

heartfailure

Autogenerated data summary from dataMaid

2021-08-22 03:09:49

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	299
Number of variables	13

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.

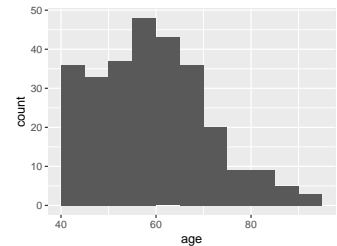
Summary table

	Variable class	# unique values	Missing observations	Any problems?
age	integer	47	6.69 %	
anaemia	factor	2	0.00 %	
cpk_enzyme	numeric	208	0.00 %	×
diabetes	factor	2	0.00 %	
ejection_fraction	numeric	17	0.00 %	×
hblood_pressure	factor	2	0.00 %	
platelets	numeric	176	0.00 %	×
creatinine	numeric	40	0.00 %	×
sodium	numeric	27	0.00 %	×
sex	factor	3	3.34 %	
smoking	factor	2	0.00 %	
time	integer	148	0.00 %	
death_event	factor	2	0.00 %	

Variable list

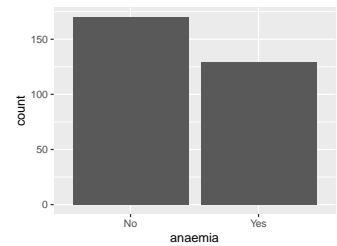
age

Feature	Result
Variable type	integer
Number of missing obs.	20 (6.69 %)
Number of unique values	46
Median	60
1st and 3rd quartiles	51; 70
Min. and max.	40; 95



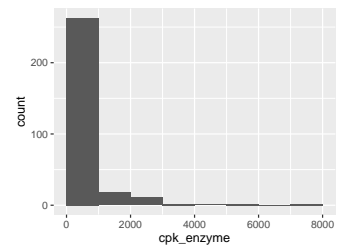
anaemia

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"No"
Reference category	No



cpk_enzyme

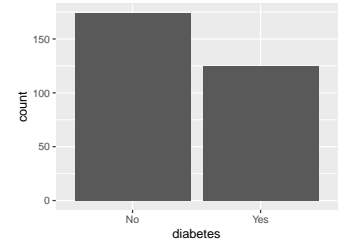
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	208
Median	250
1st and 3rd quartiles	116.5; 582
Min. and max.	23; 7861



- Note that the following possible outlier values were detected: "23", "30", "4540", "5209", "5882", "7702", "7861".

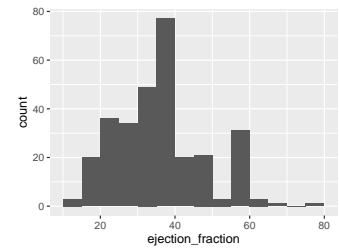
diabetes

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"No"
Reference category	No



ejection_fraction

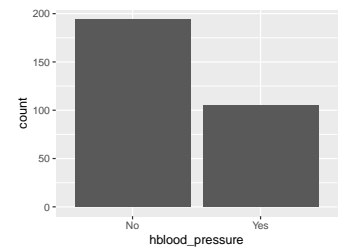
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	17
Median	38
1st and 3rd quartiles	30; 45
Min. and max.	14; 80



- Note that the following possible outlier values were detected: "65", "70", "80".

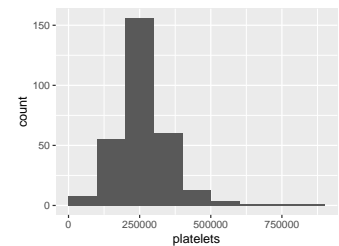
hblood_pressure

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"No"
Reference category	No



platelets

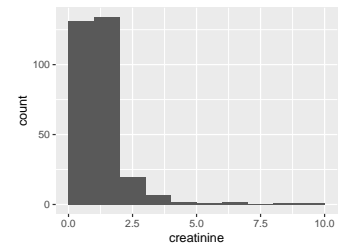
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	176
Median	262000
1st and 3rd quartiles	212500; 303500
Min. and max.	25100; 850000



- Note that the following possible outlier values were detected: "374000", "377000", "382000", "385000", "388000", "389000", "390000", "395000", "404000", "406000" (16 additional values omitted).

creatinine

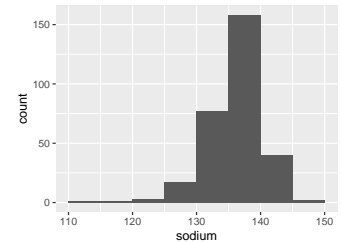
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	40
Median	1.1
1st and 3rd quartiles	0.9; 1.4
Min. and max.	0.5; 9.4



- Note that the following possible outlier values were detected: "0.5", "0.6", "0.7", "3.5", "3.7", "3.8", "4", "4.4", "5", "5.8" (4 additional values omitted).

sodium

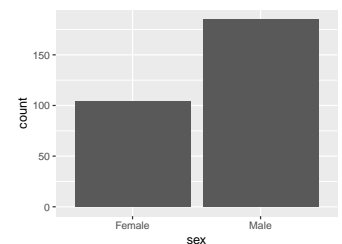
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	27
Median	137
1st and 3rd quartiles	134; 140
Min. and max.	113; 148



- Note that the following possible outlier values were detected: "113", "116", "121", "124".

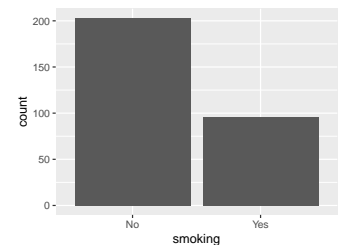
sex

Feature	Result
Variable type	factor
Number of missing obs.	10 (3.34 %)
Number of unique values	2
Mode	"Male"
Reference category	Female



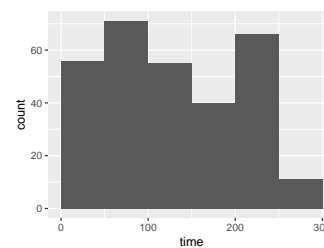
smoking

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"No"
Reference category	No



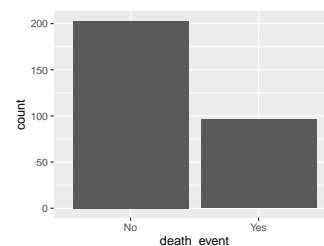
time

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	148
Median	115
1st and 3rd quartiles	73; 203
Min. and max.	4; 285



death_event

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"No"
Reference category	No



Report generation information:

- Created by: tengku-hanis (username: `tengk`).
- Report creation time: Sun Aug 22 2021 03:09:50
- Report was run from directory: `C:/Tengku/Sync_PC_Laptop/github/mywebsite/draft_blog/data exploration`
- dataMaid v1.4.0 [Pkg: 2019-12-10 from CRAN (R 4.1.0)]
- R version 4.1.1 (2021-08-10).
- Platform: x86_64-w64-mingw32/x64 (64-bit)(Windows 10 x64 (build 19042)).
- Function call: `makeDataReport(data = heartfailure, replace = T)`