sd1

Autogenerated data summary from dataReporter

2021-10-14 09:50:59

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	2154
Number of variables	29

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.

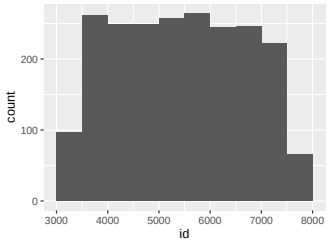
Summary table

	Variable class	# unique values	Missing observations	Any problems?
id	integer	2154	0.00 %	
exdate	Date	688	0.00 %	
age	integer	64	0.00 %	
sex	integer	2	0.00 %	
obs_bp	integer	11	0.00 %	×
obs_soma	integer	9	0.00 %	×
obs_int	integer	14	0.00 %	×
dev_bp	integer	12	0.00 %	×
dev_length	integer	4	0.00 %	×
dev_weight	integer	4	0.00 %	×
sbp1	integer	126	0.09 %	×
sbp2	integer	124	0.28 %	×
dbp1	integer	76	0.09 %	×
dbp2	integer	72	0.28 %	×
height	integer	53	0.14 %	×
weight	numeric	1764	0.19 %	×
waist	character	2106	0.14 %	×
cholesterol	numeric	1669	0.70 %	×
hdl	numeric	1112	0.74 %	×
ldl	numeric	2126	1.30 %	×
school	numeric	7	0.00 %	×
family	numeric	7	0.00 %	×
smoking	numeric	5	0.00 %	
myocard	numeric	5	0.00 %	
stroke	numeric	5	0.00 %	×
diab_known	integer	3	0.00 %	
diab_age	integer	49	0.00 %	×
contraception	numeric	4	0.00 %	×
income	integer	5	0.00 %	

Variable list

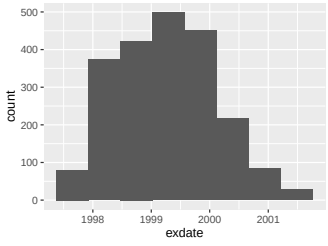
id

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	2154
Median	5428.5
1st and 3rd quartiles	4356; 6490.75
Min. and max.	3301; 7607



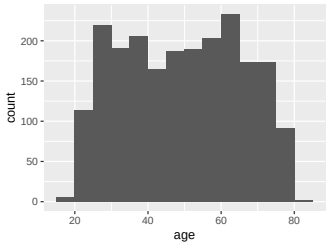
exdate

Feature	Result
Variable type	Date
Number of missing obs.	0 (0 %)
Number of unique values	688
Mode	“1999-03-26”
Min. and max.	1997-10-16; 2001-05-19
1st and 3rd quartiles	1998-08-18; 1999-11-04



age

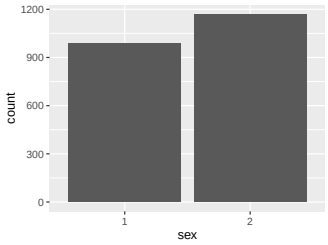
Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	64
Median	50
1st and 3rd quartiles	36; 63
Min. and max.	19; 82



sex

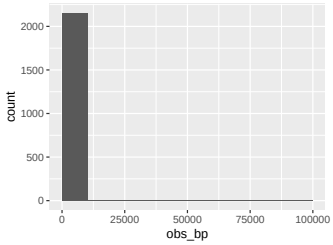
- Note that this variable is treated as a factor variable below, as it only takes a few unique values.

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



obs_bp

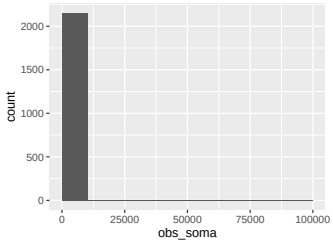
Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	11
Median	4
1st and 3rd quartiles	4; 7
Min. and max.	1; 99902



- Note that the following possible outlier values were detected: "1", "3", "99902".

obs_soma

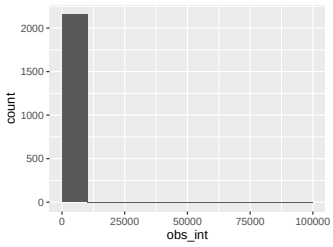
Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	9
Median	5
1st and 3rd quartiles	4; 7
Min. and max.	1; 99902



- Note that the following possible outlier values were detected: "1", "2", "99902".

obs_int

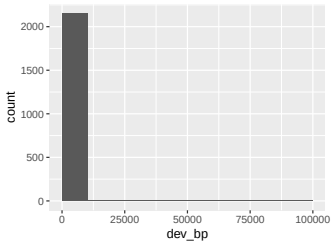
Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	14
Median	3
1st and 3rd quartiles	2; 11
Min. and max.	1; 99900



- Note that the following possible outlier values were detected: "1", "99900".

dev_bp

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	12
Median	15
1st and 3rd quartiles	10; 18
Min. and max.	7; 99902



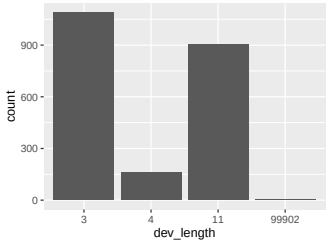
- Note that the following possible outlier values were detected: "99902".

---

dev_length

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

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	4
Mode	“3”
Reference category	3



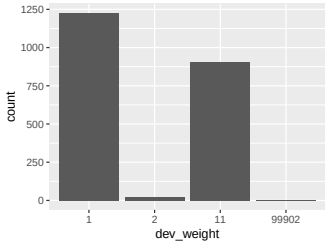
- Note that the following levels have at most five observations: "99902".

---

dev_weight

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

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

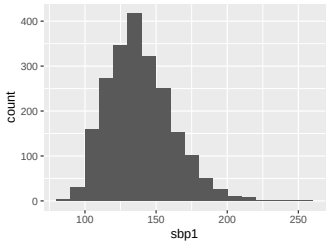


- Note that the following levels have at most five observations: "99902".

---

sbp1

Feature	Result
Variable type	integer
Number of missing obs.	2 (0.09 %)
Number of unique values	125
Median	137
1st and 3rd quartiles	123; 153
Min. and max.	83; 253

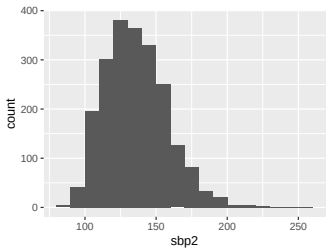


- Note that the following possible outlier values were detected: "83", "87", "88", "210", "211", . . . , "220", "222", "226", "238", "253" (2 values omitted).

---

sbp2

Feature	Result
Variable type	integer
Number of missing obs.	6 (0.28 %)
Number of unique values	123
Median	134
1st and 3rd quartiles	120; 150
Min. and max.	83; 258

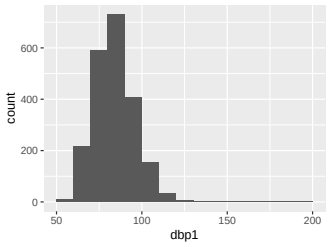


- Note that the following possible outlier values were detected: "83", "84", "87", "211", "212", . . . , "218", "228", "229", "230", "258" (2 values omitted).

---

dbp1

Feature	Result
Variable type	integer
Number of missing obs.	2 (0.09 %)
Number of unique values	75
Median	84
1st and 3rd quartiles	76; 92
Min. and max.	53; 198

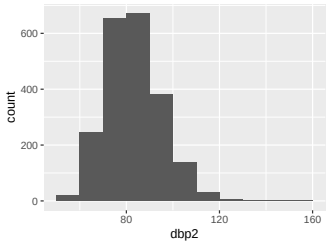


- Note that the following possible outlier values were detected: "117", "118", "119", "120", "121", . . . , "136", "141", "153", "164", "198" (3 values omitted).

---

dbp2

Feature	Result
Variable type	integer
Number of missing obs.	6 (0.28 %)
Number of unique values	71
Median	83
1st and 3rd quartiles	75; 91
Min. and max.	55; 151

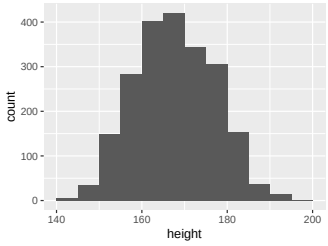


- Note that the following possible outlier values were detected: "120", "121", "122", "123", "130", "132", "151".

---

height

Feature	Result
Variable type	integer
Number of missing obs.	3 (0.14 %)
Number of unique values	52
Median	168
1st and 3rd quartiles	161; 175
Min. and max.	144; 198

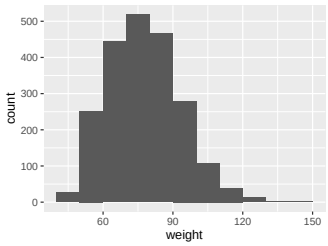


- Note that the following possible outlier values were detected: "198".

---

weight

Feature	Result
Variable type	numeric
Number of missing obs.	4 (0.19 %)
Number of unique values	1763
Median	77.04
1st and 3rd quartiles	66.16; 87.35
Min. and max.	42.6; 144.44

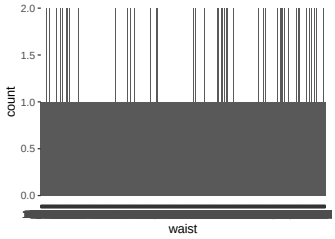


- Note that the following possible outlier values were detected: "120.52", "120.61", "120.8", "122.01", "122.07", . . . , "125.8", "126.81", "127.25", "135.48", "144.44" (5 values omitted).

---

waist

Feature	Result
Variable type	character
Number of missing obs.	3 (0.14 %)
Number of unique values	2105
Mode	“101.099”

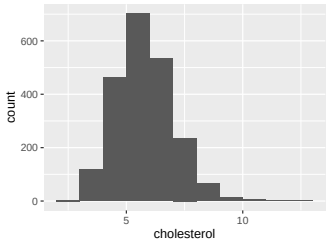


- Note that the following levels have at most five observations: "100.018", "100.027", "100.068", "100.069", "100.088", ..., "99.948", "99.956", "99.959", "99.966", "99.98" (2095 values omitted).

---

cholesterol

Feature	Result
Variable type	numeric
Number of missing obs.	15 (0.7 %)
Number of unique values	1668
Median	5.68
1st and 3rd quartiles	4.92; 6.49
Min. and max.	2.67; 12.12

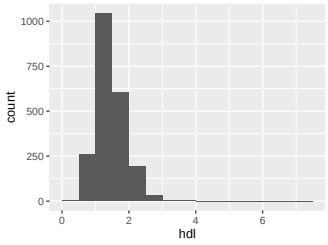


- Note that the following possible outlier values were detected: "2.67", "2.77", "2.9", "2.96", "9.3", ..., "10.11", "10.15", "10.75", "11.02", "12.12" (7 values omitted).

---

hdl

Feature	Result
Variable type	numeric
Number of missing obs.	16 (0.74 %)
Number of unique values	1111
Median	1.39
1st and 3rd quartiles	1.14; 1.7
Min. and max.	0.42; 7.2

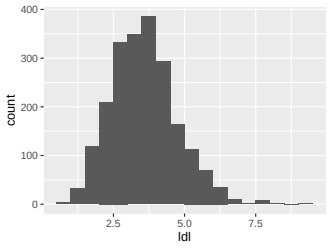


- Note that the following possible outlier values were detected: "0.42", "0.48", "0.5", "0.51", "0.64", ..., "3.24", "3.51", "3.86", "4.39", "7.2" (8 values omitted).

---

ldl

Feature	Result
Variable type	numeric
Number of missing obs.	28 (1.3 %)
Number of unique values	2125
Median	3.52
1st and 3rd quartiles	2.75; 4.24
Min. and max.	0.7; 9.24

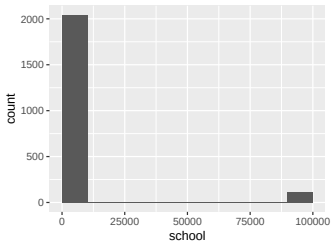


- Note that the following possible outlier values were detected: "6.69", "6.71", "6.76", "6.89", "6.95", ..., "7.82", "7.93", "8.35", "8.46", "9.24" (8 values omitted).

---

school

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	7
Median	2
1st and 3rd quartiles	0; 2
Min. and max.	0; 99914

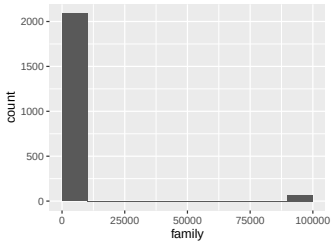


- Note that the following possible outlier values were detected: "3", "4", "99900", "99914".

---

family

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	7
Median	1
1st and 3rd quartiles	1; 3
Min. and max.	1; 99914



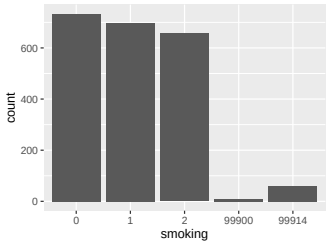
- Note that the following possible outlier values were detected: "99900", "99914".

---

smoking

- 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

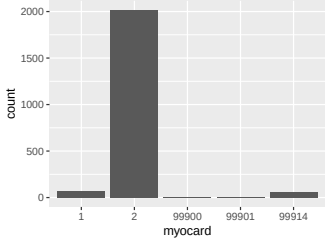


---

myocard

- 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	“2”
Reference category	1

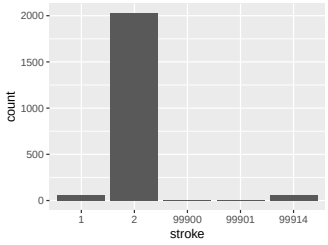


---

stroke

- 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	“2”
Reference category	1



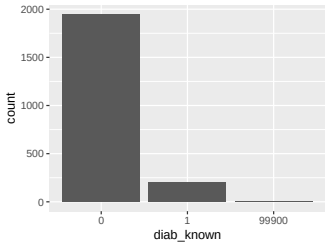
- Note that the following levels have at most five observations: "99901".

---

diab_known

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

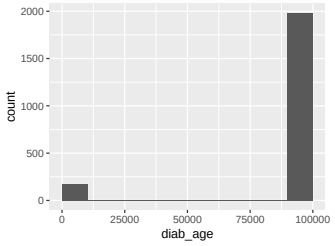
Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	3
Mode	“0”
Reference category	0



---

diab_age

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	49
Median	99801
1st and 3rd quartiles	99801; 99801
Min. and max.	12; 99900



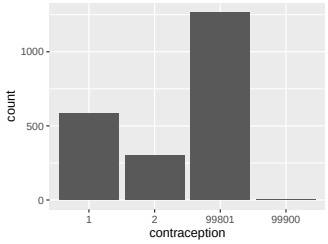
- Note that the following possible outlier values were detected: "12", "15", "18", "27", "29", . . . , "74", "75", "76", "79", "99900" (38 values omitted).

---

contraception

- 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	“99801”
Reference category	1



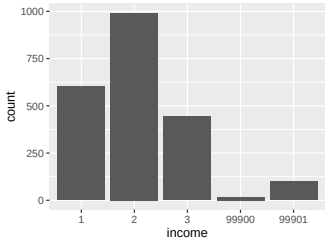
- Note that the following levels have at most five observations: "99900".

---

income

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

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	5
Mode	“2”
Reference category	1



Report generation information:

- Created by: Could not determine from system (username: Unknown)
- Report creation time: Thu Oct 14 2021 09:50:59
- Report was run from directory: `C:/Users/marinoj/Desktop/DQToolsReview`
- dataReporter v1.0.0 [Pkg: 2020-12-14 from CRAN (R 4.1.1)]
- R version 4.1.1 (2021-08-10).
- Platform: x86_64-w64-mingw32/x64 (64-bit)(Windows 10 x64 (build 17763)).
- Function call: `makeDataReport(data = sd1)`