

Heart Disease Codebook

Autogenerated data summary from dataMaid

2022-12-15 12:35:15

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	4220
Number of variables	21

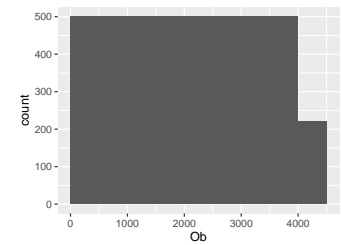
Codebook summary table

Label	Variable	Class	# unique values	Missing	Description
	Ob	integer	4220	0.00 %	
	Sex	character	5	0.00 %	
	Age	numeric	103	0.00 %	
	ChestPainType	character	4	0.00 %	
	RestingBP	integer	67	0.00 %	
	Cholesterol	integer	221	0.00 %	
	FastingBS	character	2	0.00 %	
	RestingECG	character	3	0.00 %	
	MaxHR	integer	119	0.00 %	
	ExerciseAngina	character	2	0.00 %	
	Oldpeak	numeric	53	0.00 %	
	ST_Slope	character	3	0.00 %	
	hypertension	character	2	0.00 %	
	ever_married	character	3	14.95 %	
	work_type	character	6	14.95 %	
	Residence_type	character	3	14.95 %	
	avg_glucose_level	numeric	3127	0.00 %	
	bmi	numeric	394	0.00 %	
	smoking_status	character	5	14.95 %	
	stroke	character	2	0.00 %	
	HeartDisease	character	2	0.00 %	

Variable list

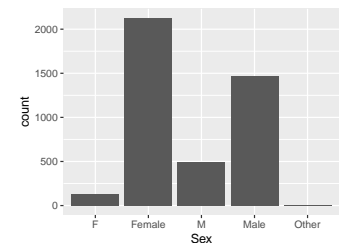
Ob

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	4220
Median	2110.5
1st and 3rd quartiles	1055.75; 3165.25
Min. and max.	1; 4220



Sex

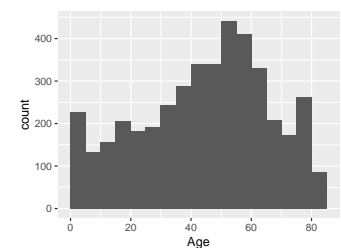
Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	5
Mode	"Female"



- Observed factor levels: "F", "Female", "M", "Male", "Other".

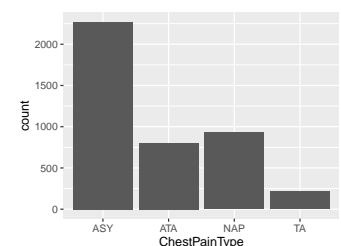
Age

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	103
Median	48
1st and 3rd quartiles	29; 61
Min. and max.	0.16; 82



ChestPainType

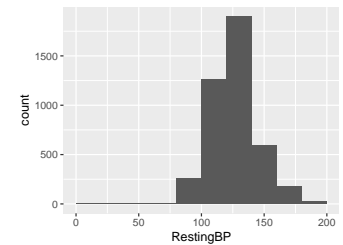
Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	4
Mode	"ASY"



- Observed factor levels: "ASY", "ATA", "NAP", "TA".

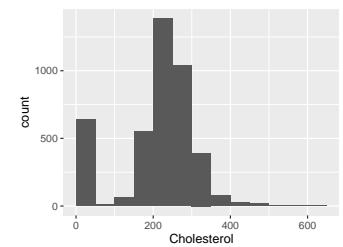
RestingBP

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	67
Median	130
1st and 3rd quartiles	120; 140
Min. and max.	0; 200



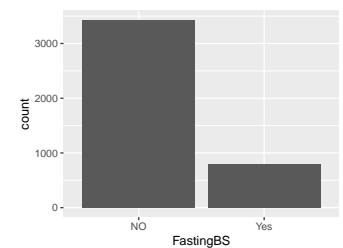
Cholesterol

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	221
Median	230
1st and 3rd quartiles	186; 273
Min. and max.	0; 603



FastingBS

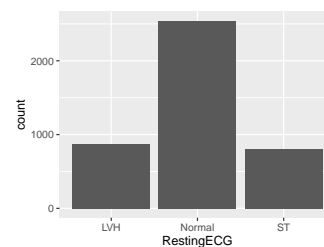
Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"NO"



- Observed factor levels: "NO", "Yes".

RestingECG

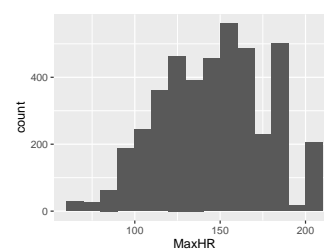
Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	3
Mode	"Normal"



- Observed factor levels: "LVH", "Normal", "ST".

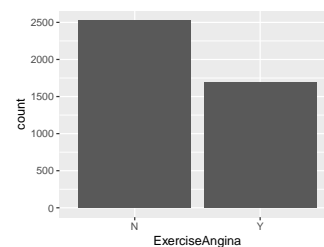
MaxHR

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	119
Median	150
1st and 3rd quartiles	125; 170
Min. and max.	60; 202



ExerciseAngina

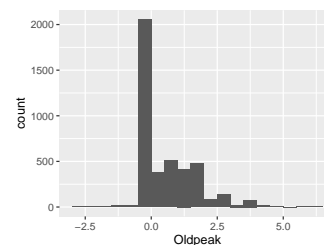
Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"N"



- Observed factor levels: "N", "Y".

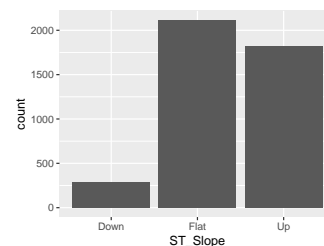
Oldpeak

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	53
Median	0.1
1st and 3rd quartiles	0; 1.4
Min. and max.	-2.6; 6.2



ST_Slope

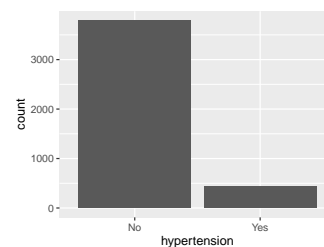
Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	3
Mode	"Flat"



- Observed factor levels: "Down", "Flat", "Up".

hypertension

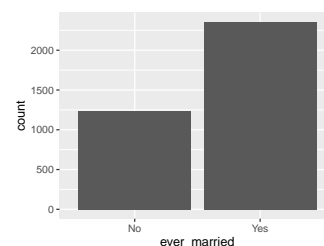
Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"No"



- Observed factor levels: "No", "Yes".

ever_married

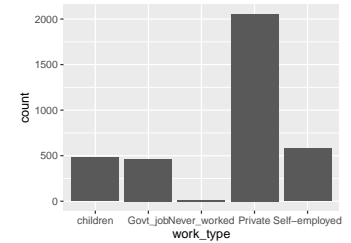
Feature	Result
Variable type	character
Number of missing obs.	631 (14.95 %)
Number of unique values	2
Mode	"Yes"



- Observed factor levels: "No", "Yes".

work_type

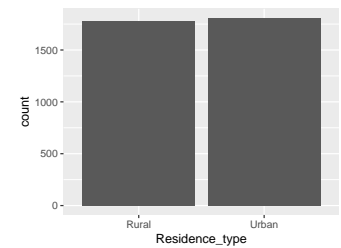
Feature	Result
Variable type	character
Number of missing obs.	631 (14.95 %)
Number of unique values	5
Mode	"Private"



- Observed factor levels: "children", "Govt_job", "Never_worked", "Private", "Self-employed".

Residence_type

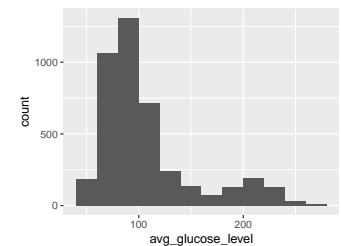
Feature	Result
Variable type	character
Number of missing obs.	631 (14.95 %)
Number of unique values	2
Mode	"Urban"



- Observed factor levels: "Rural", "Urban".

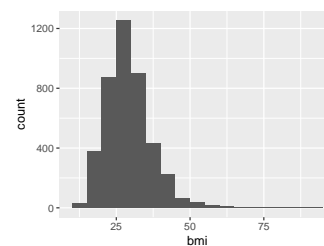
avg_glucose_level

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	3127
Median	92.42
1st and 3rd quartiles	77.48; 115.14
Min. and max.	55.12; 271.74



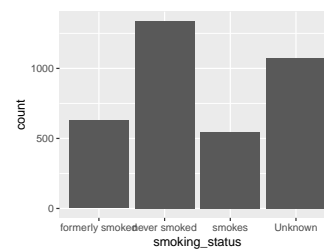
bmi

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	394
Median	28.4
1st and 3rd quartiles	24; 33.1
Min. and max.	10.3; 92



smoking_status

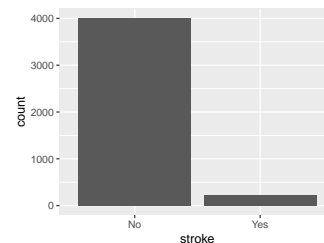
Feature	Result
Variable type	character
Number of missing obs.	631 (14.95 %)
Number of unique values	4
Mode	"never smoked"



- Observed factor levels: "formerly smoked", "never smoked", "smokes", "Unknown".

stroke

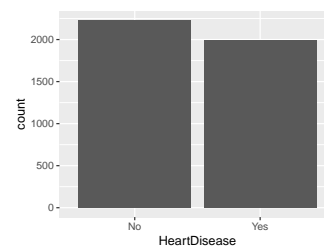
Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"No"



- Observed factor levels: "No", "Yes".

HeartDisease

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



- Observed factor levels: "No", "Yes".

Report generation information:

- Created by: Could not determine from system (username: Unknown)
- Report creation time: Thu Dec 15 2022 12:35:16
- Report was run from directory: C:/Users/domin/OneDrive/Escritorio/Stats Major Classes/Stats 101C
- dataMaid v1.4.1 [Pkg: 2021-10-08 from CRAN (R 4.1.3)]
- R version 4.1.2 (2021-11-01).
- Platform: x86_64-w64-mingw32/x64 (64-bit)(Windows 10 x64 (build 22000)).
- Function call: `dataMaid::makeDataReport(data = HD.train, mode = c("summarize", "visualize", "check"), smartNum = FALSE, file = "Heart_Disease", vol = "Heart_Disease", checks = list(character = "showAllFactorLevels", factor = "showAllFactorLevels", labelled = "showAllFactorLevels", haven_labelled = "showAllFactorLevels", numeric = NULL, integer = NULL, logical = NULL, Date = NULL), listChecks = FALSE, maxProbVals = Inf, codebook = TRUE, reportTitle = "Heart Disease Codebook")`