

recid

Autogenerated data summary from dataMaid

2019-03-10 18:48:21

Part 1

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	6537
Number of variables	12

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.

Part 2

Summary table

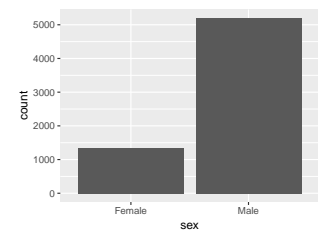
	Variable class	# unique values	Missing observations	Any problems?
sex	character	2	0.00 %	
age	numeric	65	0.00 %	×
race	factor	6	0.00 %	
juv_fel_count	numeric	11	0.00 %	×
juv_misd_count	numeric	10	0.00 %	×
juv_other_count	numeric	11	0.00 %	×
priors_count	numeric	33	0.00 %	
c_charge_degree	character	12	0.00 %	×
c_charge_desc	character	391	0.06 %	×
is_recid	factor	2	0.00 %	
jail_days	numeric	384	0.00 %	×
age_at_offense	numeric	4523	12.99 %	×

Part 3

Variable list

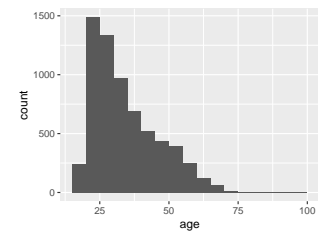
sex

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	“Male”



age

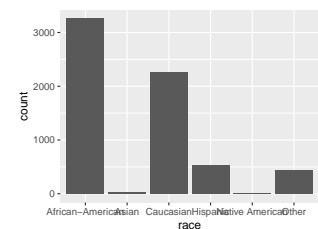
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	65
Median	31
1st and 3rd quartiles	25; 42
Min. and max.	18; 96



- Note that the following possible outlier values were detected: "18".

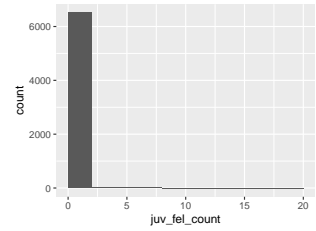
race

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	6
Mode	“African-American”
Reference category	African-American



juv__fel__count

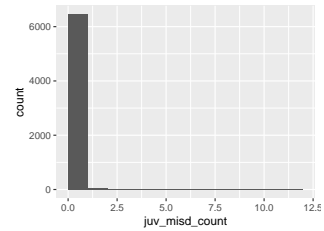
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	11
Median	0
1st and 3rd quartiles	0; 0
Min. and max.	0; 20



- Note that the following possible outlier values were detected: "1", "2", "3", "4", "5", "6", "8", "10", "13", "20".

juv__misd__count

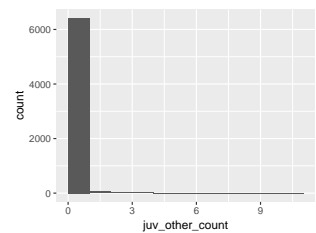
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	10
Median	0
1st and 3rd quartiles	0; 0
Min. and max.	0; 12



- Note that the following possible outlier values were detected: "1", "2", "3", "4", "5", "6", "7", "8", "12".

juv__other__count

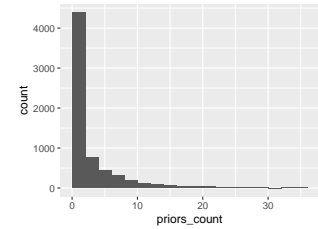
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	11
Median	0
1st and 3rd quartiles	0; 0
Min. and max.	0; 11



- Note that the following possible outlier values were detected: "1", "2", "3", "4", "5", "6", "7", "9", "10", "11".

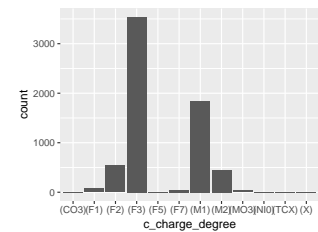
priors_count

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	33
Median	1
1st and 3rd quartiles	0; 4
Min. and max.	0; 36



c_charge_degree

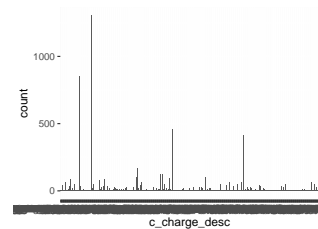
Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	12
Mode	“(F3)”



- Note that the following levels have at most five observations: "(CO3)", "(F5)", "(NI0)", "(TCX)", "(X)".

c_charge_desc

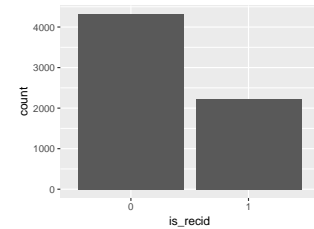
Feature	Result
Variable type	character
Number of missing obs.	4 (0.06 %)
Number of unique values	390
Mode	“Battery”



- Note that the following levels have at most five observations: "Abuse Without Great Harm", "Agg Battery Cause Bodily", "Agg Fleeing and Eluding", "Agg Fleeing/Eluding High Speed", "Aggr Child Abuse-Torture,Punish", "Aggrav Child Abuse-Agg Battery", "Aggrav Child Abuse-Causes Harm", "Aggravated Animal Cruelty", "Aggravated Assault W/o Firearm", "Aggravated Battery (Firearm)" (263 additional values omitted).
- Note that there might be case problems with the following levels: "Aggravated Assault W/dead Weap", "Aggravated Assault W/Dead Weap", "Possession of Alcohol Under 21", "Possession Of Alcohol Under 21", "Possession of Cocaine", "Possession Of Cocaine", "Tampering with a Victim", "Tampering With A Victim".

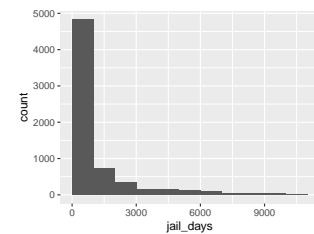
is_recid

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



jail_days

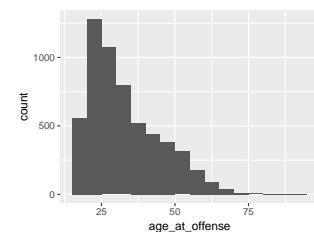
Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	384
Median	366
1st and 3rd quartiles	365; 1096
Min. and max.	29; 10592



- Note that the following possible outlier values were detected: "29", "30", "31", "59", "60", "61", "90", "91", "92", "122" (9 additional values omitted).

age_at_offense

Feature	Result
Variable type	numeric
Number of missing obs.	849 (12.99 %)
Number of unique values	4522
Median	29.68
1st and 3rd quartiles	23.29; 40.47
Min. and max.	15.04; 94.52



- Note that the following possible outlier values were detected: "15.04".

Report generation information:

- Created by Mateusz Staniak (username: `mstaniak`).
- Report creation time: niedz. mar 10 2019 18:48:23
- Report was run from directory: `C:/Users/mstaniak/Projekty/MI2DataLab/autoEDA-resources/usecase`
- dataMaid v1.2.0 [Pkg: 2018-10-03 from CRAN (R 3.5.2)]
- R version 3.5.1 (2018-07-02).
- Platform: `x86_64-w64-mingw32/x64 (64-bit)(Windows >= 8 x64 (build 9200))`.

- Function call: `makeDataReport(data = recid)`