final

Autogenerated data summary from dataMaid

2018-10-01 17:11:30

Part 1

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	26550
Number of variables	5

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.
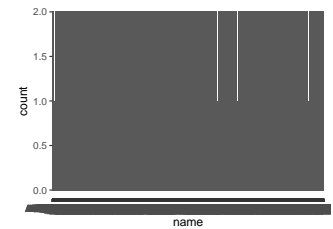
Part 2

Summary table

	Variable class	# unique values	Missing observations	Any problems?
name	character	24553	0.00 %	×
gender	factor	2	0.00 %	
count2015	integer	1397	0.00 %	×
count2016	integer	1385	0.00 %	×
Total	integer	1884	0.00 %	×

Part 3

Variable list

name

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	24553
Mode	"Aalijah"

- Note that the following levels have at most five observations: "Aaban", "Aabha", "Aabriella", "Aadam", "Aadarsh", "Aaden", "Aadhav", "Aadhavan", "Aadhi", "Aadhira" (24543 additional values omitted).
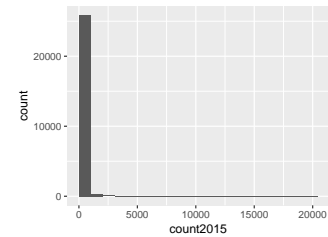
gender

Feature	Result
Variable type	factor
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"F"
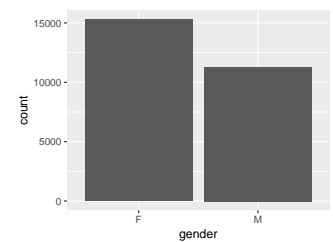
count2015

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	1397
Median	16
1st and 3rd quartiles	8; 41
Min. and max.	5; 20415

- Note that the following possible outlier values were detected: "375", "376", "377", "378", "379", "380", "381", "382", "383", "385" (1019 additional values omitted).

count2016

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	1385
Median	15
1st and 3rd quartiles	8; 41
Min. and max.	5; 19414
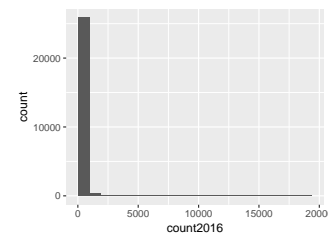
- Note that the following possible outlier values were detected: "386", "387", "388", "389", "390", "391", "392", "394", "395", "396" (997 additional values omitted).

Total

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	1884
Median	31
1st and 3rd quartiles	17; 81
Min. and max.	10; 39829
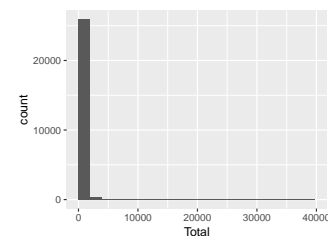
- Note that the following possible outlier values were detected: "10", "768", "769", "771", "772", "773", "774", "775", "777", "779" (1163 additional values omitted).

- Report Was run from directory: `/Users/temp/dev/SMU/MSDS6306/assignments`

- dataMaid v1.1.2 [Pkg: 2018-05-03 from CRAN (R 3.5.0)]

- R version 3.5.1 (2018-07-02).

- Platform: x86_64-apple-darwin15.6.0 (64-bit)(macOS High Sierra 10.13.6).

- Function call: `makeDataReport(data = final, replace = TRUE)`