84e+0        0
## 11 S_childpa            2601   0.611    7.80e-1 4.14e-1        0
## 12 S_childpaage         3403   0.492    3.23e+1 1.02e+1       13
## 13 S_childpaeduc        3414   0.490    6.74e+0 3.20e+0        0
## 14 S_HHpeople              0   1        5.63e+0 2.17e+0        0
## 15 S_rooms                 0   1        2.04e+0 1.09e+0        1
## 16 S_parcount           3565   0.467    3.02e-1 6.25e-1        0
## 17 S_diarrhea           2618   0.609    1.33e-1 3.40e-1        0
## 18 S_anemia             2935   0.561    3.86e-1 4.87e-1        0
## 19 S_mccdts             6092   0.0898   1.54e+1 2.01e+1        0
## 20 S_pbdypct            5104   0.237    3.19e+1 2.54e+1        3
## 21 S_haz                2735   0.591   -6.02e-1 1.11e+0    -4.78
## 22 S_whz                2715   0.594    1.31e-1 1.13e+0       -4
## 23 S_respira            2616   0.609    3.65e-1 4.82e-1        0
## 24 S_skin               2621   0.608    1.01e-1 3.02e-1        0
## 25 S_otherdis           2617   0.609    4.34e-2 2.04e-1        0
## 26 S_malincom           6140   0.0826   7.77e+0 6.06e-1     4.83
## 27 S_palincom           4645   0.306    8.11e+0 5.97e-1     3.22
## 28 dtriage_0_02_male       0   1        7.92e-3 8.86e-2        0
## 29 dtriage_0_35_male       0   1        1.37e-2 1.16e-1        0
## 30 dtriage_0_68_male       0   1        1.21e-2 1.09e-1        0
## 31 dtriage_0_911_male      0   1        1.30e-2 1.13e-1        0
## 32 dtriage_0_02_female     0   1        7.02e-3 8.35e-2        0
## 33 dtriage_0_35_female     0   1        1.26e-2 1.11e-1        0
## 34 dtriage_0_68_female     0   1        1.36e-2 1.16e-1        0
## 35 dtriage_0_911_female    0   1        1.23e-2 1.10e-1        0
## 36 dtriage_1_02_male       0   1        6.57e-3 8.08e-2        0
## 37 dtriage_1_35_male       0   1        1.36e-2 1.16e-1        0
## 38 dtriage_1_68_male       0   1        1.24e-2 1.11e-1        0
## 39 dtriage_1_911_male      0   1        1.46e-2 1.20e-1        0
## 40 dtriage_1_02_female     0   1        8.82e-3 9.35e-2        0
## 41 dtriage_1_35_female     0   1        1.42e-2 1.18e-1        0
## 42 dtriage_1_68_female     0   1        1.23e-2 1.10e-1        0
## 43 dtriage_1_911_female    0   1        1.39e-2 1.17e-1        0
## 44 dtriage_2_02_male       0   1        8.37e-3 9.11e-2        0
## 45 dtriage_2_35_male       0   1        1.36e-2 1.16e-1        0
## 46 dtriage_2_68_male       0   1        1.51e-2 1.22e-1        0
## 47 dtriage_2_911_male      0   1        1.43e-2 1.19e-1        0
## 48 dtriage_2_02_female     0   1        7.47e-3 8.61e-2        0
```

```
## 49 dtriage_2_35_female              0          1       1.73e-2 1.31e-1          0
## 50 dtriage_2_68_female              0          1       1.42e-2 1.18e-1          0
## 51 dtriage_2_911_female             0          1       1.34e-2 1.15e-1          0
## 52 dtriage_3_02_male                0          1       9.11e-3 9.50e-2          0
## 53 dtriage_3_35_male                0          1       1.48e-2 1.21e-1          0
## 54 dtriage_3_68_male                0          1       1.23e-2 1.10e-1          0
## 55 dtriage_3_911_male               0          1       1.24e-2 1.11e-1          0
## 56 dtriage_3_02_female              0          1       8.52e-3 9.19e-2          0
## 57 dtriage_3_35_female              0          1       1.43e-2 1.19e-1          0
## 58 dtriage_3_68_female              0          1       1.49e-2 1.21e-1          0
## 59 dtriage_3_911_female             0          1       1.14e-2 1.06e-1          0
## 60 dtriage_4_02_male                0          1       9.11e-3 9.50e-2          0
## 61 dtriage_4_35_male                0          1       1.61e-2 1.26e-1          0
## 62 dtriage_4_68_male                0          1       1.64e-2 1.27e-1          0
## 63 dtriage_4_911_male               0          1       1.34e-2 1.15e-1          0
## 64 dtriage_4_02_female              0          1       7.92e-3 8.86e-2          0
## 65 dtriage_4_35_female              0          1       1.33e-2 1.15e-1          0
## 66 dtriage_4_68_female              0          1       1.58e-2 1.25e-1          0
## 67 dtriage_4_911_female             0          1       1.39e-2 1.17e-1          0
## 68 dtriage_5_02_male                0          1       1.05e-2 1.02e-1          0
## 69 dtriage_5_35_male                0          1       1.55e-2 1.24e-1          0
## 70 dtriage_5_68_male                0          1       1.67e-2 1.28e-1          0
## 71 dtriage_5_911_male               0          1       1.66e-2 1.28e-1          0
## 72 dtriage_5_02_female              0          1       9.26e-3 9.58e-2          0
## 73 dtriage_5_35_female              0          1       1.64e-2 1.27e-1          0
## 74 dtriage_5_68_female              0          1       1.51e-2 1.22e-1          0
## 75 dtriage_5_911_female             0          1       1.51e-2 1.22e-1          0
## 76 S_waterland                      0          1       9.73e-1 1.63e-1          0
## 77 S_waterhouse                     0          1       5.23e-1 4.99e-1          0
## 78 S_electricity                    0          1       9.90e-1 1.02e-1          0
## 79 S_milkprogram                    0          1       6.89e-2 2.53e-1          0
## 80 S_foodprogram                    0          1       2.88e-2 1.67e-1          0
## 81 S_seguropopular                  0          1       1.42e-2 1.18e-1          0
## 82 S_hasanimals                     0          1       4.93e-1 5.00e-1          0
## 83 S_animalsinside                  0          1       1.93e-1 3.94e-1          0
## 84 S_garbage                        0          1       8.19e-1 3.85e-1          0
## 85 S_washhands                      0          1       3.74e+0 1.49e+0          0
## 86 S_incomepc                       7      0.999       1.06e+3 3.47e+3          0
## 87 S_cashtransfers                  5      0.999       1.31e+1 3.42e+1          0
## 88 S_assetspc                       3       1.00       2.14e+4 7.28e+3      5574.
##             p25        p50         p75         p100 hist
## 1            0          0            1            1  ▄▄ ▄▄▄
## 2     70001446      1.20e+8 350002560     353220000  ▄▄ ▄▄▄
## 3        -103.      -1.03e+2      -103.         -103. ▄▄▄▄▄
## 4         25.5      2.56e+1       25.6          25.6  ▄▄▄▄▄
## 5            7      1.2 e+1       35            35  ▄▄ ▄▄▄
## 6           12      2.4 e+1       38           999  ▄▄▄
```

```
##  7       2        5    e+0        27        77
##  8       1        1    e+0         1         1
##  9      23      2.8    e+1        33        79
## 10       6        6    e+0         9        16
## 11       1        1    e+0         1         1
## 12      26        3    e+1        36        87
## 13       6        6    e+0         9        16
## 14       4        5    e+0         6        19
## 15       1        2    e+0         3        10
## 16       0        0               0         5
## 17       0        0               0         1
## 18       0        0               1         1
## 19       2        7    e+0        20        96
## 20      12      2.5    e+1        47        97
## 21   -1.33   -6.00e-1          0.100      2.99
## 22  -0.560    9.00e-2          0.720      4.94
## 23       0        0               1         1
## 24       0        0               0         1
## 25       0        0               0         1
## 26    7.60   7.78e+0            8.02      14.8
## 27    7.85   8.07e+0            8.29      14.5
## 28       0        0               0         1
## 29       0        0               0         1
## 30       0        0               0         1
## 31       0        0               0         1
## 32       0        0               0         1
## 33       0        0               0         1
## 34       0        0               0         1
## 35       0        0               0         1
## 36       0        0               0         1
## 37       0        0               0         1
## 38       0        0               0         1
## 39       0        0               0         1
## 40       0        0               0         1
## 41       0        0               0         1
## 42       0        0               0         1
## 43       0        0               0         1
## 44       0        0               0         1
## 45       0        0               0         1
## 46       0        0               0         1
## 47       0        0               0         1
## 48       0        0               0         1
## 49       0        0               0         1
## 50       0        0               0         1
## 51       0        0               0         1
## 52       0        0               0         1
## 53       0        0               0         1
```

```
## 54          0      0                  0            1
## 55          0      0                  0            1
## 56          0      0                  0            1
## 57          0      0                  0            1
## 58          0      0                  0            1
## 59          0      0                  0            1
## 60          0      0                  0            1
## 61          0      0                  0            1
## 62          0      0                  0            1
## 63          0      0                  0            1
## 64          0      0                  0            1
## 65          0      0                  0            1
## 66          0      0                  0            1
## 67          0      0                  0            1
## 68          0      0                  0            1
## 69          0      0                  0            1
## 70          0      0                  0            1
## 71          0      0                  0            1
## 72          0      0                  0            1
## 73          0      0                  0            1
## 74          0      0                  0            1
## 75          0      0                  0            1
## 76          1      1    e+0            1            1
## 77          0      1    e+0            1            1
## 78          1      1    e+0            1            1
## 79          0      0                  0            1
## 80          0      0                  0            1
## 81          0      0                  0            1
## 82          0      0                  1            1
## 83          0      0                  0            1
## 84          1      1    e+0            1            1
## 85          3      3    e+0            4           10
## 86       533.   7.55e+2           1086.       127267.
## 87          0      0                  0          950
## 88     16844.   2.11e+4          26313.        64594.
```

```
tt <- skim(house)
print(tt)
```

```
## ── Data Summary ────────────────────────────
##                                Values
## Name                           house
## Number of rows                 2783
## Number of columns              78
## _____
## Column type frequency:
```

```
##    numeric                     78
## ──────────────────────────
## Group variables             None
##
## ── Variable type: numeric ──────────────────────────────────────────
##    skim_variable       n_missing complete_rate      mean       sd        p0        p25
##  1 dpisofirme                  0           1      4.99e-1 5.00e-1   0          0
##  2 idcluster                  28           0.990  2.12e+8 1.37e+8   7.00e+7    7.00e+7
##  3 coord_x                   203           0.927 -1.03e+2 5.59e-2  -1.04e+2   -1.03e+2
##  4 coord_y                   203           0.927  2.56e+1 3.74e-2   2.54e+1    2.55e+1
##  5 idmun                       0           1      2.13e+1 1.37e+1   7    e+0   7    e+0
##  6 idmza                       0           1      1.01e+2 2.62e+2   1    e+0   1.2 e+1
##  7 C_blocksdirtfloor         203           0.927  6.18e-1 2.25e-1   0          4.65e-1
##  8 C_HHdirtfloor             203           0.927  1.98e-1 1.57e-1   0          8.33e-2
##  9 C_child05                 203           0.927  7.26e-1 9.11e-2   3.61e-1    6.61e-1
## 10 C_households              203           0.927  5.10e+2 3.09e+2   3    e+0   3.04e+2
## 11 C_people                  203           0.927  2.27e+3 1.40e+3   1    e+1   1.25e+3
## 12 C_rooms                   203           0.927  2.30e+0 4.83e-1   1    e+0   1.94e+0
## 13 C_HHpersons               203           0.927  4.44e+0 5.00e-1   3.33e+0    4.20e+0
## 14 C_waterland               203           0.927  6.55e-2 1.03e-1   0          8.75e-3
## 15 C_waterhouse              203           0.927  3.66e-1 2.29e-1   1.57e-2    1.76e-1
## 16 C_waterbath               203           0.927  4.90e-1 2.32e-1   3.18e-3    3.08e-1
## 17 C_gasheater               203           0.927  3.15e-2 2.81e-2   0          1.50e-2
## 18 C_refrigerator            203           0.927  2.17e-1 1.01e-1   0          1.48e-1
## 19 C_washing                 203           0.927  3.83e-1 1.14e-1   9.88e-2    2.97e-1
## 20 C_telephone               203           0.927  8.05e-1 1.16e-1   2.43e-1    7.33e-1
## 21 C_vehicle                 203           0.927  7.45e-1 1.05e-1   2.80e-1    6.79e-1
## 22 C_overcrowding            203           0.927  2.36e+0 4.31e-1   1.17e+0    2.02e+0
## 23 C_poverty                 203           0.927  1.56e-1 8.83e-2   0          9.31e-2
## 24 C_illiterate              203           0.927  6.18e-2 2.47e-2   0          5.24e-2
## 25 C_headeduc                203           0.927  6.42e+0 1.11e+0   3.12e+0    5.86e+0
## 26 C_dropouts515             203           0.927  1.82e-1 6.21e-2   0          1.42e-1
## 27 C_employment              203           0.927  1.57e+0 3.26e-1   1    e+0   1.42e+0
## 28 C_earnincome              203           0.927  1.48e+0 3.06e-1   1    e+0   1.34e+0
## 29 S_HHpeople                  0           1      5.34e+0 1.88e+0   0          4    e+0
## 30 S_headage                   0           1      3.73e+1 1.27e+1   1.6 e+1    2.8 e+1
## 31 S_spouseage                 0           1      2.92e+1 1.46e+1   0          2.4 e+1
## 32 S_headeduc                  4           0.999  6.26e+0 3.40e+0   0          4    e+0
## 33 S_spouseeduc              339           0.878  6.39e+0 2.96e+0   0          5    e+0
## 34 S_rooms                     0           1      2.03e+0 1.08e+0   1    e+0   1    e+0
## 35 S_waterland                 0           1      9.72e-1 1.64e-1   0          1    e+0
## 36 S_waterhouse                0           1      5.29e-1 4.99e-1   0          0
## 37 S_electricity               0           1      9.89e-1 1.03e-1   0          1    e+0
## 38 S_cementfloor2000           0           1      3.28e-1 3.74e-1   0          0
## 39 S_hasanimals                0           1      4.98e-1 5.00e-1   0          0
## 40 S_animalsinside             0           1      1.91e-1 3.93e-1   0          0
## 41 S_garbage                   0           1      8.16e-1 3.87e-1   0          1    e+0
```

```
## 42 S_washhands               0           1      3.73e+0 1.48e+0  0        3     e+0
## 43 S_incomepc                3           0.999  1.04e+3 3.54e+3  0        5.20e+2
## 44 S_assetspc                1           1.00   2.22e+4 7.40e+3  5.57e+3  1.75e+4
## 45 S_shpeoplework            1           1.00   2.63e-1 1.28e-1  0        2.00e-1
## 46 S_microenter             11           0.996  1.01e-1 3.01e-1  0        0
## 47 S_hrsworkedpc             1           1.00   1.40e+1 8.39e+0  0        9.53e+0
## 48 S_consumptionpc           4           0.999  7.56e+2 1.15e+3  0        4.63e+2
## 49 S_cashtransfers           2           0.999  1.45e+1 3.92e+1  0        0
## 50 S_milkprogram             0           1      7.04e-2 2.56e-1  0        0
## 51 S_foodprogram             0           1      2.87e-2 1.67e-1  0        0
## 52 S_dem1                    1           1.00   1.46e-1 1.42e-1  0        0
## 53 S_dem2                    1           1.00   1.18e-1 1.37e-1  0        0
## 54 S_dem3                    1           1.00   1.99e-1 1.03e-1  0        1.67e-1
## 55 S_dem4                    1           1.00   2.49e-2 6.36e-2  0        0
## 56 S_dem5                    1           1.00   1.41e-1 1.36e-1  0        0
## 57 S_dem6                    1           1.00   1.15e-1 1.35e-1  0        0
## 58 S_dem7                    1           1.00   2.31e-1 8.44e-2  0        2.00e-1
## 59 S_dem8                    1           1.00   2.50e-2 6.56e-2  0        0
## 60 S_seguropopular           0           1      1.47e-2 1.21e-1  0        0
## 61 S_shcementfloor           0           1      8.28e-1 2.99e-1  0        7.5 e-1
## 62 S_cementfloorkit          0           1      7.98e-1 4.01e-1  0        1     e+0
## 63 S_cementfloordin          0           1      8.14e-1 3.90e-1  0        1     e+0
## 64 S_cementfloorbat          0           1      8.53e-1 3.54e-1  0        1     e+0
## 65 S_cementfloorbed          0           1      7.87e-1 4.10e-1  0        1     e+0
## 66 S_satisfloor              0           1      6.22e-1 4.85e-1  0        0
## 67 S_satishouse              0           1      6.53e-1 4.76e-1  0        0
## 68 S_satislife               0           1      6.56e-1 4.75e-1  0        0
## 69 S_cesds                  13           0.995  1.74e+1 8.96e+0  0        1.1 e+1
## 70 S_pss                     9           0.997  1.57e+1 6.73e+0  0        1.1 e+1
## 71 S_instcement              1           1.00   7.18e-1 4.50e-1  0        0
## 72 S_instsanita              3           0.999  9.28e-2 2.90e-1  0        0
## 73 S_restsanita              2           0.999  4.39e-2 2.05e-1  0        0
## 74 S_constceili              2           0.999  1.74e-1 3.79e-1  0        0
## 75 S_restowalls              1           1.00   1.17e-1 3.21e-1  0        0
## 76 S_improveany              0           1      3.00e-1 4.58e-1  0        0
## 77 S_logrent               187           0.933  5.94e+0 7.65e-1  3.00e+0  5.70e+0
## 78 S_logsell               294           0.894  1.05e+1 1.17e+0  3.00e+0  9.90e+0
##         p50       p75          p100 hist
##  1  0          1    e+0           1      ▁█▁▁▁█
##  2  1.21e+8   3.50e+8  353220000        ▁█▁▁▁█
##  3 -1.03e+2  -1.03e+2      -103.        ▁█▇▃▁▁
##  4  2.56e+1   2.56e+1        25.6        ▁▁▃█▁▁
##  5  1.2 e+1   3.5 e+1          35        ▁█▁▁▁█
##  6  2.4 e+1   3.8 e+1         999        ▁█▁▁▁▁
##  7  6.45e-1   8.16e-1           1        ▁▃▅█▇▇
##  8  1.62e-1   2.75e-1       0.757        ▁█▃▂▁▁
##  9  7.22e-1   7.79e-1        1.20        ▁▁█▃▁▁
```

```
## 10   4.75e+2   6.78e+2       1598
## 11   2.20e+3   3.11e+3       7285
## 12   2.31e+0   2.63e+0       4.36
## 13   4.36e+0   4.53e+0       8.10
## 14   3.01e-2   8.30e-2      0.667
## 15   3.39e-1   5.55e-1          1
## 16   4.66e-1   6.08e-1          1
## 17   2.48e-2   4.20e-2      0.188
## 18   2.05e-1   2.73e-1      0.570
## 19   3.53e-1   4.68e-1      0.769
## 20   8.33e-1   8.95e-1          1
## 21   7.54e-1   8.35e-1          1
## 22   2.34e+0   2.67e+0       3.33
## 23   1.45e-1   1.98e-1      0.438
## 24   6.04e-2   7.10e-2      0.211
## 25   6.08e+0   6.48e+0       11.4
## 26   1.76e-1   2.14e-1      0.577
## 27   1.53e+0   1.63e+0       4.88
## 28   1.45e+0   1.53e+0       4.58
## 29   5    e+0   6    e+0         19
## 30   3.4 e+1   4.4 e+1         98
## 31   3    e+1   3.6 e+1         81
## 32   6    e+0   9    e+0         16
## 33   6    e+0   9    e+0         16
## 34   2    e+0   3    e+0         10
## 35   1    e+0   1    e+0          1
## 36   1    e+0   1    e+0          1
## 37   1    e+0   1    e+0          1
## 38   2.00e-1   7.14e-1          1
## 39   0          1    e+0          1
## 40   0          0                1
## 41   1    e+0   1    e+0          1
## 42   3    e+0   4    e+0         10
## 43   7.5 e+2   1.08e+3    127267.
## 44   2.12e+4   2.65e+4     64594.
## 45   2.5 e-1   3.33e-1      0.800
## 46   0          0                1
## 47   1.2 e+1   1.75e+1         72
## 48   6.26e+2   8.43e+2     35460.
## 49   0          0              950
## 50   0          0                1
## 51   0          0                1
## 52   1.67e-1   2.5 e-1      0.667
## 53   0          2.00e-1      0.571
## 54   2.00e-1   2.5 e-1      0.600
## 55   0          0              0.5
## 56   1.67e-1   2.5 e-1      0.667
```

```
## 57   0          2.00e-1        0.571  ▆▁_
## 58   2.00e-1   2.5 e-1        0.667  ▆█_
## 59   0          0             0.5    █▁_
## 60   0          0             1      █▁_
## 61   1   e+0   1    e+0       1       _█
## 62   1   e+0   1    e+0       1       _█
## 63   1   e+0   1    e+0       1       _█
## 64   1   e+0   1    e+0       1       _█
## 65   1   e+0   1    e+0       1       _█
## 66   1   e+0   1    e+0       1      █_█
## 67   1   e+0   1    e+0       1      █_█
## 68   1   e+0   1    e+0       1      █_█
## 69   1.7 e+1   2.3 e+1        56     ▄█▁
## 70   1.6 e+1   2    e+1       39     ▄█_
## 71   1   e+0   1    e+0       1      ▁_█
## 72   0          0             1      █▁_
## 73   0          0             1      █▁_
## 74   0          0             1      █▁ _
## 75   0          0             1      █▁_
## 76   0          1    e+0       1      █▁ _
## 77   5.99e+0   6.21e+0        11.4   ▄█_
## 78   1.06e+1   1.12e+1        16.0    _██
```