dane

Autogenerated data summary from dataMaid

2018-06-13 11:30:50

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	14999
Number of variables	10

Checks performed

The following variable checks were performed, depending on the data type of each variable:

	character	factor	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.

Codebook summary table

Label	Variable	Class	# unique values	Missing	Description
	satisfaction_level	numeric	92	0.00 %	
	last_evaluation	numeric	65	0.00 %	
	number_project	integer	6	0.00 %	
	average_montly_hours	integer	215	0.00 %	
	time_spend_company	integer	8	0.00 %	
	Work_accident	integer	2	0.00 %	
	left	factor	2	0.00 %	
	promotion_last_5years	integer	2	0.00 %	
	sales	factor	10	0.00 %	
	salary	factor	3	0.00 %	

Variable list

satisfaction_level

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	92
Median	0.64
1st and 3rd quartiles	0.44; 0.82
Min. and max.	0.09; 1
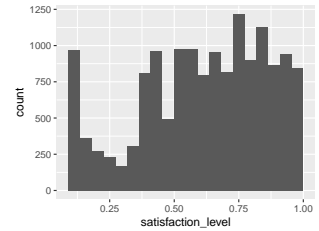
last_evaluation

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	65
Median	0.72
1st and 3rd quartiles	0.56; 0.87
Min. and max.	0.36; 1
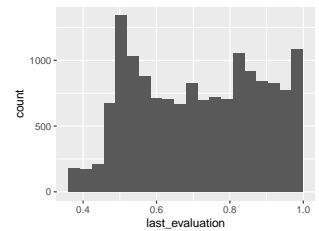
number_project

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	6
Median	4
1st and 3rd quartiles	3; 5
Min. and max.	2; 7
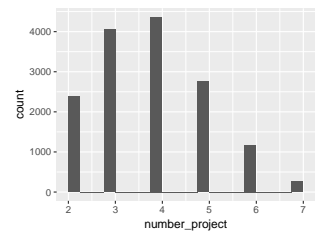
average_montly_hours

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	215
Median	200
1st and 3rd quartiles	156; 245
Min. and max.	96; 310
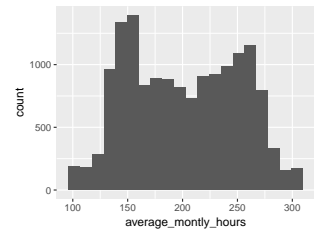
time_spend_company

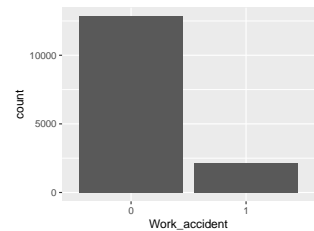| Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	8
Median	3
1st and 3rd quartiles	3; 4
Min. and max.	2; 10
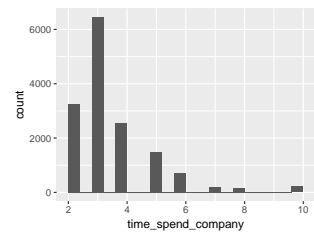
- Note that the following possible outlier values were detected: "2".
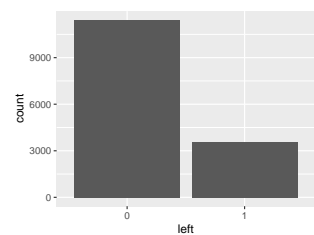
Work_accident

- 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	"0"

left

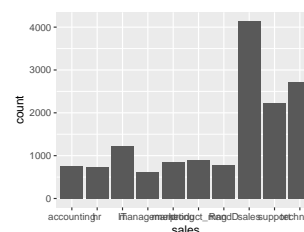| Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"0"

3

promotion_last_5years

- 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	"0"
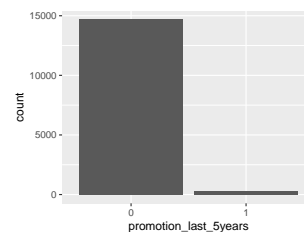
sales

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	10
Mode	"sales"

salary

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	3
Mode	"low"
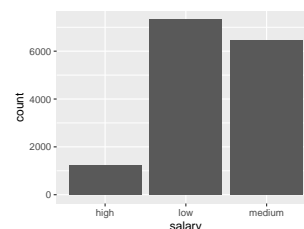
Report generation information:

- Created by chudekm (username: chudekm).

- Report creation time: Śr cze 13 2018 11:30:52

- Report Was run from directory: H:/Windows7/Documents

- dataMaid v1.1.2 [Pkg: 2018-05-03 from CRAN (R 3.4.4)]

- R version 3.4.1 (2017-06-30).

- Platform: x86_64-w64-mingw32/x64 (64-bit)(Windows 7 x64 (build 7601) Service Pack 1).

- Function call: makeDataReport(data = data, output = "pdf", smartNum = TRUE, file = "H:\\\\Windows7\\\\Documents\\\\hugo_investigation_2018_06_13\\\\gallery/hugo_summary_dane.Rmd", replace = FALSE, quiet = "silent", openResult = TRUE, maxDecimals = 2, codebook = TRUE, reportTitle = "dane")