fin\_df

Autogenerated data summary from dataMaid

2024-07-30 23:28:12.938518

# Data report overview

The dataset examined has the following dimensions:

| Feature | Result |
| --- | --- |
| Number of observations | 1998 |
| Number of variables | 66 |

### Checks performed

The following variable checks were performed, depending on the data type of each variable:

|  | character | factor | labelled | haven labelled | numeric | integer | logical | Date |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Identify miscoded missing values | × | × | × | × | × | × |  | × |
| Identify prefixed and suffixed whitespace | × | × | × | × |  |  |  |  |
| Identify levels with < 6 obs. | × | × | × | × |  |  |  |  |
| Identify case issues | × | × | × | × |  |  |  |  |
| Identify misclassified numeric or integer variables | × | × | × | × |  |  |  |  |
| Identify outliers |  |  |  |  | × | × |  | × |

Please note that all numerical values in the following have been rounded to 2 decimals.

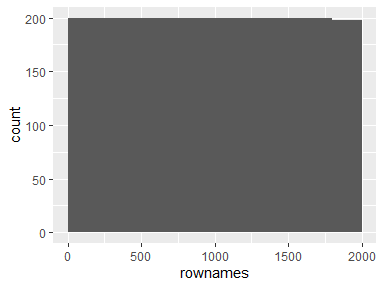
# Summary table

|  | Variable class | # unique values | Missing observations | Any problems? |
| --- | --- | --- | --- | --- |
| [rownames](#rownames) | integer | 1998 | 0.00 % |  |
| [num\_incomes](#num_incomes) | character | 5 | 0.00 % | × |
| [pan\_inc\_chg](#pan_inc_chg) | character | 4 | 0.00 % | × |
| [pan\_inc\_chg\_pct](#pan_inc_chg_pct) | character | 12 | 0.00 % | × |
| [pan\_exp\_chg](#pan_exp_chg) | character | 4 | 0.00 % | × |
| [pan\_exp\_chg\_pct](#pan_exp_chg_pct) | character | 12 | 0.00 % | × |
| [pan\_fi\_chg](#pan_fi_chg) | character | 4 | 0.00 % | × |
| [pan\_ret\_date\_chg](#pan_ret_date_chg) | character | 4 | 0.00 % | × |
| [pan\_financial\_impact](#pan_financial_impact) | character | 4 | 0.00 % | × |
| [political](#political) | character | 25 | 0.00 % | × |
| [race\_eth](#race_eth) | character | 8 | 0.00 % | × |
| [gender](#gender) | character | 7 | 0.00 % | × |
| [age](#age) | character | 12 | 0.00 % | × |
| [edu](#edu) | character | 12 | 0.00 % | × |
| [rel\_status](#rel_status) | character | 9 | 0.00 % | × |
| [children](#children) | character | 5 | 0.00 % | × |
| [country](#country) | character | 54 | 0.00 % | × |
| [fin\_indy](#fin_indy) | character | 2 | 0.00 % |  |
| [fin\_indy\_num](#fin_indy_num) | numeric | 156 | 11.01 % | × |
| [fin\_indy\_pct](#fin_indy_pct) | numeric | 244 | 11.31 % | × |
| [retire\_invst\_num](#retire_invst_num) | numeric | 153 | 14.31 % | × |
| [tgt\_sf\_wthdrw\_rt](#tgt_sf_wthdrw_rt) | numeric | 53 | 10.86 % | × |
| [max\_retire\_sup](#max_retire_sup) | numeric | 174 | 26.53 % | × |
| [retire\_exp](#retire_exp) | numeric | 142 | 13.11 % | × |
| [whn\_fin\_indy\_num](#whn_fin_indy_num) | integer | 46 | 92.94 % | × |
| [fin\_indy\_lvl](#fin_indy_lvl) | character | 6 | 0.00 % | × |
| [retire\_age](#retire_age) | character | 12 | 0.00 % | × |
| [stp\_whn\_fin\_indy](#stp_whn_fin_indy) | character | 5 | 0.00 % | × |
| [industry](#industry) | character | 20 | 0.00 % | × |
| [employer](#employer) | character | 6 | 0.00 % | × |
| [role](#role) | character | 13 | 0.00 % | × |
| [ft\_status](#ft_status) | character | 5 | 0.00 % | × |
| [pt\_status](#pt_status) | character | 5 | 0.00 % | × |
| [gig\_status](#gig_status) | character | 6 | 0.00 % | × |
| [ne\_status](#ne_status) | character | 7 | 0.00 % | × |
| [edu\_status](#edu_status) | character | 4 | 0.00 % | × |
| [housing](#housing) | character | 20 | 0.00 % | × |
| [home\_value](#home_value) | integer | 261 | 18.77 % | × |
| [brokerage\_accts\_tax](#brokerage_accts_tax) | numeric | 628 | 6.26 % | × |
| [retirement\_accts\_tax](#retirement_accts_tax) | numeric | 706 | 2.90 % | × |
| [cash](#cash) | numeric | 427 | 3.30 % | × |
| [invst\_accts](#invst_accts) | numeric | 279 | 23.47 % | × |
| [spec\_crypto](#spec_crypto) | numeric | 181 | 28.28 % | × |
| [invst\_prop\_bus\_own](#invst_prop_bus_own) | integer | 134 | 36.04 % | × |
| [other\_val](#other_val) | integer | 102 | 50.00 % | × |
| [student\_loans](#student_loans) | numeric | 157 | 82.78 % | × |
| [mortgage](#mortgage) | numeric | 425 | 57.76 % | × |
| [auto\_loan](#auto_loan) | numeric | 137 | 80.43 % | × |
| [credit\_personal\_loan](#credit_personal_loan) | numeric | 123 | 80.48 % | × |
| [medical\_debt](#medical_debt) | numeric | 17 | 95.30 % | × |
| [invst\_prop\_bus\_own\_debt](#invst_prop_bus_own_debt) | numeric | 103 | 90.04 % | × |
| [other\_debt](#other_debt) | numeric | 38 | 94.99 % |  |
| [X2020\_gross\_inc](#x2020_gross_inc) | numeric | 613 | 2.80 % | × |
| [X2020\_housing\_exp](#x2020_housing_exp) | numeric | 577 | 7.96 % | × |
| [X2020\_utilities\_exp](#x2020_utilities_exp) | numeric | 439 | 10.66 % | × |
| [X2020\_transp\_exp](#x2020_transp_exp) | numeric | 428 | 13.26 % | × |
| [X2020\_necessities\_exp](#x2020_necessities_exp) | numeric | 364 | 12.26 % | × |
| [X2020\_lux\_exp](#x2020_lux_exp) | numeric | 359 | 14.61 % | × |
| [X2020\_child\_exp](#x2020_child_exp) | numeric | 119 | 81.68 % | × |
| [X2020\_debt\_repay](#x2020_debt_repay) | numeric | 126 | 84.38 % | × |
| [X2020\_invst\_save](#x2020_invst_save) | numeric | 436 | 26.23 % | × |
| [X2020\_charity](#x2020_charity) | numeric | 188 | 59.36 % | × |
| [X2020\_healthcare\_exp](#x2020_healthcare_exp) | numeric | 292 | 44.89 % | × |
| [X2020\_taxes](#x2020_taxes) | numeric | 537 | 26.88 % | × |
| [X2020\_edu\_exp](#x2020_edu_exp) | numeric | 107 | 85.89 % | × |
| [X2020\_other\_exp](#x2020_other_exp) | numeric | 96 | 90.94 % | × |

# Variable list

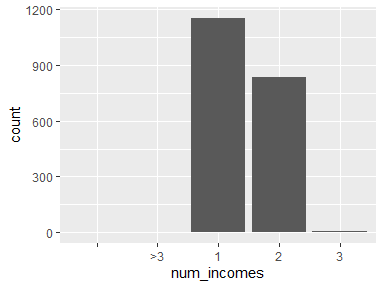
## rownames

| Feature | Result |
| --- | --- |
| Variable type | integer |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 1998 |
| Median | 999.5 |
| 1st and 3rd quartiles | 500.25; 1498.75 |
| Min. and max. | 1; 1998 |



## num\_incomes

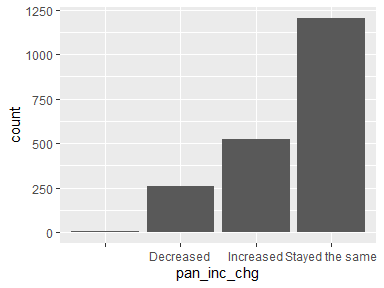
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 5 |
| Mode | “1” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "", ">3", "3".

## pan\_inc\_chg

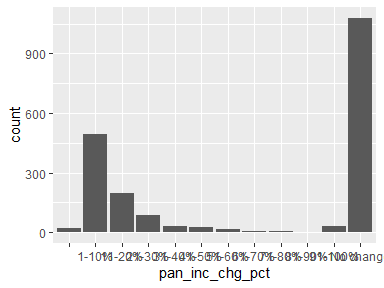
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 4 |
| Mode | “Stayed the same” |



* The following suspected missing value codes enter as regular values: "".

## pan\_inc\_chg\_pct

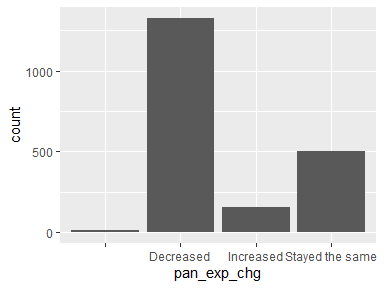
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 12 |
| Mode | “No change” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "71-80%", "81-90%".

## pan\_exp\_chg

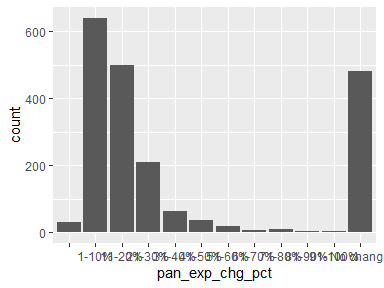
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 4 |
| Mode | “Decreased” |



* The following suspected missing value codes enter as regular values: "".

## pan\_exp\_chg\_pct

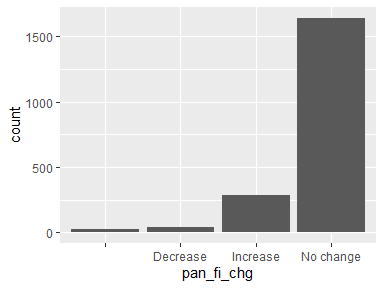
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 12 |
| Mode | “1-10%” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "81-90%", "91-100%".

## pan\_fi\_chg

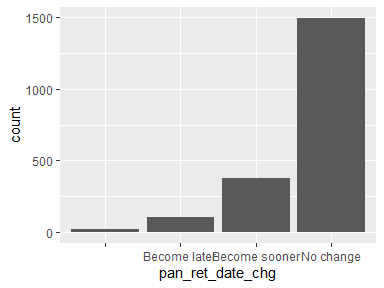
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 4 |
| Mode | “No change” |



* The following suspected missing value codes enter as regular values: "".

## pan\_ret\_date\_chg

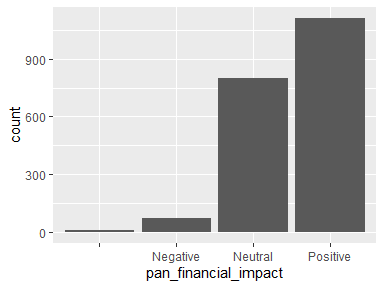
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 4 |
| Mode | “No change” |



* The following suspected missing value codes enter as regular values: "".

## pan\_financial\_impact

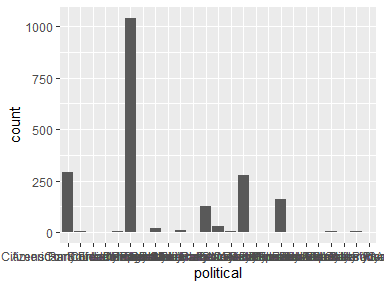
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 4 |
| Mode | “Positive” |



* The following suspected missing value codes enter as regular values: "".

## political

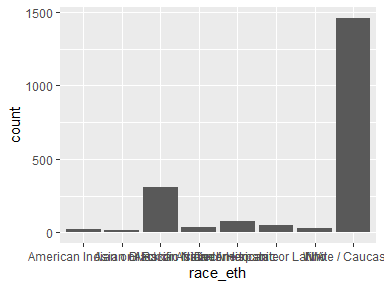
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 25 |
| Mode | “Democrat” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "American Solidarity Party", "Citizens Party of the United States", "Communist Party USA", "Constitution Party", "Freedom Socialist Party", …, "Socialist Equality Party", "Socialist Party USA", "United States Marijuana Party", "United States Pirate Party", "Workers World Party" (7 values omitted).

## race\_eth

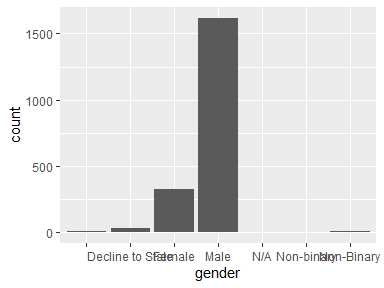
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 8 |
| Mode | “White / Caucasian” |



* The following suspected missing value codes enter as regular values: "".

## gender

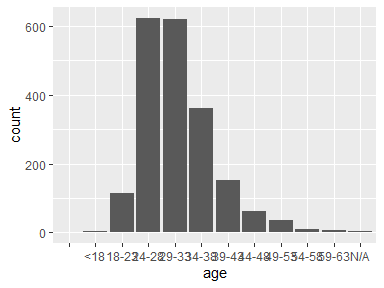
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 7 |
| Mode | “Male” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "N/A", "Non-binary".
* Note that there might be case problems with the following levels: "Non-binary", "Non-Binary".

## age

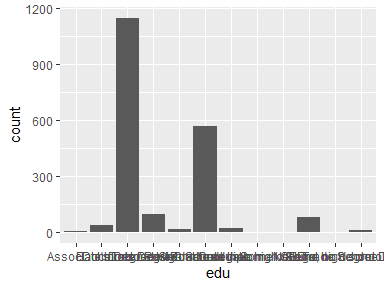
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 12 |
| Mode | “24-28” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "", "<18", "N/A".

## edu

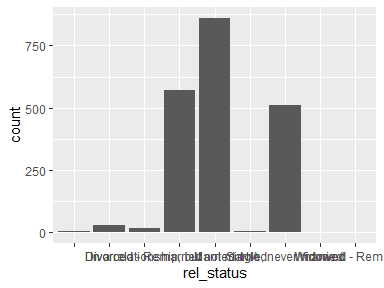
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 12 |
| Mode | “Bachelor’s Degree” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "", "Less than high school", "N/A", "Some high school".

## rel\_status

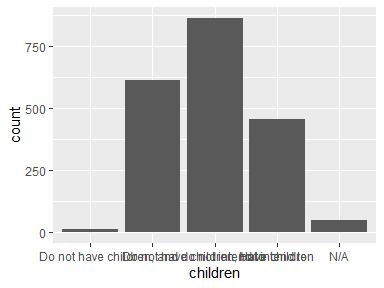
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 9 |
| Mode | “Married” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "", "N/A", "Widowed", "Widowed - Remarried".

## children

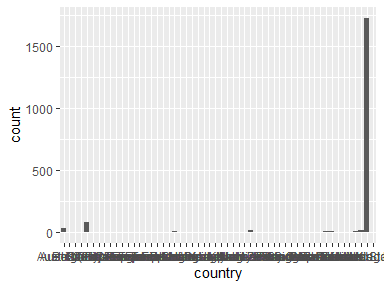
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 5 |
| Mode | “Do not have children, but intend to” |



* The following suspected missing value codes enter as regular values: "".

## country

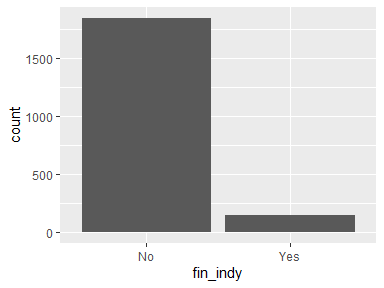
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 54 |
| Mode | “United States” |



* Note that the following levels have at most five observations: "Austria (Eu)", "Belgium", "Brazil", "China", "Colombia", …, "South Korea", "Taiwan", "Thailand", "Uae", "Us Armed Forces" (35 values omitted).

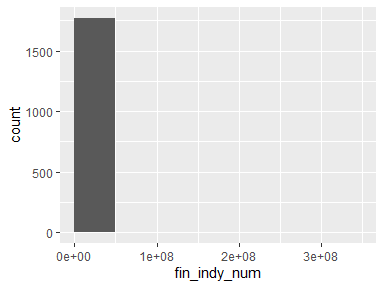
## fin\_indy

| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 2 |
| Mode | “No” |



## fin\_indy\_num

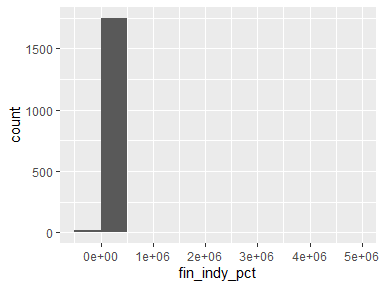
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 220 (11.01 %) |
| Number of unique values | 155 |
| Median | 1600000 |
| 1st and 3rd quartiles | 1e+06; 2500000 |
| Min. and max. | 1; 3.5e+08 |



* Note that the following possible outlier values were detected: "1", "50", "800", "10000", "44000", …, "5.5e+07", "6e+07", "1e+08", "1.8e+08", "3.5e+08" (25 values omitted).

## fin\_indy\_pct

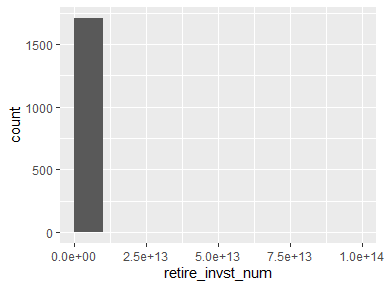
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 226 (11.31 %) |
| Number of unique values | 243 |
| Median | 21.8 |
| 1st and 3rd quartiles | 10; 40 |
| Min. and max. | -5; 4800000 |



* Note that the following possible outlier values were detected: "-5", "400", "95000", "2e+05", "260000", "4800000".

## retire\_invst\_num

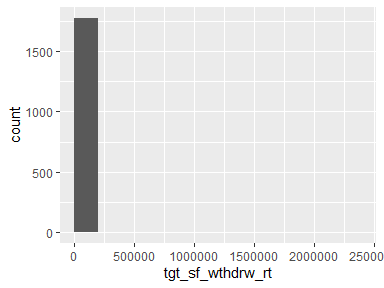
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 286 (14.31 %) |
| Number of unique values | 152 |
| Median | 2e+06 |
| 1st and 3rd quartiles | 1400000; 3e+06 |
| Min. and max. | 1; 1e+14 |



* Note that the following possible outlier values were detected: "1", "2.8", "6.67", "16", "20", …, "1e+08", "1.5e+08", "3.5e+08", "1e+09", "1e+14" (39 values omitted).

## tgt\_sf\_wthdrw\_rt

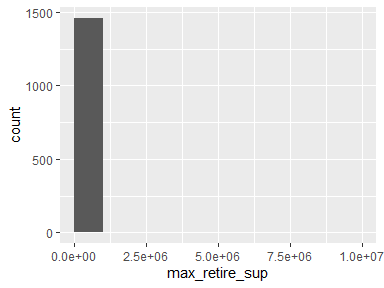
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 217 (10.86 %) |
| Number of unique values | 52 |
| Median | 3.5 |
| 1st and 3rd quartiles | 3.3; 4 |
| Min. and max. | 1; 2300000 |



* Note that the following possible outlier values were detected: "1", "2", "5.12", "5.5", "6", …, "83", "50000", "5e+05", "1500000", "2300000" (9 values omitted).

## max\_retire\_sup

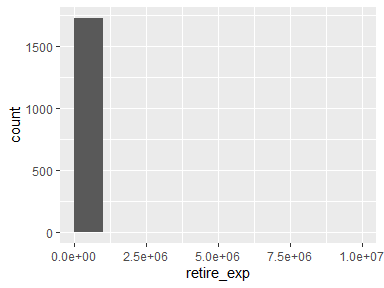
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 530 (26.53 %) |
| Number of unique values | 173 |
| Median | 20000 |
| 1st and 3rd quartiles | 4000; 40000 |
| Min. and max. | 0; 1e+07 |



* Note that the following possible outlier values were detected: "2e+05", "210000", "240000", "250000", "275000", …, "9e+05", "1250000", "2e+06", "2100000", "1e+07" (10 values omitted).

## retire\_exp

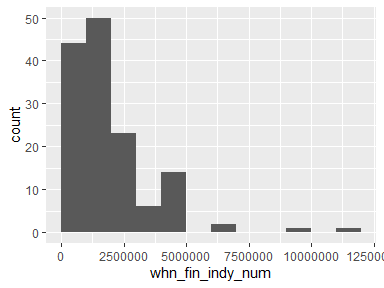
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 262 (13.11 %) |
| Number of unique values | 141 |
| Median | 60000 |
| 1st and 3rd quartiles | 40000; 90000 |
| Min. and max. | 0; 1e+07 |



* Note that the following possible outlier values were detected: "0", "2.5", "65", "1500", "2000", …, "1080000", "2100000", "3e+06", "3120000", "1e+07" (20 values omitted).

## whn\_fin\_indy\_num

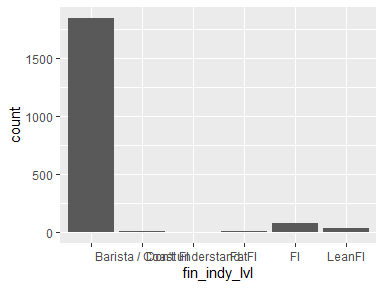
| Feature | Result |
| --- | --- |
| Variable type | integer |
| Number of missing obs. | 1857 (92.94 %) |
| Number of unique values | 45 |
| Median | 1500000 |
| 1st and 3rd quartiles | 1e+06; 2500000 |
| Min. and max. | 2500; 1.2e+07 |



* Note that the following possible outlier values were detected: "2500", "1e+05", "2e+05", "3e+05", "4e+05", "5e+05", "1.2e+07".

## fin\_indy\_lvl

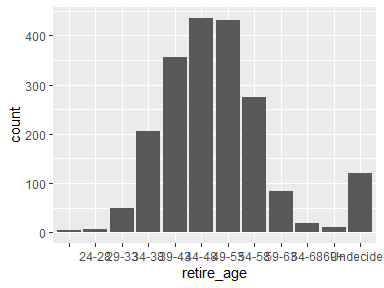
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 6 |
| Mode | “” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "Don’t understand".

## retire\_age

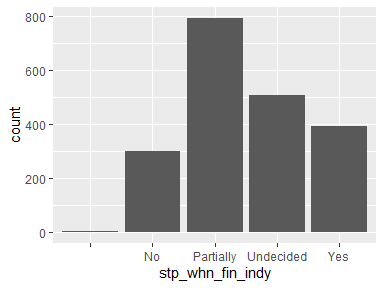
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 12 |
| Mode | “44-48” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "".

## stp\_whn\_fin\_indy

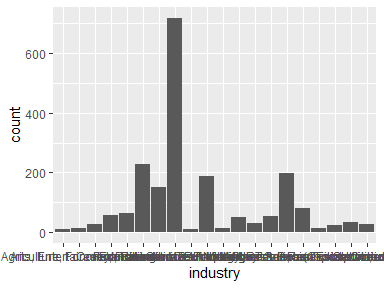
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 5 |
| Mode | “Partially” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "".

## industry

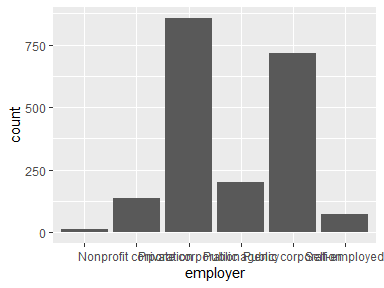
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 20 |
| Mode | “Information Technology” |



* The following suspected missing value codes enter as regular values: "".

## employer

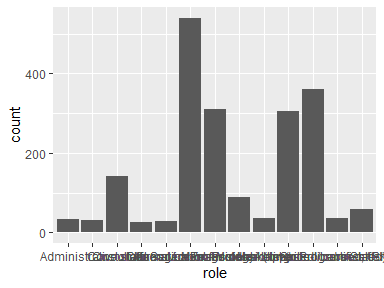
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 6 |
| Mode | “Private corporation” |



* The following suspected missing value codes enter as regular values: "".

## role

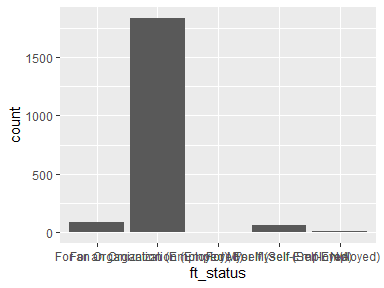
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 13 |
| Mode | “Information Technology” |



* The following suspected missing value codes enter as regular values: "".

## ft\_status

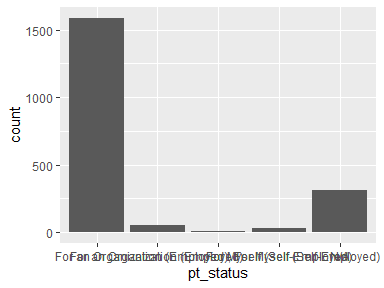
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 5 |
| Mode | “For an Organization (Employed)” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "For an Organization (Employed), For Myself (Self-Employed)".

## pt\_status

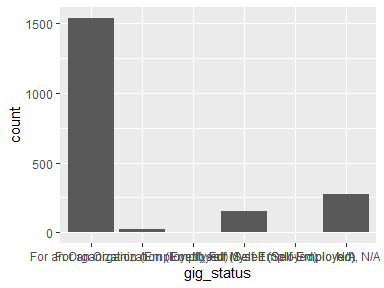
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 5 |
| Mode | “” |



* The following suspected missing value codes enter as regular values: "".

## gig\_status

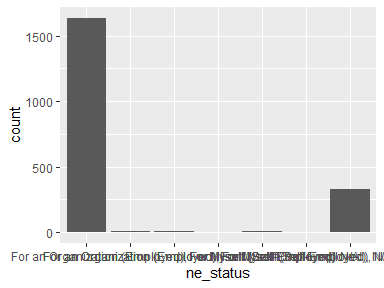
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 6 |
| Mode | “” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "For an Organization (Employed), For Myself (Self-Employed)", "For Myself (Self-Employed), N/A".

## ne\_status

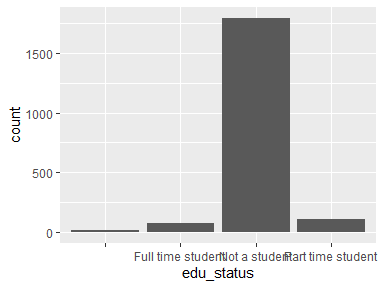
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 7 |
| Mode | “” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "For an Organization (Employed), For Myself (Self-Employed), N/A", "For Myself (Self-Employed), N/A".

## edu\_status

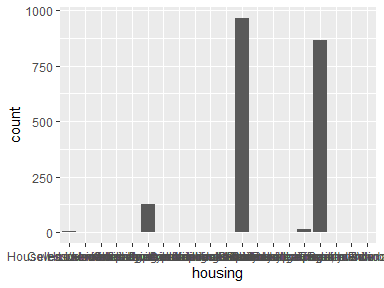
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 4 |
| Mode | “Not a student” |



* The following suspected missing value codes enter as regular values: "".

## housing

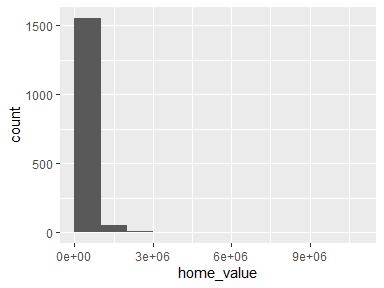
| Feature | Result |
| --- | --- |
| Variable type | character |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 20 |
| Mode | “Own” |



* The following suspected missing value codes enter as regular values: "".
* Note that the following levels have at most five observations: "", "Government housing", "Homeless", "House owned by parents", "Houseless/nomadic - living in a car voluntarily", …, "Own a house that I rent out but I rent an apartment in another state", "Own and live in a van", "Rent and own", "Rent my apartment, but own a home I rent out", "Rent temporarily while building new home" (6 values omitted).

## home\_value

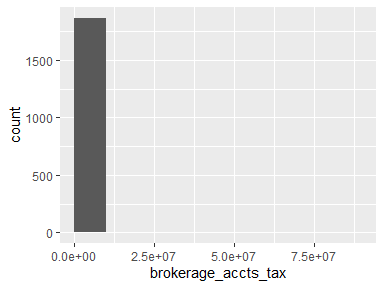
| Feature | Result |
| --- | --- |
| Variable type | integer |
| Number of missing obs. | 375 (18.77 %) |
| Number of unique values | 260 |
| Median | 224000 |
| 1st and 3rd quartiles | 0; 450000 |
| Min. and max. | 0; 1.1e+07 |



* Note that the following possible outlier values were detected: "1400000", "1500000", "1550000", "1600000", "1660000", …, "5e+06", "5300000", "6400000", "1e+07", "1.1e+07" (8 values omitted).

## brokerage\_accts\_tax

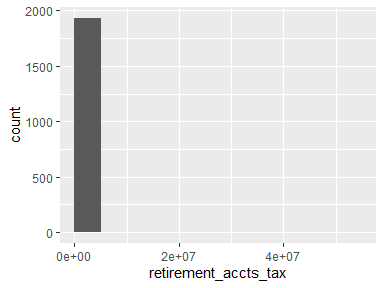
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 125 (6.26 %) |
| Number of unique values | 627 |
| Median | 78000 |
| 1st and 3rd quartiles | 15000; 267000 |
| Min. and max. | 0; 8.5e+07 |



* Note that the following possible outlier values were detected: "2600000", "3e+06", "3400000", "3440820", "3600000", …, "8e+06", "10100000", "1.2e+07", "20500000", "8.5e+07" (12 values omitted).

## retirement\_accts\_tax

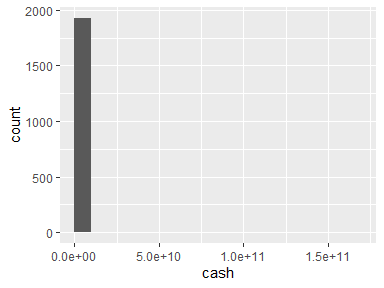
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 58 (2.9 %) |
| Number of unique values | 705 |
| Median | 180000 |
| 1st and 3rd quartiles | 80000; 4e+05 |
| Min. and max. | 0; 51500000 |



* Note that the following possible outlier values were detected: "2500000", "2518661", "2700000", "2800000", "2900000", …, "5100000", "1e+07", "1.3e+07", "2e+07", "51500000" (2 values omitted).

## cash

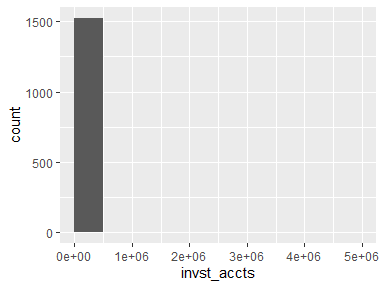
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 66 (3.3 %) |
| Number of unique values | 426 |
| Median | 26600 |
| 1st and 3rd quartiles | 11650; 55000 |
| Min. and max. | 0; 1.67e+11 |



* Note that the following possible outlier values were detected: "0", "5", "10", "30", "45", …, "2e+06", "3070000", "4500000", "3e+07", "1.67e+11" (22 values omitted).

## invst\_accts

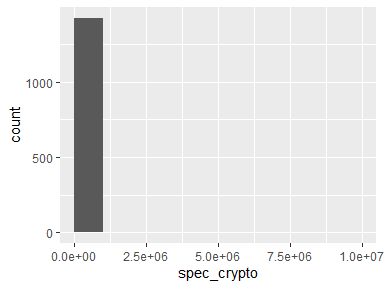
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 469 (23.47 %) |
| Number of unique values | 278 |
| Median | 4500 |
| 1st and 3rd quartiles | 0; 15000 |
| Min. and max. | 0; 5e+06 |



* Note that the following possible outlier values were detected: "130000", "140000", "144000", "150000", "155000", …, "4e+05", "5e+05", "685000", "1500000", "5e+06" (11 values omitted).

## spec\_crypto

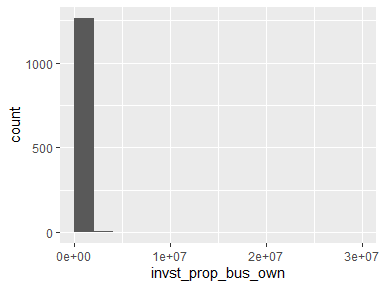
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 565 (28.28 %) |
| Number of unique values | 180 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 2000 |
| Min. and max. | 0; 1e+07 |



* Note that the following possible outlier values were detected: "65000", "70000", "73000", "75000", "80000", …, "3e+06", "3300000", "5e+06", "5500000", "1e+07" (43 values omitted).

## invst\_prop\_bus\_own

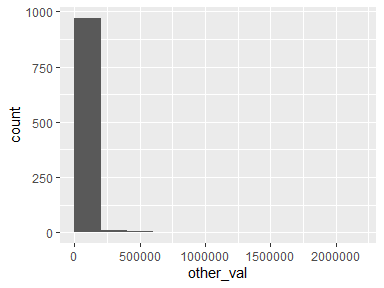
| Feature | Result |
| --- | --- |
| Variable type | integer |
| Number of missing obs. | 720 (36.04 %) |
| Number of unique values | 133 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 0 |
| Min. and max. | 0; 3e+07 |



* Note that the following possible outlier values were detected: "1", "500", "790", "1000", "1100", …, "3500000", "4300000", "5e+06", "7500000", "3e+07" (122 values omitted).

## other\_val

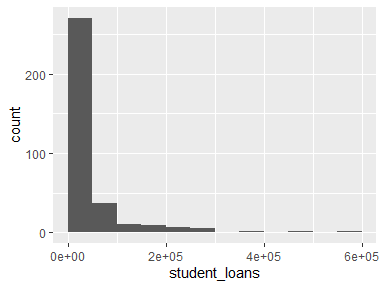
| Feature | Result |
| --- | --- |
| Variable type | integer |
| Number of missing obs. | 999 (50 %) |
| Number of unique values | 101 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 0 |
| Min. and max. | 0; 2040000 |



* Note that the following possible outlier values were detected: "300", "500", "1000", "1200", "2000", …, "1191000", "1500000", "1680000", "2e+06", "2040000" (90 values omitted).

## student\_loans

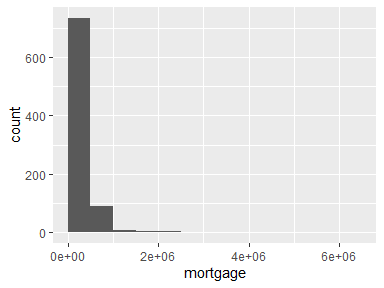
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1654 (82.78 %) |
| Number of unique values | 156 |
| Median | 19875 |
| 1st and 3rd quartiles | 9950; 41375 |
| Min. and max. | 0; 580000 |



* Note that the following possible outlier values were detected: "0", "0", "0.01", "1", "100", …, "280000", "294760", "370000", "5e+05", "580000" (13 values omitted).

## mortgage

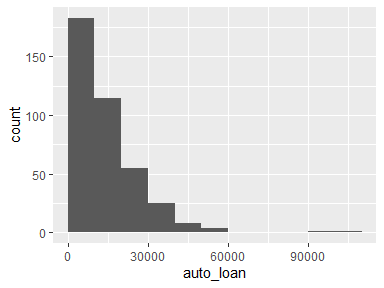
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1154 (57.76 %) |
| Number of unique values | 424 |
| Median | 240000 |
| 1st and 3rd quartiles | 150000; 374125 |
| Min. and max. | 0; 6200000 |



* Note that the following possible outlier values were detected: "0", "0", "0.01", "1", "95", …, "2260000", "2500000", "3e+06", "3200000", "6200000" (20 values omitted).

## auto\_loan

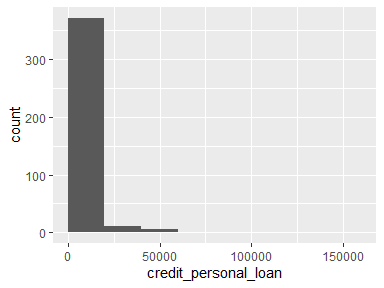
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1607 (80.43 %) |
| Number of unique values | 136 |
| Median | 12000 |
| 1st and 3rd quartiles | 5000; 20000 |
| Min. and max. | 0; 100125 |



* Note that the following possible outlier values were detected: "1e+05", "100125".

## credit\_personal\_loan

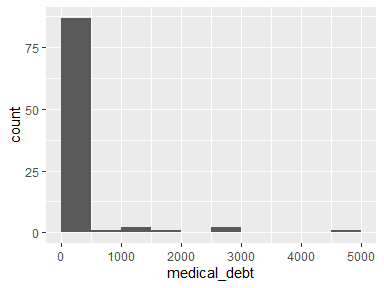
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1608 (80.48 %) |
| Number of unique values | 122 |
| Median | 2000 |
| 1st and 3rd quartiles | 908.75; 4045 |
| Min. and max. | 0; 150000 |



* Note that the following possible outlier values were detected: "18300", "19000", "20000", "22035", "23000", …, "58000", "60000", "1e+05", "110000", "150000" (6 values omitted).

## medical\_debt

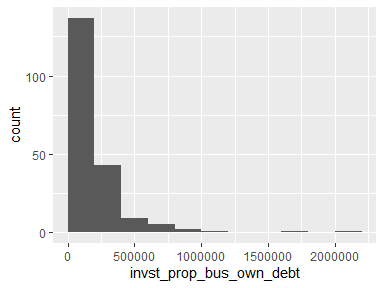
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1904 (95.3 %) |
| Number of unique values | 16 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 1 |
| Min. and max. | 0; 5000 |



* Note that the following possible outlier values were detected: "100", "225", "250", "500", "1000", "1200", "1300", "2000", "3000", "5000".

## invst\_prop\_bus\_own\_debt

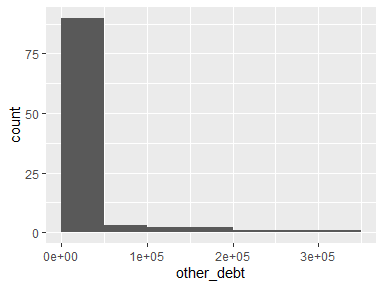
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1799 (90.04 %) |
| Number of unique values | 102 |
| Median | 96000 |
| 1st and 3rd quartiles | 0; 240500 |
| Min. and max. | 0; 2130000 |



* Note that the following possible outlier values were detected: "1624704", "2130000".

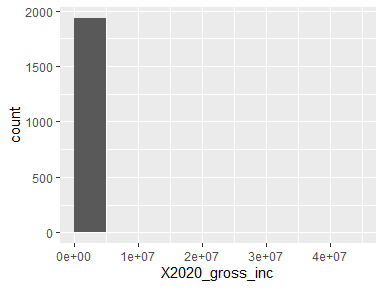
## other\_debt

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1898 (94.99 %) |
| Number of unique values | 37 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 11250 |
| Min. and max. | 0; 347000 |



## X2020\_gross\_inc

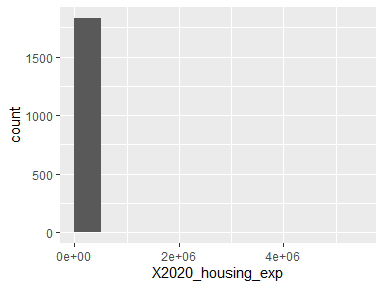
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 56 (2.8 %) |
| Number of unique values | 612 |
| Median | 140000 |
| 1st and 3rd quartiles | 95000; 226314.25 |
| Min. and max. | 0; 4.5e+07 |



* Note that the following possible outlier values were detected: "0", "100", "1400", "4500", "6500", …, "4800000", "7700000", "1e+07", "1.5e+07", "4.5e+07" (83 values omitted).

## X2020\_housing\_exp

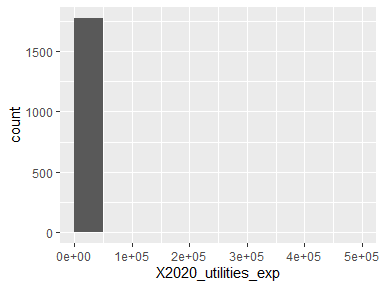
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 159 (7.96 %) |
| Number of unique values | 576 |
| Median | 17712 |
| 1st and 3rd quartiles | 10000; 26000 |
| Min. and max. | 1; 5500000 |



* Note that the following possible outlier values were detected: "63600", "64000", "65000", "65400", "65500", …, "1100000", "1200000", "1620000", "3300000", "5500000" (25 values omitted).

## X2020\_utilities\_exp

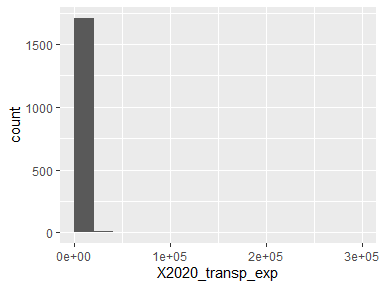
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 213 (10.66 %) |
| Number of unique values | 438 |
| Median | 2400 |
| 1st and 3rd quartiles | 1200; 4000 |
| Min. and max. | 1; 480000 |



* Note that the following possible outlier values were detected: "13000", "14000", "14300", "14400", "15000", …, "144000", "179013", "3e+05", "360000", "480000" (13 values omitted).

## X2020\_transp\_exp

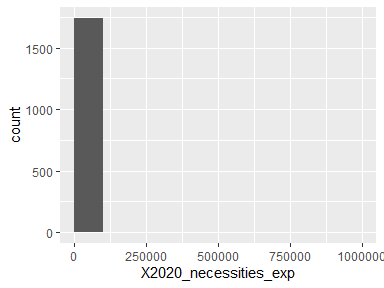
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 265 (13.26 %) |
| Number of unique values | 427 |
| Median | 2200 |
| 1st and 3rd quartiles | 1000; 5000 |
| Min. and max. | 1; 3e+05 |



* Note that the following possible outlier values were detected: "29899", "30000", "32227", "43000", "45000", …, "50000", "56235", "75000", "160000", "3e+05" (1 values omitted).

## X2020\_necessities\_exp

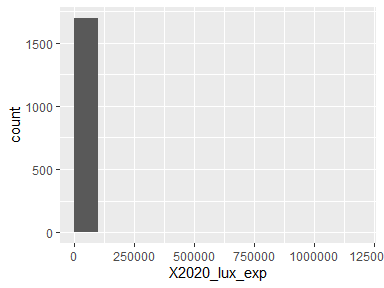
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 245 (12.26 %) |
| Number of unique values | 363 |
| Median | 6000 |
| 1st and 3rd quartiles | 3500; 10000 |
| Min. and max. | 18; 1e+06 |



* Note that the following possible outlier values were detected: "18", "20", "30", "100", "150", …, "160000", "350000", "360000", "476807", "1e+06" (29 values omitted).

## X2020\_lux\_exp

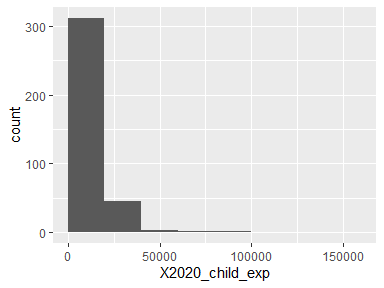
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 292 (14.61 %) |
| Number of unique values | 358 |
| Median | 5000 |
| 1st and 3rd quartiles | 2000; 10000 |
| Min. and max. | 25; 1150018 |



* Note that the following possible outlier values were detected: "50000", "60000", "60800", "64000", "70000", …, "2e+05", "348950", "6e+05", "8e+05", "1150018" (5 values omitted).

## X2020\_child\_exp

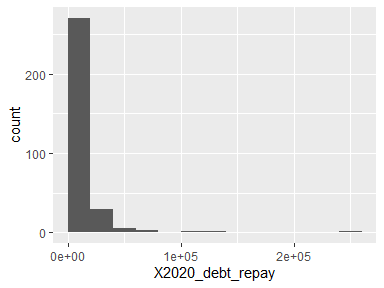
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1632 (81.68 %) |
| Number of unique values | 118 |
| Median | 5000 |
| 1st and 3rd quartiles | 1791; 12827.5 |
| Min. and max. | 0; 160000 |



* Note that the following possible outlier values were detected: "160000".

## X2020\_debt\_repay

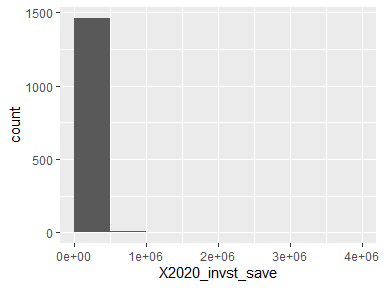
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1686 (84.38 %) |
| Number of unique values | 125 |
| Median | 2400 |
| 1st and 3rd quartiles | 0.01; 10611 |
| Min. and max. | 0; 250000 |



* Note that the following possible outlier values were detected: "120000", "140000", "250000".

## X2020\_invst\_save

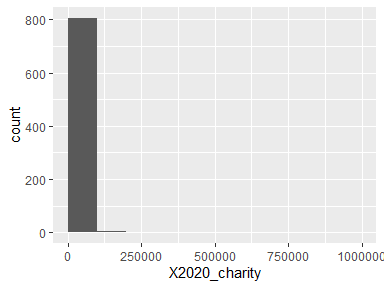
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 524 (26.23 %) |
| Number of unique values | 435 |
| Median | 45346 |
| 1st and 3rd quartiles | 25000; 80000 |
| Min. and max. | 1; 4e+06 |



* Note that the following possible outlier values were detected: "1", "5", "35", "50", "100", …, "850000", "1200000", "1500000", "2200000", "4e+06" (27 values omitted).

## X2020\_charity

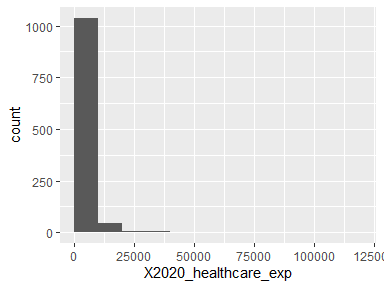
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1186 (59.36 %) |
| Number of unique values | 187 |
| Median | 1000 |
| 1st and 3rd quartiles | 357.5; 3600 |
| Min. and max. | 0; 1e+06 |



* Note that the following possible outlier values were detected: "46000", "50000", "55000", "1e+05", "120000", "129000", "190000", "1e+06".

## X2020\_healthcare\_exp

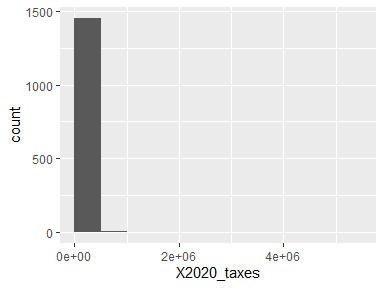
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 897 (44.89 %) |
| Number of unique values | 291 |
| Median | 1850 |
| 1st and 3rd quartiles | 600; 4000 |
| Min. and max. | 0; 115000 |



* Note that the following possible outlier values were detected: "20000", "20665", "21578", "24000", "25000", …, "48000", "50000", "54000", "63750", "115000" (6 values omitted).

## X2020\_taxes

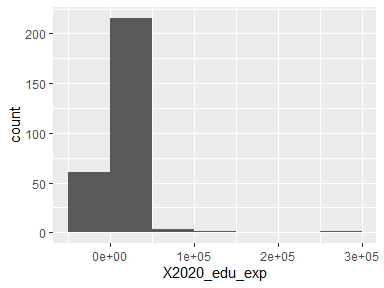
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 537 (26.88 %) |
| Number of unique values | 536 |
| Median | 30000 |
| 1st and 3rd quartiles | 16000; 55000 |
| Min. and max. | 1; 5300000 |



* Note that the following possible outlier values were detected: "1", "10", "120", "150", "300", …, "7e+05", "8e+05", "9e+05", "5e+06", "5300000" (31 values omitted).

## X2020\_edu\_exp

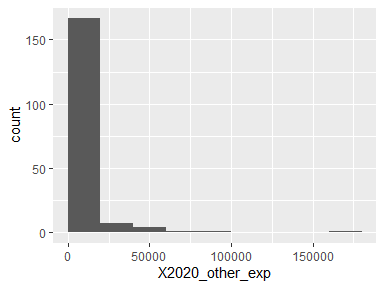
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1716 (85.89 %) |
| Number of unique values | 106 |
| Median | 1200 |
| 1st and 3rd quartiles | 18.5; 5000 |
| Min. and max. | -245; 279869 |



* Note that the following possible outlier values were detected: "70000", "95000", "1e+05", "150000", "279869".

## X2020\_other\_exp

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1817 (90.94 %) |
| Number of unique values | 95 |
| Median | 2500 |
| 1st and 3rd quartiles | 1; 8200 |
| Min. and max. | 0; 170000 |



* Note that the following possible outlier values were detected: "67165", "1e+05", "170000".

Report generation information:

* Created by: Vishwanath Guruvayur (username: vishg).
* Report creation time: Tue Jul 30 2024 23:28:15
* Report was run from directory: C:/Users/vishg/Documents/MSDS/STAT6021 - Linear Models for Data Science/financial\_independence/codes
* dataMaid v1.4.1 [Pkg: 2021-10-08 from CRAN (R 4.4.1)]
* R version 4.4.0 (2024-04-24 ucrt).
* Platform: x86\_64-w64-mingw32/x64(America/New\_York).
* Function call: makeDataReport(data = fin\_df, replace = T)