WHI.TotalPA.an\_samp

Autogenerated data summary from dataMaid

2023-11-14 14:54:55

# Data report overview

The dataset examined has the following dimensions:

| Feature | Result |
| --- | --- |
| Number of observations | 83410 |
| Number of variables | 75 |

### 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

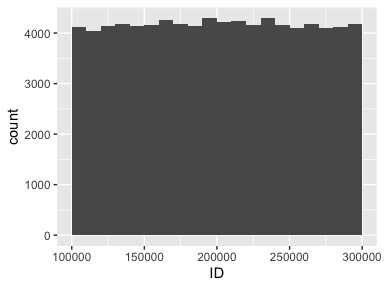
|  | Variable class | # unique values | Missing observations | Any problems? |
| --- | --- | --- | --- | --- |
| [ID](#id) | numeric | 83410 | 0.00 % |  |
| [OSFLAG](#osflag) | numeric | 2 | 0.00 % |  |
| [AGE](#age) | numeric | 33 | 0.00 % |  |
| [ETHNIC](#ethnic) | numeric | 7 | 0.22 % | × |
| [HRTARM](#hrtarm) | numeric | 5 | 0.00 % |  |
| [DMARM](#dmarm) | numeric | 3 | 0.00 % |  |
| [mi\_f2](#mi_f2) | numeric | 2 | 0.03 % | × |
| [DIABTRT](#diabtrt) | numeric | 3 | 0.06 % |  |
| [NOMAM2YR](#nomam2yr) | numeric | 3 | 2.84 % |  |
| [TIMELSTS](#timelsts) | hms | 3 | 3.31 % | × |
| [ANYINS](#anyins) | numeric | 3 | 0.86 % |  |
| [nses](#nses) | numeric | 36356 | 10.53 % | × |
| [HICHOLRP](#hicholrp) | numeric | 3 | 1.50 % |  |
| [HEI2005](#hei2005) | numeric | 82311 | 0.16 % | × |
| [F60SODUM](#f60sodum) | numeric | 83263 | 0.16 % | × |
| [PULSE30](#pulse30) | numeric | 75 | 0.17 % | × |
| [WAIST](#waist) | numeric | 788 | 0.36 % | × |
| [SYST](#syst) | numeric | 115 | 0.09 % |  |
| [DIAS](#dias) | numeric | 91 | 0.11 % | × |
| [BMI](#bmi) | numeric | 50373 | 0.87 % | × |
| [WALK](#walk) | numeric | 6 | 0.00 % |  |
| [WALKMIN](#walkmin) | numeric | 5 | 14.01 % |  |
| [WALKSPD](#walkspd) | numeric | 6 | 14.00 % | × |
| [HRDEX](#hrdex) | numeric | 7 | 0.82 % |  |
| [HRDEXMIN](#hrdexmin) | numeric | 5 | 72.23 % |  |
| [MODEX](#modex) | numeric | 7 | 0.82 % |  |
| [MODEXMIN](#modexmin) | numeric | 5 | 52.76 % |  |
| [MLDEX](#mldex) | numeric | 7 | 1.69 % |  |
| [MLDEXMIN](#mldexmin) | numeric | 5 | 72.50 % |  |
| [HRDEX18](#hrdex18) | numeric | 3 | 3.07 % |  |
| [HRDEX35](#hrdex35) | numeric | 3 | 2.72 % |  |
| [HRDEX50](#hrdex50) | numeric | 3 | 2.05 % |  |
| [TEPIWK](#tepiwk) | numeric | 43 | 0.00 % | × |
| [LEPITOT](#lepitot) | numeric | 43 | 0.00 % | × |
| [MSEPIWK](#msepiwk) | numeric | 31 | 0.00 % |  |
| [XLMSEPI](#xlmsepi) | numeric | 31 | 0.00 % |  |
| [LMSEPI](#lmsepi) | numeric | 4 | 0.00 % |  |
| [SEPIWK](#sepiwk) | numeric | 6 | 0.00 % |  |
| [TMINWK](#tminwk) | numeric | 330 | 0.00 % | × |
| [MSMINWK](#msminwk) | numeric | 235 | 0.00 % | × |
| [SMINWK](#sminwk) | numeric | 19 | 0.00 % |  |
| [TEXPWK](#texpwk) | numeric | 1341 | 0.00 % | × |
| [WALKEXP](#walkexp) | numeric | 41 | 0.00 % | × |
| [SMOKING](#smoking) | numeric | 4 | 0.93 % |  |
| [HARDEXP](#hardexp) | numeric | 19 | 0.00 % |  |
| [MODEXP](#modexp) | numeric | 19 | 0.00 % |  |
| [MILDEXP](#mildexp) | numeric | 19 | 0.00 % |  |
| [ALCOHOL](#alcohol) | numeric | 7 | 0.47 % |  |
| [PACKYRS](#packyrs) | numeric | 203 | 3.18 % |  |
| [INDRCHOR](#indrchor) | numeric | 6 | 39.68 % |  |
| [YRDWORKM](#yrdworkm) | numeric | 6 | 39.83 % |  |
| [YRDWRKH](#yrdwrkh) | numeric | 6 | 63.15 % |  |
| [SITTING](#sitting) | numeric | 9 | 39.79 % |  |
| [SLEEPING](#sleeping) | numeric | 9 | 39.69 % | × |
| [YRDMINWK](#yrdminwk) | numeric | 15 | 40.00 % |  |
| [YARDEXP](#yardexp) | numeric | 15 | 40.00 % |  |
| [INHEAEXP](#inheaexp) | numeric | 6 | 39.68 % |  |
| [INACT](#inact) | numeric | 18 | 39.88 % | × |
| [TOTESTAT](#totestat) | numeric | 4 | 0.06 % |  |
| [TOTPSTAT](#totpstat) | numeric | 4 | 0.04 % |  |
| [TOTHSTAT](#tothstat) | numeric | 4 | 0.09 % |  |
| [NIGHTSWT](#nightswt) | numeric | 5 | 0.48 % |  |
| [HOTFLASH](#hotflash) | numeric | 5 | 0.31 % |  |
| [ACTDLY](#actdly) | numeric | 9 | 3.25 % | × |
| [PHYSFUN](#physfun) | numeric | 20 | 0.80 % | × |
| [PSHTDEP](#pshtdep) | numeric | 2652 | 1.89 % | × |
| [SOCFUNC](#socfunc) | numeric | 10 | 0.28 % |  |
| [HTNrx](#htnrx) | numeric | 3 | 0.00 % |  |
| [F33PILLSHYP](#f33pillshyp) | numeric | 2 | 0.00 % |  |
| [timedy](#timedy) | numeric | 6547 | 0.00 % |  |
| [timeyr](#timeyr) | numeric | 6547 | 0.00 % |  |
| [educcat](#educcat) | numeric | 4 | 0.74 % |  |
| [HxCHD](#hxchd) | numeric | 1 | 0.00 % | × |
| [prevCanc](#prevcanc) | numeric | 2 | 0.00 % |  |
| [stroke\_f2](#stroke_f2) | numeric | 1 | 0.00 % | × |

# Variable list

## ID

*WHI Participant Common ID*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 83410 |
| Median | 200093.5 |
| 1st and 3rd quartiles | 150549.5; 249564.75 |
| Min. and max. | 100001; 299999 |

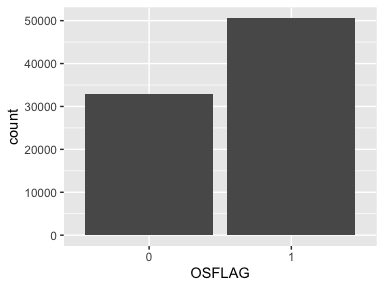


## OSFLAG

*OS Participant*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

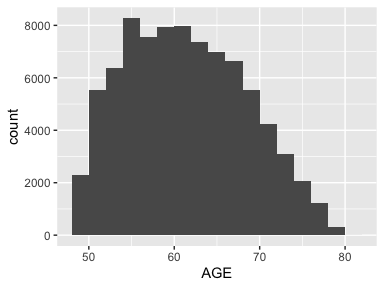
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 2 |
| Mode | “1” |
| Reference category | 0 |



## AGE

*Age at screening*

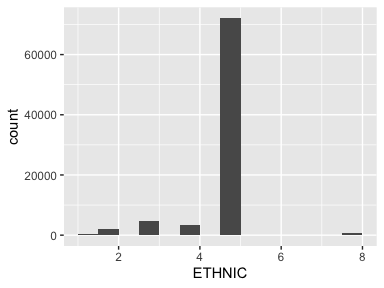
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 33 |
| Median | 61 |
| 1st and 3rd quartiles | 56; 67 |
| Min. and max. | 49; 81 |



## ETHNIC

*Ethnicity*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 186 (0.22 %) |
| Number of unique values | 6 |
| Median | 5 |
| 1st and 3rd quartiles | 5; 5 |
| Min. and max. | 1; 8 |



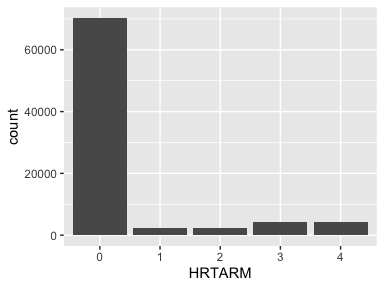
* The following suspected missing value codes enter as regular values: "8".
* Note that the following possible outlier values were detected: "1", "2", "3", "4", "8".

## HRTARM

*HRT Arm*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

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

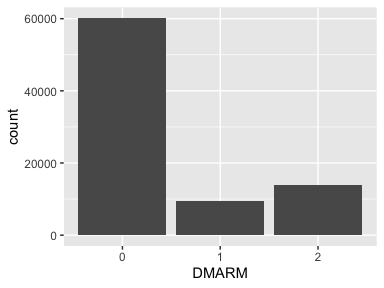


## DMARM

*DM Arm*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 3 |
| Mode | “0” |
| Reference category | 0 |



## mi\_f2

*MI ever*

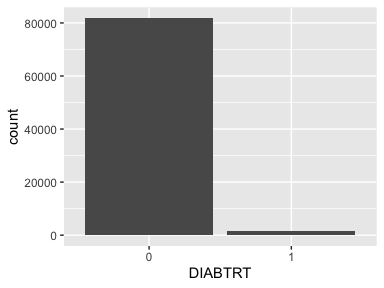
* The variable only takes one (non-missing) value: "0". The variable contains 0.03 % missing observations.

## DIABTRT

*Diabetes treated (pills or shots)*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 51 (0.06 %) |
| Number of unique values | 2 |
| Mode | “0” |
| Reference category | 0 |

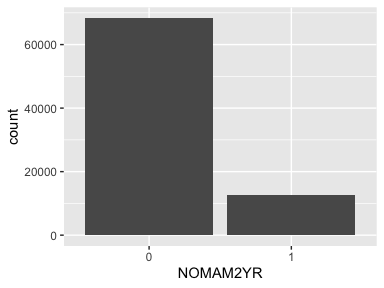


## NOMAM2YR

*No mammogram in last 2 years*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 2365 (2.84 %) |
| Number of unique values | 2 |
| Mode | “0” |
| Reference category | 0 |



## TIMELSTS

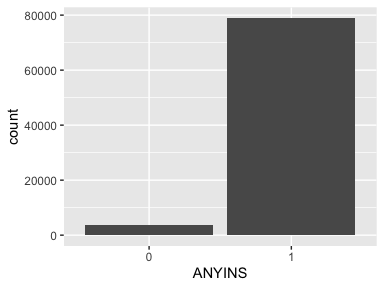
* The variable has class hms which is not supported by dataMaid.

## ANYINS

*Any Insurance*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

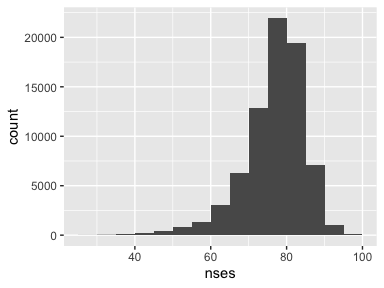
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 718 (0.86 %) |
| Number of unique values | 2 |
| Mode | “1” |
| Reference category | 0 |



## nses

*Neighborhood SES*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 8785 (10.53 %) |
| Number of unique values | 36355 |
| Median | 77.95 |
| 1st and 3rd quartiles | 72.99; 81.9 |
| Min. and max. | 27.83; 98.37 |



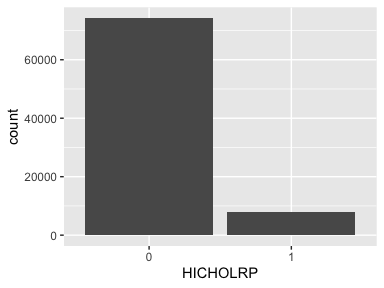
* Note that the following possible outlier values were detected: "27.83", "27.92", "28.01", "28.47", "28.56", …, "97.87", "98.07", "98.18", "98.22", "98.37" (1419 values omitted).

## HICHOLRP

*High cholesterol requiring pills ever*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

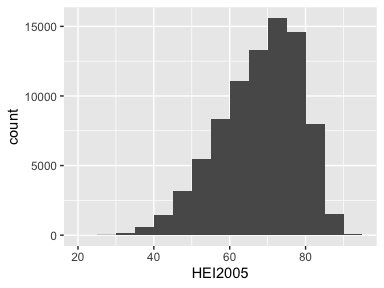
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1250 (1.5 %) |
| Number of unique values | 2 |
| Mode | “0” |
| Reference category | 0 |



## HEI2005

*TOTAL HEI-2005 SCORE*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 136 (0.16 %) |
| Number of unique values | 82310 |
| Median | 69.32 |
| 1st and 3rd quartiles | 60.8; 76.08 |
| Min. and max. | 22.72; 93.44 |

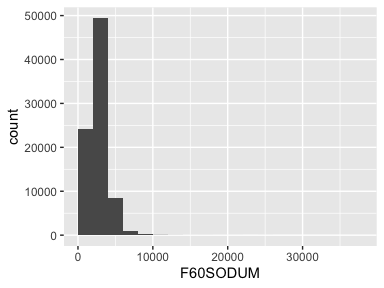


* Note that the following possible outlier values were detected: "22.72", "23.65", "24.75", "25.79", "25.86", …, "92.23", "92.44", "93.02", "93.38", "93.44" (93 values omitted).

## F60SODUM

*Dietary Sodium (mg)*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 136 (0.16 %) |
| Number of unique values | 83262 |
| Median | 2523.43 |
| 1st and 3rd quartiles | 1896.51; 3278.58 |
| Min. and max. | 15.06; 36443.54 |

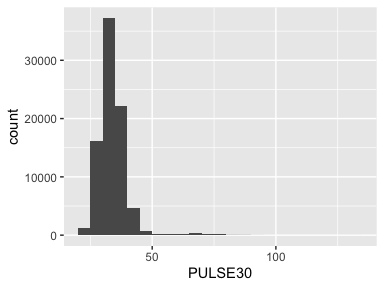


* Note that the following possible outlier values were detected: "15.06", "26.08", "69.75", "83.12", "160.6", …, "22460.73", "24560.67", "25682.45", "27307.56", "36443.54" (1437 values omitted).

## PULSE30

*Resting pulse in 30 seconds*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 142 (0.17 %) |
| Number of unique values | 74 |
| Median | 34 |
| 1st and 3rd quartiles | 31; 37 |
| Min. and max. | 20; 135 |

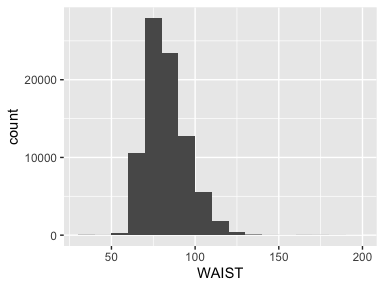


* Note that the following possible outlier values were detected: "20", "21", "47", "48", "49", …, "114", "128", "129", "133", "135" (39 values omitted).

## WAIST

*Waist, cm*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 302 (0.36 %) |
| Number of unique values | 787 |
| Median | 81 |
| 1st and 3rd quartiles | 74; 90 |
| Min. and max. | 35; 199 |

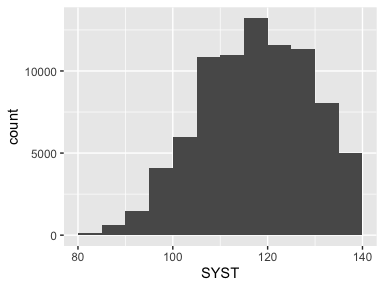


* Note that the following possible outlier values were detected: "35", "35.5", "36", "36.5", "37", …, "191.8", "192", "195", "196", "199" (156 values omitted).

## SYST

*Systolic BP*

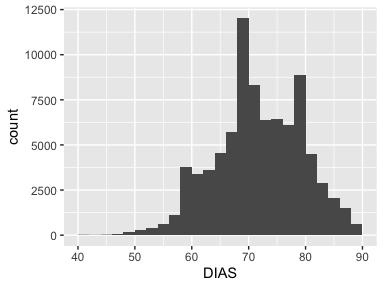
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 79 (0.09 %) |
| Number of unique values | 114 |
| Median | 119 |
| 1st and 3rd quartiles | 110; 127 |
| Min. and max. | 80; 139.5 |



## DIAS

*Diastolic BP*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 91 (0.11 %) |
| Number of unique values | 90 |
| Median | 72 |
| 1st and 3rd quartiles | 68; 78 |
| Min. and max. | 40; 89.5 |

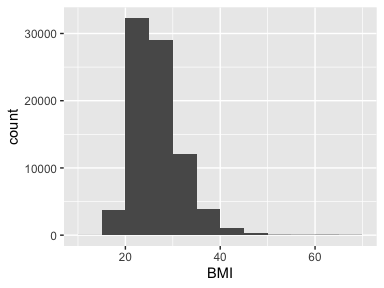


* Note that the following possible outlier values were detected: "40", "42", "42.5", "43", "44", …, "55.5", "56", "56.5", "57", "57.5" (16 values omitted).

## BMI

*Body-mass Index (BMI), kg/m2*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 726 (0.87 %) |
| Number of unique values | 50372 |
| Median | 25.71 |
| 1st and 3rd quartiles | 23.02; 29.27 |
| Min. and max. | 11.53; 69.88 |

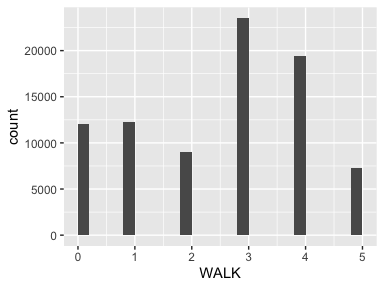


* Note that the following possible outlier values were detected: "11.53", "11.86", "11.97", "12.01", "12.11", …, "69.33", "69.4", "69.55", "69.59", "69.88" (1248 values omitted).

## WALK

*Times walk for > 10 min*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 6 |
| Median | 3 |
| 1st and 3rd quartiles | 1; 4 |
| Min. and max. | 0; 5 |

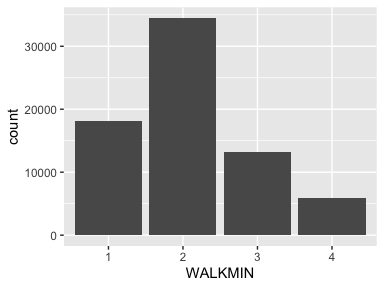


## WALKMIN

*Duration of walks when >10 min*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 11682 (14.01 %) |
| Number of unique values | 4 |
| Mode | “2” |
| Reference category | 1 |

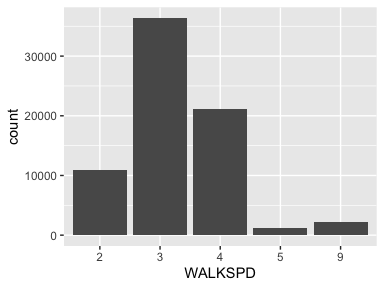


## WALKSPD

*Walking speed when walking for >10 min*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 11677 (14 %) |
| Number of unique values | 5 |
| Mode | “3” |
| Reference category | 2 |

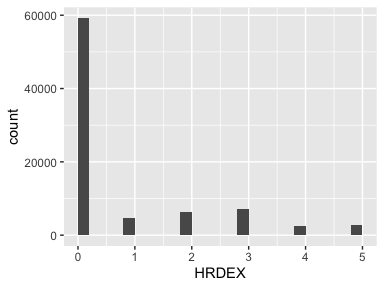


* The following suspected missing value codes enter as regular values: "9".

## HRDEX

*Times per week of very hard exercise*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 681 (0.82 %) |
| Number of unique values | 6 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 1 |
| Min. and max. | 0; 5 |

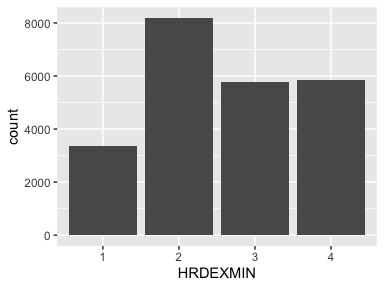


## HRDEXMIN

*Duration per time of very hard exercise*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

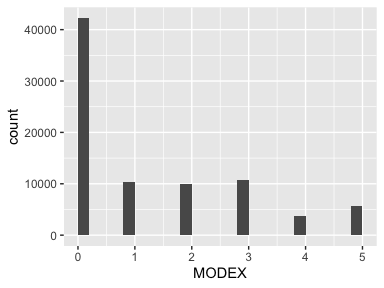
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 60246 (72.23 %) |
| Number of unique values | 4 |
| Mode | “2” |
| Reference category | 1 |



## MODEX

*Times per week of moderate exercise*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 688 (0.82 %) |
| Number of unique values | 6 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 2 |
| Min. and max. | 0; 5 |

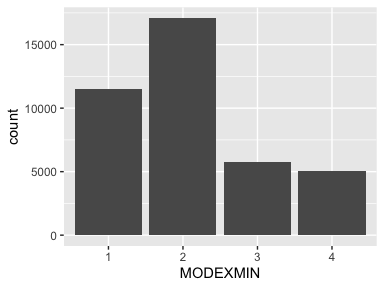


## MODEXMIN

*Duration per time of moderate exercise*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

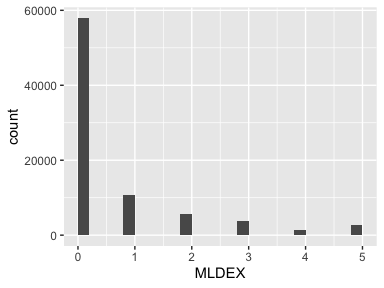
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 44009 (52.76 %) |
| Number of unique values | 4 |
| Mode | “2” |
| Reference category | 1 |



## MLDEX

*Times per week of mild exercise*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1413 (1.69 %) |
| Number of unique values | 6 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 1 |
| Min. and max. | 0; 5 |

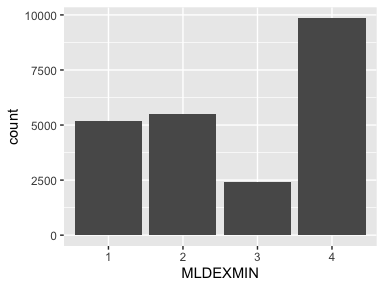


## MLDEXMIN

*Duration per time of mild exercise*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 60472 (72.5 %) |
| Number of unique values | 4 |
| Mode | “4” |
| Reference category | 1 |

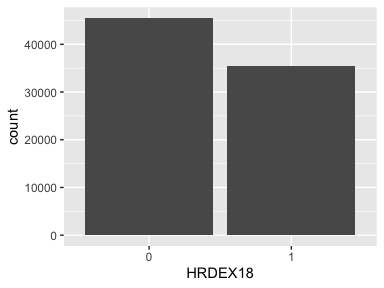


## HRDEX18

*Very hard exercise 3 times/wk at age 18*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 2564 (3.07 %) |
| Number of unique values | 2 |
| Mode | “0” |
| Reference category | 0 |

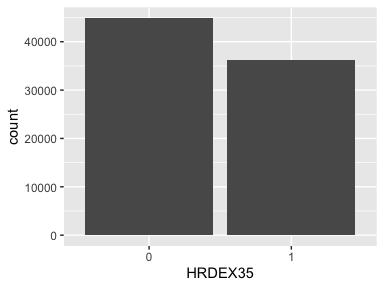


## HRDEX35

*Very hard exercise 3 times/wk at age 35*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 2270 (2.72 %) |
| Number of unique values | 2 |
| Mode | “0” |
| Reference category | 0 |

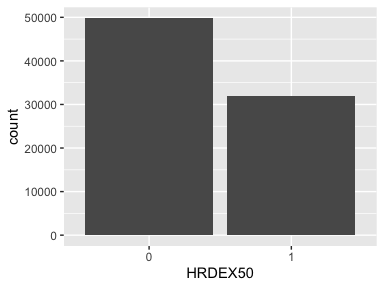


## HRDEX50

*Very hard exercise 3 times/wk at age 50*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

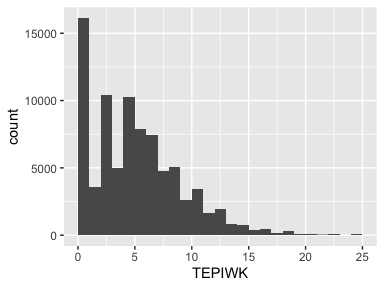
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1706 (2.05 %) |
| Number of unique values | 2 |
| Mode | “0” |
| Reference category | 0 |



## TEPIWK

*Episodes recreational phys activity per week*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 43 |
| Median | 5 |
| 1st and 3rd quartiles | 2.5; 8 |
| Min. and max. | 0; 25 |

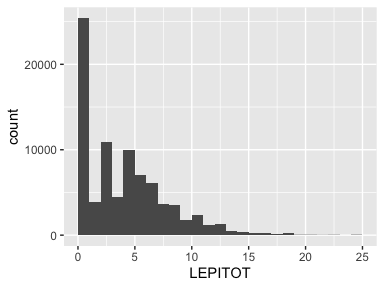


* Note that the following possible outlier values were detected: "16.5", "17", "17.5", "18", "18.5", …, "20.5", "21", "22", "23", "25" (2 values omitted).

## LEPITOT

*Episodes recreational phys activity per week >= 20 Min*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 43 |
| Median | 3.5 |
| 1st and 3rd quartiles | 1; 6.5 |
| Min. and max. | 0; 25 |

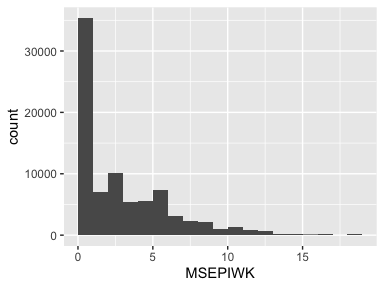


* Note that the following possible outlier values were detected: "21", "22", "23", "25".

## MSEPIWK

*Episodes moderate to strenuous phys activity per week*

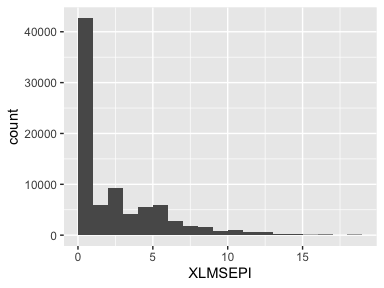
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 31 |
| Median | 2 |
| 1st and 3rd quartiles | 0; 5 |
| Min. and max. | 0; 19 |



## XLMSEPI

*Episodes moderate to strenuous activity >= 20 min/week*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 31 |
| Median | 1 |
| 1st and 3rd quartiles | 0; 4.5 |
| Min. and max. | 0; 19 |

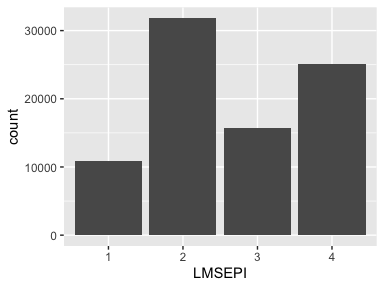


## LMSEPI

*Episodes moderate to strenuous activity >=20 min/week (categorical)*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

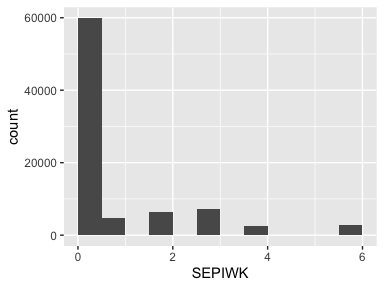
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 4 |
| Mode | “2” |
| Reference category | 1 |



## SEPIWK

*Strenuous activity episodes per week*

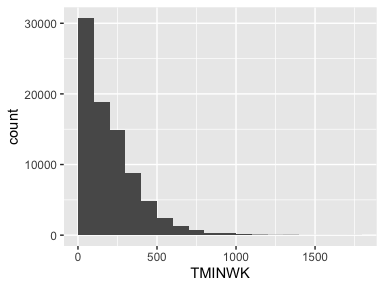
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 6 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 1 |
| Min. and max. | 0; 6 |



## TMINWK

*Minutes of recreational phys activity per week*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 330 |
| Median | 150 |
| 1st and 3rd quartiles | 60; 285 |
| Min. and max. | 0; 1750 |

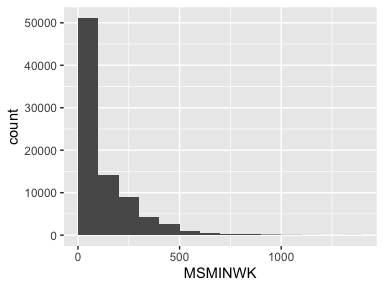


* Note that the following possible outlier values were detected: "960", "965", "967.5", "970", "980", …, "1510", "1540", "1610", "1630", "1750" (56 values omitted).

## MSMINWK

*Minutes of moderate to strenuous activity per week*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 235 |
| Median | 60 |
| 1st and 3rd quartiles | 0; 180 |
| Min. and max. | 0; 1330 |

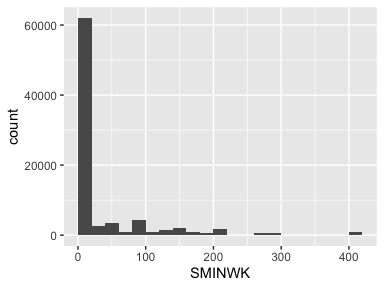


* Note that the following possible outlier values were detected: "1190", "1210", "1330".

## SMINWK

*Minutes of strenuous phys activity per week*

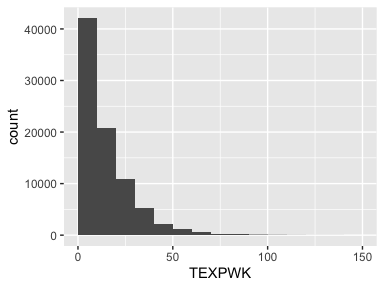
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 19 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 30 |
| Min. and max. | 0; 420 |



## TEXPWK

*Total energy expend from recreational phys activity (MET-hours/week)*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 1341 |
| Median | 10 |
| 1st and 3rd quartiles | 3.5; 20 |
| Min. and max. | 0; 142.33 |

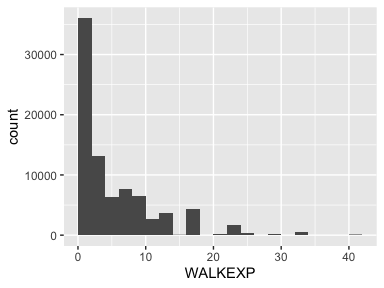


* Note that the following possible outlier values were detected: "74", "74.08", "74.17", "74.25", "74.33", …, "130.33", "130.67", "134.17", "136.33", "142.33" (192 values omitted).

## WALKEXP

*MET-hours per week from walking*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 41 |
| Median | 3.75 |
| 1st and 3rd quartiles | 0; 7.5 |
| Min. and max. | 0; 40.83 |



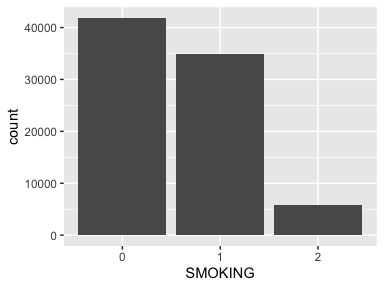
* Note that the following possible outlier values were detected: "32.67", "40.83".

## SMOKING

*Smoking status*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

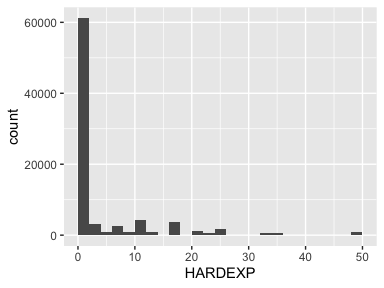
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 773 (0.93 %) |
| Number of unique values | 3 |
| Mode | “0” |
| Reference category | 0 |



## HARDEXP

*Energy expenditure from hard exercise (MET-hours/week)*

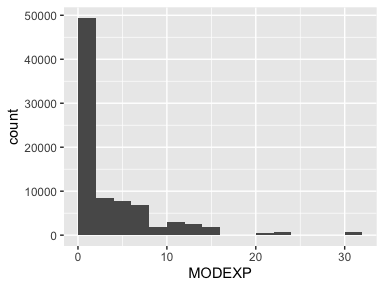
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 19 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 3.5 |
| Min. and max. | 0; 49 |



## MODEXP

*Energy expend from moderate exercise (MET-hours/week)*

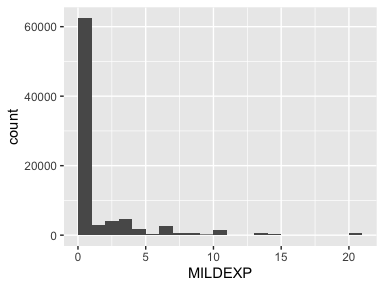
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 19 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 4.5 |
| Min. and max. | 0; 31.5 |



## MILDEXP

*Energy expenditure from mild exercise (MET-hours/week)*

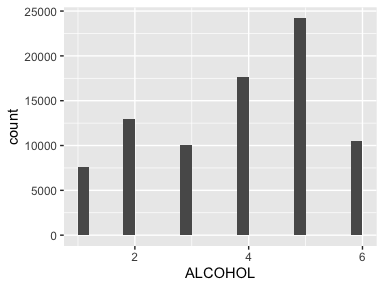
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 19 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 1.5 |
| Min. and max. | 0; 21 |



## ALCOHOL

*Alcohol intake*

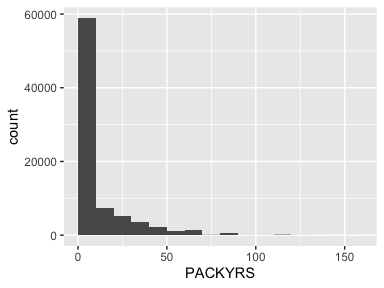
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 389 (0.47 %) |
| Number of unique values | 6 |
| Median | 4 |
| 1st and 3rd quartiles | 3; 5 |
| Min. and max. | 1; 6 |



## PACKYRS

*Pack years of smoking*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 2653 (3.18 %) |
| Number of unique values | 202 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 12.5 |
| Min. and max. | 0; 158.75 |



## INDRCHOR

*Hours doing heavy chores*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 33093 (39.68 %) |
| Number of unique values | 5 |
| Mode | “2” |
| Reference category | 1 |

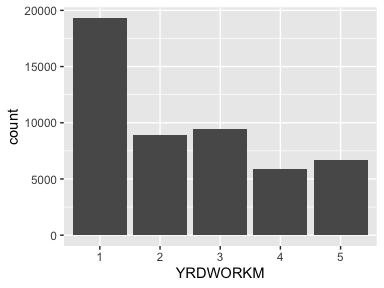


## YRDWORKM

*Months doing yard work*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 33219 (39.83 %) |
| Number of unique values | 5 |
| Mode | “1” |
| Reference category | 1 |

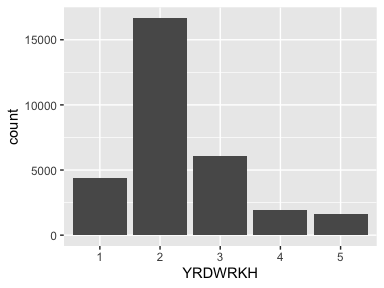


## YRDWRKH

*Hours in week doing yard work*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

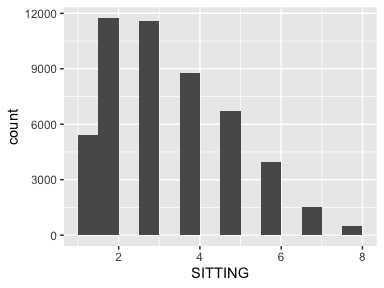
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 52673 (63.15 %) |
| Number of unique values | 5 |
| Mode | “2” |
| Reference category | 1 |



## SITTING

*Number of hours spent sitting*

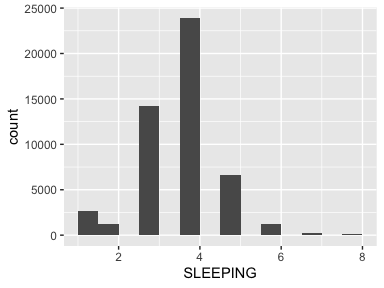
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 33193 (39.79 %) |
| Number of unique values | 8 |
| Median | 3 |
| 1st and 3rd quartiles | 2; 5 |
| Min. and max. | 1; 8 |



## SLEEPING

*Number of hours spent sleeping*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 33103 (39.69 %) |
| Number of unique values | 8 |
| Median | 4 |
| 1st and 3rd quartiles | 3; 4 |
| Min. and max. | 1; 8 |

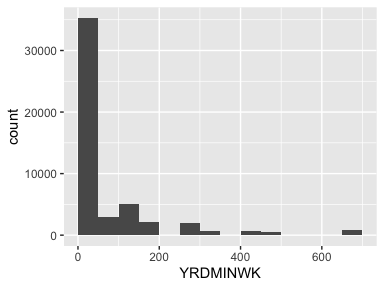


* Note that the following possible outlier values were detected: "5", "6", "7", "8".

## YRDMINWK

*Min/wk of yard work*

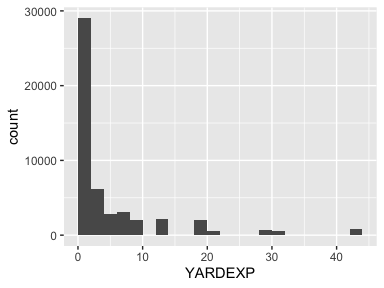
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 33367 (40 %) |
| Number of unique values | 14 |
| Median | 20 |
| 1st and 3rd quartiles | 0; 80 |
| Min. and max. | 0; 660 |



## YARDEXP

*Energy expend from yard work*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 33367 (40 %) |
| Number of unique values | 14 |
| Median | 1.33 |
| 1st and 3rd quartiles | 0; 5.33 |
| Min. and max. | 0; 44 |

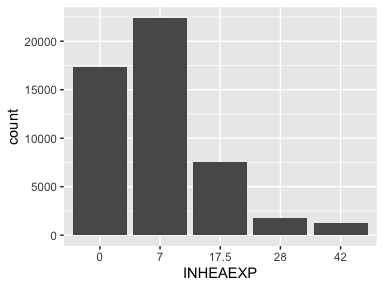


## INHEAEXP

*Energy expend heavy indoor chores*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

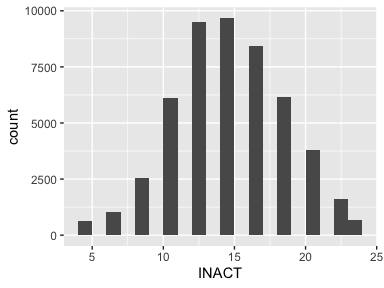
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 33093 (39.68 %) |
| Number of unique values | 5 |
| Mode | “7” |
| Reference category | 0 |



## INACT

*Hrs/day spent sitting or lying*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 33264 (39.88 %) |
| Number of unique values | 17 |
| Median | 15 |
| 1st and 3rd quartiles | 13; 17 |
| Min. and max. | 4; 24 |



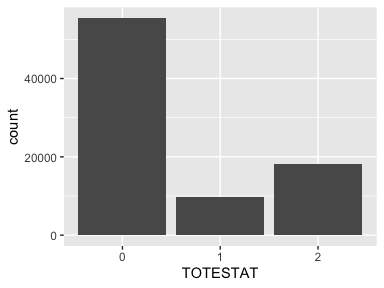
* Note that the following possible outlier values were detected: "4", "6.5", "24".

## TOTESTAT

*Unopposed estrogen usage status*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 49 (0.06 %) |
| Number of unique values | 3 |
| Mode | “0” |
| Reference category | 0 |

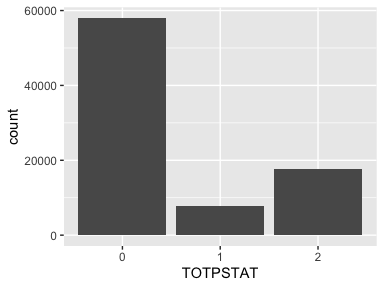


## TOTPSTAT

*Estrogen + progesterone usage status*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 34 (0.04 %) |
| Number of unique values | 3 |
| Mode | “0” |
| Reference category | 0 |

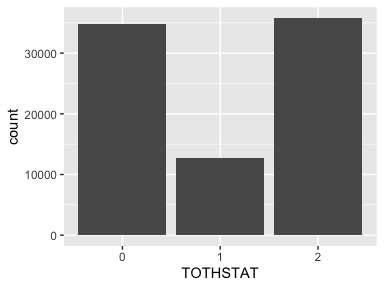


## TOTHSTAT

*HRT usage status*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 72 (0.09 %) |
| Number of unique values | 3 |
| Mode | “2” |
| Reference category | 0 |

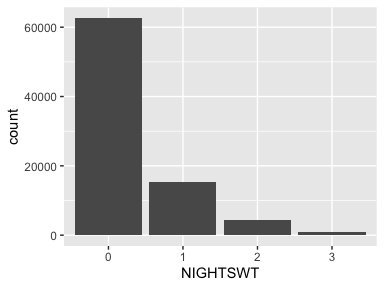


## NIGHTSWT

*Night sweats*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 399 (0.48 %) |
| Number of unique values | 4 |
| Mode | “0” |
| Reference category | 0 |

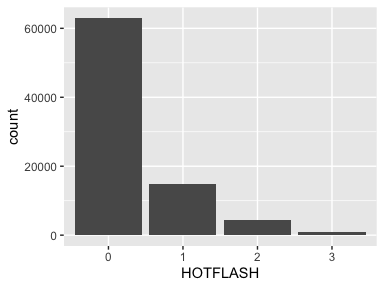


## HOTFLASH

*Hot flashes*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

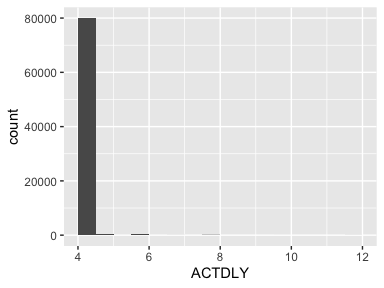
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 257 (0.31 %) |
| Number of unique values | 4 |
| Mode | “0” |
| Reference category | 0 |



## ACTDLY

*Activities of Daily Living Construct*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 2707 (3.25 %) |
| Number of unique values | 8 |
| Median | 4 |
| 1st and 3rd quartiles | 4; 4 |
| Min. and max. | 4; 12 |

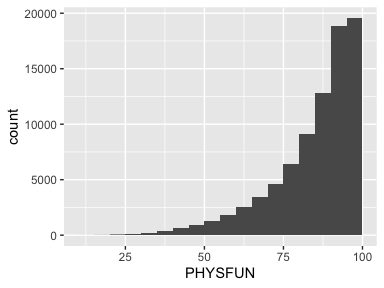


* Note that the following possible outlier values were detected: "5", "6", "7", "8", "9", "10", "12".

## PHYSFUN

*Physical Functioning Construct*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 671 (0.8 %) |
| Number of unique values | 19 |
| Median | 90 |
| 1st and 3rd quartiles | 80; 95 |
| Min. and max. | 10; 100 |

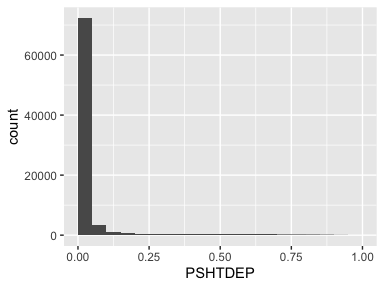


* Note that the following possible outlier values were detected: "10", "15", "20", "25", "30", "35".

## PSHTDEP

*Shortened CES-D/DIS Screening Instrument*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1573 (1.89 %) |
| Number of unique values | 2651 |
| Median | 0 |
| 1st and 3rd quartiles | 0; 0.01 |
| Min. and max. | 0; 0.96 |

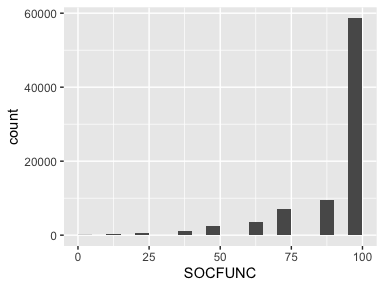


* Note that the following possible outlier values were detected: "0", "0", "0", "0", "0", …, "0.93", "0.94", "0.95", "0.95", "0.96" (1277 values omitted).

## SOCFUNC

*Social Functioning*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 235 (0.28 %) |
| Number of unique values | 9 |
| Median | 100 |
| 1st and 3rd quartiles | 87.5; 100 |
| Min. and max. | 0; 100 |

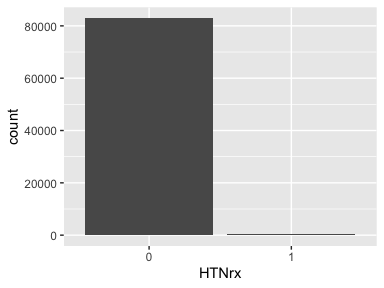


## HTNrx

*Anti-hypertensives*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 1 (0 %) |
| Number of unique values | 2 |
| Mode | “0” |
| Reference category | 0 |

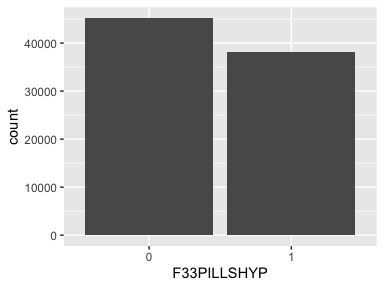


## F33PILLSHYP

*F33 Pills For Hypertension*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

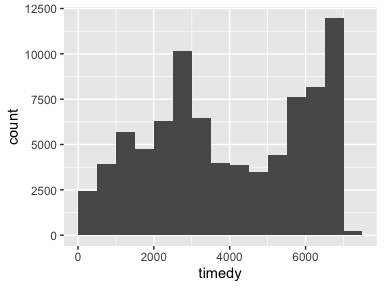
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 2 |
| Mode | “0” |
| Reference category | 0 |



## timedy

*Days to HTN or censor*

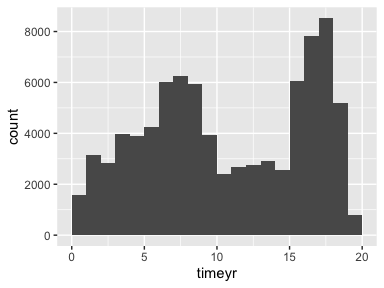
| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 6547 |
| Median | 3649 |
| 1st and 3rd quartiles | 2266; 5925 |
| Min. and max. | 123; 7295 |



## timeyr

*Years to HTN or censor*

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 6547 |
| Median | 9.99 |
| 1st and 3rd quartiles | 6.2; 16.22 |
| Min. and max. | 0.34; 19.97 |

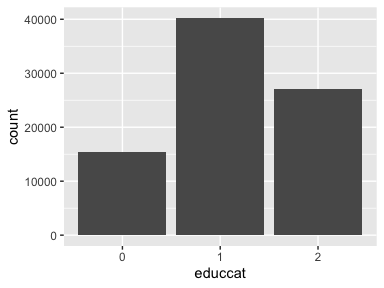


## educcat

*Education categories*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 615 (0.74 %) |
| Number of unique values | 3 |
| Mode | “1” |
| Reference category | 0 |



## HxCHD

*Hx CHD (MI or Revasc)*

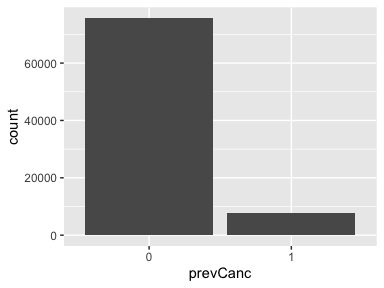
* The variable only takes one (non-missing) value: "0". The variable contains 0 % missing observations.

## prevCanc

*Hx Cancer*

* Note that this variable is treated as a factor variable below, as it only takes a few unique values.

| Feature | Result |
| --- | --- |
| Variable type | numeric |
| Number of missing obs. | 0 (0 %) |
| Number of unique values | 2 |
| Mode | “0” |
| Reference category | 0 |



## stroke\_f2

*Stroke ever*

* The variable only takes one (non-missing) value: "0". The variable contains 0 % missing observations.

Report generation information:

* Created by: Connor (username: Connor).
* Report creation time: Tue Nov 14 2023 14:55:19
* Report was run from directory: /Users/Connor/Library/CloudStorage/OneDrive-Personal/Desktop/WHI Total PA & HTN R/whi-totalPA-htn
* dataMaid v1.4.1 [Pkg: 2021-10-08 from CRAN (R 4.2.0)]
* R version 4.2.0 (2022-04-22).
* Platform: x86\_64-apple-darwin17.0 (64-bit)(macOS Monterey 12.4).
* Function call: makeDataReport(data = WHI.TotalPA.an\_samp, output = "html", replace = TRUE)