sal

Autogenerated data summary from dataMaid

2022-04-04 11:00:09

Data report overview

The dataset examined has the following dimensions:

Feature	Result
Number of observations	397
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.

Summary table

	Variable class	# unique values	Missing observations	Any problems?
rank	character	3	0.00 %	
discipline	character	2	0.00 %	
yrs.since.phd	integer	54	0.25 %	
yrs.service	integer	53	0.25 %	
sex	character	2	0.00 %	
salary	integer	371	0.00 %	×

Variable list

rank

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	3
Mode	"Prof"
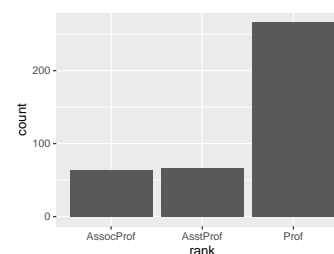
discipline

Feature	Result
Variable type	character
Number of missing obs.	0 (0 %)
Number of unique values	2
Mode	"B"
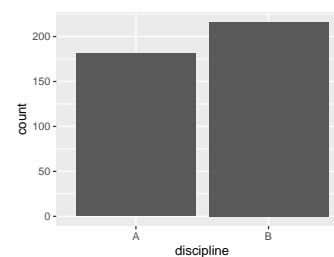
yrs.since.phd

Feature	Result
Variable type	integer
Number of missing obs.	1 (0.25 %)
Number of unique values	53
Median	21
1st and 3rd quartiles	12; 32
Min. and max.	1; 56
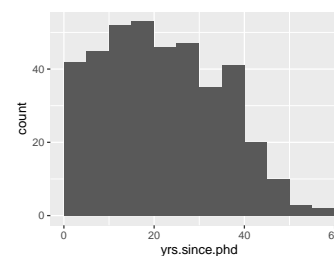
yrs.service

Feature	Result
Variable type	integer
Number of missing obs.	1 (0.25 %)
Number of unique values	52
Median	16
1st and 3rd quartiles	7; 26.25
Min. and max.	0; 60
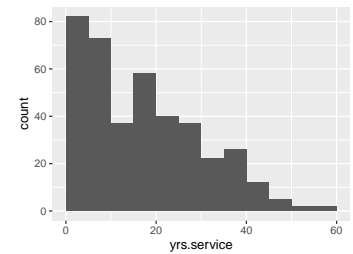
sex

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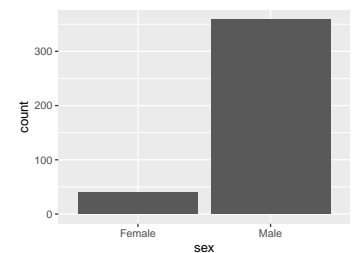
salary

Feature	Result
Variable type	integer
Number of missing obs.	0 (0 %)
Number of unique values	371
Median	107300
1st and 3rd quartiles	91000; 134185
Min. and max.	57800; 231545
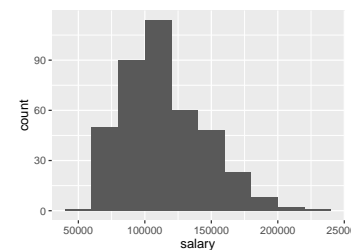
- Note that the following possible outlier values were detected: "57800", "62884", "63100", "63900".

Report generation information:

- Created by: Miguel Portela (username: `miguelportela`).

- Report creation time: Seg Abr 04 2022 11:00:09

- Report was run from directory: `/Users/miguelportela/Documents/GitHub/R_Training/EDA`

- dataMaid v1.4.1 [Pkg: 2021-10-08 from CRAN (R 4.1.0)]

- R version 4.1.3 (2022-03-10).

- Platform: x86_64-apple-darwin17.0 (64-bit)(macOS Big Sur/Monterey 10.16).

- Function call: `makeDataReport(data = sal, output = "gg.html", replace = TRUE)`