

miniNHANES

Autogenerated data summary from dataMaid

2020-12-02 18:04:33

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	10000
Number of variables	18

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.

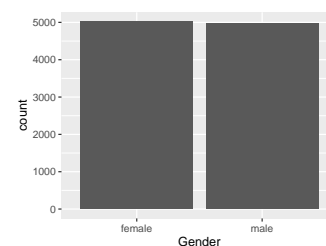
Summary table

	Variable class	# unique values	Missing observations	Any problems?
Gender	factor	2	0.00 %	
Age	integer	81	0.00 %	
Race1	factor	5	0.00 %	
Education	factor	6	27.79 %	
MaritalStatus	factor	7	27.69 %	
HHIncomeMid	integer	13	8.11 %	
Weight	numeric	1291	0.78 %	×
Height	numeric	981	3.53 %	×
BMI	numeric	2074	3.66 %	×
BPSys1	integer	71	17.63 %	×
BPDia1	integer	52	17.63 %	×
DirectChol	numeric	107	15.26 %	×
TotChol	numeric	251	15.26 %	×
Diabetes	factor	3	1.42 %	
SleepHrsNight	integer	12	22.45 %	×
PhysActiveDays	integer	8	53.37 %	
SmokeNow	factor	3	67.89 %	
HardDrugs	factor	3	42.35 %	

Variable list

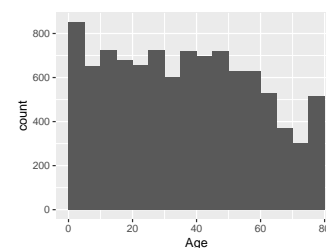
Gender

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



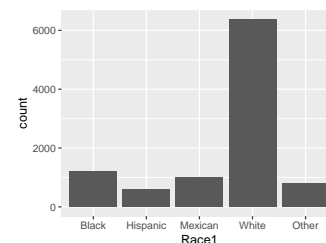
Age

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	81
Median	36
1st and 3rd quartiles	17; 54
Min. and max.	0; 80



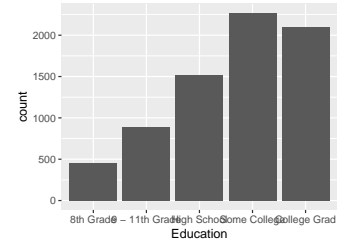
Race1

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	5
Mode	“White”
Reference category	Black



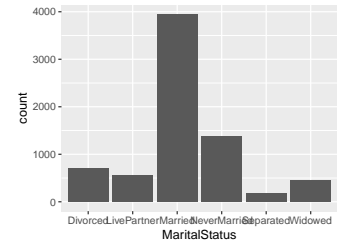
Education

Feature	Result
Variable type	factor
Number of missing obs.	2779 (27.79 %)
Number of unique values	5
Mode	"Some College"
Reference category	8th Grade



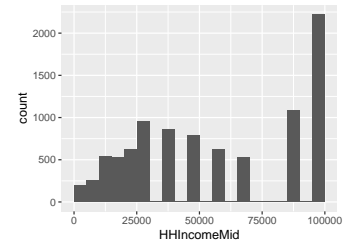
MaritalStatus

Feature	Result
Variable type	factor
Number of missing obs.	2769 (27.69 %)
Number of unique values	6
Mode	"Married"
Reference category	Divorced



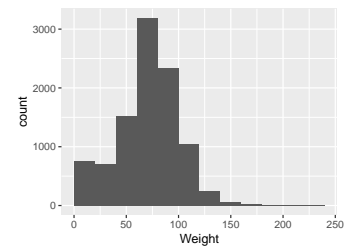
HHIncomeMid

Feature	Result
Variable type	integer
Number of missing obs.	811 (8.11 %)
Number of unique values	12
Median	50000
1st and 3rd quartiles	30000; 87500
Min. and max.	2500; 1e+05



Weight

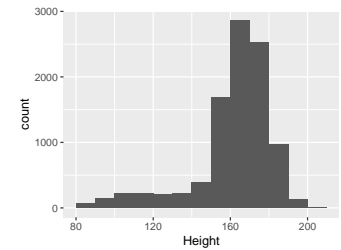
Feature	Result
Variable type	numeric
Number of missing obs.	78 (0.78 %)
Number of unique values	1290
Median	72.7
1st and 3rd quartiles	56.1; 88.9
Min. and max.	2.8; 230.7



- Note that the following possible outlier values were detected: "126.7", "127.3", "127.4", "127.7", "127.8", "127.9", "128", "128.2", "128.3", "128.4" (121 additional values omitted).

Height

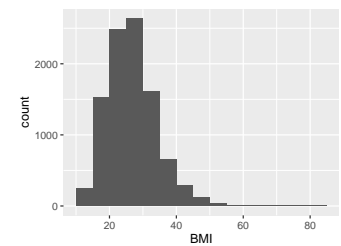
Feature	Result
Variable type	numeric
Number of missing obs.	353 (3.53 %)
Number of unique values	980
Median	166
1st and 3rd quartiles	156.8; 174.5
Min. and max.	83.6; 200.4



- Note that the following possible outlier values were detected: "83.6", "83.8", "84.5", "84.8", "84.9", "85", "85.1", "85.4", "85.5", "85.7" (286 additional values omitted).

BMI

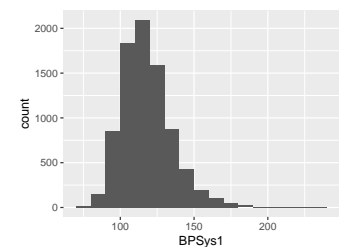
Feature	Result
Variable type	numeric
Number of missing obs.	366 (3.66 %)
Number of unique values	2073
Median	25.98
1st and 3rd quartiles	21.58; 30.89
Min. and max.	12.88; 81.25



- Note that the following possible outlier values were detected: "48.09", "48.12", "48.2", "48.22", "48.28", "48.3", "48.6", "48.68", "48.7", "48.8" (53 additional values omitted).

BPSys1

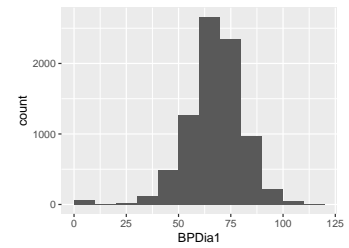
Feature	Result
Variable type	integer
Number of missing obs.	1763 (17.63 %)
Number of unique values	70
Median	116
1st and 3rd quartiles	106; 128
Min. and max.	72; 232



- Note that the following possible outlier values were detected: "72", "74", "76", "78", "80", "82", "84", "86", "180", "182" (14 additional values omitted).

BPDia1

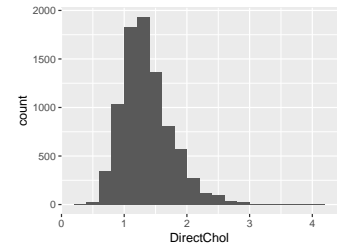
Feature	Result
Variable type	integer
Number of missing obs.	1763 (17.63 %)
Number of unique values	51
Median	70
1st and 3rd quartiles	62; 76
Min. and max.	0; 118



- Note that the following possible outlier values were detected: "0", "10", "16", "20", "22", "24", "26", "28", "30", "32" (13 additional values omitted).

DirectChol

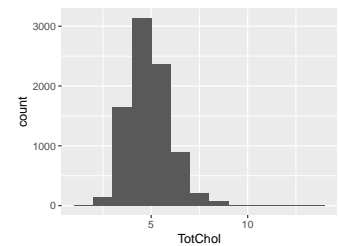
Feature	Result
Variable type	numeric
Number of missing obs.	1526 (15.26 %)
Number of unique values	106
Median	1.29
1st and 3rd quartiles	1.09; 1.58
Min. and max.	0.39; 4.03



- Note that the following possible outlier values were detected: "0.39", "0.41", "0.47", "0.49", "0.52", "0.54", "0.57", "0.59", "0.62", "0.65" (16 additional values omitted).

TotChol

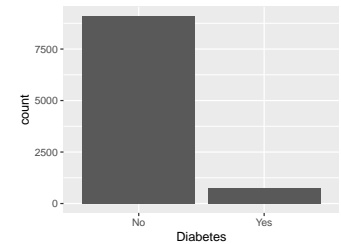
Feature	Result
Variable type	numeric
Number of missing obs.	1526 (15.26 %)
Number of unique values	250
Median	4.78
1st and 3rd quartiles	4.11; 5.53
Min. and max.	1.53; 13.65



- Note that the following possible outlier values were detected: "1.53", "2.17", "2.28", "2.3", "2.33", "2.35", "2.38", "2.4", "2.43", "2.46" (27 additional values omitted).

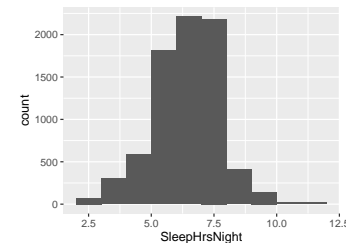
Diabetes

Feature	Result
Variable type	factor
Number of missing obs.	142 (1.42 %)
Number of unique values	2
Mode	"No"
Reference category	No



SleepHrsNight

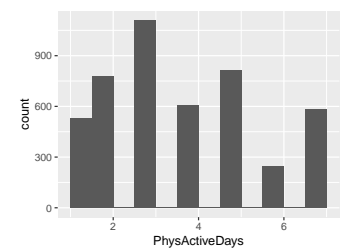
Feature	Result
Variable type	integer
Number of missing obs.	2245 (22.45 %)
Number of unique values	11
Median	7
1st and 3rd quartiles	6; 8
Min. and max.	2; 12



- Note that the following possible outlier values were detected: "2", "12".

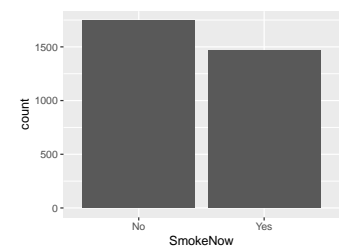
PhysActiveDays

Feature	Result
Variable type	integer
Number of missing obs.	5337 (53.37 %)
Number of unique values	7
Median	3
1st and 3rd quartiles	2; 5
Min. and max.	1; 7



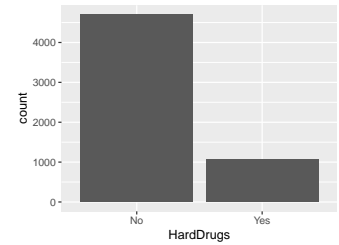
SmokeNow

Feature	Result
Variable type	factor
Number of missing obs.	6789 (67.89 %)
Number of unique values	2
Mode	"No"
Reference category	No



HardDrugs

Feature	Result
Variable type	factor
Number of missing obs.	4235 (42.35 %)
Number of unique values	2
Mode	"No"
Reference category	No



Report generation information:

- Created by: Sergio Uribe (username: `sergio`).
- Report creation time: Wed Dec 02 2020 18:04:34
- Report was run from directory: `/home/sergio/Dropbox/Public/Practico R/Introduction-to-Data-Science-RSU`
- dataMaid v1.4.0 [Pkg: 2019-12-10 from CRAN (R 4.0.2)]
- R version 4.0.2 (2020-06-22).
- Platform: x86_64-pc-linux-gnu (64-bit)(Ubuntu 20.10).
- Function call: `dataMaid::makeDataReport(data = miniNHANES)`