toyData

Autogenerated data summary from dataMaid

2020-03-31 13:32:37

Part 1

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	15
Number of variables	6

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.
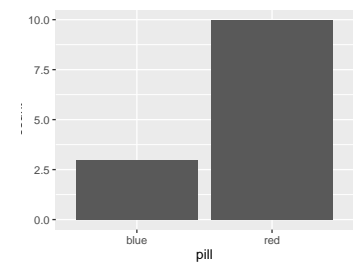
Part 2

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

Feature	Result
Variable type	factor
Number of missing obs.	2 (13.33 %)
Number of unique values	2
Mode	"red"
Reference category	blue

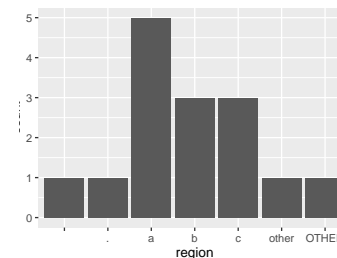- Note that the following levels have at most five observations: "blue".

events

Feature	Result
Variable type	numeric
Number of missing obs.	3 (20 %)
Number of unique values	8
Median	4.5
1st and 3rd quartiles	1.75; 6
Min. and max.	1; 999
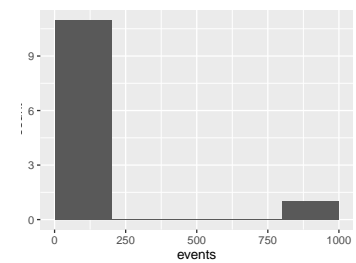
- The following suspected missing value codes enter as regular values: "999", "NaN".

- Note that the following possible outlier values were detected: "82", "999".

region

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	7
Mode	"a"
Reference category	

- 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

Feature	Result
Variable type	numeric
Number of missing obs.	0 (0 %)
Number of unique values	15
Median	0.33
1st and 3rd quartiles	-0.62; 0.66
Min. and max.	-2.21; 1.6
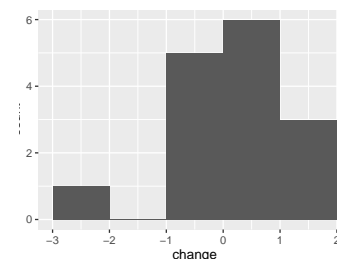
- Note that the following possible outlier values were detected: "1.12", "1.51", "1.6".

---

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 13:32:38

- 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, vol = paste("_", format(Sys.time(),  "%m-%d%-%y_%H.%M"), sep = ""), quiet = "silent", openResult = FALSE)`