

toyData

Autogenerated data summary from dataMaid

2020-03-31 13:32:42

Part 1

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	15
Number of variables	6

- Only variables that were deemed potentially problematic are included in this summary.

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?
pill	factor	3	13.33 %	×
events	numeric	9	20.00 %	×
region	factor	7	0.00 %	×
change	numeric	15	0.00 %	×
id	factor	15	0.00 %	×
spotifysong	factor	1	0.00 %	×

Part 3

Variable list

pill

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

events

- The following suspected missing value codes enter as regular values: "999", "NaN".
 - Note that the following possible outlier values were detected: "82", "999".
-

region

- The following suspected missing value codes enter as regular values: " ", ".".
 - The following values appear with prefixed or suffixed white space: " ".
 - Note that the following levels have at most five observations: " ", ".", "a", "b", "c", "other", "OTHER".
 - Note that there might be case problems with the following levels: "other", "OTHER".
-

change

- Note that the following possible outlier values were detected: "1.12", "1.51", "1.6".
-

id

- The variable is a key (distinct values for each observation).
-

spotifysong

- The variable only takes one (non-missing) value: "Irrelevant". The variable contains 0 % missing observations.
-

Report generation information:

- Created by: Piotr Piątyszek (username: `piotrp`).
- Report creation time: Tue Mar 31 2020 13:32:42
- Report was run from directory: `/home/piotrp/projects/warszt/pd1/pd1-pp`
- dataMaid v1.4.0 [Pkg: 2019-12-10 from CRAN (R 3.6.2)]
- R version 3.6.2 (2019-12-12).
- Platform: x86_64-suse-linux-gnu (64-bit)(openSUSE Leap 15.0).
- Function call: `makeDataReport(data = toyData, onlyProblematic = TRUE, mode = "check", replace = TRUE)`